

2001

## Molecular genetic investigation of autosomal dominant muscular dystrophy

Christopher Meredith  
*Edith Cowan University*

Follow this and additional works at: <https://ro.ecu.edu.au/theses>



Part of the [Medical Genetics Commons](#)

---

### Recommended Citation

Meredith, C. (2001). *Molecular genetic investigation of autosomal dominant muscular dystrophy*. Edith Cowan University. Retrieved from <https://ro.ecu.edu.au/theses/1509>

This Thesis is posted at Research Online.  
<https://ro.ecu.edu.au/theses/1509>

*Theses*

*Theses: Doctorates and Masters*

---

*Edith Cowan University*

*Year 2001*

---

Molecular Genetic Investigation Of  
Autosomal Dominant Muscular  
Dystrophy

Christopher Meredith  
Edith Cowan University

This paper is posted at Research Online.  
<http://ro.ecu.edu.au/theses/1509>

# Edith Cowan University

## Copyright Warning

You may print or download ONE copy of this document for the purpose of your own research or study.

The University does not authorize you to copy, communicate or otherwise make available electronically to any other person any copyright material contained on this site.

You are reminded of the following:

- Copyright owners are entitled to take legal action against persons who infringe their copyright.
- A reproduction of material that is protected by copyright may be a copyright infringement. Where the reproduction of such material is done without attribution of authorship, with false attribution of authorship or the authorship is treated in a derogatory manner, this may be a breach of the author's moral rights contained in Part IX of the Copyright Act 1968 (Cth).
- Courts have the power to impose a wide range of civil and criminal sanctions for infringement of copyright, infringement of moral rights and other offences under the Copyright Act 1968 (Cth). Higher penalties may apply, and higher damages may be awarded, for offences and infringements involving the conversion of material into digital or electronic form.

## USE OF THESIS

The Use of Thesis statement is not included in this version of the thesis.



**MOLECULAR GENETIC INVESTIGATION  
OF  
AUTOSOMAL DOMINANT MUSCULAR DYSTROPHY**

**CHRISTOPHER MEREDITH**

**BSc Hons *Edln*, Cert Ed *Lond*, MSc *Durh***

**DOCTOR OF PHILOSOPHY THESIS**

**FACULTY OF COMMUNICATIONS, HEALTH AND SCIENCE**

**EDITH COWAN UNIVERSITY**

## **Abstract**

This thesis contributes to the Human Genome Project by adding detail to the physical and genetic maps of the human genome, and by identifying a strong candidate gene for a form of distal myopathy.

Genomic clones for the human skeletal muscle genes slow troponin (*TNNI1*), alpha actin (*ACTA1*), and  $\beta$ -tropomyosin (*TPM2*) were isolated for use in the fluorescent *in situ* hybridisation localisation of these genes on the cytogenetic map of the human genome. The localisation of these genes made them potential candidates for inherited skeletal muscle diseases, including the muscular dystrophies investigated here.

Microsatellite, VNTR and RFLP markers were used in a search for linkage to a novel form of distal myopathy segregating in a Western Australian family. The decadic logarithm of the likelihood ratio, or 'lod score' method, was used to determine linkage between markers and this distal myopathy gene. A 22.4 cM candidate region was identified at 14q11.2. This was the first localisation of a distal myopathy gene. The Human Genome Organisation Nomenclature Committee reserved MPD1, 'myopathy, distal 1', for this form of distal myopathy, now known as Laing myopathy.

The MPD1 candidate region was excluded as the disease gene location for two other forms of distal myopathy. Silburn myopathy in 1994, which established the genetic heterogeneity of distal myopathy, and Felice myopathy in 1996. The exclusion of the MPD1 and French-Canadian OPMD candidate regions as disease gene locations for a putative-OPMD segregating in a Western Australian family, proved that this disease gene did not lie at 14q11.2.

Testing an MPD1 muscle-specific candidate gene for the Laing myopathy mutation, the myosin heavy polypeptide 7 gene (*MYH7*), identified seven base changes between the MPD1 proband sequence and the published *MYH7* cDNA sequence. All of these base changes were found in eight unrelated, unaffected Western Australians, therefore none of them were the Laing myopathy mutation. Two further differences to the published *MYH7* sequence segregated exclusively with the MPD1 proband. One of these, the *MYH7* G5073C (cDNA)/G23628C (gDNA) base change, caused a critical change to the *MYH7*  $\beta$ -myosin heavy chain polypeptide product ( $\beta$ -MyHC). An A1663P  $\beta$ -MyHC substitution. G23628/C23628 segregated with Laing myopathy in the Western Australian distal myopathy family. This segregation was confirmed by a single-strand conformation polymorphism test, then used to test 256 unaffected chromosomes. None possessed *MYH7* C23628.

Two patients from European distal myopathy families phenotypically similar to Laing myopathy, the Voit and Scoppetta families, were tested for the presence of *MYH7* gDNA G23628/C23628 heterozygosity. Both were homozygous *MYH7* G23628. One of these patients (Voit) was also tested for *MYH7* cDNA G5073/C5073 heterozygosity. She was homozygous *MYH7* G5073.

An analysis of the effect of the  $\beta$ -MyHC A1663P substitution at various levels of protein structure strengthened the candidature of *MYH7* G5073C as the Laing myopathy mutation. It demonstrated the extreme rarity of the  $\beta$ -MyHC A1663P substitution; it showed that this substitution did have a detrimental effect on coiled-coil formation; and it identified ways in which the  $\beta$ -MyHC A1663P substitution could disrupt myofibrillogenesis or contractility. Future research directions are identified and the contribution of this work to evolving concepts in muscular dystrophy is evaluated.

## **DECLARATION**

To the best of my knowledge and belief, I certify that this thesis:

- (i) does not incorporate without acknowledgement any material previously submitted for a degree or diploma in any institution of higher education;
- (ii) does not contain any material previously published or written by another person except where due reference is made in the text;
- (iii) reports only research actually carried out by this candidate except where due reference is made in the text;
- (iv) does not contain any defamatory material;
- (v) conforms to guidelines concerning text length and format.

C. Meredith  
Doctoral Candidate

January 28, 2001

## **ACKNOWLEDGEMENTS**

I gratefully acknowledge help from the following people.

Professor Byron Kakulas, Medical Director of the Australian Neuromuscular Research Institute, for his kind hospitality throughout the duration of this thesis.

Professor Frank Mastaglia, Director of the Neuromuscular Clinic of the Australian Neuromuscular Research Institute (ANRI), Dr. Beverly Phillips and Dr. Timothy Day who clinically ascertained the Western Australian families in this investigation.

Dr. Peter Silburn, Dr. R.J. Conrad and Dr. A.E.G. Tannenberg of the Department of Neurology, Princess Alexandra Hospital, Brisbane who clinically ascertained the Queensland distal myopathy family. Dr. Kevin Felice, Department of Neurology, University of Connecticut School of Medicine, Farmington, USA, who clinically ascertained the Connecticut distal myopathy family.

Professor Alan Bittles, the Director of the Centre for Human Genetics at Edith Cowan University, for his advice and support throughout this project.

Dr Angus Stewart, Senior Lecturer at Edith Cowan University who first introduced me to ANRI and the possibility of conducting research within this Research Institute.

Associate Professor Nigel Laing, Chief, Molecular Neurogenetics Laboratory at ANRI, who conceived and directed this project. I was fortunate to work in Nigel's laboratory, the best molecular neurogenetics school in the world. Thanks NGL.

The molecular neurogenetics group at ANRI: in the early years Dr. Anthony Akkari, Bernadette Majda and Dr. Jo Stanton, and more recently Lori Blechynden, Dr. Nici Binz, Hayley Durling, Dr. Kerry Garret, Kristen Nowak, Alison Oddy, and Kendall Walker. Your encouragement and assistance on a wide range of topics over such a long time was magnificent. I could not have wished for a better working environment.

Dr. Steve Wilton, Dr. Sue Fletcher and Christo Mann of the gene therapy group at ANRI always provided valuable advice, and Kaite Honeyman as an undergraduate work-experience student assisted with the localisation of Laing myopathy.

Dr. Michelle Byrnes and Dr. Bruno Meloni for advice and proof reading, and Brendan Barlow, Jude Newberry and Margaret Taylor for their constant encouragement.

Finally to my family, Brenda, Alexander, Heather and Andrew. Their sustained support was invaluable throughout this thesis. You can take my photo off the fridge now. I'll be coming home soon.

# CONTENTS

Use of Thesis	(i)
Title Page	(ii)
Abstract	(iii)
Declaration	(v)
Acknowledgements	(vi)
Contents	(viii)
<b>1.0 INTRODUCTION</b>	<b>1</b>
1.1 AIMS	3
1.2 RATIONALE	4
1.2.1 Clinical Heterogeneity	4
1.2.2 Muscular Dystrophy Nosology	7
1.2.3 Molecular Pathology, Diagnosis, Prognosis, Prevention and Treatment	10
1.2.4 Ascertainment	12
<b>2.0 HUMAN DISEASE GENES</b>	<b>14</b>
2.1 HUMAN INHERITANCE	15
2.1.1 Patterns of Inheritance	15
2.1.2 Inherited Diseases	16
2.2 HUMAN GENOME PROJECT	18
2.3 GENETIC MAP OF THE HUMAN GENOME	19
2.3.1 Recombination Fraction	20
2.3.2 Genetic Markers	22
2.3.3 Polymerase Chain Reaction	23
2.3.4 Unity of Gene Mapping	26
2.3.5 Connecting Meioses ( <i>m</i> )	26

## Contents

2.3.6	Likelihood	28
2.3.7	Two-point Lod Score	31
2.3.8	Multipoint Lod Score	32
2.3.9	Parametric Methods: Co-segregation Linkage Analysis	33
2.3.10	Nonparametric Methods: Allele-Sharing	35
2.3.11	Nonparametric Methods: Allelic Association	37
2.3.12	Resolving Power	40
2.3.13	Combined Mapping Strategies	41
2.3.14	Genetic Map Limitations	43
2.4	PHYSICAL MAP OF THE HUMAN GENOME	47
2.4.1	Fluorescent <i>in situ</i> Hybridisation	47
2.4.2	Clone Contigs	49
2.4.2.1	Sequence Tagged Site Content Mapping	52
2.4.3	Multi-Level Physical Maps	53
2.5	CLONING HUMAN DISEASE GENES	53
2.5.1	Functional Cloning	54
2.5.1.1	Oligonucleotides and Antibodies	54
2.5.1.2	Functional Complementation	55
2.5.1.3	Muscular Dystrophy Disease Genes	56
2.5.2	Position-Independent Candidate Approach	56
2.5.3	Positional Cloning	57
2.5.3.1	Transcript Map of Candidate Regions	58
2.5.3.2	Expressed Sequence Tags	59
2.5.4	Positional Candidate Approach	62
2.5.5	Mutation Detection	64
2.5.5.1	Direct Sequencing of Transcripts	65
2.5.5.2	DNA Single-Strand Conformation Polymorphism	66
2.5.5.3	Repeat Expansion Detection	67
2.6	CONCLUSION	69



## *Contents*

<b>3.0</b>	<b>MUSCULAR DYSTROPHY: A CANDIDATE APPROACH</b>	<b>71</b>
3.1	GENOMIC LIBRARIES	73
3.1.1	Bacterial Host Strains	74
3.1.2	Growth Media	74
3.1.3	Glycerol Stocks and Inoculation	75
3.1.4	Plating Genomic Libraries	75
3.1.5	Titering Genomic Libraries	76
3.1.6	Determining Insert Size and pfu	77
3.2	PROBING GENOMIC LIBRARIES	77
3.2.1	Lifting Libraries	77
3.2.2	Labelling Probes	78
3.2.3	Pre-hybridisation	81
3.2.4	Hybridisation	82
3.2.5	Positive Primary Clones	84
3.2.6	Secondary and Subsequent Screening	85
3.3	RESULTS	86
3.3.1	Genomic Libraries and Probes	86
3.3.2	Hybridisation	87
3.3.3	FISH Localisation	87
3.4	CONCLUSION	90
<b>4.0</b>	<b>THE LOCALISATION OF AN AUTOSOMAL DOMINANT DISTAL MYOPATHY GENE</b>	<b>92</b>
4.1	DISTAL MYOPATHY NOSOLOGY	93
4.2	ASCERTAINMENT OF WESTERN AUSTRALIAN FAMILY SEGREGATING DISTAL MYOPATHY	95
4.2.1	Clinical Features	95
4.2.2	DNA Extraction	102

## *Contents*

4.3	CO-SEGREGATION LINKAGE ANALYSIS	102
4.3.1	Genetic Markers	103
4.3.2	Two-point Linkage Analysis	106
4.3.3	Marker Haplotypes from D14S72 (q11) to D14S76 (q11-q32) for the WAMPD Family	109
4.3.4	Multipoint Linkage Analysis	111
4.2.6	EXCLUDE	112
4.4	REFINEMENT OF THE MPD1 LINKAGE REGION	115
4.4.1	New Genetic Markers	115
4.4.2	Final Two-point Linkage Analysis	117
4.5	REPEAT EXPANSION DETECTION	122
4.6	CONCLUSION	122
5.0	<b>EXCLUSION OF THE MPD1 CANDIDATE REGION FOR TWO OTHER AUTOSOMAL DOMINANT DISTAL MYOPATHIES</b>	 <b>124</b>
5.1	AUTOSOMAL DOMINANT DISTAL MYOPATHY SEGREGATING IN A QUEENSLAND FAMILY	 127
5.1.1	Ascertainment	127
5.1.2	Genetic Markers	132
5.1.3	Chromosome 14 Two-point Linkage Analysis	132
5.1.4	Partial Genome Screen	135
5.1.5	EXCLUDE	137
5.2	AUTOSOMAL DOMINANT DISTAL MYOPATHY SEGREGATING IN A CONNECTICUT FAMILY	 140
5.2.1	Ascertainment	140
5.2.2	Genetic Markers	145
5.2.3	Chromosome 14 Two-point Linkage Analysis	145
5.2.4	Multipoint Linkage Analysis: MPD1 Candidate Region	147

## *Contents*

5.3	LAING, SILBURN AND FELICE MYOPATHIES	150
5.4	THREE OTHER DISTAL MYOPATHY LOCI	150
5.5	CONCLUSION	153
<b>6.0</b>	<b>EXCLUSION OF THE MPD1 AND OPMD CANDIDATE REGIONS FOR A PUTATIVE-OPMD SEGREGATING IN A WESTERN AUSTRALIAN FAMILY</b>	<b>155</b>
6.1	OPMD NOSOLOGY	157
6.1.1	Oculopharyngodistal Myopathy	159
6.1.2	Molecular Genetics and OPMD Nosology	162
6.2	AN AUTOSOMAL DOMINANT PUTATIVE-OPMD SEGREGATING IN A WESTERN AUSTRALIAN FAMILY	163
6.2.1	Ascertainment	163
6.2.2	Genetic Markers	165
6.2.3	Chromosome 14 Two-point Linkage Analysis	166
6.3	CONCLUSION	171
<b>7.0</b>	<b>LAING MYOPATHY: POSITIONAL CANDIDATE APPROACH</b>	<b>173</b>
7.1	<i>MYH6</i> LAING MYOPATHY CANDIDATURE	174
7.1.1	<i>MYH6</i> Expression	175
7.1.2	<i>MYH6</i> Mutations and Disease	175
7.2	<i>MYH7</i> LAING MYOPATHY CANDIDATURE	176
7.2.1	<i>MYH7</i> Expression	177
7.2.2	<i>MYH7</i> Mutations and Disease	178

## Contents

7.3	LAING MYOPATHY MUTATION: TESTING <i>MYH7</i>	179
7.3.1	Methods	
7.3.1.1	RNA Extraction from Skeletal Muscle	179
7.3.1.2	Skeletal Muscle cDNA Manufacture	181
7.3.1.3	<i>MYH6</i> and <i>MYH7</i> Primer Design	183
7.3.1.4	<i>MYH6</i> Primers	185
7.3.1.5	<i>MYH7</i> Primers	185
7.3.1.6	DNA Amplification Conditions	187
7.3.1.7	DNA Purification	189
7.3.1.8	DNA Sequencing Reactions	190
7.3.1.9	SSCP Screening of <i>MYH7</i> gDNA	191
7.3.2	Results	192
7.3.2.1	Skeletal Muscle RNA	192
7.3.2.2	Skeletal Muscle cDNA	194
7.3.2.3	<i>MYH6</i> Amplification	195
7.3.2.4	<i>MYH6</i> Amplicon Sequencing	195
7.3.2.5	<i>MYH7</i> Amplification	197
7.3.2.6	<i>MYH7</i> Amplicon Sequencing	200
7.3.2.7	Base Differences in Western Australian <i>MYH7</i> cDNA	200
7.3.2.8	Base Differences in WAMPD Proband <i>MYH7</i> cDNA	205
7.3.2.9	<i>MYH7</i> gDNA G23628C Segregation in WAMPD Family	207
7.3.2.10	Testing Two European Distal Myopathy Families for <i>MYH7</i> cDNA G5073C/gDNA G23628C	211
7.3.2.11	SSCP Screening of <i>MYH7</i> gDNA for G23628C	212
7.4	CONCLUSION	215

## Contents

<b>8.0</b>	<b>MYH7 G5073C: THE LAING MYOPATHY MUTATION OR A RARE POLYMORPHISM?</b>	<b>217</b>
<b>8.1</b>	<b>LAING MYOPATHY: SARCOMERIC GENE MUTATIONS</b>	<b>218</b>
8.1.1	Do <i>MYH7</i> Mutations Have a Cardiac Phenotype?	219
8.1.2	Do <i>MYH7</i> Mutations Cause Skeletal Muscle Myopathy?	223
8.1.3	Which CMH Disease Genes are Associated with Skeletal Muscle Myopathy?	224
8.1.4	Do Substitutions in the $\beta$ -MyHC LMM Rod Region Cause Defects in Specific Muscles?	227
8.1.5	Does the LMM Rod Region of Human $\beta$ -MyHC Interact with Other Molecules?	229
8.1.6	Do Other MYH Mutations Cause Skeletal Muscle Myopathy?	230
8.1.7	Sarcomeric Gene Mutations and Laing Myopathy: An Evaluation	230
<b>8.2</b>	<b>A1663P AND THE PRIMARY STRUCTURE OF THE <math>\beta</math>-MyHC ROD</b>	<b>232</b>
8.2.1	Derived Primary Structure of $\beta$ -MyHC	233
8.2.2	Comparative Analysis of MyHC Rod Primary Structure	236
<b>8.3</b>	<b>SECONDARY STRUCTURE OF THE <math>\beta</math>-MyHC ROD</b>	<b>239</b>
8.3.1	The $\alpha$ -helix: Structure, Direction and Sense of Twist	239
8.3.2	Proline the $\alpha$ -helix Breaker: Rigid-Geometry Analysis	241
8.3.3	The Validity of Rigid-Geometry Calculations	245
8.3.4	Secondary Structure of $\alpha$ -helices with Internal Proline	248
<b>8.4</b>	<b>QUATERNARY STRUCTURE OF MYOSIN II</b>	<b>252</b>
8.4.1	Deformed $\alpha$ -helices and Coiled-Coils	252
8.4.2	MyHC Dimerisation Geometry: Heptad Repeats	256

## *Contents*

8.4.3	A1663P Distorts Perpendicular and Parallel Geometry	259
8.4.4	Proline Effect on MyHC Dimerisation: COILS Analysis	262
8.5	SUPERSTRUCTURE OF MYOSIN II	265
8.5.1	Thick Filament Formation: Charged-Repeat Attraction	266
8.5.2	Thick Filament Formation: Charge-Phase Distortion	268
8.5.3	$\beta$ -MyHC A1663P and Azimuthal Charge Disruption	270
8.5.4	$\beta$ -MyHC A1663P and the Assembly Conserved Domain	272
8.6	CONCLUSION	272
<b>9.0</b>	<b>DISCUSSION</b>	<b>274</b>
9.1	EVOLVING CONCEPTS IN MUSCULAR DYSTROPHY	276
9.1.1	Molecular Pathology has Refined MD Nosology	277
9.2	EVOLVING CONCEPTS IN DISTAL MYOPATHY	280
9.2.1	Molecular Genetics Confirms and Refines Traditional Distal Myopathy Nosology	281
9.2.2	Nonaka Distal Myopathy and a form of h-AR-IBM may have a Common Disease Gene	287
9.2.3	Different Aetiologies Can Produce Similar Pathogenic Cascades	289
9.2.4	Markesbery Distal Myopathy and TMD may have the Same Disease Gene	291
9.2.5	The Same Mutation can produce different Pathogenesis	292
9.2.6	Distal Myopathy May be Caused by Mutations in Genes Encoding Contractile Proteins	295
9.2.7	Childhood Onset Distal Myopathy	295
9.3	EVOLVING CONCEPTS IN OPMD	298
9.3.1	Locus Heterogeneity Unlikely for OPMD	299
9.4	CONCLUDING COMMENTS	301

## Contents

REFERENCES		304
APPENDIX A	Thesis Publications	353
APPENDIX B	Lod Scores for the Segregation of Marker Loci with WAMPD and QMPD Disease Loci	375
APPENDIX C	<i>MYH6</i> , and <i>MYH7</i> ; $\alpha$ -MyHC and $\beta$ -MyHC	382
APPENDIX D	Abbreviations and Symbols; Tables and Figures	399

## **CHAPTER 1**

### **INTRODUCTION**



## **1.0 INTRODUCTION**

Clinical medicine has undergone significant changes during the past ten years as molecular genetics has uncovered many mutations causing inherited human diseases. Objective confirmation of diagnosis and accurate disease risk-assessment for asymptomatic individuals have become increasingly available for families segregating inherited diseases. Based on these developments, this thesis investigated autosomal dominant (AD) muscular dystrophies classified as distal myopathy and oculopharyngeal muscular dystrophy (OPMD).

The distal myopathies comprise a heterogeneous group of disorders showing both AD and autosomal recessive (AR) inheritance, with diverse phenotypic features and pathological changes (Mastaglia, 1991, pp. 1300-1301). They share the clinical feature of myopathic weakness beginning in the distal muscles of the lower and/or upper extremities (Walton & Gardner-Medwin, 1981, pp. 505-506). OPMD was first described in 1915 by Taylor (cited in Brais et al., 1995) who noted progressive vagus-glossopharyngeal paralysis with ptosis (eyelid droop) of familial nature. OPMD is now known to have world-wide incidence, having been reported in more than 20 countries (Brais et al., 1995).

Some inherited forms of muscular dystrophy display both OPMD and distal myopathy symptoms (Satoyoshi & Kinoshita, 1977). The possibility of allelic heterogeneity and/or variable expressivity was one reason for the joint investigation of these two types of muscular dystrophy. Gelehrter and Collins, (1990, pp. 299 & 311) define allelic heterogeneity as "similar or identical phenotypes caused by different mutant alleles at the same genetic locus", and variable expressivity as "the variable severity of a genetic trait".

This thesis has five aims. To:

- (i) isolate genomic clones for human skeletal muscle slow troponin (*TNNI1*), alpha actin (*ACTA1*) and  $\beta$ -tropomyosin (*TPM2*) genes, for use in the *in situ* cytogenetic mapping of these skeletal muscle contractile protein genes (a candidate approach);
- (ii) localise the disease gene causing a form of AD-distal myopathy segregating in a Western Australian family;
- (iii) test the Western Australian distal myopathy candidate region as the possible location for two other phenotypically different forms of AD-distal myopathy, one segregating in a Queensland family, the other in a Connecticut family;
- (iv) test the Western Australian distal myopathy candidate region and the French Canadian OPMD candidate region, as possible locations for a form of AD-OPMD segregating in a Western Australian family;
- (v) test muscle-specific candidate genes within the Western Australian distal myopathy candidate region for a possible disease-causing mutation (the positional candidate approach).

## **1.2**

## **RATIONALE**

The rationale of this thesis is, "Knowledge of the structure and function of the human genome will improve the nosology, diagnosis, prognosis, prevention, treatment and future research into the molecular pathology of AD-muscular dystrophy".

### **1.2.1 Clinical Heterogeneity**

Until recently, nosology, diagnosis, prognosis and research into the molecular pathology of muscular dystrophy (MD) was based almost entirely on clinical phenotype. Gelehrter and Collins (1990, p. 307) define phenotype as "the observed result of the interaction of genotype with environmental factors, the observable expression of a particular gene or genes". These same authors (1990, p. 303) define genotype as "the genetic constitution of an individual or more specifically, the alleles [alternative forms of genes] at specific loci".

Clinical phenotype usually develops its characteristic signs and symptoms after a long chain of causation whose individual events and timing are unknown. Under these circumstances, clinical heterogeneity is inevitable with genetically identical diseases displaying different symptoms or, conversely, with genetically different diseases displaying similar symptoms. In all cases, clinical heterogeneity reduces nosologic, diagnostic and prognostic accuracy and diminishes research accuracy.

Until recently, even comprehensive diagnostic testing could not overcome the difficulty of clinical heterogeneity. Before the cloning of the gene responsible for the human muscular dystrophies, Duchenne and Becker MDs (Koenig et al.,

1987), diagnosis of these diseases was based on family history, clinical findings, serum muscle enzyme levels, muscle biopsy, electromyography and electrocardiography (Kakulas & Adams, 1985 cited in Laing, 1993, p. 45). Despite the thoroughness of this protocol, its major weakness was that all the elements relied on phenotypic tests of genotypic status, and phenotype can be highly variable. At the commencement of this thesis in 1993, all inherited forms of distal myopathy and OPMD were diagnosed on the basis of phenotype. Diagnostic accuracy for these forms of MD was inevitably confounded by clinical heterogeneity. Two approaches had been developed to address clinical heterogeneity. The first involved a detailed mathematical analysis of clinical observations, the second was based on molecular genetics.

There have been several mathematical attempts to provide a degree of objectivity to clinical diagnosis based on phenotype (Preus & Ayme, 1983; van den Anker et al., 1993; Verloes 1995). This later attempt (Verloes, 1995, p. 433) submitted fifty-five published cases, belonging to five overlapping phenotypically related multiple congenital anomaly syndromes, to principal factor analysis, hierarchical clustering and graphical scaling, due to:

the inability of the human brain to do 'unweighted' and unbiased multivariate analysis i.e., the impossibility of analyzing a phenotype without some preconceived view of what is discriminant and what is secondary and also the inability to perceive similarities between cases compared for dozens of manifestations.

Twenty-nine cardinal signs used to define each syndrome were selected for analysis. Dendograms derived from cluster analysis showed that the five selected syndromes, hydroletharus, Smith-Lemli-Opitz type II, orofaciogigital

type VI, holoprosencephaly-polydactyly and Pallister-Hall syndromes, evidently constituted independent phenotypic entities as clinically described. Overlapping cases could usually be placed, unambiguously, into one of these categories. In three cases this required revision of the previous Pallister-Hall syndrome diagnosis since they "have probably been abusively diagnosed because of an undue stress placed on the presence of a CNS hamartoma," (Verloes, 1995, p. 441).

This type of analysis is called numerical syndromology. It produces results comparable to the subjective pattern recognition process characterising MCA syndrome diagnoses by physicians. Some objective patterns of anomalies differing between closely similar syndromes may have been identified. Nevertheless, numerical syndromology seems inordinately complex and, from the outset, quite limited in expectation. It acknowledges; that it is a heuristic process not supported by statistical reasoning, that biased selection of data rather than causal entity may account for clusters, that it cannot deliver any hierarchy of syndromes and that it cannot be expected to solve the problem of genetic heterogeneity among phenotypically related cases.

Numerical syndromology has had limited success at introducing objectivity to clinical diagnosis. It can improve the accuracy with which syndromes are categorised and described, but that is all. No matter what mathematical model it adopts, this approach can never address the underlying causes of the syndromes. It can never identify disease genes, those genes that when mutated give rise to human disease (Collins, 1992, p. 3). In contrast, molecular genetics can identify disease genes, improving disease classification, diagnosis and, sometimes, prognosis.

The process of linking genetic change to clinical phenotype commenced with the discovery of disease-causing mutations by Pauling et al. (1949), with each newly-discovered mutation suggesting further avenues of research into the molecular pathology of the disease that it caused. For these reasons, molecular genetics was the approach adopted in this thesis. It provided the opportunity to make the complexities of clinical heterogeneity less refractory.

### **1.2.2 Muscular Dystrophy Nosology**

Human MDs are defined as "those conditions in man in which there is a primary, progressive degenerative disease of skeletal muscle. The characteristic features include a symmetric distribution of muscular atrophy, an intact nerve supply, and the liability to heredo-familial incidence." (Adams, et al., 1962, p. 78). The distal myopathies and the OPMD investigated in this thesis are MDs.

MD has become a general term to describe "a group of inherited and progressively debilitating myogenic disorders which can be classified into more than 20 diseases on the basis of genetic as well as clinical features." (Ozawa et al., 1998, p. 421).

The power of molecular genetics to contribute to MD nosology became apparent with the cloning of the first major human disease gene identified using the 'new genetics', the dystrophin gene (Koenig et al., 1987). Four technological advances have contributed most to the identification of disease genes: DNA banking technology providing a permanent supply of a family's DNA (Neitzel, 1986); the polymerase chain reaction (PCR) (Saiki et al., 1985); rapid PCR based linkage studies (Weber & May, 1989); and rapid mutation screening technologies (Saiki et al., 1988; Vosberg, 1989; Williams, 1989; Bloch, 1991).

The first MDs to be reclassified by molecular genetics were dystrophinopathies, the MDs caused by mutations in the dystrophin gene. The phenotypes of dystrophinopathies occur "with a wide range of neuromuscular symptoms, and neither male sex nor proximal weakness are diagnostic prerequisites for consideration of an underlying dystrophin abnormality." (Hoffman & Schwartz, 1991, p. 175). Table 1.1 (p. 9) shows the early impact of molecular genetics on MD nosology.

Apart from the dystrophinopathies, by 1993 only myotonic dystrophy diagnosis could be confirmed by mutation type (Brook et al., 1992). In addition, genetic location could be used to confirm disease diagnosis for X-linked myotubular myopathy (MTM) (Liechti-Gallati et al., 1991), Emery-Dreifuss MD (Consalez et al., 1991), facioscapulohumeral MD (FSHD) (Wijmenga et al., 1992), severe childhood AR-MD (SCARMD) (Ben Othmane et al., 1992) and limb girdle MD (LGMD) (Beckmann et al., 1991; Speer et al., 1992). SCARMD has been classified as a subgroup of the AR-LGMD group (Bushby & Beckmann, 1995), but debate continues about SCARMD nosology (Ozawa et al. 1998, p. 433).

Table 1.1 also shows that genetic linkage had not been demonstrated in any form of distal myopathy or OPMD at the commencement of this project, nor had any candidate gene been identified for either of these disorders (MIM 160500 & MIM 164300; McKusick, 1992).

Table 1.1

**Muscular Dystrophy Nosology, 1993**

Dystrophy type	Identification	Cytogenetic Location		Gene	
<b><u>X-linked recessive</u></b>					
Dystrophinopathies	Koenig et al., 1987				
Duchenne	Meryon, 1852 (cited in Partridge, 1993, p. xiii)	Xp21.2	Boyd & Buckle, 1986	Dystrophin ( <i>DMD</i> )	Koenig et al., 1987
Becker	Becker, 1955 (cited in Ozawa, et al., 1998, p. 421)	Xp21.2	Koenig et al., 1987	Dystrophin ( <i>DMD</i> )	Koenig et al., 1987
Quadriceps myopathy	Thage, 1965	Xp21.2	Sunohara et al., 1990	Dystrophin ( <i>DMD</i> )	Koenig et al., 1987
Myalgia and cramps	Kuhn et al., 1979	Xp21.2	Gospe et al., 1989	Dystrophin ( <i>DMD</i> )	Koenig et al., 1987
Emery-Dreifuss MD	Emery & Dreifuss, 1966	Xq28	Consalez et al., 1991		
Myotubular myopathy (MTM1)	Engel et al., 1968	Xq28	Liechti-Gallati et al., 1991		
<b><u>Autosomal dominant</u></b>					
Myotonic dystrophy (DM1)	Bell, 1947	19q13.2-q13.3	Harley et al., 1991	CTG repeat expansion	Brook et al., 1992
Facioscapulohumeral MD	Tyler & Wintrobe, 1950				
FSHD	Wijmenga et al., 1992	4q35	Wijmenga et al., 1992		
Limb-girdle MD (LGMD)	Walton & Natrass, 1954				
LGMD	Gilchrist et al., 1988	5q31-q33	Speer et al., 1992		
LGMD	Panegyres et al., 1990				
Bethlem myopathy	Bethlem et al., 1976				
Distal myopathy	Gowers, 1902				
Welander (WMD)	Welander, 1951				
Finnish Tibial	Udd et al., 1991				
Markesbery	Markesbery et al., 1974				
Oculopharyngeal MD	Taylor, 1915 (cited in Brais et al., 1995)				
OPMD	Barbeau, 1965 (cited in Brais et al., 1995)				
Oculopharangodistal myopathy	Satoyoshi & Kinoshita (1977)				
<b><u>Autosomal recessive</u></b>					
Severe Childhood AR-MD	Kloepfer & Talley, 1958				
SCARMD	Passos-Bueno et al., 1991a & b	13q	Ben Othmane et al., 1992		
Limb girdle MD	Walton & Natrass, 1954				
LGMD	Shokeir & Kobrinsky, 1976				
LGMD	Panegyres et al., 1990				
LGMD	Beckmann et al., 1991	15q15.1-q21.1	Beckmann et al., 1991		
LGMD	Mahjneh et al., 1992				
Congenital MD	Batten, 1909 (cited in Markesbury, et al., 1974, p. 127)				
*Fukuyama	Fukuyama, et al., 1981				
Merosin deficiency	Tome et al., 1994				
*Walker-Warburg syndrome	Warburg, 1976				
*Muscle, eye and brain (MEB) disease	Santavuori & Leisti 1980				
Distal myopathy	Gowers, 1902				
Miyoshi (MM)	Miyoshi et al., 1975 & 1986				
Nonaka	Nonaka et al., 1981 & 1985				
Hereditary inclusion body myopathy (h-IBM)	Askanas & Engel 1993				

\* with neurogenic involvement



### **1.2.3 Molecular Pathology, Diagnosis, Prognosis, Prevention and Treatment**

Nosology is a means, not an end. It groups diseases according to perceived similarities in pathology. Its precision depends upon knowledge of pathological mechanisms. Molecular genetics has been able to identify primary causes of inherited disease, the first step in their molecular pathology, by identifying disease gene locations and disease-causing mutations. But improved knowledge of molecular pathology has not only lead to refined nosology. More importantly, it has lead to better diagnosis and prognosis and better methods of prevention. This became evident shortly after the cloning of the dystrophin gene.

The dystrophinopathies belong to an allelic series, a term first used by McKusick (1973, p. 446) to describe phenotypic diversity of human diseases caused by mutations in the same gene. Monaco et al. (1988) demonstrated that mutation type could contribute to the severity of dystrophinopathy. They linked disease severity and mutation type by discovering that the size or location of dystrophin mutations did not correlate with the clinical severity of DMD or BMD. In an attempt to explain these observations, Monaco et al. (1988) proposed the 'reading frame theory' which hypothesised that disease severity was directly correlated to the effect of the genetic lesions on the open reading frame (ORF) of messenger RNA (mRNA) transcript triplet codons.

Analysis of the mutations in lymphocyte DNA from many hundreds of DMD and BMD patients has shown that severe DMD is predominantly a consequence of intragenic deletions that cause translational reading frame shifts in the mRNA (Koenig et al., 1989; den Dunnen et al., 1989). Stop codons generated by these frameshift mutations result in the premature termination of

protein synthesis during translation, leading to the production of truncated dystrophin molecules lacking a carboxyl terminus. These truncated molecules are non-functional and rapidly degraded, explaining the lack of dystrophin in these DMD patients.

Conversely, the deletions or duplications resulting in the less severe BMD patients bring together exons that maintain the reading frame, resulting in the synthesis of shorter, lower molecular weight semi-functional dystrophin (Koenig, et al., 1989; den Dunnen et al., 1989). A large study of 256 deletions along the entire length of the dystrophin gene, found that the 'ORF frameshift theory' of Monaco et al. (1988) was consistent with phenotype in 92% of cases (Koenig et al., 1989).

Beggs et al. (1991) confirmed that this was the molecular basis for phenotype variability among patients suffering from dystrophinopathies. In addition, these authors established that deletions within the amino-terminal domain tended to result in low levels of dystrophin and, consequently, a more severe phenotype. The phenotypes of patients with deletions or duplications in the central rod domain were found to be more variable. While agreeing that the discovery of disease-causing mutations could improve diagnosis and prognosis, these authors also added a cautionary note, "... however, phenotypic variability among patients with similar mutations suggests that epigenetic and/or environmental factors play an important role in determining the clinical progression." (Beggs et al., 1991, p. 54). This caution is revisited several times throughout this thesis.

Prevention of dystrophinopathies, particularly DMD, is achieved by prenatal diagnosis allowing known carriers the choice of terminating affected foetuses.

The status of male foetuses is most frequently determined by multiplex PCR (Chamberlain et al., 1988) because this has the advantage of detecting 98% of all dystrophin deletions in about 24 hours (Beggs et al., 1990).

Dystrophinopathy treatments based on replacement of damaged dystrophin have been under development for several years, (Anderson & Kunkel, 1992). There are three general approaches: myoblast transplantation, gene therapy and genetic therapy. The objective of all these approaches has been to increase the quantity of functional dystrophin in affected DMD tissues to levels of therapeutic value, effectively ameliorating DMD symptoms to that of BMD.

The cloning of the dystrophin gene, the substantiation of the 'ORF frameshift theory' of Monaco et al. (1988), improved diagnosis, prognosis and prevention for the dystrophinopathies and the promise of dystrophin replacement therapy, launched an exponential expansion of molecular genetic investigation into the MDs.

The rationale of this thesis was based on the success of these early investigations into the molecular genetics of the MDs. The expectation was that their success could be repeated, at least in part, for three forms of AD-distal myopathy and one form of AD-OPMD. Refinement of distal myopathy and OPMD nosology, diagnosis and prognosis appeared an achievable goal. Local factors also made the MD investigations in this thesis viable.

#### **1.2.4 Ascertainment**

Ascertainment of families segregating inherited disorders and the banking of their DNA is the first crucial step in any attempt to identify human disease genes.

Without families containing sufficient numbers of affected individuals, the cloning of human disease genes cannot proceed. Geographic isolation and an associated low rate of emigration, has led to the establishment of large families in Western Australia. Having family members all within one state is an important advantage for human disease gene cloning, because ascertainment and the banking of DNA is frequently the rate-limiting step of gene localisation.

Ascertainment in Western Australia is also greatly assisted by world-class diagnostic facilities and excellent medical records stretching back to the early years of colonisation. The identification of a proband in a large Western Australian family presented the possibility of tracing their inherited disease backwards and/or forwards through three, four or five generations. Exploiting these local advantages, this thesis investigated a large Western Australian family segregating a form of AD-distal myopathy over three generations. Medical records of this kinship stretched back more than 30 years, all living members resided in Western Australia, and, most importantly, this kinship contained nine affected members, sufficient to be confident of localising the responsible disease gene.

The ultimate rationale of all molecular genetic investigations into inherited human disease is to improve knowledge of their molecular pathology and so develop better nosology, diagnosis, prognosis, prevention, and treatments for each disease. This investigation into three forms of AD-distal myopathy and one form of AD-OPMD was based on this rationale.

## **CHAPTER 2**

### **HUMAN DISEASE GENES**

## **2.0 HUMAN DISEASE GENES**

The search for human disease genes using 'new genetics' began during the 1980s, driven by a desire to influence the course of genetic disease. As a direct result of this ongoing search, high-resolution genetic and physical maps of the human genome are now available, and it was predicted that the entire sequence of the human genome would be known by the year 2003 (G. Sutherland, personal communication, 1997; Collins et al., 1998, p. 687). In fact, 97% of the human genome sequence became available in July, 2000 (Golden & Lemonick, 2000).

### **2.1 HUMAN INHERITANCE**

Human phenotype is determined by a mixture of major and minor genetic determinants, together with environmental influences (Gelehrter & Collins, 1990, pp. 57 & 64; Strachan & Read, 1996, p. 61).

#### **2.1.1 Patterns of Inheritance**

The genotype at a locus often influences human phenotype to the extent that non-Mendelian patterns of inheritance result (Ott, 1997, p. 23). Traditionally, non-Mendelian human traits have been described as multifactorial, resulting from the interplay of multiple environmental factors with multiple genes. Multifactorial inheritance has been further subdivided into polygenic and oligogenic inheritance (Gelehrter & Collins, 1990, p. 57). Polygenic traits are those determined by many genes, called polygenes, each locus having a small effect on the determination of a trait (Strachan & Read, 1996, p. 81). Oligogenic traits are those determined by a small number of genes, called oligogenes, originally called 'leading factors' by Wright (1968, cited by Morton & Lio, 1997, p. 17),

each having a larger effect on the determination of a trait (Strachan & Read, 1996, p. 81).

### **2.1.2 Inherited Disease**

Human diseases with an inherited component (inherited human diseases) are classified in the same way as normal traits. Like these traits, most inherited human diseases do not display Mendelian segregation (Lander & Schork, 1994, p. 2037). Even when Mendelian inheritance is apparent, three common confounders of this pattern of disease segregation have been identified: locus heterogeneity, where similar phenotypes are caused by mutations at different genetic loci, e.g. retinitis pigmentosa (Kajiwara et al., 1991; Dryja, et al., 1990); variable expressivity, where individuals with the same mutant gene show variable clinical phenotype, e.g. peripheral neurofibromatosis (von Recklinghausen disease) (Collins et al., 1989); and reduced penetrance, where only a proportion of those with a disease-causing mutation actually express clinical phenotype, e.g. *RB1* gene mutations in retinoblastoma (Lohmann, 1999).

The majority of common inherited human diseases display a genetic component but their pattern of inheritance is unclear. They appear to have their clinical phenotype determined by multiple environmental and genetic factors (Ott, 1997, p. 23). The extent to which these diseases run in families is measured by the ratio of recurrence risk in monozygotic twins versus that in dizygotic twins, or the ratio of recurrence risk in a sibling versus population incidence (Plomin et al., 1994).

If individual disease liability is normally distributed with higher average liability for sibs of affected parents, a disease is classified as polygenic (Falconer, 1981,

cited by Strachan & Read, 1996, pp. 486-488), each polygene having a small effect on the determination of the disease. The locations of polygenes contributing to the expression of polygenic disease are called susceptibility loci (Strachan & Read, 1996, pp. 486-491). A recent genomic screen involving 519 markers and 147 independent affected sib pairs provided evidence that one form of autism was a polygenic disease, not an oligogenic one, with more than 15 susceptibility loci involved (Risch et al., 1999). Supporting this conclusion, other studies involving these same affected sib pairs have excluded the suggested oligogenic role of the chromosome 6p histocompatibility complex (HLA) loci in the genetic aetiology of autism (Rogers et al., 1999), and have excluded the chromosomal region 15q11-q13, also implicated as a region harbouring major susceptibility autism gene(s) (Salmon, et al., 1999).

Inherited human diseases are classified as oligogenic if each locus has sufficient effect to be detected by linkage analysis (Morton & Lio, 1997, p. 17). The locations of oligogenes contributing to the expression of oligogenic disease are called major susceptibility loci (Strachan & Read, 1996, pp. 486-491). Some major susceptibility loci for oligogenic diseases have been identified: HLA-B susceptibility locus for ankylosing spondylitis (Schlosstein et al., 1973), apolipoprotein E (APOE) susceptibility locus for late onset Alzheimer's (Corder et al., 1993), and Vitamin D Receptor (VDR) susceptibility locus for osteoporosis (Morrison et al., 1994).

Classification of human inheritance patterns is further complicated by observations that oligogenes may operate over a polygenic background and that some genes may appear as a Mendelian determinant for one character but as a polygene for another. Strachan and Read, (1996, p. 490) have suggested that, "between pure Mendelian and pure polygenic characters there



lies a whole spectrum of traits governed by major susceptibility loci, possibly operating over a polygenic background and sometimes subject to major environmental determinants".

This view sees all human characters and inherited diseases lying along a Mendelian - oligogenic - polygenic - environmental continuum. It is a panoramic view, but one which lacks important detail because it captures interacting environmental and genotypic events through a wide-angled phenotypic lens. Through this lens, developed pictures of inherited disease are inevitably blurred by clinical heterogeneity. Human molecular genetics has started to improve focus by adding genotypic lenses.

The cloning of human disease genes is the first step towards uncovering the molecular events determining clinical phenotype. Their cloning and analysis have already added, and will continue to add, functional detail to contemporary phenotype definitions of Mendelian, oligogenic and polygenic inheritance. As the cloning of human disease genes progresses, phenotypic descriptions of inherited human disease will be supplemented with patho-physiology details, based upon improved understanding of mutated gene action. Mechanisms underlying the clinical heterogeneity of inherited disease will begin to emerge, promoting further research into the molecular pathology of human genetic disease. Better treatments for affected individuals will develop at an increasing pace.

## **2.2 HUMAN GENOME PROJECT**

The cloning of human disease genes in the years 1984-1989 led to three observations: that altered gene function was an effective way to identify human genes; that the identification of disease-causing mutations improved both the

genetic and physical map of the human genome; and the efforts of individual laboratories to clone genes, 'one-at-a-time', was wasteful of time and resources. These observations first began to emerge in 2 reports published by the Department of Energy (DOE) in the USA; "Technologies for Detecting Heritable Mutations in Human Beings" (1985) and, "Report on the Human Genome Initiative" (1987). It was these reports that launched the idea of a dedicated Human Genome Project (HGP).

After a series of pilot projects and the involvement of the United States Office of Technology Assessment and the United States National Institutes of Health, HGP commenced in 1991, governed by the Human Genome Organisation (HUGO) (Strachan & Read, 1996, p. 336). Its major goal was to understand the structure of the human genome, its major rationale was the expectation that new advances in molecular genetics, would find application in the fields of diagnostic and therapeutic medicine (Collins, 1992, pp. 5-6).

From the DOE reports and pilot projects, three things became apparent: that the most effective way to track down the vast majority of human disease genes was to develop a high-resolution physical map of the human genome; that this could only be built from detailed genetic and physical maps; and that the localisation of disease genes on these maps would simultaneously improve both maps and lead to the cloning of disease genes.

## **2.3 GENETIC MAP OF THE HUMAN GENOME**

For the great majority of disease genes, the first step towards their cloning has been their localisation on the genetic map of the human genome. The

foundations for genetic mapping were laid when Alfred Sturtevant (1913) charted the locations of mutations affecting six *Drosophila melanogaster* developmental genes. He understood neither the biochemical basis of these defects nor even that genes were made from DNA. By correlating the co-inheritance patterns of six X-linked mutations he developed a genetic map based solely on abstract mathematical reasoning and not biochemistry: the closer two genetic markers were on a chromosome, the more frequent was their co-inheritance.

Sturtevant's abstract mathematical reasoning was the first demonstration of what is now called parametric linkage analysis. The HGP is developing a genetic map of the human genome based on parametric and nonparametric linkage analysis and by identification of allelic association.

### **2.3.1 Recombination Fraction**

Recombination fraction, or theta ( $\theta$ ), measures the frequency of crossovers between loci, and is defined as the fraction of meiotic events that show a recombination between two loci (Gelehrter & Collins, 1990, p. 309). It provides a measure of the genetic distance between two markers. If two markers are close together,  $\theta$  will be small with a value of zero attained when both markers coincide. If they lie further apart,  $\theta$  will be larger with an upper limit of 0.5, corresponding to 'recombination' between markers located on non-homologous chromosomes. Recombination fractions can never exceed 0.5 no matter how great the physical distance between them. This is immediately obvious for a single crossover and is also true for multiple crossovers (Strachan & Read, 1997, pp. 314-315).

Haldane (1919) showed that, if crossovers were distributed independently with respect to one another, occurring randomly along synapsed diplonema chromosomes and having no influence on each other, the number of crossovers between two markers was Poisson distributed. By summing alternative terms of the Poisson distribution, the probability of an odd number of crossovers (evidence of recombination) became:

$$\theta = (1 - e^{-2d}) / 2$$

where  $e$  is the base of natural logarithms and  $d$  is the map distance between the markers. This equation became known as Haldane's mapping function because it expressed the mathematical relationship between recombination fraction and genetic map distance. For small distances, this mapping function reduces to:

$$\theta = d$$

reflecting the fact that the possibility of more than one crossover can be neglected for closely linked markers.

Haldane (1919) defined two loci showing 1% recombination (a recombination fraction of 0.01) as separated by 1 centiMorgan (cM), in recognition of the genetic mapping of the *Drosophila sp.* fruit fly being carried out by Morgan's group at Columbia University. In 1916 Muller wrote a four-part review based on his doctoral thesis, "The mechanism of crossing over". Edwards (1997, p. 11), considered that Morgan's group may have been aware of Haldane's mapping function before 1919, and substantiates this view by quoting from the second-part of Muller's report:

In the case of higher percents of separation (long distances), the highest of three frequencies (let us call it  $AC$ ) falls short of the sum of the other two ( $AB + AC$ ), and so it is a smaller number than the distance representing it on the diagram, but nevertheless (within the normal limits of error) can be calculated from this diagram distance  $AC$ , for a constant relation was discoverable between this hypothetical distance and the actual frequency.

Since Muller's reviews focussed significantly on the work done by Morgan's group, it may be that this group actually laid the foundations for linkage estimation theory. Unfortunately, Haldane does not refer to Muller's work in his paper and Muller's doctoral thesis cannot be found. Perusal of Muller's thesis might have settled the issue of who first developed a genetic mapping function (Edwards, 1997, p. 11).

### **2.3.2 Genetic Markers**

All forms of genetic mapping use genetic markers, which are defined as, "mendelian [sic] characters which are sufficiently polymorphic to give a reasonable chance that a randomly selected person will be heterozygous" (Strachan & Read, 1996, p 317).

Five types of polymorphic DNA marker have contributed most to the development of the human genetic map: restriction fragment length polymorphisms (RFLPs) (Botstein et al., 1980), variable number tandem repeat (VNTR) minisatellite polymorphisms (Jeffreys et al., 1985; Nakamura et al., 1987), microsatellite polymorphisms (Beckmann, 1988; Weber & May, 1989) and single nucleotide polymorphisms (SNPs) (Kwok & Gu, 1999). All DNA polymorphisms can be associated with genes or can be anonymous DNA fragments (D-segments) without gene association.

RFLPs and minisatellite VNTRs are now seldom used for genetic mapping. RFLPs have only two alleles (the restriction site is present or is not) and so lack informativeness. Minisatellite VNTRs have preferential sub-telomeric location (de Lange et al., 1990), which excludes large parts of the genome from analysis. Also, since minisatellite size ranges from 0.5 kb to 40 kb (Armour, 1996, p. 172), large alleles will not amplify under standard PCR conditions (Saiki et al., 1990).

For these reasons, microsatellites became the standard tool for developing the genetic map of the human genome. They are evenly distributed throughout the human genome, are highly polymorphic and arrays of the most commonly used sequence for genetic studies,  $(CA/GT)_n$  where  $n$  is greater than 12, are estimated to occur approximately 35,000 times or once every 100 kilobase (kb) (Weber, 1990).

More recently, a world-wide effort has commenced to discover several hundred thousand SNPs throughout the human genome, each a variation of one nucleotide between the DNA sequence of individuals (Collins et al., 1997). SNP databases have been developed (Smigielski et al., 2000; <http://www.ncbi.nlm.nih.gov/SNP>), in anticipation that the identification of these polymorphic sites will revolutionise molecular diagnosis, accelerate the discovery of genes contributing to disease and contribute to pharmacogenomics (Pfoest et al., 2000).

### **2.3.3 Polymerase Chain Reaction**

PCR amplifies short sequences of DNA, *in vitro*, from template DNA between 100 base pairs (bp) to 6 kb in length (Saiki et al., 1990). It amplifies template DNA approximately one billion times (Mullis, et al., 1986). Standard PCR

amplifies picograms ( $\mu\text{g}$ ) of template DNA ( $10^{-12}$  g) to produce micrograms ( $\mu\text{g}$ ) of DNA product ( $10^{-6}$  g), which can be routinely sized by agarose gel electrophoresis.

The first step of PCR is the denaturation of template DNA by heating to about  $94^{\circ}\text{C}$  for approximately 3 minutes on the first cycle, 30-60 seconds on subsequent cycles. After the initial denaturation, the reaction is cooled, for example, to  $50^{\circ}\text{C}$ - $65^{\circ}\text{C}$  for 1 minute and the primers, present in high concentration, anneal to their complementary sequence at the 3' ends of each template DNA strand. Optimum annealing temperatures vary for different primers and template DNA.

Primers are chemically synthesised oligonucleotides, usually 20-25 nucleotides long with a free hydroxyl group present on their 3' ends. They are designed to target specific regions in the genome through the laws of probability. If a primer has a random sequence of 20 nucleotides, the number of possible sequences for a primer of that length is  $(1/4)^{20}$ , given that any one of four nucleotides can occupy each site. This means that the chance of any 20 nucleotide sequence appearing in a  $3 \times 10^9$  bp double stranded length of DNA is:  $2 \times (1/4)^{20} \times (3 \times 10^9)$ , making the chance that it will appear twice equal to:  $2 \times (1/4)^{20} \times (1/4)^{20} \times (3 \times 10^9)$ , or approximately  $5 \times 10^{-15}$ . Based on these calculations, each primer is unique in the genome by a factor of  $10^{15}$ - $10^{16}$ . In reality, the bulk of the human genome comprises DNA of varying degrees of repetitivity, and only 10% of it is single copy sequence (Cooper, 1995a, p. 1). For primers targeting single copy sequence, the uniqueness of each primer becomes  $10^{14}$ - $10^{15}$ . Notwithstanding these impressive statistics, false priming can still occur.

PCR reactions are then heated from their primer annealing temperatures to 72°C for a length of time which is dependent on amplicon size. Nucleotides are present in the reaction mixture. This allows template DNA amplification to occur. Nucleotides attach to the 3' hydroxyl group of the primer and extend along each template DNA strand by adding complementary nucleotides, and so make each strand of the template double stranded. DNA synthesis proceeds in a 5' to 3' direction while template DNA strands are 'read' in a 3' to 5' direction (Mullis et al., 1986; Saiki et al., 1989).

Amplification is achieved using thermostable DNA polymerases that withstand high temperatures. These are isolated and purified from organisms adapted to hot environments, such as the bacterium *Thermus aquaticus*. The optimum amplification temperature for thermostable PCR DNA polymerases is 72°C, and they can withstand the repeated 94°C DNA denaturation cycles which follow each amplification cycle (Hayashi, 1994).

After the initial denaturation, the cycle, 'primer annealing -> DNA amplification -> denaturation' is repeated, frequently up to 40 times. Template DNA is continuously denatured and amplified, with the new product acting as a template for the next. Unless the template DNA is exactly the same size as the product, as happens when a gel-purified PCR product is re-amplified, DNA polymerase will extend past the other primer but in an arithmetic fashion. In contrast, copies of template DNA between the primers increase exponentially and most PCR DNA-product is of this defined length. PCR DNA polymerases synthesise DNA at a rate of approximately 1000 bases per minute, achieving a billion fold increase in template DNA between the primers after 30-35 PCR cycles (Saiki, 1990).



#### **2.3.4 Unity of Gene Mapping**

All methods of human disease gene localisation share the following similarities. All depend upon affected individuals' posterior probabilities (probability after pedigree analysis) of being identical by descent (IBD) at genetic markers exceeding the prior probabilities (probability before pedigree analysis) of the same genetic markers being IBD. All rely upon the decline of IBD segments between individuals as genetic relationship diminishes. All are based on models, explicit in the case of parametric methods and implicit in nonparametric methods. All establish linkage between markers and disease genes by identifying a series of recombination fractions equal to zero for the segregation of a disease gene and a group of markers. These similarities suggest an underlying unity to human gene mapping. This unity was described eleven years before HGP got underway, then apparently overlooked (Thompson, 1997, p. 139).

#### **2.3.5 Connecting Meioses ( $m$ )**

The basic unity of genetic mapping was first emphasised by Edwards (1980) who defined genealogy as the number of meioses ( $m$ ) connecting individuals. Defining genealogy in this way, placed regular families, isolated pairs of relatives, populations of individuals related in unknown ways and even whole species (huge numbers of families of 1 connected by unknown numbers of meioses) all on the same  $m$ -continuum. It also linked genealogy to genetic distance. Since  $m$  determines IBD marker prior probabilities, and all disease gene mapping depends upon comparing IBD marker prior probabilities to IBD marker posterior probabilities in affected individuals,  $m$  is the fundamental unit of linkage analysis.

Edwards (1997, p. 34) cites Robbins (1918) as having demonstrated that allelic association decay was given by  $(1-t)^n$ , where  $t$  was the recombination fraction (now given the symbol  $\theta$ ) and  $n$  was the number of generations. Edwards (1980) used Robbins' allelic association decay formula to determine that 1 cM DNA lengths retain approximately 50% continuity after 70 generations or after about 2100 years if generation time was 30 years. Defining haplotype half-life as the number of segregations required to make prior haplotype probability approximately 0.5, the haplotype half-life for 1 cM DNA was about 70 generations for  $t$  (or  $\theta$ ) = 0.01. This was obtained by substituting these values in Robbins allelic association decay formula:

$$(1 - 0.01)^{70} = 0.99^{70} = 0.495$$

van der Meulen and te Meerman (1997) provided theoretical support for Edwards' analysis. Assuming a population growth of 5.5% per generation and a quasi-exponential distribution for the number of offspring, the probability distribution for the size of the area where no recombination occurs, could be calculated from:

$$N \times (1-t)^{N-1}$$

for  $N$  meioses and a probability of recombination,  $t$ . When the expected size of this area and its variance are calculated by partial integration, carriers of human disease genes that appeared 60 generations ago (approximately 1500 years) share, on average, about 5cM DNA with a standard deviation of 8 cM.

Ten years after Edwards' discussion of allelic association in man, Bishop and Williamson (1990) examined the conditional probability of gene IBD:

$P$  (relatives IBD at disease locus| relatives both affected)

They considered this as the conditional probability of gene IBD because all linkage detection methods are based on the premise that marker loci surrounding loci causing or contributing to a disease, are likely to be genotypically similar in affected relatives. Bishop and Williamson (1990) showed that the probability of gene identity between single loci surrounding disease loci, given two affected relatives, depended upon genealogical relationship, disease allele frequency and the risk-ratio of the disease susceptibility allele as calculated from penetrance probability. The overall result of including more parameters in their calculations was the same as Edwards' (1980) earlier conclusions. The greater the genealogical distance to a common ancestor, the greater the number of meiotic steps, the smaller the conserved ancestral segments.

### 2.3.6 Likelihood

One statistic dominates all methods of gene mapping: likelihood (Fisher, 1912, 1922a). In genetic mapping, probability refers to events yet to be observed, likelihood refers to events that have already been observed. In 1912, Fisher introduced the concept of likelihood and then developed it into the method of maximum likelihood (Fisher, 1922a). Likelihood is the probability of the occurrence of results *had* this probability been calculated *before* the observations were made. In genetics, the value of likelihood depends upon the assumed value of the recombination fraction, so likelihood is a function of  $\theta$  and can be plotted against it. Fisher's method of maximum likelihood was based on his observations that the likelihood function (or graph) contained within itself all the relevant information provided by the observations, and that the value of  $\theta$  at which likelihood attained its maximum value was a good estimate of

the actual value of  $\theta$ . Because likelihoods can reach very high values and are difficult to represent on a graph, Fisher chose to plot the natural logarithm of the likelihood as a function of  $\theta$  (Fisher, 1922a).

Fisher (1922b) tested his maximum likelihood method by calculating maximum likelihood estimates for eight X-linked *Drosophila* loci whose order was assumed known. The possibility of multiple crossing-over was ignored, making Fisher's estimates approximate. Edwards (1989) derived the exact maximum likelihood solution for Fisher's work, confirming the power of the original calculations.

There are two types of likelihood calculations. First, when it is possible to specify parameters of disease inheritance, it is also possible to predict from pedigrees the prior probability that two loci should be linked. The widely accepted prior probability that any two loci are linked on the human genome is 1 in 50, giving a prior probability that they are not linked of 49 in 50 (Strachan and Read, 1996. p. 320). Bayes' theorem provides a general method for combining probabilities into joint probabilities, allowing the determination of posterior probabilities that indicate linkage, conditional on the above accepted prior probabilities (Nicholls & Stark, 1971). A formal statement of Bayes' theorem is:

$$P(H_i|E) = P(H_i) P(E|H_i) / \sum P(H_j) P(E|H_j)$$

where  $P(H_i)$  means the probability of the  $i$ th hypothesis and  $P(E|H_i)$  the probability of the evidence  $E$ , given the hypothesis  $H_i$  (Krawczak & Schmidtke, 1994, pp. 66-68). In this case, calculation of linkage threshold yields

a joint probability of linkage (the overall odds in favour of linkage) as the product of prior and posterior linkage probabilities. Posterior probabilities are empirically determined,  $E$ , in this sampling design.

Calculating acceptable posterior probabilities from Bayes' theorem for both linkage and non-linkage requires specification of an acceptable level of significance for joint probability of linkage. The conventional  $p=0.05$  threshold of significance, or odds of 20:1, for joint probability of linkage versus non-linkage, is universally accepted. Therefore, calculating posterior probabilities from Bayes' theorem for linkage and non-linkage is a simple matter of substituting the appropriate linkage and non-linkage prior probabilities (1/50 and 49/50) and the linkage and non-linkage joint probabilities (20 and 1, as odds). To achieve an overall linkage threshold of significance  $p=0.05$ , conditional upon the stated prior probabilities, the minimum required posterior probability odds that emerge are, 1000:1/linkage:non-linkage. These are the odds against which genome testing data is measured (Blank et al., 1988).

Second, when it is not possible to specify an explicit disease model, the posterior probability of shared genome at the marker loci is compared to the prior probability. The probability of obtaining the observed posterior probabilities, had the prior probabilities been correct, is a measure of relative likelihood since it is derived from the ratio of relative posterior and prior probabilities (Smith & Stephens, 1997). Since this a 'model-free' analysis, prior probabilities are estimated from what would be expected by random segregation. Absolute probabilities do not change likelihood estimates because they are obtained by adjusting the sum of relative posterior and prior probabilities to equal 1. For example, the ratio 3:1 could become become 3/4:1/4. This logic underlies the statistics of all nonparametric linkage detection methods.

### 2.3.7 Two-point Lod Score

The theory of detection and estimation of linkage was revolutionised by Morton's lod score (Morton, 1955). It is based on the Bayesian calculations described above and begins by determining  $\theta_{\max}$ , the value of  $\theta$  at which the likelihood of obtaining the data,  $L$ , has its maximum value, or equivalently, at which  $dL/d\theta = 0$ . This maximum likelihood is expressed as  $L(\text{data}|\theta = \theta_{\max})$ . Maximum likelihood analysis next calculates the odds in favour of linkage versus non-linkage by computing the likelihood ratio:

$$\text{Likelihood ratio } (\theta) = L(\text{data}|\theta = \theta_{\max})/L(\text{data}|\theta = 1/2)$$

where  $L(\text{data}|\theta = 1/2)$  is the likelihood of obtaining the data when  $\theta = 1/2$  or, equivalently, when the disease gene and a marker are unlinked. This ratio gives the odds in favour of linkage versus non-linkage with a recombination fraction  $\theta_{\max}$ . After Morton (1955), likelihood ratios are reported as lod scores,  $Z$ , the decadic logarithm of:

$$L(\text{data}|\theta = \theta_{\max})/L(\text{data}|\theta = 1/2)$$

Critical thresholds for a single test between the disease locus and a marker, a two-point lod score, are  $Z = 3.0$  and  $Z = -2.0$ .  $Z = 3.0$  corresponds (approximately) to 1000:1 odds that the disease gene and a marker are linked, setting  $p=0.05$ . Linkage can be rejected if  $Z < -2$  and values of  $Z$  between  $-2$  and  $+3$  are inconclusive. Specifying parameters for linkage analysis is complicated, and the extraction of linkage information from human pedigrees is entirely dependent upon computer programs such as LIPED (Ott, 1974), LINKAGE (Lathrop *et al.*, 1985) and ILINK and MLINK (Schaffer, 1996).

### 2.3.8 Multipoint Lod Score

Almost since its inception, classical genetics has used three-point crosses to order linked loci on chromosomes (Sturtevant, 1913). It is based on the established premise that the rarest recombinant class is the one that requires the most recombinations. For  $n$  linked markers,  $(n-1)$  crossovers always produces the rarest recombinant class which is, therefore, two crossovers for a three-point cross. Chromosomal order can be deduced from such linked multilocus/multipoint inheritance patterns. An advantage of multipoint linkage analysis is that it overcomes limited marker informativeness. Highly informative microsatellite markers have reduced this problem, but it still persists in pedigrees lacking optimum structure.

Like two-point lod scores, multipoint lod scores are totally dependent on computer programs like VITESSE (O'Connell & Weeks, 1995) and LINKMAP (Schaffer, 1996) that move the unmapped disease locus across the linkage region from one marker location to the next, calculating the likelihood of obtaining the pedigree data, its lod, at and between each marker had the disease gene actually been at each point of calculation. Plotting likelihood/multipoint lod as a function of map location, produces a curve whose highest peak marks the most likely position of a disease locus (Morton, 1995). Conversely, exclusion is indicated should a multipoint curve remain below a lod of -2 across a linkage region. The same two-point critical thresholds operate for multipoint lod analysis.

The EXCLUDE program (Edwards, 1987) turns negative linkage data into a diagram of remaining candidate locations. The value of EXCLUDE is that it progressively narrows the range of possible candidate locations during genomic testing.

### **2.3.9 Parametric Methods: Co-segregation Linkage Analysis**

Parametric co-segregation linkage analysis involves Mendelian segregation within pedigrees, establishes relationships between loci, specifies disease inheritance parameters and employs parametric statistics to compare prior and posterior probabilities of IBD markers. Disease inheritance parameters specified are: mode of inheritance; phase (the combination of alleles, or haplotypes, inherited from each parent); mutation rates; allele frequencies; extent of phenocopy (environmentally determined phenotypes mimicking a genetically determined trait); degrees of penetrance (when 'non-affecteds' are included); locus heterogeneity; allelic heterogeneity; variable expressivity; interactions for more than one contributory locus; and incomplete ascertainment.

The greater the degree of specification, the greater the power of parametric linkage analysis. For example, Grimm et al. (1994) showed that the male mutation rate was about ten times higher than the female mutation rate in families segregating DMD point mutations. In deletion DMD families the situation was different. DMD deletion mutation rate was about twice as high in females as it was in males. Consequently, the specification of mutation rate by sex (in this case) increased the power of linkage analysis, improving estimates of wild-type and disease allele frequencies.

Conversely, inability to specify any disease inheritance parameters reduces the power of parametric methods by reducing the ability to extract posterior probabilities from pedigrees. Inability to specify disease inheritance parameters stems from small human pedigree size (insufficient affected individuals), imperfect human pedigree structure (missing individuals) and the necessary retrospective analysis of all human pedigrees. All of these factors



confound the specification of disease parameters, but sometimes their effect can be minimised. For example, when the mode of inheritance of *one form* of severe childhood autosomal recessive muscular dystrophy (SCARMD) was uncertain, lod scores were calculated under pure autosomal recessive and co-dominant models before establishing linkage of autosomal recessive SCARMD to 17q12-q21.33 (Roberds et al., 1994). This form of SCARMD is also called LGMD2D.

Major susceptibility loci of oligogenic disease can be identified by co-segregation linkage analysis (Morton & Lio, 1997, p. 18). Heterogeneous and essentially non-Mendelian breast cancer was subjected to this approach. A subset of families with an unusually high background level of susceptibility was identified and, in these families, breast cancer followed reduced-penetrance Mendelian inheritance. Co-segregation linkage analysis mapped a major breast cancer susceptibility locus to 17q21 (Hall et al., 1990).

Mapping the breast cancer major susceptibility locus to 17q21 highlighted major difficulties confronting this type of complex segregation analysis. First, 1500 families were screened to identify the 4-5% of these families with high background level of susceptibility and seemingly reduced-penetrance Mendelian inheritance (Newman et al., 1988). Genetic versus environmental determinants, shared environment versus shared genes and bias of ascertainment all confounded heritability estimates and, therefore, the proposed co-segregation linkage analysis for these families. Maximum likelihood programs struggled to include the correct mix of parameter values and, until it succeeded, it was doubtful if the greatest overall likelihood for the observed data would emerge (Hall et al., 1990).

Ascertainment for breast cancer was difficult but not impossible. For other complex diseases like schizophrenia, ascertainment has proven to be the greatest stumbling block (Byerley, 1989). The oligogenic diseases most amenable to co-segregation linkage analysis are those whose major susceptibility locus reveals itself in a subset of affected families, segregating the disease in reduced-penetrance Mendelian mode.

### **2.3.10 Nonparametric Methods: Allele-Sharing**

Nonparametric linkage analysis, also called shared segment linkage analysis, is not entirely model free, assuming the linear arrangement of genes on chromosomes, the Mendelian segregation of loci (although precise mode of inheritance is not specified), and assuming that genes affect traits. Shared segment linkage analysis is similar to co-segregation analysis in that it focusses upon affected individuals in pedigrees, but it is different in that it localises both Mendelian and non-Mendelian conditions and never overtly specifies disease parameters.

Nonparametric linkage analysis must be adopted for the analysis of high prevalence, genetically influenced human diseases such as asthma, diabetes mellitus or psychiatric disorders. Modelling disease parameters for these diseases is difficult. Nonparametric linkage analysis compares the observed agreement of alleles at a marker locus with the agreement expected by chance. The degree of agreement at a marker locus can be measured either by the number of alleles identical by state (IBS) or the number of alleles IBD. IBD affected sib pair (ASP) (Suarez et al., 1978), and IBS affected pedigree member (APM) (Weeks & Lange, 1988, 1992) strategies and their extensions are the most commonly used IBD and IBS methods.

An ideal sample for ASP analysis consists of  $n$  nuclear families, each family comprising two affected sibs and their parents, where both parents possess a different heterozygous genotype at each tested marker locus. In this sample, sib IBD can be determined unambiguously and, for each marker tested, two sibs will have an IBD score of 2 (identical alleles) or 0 (non-identical alleles) with a probability of 0.25 and an IBD score of 1 (one identical allele) with a probability of 0.5. From this starting position, a range of different ASP statistical tests have been developed to distinguish between IBD distributions when disease and marker locus are unlinked ( $H_0$ ) or linked ( $H_1$ ). These methods include: IBD proportions, means or goodness-of-fit tests (Blackwelder & Elston, 1985); correction for incomplete information (Holmans, 1993); inclusion of unaffected sibs (Green & Montasser 1988); and for the inclusion of two linked marker loci (Knapp et al., 1994) or two unlinked marker loci (Cordell et al., 1995).

Todd et al., (1987, 1998) used ASP to show that over half of the inherited predisposition to insulin-dependent diabetes mellitus (IDDM) (type 1) was associated with chromosome 6p21 'high risk' HLA haplotypes. Genome-wide ASP scans for linkage of chromosome regions to type 1 diabetes in affected families have confirmed this locus and the 11p15 region, containing the insulin gene, as type 1 diabetes susceptibility loci (Davies et al., 1994; Todd & Farrall, 1996). ASP analysis has also located four additional but equivocal type 1 diabetes susceptibility loci at 11q13, 6q22, 6q27 and 2q33 (Todd & Farrall, 1996).

The APM method can be used with arbitrary pedigrees and it only requires the marker genotypes of the affected individuals within the family. The APM method tests whether IBS sharing between all affected relative pairs is significantly increased compared to that expected under the null hypothesis.

APM has an advantage over ASP methods because it uses information obtained from the entire pedigree, but because it does not uniquely determine IBD, it has lower power than ASP and other IBD methods (Davis et al., 1996). APM is usually combined with other genetic mapping methods as discussed in, 'Combined Mapping Strategies' (pp. 41-43).

### **2.3.11 Nonparametric Methods: Allelic Association**

The fourth type of segment sharing analysis is allelic association. This method of genetically mapping diseases examines allelic association in populations. Allelic association is defined as any significant association between specific alleles at two or more neighbouring loci (Strachan & Read, 1996, p 589). It makes no assumptions about disease inheritance and is therefore a nonparametric method of gene mapping. It is a method of analysis that only applies to diseases and populations where most affected people are descended from a common 'carrier' ancestor. Possible disease allele locations are identified by finding series of alleles in affected individuals, measured two at a time, at a frequency above that expected by random segregation.

Establishing relationships between alleles, and not loci, distinguishes allelic association from linkage analysis mapping. This distinction is important because allelic association can happen for reasons other than disease allele linkage; natural selection favouring a particular combination of alleles at two loci or population stratification where those with and those without a disease may, by history, be genetically different subsets of their population (Lander & Schork, 1994, pp. 2041-2042). For these reasons, allelic association analysis is not linkage analysis.

Allelic association studies can also be distinguished by the fact that linkage disequilibrium need not occur around a disease locus. Linkage disequilibrium (also sometimes called allelic association) is defined as the non-random association of alleles at linked loci (Strachan & Read, p. 593). Strong selection against disease phenotype, and/or a high gene mutation rate creating an allelic series of disease-causing mutations in a population, will prevent linkage disequilibrium. DMD is an example of a disease that does not show linkage disequilibrium for both of these reasons.

Allelic association is rarely detectable beyond 1 cM, even for major genes of larger effect (Morton & Lio, 1997, p. 20), therefore it is of uncertain value for genome scans (Lonjou et al., 1999, pp. 1625-1626). Its major application is to increase map resolution for Mendelian genes and oligogenes within candidate regions, and it remains the only practical method to locate polygenes (Morton, 1998, p. 690). The case-control design study is one allelic association strategy to localise Mendelian genes, oligogenes and polygenes. This method tests specific candidate alleles in two random samples of unrelated affecteds and unrelated controls, defining each group by current and antecedent health states.

Case-control design studies have repeatedly established an association between the *HLA-DR2* allele within the major histocompatibility complex (MHC) and multiple sclerosis (MS) in different populations (Haines et al., 1998). Repeating case-control studies in different populations is important because the associations they establish may arise as an artifact of migration and population admixture (Lander & Schork, 1994, pp. 2041-2042; Morton & Collins, 1998, p. 11389). The greatest difficulty confronting this type of association study is the risk that the study and control groups are drawn from genetically distinct populations.

A series of methods have been developed to to circumvent the 'choice of control' difficulty posed by population stratification. All of these use parental genotypes in place of non-related controls. By this approach, nuclear families with one or more affected child are collected and typed at the marker locus. The two parental alleles that have not been transmitted to the affected child are combined to form the marker genotype of the 'control individual'. The two principal methods based on this logic are the haplotype relative risk (HRR) test (Falk & Rubenstein, 1987) and the transmission disequilibrium test (TDT) (Spielman et al., 1993).

HRR operates on the basis that the distribution of marker genotypes obtained from the non-transmitted parental alleles, in families with a single affected child, is identical to the distribution of marker genotypes in the population (Falk & Rubenstein, 1987). This is the 'control', and marker frequency in probands and controls are compared and tested for significance with a  $\chi^2$  test.

TDT compares the alleles transmitted by heterozygous parents to their affected children with the non-transmitted alleles and permits simultaneous testing for linkage and association. The alleles not transmitted to the affected child(ren) create an internal 'family-based' control permitting a  $\chi^2$  test estimate of relative risk. Thomson (1995, pp. 492-493) has shown that the non-transmitted allele sample provides unbiased allele frequency estimates. TDT has been extended to the situation of multi-allele marker loci (Sham & Curtis, 1995; Bickeboller & Clerget-Darpoux, 1995) and to 'sib TDT' that uses marker data from unaffected sibs instead of from parents (Spielman & Ewens, 1998). Both HRR and TDT are applied when other forms of genetic mapping or other types of evidence have given some indication of disease gene location (see Section 2.3.13, pp. 41-43).

### 2.3.12 Resolving Power

Following Edwards (1980) and Bishop and Williamson (1990), Houwen et al. (1994) identified the 'number-of-meiotic-steps' encompassed by different sampling designs as: population allelic association studies, 34-100 meiotic steps; shared segment analysis in groups of families with common ancestor, 9-34 meiotic steps; homozygosity mapping first and second cousins, 6-9 meiotic steps; and sib pair studies, 2 meiotic steps. Effectively, family linkage and population association analyses were at opposite ends of a genetic mapping *m*-continuum.

Close genetic relationships, therefore, provide greater power over larger genetic distances because the probability of IBD decline is smaller at a disease locus. Linkage detection power is good but map resolving power is necessarily low (Thompson, 1997, pp. 137 & 139-140). When disease inheritance parameter specification is accurate, human genetic mapping by co-segregation analysis normally has a highest statistical resolution of 1 cM, which is equivalent to a 1% recombination probability (Strachan & Read, 1996, pp. 327-328).

Conversely, more distant genetic relationships provide more power over smaller genetic distance because the probability of IBD decline is greater at a disease locus. With increasing *m*, linkage detection power declines (smaller and smaller conserved sequences) as map resolving power increases (Thompson, 1997, p. 137). 1 cM is the genetic distance at which allelic association due to sharing of ancestral chromosomes starts to become noticeable (Morton & Lio, 1997, p. 20). Therefore, allelic association is only useful for gene mapping over distances of 1 cM and less.

### 2.3.13 Combined Mapping Strategies

All genetic mapping strategies aim to detect linkage. In the balance between robustness, power and resolution all have discreet advantages. Attempts to analyse pedigrees at  $\theta=1 \times 10^{-3}$  would be pointless (Boehnke, 1994) as would genome scan disequilibrium mapping at  $\theta=1 \times 10^{-1}$  (Thompson & Neel, 1997, p. 201). Different mapping strategies are exploited concurrently or sequentially to refine disease gene localisation.

Allelic association has proven useful to refine the location of major disease genes established by co-segregation analysis (Lonjou et al., 1999, p. 1621). The localisation of disease genes for Huntington disease (HD) (The Huntington's Disease Collaborative Research Group, 1993) and cystic fibrosis (Riordan, et al., 1989) were refined by the employment of allelic association analysis after co-segregation linkage analysis and prior to the positional cloning of these disease genes.

Conversely, linkage analysis can be used to legitimise an allelic association. Working with 98 MS families Haines et al. (1998) showed that three highly polymorphic markers, HLA-DR, D6S273 and TNF $\beta$ , in the MHC demonstrated strong genetic linkage (respective parametric lod scores of 4.60, 2.20 and 1.24). This confirmed specific association with the *HLA-DR2* allele and MS. The *HLA-DR2* allele, or a nearby disease allele in linkage disequilibrium with the *HLA-DR2* allele, is, therefore, one of the major susceptibility alleles for oligogenic MS.

Linkage analysis can also refine a candidate region defined by allelic association. Houwen et al. (1994) assigned the gene for benign recurrent intrahepatic cholestasis (BRIC) to chromosome 18q21 by searching for chromosome



segments shared by only four patients, two sibs and two unrelated individuals. BRIC is an autosomal recessive liver disease characterised by multiple episodes of cholestasis (bile obstructive jaundice) without progression to chronic liver disease. Such use of allelic association to localise private diseases of genetically isolated populations, is a powerful empirical method for mapping disease genes. Sinke et al. (1997) performed co-segregation linkage analysis of an expanded sample of 14 BRIC families, using 15 microsatellite markers from the 18q21 region. Obligate recombinants in two families placed the *BRIC* gene in a 7cM interval, between markers D18S69 and D18S64.

Sib pair analysis has difficulty distinguishing true susceptibility loci from false positives because it is a low resolution form of analysis (Houwen et al., 1994). It can seldom identify chromosomal regions small enough for the identification of disease genes. Population allelic associations studies have the potential to confirm 'sib pair' susceptibility loci and reduce the size of location, but are only justified if there is sufficient evidence from different linkage analyses (Todd & Farrall, 1996, pp. 1443-1445).

APM has been applied to linkage analysis of complex diseases after co-segregation linkage analysis on informative pedigrees has indicated possible locations of (major) susceptibility loci. It has been used to support the localisation of (major) susceptibility loci for late-onset familial Alzheimer disease (Schellenberg et al., 1993), bipolar disorder (Berrettini et al., 1997; Straub et al., 1994), pre-eclampsia (high blood pressure during pregnancy) and eclampsia (convulsions related to pre-eclampsia) (Harrison, et al., 1997), and schizophrenia (Wang et al., 1995).

HRR and TDT are both used to confirm genetic location or associations. HRR

has implicated the dopamine D4 receptor as a susceptibility gene in attention deficit hyperactivity disorder (ADHD) (Smalley et al., 1998; Swanson et al., 1998 ) and excluded dopamine *DRD2 -141 delta C* allele as a polymorphism associated with schizophrenia (Li et al., 1998). The dopamine system has long been suspected of aetiological involvement in ADHD and schizophrenia, thus focusing these HRR tests. TDT has: confirmed the *HLA-DR2* allele association with MS (Haines et al., 1998); identified increased transmission of allele 4 of D18S487 to affected children in type 1 diabetic families (Merriman et al., 1998); and verified a population association between IDDM (type 1) and the 'class 1' alleles of the region of a 5' flanking tandem-repeat DNA polymorphism adjacent to the insulin gene on chromosome 11p (Spielman et al., 1993).

#### **2.3.14 Genetic Map Limitations**

The methods developed for the analysis of complex traits have not been as successful as the parametric methods developed for the linkage analysis of Mendelian traits. Thompson (1997, p. 276) has suggested that the increasing availability of marker data, and the increasing focus on linkage detection by association studies among relatives, or at the population level, may improve the discovery of susceptibility loci. However this may not happen. High density genetic maps serve to increase haplotype half-life over the shorter genetic distances that they can resolve, because the chance of recombination simultaneously reduces with genetic distance reduction. Consequently, diminishing recombination means that higher and higher density genetic maps are not nearly as useful as often argued, requiring an unrealistic number of connecting meioses for recombination to occur and therefore add little relevant information (J.H. Edwards, personal communication, 1996).

High-resolution mapping is also constrained by the fact that small conserved segments are harder to find than larger conserved segments and require more markers for identification. Since the risk of false positives is related to the number of markers needed to identify conserved segments, higher resolution mapping carries a higher risk of false positives (Strachan & Read, 1996, p. 501). This is a manageable problem for Mendelian conditions which must map somewhere on the genetic map, but for association studies there may be no association to find, so the risk of false positives increases as the search continues. Correction has to be applied to the threshold of significance which is set at:

$$p = 0.05/n$$

where  $n$  is the number of independent associations checked (Grove & Andreasen, 1982, p. 7). With this correction, the detection ability of association studies is reduced to detecting only the strongest of associations.

There is also evidence that genetic mapping cannot unravel all evolutionary complexity, thereby limiting its resolving power. Disease haplotypes larger than predicted by Edwards (1980) and van der Meulen and te Meerman (1997) have been identified. Evidence of greater extended genomic sharing comes from the A455E cystic fibrosis mutation's 25 cM common affected haplotype in French Canadian and Dutch populations and the 621+1G-->T cystic fibrosis mutation's 14 cM common affected haplotype in French Canadian populations (de Vries et al., 1996). This extended genomic sharing may reflect too large a recombination fraction in the calculations of Edwards (1980) and van der Meulen and te Meerman (1997) or undetected consanguinity in empirical studies.

Alternatively, the creation and disappearance of haplotypes through repeated introduction (de Vries et al., 1996) or during haplotype-recombination-drift equilibria (van der Meulen & te Meerman, 1997), may be more complex than present models allow. Certainly, the allelic association found around the locus for HD paints a complex picture of linkage disequilibrium in one of the best characterised human chromosomal regions (MacDonald et al., 1991).

For any two loci, linkage disequilibrium is commonly estimated by the Yule coefficient ( $\Phi$ ) (Yule & Engledow, 1914 and Engledow & Yule, 1915 as cited in Edwards, 1997, pp. 9-10 and in Krawczak & Schmidtke, 1994, pp. 64-66). The Yule coefficient was defined as the absolute value:

$$\Phi = (p_{1,1} - p_{1,2}) / (p_{1,1} + p_{1,2} - 2 \cdot p_{1,1} \cdot p_{1,2})$$

where  $p_{1,1}$  and  $p_{1,2}$  denote the frequency of allele  $A_1$  on chromosomes bearing allele  $B_1$  and  $B_2$ , respectively. When  $\Phi$  equals 1, maximum possible linkage disequilibrium is indicated. Conversely, when  $\Phi$  equals 0, linkage disequilibrium is absent. However when  $\Phi$  is between 1 and 0, discontinuities can arise that are not fully understood.

In a region spanning 2500 kb around the HD locus, one locus D4S95 (defined by the probe L19ps11) showed almost perfect *TaqI* RFLP equilibrium with the HD locus ( $\Phi \sim 0$ ) despite being within 250kb of the HD locus. Yet the same D4S95 locus yielded  $\Phi = 0.35$  and  $\Phi = 0.56$  for *Acl* and *Mbol* RFLPs respectively and a more distant locus D4S98 (defined by probe BS731) had a *SacI* RFLP yielding  $\Phi = 0.46$  (MacDonald et al., 1991).

The lack of steady linkage disequilibrium gradients of allelic association at the HD locus do not necessarily invalidate the relationship between genetic distance and number of connecting meioses ( $m$ ). Probably they reflect a complex history, involving any one or all of several independent mutations, early recombination between disease chromosomes or possibly the origin of some marker polymorphisms after the appearance of disease mutations (Strachan & Read, 1996, p. 333).

The resolving power of genetic mapping is limited, but there is a way to overcome these limitations. Integration of the genetic and physical map of the human genome. Six years after the commencement of HGP, the importance of map integration was emphasised by Morton and Lio (1997, pp. 19-20):

The need for better maps becomes more critical as the most striking major loci are cloned. The remaining major genes are rarer and may have less obvious candidates. Complex inheritance adds the additional complications of low penetrance and high heterogeneity. Allelic association (linkage dis-equilibrium) can be a powerful tool in a sufficiently small region, but not in a larger region or with a genome scan unless the density of markers is high, since allelic association is rarely detectable beyond 1 cM even for major genes of larger effect, short history, and less heterogeneity. Positional cloning for disease genes under complex inheritance cannot succeed unless map quality is improved. Linkage cannot determine order within 1 cM except in enormous samples, and so integration of data from physical and radiation hybrid maps is essential.

## **2.4 PHYSICAL MAP OF THE HUMAN GENOME**

The primary goal of all mapping techniques is to identify the locations of genes. The genetic map of the human genome does not allow the isolation of genes. It gives the relative locations of detectable variations in DNA base sequence. Although this allows the positioning of polymorphisms associated with genes between flanking markers, it is insufficient information to clone individual genes.

Unlike the genetic map, the physical map specifies physical distances, not statistical distances between markers. HGP is developing a physical map of the human genome. At lower levels of resolution its unit of measurement is the megabase (Mb), at mid-levels the kb, and at the highest level the bp. The size of the autosomal genetic map has been estimated as 2809 cM in males and 4782 cM in females (Morton, 1991). Since the size of the haploid human genome is estimated at 3.2 billion nucleotides (3,200 Mb) (Bodmer, 1981), 1 cM on the sex-averaged genetic map is roughly equivalent to 1 Mb on the physical map. Table 2.1 (p. 48) shows methods of physical mapping. Three methods were employed and are reviewed.

### **2.4.1 Fluorescent *in situ* Hybridisation**

Genomic fragments can be physically ordered on a chromosome by Fluorescent *in situ* Hybridisation (FISH) (Montanaro et al., 1991; Korenberg et al., 1992). After hybridization of specific DNA probes to chromosomes, reporter molecules attached to these probes are detected by fluorescently labelled affinity molecules. Table 2.1 shows that FISH has been applied to metaphase, prometaphase and interphase chromosomes, artificially extended chromatin fibre and DNA with increasing resolution.

**Table 2.1*****Physical Mapping***

<b>Method</b>	<b>Highest Resolution</b>	<b>Authors</b>
Somatic cell panels	chromosome	Dubois & Naylor, 1992
Flow cytometry	chromosome	Bartholdi et al., 1987
Chromosome painting	sub-chromosome	Dauwerse et al., 1992
Chromosome staining	10 Mb	Buckton et al., 1976
Comparative mapping	10 Mb	O'Brien & Graves, 1991
Translocation and deletion hybrid panels	8 Mb	Abrams et al., 1995
Cytogenetic breakpoint mapping	2 Mb	Chumakov et al., 1992a
* Fluorescence in situ hybridization:		
* metaphase chromosomes	5 Mb	Ireland et al., 1995
prometaphase chromosomes	2 Mb	Narahara et al., 1997
interphase chromosomes	50 kb	Wilke et al., 1994
artificially extended chromatin fibre	<5 kb	Haaf & Ward, 1994
artificially extended DNA	<5 kb	van Ommen et al., 1995
Radiation hybrid panels	60 kb	Stanford TNG4 panel, 1997
* Clone contigs:		
clone fingerprinting:	clone size	
repetitive DNA fingerprinting		Bellanne-Chantelot et al., 1992
IRE-PCR hybridisation		Nelson et al., 1989
* STS content mapping		Doggett et al., 1995
chromosome walking:	clone size	
bidirectional hybridisation		Riley et al., 1990
whole clone-clone hybridisation		Nizetic et al., 1994
* Automated DNA cycle sequencing	1 bp	Sears et al., 1992
<p>* Reviewed in this Section, pp. 47-52.      * Reviewed in Section 2.5.5.1, 'Mutation Detection', pp. 65-66.</p>		

FISH was a significant advance on earlier  $^3\text{H}$  *in situ* hybridisation methods, which suffered from low intensity signal, and were difficult to distinguish from background noise. FISH probes can be PCR-amplified genomic STSs of >1.0 kb to avoid hybridization to isoforms or pseudogenes (Cayanis et al., 1998). Alternatively, STSs are used to isolate larger clones from genomic libraries for use in FISH (Callen et al., 1990). These larger probes, containing up to 40 kb of insert, are preferred because they intensify signal. STSs were proposed by Olson et al. (1989) as the common language of the genetic and physical map. Each STS is a single-copy, unique sequence of DNA between 100-250 bp. These sequences are easily amplified and unequivocally recognised by PCR. They are independent of probe used or mapping strategy employed, and have been adopted by human genome mapping groups as their universal language.

#### **2.4.2 Clone Contigs**

Partial digestion of purified human DNA with restriction endonucleases results in a series of overlapping fragments whose size is determined by the restriction enzyme cutting frequency. Small fragments (>10 kb) are separated by agarose electrophoresis, large fragments by pulse field gel electrophoresis (PFGE). To provide a framework for the sequencing of the human genome, series of overlapping DNA fragments have been cloned to form DNA libraries. Table 2.2 (p. 50) shows commonly used cloning vectors with insert sizes.

Cloned DNA may be genomic DNA (gDNA) or already part of a DNA library such as a chromosome specific-library prepared by flow cytometry (Cotter et al., 1989). Purified vectors are digested with the same restriction endonuclease and ligated to the overlapping human DNA fragments.



<b>Table 2.2 Cloning Vectors (Strachan &amp; Read, 1997, p. 101)</b>		
<b>Cloning Vector</b>	<b>Size of Insert</b>	<b>Authors</b>
Plasmid vectors	0-10 kb	Okayama & Berg, 1983
Bacteriophage $\lambda$ replacement vectors	9-23 kb	Sambrook et al., 1989
Cosmid vectors	30-44 kb	Bellanne-Chantelot, 1991
Bacteriophage P1	70-100 kb	Sternberg, 1992
PAC vectors	130-150 kb	Ioannou et al., 1994
BAC vectors	up to 300 kb	Shizuya et al., 1992
YAC vectors	0.2-2.0 Mb	Albertstein et al., 1990
PAC = P1-derived artificial chromosomes; BAC = bacterial artificial chromosome; YAC = yeast artificial chromosomes		

The final stage in the creation of clones for contig construction, is the transformation of the host cell with vectors carrying inserted foreign DNA. YACs are a commonly used eukaryotic vector. They consist of two arms that are propagated independently in *E. coli* as plasmids, pJS97 and pJS98 (Riley et al., 1992). Each contains an ampicillin resistance gene, an *E. coli* replicon to propagate as a plasmid, a yeast origin of replication, a yeast telomere and several restriction sites including one for *Cla*I at the end of the telomere. pJS97 also contains a yeast centromere and a pigment suppressor gene that suppresses the host strain's production of a red pigment. *Cla*I digestion linearises the plasmids and fragments the target DNA. The vector arms

therefore each have a *Cla*I tail at one end and a telomere at the other end. Ligation of the target DNA and the vector arms produces a YAC which can transform yeast cells. The presence of one pigment suppressor gene, or one YAC, per cell makes colonies pink, two or more make them white. Yeast colour change is used to monitor the number of YACs in a cell and their stability against deletions. Unstable colonies have red and pink sectors as some cells lose their YACs as they divide.

Cosmids are synthetic cloning vectors possessing desirable features of both plasmid and  $\lambda$  phage cloning vectors. The bacteriophage P1 plasmid vector comprises P1 lytic genes and a plasmid replicon (Sternberg, 1992). DNA is inserted between two vector arms and packaged into a P1 protein coat *in vitro*. Upon injection into a host, the recombinant P1 circularises and can be induced to replicate. Recombinant P1 exploits P1's large size of 110-115 kb linear DNA, replacing most of this with insert DNA. PACs combine some features of the F-factor (Hayes, 1953) into the P1 vector (Ioannou et al., 1994), enabling them to carry larger inserts.

The ability of F-factors to carry large foreign DNA fragments is also exploited in BAC vectors as are the two F-factor genes, *parA* and *parB*, which keep copy number down to 1 or 2 per cell (Shizuya et al., 1992). In this way, recombination between inserts is reduced, greatly assisting the maintenance of large intact eukaryotic DNA in bacterial cells. The frequent repetitive sequences found in eukaryotic DNA and the high copy number mean that most other bacterial vectors are plagued with structural instability of inserts. Low copy number does, however, mean low yield of recombinant DNA.

Building contigs by linking clones depends upon establishing overlap by

common fragment. There are two general ways to develop clone contigs; chromosome walking and clone fingerprinting (Table 2.1, p. 48). Only the clone fingerprinting STS content mapping has relevance to this project.

#### **2.4.2.1 Sequence Tagged Site Content Mapping**

In developing clone contigs, STS content mapping usually involves sequencing approximately 300 bases at the end of isolated clones. Candidate primer sequences are identified without GC-rich regions, primers synthesised and an STS is identified if only one region of total gDNA is PCR amplified.

The STS screening of chromosome 16 (Doggett et al., 1995) used YAC library pools based on a YAC library pooling scheme designed by Balding and Torney (1991). This scheme maximised the chance of detecting positive clones, false positives and false negatives. Primary pools of clones were screened for each STS then secondary pools of STS-positive primary pools were screened to identify overlapping clones containing the same STS. Total gDNA was used as a positive control for each STS identified.

550 Los Alamos chromosome 16-specific YACs were used in the STS screening of chromosome 16. Primary pools of YACs contained 36 chromosome 16-specific YACs, 10 secondary pools contained 18 of the primary pools of YAC such that each individual YAC was present in only 5 of these 10 secondary pools. By identifying STSs present on both YACs and cosmids, it became possible to close the gaps between the cosmid contigs which were, on average, 65 kb before this YAC STS screening project was undertaken (Doggett et al., 1995).

### **2.4.3 Multi-Level Physical Maps**

The integration of genetic maps with low-resolution and high-resolution physical maps has resulted in multi-level maps of the human genome accessible at different levels of resolution. Doggett et al., (1995) developed an integrated low-resolution mega-YAC with a high-resolution cosmid contig mini-YAC map. They measured the power of their integrated physical, genetic and cytogenetic map of human chromosome 16 as 'completeness at resolution', the fraction of the chromosome within a fixed distance of mapped points. At resolution of 200 kb it was 0.88 complete and at resolution of 500 kb it was 0.97 complete. The average density or spacing was 1 STS every 160 kb.

Multi-level physical maps are important for the cloning of human disease genes because they are less sensitive to the vagaries of individual maps, each map acting as a reference for others (Cohen et al., 1993). The more accurate the physical map, the easier it is to isolate transcribed sequences and their controlling elements. Multi-level physical maps are accurate physical maps of the human genome. They are available at the National Centre for Biotechnology Information (NCBI) site, under 'Hot Spot', Human Genome Resources (<http://www.ncbi.nlm.nih.gov/genome/guide/>).

## **2.5 CLONING HUMAN DISEASE GENES**

There are four general ways to clone human disease genes: functional cloning, position-independent candidate approach, positional cloning, and positional candidate approach. Only the last two methods depend upon the genetic and physical maps of the human genome, however confidence in functionally cloned genes and position independent candidate genes as disease genes is

substantially increased if they can be shown to map to the same chromosomal region as a disease gene. All four strategies conclude with the identification of mutations in cloned genes. All these strategies had to be considered for inclusion in this project. All are reviewed. Two were employed.

## **2.5.1 Functional Cloning**

Functional cloning requires knowledge of biochemical basis of an inherited disease or evidence from functional complementation assay about altered gene function.

### **2.5.1.1 Oligonucleotides and Antibodies**

Two approaches are possible when the biochemical basis of an inherited disease is known. The first involves determining part of the amino acid sequence of a defective/deficient protein from partially degenerate oligonucleotides corresponding to all possible codon permutations. Screening an appropriate complementary DNA (cDNA) library with labelled probe comprising a mix of all partially degenerate oligonucleotides enables correct cDNA clones to be isolated. These are then used to probe genomic libraries for gDNA clones, allowing the complete characterisation of the gene. The factor VIII disease gene (*F8*) responsible for haemophilia A was cloned and then characterised in this way (Gitschier et al., 1984).

The second approach involves raising specific antibodies against the normal protein product of a disease gene. These antibodies immuno-precipitate polysomes containing desired mRNA in *in vitro* protein synthesising systems. Converting such mRNA to cDNA allows specific cDNA clones to be isolated

and the gene characterised. This was how the phenylketonuria (PKU) disease gene, phenylalanine hydroxylase (*PAH*), was cloned. Low tyrosine and high phenylalanine plasma concentrations identified hepatic phenylalanine hydroxylase as the deficient enzyme in PKU patients (Levy, 1989). Antibodies raised against this enzyme led to the full characterisation of this gene and the discovery that PKU can be caused by at least 75 allelic variants of the phenylalanine hydroxylase (*PAH*) gene (Cooper & Krawczak, 1994, p. 52).

#### **2.5.1.2 Functional Complementation**

Functional complementation has been employed to identify human disease genes, particularly those involved in DNA repair. Fanconi anaemia (FA) is characterised by cellular hypersensitivity to DNA cross-linking agents. On this basis it had been tentatively classified as a DNA-repair disorder (Setlow, 1978). Sensitivity to the DNA cross-linking agents mitomycin C and diepoxybutane was exploited by Strathdee et al. (1992a) to identify three new FA complementation groups, FA(B), FA(C) and FA(D) from human/hamster somatic hybrid panels. A functional complementation method was employed to clone cDNAs that corrected the defect of group C cells (Strathdee et al. 1992b). The cDNA defective in this anaemia was cloned by transfecting the FA(C) cell line with a cDNA library inserted into a cDNA expression shuttle vector, pREP4. Selection with mitomycin C and diepoxybutane led to the identification of eight FA(C) candidates, three of which specifically complemented the FA(C) defect.

These cDNAs encoded alternatively processed transcripts of a new gene, *FACC*, which was mutated in group C patients. This *FACC* polypeptide did not contain any motifs common to other proteins and so represented a new

gene involved in the cellular response to DNA damage. The function of *FACC* led directly to its cloning.

#### **2.5.1.3 Muscular Dystrophy Disease Genes**

Functional cloning has only been useful in a few cases of muscular dystrophy, one of which was the cloning of the adhalin gene responsible for LGMD2D. It was the knowledge that mutated dystrophin-associated-sarcoglycan proteins caused forms of muscular dystrophy that led Roberds et al. (1994) to look for a mutated protein from this complex in the skeletal muscles of LGMD2D patients. A deficiency of the sarcoglycan complex protein adhalin (also called  $\alpha$ -sarcoglycan) was identified, the human adhalin gene mapped to chromosome 17q12-q21.33 and adhalin cDNA cloned and sequenced to show missense mutations segregating with LGMD2D (Roberds et al., 1994).

#### **2.5.2 Position-Independent Candidate Approach**

Molecular pathogenesis or phenotype similarity (to animals or humans) can suggest candidate genes for human disorders independent of knowledge about chromosomal locations of disease genes. Case-control allelic association studies use this approach to identify suitable candidate alleles for testing as oligogenes and polygenes. Disease genes segregating in a Mendelian fashion can also be identified in this way. Goebel et al., (1997) described three unrelated 'floppy' children with myopathic fibres containing large masses of thin myofilaments. The immunological detection of actin in these thin myofilament masses suggested that this was a discrete form of congenital myopathy, an 'actin myopathy' of either sporadic or hereditary type. Recently, mutations in the alpha actin gene have been identified in patients with this actin myopathy and

with another muscle disease, nemaline myopathy (Nowak, et al., 1999).

### 2.5.3 Positional Cloning

Positional cloning is the strategy employed when nothing is known about the altered function of a disease gene. It commences with one of five different genetic mapping methods to determine sub-chromosomal localisation of a disease gene: cytogenetic deletions or disrupted chromosomal regions; loss of heterozygosity (LOH) in somatic disease; co-segregation linkage analysis; shared segment linkage analysis; and allelic association between marker genotype and disease locus in a population. Often this first stage of positional cloning involves application of more than one of these methods.

The association of cytogenetic deletions and translocations (visible genetic markers) with disease has lead to the cloning of disease genes by positional cloning. Subtraction cloning of *test* DNA sequences corresponding to deleted Xp21 sequence in a DMD patient were used to identify dystrophin cDNA clones (Koenig et al., 1987). Similarly, cloning an Xp21 translocation breakpoint in a DMD patient also led to the isolation of the DMD locus (Worton & Thompson, 1988).

LOH allelic imbalance of markers adjacent to deleted disease genes has been used to identify chromosomal regions deleted in somatic disease. It has been most successful in identifying familial cancers such as oral cancers (Roz et al., 1996). However, even within its limited area of application, LOH disease gene localisation does have its limitations. For example, in the LOH identification of the retinoblastoma 13q14 LOH region (Cavenee et al., 1983; Friend et al., 1986; Lee et al., 1987), homozygous patients provided no information and



most tumour samples contained a mixture of tumour and non-tumour tissue, so band intensity of one marker allele was decreased rather than lost entirely. In addition, the retinoblastoma 13q14 deletion was not retinoblastoma-specific, also associating with osteosarcomas (Friend et al., 1987), breast cancer (Lee et al., 1988), and small-cell lung cancer (Yokota et al., 1987).

The cloning of fully penetrant, uniformly expressed, autosomal dominant and recessive X-linked disease genes usually commences with co-segregation analysis. For these diseases, mutation segregation can be deduced from clinical phenotypes and, unlike autosomal recessive diseases, linkage can be established in relatively small, non-consanguineous pedigrees. Disease genes of this type feature prominently in the human genome map (Bassett et al., 1997).

Shared-segment linkage analysis and allelic association analysis are often used as the first step in the positional cloning of autosomal recessive and non-Mendelian disease genes contributing to complex inherited diseases.

#### **2.5.3.1 Transcript Map of Candidate Regions**

Genetic mapping only gives a crude estimate of candidate region size because genetic distances between markers are uncertain. To identify disease gene candidates, an ordered contig of clones has to be built across the candidate region. The next step of positional cloning is to construct a transcript map of all relevant expressed sequences within candidate region genomic clone contigs.

Genes can be identified because they make RNA, their sequences are conserved, they have relatively long open reading frames (ORFs) and they are

often associated with CpG islands (Bird, 1986). Based on these characteristics, a variety of methods have been developed to clone genes in cloned DNA. Collins (1992, p. 6) and Cooper (1995b, p. 58) have identified three general approaches for the isolation of novel transcribed sequences: hybridization-based, function-based and sequence-based. The methods within each of these general approaches are identified in Table 2.3 (p. 60). Only one of these methods was used in this thesis.

#### **2.5.3.2 Expressed Sequence Tags**

Adams et al. (1995) reported the discovery of 58,384 non-overlapping human expressed sequence tags ESTs from the analysis of 300 poly-A selected cDNA libraries constructed from 37 distinct organs and tissues. Since the identification of such a large number of ESTs necessarily required large scale sequencing, cDNA library quality was carefully evaluated. To achieve high quality control, the sequences of between 100-200 clones from each library were assessed on seventeen counts. The libraries selected for large scale EST analysis were those with a broad diversity of transcripts, no evidence of DNA from another organism, no mitochondrial DNA, no ribosomal DNA, and a low percentage of clones without insert.

Sequencing was done from the 5' end of each cDNA clone rather than the 3' end since each 5' end was more likely to contain a protein coding sequence, enabling database searches to assign putative identifications. This was achieved for 11.6% of the identified ESTs. Three genes associated with colon cancer were identified in this EST dataset.

**Table 2.3**      ***Novel Genomic DNA Transcript Identification***  
*(from Collins, 1992, p. 6 and Cooper, 1995b, p. 58)*

Methods	References
<b><i>Hybridisation-based</i></b>	
Northern Blotting	Alwine et al., 1977.
Inter-species cross hybridisation ('zoo blotting')	Monaco et al., 1986
Identification of CpG islands	Bird, 1986; Brown & Bird, 1987
CpG Island Rescue PCR (IRP)	Valdes et al., 1994.
Direct cDNA library screening/sequencing	Hochgeschwender et al., 1989; Elvin et al., 1990; Geraghty et al., 1993; Adams et al., 1991; Soeda et al., 1995
Subtractive hybridisation	Duguid et al., 1988
Heteronuclear cDNA library screening	Whitmore et al., 1994
Homologous recombination	Vollrath et al., 1988; Kurnit & Seed, 1990
Direct cDNA selection using cosmids/YAC immobilised arrays	Lovett et al., 1991
Reverse transcript PCR	Cheng & Zhu, 1994
<b><i>Function-based</i></b>	
ORF identification by lac-z fusion protein expression	Gray et al., 1982
Exon amplification/trapping	Duyk et al., 1990; Buckler et al., 1991
Enhancer/poly A signal trapping	Weber et al., 1984; Rosenthal, 1987
Gene transfer and transcript identification	Strathdee et al., 1992b
Promoter trapping	von Melchner & Ruley, 1989; von Melchner et al., 1990, 1995
<b><i>Sequence-based</i></b>	
Exon prediction	Fearon et al., 1990
Differential display	Liang & Pardee, 1992
Interspecific comparison	Sargent et al., 1993
Database similarity searching	Gish & States, 1993
*Expressed sequence tag PCR mapping (EST-PCR)	Adams et al., 1993 (a & b), 1995; Pallavicini et al., 1997
*Relevant to this thesis and reviewed on pages 59-61.	

A preliminary EST transcript map of human skeletal muscle has also been provided (Pallavicini et al., 1997). This map required the design of 267 pairs of specific primers, corresponding to newly identified muscle genes. PCR amplification of the GeneBridge 4 radiation hybrid (RH) panel (Gyapay et al., 1996; <http://www-genome.wi.mit.edu/cig-bin/contig/rhmapper.pl>), gave a precise map position for 230 skeletal muscle ESTs.

Diverse clone collections/contigs from many physical mapping groups around the world are currently being typed by EST content. An international EST mapping consortium has been formed to develop a continuum of overlapping clones (Collins, 1998). More than 50,000 ESTs are currently being mapped to 0.1-0.5 Mb intervals, utilising the whole genome RH and YAC panel approaches. The efforts of the international EST mapping consortium will undoubtedly lead to the discovery of hundreds (perhaps thousands) of as yet unidentified disease genes associated with ESTs. Sequencing overlapping clones confirm the localisation of each transcript (Collins, 1998).

EST identification from cDNA libraries has some limitations. It can be biased towards highly expressed genes, fail to distinguish between pseudogenes, multigene family members and functional genes and, if poly A purified cDNA is used, ESTs in 5' exons can be poorly represented (Towner 1995, pp. 79-80). Promoter-tagged site (PTS) databases avoid these limitations by using recombinant retroviral gene trap vectors targeting promoter regions (von Melchner & Ruley, 1989). Recombinant retroviral gene trap vectors contain a reporter gene without a promoter or enhancer. Reporter gene activation therefore requires integration into expressed genes in the same transcriptional orientation relative to the gene, allowing selection for 5' retrovirus gene trap vector integration (von Melchner, et al., 1990; von Melchner & Ruley, 1995).

#### **2.5.4 Positional Candidate Approach**

Increasingly, combinations of functional, position-independent and positional cloning strategies are being adopted for disease gene isolation. This combined strategy was first termed the positional candidate approach by Ballabio (1993). It is distinguished from functional, position-independent and positional cloning because it identifies disease genes on the basis of both position and function.

In reality, most of the 84 positional cloning successes listed by Bassett et al. (1997) relied upon something other than pure positional cloning. For example, the large-scale deletions associated with LOH localisations suggested that likely disease gene candidates might be DNA repair enzymes. They were.

Similarly, it was the knowledge that three forms of autosomal recessive LGMD were associated with mutations in various sarcoglycans of the dystrophin-associated sarcoglycan complex, that led Nigro et al. (1996) to look for, and find, a sarcoglycan candidate in the 5q33-34, LGMD2F linkage region previously identified by Passos-Bueno et al. (1996). This gene turned out to contain a single base deletion. This single base deletion segregated with LGMD2F in AR mode (Nigro et al., 1996).

The current rapid localisation of human genes has greatly increased the density of the human transcript map and accelerated the movement of disease gene cloning towards the positional candidate approach. It has already assumed a dominant position and with the projected completion of the human genome sequence in 2003 (Collins, 1998), the positional candidate approach will increasingly dominate human disease gene cloning.

Collins overestimated the time required for the completion of the sequence of the human genome. The completion of 97% of the sequence of the human genome was announced in July 2000 by Craig Venter, the Chief Executive Officer of Celera Genomics, and Francis Collins, director of the US National Institutes of Health's National Human Genome Research Institute (and as such the unofficial head of the publicly funded HGP). The positional candidate approach will be one of many beneficiaries of human genome sequence data.

Gerald Rubin, the vice president for biomedical research at the Howard Hughes Medical Institute, describes the human genome sequence as, "written in a foreign language. It's a very complicated problem. It's going to be a long time coming" (cited by Lemonick, 2000). Even when the human DNA code is readable, the enormous problem of discovering what each protein does and how it interacts with other proteins and cell components will still remain. Progress will be slow, but the implications of this milestone towards genotypic medicine are significant.

As far as the positional candidate approach is concerned, the human genome sequence should accelerate the identification of susceptibility loci and the contributing polygenes (defined, p.15) within these loci. This is significant because polygenes are involved in the great majority of inherited human disease. Yet individually each polygene contributes little to the determination of a disease, cannot be localised by co-segregation linkage analysis and presents considerable difficulties for segment-sharing linkage analysis and association studies. Identifying polygenes is important not only for polygenic disease, but also for Mendelian and oligogenic disease operating over polygenic background. With the sequence of every human gene available, associations between diseases and alleles of candidate genes within susceptibility regions will become easier to establish in families, and in the general population.

This approach has already demonstrated its usefulness. An association between the epsilon 4 allele of the apolipoprotein E gene (*APOE*) and familial late-onset Alzheimer's disease was demonstrated in 15 Swedish late-onset Alzheimer's families ( $p = 0.01$ ), despite the absence of AD linkage to the *APOE* region in these families (two-point linkage analysis), and only weak evidence of linkage from APM (Liu *et al.*, 1996).

#### **2.5.5 Mutation Detection**

The final step of all human disease gene cloning strategies requires the identification of the mutations contributing to each disease. Mutations can be translocations, deletions, duplications, repeat expansions, insertions, splice site, missense, frameshift or stop codon mutations. They can affect transcribed and non-transcribed portions of the genome.

To detect gross alterations or rearrangements of genes, transcripts are screened for mutations by Southern blotting (Southern, 1975), which is often combined with PFGE (Fountain *et al.*, 1989), and Northern blotting (Alwine *et al.*, 1977). Deletions that include markers adjacent to the disease gene can be recognised from marker segregation in pedigrees for both autosomal dominant and recessive diseases. PCR methods detect smaller mutations in cloned DNA lying inside or outside coding regions. Commonly used mutation detection methods are listed in Table 2.4 (p. 65).

Three of these were relevant to this thesis.

**Table 2.4                      Common PCR Mutation Detection Methods**

<b>Method</b>	<b>Authors</b>
*Direct Sequencing of Transcripts	Sanger, 1981.
Denaturing Gradient Gel Electrophoresis	Fischer & Lerman, 1983.
RNase Cleavage	Myers et al., 1985.
Deletion screening ('multiplex PCR')	Chamberlain et al., 1988.
Chemical Mismatch Cleavage	Cotton et al., 1988.
*DNA Single-strand Conformation Polymorphism (SSCP)	Orita et al., 1989.
RNA SSCP (rSSCP)	Sarkar et al., 1992a.
Dideoxy Fingerprinting	Sarkar et al., 1992b.
*Repeat Expansion Detection (RED)	Schalling et al., 1993.
Restriction Endonuclease Fingerprinting (REF)	Liu & Sommer, 1995.
Heteroduplex Cleavage with Bacterial Resolvases	Youil et al., 1995.
*Relevant to this thesis and reviewed on pages 65-69	

#### **2.5.5.1      Direct Sequencing of Transcripts**

Automated DNA sequencing based on the dideoxy sequencing method (Sanger, 1981) has now largely replaced the traditional radioisotope versions of this sequencing method. Each sequencing reaction contains: the four types of deoxynucleoside triphosphates (dNTPs) deoxyadenosine triphosphate (dATP), deoxythymidine triphosphate (dTTP), deoxyguanosine triphosphate



dGTP, deoxycytidine triphosphate (dCTP) each labelled with a different fluorophore; one primer; magnesium; a buffer; and a thermostable DNA polymerase. Chain termination occurs when ddNTPs are incorporated because they lack a hydroxyl group at the 3' carbon atom of the ribose sugar. The concentration of ddNTPs is much lower than that of the dNTPs, consequently chain termination occurs randomly. Cycle sequencing (Sears et al., 1992) is a PCR type of sequencing reaction that produces good results for double stranded DNA (dsDNA). During electrophoresis of sequencing product, a monitor records the four different fluorophore signals as a chromatographic intensity profile with ATGC translation.

#### **2.5.5.2 DNA Single-Strand Conformation Polymorphism**

Size and nucleotide sequence determine the folding structure(s) of single-stranded DNA molecules. In conjunction with surface charge density, these parameters contribute to electrophoretic mobility and form the basis of a mutation detection method called single-strand conformation polymorphism (SSCP) analysis (Orita et al., 1989). PCR-amplified ssDNA fragments up to 200 bases in length (Sheffield et al., 1993) are denatured and snap-chilled.

These DNA molecules are separated on a non-denaturing polyacrylamide gel under constant voltage and controlled ionic strength and temperature to reduce conformation changes in polymorphisms. Label may be incorporated during PCR amplification (Orita et al., 1989) or polymorphisms causing a mobility shift can be detected by ethidium bromide (Hayashi, 1991) or by silver staining acrylamide gels (Hoshino et al., 1992). If the underlying reason for an SSCP is not known, bands displaying mobility-shift are cut out, re-amplified by PCR, and sequenced to detect nucleotide sequence differences between individuals.

### 2.5.5.3 Repeat Expansion Detection

Allowing for cyclical permutations, DNA can only have 10 discrete trinucleotide sequences. Repeats of all these trinucleotide sequences occur throughout the human genome, the majority without pathological effect (Fu et al., 1991). There are two ways in which trinucleotide repeats are known to cause disease. First, short meiotically stable repeats have been shown to cause both AD- and AR-OPMD. Expansion of the normal (GCG)<sub>6</sub> repeat to a (GCG)<sub>9</sub> repeat in exon 1 of the poly(A) binding protein 2 gene (*PABP2*) causes AD-OPMD (Brais et al., 1998). The pathological association of this type of repeat expansion can only be established when detected during the sequencing of candidate genes.

The second type of disease-causing repeat expansion are the meiotically unstable repeat expansions involving much larger expansions. Some trinucleotide repeats become extremely unstable above a certain threshold length. Their expansion beyond this 'upper maximum limit' causes disease. These unstable repeats are seldom transmitted unchanged from parent to child. Their tendency is to increase rather than to decrease repeat copy number. For example, the expansion:contraction odds ratio for the unstable CGG expansion in the *FMR1* fragile X gene, has been calculated as 64:3 (Fu et al., 1991). These 'dynamic mutations' can cause anticipation (Richards et al., 1992).

Anticipation describes the tendency of some dominant conditions to become more severe in successive generations, often lowering the age at which clinical symptoms onset. It is a type of variable expression. The discovery that repeat length correlated with severity for many trinucleotide diseases provided a powerful diagnostic and prognostic tool for these diseases, including the nine CAG repeat expansion neurodegenerative disorders: HD; dentatorubral

pallidolusian atrophy (DRPLA); spinal and bulbar muscular atrophy (SBMA); dentatorubral-pallidolusian atrophy (DRPLA); and spinocerebellar ataxia (SCA) types 1, 2, 3, 6 and 7 (Koshy & Zoghbi, 1997).

The repeat expansion detection (RED) technique (Shalling et al., 1993) can detect trinucleotide repeats throughout the human genome, including pathological trinucleotide repeat expansions. It does not require knowledge of affected gene location. Trinucleotide repeat expansions can cause disease within coding regions, within transcribed but not translated regions and by positional effects on nearby genes. HD is an example of a coding region trinucleotide repeat expansion (The Huntington's Disease Collaborative Research Group, 1993); myotonic dystrophy, an example of a 3' untranslated region (UTR) trinucleotide repeat expansion (although this expansion may not only be affecting the *DMPK* gene in whose 3' UTR it is situated) (Boucher et al., 1995); and the fragile-X syndromes FRAXA (Pintado et al., 1995) and FRAXE (Knight et al., 1993), are examples of trinucleotide repeat expansions having positional effects and influencing adjacent genes.

The RED technique works by hybridising trinucleotide repeat oligonucleotides to ssDNA. Trinucleotide repeat oligonucleotides hybridised to genomic trinucleotide repeats are ligated to neighbouring oligonucleotides, producing a family of oligonucleotides representative of the trinucleotide repeats present in the genomic DNA. Trinucleotide repeat size is determined by electrophoresis on a 6% acrylamide gel. Potentially pathological trinucleotide repeat expansions (of any motif) are detected by comparison of affected and unaffected pedigree members. Assuming full penetrance, pathological trinucleotide repeat expansions would be present in the former but not the latter for all individuals over age of onset.

## 2.6 CONCLUSION

Research on mapping and identifying human disease genes interacts with the recent thrust of the HGP: the isolation of human gDNA and cDNA clones for the construction of high-resolution physical maps. The highest resolution physical map, the complete nucleotide sequence of the 3000 Mb nuclear human genome, is now near completion.

This chapter focused on those principles and strategies behind the identification of human disease genes that were employed in this thesis. It showed how this research contributes to the HGP, through its contribution to the development of genetic and physical maps of the human genome. It also showed that the final identification of disease genes depends upon the physical map of the human genome, a map derived from human cDNA and gDNA clones that, predominantly, do not have disease association. This is a timely reminder that the term 'human disease gene', is a term derived from the contribution inherited disease has made to the identification of human genes. A term, according to Mat Ridley (2000, pp. 54-55), with severe limitations:

Open any catalogue of the human genome and you will be confronted not with a list of human potentialities, but with a list of diseases, mostly named after pairs of obscure central-European doctors. This gene causes Niemann-Pick disease; that one causes Wolf-Hirschhorn syndrome. The impression given is that genes are there to cause diseases. . . . Yet to define genes by the diseases they cause is about as absurd as defining organs of the body by the diseases they get: livers are there to cause cirrhosis, hearts to have heart attacks and brains to have strokes. It is a

measure, not of our knowledge but of our ignorance that this is the way the genome catalogues read. It is literally true that the only thing we know about some genes is that their malfunction causes a particular disease. This is a pitifully small thing to know about genes, and a terribly misleading one.

All of this is true. Genes are positive forces. It is mutations that cause disease, not genes. Nevertheless, 'this pitifully small thing to know about genes' is what this thesis seeks to know, because it is the starting point for far more important knowledge. A full understanding of normal gene function. Sometime in the future, all genes and their products may be named by what they do, and diseases identified by mutation. But for the present, the term 'human disease gene' accurately describes what we know, and what we do not know, about the majority of human genes.

## **CHAPTER 3**

# **MUSCULAR DYSTROPHY: A CANDIDATE APPROACH**

### **3.0 MUSCULAR DYSTROPHY: A CANDIDATE APPROACH**

In the early 1990s, the molecular neurogenetic team at ANRI took the view that any of the contractile proteins could be disease genes for forms of muscular dystrophy. They could easily be mutated in muscular dystrophies not yet localised in the genome, or in localised muscular dystrophies whose precise genes had not been identified. Of particular interest were the seven contractile muscle proteins which account for 20% of all mRNA in mature muscle fibres (Paterson and Bishop, 1977). The high level of muscle-specific expression of these genes, and their central contractile role, made them promising candidate genes for muscular dystrophy. The task was to place them on the physical map of the human genome as 'candidates in waiting', or as candidates for co-localised disease genes of interest to the molecular neurogenetic team at ANRI. Predominantly, these were the disease genes for forms of nemaline myopathy and distal myopathy.

This approach had elements of the position independent candidate approach (p. 56-57) and the positional candidate approach (pp. 62-64). Like the position independent candidate approach, it selected for study genes of perceived possible relevance to muscular dystrophy, and without knowledge of their location. But then it localised these genes on the human physical map, in this case the cytogenetic map. This is, of course, the first crucial step of the positional candidate approach. Genes must be localised before they can become candidates in identified linkage regions. So the approach adopted by the ANRI molecular neurogenetic group was a candidate approach. A unique candidate approach in that it did not quite fit the definition of the position independent candidate approach or the positional candidate approach, yet it had some of the essentials of each. With hindsight, this was not surprising. It was developed

and adopted when 'Reverse Genetics' was first being renamed 'Positional Cloning' (Collins, 1992), and before the terms 'position independent candidate approach' and 'positional candidate approach' had been coined.

The work reported in this chapter was done as a part of this team's effort to identify muscular dystrophy disease genes. Of special interest to this thesis was the possibility that this work might uncover contractile candidate gene(s) for a Western Australian form of distal myopathy discussed in Chapter 4.

Those involved in this research are identified in the three publications produced by this work (see Appendix A). The contribution made in this thesis to the cytogenetic localisation of these skeletal muscle contractile genes was the isolation of genomic clones for the slow skeletal muscle troponin gene (*TNNI1*), the human skeletal muscle alpha actin gene (*ACTA1*) and the  $\beta$ -tropomyosin gene (*TPM2*). This work was done in collaboration with Dr. Anthony Akkari and Mr. Clive Hunt, both of ANRI. Positive clones were then assayed and purified by Dr. Anthony Akkari (*TNNI1* and *ACTA1*) and Mr. Clive Hunt (*TPM2*) for use in the *in situ* cytogenetic mapping of these skeletal muscle contractile protein genes.

### **3.1 GENOMIC LIBRARIES**

Genomic clones were isolated by amplifying genomic libraries in suitable bacterial host strains followed by plating, lifting, probing, picking and re-amplification. *TNNI1* and *ACTA1* positive clones were isolated from the  $\lambda$ Gem 11 (Promega) genomic library (a gift from Dr. L. Abrahams). This was available at ANRI. *TPM2* positive clones were isolated from a  $\lambda$ EMBL3 SP6/T7 genomic library (CLONTECH) made from partially digested *Sau3A* I genomic



DNA separated on a sucrose gradient. Fragments in the range 8 to 22 kb were cloned into the *Bam*H I site of  $\lambda$ EMBL3 SP6/T7. Both *Sal*BA I and *Bam*H I generate the same GATC sticky-end overhang.

### 3.1.1 Bacterial Host Strains

The bacterial host strain for the  $\lambda$ Gem 11 library was *Escherichia coli* LE392 and for the  $\lambda$ EMBL3 SP6/T7 library it was *Escherichia coli* K802. For genomic library amplification, both these libraries exploit the fact that  $\lambda$  bacteriophage adsorbs through maltose receptors in the outer membrane of *E. coli*, receptors which normally transport this oligosaccharide into the cell. Adsorption of  $\lambda$  bacteriophage through maltose receptors is facilitated by magnesium ions.

### 3.1.2 Growth Media

Both host strains of *E. coli* were grown on Luria-Bertani (LB) broth and LB agar plates supplemented with maltose and  $\text{MgSO}_4$ .

LB Broth:	Bactotryptone	10 g
	Bacto yeast extract	5 g
	NaCl	10 g
	ddH <sub>2</sub> O to	980 ml
	Autoclaved	20 psi for 20 minutes

Before inoculating with either *E. coli* LE392 or K802, 1 ml 20% maltose ( $\text{C}_{12}\text{H}_{22}\text{O}_{11}$ ) and 1 ml 1 M  $\text{MgSO}_4$  were added to 100 ml of autoclaved LB buffer at 37°C. Maltose and  $\text{MgSO}_4$  were filter-sterilised as both breakdown at autoclave temperature.

LB Agar:	Bactotryptone	10 g
	Bacto yeast extract	5 g
	NaCl	10 g
	Bacto agar	15 g
	ddH <sub>2</sub> O to	965 ml
	Autoclaved	20 psi for 20 minutes

1 ml 20% maltose and 1 ml 1 M MgSO<sub>4</sub> were added per 100 ml LB agar broth, just before plates were poured and when its temperature was approximately 65°C. The recipe was the same for top agarose except that 7 g of agarose was added in place of the 15 g agar.

### **3.1.3 Glycerol Stocks and Inoculation**

Glycerol stocks of each strain were prepared by adding 8.5 ml of stationary phase broth to 1.5 ml of glycerol. These were stored at -80°C. All working cultures were taken from the glycerol stocks, grown overnight in LB broth with maltose and MgSO<sub>4</sub>, and subcultured in the same medium, and allowed to achieve high-density exponential growth at 2X10<sup>8</sup> bacteria per ml. All cultures were incubated at 37°C and aerated by shaking at 200 rpm (Centomat H shaker), until they had reached the desired cell density.

### **3.1.4 Plating Genomic Libraries**

15 cm LB agar plates, containing 30 ml LB agar with final concentrations of 10 mM MgSO<sub>4</sub> and 0.2% maltose, were left to set for approximately 15 minutes after flaming to remove any bubbles. 10 ml maltose/MgSO<sub>4</sub> supplemented LB broth was inoculated with the host strains, LE392 or K802, from glycerol stocks and incubated on a shaker at 37°C overnight, then subcultured into 10 ml

of the same maltose/MgSO<sub>4</sub> LB broth and brought to a cell density of  $2 \times 10^8$  bacteria per ml. In 50 ml capped polypropylene tubes, 200  $\mu$ l of these exponential cultures and 1  $\mu$ l of diluted stock genomic library were incubated at 37°C for 20 minutes. The stock genomic library dilution to use was determined by a library's titre (see the next section). 10 ml of melted 0.07% maltose/MgSO<sub>4</sub> top agarose, at approximately 45°C, was added to each tube and quickly mixed by inversion. Entire contents were poured on to separate maltose/MgSO<sub>4</sub> LB bottom agar plates. These were quickly swirled to allow even spreading of the agarose. Bubbles were removed by flaming and plates were allowed to cool to room temperature for 10 minutes to let them set and for the inoculum to soak into the bottom agar.

All plates were then inverted to prevent condensation smearing, and incubated at 37°C overnight, or until the plaques were just about to make contact with each other. This time varied from 8-16 hours. Plates were sealed in parafilm, inverted and stored at 4°C. This had the desired effect of arresting lysis before scoring.

### **3.1.5 Titering Genomic Libraries**

To be confident of identifying a positive genomic clone, screening had to be redundant. This meant that sufficient copies of the human genome had to be screened to overcome the possibility that, by chance, positive clones were missing from any particular aliquot of a library. Six-fold redundancy was adopted as an appropriate screening precaution, and calculated according to the formula:

$$\text{human genomes per } \mu\text{l} = \frac{\text{average number bp per insert} \times \text{number pfu}/\mu\text{l}}{\text{bp in human genome}}$$

In this formula, pfu stands for plaque forming units. The pfu titre was determined empirically following the plating protocol described in Section 3.1.4 (pp. 75-76).  $10^{-6}$ ,  $10^{-7}$ ,  $10^{-8}$ , and  $10^{-9}$  dilutions of the original library were plated and scored to determine pfu.

### 3.1.6 Determining Insert Size

The average size of the human insert for the  $\lambda$ Gem 11 library had been determined to be 15 kb (personal communication Dr. A. Akkari, 1993). Insert size for the  $\lambda$ EMBL3 SP6/T7 library had to be confirmed from the stated CLONTECH range of 8-22 kb. This was achieved by comparison of two uncut  $\lambda$ EMBL3 SP6/T7 clones with *Bam*H I, *Eco*R I and *Hind* III digests. It was possible to calculate insert size because *Bam*H I, *Eco*R I and *Hind* III did not cut the vector or the TPM2 STS probe sequence. Therefore, any lane with two or more fragments was due to one of these enzymes cutting the non-STS part of the insert. Lower molecular weight bands on digest gels were summed to give approximate insert sizes.

## 3.2 PROBING GENOMIC LIBRARIES

### 3.2.1 Lifting Libraries

Using sterile forceps, labelled 137 mm diameter nylon nucleic acid transfer membranes (Hybond-N<sup>+</sup>, Amersham) were placed on to the top of plated genomic libraries. Trapping air bubbles was avoided by allowing the

membrane to 'peel down', and spread across the library from its initial point of contact. Once in place, membranes were marked by stabbing through them into the agar with an 18-gauge needle dipped in printer's black ink. This allowed the alignment of the membrane autoradiographs with plates for the identification of positive clones.

After 2 minutes, membranes were carefully peeled off genomic libraries, placed plaque-side-up for 2 minutes on 15 x 15 cm Whatman 3MM chromatography paper soaked in NaOH denaturing solution, which comprised 1.5 M NaCl in 0.5 M NaOH. Membranes were then placed, again plaque-side-up, for 2 minutes on another 15 x 15 cm Whatman 3MM paper soaked in Tris neutralising solution which comprised 1.5 M NaCl in 0.5 M Tris-HCl (pH 7.5). Finally, membranes were placed between two Whatman 3MM papers to blot dry.

Duplicate lifts were made for all plates by placing a second Hybond-N<sup>+</sup> membrane on to each plate and marking them with ink in the same locations as the first membrane. The lifting protocol was the same for primary and duplicate lifts. All membranes were left to dry at room temperature, overnight and in a dark place, stacked between Whatman 3MM papers.

### **3.2.2 Labelling Probes**

The probes employed to screen the  $\lambda$ Gem 11 and  $\lambda$ EMBL3 SP6/T7 genomic libraries were available at ANRI. They were: LK603 (Wade et al., 1990) (a gift from Dr. Larry Kedes) for *TNNI1*; a PCR amplified STS of the *ACTA1* gene (Freeman and States, 1991); and a 1.5 kb *TPM2* STS manufactured with a common exon IV forward primer and a muscle specific exon VIIIsk reverse primer (Hunt et al., 1995).

The probe labelling system employed was developed by Feinberg and Vogelstein (1984). Random hexanucleotides were used to prime the Klenow DNA polymerase extension of each probe strand, with  $\alpha$ - $^{32}\text{P}$  dCTP to internally label probes. This method generated probes of up to 50% of the length of the original template, and with specific activity greater than  $10^9$  cpm/ $\mu\text{g}$  (Feinberg & Vogelstein, 1984).

25 ng of DNA template was denatured by heating at 95-100°C for 2 minutes. Denatured probes were snap-chilled in an ice-bath to prevent renaturation. The labelling reaction was assembled on ice and behind a perspex screen by adding the following reagents to a 0.6 ml eppendorf tube (final reaction concentrations are shown in brackets): 10  $\mu\text{l}$  labelling 5X buffer (1X); 2  $\mu\text{l}$  mixture of unlabelled dNTP (20  $\mu\text{M}$  each); 25 ng denatured DNA template (0.5 ng/ $\mu\text{l}$ ); 5  $\mu\text{l}$   $\alpha$ - $^{32}\text{P}$  dCTP (Bresatec) (50  $\mu\text{Ci}$  where Ci stands for Curie); 2  $\mu\text{l}$  acetylated, nuclease free, 10 mg/ml BSA (0.4  $\mu\text{g}/\mu\text{l}$ ); 5 units (u) Klenow enzyme (0.1 u/ $\mu\text{l}$ ); sterile ddH<sub>2</sub>O (Baxter) to final volume 50 $\mu\text{l}$ . This reaction was mixed gently and incubated at 25°C for 60 minutes in a thermal cycler (MJ Bresatec), then terminated by heating to 95-100°C for 2 minutes. Ice-bath snap-chilling kept probes single-stranded. EDTA was added to 20 mM.

Probes manufactured in this way were purified using Sephadex G-50 size-exclusion chromatography spin columns. The purpose of the size-exclusion chromatography spin column was to remove the unincorporated  $\alpha$ - $^{32}\text{P}$  dCTP. Spin columns were prepared by plugging a 3 ml syringe with a small piece of sterile non-absorbent cotton wool. The 3 ml syringe was filled with Sephadex

G-50 and compacted until it measured about 2 ml. The plunger was removed and discarded. 1 ml STE buffer was added to the top of the Sephadex G-50 column and the syringe placed in a 15 ml polypropylene conical tube. STE buffer comprised: 0.1 M NaCl; 10 mM Tris.Cl (pH 8.0); and 1 mM EDTA (pH 8.0). Spin columns were centrifuged at 2000 rpm for 5 minutes in a Beckman TJ-6 centrifuge. Syringes were placed in new 15 ml tubes and labelled probes were added to the top of the Sephadex column. These were centrifuged at 2000 rpm for 5 minutes in an International Equipment Company (IEC) HN-5 'radioisotope' centrifuge.

Purified probes, the eluents, were pipetted from the 15 ml tubes into 0.6 ml eppendorf tubes and denatured by heating to 94°C for 5 minutes. Snap-chilling on ice for 1 minute kept labelled probes single-stranded.

The percentage incorporation of  $\alpha$ -<sup>32</sup>P into the probe by random priming was checked in two ways. First, the fraction, 'eluent cpm/total cpm' gave an approximate percentage value for incorporation efficiency. Total cpm was calculated from: Sephadex column cpm plus eluent cpm. It was a simple matter to determine the actual counts incorporated by multiplying the original counts added to labelling reaction with incorporation percentage. Since 1  $\mu$ Ci contains  $2.2 \times 10^6$  counts, the added 50  $\mu$ Ci  $\alpha$ -<sup>32</sup>P dCTP contained  $1.1 \times 10^8$  counts.

The second estimate of radioisotope incorporation efficiency used PEI cellulose thin layer chromatography. An 8 cm strip of PEI paper was cut an 'x' was marked, in pencil, in the middle of this strip, 1 cm from one end. This end was designated the bottom of the PEI strip. 0.5 ml of the random priming reaction was loaded on to this 'x'. 5 ml of 0.75 M K<sub>2</sub>H<sub>2</sub>PO<sub>4</sub> (pH 3.0) was pipetted into

a Coplin jar, and the loaded PEI strip was placed in this jar so that its bottom end was in the  $K_2H_2PO_4$ , with the 'x' just above the solvent level. Only the top of the strip touched the side of the jar. The  $K_2H_2PO_4$  solvent moved up to the top of the PEI strip. The front of the  $K_2H_2PO_4$  solvent could be seen quite clearly, and the PEI strip was removed from the Coplin jar when this front had migrated to within 1 cm of the top end of the PEI strip. cpm was measured at 'x', then up to the solvent front. Since unincorporated  $\alpha$ - $^{32}P$  dCTP was dissolved in the  $K_2H_2PO_4$ , while labelled probe was not, probe labelling efficiency could be calculated from:

$$\text{Probe labelling efficiency} = \text{cpm at 'x'} / (\text{cpm up the strip} + \text{cpm at 'x'})$$

If the labelling efficiency was too low, 0-20%, then the random priming reaction was repeated using more than the original 25 ng DNA template.

### **3.2.3 Pre-hybridisation**

Membranes were placed in a round plastic container with pre-hybridisation solution and incubated at 42°C for at least four hours with shaking (Stuart Scientific Hybridisation Oven/Shaker). 15 ml of pre-hybridisation solution was added per membrane. The ingredients and final concentrations of the pre-hybridisation solution were: 1 M NaCl; 1 M de-ionised formamide; 50 mM Tris pH 7.5; 0.1 mg/ml salmon sperm DNA; and 10% dextran sulphate. Autoclaved salmon sperm DNA was boiled at 100°C for 5 minutes, snap-chilled on ice then added to the pre-hybridising solution. This procedure fragmented the salmon sperm DNA to a size range of 500-700 bp (Sambrook,



1989, B.15). Formamide was deionised by mixing 5 g AG 501-X8 mixed bed resin (*Bio-Rad*) with every 100 ml formamide. This was gently shaken for one hour at room temperature then filtered through 305 mm diameter Postlip Paper (Hollingsworth & Vose Co. Ltd).

The salmon sperm blocks non-specific binding of the probe to the membrane (Sambrook et al., 1989, 9.48-9.50); formamide keeps the DNA probe single stranded; 10% dextran sulphate increases the rate of hybridisation ten-fold, because nucleic acids are excluded from the volume of the solution occupied by this polymer, effectively increasing probe concentration; 1 M NaCl maximises the rate of annealing of the probe with its target as solutions of high ionic strength increase hybridisation; 1% SDS solution is a surfactant that facilitates probe-target annealing (Sambrook, et al., 1989, 9.47-9.55). Pre-hybridising solution was stored at -20°C in 40 ml aliquots which were wrapped in alfoil. It was considered suitable for use as long as it remained frozen.

#### **3.2.4 Hybridisation**

After 4 hours of pre-hybridisation at 42°C, *TNNI1* and *ACTA1* radiolabelled probes were added to the pre-hybridised  $\lambda$ Gem 11 membranes and the *TPM2* radiolabelled probe was added to the pre-hybridised  $\lambda$ EMBL3 SP6/T7 membranes. Labelled probes were again denatured, this time by boiling with 6  $\mu$ g salmon sperm DNA. After snap-chilling on ice, this mixture was added to the pre-hybridisation solution containing membranes, so that the final activity of the hybridisation solution was  $1 \times 10^6$  counts per ml. Hybridisation took place at 42°C for between 16-20 hours, with shaking.

At the end of hybridisation, 2XSSC was added to the hybridisation solution to triple its volume and this mixture was drained into a flushing sink. The added 2XSSC reduced the viscosity of the hybridisation solution, thus facilitating its removal. SSC was made up as a stock 20XSSC: 175.3 g NaCl; 88.2 g sodium citrate; 800 ml ddH<sub>2</sub>O; pH adjusted to 7.0 with a few drops of a 10 M NaOH; and final volume set at 1 litre with addition of ddH<sub>2</sub>O.

Membranes were washed at increasing stringency in the hybridisation tubs until membrane cpm fell to 10-20 cpm/membrane. This was determined by removing membranes and testing them individually with a Geiger counter. The first and least stringent wash was in 2XSSC for 15 minutes in a shaking water bath at 42°C. 75 ml 2XSSC was added per membrane. The second wash was in 2X SSC and 1% SDS for 15 minutes at 42°C, in a shaking waterbath.

The third wash, only conducted if activity had not dropped sufficiently, was the most stringent wash and carried out in 0.1X SSC and 0.1% SDS, for 15 minutes in a shaking water bath at 42°C. Washing temperatures were increased for membranes with high to very high activity, as this also increased washing stringency.

Membranes were blot dried with Whatman 3MM paper and wrapped in Glad Wrap. The duplicate membranes were paired and taped to a large piece of Whatman 3MM paper, that was cut to fit the size of the autoradiograph cassette. This anchored membranes, preventing movement. The Whatman 3MM backing paper was fluorescently and asymmetrically marked with three distinctly different symbols. The fluorescent marker exposed the X-ray film. Membranes were placed in a cassette for autoradiography with Dupont Cronex

4 X-ray film and Quanta III autoradiography intensifying screens, at -80°C. Exposure time was overnight, precise times determined by the activity level of membranes.

### **3.2.5 Positive Primary Clones**

After developing, films were aligned with membranes to ensure that they had not moved. The fluorescent marks were useful in this alignment. Permanent ink marks were added to the autoradiographs to match the Indian ink marks on the membranes. These permanent ink marks were then used to align autoradiographs with agar plates and so identify positive plaques. Only plaques that were positive on duplicate lifts were considered for picking. When possible, single isolated plaques were picked. In the initial screens this was not usually possible due to confluence. Agar plugs normally contained several plaques, only one of which was positive.

Plaques, were picked over a light box and using sterile, Pasteur pipettes. Transparent glass pipettes were preferred to disposable tips with their ends cut off. The tips of Pasteur pipettes were pushed through the 0.7% agarose top agar and into the bottom agar to select positive plaques. Agar plugs were removed by suction with a moistened finger held over the open 'bulb end' of the pipette. Bulbs had a tendency to suck agar plugs too far into the pipette. Agar plugs were placed into a sterile 1.5 ml eppendorf with 200 µl autoclaved SM buffer. Phage eluted from agar plugs into SM buffer. Eluted phage and agar plugs were stored at 4°C. SM buffer comprised: 5.8 g NaCl; 2 g MgSO<sub>4</sub>·7H<sub>2</sub>O; 50 ml 1 M Tris Cl (pH 7.5); 5 ml 2% gelatin solution; ddH<sub>2</sub>O to 1 litre. These were autoclaved and stored in 50 ml aliquots at room temperature (Sambrook et al., 1989, A7).

### 3.2.6 Secondary and Subsequent Screening

Using the eluted phage from each selected positive plaque, a secondary screen was conducted. Phage were re-plated and titered to obtain 200-1000 plaques on a 15 cm petri dish. This was achieved by placing 1, 10, and 100  $\mu$ l of the eluted phage from the agar plug into 200  $\mu$ l of the overnight broth. Then, 1  $\mu$ l of these phage eluates were added to 200  $\mu$ l of exponential growing *E. coli* LE392 or K802 host bacteria. Suitable density plates were lifted using the nylon membranes and rescreened using the same radiolabelled probes and same procedures employed in the primary screen.

Rescreening was repeated until 100% of the plaques showed up as positive. For some primary plaque clones this would occur after the tertiary screen; for all clones this was achieved after four rounds of screening. In the final screen, low density plates were prepared with 10-50 plaques, as well as plates with 200-1000 plaques. The lower density plates were used to pick positive clones for stock, while higher density plates were used to confirm that clones were pure clones. Well-isolated positive plaques selected for stock, were picked and placed into 200  $\mu$ l SM buffer. These were stored at 4°C.

The PCR assay of positive stock clones and the preparation and purification of large quantities of high titre positive clone DNA by plate lysate method, was conducted by others. Approximately 10  $\mu$ g of purified positive clones were forwarded to Department of Cytogenetics and Molecular Genetics, Adelaide Children's Hospital, for use in FISH.

### 3.3 RESULTS

#### 3.3.1 Genomic Libraries and Probes

The  $\lambda$ Gem 11 had an original activity of approximately  $1 \times 10^{10}$  pfu/ $\mu$ l (personal communication, Dr. L. Abrahams 1993). Repeated titrations, however, showed this library to have an average activity of  $2.4 \times 10^8$  pfu/ $\mu$ l. Genomic libraries lose activity over time. The  $\lambda$ EMBL3 SP6/T7 library had a stated activity of  $5 \times 10^6$  pfu/ $\mu$ l (CLONTECH, 1994). This was its measured activity. It was a recently purchased genomic library.

$\lambda$ Gem 11 insert was, on average, 15 kb (personal communication, Dr. A. Akkari, 1993). The *TPM2* insert was also approximately this size. Restriction digestion showed one positive *TPM2* clone insert to be 15 kb while another was 16-18 kb. This was in agreement with the CLONTECH range of 8-22 kb for genomic cloned fragments of average size 15 kb. For both the  $\lambda$ Gem 11 and the  $\lambda$ EMBL3 SP6/T7 genomic libraries, an average insert size of 15 kb was used to calculate the number of plates needed to achieve six-fold redundancy screening.

Labelling the LK603 probe (Wade et al., 1990) and the *ACTA1* STS probe (Freeman and States, 1991) was successful with 25 ng DNA template as described in Section 3.2.2 (pp. 78-81). However the incorporation of  $\alpha$ - $^{32}$ P dCTP was too low when 25 ng of purified *TPM2* STS was used as labelling template DNA. It became necessary to increase the *TPM2* STS purified DNA template five-fold, to 125 ng, to obtain a satisfactory level of  $\alpha$ - $^{32}$ P dCTP

incorporation into this probe. Probes used in hybridisation all had  $\alpha$ - $^{32}\text{P}$  dCTP incorporation greater than 50%.

### **3.3.2 Hybridisation**

Primary plates were always allowed to grow close to confluence with an average of 50 000 pfu/plate. These were ideal for the first screening of the duplicate membrane lifts because at this confluence eight plates contained the equivalent of one diploid genome ( $50\,000 \times 8 \times 15\,000 \text{ bp} = 6 \times 10^9 \text{ bp}$ ).

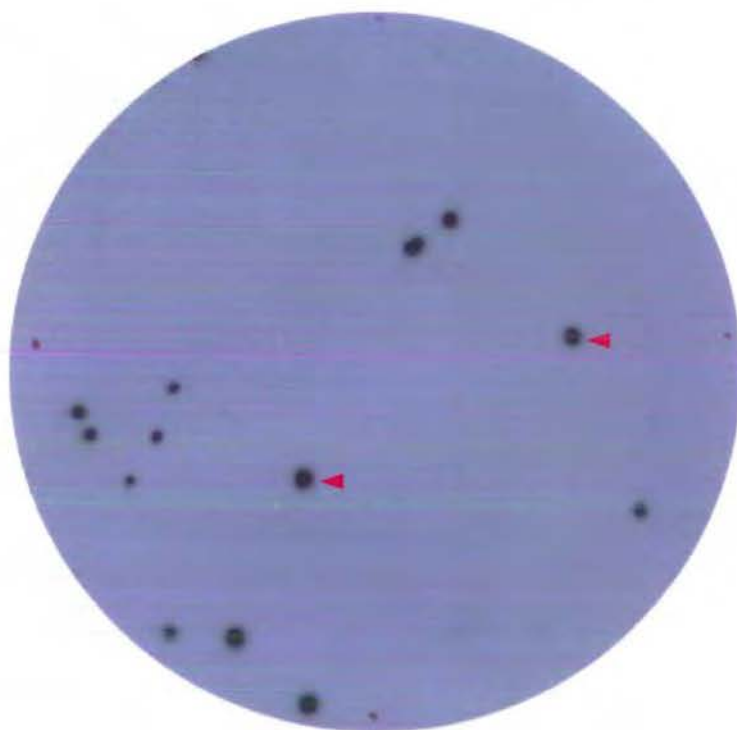
This near-confluence meant that secondary screening always showed that primary 'positive plaque' agar plugs contained several plaques, only one of which was positive. Figure 3.1 (p. 88) shows the result of a duplicate quaternary screen of the  $\lambda$ Gem 11 for positive *TNNI1* plaques.

### **3.3.3 FISH Localisation**

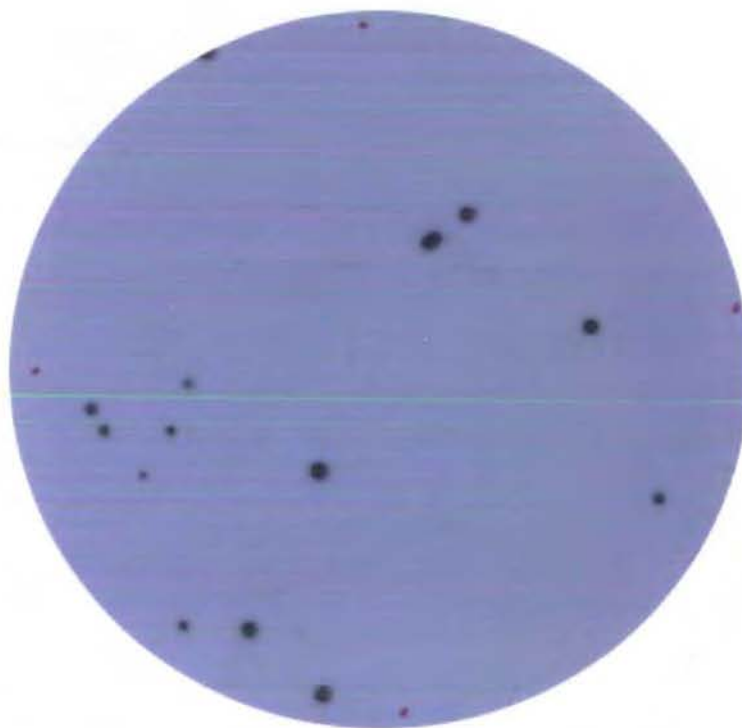
Under the direction of Dr. David Callen at Department of Cytogenetics and Molecular Genetics, Adelaide Children's Hospital, the *TNNI1*, *ACTA1* and *TPM2* probes provided by ANRI were used to localise these genes by FISH. FISH expertise was not available in Western Australia at that time, so the world-wide reputation of Dr. Callen's team as FISH localisation experts, made them an obvious choice to complete this section of the research.

The first of these genes to be localised by FISH to 1q32 was *TNNI1* (Eyre et al., 1993). Figure 3.2 (p. 89) shows the FISH localisation of *TNNI1* to demonstrate the end-point of this research. Subsequently, *ACTA1* was localised to 1q42 (Akkari et al., 1994) and *TPM2* to 9p13 (Hunt et al., 1995).

**A.**

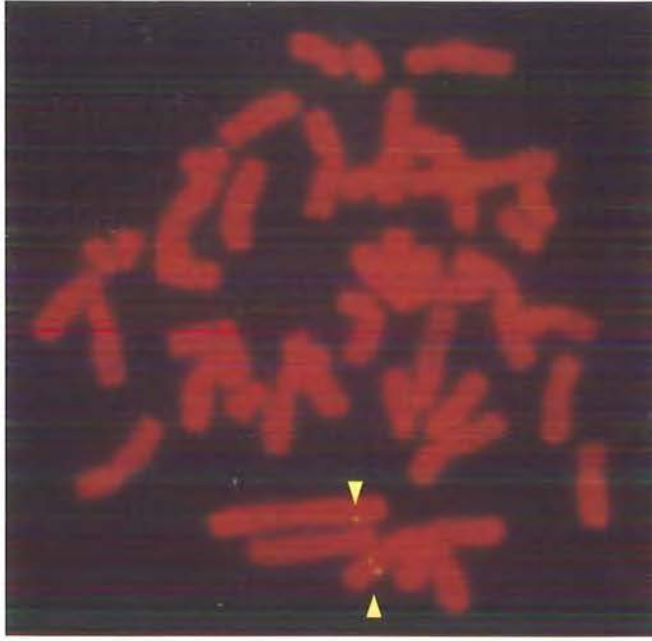


**B.**



**Figure 3.1** Autoradiographs of a  $\lambda$ Gem 11 genomic library quaternary screen for positive *TNNI1* clones. **A** is the first lift, **B** the duplicate lift. Red arrows identify the positive *TNNI1* clones that were picked. Low density plates like this one were preferred for final selection of 'stock' positive clones. Perimeter 'aligning' dots can be seen on each autoradiograph.

**A.**



**B.**



**Figure 3.2** A whole metaphase showing *in situ* hybridisation of the biotinylated *TNNI1* probe to 1q32. **A.** Normal male chromosomes stained with propidium iodide. Hybridisation sites are indicated by arrows. **B.** The same metaphase stained with DAPI for chromosome identification. Photographs kindly provided by H.J. Eyre, Department of Cytogenetics and Molecular Genetics, Adelaide Children's Hospital.



For FISH, all ANRI probes were nick-translated with biotin-14-dATP and hybridised *in situ* to metaphases of two or three normal males. Probe concentration was 15 ng/ml. The FISH method was modified from that previously described (Callen et al., 1990) in that only two rounds of amplification with fluorescein-conjugated avidin and biotinylated goat anti-avidin were employed. Chromosomes were stained before analysis with both propidium iodide (as counterstain) and, for chromosome identification, with 4', 6 diamidino-2-phenylindole (DAPI).

In this thesis, all figures which involved scanning, like Figures 3.1 and 3.2, were scanned using a MicroteckX6 scanner, run by Microteck *Scan Suite* 1.0.1 software (1998). Images were saved in either *MicroFrontier ColourIt* (1997) or AdobePhotoshop 5.0.2 (1998) using either PICT or JPEG formats. To improve the printed image quality of scanned developed X-ray films the, brightness/contrast in ColourIt was set at +11/+22, effectively altering the background from smoke-grey to blue.

### **3.4 CONCLUSION**

The human contractile skeletal muscle genes *TNNI1*, *ACTA1*, and *TPM2* were localised to 1q32 (*TNNI1*), 1q42 (*ACTA1*), and 9p13 (*TPM2*) on the cytogenetic map of the human genome. This made them candidate genes for inherited skeletal muscle diseases whose candidate region included one of these locations. At the time of completion of this work no muscle diseases had been localised to these regions. They were, therefore, candidates 'in waiting' for inherited muscle diseases yet to be localised. This included the forms of distal myopathy and OPMD discussed in Chapters 4, 5, and 6 of this thesis.

Subsequently, mutations have been identified in *ACTA1* (Nowak et al., 2000) and *TPM2* (Donner et al., 2000), all associated with congenital myopathies, especially nemaline myopathy. No disease causing mutations have been identified in *TNNI1*.

## **CHAPTER FOUR**

# **THE LOCALISATION OF AN AUTOSOMAL DOMINANT DISTAL MYOPATHY GENE**

## **4.0 THE LOCALISATION OF AN AUTOSOMAL DOMINANT DISTAL MYOPATHY GENE**

The localisation of a form of AD-distal myopathy segregating in a Western Australian kindred is reported here (Aim (ii), p. 3). It launched the molecular genetic investigation of the distal myopathies. In line with the rationale of this thesis (p. 4), there was the expectation that this work would contribute to progressive improvements in the nosology, diagnosis, research, prognosis, prevention, and treatment of the distal myopathies.

### **4.1 DISTAL MYOPATHY NOSOLOGY**

Gowers (1902) described a 10 year old patient with weakness and wasting of the hand, forearm and anterior tibial muscles. By the age of 18, severe wasting of the sternomastoids, wasting of the tongue and weakness of the facial muscles were also evident. This was the first description of hereditary distal myopathy (MIM160500) (McKusick, 1992).

At the start of this thesis, Markesbery and Griggs (1986, p. 1322) and Barohn et al. (1991, p. 1369) had presented the most recent nosologies for distal myopathy. Their classifications were similar, both based on the same clinical and pathological features and mode of inheritance. They differed in that Barohn et al. (1991) expanded the Markesbery and Griggs (1986) classification from two late-adult onset forms and one early-adult onset form of distal myopathy, to two late-adult onset forms and two early-adult onset forms of distal myopathy. As well as grouping distal myopathies by common age of onset, the classification of Barohn also grouped distal myopathies by commonality of onset-site, mode of inheritance, creatine kinase (CK) levels and muscle biopsy results. Table 4.1 (p. 94) groups distal myopathies after Barohn's classification.

TABLE 4.1

*Nosology of the Hereditary Distal Myopathies (after Barohn et al., 1991, p. 1369)*

Clinical phenotype	Late-Adult Onset, Type 1	Late-Adult Onset, Type 2	Early-Adult Onset, Type 1	Early-Adult Onset, Type 2
Identification:	Welander, 1951 Edström et al., 1975	Sumner et al., 1971 (F) Markesbery et al., 1974 (AD)	Markesbery et al., 1977 (S) Nonaka et al., 1981 & 1985 (AR) Scoppetta et al., 1984 (AR)	Miyoshi et al., 1986 (AR) Kuhn & Schroder 1981 (AR/F) Galassi et al., 1987 (AR/F+S)
Inheritance:	AD	AD & F	AR & S	AR, AR or F & S
Age of onset:	Over age 40	Over age 40	20-30 years	15-30 years
First symptoms:	Hand muscles	Distal leg muscles, anterior compartment	Distal leg muscles, anterior compartment	Distal leg muscles, posterior compartment
Progression:	Very slow - distal flexors and extensors	Slow - upper extremities, proximal limb muscles	Rapid - proximal leg muscles, arms, hands, neck then trunk. Tendon reflexes diminished	Medium - proximal leg muscles. No atrophy or weakness in upper extremities
Cardiomyopathy:	No	Yes	No	No
Serum CK:	Normal or slightly elevated	Normal or slightly elevated	Increased, usually <10X normal, often 2-5X normal	10-150X normal
Electromyography:	Myopathic pattern	Myopathic pattern	Myopathic pattern	Myopathic pattern
Muscle biopsy :	Variable, some rimmed vacuoles	Rimmed vacuoles	Rimmed vacuoles	Dystrophy, no vacuoles, some gastrocnemius weakness
F = Familial;    S = Sporadic;    X = 'times'.				

## **4.2 ASCERTAINMENT OF A WESTERN AUSTRALIAN FAMILY SEGREGATING DISTAL MYOPATHY**

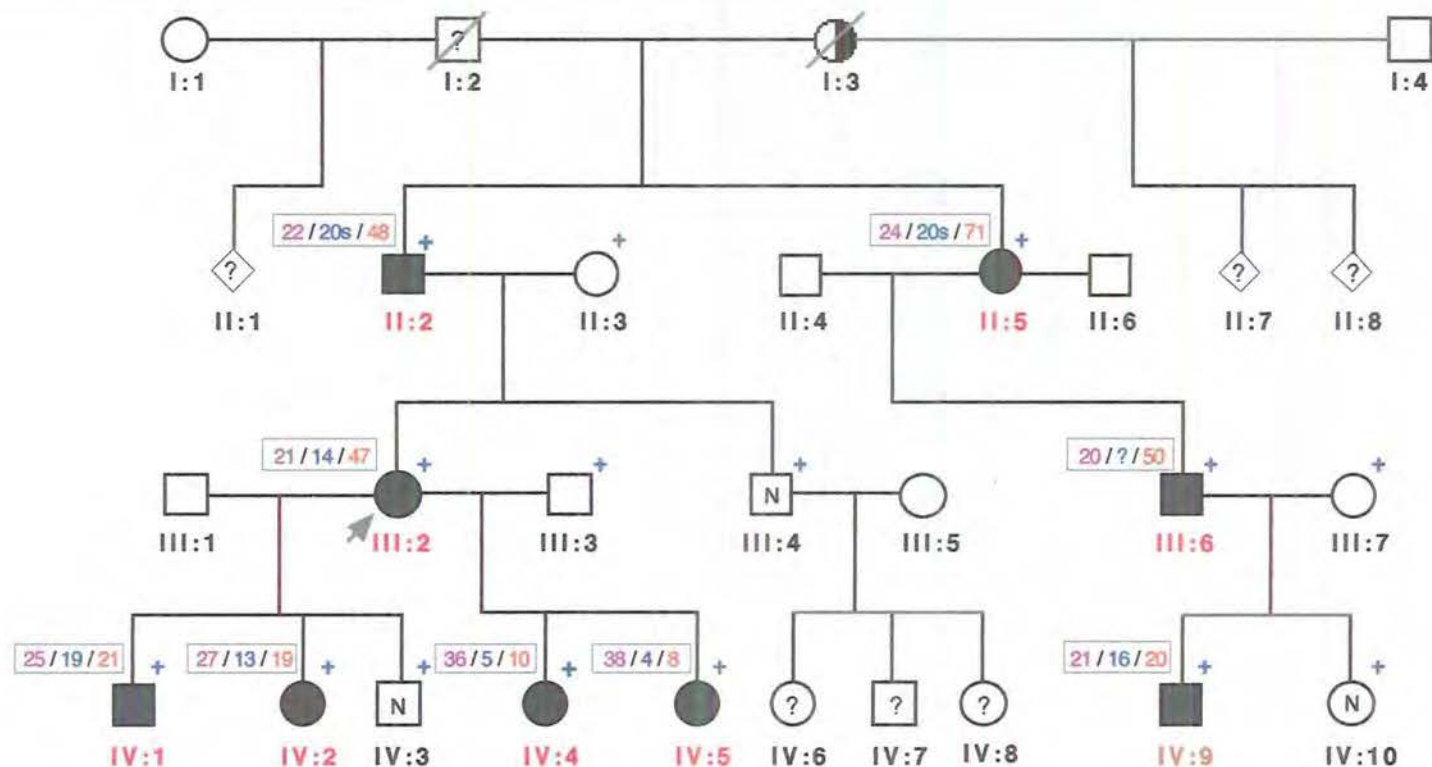
Ascertainment of a Western Australian (WA) family segregating a form of distal myopathy was conducted in the Neuromuscular Clinic of the Australian Neuromuscular Research Institute (ANRI), under the supervision of Professor Frank Mastaglia, Professor of Neurology at the University of Western Australia. This family was of English/Welsh origin and was described as the WA Myopathy Distal family, or the WAMPD family. Nine affected and six unaffected WAMPD family members agreed to take part in this study. The four generation WAMPD pedigree is consistent with a highly penetrant AD form of distal myopathy (Figure 4.1, p. 96).

### **4.2.1 Clinical Features**

Onset varied from 4-25 years, with indication of earlier onset in generations III and IV. This indication may be ascertainment bias coming from two sources. First, had the onset of distal myopathy been earlier in II:2 and II:5 (red means affected), they might not have reproduced. The very existence of this distal myopathy family could be a source of ascertainment bias.

Second, the earlier onset of distal myopathy in III:2 and in her children, IV:1, IV:2, IV:4 and IV:5, made the segregation of this disease quite apparent and caused III:2 to draw clinical attention to her family's condition. If distal myopathy onset had been later in III:2 and her children, this family may not have identified itself.

Either or both of these sources of ascertainment bias could explain the apparent earlier onset of distal myopathy in generations III and IV.



**Figure 4.1** Pedigree of a Western Australian family segregating distal myopathy (the WAMPD family). Squares denote males, circles females, lines through squares or circles indicate 'deceased', and diamonds identify individuals of unknown sex. Blackened symbols and red numbers indicate individuals affected by distal myopathy. The half-blackened symbol shows an individual who, by history, was probably affected by distal myopathy. An 'N' in the centre of the symbol denotes 'unaffected by distal myopathy' while a '?' in the centre of the symbol denotes 'unknown status'. The arrow identifies the proband, III:2. Each individual marked with a cross (+) was examined clinically, had an immortalised lymphocyte cell line established, and contributed DNA for linkage analysis. The boxes at the top left-hand side of affected participants contain: their mother's age at the time of their conception/ the reported age of onset of their distal myopathy symptoms/ and their age at ascertainment. AD inheritance of high penetrance was indicated.



Excluding ascertainment bias, there are four other explanations for the apparent earlier onset of distal myopathy in generations III and IV: anticipation caused by an unstable trinucleotide repeat expansion (Fu et al., 1991); anticipation caused by meiotically stable trinucleotide expansion (Brais et al., 1998); maternal age effect; and effects of other genes. Two events in the WAMPD pedigree indicate possible effects of maternal age and/or other genes on the age of distal myopathy onset.

The progressively earlier age of distal myopathy onset in the siblings/step-siblings IV:1, IV:2, IV:4 and IV:5, correlate with increasing maternal age of III:2, as shown in Figure 4.1. The mother's age at conception versus the age of distal myopathy onset for these individuals, was: IV:1 (25/19), IV:2 (27/13), IV:4 (36/5) and IV:5 (38/4). From a study of 17 affected sibling pairs born to mothers with myotonic dystrophy, Andrews and Wilson (1992) discovered that, in 13 of the 17 sibling pairs, later-born affected children suffered more severe disease than their first-born siblings. They displayed significantly more neonatal feeding difficulties, a later age when first sitting alone, a later age when first walking alone, and a higher incidence of scoliosis. The greater the age difference between affected sibs, the greater the difference in their disease severity, suggesting that increasing maternal age was a factor in the relative disease severity of affected children.

The two older step-siblings, IV:1 and IV:2, and the two younger step-siblings, IV:4 and IV:5, had different fathers, III:1 and III:3 respectively. These older step-siblings did not develop symptoms until their mid- to late-teens, the age at which their affected mother became symptomatic. Because III:2 had two partners, each of whom fathered two affected children, there is a potential link between paternal genetic contribution and age of onset of distal myopathy.



Unfortunately, the putative maternal age effect and paternal genetic contribution, confound each other in this pedigree.

Muscle weakness first presented with selective weakness of the toe extensors and ankle dorsiflexors, resulting in bilateral foot drop. This form of distal myopathy onsets in the anterior compartment of the leg which contains: the tibialis anterior (dorsiflexes and inverts the foot); the extensor hallucis (dorsiflexes and inverts the foot, extends the great toe); extensor digitorum longus (extends the toes II-V); peroneus tertius (dorsiflexes and inverts the foot).

The degree of weakness in the anterior tibial muscle groups was variable. Weakness of the ankle evertors and invertors was difficult to evaluate in the presence of weak dorsiflexion. Some degree of weakness of neck flexors was present in all cases and was associated with mild to moderate atrophy of the sternomastoid muscles, particularly the sternal head. Selective weakness of the long finger extensor muscles was present. The finger flexors and intrinsic hand muscles were relatively unaffected, but certain proximal muscle groups such as the hip abductors and external rotators and shoulder abductors were mildly affected.

Table 4.2 (p. 99) shows the pattern of muscle weakness in the nine affected members of this family. Although progression was gradual, there was eventually a moderate degree of incapacity. The oldest living affected family member, II:2, was initially diagnosed as suffering from distal myopathy in 1969, yet was still walking when re-examined in 1992, although he had developed difficulty maintaining an erect posture. The pattern of muscle weakness in the WAMPD family most closely resembles that described by Gowers (1902) in his original report of distal myopathy.

**Table 4.2**     *Distribution of muscle weakness in 9 distal myopathy individuals in the WAMPD family*

Muscle Group		* MRC Grades
<b>Neck</b>	sternomastoids	2 - 4+
<b>Shoulder</b>	flexors	3+ - 5
	abductors	3+ - 5
<b>Wrist</b>	flexors	4+ - 5
<b>Finger</b>	flexors	4 - 5
	extensors	1+ - 5
	abductors	3+ - 5
	adductors	4 - 5
<b>Thumb</b>	flexors	4 - 5
	extensors	3+ - 5
	abductors	4 - 5
	opposition	3+ - 5
<b>Trunk</b>	flexors	2+ - 5
<b>Hip</b>	flexors	4 - 5
	extensors	4 - 5
	abductors	4+ - 5
	adductors	4 - 5
	external rotators	4 - 5
<b>Ankle</b>	dorsiflexors	0 - 4
	invertors	4 - 4+
	evertors	2+ - 5
<b>Toe</b>	extensors	0 - 4
	flexors	3+ - 5
Remaining muscle groups were grade 5 in all subjects.		

*\* Medical Research Council Grades: 0 = no contraction; 1 = flicker or trace of contraction; 2 = active movement with gravity eliminated; 3 = active movement against gravity; 4 = active movement against gravity and resistance; 5 = normal power (Brooke et al., 1983).*

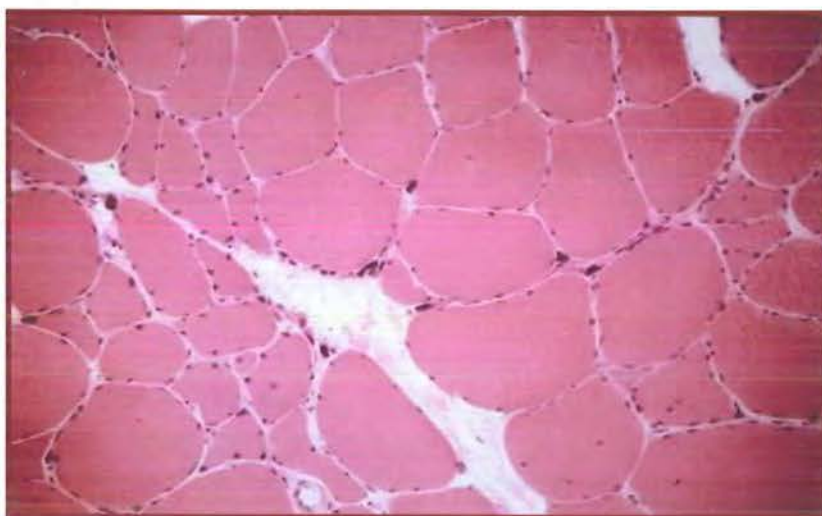
Electromyography and nerve conduction studies were performed in III:2, II:2 and IV:2 by Dr Tim Day of the Neuromuscular Clinic, ANRI. Recordings showed striking myopathic motor unit potential changes; low amplitude, brief duration units, with many polyphasic units. A full, low amplitude (<0.5mV) interference pattern was recorded, particularly in affected distal limb muscles and to a lesser extent in some proximal muscles. Occasional spontaneous fibrillation potentials and positive waves were present in some affected muscles in III:2. In all affected individuals, deep tendon reflexes were preserved and the plantar responses were flexor. There was no myotonia (a disorder of the muscle fibres that results in abnormally long contractions), no sensory impairment or other neurological abnormalities. Motor and sensory nerve conduction studies in the upper and lower limbs were normal. Serum CK levels were elevated in three affected individuals (216-531U/l), according to the Gilboa and Swanson (1976) definition of normal serum CK level as, 'lower than 180U/l'. U stands for unit.

An open biopsy taken from the left vastus lateralis muscle in III:2 showed occasional necrotic and regenerating fibres, excessive variation in fibre size, increased numbers of fibres with internal nuclei, occasional angulated atrophic fibres and nuclear clumps (Figure 4.2A, p. 101). One feature characteristic of some other forms of distal myopathy was *not* observed. This was rimmed vacuoles (RVs). Most forms of distal myopathy have RVs (Table 4.1, p. 94). The failure to observe RVs could have been due to sampling limitations.

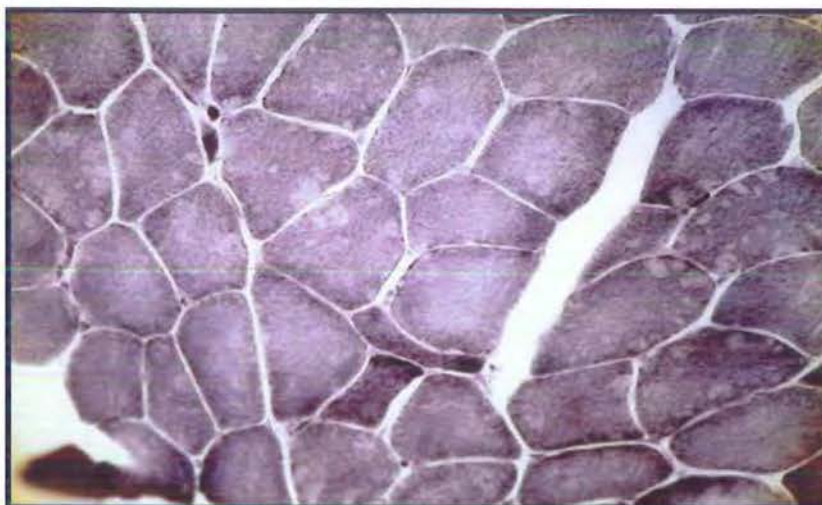
The NADH-TR histochemistry preparations showed only non-specific changes in the more lightly and variably stained moth-eaten type I fibres, but increased enzyme activity in the more darkly stained angulated atrophic type II fibres (Figure 4.2B).



**A.**



**B.**



**Figure 4.2** Vastus lateralis biopsy from III:2 showing:  
**A.** Excessive variation in myofibril size and nuclear clumps (haematoxylin and eosin X400). **B.** Darkly stained angulated atrophic type II fibres and moth-eaten type I fibres (NADH-TR X400). The vastus lateralis predominantly consist of type II, fast-twitch fibres.

Dystrophin and desmin immuno-histochemistry were normal. Desmin storage has been described in a form of AD-distal myopathy (Horowitz & Schmalbruch, 1994). A 1969 biopsy of the left tibialis anterior muscle from II:2 showed evidence of end-stage disease with extreme muscle fibre atrophy, many central nuclei and nuclear clumps, and increase or replacement of endomysial connective tissue.

#### **4.2.2 DNA Extraction**

Blood samples were taken and immortalised cell lines established for the 15 participating family members, marked + in Figure 4.1 (p. 96). Venous blood samples were split 50:50 into lithium-heparin and EDTA tubes for immediate DNA isolation by conventional phenol-chloroform or salt precipitation techniques (Miller et al., 1988), and for lymphocyte isolation and immortalisation with Epstein-Barr virus (Neitzel, 1986). DNA extraction and lymphocyte immortalisation was performed by medical scientists at the Department of Neuropathology, Royal Perth Hospital.

### **4.3 CO-SEGREGATION LINKAGE ANALYSIS**

Table 1.1 (p. 9) showed that genetic linkage had not been demonstrated in any form of distal myopathy at the start of this study, nor had distal myopathy candidate gene(s) been identified (McKusick, 1992). Since information necessary for functional cloning was unavailable for all forms of distal myopathy, co-segregation linkage analysis was the chosen method for disease gene localisation. The WAMPD pedigree contained enough affected members to be reasonably confident of achieving linkage.

#### 4.3.1 Genetic Markers

Genetic markers were identified from a comprehensive linkage map of the human genome (NIH/CEPH Collaborative Mapping Group, 1992), a second-generation linkage map of the human genome (Weissenbach et al., 1992), the Cooperative Human Linkage Center (CHLC), 'CHLC REPORT' (Murray et al., 1993), and from the Genome Database (GDB, 1993; <http://gdbwww.gdb.org/>). 97 polymorphic markers spread over all autosomes were tested for the fifteen WAMPD family members taking part in the study. DNA polymorphisms with a D-number were anonymous DNA fragments (D-segments), other markers were associated with genes. The legend of Table B4.1 (Appendix B) defines acronyms for marker associated with genes. From chromosome 1 to 22, the markers tested against the WAMPD pedigree were:

D1S80, D1S64, D1S223, D1S185, D1S176, CRP, D1S104, AT3, and ACTN2; APOB, D2S44, HOXD4, and D2S102; D3S1307, D3S1304, D3S1285, D3S1278, D3S1279, D3S1282, D3S1262, and D3S1311; D4S45(HD), D4S174, and D4S171; D5S268, D5S112, D5S39, APC, and D5S210; DMDL; D7S472, D7S435, D7S474, GCK, D7S440, and CF; D8S166, D8S84, D8S198, SCA1, and D8S200; D9S104, D9S15, and D9S53; D10S28; D11S875, D11S873, CD3D, and D11S836; D12S43, and D12S60; D13S71, and RB1; D14S72, D14S50, MYH7, D14S64, D14S54, D14S49, D14S52, D14S76, D14S53, D14S74, D14S48, D14S81, D14S45, D14S51, and D14S13; TPM1, and D15S87; D16S291, D16S292, D16S287, D16S295, D16S298, D16S300, D16S308, D16S265, D16S186, D16S301, D16S260, D16S266, and D16S305; D17S30, D17S122, GX-Alu, HOXB6, and D17S26; D18S40, and D18S51; D19S75, APOC2, and DM; D20S66; D21S210, and D21S213; D22S258.

Probes for genotyping the 4 VNTR markers D2S44, D10S28, D14S13 and D17S26 were purchased from Promega. They were labelled with  $\alpha$ -<sup>32</sup>P-dCTP by random priming method as described in Section 3.2.2 (pp. 78-81). For each participating family member, 4  $\mu$ g of gDNA was digested overnight with *Hae*III (Amersham) and digest fragments separated by 0.8% agarose gel electrophoresis. *Bio-Rad* Mini-Sub Cells or Wide Mini-Sub Cells were used to run gels. A *Bio-Rad* PowerPac 1000 generated electrophoresis power.

Gels were alkali transferred to Hybond N<sup>+</sup> nylon membranes (Amersham) using a TE 80 TransVac vacuum blotter (Hoefer) attached to a PV 100 Red-Evac Vacuum/Pressure Module (Hoefer). The pre-hybridisation of these vacuum blots and subsequent hybridisation to labelled probes followed the protocols described in Sections 3.2.3 and 3.3.3 (pp. 81-83). Dried and probed vacuum blots were placed in radiography hypercassettes (Amersham) with Cronex 4 film (Du Pont) and exposed for 7-10 days at -80°C with single intensifying screens.

A Cyclone Plus DNA Synthesiser (Milligen/Biosearch) was used to synthesise primers for the 89 microsatellite (Weber & May, 1989) and 4 non-microsatellite PCR polymorphisms (APOB, AT3, HD, D17S30). Primer sequences were obtained from the GDB (1993; <http://gdbwww.gdb.org/>). Internal labelling was the chosen method for labelling microsatellites. It was simpler, and kinasing trials did not improve low repeat microsatellites autoradiographs, the hardest microsatellites to score.

Microsatellite PCR conditions for a reaction volume of 20  $\mu$ l were as follows: 4  $\mu$ l 5 X buffer (335 mM Tris-HCl [pH 8.8] at 25°C; 83 mM (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>; 100  $\mu$ M of each of the the four dNTPs; 0.2  $\mu$ l 25 mM MgCl<sub>2</sub>; 50 ng of each primer; 0.1  $\mu$ l

$\alpha$ -<sup>32</sup>P dCTP (3,000 Ci/mM, Amersham); 50 ng target DNA; and 0.5 unit (u) *Tth* polymerase (Biotech International). The reactions were overlaid with mineral oil and the cycling conditions were based on two patterns. Using a GENATQ thermal cycler (Pharmacia): 94°C for 5 minutes to denature; 58°C for 6 minutes to anneal and elongate. This initial cycle was followed by 34 cycles of 94°C for 1 minute and 58°C for 6 minutes. Using an MJ thermal cycler (Bresatec): 94°C for 3.5 minutes to denature; 55°C for 1 minute to anneal; 72°C for 1 minute to elongate. This initial cycle was followed by 34 cycles of 94°C for 30 seconds, 55°C for 1 minute and 72°C for 1 minute. When necessary, annealing temperatures were increased to improve primer specificity and yield. PCR reaction conditions for non-microsatellite polymorphisms were similar to those described for microsatellite PCR polymorphisms, except  $\alpha$ -<sup>32</sup>P dCTP was not included in the non-microsatellite PCR reaction mix.

Depending on comb size, 4-8  $\mu$ l aliquots of microsatellite PCR products were mixed with an equal volume of formamide loading buffer and electrophoresed on standard 4% or 6% acrylamide denaturing sequencing gels (Sambrook et al., 1989, 6.49 & 13.47-13.53). Allele size determined gel acrylamide percentage, 4% used for alleles averaging <250 bp. Either an SE 1600 Poker Face II Nucleic Acid Sequencer (Hoeffer) or a Vertical Sequencing Gel (*Bio-Rad*) were used to run gels. Electrophoresis power was supplied by a *Bio-Rad* PowerPac 3000. Gels were washed in a 4:1 (v:v) methanol:glacial acetic acid fixative; partially dried at room temperature; lifted with Whatman 3MM paper cut to size; and completely dried at 80°C in a vacuum Slab Gel Dryer (Savant). Levels of  $\alpha$ -<sup>32</sup>P dCTP incorporation were measured by passing a Geiger



counter over dried gels. Gels were placed in 24 cm X 30 cm autoradiography hypercassettes (Amersham, Australia) with Cronex 4 film (Du Pont). Films were exposed at -80°C for between two hours to several days with a single intensifying screen. Exposure time was determined by levels of  $\alpha$ -<sup>32</sup>P dCTP.

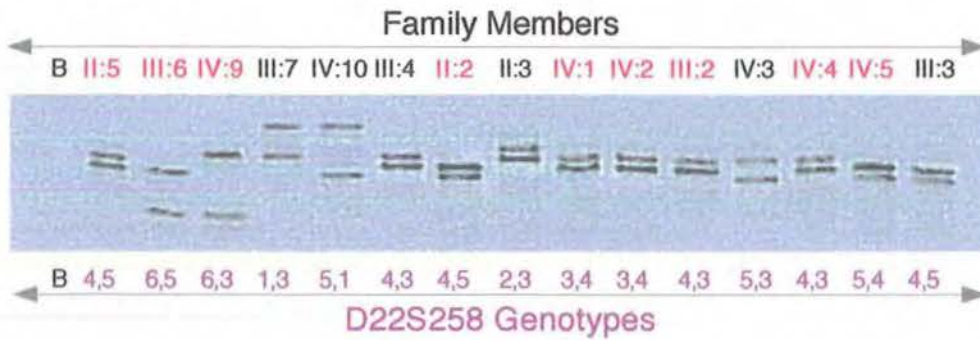
5  $\mu$ l aliquots of the 4 non-microsatellite PCR products (APOB, AT3, HD, D17S30) were electrophoresed in 2% agarose. *Bio-Rad* Mini-Sub Cell or Wide Mini-Sub Cell were used to run gels. Electrophoresis power was supplied by a *Bio-Rad* PowerPac 1000. Gels were stained with ethidium bromide (EtBr) (1  $\mu$ g/ $\mu$ l); viewed with a UVTM-40 Mighty Bright UV Transilluminator (Hoeffer); and photographed with a DS-34 fixed focal length Polaroid camera loaded with Polaroid 667 film.

Figure 4.3 (p. 107) shows the microsatellite D22S258 gel and the VNTR, D17S26 gel. Screening the WAMPD genome ran from chromosome 22 to chromosome 1, so the 2 markers in Figure 4.3 were among the first markers tested. Figure 4.4 (p. 108) shows the microsatellite D4S174 gel and the VNTR D2S44 gel, which were among the last markers to be tested. All autoradiograph gel images in figures throughout this thesis were enlarged horizontally to accommodate pedigree numbers and alleles.

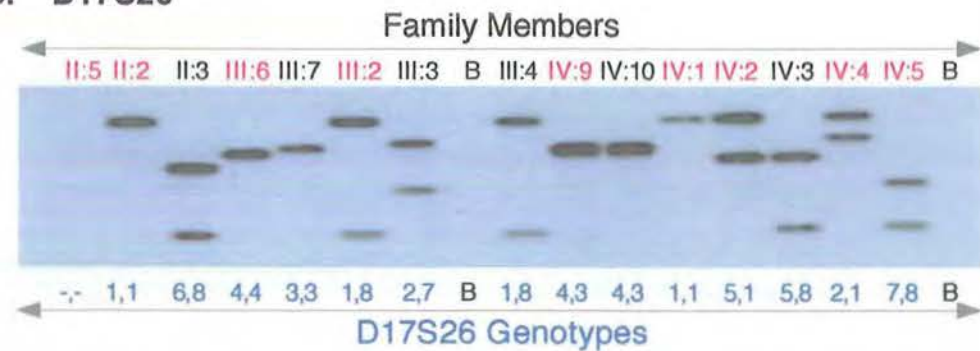
#### **4.3.2 Two-point Linkage Analysis**

Two-point linkage analysis was carried out for all family members on these 97 marker loci using the computer program LIPED (Ott, 1974) to apply the method of maximum likelihood and lod scores (Morton, 1955). All family members were included in this analysis because the pedigree indicated high penetrance of this form of distal myopathy (Figure 4.1, p. 96).

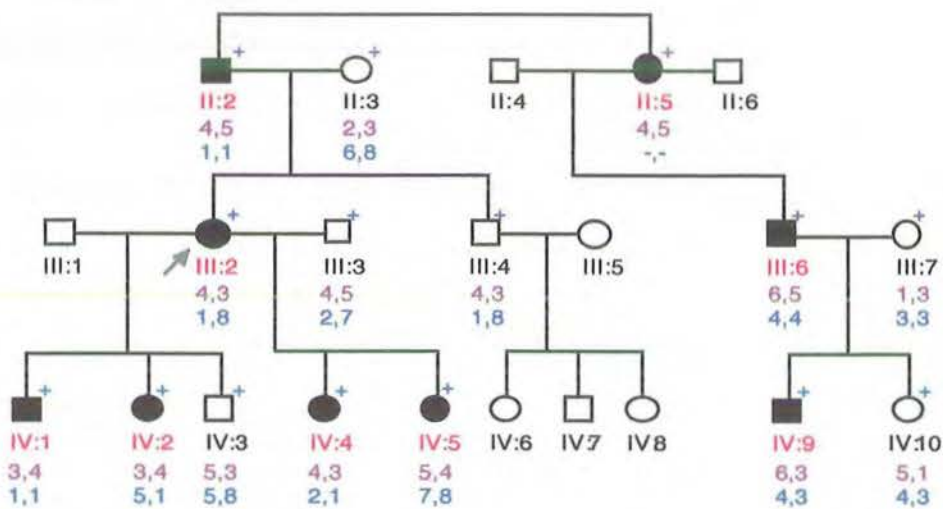
### A. D22S258



### B. D17S26

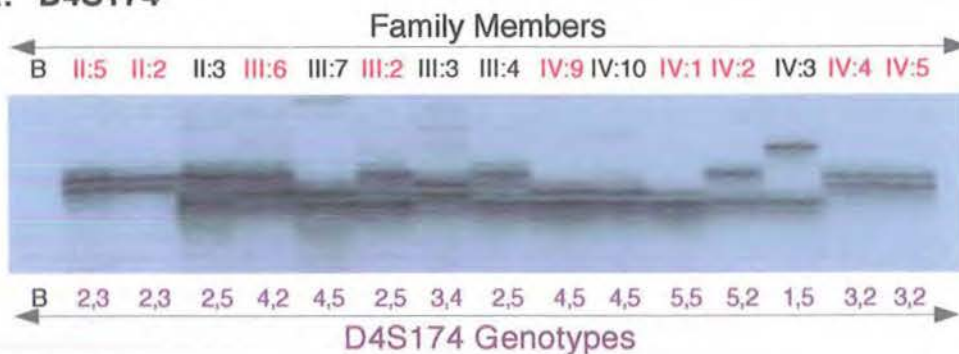


### C. WAMPD Pedigree

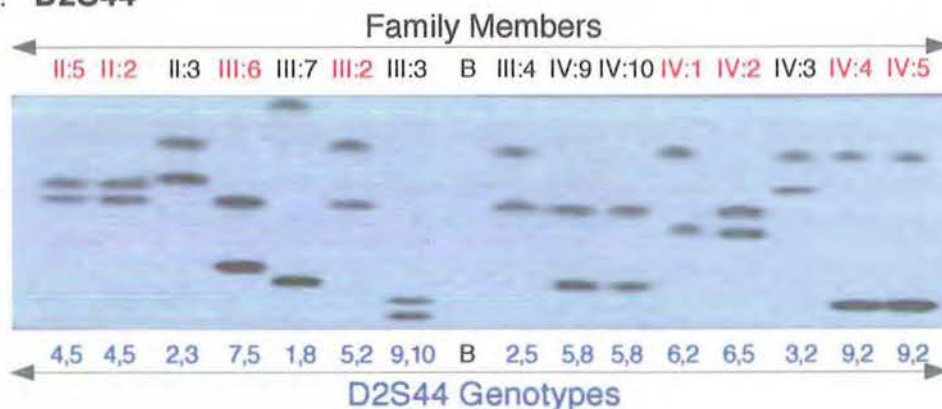


**Figure 4.3** Microsatellite **D22S258** and VNTR **D17S26** segregation in the WAMPD pedigree. **A.** **D22S258** autoradiograph. **B.** **D17S26** autoradiograph. Both **A** and **B** incorporated  $\alpha$ - $^{32}\text{P}$  dCTP and were electrophoresed on 6% acrylamide gels. Red numbers indicate affected individuals; B means 'blank'. On the **D17S26** gel and the WAMPD pedigree, the genotype -,- indicates a failed marker. **C.** The segregation of **D22S258** and **D17S26** genotypes in the 15 participating (+) WAMPD pedigree members. Where possible, paternal alleles are shown on the left.

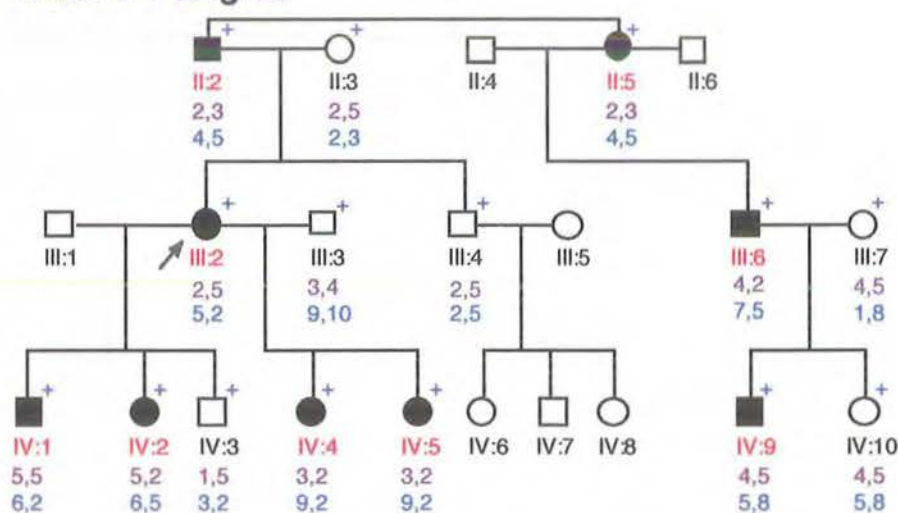
### A. D4S174



### B. D2S44



### C. WAMPD Pedigree



**Figure 4.4** Microsatellite D4S174 and VNTR D2S44 segregation in the WAMPD pedigree. **A.** D4S174 autoradiograph showing (CA)<sub>n</sub> replication slippage 'stutter bands'. These can make the results for dinucleotide repeat markers hard to read. **B.** D2S44 autoradiograph. **A** and **B** incorporated  $\alpha$ -<sup>32</sup>P dCTP and were electrophoresed on 6% acrylamide gels. Red numbers indicate affected individuals; B means 'Blank'. **C.** The segregation of D4S174 and D2S44 genotypes in 15 participating (+) WAMPD pedigree members. Where possible, paternal alleles are shown on the left.

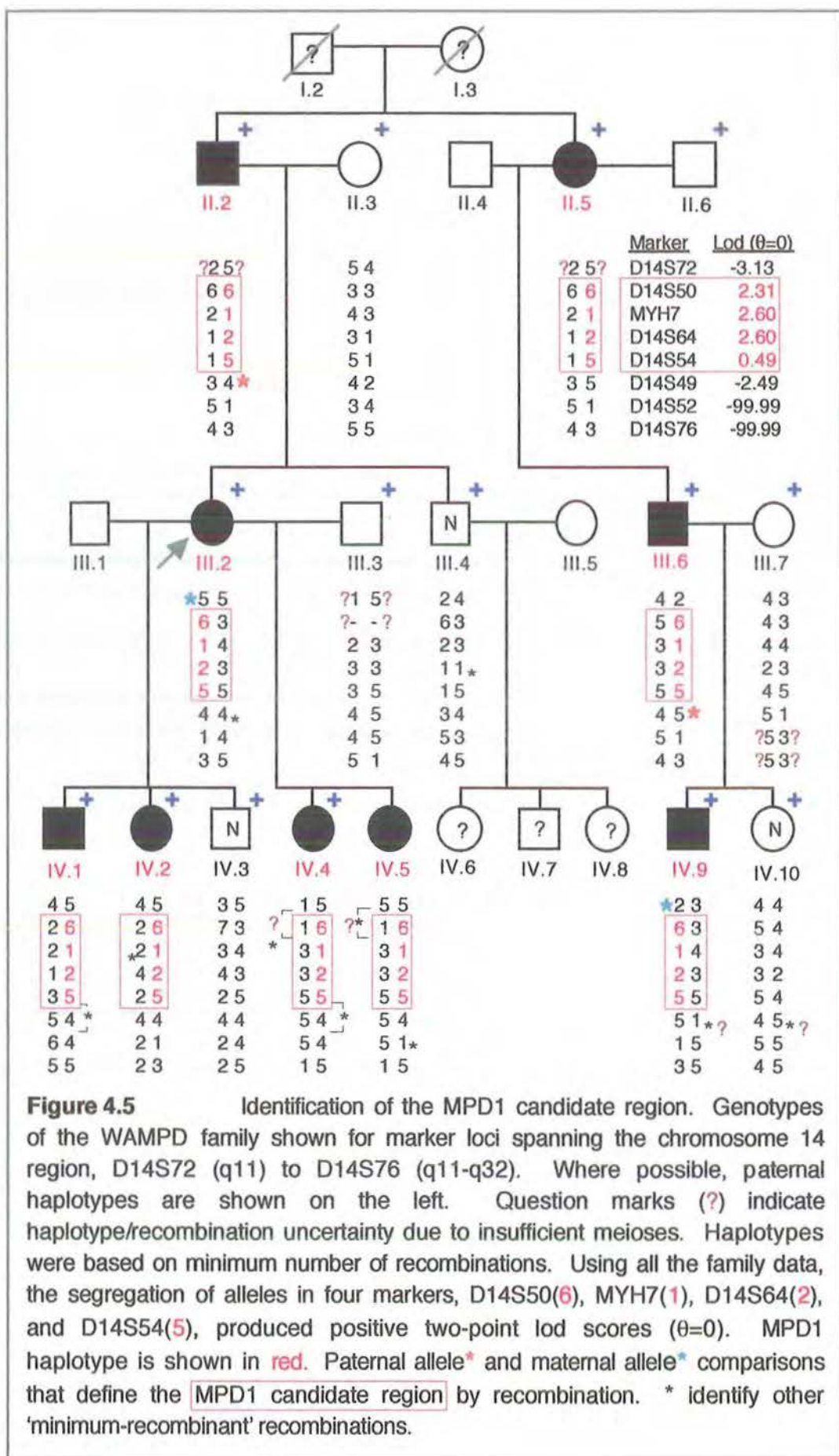


The chromosomal positions of loci were estimated from NIH/CEPH Collaborative Mapping Group (1992) linkage maps, and data from the 'CHLC REPORT' (Murray et al., 1993 & 1994). Two markers on chromosome 8 gave positive two-point lod scores of over 1.00. D8S84 gave a lod score of 1.11 ( $\theta = 0.00$ ) and D8S200 a lod score of 1.09 ( $\theta = 0.00$ ). These two markers were, however, only informative in one branch of the family and other markers on chromosome 8, D8S166 and D8S198, showed multiple recombinants for both branches of the family. Thus, significantly negative lod scores were obtained for these markers, making chromosome 8 unlikely as the site of this distal myopathy disease locus.

More strongly positive two-point lod scores were obtained with markers on chromosome 14. D14S50, MYH7 and D14S64 showed no recombinants and yielded respective two-point lod scores of 2.31, 2.60 and 2.60. The cardiac myosin beta chain gene, *MYH7*, has two microsatellite markers, MYH7.1 and MYH7.2. These microsatellites were informative for different parts of the pedigree, so MYH7.1 and MYH7.2 were haplotyped to obtain the two-point lod score of 2.60 for the 'MYH7 microsatellite'. The family structure precludes a two-point lod score of 3 or greater. Table B4.1 in Appendix B shows the two-point LIPED generated lod scores for all family members.

#### **4.3.3 Marker Haplotypes from D14S72 (q11) to D14S76 (q11-q32) for the WAMPD Family**

Figure 4.5 (p. 110) shows the WAMPD family's marker haplotypes from D14S72 (q11) to D14S76 (q11-q32) requiring the least number of recombinations. Paternal haplotypes are shown on the left. Because DNA from I:2 and I:3 was not available, phase was uncertain for II:2 and II:5.



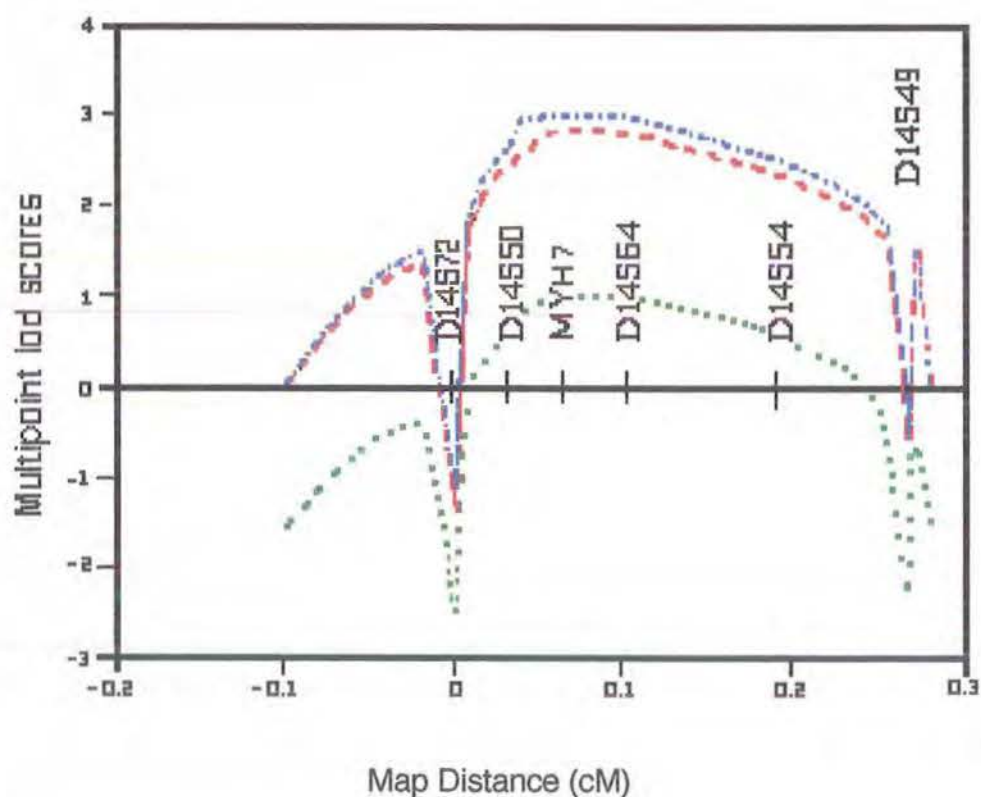


Based on a report from II:2 that his mother, I:3, had the 'family walk' and from her death certificate listing, 'Muscular dystrophy' as a cause of death (dated and signed 22 April 1965, by G.B. Mortlock, Registrar, Adelaide, South Australia), the right-hand haplotypes of II:2 and II:5 are assumed to be from I:3. Phase was also uncertain for individuals II:3, III:3 and III:7 who married into this pedigree.

The co-segregating alleles D14S50 (6), MYH7 (1), D14S64 (2), and D14S54 (5) and the positive two-point lod scores for these markers (shown in Figure 4.5 & Table B4.1, Appendix B), indicated that this haplotype was segregating with the disease. Comparison of the alleles at the centromeric marker D14S72 in III:2 and IV:9 (marked \*), and the alleles at the telomeric marker D14S49 in II:2 and III:6 (marked \*), identify the recombinations that define the candidate region for this form of distal myopathy.

#### 4.3.4 Multipoint Linkage Analysis

Multipoint linkage analysis at 100% 80% and 50% penetrance was performed by Helen Kozman, Department of Cytogenetics and Molecular Genetics, Adelaide Children's Hospital, using the CRIMAP version 2.4 package as previously described by Mulley et al. (1993). All the data for D14S72, D14S50, MYH7, D14S64, D14S54 and D14S49, and assuming 100% penetrance, yielded multipoint lod scores of 3.00 for the 'MYH7 microsatellite' (MYH 7.1 & 7.2 haplotyped) and 2.99 for D14S64. A multipoint lod score of 2.80 was obtained for MYH7 at a penetrance of 80%. These multipoint lod scores, shown in Figure 4.6 (p. 112), along with the lack of significant linkage elsewhere in the genome, confirm the region from D14S72 to D14S49 as the likely location for this distal myopathy disease gene.

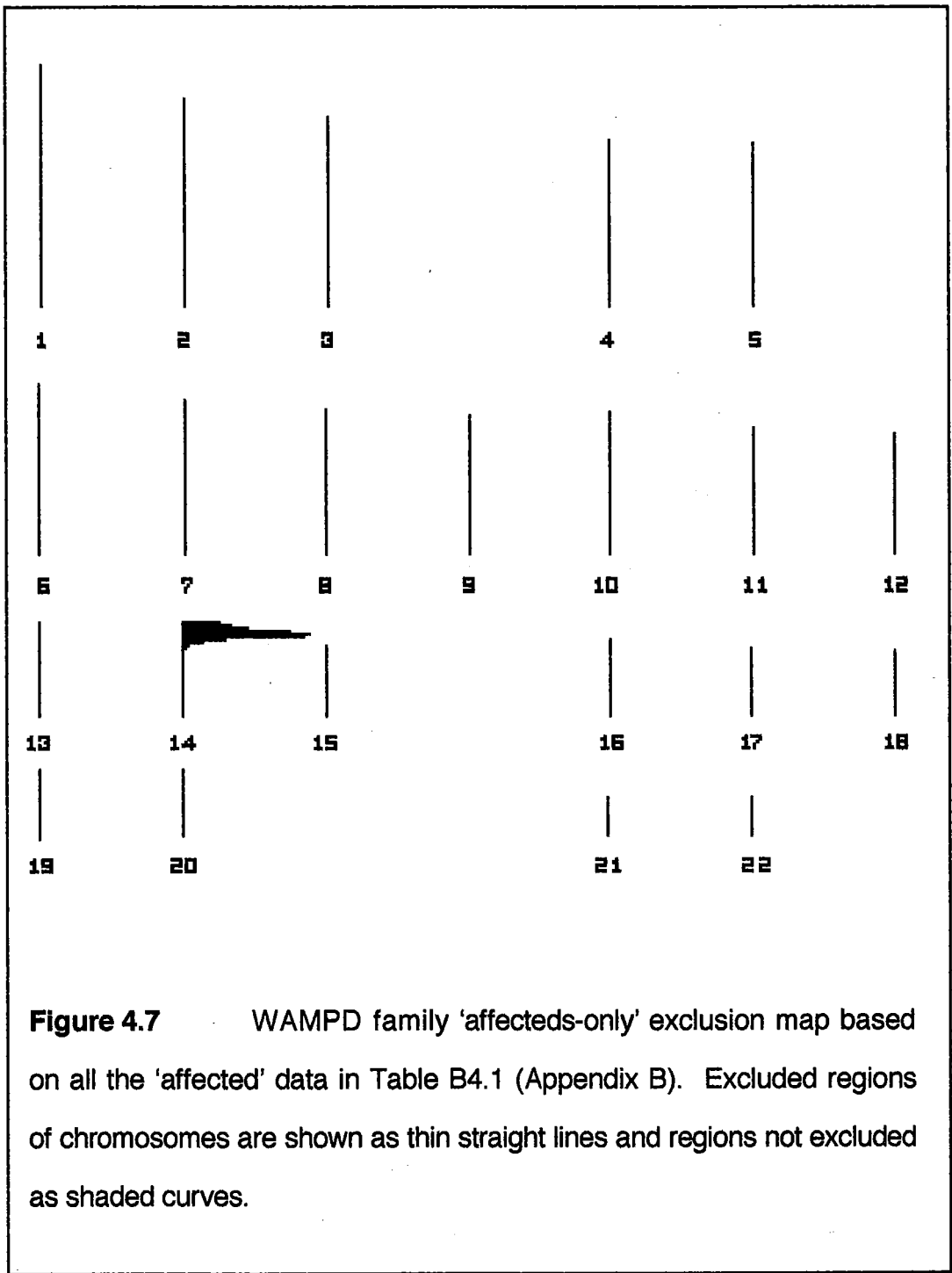


**Figure 4.6** Multipoint analysis for the markers D14S72, D14S50, MYH7, D14S64, D14S54 and D14S49 assuming 100%, 80% and 50% penetrance.

#### 4.2.6 EXCLUDE

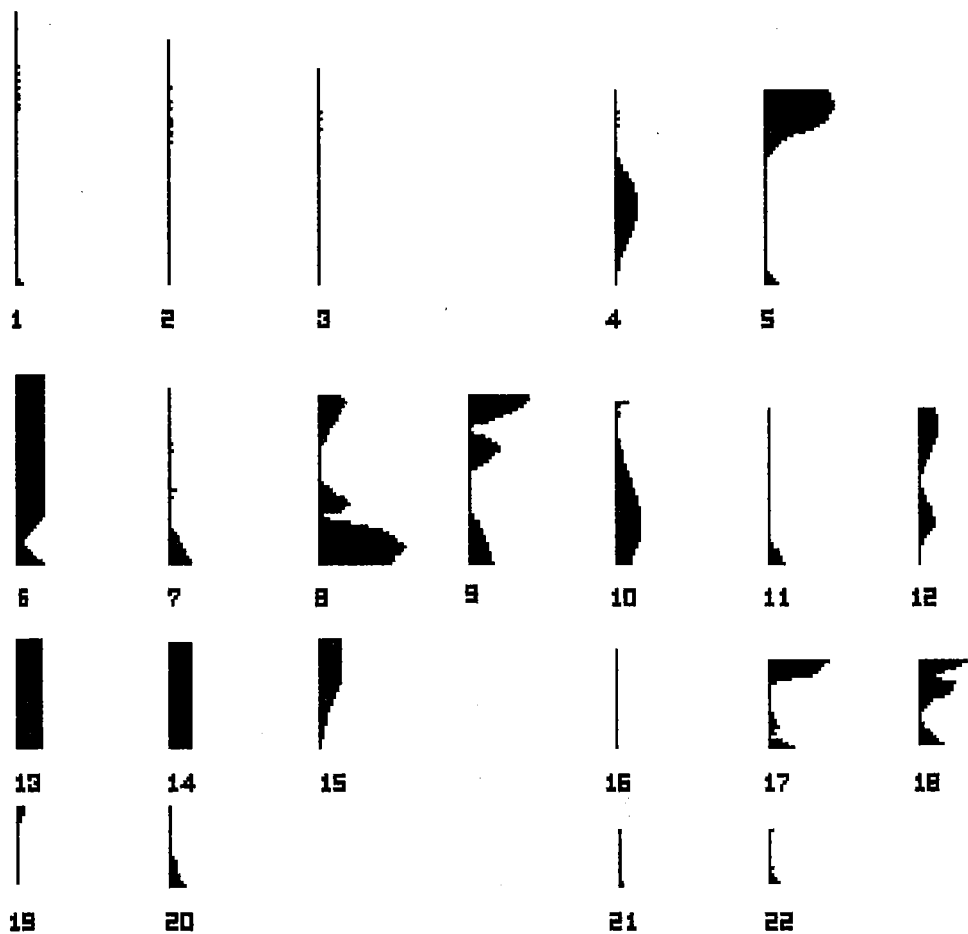
Analysis of the combined LIPED data in Table B4.1 (Appendix B) for all markers and all family members was performed using the EXCLUDE program (Edwards, 1987). This was performed to determine the likelihood that the distal myopathy gene in this family lay on chromosome 14, and to estimate the proportion of the genome excluded as the possible site for the disease gene. Running the EXCLUDE program for the entire data set gave a probability of 99.9% that this distal myopathy disease gene lay on chromosome 14.

Running the EXCLUDE program for the affecteds-only data set, controlled for the possibility of reduced penetrance. It too gave a probability of 99.9% that the gene responsible for the WAMPD AD-distal myopathy lay on chromosome 14. The WAMPD affecteds-only EXCLUDE results are presented in Figure 4.7.





Running the EXCLUDE program for the data set minus the data for chromosome 14 showed, in an affecteds only analysis, that at least 50% of the genome had been excluded as the possible site of the disease gene in this family. This is shown in Figure 4.8.



**Figure 4.8** WAMPD family 'affecteds-only' exclusion map with the chromosome 14 'affected data' removed from the data in Table B4.1 (Appendix B). Excluded regions of chromosomes are shown as thin straight lines and regions not excluded as shaded curves and rectangles.

On the basis of this EXCLUDE data, and the multipoint lod scores, the HUGO Nomenclature Committee reserved MPD1, 'myopathy, distal 1', for this form of distal myopathy. It is now called 'Laing myopathy' (Voit et al., 1998; Barohn et al., 1999). These results have been reported in brief (Meredith et al., 1994a & b) and in full (Laing et al., 1995b).

#### **4.4 REFINEMENT OF THE MPD1 LINKAGE REGION**

At the time of localisation, the sex-averaged distance of the MPD1 linkage interval was estimated to be 28.5 cM from the 'CHLC REPORT' (Murray et al., 1993) and as 31 cM on the CEPH Consortium map (Cox et al., 1995) (Figure 4.9, p. 116). If the estimated 2500 human muscle genes (Paterson & Bishop, 1977) are evenly spread through the 3000 cM human genome, then approximately 25 will be located in the 28.5-31 cM MPD1 candidate region. Refinement of the initial MPD1 candidate region had the potential to reduce the number of muscle candidate genes for Laing myopathy.

##### **4.4.1 New Genetic Markers**

Seven new microsatellite markers within the MPD1 candidate region were identified from the 'CHLC REPORT' (Murray et al., 1994), the 1993-94 Genethon human genetic linkage map (Gyapay et al., 1994) and from GDB (1994; <http://gdbwww.gdb.org/>). One marker, D14S283, was at the centromeric end of the MPD1 candidate region between D14S72 and D14S50. The remaining six markers, D14S264, D14S80, D14S275, D14S262, D14S252, and D14S257 were at the telomeric end of the MPD1 candidate region, between D14S49 and D14S64. These markers were tested against the WAMPD family using the microsatellite protocol described in Section 4.3.1 (pp. 104-106).

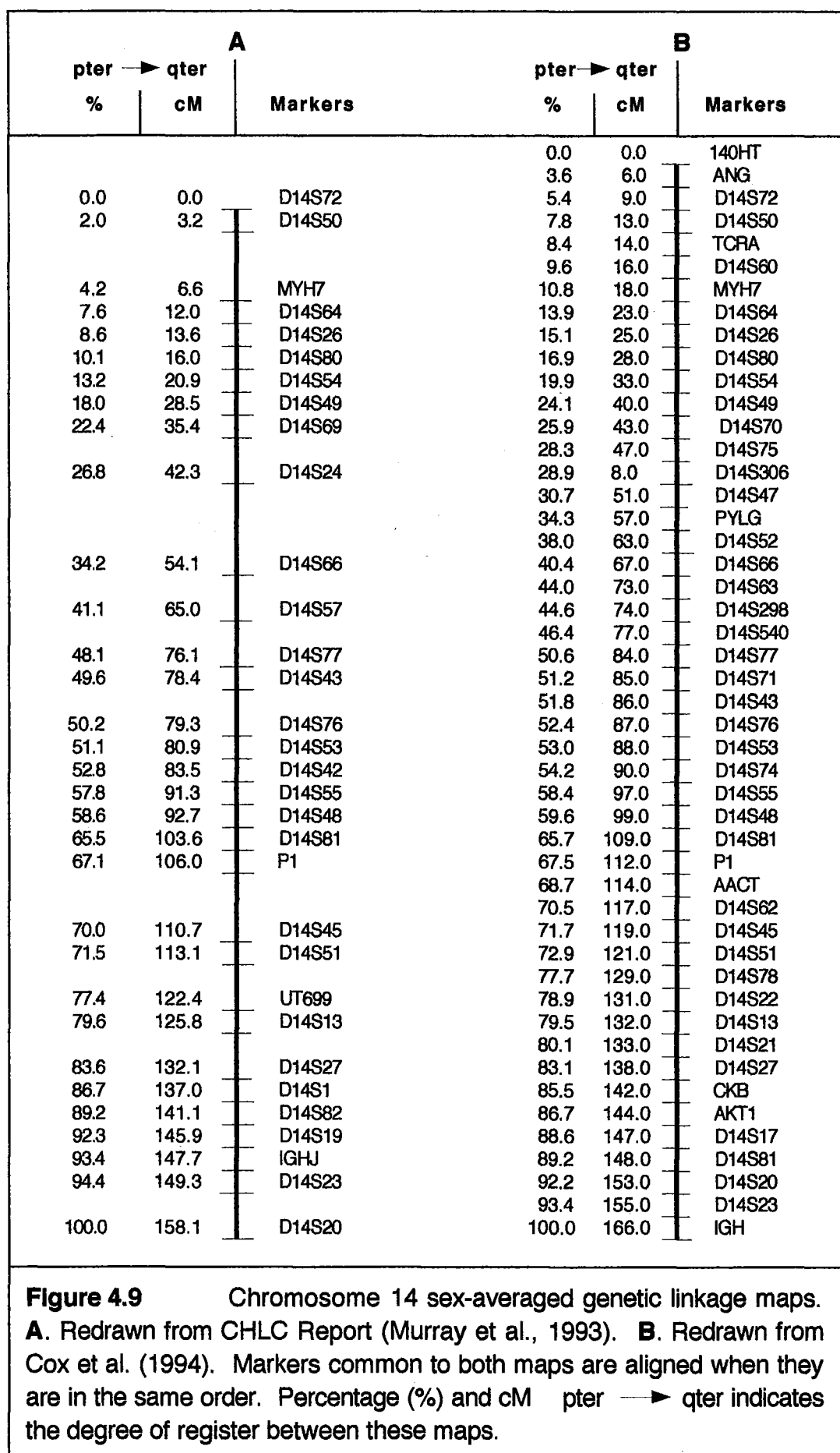


Figure 4.10 (p. 118) shows the electrophoresis autoradiographs for D14S257 (the most distal marker within the MPD1 candidate region) and D14S262 (the most informative marker with the highest two-point lod score of 2.9 at  $\theta=0$ ). Figure 4.11 (p. 119) shows the 2 microsatellites, D14S283 (proximal) and D14S49 (distal), now defining the MPD1 candidate region.

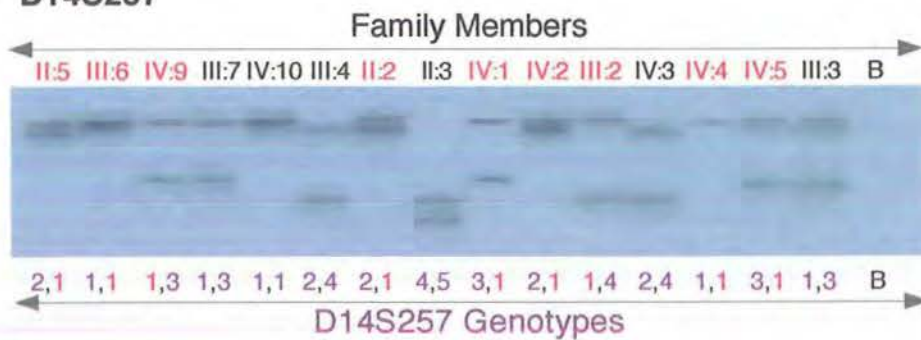
#### 4.4.2 Final Two-point Linkage Analysis

The WAMPD family's marker haplotypes from D14S72 (q11) to D14S76 (q11-q32) are presented in Figure 4.12 (p. 120). Recombination was detected at the centromeric marker D14S283 in III:2 and IV:9. This excluded that portion of the original MPD1 candidate region between D14S72 and D14S283. Linkage analysis with the remaining 6 MPD1 candidate region markers did not reveal further recombination. Figure 4.12 shows the MPD1 disease haplotype as: 6-1-2-1-1-2-1-3-5-1.

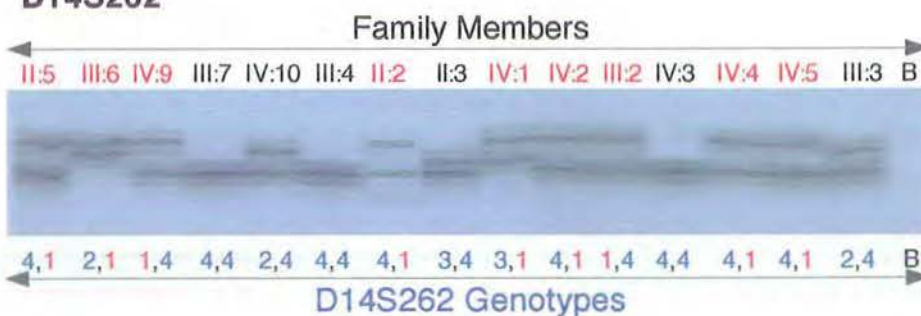
All linkage data available for D14S72 (q11) to D14S76 (q11-q32) is presented in Table 4.3 (p. 121). Microsatellite D14S262 was fully informative and achieved this family's theoretical maximum two-point lod score of 2.9.

The MPD1 candidate region between D14S283 and D14S49 is currently estimated to be 22.4 cM GB4 GeneMap'98 (NCBI, 1999; <http://www.ncbi.nlm.nih.gov/>). Since this map estimates that the original D14S72 to D14S49 MPD1 candidate region was 27.8 cM, running 7 new markers in this interval effected a 5.4 cM reduction of the original MPD1 candidate region.

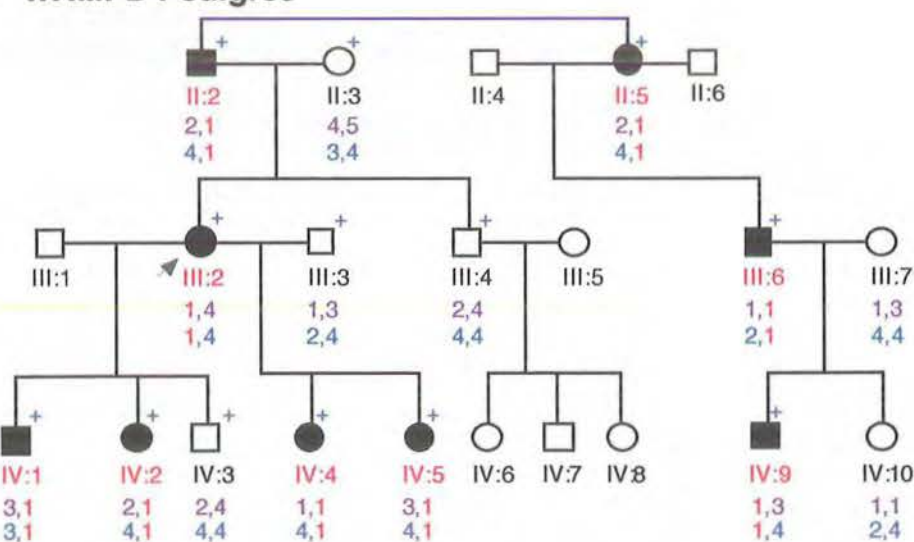
### A. D14S257



### B. D14S262

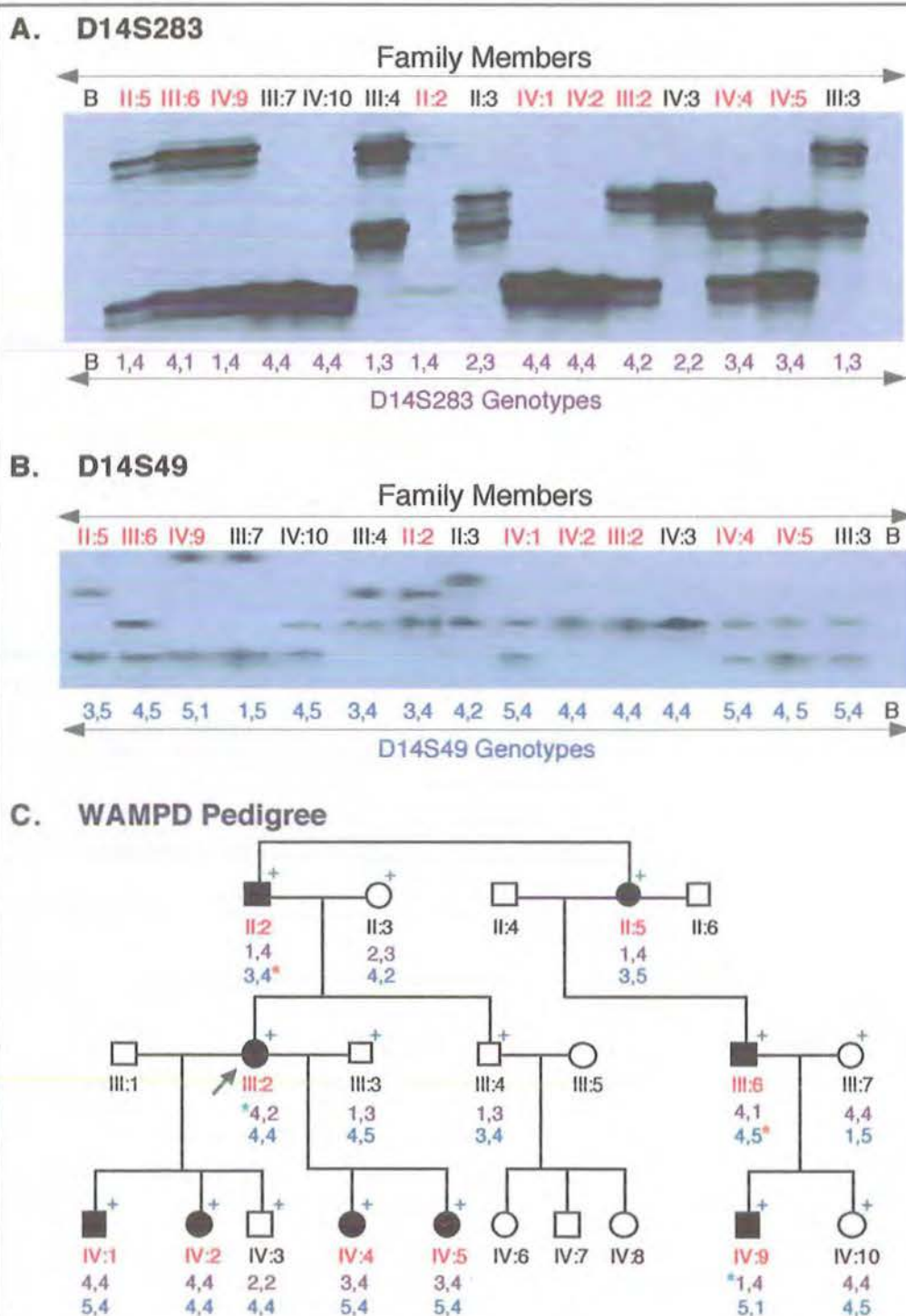


### C. WAMPD Pedigree



**Figure 4.10** Microsatellites D14S257 and D14S262 segregation in the WAMPD pedigree. **A.** D14S257 autoradiograph, the most distal marker within the linkage region. **B.** D14S262 autoradiograph, the most informative marker within the linkage region. Both **A** and **B** incorporated  $\alpha$ - $^{32}$ P dCTP and were electrophoresed on 6% acrylamide gels. Red numbers indicate affected individuals; disease haplotype alleles are printed in red; B means 'Blank'. **C.** Segregation of D14S257 and D14S262 alleles in 15 participating (+) WAMPD pedigree members. Where possible, paternal alleles are shown on the left.



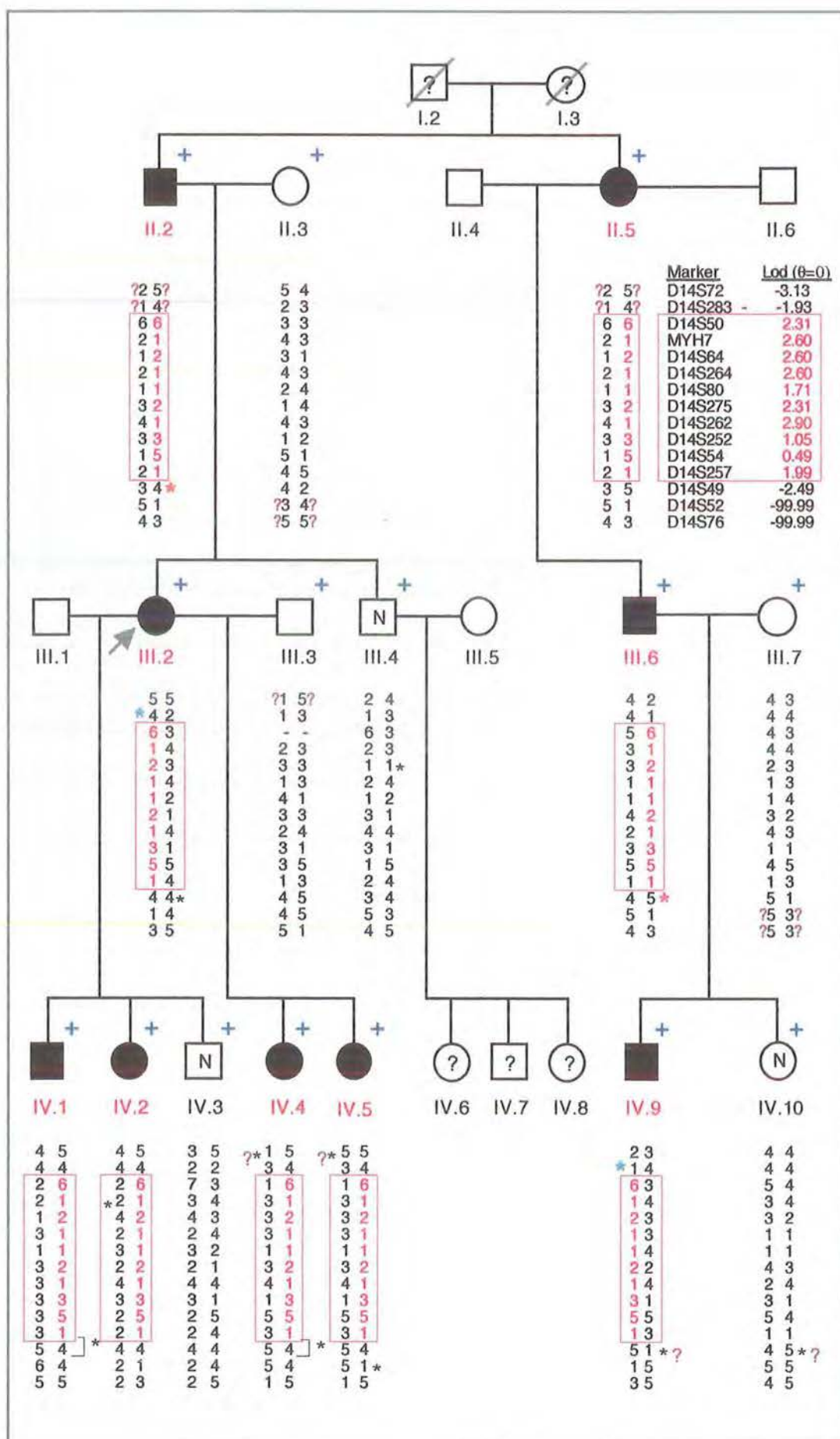


**Figure 4.11** Microsatellite markers defining the MPD1 candidate region. **A.** a deliberately over-exposed **D14S283** autoradiograph to determine **II:2** genotype. Other genotypes determined by shorter exposure. **B.** **D14S49** autoradiograph. **A** and **B** incorporated  $\alpha$ - $^{32}\text{P}$ -dCTP and electrophoresed on 6% acrylamide gels. **Red** numbers indicate affected individuals; B means 'Blank'. **C.** segregation of **D14S283** and **D14S49** alleles in participating (+) WAMPD pedigree members. Paternal alleles are shown on the left. Paternal allele\* and maternal allele\* comparisons that define MPD1 candidate region by recombination.

**Figure 4.12** Refinement of the MPD1 candidate region.

Participating individuals<sup>+</sup>. Genotypes of the Laing myopathy family are shown for marker loci spanning the chromosome 14 region, D14S72 (q11) to D14S76 (q11-q32). Where possible, paternal alleles are shown on the left. Using all the family data, the segregation of alleles in ten markers, D14S50(6), MYH7(1), D14S64(2), D14S264(1), D14S80(1), D14S275(2), D14S262(1), D14S252(3), D14S54(5) and D14S257(1), produced positive two-point lod scores ( $\theta=0$ ). Haplotypes were based on minimum number of recombinations. Question marks (?) indicate haplotype/recombination uncertainty due to insufficient meioses. Paternal allele<sup>\*</sup> and maternal allele<sup>\*</sup> comparisons that define, the MPD1 candidate region by recombination. \* identify other 'minimum-recombinant' recombinations.

'Minimum-recombination' haplotypes are not all the same as shown in Figure 4.5 (p. 110). This happens because more information was available for haplotype construction in Figure 4.10 than was available for this task in Figure 4.5. Individual III.3 exemplifies this effect.





**Table 4.3** *LIPED generated two-point lod scores between the MPD1 disease locus and loci lying within and flanking the MPD1 candidate region on chromosome 14. Analysis includes data from all family members.*

Chromosome		Recombination Fraction ( $\theta$ )							
%pter-qter	Locus	0.00	0.001	0.01	0.05	0.1	0.20	0.3	0.4
14.08	D14S72	-3.13	-1.56	-0.59	0.030	0.23	0.31	0.26	0.15
14.10	D14S283	-1.93	-0.35	0.60	1.14	1.24	1.11	0.81	0.43
14.10	D14S50	2.31	2.31	2.27	2.12	1.93	1.51	1.06	0.57
14.11	MYH7	2.60	2.59	2.55	2.36	2.12	1.61	1.08	0.53
14.15	D14S64	2.60	2.59	2.55	2.36	2.12	1.61	1.08	0.53
14.15	D14S264	2.60	2.59	2.55	2.36	2.12	1.61	1.08	0.53
14.17	D14S80	1.71	1.71	1.68	1.57	1.42	1.11	0.77	0.41
14.18	D14S275	2.31	2.31	2.27	2.12	1.93	1.51	1.06	0.57
14.18	D14S262	2.90	2.89	2.85	2.64	2.38	1.82	1.22	0.61
14.18	D14S252	1.05	1.05	1.03	0.95	0.86	0.65	0.45	0.23
14.20	D14S54	0.49	0.49	0.47	0.41	0.34	0.21	0.1	0.03
14.20	D14S257	1.99	1.99	1.96	1.81	1.61	1.21	0.79	0.37
14.22	D14S49	-2.49	-1.25	-0.30	0.30	0.47	0.49	0.36	0.19
14.37	D14S52	-99.99	-3.41	-1.45	-0.22	0.17	0.32	0.23	0.10
14.50	D14S76	-99.99	-6.39	-3.42	-1.46	-0.72	-0.17	0.01	0.05
14.52	D14S53	-99.99	-3.31	-1.38	-0.14	0.27	0.49	0.44	0.27
14.54	D14S74	-99.99	-99.99	-0.02	0.55	0.69	0.66	0.52	0.30
14.58	D14S48	-99.99	-0.47	0.49	1.00	1.06	0.87	0.56	0.26
14.63	D14S81	0.20	0.20	0.20	0.15	0.11	0.04	0.01	0.00
14.69	D14S45	-99.99	-2.20	-1.21	-0.57	-0.34	-0.15	-0.07	-0.02
14.70	D14S51	-99.99	-3.77	-1.80	-0.54	-0.10	0.15	0.16	0.08
14.78	D14S13	-99.99	-1.61	-0.63	-0.03	0.15	0.20	0.13	0.04

%pter-qter originally determined from from NIH/CEPH Collaborative Mapping Group (1992), then progressively updated from CHLC Reports (Murray et al., 1993 & '94), Cox et al. (1994) and "The National Centre For Biotechnology Information" GB4 GeneMap'98 (<http://www.ncbi.nlm.nih.gov/>)

#### **4.5 REPEAT EXPANSION DETECTION**

In 1994, a fellow doctoral candidate, David Chandler, visited Dr. Martin Schalling's Clinical Genetics Laboratory at the Karolinska Hospital in Stockholm, to receive instruction in the RED technique, first published by Schalling et al. (1993). With purified DNA provided by this doctoral candidate, the WAMPD family was screened in Schalling's laboratory for all possible 10 trinucleotide repeat expansion motifs. All were excluded. A CAG expansion was detected but it was present in all 15 family members (personal communication, David Chandler, 1994).

#### **4.6 CONCLUSION**

The distal myopathy segregating in this Western Australian family does not fit the distal myopathy nosologies proposed by Markesbery and Griggs (1986, p. 1322) or by Barohn et al. (1991, p. 1369) (Table 4.1, p. 94). Under these nosologies, the earliest onset of distal myopathy was, 'Early-Adult Onset Types I and II'. In the WAMPD family, onset has been detected as early as age 4 (IV:5, Figure 4.1, p. 96). Also, in the Barohn et al. (1991) distal myopathy nosology, both early-adult onset types were autosomal recessive whereas the WAMPD family displays AD inheritance. Furthermore, the clinical pattern and rate of progression of Laing myopathy was unique, showing elements of both 'Late-Adult Onset Types I and II' and 'Early-Adult Onset Type I' of the Barohn nosology. Chapters 4 and 9 also consider distal myopathy nosology.

Indications that the onset of this form of distal myopathy is earlier with the passage of generations were discussed. Six possible causes were noted, two relating to ascertainment bias and four other causes. Only one of these possible causes has been excluded; anticipation caused by an

unstable trinucleotide repeat expansion. The apparent earlier onset of distal myopathy in generations III and IV shown in Figure 4.1 (p. 96) remains unexplained.

The linkage of MPD1 to a 27.8 cM interval at 14q11.2 was the first localisation of a distal myopathy (Meredith et al., 1994; Laing et al., 1995b). The identification of recombination(s) that would significantly reduce the MPD1 candidate region, required more affected WAMPD family members or more families segregating Laing myopathy. Such family members or families were not available. Consequently, only small refinements of MPD1 were possible. These were entirely dependent upon the generation of new markers between the two proximal and two distal candidate region markers, D14S283 and D14S50, and D14S49 and D14S257 respectively.

The refinement of the MPD1 region was achieved by running 7 new microsatellite markers within the MPD1 candidate region. These new markers reduced the size of this candidate region to 22.4 cM, a 5.4 cM reduction of the original MPD1 candidate region. This was equivalent to a reduction of approximately 5 muscle-specific MPD1 candidate genes (Paterson & Bishop 1977), from about 25 genes to about 20 genes.

The linkage analysis performed on this family segregating Laing myopathy had approached its limit of refinement.

---

#### FOOTNOTE

As distal myopathy linkage regions were identified, medical scientists at the Department of Neuropathology, Royal Perth Hospital, tested the WAMPD family with markers located within these regions: the chromosome 2 Miyoshi candidate region; the chromosome 9 Nonaka myopathy candidate region. The Miyoshi and Nonaka candidate region markers are listed in Chapter 5 (p. 152). In addition, they tested the WAMPD family with markers within the chromosome 5q31 candidate region of an AD-distal myopathy with vocal cord and pharyngeal weakness (AD-VCPMD) (MPD2) (Feit et al., 1998). The chromosome 5 markers were: D5S1995, D5S658, and D5S1972. All of the Miyoshi, Nonaka and Feit candidate region markers showed recombination in the WAMPD family. All generated significantly negative lod scores. The MPD1 gene did not reside in any of these distal myopathy candidate regions, strengthening its localisation to chromosome 14q.

## **CHAPTER FIVE**

### **EXCLUSION OF THE MPD1 CANDIDATE REGION FOR TWO OTHER AD-DISTAL MYOPATHIES**

## **5.0 EXCLUSION OF THE MPD1 CANDIDATE REGION FOR TWO OTHER AD-DISTAL MYOPATHIES**

The MPD1 candidate region could be further refined by testing for linkage in other pedigrees segregating Laing myopathy. Refinement would occur if recombination was detected within the MPD1 candidate region in affected individuals. There was a practical difficulty with this approach: identifying sufficient pedigrees segregating Laing myopathy. Since linkage detection power in pedigree analysis is good, but map resolving power is low, several large pedigrees of this type would be required to significantly reduce the MPD1 candidate region.

Laing et al. (1995b, p. 422) described Laing myopathy as quite distinct from the better known distal myopathies found in Scandinavia and Japan, "with phenotypic features closely resembling those described in Gowers's [sic] original report (Gowers 1902)". Likening Laing myopathy's clinical phenotype to a 1902 report of this disease, and only to this 1902 report, suggested that Laing myopathy was a rare form of distal myopathy. The best hope of refining the MPD1 candidate region would be to discover sufficient Laing myopathy pedigrees, preferably through a founder effect some 200-300 years ago followed by restricted emigration.

Although geographical isolation restricts emigration from Western Australia, colonisation of this state did not begin until 1829 and the Laing myopathy family moved to Western Australia during the early years of the twentieth century. The excellent medical records of this state did not contain evidence of further individuals in Western Australia suffering from Laing myopathy. The possibility of refining the MPD1 candidate region with other Laing myopathy Western Australian pedigrees appeared unlikely.

In addition, the apparent rareness of Laing myopathy made the task of identifying and accessing Laing myopathy families living in other parts of the world a formidable task. As briefly noted in Section 4.6 (p. 123), when the MPD1 candidate region was first identified we were not aware of any descriptions of other families segregating distal myopathy phenotypically similar to Laing myopathy.

The genetic testing of further distal myopathy pedigrees was, therefore, restricted to two outcomes. First, if distal myopathy pedigree(s) containing sufficient affected individuals became available for genetic testing, then the discovery of novel distal myopathy candidate regions would be possible. Second, if distal myopathy pedigree(s) became available for genetic testing but lacked sufficient affected members for disease gene localisation, then the exclusion of the MPD1 candidate region might be possible. Either way, it was possible that the genetic heterogeneity of the distal myopathies could be established.

Two further families segregating AD-distal myopathy were sent to us for analysis once the MPD1 localisation was published. One of these families resided in Queensland, Australia, the QMPD family, the other in Connecticut, USA, the CMPD family. Neither pedigree had sufficient affected members to establish linkage with the universally accepted  $p=0.05$  threshold of significance, or odds of 20:1 for the joint probability of linkage versus non-linkage.

The genetic testing of these families was, therefore, restricted to the possibility of establishing genetic heterogeneity for the distal myopathies.



## 5.1 AUTOSOMAL DOMINANT DISTAL MYOPATHY SEGREGATING IN A QUEENSLAND FAMILY

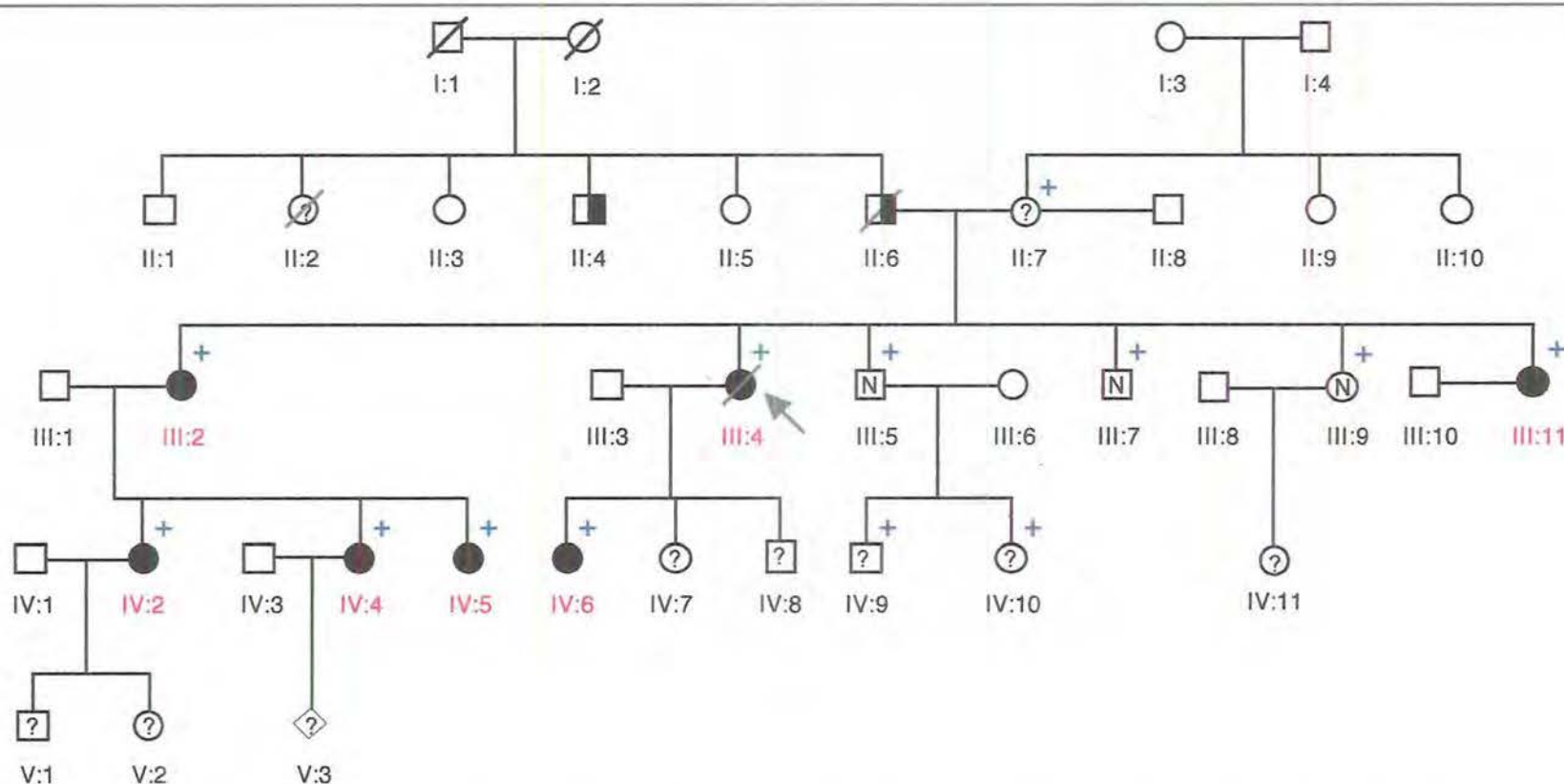
Seven affected and six unaffected members of the QMPD family agreed to participate in genetic testing (Figure 5.1, p. 128). The QMPD pedigree indicated that this family was segregating an AD form of distal myopathy of relatively high penetrance. Those participating members of the QMPD family were ascertained and genetically tested as a part of this distal myopathy exclusion project.

### 5.1.1 Ascertainment

Clinical features of the QMPD family were ascertained by Dr. Peter Silburn, Department of Neurology, Princess Alexandra Hospital, Brisbane, Queensland, Australia. Onset was in the second decade, always being clinically apparent by the time affected individuals reached the early part of their third decade.

In the seven affected individuals, there was a fixed pattern of cramps then weaknesses, progressing: deltoid weakness, weak dorsiflexion of wrists; mild weaknesses of finger flexion; and variable weakness of facial musculature. In addition, four of the seven affected individuals, III:4, IV:2, IV:4, and IV:6, spanning the age range 23-46 years, showed signs of cardiomyopathy. Variable expressivity was also apparent in other ways.

The proband, III:4, who had a history of progressive weakness and lethargy, presented with recent onset of dysphagia for solids and liquids. She was unaware of any other specific muscle problems. Examination revealed marked bilateral wasting of the hands and feet and sternocleido-mastoids.



**Figure 5.1** Pedigree of a Queensland family (QMPD) segregating distal myopathy. Squares denote males, circles females, lines through squares or circles indicate 'deceased', and diamonds individuals of unknown sex. Shaded symbols and red numbers indicate individuals affected by distal myopathy, half-shaded symbols show individuals who, by history, were 'probably affected' by distal myopathy. An 'N' in the centre of the symbol denotes 'unaffected by distal myopathy', while a '?' in the centre of the symbol denotes 'unknown disease status'. The proband (III:4) is indicated by an arrow. Each individual marked with a cross (+) had an immortalised lymphocyte cell line established and contributed DNA for exclusion analysis. AD inheritance of relatively high penetrance was indicated.



Mild weakness was present in wrist dorsiflexion and finger extension but with preservation of finger flexion. There was also: mild weakness of dorsiflexion of the feet, with weakness of plantar and mid-hip flexion; mild weakness of shoulder abduction and neck flexion and severe symmetrical facial weakness; depressed deep tendon reflexes in the upper limbs and absent in the lower limbs. Plantar responses were flexor.

CK levels were normal on multiple occasions. Nerve conduction studies of sensory and motor nerves were also normal. EMG, particularly of the first dorsal interosseous and extensor digitorum brevis (EDB), demonstrated small amplitude, polyphasic units with an early full interference pattern. There was no spontaneous activity and no myotonia was present.

Proband biopsy samples were obtained from the EDB and deltoid muscle. Histologically, EDB showed intermixed hypertrophic and atrophic fibres with rounded contours and increased numbers of central nuclei. There was preservation of the normal proportion and distribution of fibre types. No angulated fibres or other changes of denervation were seen. The interstitium showed increased fibrous tissue and fatty infiltration. There were no RVs or inclusion bodies and no significant inflammation. Deltoid muscle appeared normal. The pathologists, Dr. R.J. Conrad and Dr. A.E.G. Tannenberg (also of the Department of Neurology, Princess Alexandra Hospital, Brisbane) concluded that the histology of the EDB muscle was consistent with a diagnosis of, "Non-Scandinavian distal myopathy, late-adult-onset type" (personal communication, 1994).

A transthoracic echocardiogram (TTE) of the proband demonstrated marked left ventricular dilation with septal thinning and hypokinesis (reduced movement). Atria were normal. Cardiac conduction defects were not found. Unfortunately

this patient died of multiple organ failure in an intensive care unit, due to a self-administered overdose of digoxin and an angiotension converting enzyme inhibitor. Post-mortem biopsy samples were obtained from the quadriceps, gastrocnemius, biceps brachialis, and EDB. Biopsy samples were not obtained for the heart.

In the quadriceps, type II muscle fibre atrophy was present. A rare atrophic fibre with small RVs was also present. In the biceps, the findings were similar although more atrophic fibres containing RVs were present.

In the gastrocnemius, there was increased variation of muscle fibre diameter with hypertrophied, normal and atrophic fibres in a haphazard distribution. Interstitial fibrosis was present around some severely atrophic fibres and the atrophic fibres were of both types. A rare fibre showing muscle fibre splitting was present. No inflammation was seen and there was no evidence of denervation.

Severe signs of myopathy were seen in EDB and foot muscles showing fibrous and fatty replacement, more numerous RVs, small cytoplasmic inclusion bodies and internal nuclei. Immunoperoxidase staining for desmin bodies was negative and ragged red fibres were not seen. as previously noted (p 102), desmin bodies have been associated with AD-distal myopathy, particularly when distal myopathy was associated with cardiomyopathy (Horowitz & Schmalbruch, 1994).

For individuals, IV:2, IV:4, IV:5 and IV:6, EMG confirmed the myopathic pattern described for the proband III.4 with early, full interference pattern on minimal effort. TTE also showed that cardiomyopathy was significantly present in IV:2, revealing a thin hypokinetic septum, left ventricular dilatation and moderate left

ventricular dysfunction. The remainder of the heart of IV:2 was preserved. Finally, TTE revealed mild septal hypokinesis and mild reduction in left ventricular function in IV:4 and mild mid-basal septal hypokinesis in IV:6.

Atypical of other affected family members, the proband III:4 also manifested possible myopathy of external ocular muscles as determined by slow saccadic eye movements in the vertical and horizontal planes, and incomplete saccades in the horizontal plane. CNS disturbances would have to be excluded before these abnormal saccadic movements could be taken as clear evidence of ocular myopathy. Swallow assessment with video fluoroscopy demonstrated reduced propulsion in the pharynx and oesophagus with decreased peristalsis consistent with muscular weakness. These atypical symptoms of the proband are also associated with OPMD (Victor et al., 1962) and oculopharyngodistal myopathy (Satoyoshi & Kinoshita, 1977).

In this thesis, the AD-distal myopathy segregating in the QMPD family is described as 'Silburn myopathy'.

The publication of the 14q11.2-q13, OPMD candidate region spanning the markers D14S283-D14S50-MYH7.1 for a French-Canadian form of oculopharyngeal MD (OPMD) (Brais et al., 1995), revealed an overlap of this candidate region with the D14S283-D14S49 MPD1 candidate region (Laing et al. 1995b). This report, and phenotype overlap between Silburn myopathy and OMPD, strengthened the perception that a form or forms of distal myopathy, OPMD or oculopharyngodistal myopathy might belong to allelic series. Therefore, the OPMD-like symptoms of the QMPD proband, III:4, effectively increased the incentive to conduct the Silburn myopathy MPD1 exclusion project.

### **5.1.2 Genetic Markers**

Blood samples were taken and immortalised cell lines were established for the 13 family members who volunteered to participate in the study. DNA extraction was identical to that described in Section 4.2.2 (p. 102). The segregation of 11 chromosome 14 genetic markers was determined for the QMPD family: D14S72, MYH7.1, D14S64, D14S80, D14S54, D14S49, D14S76, D14S53, D14S48, D14S45 and D14S51. Primer synthesis,  $\alpha$ -<sup>32</sup>P dCTP labelling, PCR conditions, thermal cyclers used, electrophoresis and gel exposure to film were described in Section 4.3.1 (pp. 104-106).

Figure 5.2 (p. 133) shows an MPD1 candidate region microsatellite marker, D14S64, as tested on the Silburn myopathy family.

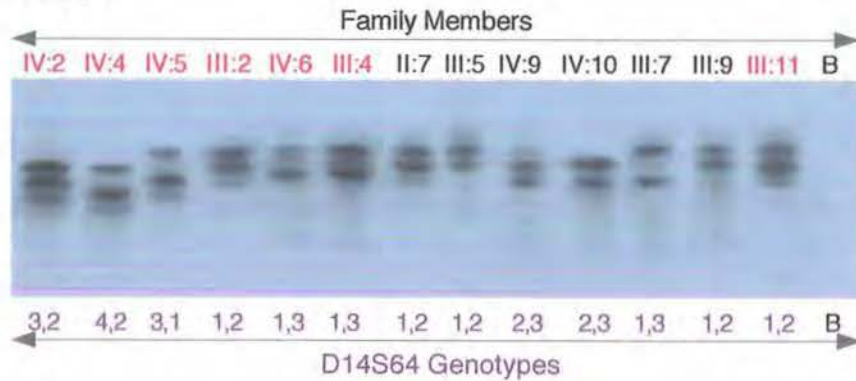
### **5.1.3 Chromosome 14 Two-point Linkage Analysis**

Two-point linkage analysis was carried out for the QMPD family as described for the WAMPD family in Section 4.3.2 (pp. 106-109). Table 5.1 (p. 134) shows QMPD 'affecteds-only' two-point LIPED (Ott, 1974) lod scores for all of the tested chromosome 14 markers. An 'affecteds-only' analysis was adopted to control for reduced penetrance. The maximum possible 'affecteds-only' two-point lod score for this family was 1.5.

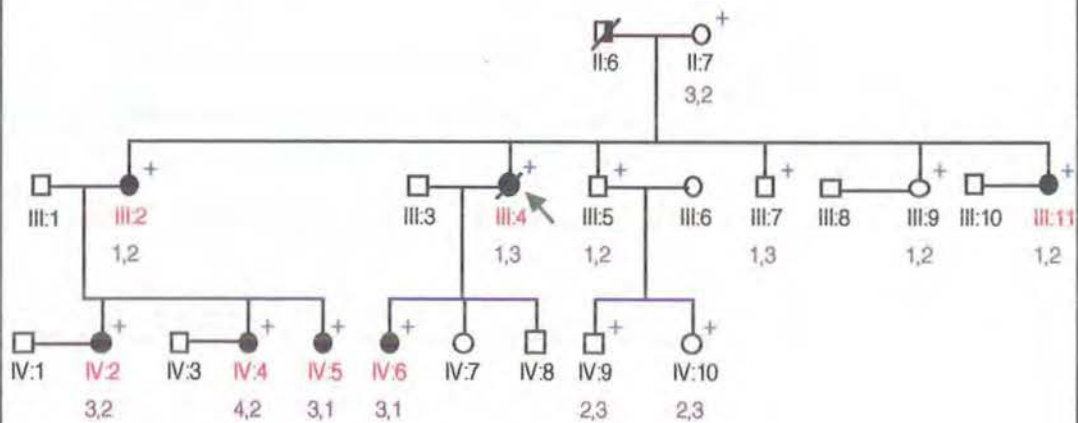
Significantly negative lod scores ( $<-2.0$  at  $\theta=0$ ) were obtained for the genetic markers: D14S72, MYH7.1, D14S64, D14S80, D14S49, D14S53, D14S48, D14S45 and D14S51. Three of these genetic markers, MYH7.1, D14S64, D14S80, were within the MPD1 candidate region.



**A. D14S64**



**B. QMPD Pedigree**



**Figure 5.2** Segregation of the MPD1 candidate region microsatellite, D14S64, in the QMPD pedigree. **A.** Microsatellite D14S64 autoradiograph. D14S64 incorporated  $\alpha$ - $^{32}\text{P}$  dCTP and was electrophoresed on 6% acrylamide gel. **B.** The segregation of D14S64 alleles in 13 participating (+) QMPD pedigree members. Where possible, paternal alleles are shown on the left.

Comparison of affected genotypes shows why an 'affecteds-only' negative lod score of -99.99 ( $\theta=0$ ) was obtained for co-segregation of D14S64 with the Silburn myopathy disease gene. The affected individuals III:2, III:4, III:11, IV:2, IV:4, IV:5 and IV:6 do not share a common allele. In fact, only one of the four affected grandchildren, IV:5, definitely received a D14S64 allele (1) from their 'probably affected' grandfather, II:6. Grandchild IV:6 may have received this D14S64 allele from II:6.

**Table 5.1** *LIPED generated two-point lod scores between QMPD disease locus and loci lying within and flanking the **MPD1 candidate region** on chromosome 14. Analysis includes data from affected individuals only.*

Chromosome		Recombination Fraction ( $\theta$ )							
%pter-qter	Locus	0.00	0.001	0.01	0.05	0.1	0.20	0.3	0.4
14.08	D14S72	-99.99	-7.45	-4.5	-2.44	-1.6	-0.79	-0.37	-0.13
14.11	MYH7.1	-99.99	-1.99	-1.01	-0.36	-0.12	0.05	0.09	0.07
14.15	D14S64	-99.99	-4.99	-3.01	-1.64	-1.08	-0.55	-0.28	-0.11
14.17	D14S80	-3.50	-2.45	-1.5	-0.83	-0.56	-0.29	-0.14	-0.04
14.20	D14S54	0.13	0.12	0.12	0.09	0.05	0.01	-0.01	-0.01
14.22	D14S49	-99.99	-1.70	-0.71	-0.08	0.13	0.25	0.23	0.15
14.50	D14S76	0.13	0.13	0.12	0.10	0.08	0.05	0.02	0.01
14.52	D14S53	-99.99	-7.69	-4.70	-2.64	-1.78	-0.95	-0.50	-0.21
14.58	D14S48	-99.99	-7.50	-4.51	-2.47	-1.63	-0.88	-0.50	-0.24
14.69	D14S45	-99.99	-10.21	-6.26	-3.50	-1.23	-1.23	-0.63	-0.25
14.70	D14S51	-2.80	-1.76	0.81	-0.18	0.27	0.14	0.14	0.08

%pter-qter originally determined from from NIH/CEPH Collaborative Mapping Group (1992), then progressively updated from CHLC Reports (Murray et al., 1993 & 1994), Cox et al. (1994) and "The National Centre For Biotechnology Information" GB4 GeneMap'98 (1998). An 'affecteds-only' EXCLUDE map for the QMPD family (see, Figure 5.5, p. 139) was generated from this data and the data in Appendix B, Table B5.1.

Both D14S54 and D14S76 achieved the positive lod score of 0.13 ( $\theta=0$ ). D14S54 lies within the MPD1 candidate region while D14S76 lies distal to D14S49, the marker defining the distal end of the MPD1 candidate region. The positive lod scores for D14S54 and D14S76 arose because they were uninformative markers, each with 0.0 equivalent number of recombinants and 0.4 equivalent number of informative meioses. The informativeness of all the QMPD markers and the lod scores for linkage between the Silburn myopathy disease locus and other loci in the genome, are presented in Table B5.1, Appendix B.

Figure 5.3 (p. 136) shows QMPD chromosome 14 haplotypes based on minimum recombination. These haplotypes do not support the linkage of any chromosome 14 markers to the Silburn myopathy disease locus. It is noticeable that both the markers with positive lod scores, D14S54 and D14S76, were close to markers generating significantly negative two-point lod scores for co-segregation with the Silburn myopathy disease locus (Table 5.1, p. 134).

#### **5.1.4 Partial Genome Screen**

Further exclusion of the QMPD from regions of the genome was undertaken in preparation for future distal myopathy gene localisation. This work was completed knowing the Laing myopathy MPD1 location but before this location, or any other distal myopathy disease gene locations, were published. Accordingly, the segregation of polymorphic markers was not targeted at genetic loci which have subsequently been identified as distal myopathy candidate regions. These are discussed in Section 5.2.5 (pp. 150-152).



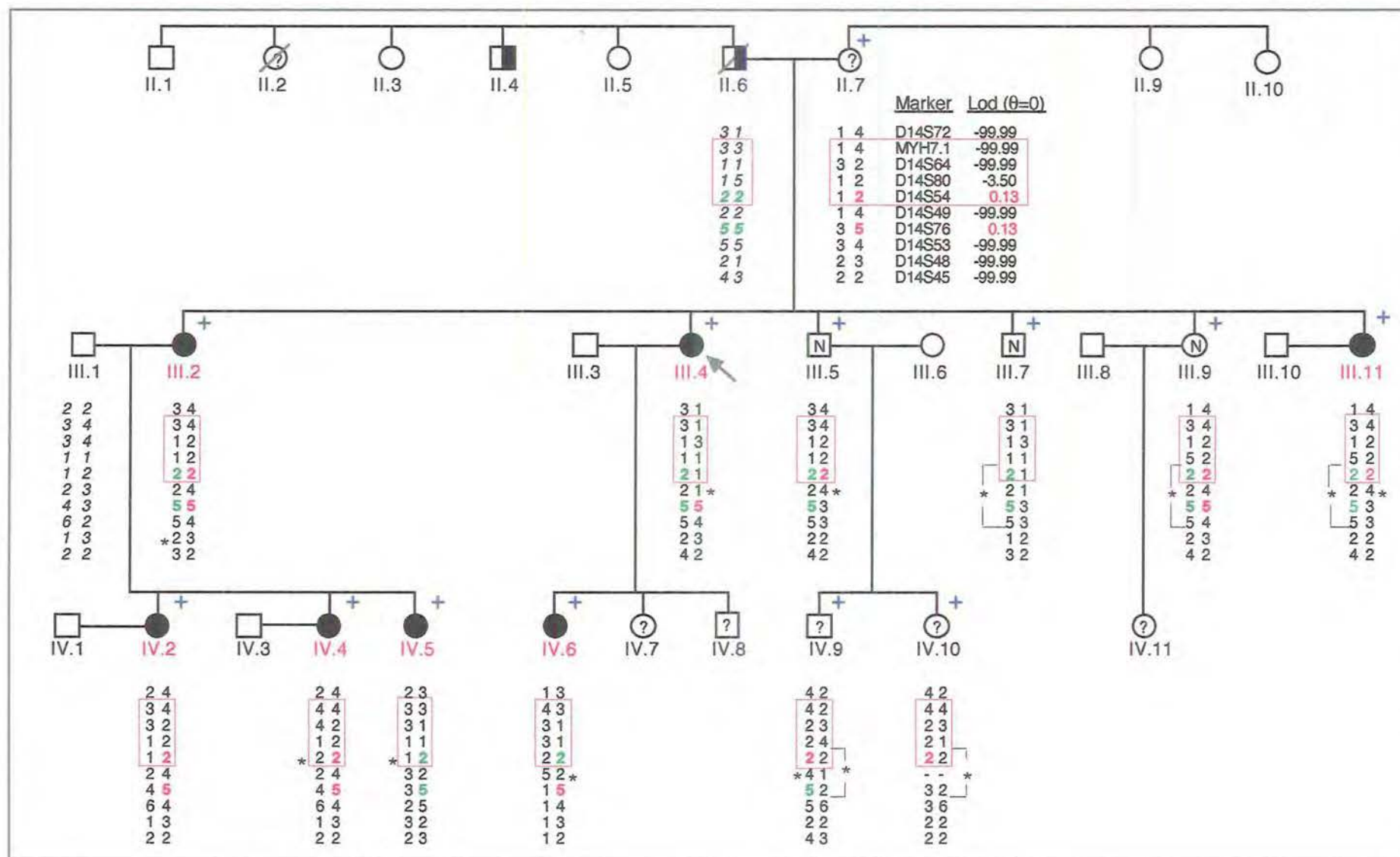
**Figure 5.3** Genotypes of participating (+) QMPD pedigree members for marker loci spanning MPD1 candidate region. Haplotypes were based on minimum number of recombinations. Where possible, paternal haplotypes are shown on the left. Italicised haplotypes were deduced, not determined. \* indicate 'minimum-recombinant' recombinations. Question marks (?) indicate haplotype/recombination uncertainty due to insufficient meioses.

In an 'affecteds-only' analysis, the segregation of D14S54, a MPD1 candidate region microsatellite, and D14S76, a microsatellite distal to the MPD1 candidate region, both produced the same positive two-point lod score of 0.13 ( $\theta=0$ ). This happened because both the D14S54(2) and the D14S76(5) alleles were present in every affected person, sometimes in homozygous genotypes (2,2) and (5,5). However, inspection revealed that: (i) the co-segregation of the D14S54(2) and D14S76(5) alleles in affected individuals and without recombination (on the same haplotype) was, from II.7 (unknown disease status) to III.2, and then from III.2 to IV.2 and IV.4; (ii) similarly, the co-segregation of the D14S54(2) and D14S76(5) alleles in affected individuals and on the same haplotype was, from II.6 (probably affected) to III.2 and III.4 and from III.2 to IV.5. Therefore, the D14S54(2) and D14S76(5) alleles do co-segregate to and from affected individuals but do so on two different haplotypes. In addition, this cosegregation was not exclusive to affected individuals and also included those of 'unknown disease status', and 'probably affected individuals', a fact not included in the affecteds-only analysis.

The possibility of linkage of the Silburn myopathy disease gene to either or both of these markers was further dispelled by inspection of unaffected haplotypes. D14S54(2) and D14S76(5) co-segregate from II.6 (probably affected) to III.5 (ascertained unaffected) while D14S54(2) and D14S76(5) co-segregate from II.7 (unknown disease status) to III.9 (ascertained unaffected).

Linkage of the Silburn myopathy gene to either of these markers or to the MPD1 candidate region were both refuted.





The segregation of 29 polymorphic markers spread over the autosomes 1, 2, 4, 5, 6, 7, 8, 13, 15, 17, 19, 20 and 21 was recorded for the 13 QMPD family members taking part in the study. DNA polymorphisms with a D-number were anonymous DNA fragments (D-segments), other markers were associated with genes. The Table B5.1 legend (Appendix B) defines the acronyms for markers associated with genes. From chromosome 1 to 21, the markers tested against the QMPD pedigree were: D1S164, D1S223, D1S187 D1S185, D1S242, and ACTN2; APOB, and D2S102; D4S174, and D4S171; D5S268, D5S112, D5S127, and D5S399; SCA1, and DMDL; D7S531, D7S472, D7S517, and D7S483; D8S198; D13S71; TPM1, and D15S87; GX-Alu; DMCTG; D20S66; D21S213; and D21S171.

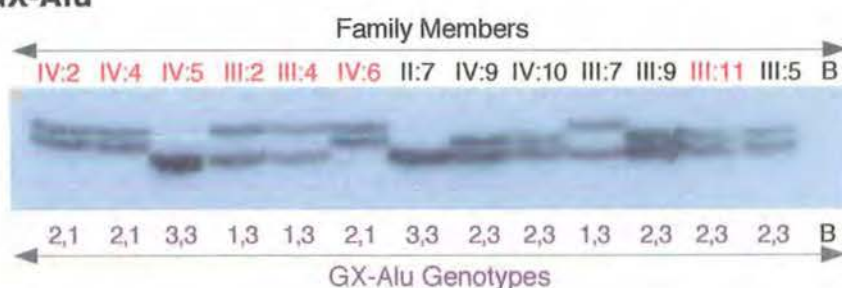
Primer synthesis,  $^{32}\text{P}$  dCTP labelling, PCR conditions, thermal cyclers used, electrophoresis and gel exposure to film for the partial genome screen of QMPD were as described in Section 4.3.1 (pp. 104-106). Figure 5.4 (p. 138) shows a representative microsatellite marker, GX-Alu on chromosome 17, from this partial genome screen of the QMPD family.

Table B5.1 (Appendix B) shows 'affecteds-only' two-point LIPED generated lod scores for linkage between the Silburn myopathy disease gene and all the 40 markers tested on the QMPD pedigree.

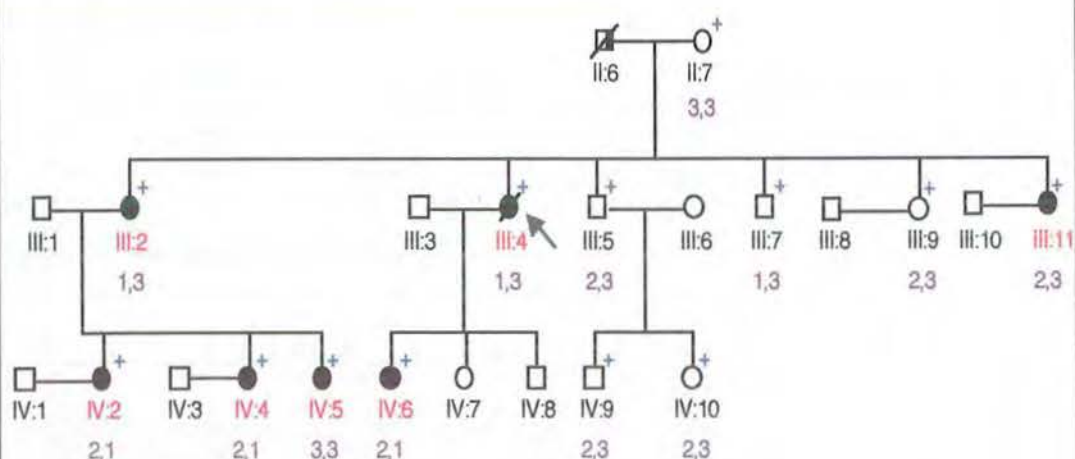
### **5.1.5 EXCLUDE**

Analysis of the combined LIPED data for all 40 markers for the affected family members was performed using the EXCLUDE program (Edwards, 1987; Sarfarazi et al. 1989). The 'affecteds-only' EXCLUDE map for the Silburn myopathy family, presented in Figure 5.5 (p.139), was generated from the 'affecteds' data in Table B5.1 (Appendix B).

### A. GX-Alu



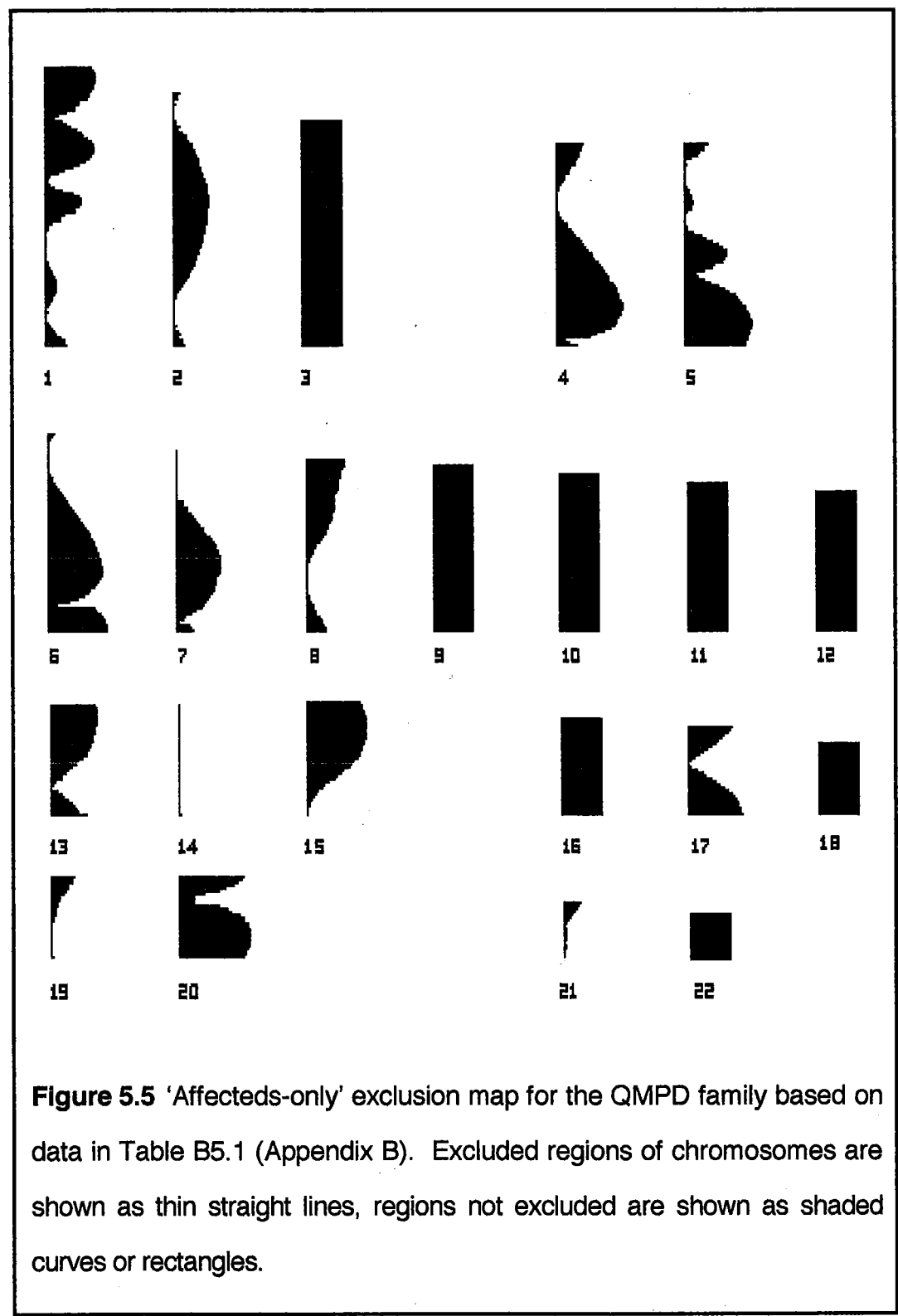
### B. QMPD Pedigree



**Figure 5.4** Segregation of chromosome 17 microsatellite GX-Alu in QMPD pedigree. **A.** GX-Alu microsatellite autoradiograph. The GX-Alu microsatellite incorporated  $\alpha$ - $^{32}\text{P}$  dCTP and was electrophoresed on 6% acrylamide gel. **B.** Segregation of GX-Alu alleles in participating (+) QMPD pedigree members. Where possible, paternal alleles are shown on the left.

Comparison of affected genotypes shows why an affecteds-only negative lod score of  $-99.99$  ( $\theta=0$ ) was obtained for co-segregation of GX-Alu with the Silburn myopathy disease gene. Grandchildren **IV:2**, **IV:4** and **IV:6** all received GX-Alu allele 1 from their 'probably affected' maternal grandfather, II:6, while grandchild **IV:5** did not receive a GX-Alu allele from this grandfather. In addition, daughter **III:11** received the GX-Alu allele 2 and not allele 1 from her 'probably affected' father, II:6.

The MPD1 candidate region was excluded as a location for the gene causing Silburn myopathy, confirming the genetic heterogeneity of the distal myopathies first suggested by the two-point lod scores presented in Table 5.1 (p. 134).



## 5.2 AUTOSOMAL DOMINANT DISTAL MYOPATHY SEGREGATING IN A CONNECTICUT FAMILY

Four affected and five unaffected members of the CMPD pedigree agreed to participate in genetic testing as shown in Figure 5.6 (p. 141). Like the QMPD family, there were too few affected individuals for localisation of the CMPD disease gene, but exclusion of the MPD1 candidate region was possible.

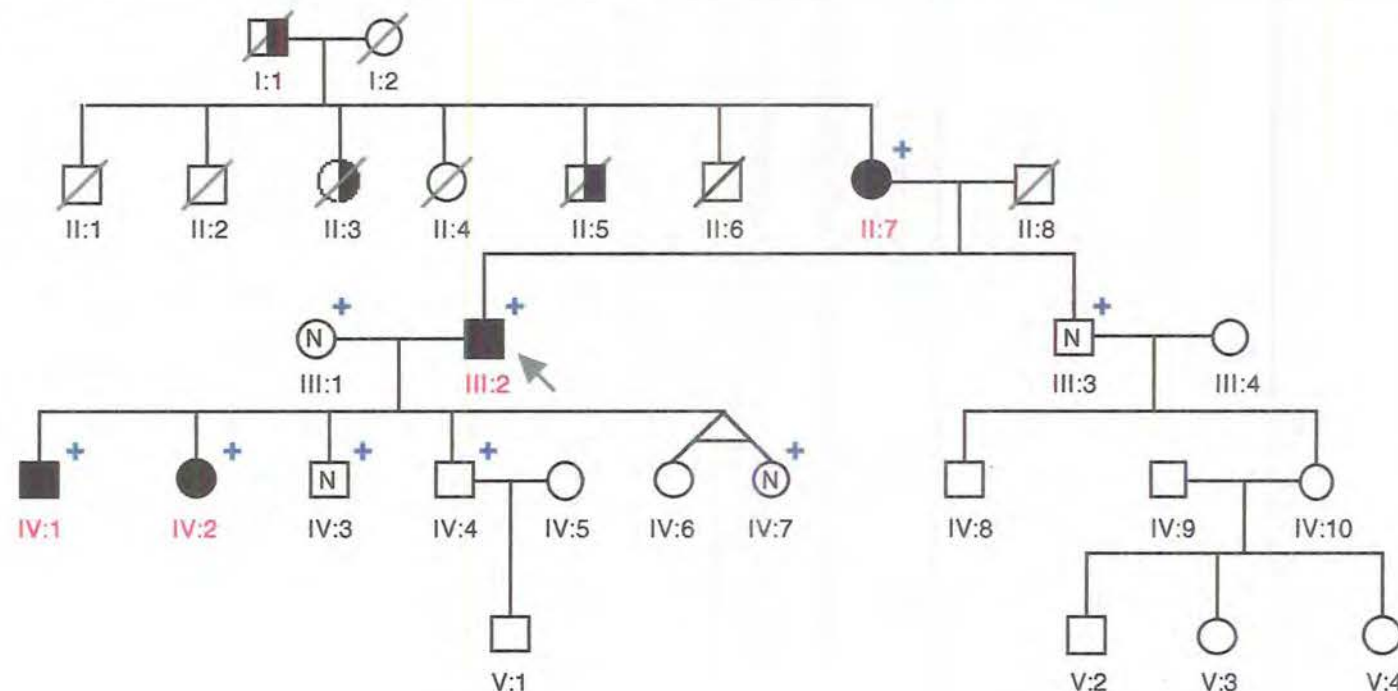
### 5.2.1 Ascertainment

The four affected and five unaffected participants from this autosomal dominant distal myopathy kindred were examined by Dr. Kevin Felice, Department of Neurology, University of Connecticut School of Medicine, Farmington, USA (Felice et al., 1999). Those affected showed similar patterns of weakness, but displayed varying severity.

The proband, III:2, in Figure 5.6 was a male with progressive foot drop first noticed at age 35 years. The initial problems included frequent tripping and difficulty with stair climbing. This is usually a sign of proximal muscle weakness and not due to the onset of foot drop. By the time of clinical ascertainment, he was using a cane and bilateral ankle foot orthoses. Other problems included progressive exertional dyspnoea (laboured breathing), bilateral cataract surgery at age 44 (capsular cataracts), hypertension and gouty arthritis (deposition of uric acid and urates in joints).

The examination of III:2 at age 57 revealed that muscle bulk and tone were normal in the upper extremities. Strength of III:2 at age 57 was determined according to MRC Grades (Brooke et al., 1983) a scale running from 0 (no contraction) to 5 (normal power) (see, Table 4.2, p. 99, for all MRC Grades).





**Figure 5.6** Pedigree of a Connecticut family (CMPD) segregating distal myopathy. Squares denote males, circles females, and lines through squares or circles indicate 'deceased'. Shaded symbols and red number indicate individuals affected by distal myopathy, half-shaded symbols show individuals who, by history, were 'probably affected' by distal myopathy. An 'N' in the centre of the symbol denotes 'unaffected by distal myopathy'. The proband (III:2) is indicated by the arrow. Each individual marked with a cross (+) had an immortalised cell line established and contributed DNA for exclusion analysis. Individuals II:7, III:2, IV:1, IV:2, III:1, III:3, IV:3 and IV:7 were clinically ascertained. Individuals I:2, II:1, II:2, II:4, II:6, II:8, IV:6, IV:8, IV:9, V:1, V:2, V:3 and V:4 were not examined but, by history, appeared to be, or were, unaffected by distal myopathy. Similarly, history indicates that individuals I:1, II:3 and II:5 were probably affected by distal myopathy. AD inheritance was indicated.

Strength was found to be normal in the upper extremities except for weakness of the arm extensors (3), finger extensors (4), and the part of the flexor digitorum profundus controlling digits V (2). The forced vital capacity was 1.7 litre (30% of predicted). There was marked foreleg and foot muscle atrophy. III:2 needed to push off with both hands to stand from a chair. The gait was unsteady with a combined waddle and steppage quality. The MRC-graded lower limb muscle strength was: hip flexors (3); knee extensors (4); foot dorsiflexors (1); plantar flexors (2); toe flexors (1); and toe extensors (1). CK concentration was 534 U/l. This was raised above 'normal' <180 U/l level as defined by Gilboa and Swanson (1976).

Concentric needle electromyography (CNE) revealed slightly increased insertional activity, rare fibrillations and positive waves, and increased recruitment of short duration, low amplitude, polyphasic motor unit action potentials (i.e. myopathic recruitment). A left vastus lateralis muscle biopsy disclosed non-specific myopathic changes including increased variation in fibre size, increased internalised nuclei, and mildly increased endomysial fat and connective tissue. Rare necrotic fibres were present. No abnormal inclusions or vacuoles were present.

Many clinical measurements in this patient were normal. He denied double vision, swallowing problems, weight loss, muscle cramps, sensory loss or paresthesias, passage of dark coloured urine, exercise or cold-induced weakness, or voiding difficulties. The heart and lung sounds were normal. There were no dysmorphic features, scapular winging, scoliosis or foot deformities. The cranial nerve examination was normal with no weakness of eyelid, ocular, facial, pharyngeal, lingual (tongue) or neck muscles. There was no sternocleidomastoid muscle atrophy. There was no myotonia, and pathological reflexes

were not present. Anti-nuclear antibody titre and thyroid stimulating hormone, carnitine, lactate and aldolase levels were all normal. Sensory examination, electrocardiogram (ECG), sensory and motor nerve conduction, repetitive motor nerve stimulation studies and chest X-ray were all normal. The vastus lateralis muscle biopsy revealed no RVs, inclusions, cores or rods or abnormal storage of lipid or glycogen. EM showed normal looking mitochondria with no abnormal inclusions or filaments.

The son of III.2, IV.1 (Figure 5.6) appeared to have an earlier onset of this family's form of distal myopathy and began to notice weakness at age 20, while he was in the army. He recalled having difficulty walking on his heels, performing sit-ups and push-ups and having difficulty in running. The leg weakness progressed slowly to the point that he reported recent occasional falls due to leg and foot weakness. Other problems included a protuberant abdomen, exaggerated lumbar lordosis and small bilateral capsular cataracts.

The examination of IV.1 at age 35 revealed a normal upper extremity muscle strength except for mild finger extensor weakness. The forced vital capacity was 4.5 litre (82% of predicted). The gait had both a waddling and steppage quality. There was moderate foreleg and foot muscle atrophy. The MRC-graded lower limb muscle strength was as follows: hip flexors (3); knee flexors (5); knee extensors (5); foot dorsiflexors (3); plantar flexors (4); toe flexors (2); and toe extensors (2). He was able to stand on his toes but not his heels. Nerve conduction studies revealed absent peroneal compound muscle action potentials with otherwise normal sensory and motor studies. CNE revealed rare fibrillations, positive waves, and myopathic-type recruitment in proximal and distal limb muscles.



Like his father, IV:1 had many normal clinical measurements. In particular dysmorphic features and scapular winging were absent and there was no ocular, facial, lingual, pharyngeal or neck muscle weakness. Electrocardiogram (ECG) and sensory examinations were normal.

The daughter of the III:2, IV:2 (Figure 5.6) presented with a milder form of the family's distal myopathy. She was remarkable only in that she could not walk on her heels, her foreleg and foot muscles were thin, and, in the lower extremities, there was a mild weakness of the hip flexors and foot and toe extensors. All the other phenotype variables ascertained for III:2 and IV:1, were found to be normal for IV:2.

II:7, the mother of III:2 and grandmother of IV:1, and IV:2 was 83 when ascertained and admitted to a mild difficulty climbing stairs for the past 2-3 years. The examination was remarkable for moderate difficulty standing from a chair, inability to heel-walk and mild to moderate weakness of the hip flexors and foot dorsiflexors.

Although not examined, history indicates that individuals I:1, II:3 and II:5 (Figure 5.6) were probably affected by distal myopathy. Individuals III:1, III:3, IV:3 and IV:7 were clinically ascertained and found to be unaffected by distal myopathy. In addition, individuals I:2, II:1, II:2, II:4, II:6, II:8, IV:6, IV:8, IV:9, V:1, V:2, V:3 and V:4 were not examined but, by history, appeared to be, or were, unaffected by distal myopathy.

In this thesis, the AD-distal myopathy segregating in the CMPD family is described as 'Felice myopathy'.

### **5.2.2 Genetic Markers**

Blood samples were taken and immortalised cell lines were established for the 9 family members who volunteered to participate in the study. DNA extraction was identical to that described in Section 4.2.2 (p. 102).

DNA studies for the myotonic dystrophy CTG repeat expansion on 19q13 were negative. These were performed by Athena Laboratories, Worcester, MA (Felice et al., 1999). The chromosome 14 markers run for this CMPD family were: D14S72, D14S283, D14S264, D14S80, D14S262, D14S252, D14S257, D14S49 and D14S253. Primer synthesis, <sup>32</sup>P dCTP labelling, PCR conditions, thermal cyclers, electrophoresis, and gel exposure to film were all conducted as described in Section 4.3.1 (pp. 104-106).

### **5.2.3 Chromosome 14 Two-point Linkage Analysis**

Two-point linkage analysis was carried out for the CMPD family as described for the WAMPD and QMPD families. Table 5.2 (p. 146) shows the CMPD 'affecteds-only' two-point LIPED generated lod scores for all tested chromosome 14 markers. An 'affecteds-only' analysis was adopted to control for reduced penetrance.

Significantly negative lod scores ( $<-2.0$ ) were obtained for 7 of the 9 genetic markers: D14S72, D14S283, D14S80, D14S262, D14S257, D14S49, and D14S253. Three of these markers, D14S80, D14S262 and D14S257, were within the MPD1 candidate region.

**Table 5.2** *LIPED generated two-point lod scores between CMPD disease locus and loci lying within and flanking the **MPD1 candidate region** on chromosome 14. Analysis includes data from affected individuals only.*

Chromosome		Recombination Fraction ( $\theta$ )							
%pter-qter	Locus	0.00	0.001	0.01	0.05	0.1	0.20	0.3	0.4
14.08	D14S72	-3.70	-2.38	-1.40	-0.72	-0.44	-0.19	-0.08	-0.02
14.10	D14S283	-3.40	-3.39	-3.10	-1.98	-1.39	-0.80	-0.44	-0.19
14.15	D14S264	0.30	0.30	0.29	0.26	0.22	0.13	0.06	0.02
14.17	D14S80	-3.40	-3.39	-3.10	-1.98	-1.39	-0.80	-0.44	-0.19
14.18	D14S262	-3.30	-3.30	-3.05	-1.98	-1.39	-0.80	-0.44	-0.19
14.18	D14S252	0.30	0.30	0.29	0.26	0.22	0.13	0.06	0.02
14.20	D14S257	-3.40	-3.39	-3.10	-1.98	-1.39	-0.80	-0.44	-0.19
14.22	D14S49	-3.40	-3.39	-3.10	-1.98	-1.39	-0.80	-0.44	-0.19
14.25	D14S253	-3.00	-3.00	-2.85	-1.96	-1.39	-0.79	-0.44	-0.19

%pter-qter originally determined from from NIH/CEPH Collaborative Mapping Group (1992), then progressively updated from CHLC Reports (Murray et al., 1993 & 1994), Cox et al. (1994), Généthon Human Genome Research Centre Genetic Maps (1993-1996) and "The National Centre For Biotechnology Information" GB4 GeneMap'98 (1998).

Two of the 9 markers run against the CMPD family, D14S264 and D14S252, achieved the positive lod score of 0.3 ( $\theta=0$ ). Both these markers lay within the MPD1 candidate region. The positive lod scores for D14S264 and D14S252 arose because they were uninformative markers, each with 0.0 equivalent number of recombinants and 0.9 equivalent number of informative meioses (see Table B5.1, Appendix B). Figure 5.7 (p. 148) shows the genotypes of CMPD pedigree members for marker loci spanning the MPD1 candidate region. The legend to Figure 5.7 argues that that haplotypes based on minimum recombination do not support linkage of these markers to the Felice myopathy disease locus. In addition, Figure 5.7 and Table 5.2 show that these markers are close to, or coincident with, markers generating significantly negative two-point lod scores for co-segregation with the Felice myopathy disease locus.

#### **5.2.4 Multipoint Linkage Analysis: MPD1 Candidate Region**

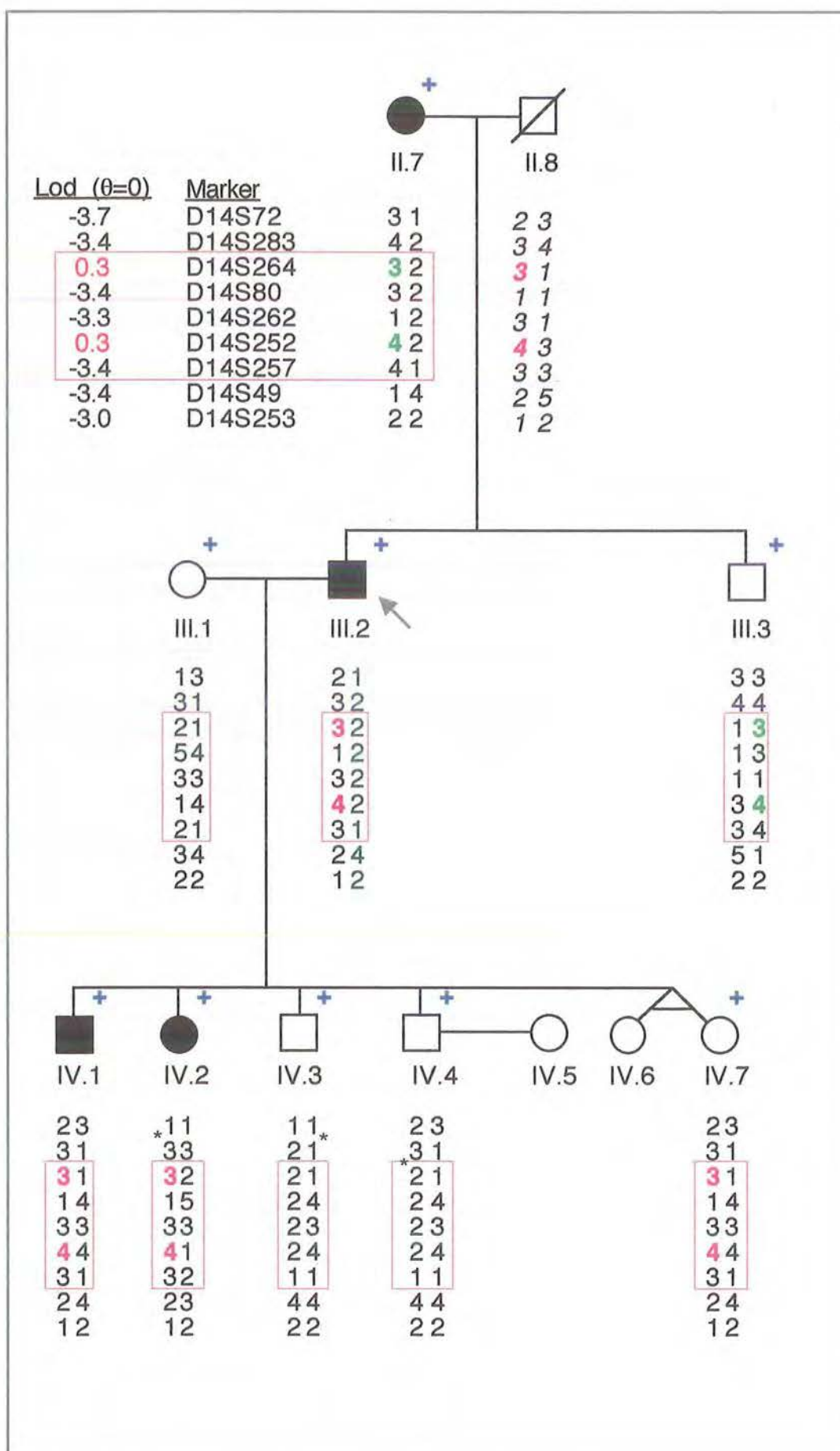
Multipoint linkage analysis (Mulley et al., 1993) was performed using the affected data presented in Table 5.2. This was conducted by Dr. Joachim Hallmayer, Centre for Clinical Research in Psychiatry, Department of Psychiatry, University of Western Australia, Graylands Hospital, Mount Claremont, Western Australia. Multipoint lod scores were calculated between a putative disease locus and a marker locus using the program GENEHUNTER (Kruglyak et al., 1996).

Distances on chromosome 14 were obtained from the maps in Gyapay et al. 1994, and the on-line map of the Co-operative Human Linkage Centre (1997; <http://lpg.nci.nih.gov/CHLC/>). In the multipoint linkage calculations, the disease gene frequency was assumed to be 0.001. Since the number of affected family members was not large enough to draw any firm conclusions in regard to the penetrance, penetrance was set to 70%.

**Figure 5.7** Genotypes of participating (+) CMPD pedigree members for marker loci spanning the MPD1 candidate region. Haplotypes were based on minimum number of recombinations. Where possible, paternal haplotypes are shown on the left. Italicised haplotypes were deduced, not determined. \* indicate 'minimum-recombinant' recombinations.

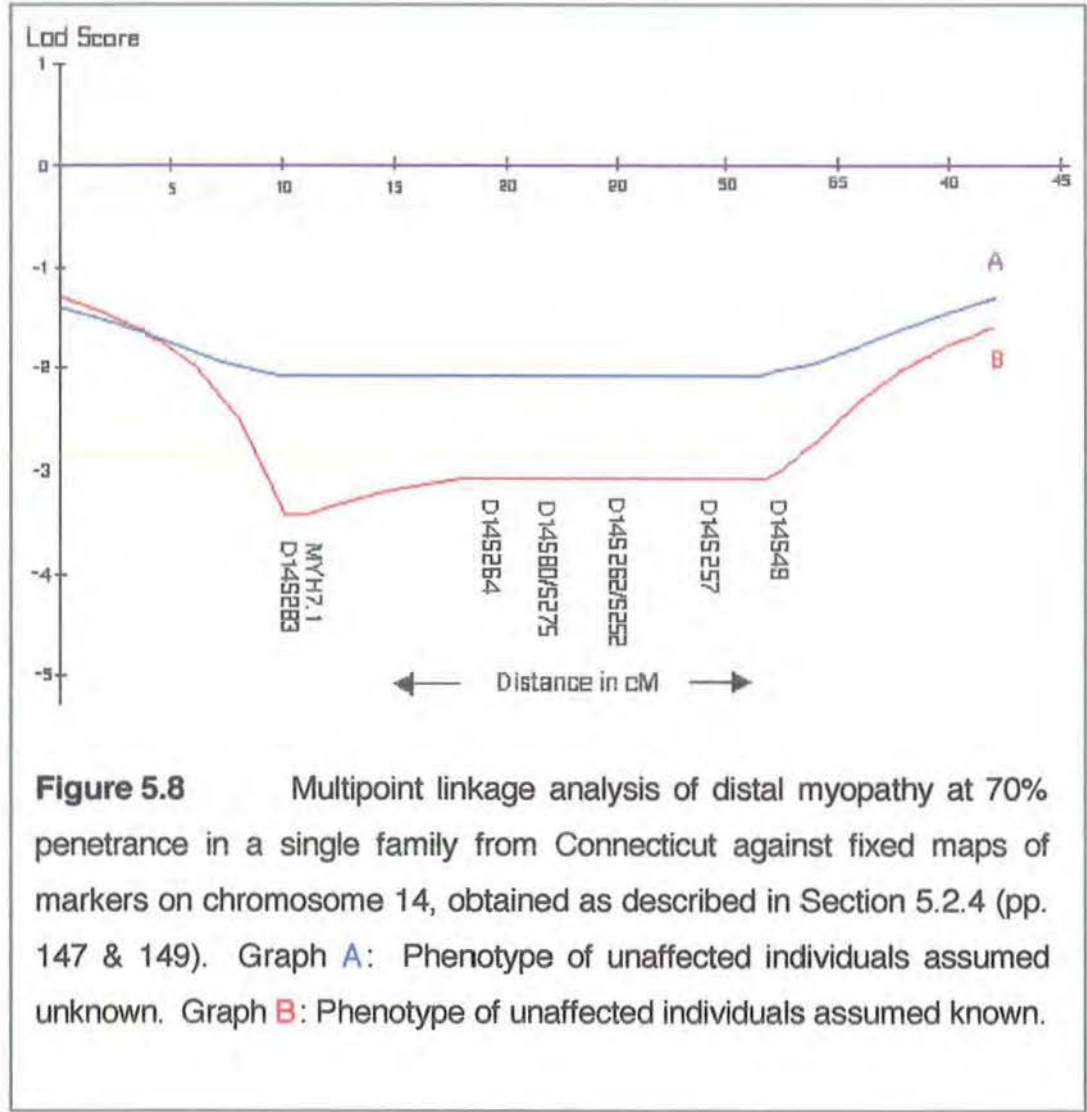
In an 'affecteds-only' analysis, the segregation of the candidate region markers D14S264 and D14S252, produced a positive two-point lod score of 0.30 ( $\theta=0$ ). This happened because both the D14S264(3) and the D14S252(4) alleles were present in every affected person. However, inspection revealed that, without recombination: (i) the co-segregation of D14S264(3) and D14S252(4) on the same haplotype was, from II.8 (who married into the family and by history was unaffected) to III.2, and then from III.2 to IV.1 and IV.2 and to IV.7 (ascertained unaffected); (ii) similarly the co-segregation of the D14S264(3) and D14S252(4) alleles co-segregate from II.7 to III.3 (ascertained unaffected). Therefore, the D14S264(3) and D14S252(4) alleles do co-segregate but do so on two different haplotypes and to-and-from affected and unaffected individuals. This could only be determined by haplotype analysis, and was obscured by 'affecteds-only' analysis

Linkage of the Felice myopathy gene to either of these markers or to the MPD1 candidate region were both refuted.





Two sets of multipoint linkage analysis calculations were performed: either the phenotypic status of unaffected individuals was considered unknown; or phenotypic information of unaffected individuals was included. All informative MPD1 candidate region markers demonstrated recombinations with affected individuals in the pedigree (Felice et al., 1999, p. 63). Figure 5.8 shows that the lod scores obtained from the multipoint analyses, under either model, were less than -2.0 for the entire MPD1 candidate region. The MPD1 candidate region was excluded by multipoint linkage analysis as the Felice myopathy disease gene location, confirming the indication of distal myopathy's genetic heterogeneity implicit in the data already presented in Table 5.2 (p. 146).



### **5.3 LAING, SILBURN AND FELICE MYOPATHIES**

Table 5.3 (p. 151) compares the essential features of Laing, Silburn and Felice myopathies. It shows that the clinical features of Laing myopathy were more similar to Felice myopathy than to Silburn myopathy. Mode of inheritance, first symptoms, EMG, lack of cardiomyopathy, progression and biopsy findings were all similar for Laing and Felice myopathies. Onset age, pattern of weakness and CK levels differed.

Silburn myopathy and Laing myopathy were similar in their mode of inheritance, age of onset and EMG data but different in their first symptoms, pattern of weakness, rate of progression, CK levels, biopsy findings the presence of cardiomyopathy and in the OPMD symptoms in the Silburn myopathy proband (see p. 131).

### **5.4 THREE OTHER DISTAL MYOPATHY LOCI**

Six other forms of distal myopathy have been assigned to three distinct loci since the localisation of Laing myopathy to the MPD1 candidate region (Laing et al., 1995b). Linkage of the AR Miyoshi myopathy (MM) to chromosome 2p12-p14 was achieved shortly after MPD1 localisation (Bejaoui et al., 1995). More recently, a newly described distal anterior compartment myopathy has been linked to the same 2p12-p14 locus (Illa et al., 1998). Tibial muscular dystrophy (TMD) and late-onset distal myopathy of Markesbury have both localised to 2q31-33 (Haravuori et al., 1998a & 1998b). Finally, an AR distal myopathy, Nonaka myopathy, has been localised to 9p1-q1 (Ikeuchi et al., 1997)..deltoid weakness, weak dorsiflexion of wrists; mild weaknesses of finger flexion; and variable weakness of facial musculature



**Table 5.3*****Laing, Silburn and Felice myopathies.***

Distal myopathy	Infant/early-childhood onset ( <i>Laing myopathy, 1995</i> )	Early-adult onset ( <i>Silburn myopathy</i> )	Early-adult onset ( <i>Felice myopathy</i> )
Inheritance:	AD	AD	AD
Earliest onset:	4 years	Second decade	Third decade
First symptoms:	Distal leg muscles in the the anterior compartment	deltoid weakness	Distal leg muscles in the anterior compartment
Progression:	Slow: to neck flexors, finger extensors, then to proximal hip abductors, external rotators and shoulder abductors	Variable, slow <u>or</u> fast: to wasting of hands, feet and sternocleido-mastoids, then to weakness of external ocular and bulbar muscles	Slow: to knee and finger extensors, then to proximal hip flexors
Cardiomyopathy:	No: but one autopsy report of 'Myocardial degeneration'	Yes: biventricular failure	No
Serum CK:	Normal or slightly elevated	1.3-6.3X normal	
EMG:	Myopathic pattern	Myopathic pattern	Myopathic pattern
Muscle biopsy:	Variable, no vacuoles observed	Hyper & atrophic, fibres, fibrous & fatty infiltration	Variable, no vacuoles observed
Locus:	14q11-q13	NOT at: 14q11-q13	NOT at: 14q11-q13, 9p1-q1, 2p12-p14, or 2q31-33

The CMPD family was tested by Medical Scientists at the Department of Neuropathology, Royal Perth Hospital, with genetic markers covering each of these distal myopathy candidate regions. Their results were reported with the work described above (Felice et al., 1999, pp. 63-65) and are briefly detailed here due to their relevance to distal myopathy nosology. The markers run were: D2S286, D2S145, D2S2111, D2S291, and D2S303 for the 2p12-p14 Miyoshi distal myopathy candidate region; D2S389, D2S2273, D2S364, D2S148, and D2S300 for the 2q31-33 TMD and late-onset Markesbury distal myopathy candidate region; and D9S276, D9S248, D9S43, and D9S319 for the 9p1-q1 Nonaka myopathy candidate region.

Multipoint linkage analysis between Felice myopathy disease locus and genetic marker data in these locations was similar to that performed for the MPD1 candidate region; the disease gene frequency was assumed to be 0.001, penetrance was set to 70% and calculations were carried out by alternately including or ignoring unaffected phenotypic status. Although the distances between markers on chromosome 9 were fixed in the same way as those in the MPD1 candidate region, the distances between markers on chromosome 2 were fixed according to the maps given in Gyapay et al., (1994), Bashir et al. (1996) and the on-line map of the Généthon Human Genome Research Centre (1993-1996).

Lod scores obtained from the multipoint analyses under either model, were less than -2 for virtually the entire 2p12-p14, 2q31-33 and 9p1-q1 distal myopathy candidate regions (Felice et al., 1999, p. 63). With the MPD1 candidate region, these distal myopathy candidate regions were, therefore, excluded as locations for the disease gene causing Felice myopathy. The existence of a further distal myopathy locus for this form of distal myopathy seemed likely, further increasing the extent of distal myopathy genetic heterogeneity (Felice et al., 1999, p. 65).

## 5.5 CONCLUSION

Clinical and pathological features of Silburn and Felice myopathies differed in some respects from each other and from Laing myopathy. However, as the dystrophinopathies first showed, allelic series can be quite diverse in clinical and pathological features. This was discussed in, 'Muscular Dystrophy Nosology' (Section 1.2.2, pp. 7-9). Therefore, the clinical and pathological distinctions between these three myopathies did not exclude the possibility that all of them, or any two of them, belonged to an allelic series. It did not even exclude the possibility that all of them, or any two of them, were caused by the same mutation.

Liu et al. (1998) found that an ATC->GTC missense at 4265 mutation in the 2p13 dysferlin gene (*DYSF*), gave rise to MM and AR-distal anterior compartment (tibial) myopathy (DMAT) in the same Italian family. Moreover, in two Spanish families the same *DYSF* mutation, a deletion of G at 5966, resulted in MM in one of these families but DMAT in the other. DMAT was first described by Illa et al. (1998). The implications of the dysferlin findings for MD nosology, and nosology in general, are discussed further in Chapter 9.

The exclusion of the MPD1 candidate region as the location for both the Silburn and Felice myopathy disease gene(s), established unequivocally that Laing myopathy did not belong to an allelic series which included either or both the Silburn myopathy or the Felice myopathy. The simultaneous exclusion of the Felice myopathy disease gene by fellow workers from the candidate regions for Miyoshi myopathy, tibial muscular dystrophy and Nonaka myopathy, indicated the existence of a further distal myopathy locus (Felice et al., 1999, p. 65).

These discoveries launched a revised genetic classification of the distal myopathies (Felice et al., 1999, pp. 64-65) which had some differences to the one presented in Table 4.1 (p. 94). The genetic classification of the distal myopathies will be further considered in Chapter 9.

## **CHAPTER SIX**

# **EXCLUSION OF THE MPD1 AND OPMD CANDIDATE REGIONS FOR A PUTATIVE- OPMD SEGREGATING IN A WESTERN AUSTRALIAN FAMILY**

## **6.0 EXCLUSION OF THE MPD1 AND OPMD CANDIDATE REGIONS FOR A PUTATIVE-OPMD SEGREGATING IN A WESTERN AUSTRALIAN FAMILY**

On-line Mendelian Inheritance in Man (OMIM, 164300) (1999) cites Taylor (1915) as the first to describe progressive vagus-glossopharyngeal paralysis with ptosis (eyelid droop) characteristic of OPMD. This was observed in a family of French-Canadian extraction. Taylor's report went largely unnoticed for about 50 years during which time those with similar clinical phenotype were classified as 'ocular myopathy' cases, particularly after the extensive review of ocular myopathies by Kiloh and Nevin, (1951).

The modern study of OPMD began when Victor et al. (1962) described a Jewish family from Eastern Europe segregating an AD condition, presenting late in life and characterised by dysphagia caused by swallowing muscle failure and progressive ptosis with predominant involvement of the levator palpebrae. These authors recognised the primary myopathic nature of the disease by giving it its present name. Hayes (1963) confirmed the observations of Taylor (1915) and Victor et al. (1962) by adding two generations to Taylor's 1915 pedigree. Nevertheless, confusion between ocular myopathies and OPMD continued. In 1965 Bray et al. reviewed the literature from 1948-1964 on 'ocular myopathies with and without dysphagia'. Forty-three of 148 cases of ocular myopathy discovered by these authors, had dysphagia. Although the age of onset, familial incidence and weakness of muscle groups of the ocular myopathies with and without dysphagia were quite different, all continued to be described as forms of ocular myopathy.

Unequivocal confirmation that OPMD was a distinct disease entity, was finally made by a full autopsy study of an Italian OPMD patient (Rebeiz et al., 1969

cited in Bouchard et al., 1997, pp. 157-158). In 1980, Tome and Fardeau strengthened OPMD as a distinct disease entity by identifying 8.5 nm intranuclear tubular filament inclusions (INI) in OPMD skeletal muscle. These authors asserted that INI had to be present for OPMD diagnosis. Subsequently, INI have become the specific skeletal muscle, histological marker for OPMD (Tome et al., 1993).

From the first modern description of OPMD by Victor et al., 1962 to the identification of a specific OPMD INI (Tome & Fardeau, 1980), OPMD gradually became recognised as a unique classical muscular dystrophy whose diagnosis depended upon the presence of AD segregation, late-onset dysphagia and progressive ptosis, generally after age 45, and 8.5 nm diameter skeletal muscle INI (Tome & Fardeau, 1980). Based on these diagnostic criteria, OPMD was subsequently reported in more than 20 countries and was considered by Brais et al. (1995) to have a world-wide incidence. The largest ethnic group segregating OPMD are a French-Canadian OPMD population with a highest frequency of 1/7500 in the Saguenay-Lac St Jean region (De Braekeleer, 1991). Bouchard (1997) cites the demographic analysis of Barbeau (1966) as revealing that this OPMD population was founded by a single ancestor who migrated during the 1600's.

## **6.1 OPMD NOSOLOGY**

Throughout the establishment of OPMD as a unique classical muscular dystrophy with stringent diagnostic requirements, questions about its homogeneity have been raised. First, it has been frequently noted that: OPMD is a rather generalised dystrophy, with broader clinical phenotype than its diagnostic criteria; RVs of likely autophagic nature may or may not be present

(Tome et al., 1997); mitochondrial abnormalities can be associated with some stages of OPMD (Pratt & Meyers, 1986; Goas et al., 1991); and the issue of neurogenic involvement in the aetiology of this disease remains unsettled (Hardiman et al., 1993).

Second, reports of 'OPMD' that do not meet all OPMD diagnostic criteria include those: with mitochondrial abnormalities but lacking INI, AD (Pauzner et al., 1991) and AR (Schroder et al., 1995); with myopathic dystrophic features and associated neurogenic involvement (Linoli et al., 1991); with AR inheritance and adult onset (Nishimura et al., 1991); with AR inheritance and childhood onset (Lacomis et al., 1991); and with AR inheritance in an Ashkenazi Jewish family of Hungarian descent (Fried et al., 1975). In this thesis, these forms of OPMD are described as putative-OPMDs because they are repeatedly described by authors as forms of OPMD, despite the fact that they did not fulfil the accepted OPMD diagnostic criteria (Tome et al., 1993).

Third, clinical phenotypes have been recognised which manifest both OPMD and distal myopathy symptoms. All taxonomies have difficulty with 'intermediate' forms like these 'OPMD-distal myopathy phenotypes' and eventually respond by evolving new classifications with inclusive categories for persistent 'intermediates'. Nosology is no exception. Oculopharyngodistal myopathy was proposed as a separate disease entity to describe the intermediate OPMD-distal myopathy phenotypes (Satoyoshi & Kinoshita, 1977).

The OPMD-distal myopathy phenotype was occasionally present in the investigations reported here, never to the extent that the proposed classification of oculopharyngodistal myopathy would apply, but sufficiently prominent to



highlight the overlap of the distal myopathy and OPMD phenotypes. In this context, oculopharyngodistal myopathy is reviewed.

### **6.1.1 Oculopharyngodistal Myopathy**

Satoyoshi and Kinoshita (1977) reported an autosomal dominant myopathy consisting of slowly progressive ptosis and extraocular palsy, weakness of the masseter, facial, and bulbar muscles, as well as distal involvement of the limbs starting around 40 years of age or later. Autopsy disclosed no remarkable change in the central or peripheral nervous system, and muscle biopsy specimens from all patients showed myopathic patterns without any specific change. To distinguish this disease, and all diseases displaying both OPMD and distal myopathy symptoms, Satoyoshi and Kinoshita proposed the term 'oculopharyngodistal myopathy'.

Whether the oculopharyngodistal myopathy classification survives or disappears as a part of MD nosology, will depend on the strength of supporting clinical and molecular evidence. Retrospective support can be found in the literature. Distal myopathies with clinical phenotype overlapping that of OPMD had been reported before 1977. Two examples are given.

Gowers' original description of distal myopathy, which most closely resembled Laing myopathy (Laing et al., 1995b), displayed the OPMD characteristics of an inability to raise eyebrows and a weakness in closing eyes. Frontales were powerless, orbiculares were much weaker than normal (Gowers, 1902). Schotland and Rowland (1964) described a family displaying ptosis, dysphagia and weakness and wasting of face, neck and distal limb muscles. Both OPMD and distal myopathy symptoms were present. Amato et al.

(1995) consider that Schotland and Rowland's cases were suffering from what Satoyoshi and Kinoshita (1977) later called oculopharyngodistal myopathy.

Subsequent to the Satoyoshi and Kinoshita (1977) oculopharyngodistal myopathy proposal, reports of OPMD with distal myopathy characteristics are also represented in the literature. Vita et al. (1983) described an OPMD with distal spread involving twelve individuals over three generations of an Italian family. They were remarkable in that all had ptosis, dysphagia, nasal voice and difficulty in walking. The distribution of muscle weakness in a 50 year old male proband and in one of his sisters, was proximal in the upper but distal in the lower limbs, confirming the existence of a relationship between OPMD and distal myopathy.

Distal limb myopathy with onset at 30 years, followed by the development of progressive ptosis, external ophthalmoplegia (affecting muscles that move the eye) and pharyngeal myopathy was reported in a 37-year-old Melanesian man from the Gulf Province of Papua New Guinea (Scrimgeour & Mastaglia, 1984). Two of this patient's sisters were affected, exhibiting ptosis and external ophthalmoplegia but without apparent distal muscle involvement or dysphagia. Their ages of onset were 35 and 25 years. A familial myopathy with variable expressivity was suggested.

Amato et al. (1995) describe a 19 year old female patient with early onset dysarthria (poor pronunciation) and dysphagia followed by distal extremity weakness, ptosis, progressive external ophthalmoplegia and finally recurrent episodes of intestinal pseudo-obstruction. Autosomal dominant inheritance was suggested on the basis of bilateral ptosis in the paternal grandfather. Her gait was remarkable for bilateral foot drop with ankle dorsiflexors muscle strength

scoring 3- on the Medical Research Council's scale of 0 (no strength) to 5 (full strength) (Brooke et al., 1983; Table 3.2, p. 82). In Laing myopathy patients, dorsiflexors MRC muscle strength range was 0-4.

The evaluation of this patient by Amato et al. (1995, pp.844-845), excluded diseases characterised by ptosis with or without ophthalmoplegia; desmin storage myopathy, mitochondrial myopathies, myotonic dystrophy, neuromuscular junction disorders, dysthyroid myopathy and the congenital centronuclear, nemaline rod and central core myopathies. It also excluded diseases characterised by intestinal pseudo-obstruction; DMD, facioscapulohumeral MD and familial visceral myopathies. Unlike Laing myopathy, this patient had RVs but lacked 13-18 nm cytoplasmic tubulofilaments accumulations sometimes found in the distal vacuolar myopathies (Amato et al., 1995, p. 845). She also lacked the amyloid deposition and 8.5 nm tubulofilaments specific for OPMD (Tome et al., 1993). Amato et al., (1995, p. 846) concluded:

Whether oculopharyngodistal myopathy is a variant of OPMD or a distinct neuromuscular disorder that merges in the spectrum of the other vacuolar myopathies is not known. Rimmed vacuoles and 13-18 nm cytoplasmic tubulo-filaments have been observed in some but not all cases.

Overlap of OPMD and distal myopathy clinical phenotypes reported in the literature have also been encountered in this thesis. For example, the proband III.4 of the Silburn myopathy family (Figure 5.1, p. 128) displayed weakness of external ocular muscles much like that reported by Gowers (1902) in his first report of distal myopathy. This patient also possessed reduced propulsion in the pharynx and oesophagus with decreased peristalsis consistent with muscular weakness (p. 131). Both symptoms are characteristic of OPMD.

The slow emergence of OPMD as a distinct disease entity, its generalised dystrophic nature, persistent reports of conditions resembling, but failing to meet all OPMD diagnostic criteria, and the identification of intermediate OPMD-distal myopathy phenotypes, all led to nosological uncertainty. The fact that there is only one accepted MD classification, OPMD, and one tentative MD classification, oculopharyngodistal myopathy, to describe these clinically diverse MD diseases, clearly indicates the incompleteness of this part of MD nosology. The growing number of putative-OPMD and oculopharyngodistal myopathy diseases in the clinical literature serves only to exacerbate matters. Genetic analysis offers a way to break this nosological impasse.

#### **6.1.2 Molecular Genetics and OPMD Nosology**

The identification of the Laing myopathy MPD1 candidate region and the publication of the 14q11.2-q13, 5cM OPMD candidate region, spanning the markers D14S283-D14S50-MYH7.1 for a French-Canadian form of OPMD (Brais et al., 1995), made it possible to test these candidate regions as possible candidate regions for the putative-OPMD segregating in a Western Australian family. The candidate region for the French Canadian form of OPMD was subsequently reduced to a 1.5 cM region between D14S283 and MYH7.1 (Brais et al., 1998). Overlap between the clinical phenotypes of the French Canadian OPMD and WAOPMD's putative-OPMD, and between this putative-OPMD and Laing myopathy made the testing of these candidate regions appear particularly worthwhile.

The genetic classification of distal myopathy and OPMD had begun. There was no reason why the genetic differentiation of distal myopathy, OPMD, putative-OPMD and oculopharyngodistal myopathy should not also begin.

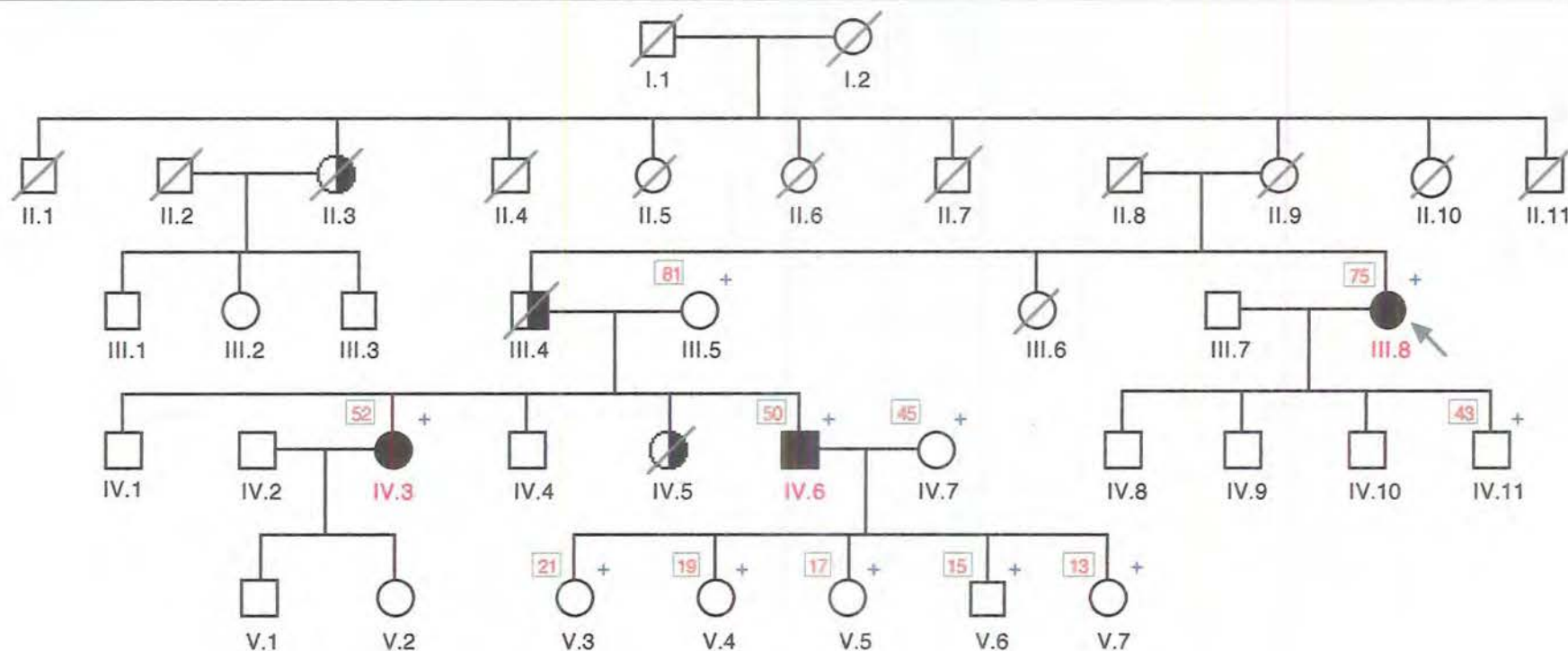
## **6.2 AN AUTOSOMAL DOMINANT PUTATIVE-OPMD SEGREGATING IN A WESTERN AUSTRALIAN FAMILY**

Three affected and eight unaffected members of the WAOPMD pedigree agreed to participate in genetic testing (Figure 6.1, p. 164). The pedigree indicated AD-OPMD segregation. Like the QMPD and CMPD families, there were too few affected individuals for the localisation of the WAOPMD disease gene, but exclusion of the French-Canadian OPMD and the MPD1 candidate regions were possible.

### **6.2.1 Ascertainment**

Ascertainment of the WAOPMD family was conducted in the Neuromuscular Clinic of the Australian Neuromuscular Research Institute, Nedlands, Western Australia, under the supervision of Dr. T. Day. Individuals III.8, IV.3 and IV.6 of the WAOPMD family were clinically examined and diagnosed as affected with OPMD symptoms.

The proband, III.8 was aged 75 years old at the time of her most recent examination. The proband was remarkable with: severe weakness of facial and palatal muscles; moderate to severe restriction of ocular movements with slow saccadic velocities; bilateral ptosis (which had been partially treated by surgery); neck flexion weakness; bilateral distal weakness; wasting in both upper and lower limbs; and bilateral foot drop like that found in Laing myopathy (p. 98). This last symptom again reinforced the frequently recorded overlap between distal myopathy and of OPMD. III.8 was unstable when walking and liable to fall. Since proximal muscle strength was normal, this may indicate the presence of CNS or neuropathic problems. Electron microscopy (EM) was performed on the muscle biopsy of this patient but INIs were not detected.



**Figure 6.1** Pedigree of a Western Australian family (WAOPMD) segregating putative-OPMD. Squares denote males, circles females, lines through squares or circles indicate 'deceased'. Shaded symbols and red numbers indicate individuals affected by OPMD, half-shaded symbols indicate individuals 'probably affected' by OPMD. The proband, III.8, is indicated by an arrow. The box at the top left-hand side of participant's symbols contains their age at ascertainment. III.8, IV.3 and IV.6 were clinically examined and diagnosed as affected. Ages at which affected individuals first noticed OPMD symptoms were not ascertained. History indicates that II.3, III.4 and IV.5 were 'probably affected' by putative-OPMD. Each individual marked with a cross (+) had an immortalised cell line established and contributed DNA for exclusion analysis. AD inheritance was indicated. Since AD-OPMD generally onsets after age 45, this relatively late onset effectively conferred 'unknown disease status' on V.3, V.4, V.5, V.6, V.7 and, possibly, on IV.11. Thus, the resolving power of this exclusion analysis was diminished.



The failure to detect INIs may have been due to sampling problems as INI detection is a difficult and time-consuming task requiring considerable experience (Bouchard et al., 1998). Until INI are identified in WAOPMD muscle, this disease can only be classified as 'putative-OPMD'.

At age 43 individual IV.3 noticed weakness of both eyes with bilateral droopiness. Within a year she began to complain of swallowing difficulty. This was followed by weakness of lower limbs and difficulty climbing stairs. It is probable that the 'weakness of lower limbs' and the 'difficulty in climbing stairs' are not related since the latter is usually a sign of proximal muscle weakness. By age 49 her voice became hoarse and sometimes unintelligible. There was occasional blurring of vision and double vision. Upon examination at age 52, bilateral ptosis, nasal voice, weakness of pharyngeal muscles and mild weakness of mainly proximal musculature, were all identified. EMG recorded small short duration units, myopathic potentials, in both upper and lower limbs. EM was not performed on the IV.3 muscle biopsy.

From age 47, individual IV.6 noticed increasing ptosis of both eyelids. Two years later he noticed a slightly nasal voice and possible droop in the corners of the mouth. Examination at age 50 noted: bilateral ptosis of moderate degree, mild restriction of eye movements in all directions with gaze-paretic nystagmus, slightly slowed saccadic movements, and a nasal voice. There was minimal weakness of lips and no muscle wasting or weakness in upper and lower limbs. Distal musculature was quite normal. EM was not done on IV.6 muscle biopsy.

### **6.2.2 Genetic Markers**

Blood samples were taken and immortalised cell lines were established for the 11 family members who volunteered to participate in the study (Figure 6.1, p.

164). DNA extraction was identical to that described in Section 4.2.2 (p. 102). For the WAOPMD family, the segregation of 14 chromosome 14 genetic markers was determined: D14S72, D14S283, MYH7.1, MYH7.2, D14S64, D14S264, D14S80, D14S275, D14S262, D14S252, D14S54, D14S257, D14S49 and D14S53

Primer synthesis,  $\alpha$ -<sup>32</sup>P dCTP labelling, PCR conditions, thermal cyclers used, electrophoresis and gel exposure to film were described in Section 4.3.1 (pp. 104-106). Figure 6.2 (p. 167) shows the markers lying at the proximal and distal ends of the MPD1 candidate region, D14S283 and D14S49. D14S283 also lies at the proximal end of the French-Canadian OPMD candidate region. Figure 6.3 (p. 168) shows D14S252, a marker lying within the MPD1 candidate region. Figure 6.4 (p. 169) displays the genotypes of WAOPMD pedigree members for 14 tested marker loci spanning the MPD1 and French-Canadian OPMD candidate regions.

### **6.2.3 Chromosome 14 Two-point Linkage Analysis**

Two-point linkage analysis was carried out for the WAOPMD family as described for the Laing myopathy family (p. 89). Table 6.1 (p. 170) contains chromosome 14 two-point LIPED generated lod scores for all affected WAOPMD family members for all of the markers run on this family. An 'affecteds-only' analysis was adopted to control for reduced penetrance.

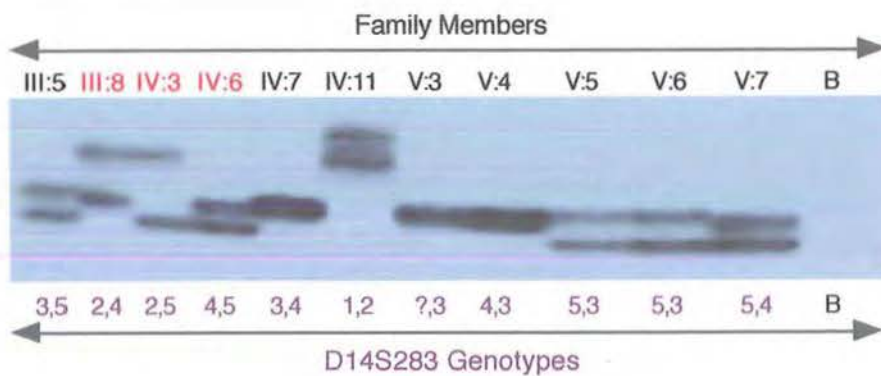
Significantly negative lod scores ( $<-2.0$  at  $\theta=0$ ) were obtained for 12 of the 14 genetic markers tested, 8 of which (those marked \*) were within the MPD1 candidate region; D14S72, D14S283, MYH7.1\*, MYH7.2\*, D14S64\*, D14S264\*, D14S80\*, D14S275\*, D14S252\*, D14S257\*, D14S49 and D14S53. The two positive lod scores were for D14S262\* and D14S54\*.



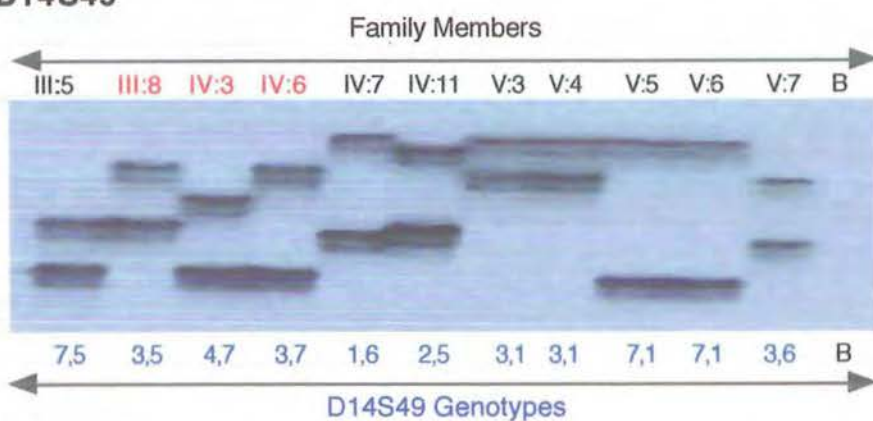
**Figure 6.2** Segregation of the microsatellite markers D14S283 and D14S49 in the WAOPMD family. D14S283 and D14S49 lie at the proximal and distal ends of the MPD1 candidate region. **A.** D14S283 autoradiograph. **B.** D14S49 autoradiograph. **A** and **B** incorporated  $\alpha$ -<sup>32</sup>P dCTP and were electrophoresed on 6% acrylamide gels. **Red** numbers indicate affected individuals; B means 'Blank'. **C.** Segregation of D14S283 and D14S49 alleles in the 11 participating (+) WAOPMD pedigree members. Where possible, paternal alleles are shown on the left.

Comparison of affected genotypes shows why an 'affecteds-only' negative lod scores of -2.83 ( $\theta=0$ ) were obtained for co-segregation of both D14S283 and D14S49 alleles with the WAOPMD disease gene. Affected individuals **III:8**, **IV:3** and **IV:6** do not all share a common allele for either D14S283 or D14S49. In addition, individuals **IV:3** and **IV:6** inherited different alleles from their 'probably affected' father, **III.4**, for both of these markers.

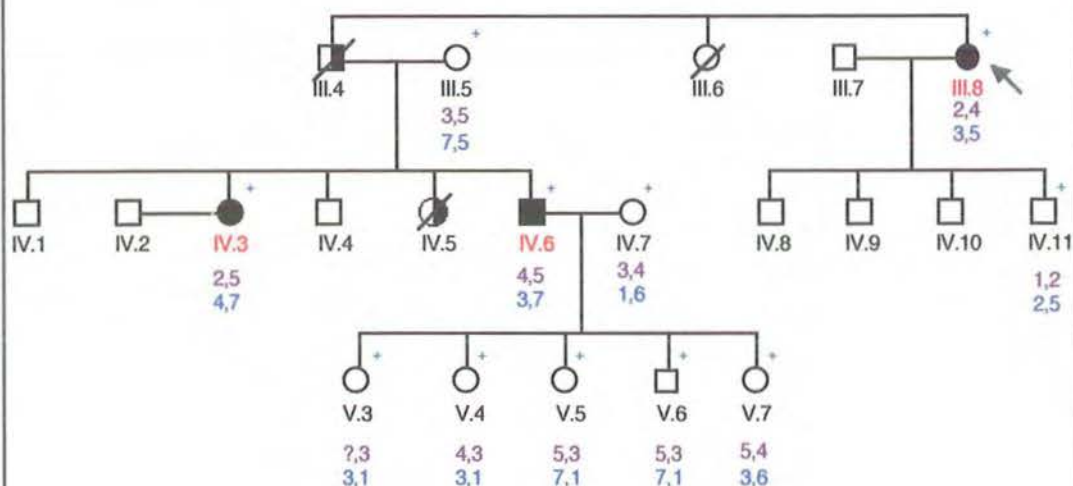
### A. D14S283



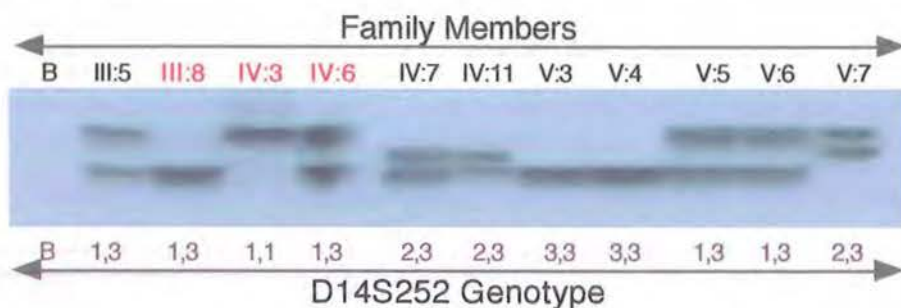
### B. D14S49



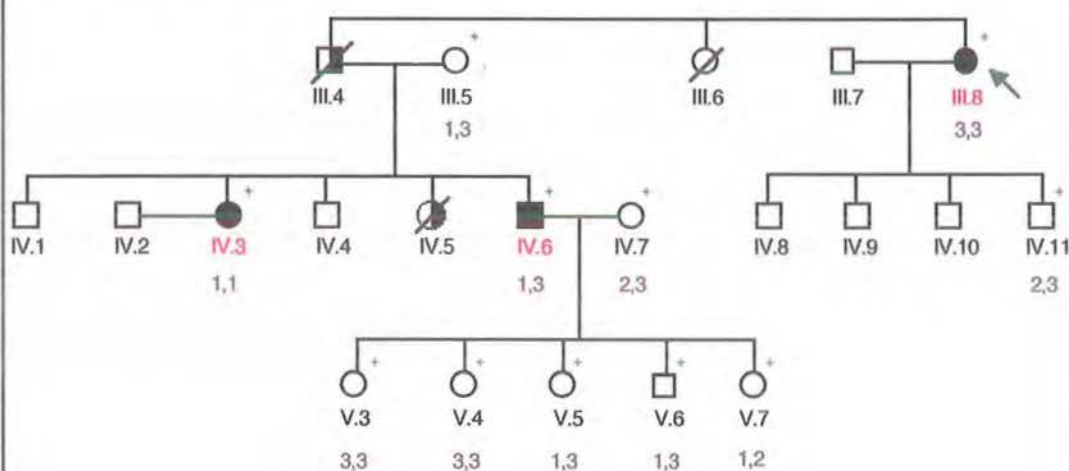
### C. WAOPMD Pedigree



### A. D14S252



### B. WAOPMD Pedigree



**Figure 6.3** Segregation of the microsatellite marker D14S252 in the WAOPMD family. **A.** D14S252 autoradiograph. The D14S252 microsatellite incorporated  $\alpha$ - $^{32}\text{P}$  dCTP and was electrophoresed on a 6% acrylamide gel. Red numbers indicate affected individuals; B means 'Blank'. **B.** The segregation of D14S252 alleles in the 11 participating (+) WAOPMD pedigree members. Where possible paternal alleles are shown on the left.

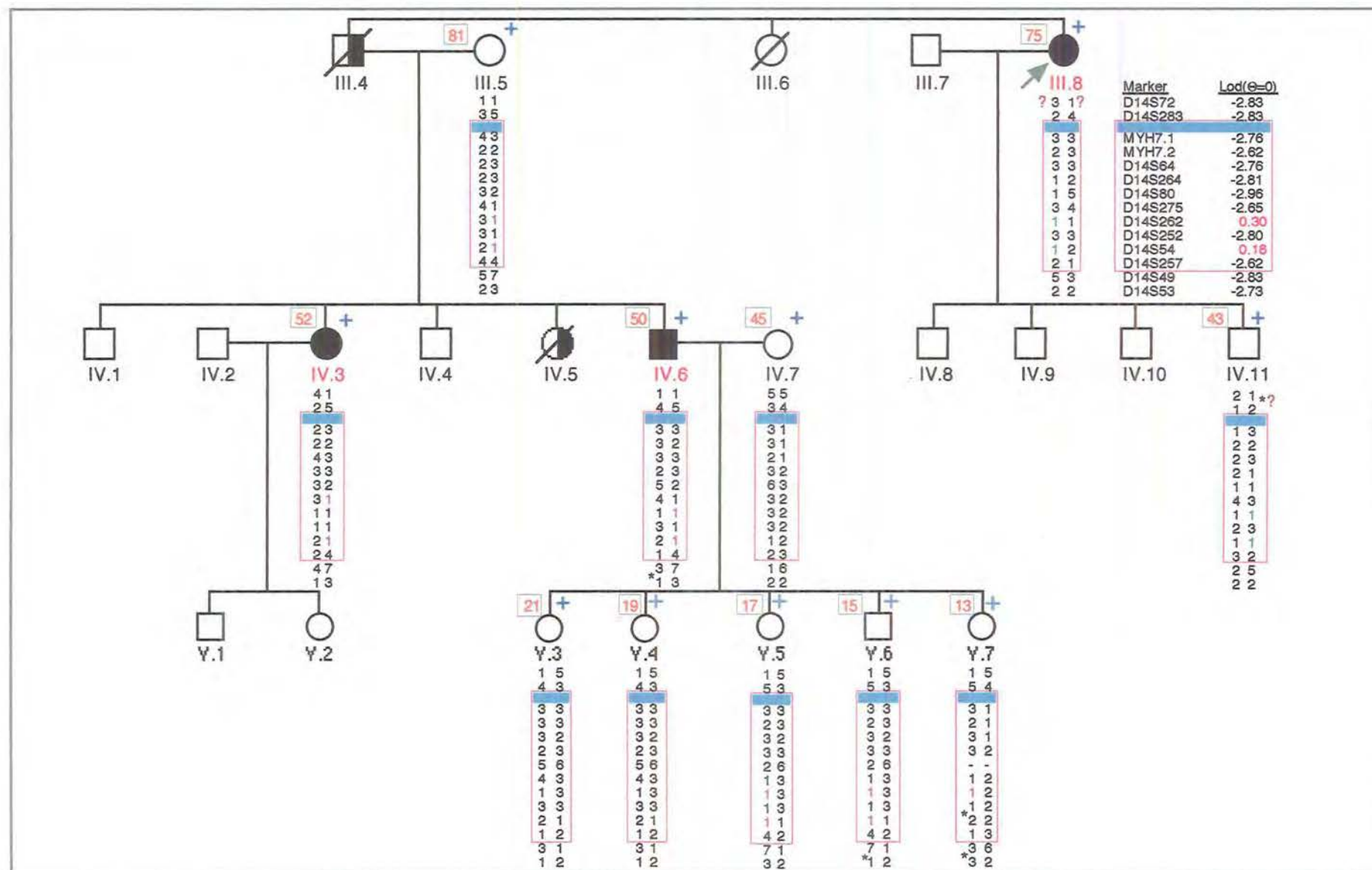
Comparison of affected genotypes shows why an affecteds-only negative lod score of -2.8 ( $\theta=0$ ) was obtained for co-segregation of D14S252 with the WAOPMD disease gene. Affected individuals III:8, IV:3 and IV:6 do not all share a common D14S252 allele.



**Figure 6.4** Genotypes of participating (+) WAOPMD pedigree members for marker loci spanning the **MPD1** candidate region (Laing et al., 1995) and the refined French-Canadian OPMD candidate region (Brais et al., 1998). The box at the top left-hand side of participant's symbols contains their **age at ascertainment**. Haplotypes are based on minimum number of recombinations. Where possible paternal haplotypes are shown on the left. \* indicate possible recombinations. Question marks (?) indicate haplotype/recombination uncertainty due to insufficient meioses.

In an 'affecteds-only' analysis, the co-segregation of the microsatellite marker alleles D14S262(1) and D14S54(1), produced the positive two-point lod scores of **0.30** and **0.18** respectively ( $\theta = 0$ ) in the MPD1 candidate region. Both the D14S262(1) and the D14S54(1) alleles were present in every affected person. Nevertheless, these positive lod scores are artifacts of the 'affecteds-only' analysis.

Haplotypes based on minimum recombination do not support linkage of these markers to the WAOPMD disease locus. Inspection reveals that, without recombination, the key co-segregations were: (i) D14S262(1) and D14S54(1) from III.8 (affected) to IV.11 (so far unaffected) and (ii) D14S262(1) and D14S54(1) from III.5 (unaffected) to IV.3 and IV.6 (both affected) and then on to V.5 and V.6 (so far unaffected). Affected individuals **III.8**, **IV.3** and **IV.6** and the 'probably affected' III.4, do not share a common D14S262-D14S54 haplotype in the MPD1 candidate regions. Therefore, positive 'affecteds-only' positive lod scores for MPD1 markers D14S262 and D14S54, do not indicate linkage to the WAOPMD disease locus.



**Table 6.1** *LIPED generated two-point lod scores between the WAOPMD disease locus and loci lying within and flanking the **MPD1 candidate region** on chromosome 14. Analysis includes data from affected individuals only.*

Chromosome		Recombination Fraction ( $\theta$ )							
%pter-qter	Locus	0.00	0.001	0.01	0.05	0.1	0.20	0.3	0.4
14.08	D14S72	-2.83	-2.38	-1.55	-0.84	-0.54	-0.25	-0.1	-0.02
14.10	D14S283	-2.83	-2.38	-1.55	-0.84	-0.54	-0.25	-0.1	-0.02
14.11	MYH7.1	-2.76	-2.31	-1.47	-0.79	-0.50	-0.23	-0.09	-0.02
14.11	MYH7.2	-2.62	-2.07	-1.21	-0.56	-0.31	-0.12	-0.04	-0.01
14.15	D14S64	-2.76	-2.31	-1.47	-0.79	-0.50	-0.23	-0.09	-0.02
14.15	D14S264	-2.81	-2.12	-1.22	-0.57	-0.32	-0.11	-0.03	-0.01
14.17	D14S80	-2.96	-2.51	-1.67	-0.94	-0.61	-0.28	-0.11	-0.03
14.18	D14S275	-2.65	-2.11	-1.25	-0.60	-0.35	-0.14	-0.5	-0.01
14.18	D14S262	0.30	0.29	0.28	0.25	0.2	0.12	0.05	0.01
14.18	D14S252	-2.80	-2.13	-1.24	-0.59	-0.33	-0.12	-0.04	-0.01
14.20	D14S54	0.18	0.18	0.18	0.15	0.12	0.07	0.03	0.01
14.20	D14S257	-2.62	-2.07	-1.21	-0.56	-0.31	-0.12	-0.04	-0.01
14.22	D14S49	-2.83	-2.38	-1.55	-0.84	-0.54	-0.25	-0.1	-0.02
14.52	D14S53	-2.73	-2.05	-1.15	-0.50	-0.26	-0.07	-0.01	0.00

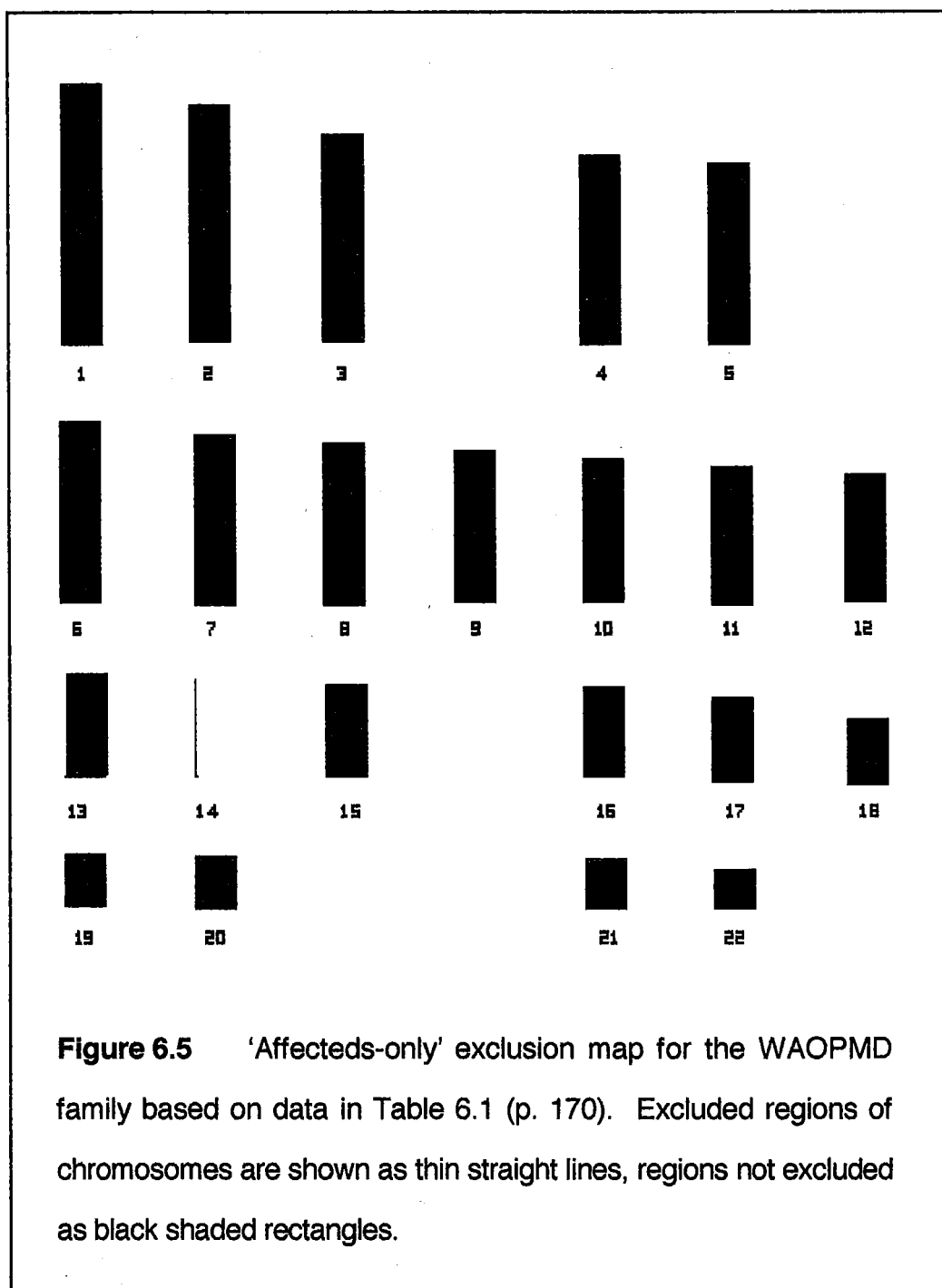
%pter-qter originally determined from from NIH/CEPH Collaborative Mapping Group (1992), then progressively updated from CHLC Reports (Murray et al., 1993 & 1994), Cox et al. (1994) and "The National Centre For Biotechnology Information" GB4 GeneMap'98 (<http://www.ncbi.nlm.nih.gov/>)

The MPD1 candidate region markers D14S262 and D14S54 achieved positive lod scores of 0.30 and 0.18 respectively ( $\theta=0$ ). The positive lod scores for D14S262 and D14S54 arose because they were uninformative markers, with 0.0 equivalent number of recombinants and, respectively, 0.9 and 0.5 equivalent number of informative meioses. Figure 6.4 (p. 169) shows that haplotypes based on minimum recombination do not support linkage of these markers to the WAOPMD disease locus. In addition, both these markers were close to markers generating significantly negative two-point lod scores for co-segregation with the WAOPMD disease locus (Table 6.1, p. 170). The data presented in Table 6.1 could only be considered as the first indication of OPMD genetic heterogeneity, if INI are found in skeletal muscle biopsies from affected WAOPMD family members.

Figure 6.5 (p. 172) presents the EXCLUDE (Edwards, 1987) results for the WAOPMD family. It confirms that the WAOPMD disease locus does not lie in either the French-Canadian OPMD or MPD1 candidate regions.

### **6.3 CONCLUSION**

The relatively late onset of this disease reduced the number of available informative meioses. This diminished the resolving power of the WAOPMD pedigree analysis. Nevertheless, it was possible to establish that the disease gene causing this putative-OPMD is not located in the Laing myopathy MPD1 or French-Canadian OPMD candidate regions. This result cannot be taken to demonstrate locus heterogeneity for AD-OPMD until it has been demonstrated unequivocally that this family segregates OPMD. For this to happen, INI must be shown to be associated with this family's disease. Conclusions about OPMD locus heterogeneity cannot be drawn from the work reported here.



Although the genetic heterogeneity of OPMD could not be established, the exclusion of chromosome 14 as the disease gene location for the WAOPMD form of putative-OPMD has started to unravel the complex genotype-phenotype relationships between inherited distal myopathy, oculopharyngo-distal myopathy, putative-OPMD and OPMD with INI.



## **CHAPTER SEVEN**

### **LAING MYOPATHY: POSITIONAL CANDIDATE APPROACH**

## **7.0 LAING MYOPATHY: POSITIONAL CANDIDATE APPROACH**

Linkage region genes become candidate genes if they are expressed in affected tissue before the onset of disease, or if uncertainty exists over the timing and/or place of their expression. Tissue-specific genes are the most conspicuous candidate genes, but disease genes need not have tissue-specific expression. The expansion of the normal (GCG)<sub>6</sub> repeat to a (GCG)<sub>9</sub> repeat in exon 1 of the *PABP2* gene to cause AD-OPMD (Brais et al., 1998) is one example of a muscular dystrophy disease gene that is expressed in many other tissues.

At the time of its identification, the  $\alpha$ - and  $\beta$ -cardiac myosin heavy chain genes (MYH genes), *MYH6* (MIM 160710) and *MYH7* (MIM 160760), were the only two muscle-specific genes known to lie within the MPD1 candidate region. These are still the only two muscle-specific genes listed in the MPD1 candidate region by 'Genomic Research', University of Padua (2000; <http://grup.cribi.unipd.it/muscle/index.html>). The candidature of *MYH6* and *MYH7* for Laing myopathy was assessed by considering the expression patterns and disease associations of each of these genes. Investigating disease associations can give insights into how proteins work, and so assist in the evaluation of disease candidature.

### **7.1 MYH6 LAING MYOPATHY CANDIDATURE**

Laing myopathy onsets in the anterior compartment of the leg. The earliest recorded onset was in early childhood (see, pp. 95-98). For *MYH6* to be a strong candidate gene for this disease, it must be expressed in affected muscles during the embryonic, foetal, infant or early childhood stages of human development.

### 7.1.1 ***MYH6* Expression**

The candidature of *MYH6* for Laing Myopathy is supported by the timing and location of its expression in skeletal muscle. Pedrosa-Domellof and Thornell (1994) collected samples from upper and lower limbs from two fetuses at each of 10, 11, 14, 16, 18, 20 and 25 weeks of gestation (wG). Serial sections of whole limb (10-14 wG) and the dissected biceps and triceps brachii, quadriceps, tibialis anterior, extensor digitorum longus, soleus and gastrocnemius muscles (16-25 wG) were investigated by light microscope immunocytochemistry. Using a monoclonal antibody specific for human and rat  $\alpha$ -cardiac myosin heavy chain (MyHC), Pedrosa-Domellof and Thornell (1994, p. 77), identified  $\alpha$ -cardiac MyHC in human muscle spindle intrafusal fibres. They concluded that *MYH6* is expressed in the tibialis anterior and the extensor digitorum longus, during weeks 10-25 of gestation. This work was not confirmed by the identification of cardiac  $\alpha$ -MyHC mRNA in affected tissue. Nevertheless, it established *MYH6* as a candidate gene for Laing Myopathy. *MYH6* was expressed before the first onset of Laing myopathy, and it was expressed in affected muscles, the tibialis anterior and the extensor digitorum longus.

### 7.1.2 ***MYH6* Mutations and Disease**

An examination of how mutations in *MYH6* might cause Laing myopathy, showed that its expression in intrafusal fibres was unlikely to cause this disease. These encapsulated bundles of small-diameter muscle fibres receive both sensory and motor innervation. They monitor changes in muscle length which are transduced into proprioceptive signals by afferent neurones and transmitted to the central nervous system (CNS) (Matthews, 1981).

Myopathy has never been associated with mutations in human muscle spindle genes. Muscle spindle absence or failure in mammals causes loss of proprioception, not myopathy. Ernfors et al., (1994) used immunocytochemistry to detect muscle spindles in wild type mice, and neurotrophin-3-deficient (NT-3) mice. Homozygous NT-3 mutant mice lacked muscle spindles, while heterozygous NT-3 mutant mice contained approximately half the wild type complement of muscle spindles.

Significantly, the histology of the soleus muscles of normal mice, heterozygous NT-3 mutant mice and homozygous NT-3 mutant mice were identical in all other respects. There were no signs of myopathy in the soleus sections of the homozygous NT-3 mutant mice (Ernfors et al., 1994, Figure 2, p. 506). The limb ataxia and inability to position the extremities properly in homozygous NT-3 mutant mice was due to loss of proprioception, not myopathy.

Since a lack of muscle spindles did not result in myopathy in NT-3 homozygous mice, it seems improbable that the malfunctioning of one muscle spindle gene, *MYH6*, could cause myopathy in humans. This, and the fact that  $\alpha$ -MyHC has never been detected in the mammalian extrafusal fibres (Walro & Kucera, 1999, p. 182), made *MYH6* a weak candidate gene for Laing myopathy.

## **7.2 MYH7 LAING MYOPATHY CANDIDATURE**

For *MYH7* to be a candidate gene for Laing myopathy, its product,  $\beta$ -cardiac MyHC, would have to be present in muscle tissue of the anterior leg compartment before or during early childhood.

### 7.2.1 *MYH7* Expression

There are no reports of  $\beta$ -cardiac MyHC expressed in human muscle spindles, even in a recent publication investigating MyHC expression patterns in unusual intrafusal fibres (Soukup & Thornell, 1999).

The expression of *MYH7* in the extrafusal fibres in primary and secondary myotubes during human embryogenesis and foetal development has been investigated in selected upper leg muscles (Hughes et al., 1993). Its expression in the anterior compartment of the leg during embryogenesis and foetal development has not been investigated.

Using immunocytochemistry (Hughes et al., 1993) detected a C-terminal  $\beta$ -cardiac MyHC epitope while primary myotubes were forming during weeks 6 to 8 of embryogenesis. In addition, when serial sections of week 17 human foetal vastus lateralis muscle were exposed under hybridisation conditions to random-primed, digoxigenin labelled, 3' untranslated human  $\beta$ -cardiac MyHC cDNA probes, mRNA containing the unique 3' untranslated region of the human  $\beta$ -cardiac MyHC was only present in the same extrafusal fibres that contained C-terminal  $\beta$ -cardiac MyHC (Hughes et al., 1993, p. 190).

The detection of  $\beta$ -MyHC epitopes in upper leg skeletal muscle during human gestation by Hughes showed that *MYH7* was expressed before the onset of Laing myopathy in the type of tissue affected by this disease. This finding did not positively identify *MYH7* as a candidate gene for Laing myopathy. Nevertheless, the lack of information about its expression in affected tissue before the onset of this disease meant that *MYH7* could not be excluded as a candidate gene for Laing myopathy.

### 7.2.2 *MYH7* Mutations and Disease

Four observations are relevant to *MYH7* candidature for Laing myopathy. First, numerous reports have associated *MYH7* mutations with hypertrophic cardiomyopathy (CMH) (Vikstrom & Leinwand, 1996; Redwood et al., 1999). Laing myopathy patients do not have CMH. Second, some mutations that cause CMH also have a skeletal muscle phenotype (Lankford et al. 1995). Third, *MYH7* mutations have been found to show variable expressivity, causing central core disease (CCD) skeletal muscle pathology, with or without CMH (Fananapazir et al., 1993). Thus, mutations in *MYH7* could cause skeletal muscle disease without CMH.

Fourth, Laing myopathy onsets in muscles in which fast-twitch fibres predominate. In this study, myopathy was detected in the vastus lateralis, a muscle dominated by type II, fast-twitch fibres (Figure 4.2, p. 101). However, the expression of *MYH7* appears to occur predominately in slow-twitch fibres (Jandreski et al., 1987). Alternatively, Izumo et al. (1986, p. 597-598) reported high levels of  $\beta$ -MyHC mRNA in the fast-twitch diaphragm muscle of the rat, *Rattus norvegicus*. The possibility that *MYH7* is also selectively expressed in human fast-twitch fibres cannot be excluded.

Consideration of these points left uncertainty about the Laing myopathy candidature of *MYH7*, with one important exception. The association of mutated forms of *MYH7* with skeletal muscle pathology, made the candidature of *MYH7* stronger than that of *MYH6*. In the absence of any other muscle-specific genes, the *MYH7* cDNA of the WAMPD proband, III.2, was tested for the Laing myopathy mutation.

### **7.3 LAING MYOPATHY MUTATION: TESTING *MYH7***

The detection of disease-causing mutations often involves the examination of exonic sequences. The simultaneous amplification of multigene family members can confound this examination. Epp et al. (1993) determined the entire 4,484 bp intergenic region between *MYH6* and *MYH7*. They showed that the *MYH6* and *MYH7* genes were arranged head-to-tail, with the *MYH7* gene located 5' to the *MYH6* gene. *MYH6* and *MYH7* were probably formed by the duplication of an ancestral gene. They retained significant homology, so their simultaneous amplification was a distinct possibility.

#### **7.3.1 Methods**

Testing *MYH7* for a Laing myopathy mutation involved: RNA extraction from skeletal muscle; skeletal muscle cDNA manufacture; *MYH7* amplification, purification and sequencing from this template cDNA; *MYH7* amplification, purification and sequencing from template lymphocyte gDNA; and the SSCP screening of *MYH7* exon 35 amplified from lymphocyte gDNA.

##### **7.3.1.1 RNA Extraction from Skeletal Muscle**

RNA was isolated from a quadriceps biopsy taken from the proband, III.2, of the WAMPD family (Figure 3.1, p. 78). This biopsy (D4292522) was collected at the Department of Pathology, Sir Charles Gairdner Hospital, Perth, Western Australia. Quadriceps biopsies from eight unrelated individuals who were unaffected by Laing myopathy were supplied by the Department of Neuropathology, Royal Perth Hospital. Demographic information was not provided, so these individuals were broadly classified as 'Western Australian residents'. All biopsy muscle tissue was placed in liquid nitrogen immediately upon collection, then stored at -80°C.

Two grinding methods were used. Initially, a frozen muscle biopsy was placed under liquid nitrogen and approximately 50 mg of muscle tissue was cut from it. Fatty or connective tissue was avoided. The muscle samples were ground to a powder in the mortars, using an electric screwdriver to drive steel pestles. The steel mortars were engineered in a steel block, while the steel pestles were carefully engineered to fit into these steel mortars. 0.1% (v/v) Valcorin (dimethylcarbonate; Bayer AG) was used to wash all implements that were not autoclaved, including the mortars and pestles. The mortars and pestles were cooled to -80°C and placed under liquid nitrogen upon removal from the -80°C freezer, prior to their use in grinding muscle tissue. The second grinding method used the same technique to obtain 50 mg samples of frozen muscle tissue, but these were then placed in Valcorin treated glass test tube mortars on dry ice. Hand-operated sintered-glass pestles were used to disrupt the muscle tissue. This was an easier method and was used for approximately 50% of the muscle RNA extractions.

At the completion of grinding by either method, powdered muscle was transferred into 0.75 ml TRI-REAGENT (Molecular Research Centre, Inc.) in a 1.5 ml eppendorf tube. The TRI-REAGENT RNA extraction method is a modification of the guanidinium thiocyanate and acidic phenol method (Chomczynski, 1993). The homogenised muscle tissue was left to stand in TRI-REAGENT for five minutes at room temperature. 0.2 ml chloroform was then added to each tube and all were shaken vigorously for 15 seconds. These samples were allowed to stand for 15 minutes at room temperature before centrifugation at 13000 rpm for 15 minutes at 4°C. Centrifugation throughout RNA and DNA preparation was done in an Eppendorf Centrifuge 5415C, tube lid-hinges were always upwards. An 18-tube rotor was used for 1.5 ml tubes while a 24-tube rotor was used for 0.65 ml tubes.



The mixture of homogenised muscle tissue, TRI-REAGENT and chloroform separated into three layers: a lower, red, phenol-chloroform organic phase; an interphase; and a colourless aqueous upper phase. DNA and proteins were in the organic phase and interphase, while RNA remained exclusively in the aqueous phase. The aqueous phase was transferred into a new 1.5 ml eppendorf tube and mixed with 0.5 ml isopropyl alcohol. These samples were stored at room temperature for 10 minutes, then centrifuged at 13000 rpm for 20 minutes at 4°C. RNA appeared as a gel-like pellet at the bottom each eppendorf tube, on the side directly under the lid-hinge.

The supernatant was carefully removed from the tube and the RNA pellet vortexed in 1 ml 75% ethanol for 15 seconds, then centrifuged at 7500 rpm for 5 minutes at 4°C. The supernatant was again removed and RNA pellets allowed to 'almost dry' at room temperature with eppendorf tube lids open. The drying of pellets was monitored throughout to avoid complete drying, as this greatly decreases RNA solubility. RNA pellets were dissolved in 50 µl ddH<sub>2</sub>O (Baxter) by incubating for 12 minutes at 60°C, then stored in a freezer at -80°C.

#### **7.3.1.2 Skeletal Muscle cDNA Manufacture**

The enzyme reverse transcriptase uses the 3' end of short double stranded regions on mRNA to initiate the synthesis of complementary DNA (Retzel et al, 1980). The double stranded region was provided by adding hexamer oligodeoxyribonucleotides of random sequence (Promega), and allowing them to anneal to the mRNA. Their random sequence meant that these random hexamers annealed at many sites along the the mRNA sequence.

Muscle cDNA was manufactured as follows: 5  $\mu$ l template mRNA, 4  $\mu$ l of random hexamers (250 ng/ $\mu$ l) and 7  $\mu$ l ddH<sub>2</sub>O, were heated to 90°C, snap-chilled and condensation spun down with a pulse spin at 13000 rpm. This step removed all double stranded RNA loops. 6  $\mu$ l Promega RT (Reverse Transcriptase) buffer, 6  $\mu$ l dNTPs (5 mM) and 2  $\mu$ l Maloney murine Leukaemia virus-RT (MMLV-RT) were then added. The reaction tube was held at room temperature for 5 minutes, 37°C for 1 hour and 90°C for 2 minutes (to make sscDNA). It was snap-chilled on ice to retain single-strandedness, then condensation was spun down with a pulse spin at 13000 rpm. Since MMLV-RT (Promega) is RNase H minus, the product of this cDNA reaction contains template mRNA.

The presence of cDNA in these preparations was assayed by amplifying a section of the copper zinc superoxide dismutase (SOD1) cDNA. A strong signal indicated that a cDNA preparation was suitable template for *MYH6* and *MYH7* amplification. A 20  $\mu$ l, standard SOD PCR reaction comprised: 4  $\mu$ l 5 X buffer; 50 ng each SOD primer (1  $\mu$ l); 5  $\mu$ l template cDNA; 8.9  $\mu$ l ddH<sub>2</sub>O; and 0.5 u *Tth* polymerase (0.1  $\mu$ l) (Biotech International). Each ml of 5X buffer contained: 500  $\mu$ l 10X *Tth* buffer (supplied with the enzyme); 400  $\mu$ l 25 mM MgCl<sub>2</sub>; 40  $\mu$ l 25 mM dNTPs; and 60  $\mu$ l ddH<sub>2</sub>O. This buffer provided the 'low-stringency' reaction concentrations of 2 mM MgCl<sub>2</sub> and 200  $\mu$ M dNTP, considered appropriate for this assay.

SOD1 PCR reactions were initially denatured at 94°C for 2 minutes 30 seconds. Followed by 34 cycles of: 30 seconds denaturation at 94°C; 45 seconds primer annealing at 58°C; and 45 seconds amplification at 72°C.

### 7.3.1.3 *MYH6* and *MYH7* Primer Design

Table C7.1 in Appendix C presents a comparison of the *MYH6* (Epp et al., 1993) and the *MYH7* (Jaenicke et al., 1990) exon sequences, both of which were derived from their published genomic sequences. The application program used to make this comparison was *SeqEd*, Version 1.0.3 (Myers, 1992: Applied Biosystems International) (ABI). *MYH6* and *MYH7* cDNAs do not have coincident base numbers because these genes have different lengths of 5' and 3' untranslated regions, and *MYH7* has four deleted codons compared to *MYH6*. However, for the purposes of comparison, they are aligned in Table C7.1, even although this has the effect of 'misnumbering' the bases of both these genes. The complete exonic sequence of *MYH7* is 6008 bp while that of *MYH6* is 5939 bp.

Comparison of the 5939 bp of the *MYH6* cDNA to the 6008 bp of the *MYH7* cDNA revealed 635 bp differences where these sequences overlapped, equivalent to a 10.7% difference (635/5939). Similarly, comparison of the 5820 bp translated region of *MYH6* (including the stop codon) to the 5808 bp translated region of *MYH7* revealed 554 bp differences, equivalent to a 9.52% difference (554/5820). The 12 bases in the four deleted codons of *MYH7* were included in this comparison, because they were important for transcript identification and primer design. A comparison of the 1939 amino acids of  $\alpha$ -MyHC coded by *MYH6* to the 1935 amino acids of  $\beta$ -MyHC coded by *MYH7*, also presented in Table C7.1, showed there to be 142 different residues, equivalent to a 7.32% difference (142/1939).

The 92.68% homology between human  $\alpha$ - and  $\beta$ -MyHC determined in this way was close to the 93.2% identity reported for mammalian cardiac  $\alpha$ - and

$\beta$ -MyHC isoforms (Weiss, 1999, et al., p. 63). These authors also report that mammalian cardiac MyHC isoforms have only 77-81% identity with the MyHC skeletal isoforms. *MYH6* and *MYH7* had greater homology with each other than to the other known MYH genes. The design of primers around differences between *MYH6* and *MYH7* was the most stringent approach possible.

The genetic code and the universal single-letter amino acid code in Table C7.2 in Appendix C, allowed a detailed examination of Table C7.1. This showed that only 196 of the 635 bp differences between *MYH6* and *MYH7* caused amino acid codon changes. There were more base changes (196) than actual amino acid alterations (142) because the four deleted codons in *MYH7* contributed 12 bases to the 196 base differences between these genes but only 'changed' 4 amino acids, and because some codons contained more than one base change, each of which independently altered amino acid sequence.

It was possible that the Western Australian individuals under investigation might not possess all of the 493 'neutral' nucleotide differences and that some novel 'neutral' nucleotide differences might be present in this population. Accordingly, primers for PCR amplification and sequencing for both *MYH6* and *MYH7* were designed around the maximum number of sequence differences between these genes. In particular, each primer was designed such that its 3' base differed to the equivalent base in the other cardiac MYH gene, and wherever practicable, *MYH6*/*MYH7* amino acid codon changes were included in exonic primers. It was thought unlikely that exonic primers designed in this way would amplify other members of the MYH multigene family. In addition to these cDNA exonic *MYH6* and *MYH7* primers, two intronic *MYH7* primers were designed to amplify *MYH7* exon 35 from lymphocyte gDNA.

7.3.1.4 MYH6 Primers

To ascertain the risk of simultaneously amplifying MYH6 and MYH7, the possible expression of the MYH6 gene in biopsied muscle spindles was tested by designing two primers specific for MYH6. Table 7.1 identifies the two MYH6 primers designed to specifically amplify MYH6. Both the forward primer, MYH6 1838F, and the reverse primer, MYH6 2127R, each had five differences to the equivalent sequences in MYH7, and both their 3' bases were different to equivalent bases in MYH7. The MYH6 1838F<--->2127R amplicon was called amplicon 12.

Table 7.1 MYH6 cDNA/Exonic Primers	
Primer	Sequence
1838F	5'-C ATC CTG GGC TGG CTG GAA AAA-3'
2127R	5'-C CAT CAC CCC TGG AGC CTT CCG-3'
All primers are identified by their 5' base number. This is the same number as the complementary MYH6 exonic base number. Translated MYH6 primer sequences are shown in DNA 'codons', referred to as codons throughout the rest of this thesis since RNA codons are not discussed. Black identifies identical bases in MYH6 (Epp et al., 1993) and MYH7 (Jaenicke et al., 1990). Red identifies base differences resulting in different amino acids in $\alpha$ - and $\beta$ -MyHC. Blue identifies base differences not altering amino acids in $\alpha$ -MyHC when compared to $\beta$ -MyHC.	

7.3.1.5 MYH7 Primers

Table C7.3 in Appendix C presents the double-stranded sequence of MYH7 and locates the primers used to manufacture amplicons of this gene. Table 7.2 (p. 186) shows the exonic MYH7 primers used. It identifies base differences between MYH6 and MYH7 which altered amino acid codons and those base changes which did not change the amino acid sequence.



**Table 7.2 MYH7 cDNA/Exonic Primers**

Amplicon	Primers	Primer Sequences
1	42F	5'- <i>ccctt</i> <b>ccctcatctg</b> tagacac-3'
	1041R	5'- <b>TGGT</b> CTC TCC TTG <b>GGA GAT</b> GAA T-3'
2	887F	5'-T CTG GAA AAA TCC <b>AGA GTT</b> ATT-3'
	2015R	5'-GGA <b>CGA GCC TTT CTT</b> <b>GGC CTT</b> G-3'
3	1943F	5'- <b>C</b> CTG TTT <b>GCC AAC</b> TAT GCT <b>GGG</b> -3'
	3404R	5'-CTT CTG <b>CAG CTG GCT</b> <b>GCC GAG</b> G-3'
4	3335F	5'- <b>A</b> GAC TTT GAG <b>CTG AAT</b> <b>GCT CTC</b> -3'
	3894R	5'- <b>C</b> AGA <b>ACG CTG GGT</b> CTC <b>CTC CGC</b> -3'
5	3832F	5'- <b>GC</b> CGG ACC TTG GAA GAC CAG <b>ATG</b> -3'
	4659R	5'-C ATG <b>GAT AGT</b> CTT TCC <b>GCT GGA</b> -3'
6	4610F	5'- <b>G</b> ATC TCC GAC <b>TTG</b> ACT GAG CAG <b>TTG</b> -3'
	5975R	5'- <i>ctccaaggagctgttacacag</i> -3'
7	1145F	5'- <b>A</b> GGC <b>GCC</b> ATC ATG CAC <b>TTT</b> GGA-3'
	1589R	5'-GTA CTC CTC CTG CTC CAG CAC <b>A</b> -3'
8	1472F	5'-C GCT GGC TTC GAG ATC TTC <b>GAT</b> -3'
	2091R	5'-A GTG <b>GGG</b> ATG GGT <b>GGA GCG</b> CAA-3'
9	4550F	5'- <b>T</b> GAG GAG TCC CTG GAA CAT <b>CTG</b> -3'
	5072R	5'-ACG GAC <b>TGC</b> ATC <b>GTC</b> CAG CTG <b>A</b> -3'
10	4897F	5'-AG ACA CGC AGC CGC AAC GAG <b>GC</b> -3'
	5354R	5'-CTT GGC CTT CTC CTC <b>AGC ATT</b> C-3'
	396F	5'-AAC CTC AAG <b>GAT</b> CGC TAC <b>GGC</b> T-3'
	2248F	5'-GG TAT CGC ATC CTG AAC CCA <b>GC</b> -3'
	2492F	5'- <b>A</b> AAG CT <b>G</b> CTG GAA CGT AGA <b>GAC</b> T-3'
	2782F	5'- <b>CA</b> GAT GCT GAG GAG CGC TGT <b>GAT</b> -3'
	3137F	5'-C AAC <b>ACC</b> CTG <b>ACT</b> AAG <b>GCC</b> AAA-3'
	3335F	5'- <b>A</b> GAC TTT GAG <b>CTG AAT</b> <b>GCT CTC</b> -3'
	5268F	5'-AAG AAG ATG <b>GAT</b> <b>GCT</b> GAC CTG <b>T</b> -3'
	365R	5'-T CAG CAT GGC CAT GTC CTC <b>C</b> <b>GAT</b> T-3'
	611R	5'- <b>C</b> AG GAT GGA CTG GTT <b>TTC</b> <b>TCT</b> G-3'
	2639R	5'-GG <b>A</b> GGC CAT CTC CTT CTC <b>TCT</b> T-3'
	3143R	5'-GGT GTT GAC CTT GTC <b>CTC</b> CTC <b>GG</b> -3'
	3998R	5'- <b>GCC</b> TCG GGT CAG CTG <b>GGA</b> GAT <b>C</b> -3'
	5680R	5'-TG <b>C</b> AG CCG CAG CAG GTT <b>TTT</b> <b>CC</b> -3'

5' MYH7 exonic base numbers used to identify primers. Lower-case identifies 5'-UTR and 3'-UTR primer sequences. Translated MYH7 primer sequences shown in uppercase and in codons. Black identifies bases that were identical in published MYH6 sequence (Epp et al., 1993) and published MYH7 sequence (Jaenicke et al., 1990). Red identifies base differences between MYH6 and MYH7 that cause different amino acids to be present in  $\alpha$ -MyHC and  $\beta$ -MyHC. Blue identifies base differences coding for the same amino acids in  $\alpha$ -MyHC and  $\beta$ -MyHC.

Primers shown above the **grey line** were used for amplicon manufacture and sequencing, while those below the line were used only for sequencing.

The MYH7 primers in Table 7.2 contained the maximum MYH6/MYH7 sequence differences, given the need to space primers appropriately and to avoid complications such as intra- and/or inter-primer complementation.

The segregation of a base change in MYH7 exon 35 in WAMPD family members was determined using lymphocyte gDNA as a template. This was available for all the participants in this study as identified in Figure 3.1 (p. 78). Table 7.3 shows the MYH7 gDNA intronic primers designed to target intron 34, primer MYH7 23531F, and intron 35, primer MYH7 23825R. The product of MYH7 23531F<->23825R was called amplicon 11. It contained MYH7 exon 35

Table 7.3 MYH7 gDNA Intronic Primers	
Primer	Sequence
23531F	5'-aggtgcccattagtgagggaac-3'
23825R	5'-tcagggaatgagcaggggagctg-3'
5' MYH7 intronic base numbers used to identify primers. Lower-case was employed for intronic sequences. Black identifies identical bases in MYH6 (Epp et al., 1993) and MYH7 (Jaenicke et al., 1990). Blue identifies base differences between MYH6 and MYH7.	

### 7.3.1.6 DNA Amplification Conditions

MYH6 amplicons were obtained using standard-PCR, while MYH7 amplicons were prepared using both standard and XL-PCR conditions. A 20 µl, standard MYH6/MYH7 PCR reaction comprised: 4 µl 5 X buffer; 50 ng each primer (1 µl of each); 5 µl template cDNA or 5 µl of a 1/40 dilution of an MYH7 amplicon

(when this was used as starting template DNA), or 5  $\mu$ l template gDNA (50 ng); 8.9  $\mu$ l ddH<sub>2</sub>O; and 0.5 u Tth polymerase (0.1  $\mu$ l) (Biotech International). Each ml of 5X buffer contained: 500  $\mu$ l 10X *Tth* buffer (supplied with the enzyme); 200  $\mu$ l 25 mM MgCl<sub>2</sub>; 10  $\mu$ l 25 mM dNTPs; and 290  $\mu$ l ddH<sub>2</sub>O. This buffer provided the 'medium-stringency' reaction concentrations of 1 mM MgCl<sub>2</sub> and 50  $\mu$ M dNTP, considered appropriate for the specific amplification of *MYH6* and *MYH7* amplicons.

All standard-PCR reactions were initially denatured at 94°C for 2 minutes 30 seconds. This was followed by 34 cycles of: 30 seconds denaturation at 94°C; 45 seconds at the chosen annealing temperature; and 45 seconds amplification at 72°C. Annealing temperatures were determined from *Oligo*, Version 5.1 (Rychlik, 1996: National Biosciences Inc., Plymouth, USA) (NBI). If the *Oligo* program recommended a two-step protocol, the annealing/amplification time was increased to 1 minute.

For reaction volumes of 25  $\mu$ l, extra long PCR (XL-PCR) were: 3  $\mu$ l 3.3X reaction buffer; 2  $\mu$ l 5 mM dNTPs; 50 ng of forward and reverse primers (1  $\mu$ l of each); 1.1  $\mu$ l magnesium acetate (MgOAc) (25 mM); 6.9  $\mu$ l ddH<sub>2</sub>O; and 5  $\mu$ l template cDNA. The MgOAc and 3.3X buffer were provided by Perkin Elmer. A drop of paraffin was placed on to the top of each reaction mix. Towards the end of the primary denaturation step, a mixture of 0.5  $\mu$ l XL-r(recombinant)*Tth* polymerase (Perkin Elmer); and 4.5  $\mu$ l 3.3X buffer (Perkin Elmer) was added through the paraffin to each reaction. This 'hot-start' method was employed to preserve the XL-r*Tth* polymerase.



All XL-PCR reactions were initially denatured at 94°C for 2 minutes 30 seconds. This was followed by two sets of 20 cycles. The first set of 20 cycles comprised 15 seconds at 94°C and 8 minutes at the chosen annealing/amplification temperature. The second set of 20 cycles comprised 15 seconds at 94°C and 8 minutes plus a 15 second increment per cycle at the chosen annealing/amplification temperature. The annealing/amplification temperature was initially set at 68°C, but this was lowered to improve yield where necessary.

All amplification conditions are presented in Table 7.5 in Section 7.3.2.7 (p. 201), where it was possible to align amplification conditions, amplicons, sequencing primers and bases sequenced. All PCR products were assayed by electrophoresis on 2% agarose gels, run at 80V for 1 hour.

#### **7.3.1.7 DNA Purification**

QiaQuick buffers and spin-columns (Qiagen) were used to purify PCR products. A volume of PB buffer 5 times the volume of PCR was added to each completed PCR. This mixture, generally 120 µl, was loaded on to a spin-column sitting in a collecting tube and centrifuged at 13000 rpm for 1 minute. Since DNA binds to the spin-column matrix, the flow-through was discarded. 700 µl of 20% PE buffer (in ethanol) was placed on top of each spin-column and these were centrifuged for 1 minute at 13000 rpm. The flow-through was discarded and the columns were centrifuged again for 1 minute at 13000 rpm to collect and discard any remaining PE buffer.

Spin-columns were then transferred into labelled 1.5 ml eppendorf tubes, 30-

50  $\mu$ l of ddH<sub>2</sub>O added to the centre of each spin-column, and the columns centrifuged for 1 minute at 13000 rpm for the last time. The volume of ddH<sub>2</sub>O added was determined by the concentration of DNA estimated from gel band intensity on PCR assay gels. 30  $\mu$ l ddH<sub>2</sub>O was added for lowest DNA concentrations increasing to 50  $\mu$ l ddH<sub>2</sub>O for the highest DNA concentrations. DNA bound to the spin-column matrix dissolved in this ddH<sub>2</sub>O as it moved through the column into the supporting eppendorf tube in the last centrifugation. This purification method removed all PCR reagents and DNA fragments smaller than 100 bases in length.

#### **7.3.1.8 DNA Sequencing Reactions**

The ABI PRISM BigDye Terminator Cycle Sequencing Ready Reaction Kit (PE Biosystems) was used to sequence amplicons. This kit, based on the dideoxy sequencing method (Sanger, 1981), provided AmpliTaqDNA Polymerase FS, dye terminators, deoxynucleoside triphosphates, MgCl<sub>2</sub> and buffer, all in a single tube. A 10  $\mu$ l, BigDye sequencing reaction comprised: 4  $\mu$ l Ready Reaction (also known as 'terminator mix'); 50 ng (1  $\mu$ l) of either a forward or a reverse primer; 2-4  $\mu$ l of template dsDNA; and 1-3  $\mu$ l ddH<sub>2</sub>O. DNA sequencing reactions require 100-200 ng template dsDNA (BigDye Terminator Cycle Sequencing Protocol (nd), p. 3-2). The concentration of purified dsDNA *MYH6* and *MYH7* was estimated from band intensity on agarose gels, measured against standards of known concentration. The volumes of template DNA and ddH<sub>2</sub>O added to sequencing reactions were adjusted according to these estimated DNA concentrations. When necessary, purified dsDNA was concentrated by vacuum centrifugation to a minimum volume of 10  $\mu$ l, enough for one forward and one reverse reaction.

Excess dye terminators make base calling difficult during electrophoretic analysis of sequencing reactions. Unincorporated dye terminators were removed from sequencing extension products by ethanol/sodium acetate precipitation. 2 µl of 3 M sodium acetate (pH 4.6) and 50 µl of 95% ethanol were added to each completed sequencing reaction. Tubes were shaken and left to stand at room temperature for 5 minutes, then centrifuged for 45 minutes at 13000 rpm. Supernatants were discarded, 150 µl of 70% ethanol added to each pellet and tubes centrifuged at 13000 rpm for ten minutes. Supernatants were again discarded and pellets briefly dried at 65°C in a thermal cycler. These were forwarded to the research assistant in charge of sequencing at ANRI, Lori Blechynden. Extension products were analysed by electrophoresis on an ABI 373 Automated DNA Sequencer with a BigDye Filter Wheel Upgrade. Sequencing chromatograms were then imported into *SeqEd*, Version 1.0.3 (Myers, 1992: ABI) for further analysis.

#### **7.3.1.9 SSCP Screening of *MYH7* gDNA**

*MYH7* amplicon 11, 23531F<->23825R, was amplified from lymphocyte gDNA template, using standard PCR (Section 7.3.1.6, pp. 187-188) with an annealing temperature of 64°C. Eight trial SSCP gels established the most suitable conditions for screening an *MYH7* amplicon 11 polymorphism segregating with Laing myopathy. For 12% and 8% gels, the trials were: 19:1 (acrylamide:bisacrylamide) at 4°C and 350V; 29:1 at 4°C and 280V; 75:1, at room temperature (~22°C) and 200V; 99:1 at room temperature and 130V. All SSCP gels were run over four hours in TBE (pH8.3) gel-running buffer. The gel most clearly discriminating the G23628/C23628 SSCP was the 8%, 19:1

gel run at 350V in the cold room (~4°C). Further optimisation was achieved by buffer replacement to maintain electrophoresis temperature below 22°C. Electrophoresis tanks were XCell Sure Lock (Novex) and gels were either silver-stained (Bassam et al., 1991) or stained with Sybr Gold (Molecular Probes Inc. USA).

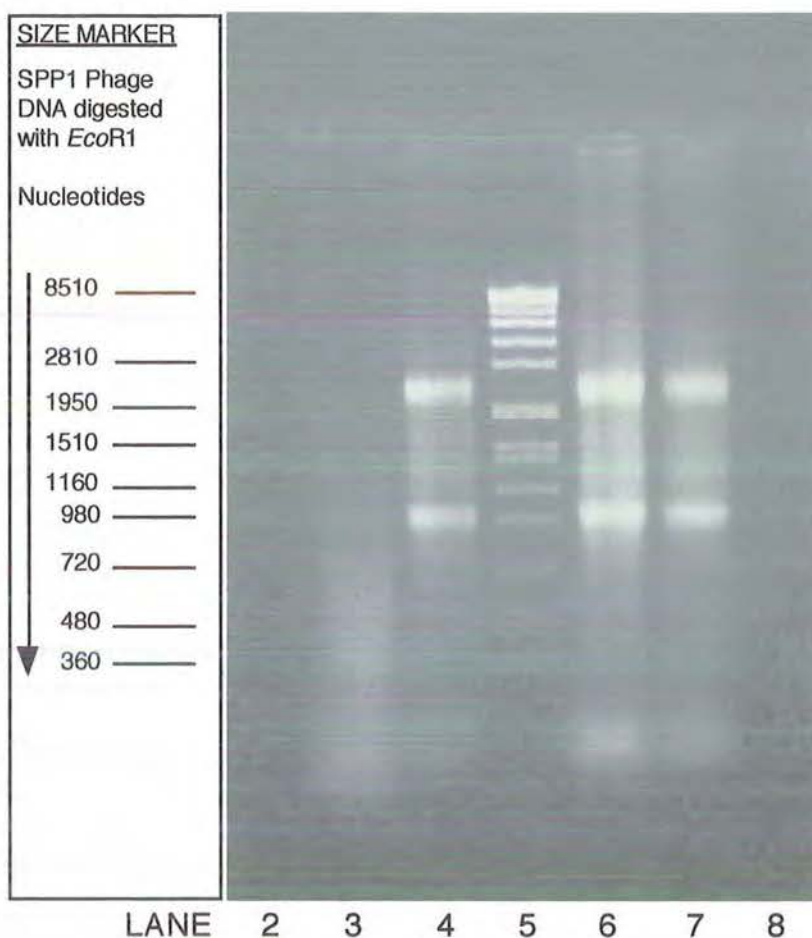
### **7.3.2 Results**

This section reports the isolation of skeletal muscle RNA, the PCR amplification of *MYH6* and *MYH7*, the sequences of *MYH6* and *MYH7* amplicons and the differences between the published and the determined sequence of the WAMPD, III:2 *MYH7*.

#### **7.3.2.1 Skeletal Muscle RNA**

The success of RNA extraction was estimated from the bands on an agarose gel. This check was necessary because, unlike DNA, RNA is unstable due to its adjacent 2' OH group and 3'-5' phosphodiester bond. At alkaline pH, the 2' hydroxyl group acts as a nucleophile, transesterifying the 5' hydroxyl and 2'-3' phosphodiester bond, which breaks the RNA polymer into fragments (Sambrook et al., 1989, 7.3 & 7.6). At neutral pH, RNA is also degraded by RNases, present through contamination or within the starting tissue.

Good preparations of RNA were distinguished by the presence of two clearly defined bands of ribosomal RNA (rRNA) and another low molecular weight broader band of transfer RNA (tRNA) at 75-120 nucleotides. These bands were used to determine the quality of an RNA sample. Figure 7.1 shows representative TRI-REAGENT RNA extractions from human muscle.

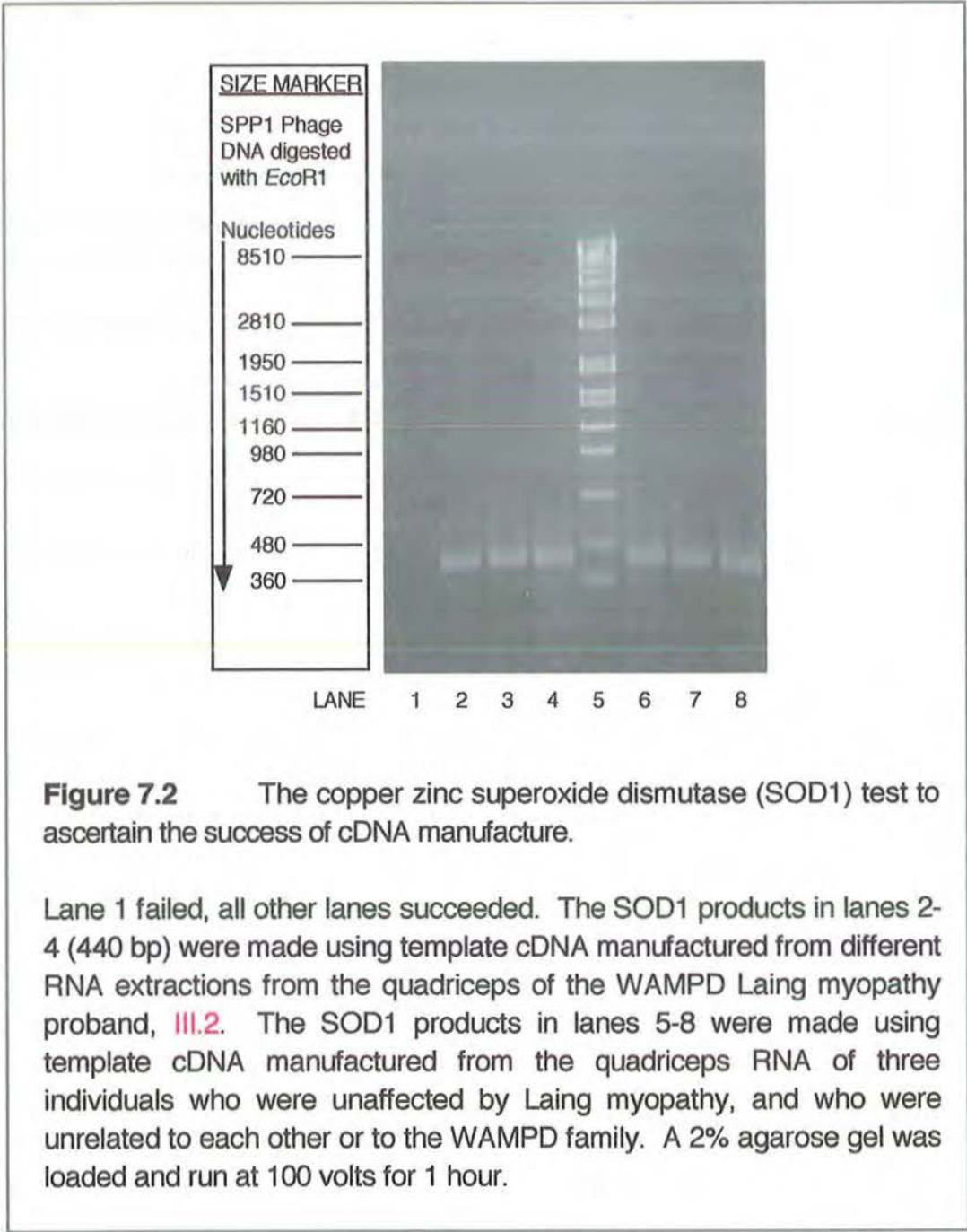


**Figure 7.1** RNA extractions from human muscle using the TRI-REAGENT method.

The presence of three bands in RNA extractions is accepted as an indication that all forms of RNA are intact, an important consideration since mRNA does not appear as a discrete band. Its variable molecular weight and its low representation (1-5%) in the total RNA present, make mRNA only visible as a background smear. The first of these extractions, in lane 3, was considered to be of insufficient quality to use as a template for cDNA manufacture, even although there was a background smear. The other three RNA extractions, in lanes 4, 6 and 7, were used to manufacture cDNA. In these lanes, mRNA can be seen as a background smear. A 2% agarose gel was loaded and run at 100 volts for 1 hour. Lanes 1, 2 and 8 were not loaded.

### 7.3.2.2 Skeletal Muscle cDNA

Figure 7.2 shows representative copper zinc superoxide dismutase (SOD1) products amplified from cDNA manufactured from Laing myopathy quadriceps muscle biopsy and unaffected quadriceps muscle biopsy. The SOD1 test was the only method used to determine the suitability of cDNA preparations as template for *MYH6* and *MYH7* amplifications. They were not quantified spectrophotometrically.



### 7.3.2.3 *MYH6* Amplification

The *MYH6* 1838F and 2127R primers (Table 7.1, p. 185) were selected because the 246 bp *MYH6* sequence between the 3' ends of these primers, amplicon 12, contained 62 differences to the equivalent bases in *MYH7*, including one deleted *MYH7* codon. These 62 differences are shown in Table 7.4 (p. 196). 1838F*MYH6* and 2127R*MYH6* primers gave amplicons of the right size (290 bp) from template muscle cDNA for III.2, the WAMPD proband, and from muscle cDNA for 8 unrelated, unaffected individuals.

### 7.3.2.4 *MYH6* Amplicon Sequencing

Table 7.4 shows four sequences: the published sequences for *MYH6* amplicon 12; the equivalent published *MYH7* sequence; the determined *MYH6* amplicon 12 sequence for eight unrelated, unaffected individuals; and the determined *MYH6* amplicon 12 sequence for the WAMPD proband, III.2. All *MYH6* amplicon 12 sequences were identical. Figure 7.3 (p. 197) presents part of the sequencing chromatograms for the *MYH6* amplicon 12. It shows the chromatograms for the bases 1947-1982 for the WAMPD proband, III.2, and for an unaffected, unrelated individual.

Table 7.4 and Figure 7.3 prove that *MYH6* was expressed in the biopsy of quadriceps muscle used in this project. The *MYH6* amplicon 12 was amplified from this starting source of cDNA, proving that *MYH6* mRNA was present. Real differences between affected and unaffected *MYH7* sequences could be obscured and/or apparent differences created by the simultaneous amplification of *MYH6* and *MYH7*. The *MYH7* primers listed in Table 7.2 (p. 186) were designed with as many *MYH6*/*MYH7* sequence differences as practicable, specifically to avoid this very real possibility.

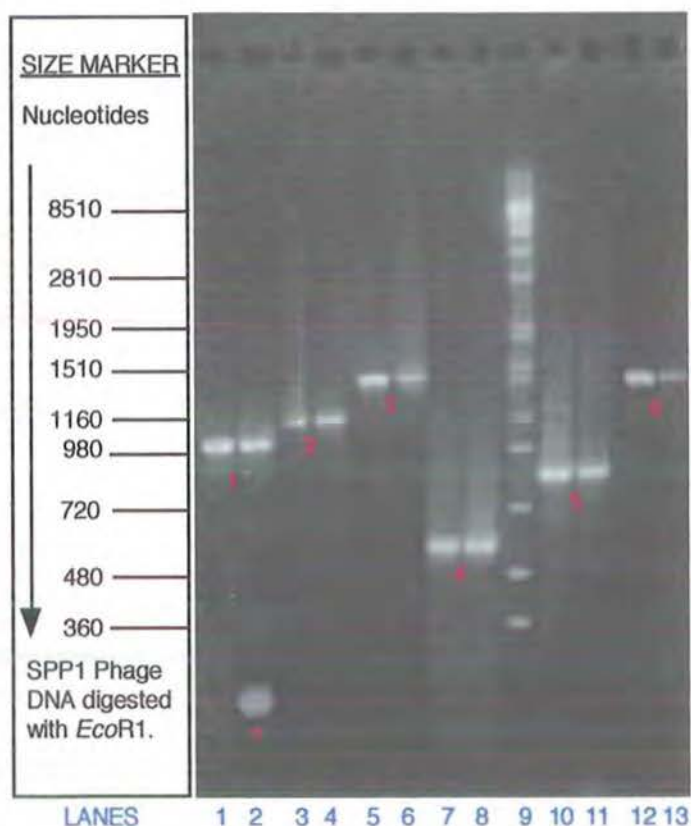


**Table 7.4 MYH6 amplicon 12 sequences. The sequence of the WAMPD proband, III.2, and 8 unaffected individuals compared to the equivalent published sequences for MYH6 and MYH7.**

	5' -----Primer 1838FMYH6 -----> 3'		
	C ATC CTG GGC TGG CTG GAA AAA		Base
Affected(III.2)MYH6	Primer sequence not determined	AAC AAG GAT CCT CTC	1874
8 Unaffected MYH6	from the chromatogram	AAC AAG GAT CCT CTC	
Published MYH6	C ATC CTG GGC TGG CTG GAA AAA	AAC AAG GAT CCT CTC	
Published MYH7	C ATC ATT GGC TGG CTG CAG AAG	AAC AAG GAT CCT CTC	1889
Affected(III.2)MYH6	AAC GAG ACT GTT GTG GCC CTG TAC CAG AAG TCC TCC		1910
8 Unaffected MYH6	AAC GAG ACT GTT GTG GCC CTG TAC CAG AAG TCC TCC		
Published MYH6	AAC GAG ACT GTT GTG GCC CTG TAC CAG AAG TCC TCC		
Published MYH7	AAT GAG ACT GTC GTG GGC TTG TAT CAG AAG TCT TCC		1935
Affected(III.2)MYH6	CTC AAG CTC ATG GCC ACT CTC TTC TCC TCC TAC GCA		1946
Unaffected MYH6	CTC AAG CTC ATG GCC ACT CTC TTC TCC TCC TAC GCA		
Published MYH6	CTC AAG CTC ATG GCC ACT CTC TTC TCC TCC TAC GCA		
Published MYH7	CTC AAG TTG CTC AGC ACC CTG TTT GCC AAC TAT GCT		1971
MYH6 sequence shown in Figure 7.3 (p. 197)			
Affected(III.2)MYH6	ACT GCC GAT ACT GGG GAC AGT GGT AAA AGC AAA GGA		1982
8 Unaffected MYH6	ACT GCC GAT ACT GGG GAC AGT GGT AAA AGC AAA GGA		1982
Published MYH6	ACT GCC GAT ACT GGG GAC AGT GGT AAA AGC AAA GGA		1982
Published MYH7	GGG GCT GAT GCG CCT +++ ATT GAG AAG GGC AAA GGC		2004
Affected(III.2)MYH6	GGC AAG AAA AAG GGC TCA TCC TTC CAG ACG GTG TCG		2018
8 Unaffected MYH6	GGC AAG AAA AAG GGC TCA TCC TTC CAG ACG GTG TCG		
Published MYH6	GGC AAG AAA AAG GGC TCA TCC TTC CAG ACG GTG TCG		
Published MYH7	AAG GCC AAG AAA GGC TCG TCC TTT CAG ACT GTG TCA		2040
Affected(III.2)MYH6	GCT CTC CAC CGG GAA AAT CTC AAC AAG CTA ATG ACC		2054
Unaffected MYH6	GCT CTC CAC CGG GAA AAT CTC AAC AAG CTA ATG ACC		
Published MYH6	GCT CTC CAC CGG GAA AAT CTC AAC AAG CTA ATG ACC		
Published MYH7	GCT CTG CAC AGG GAA AAT CTG AAC AAG CTG ATG ACC		2076
Affected(III.2)MYH6	AAC CTG AGG ACC ACC CAT CCT CAC TTT GTG CGT TGC		2090
8 Unaffected MYH6	AAC CTG AGG ACC ACC CAT CCT CAC TTT GTG CGT TGC		
Published MYH6	AAC CTG AGG ACC ACC CAT CCT CAC TTT GTG CGT TGC		
Published MYH7	AAC TTG CGC TCC ACC CAT CCC CAC TTT GTA CGT TGT		2112
	3'<----- Primer 2127RMYH6 -----5'		
	GCC TTC CGA GGT CCC CAC TAC C		
Affected(III.2)MYH6	ATC ATC CCC AAT GAG	Primer sequence not determined	
8 Unaffected MYH6	ATC ATC CCC AAT GAG	from the chromatogram	
Published MYH6	ATC ATC CCC AAT GAG CGG AAG GCT CCA GGG GTG ATG G		2127
Published MYH7	ACT ATC CCT AAT GAG ACA AAG TCT CCA GGC GTG ATG G		2149
MYH6 and MYH7 sequences presented in codons. Blue: sequence differences between MYH6 and MYH7 which do not alter amino acids. Red: sequence differences between MYH6 and MYH7 which alter amino acids. III.2: WAMPD proband. Unaffected: by Laing myopathy and unrelated to III.2. 8 'unaffecteds' had the same MYH6 sequence.			



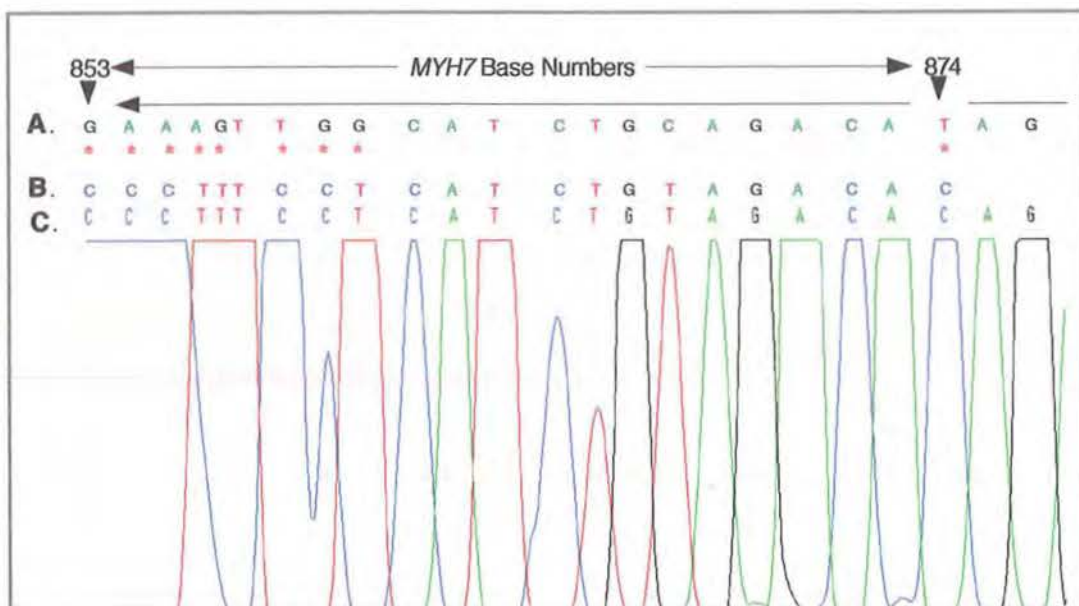




**Figure 7.4** The original six *MYH7* amplicons (p. 197) are identified by red numbers on the gel. XL-PCR and standard-PCR amplicons were run in pairs with XL products always on the left of each amplicon pair. Standard-PCR did not amplify amplicon 6 very successfully. An illegitimate amplicon (?) was identified in the standard-PCR amplicon 1, lane 2, which was not present in the XL-PCR amplicon 1 in lane 1. This illegitimate amplicon persisted up to a standard-PCR annealing temperature of 64°C, at which annealing temperature it had a *stronger* signal than that of the legitimate 1 kb amplicon 1.

The illegitimate amplicon in lane 2 was band-stabbed (Wilton, et al., 1997), re-amplified using the *MYH7* primers 42F<->1041R at a standard-PCR annealing temperature of 64°C. The product was QiaQuick purified (Qiagen) and sequenced using the reverse primer 1041R. Figure 7.5 (p. 199) shows the chromatogram of the last twenty-two 3' bases of this illegitimate 42F<->1041R amplicon.





**Figure 7.5** The reverse sequence chromatogram of the illegitimate 42F<->1041R amplicon over its last twenty-two 3' bases.

For comparison, three sequences are aligned along the top of this chromatogram. **A.** The published sequence of *MYH7* for bases 853-874. **B.** The sequence of primer 42F. **C.** The reverse sequence chromatogram of the illegitimate 42F<->1041R amplicon. \* identify base differences between the published *MYH7* sequence (**A**) (Jaenicke et al., 1990), the primer *MYH7* 42F sequence (**B**) and the determined reverse sequence of the illegitimate amplicon (**C**). It can be seen that *MYH7* primer 42F exactly matches the 42F<->1041R illegitimate amplicon sequence over its last twenty-two 3' bases.

The sequence of the illegitimate amplicon in lane 2 (Figure 7.4) showed that it was 189 bases long. Apart from its 22, 3' bases, it had the same sequence as *MYH7* from base 875 to the limit of good sequence at base 1021. This fact, and the fact that the last 22, 3' bases matched perfectly with the primer *MYH7* 42F, proved that the illegitimate 42F<->1041R amplicon, was produced by the 'false-priming' of primer *MYH7* 42F at bases 853-874. Given that primer *MYH7* 42F had only 12 complementary bases with the published 853-874 *MYH7* base sequence, this false-priming was quite remarkable. It was even more remarkable in that the 3' base of primer *MYH7* 42F was not complementary to the *MYH7* base 874.

#### **7.3.2.6      *MYH7* Amplicon Sequencing**

Table 7.5 (p. 201) lists the *MYH7* amplicons selected for sequencing, the annealing and amplification temperatures used to make each amplicon and the bases successfully sequenced in each amplicon. It includes the original six amplicons shown in Figure 7.4 (p. 198) and four further amplicons, 1145-1589 (7), 1472-2091 (8), 4550-5072 (9) and 4897-5354 (10). Two of these amplicons, 7 and 9 were necessary to clarify sequence in the areas of overlap between amplicon 2 and 3 and between amplicons 5 and 6. The other two amplicons, 8 and 9, were necessary to clarify sequence within amplicons 2 and 6. All amplicons were purified using the QiaQuick spin-column method before setting up sequencing reactions.

Internal sequencing primers were designed as required. These were identified in Table 7.2 (p. 186). The purpose of using ten overlapping amplicons as sequencing templates was to reduce the risk of simultaneous amplification of *MYH6* amplicons or other *MYH* amplicons. The presence of *MYH6* transcripts and the false-priming of 42F (Figures 7.3 & 7.5) had shown how easily this could occur. The fact that many regions of *MYH6* and *MYH7* were conserved added to this concern. Once the sequence of the 5' and 3' ends of *MYH7* amplicons 1-10 had been determined, it was clear that they were *MYH7* amplicons, therefore suitable as sequencing templates for all *MYH7* primers.

#### **7.3.2.7      Base Differences in Western Australian *MYH7* cDNA**

Table 7.6 (p. 202) lists seven base differences between the published cDNA sequence for *MYH7* and that determined for nine individuals: the Laing myopathy proband (WAMPD III.2), and eight unrelated, unaffected Western Australian residents.

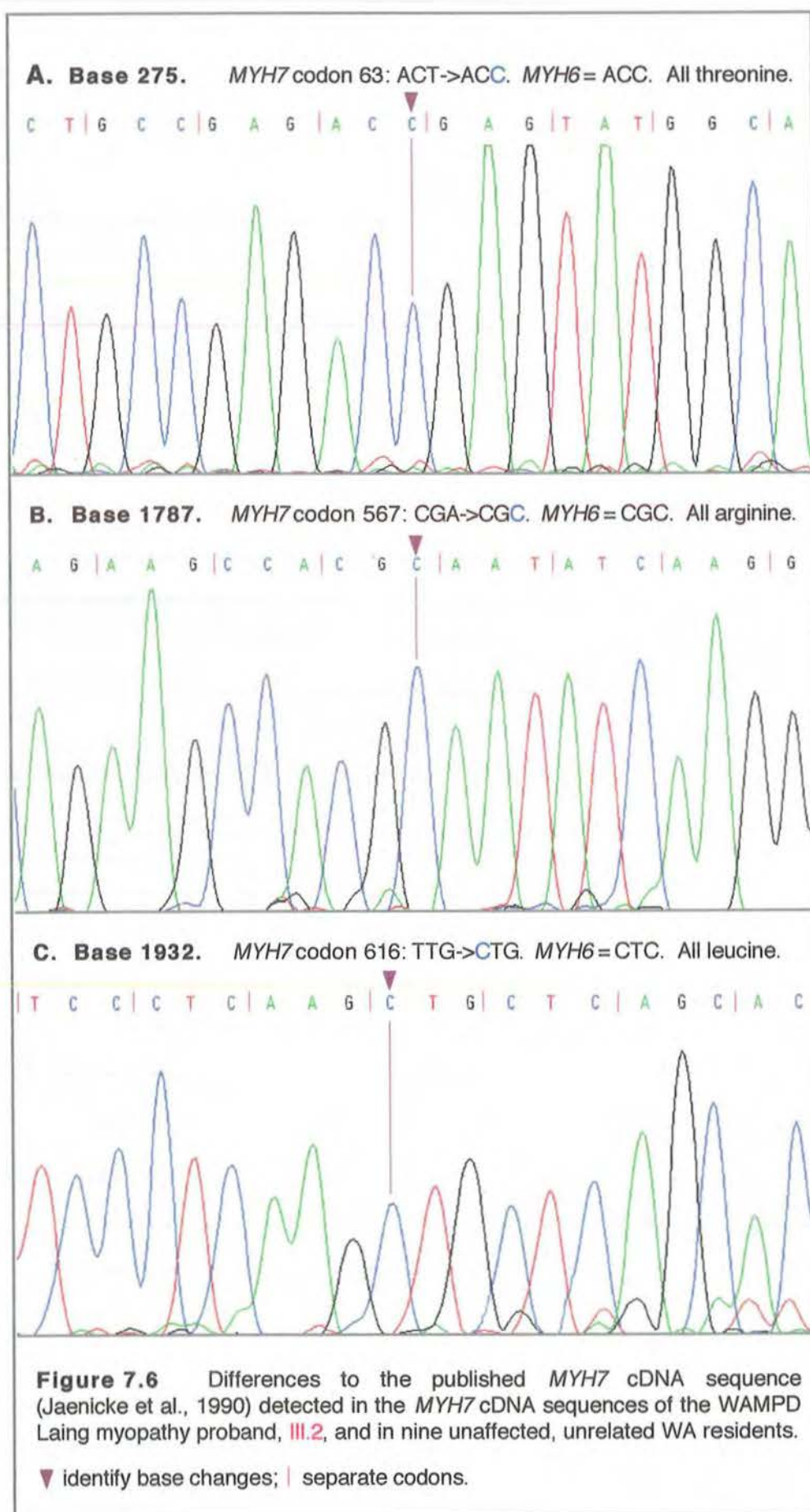
<b>Table 7.5      <i>MYH7</i> Amplicon Sequencing</b>			
<b>Template DNA (amplicon)</b>	<b>PCR Annealing/ Amplification (°C)</b>	<b>Sequencing Primer</b>	<b>Bases Sequenced</b>
42-1041 (1)	XL-68°C	365R	42-335
42-1041 (1)	XL-68°C	42F	74-501
42-1041 (1)	XL-68°C	611R	402-590
42-1041 (1)	XL-68°C	396F	437-866
42-1041 (1)	XL-68°C	1041R	589-1001
887-2015 (2)	XL-66°C	887F	930-1481
1145-1589 (7)	59°C/72°C	1145F	1175-1651
887-2015 (2)	XL-66°C	1589R	1186-1553
887-2015 (2)	XL-66°C	2015R	1470-1974
1472-2091 (8)	68°C (2 step)	1472F	1512-1602
1472-2091 (8)	68°C (2 step)	2091R	1779-2061
1943-3404 (3)	XL-68°C	1943F	1969-2493
1943-3404 (3)	XL-68°C	2639R	2185-2613
1943-3404 (3)	XL-68°C	2248F	2377-2645
1943-3404 (3)	XL-68°C	2492F	2517-2979
1943-3404 (3)	XL-68°C	3143R	2722-3102
1943-3404 (3)	XL-68°C	2782F	2815-3080
1943-3404 (3)	XL-68°C	3404R	2904-3380
1943-3404 (3)	XL-68°C	3137F	3182-3403
3335-3894 (4)	60°C/72°C	3335F	3365-3812
3335-3894 (4)	60°C/72°C	3894R	3418-3855
3832-4659 (5)	61°C/72°C	3832F	3879-4270
3832-4659 (5)	61°C/72°C	3998R	3832-3962
3832-4659 (5)	61°C/72°C	4659R	4187-4617
4550-5072 (9)	59°C/72°C	4550F	4577-4995
4550-5072 (9)	59°C/72°C	5072R	4610-5050
4610-5972 (6)	XL-66°C	4610F	4636-5063
4897-5354 (10)	60°C/72°C	5354R	4910-5326
4897-5354 (10)	60°C/72°C	4897F	4934-5224
4610-5972 (6)	XL-66°C	5680R	5277-5624
4610-5972 (6)	XL-66°C	5268F	5308-5658
4610-5972 (6)	XL-66°C	5975R	5533-5934

<b>Table 7.6</b> <b><i>MYH7 cDNA Sequence Differences in Nine Western Australian Residents</i></b>				
<b>Altered MYH7 Base/Codon</b>	<b>Published MYH7 Codon</b>	<b>Determined MYH7 Codon</b>	<b>Equivalent MYH6 Codon</b>	<b>Amino Acid</b>
275/63	ACT	ACC	ACC	All Threonine
1787/567	CGA	CGC	CGC	All Arginine
1932/616	TTG	CTG	CTC	All Leucine
2132/682	GGC	GGG	GGC	All Glycine
2837/917	GCT	GCC	GCC	All Alanine
2913/943	TTG	CTG	TTG	All Leucine
3456/1124	TCC	GCC	GCC	Serine (TCC) Alanine (GCC)
Neutral changes to published MYH7 bases (Jaenicke et al., 1990) are indicated in blue. The T3456G MYH7 base change causes the $\beta$ -MyHC S1124A amino acid change. Published MYH6 bases, Epp et al. (1993).				

Since the primary objective of this sequencing project was to determine the MYH7 cDNA sequence of WAMPD III.2, unaffected, unrelated individuals were only sequenced over regions where differences were observed between the published MYH7 cDNA sequence and that determined for WAMPD III.2.

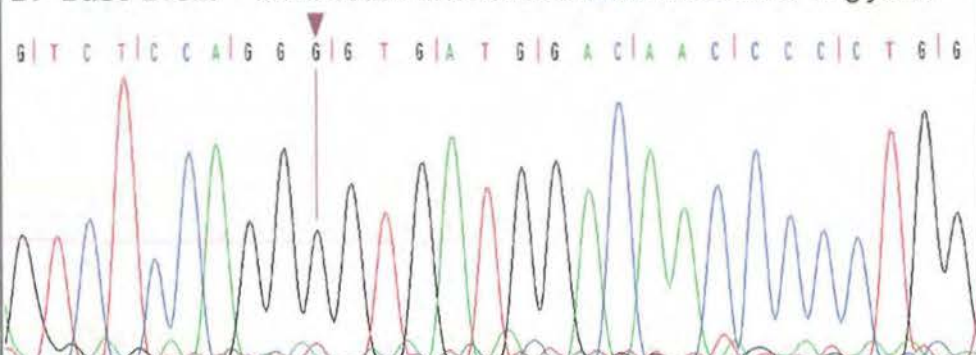
Figure 7.6 (pp. 203-204) shows that all seven base changes listed in Table 7.6 were found as homozygotes. The T3456G difference to the published MYH7 cDNA sequence was found in all the Western Australian residents tested and so could not be a disease-causing mutation. This base difference in MYH7 was the same as the published sequence for MYH6 and meant that the published  $\beta$ -MyHC had an uncharged polar serine at residue 1124 while the Western Australian form of  $\beta$ -MyHC (and  $\alpha$ -MyHC) had non-polar alanine at this residue. The change was  $\beta$ -MyHC S1124A.



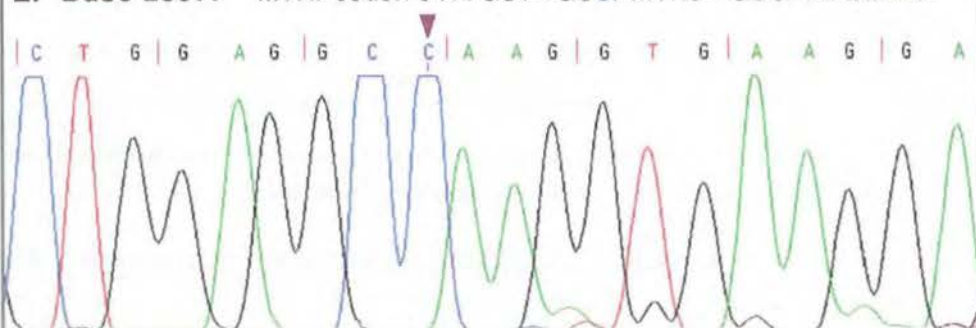


**Figure 7.6 (continued)**

**D. Base 2132.** MYH7 codon 682: GGC->GGG. MYH6 = GGC. All glycine.



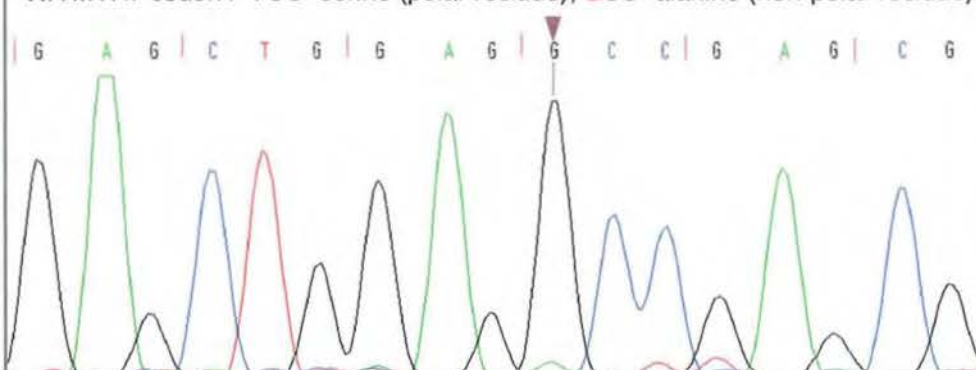
**E. Base 2837.** MYH7 codon 917: GCT->GCC. MYH6 = GCC. All alanine.



**F. Base 2913.** MYH7 codon 943: TTG->CTG. MYH6 = TTG. All leucine.



**G. Base 3456.** MYH7 codon 1124: TCC->GCC. MYH6 = GCC = same as the 'WA MYH7 codon'. TCC=serine (polar residue); GCC=alanine (non-polar residue).



▼ identifies base changes; | separates codons



Apart from the T3456G base change, all the other base changes in Table 7.6 were neutral DNA polymorphisms. Their neutrality meant that it was unlikely that they could be disease causing mutations, a fact confirmed by their presence in eight unrelated, unaffected individuals. This confirmation was important because neutral base changes can cause disease, if, for example, they create cryptic splice sites (Richard & Beckmann, 1995).

The determined base change at *MYH7* position 1932, *MYH7* codon 616, was of interest because it identified itself as the likely ancestral '616' codon for *MYH6* and *MYH7*. The Western Australian *MYH7* 616 codon (CTG) had only one base change to the published *MYH7* 616 codon (TTG), while the published equivalent *MYH6* codon (CTC) had two base changes. Given that *MYH6* and *MYH7* almost certainly arose by duplication of a common ancestor (Epp et al., 1993), and that these are highly conserved genes (Weiss et al., 1999), the ancestral '616' codon was probably the codon found in the Western Australian group. This conclusion follows from the observation that only the Western Australian CTG codon connects these three codons with just two neutral base changes: TTG--CTG-->CTC. It is also possible that one, or both, of the published codons are in error.

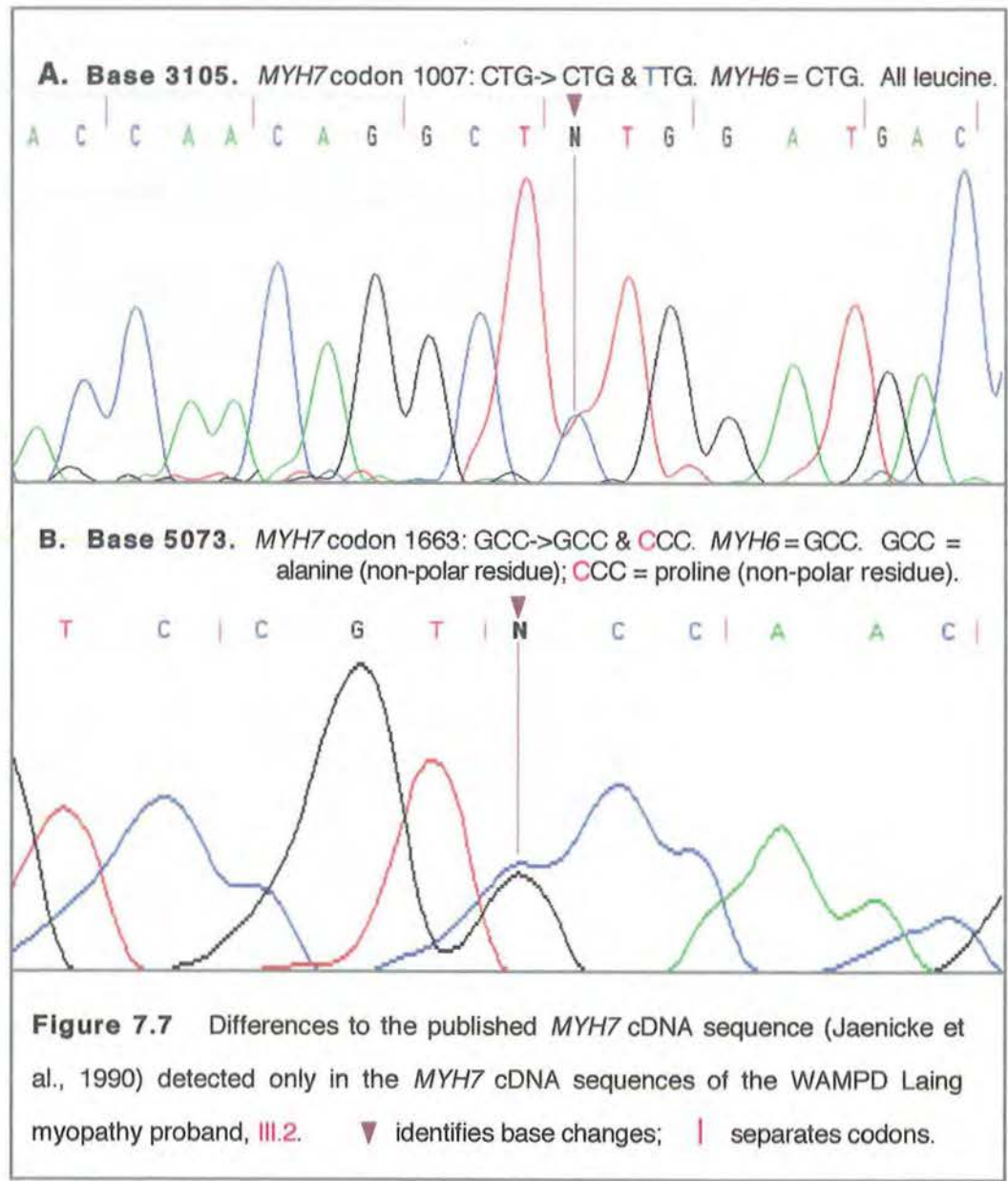
#### 7.3.2.8 Base Differences in WAMPD Proband *MYH7* cDNA

Table 7.7 (p. 206) identifies two differences to the published *MYH7* sequence, which were only found in the cDNA of WAMPD Laing Myopathy proband. Figure 7.7 (p. 206) shows that both these changes were found as heterozygotes with the published sequence. The base change at position 3105 altered codon 1007 in WAMPD III.2 from the published CTG to TTG but this codon change did not alter amino acid 1007, leucine, in  $\beta$ -MyHC.

**Table 7.7**      **MYH7 Sequence Differences in Proband III.2**

Altered MYH7 Base/Codon	Published MYH7 Codon	Determined MYH7 Codon	Equivalent MYH6 Codon	Amino Acid
3105/1007	CTG	CTG/TTG	CTG	All Leucine
5073/1663	GCC	GCC/CCC	GCC	Alanine (GCC) Proline (CCC)

Neutral changes to published MYH7 bases (Jaenicke et al., 1990) are indicated in blue. The MYH7 base change causing  $\beta$ -MyHC A1663P indicated in red. Published MYH6 bases from Epp et al. (1993).



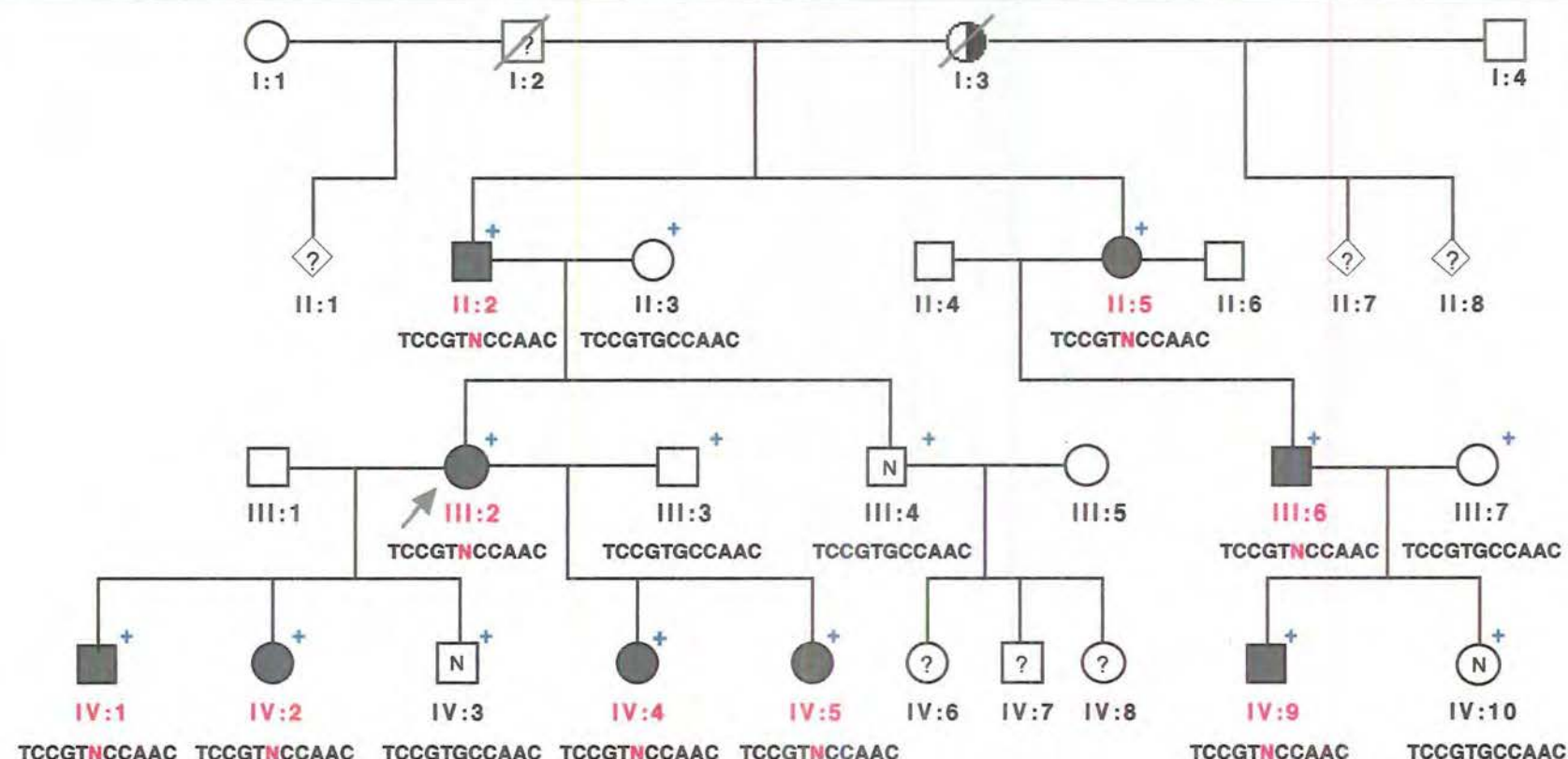


The *MYH7* C3105T change was not observed in unaffected, unrelated Western Australians. However, it was a neutral change and it did not create a cryptic splice site since it occurred in the sequence GCT TTG GAT. It was unlikely to be a disease causing mutation and was not investigated further. Its discovery was significant because it was the first heterozygous base encountered in the WAMPD III.3 *MYH7* sequence. It was the first indication that two alleles were being sequenced. The *MYH7* G5073C base change in exon 35, was not present in any of the tested unrelated, unaffected cDNA. It caused the  $\beta$ -MyHC amino acid substitution, A1663P. This was a significant change. Residue 1663 was in a  $\alpha$ -helical section of  $\beta$ -MyHC. Proline has a reputation as an ' $\alpha$ -helix breaker' (Schimmel & Florey, 1968).

#### 7.3.2.9 *MYH7* gDNA G23628C Segregation in WAMPD Family

In gDNA, *MYH7* G23628C is equivalent to *MYH7* cDNA G5073C. To establish whether *MYH7* gDNA G23628C segregated with all WAMPD family members or with the Laing myopathy affected family members alone, the entire WAMPD family was tested, by sequencing, for *MYH7* exon 35 G23628C heterozygosity. gDNA was available for all WAMPD family members. *MYH7* G23628C is located 12 bases into exon 35. Exon 35 is 204 nucleotides in length, from base 23595 to base 23798 in *MYH7* gDNA. The amplification and sequencing of the *MYH7* exon 35 PCR product, amplicon 11, was carried out with the intronic primers 23531F and 23825R (Table 7.3, p. 187), at an annealing temperature of 64°C.

Figure 7.8 (p. 208), Figure 7.9 (p. 209) and Figure 7.10 (p. 210) all show the segregation of *MYH7* G23628C heterozygosity with Laing myopathy.

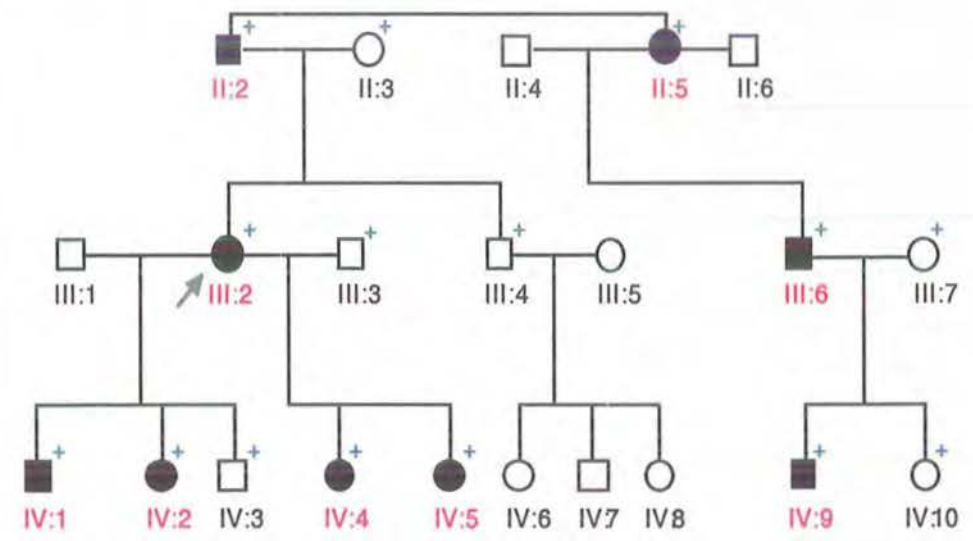


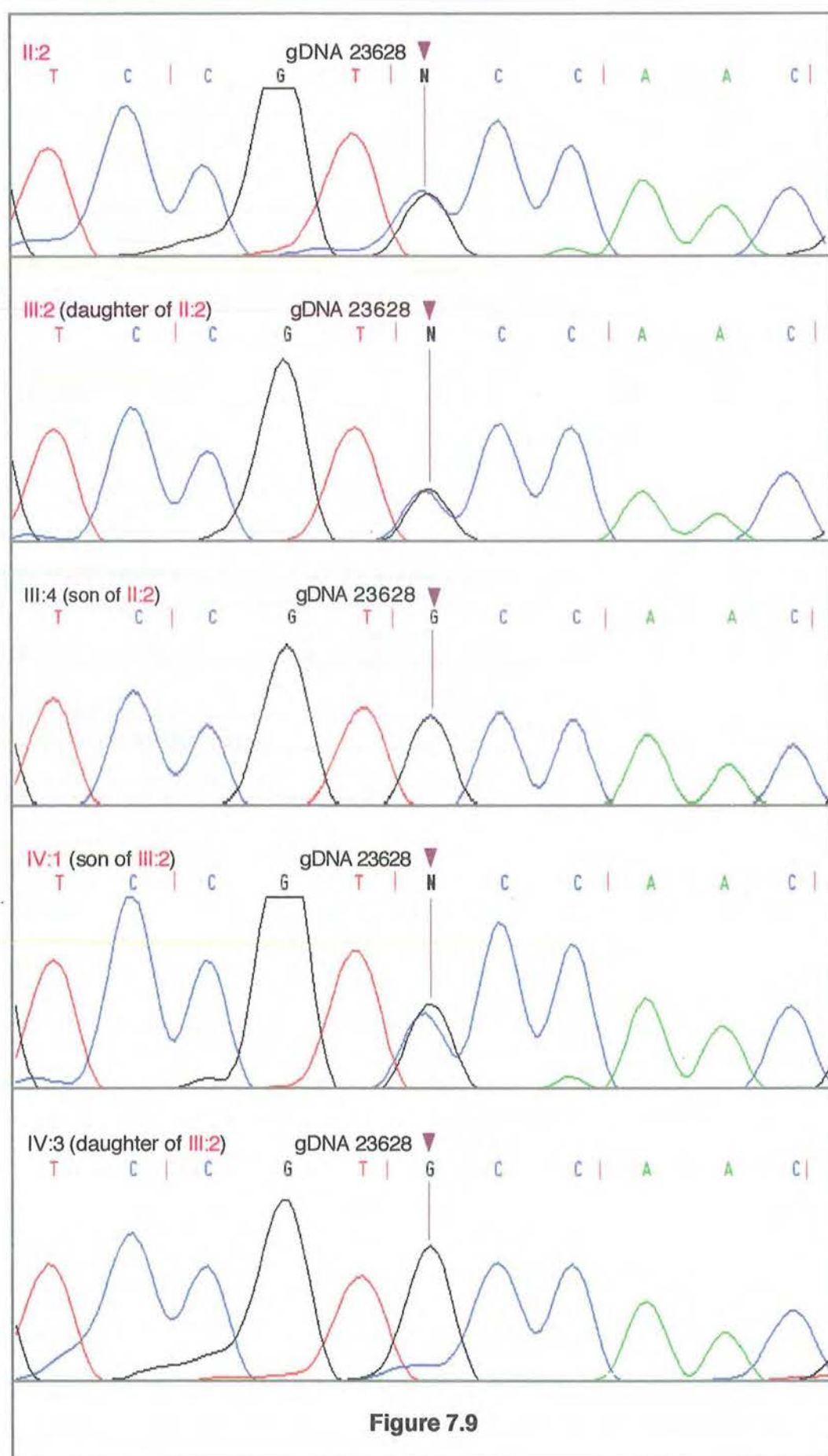
**Figure 7.8** The segregation of the gDNA, *MYH7* 23623-23633 base sequence in the WAMPD pedigree. *SeqEd*, Version 1.0.3 (Myers, 1992; ABI) was used to convert chromatograms into nucleotide sequences. The *N*23628 base segregating with Laing myopathy was a G/C heterozygote which the ABI 373 DNA Sequencer could not call. Figure 7.9 (pp. 209) shows five WAMPD *MYH7* chromatograms from bases 23623 to 23633, while Figure 7.10 (p. 210) shows the segregation of *MYH7* 23623-23633 in the WAMPD pedigree, from III:2 to IV:1 and IV:3.

**Figure 7.9** The chromatograms of *MYH7* gDNA sequence 23623<->23633 in selected members of the WAMPD family. **II.2**, **III.2**, **III.4**, **IV.1** and **IV.3** are shown. Red pedigree numbers identify Laing myopathy individuals. Sequencing template was made from lymphocyte gDNA with primers 23531F and 23825R.

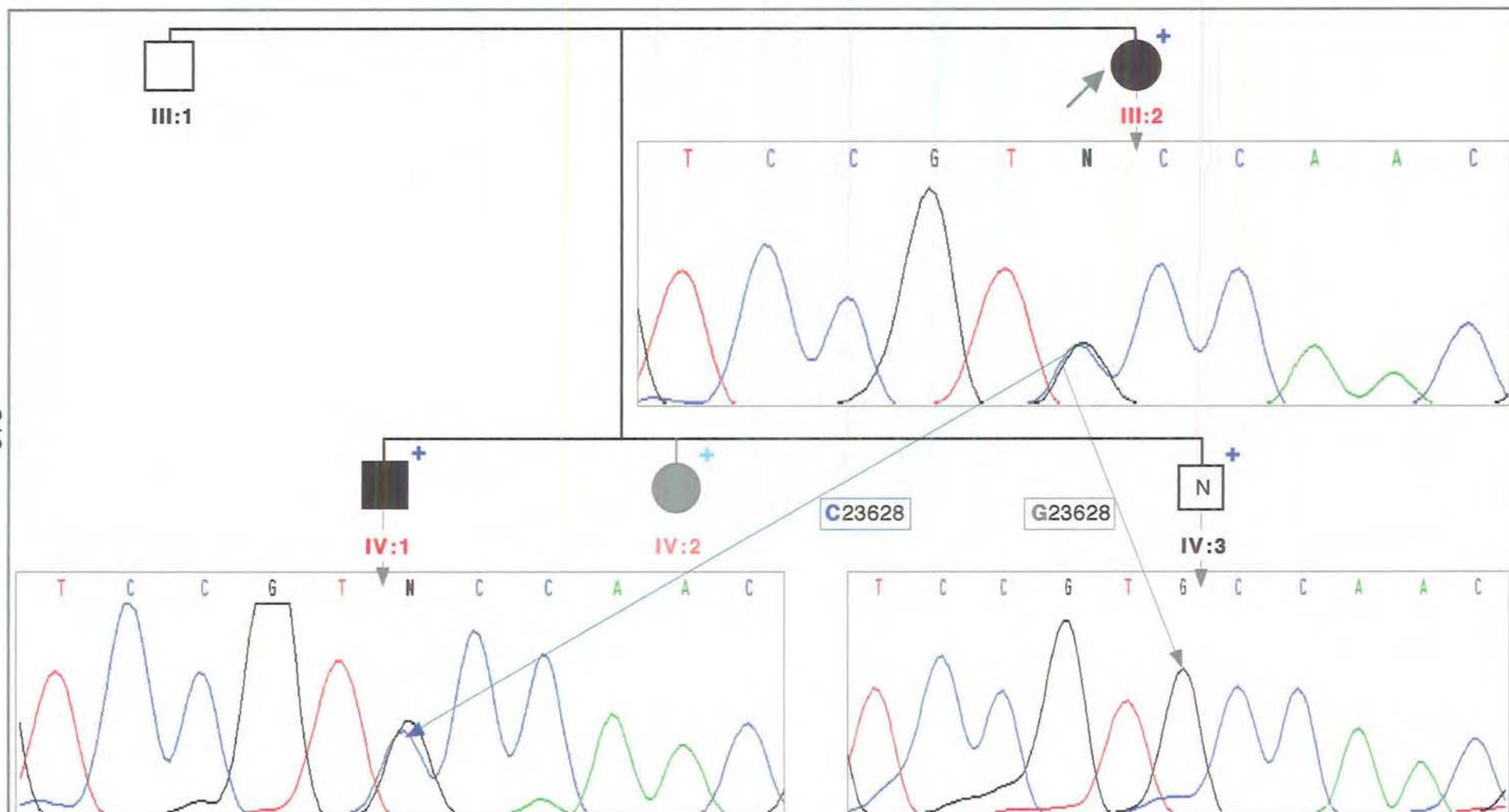
▼ identifies heterozygous gDNA base 23628, in affected individuals;  
| separates codons.

**WAMPD Pedigree**









**Figure 7.10** The segregation of *MYH7* 23623-23633 in the WAMPD pedigree, from III:2 to IV:1 and IV:3. The IV:2 chromatogram is not shown.

### 7.3.2.10 Testing Two European Distal Myopathy Families for *MYH7* cDNA G5073C/gDNA G23628C

The discovery of *MYH7* G23628C segregating in other distal myopathy families would strengthen its claim to being the Laing myopathy mutation. At the time of its discovery we were aware of European distal myopathy families with phenotypes similar to Laing myopathy. These were the five Italian families reviewed by Scoppetta et al. (1995), and the German family reported by Voit et al. (1998). The Voit family and one of the Scoppetta families have also been reviewed more recently (Voit et al., 2001). The five Scoppetta distal myopathy families had a remarkably homogeneous clinical picture with AD inheritance, onset in infancy or childhood with peroneal muscles weakness, benign progression, possible late involvement of muscles other than tibio-peroneal muscles, usually normal serum CK, EMG evidence of primary myogenic damage and morphological findings of non-specific myopathy. These authors described this form of distal myopathy as infantile AD-distal myopathy.

The Voit distal myopathy family comprised a father and his three children from two marriages. All were affected by distal myopathy. The father's distal myopathy onset at 15 months when it was noticed that he fell frequently. Subsequently, he could walk and run but never walk on his heels. In the third decade, weakness of hand extensors and finger extensors developed followed by hip extensor and generalised muscle weakness in the fourth decade. All his three children had infantile onset of distal myopathy, marked by weakness of the tibialis anterior and toe extensor muscles. Magnetic Resonance Imaging revealed complete wasting of the tibialis anterior and toe extensor muscles in his 15 year old daughter. These authors described this form of distal myopathy as AD infantile onset tibial myopathy.



The AD mode of inheritance, the age of onset and the pattern of muscle involvement of these European distal myopathy families was similar to that of Laing myopathy. Testing them for *MYH7* G23628C and /or gDNA/G5073C cDNA heterozygosity appeared worthwhile. In addition, and importantly, these families did not show exclusion of the MPD1 region of chromosome 14 by linkage analysis (Professor Dr. Thomas Voit, personal communication, 2000). Professor Voit, Director, Paediatric Clinic of the University of Essen, Germany, forwarded the gDNA of one affected individual in one of the Italian (Scoppetta) distal myopathy families and the gDNA and cDNA from one affected member of the German (Voit) distal myopathy family. These were tested in the same way as Laing myopathy DNA was tested for the *MYH7* cDNA G5073C, and gDNA G23628C. Neither co-segregated cDNA G5073C/gDNA G23628C with their distal myopathy (Figure 7.11, p. 213).

#### **7.3.2.11 SSCP Screening of *MYH7* gDNA for G23628C**

Figure 7.12A (p. 214) shows the WAMPD family SSCP gel for *MYH7* gDNA amplicon 11. The conditions employed were described in Section 7.3.1.9 (pp. 191-192). The segregation of an SSCP with Laing myopathy was evident. A comparison of all the WAMPD sequences for the *MYH7* gDNA amplicon 11, showed G23628C to be the only sequence difference between affected and unaffected WAMPD family members. The observed SSCP was caused by G23628C segregating with Laing myopathy.

A further 122 unrelated unaffected individuals were tested by SSCP for *MYH7* G23628C. Figure 7.12 B and 7.12 C (both p. 214) show two typical SSCP gels loaded with 23531F<->23825R *MYH7* gDNA amplicon 11 from unrelated individuals, all of whom were unaffected by Laing myopathy.

**Figure 7.11** *MYH7* cDNA amplicon 10 and *MYH7* gDNA amplicon 11 sequences for two European families segregating early onset AD-MPD. Chromatograms for the *MYH7* cDNA bases 5068-5078 and for the *MYH7* gDNA bases 23623-23633 are shown.

**A.** The *MYH7* cDNA amplicon 10 sequence of one MPD affected member in a German distal myopathy family (Voit) showed that this individual was homozygous *MYH7* G5073. This agreed with the published sequence (Jaenicke et al., 1990). Reverse sequence is shown.

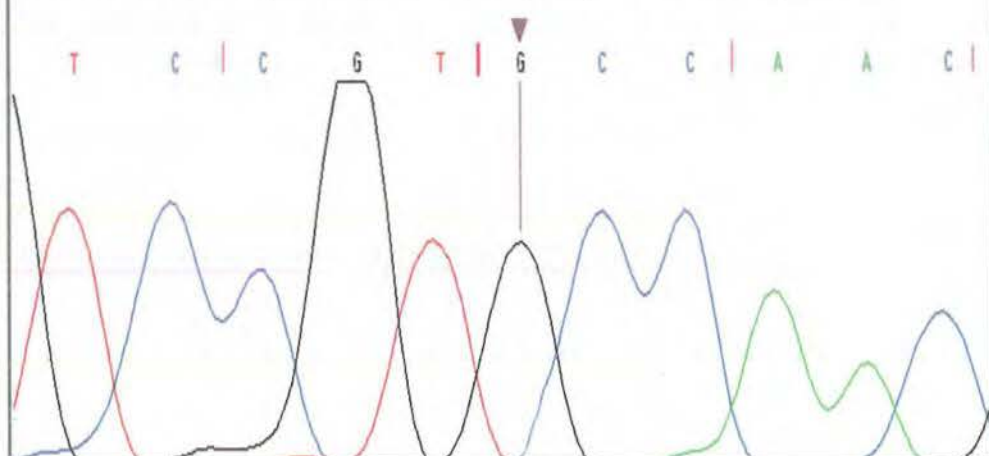
**B.** The *MYH7* gDNA amplicon 11 sequence using template gDNA from the same affected German family member, confirmed that this individual was homozygous *MYH7* G23628. Reverse sequence is shown.

**C.** The *MYH7* gDNA amplicon 11 sequence from an MPD affected member of an Italian family (Scoppetta), also showed that this individual was homozygous *MYH7* G23628. Reverse sequence is shown.

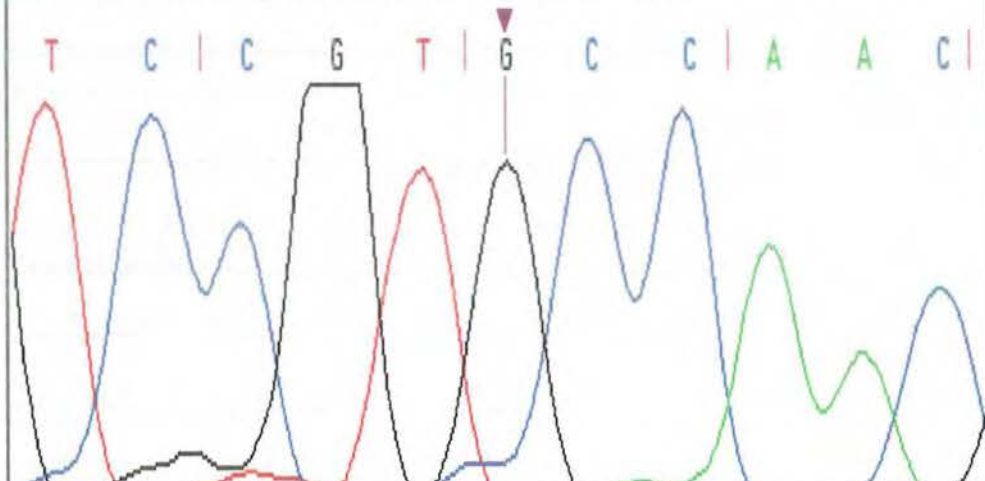
From these results it was concluded that neither this German distal myopathy family (Voit et al., 1998), nor this Italian distal myopathy family (Scoppetta et al., 1995), segregated the *MYH7* cDNA G5073C/gDNA G23628C heterozygosity with their form(s) of distal myopathy.

▼ identifies base changes; | separates codons.

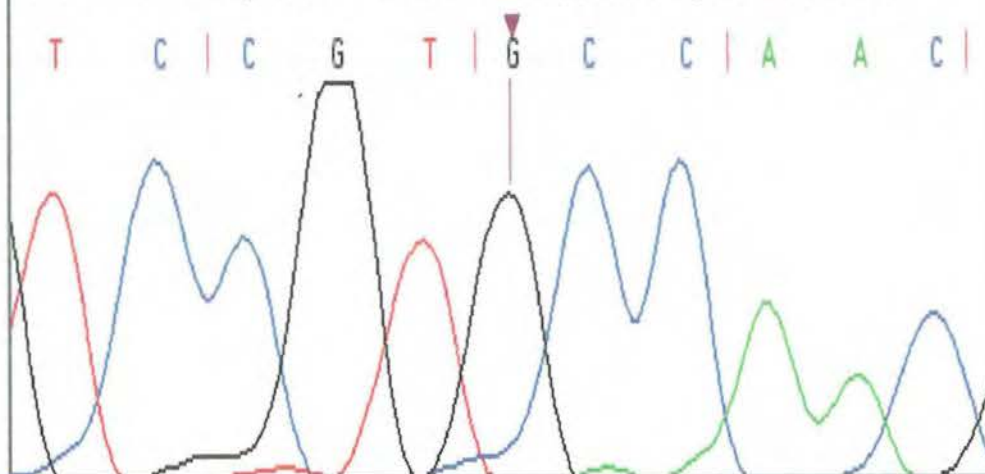
**A. German Early-Onset AD-MPD (Voit) - cDNA amplicon 10.**



**B. German Early-Onset AD-MPD (Voit) - gDNA amplicon 11**



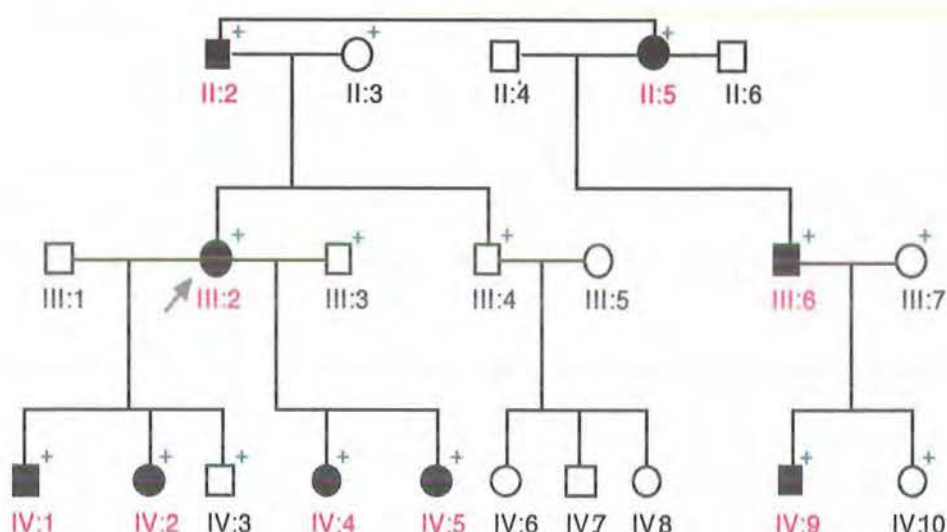
**C. Italian Early-Onset AD-MPD (Scoppetta) - gDNA amplicon 11**



**Figure 7.11**

**Figure 7.12** Screening *MYH7* gDNA 23531F<->23825R amplicon 11 for the G23628C SSCP. This SSCP is visible in the bottom band of each of these gels.

**A.** Shows the segregation of the *MYH7* G23628C SSCP in the WAMPD family with Laing myopathy (red pedigree numbers). This gel was silver stained. **IV:9** was of interest because it appeared as if he may have been homozygous C23628, even although sequencing had previously shown this individual to be heterozygous G23628C. PCR mutations are rare, but they do occur. However, sequencing the actual **IV:9** amplicon 11 (23531F<->23825R) used on this SSCP gel showed that **IV:9** was G23628C heterozygous. The entire sequence of this **IV:9** amplicon was identical to all other Laing myopathy individuals.



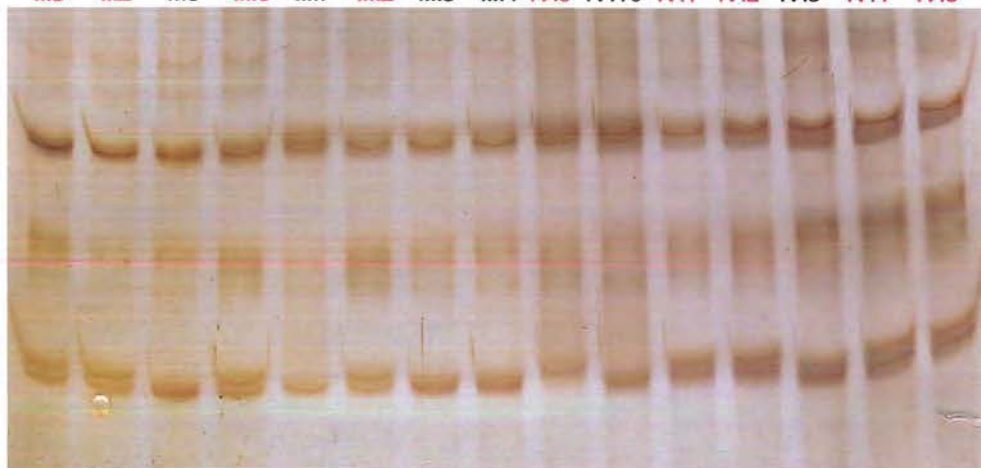
**B.** An *MYH7* G23628C SSCP gel of the *MYH7* gDNA amplicon 11 showing 12 unaffected, unrelated individuals. Unaffected individuals were numbered from 1N to 12N. This was the first 'unaffected' gel, showing the SSCP results for individuals 1N-12N. The positive control was WAMPD Laing myopathy individual **II:2**. This gel was stained with Sybr Gold.

**C** An *MYH7* G23628C gel showing the SSCP results for individuals 80N-89N. This gel was also stained with Sybr Gold. Initially one positive control was loaded on each SSCP gel, but this was expanded to 3 positive controls to increase confidence. One of the positive controls used on this gel was individual **IV:9**. This time, the G23628C SSCP could be identified in **IV:9**. Such is SSCP.



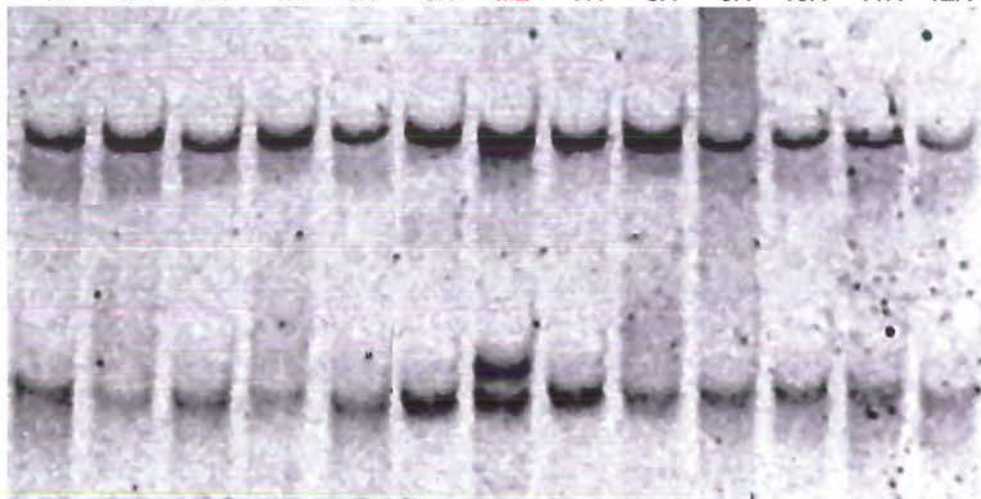
**A.**

II:5 II:2 II:3 III:6 III:7 III:2 III:3 III:4 IV:9 IV:10 IV:1 IV:2 IV:3 IV:4 IV:5



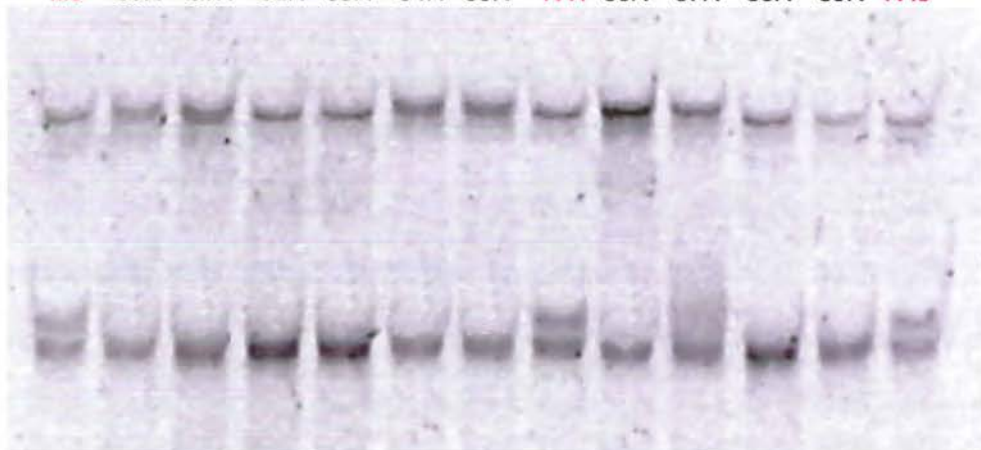
**B.**

1N 2N 3N 4N 5N 6N II:2 7N 8N 9N 10N 11N 12N



**C.**

II:5 80N 81N 82N 83N 84N 85N IV:1 86N 87N 88N 89N IV:9



A total of 256 unaffected chromosomes were tested for the *MYH7* SSCP G23628C segregating with Laing myopathy. This sample comprised 122 unrelated, unaffected individuals and the six unaffected WAMPD family members. None were found to have this SSCP.

In addition, ten of the unaffected, unrelated individuals were sequenced over the *MYH7* amplicon 11 from base 23531 to 23825. These were individuals 2N, 6N, 12N, 21N, 39N, 53N, 76N, 82N, 92N, and 100N. None showed G23628C heterozygosity, confirming the validity of the SSCP test.

#### 7.4 CONCLUSION

Testing *MYH7* cDNA of the WAMPD proband, III.2, for a Laing myopathy mutation, identified seven base differences between the WAMPD III.2 sequence and the published cDNA sequence for *MYH7*. All of these differences were found in eight unrelated, unaffected Western Australian residents. None of them could be disease-causing.

Two further differences to the published *MYH7* sequence segregated exclusively with cDNA of WAMPD III.2. An *MYH7* C3105T base change and an *MYH7* G5073C. Only the *MYH7* G5073C in exon 35, caused a significant change to  $\beta$ -MyHC, resulting in an A1663P substitution in a repeating  $\alpha$ -helical section of the  $\beta$ -MyHC rod. The significance of this change was that proline has been described as an ' $\alpha$ -helix breaker' (Schimmel & Florey, 1968) Sequencing demonstrated that gDNA G23628/C23628 heterozygosity segregated with Laing myopathy in the WAMPD pedigree. This segregation was confirmed by an SSCP test which was then used to test a total of 256 unaffected chromosomes. None of these controls showed G23628/C23628 heterozygosity.

Two patients from European distal myopathy families, phenotypically similar to Laing myopathy (Voit and Scoppetta), were tested for the presence of *MYH7* gDNA G23628/C23628 heterozygosity. One of these patients (Voit) was also tested for *MYH7* cDNA G5073/C5073 heterozygosity. The gDNA of both patients was homozygous G23628. The cDNA of the Voit family patient was homozygous G5073.

Chapter 8 evaluates the *MYH7* G5073C base change as the Laing myopathy mutation.

## **CHAPTER EIGHT**

### **MYH7 G5073C: THE LAING MYOPATHY MUTATION OR A RARE POLYMORPHISM?**



## **8.0 MYH7 G5073C: THE LAING MYOPATHY MUTATION OR A RARE POLYMORPHISM?**

The  $\beta$ -MyHC A1663P substitution segregating with Laing myopathy raises the question, 'Could the  $\beta$ -MyHC A1663P substitution cause Laing myopathy?' It was the most important question raised by this thesis. It is addressed in this chapter.

The effect of the  $\beta$ -MyHC A1663P substitution will be analysed in two ways. First, a survey of disease-causing mutations in MYH genes and other relevant sarcomeric genes, will be conducted to gauge the likelihood that this substitution could cause Laing myopathy. Difficulties confronting the  $\beta$ -MyHC A1663P Laing myopathy candidature and possible disease mechanisms will be identified and analysed. Second, the possible effect(s) of the  $\beta$ -MyHC A1663P substitution will be deduced from first principles: at the primary, secondary, quaternary structural levels of  $\beta$ -MyHC and at superstructural level of the thick filament. Future research directions for Laing myopathy may be indicated by either, or both, of these analyses.

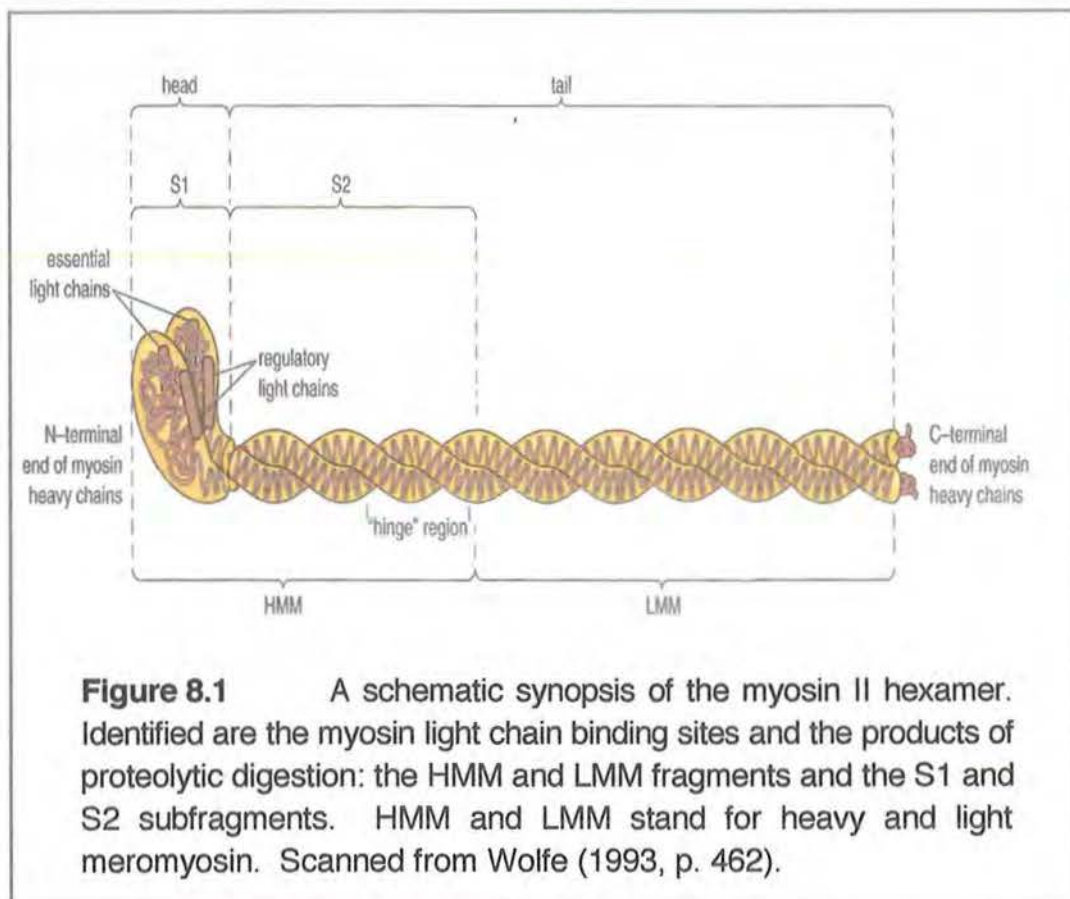
### **8.1 LAING MYOPATHY: SARCOMERIC GENE MUTATIONS**

Laing myopathy has three important identifying characteristics relevant to this analysis. It is not associated with heart disease; it is a skeletal muscle myopathy; and it affects specific muscles. MYH disease-causing mutations with any of these characteristics are of most interest, especially if they occur in the LMM region of the MyHC molecule. This analysis starts, therefore, by asking a series of questions about mutations in *MYH7*, then it broadens its enquiry to include other myosins and other sarcomeric genes.

### 8.1.1 Do *MYH7* Mutations Have a Cardiac Phenotype?

$\beta$ -MyHC is the predominant ventricular isoform, and mutations in  $\beta$ -MyHC cause hypertrophic cardiomyopathy (CMH) (Redwood et al., 1999). Any aetiology for Laing myopathy based on the *MYH7* G5073C mutation, must be able to explain why heart disease has not been detected in affected WAMPD individuals.

CMH is a genetically and phenotypically heterogeneous disease associated with mutations in nine sarcomeric protein genes (listed in Table 8.1 pp. 225, where these genes are more fully considered). Mutations in *MYH7* cause CMH1 (MIM:192600). All CMH1 mutations segregate in an AD fashion, and all lie in the first 2847 bases (949 codons) of the 5805 coding bases in *MYH7*. The structure of myosin II helps to explain why this is observed.

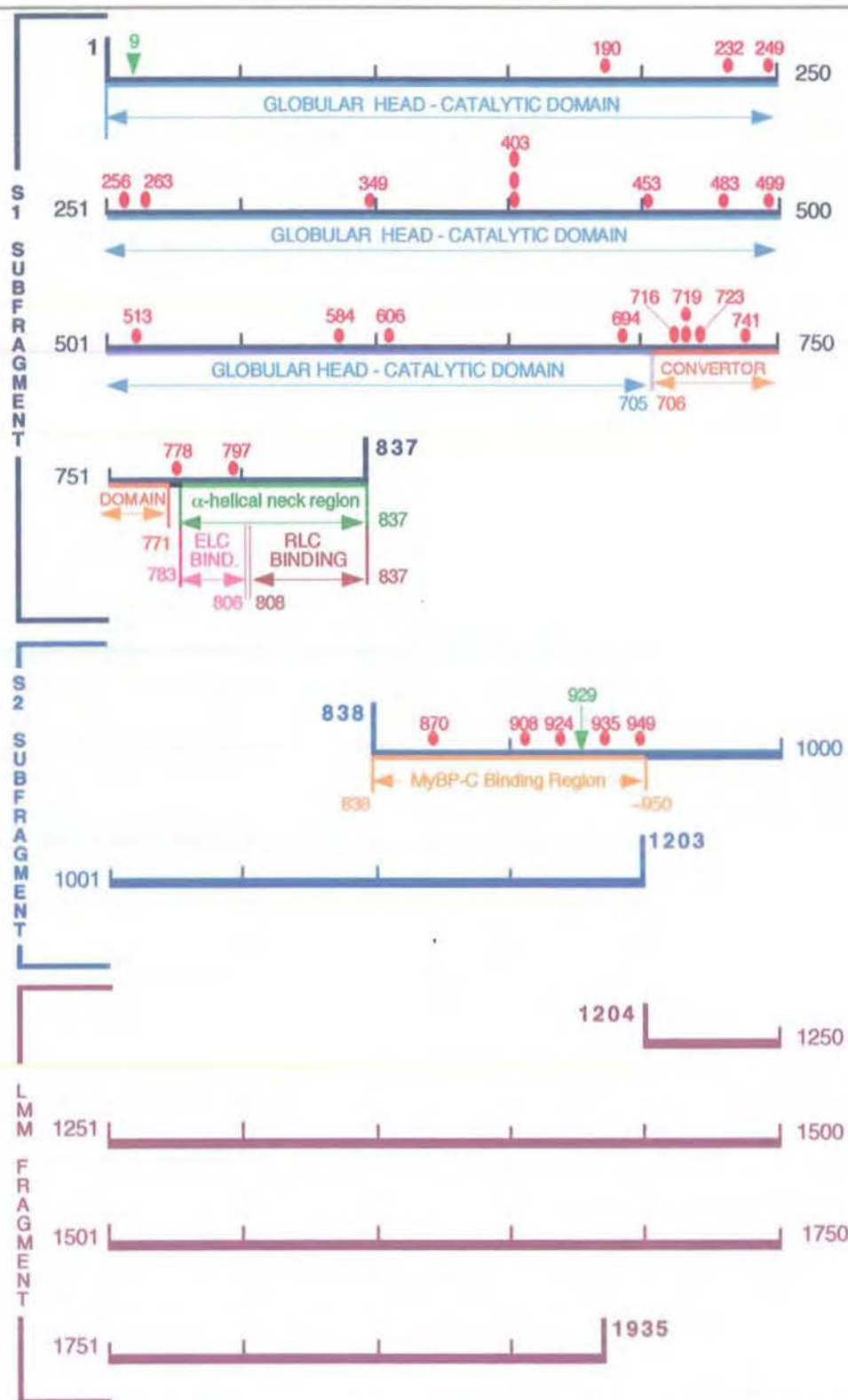


The two  $\beta$ -MyHC monomers in MyHCs comprise one proteolytic fragment, the LMM fragment and two proteolytic subfragments S1 and S2. These fragments/subfragments are not functionally significant, but are still widely used for descriptive purposes.

The S1 subfragment consists of the globular head, the convertor domain and an  $\alpha$ -helical neck region. The globular head is the catalytic/motor domain of the MyHCs. It contains sites for binding actin, ATPase, a nucleotide binding cleft for binding ATP, and a long narrow D464-M495 communication/transducer domain between the ATP and actin binding sites (Weiss et al., 1999, p. 64). The globular head connects to the second communication/transducer domain, the convertor domain, which runs from residue 706 to 771 and appears to communicate activity at the nucleotide binding site to the next section of MyHC molecule, the 783-843 residue  $\alpha$ -helical neck region. The essential light chain (ELC), and the regulatory light chain, (RLC) bind to the  $\alpha$ -helical neck region. The neck region functions as a lever arm, the length of which is correlated with the power stroke size and velocity of movement (Uyeda et al., 1996).

The S2 subfragment comprises the N-terminal one-third of the rod and contains a myosin binding protein-C binding site at its N-terminal end, and a hinge-like region at its C-terminal end. The hinge-like region facilitates movement of the neck and head during a power stroke (Weiss et al., 1999, p. 70). The C-terminal two-thirds of the rod, the LMM fragment, is responsible for MyHC self-association and forms the backbone of the thick filament (Weiss et al., 1999, p. 70). It commences between residues 1203 and 1282, its starting point depending on specificity of cleaving enzyme.

Figure 8.2 superimposes CMH1 mutations on an *MYH7* cartoon to produce a codon mutation map.



**Figure 8.2** Superimposition of CMH1 mutations on a cartoon of the *MYH7* gene. Mutations were identified from the Human Gene Mutation Database (2000; <http://archive.uwcm.ac.uk/uwcm/mg/hgmd0.html>). In this codon mutation map, red numbers identify codons with missense mutations while green numbers identify the two known CMH1 codon deletions. All other numbers are *MYH7* codon/ $\beta$ -cardiac myosin residue numbers.

Figure 8.2 shows that 27 different CMH1 mutations can result in heart disease. Since 21 of these CMH1 mutations lie in the head of  $\beta$ -MyHC, between *MYH7* cDNA bases 25-2847, it follows that CMH is predominantly caused by disruption of the functions carried out by the head, or S1 subfragment, of  $\beta$ -MyHC (Vikstrom & Leinwand, 1996). Kelly and Strauss (1994) suggested that *MYH7* mutations represent dominant negatives by disturbing contractile function, despite the remaining normal allele producing normal protein. Nishi et al., (1995) showed that while an *MYH7* missense mutation caused CMH1, a *MYH7* nonsense mutation segregating in the same pedigree had no effect. This showed that null-*MYH7* alleles do not produce CMH1.

On the basis of many *in vivo* and *in vitro* motility assays of  $\beta$ -MyHC A403Q and *in vitro* motility assays of many other CMH1 mutations, Redwood et al. (1999, p. 25) concluded that CMH1 mutations appear to lead, "to a primary hypocontractile state, which, in turn, provides a stimulus for the development of compensatory hypertrophy. . . . The most straightforward hypothesis for the pathogenesis of the disease is that the reduced force provides a stimulus for compensatory hypertrophy."

This analysis was encouraging for the candidature of *MYH7* G5073C as the Laing myopathy mutation, because all but six of these CMH mutations were in the globular head of the  $\beta$ -MyHC molecule. The six CMH mutations that were in the rod region could also be explained in a way that did not compromise the likelihood of *MYH7* G5073C as the Laing myopathy mutation. These mutations were at codons 870, 908, 924, 929, 935 and 949. Gruen and Gautel (1999) discovered a binding site towards the N-terminal end of the myosin binding protein-C (MyBP-C), which had high affinity for the S2 region of the MyHC rod, in the 126 residues immediately following the head tail junction. All of the six  $\beta$ -MyHC S2 rod region CMH1 mutations lie in the MyBP-C binding

region. Binding studies involving wild type MyBP-C and *MYH7* with the mutations R870H and E924K showed that these *MYH7* mutations greatly reduced binding of MyBP-C to the  $\beta$ -MyHC S2 rod region. Gruen and Gautel concluded that the disruption of the MyBP-C binding site by the six S2 rod region CMH1 mutations, was the cause of CMH1 in these patients.

### **8.1.2 Do *MYH7* Mutations Cause Skeletal Muscle Myopathy?**

A myopathy was discovered during the histological examination of soleus biopsies taken from CMH1 patients in five kindreds. These kindreds were segregating four different *MYH7* mutations (Fananapazir et al., 1993, pp. 3994-3997). Twelve patients showed mild myopathic changes, four showed moderate myopathic changes: variation in fibre size, hypertrophic fibres with internal nuclei, increased connective tissue, splitting of muscle fibres, and occasional necrotic fibres. Clinical or electrophysiological signs of myopathy were not described. The  $\beta$ -MyHC amino acid substitutions associated with soleus muscle myopathy were, L908V (in the S2 rod), R403Q, and G256E.

Of special interest to this project was the observation that the  $\beta$ -MyHC S2 rod substitution, L908V, segregated from a father with CMH1 phenotype and with soleus muscle myopathic changes to his daughter and son who were 33 and 27 years old respectively at the time of their ascertainment. Both the daughter and the son showed muscle myopathic changes, but not CMH1 phenotype. The son had three children, two daughters and one son aged 5, 4, and 2. All these children had soleus muscle central core phenotype. Muscle central core phenotype is also a variably expressed part of nemaline myopathy (Afifi et al., 1965) and malignant hyperthermia (Frank et al., 1978). In this CMH1 kindred it is therefore apparent that the  $\beta$ -MyHC L908V substitution has co-segregated with unspecified muscle myopathic changes and CMH1 phenotype, with

unspecified muscle myopathic changes alone and with soleus muscle central core phenotype. Echocardiogram was used to determine the presence or absence of CMH1 phenotype. Fananapazir (1993) also found that soleus muscle central core phenotype was present in a further 7 of 8 patients with the L908V substitution, 5 of 8 patients with the R403Q substitution, 1 of 3 patients with the G741R substitution, and in the 1 patient with the G256E substitution. Although Laing myopathy has defined myopathic features and has not been associated with muscle central core phenotype, the investigations of Fananapazir (1993) show that *MYH7* mutations can cause skeletal muscle pathology without CMH1 phenotype.

Lankford et al. (1995) compared the contractile properties of single soleus muscle slow-twitch muscle fibers in normal controls to those of patients with 2 distinct CMH1 mutations. They found that fibers from  $\beta$ -MyHC G741R patients, had decreased maximum velocity of shortening (39% of normal) and decreased isometric force generation (42% of normal). Fibers from  $\beta$ -MyHC R403Q patients showed lower force/stiffness ratio (56% of normal) and depressed velocity of shortening (50% of normal). Both of these mutation-containing fibers displayed abnormal force-velocity relationships and reduced power output. Therefore, mutations in *MYH7* affect the contractile properties of skeletal muscle. This agrees with the later conclusion of Redwood et al. (1999; cited on p. 222) that CMH1 is a contractile disorder. The observed variably expressed myopathy and central core histology may be secondary consequences of altered contractile properties.

### **8.1.3 Which CMH Disease Genes are Associated with Skeletal Muscle Myopathy?**

Nine CMH genes have been identified and are listed in Table 8.1.



**Table 8.1      *Cardiomyopathy Hypertrophic, Familial (CMH)  
and Associated Disease Genes.***

<b>Gene (CMH)</b>	<b>Protein/ Polypeptide</b>	<b>Number of Mutations*</b>	<b>First Reports</b>
<i>MYH7</i> (CMH1)	Myosin heavy polypeptide 7, cardiac muscle, $\beta$	27	Geisterfer-Lowrance et al. (1990)
<i>TNNT2</i> (CMH2)	Troponin T2, cardiac	11	Thierfelder et al. (1994)
<i>TPM1</i> (CMH3)	Tropomyosin 1 alpha	4	Thierfelder et al. (1994)
<i>MYBPC3</i> (CMH4)	Myosin binding protein, cardiac	28	Bonne et al. (1995) Watkins et al. (1995)
<i>TNNI3</i> (CMH7)	Troponin I, cardiac muscle isoform	5	Kimura et al. (1997)
<i>MYL3</i> (CMH8)	Myosin, light chain, alkali; ventricular and skeletal slow	2	Poetter et al. (1996)
<i>TTN</i> (CMH9)	Titin	1	Satoh et al. (1999)
<i>MYL2</i>	Myosin, light chain 2; regulatory ventricular	5	Poetter et al. (1996)
<i>ACTC</i>	Actin, $\alpha$ , cardiac muscle	1	Mogensen et al. (1999)

\* as listed in HGMD (2000)

Note that CMH5 was assigned to a 'wastebasket' category for forms of CMH not linked to any known loci; CMH6 was assigned to hypertrophic cardiomyopathy associated with Wolff-Parkinson-White Syndrome; CMH associated with mutations in *MYL2* and *ACTC* have not yet been given CMH numbers.

A PubMed search (<http://www.ncbi.nlm.nih.gov:80/entrez/query.fcgi?>) linking all CMH genes with 'myopathy' produced only two positive responses. The first was eight articles linking *MYH7* and myopathy, the second, four articles linking *TTN* and myopathy. An examination of the *MYH7* articles showed that six of them related to the localisation of the French Canadian form of OPMD, one was the Laing et al. (1995b) localisation of MPD1 and one was the Fananapazir et al. (1993) report discussed above. An examination of the four titin references found that three did not directly link titin to myopathy (Fougerousse et al., 1998; Pelin et al., 1997; Muller-Seitz et al., 1993). The only article which linked titin to

myopathy was one authored by Udd et al. (1998), where titin was identified as a 'superior TMD candidate'. Three months prior to the Udd article, Haravuori et al., (1998a) had reported the results of linkage analysis on four different Finnish AD-TMD families. They found a common chromosome 2q31 core haplotype segregating with affecteds. This was indicative of one major ancient founder mutation for AD-TMD in Finland. Multipoint likelihood calculations, combined with haplotype and recombination analyses, restricted the TMD locus to a chromosomal region of ~1cM. Udd et al. (1998, p.331) comment:

There are not many interesting genes in the 2q31 locus, which is about 1 cM in size. One known gene is the major candidate, namely the gene encoding the giant muscle gene titin (former name connectin) (Labeit et al., 1995). One of the main functions of titin is to keep the contractile elements of the sarcomere in place during the variations of the functional state of the muscle fibre. It also binds a large number of other muscle proteins and is responsible for muscle elasticity. . . . Titin has many properties making it a highly interesting candidate for TMD. It has several sites for differential splicing causing different isoforms to appear in different muscles (Labeit et al., 1997). This might be relevant for the selective involvement of muscles in anatomically restricted myopathies.

The last sentence of this quotation was most apposite. In 1999 Satoh et al. identified in a CMH patient, a G to T transversion in titin codon 740, changing this codon from CGC (arginine) to CTC (leucine). This mutation increased the binding affinity of titin to alpha-actinin in the yeast two-hybrid assay, suggesting that the G2219T titin mutation may cause CMH in this patient by altering its affinity to alpha-actinin. Characterisation of titin transcripts from cardiac and skeletal muscles show that cascades of exon-skipping events produce titins with distinct structures, which confer distinct elastic properties on muscle types (Freiburg et al., 1999; Centner et al., 2000). So it appears quite possible that

the differential splicing of titin might not only explain the regional distribution of this distal myopathy, but could also explain the lack of CMH in TMD patients. The splicing-out of the putative TMD mutation in cardiac isoforms could explain the absence of CMH, or other cardiac phenotypes, in TMD patients.

There is no evidence to suggest that differential splicing could explain the absence of CMH in Laing myopathy patients. A PubMed search for  $\beta$ -cardiac myosin identified 720 articles, none describing  $\beta$ -cardiac myosin genes as differentially spliced in any of the many organisms studied. There is only one report of alternate splicing occurring in  $\alpha$ -cardiac MyHC and that involved the inclusion or exclusion of the glutamine residue at position 1931 in rat  $\alpha$ -cardiac MyHC (Sindhwani et al., 1994). Differential splicing does not contribute to the anatomical restriction of Laing myopathy and cannot be invoked to explain the absence of CMH1 in these patients. However, the restricted phenotype of a mutation can be explained in other ways such as the timing of gene expression or tissue-specific factors. The latter is discussed in the next section.

#### **8.1.4 Do Substitutions in the $\beta$ -MyHC LMM Rod Region Cause Defects in Specific Muscles?**

Kronert et al. (1995) determined the molecular and ultrastructural defects associated with three homozygous-viable MyHC mutations of *Drosophila melanogaster*. These mutations caused a dominant flightless phenotype but allowed relatively normal assembly of indirect flight muscle myofibrils. As a result of myofibril 'hypercontraction', the contents of the indirect flight muscle myofibers were pulled to one end of the thorax as the sarcomere progressively ruptured with adult aging. All three mutations caused single amino acid changes in the LMM region of the myosin rod. Two changed the same GAG codon (glutamic acid) to an AAG codon (lysine) while the third affected an amino acid

five residues away, changing CGC codon (arginine) to CAC codon (histidine). The amino acid substitutions were E1555K and R1561H. Both of these altered residues were conserved in nematode, chicken, and rat muscle myosins, in one human non-muscle myosin (now called myosin 9), and in human  $\beta$ -MyHC (see Table 8.2, p. 234, where the organisation of the MyHC rod is presented). In  $\beta$ -MyHC these residues were E1553 and R1559 and were positioned 102 and 107 residues N-terminal to the P1663 substitution segregating with Laing myopathy. *Drosophila melanogaster* and human LMM rods are two residues 'out-of-register' at this part of their amino acid sequence.

*Drosophila melanogaster* is unique in that its multiple forms of muscle MyHC are encoded by a single copy gene (Bernstein et al, 1983). *Drosophila* uses alternative RNA splicing to produce all the isoforms required by the functionally different muscles in *Drosophila* (George et al., 1989). The mutations described by Kronert (1995) were in exon 16 of the *Drosophila* MyHC gene and were found to be associated with age-dependent, site-specific degradation of MyHC. They were also associated with a failure to accumulate phosphorylated forms of flightin, a specific protein found in indirect flight muscle previously localised to the thick filament. Kronert et al. considered it remarkable that these single amino acid changes in the LMM section of the MyHC rod could cause severe defects *and* that these severe defects were only apparent in indirect flight muscles. All other muscles were normal. They concluded (pp. 121-122):

The simplest explanation is that flightin binds to myosin where the mutations described in this paper have been mapped, and that binding is required for phosphorylation. Ferguson et al. (1994) recently showed that a possible analogue of flightin in *Lethocerus indicus* [a water bug] does indeed bind to thick filaments, most likely near the end of the myosin rod. Flightin interaction with myosin might permit proper muscle contraction/relaxation to occur, a process that is disrupted in the *Drosophila* rod mutants.

### **8.1.5 Does the LMM Rod Region of Human $\beta$ -MyHC Interact with Other Molecules?**

Okagaki et al., (1993) established that the skeletal muscle MyBP-C, bound to the LMM section of the myosin rod. The last 102 amino acids of MyBP-C was involved in LMM binding. This means that there are at least two regions on the  $\beta$ -MyHC rod to which MyBP-C binds; at this site and at the myosin S2 site associated with CMH1 which was discussed in Section 8.1 (pp. 222-223). Gilbert et al. (1996) found that one truncated mutant of MyBP-C, only lacking the myosin binding domain, produced a clear dominant-negative disruption of myofibril assembly. These authors performed transfection studies in skeletal muscle myoblasts with constructs encoding truncation mutants of MyBP-C. This finding was relevant to the present study for three reasons.

First, it opened up the possibility that mutations in the MyHC rod might have the same dominant negative effect as the truncated MyBP-C. The site(s) on the myosin LMM rod to which MyBP-C binds, is/are unknown. Second, the cardiac MyBP-C isoform contains a unique 28 residue insertion, not found in skeletal muscle MyBP-C, which is abundant in proline and charged residues (Redwood et al., 1999; Gruen & Gautel, 1999). Were MyBP-C involved in Laing myopathy pathogenesis, the absence of a cardiac phenotype in this disease could be a direct result of the differences between skeletal and cardiac muscle MyBP-C isoforms.

Finally, the same MyBP-C high-affinity LMM-binding domain also binds to titin but with lower affinity (Gruen & Gautel, 1999). Thus the rod region of  $\beta$ -MyHC molecules is closely bound to titin by MyBP-C. Could these two distal myopathies be linked by the failure of MyBP-C to bind to  $\beta$ -MyHC in Laing myopathy and to titin in TMD? Other possible connections between the distal myopathies are considered in Chapter 9, Section 9.2.

### **8.1.6 Do Other MYH Mutations Cause Skeletal Muscle Myopathy?**

The genes encoding embryonic, IIa, IIx/d, IIb, perinatal and extraocular MyHCs are located in a cluster on chromosome 17 (Weiss et al., 1999). By linkage analysis and radiation hybrid mapping, Martinsson et al. (1999) localised an AD hereditary inclusion-body myopathy gene (Human Genome Map locus name: *IBM3*) to chromosome region 17p13.1. An exon 17 G2116A mutation in the chromosome 17p13.1 *MyHC-IIa* gene, has been found to segregate with *IBM3* hereditary inclusion-body myopathy (Martinsson et al, 2000). Therefore, a mutation in this chromosome 17 MYH gene is associated with a human skeletal muscle myopathy.

### **8.1.7 Sarcomeric Gene Mutations and Laing Myopathy: An Evaluation**

In the introduction to this analysis it was stated that, 'difficulties confronting the  $\beta$ -MyHC A1663P Laing myopathy candidature and possible disease mechanisms will be identified and analysed' (p. 218). Three difficulties confronting the A1663P Laing myopathy candidature were identified and discounted.

First, although *MYH7* mutations do cause CMH, none of these mutations lie in the LMM region of the  $\beta$ -MyHC molecule. Furthermore, TMD shows no signs of CMH yet, like Laing myopathy, it also has a CMH gene, titin (CMH9), as its candidate gene. Although differential splicing may explain the absence of CMH from TMD, it does not do so for Laing myopathy. However, skeletal muscle myopathy and central core phenotype do segregate with *MYH7* mutations, without CMH1 phenotype. *MYH7* mutations can cause skeletal muscle myopathy, restricted to specific muscles and without cardiac phenotype.

Second, mutation to a human MYH gene, *MyHC-IIa*, can cause myopathy (Martinsson et al, 2000). Third, mutations to the MyHC LMM rod region can have selective phenotypic expression. Mutations to the LMM rod region of the *Drosophila melanogaster* MYH gene causes a hypercontractile state in only one type of muscle, the indirect flight muscles.

This analysis also gave two insights into disease mechanisms which may be of relevance to Laing myopathy. First, CMH1 mutations in the globular head of the  $\beta$ -MyHC molecule cause hypocontractility (Redwood et al., 1999). Second, mutations in the LMM rod region of the *Drosophila melanogaster* MYH gene cause hypercontractility and this has been explained by the failure of flightin to attach to the MyHC LMM rod region (Kronert et al., 1995). In addition, a truncated mutant of MyBP-C without the domain for binding to MyHC LMM rod region, produced a clear dominant-negative disruption of myofibril assembly (Gilbert et al., 1996). Therefore, from this analysis of *MYH7* mutations, two possible directions for future Laing myopathy have emerged. Research focussed on the contractility of myosin II containing both A1663 and P1663  $\beta$ -MyHC monomers, and research focussed on the interaction of the  $\beta$ -MyHC LMM rod region with other macromolecules, involving A1663 and P1663  $\beta$ -MyHC monomers if appropriate.

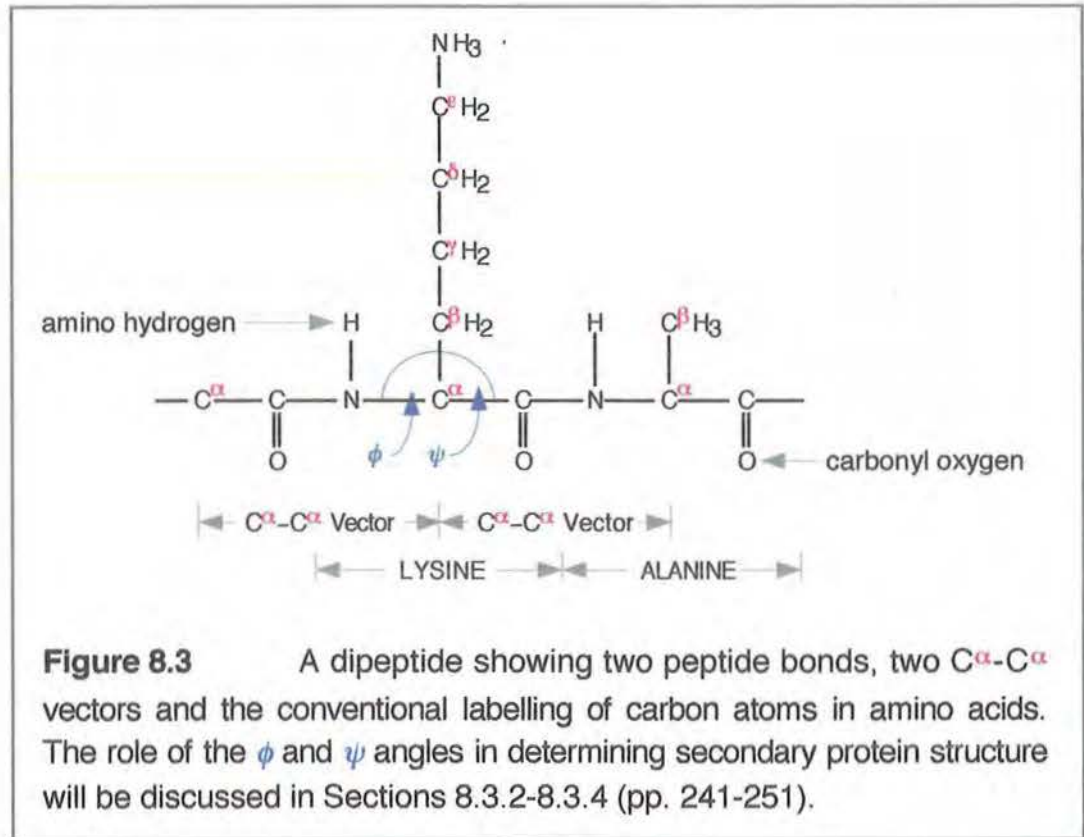
By resolving some difficulties confronting *MYH7* G5073C as the Laing myopathy mutation, the candidature of *MYH7* G5073C has been strengthened. But the reality remains.  $\beta$ -MyHC A1663P is a novel substitution, of effect(s) unknown. After *MYH7* mutation analysis, the answer to question, "Could the  $\beta$ -MyHC A1663P substitution cause Laing myopathy?" was only, "It could". Further analysis was necessary.



The next four sections continue considering the likelihood of the  $\beta$ -MyHC A1663P Laing myopathy candidature, and continue looking for possible disease causing mechanisms. However, these sections adopt a different approach: the examination of the possible effect(s) of the  $\beta$ -MyHC A1663P substitution at the primary, secondary, and quaternary structural levels of  $\beta$ -MyHC, and then at the superstructural level of the thick filament.

## 8.2 A1663P AND THE PRIMARY STRUCTURE OF THE $\beta$ -MyHC ROD

The primary structure of proteins is defined as their sequence of amino acids. Peptide bonds, also called vectors, link amino acids together to form polypeptides. Carbon atoms in amino acids are conventionally labelled  $\alpha$  for the centre carbon atom and  $\beta$ ,  $\gamma$ ,  $\delta$  and  $\epsilon$  for the subsequent side chain carbon atoms. Since this nomenclature is employed throughout the remainder of this chapter, Figure 8.3 provides a reference point.



The primary structure of the MyHC rod region is dominated by two repeat patterns, the heptad-residue repeat pattern and the 28-residue repeat pattern. The heptad repeat pattern is a regular, 7-residue pattern, conventionally labelled *a-b-c-d-e-f-g*, in which hydrophobic residues are concentrated at intervals of 3 and 4 residues in positions *a* and *d* (McLachlan and Karn 1982, p.227). Sellers (1999, p. 52) also notes that the outermost positions *b*, *c*, and *f* contain the greatest number of charged, hydrophilic polar residues. The heptad distribution of hydrophobic and hydrophilic residues determines the quaternary structure of the MyHCs. Heptad repeats are only found in  $\alpha$ -helices that form quaternary coiled-coil structures. Discussion of heptad repeats continues throughout this chapter, but particularly in the section considering quaternary structure.

Internal comparison of the amino acid sequence of the *C. elegans* myosin rod region of the nematode shows that the entire rod sequence could be mapped out as a cyclical 28-residue charge-repeat pattern. This charge-repeat pattern controls thick filament formation (McLachlan & Karn, 1982, p.227). Discussion of 28-residue, also continues throughout this chapter, but particularly in the section considering thick filament formation.

### **8.2.1      Derived Primary Structure of $\beta$ -MyHC**

Table 8.2 (p. 234) shows the derived primary structure for  $\beta$ -MyHC rod of WAMPD Laing myopathy proband, III.2. Four observations and one deduction could be made from the primary  $\beta$ -MyHC structure presented in this table. First, proline was not part of the rod sequence after position 838. It has never been identified between the residues 838 and 1935 in human  $\beta$ -MyHC sequence.

**Table 8.2    *The Derived  $\beta$ -MyHC Rod Region Amino Acid Sequence for WAMPD Laing Myopathy Proband, III.2.***

The WAMPD III.2 sequence for  $\beta$ -MyHC was derived from the determined MYH7 cDNA sequence, using SeqEd, Version 1.0.3 (Myers, 1992). Only the amino acid sequence for the  $\beta$ -MyHC rod region is shown. The heptad repeats of the human  $\beta$ -MyHC rod commence with an invariant proline residue (Offer and Knight, 1996, p. 408), located at position 838, and both the rod and the  $\beta$ -MyHC molecule terminate at residue 1935. Since the myosin rod starts with a proline in position *d*, the MyHC 28-residue repeats start at heptad position *d*. The amino acid sequence is presented in heptad and 28-residue repeats. In the human  $\beta$ -MyHC rod, there are 38, full, 28-residue repeats and 2 half 28-residue repeats. For the purpose of identification, these repeats were numbered 1-40. The 28-residue repeat pattern is broken by four skip residues (boxed) at positions 1188, 1385, 1582, and 1807. In addition, there is a 2 residue tailpiece occupying the final two positions, 1934 and 1935.

Residues are highlighted in red for charged hydrophilic polar residues; black for uncharged hydrophilic polar residues; and blue for hydrophobic nonpolar residues. The heptad positions *d, e, f, g, a, b, c* are also highlighted with the same colour code to indicate the predicted, 'most common type of residue' in each position (Sellers, 1999, p. 52). Positions *e* and *g* do not have a 'most common type of residue'. The numbers at the bottom of the *d, e, f, g, a, b, c* columns indicate the frequency of each type of residue in each column. The maximum number of residues in each of these columns was 39, with the exception of the *d* and *e* column starting at positions 838 and 839. These columns had 40 residues because they each included one residue of the tailpiece. Heptads within each 28-repeat were labelled according to convention - heptads 1, 2, 3, and 4. LMM commences between residues 1203-1282.

Two changes to the published  $\beta$ -MyHC sequence were detected (Jaenicke et al., 1990). The first of these was the S1124A substitution. This was present in the WAMPD proband III.2 and all eight tested, unrelated, unaffected individuals. It was a homozygous change. The second amino acid change to be detected was the A1663P substitution. This segregated with Laing myopathy. It was a heterozygous change.

Universal, single letter codes were used to designate amino acid residues (see also Appendix C). Charged hydrophilic, polar residues: **D** = aspartic acid; **E** = glutamic acid; **H** = histidine; **K** = lysine; **R** = arginine. Uncharged hydrophilic, polar residues: **C** = cysteine; **N** = asparagine; **Q** = glutamine; **S** = serine; **T** = threonine; **Y** = tyrosine. Hydrophobic, nonpolar residues: **A** = alanine; **F** = phenylalanine; **G** = glycine; **I** = isoleucine; **L** = leucine; **M** = methionine; **P** = proline; **W** = tryptophan.



Heptad Number																												Amino Acid Number / 28-Repeat Number		
1							2							3							4									
d	f	g	a	b	c	d	f	g	a	b	c	d	f	g	a	b	c	d	f	g	a	b	c	d	f	g	a	b	c	851/1
M	K	E	E	F	T	R	L	K	E	A	L	E	K	S	E	A	R	R	K	E	L	E	E	K	M	V	S			
L	L	Q	E	K	N	D	L	Q	L	Q	V	Q	A	E	Q	D	N	L	A	D	A	E	E	R	C	D	Q			
L	I	K	N	K	I	Q	L	E	A	K	V	K	E	M	N	E	R	L	E	D	E	E	E	M	N	A	E			
L	T	A	K	K	R	K	L	E	D	E	C	S	E	L	K	R	D	I	D	D	L	E	L	T	L	A	K			
V	E	K	E	K	H	A	T	E	N	K	V	K	N	L	T	E	E	M	A	G	L	D	E	I	I	A	K			
L	T	K	E	K	K	A	L	Q	E	A	H	Q	Q	A	L	D	D	L	Q	A	E	E	D	K	V	N	T			
L	T	K	A	K	V	K	L	E	Q	Q	V	D	D	L	E	G	S	L	E	Q	E	K	K	V	R	M	D			
L	E	R	A	K	R	K	L	E	G	D	L	K	L	T	Q	E	S	I	M	D	L	E	N	D	K	Q	Q			
L	D	E	R	L	K	K	K	D	F	E	L	N	A	L	N	A	R	I	E	D	E	Q	A	L	G	S	Q			
L	Q	K	K	L	K	E	L	Q	A	R	I	E	E	L	E	E	L	E	A	E	R	T	A	R	A	K		1131/11		
V	E	K	L	R	S	D	L	S	R	E	L	E	E	I	S	E	R	L	E	E	A	G	G	A	T	S	V			
Q	I	E	M	N	K	K	R	E	A	E	F	Q	K	M	R	R	D	L	E	E	A	T	L	Q	H	E	A	T	1188/13	
A	A	A	L	R	K	K	H	A	D	S	V	A	E	L	G	E	Q	I	D	N	L	Q	R	V	K	Q	K			
L	E	K	E	K	S	E	F	K	L	E	L	D	D	V	T	S	N	M	E	Q	I	I	K	A	K	A	N			LMM START
L	E	K	M	C	R	T	L	E	D	Q	M	N	E	H	R	S	K	A	E	E	T	Q	R	S	V	N	D			
L	T	S	Q	R	A	K	L	Q	T	E	N	G	E	L	S	R	Q	L	D	E	K	E	A	L	I	S	Q			
L	T	R	G	K	L	T	Y	T	Q	Q	L	E	D	L	K	R	Q	L	E	E	E	V	K	A	K	N	A			
L	A	H	A	L	Q	S	A	R	H	D	C	D	L	L	R	E	Q	Y	E	E	E	T	E	A	K	A	E			
L	Q	R	V	L	S	K	A	N	S	E	V	A	Q	W	R	T	K	Y	E	T	D	A	I	Q	R	T	E	1385/20		
L	E	E	A	K	K	L	A	Q	R	L	Q	E	A	E	E	A	V	E	A	V	N	A	K	C	S	S				
L	E	K	T	K	H	R	L	Q	N	E	I	E	D	L	M	V	D	V	E	R	S	N	A	A	A	A	A			
L	D	K	K	Q	R	N	F	D	K	I	L	A	E	W	K	Q	K	Y	E	E	S	Q	S	E	L	E	S			
S	Q	K	E	A	R	S	L	S	T	E	L	F	K	L	K	N	A	Y	E	E	S	L	E	H	L	E	T			
F	K	R	E	N	K	N	L	Q	E	E	I	S	D	L	T	E	Q	L	G	S	S	G	K	T	I	H	E			
L	E	K	V	R	K	Q	L	E	A	E	K	M	E	L	Q	S	A	L	E	E	A	E	A	S	L	E	H			
E	E	G	K	I	L	R	A	Q	L	E	F	N	Q	I	K	A	E	I	E	R	K	L	A	E	K	D	E	1582/27		
M	E	Q	A	K	R	N	H	L	R	V	V	D	S	L	Q	T	S	L	D	A	E	T	R	S	R	N	E			
A	L	R	V	K	K	K	M	E	G	D	L	N	E	M	E	I	Q	L	S	H	A	N	R	M	A	A	E			
A	Q	K	Q	V	K	S	L	Q	S	L	L	K	D	T	Q	I	Q	L	D	D	A	V	R	A	P	N	D	D	1666/30	
L	K	E	N	I	A	I	V	E	R	R	N	N	L	L	Q	A	E	L	E	E	L	R	A	V	V	E	Q			
T	E	R	S	R	K	L	A	E	Q	E	L	I	E	T	S	E	R	V	Q	L	L	H	S	Q	N	T	S			
L	I	N	Q	K	K	K	M	D	A	D	L	S	Q	L	Q	T	E	V	E	E	A	V	Q	E	C	R	N			
A	E	E	K	A	K	K	A	I	T	D	A	A	M	M	A	E	E	L	K	K	E	Q	D	T	S	A	H			
L	E	R	M	K	K	N	M	E	Q	T	I	K	D	L	Q	H	R	L	D	E	A	E	Q	I	A	L	K	G	1807/35	
G	K	K	Q	L	Q	K	L	E	A	R	V	R	E	L	E	N	E	L	E	A	E	Q	K	R	N	A	E			
S	V	K	G	M	R	K	S	E	R	R	I	K	E	L	T	Y	Q	T	E	E	D	R	K	N	L	L	R			
L	Q	D	L	V	D	K	L	Q	L	K	V	K	A	Y	K	R	Q	A	E	E	A	E	E	Q	A	N	T			
N	L	S	K	F	R	K	V	Q	H	E	L	D	E	A	E	E	R	A	D	I	A	E	S	Q	V	N	K			
L	R	A	K	S	R	D	I	G	T	K	G	L	N	E																
33	20	30	16	21	26	24	32	20	14	28	33	19	25	33	18	20	22	32	32	26	19	18	21	17	19	15	20			
6	10	5	15	13	7	11	4	14	13	6	4	12	7	6	17	10	14	6	5	7	11	13	11	12	11	14	15			Residue Proportions
		9	4	8	5	6	4	3	5	12	5	2	8	7	1	6	9	3	1	2	6	9	8	7	10	9	10	4		
d	f	g	a	b	c	d	f	g	a	b	c	d	f	g	a	b	c	d	f	g	a	b	c	d	f	g	a	b	c	

Residue  
Proportions

Second, the distribution of hydrophobic and hydrophilic residues showed trends towards the predicted distributions, sometimes matching them quite closely, sometimes not so closely. For example, the figures at the bottom of position 1 *d* (column 1 in Table 8.2) show that this heptad position had 33 hydrophobic and 6 uncharged hydrophilic non-polar residues. This was a close approximation to the predicted occupancy for a *d* heptad position. However, Table 8.2 also shows that the other hydrophobic position in heptad 1, 1 *a*, had an amino acid occupancy far closer to that predicted for the *b*, *c*, and *f* heptad positions. 1 *a* had 21 charged hydrophilic polar residues, 13 hydrophobic residues and 5 uncharged hydrophilic non-polar residues. Excluding the non-helical tailpiece, 99 of the 312 *a* and *d* residues in the  $\beta$ -MyHC rod are polar residues, and 50 of these are charged polar residues. The tendency for hydrophobic residues to occupy *a* and *d* positions is just that, a tendency.

From this observation it was possible to make a deduction. The moderate stringency with which heptad positions *a* and *d* were occupied by hydrophobic residues, suggested that heptad positions *a* and *d* were subjected to further selection force(s), in addition to 'hydrophobic selection'. Lupas (1996, p. 377) argues that core polar residues could be partly solvated in dimers, more so than in trimers and tetramers, so the presence of higher numbers of core polar residues would favour dimerisation. Exactly what is observed in  $\beta$ -MyHC.

The third observation to be made from Table 8.2 is that, for each *a*, *b*, *c*, *d*, *e*, *f*, or *g* heptad position, the residue occupancy displayed considerable variation. This can be observed by comparing the figures at the bottom of all the *a* columns, then all the *b* columns, and so on.

The final observation to be made from Table 8.2, is that there are four skip residues disrupting both the heptad and 28-residue repeats. The role of skip

residues will be considered further when the thick filament assembly is discussed in Section 8.5.2 (pp. 268-273). These last two observations support the deduction presented above. The variable occupancy for each heptad position and the presence of skip residues, both indicate that selection forces other than hydrophobic and hydrophilic selection must also be operating on the  $\beta$ -MyHC rod region. The  $\beta$ -MyHC A1663P substitution could, therefore, potentially affect  $\beta$ -MyHC function in a variety of ways.

These four observations and one deduction made from the published and derived primary structure of  $\beta$ -MyHC, did not provide enough information to predict why proline was so rare or the effect(s) it might have on  $\beta$ -MyHC dimerisation or thick filament formation. Limited sampling, a lack of proline phenotype or the lethality of proline substitution (with WAMPD A1663P being a very rare exception), could all explain why proline had not been identified in this section of  $\beta$ -MyHC. These possibilities could not be discriminated.

### **8.2.2 Comparative Analysis of MyHC Rod Primary Structure**

Table 8.3 (p. 237) compares the published sequence for the thirtieth 28-residue repeat in the human  $\beta$ -MyHC rod sequence to the WAMPD sequence, and to the equivalent 28-residue repeat sequences for all MyHCs found by a BLAST search based on the complete  $\beta$ -MyHC rod sequence (<http://www.ncbi.nlm.nih.gov/BLAST/>). The thirtieth 28-residue repeat in the human  $\beta$ -MyHC rod contains the A1663P segregating with Laing myopathy. Apart from the last two entries in Table 8.3, all of the MyHC molecules identified by BLAST were muscle MyHCs. This table shows that proline was not present in the 30.4g position, the  $\beta$ -MyHC A1663P substitution heptad position, and it was not present in *any* of the heptad presented in this table.



**Table 8.3    *Comparison of Number 30 Human  $\beta$ -MyHC 28-Residue Repeat to Equivalent Regions in other MyHCs.***

The A1663P substitution segregating with Laing myopathy in the WAMPD pedigree lay in the  $\beta$ -MyHC 28-residue repeat number 30, the top sequence in the table. The last 20, complete 28-residue repeats of the published human  $\beta$ -MyHC sequence (Jaenicke et al., 1990), from 28-residue repeat number 20 to 39, were entered into BLAST to identify homologous sequences and to determine their degree of homology. This totalled 563 residues as 3 skip residues were included. The degree of homology of each sequence, over these 563 residues, to human  $\beta$ -MyHC was shown in the '%' column. The 'Source' column employed common names to identify those organisms whose MyHCs were compared to human  $\beta$ -MyHC. The exception was the 'WAMPD' label which identified the human  $\beta$ -MyHC 28-residue repeat number 30 segregating with Laing myopathy.

The abbreviations used in the 'Tissue' column to identify the site of MyHC expression, were: *C* for cardiac; *S* for skeletal; *St* for striated; *Emb* for embryonic muscle; *St* muscle; *Sm* for smooth; *NM* for non-muscle. Apart from human  $\beta$ -MyHC, only the first reported site of expression was listed.

As for Table 8.2 (p. 234), **red** is used in this table to identify charged hydrophilic polar residues, **black** to identify uncharged hydrophilic polar residues, and **blue** to identify hydrophobic nonpolar residues. Heptad repeat positions, *d*, *e*, *f*, *g*, *a*, *b*, *c* are coloured to indicate the predicted, 'most common type of residue' in each position (Sellers, 1999, p. 52). Positions *e* and *g* do not have an identifiable residue pattern.

**Bolding** and a change of FONT identify where residues differ from the published human  $\beta$ -MyHC residues (Jaenicke et al., 1990).





An examination of the entire rod sequences of all of these MyHC failed to identify a single proline residue within their rod region. All rods started with a capping proline, and proline was sometimes present at the end of the  $\alpha$ -helical region, marking the start of a non-helical tailpiece. For example, nematode MyHC had a proline starting its tailpiece and another within this tailpiece (McLachlan & Karn, 1982, p. 228).

The comparative analysis of MyHC rod regions strengthened the case against proline. Its complete absence within these rod regions implies that its presence is not advantageous. High conservation of a region indicates selection rigour. From Table 8.3 it can be seen that position 30.3e was 100% conserved and therefore presumably under strong selection. With eight different residues occupying position 30.4g, the position containing the A1663P substitution, 30.4g is not as conserved as 30.3e. Position 30.4g was, however, conserved until chicken skeletal muscle MyHC, where serine is substituted for alanine. This chicken skeletal muscle MyHC had only 73% homology to human  $\beta$ -MyHC over the 563 residues compared.

The MyHC phase shifts caused by skip residues are also highly conserved. Skip residues distort the electrostatic charge pattern determined by the 28-residue repeats. The nematode MyHC rod has skip residues in the same positions as the human skip residues. The only difference between nematode and human MyHC skip residues is in the first skip residue, which, in the nematode, is the uncharged polar residue glutamine (McLachlan & Karn, p. 228). In humans, the uncharged polar residue threonine occupies the first skip residue position (Table 8.2, p. 234).

From the data shown in Table 8.3, and from this discussion of that data, it follows that the  $\beta$ -MyHC A1663P substitution occurs in a highly conserved region of a highly conserved molecule. The absence of proline within the rod regions of all

MyHCs and the high degree of conservation of the MyHCs, particularly over the region containing the A1663P substitution in human  $\beta$ -MyHC, further indicates that proline within the rod region is not advantageous. In fact, the supposition that proline could be deleterious to the structure and functioning of the MyHC rod region, now appear tenable.

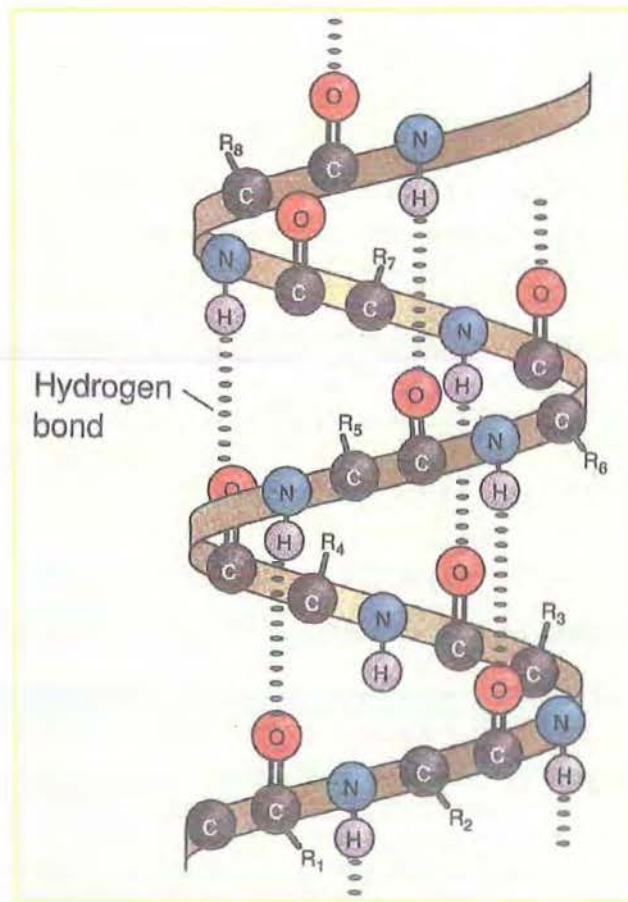
### **8.3 SECONDARY STRUCTURE OF THE $\beta$ -MyHC ROD**

The secondary structure of a protein has been traditionally defined as the folding of amino acid chain into motifs such as  $\alpha$ -helices,  $\beta$ -pleated sheets, loops, or random coils. The fundamental secondary structure of the MyHC rod region is the  $\alpha$ -helix, a structure first proposed by Pauling and Corey (1951) to explain the X-ray diffraction patterns obtained for  $\alpha$ -keratin (cited in Crick, 1953, p. 689).

#### **8.3.1 The $\alpha$ -helix: Structure, Direction and Sense of Twist**

The  $\alpha$ -helix involves the formation of a rigid cylinder. The structure is dominated by hydrogen bonding between carbonyl oxygens and the hydrogen atoms of the amino nitrogen. Stabilising hydrogen bonds form between carbonyl oxygens and amino nitrogen hydrogens in peptide bonds separated by three amino acids. That is, hydrogen bonds form between every fourth peptide bond. Consecutive  $\alpha$ -carbons and peptide linkages follow a regular spiral path, producing an  $\alpha$ -helix with 3.6 residues per turn. Amino acid side chains extend outwards from the axis of the spiral.

Figure 8.4 (p. 240) presents a short segment of a polypeptide  $\alpha$ -helix.



**Figure 8.4** Section of a right-handed polypeptide  $\alpha$ -helix. It has 3.6 residues per  $\alpha$ -helix turn. Dashed lines identify stabilising hydrogen bonds between carbonyl oxygens and hydrogen atoms of amino nitrogens. Scanned from Kleinsmith and Kish (1995, p. 24).

Crick (1953) first noted that four types of  $\alpha$ -helix were possible, each described by its sense of twist, right-handed or left-handed, and by its direction, up or down. The sense of twist or 'handedness' of an  $\alpha$ -helix and its direction are integral, unchanging properties of each helix. Right-handed  $\alpha$ -helices always turn in a clockwise direction along their axis. Visually, they appear to spiral in a clockwise manner when viewed, from *either* end, along the direction of the main chain carbon atom sequence. Left-handed  $\alpha$ -helices always spiral in an anti-clockwise direction along their axis, irrespective of viewing 'end'.

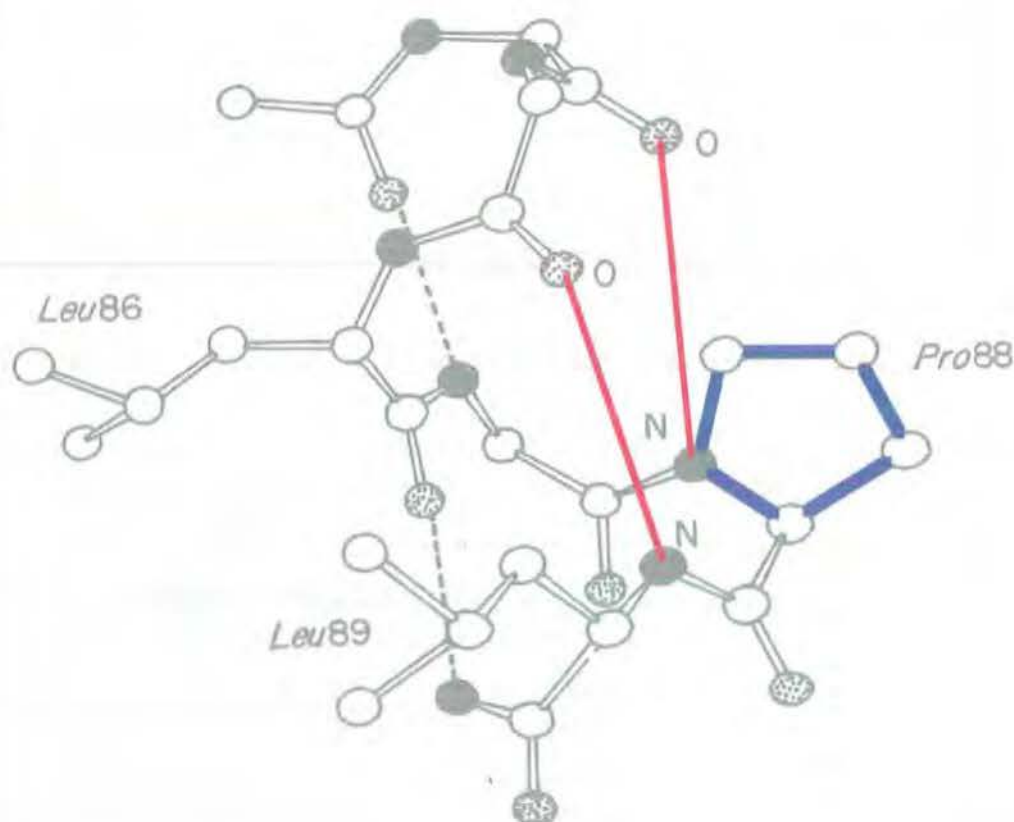
### 8.3.2 Proline the $\alpha$ -helix Breaker: Rigid-Geometry Analysis

Proline is unique among the 20 naturally occurring amino acids in mammalian systems in that its side chain cyclizes back to the backbone amide/amino nitrogen, leaving one of its dihedral angles ( $\phi$ ) fixed at approximately  $-65^\circ$  (MacArthur & Thornton, 1991, p. 398). Two dihedral angles,  $\phi$  and  $\psi$ , are called the main chain torsion angles and they measure the angle that the  $C\alpha$ - $C\beta$  bond makes to adjacent  $C\alpha$ - $C\alpha$  vectors (diagrammed in Figure 8.3, p. 232).

Rigid-geometry analysis of proline within  $\alpha$ -helices assumes that all bond lengths and bond angles have fixed, ideal values and that  $\alpha$ -helices are always regular (Weiner et al., 1986). Rigid-geometry calculations show that proline is detrimental to the  $\alpha$ -helical conformation for three reasons. First, proline lacks an amide proton to form an  $\alpha$ -helix stabilising intramolecular hydrogen bond with the carbonyl oxygen of a peptide bond located three amino acids away (Williams & Deber, 1991). Figure 8.5 (p. 242) identifies this and another 'lost' hydrogen bond.

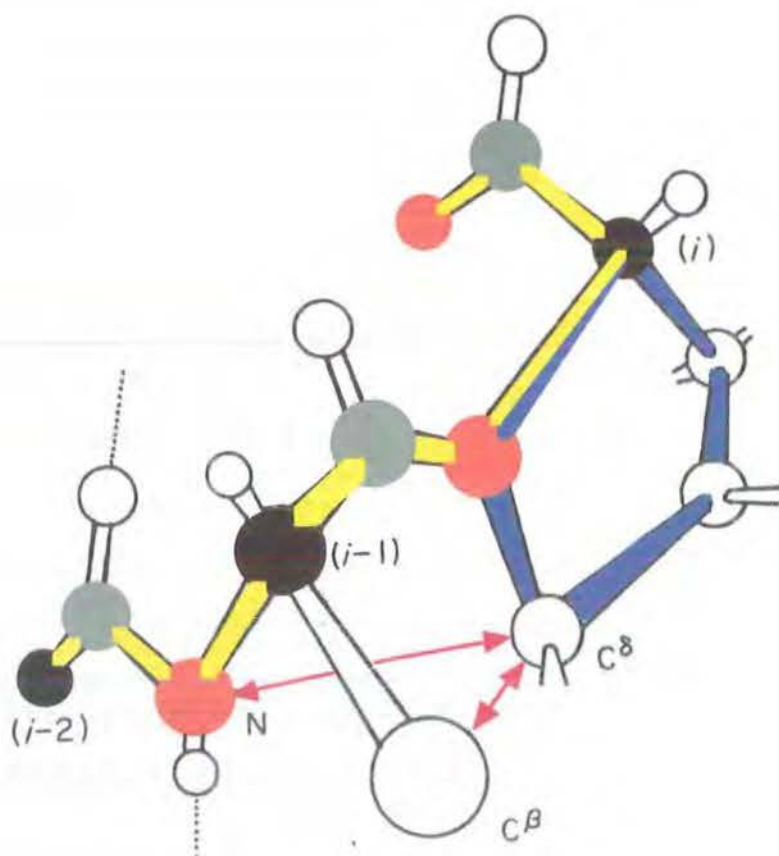
Second, the bulkiness of its pyrrolidine ring places steric constraints on the conformation of the preceding residue in the  $\alpha$ -helix. Schimmel and Florey (1968) cited in MacArthur and Thornton (1991, p. 401) showed that the space available to the preceding residue was severely curtailed by steric conflicts between the  $C\beta H_2$  attached to the pyrrolidine amino nitrogen and the  $C\delta H_2$  of the preceding residue, and between this same  $C\beta H_2$  and NH amide group in the peptide bond joining the two preceding residues. Figure 8.6 diagrams steric clashes caused by proline residues.





**Figure 8.5** Molecular model of proline in a polypeptide  $\alpha$ -helix. The bonds of proline's pyrrolidine ring are coloured blue, 'lost' hydrogen bonds are indicated by red lines, and existing hydrogen bonds by broken lines. The section of  $\alpha$ -helix illustrated, shows residues 83-89 in myoglobin. Oxygen atoms are shown as stippled ellipses, nitrogen atoms as filled ellipses and carbon atoms as unshaded ellipses. Scanned from Barlow and Thornton (1988, p. 613), then coloured.

The lack of a hydrogen bond involving the amino nitrogen of P88 is, as discussed, due to a lack of a proton on this nitrogen. The lack of a hydrogen bond involving the amino hydrogen of L89, is due to a steric effect of P88 (see Figure 8.5, p. 243). The second hydrogen bond is 'lost' because P88 forces a 4.1 Å gap between the L89 amide hydrogen and the N-terminal carbonyl oxygen four peptide bonds away. 4.1 Å is longer than hydrogen bond range.



**Figure 8.6** Steric clashes between the proline  $C\beta H_2$  group and preceding  $C\delta H_2$  and NH groups. Clashes between the proline residue  $C\delta$  and both the  $C\beta$  and amide nitrogen of the preceding residue are shown by red lines.  $i = C\alpha$  of proline;  $(i-1) = C\alpha$  of residue preceding proline;  $C\alpha-C\alpha$  vectors shown as —; ●  $C\alpha$ ; ● C-carbonyl; ● N-amide; pyrrolidine bonds are shown in blue. Note that part of the  $C\alpha-C\alpha$  vector is also part of the pyrrolidine ring. Black and white version scanned from MacArthur and Thornton (1991, p. 401), then coloured.

Proline is also detrimental to the  $\alpha$ -helical conformation because, as a secondary amide, proline is a 'relatively polar' non-polar residue (its classification). As such, it forms strong hydrogen bonds with neighbouring amino acid residues, causing structural instability of  $\alpha$ -helix segments. A multiplicity of local conformations become possible (Deber et al., 1990, p. 155).



Because of their importance to the final conformation and functioning of proteins, steric interactions between amino acids in polypeptides have been examined extensively. One further rigid geometry analysis is considered here, the conformation adopted by amino acid side chains. Ramachandran and Lakshminarayanan (1966) established that amino acid side chains could adopt one of two minimum energy conformations, the  $\alpha$ -conformation or the  $\beta$ -conformation. Zimmerman et al. (1977) employed the Empirical Conformational Energy Program for Peptides to refine amino acid conformation classification by including side chain bond angles, intramolecular hydrogen bonds, entropy states and intrinsic torsional potentials as conformation determinants. Thomasson and Applequist (1990) described the  $\alpha$ - and  $\beta$ -conformations adopted by the pyrrolidine ring of proline as the puckered 'up' and the puckered 'down' conformations.

Summers and Karplus (1990) calculated that the  $\alpha$ -conformation of a residue preceding a proline was 7 kcal/mol less favourable than the  $\beta$ -conformation. Since the  $\alpha$ - and  $\beta$ -conformations are found equally among known protein sequences, this meant that, not only were two hydrogen bonds missing with proline substitutions in  $\alpha$ -helices, but that in approximately half such events an extra 7 kcal/mol  $\alpha$ -helix destabilisation energy was involved.

MacArthur and Thornton (1991, p. 401) summarised the prevailing 'rigid-geometry' view of proline within  $\alpha$ -helices, "These conformational constraints on residues preceding proline have a most important consequence. Location within an  $\alpha$ -helix is in theory impossible, not just because of the lack of hydrogen bonds but also because of this steric constraint". It was largely the calculations of Schimmel and Florey (1968) that established the, "widespread belief that proline residues act as  $\alpha$ -helix 'breakers'." (Barlow & Thornton, 1988,

p. 612). Proline residues were seen to cause an unacceptable disruption to  $\alpha$ -helix hydrogen bonding and to the packing of side chains.

Rigid-geometry calculations conveniently explained the apparent lack of proline in the centre of periodic secondary structures such as the  $\alpha$ -helix or  $\beta$ -pleated sheet. When proline did occur in  $\alpha$ -helices, it was usually located in the first turn, acting as an N-capping residue, marking the start of the main chain hydrogen bonds that characterise the  $\alpha$ -helix (Richardson et al., 1988, p. 1650).

### **8.3.3 The Validity of Rigid-Geometry Calculations**

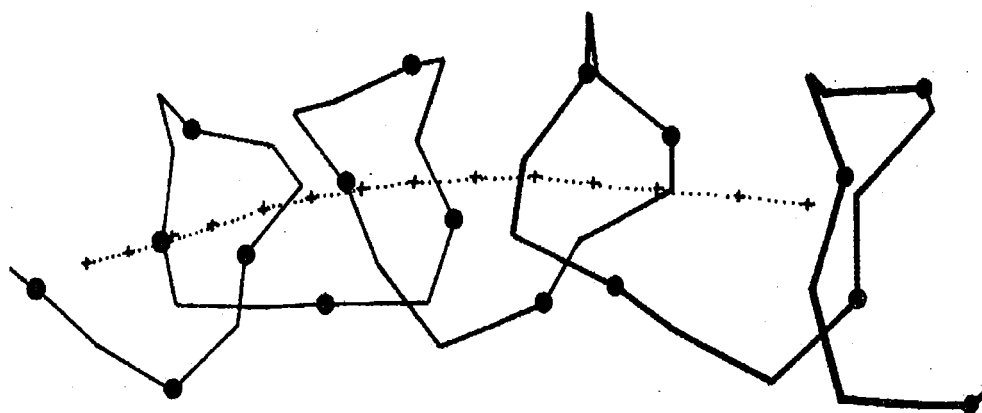
Observations have subsequently shown that proline does occur within  $\alpha$ -helices (Barlow & Thornton, 1988; Kumar & Bansal, 1996). The presence of proline in  $\alpha$ -helices was sufficient for these and other workers to reappraise the doctrine of 'proline the  $\alpha$ -helix breaker' as developed from rigid-geometry calculations. Two rigid-geometry tenets came under scrutiny: that all bond lengths and bond angles have fixed, ideal values; and that  $\alpha$ -helices are always regular.

Hurley et al. (1992) challenged the notion that all bond lengths and bond angles have fixed, ideal values. They reconsidered the energy calculations of Summers and Karplus (1990) using flexible-geometry, which differed from rigid-geometry only in that bond lengths and angles were allowed to relax to the nearest energy minimum. The main chain torsion angles,  $\phi$  and  $\psi$  and the conformation of residues were not fixed or independent of adjacent residues. The flexible-geometry calculations of Hurley et al. (1992, p. 1444), found that the energy cost of a residue in  $\alpha$ -conformation preceding a proline was only 1.1 kcal/mol, not the 7 kcal/mol determined by Summers and Karplus (1990).

Barlow and Thornton (1988, pp. 606-612) challenged the second tenet of rigid-geometry, that all  $\alpha$ -helices were regular. These authors analysed 48  $\alpha$ -helices in globular proteins and calculated the axis of each helix using a purpose built computer program HBEND. Twenty-eight (58.3%) of these  $\alpha$ -helices were curved, seven (14.6%) were linear, five (10.4%) were malformed throughout their length and eight (16.7%) were kinked. For curved helices, their radius of curvature was approximately 60 Å, corresponding to a deviation of about 1° of arc per residue from the arc defined by a linear  $\alpha$ -helix. Figure 8.7 (p. 247) shows the curvature of an  $\alpha$ -helix. The energy penalty calculated for a curvature deviation of one arc degree per residue was approximately 0.1 kilocalories/mole/residue, equivalent to 0.4184 Joules/mole/residue. This penalty was so small that Barlow and Thornton (p. 616) considered that it could:

account for the predominance of curved helices. . . . Although the perturbation required to cause curvature of the helix are small at the level of the individual residues, they can be greatly amplified in a long helix. . . . Thus the relative positions of the side-chains in a curved  $\alpha$ -helix, are quite different from those in a linear helix. This may be of benefit during protein evolution, since mutations that lead to changes in side-chain volume can be accommodated by minor changes in the main chain conformation e.g. a linear helix may become curved (or a curved helix, more curved) in order to preserve an optimal packing of side-chains.

MacArthur and Thornton (1991, pp. 398-399) identified three further mechanisms operating at the secondary level of protein structure, that could attenuate the ' $\alpha$ -helix-breaking' reputation of proline located within  $\alpha$ -helices. Proline isomerization,  $\alpha$ - and  $\beta$ -conformation of residues, and the presence of proline-stabilising hydrogen bonds.



**Figure 8.7** A curved helix involving residues 14 to 32 in avian pancreatic polypeptide. The axis generated by HBEND program is represented by a series of broken lines joining consecutive axis points which are shown as pluses (+). Main-chain  $C^\alpha$  atoms are shown as filled circles. Scanned from Barlow and Thornton (1988, p. 601).

In proteins, isomerization is measured relative to preceding amino acid conformation. With the exception of proline, the frequency of *cis*-amino acid isomers in globular proteins is  $<0.001$  (Brandts et al., 1975). Descriptively, the *cis*-configuration means that the side chains of adjacent amino acids lie on the same side of the  $C^\alpha$ - $C^\alpha$  vector joining them. Proline adopts the *cis*-isomer of the preceding peptide bond with a frequency of 0.057 (58/1021), 57 times more frequently than any other amino acid (MacArthur & Thornton, 1991, p. 400). However, out of the 166  $\alpha$ -helices containing proline, only two prolines adopted the *cis*-configuration. Both of these prolines were the third residue in their respective  $\alpha$ -helix, P89 in cytochrome  $P_{450}$  and P131 in malate dehydrogenase (MacArthur & Thornton, 1991, p.407). Within  $\alpha$ -helices, therefore, proline almost always adopts the *trans*-configuration.

These authors also noted that the  $\alpha$ - or  $\beta$ -conformation adopted by proline in an  $\alpha$ -helix was influenced by the nature of the preceding residue. For example, when arginine was the preceding residue (as is the case for the A1663P substitution in  $\beta$ -MyHC), the proline conformational ratio was 24  $\alpha$ -conformation: 76  $\beta$ -conformation. The  $\beta$ -conformation was three times more likely than the  $\alpha$ -conformation. Chothia (1984) had earlier suggested that the pressures on an  $\alpha$ -helix that arise from packing effects could be relieved by changes in side chain conformation of all residues. Selectively adopting the  $\alpha$ - or  $\beta$ -conformation was a general phenomenon, but one which could help  $\alpha$ -helices accommodate proline residues.

Finally, MacArthur and Thornton (p. 401) pointed out that the steric conflicts between the  $C^{\delta}H_2$  attached to the pyrrolidine amide nitrogen and the  $C^{\beta}H_2$  in the preceding residue and also between the NH of the peptide bond connecting the two preceding residues (Figure 8.6, p. 243), "may be stabilised by hydrogen bonding to either the amide of residue  $X(i-1)$  [the residue preceding proline] or the carbonyl of residue  $(i+2)$ ".

The theoretical analyses of proline in  $\alpha$ -helices by Barlow and Thornton (1988), MacArthur and Thornton (1991), Chothia (1984) and Hurley et al. (1992), identified mechanisms by which proline within  $\alpha$ -helices might not cause the level of distortion predicted by rigid-geometry analysis. An examination of the observed secondary structure of  $\alpha$ -helices containing internal proline will now be used to test the accuracy of these analyses.

#### **8.3.4 Secondary Structure of $\alpha$ -helices with Internal Proline**

Fibrous proteins like myosin II are only a tiny fraction of all proteins. They do not function as monomers *in vivo*. Fibrous protein monomers are very long and

always aggregate to form dimers, trimers, tetramers or pentamers, (Lupas, 1996, p. 375). Therefore, the conformation of each constituent  $\alpha$ -helix is, in part, a function of aggregation/quaternary structure. This makes it impossible to determine the effect of proline substitutions on secondary structure. Proline does, however, exist in 'monomer' sections of  $\alpha$ -helices within globular proteins. The effect of internal proline on the secondary structure of globular protein  $\alpha$ -helices was examined, therefore, as 'the best available option'. The major differences between globular and fibrous proteins are the obvious difference in length and the decreased solubility of fibrous proteins. Protein solubility decreases with size.

In a survey of 291  $\alpha$ -helices in 57 proteins, Barlow and Thornton (1988, p. 612-614) discovered ten internal prolines. They found that proline only occurred in relatively long helical sections of mean length of greater than 5 turns. These authors concluded that proline had been selected to kink the helix. They based this conclusion on the fact that the  $\alpha$ -helices were all kinked in the vicinity of the proline, that these proline were highly conserved, and that they lay within sequences which had no common pattern. Each of these  $\alpha$ -helices was kinked because of the presence of the proline, and only because of the proline.

Barlow and Thornton (1988, pp. 616-618) also noted that proline-induced kinks occurred near residues close to active sites, from which they developed a list of proposed functions for proline-induced kinks. All of the functions they identified caused only very local effects. Of interest to this project was the Barlow and Thornton proviso (p. 618) that, "The kinks that are caused by proline residues are tolerated only by long helices that are packed against several other secondary structures". The myosin rod is extremely long and fits the description of being 'packed against several other secondary structures' extremely well.

More proline residues have been uncovered in globular proteins since the work of Barlow and Thornton. Kumar and Bansal (1998) analysed 205 non-homologous globular proteins whose three-dimensional structures had been solved by X-ray crystallography. These authors identified 1131 non-identical  $\alpha$ -helical sequences of 9 or more amino acid residues, the minimum helix length required by their helix analysing program, HELANAL (Kumar & Bansal, 1996).

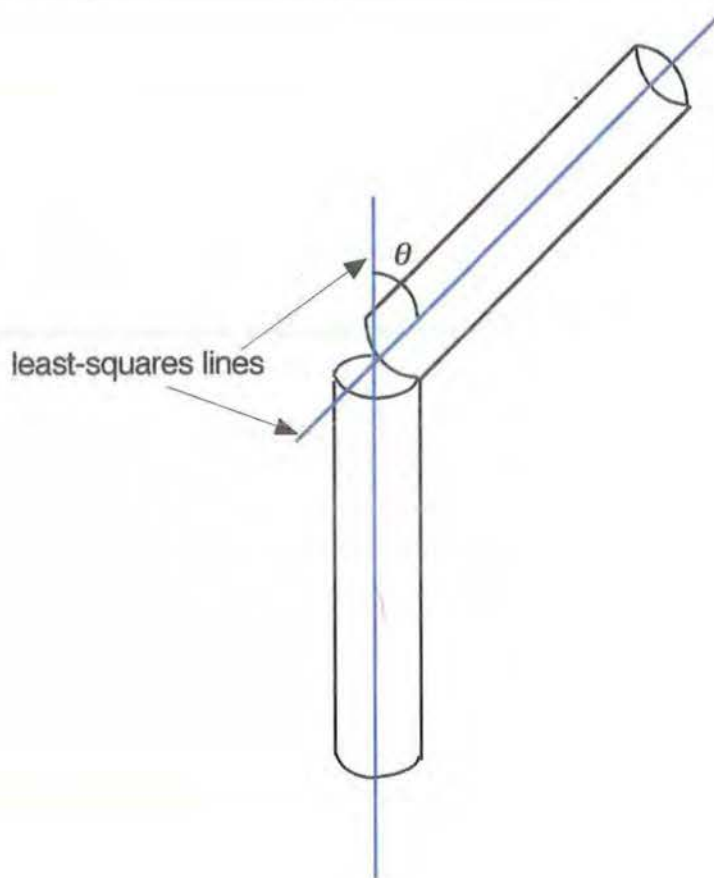
HELANAL identified a total of 45 kinks in the 1131  $\alpha$ -helical regions. Significantly, 25 of these 45 kinks were caused by the presence of proline, which created an average kink angle of  $32^\circ \pm 8^\circ$ , with a range of  $17^\circ$ - $58^\circ$ . The kink angle was calculated as the angle between the least-squares lines, fitted to the local helix fragments before and after the region of kink. On average, two main-chain hydrogen bonds were 'lost' in proline-induced  $\alpha$ -helical kinks (Kumar & Bansal, 1998, pp. 1937-1940). 'Lost' hydrogen bonds in proline-induced kinks were identified in Figure 8.5 (p. 242). The average proline kink angle of  $32^\circ \pm 8^\circ$  determined by Kumar and Bansal agreed with an earlier estimate of  $26^\circ \pm 5^\circ$  by Barlow and Thornton (1988, p. 612-613) who commented:

The conformational changes that take place to accommodate proline vary from one structure to another. In most cases the changes are relatively minor, so that all residues retain the  $\phi/\psi$  angles appropriate for a right-handed helix. . . . All ten of the helices show surprisingly little disruption of the hydrogen bonding.

The average length of the ten proline-containing  $\alpha$ -helices identified by Barlow and Thornton, was 19.9 residues with proline positioned, on average, 10.5 residues from the start of the helical region. These were, therefore, relatively short  $\alpha$ -helices with truly internal prolines.



In comparison, the  $\beta$ -MyHC  $\alpha$ -helical rod contains 1096 residues. It is remarkable that even in the short  $\alpha$ -helical regions described by Barlow and Thornton and Kumar and Bansal, the degree of distortion caused by internal proline is relatively small. Distortion to an  $\alpha$ -helix caused by proline is illustrated in Figure 8.8.



**Figure 8.8** A schematic diagram showing how the  $\alpha$ -helical kink angle ( $\theta$ ) is defined. The  $\alpha$ -helix is represented by a broken cylinder, and the axes, or 'least-squares lines' of the N- and C-terminal 'halves' are shown as blue lines. The analysis of globular protein  $\alpha$ -helical regions containing internal proline, indicates that proline residues introduce only moderate kinks. Redrawn from Barlow and Thornton (1988, p. 613).

## 8.4 QUATERNARY STRUCTURE OF MYOSIN II

The quaternary structure of myosin II is the complete three-dimensional folding pattern of the two MyHC polypeptides, the two essential light chains and the two regulatory polypeptides. The rod section of myosin II comprises two  $\beta$ -MyHC molecules wound together in a coiled-coil dimer superhelix. Figure 8.1 (p. 219) is a two dimensional representation of the six polypeptides in a myosin II molecule. It shows the coiled-coil section of myosin II.

The coiled-coil structure was first proposed by Crick (1952, p. 882; 1953, p. 689) when he noticed that some characteristics of the  $\alpha$ -keratin X-ray diffraction pattern did not agree with that calculated for a perfect  $\alpha$ -helix. In particular, the presence of meridional arcs at spacings of about 5.15 Angstroms ( $\text{\AA}$ ) disagreed with a predicted  $\alpha$ -helix intensity of zero. To explain the observed 5.15  $\text{\AA}$  meridional arcs of the  $\alpha$ -keratin X-ray diffraction pattern, Crick (1952) suggested that the  $\alpha$ -keratin might be deformed so that two identical keratin  $\alpha$ -helices could wind round each other to form an  $\alpha$ -keratin dimer in the form of a coiled-coil.

### 8.4.1 Deformed $\alpha$ -helices and Coiled-Coils

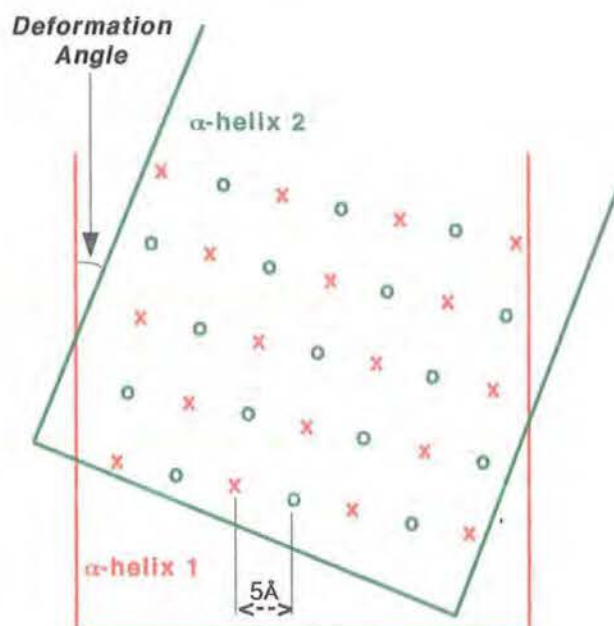
Deformation of  $\alpha$ -helices was necessary because the coiled-coil structure proposed by Crick required the regular meshing of the amino acid side chains of each  $\alpha$ -helix as they wound around each other. In the perfect  $\alpha$ -helix, the position of amino acid side chains repeat only every 18 residues, or after 5 turns ( $5 \times 3.6 = 18$ ). Although two keratin  $\alpha$ -helices start by meshing their side chains, this would not continue for long. To avoid separation of  $\alpha$ -helices, side chains must occupy equivalent positions turn after turn after turn. This was an impossibility for perfect, undistorted, 3.6 residues per turn  $\alpha$ -helices.

Crick (1953, pp.690-691) developed the basic coiled-coil packing scheme for two  $\alpha$ -helices of the same sense and direction. Starting with an idealised  $\alpha$ -helix, with side chains of uniform size and shape:

Let us pretend that we wrap around it at a radius of about 5 Å, a cylinder of paper, and mark on the paper the point where each side chain comes. We then open the paper until it is flat again, and examine the pattern we have drawn. . . . Let us do the same thing for a second  $\alpha$ -helix. Naturally we get the same pattern as before. If we now try to pack our two helices together, the two pieces of paper will come into contact, and the condition that the knobs of one helix fall between the knobs of the other is equivalent to saying that where the two pieces of paper are in contact - or nearly in contact - the marks on one must fall between the marks on the other. If we do this with the two bits of paper opened out flat, however, we must turn one of them over before we superimpose them, since the 'outside' of both sheets must come into contact. It is immediately found that we can only do this in one way if the chains are to be approximately parallel.

Figure 8.9 (p. 254) illustrates how the two pages have to be aligned (deformed) so that the knobs on one  $\alpha$ -helix (marked as crosses) fall between the knobs on the second  $\alpha$ -helix (marked as circles). Crick (1953) calculated that a left-handed twist to right handed  $\alpha$ -helices, would reduce the number of residues per  $\alpha$ -helical turn to 3.5, and allow the position of side chains to repeat after two turns (or seven residues).

The deformation angle shown in Figure 8.9 represented (in two dimensions) the left-handed twist to right handed  $\alpha$ -helices necessary for the formation of left-handed coiled-coils.



**Figure 8.9** The pattern formed by the side chains of two superimposed  $\alpha$ -helices. Side chains are marked on two separate pieces of paper wrapped around each  $\alpha$ -helix. In this representation, the side chains of the  $\alpha$ -helix are marked with orange crosses on one piece of 'wrapping' paper, and with green circles on the other piece of 'wrapping' paper. The pages are arranged so that when helices come into contact, the side chains of one  $\alpha$ -helix fall between those of the other. Redrawn from Crick (1953, p. 691).

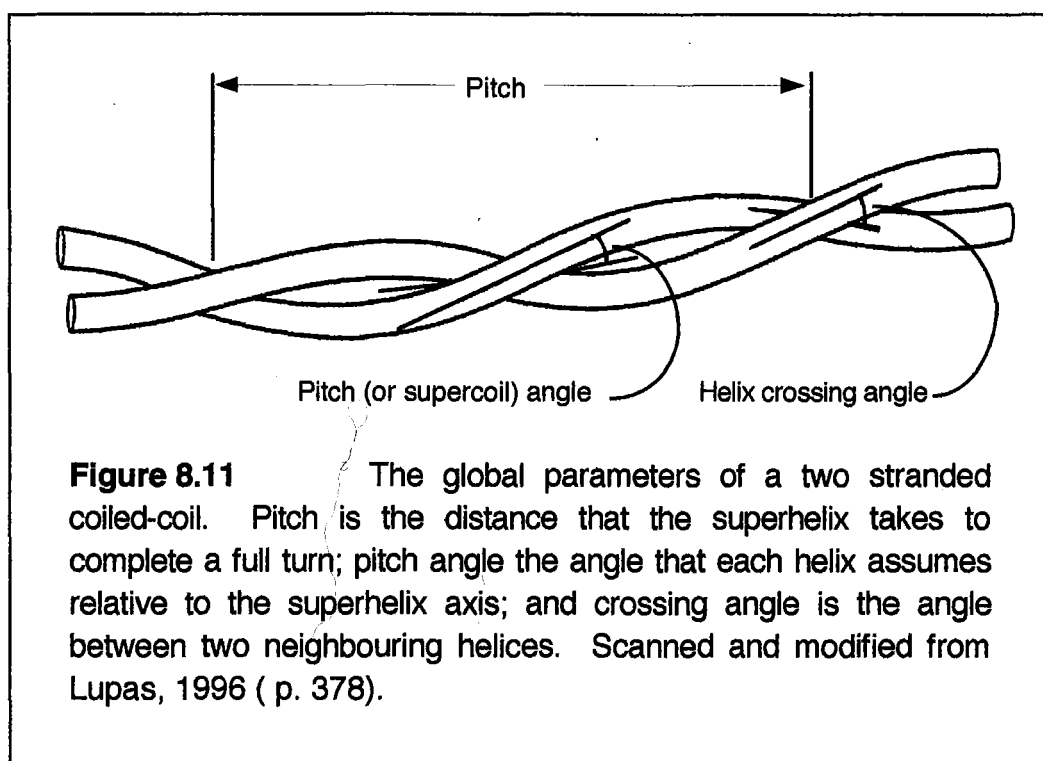
Residues are about 5 Å apart. Crick (1952, p. 882) calculated that the necessary deformation required little energy, approximately 0.1 kilocalories/mole/residue (0.4184 J), and speculated that the driving force behind  $\alpha$ -helical deformation and coiled-coil assembly might simply be the need to remove apolar residues from the aqueous medium. In support, Crick (1953, p. 696) cites an analysis of tropomyosin (Bailey, 1948) which showed that hydrophobic alanine, valine and leucine made up 0.29 ( $\sim 2/7$ ) of the total tropomyosin residues. This was consistent with, but did not prove, Crick's deduction that the apolar residues in coiled-coils occurred at intervals of 3.5

residues, always pointed inwards, and fitted like 'knobs-into-holes' between the 'core' apolar residues on the other  $\alpha$ -helical strand(s) of the coiled-coil (Crick, 1952, p. 882). Fifty years later, it is known that all dimer coiled-coils like those found in the MyHCs are left-handed superhelices, winding around each other in the opposite direction to their constituent right handed helices (Lupas, 1996, p. 379). Crick's coiled-coil analysis was correct. Figure 8.10 shows a representation of a coiled-coil dimer.



**Figure 8.10** Two right-handed  $\alpha$ -helices forming a left-handed, coiled-coil superhelix. Note that each  $\alpha$ -helix turns in a clockwise direction while winding around the other in an anti-clockwise direction. Crick (1953) proposed that the hydrophobic side chains, shown here in blue, were positioned such that they would form hydrophobic cores in the centre of coiled-coils as the two  $\alpha$ -helices wrapped around each other. An orange and green hetero-dimer is shown (to match the colours used in Figure 8.9). The MyHCs form homo-dimers, built from two identical monomer helices. Scanned and modified from Alberts et al. (1994, p. 125).

Coiled-coils can comprise two, three, four, or (rarely) five strands (Lupas, 1996, p.375). Coiled-coils have two geometries: the geometry of the constituent  $\alpha$ -helices and the geometry of the whole, called a superhelix. Figure 8.11 shows the global parameters of the two-stranded superhelix. The values of pitch, the pitch angle and crossing angle shown in this figure, follow directly from the distortion necessary to achieve the critical coiled-coil 3.5 residue repeat.



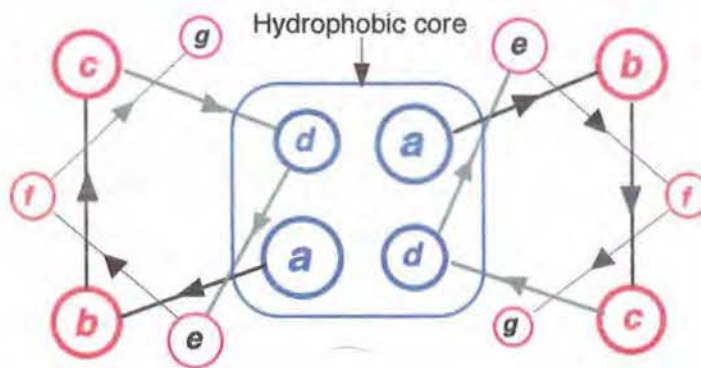
The  $\alpha$ -helices of almost all natural proteins are right-handed (Wolfe, 1993, p. 63). The two  $\alpha$ -helices comprising the MyHC rod are right-handed and deformed in accordance with Crick's calculations, to form a dimer coiled-coil (Brown et al., 1996 p. 136).

#### 8.4.2 MyHC Dimerisation Geometry: Heptad Repeats

The characteristic, 7-residue repeating pattern, *a-b-c-d-e-f-g* (see p. 233), in which hydrophobic residues are concentrated at intervals of 3 and 4 residues at



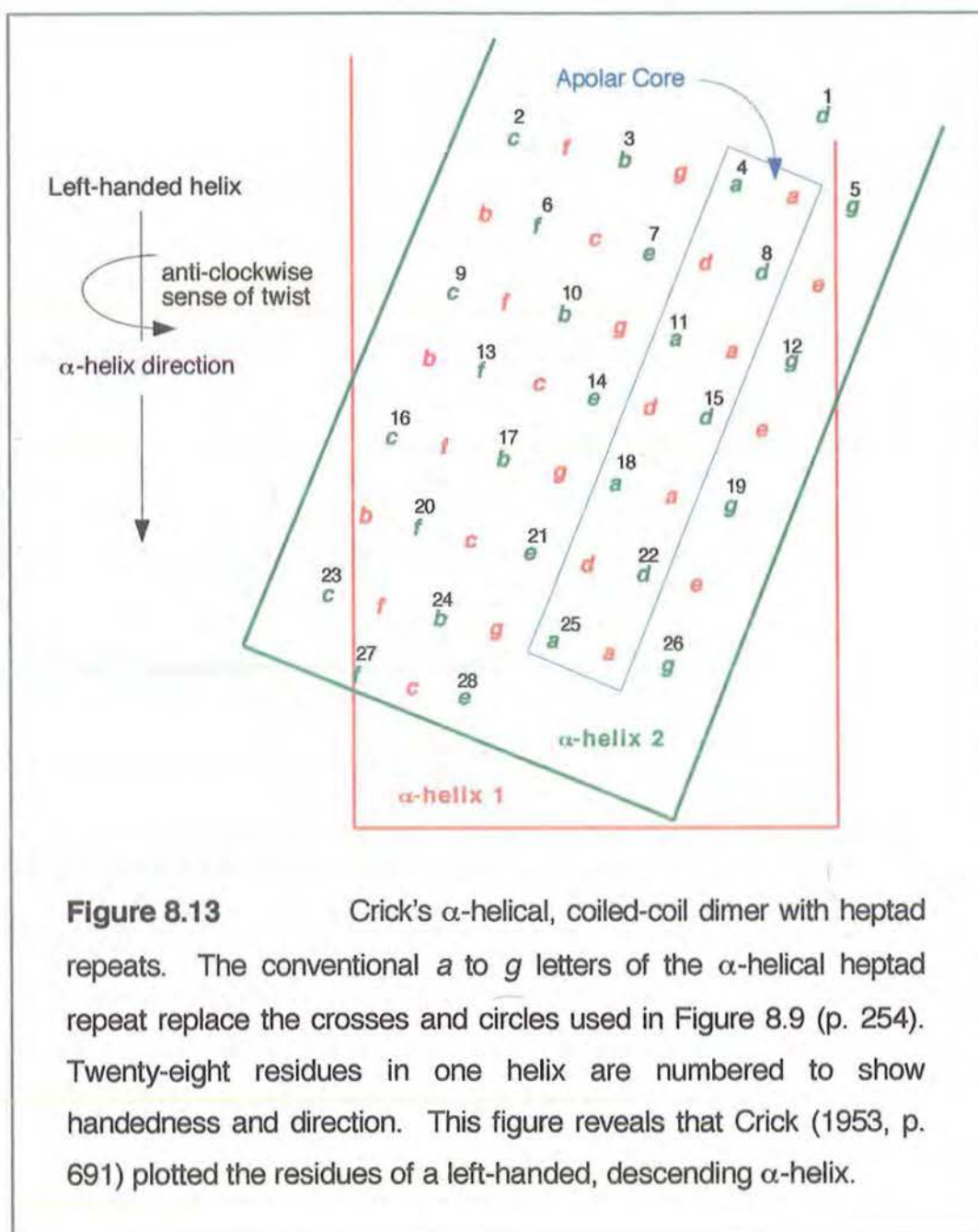
positions *a* and *d* was first clearly identified by McLachlan and Karn (1982, p.227). Although Crick had articulated this possibility thirty years earlier, the structure of sufficient  $\alpha$ -helical coiled-coil proteins had to be determined before a clear pattern was established. The structure of the coiled-coil domains in tropomyosin, fibrinogen,  $\alpha$ -keratin and nematode MyHC provided proof that the heptad repeat was the basic unit of the  $\alpha$ -helices that formed coiled-coil. In globular proteins  $\alpha$ -helices do not have heptad repeats. Figure 8.12 presents a transverse view through two  $\alpha$ -helical turns of a coiled-coil dimer. This is known as a heptad wheel representation.



**Figure 8.12** Two heptad wheels showing the sequence motif of an  $\alpha$ -helical, coiled-coil dimer. This representation descends into the page from position *a*. Note the clockwise arrangement of the heptad letters characteristic of right-handed helices. The letters *a* to *g* mark positions along the  $\alpha$ -helix. Positions *a* and *d* are typically occupied by hydrophobic residues and constitute the core of the coiled-coil dimer. Positions *b*, *c* and *f* constitute the 'outer' surface of the coiled-coil dimer and are typically occupied by polar hydrophilic residues, which are often charged.

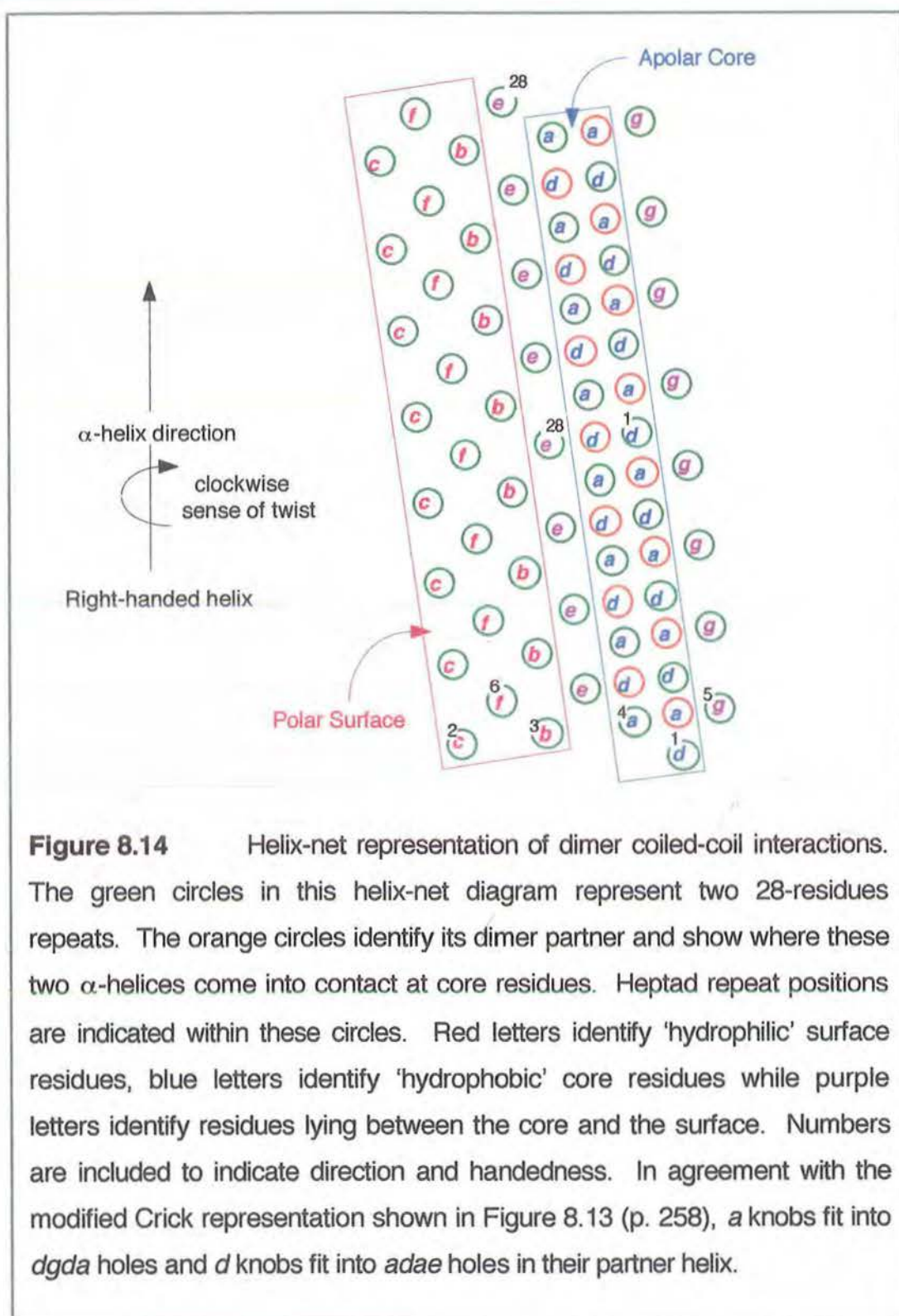
Figure 8.13 (p. 258) adds the essential features of the  $\alpha$ -helical heptad repeat to Crick's model for two interacting deformed  $\alpha$ -helices.





**Figure 8.13** Crick's  $\alpha$ -helical, coiled-coil dimer with heptad repeats. The conventional *a* to *g* letters of the  $\alpha$ -helical heptad repeat replace the crosses and circles used in Figure 8.9 (p. 254). Twenty-eight residues in one helix are numbered to show handedness and direction. This figure reveals that Crick (1953, p. 691) plotted the residues of a left-handed, descending  $\alpha$ -helix.

The advantage of replacing the crosses and circles used by Crick with heptad letters is that it shows which knobs fit into which holes. Clearly, *a* knobs on one  $\alpha$ -helix fit into *dgda* holes on the other  $\alpha$ -helix (reading clockwise from the top), while *a* knobs on the other  $\alpha$ -helix fit into *dgda* holes on the first helix (reading anti-clockwise from the top); similarly *d* and *d* knobs fit into *adae* and *adae* holes on their partner  $\alpha$ -helix. Figure 8.14 also demonstrates this phenomenon, but on the now commonly used helix-net diagram.

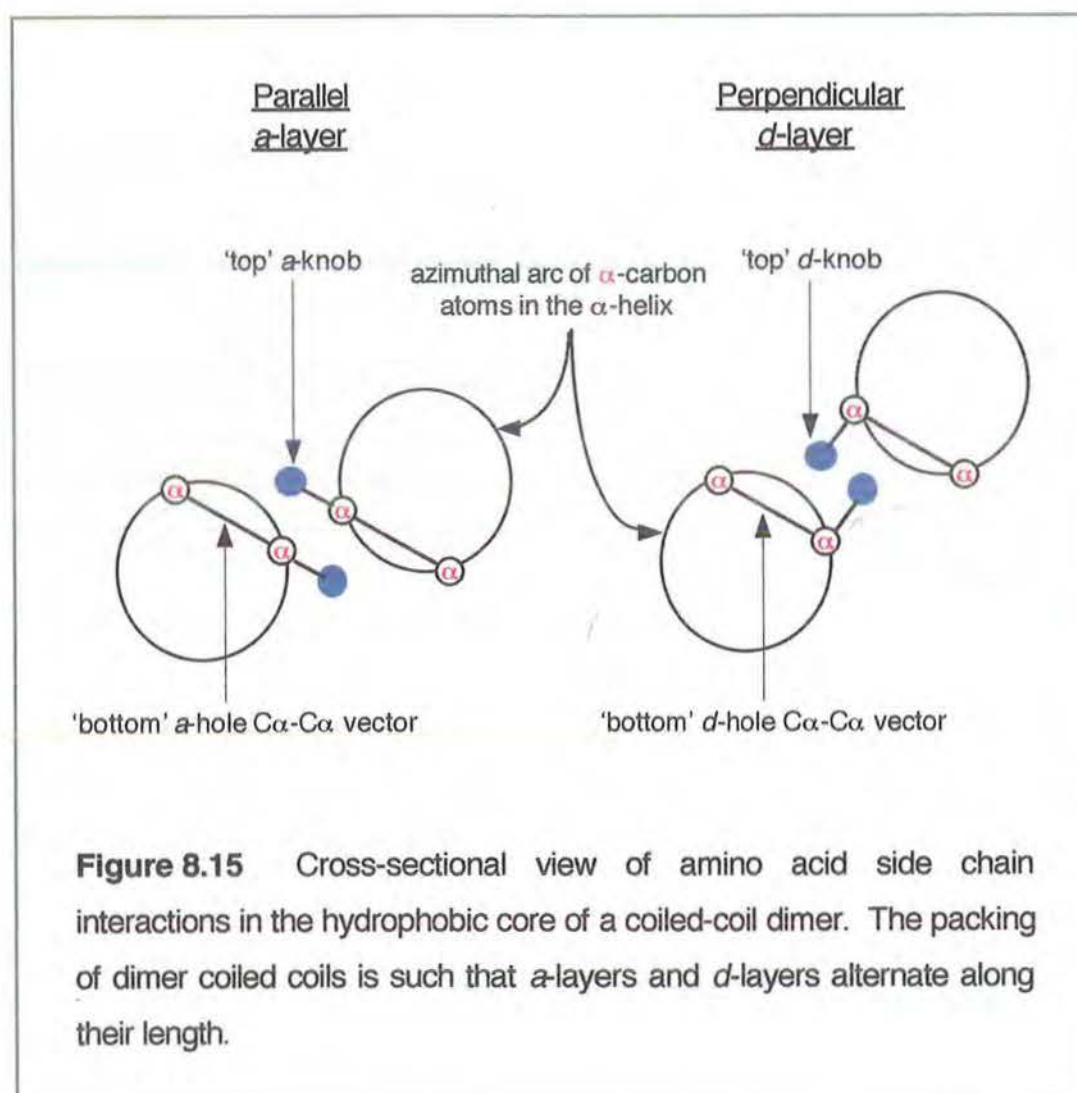


### 8.4.3 A1663P Distorts Perpendicular and Parallel Geometry

The  $\beta$ -MyHC A1663P substitution distorts the  $\alpha$ -helix in which it lies by altering the geometry of dimerisation. Table 8.2 (p. 233) shows that the  $\beta$ -MyHC

A1663P substitution occurs in position *g* in heptad 30.4. This position was part of a *dgda* hole into which fitted an *a* knob from its dimer partner. Not only was the geometry of this hole altered, but the geometry of the *a* knob adjacent to position 30.4*g* was also altered by the  $\beta$ -MyHC A1663P substitution.

Figure 8.15 presents an cross-sectional view of the geometry for the MyHC dimer coiled-coil.



In two-stranded coiled-coils, the  $C\alpha$ - $C\beta$  bond of an *a* core residue is parallel to the  $C\alpha$ - $C\alpha$  vector at the bottom of the facing hole, but at right angles to the  $C\alpha$ - $C\beta$  bond of a *d* core. Residues and holes are described in relation to the

alignment of carbon atoms in amino acids, numbered  $\alpha$  for the centre carbon atom and  $\beta$ ,  $\gamma$ ,  $\delta$  and  $\epsilon$  as shown in Figure 8.3 (p. 232).

Selection will favour the most appropriate amino acids in positions *d* and *a*. Therefore the perpendicular geometry of the *d* position should favour leucine, but not isoleucine, while parallel geometry of the *a* position should favour isoleucine (Harbury et al., 1994).

An inspection of the amino acid sequence determined for WAMPD Laing myopathy proband, III.2. in Table 8.2 (p. 234) reveals that the *d*-heptad position was occupied by 66 leucines and 4 isoleucines. Given the perpendicular side chain geometry of leucine, and therefore its penchant for perpendicular holes, the preponderance of leucine in the *d*-heptad position of the  $\beta$ -MyHC rod, agreed with that predicted by knobs-into-holes geometry.

On the other hand, the *a*-heptad position of the  $\beta$ -MyHC rod was occupied by 15 isoleucines and 45 leucines, only showing a trend towards the occupancy predicted by knobs-into-holes geometry. Since the  $\beta$ -MyHC A1663P substitution is in a *g* position, proline steric clash will disrupt the adjacent *a* knob 'parallel geometry'. In addition, it will disrupt the geometry of the '*a* accepting' *dgda* holes, which contain the altered *g* position. How serious this disruption is, depends on the importance of local knob and hole geometry to the functioning of the myosin II hexamer. There are 312 knobs-into-holes interactions in the myosin II rod. The low isoleucine occupancy of *a*-heptad positions in the  $\beta$ -MyHC rod, and the uncharged hydrophilic, polar residue R in hydrophobic 30.4a position (residue 1664), suggest relaxed parallel-geometry selection at *a* positions, and specifically at 30.4a. Distortion caused by  $\beta$ -MyHC P1663 may be better tolerated at 30.4a than in other positions.

#### **8.4.4 Proline Effect on MyHC Dimerisation: COILS Analysis**

The effect of proline on the ability of MyHC monomers to form multicoil-dimers was analysed using the COILS program, version 2.2 with MTK matrix (Russell & Lupas, 1999). The COILS program, takes into account the residue distribution in known coiled-coils to evaluate submitted sequences. It relies on determined sequence profiles. Table 8.4 (p.263) presents the results of COILS analysis for five peptide sequences. The legend for this table explains how to interpret COILS results and the factors used to select peptides for comparison.

There was one further consideration, relevant to the interpretation of COILS results. Window size 14 was the best window size for identifying local discontinuities, as the status of each residue contributed one fourteenth of the total to determination of probabilities. However, window size 14 was confounded by the fact that COILS prediction reliability decreases strongly with decreasing window size. The combined effects of lower statistical significance of smaller samples and the increasing number of packing options for shorter peptides, even for those with excellent coiled-coil forming potential, means that, "reliable coiled-coil predictions could only be made from 14-residue window if a coiled-coil was consistently predicted in a family of divergent sequences with less than 50% identity" (Lupas, 1996, p. 377). Larger window size reduced this problem, but necessarily diminished the effect of individual residues and, therefore, of individual sequence alterations.

Comparison of human  $\beta$ -MyHC to fruit fly MyHC-A (with 52% homology) showed their quaternary structures both to be dominated by coiled-coil dimers over the entered sequences. The same pattern of coiled-coil dimer also characterised human SMMMyHC, which had only 33.1% homology.



Table 8.4

***COILS Analysis of Polypeptides with Varying Degrees of Homology to Human  $\beta$ -MyHC.***

Results of the COILS program (Russell & Lupas, 1999) were compared from residue 1635M to 1691V for five proteins over the published human  $\beta$ -MyHC sequence (Jaenicke et al., 1990). This included 28 residues on either side of the  $\beta$ -MyHC residue of interest, 1663. To avoid spurious low coiled-coil probabilities at either end of this sequence, the actual sequence entered into COILS ran from residue 1583 to 1750 in human  $\beta$ -MyHC and over the equivalent 168 residues (6 X 28-residue repeats) in the other proteins. The degree of homology between each sequence and human  $\beta$ -MyHC, was also determined from a comparison of these 168 residue sequences.

The sequences compared were: wild type  $\beta$ -MyHC (P12883 - SwissProt Accession Number); WAMPD A1663P  $\beta$ -MyHC; fruit fly (*Drosophila melanogaster*) embryonic muscle MyHC-A (P05661); human smooth muscle SMMMyHC (P35749); and human centromeric protein-E, CENP-E (Q02224).

Row 2 presents the homology of each of these proteins to human  $\beta$ -MyHC. They were selected for comparison to human  $\beta$ -MyHC because the first three of these proteins were all muscle MyHCs, with the third, SMMMyHC, having the lowest homology to  $\beta$ -MyHC of all the human MyHCs. CENP-E was selected as the 'highest homology to  $\beta$ -MyHC' protein that contained a naturally selected proline, designated as **P**.

In its analyses, COILS used the 3 different window sizes of **14**, **21** and **28** adjacent residues. Limitations and advantages of window size are considered in the text. This table presents the probabilities of a residue forming a multicoil-dimer based on their status, the status of adjacent residues and as a function of window size. Probabilities were presented on a scale of 1 to 9, where 1 stood for a probability in the range, 0.1-0.2 and 9 stood for a probability range, 0.9-1.0.

		Cardiac & Skeletal Muscle				Embry. Muscle				Oesophagus				HeLa Cells			
		HOMOLOGY → 98%				52%				33.1%				26%			
		HUMAN β-MyHC				WAMPD β-MyHC				FRUIT FLY MyHC-A				HUMAN SMMMyHC			
		WINDOWS				WINDOWS				WINDOWS				WINDOWS			
		14	21	28		14	21	28		14	21	28		14	21	28	
g	M	5	9	9		M	5	9	9	A	8	9	9	G	3	9	9
a	A	8	9	9		A	8	9	9	N	9	9	9	R	9	9	9
b	A	8	9	9		A	8	9	9	A	9	9	9	E	9	9	9
c	E	8	9	9		E	8	9	9	E	9	9	9	E	9	9	9
d	A	8	9	9		A	8	9	9	A	9	9	9	A	9	9	9
e	Q	8	9	9		Q	8	9	9	Q	9	9	9	I	9	9	9
f	K	8	9	9		K	8	9	9	K	9	9	9	K	9	9	9
g	Q	8	9	9		Q	8	9	9	N	9	9	9	Q	9	9	9
a	V	8	9	9		V	8	9	9	I	9	9	9	L	9	9	9
b	K	8	9	9		K	8	9	9	K	9	9	9	R	9	9	9
c	S	8	9	9		S	8	9	9	R	9	9	9	K	9	9	9
d	L	8	9	9		L	8	9	9	Y	9	9	9	L	9	9	9
e	Q	8	9	9		Q	8	9	9	Q	9	9	9	Q	9	9	9
f	S	8	9	9		S	8	9	9	Q	9	9	9	A	9	9	9
g	L	8	9	9		L	8	9	9	Q	9	9	9	Q	9	9	9
a	L	8	9	9		L	8	9	9	L	8	9	9	M	9	9	9
b	K	3	9	9		K	3	9	9	K	8	9	9	K	9	9	9
c	D	3	9	9		D	3	9	9	D	8	9	9	D	9	9	9
d	T	3	9	9		T	3	9	9	I	8	9	9	F	9	9	9
e	Q	3	9	9		Q	3	9	9	Q	8	9	9	Q	9	9	9
f	I	3	9	9		I	3	9	9	T	8	9	9	R	9	9	9
g	Q	3	9	9		Q	3	9	9	A	8	9	9	E	9	9	9
a	L	7	9	9		L	-	2	8	L	9	9	9	L	9	9	9
b	D	7	9	9		D	-	2	8	E	9	9	9	E	9	9	9
c	D	7	9	9		D	-	2	8	E	9	9	9	D	9	9	9
d	A	7	9	9		A	-	2	8	E	9	9	9	A	9	9	9
e	V	7	9	9		V	-	2	8	Q	9	9	9	R	9	9	9
f	R	7	9	9		R	-	2	8	R	9	9	9	A	9	9	9
g	A	7	9	9		P	-	2	8	A	9	9	9	S	9	9	9
a	N	8	9	9		N	8	9	9	R	9	9	9	R	-	6	9
b	D	8	9	9		D	8	9	9	D	9	9	9	D	-	6	9
c	D	8	9	9		D	8	9	9	D	9	9	9	E	-	6	9
d	L	8	9	9		L	8	9	9	A	9	9	9	I	-	6	9
e	K	8	9	9		K	8	9	9	R	9	9	9	F	-	6	9
f	E	8	9	9		E	8	9	9	E	9	9	9	A	-	6	9
g	N	8	9	9		N	8	9	9	Q	9	9	9	T	-	6	9
a	I	8	9	9		I	8	9	9	L	2	9	9	A	6	9	9
b	A	8	9	9		A	8	9	9	G	2	9	9	K	6	9	9
c	I	8	9	9		I	8	9	9	I	2	9	9	E	6	9	9
d	V	8	9	9		V	8	9	9	S	2	9	9	N	6	9	9
e	E	8	9	9		E	8	9	9	E	2	9	9	E	6	9	9
f	R	8	9	9		R	8	9	9	R	2	9	9	K	6	9	9
g	R	8	9	9		R	8	9	9	R	2	9	9	K	6	9	9
a	N	5	9	9		N	5	9	9	A	4	9	9	A	9	9	9
b	N	5	9	9		N	5	9	9	N	4	9	9	K	9	9	9
c	L	5	9	9		L	5	9	9	A	4	9	9	S	9	9	9
d	L	5	9	9		L	5	9	9	L	4	9	9	L	9	9	9
e	Q	5	9	9		Q	5	9	9	Q	4	9	9	E	9	9	9
f	A	5	9	9		A	5	9	9	N	4	9	9	A	9	9	9
g	E	5	9	9		E	5	9	9	E	4	9	9	D	9	9	9
a	L	8	9	9		L	8	9	9	L	4	7	9	L	9	9	9
b	E	8	9	9		E	8	9	9	E	4	7	9	M	9	9	9
c	E	8	9	9		E	8	9	9	E	4	7	9	Q	9	9	9
d	L	8	9	9		L	8	9	9	S	4	7	9	L	9	9	9
e	R	8	9	9		R	8	9	9	R	4	7	9	Q	9	9	9
f	A	8	9	9		A	8	9	9	T	4	7	9	E	9	9	9
g	V	8	9	9		V	8	9	9	L	4	7	9	D	9	9	9



Since human SMMMyHC was the human MyHC with the least homology to human  $\beta$ -MyHC (33.1%), it appeared that the coiled-coil dimer dominated the quaternary structure of the rod regions of all known human MyHCs.

The COILS predicted effect of the  $\beta$ -MyHC A1663P substitution on  $\beta$ -MyHC rod region, can be seen in the WAMPD  $\beta$ -MyHC column. Coiled-coil probabilities were reduced only in the heptad on the N-terminal side of P1663, by an amount that was a function of window size. Importantly, all window sizes predicted an effect over one heptad only. COILS predicted a local effect. This limited effect agreed the findings of Barlow and Thornton (1988, p. 612; discussed on p. 250). The N-terminal effect of P1663 was also in agreement with the arguments already presented in Section 8.3.2 (p. 241-243), that proline-induced steric clashes and 'lost' hydrogen bonds lay on the N-terminal side of proline within  $\alpha$ -helices.

Human centromeric protein-E (CENP-E) appears in Table 8.4 because it was the only coiled-coil dimer with an internal proline. CENP-E has only 26% homology to the  $\beta$ -MyHC. This protein is located in the kinetocore complex, a macromolecular structure within the centromeric heterochromatin of mitotic chromosomes. Like myosin II, it is a 'molecular motor', but is involved in the movement of chromosomes not muscle.

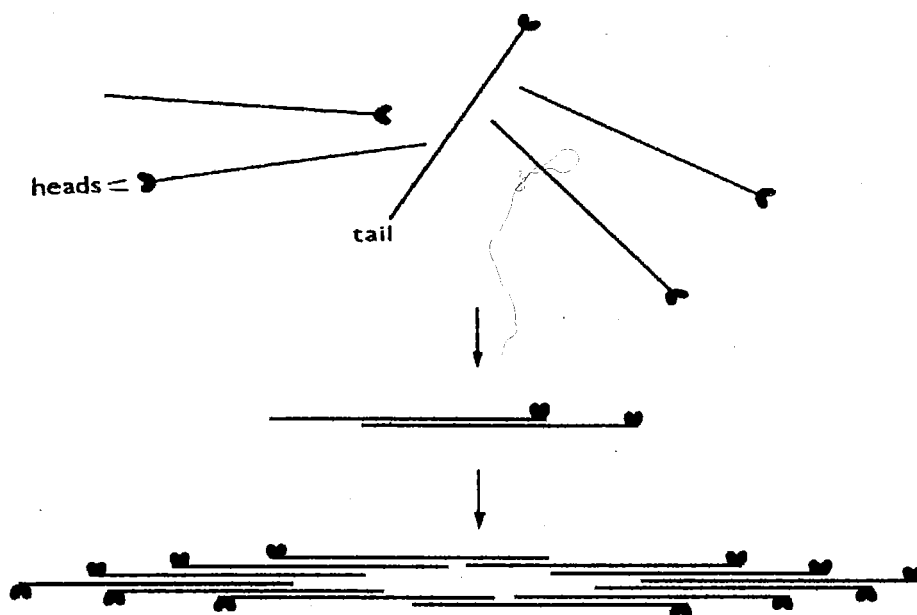
COILS predicted that the effect of the CENP-E proline and the  $\beta$ -MyHC P1663 were similar. Both produced a mild-kink on N-terminal side, as predicted by Barlow and Thornton (1988) and Kumar and Bansal (1996). It was reassuring that the COILS program gave essentially the same result for the  $\beta$ -MyHC P1663 substitution and for CENP-E, with its naturally selected proline. The importance of this result is that it partly assuaged doubts that always linger

with program analyses. They analyse only according to their settings. In 1996 Lupas (p.377), the original author of the COILS program, expressed this doubt, "...never simply believe a computer output. Programs are a great help, but nothing replaces the informed opinion of the researcher."

The opinion of Dr John Sparrow, Department of Biology, York University was sought regarding the likely effect of  $\beta$ -MyHC A1663P. In a personal communication (2000) he commented that, "such a mutation in the rod may lead to disruption of rod formation and/or dimer-dimer interactions [necessary] to form a thick filament" and that, "mutant/+ or mutant/mutant dimers might additionally interfere with higher order processes in thick filament assembly e.g. dimer-dimer binding. It might be especially potent at the outset where the myosins assemble at what will be the centre of the thick filament in an anti-parallel manner."

## **8.5 SUPERSTRUCTURE OF MYOSIN II**

The superstructure of myosin II is the thick filament. Thick filaments are bipolar spindle shaped structures, 1.6  $\mu\text{m}$  in length and 15 nm in diameter, which dissolve when treated with high ionic solution (Huxley, 1963). EM revealed that myosin molecules treated in this way were myosin dimers and that each thick filament was a polymer of about 300 molecules. Negatively stained EM preparations of isolated filaments had a rough appearance along their length except for a centrally located 'bare zone'. Huxley (1963) proposed a simple thick filament assembly model to account for the appearance of the thick filaments preparations he observed under EM. Figure 8.16 (p. 266) shows Huxley's thick filament assembly model.



**Figure 8.16** Model of assembly of myosin molecules to form the thick filament. Molecules initially associated in a head-to-tail, antiparallel array with the 'tail-only' region accounting for the bare zone. The filaments then elongate by staggered, head to tail, parallel addition of molecules at each end. Scanned from R. Craig in Engel (1994, p. 141).

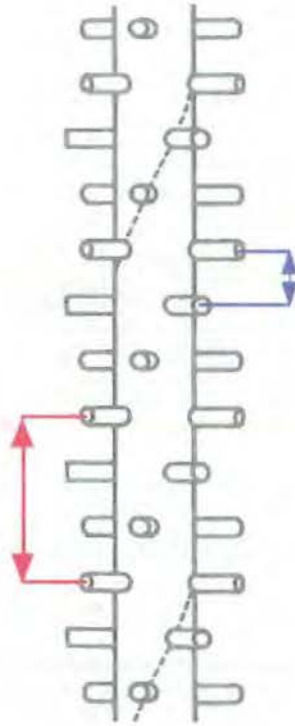
### 8.5.1 Thick Filament Formation: Charged-Repeat Attraction

Heptad repeats are not the only repeats found in coiled-coils. Internal comparison of the amino acid sequence of the *C. elegans* myosin rod region of the nematode showed that the entire helical sequence of each coil could be mapped out as a repeating pattern of 28 residues. A typical MyHC rod 28-residue zone divides into eight hydrophobic zones and six alternating bands of positive and negative charge which are spaced so that the strongest position of the positive charge (position 1*b*) is exactly 14 residues (or one half zone) away from the strongest peak of negative charge (position 3*b*) (McLachlan & Karn, 1982, p. 227).

The high density of charged groups and the regular bands of charge on the on the outer surface of the myosin rod, led McLachlan and Karn (1982, pp. 228-230) to suggest that ionic attractions were most important in the assembly of thick filaments. They suggested that rod sections of myosin molecules had staggered alignment so that positive and negative charges came together. Strong attractions were the organising force of thick filament formation.

The distribution of charge on the MyHC rod indicated that the coils in the MyHC dimer coiled-coil would most strongly attract each other when they were staggered by odd multiples of half-repeats, or 14 residues. Particularly strong attractions were predicted for staggers of 7 'half-repeats' or 98 residues, and for 21 'half repeats' or 294 residues (McLachlan & Karn, 1982, pp. 230-231). The charge-repeat periodicity ratio of 3:1 (294:98) was the same as the periodicity ratio of axial spacings in muscle. X-ray diffraction patterns had shown the axial distance between adjacent actin-myosin cross-bridges was 143 Å, and the repeat distance of the helix was 430 Å ( $3 \times 143 \text{ Å} = 429 \text{ Å}$ ) (Poulsen & Lowy, 1983). These periodicities are diagrammed in Figures 8.17 (p. 268).

It was the coincidence of charge-repeat periodicity and helix : cross-bridge periodicity that suggested to McLachlan and Karn that the rod charge-repeat periodicities were responsible for the arrangements of cross-bridges on the surface of the thick filament. It elegantly explained the same periodicities in the rod sequence, X-ray diffraction patterns and myosin electron micrographs (Craig, 1994). Since the A1663P substitution in  $\beta$ -MyHC replaced one non-polar amino acid with another non-polar amino acid, there was only one way in which this substitution could affect thick filament formation: by altering charge distribution through azimuthal distortion of the rod.



**Figure 8.17** Schematic diagram of cross-bridge and helical periodicity in thick filaments as determined from X-ray diffraction patterns. The 'rod' is the myosin II thick filament backbone, the 'pegs' are the cross-bridges to actin. Blue identifies the 143 Å axial periodicity, while red identifies the 430 Å axial periodicity. Scanned and retouched from R. Craig in Engel (1994, p. 142).

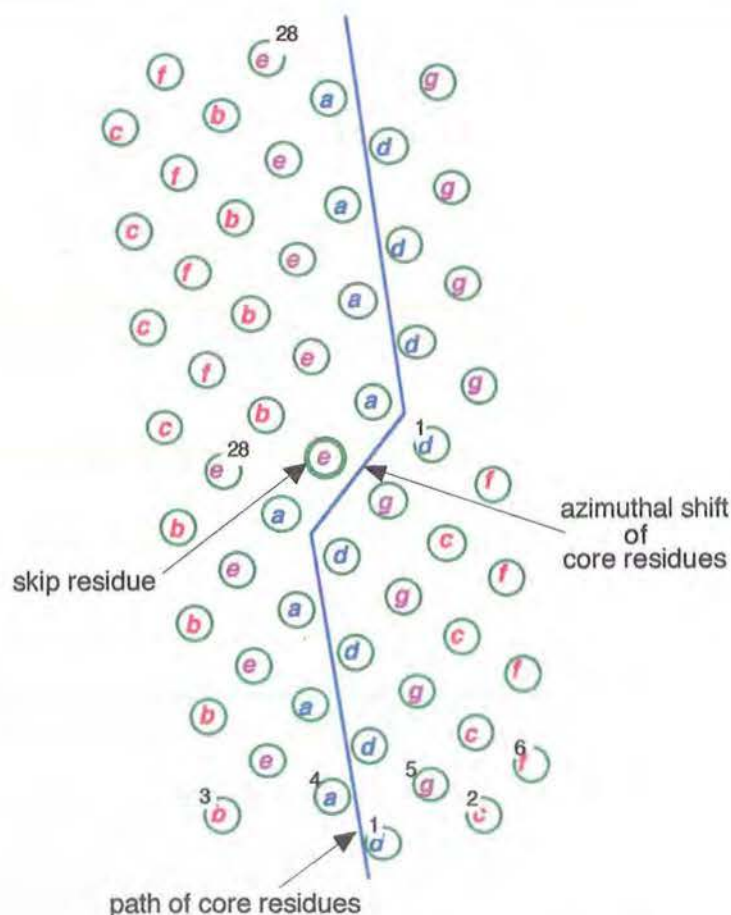
### 8.5.2 Thick Filament Formation: Charge-Phase Distortion

Skip residues distort the charge-phase of the MyHC rod. Only a few coiled-coils are without skip residue, charge-phase distortions. An analysis of their effect may indicate if, and how, the  $\beta$ -MyHC A1663P substitution could do likewise. The four skip residues in human  $\beta$ -MyHC rod all occur within four 28-residue repeat. Conventionally they are placed at the end of that repeat. The first skip residue is the uncharged polar residue, threonine, at position 1188, the

next two skip residues are the charged polar residue, glutamic acid, at positions 1385 and 1582, and the fourth skip residue is the uncharged non-polar residue, glycine at position 1807 (Table 8.2, p. 234). The high degree of  $\beta$ -MyHC skip-residue conservation from nematodes to humans indicates that skip residues are likely to have global as well as local consequences, possibly affecting thick filament formation (Brown et al., 1996, p. 143).

As early as 1982, McLachlan and Karn (p. 228) suggested that skip residues might have important effects on the way myosin molecules packed in the cross-sections of the thick filaments. Although two of the nematode MyHC rod region skip residues were charged polar residues, and therefore altered the charge on the MyHC rod, these authors focussed on the azimuthal effect of skip residues on charge-phase distortion, and reported that skip residues cause: the  $\alpha$ -helix to have a wider turn because the twist of the supercoil was under-wound for a few turns; and the seam of hydrophobic residues to shift across the helix surface for 14-21 residues.

Brown et al. (1996, pp.135-136) reported that the movement, or azimuthal shift, caused by skip residues was about  $+100^\circ$ , and the displacement caused by each skip residue was resolved within 2 to 3 heptads, at which point the regular coiled-coil pattern resumed. In each skip region, the supercoil accommodated the extra residue by yielding to local distortions without severe alteration to the structure (Lupas, 1996, p. 378-379). Figure 8.18 (p. 270) shows the effect of a skip residue in an  $\alpha$ -helix net diagram. Brown et al. (p. 144) noted that an important consequence of skip residues, "would be the alteration of the azimuthal position of the periodic patches of charged and/or polar residues along the coiled-coil that are known to play a primary role in the aggregation of some  $\alpha$ -fibrous proteins".



**Figure 8.18** A single skip residue causing a 100° azimuthal shift of MyHC hydrophobic core residues. Core residues in this helix-net diagram were positioned more centrally than in Figure 8.14 (p. 259). This was a diagrammatic device to prevent the azimuthal shift moving core residues from one side of the figure to the other, obscuring the effect of the skip residue. Red letters identify 'hydrophilic' surface residues, blue letters identify 'hydrophobic' core residues while purple letters identify residues lying between the core and the surface.

### 8.5.3 $\beta$ -MyHC A1663P and Azimuthal Charge Disruption

Only one analysis has proposed a model for thick filament aggregation which suggests a mechanism whereby the apparently private WAMPD  $\beta$ -MyHC A1663P substitution could cause azimuthal charge-phase distortion. If correct, such charge-phase distortion could, theoretically, disrupt thick filament assembly.



Atkinson and Stewart (1992) considered the possibility that the driving force for myosin assembly was a favourable entropy change as myosin II moved from solvated state to assembled state. If this was so, the complementation of charged zones would not make a major contribution to the free energy difference between molecules in solution and aggregates. The essential role of the charged zones were to block 'inappropriate' and 'energetically less favourable' assembly geometries. 'Appropriate' staggers had a lower free energy than any other arrangements. Atkinson and Stewart (p. 11) developed a hypothesis linking small MyHC sequence changes to thick filament disruption:

There are probably a large number of ways in which LMM molecules can interact that have nearly the same energy. In other words, there are probably a number of potential polymer states that differ only very slightly in their free energy from the "normal", polymer state usually observed. Such "cryptic" states would not be observed if their critical concentration for polymerisation was only marginally higher than that of the "normal", because the concentration of free monomers would never reach that required for their formation. Even small changes to the protein (by mutation for example) might raise the critical concentration for the "normal" form above that of the "cryptic", so that formation of the latter becomes possible.

According to this hypothesis, an azimuthal charge-phase distortion caused by the  $\beta$ -MyHC A1663P substitution might raise the critical concentration of the dimer required for normal  $\beta$ -MyHC thick filament formation, *above* that required for the cryptic aggregation of mutated dimers. The cryptic aggregation of  $\beta$ -MyHC coiled-coils would underlie Laing myopathy. In affected muscles, the  $\beta$ -MyHC A1663P substitution would have made misalignment of  $\beta$ -MyHC monomers energetically more favourable than the correct alignment.

#### 8.5.4 $\beta$ -MyHC A1663P and the Assembly Conserved Domain

Thick filament formation is critically dependent upon a small, highly conserved, negatively charged C-terminal residue sequence (Sohn et al., 1997). This region has been termed the ACD for assembly competence domain, and it runs from residue 1871 to residue 1899 in the human  $\beta$ -MyHC. Cohen and Parry (1998, p. 324) identified an extended ACD running from 1860 to residue 1922 and argued that, "the ACD would act as an additional factor to insure the selection of the correct molecular stagger for assembly". Were the  $\beta$ -MyHC A1663P substitution to interfere with the ACD it would have to have a long-range effect, spanning a minimum of 227 residues. Of interest, was the Cohen and Parry observation (p. 324), "that the ACD is not sufficient for assembly if the rod fragment is below a certain length". The shortest human fast IId skeletal MyHC rod in which they observed ACD functioning normally, was a fragment running from residue 1661-1902. The  $\beta$ -MyHC A1663P substitution falls within this region and so it may, or it may not, be close enough to the ACD to influence its function. Similar *in vitro* experiments with A1663P  $\beta$ -MyHC would be informative. Since neither Sohn or Cohen and Parry investigated the ACD effect in heterozygous combination, this would be a useful extension to their work, and of direct importance to dominant-negative diseases like Laing myopathy.

#### 8.6 CONCLUSION

An analysis of sarcomeric gene mutations, particularly mutations of the MYH genes, revealed that mutations in these genes can have skeletal muscle phenotype alone, even when a gene is also expressed in the heart. CMH1 mutations are variably expressed, leaving some patients with skeletal muscle

myopathy but no cardiac phenotype. This analysis also found that mutations to the *Drosophila* MYH gene, which altered conserved residues in the LMM rod region, only developed a phenotype in one type of muscle, the indirect flight muscles. In a general sense, the analysis of sarcomeric gene mutations pointed to changes in contractility and to the interaction of the  $\beta$ -MyHC rod region with other macromolecules as promising areas of future research into Laing myopathy.

The analysis of the effect of the  $\beta$ -MyHC A1663P substitution at various levels of protein structure also strengthened the candidature of *MYH7* G5073C as the Laing myopathy mutation. It demonstrated the extreme rarity of this type of substitution, it showed that the  $\beta$ -MyHC A1663P substitution did have a detrimental effect on coiled-coil formation, and it identified several ways in which  $\beta$ -MyHC A1663P *could* disrupt myofibrillogenesis. However, predicting the precise effect the  $\beta$ -MyHC A1663P substitution might have on myofibrillogenesis, is not possible from this theoretical analysis.

Future research on myofibrillogenesis involving both the A1663 and the P1663  $\beta$ -MyHC monomers would be worthwhile. Transfection of cultured-MYH mouse myoblasts with different CMH1 mutants, followed by the stable expression of  $\beta$ -MyHC, is already being employed by a group investigating myofibrillogenesis in CMH1 patients (personal communication, Dr. Michelle Peckham, School of Biological Sciences, University of Leeds, UK). Such a system would be very suitable for the investigation of myofibrillogenesis in Laing myopathy patients.

## **CHAPTER 9**

## **DISCUSSION**

## 9.0 DISCUSSION

The rationale of this thesis was, 'Knowledge of the structure and function of the human genome will improve the nosology, diagnosis, prognosis, prevention, treatment and future research into the molecular pathology of MD' (p. 3). This rationale has been supported by the work reported here:

The isolation of genomic clones for the slow skeletal muscle troponin gene (*TNNI1*), the human skeletal muscle alpha actin gene (*ACTA1*) and the  $\beta$ -tropomyosin gene (*TPM2*) for use in the *in situ* cytogenetic mapping of the *TNNI1*, *ACTA1* and *TPM2* skeletal muscle contractile protein genes.

The first localisation of an AD-distal myopathy, Laing myopathy, to the MPD1 candidate region.

The first demonstration of AD-distal myopathy genetic heterogeneity with the exclusion of the MPD1 candidate region for Silburn myopathy.

Confirmation of the genetic heterogeneity of AD-distal myopathy by the exclusion of the MPD1 candidate region for Felice myopathy.

The exclusion of the French Canadian 14q AD-OPMD candidate region for an AD putative-OPMD segregating in a Western Australian family.

The discovery of neutral polymorphisms in *MYH7*.

The discovery of a possible disease-causing mutation, *MYH7* cDNA C5073/gDNA C23628, segregating with Laing myopathy.

The publications emanating from this work (Eyre et al., 1993; Akkari et al., 1993; Hunt et al., 1994; Laing et al., 1995b; Felice et al., 1999), show that this project was a collaborative project. It was also part of a much larger collaborative project, the HGP, to which all of the above publications contributed.

An 'evolving concepts' approach has been adopted in this Chapter, to reflect the dynamic state of MD molecular pathology.

## 9.1 EVOLVING CONCEPTS IN MUSCULAR DYSTROPHY

During the rapid development of molecular genetics throughout this thesis, one discipline has been given substantial impetus; molecular pathology. "Molecular pathology seeks to explain why a given genetic change should result in a particular clinical phenotype." (Strachan and Read, 1996, p. 401). The definition of sickle cell anaemia as a molecular disease (Pauling et al., 1949), launched molecular pathology. Pauling et al. (1949) used aspects of the the clinical phenotype to determine the biochemical basis of sickle cell anaemia. It was the first example of functional cloning. Collins (1992, p. 3) makes the point that very few human genes are amenable to functional cloning:

While these strategies, referred to here as 'functional cloning', have been remarkably successful, they are limited to disorders for which biological information about the basic genetic defect exists. Unfortunately, for the vast majority of single gene disorders listed in McKusick's catalog, *Mendelian Inheritance In Man* (1991), no such functional detail exists.

Molecular genetics' contribution to molecular pathology would have been considerably diminished had it remained dependent on functional cloning. It did not. The position-independent candidate approach, positional cloning and the positional candidate approach (reviewed pp. 56-64), have greatly increased the contribution of molecular genetics to molecular pathology. The founding pathological events for the majority of human diseases will be identified within five years (Collins et al., 1998, p. 682). Molecular genetics has become the foundation of the molecular pathology of inherited diseases.

### **9.1.1 Molecular Pathology has Refined MD Nosology**

Since, the identification of the mutated dystrophin gene as the disease gene causing DMD and BMD, a further 33 genes causing various forms of MD have been localised. Nineteen of these genes have been cloned (Tables 9.1 & 9.2, pp. 278 & 279), and the investigation of their molecular pathology commenced.

Molecular pathology is a very incomplete discipline, but it has already made a significant change to one part of clinical genetics; nosology. In slightly more than a decade, MD nosology has moved from clinical presentation and pathology to a base which also includes molecular genetics. Evidence of this shift can be discerned from a comparison of Table 1.1 (p. 9) with Tables 9.1 and 9.2. Although Tables 9.1 and 9.2 are not intended to be complete, they do show that many new subdivisions within the MDs have appeared since this thesis started. Refinement has occurred with the genetic localisation of disease genes and with the identification of disease-causing mutations. The criteria used to define MD has expanded, improving both diagnostic accuracy and precision.

Mastaglia and Laing (1996, pp. 263-265) considered that the application of molecular genetic methods permitted the classification of the MDs into: dystrophinopathies, LGMDs, sarcoglycanopathies, congenital MDs (CMD), facioscapulo-humeral MD (FSHD), OPMDs, Emery-Dreifuss MD and the distal myopathies. Tables 9.1 and 9.2 include these MD groups.

The developing nosology of the LGMDs illustrates the impact of molecular genetics on MD nosology. By 1993, disease gene location could only confirm the diagnosis of one form of AD-LGMD and one form of AR-LGMD. Clinical phenotype and mode of inheritance remained the only way to distinguish between the other form of AD-LGMD and the three other forms of AR-LGMD identified at that time (Table 1.1, p.9).



Table 9.1

**Molecular Genetics and Autosomal Dominant Muscular Dystrophy Nosology**

Dystrophy type	Identification	Cytogenetic Location		Gene	
<b><u>Autosomal Dominant</u></b>					
Myotonic dystrophy	Bell, 1947				
DM1	Bell, 1947	19q13.2-q13.3	Harley et al., 1991	Protein kinase ( <i>DMPK</i> )	Brook et al., 1992
DM2 (PROMM)	Day et al., 1999	3q	Ranum et al., 1998		
Facioscapulohumeral MD	Tyler & Wintrobe, 1950				
FSHD		4q35	Wijmenga et al., 1992	non-transcribed p13E-11 del.	Lemmers et al., 1998
Limb-girdle MD					
LGMD1A	Gilchrist et al., 1988	5q31-q33	Speer et al., 1992	Myotilin	Hauser et al., 2000
LGMD1B	van der Kooi et al., 1997	1q11-q21	van der Kooi et al., 1997		
FDC-CDM	Messina et al., 1997	6q23	Messina et al., 1997		
LGMD1C	Minetti et al., 1998	3p25	Minetti et al., 1998	Caveolin-3 ( <i>CAV3</i> )	Minetti et al., 1998
LGMD1?	Speer et al., 1999	5q	Speer et al., 1999		
Bethlem myopathy	Bethlem et al., 1976				
COL6A1	Jobsis et al., 1996a	21q22.3	Jobsis et al., 1996a	Collagen6A1 ( <i>COL6A1</i> )	Jobsis et al., 1996b
COL6A2	Jobsis et al., 1996a	21q22.3	Jobsis et al., 1996a	Collagen6A2 ( <i>COL6A2</i> )	Jobsis et al., 1996b
COL6A3	Mohire et al., 1988	2q37	Speer et al., 1996	Collagen6A3 ( <i>COL6A3</i> )	Pan et al., 1998
Distal myopathy					
Welander	Welander, 1951	2p13	Ahlberg et al., 1998 & 1999		
Finnish Tibial	Udd et al., 1991 & 1993	2q31	Haravuori et al., 1998a		
Markesbery	Markesbery et al., 1974	2q	Haravuori et al., 1998b		
Laing (MPD1)	Laing et al., 1995(b)	14q	Laing et al., 1995(b)		
Feit (VCPDM) (MPD2)	Feit et al., 1998	5q31	Feit et al., 1998		
Oculopharyngeal MD	Taylor, 1915				
	(cited in Brais et al., 1995)				
OPMD( <i>PABP2</i> )	Barbeau, 1965	14q11.2-q13	Brais et al., 1995 & 1997	Poly(A)Bind. Protein-2 ( <i>PABP2</i> )	Brais et al., 1998
	(cited in Brais et al., 1995)				
Oculopharangodistal MPD	Satoyoshi & Kinoshita (1977)				

**Table 9.2** *Molecular Genetics and X-linked Recessive and Autosomal Recessive Muscular Dystrophy Nosology*

Dystrophy type	Identification	Cytogenetic Location		Gene	
<b><u>X-linked Recessive</u></b>					
Dystrophinopathies	Koenig et al., 1987				
Duchenne	Meryon, 1852 (c. Partridge, 1993, p. xiii)	Xp21.2	Boyd & Buckle, 1986	Dystrophin ( <i>DMD</i> )	Koenig et al., 1987
Becker	Becker, 1955 (c. Ozawa, et al., 1998, p. 421)	Xp21.2	Koenig et al., 1987	Dystrophin ( <i>DMD</i> )	Koenig et al., 1987
Quadriceps myopathy	Thage, 1965	Xp21.2	Sunohara et al., 1990	Dystrophin ( <i>DMD</i> )	Koenig et al., 1987
Myalgia and cramps	Kuhn et al., 1979	Xp21.2	Gospe et al., 1989	Dystrophin ( <i>DMD</i> )	Koenig et al., 1987
Emery-Dreifuss MD (EDMD)	Emery & Dreifuss, 1966	Xq28	Consalez et al., 1991	Emerin ( <i>EMD</i> )	Bione et al., 1994
<b><u>Autosomal Recessive</u></b>					
Limb girdle MD					
LGMD2A	Beckmann et al., 1991	15q15.1-q21.1	Beckmann et al., 1991	Calpain-3 ( <i>CANP3</i> )	Richard et al., 1995
LGMD2B	Passos-Bueno et al., 1991a & b	2p13.3	Bashir et al, 1994	Dysferlin ( <i>DYSF</i> )	Liu et al., 1998
LGMD2C ( $\gamma$ -SG SCARMD)	Ben Othmane et al., 1995	13q12	Ben Othmane et al., 1995	$\gamma$ -sarcoglycan ( $\gamma$ -SG)	Noguchi et al., 1995
LGMD2D ( $\alpha$ -SG SCARMD)	Roberds et al., 1994	17q12-q21.33	Roberds et al., 1994	$\alpha$ -sarcoglycan (adhalin)	Roberds et al., 1994
LGMD2E ( $\beta$ -SG SCARMD)	Bonnemann et al., 1995	4q12	Bonnemann et al., 1995	$\beta$ -sarcoglycan ( $\beta$ -SG)	Lim et al., 1995
LGMD2F ( $\delta$ -SG SCARMD)	Passos-Bueno et al., 1991a & b	5q33-q34	Passos-Bueno et al., 1996	$\delta$ -sarcoglycan ( $\delta$ -SG)	Nigro et al., 1996
LGMD2G	Passos-Bueno et al., 1991a & b	17q11-q12	Moreira et al., 1997		
LGMD2H	Shokeir & Kobrinsky, 1976	9q31-q33	Weiler et al., 1996		
Congenital MD					
Merosin deficient	Tome et al., 1994	6q22-q24	Hillaire et al., 1994	$\alpha$ -2 Laminin ( <i>LAMA2</i> )	Helbling-Leclerc, 1995
Fukuyama	Fukuyama et al., 1981	9q31-q33	Toda et al., 1993		
Congenital MD ( <i>ITGA7</i> )	Hayashi et al., 1998	12q13	Wang et al., 1995	Integrin $\alpha 7$ ( <i>ITGA7</i> )	Hayashi, 1998
Walker-Warburg syndrome	Warburg, 1976				
Muscle, eye and brain disease (MEB)	Santavuori & Leisti 1980	1p32-p34	Cormand et al., 1999		
Distal myopathy					
Miyoshi (MM)	Miyoshi, et al., 1986	2p12-p14	Bejaoui et al., 1995	Dysferlin ( <i>DYSF</i> )	Liu et al., 1998
Nonaka	Nonaka, et al., 1981 & 1985	9p1-q1	Ikeuchi et al., 1997		
Oculopharyngeal MD					
OPMD ( <i>PABP2</i> )	Brais et al., 1998	14q11.2-q13	Brais et al., 1998	<i>PABP-2</i>	Brais et al., 1998
Hered. Inc. BodyMP (h-IBM)	Argov & Yarom (1984)	9p1-q1	Mitrani-Rosenbaum, 1996		

Subsequently, increased knowledge of LGMD disease gene locations and the identification of disease-causing mutations has made it possible to distinguish five forms of AD-LGMD (Table 9.1) and eight forms of AR-LGMD (Table 9.2). Linkage analysis has excluded the eight AR-LGMD genes in two Brazilian families segregating a form of AR-LGMD, indicating that there is at least one more AR-LGMD gene (Passos-Bueno et al., 1999).

At all levels, MD nosology now includes molecular pathology. Together, Tables 9.1 and 9.2 comprise an MD nosology based on clinical phenotype and founding molecular event.

## **9.2 EVOLVING CONCEPTS IN DISTAL MYOPATHY**

During the course of this thesis, the genetic location of eight forms of distal myopathy were identified, some of which may be allelic. The first genetic location of the distal myopathies was reported from work conducted in this thesis by Laing et al. (1995b). By 1999, the genetic locations of four of the distal myopathies listed in Table 4.1 (p. 94) had been identified. These were those first described by: Welander (1951) (WDM) - localised by Ahlberg et al. (1998 & 1999); Markesbery et al. (1974) - localised by Haravuori et al. (1998b); Nonaka et al. (1981 and 1985) - localised by Ikeuchi et al. (1997); and Miyoshi et al. (1986) (MM) - localised by Bejaoui et al. (1995).

In addition to these localisations, four more forms of distal myopathy were identified and localised: tibial muscular dystrophy, Udd et al. (1993) - localised by Haravuori et al. (1998a); Laing myopathy - described and localised simultaneously by Laing et al. (1995b); distal anterior compartment (tibial) myopathy (DMAT) - described and localised simultaneously by Illa et al.

(1998); distal myopathy with vocal cord and pharyngeal weakness (ADVCPMD) (MPD2) - described and localised simultaneously by Feit et al. (1998). In addition to these localisations, a distal myopathy disease gene, dysferlin (*DYSF*), has been cloned from within the MM and DMAT overlapping candidate regions (Liu et al. 1998; Bashir et al., 1998). Yet despite these successes, the molecular pathology of the distal myopathies is still poorly understood. Currently, old concepts are being reviewed and new concepts developed from molecular genetic discoveries. Only those of relevance to the distal myopathies investigated in this thesis are reviewed.

### **9.2.1 Molecular Genetics Confirms and Refines Traditional Distal Myopathy Nosology**

In 1994 Griggs and Markesbery published a revised nosology of the distal myopathies and this was followed by the publication of a similar classification from discussions at the 25th European Neuromuscular Centre (ENMC), 'International Workshop: Distal Myopathies', held in February 1994 (Somer, 1994). Subsequently, a refinement of distal myopathy nosology was proposed (Somer et al., 1998). It retained all the essential structure of Griggs and Markesbery (1994) and Somer (1994) nosologies, and continued to be built on traditional clinical and genetic grounds.

Table 9.3 (p. 282) shows the distal myopathy nosology developed by these authors. It was remarkable in its similarity to that of Barohn et al. (1991, p. 1369) (Table 4.1, p. 94). It had three changes to that suggested by Barohn: minor name changes to the four forms of distal myopathy; the exclusion of sporadic and familial reports of distal myopathy, including only forms displaying unequivocal AR or AD inheritance; and the inclusion of tibial muscular dystrophy (TMD) (Udd et al., 1993) in, 'Category II: Late Adult-Onset in Legs'.

**Table 9.3*****Nosology of the Hereditary Distal Myopathies (after Griggs & Markesbery, 1994; Somer et al., 1998)***

Clinical phenotype	I Late Adult-Onset in Hands. AD.	II Late Adult-Onset in Legs. AD.	III Early Adult-Onset in Posterior Leg Compartment. AR.	IV Early Adult-Onset in Anterior Leg Compartment. AR.
Identification:	Welander, 1951 Dahlgaard, 1960 (cited in Somer et al., 1998, p. 183)	Markesbery et al., 1974 Udd et al., 1993 (tibial muscular dystrophy-TMD)	Markesbery et al., 1977 Miyoshi et al., 1986	Nonaka et al., 1981 & 1985
Inheritance:	AD	AD	AR	AR
Age of onset:	Over age 40	Over age 40 (Markesbery) Over age 35 (Udd)	15-30 years	20-30 years
First symptoms:	Hand muscles	Distal leg muscles, anterior compartment	Distal leg muscles, posterior compartment	Distal leg muscles, anterior compartment
Progression:	Very slow to distal flexors and extensors	Markesbery: slow to upper extremities and proximal limb muscles. Udd: slow, does not progress to upper extremities or to proximal limb muscles	Markesbery: Medium to proximal leg muscles. Forearm muscles become mildly atrophic. Miyoshi: progression towards a generalised atrophy	Rapid to proximal leg muscles, hands, neck then trunk. Tendon reflexes diminished
Cardiomyopathy:	No	Markesbery: one patient Udd: none	No	No
Serum CK:	Normal or slightly elevated	Normal or slightly elevated	10 to 150X normal	Increased, usually < 10X normal, often 2-5X normal
Electromyography:	Myopathic pattern	Markesbery: not recorded Udd: myopathic pattern	Myopathic pattern	Myopathic pattern
Muscle biopsy:	Variable, some rimmed vacuoles	Rimmed vacuoles	Dystrophic changes, no vacuoles, some gastrocnemius weakness	Rimmed vacuoles

The 'Category II: Late Adult-Onset in Legs' classification had been previously called, 'Late Adult-Onset, Type 2' in the nosology of Barohn et al. (1991). Table 9.3 also shows that TMD and the Markesbery distal myopathy (1974) have significant differences in disease progression. In grouping of TMD and the Markesbery distal myopathy (1974) together in, 'Category II: Late Adult-Onset in Legs', Somer et al. (1998, p. 186) expressed some concern, "Whether TMD patients have the same disease as those described by Markesbery et al. (1974) cannot be settled at the moment."

It is evident from Table 9.3 that genetic location was not employed to classify the distal myopathies. Somer (1994, p. 4) did record the localisation of MPD1 in the Workshop Report of the, '25th ENMC International Workshop: Distal Myopathies':

Emery informed about a breakthrough achieved by Nigel Laing and collaborators in Australia in mapping one form of distal myopathy to chromosome 14. Patients with this myopathy have a phenotype resembling that described by Gowers already in 1902 showing onset in childhood and early adulthood, and distal muscle weakness predominantly in the upper extremities. Linkage was obtained by analysis of one large family. This clinical entity was previously considered somewhat uncertain, and has not been included in recent classifications. This now appears to be the first distal myopathy with established genomic location.

The fact that MPD1 was the only distal myopathy locus identified in 1994, is possibly why Griggs and Markesbery (1994) and Somer (1994) did not use genetic location of MPD1 in their classifications of distal myopathies. However, distal myopathy nosology based on traditional clinical and genetic grounds was continued by Somer et al. (1998), despite their recognition of the potential

usefulness for classification purposes of the four distal myopathy gene locations that were known at that time. These were: MPD1 to 14q (Laing et al., 1995b); MM to chromosome 2p12-p14 (Bejaoui et al., 1995); Nonaka myopathy to 9p1-q1 (Ikeuchi et al., 1997); TMD to 2q31-33 (Haravuori et al., 1998a).

Laing myopathy reported in this thesis and by Laing et al. (1995b), continued to be excluded, seemingly due to its rarity and because insufficient clinical detail was available for its status as a distal myopathy to be firmly established. It was both rare and equivocal. Somer et al. (1998, p. 195) comment, "Histopathological characterisation of Gowers' phenotype in the Australian family (Laing et al., 1995) is still unknown". In more general terms Somer et al. (p. 190) state that, "Laing et al. (1995) reported genetic linkage of a myopathy inherited in autosomal dominant fashion to chromosome 14 in an Australian family, where the phenotype is said to resemble that described by Gowers (1902). A more detailed clinical description is still lacking".

Lately, a more detailed clinical ascertainment of the Laing myopathy family has been conducted (personal communication, Dr. Beverly Phillips, ANRI, Perth, Western Australia, 1999). This review classifies Laing myopathy as 'early-onset'.

Two recent distal myopathy classifications were presented by Felice et al. (1999, p. 64) and Barohn and Amato et al., (1999, p. 47). They were the first attempts to develop a classification of the distal myopathies based on gene location. Predictably, they were very similar. They differed only in that the classification of Barohn and Amato, included myofibrillar (desmin) myopathy. Table 9.4 (p. 285) shows the Felice distal myopathy classification.



<b>Table 9.4    <i>Classification of Distal Myopathies based on Gene Location</i> (modified from Felice et al., 1999)</b>			
<b>Myopathy</b>	<b>AR/AD</b>	<b>Genetic Location</b>	<b>Authors (localisation)</b>
AR distal myopathy (Miyoshi myopathy) (MM)	AR	2p12-14	Bejaoui et al. (1995)
Distal anterior compartment (tibial) myopathy (DMAT)	AR	2p13	Illa et al. (1998)
Tibial MD (TMD) (Udd myopathy)	AD	2q31	Haravuori et al. (1998a)
Late-onset distal myopathy (Markesbery myopathy)	AD	2q31	Haravuori et al. (1998b)
Nonaka myopathy	AR	9p1-q1	Ikeuchi et al. (1997)
Hereditary inclusion body myopathy (h-IBM)	AR	9p1-q1	Mitrani-Rosenbaum et al. (1996)
AD distal myopathy (Laing myopathy) (MPD1)	AD	14q	Laing et al. (1995b)
Welander myopathy (WDM)	AD	?	

A comparison of Table 9.4 with Table 4.1 (p. 94) and Table 9.3 (p. 282), shows that molecular genetics has confirmed the traditional clinical and genetic nosology of the distal myopathies as proposed by Somer et al. (1998). Notwithstanding its cautious approach, this traditional distal myopathy classification did successfully discriminate MM, Nonaka and TMD forms of distal myopathy.

Compared to the Somer nosology (1998), the Felice nosology (1999) contained new categories and new groupings. Laing myopathy was included as a new category, DMAT was included and grouped with MM, and h-IBM was included and grouped with Nonaka myopathy. One advantage of developing nosology based on genetic location was its ability to diminish uncertainties based on clinical phenotype. In this context, the Felice nosology: confirmed a connection between distal myopathy and h-IBM as suggested by Askanas (1997, p. 422); it diminished the uncertainties expressed by Somer et al. (1998, p. 186) about grouping TMD and Markesbery distal myopathy together; it placed Laing myopathy in a class by itself; and it associated MM with DMAT. Distal myopathies with overlapping genetic locations may have the same starting point for their pathology, those with different genetic locations have different starting points for their pathology.

Table 9.4 was the first genetic classification of distal myopathies and in appreciation of this fact Felice (p. 64) described their genetic classification of the distal myopathies as 'tentative'. There was an expectation that their proposed distal myopathy classification would continue to change, as distal myopathy disease gene locations and disease-causing mutations were discovered. In fact, between its acceptance and publication, the classification proposed by Felice et al. (1999) had already undergone refinement. Ahlberg et al. (1998, p. S6) announced the genetic localisation of WDM at the 'IX International Congress on Neuromuscular Diseases', Adelaide, Australia. Identification of a common shared haplotype in seven Swedish families segregating WDM located this distal myopathy disease gene to a 0.2-2.4 cM region on chromosome 2p between the markers D2S303 and D2S2115. This is very close to the localisation of the Miyoshi myopathy and the dysferlin gene. Dysferlin has been excluded as the gene for WMD (B.Udd, personal communication, 2001).

### **9.2.2 Nonaka Distal Myopathy and a form of h-AR-IBM may have a Common Disease Gene**

Nonaka distal myopathy is an AR-RV form of distal myopathy, while the Laing, Silburn and Felice myopathies reported here were AD, non-vacuolar forms of distal myopathy. Nevertheless, the conclusions derived in this section, and the following section, are directly relevant to the Laing, Silburn and Felice myopathies.

Askanas and Engel (1993) introduced the term 'hereditary inclusion body myopathies' (h-IBMs) to describe hereditary progressive muscle diseases with muscle pathology strikingly similar to the pathology of 'classic' sporadic inclusion-body myositis (s-IBM) (Yunis & Samaha, 1971; Fardeau & Tome, 1998, p.252), except that: h-IBM muscle biopsy specimens lacked signs of inflammation (hence 'myopathy' rather than 'myositis'); usually occurred in childhood or young adulthood; and frequently displayed an unusual distribution of atrophy, sparing quadriceps muscles (Argov & Yarom, 1984). Currently, h-IBMs have been delineated as the h-AR-IBM and h-AD-IBM genetic myopathies (Fardeau & Tome, 1998, pp. 252-256), both of which have similar muscle pathology (Askanas & Engel, 1995; Griggs et al., 1995).

A form of h-AR-IBM, first reported by Argov and Yarom (1984) in 30 members of nine Iranian Jewish families, has been associated with the Nonaka form of distal myopathy (Nonaka et al., 1981). These two myopathies were found to be strikingly similar. Both displayed AR inheritance; both were quadriceps-sparing (although the quadriceps muscle was slightly affected in some Nonaka myopathy patients); both were RV myopathies with multiple or single RVs that were more common in type I muscle fibres (Argov et al., 1997); both showed marked variation in muscle fibre size; both had myogenic and

neurogenic EMGs; both were associated with lower limb distal weakness causing foot-drop and steppage-gait; and both progressed slowly to proximal muscles (Sunohara et al., 1989; Serratrice, 1998).

Muscle pathologies of these myopathies were also found to be strikingly similar, both possessing paired helical filaments and both possessing the same Alzheimer-characteristic protein accumulations within vacuolar muscle fibres:  $\beta$ -amyloid protein, two  $\beta$ -amyloid precursor protein epitopes, apolipoprotein E, hyperphosphorylated  $\tau$ ,  $\alpha_1$ -antichymotrypsin, and ubiquitin (Murakami et al., 1995; Askanas, 1997). As disease phenotype similarities between Nonaka myopathy and Iranian h-AR-IBM grew, Askanas and Engel (1995) proposed that the h-AR-IBM of Iranian Jews and Nonaka distal myopathy might be the same entity. Subsequently, Satoyoshi et al. (1998, p. 249) have indicated that the disease phenotypes of these two disease are so similar that misdiagnosis could occur:

In Japan, however, IBM is reported much less frequently . . . there were only 5 patients with IBM among 143 patients with inflammatory myopathies, whereas there were 18 patients with DMRV. One might raise the possibility that IBM was misdiagnosed as DMRV because RV formation is common to both diseases.

The localisation to chromosome 9p1-q1 of the h-AR-IBM disease gene in: Iranian-Jews (Mitrani-Rosenbaum et al., 1996), Afghani-Jews, Egyptian-Jews, Iraqi-Jews, non-Jews from India (Argov et al., 1997; Argov & Mitrani-Rosenbaum, 1998), and non-Jews from Mexico (Askanas, 1997; Askanas & Engel, 1998) followed by the the localisation of Nonaka myopathy to the same 9p1-q1 (Ikeuchi et al., 1997) region, led Argov and Mitrani-Rosenbaum (1998, p. 209) to conclude that:

A recent report suggests that a Japanese form of distal myopathy also maps to a region of chromosome 9p1-q1 (Ikeuchi et al., 1997). These findings suggest that in addition to the disease in Jewish communities related to Persia, a similar (allelic?) disease exists in non-Jews, most probably resulting from defects in the same gene.

Askanas (1997, p. 422) had already reached the same conclusion, "It is apparent that the Japanese 'distal myopathy with RVs' and the above examples of 'quadriceps-sparing h-IBM' are the same disorder or very closely related disorders." The eventual cloning of the disease gene responsible for either h-AR-IBM or Nonaka myopathy will allow this conclusion to be tested. Askanas (1997, p. 422) argues for the establishment of an h-IBM classification and nomenclature to include some 'distal myopathies'. Others argue that distal myopathy classification and nomenclature should include some 'h-IBMs' (Felice et al., 1999, p. 64).

### **9.2.3 Different Aetiologies Can Produce Similar Pathogenic Cascades**

Askanas and Engel (1995, pp. 490-494) suggested h-AR-IBM and h-AD-IBM were probably caused by different genetic defects because they had different genetic transmission. They must also have a different aetiology to s-IBM, which could be viral in origin (Askanas, 1997, p. 421). Yet in many clinical respects h-AR-IBM, h-AD-IBM and s-IBM are quite similar (Askanas & Engel, 1995). This led these authors to ask whether various causes of IBM could lead to the same IBM pathological process, or as Askanas (1997, p. 422) suggests, "are the result of the same final pathogenic cascade."

The AD inheritance of WDM and its localisation to 2p13 (Ahlberg et al., 1998) means that it has a different mode of inheritance and a different disease gene to the 9p1-q1 Nonaka distal myopathy (Ikeuchi et al., 1997) and to the 9p1-q1 form of h-AR-IBM (Mitrani-Rosenbaum et al., 1996). The fact that WDM pathogenesis is similar to a very slowly progressive form of h-IBM and the fact that h-AR-IBM, h-AD-IBM and s-IBM can be hard to distinguish (Askanas & Engel, 1995, pp. 490-494; Serratrice, 1998, pp. 95-97), supports the concept that different aetiologies can produce the same final pathogenic cascade.

At present, the IBM/distal myopathy pathogenic pathways are not well understood. As genetic loci for h-IBMs and distal myopathies are identified and disease genes cloned, their effect on the muscle fibre will be delineated and the pathogenic pathway of each disease will begin to emerge. If this concept is correct, it predicts that points of convergence in the molecular pathology of each disease will be revealed. Uncovering the factors that cause s-IBM may also start to uncover its pathogenic pathway. Further points of convergence between s-IBM, h-IBM and some of the distal myopathies may then be identified.

The implication for the Laing, Silburn and Felice myopathies is clear. Pathognomic similarities may not reflect events precipitating disease onset, but rather reflect a final common pathogenic cascade. Clinical ascertainment cannot resolve this issue, highlighting the danger of combining families of similar clinical phenotype for the purposes of linkage analysis. Exclusion analysis was able to establish that Laing myopathy had a different disease gene location to both Silburn and Felice myopathies and to the putative-OPMD segregating in a Western Australian family. Unequivocally, therefore, Laing myopathy has a different starting point for its pathogenic pathway to these other forms of MD.

Its pathognomic similarities to these MDs must be due to convergence of pathological pathways. It does not belong to an allelic series that includes any of these other MDs, thus its phenotypic similarities to them cannot be explained by allelic heterogeneity.

Molecular genetic investigations into the distal myopathies has started to reveal possible cases of allelism (Table 9.4, p. 285). The possibility of an allelic series containing any two or all three of the Silburn myopathy, Felice myopathy and the Western Australian putative-OPMD cannot be discounted. Only the discovery of disease-causing mutation(s) can resolve this question.

#### **9.2.4      Markesbery Distal Myopathy and TMD may have the Same Disease Gene**

TMD (Udd et al., 1993) and Markesbery distal myopathy (Markesbery et al., 1974) are clinically similar AD diseases of late adult onset. Both show onset in the tibial anterior muscles, AD-TMD between the fourth to seventh decade and Markesbery between the ages of 43-51. In TMD, muscle weakness remains confined to tibial anterior muscles in most patients while in Markesbery distal myopathy muscle weakness manifests distally in tibial anterior muscles, but then progresses to the proximal muscles.

A comparison of morphologic changes seen in muscle biopsy reveals: at the early stage of dystrophic change in both TMD and Markesbery, RVs in affected muscle fibres varied from very few to numerous; severe end-stage dystrophy, manifest in Markesbery by variability in fibre size, necrosis, central nuclei, an increase in connective tissue and fatty infiltration (Serratrice, 1998, p. 97), and manifest in TMD by variability in fibre size, fibre splitting, and internalised nuclei (Udd et al., 1998, p. 235).



AD-TMD and AD-Markesbery distal myopathy differ in their rate of clinical progression. In AD-TMD, weakness of lower leg muscles is slowly progressive, eventually causing a moderate foot drop. Walking ability is preserved throughout the patient's lifetime because muscle weakness remains confined to the tibial anterior muscles (Udd et al., 1998, p. 233). Muscle weakness is not so confined in AD-Markesbery distal myopathy and progresses to the intrinsic hand muscles and the wrist extensors then, finally, to proximal muscles (Markesbery et al., 1974, p. 127 & 129).

Haravuori et al. (1998b) tested the only AD-Markesbery distal myopathy family (LODM) to verify or exclude the AD-TMD 2q31 locus. A maximum two-point lod score of 2.57 ( $\theta = 0.00$ ) was obtained for the marker D2S2310 and a multipoint lod score analysis placed the AD-Markesbery distal myopathy locus between the markers D2S385 and D2S389 with a maximum lod score of 3.62. This showed that the AD-TMD and the AD-Markesbery distal myopathy candidate regions overlapped, leading Haravuori et al. (1998b, p. A186) to the conclusion that AD-TMD and the AD-Markesbery distal myopathy, "may be caused by the same defective gene since these diseases appear to be allelic."

There is little scope to refine the 1cM AD-TMD candidate region, so the cloning of the disease gene(s) will determine if AD-TMD and AD-Markesbery distal myopathy are allelic. This will most likely be by the positional candidate approach since, as discussed in Chapter 8, within this candidate region there is a strong candidate gene, the titin gene.

### **9.2.5 The Same Mutation can produce different Pathogenesis**

Liu et al. (1998) found that the *DYSF* A4265G mutation, causing an I1299V substitution, gave rise to MM and DMAT in homozygous G4265 individuals

belonging to the same Italian family. *DYSF* has a 5' UTR of 368 bases. Moreover, in two Spanish families the deletion of G at 5966, resulted in MM in one of these families but DMAT in the other. A single mutation causing two phenotypically different MDs cannot be described as demonstrating allelic heterogeneity (Gelehrter & Collins, 1990, p. 299) or as belonging to an allelic series (McKusick, 1973, p. 446). The term 'mutation heterogeneity' is proposed here to describe these occurrences.

The discovery of the role of the *DYSF* gene in MM by Liu et al. (1998), made it possible to test the hypothesis that the same mutation caused both MM and LGMD2B co-segregating in a large Canadian aboriginal kindred (Weiler et al., 1996). Weiler et al. (1999) found that all LGMD2B and MM patients in this kindred were homozygous for *DYSF* G2745, a C->G mutation causing a P791R dysferlin substitution. P791R resulted in similar reductions of *DYSF* expression in the two types of patients, strengthening the conclusion that the homozygous *DYSF* G2745 genotype could cause two forms of MD in the same kindred. The concept of mutation heterogeneity had been confirmed.

The *DYSF* investigations of Liu et al. (1998) and Weiler et al. (1999), have made it clear that disease classification retains some refractory aspects at the level of disease-causing mutation. Evidence of mutation heterogeneity confounded the connection between mutation and pathogenesis and immediately led to speculation that modifier gene(s) or additional factors must account for the differences in the clinical phenotype (Weiler et al., 1999, p. 871 & 875). Bashir et al. (1998, p. 40) suggested that, "Some muscular dystrophy genes may allow for variability in the patterning observed, or the involvement of another gene may modify the phenotype."

In support of MD genes giving rise to 'variability in the patterning', Anderson et al. (1999, p. 855) have offered a hypothesis about how dysferlin might produce different patterns of affected muscles. Immunolabelling with monoclonal antibody to the DYSF protein localised this protein to the muscle fibre membrane at both the light and electron microscopic level:

Analysis of human fetal tissue showed that dysferlin was expressed at the earliest stages of development examined, at Carnegie stage 15 or 16 (embryonic age 5-6 weeks). Dysferlin is present, therefore, at a time when the limbs start to show regional differentiation. Lack of dysferlin at this critical time may contribute to the pattern of muscle involvement that develops later, with the onset of a muscular dystrophy primarily affecting proximal or distal muscles.

This is still an untested hypothesis. Other possibilities exist. Dysferlin is located in the membrane and both Anderson et al. (1999) and Weiler et al. (1999) found low levels of dysferlin in MM/LGMD2B patients, reminiscent of dystrophin levels in many Duchenne patients. The dysferlinopathies may not be a separate type of MD as suggested by Liu et al. (1998), but instead be a 'loss of muscle cell membrane integrity' type of MD, similar to the MDs with defects in dystrophin and dystrophin-associated proteins.

This discussion leads to the conclusion that any two, or all of Silburn myopathy, Felice myopathy and the putative-OPMD segregating in a Western Australian family, could be caused by the same mutation. Only genetic localisation of these distal myopathies, started in this thesis, and the subsequent identification of disease-causing mutation(s), can confirm or exclude these possibilities.

### **9.2.6 Distal Myopathy May be Caused by Mutations in Genes Encoding Contractile Proteins**

Table 8.1 (p. 225) shows that all of the hypertrophic cardiomyopathy (CMH) genes encode contractile proteins. Could the same be true for the distal myopathies? Dysferlin has been identified as a distal myopathy disease gene (Liu et al., 1998; Bashir et al. 1998), Udd et al. (1998) have identified titin as a 'superior candidate' for TMD, Martinsson (unpublished) has recently found a mutation in a chromosome 17 MYH gene segregating with a hereditary inclusion-body myopathy (IBM3), and this thesis argues that *MYH7*, which lies within the MPD1 linkage region, is a strong candidate gene for Laing myopathy. The possibility that many of the distal myopathies are a new class of contractile protein disease cannot be excluded.

### **9.2.7 Childhood Onset Distal Myopathy**

One further distal myopathy concept emerged during this review. Surprisingly, it did not come from molecular genetic analysis, but from an examination of distal myopathies systematically excluded from nosologies during the past 35 years.

The refinement of distal myopathy classification by Griggs and Markesbery (1994), Somer (1994) and Somer et al. (1998) were based on traditional clinical and genetic grounds. They excluded rare or genetically undetermined distal myopathies, and those where distal weakness was just one prominent feature of a disease. Somer et al. (1998 p. 190) identified five other, "myopathies with onset in distal muscles" that did not fit their classification, "Nevertheless, their clinical picture, histopathology and genomic location provide interesting similarities to the distal myopathies described above, and are reviewed briefly here." Table 9.5 (p. 296) lists these 'excluded' distal myopathies.

**Table 9.5     *Distal Myopathies Excluded from the Distal Myopathy Nosology (after Somer et al., 1998).***

<b>Distal Myopathy</b>	<b>Ascertainment</b>	<b>Authors</b>
I. Early adult onset.	Facial, sternomastoid atrophy. AD.	Gowers (1902) Laing et al. (1995b)
II. Early adult onset.	Respiratory, bulbar, proximal muscles all affected. Desmin storage, autophagocytosis. AD.	Milhorat and Wolff, (1943); Horowitz & Schmalbruch (1994)
III. Late onset.	Thenar, hand flexors, proximal muscles all affected. Intermediate filaments, sarcoplasmic bodies. AD.	Edström et al. (1980)
IV. Infantile onset.	Peroneal, facial and trunk muscles all affected. AD.	Scoppetta et al. (1995)
V. Adult onset.	Vacuolar myopathy sparing quadriceps. Persian Jews. AR.	Argov & Yarom (1984); Sadeh et al. (1993)
Categories I to V are presented as suggested by Somer et al. (1998). They are not related to Categories I to IV in Table 9.3 (p. 282).		

Somer et al. (1998, p. 190-192) justified the exclusion of the myopathies in Table 9.5 because they: had only been detected in a single large pedigree (Milhorat and Wolff, 1943; Laing et al., 1995b); had distal muscle weakness as a predominant clinical finding only for a certain stage of their disease (Edström et al., 1980); or were 'dated' studies, lacking sufficient data for classification purposes (Gowers, 1902).

Somer et al. (1998) also excluded the report by Scoppetta et al. (1995) of AD infantile-onset in four Italian families without explanation. It did not qualify for exclusion on any of the above grounds. The Somer nosology did not include an infant-onset/childhood-onset category.

Laing myopathy was included in the distal myopathy classification of Felice et al. (1999, p. 64), on the strength of its genetic localisation, not its age of onset. Nosological inclusion or exclusion, therefore, clearly depends upon the classification criteria used. The Laing and Scoppetta distal myopathy families share clinical features with four other distal myopathy families reported by Voit et al. (1997), Bautista et al. (1978), Does de Willebois et al. (1968) and Magee & DeJong (1965). The Scoppetta and Voit distal myopathy families were discussed in Section 7.3.2.10 (pp. 211-212). Recently, the Voit family and one of the Scoppetta families have been reviewed (Voit et al., 2001).

All of these families displayed: AD inheritance; onset in infancy or childhood; initial weakness in anterior calf muscles followed by weakness in finger extensors; a lack of RVs. All of these families except the Voit family also displayed: EMG indications of a myopathic origin; an increase in connective and adipose tissue and central nuclei in affected muscle; a variation in muscle fibre size; normal or moderately elevated CK levels (maximum was 3X); and benign to mild progression.

Voit et al. (1997, p. 251) claimed that the distal myopathy in their family was distinguished by its very early age of onset, "As far as the existing literature is concerned no such family with dominant transmission and infantile onset within the first year of life has been reported to date." This claim may not be correct. Magee and DeJong (1965, p. 390) report that in their distal myopathy family, with 18 affected members, onset was evident:

no later than age 2 . . . it is possible that the weakness is present from birth and that the disorder should properly be termed "hereditary distal myopathy with onset in infancy." Because evaluation of muscle function was impossible before affected children began to walk, the term "hereditary distal myopathy with onset in infancy," although less precise, seems to be the most appropriate name for the disorder we have described.

The reports of distal myopathy by Bautista et al. (1978), Does de Willebois et al. (1968) and Magee and DeJong (1965) are seldom mentioned in the literature and, for the purposes of distal myopathy nosology, they had faded from view. A re-examination of the literature has identified traditional clinical and genetic evidence to indicate that a form, or forms, of distal myopathy with infant/childhood onset might exist. One of these distal myopathies was studied in this thesis, Laing myopathy.

### **9.3 EVOLVING CONCEPTS IN OPMD**

In contrast to distal myopathy nosology, OPMD classification has remained stable. The OPMD diagnostic criteria of AD segregation, late-onset dysphagia and progressive ptosis, generally after age 45, and 8.5 nm diameter skeletal muscle INI (Tome & Fardeau, 1980), have become firmly established as prerequisite for OPMD diagnosis. This is of relevance to the putative-OPMD reported in Chapter 6, which does not meet these strict diagnostic criteria.

#### **9.3.1 Locus Heterogeneity Unlikely for OPMD**

Since Brais et al. (1995) localised the OPMD segregating in French-Canadian



families, OPMD locus heterogeneity has not been reported. Bouchard et al., (1998) have listed 17 populations around the world that meet OPMD diagnostic criteria, and include: French-Canadians, Russian and Bukhara Jews, Armenians, English, Brazilians and Australians of German descent. An additional five ethnic groups have had their OPMD localised to 14q, overlapping the 14q11.2-q13 French Canadian OPMD candidate region (Brais et al., 1995). These were: one multigenerational family of Scottish/English ancestry (Stajich et al., 1996); one large pedigree of northern German ancestry (Friesland) (Porschke et al., 1997; Kress et al., 1998); one large pedigree of Bavarian ancestry (Kress et al., 1998); an Australian kindred of German descent (Teh et al., 1997); and three American families of Hispanic descent (Grewal et al., 1998). All of these ethnic groups had different risk haplotypes to each other and to the French-Canadian OPMD population, suggesting 6 independent mutations for this disorder. Four multi-generational American families of French-Canadian ancestry in the investigation by Stajich et al. (1996), had the same risk haplotype as the French-Canadian population studied by Brais et al. (1995).

Linkage analysis with seven chromosome 14q polymorphic markers in 11 large French Canadian OPMD families established D14S283 as the new centromeric flanking marker, reducing the previously reported 5cM candidate interval (Brais et al., 1995) to 2cM (Brais et al., 1997). Multipoint analysis suggested that the OPMD disease gene lay within a 1.5cM region around D14S990. A candidate for OPMD, poly(A) binding protein 2 gene (*PABP2*) had been mapped by *in situ* hybridization to chromosome 14q11 (Nemeth et al., 1995). *PABP2* was a good candidate because its mRNA was highly expressed in skeletal muscle and *PABP2* protein was exclusively localised to the nucleus (Krause et al., 1994) where it acted as one factor in mRNA polyadenylation (Bienroth et al., 1993). *PABP2* exon 1 contains the start codon and a (GCG)<sub>6</sub> repeat coding

for six alanine residues. Brais et al. (1998, p. 164) found that (GCG)<sub>6</sub> allele was present in 98.0% of French Canadian controls, while the (GCG)<sub>7</sub> allele was found in the remaining 2.0% of this control population (n=86). Screening 144 OPMD families from 15 different countries for a 242 bp *PABP2* exon 1 fragment, revealed a PCR product in all affected individuals that exceeded those found in controls by 6 to 21 bp (Brais et al., 1998). The disease gene causing AD-OPMD had been identified.

The length of time taken to swallow 80 ml of ice cold water was found to correlate with disease severity, the size of GCG-repeat expansion, and with mode of inheritance in the French Canadian OPMD population (Brais et al. 1998, p. 165). In the (GCG)<sub>9</sub>+(GCG)<sub>6</sub> genotype the (GCG)<sub>9</sub> allele was inherited as a dominant disease allele, causing a progressive increase in swallowing difficulty from about age 40 compared to control genotype (GCG)<sub>6</sub>+(GCG)<sub>6</sub>. The (GCG)<sub>7</sub>+(GCG)<sub>7</sub> genotype had a late-onset, biopsy-proven, form of OPMD which, from pedigree analysis, had AR inheritance. This accounts for the inclusion of a new AR form of OPMD in Table 9.2 (p. 279).

The identification of *PABP2* as the OPMD disease gene and the proposed accumulation of INI as the central event in OPMD molecular pathology, make OPMD locus heterogeneity unlikely. The fact that OPMD can also be a rather generalised dystrophy has not, and will not, influence the stringency of diagnostic criteria. Reports continue to appear of OPMD with diffuse phenotypes including: signs of distal symmetrical neuropathy (Porschke et al., 1997); paracrystalline mitochondrial inclusions (Wong et al., 1996); and genetic anticipation (Blumen et al., 1997). Wong et al. (1996) suggest that abnormal mitochondria are a non-specific epiphenomena in OPMD while Porschke et al.

(1997) continue to agree with Hardiman et al. (1993) that the association between neurogenic lesion and OPMD remains unclear. Notwithstanding this debate, all of these reports of OPMD with diffuse phenotype also displayed AD segregation, late-onset dysphagia, progressive ptosis after age 45, and 8.5 nm diameter skeletal muscle INI.

From this it follows that the exclusion of the 14q OPMD locus for the WAOPMD family (Figure 6.5, p. 172) cannot be taken as evidence for OPMD locus heterogeneity until it is shown that the affected members of this family possess skeletal muscle INI histological markers. Even before *PABP2* cloning, Meola et al. (1997, p. S53) had presented a strident argument in favour of INI as diagnostic for OPMD, "The identification of nuclear inclusions is mandatory in order to confirm the diagnosis prior to linkage analysis." The cloning of the *PABP2* as the OPMD disease gene and its emerging molecular pathology, have made an exhaustive search for INI in affected WAOPMD skeletal muscle, a top priority.

#### **9.4 CONCLUDING COMMENTS**

McKusick (1990, p. v) has defined clinical genetics as:

The science and practice ('art') of the diagnosis, prevention, and treatment of genetic disease. Some years ago I heard the British biostatistician Bradford Hill suggest that the practice of medicine consists of seeking the answers to three questions: What is wrong? (The answer is diagnosis). What is going to happen? (The answer is prognosis). What can be done about it (The answer is treatment). I would add a fourth question, Why did it happen? On the answer to this question depend prevention and the advancement of science.

Human molecular genetics has provided many answers to this last question. During the past decade it has forged numerous links between genotype and clinical phenotype. In addition, human molecular genetics has largely overcome the traditional foes of clinical genetics; long human generation time and small family size. Clinical genetics has become clinical *molecular* genetics. The expectation is that clinical molecular genetics will continue to change the medical landscape by constantly refining clinical practice.

AD-MDs were the subject of this thesis because AD-mutations have the greatest effect on disease onset, the greatest effect on disease progression, and they are the easiest to clone. This means that research into AD-diseases has the best prospect for success; to delay the onset of diseases, to slow progression and, ultimately, to neutralise the pathological effect of disease-causing mutations. A 'candidate mutation' was identified for Laing myopathy and future research directions described in Section 8.17 (p. 231). Molecular genetic analysis indicated that investigations into the the contractility of myosin II containing both A1663 and P1663  $\beta$ -MyHC monomers, and the interaction of the  $\beta$ -MyHC LMM rod region with other macromolecules, could both be productive areas of research. Productive not only for understanding Laing myopathy but also for understanding the pathological links between different forms of distal myopathy.

The very similar appearance of Laing myopathy and TMD in MRI and CT scanning (personal communication, N.G. Laing, 2000) highlight the possibility of a common final pathway for these forms of distal myopathy. In Section 8.1.5 (p. 229) the failure of MyBP-C to bind to  $\beta$ -MyHC in Laing myopathy and to titin in TMD was identified as *possible* connection between these phenotypically similar diseases. Although it is impossible to predict what

molecular connections future research might find between the distal myopathies, it will become successively easier to look in the right places, and it is almost certain that patho-physiology connections will be found. Research on one form of distal myopathy will increasingly point the way for others. As molecular pathologies of the distal myopathies emerge and converge, it will be possible to influence the progression of this type of MD. In small ways to start with, but with gradual improvement as research progresses.

Ultimately, Nagel (1999, p. 3) has argued, that as molecular genetics progresses we will no longer treating diseases but diseased individuals, "The more that we know about the genetics of each individual, the more it will become apparent that each pathological event is unique to that individual". For Nagel, the conclusion became inescapable: the traditional concepts of disease had outlived their usefulness. Each mutation acts in the 'genetic terrain' in which it finds itself. By adding genetic individuality to the disease process, molecular genetics had opened another conceptual framework for pathology. The genetic individualisation of disease; there are no diseases, only diseased individuals. Molecular genetics was leading to customised medicine.

In practice, however, a range of logistical factors will combine to limit the development of customised medicine for diseased individuals, including: statistical impasses; ascertainment difficulties; non-genetic environmental factors; and the frequency and effect of mutations. In addition, contributory disease genes will be harder and harder to detect, therefore more and more expensive to identify, as their contribution to disease diminishes. So how can research into rare diseases like Laing myopathy be encouraged? How can the most expensive option of all, 'the genetic individualisation of disease' be pursued so that medicine really starts to become customised? We need not ask. The tried and tested traditional method is at hand. The doctoral candidate!

## **REFERENCES**

## REFERENCES

- Abrams, W.R., Ma, R.I., Kucich, U., Bashir, M.M., Decker, S., Tsipouras, P., McPherson, J.D., Wasmuth, J.J., & Rosenbloom, J. (1995) Molecular cloning of the microfibrillar protein MFAP3 and assignment of the gene to human chromosome 5q32-q33.2. *Genomics*, 26(1), 47-54.
- Adams, M.D., Kelley, J.M., Gocayne, J.D., Dubnick, M., Polymeropoulos, M.H., Xiao, H., Merrill, C.R., Wu, A., Olde, B., Moreno, R.F., Kerlavage, A.R., McCombie, W.R., & Venter, J.C. (1991) Complementary DNA sequencing: expressed sequence tags and human genome project. *Science*, 252, 1651-1656.
- Adams, M.D., Kerlavage, A.R., Fields, C., & Venter, J.C. (1993a) 3,400 expressed sequence tags identify diversity of transcripts in human brain. *Nat Genet*, 4, 256-267.
- Adams, M.D., Soares, M.D., Kerlavage, A.R., Fields, C., & Venter, J.C. (1993b) Rapid cDNA sequencing (expressed sequence tags) from a directionally cloned human infant brain cDNA library. *Nat Genet*, 4, 373-380.
- Adams, M.D., Kerlavage, A.R., Fleischmann, R.D., Fuldner, R.A., Bult, C.J., Lee, N.H., Kirkness, E.F., Weinstock, K.G., Gocayne, J.D., White, O., Sutton, G., Blake, J.A., Brandon, R.C., Chiu, M-W., Clayton, R.A., Cline, R.T., Cotton, M.D., Earle-Hughes, J., Fine, L.D., Fitzgerald, L.M., Fitzhugh, W.M., Fritchman, J.L., Geoghagen, N.S.M., Glodek, A., Gnehm, C.L., Hanna, M.C., Hedblom, E., Hinkle, Jr. P.S., Kelley, J.M., Klimem, K.M., Kelley, J.C., Liu, L-I., Marmaros, S.M., Merrick, J.M., Moreno-Palanques, R.F., McDonald, L.A., Nguyen, D.T., Phillips, C.A., Ryder, S.E., Scott, J.L., Saudek, D.M., Shirley, R., Small, K.V., Spriggs, T.A., Utterback, A., Weildman, J.F., Li, Y., Barthlow, R., Bednarik, D.P., Cao, L., Cepeda, M.A., Coleman, T.A., Collins, E-J., Dimke, D., Feng, P., Ferrie, A., Fischer, C., Hastings, G.A., He, W-W., Hu, J-S., Huddleston, K.A., Greene, J.M., Gruber, J., Hudson, P., Kim, A., Kozak, D.L., Kunsch, C., Ji, H., Li, H., Meissner, P.S., Olson, H., Raymond, L., Wei, Y-F., Wiing, J., Xu, C., Yu, G-L., Ruben, S.M., Dillon, P.J., Fannon, M.R., Rosen, C.A., Haseltine, W.A., Fields, C., Fraser, C.M., & Venter, J.C. (1995) Initial assessment of human gene diversity and expression patterns based upon 83 million nucleotides of cDNA sequence. *Nature*, 377(Suppl.), 3-174.
- Afifi, A.K., Smith, J.W., & Zellweger, H. (1965) Congenital nonprogressive myopathy. Central core disease and nemaline myopathy in one family. *Neurology*, 15, 371-381.
- Ahlberg, G., Borg, K., Ansved, T., Edström, L., & Anvret, M. (1998) Welander distal myopathy: genetic linkage and candidate genes on chromosome 2. *Muscle & Nerve*, S7, S6.



- Ahlberg, G., von Tell, D., Borg, K., Edström, L., & Anvret, M. (1999) Genetic linkage of Welander distal myopathy to chromosome 2p13. *Ann Neurol*, 46(3), 399-404.
- Akkari, P.A., Eyre, H., Wilton, S.D., Callen, D.C., Meredith, C., Kedes, L., & Laing, N.G. (1994) Assignment of the human skeletal muscle alpha actin gene (*ACTA1*) to 1q42 by fluorescence *in situ* hybridisation. *Cytogenet Cell Genet*, 65, 265-267.
- Alberts, B., Bray, D., Lewis J., Raff, M., Roberts, K., & Watson J.D. (1994) *Molecular Biology of the Cell*. New York: Garland Publishing Inc.
- Albertstein, H.M., Abderrahim, H., Cann, H.M., Dausset, J., Paslier, D.L., & Cohen, D. (1990) Construction and characterisation of a yeast artificial chromosome library containing seven haploid genome equivalents. *Proc Natl Acad Sci, USA*, 87, 4256-4260.
- Alwine, J.C., Kemp, D.J., & Stark, G.R. (1977) Method for detection of specific mRNAs in gels by transfer to diazobenzyloxymethyl-paper and hybridization with DNA probes. *Proc Natl Acad Sci, USA*, 74, 5350-5354.
- Amato, A.A., Jackson, C.,E., Ridings, L.,W., & Barohn, R.,J. (1995) Childhood onset oculopharyngodistal myopathy with chronic intestinal pseudo-obstruction. *Muscle &Nerve*, 18(8), 842-847.
- Anderson, L.V., Davison, K., Moss, J.A., Young, C., Cullen, M.J., Walsh, J., Johnson, M.A., Bashir, R., Britton, S., Keers, S., Argov, Z., Mahjneh, I., Fougousse, F., Beckmann, J.S., & Bushby, K.M. (1999) Dysferlin is a plasma membrane protein and is expressed early in human development. *Hum Mol Genet*, 8(5), 855-861.
- Anderson, M.S., & Kunkel, L.M. (1992) The molecular and biochemical basis of Duchenne muscular dystrophy. *Trends Biochem Sci*, 17(8), 289-292.
- Andrews, P.I., & Wilson, J. (1992) Relative disease severity in siblings with myotonic dystrophy. *J Child Neurol*, 7(2), 161-167.
- Argov, Z., & Yarom, R. (1984) "Rimmed vacuole myopathy" sparing the quadriceps. A unique disorder in Iranian Jews. *J Neurol Sci*, 64(1), 33-43.
- Argov, Z., Tiram, E., Eisenberg, I., Sadeh, M., Seidman, C.E., Seidman, J.G., Karpati, G., & Mitrani-Rosenbaum, S. (1997) Various types of hereditary inclusion body myopathies map to chromosome 9p1-q1. *Ann Neurol*, 41(4), 548-551.

- Argov, Z., & Mitrani-Rosenbaum, S. (1998) Hereditary inclusion-body myopathy with quadriceps sparing: epidemiology and genetics. In V. Askanas, G. Serratrice, & W. King Engel (Eds.), *Inclusion-body Myositis and Myopathies* (pp. 200-210). Cambridge, UK: Cambridge University Press.
- Armour, J.A.L. (1996) Tandemly repeated minisatellites: generating human genetic diversity via recombinational mechanisms. In M.S. Jackson, T. Strachan & G. Dover (Eds.), *Human Genome Evolution* (pp. 171-190). Oxford, UK: BIOS Scientific Publisher.
- Askanas, V., & Engel, W.K. (1993) New advances in inclusion-body myositis. *Curr Opin Rheumatol*, 5(6), 732-741.
- Askanas, V., & Engel, W.K. (1995) New advances in the understanding of sporadic inclusion-body myositis and hereditary inclusion-body myopathies. *Curr Opin Rheumatol*, 7, 486-496.
- Askanas, V. (1997) New developments in heredity inclusion body myopathies. *Ann Neurol*, 41(4), 421-422.
- Askanas, V., & Engel, W.K. (1998) Overview of pathologic and pathogenic comparisons between sporadic inclusion-body myositis and hereditary inclusion-body myopathies. In V. Askanas, G. Serratrice, & W. King Engel (Eds.), *Inclusion-body Myositis and Myopathies* (pp.3-78). Cambridge, UK: Cambridge University Press.
- Atkinson, S.J., & Stewart, M. (1992) Molecular interactions in myosin assembly. Role of the 28-residue charge repeat in the rod. *J Mol Biol*, 226(1), 7-13.
- Balding, D.J. & Torney, D.C. (1991) Statistical analysis of DNA fingerprint data for ordered clone physical mapping of human chromosomes. *Bull Math Biology*, 53, 853-879.
- Ballabio, A. (1993). The rise and fall of positional cloning? *Nat Genet*, 3, 277-279.
- Barlow, D.J., & Thornton, J.M. (1988) Helix geometry in proteins. *J Mol Biol*, 201(3), 601-619.
- Barohn, R.J., Miller, R.G., & Griggs, R.C. (1991) Autosomal recessive distal myopathy. *Neurol*, 41, 1365-1370.
- Barohn, R.J. & Amato, A.A. (1999) Distal myopathies. *Seminars in Neurology*, 19(1), 45-58.

- Bartholdi, M., Meyne, J., Albright, K., Luedemann, M., Campbell, E., Chritton, D., Deaven, L.L., & Cram, L.S. (1987) Chromosome sorting by flow cytometry. *Methods Enzymol*, 151, 252-267.
- Bashir, R., Strachan, T., Keers, S., Stephenson, A., Mahjneh, I., Marconi, G., Nashef, L., & Bushby, K.M. (1994) A gene for autosomal recessive limb-girdle muscular dystrophy maps to chromosome 2p. *Hum Mol Genet*, 3(3), 455-457.
- Bashir, R., Keers, S., Strachan, T., Passos-Bueno, R., Zatz, M., Weissenbach, J., Le Paslier, D., Meisler, M., & Bushby, K. (1996) Genetic and physical mapping at the limb-girdle muscular dystrophy locus (*LGMD2B*) on chromosome 2p. *Genomics*, 33, 46-52.
- Bashir, R., Britton, S., Strachan, T., Keers, S., Vafiadaki, E., Lako, M., Richard, I., Marchand, S., Bourg, N., Argov, Z., Sadeh, M., Mahjneh, I., Marconi, G., Passos-Bueno, M.R., Moreira, E. de S., Zatz, M., Beckmann, J.S., & Bushby, K. (1998) A gene related to *Caenorhabditis elegans* spermatogenesis factor fer-1 is mutated in limb-girdle muscular dystrophy type 2B. *Nat Genet*, 20(1), 37-42.
- Bassam, B.J., Caetano-Anolles, G., & Gresshoff, P.M. (1991) Fast and sensitive silver staining of DNA in polyacrylamide gels. *Analyt Biochem*, 196, 80-83.
- Bassett, D., Boguski, M., Spencer, F., Reeves, R., Kim, S-H., Weaver, T., & Hieter, P. (1997) Genome cross-referencing and XREFdb: implications for identification and analysis of genes mutated in human disease. *Nat Genet*, 15, 339-344.
- Bautista, J., Rafel, E., Castilla, J.M., & Alberca, R. (1978) Hereditary distal myopathy with onset in early infancy. Observation of a family. *J Neurol Sci*, 37(3), 149-158.
- Beckmann, J.S. (1988). Oligonucleotide polymorphisms: a new tool for genomics genetics. *Bio/Technology*, 6, 1061-1064.
- Beckmann, J.S., Richard, I., Hillaire, D., Broux, O., Antignac, C., Bois, E., Cann, H., Cottingham, R.W. Jr., Feingold, N., & Feingold, J. (1991) A gene for limb-girdle muscular dystrophy maps to chromosome 15 by linkage. *C R Acad Sci III*, 312(4), 141-148.
- Beggs, A.H., Koenig, M., Boyce, F.M., & Kunkel, L.M. (1990) Detection of 98% of DMD/BMD gene deletions by polymerase chain reaction. *Hum Genet*, 86(1), 45-48.
- Beggs, A.H., Hoffman, E.P., Snyder, J.R., Arahata K., Specht, L., Shapiro, F., Angelini, C., Sugita, H., & Kunkel, L.M. (1991). Exploring the molecular basis for variability among patients with Becker muscular dystrophy: dystrophin gene and protein studies. *Am J Hum Genet*, 49(1), 54-67.

- Behn-Krappa, A., & Doerfler, W. (1994) Enzymatic amplification of synthetic oligodeoxyribonucleotides: implications for triplet repeat expansions in the human genome. *Hum Mutat*, 3(1), 19-24.
- Bejaoui, K., Hirabayashi, K., Hentati, F., Haines, J.L., Ben Hamida, C., Belal, S., Miller, R.G., McKenna-Yasek, D., Weissenbach, J., Rowland, L.P., Griggs, R.C., Munsat, T.L., Ben Hamida, M., Arahata, K., & Brown, R.H. (1995) Linkage of Miyoshi myopathy (distal autosomal recessive muscular dystrophy) locus to chromosome 2p12-14. *Neurol*, 45, 768-767.
- Bell, J. (1947) Dystrophia myotonica and allied disease. In: Penrose, L. S. *Treasury of Human Inheritance* (4th ed., pp. 343-410). Cambridge, UK: Cambridge University Press.
- Bellanne-Chantelot, C., Barillot, E., Lacroix, B., Le Paslier, D., & Cohen, D. (1991) A test case for physical mapping of human genome by repetitive sequence fingerprints: construction of a physical map of a 420kb YAC subcloned into cosmids. *Nucleic Acids Res*, 19, 505-515.
- Bellanne-Chantelot, C., Lacroix, B., Ougen, P., Billaut, A., Beaufrils, S., Bertrand, S., Georges, I., Gilbert, F., Gros, I., Lucotte, G., Susini, L., Codani, J.-J., Gesnoui, P., Pook, S., Vaysseix, G., Lu-kuo, J., Ried, T., Ward, D., Chumakov, I., Le Paslier, D., Barillot, E., & Cohen, D. (1992) Mapping the whole human genome by fingerprinting yeast artificial chromosomes. *Cell*, 70, 1059-1068.
- Ben Othmane, K., Ben Hamida, M., Pericak-Vance, M.A., Homida, C.B., Blal, S., Carter, S.C., Bowcock, A.M., Petrukhin, K., Gilliam, T.C., Roses, A.D., Hentati, F., & Vance, J.M. (1992) Linkage of the Tunisian autosomal recessive Duchenne-like muscular dystrophy to the pericentromeric region of chromosome 13q. *Nat Genet*, 2, 315-317.
- Ben Othmane, K., Speer, M.C., Stauffer, J., Blal, S., Middleton, L., Hamida, C., Etrib, A., Loeb, D., Hentati, F., Roses, A.D., Ben Hamida, M., Pericak-Vance, M.A., Vance, J. M. (1995) Evidence for linkage disequilibrium in chromosome 13-linked Duchenne-like muscular dystrophy (LGMD2C). *Am J Hum Genet*, 57(3), 732-734.
- Bernstein, S.I., Mogami, K., Donady, J.J., & Emerson, C.P. Jr. (1983) *Drosophila* muscle myosin heavy chain encoded by a single gene in a cluster of muscle mutations. *Nature*, 302(5907), 393-397.
- Berrettini, W.H., Ferraro, T.N., Goldin, L.R., Detera-Wadleigh, S.D., Choi, H., Muniec, D., Guroff, J.J., Kazuba, D.M., Nurnberger, J.I. Jr., Hsieh, W.T., Hoehe, M.R., & Gershon, E.S. (1997) A linkage study of bipolar illness. *Arch Gen Psychiatry*, 54(1), 27-35.

- Bethlem, J., & Wijngaarden, G.K. (1976) Benign myopathy with autosomal dominant inheritance. A report on three pedigrees. *Brain*, 99(1), 91-100.
- Bienroth, S., Keller, W., & Wahle, E. (1993) Assembly of a processive messenger RNA polyadenylation complex. *EMBO J*, 12(2), 585-594.
- Bickeboller, H., & Clerget-Darpoux, F. (1995) Statistical properties of the allelic and genotypic transmission/disequilibrium test for multi-allelic markers. *Genet Epidemiol*, 12(6), 865-870.
- Bione, S., Maestrini, E., Rivella, S., Mancini, M., Regis, S., Romeo, G., & Toniolo, D. (1994) Identification of a novel X-linked gene responsible for Emery-Dreifuss muscular dystrophy. *Nat Genet*, 8(4), 323-327.
- Bird, A.P. (1986) CpG-rich islands and the function of DNA methylation. *Nature*, 321, 209-213.
- Bishop, D.T., & Williamson, J.A. (1990) The power of identity-by-state methods for linkage analysis. *Am J Hum Genet*, 46(2), 254-265.
- Blackwelder, W.C., & Elston, R.C. (1985) A comparison of sib-pair linkage tests for disease susceptibility loci. *Genet Epidemiol*, 2(1), 85-97.
- Blank, R.D., Campbell, G.R., Pollak, M., & D'Eustachio, P. (1988) Bayesian multilocus linkage mapping. *Curr Top Microbiol Immunol*, 137, 25-32.
- Bloch, W. (1991) A biochemical perspective of the polymerase chain reaction. *Biochem*, 30, 2735-2747.
- Blumen, S.C., Nisipeanu, P., Sadeh, M., Asherov, A., Blumen, N., Wirguin, Y., Khilkevich, O., Carasso, R.L., & Korczyn, A.D. (1997) Epidemiology and inheritance of oculopharyngeal muscular dystrophy in Israel. *Neuromuscul Disord*, 7 (Suppl 1), S38-S40.
- Blumen, S.C., Brais, B., Korczyn, A.D., Medinsky, S., Chapman, J., Asherov, A., Nisipeanu, P., Codere, F., Bouchard, J.P., Fardeau, M., Tome, F.M., & Rouleau, G.A. (1999) Homozygotes for oculopharyngeal muscular dystrophy have a severe form of the disease. *Ann Neurol*, 46(1), 115-118.
- Bodensteiner, J.B. (1994) Congenital myopathies. *Muscle & Nerve*, 17(2), 131-144.

- Bodmer, W. (1981) Gene clusters, genome organisation and complex phenotypes. When the sequence is known, what will it mean? *Am J Hum Genet*, 33, 664-682.
- Boehnke, M. (1994) Limits of resolution of genetic linkage studies: implications for the positional cloning of human disease genes. *Am J Hum Genet*, 55(2), 379-390.
- Bonne, G., Carrier, L., Bercovici, J., Cruaud, C., Richard, P., Hainque, B., Gautel, M., Labeit, S., James, M., & Beckmann J. (1995) Cardiac myosin binding protein-C gene splice acceptor site mutation is associated with familial hypertrophic cardiomyopathy. *Nat Genet*, 11(4), 438-440.
- Bonnemann, C.G., Modi, R., Noguchi, S., Mizuno, Y., Yoshida, M., Gussoni, E., McNally, E.M., Duggan, D.J., Angelini, C., & Hoffman, E.P. (1995) Beta-sarcoglycan (A3b) mutations cause autosomal recessive muscular dystrophy with loss of the sarcoglycan complex. *Nat Genet*, 11(3), 266-273.
- Bonnemann, C.G., McNally, E.M., & Kunkel, L.M. (1996) Beyond dystrophin: current progress in the muscular dystrophies. *Curr Opin Pediatr*, 8(6), 569-582.
- Borg, K., Ahlberg, G., Borg, J., & Edström, L. (1991) Welander's distal myopathy: clinical, neurophysiological and muscle biopsy observations in young and middle aged adults with early symptoms. *J Neurol Neurosurg Psychiatry*, 54(6), 494-498.
- Botstein, D., White, R.L., Skolnick, M., & Davis, R.W. (1980) Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *Am J Hum Genet*, 32, 314-331.
- Bouchard, J.P. (1997) Andre Barbeau and the oculopharyngeal muscular dystrophy in French Canada and North America. *Neuromuscul Disord*, 7(Suppl 1), S5-S11.
- Bouchard, J.P., Brais, B., & Tome, F.M.S. (1998) Oculopharyngeal muscular dystrophy. In A.E.H. Emery (Ed.) *Neuromuscular Disorders: Clinical and Molecular Genetics* (pp. 155-179) West Sussex: John Wiley & Sons Ltd.
- Boucher, C.A., King, S.K., Carey, N., Krahe, R., Winchester, C.L., Rahman, S., Creavin, T., Meghji, P., Bailey, M.E., Chartier, F.L., Brown, S.D., Siciliano, M.J., & Johnson, K.J. (1995) A novel homeodomain-encoding gene is associated with a large CpG island interrupted by the myotonic dystrophy unstable (CTG)<sub>n</sub> repeat. *Hum Mol Genet*, 4(10), 1919-1925.

- Boyd, Y., & Buckle, V.J. (1986) Cytogenetic heterogeneity of translocations associated with Duchenne muscular dystrophy. *Clin Genet*, 29(2), 108-115.
- Brais, B., Xie, Y.G., Sanson, M., Morgan, K., Weissenbach, J., Korczyn, A.D., Blumen, S.C., Fardeau, M., Tome, F.M.S., Bouchard, J.P., & Rouleau, G.A. (1995) The Oculopharyngeal muscular dystrophy locus maps to the region of the cardiac alpha and beta myosin heavy chains on chromosome 14q11.2-q13. *Hum Mol Genet*, 4, 429-434.
- Brais, B., Bouchard, J.P., Gosselin, F., Xie, Y-G., Fardeau, M., Tome, F.M.S., & Rouleau, G.A. (1997) Using the full power of linkage analysis in 11 French Canadian families to fine map the oculopharyngeal muscular dystrophy gene. *Neuromuscul Disord*, 7(Suppl. 1), S70-S74.
- Brais, B., Bouchard, J.P., Xie, Y-G., Rochefort, D.L., Chretien, N., Tome, F.M.S., Lafreniere, R.G., Rommens, J.M., Uyama, E., Nohira, O., Blumen, S.C., Korczyn, A.D., Heutink, P., Mathieu, J., Duranceau, A., Codere, F., Fardeau, M., & Rouleau, G.A. (1998) Short GCG expansions in the *PABP2* gene cause oculopharyngeal muscular dystrophy. *Nat Genet*, 18, 164-167.
- Brandts, J.F., Halvorson, H.R., & Brennan, M. (1975) Consideration of the possibility that the slow step in protein denaturation reactions is due to cis-trans isomerism of proline residues. *Biochemistry*, 14(22), 4953-4963.
- Bray, G.M., Kaarsoo, M., & Ross R.T. (1965) Ocular myopathy with dysphagia. *Neurology*, 15, 678-684.
- Brook, J.D., McCurrach, M.E., Harle, H.G., Buckler, A.J., Church, D., Aburatani, H., Hunter, K., Stanton, V.P., Thirion, J-P., Hudson, T., Sohn, R., Zemelman, B., Snell, R.G., Rundle, S.A., Crow, S., Davies, J., Shelbourne, P., Buxton, J., Jones, C., Juvonen, V., Johnson, K., Harper, P.S., Shaw, D.J., & Housman, D.E. (1992) Molecular basis of myotonic dystrophy: expansion of a trinucleotide (CTG) repeat at the 3-prime end of a transcript encoding a protein kinase family member. *Cell*, 68, 799-808.
- Brooke, M., Fenichel, G., Griggs, R., Mendell, J., Moxley, R., Miller, J., & Province, M. (1983) Clinical investigation of Duchenne dystrophy: determination of the "power of" therapeutic trials based on the natural history of the disease. *Muscle & Nerve*, 6, 91-103.
- Brown, J.H., Cohen, C., & Parry, D.A. (1996) Heptad breaks in alpha-helical coiled coils: stutters and stammers. *Proteins*, 26(2), 134-145.
- Brown, R.H. Jr. (1997) Dystrophin-associated proteins and the muscular dystrophies. *Ann Rev Med*, 48, 457-466.



- Brown, W.R.A. & Bird, A.P. (1986) Long-range restriction site mapping of mammalian genomic DNA. *Nature*, 322, 477-481.
- Buckler, A.J., Chang, D.D., Graw, S.L., Brook, J.D., Haber, D.A., Sharp, P.A., & Houseman, D.E. (1991) Exon amplification: a strategy to isolate mammalian genes based on RNA splicing. *Proc Natl Acad Sci, USA*, 88, 4005-4009.
- Buckton, K.E., O'Riordan, M.L., Jacobs, P.A., Robinson, J.A., Hill, R., & Evans, H.J. (1976) C- and Q-band polymorphisms in the chromosomes of three human populations. *Ann Hum Genet, (Lond.)*, 40, 99-111.
- Bushby, K.M.D., & Beckmann, J.S. (1995) The limb-girdle muscular dystrophies - proposal for a new nomenclature. *Neuromuscul Disord*, 5, 337-343.
- Byerley, W.F. (1989) Schizophrenia. Genetic linkage revisited. *Nature*, 340(6232), 340-341.
- Callen, D.F., Baker, E., Eyre, H.J., & Lane, S.A. (1990) An expanded mouse-human hybrid cell panel for mapping human chromosome 16. *Ann Genet*, 33, 190-195.
- Chamberlain, J.S., Gibbs, R.A., Ranier, J.E., Nguyen, P.N., & Caskey, C.T. (1988) Deletion screening of the Duchenne muscular dystrophy locus via multiplex DNA amplification. *Nucleic Acids Res*, 16, 11141-11156.
- Cavenee, W.K., Dryja, T.P., Phillips, R.A., Benedict, W.F., Godbout, R., Gallie, B.L., Murphree, A.L., Strong, L.C., & White, R.L. (1983) Expression of recessive alleles by chromosomal mechanisms in retinoblastoma. *Nature*, 305(5937), 779-784.
- Cayanis, E., Russo, J.J., Kalachikov, S., Ye, X., Park, S.H., Sunjevaric, I., Bonaldo, M.F., Lawton, L., Venkatraj, V.S., Schon, E., Soares, M.B., Rothstein, R., Warburton, D., Edelman, I.S., Zhang, P., Efstratiadis, A., & Fischer, S.G. (1998) High-resolution YAC-cosmid-STS map of human chromosome 13. *Genomics*, 47(1), 26-43.
- Chamberlain, J.S., Gibbs, R.A., Ranier, J.E., Nguyen, P.N., & Caskey, C.T. (1988) Deletion screening of the Duchenne muscular dystrophy locus via multiplex DNA amplification. *Nucleic Acids Res*, 16, 11141-11156.
- Cheng, J-F., & Zhu, Y. (1994) Reverse transcription-polymerase chain reaction detection of transcribed sequences on human chromosome 21. *Genomics*, 20, 184-190.

- Chomczynski, P. (1993) A reagent for the single-step simultaneous isolation of RNA, DNA and proteins from cell and tissue samples. *Biotechniques*, 15(3), 532-537.
- Chothia, C., & Janin, J. (1982) Orthogonal packing of beta-pleated sheets in proteins. *Biochemistry*, 21(17), 3955-3965.
- Chumakov, I.M., Rigault, P., Guillou, S., Ougen, P., Billaut, A., Guasconi, G., Gervy, P., Le Gall, I., Soularue, P., Grinas, L., Bougueleret, L., Bellanne-Chantelot, C., Lacroix, B., Barillot, E., Gesnouin, P., Pook, S., Vaysseix, G., Frelat, G., Schmitz, A., Sambucy, J-L., Bosch, A., Estivill, X., Weissenbach, J., Vignal, A., Riethman, H., Cox, D., Patterson, D., Gardiner K, Hattori M, Sakaki Y, Ichikawa H, Ohki M, Le Paslier D, Heilig, R., Antonarakis, S., & Cohen, D. (1992a) Continuum of overlapping clones spanning the entire human chromosome 21q. *Nature*, 359, 380-387.
- Chumakov, I.M., Le Gall, I., Billaut, A., Ougen, P., Soularue, P., Guillou, S., Rigault, P., Guasconi, G., Bui, H., De, T.M-F., Barillot, E., Abderrahim, H., Cherif, D., Berger, R., Le Paslier, D., & Cohen, D. (1992b) Isolation of chromosome 21-specific yeast artificial chromosomes from a total human genomic library. *Nat Genet*, 1, 222-225.
- Cohen, C., & Parry, D.A. (1998) A conserved C-terminal assembly region in paramyosin and myosin rods. *J Struct Biol*; 122(1-2), 180-187.
- Cohen, D., Chumakov, I.M., & Weissenbach, J. (1993) A first-generation physical map of the human genome. *Nature*, 366, 698-701.
- Collins, F.S., Ponder, B.A., Seizinger, B.R., & Epstein, C.J. (1989) Editorial: the von Reckingham neurofibromatosis region on chromosome 17 - genetic and physical maps come into focus. *Am J Hum Genet*, 44(1), 1-5.
- Collins, F.S. (1992) Positional cloning: let's not call it reverse anymore. *Nat Genet*, 1, 3-6.
- Collins, F.S. (1995) Positional cloning moves from perditional to traditional. *Nat Genet*, 9, 347-350.
- Collins, F.S., Patrinos, A., Jordan, E., Chakravarti, A., Gesteland, R., & Walters, L. (1998) New goals for the U.S. Human Genome Project: 1998-2003. *Science*, 282(5389), 682-689.
- Consalez, G.G., Thomas, N.S.T., Stayson, C.L., Knight, S.J.L., Johnson, M., Hopkins, I.C., Harper, P.S., & Elsas, L.J. (1991) Assignment of Emery-Dreifuss muscular dystrophy: localization to the distal region Xq28: the results of a collaborative approach. *Am J Hum Genet*, 48, 468-480.

- Cooper, D.N., & Krawczak, M. (1994) *Human Gene Mutation*. Oxford, UK: BIOS Scientific Publishers Limited.
- Cooper, D.N. (1995a) Structure and function in the human genome. In F. Farzaneh & D.N. Cooper (Eds.), *Functional Analysis of the Human Genome* (p. 1-41). Oxford, UK: BIOS Scientific Publishers Limited.
- Cooper, D.N. (1995b) Mapping the human genome. In F. Farzaneh & D.N. Cooper (Eds.), *Functional Analysis of the Human Genome* (p. 43-68). Oxford, UK: BIOS Scientific Publishers Limited.
- Cordell, H.J., Todd, J.A., Bennett, S.T., Kawaguchi, Y., & Farrall, M. (1995) Two locus maximum lod score analysis of a multifactorial trait: joint consideration of IDDM2 and IDDM4 with IDDM1 in type 1 diabetes. *Am J Hum Genet*, 59(4), 920-934.
- Corder, E.H., Saunders, A.M., Strittmatter, W.J., Schmechel, D.E., Gaskell, P.C., Small, G.W., Roses, A.D., Haines, J.L., & Pericak-Vance, M.A. (1993) Gene dose of apolipoprotein E type 4 allele and the risk of Alzheimer's disease in late onset families. *Science*, 261(5123), 921-923.
- Cormand, B., Avela, K., Pihko, H., Santavuori, P., Talim, B., Topaloglu, H., de la Chapelle, A., & Lehesjoki, A.E. (1999) Assignment of the muscle-eye-brain disease gene to 1p32-p34 by linkage analysis and homozygosity mapping. *Am J Hum Genet*, 64(1), 126-135
- Cotter, F., Nasipuri, S., Lam, G., & Young, B.D. (1989) Gene mapping by enzymatic amplification from flow-sorted chromosomes. *Genomics*, 5, 470-474.
- Cotton, R.G.H., Rodrigues, N.R., & Campbell, R.D. (1988) Reactivity of cytosine and thymidine in single base-pair mismatches with hydroxylamine and osmium tetroxide and its application to the study of mutations. *Proc Natl Acad Sci, USA*, 85, 4397-4401.
- Cox, D.R., Burmeister, M., Price, E.R., Kim, S., & Myers, R.M. (1990) Radiation hybrid mapping: a somatic cell genetic method for constructing high-resolution maps of mammalian chromosomes. *Science*, 250, 245-250.
- Cox, D.W., Gedde Dahl, T., Menon, A.G., Nygaard, T.G., Tomlinson, I.M., Peters, J., & St George-Hyslop, P.H. (1995) Report of the second international workshop on human chromosome 14 mapping 1994. *Cytogenet Cell Genet*, 69, 160 -171.
- Craig, R. (1994) The structure of contractile filaments. In A. G. Engel, & Franzini-Armstrong, C. (Eds.), *Myology: Basic and Clinical* (2nd Ed.) (pp. 134-175). New York: McGraw-Hill, Inc.

- Crick, F.H.C. (1952) The packing of  $\alpha$ -helices of the same sense. *Nature Lond*, 170, 882.
- Crick, F.H.C. (1953) The packing of  $\alpha$ -helices: simple coiled-coils. *Acta Cryst*, 6, 689-697.
- Dauwerse, J.G., Wiegant, J., Raap, A.K., Breuning, M.H., & van Ommen, G.J. (1992) Multiple colors by fluorescence in situ hybridization using ratio-labelled DNA probes create a molecular karyotype. *Hum Mol Genet*, 1(8), 593-598.
- Davies, J.L., Kawaguchi, Y., Bennett, S.T., Copeman, J.B., Cordell, H.J., Pritchard, E., Reed, P.W., Gough, S.C., Jenkins, S.C., Palmer, S.M., Balfour, S.M., Rowse, B.R., Farrall, M., Barnett, A.H., Bain, S.C., & Todd, J.A. (1994) A genome-wide search for human type 1 diabetes susceptibility genes. *Nature*, 371(6493), 130-136.
- Davis, S., Schroeder, M., Goldin, L.R., & Weeks, D.E. (1996) Nonparametric simulation based statistics for detecting linkage in general pedigrees. *Am J Hum Genet* 53(1), 252-263.
- Day, J.W., Roelofs, R., Leroy, B., Pech, I., Benzow, K., & Ranum, L.P. (1999) Clinical and genetic characteristics of a five-generation family with a novel form of myotonic dystrophy (DM2) *Neuromuscul Disord.*, 9(1), 19-27.
- De Braekeleer, M. (1991) Hereditary disorders in Saguenay-Lac-St-Jean (Quebec, Canada) *Hum Hered*, 41, 141-146.
- Deber, C.M., Glibowicka, M., & Woolley, G.A. (1990) Conformations of proline residues in membrane environments. *Biopolymers*, 29(1), 149-157.
- den Dunnen, J.T., Grootsholten, P.M., Bakker, E., Blonden, L.A., Ginjaar, H.B., Wapenaar, M.C., van Paassen, H.M., van Broeckhoven, C., Pearson, P.L., & van Ommen, G.J. (1989) Topography of the Duchenne muscular dystrophy (DMD) gene: FIGE and cDNA analysis of 194 cases reveals 115 deletions and 13 duplications. *Am J Hum Genet*, 45(6), 835-847.
- de Lange, T., Shiue, L., Myers, R.M., Cox, D.R., Naylor, S.L., Killery, A.M., & Varmus, H.E. (1990) Structure and variability of human chromosome ends. *Mol Cell Biol*, 10, 518-527.
- Department of Energy (USA). (1985) *Technologies for Detecting Heritable Mutations in Human Beings*. Washington DC: DOE.
- Department of Energy (USA). (1987) *Report on the Human Genome Initiative*. Washington DC: DOE.

- de Vries, H.G., van der Meulen, M.A., Rozen, R., Halley, D.J., Scheffer, H., ten Kate, L.P., Buys, C.H., & te Meerman, G.J. (1996) Haplotype identity between individuals who share a CFTR mutation allele "identical by descent": demonstration of the usefulness of the haplotype-sharing concept for gene mapping in real populations. *Hum Genet*, 98(3), 304-309.
- Does de Willebois, A.E. va, Meyer, A.E.F.H., Simons, A.J.R., & Bethlem, J. (1968) Distal myopathy with onset in early infancy. *Neurol* 18(4), 383-390.
- Doggett, N.A., Goodwin, L.A., Tesmer, J.G., Meincke, L.J., Bruce, D.C., Clark, L.M., Altherr, M.R., Ford, A.A., Chi, H-C., Marrone, B.L., Longmire, J.L., Lane, S.A., Whitmore, S.A., Lowenstein, M.G., Sutherland, R.D., Mundt, M.O., Knill, E.H., Bruno, W.J., Macken, C.A., Torney, D.C., Wu, J-R., Griffith, J., Sutherland, G.R., Deven, L.L., Callen, D.F., & Moyzis, R.K. (1995) An integrated physical map of chromosome 16. *Nature*, 377(S), 335-365.
- Donner, K., Ollikainen, M., Pelin, K., Grönholm, M., Carpen, O., Wallgren-Pettersson, C., Ridanpää, M. (2000) Mutations in the  $\beta$ -tropomyosin (*TPM2*) gene in rare cases of autosomal dominant nemaline myopathy. *Neuromusc Disord*, 10, 342-343.
- Dryja, T.P., McGee, T.L., Reichel, E., Hahn, L.B., Cowley, G.S., Yandell, D.W., Sandberg, M.A., & Berson, E.L. (1990) A point mutation of the rhodopsin gene in one form of retinitis pigmentosa. *Nature*, 343, 364-366.
- Dubois, B.L., & Naylor, S.L. (1992) Characterisation of NIGMS human/rodent somatic cell hybrid mapping panel 2 by PCR. *Genomics*, 16, 315-319.
- Duguid, J.R., Rohwer, R.G., & Seed, B. (1988) Isolation of cDNA of scrapie-modulated RNAs by subtractive hybridisation of a cDNA library. *Proc Natl AcadSci, USA*, 85, 949-955.
- Duyk, G.M., Kim, S., Myers, R.M., & Cox, D.R. (1990) Exon trapping: a genetic screen to identify candidate transcribed sequences in cloned mammalian genomic DNA. *Proc Natl Acad Sci, USA*, 87, 8995-8999.
- Edström, L. (1975) Histochemical and histopathological changes in skeletal muscles in late-onset hereditary myopathy (Welander). *J Neurol Sci*, 26, 147-157.
- Edström, L., Thornell, L., & Eriksson, A. (1980) A new type of hereditary distal myopathy with characteristic sarcoplasmic bodies and intermediate (skeleton) filaments. *J Neurol Sci*, 47, 171-190.

- Edwards A.W.F. (1997) The early history of the statistical estimation of linkage. In I-H Pawlowitzki, J.H. Edwards & E.A. Thompson (Eds.), *Genetic Mapping of Disease Genes*, (pp 9-14). San Diego: Academic Press Inc.
- Edwards, J.H. (1980) Allelic association in man. In A.W. Erickson (Ed.), *Population Structure and Genetic Disorders*, (pp. 239-256). New York: Academic Press Inc.
- Edwards, J.H. (1987) Exclusion mapping. *J Med Genet*, 24, 539-543.
- Edwards, J.H. (1988) The importance of genetic diseases and the need for prevention. *Phil Trans R Soc Lond B*, 319, 211-227.
- Edwards, J.H. (1989) The locus positioning problem. *Ann Hum Genet*, 53, 271-275.
- Edwards J.H. (1997) Recessive disease and allelic association. In I-H Pawlowitzki, J.H. Edwards & E.A. Thompson (Eds.), *Genetic Mapping of Disease Genes*, (pp 31-57). San Diego: Academic Press Inc.
- Elvin, P., Slynn, G., Black, D., Graham, A., Butler, R., Riley, J., Anand, R., & Markham, A.F. (1990) Isolation of cDNA clones using yeast artificial chromosome probes. *Nucleic Acids Res*, 18, 3913-3197.
- Emery, A.E.H., & Dreifuss, F.E. (1966) Unusual type of benign X-linked muscular dystrophy. *J Neurol Neurosurg Psychiatry*, 29, 338.
- Emery, A.E.H. (1998) Molecular genetics of Neuromuscular disorders: applications in clinical medicine. In A.E.H. Emery (Ed.), *Neuromuscular Disorders: Clinical and Molecular Genetics* (pp. 1-20). West Sussex: John Wiley & Sons Limited.
- Epp, T.A., Dixon, I.M., Wang, H.Y., Sole, M.J., & Liew, C.C. (1993) Structural organization of the human cardiac alpha-myosin heavy chain gene (*MYH6*). *Genomics*, 18(3), 505-509.
- Ernfors, P., Lee, K.F., Kucera, J., & Jaenisch, R. (1994) Lack of neurotrophin-3 leads to deficiencies in the peripheral nervous system and loss of limb proprioceptive afferents. *Cell*, 77(4), 503-512.
- Engel, W.K., Gold, G.N., & Karpati, G. (1968) Type 1 fibre hypotrophy and central nuclei. *Arch Neurol*, 18, 435-444.
- Eyre, H.J., Akkari, P.A., Meredith, C., Wilton, S.D., Callen, D.C., Kedes, L., & Laing, N.G. (1993) Assignment of the human slow skeletal muscle troponin gene (*TNNI1*) to 1q32 by fluorescence *in situ* hybridisation. *Cytogenet Cell Genet*, 62, 181-182.

- Falk, C.T., & Rubenstein, P. (1987) Haplotype relative risks: an easy reliable way to construct a proper control sample for risk calculations. *Ann Hum Genet*, 51(3), 277-233.
- Fanning, T.G., & Singer, M.F. (1987) LINE-1: a mammalian transposable element. *Biochim Biophys Acta*, 910(3), 203-212.
- Fananapazir, L., Dalakas, M.C., Cyran, F., Cohn, G., & Epstein, N.D. (1993) Missense mutations in the beta-myosin heavy-chain gene cause central core disease in hypertrophic cardiomyopathy. *Proc Natl Acad Sci USA*, 90(9), 3993-3997.
- Fardeau, M., & Tome, F.M.S. (1998) Inclusion-body myopathies. In V. Askanas, G. Serratrice, & W. King Engel (Eds.), *Inclusion-body Myositis and Myopathies* (pp. 252-260). Cambridge, UK: Cambridge University Press.
- Fearon, E.R., Cho, K.R., Nigro, J.M., Kern, S.E., Simons, J.W., Ruppert, J.M., Hamilton, S.R., & Preisinger, A.C., Thomas, G., Kinzler, K.W., Vogelstein, B. (1990) Identification of a chromosome 18q gene that is altered in colorectal cancers. *Science*, 247, 49-56.
- Feinberg P, Vogelstein B. (1983) A technique for radiolabelling DNA restriction enzyme fragments to high specific activity. *Analytical Biochem*. 137, 266-267.
- Feit, H., Silbergleit, A., Schneide, L.B., Gutierrez, J.A., Fitoussi, R.P., Reyes, C., Rouleau, G.A., Brais, B., Jackson, C.E., Beckmann, J.S., & Seboun, E. (1998) Vocal cord and pharyngeal weakness with autosomal dominant distal myopathy: clinical description and gene localization to 5q31. *Am J Hum Genet*, 63(6), 1732-1742.
- Felice, K.J., Meredith, C., Binz, N., Butler, A., Jacob, R., Akkari, P., Hallmayer, J., & Laing, N. (1999) Autosomal distal myopathy not linked to the known distal myopathy loci. *Neuromuscul Disord*, 9, 59-65.
- Fischer, S.G., & Lerman, L.S. (1983) DNA fragments differing by by single base-pair substitutions are seperated in denaturing gradient gels: correspondence with melting theory. *Proc Natl Acad Sci, USA*, 80, 1579-1584.
- Fisher, R.A. (1912) On absolute criterion for fitting frequency curves. *Mess Math*, 41, 155-160.
- Fisher, R.A. (1922a) On the mathematical foundations of theoretical statistics. *Phil Trans R Soc, A222*, 309-368.
- Fisher, R.A. (1922b) The systematic location of genes by means of crossover observations. *Am Naturalist*, 56, 406-411.



- Fougerousse, F., Durand, M., Suel, L., Pourquoi, O., Delezoide, A.L., Romero, N.B., Abitbol, M., & Beckmann, J.S. (1998) Expression of genes (CAPN3, SGCA, SGCB, and TTN) involved in progressive muscular dystrophies during early human development. *Genomics*, 48(2), 145-156.
- Frank, J.P., Harati, Y., Butler, I.J., Nelson, T.E., & Scott, C. I. (1980) Central core disease and malignant hyperthermia syndrome. *Ann. Neurol*, 7, 11-17.
- Freeman BC, States JC. (1991) An STS in human skeletal alpha actin gene. *Nucleic Acids Res*, 19, 5086.
- Fried, K., Arlozorov, A., & Spira, R. (1975) Autosomal recessive oculopharyngeal muscular dystrophy. *J Med Genet*, 12(4), 416-441.
- Friend, S., Bernards, R., Rogelj, S., Weinberg, R.A., Rapport, J.M., Albert, D.M., & Fryja, T.P. (1986) A human DNA segment with properties of the gene that predisposes to retinoblastoma and osteosarcoma. *Nature*, 323, 643-646.
- Friend, S., Horowitz, J., Gerber, M., Wang, X-F., Bogenmann, E., Li, F., & Weinberg, R. (1987) Deletions of a DNA sequence in retinoblastoma and mesenchymal tumours: organization of the sequence and its encoded protein. *Proc Natl Acad Sci, USA*, 84, 9059-9063.
- Fu, Y-H., Kuhl, D.P., Pizzuti, A., Pieretti, M., Sutcliffe, J.S., Richards, S., Verkerk, A.J., Holden, J.J., Fenwick, R.G. Jr., Warren, S.T., Oostra, B.A., Nelson, D.L., & Caskey, C.T. (1991) Variation of the CGG repeat at the fragile X site results in genetic instability: resolution of the Sherman paradox. *Cell*, 67(6), 1047-1058.
- Fukuhara, N., Kumamoto, T., Tsubaki, T., Mayuzumi, T., & Nitta, H. (1982) Oculopharyngeal muscular dystrophy and distal myopathy. Intrafamilial difference in the onset and distribution of muscular involvement. *Acta Neurol Scand*, 65(5), 458-467.
- Fukuyama, Y., Osawa, M., & Suzuki, H. (1981) Congenital progressive muscular dystrophy of the Fukuyama type - clinical, genetic and pathological considerations. *Brain Dev*, 3, 1-29.
- Galassi, G., Rowland, L.P., Hays, A.P., Hopkins, L.C., & DiMauro, S. (1987) High serum levels of creatine kinase: asymptomatic prelude to distal myopathy. *Muscle & Nerve*, 10(4), 346-350.
- GB4 GeneMap'98. (1998-1999) The National Centre For Biotechnology Information (NCBI) Available <http://www.ncbi.nlm.nih.gov/> (1999, December 1).

- Geisterfer-Lowrance, A.A., Kass, S., Tanigawa, G., Vosberg, H.P., McKenna, W., Seidman, C.E., & Seidman, J.G. (1990) A molecular basis for familial hypertrophic cardiomyopathy: a beta cardiac myosin heavy chain gene missense mutation. *Cell*, 62(5), 999-1006.
- Gelehrter, T.D., & Collins, F.S. (1990) *Principles of Medical Genetics*. Baltimore, USA: Williams and Wilkins.
- GeneBridge RH panel. (1996) *GeneBridge 4 RH panel*. Available <http://www.hgmp.mrc.ac.uk/Registered/Biology/descriptions/genebridge.html> (1999, November, 6)
- Généthon Human Genome Research Centre. (1993-1996) Genetic map Available [http://www.genethon.fr/genethon\\_en.html](http://www.genethon.fr/genethon_en.html) (1999, December 18).
- Genome Database (GDB) (1993-1999) Available <http://gdbwww.gdb.org/> (1999, December 1).
- George, E.L., Ober, M.B., & Emerson, C.P. Jr. (1989) Functional domains of the *Drosophila melanogaster* muscle myosin heavy-chain gene are encoded by alternatively spliced exons. *Mol Cell Biol*, 9(7), 2957-2974.
- Geraghty, M.T., Brody, L.C., Martin, L.S., Marble, M., Kearns, W., Pearson, P., Monaco, A.P., Lehrach, H., & Valle, D. (1993) The isolation of cDNAs from OATL1 at Xp11.2 using a 480-kb YAC. *Genomics* 16, 440-446.
- Gilboa, N., & Swanson, J.R. (1976) Serum creatine phosphokinase in normal newborns. *Arch Dis Child*, 51(4), 283-285.
- Gilbert JR, Stajich JM, Wall S, Carter SC, Qiu H, Vance JM, Stewart CS, Speer MC, Pufky J, Yamaoka LH, Rozear, M.; Samson F, Fardeau M, Roses AD, Pericak-Vance MA. (1993) Evidence for heterogeneity in facioscapulohumeral muscular dystrophy (FSHD). *Am J Hum Genet*, 53(2), 401-408.
- Gilchrist, J.M., Pericak-Vance, M., Silverman, L., & Roses, A.D. (1988) Clinical and genetic investigation in autosomal dominant limb-girdle muscular dystrophy. *Neurol*, 38(1), 5-9.
- Gish, W., & States, D.J. (1993) Identification of protein coding regions by database similarity search. *Nat Genet*, 3, 266-269.
- Gitschier, J., Wood, W.I., Goralka, T.M., Wion, K.L., Chen, E.Y., Eaton, D.H., Vehar, G.A., Capon, D.J., & Lawn, R.M. (1984) Characterization of the human factor VIII gene. *Nature*, 312(5992), 326-330.

- Goas, J.Y., Leroy, J.P., Mocquard, Y., & Rouhart, F. (1991) A case of mitochondrial myopathy in a family with oculo-pharyngeal myopathy. *Rev. Neurol. (Paris)*, 147(6-7), 536-537. (English abstract)
- Goebel, H.H., Anderson, J.R., Hubner, C., Oexle, K., & Warlo, I. (1997) Congenital myopathy with excess of thin myofilaments. *Neuromuscul Disord*, 7, 160-168.
- Gospe, S.M.(Jr.), Lazaro, R.P., Lava, N.S., Grootsholten, P.M., Scott, M.O., & Fischbeck, K.H. (1989) Familial X-linked myalgia and cramps: a nonprogressive myopathy associated with a deletion in the dystrophin gene. *Neurol*, 39(10), 1277-1280.
- Gowers, W.R. (1902) A lecture on myopathy and a distal form. *Br Med J*, 2, 89-92.
- Gray, M.R., Colot, H.V., Guarente, L., & Rosbash, M. (1982) Open reading frame cloning: identification, cloning and expression of open reading frame DNA. *Proc Natl Acad Sci, USA*, 79, 6598-6602.
- Green, J., & Montasser, M. (1988) HLA haplotype discordance. *Biometrics*, 44(4), 941-950.
- Grewal, R.P., Cantor, R., Turner, G., Grewal, R.K., & Detera-Wadleigh, S.D., (1998) Genetic mapping and haplotype analysis of oculopharyngeal muscular dystrophy. *Neuroreport*, 9(6), 961-965.
- Griggs, R.C., & Markesbery, W.R. (1994) Distal myopathies. In A.G. Engel & C. Franzini-Armstrong (Eds.) *Myology: Basic and Clinical* (2nd Ed.) (pp. 1246-1257). New York: McGraw-Hill, Inc.
- Griggs, R.C., Askanas, V., DiMauro, S., Engel, A., Karpatis, G., Mendell, J.R., & Rowland, L.P. (1995) Inclusion body myositis and myopathies. *Ann Neurol*, 38(5), 705-713.
- Grimm, T., Meng, G., Liechti-Gallati, S., Bettecken, T., Muller, C.R., & Muller, B. (1994) On the origin of deletions and point mutations in Duchenne muscular dystrophy: most deletions and point mutations arise in oogenesis and most point mutations result from events in spermatogenesis. *J Med Genet*, 31, 183-186.
- Grove, W.M. & Andreasen, N.C. (1982) Simultaneous tests of many hypotheses in exploratory research. *J Nerv Ment Dis*, 170(1), 3-8.
- Gruen, M., & Gautel, M. (1999) Mutations in beta-myosin S2 that cause familial hypertrophic cardiomyopathy (FHC) abolish the interaction with the regulatory domain of myosin-binding protein-C. *J Mol Biol*, 286(3), 933-949.

- Gyapay, G., Morissette, J., Vignal, A., Dib, C., Fizames, C., Millasseau, P., Marc, S., Bernardi, G., Lathrop, M., & Weissenbach, J. (1994) The 1993-94 Genethon human genetic linkage map. *Nat Genet.*, 7, 246-339.
- Gyapay, G., Schmitt, K., Fizames, C., Jones, H., Vega-Czarny, N., Spillet, D., Muselet, D., Prud'Homme, J.F., Dib, C., Auffray, C., Morissette, J., Weissenbach, J., & Goodfellow, P.N. (1996) A radiation hybrid map of the human genome. *Hum Mol Genet*, 5(3), 339-346.
- Haaf, T., & Ward, D.C. (1994) Structural analysis of alpha-satellite DNA and centromere proteins using extended chromatin and chromosomes. *Hum Mol Genet*, 3(5), 697-709.
- Haines, J.L., Terwedow, H.A., Burgess, K., Pericak-Vance, M.A., Rimmler, J.B., Martin, E.R., Oksenberg, J.R., Lincoln, R., Zhang, D.Y., Banatao, D.R., Gatto, N., Goodkin, D.E., & Hauser, S.L. (1998) Linkage of the MHC to familial multiple sclerosis suggests genetic heterogeneity. *Hum Mol Genet*, 7(8), 1229-1234.
- Haldane, J.B.S. (1919) The combination of linkage values and the calculation of distances between the loci of linked factors. *J Genet*, 8, 299-309.
- Hall, J.M., Lee, M.K., Newman, B., Morrow, J.E., Anderson, L.A., Huey, B., & King, M.C. (1990) Linkage of early-onset familial breast cancer to chromosome 17q21. *Science*, 250(4988), 1684-1689.
- Haravuori, H., Makela-Bengsi, P., Udd, B., Pulkkinen, J., Partanen, L., Somer, H., & Peltonen, L. (1998a) Assignment of the tibial muscular dystrophy locus to chromosome 2q31. *Am J Hum Genet*, 62, 620-626.
- Haravuori, H., Makela-Bengsi, P., & Figlewicz, D.A. (1998b) Tibial muscular dystrophy and late onset distal myopathy are linked to the same locus on chromosome 2q. *Neurol*, 50, A186.
- Harbury, P.B., Kim, P.S., & Alber, T. (1994) Crystal structure of an isoleucine-zipper trimer. *Nature*, 371(6492), 80-83.
- Harley, H.G., Walsh, K.V., Rundle, S., Brook, J.D., Sarfarazi, M., Koch, M.C., Floyd, J.L., Harper, P.S., Shaw, D.J. (1991) Localisation of the myotonic dystrophy locus to 19q13.2-19q13.3 and its relationship to twelve polymorphic loci on 19q. *Hum Genet*, 87, 73-80.
- Hardiman, O., Halperin, J.J., Farrell, M.A., Shapiro, B.E., Wray, S.H., & Brown, R.H. (Jr.) (1993) Neuropathic findings in oculopharyngeal muscular dystrophy. A report of seven cases and a review of the literature. *Arch Neurol*, 50(5), 481-488.

- Harrison, G.A., Humphrey, K.E., Jones, N., Badenhop, R., Guo, G., Elakis, G., Kaye, J.A., Turner, R.J., Grehan, M., Wilton, A.N., Brennecke, S.P., & Cooper, D.W. (1997) A genomewide linkage study of preeclampsia/eclampsia reveals evidence for a candidate region on 4q. *Am J Hum Genet*, 60(5), 1158-1167.
- Hauser, M.A., Horrigan, S.K., Salmikangas, P., Torian, U.M., Viles, K.D., Dancel, R., Tim, R.W., Taivainen, A., Bartoloni, L., Gilchrist, J.M., Stajich, J.M., Gaskell, P.C., Gilbert, J.R., Vance, J.M., Pericak-Vance, M.A., Carpen, O., Westbrook, C.A., & Speer, M.C. (2000) Myotilin is mutated in limb girdle muscular dystrophy 1A. *Hum Mol Genet*, 9(14), 2141-2147.
- Hayashi, K. (1991) PCR-SSCP: a simple and sensitive method for detection of mutations in the genomic DNA. *PCR Methods Appl*, B1, 34-38.
- Hayashi, K. (1994) Manipulation of DNA by PCR. In K. B. Mullis (Ed.), *PCR: The Polymerase Chain Reaction* (pp. 3-13) Boston: Birkhauser.
- Hayashi, Y.K., Chou, F.L., Engvall, E., Ogawa, M., Matsuda, C., Hirabayashi, S., Yokochi, K., Ziober, B.L., Kramer, R.H., Kaufman, S.J., Ozawa, E., Goto, Y., Nonaka, I., Tsukahara, T., Wang, J.Z., Hoffman, E.P., & Arahata, K. (1998) Mutations in the integrin alpha7 gene cause congenital myopathy. *Nat Genet*, 19(1), 94-97.
- Hayes, R., London, W., Seidman, J., & Embree, L. (1963) Oculopharyngeal muscular dystrophy. *New Eng J Med*, 268, 163.
- Hayes, W. (1953) Observations on a transmissible agent determining sexual differentiation in *Bact. coli*. *J Gen Microbiol*, 8, 72.
- Helbling-Leclerc, A., Zhang, X., Topaloglu, H., Cruaud, C., Tesson, F., Weissenbach, J., Tome, F.M.S., Schwartz, K., Fardeau, M., Tryggvason, K., & Guicheney, P. (1995) Mutations in the laminin  $\alpha$ -2 chain gene (LAMA2) cause merosin deficient congenital muscular dystrophy. *Nat Genet*, 11, 216-218.
- Henry, M.D., & Campbell, K.P. (1996) Dystroglycan: an extracellular matrix receptor linked to the cytoskeleton. *Curr Opin Cell Biol*, 8(5), 625-631.
- Herrell, S., Novo, F.J., Charlton, R., & Affara, N.A. (1995) Development and physical analysis of YAC contigs covering 7 Mb of Xp22.3-p22.2. *Genomics*, 25(2), 526-537.
- Hillaire, D., Leclerc, A., Faure, S., Topaloglu, H., Chiannikulchai, N., Guicheney, P., Grinas, L., Legos, P., Philpot, J., Evangelista, T., Routon, M-C., Mayer, M., Pellissier, J-F., Estournet, B., Barois, A., Hentati, F., Feingold, N., Beckmann, J.S., Dubowitz, V., Tome, F.M.S., & Fardeau, M. (1994) Localization of merosin-negative congenital muscular dystrophy to chromosome 6q2 by homozygosity mapping. *Hum Mol Genet*, 3(9), 1657-1661.

- Hoffman, E.P., & Schwartz, L. (1991) Dystrophin and disease. *Mol Aspects Med*, 12(3), 175-194.
- Hochgeschwender, U., Sutcliffe, J.G., & Brennan, M.B. (1989) Construction and screening of a genomic library specific for mouse chromosome 16. *Proc Natl Acad Sci, USA*, 86, 8482-8486.
- Holmans, P. (1993) Asymptomatic properties of affected-sib pair linkage analysis. *Am J Hum Genet*, 52(2), 362-374.
- Horowitz, S.H., & Schmalbruch, H. (1994) Autosomal dominant distal myopathy with desmin storage: a clinicopathologic and electrophysiologic study of a large kinship. *Muscle & Nerve*, 17, 151-160.
- Hoshino, S., Kimura, A., Fukuda, Y., Dohi, K., & Sasazuki, T. (1992) Polymerase chain reaction - single-strand conformation polymorphism analysis of polymorphism in DPA1 and DPB1 genes: a simple, economical, and rapid method for histocompatibility testing. *Hum Immunol*, 33(2), 98-107.
- Houwen, R.H.J., Baharloo, S., Blankenship, K., Raeymaekers, P., Juyn, J., Sandknijl, L.A., & Freimer, N.B. (1994) Genome screening by searching for shared segments: mapping a gene for benign recurrent intrahepatic cholestasis. *Nat Genet*, 8(4), 380-386.
- Hughes, S.M., Cho, M., Karsch-Mizrachi, I., Travis, M., Silberstein, L., Leinwand, L.A., & Blau, H.M. (1993) Three slow myosin heavy chains sequentially expressed in developing mammalian skeletal muscle. *Dev Biol*, 158(1), 183-199.
- Hunt, C.C.J, Eyre, H.J., Akkari, P.A., Meredith, C., Dorosz, S.M., Wilton, S.D., Callen, D.F., Laing, N.G., & Baker E. (1995) Assignment of the beta tropomyosin gene (*TPM2*) to band 9p13 by fluorescence in situ hybridisation. *Cytogenet Cell Genet*, 71, 94-95.
- Hurley, J.H., Mason, D.A., & Matthews, B.W. (1992) Flexible-geometry conformational energy maps for the amino acid residue preceding a proline. *Biopolymers*, 32(11), 1443-1446.
- Huxley, H.E. (1953) Electron microscope studies of the organization of the filaments in striated muscle. *Biochim Biophys Acta*, 12, 387.
- Ikeuchi, T., Asaka, T., Saito, M., Tanaka, H., Higuchi, S., Tanaka, K., Saida, K., Uyama, E., Mizusawa, H., Fukuhara, N., Nonaka, I., Takamori, M., & Tsuji, S. (1997) Gene locus for autosomal recessive distal myopathy with rimmed vacuoles maps to chromosome 9. *Ann Neurol*, 41(4), 432-437.

- Illa, I., Serrano, C., Gallardo, E., Barraquer, L., Lassa, A., Gallano, P., Baiget, M., Bejaoui, K., & Brown, R.H. (1998) Distal anterior compartment myopathy: a new severe dystrophic phenotype linked to chromosome 2p13. *Neurol*, 50, A186.
- Iliaroshkin, S.N., Ivanova-Smolenskaya, I.A., Tanaka, H., Vereshchagin, N.V., Markova, E.D., Poleshchuk, V.V., Lozhnikova, S.M., Sukhorukov, V.S., Limborska, S.A., Slominsky, P.A., Bulayeva, K.B., & Tsuji, S. (1996) Clinical and molecular analysis of a large family with three distinct phenotypes of progressive muscular dystrophy. *Brain*, 119 (6), 1895-1909.
- Ioannou, P.A., Amemiya, C.T., Garnes, J., Kroisel, P.M., Shizuya, H., Chen, C., Batzer, M.A., & de Jong, P. (1994) A new bacteriophage P1 derived vector for the propagation of large human DNA fragments. *Nat Genet*, 6(1), 84-89.
- Ireland, M., English, C., Cross, I., Lindsay, S., & Strachan, T. (1995) Partial trisomy 3q and the mild Cornelia de Lange syndrome phenotype. *J Med Genet*, 32(10), 837-838.
- Izumo, S., Nadal-Ginard, B., & Mahdavi, V. (1986) All members of the MHC multigene family respond to thyroid hormone in a highly tissue-specific manner. *Science*, 231(4738), 597-600.
- Jaenicke, T., Diederich, K.W., Haas, W., Schleich, J., Lichter, P., Pfordt, M., Bach, A., & Vosberg, H.P. (1990) The complete sequence of the human beta-myosin heavy chain gene and a comparative analysis of its product. *Genomics*, 8(2), 194-206.
- Jandreski, M.A., Sole, M.J., & Liew, C-C. (1987) Two different forms of beta myosin heavy chain are expressed in human striated muscle. *Hum Genet*, 77, 127-131.
- Jeffreys, A.J., Wilson, V., & Thein, S.L. (1985) Hypervariable 'minisatellite' regions in human DNA. *Nature*, 314, 67-73.
- Jelinek, W.R., & Schmid, C.W. (1982) Repetitive sequences in eukaryotic DNA and their expression. *Ann Rev Biochem*, 51, 813-44.
- Jobsis, G.J., Bolhuis, P.A., Boers, J.M., Baas, F., Wolterman, R.A., Hensels, G.W., & de Visser, M. (1996a) Genetic localization of Bethlem myopathy. *Neurol*, 46(3), 779-782.
- Jobsis, G.J., Keizers, H., Vreijling, J.P., de Visser, M., Speer, M.C., Wolterman, R.A., Baas, F., & Bolhuis, P.A. (1996b) Type VI collagen mutations in Bethlem myopathy, an autosomal dominant myopathy with contractures. *Nat Genet*, 14(1), 113-115.

- Kajiwara, K., Hahn, L.B., Mukai, S., Travis, G.H., Berson, E.L., & Dryja, T.P. (1991) Mutations in human retinal degeneration slow gene in autosomal dominant retinitis pigmentosa. *Nature*, 354, 480-483.
- Kelly, D.P., & Strauss, A.W. (1994) Inherited cardiomyopathies. *N Engl J Med*, 330(13), 913-919.
- Kiloh, L.G., & Nevin, S. (1951) Progressive dystrophy of the external ocular muscles (ocular myopathy). *Brain*, 74, 115-143.
- Kimura, A., Harada, H., Park, J.E., Nishi, H., Satoh, M., Takahashi, M., Hiroi, S., Sasaoka, T., Ohbuchi, N., Nakamura, T., Koyanagi, T., Hwang, T.H., Choo, J.A., Chung, K.S., Hasegawa, A., Nagai, R., Okazaki, O., Nakamura, H., Matsuzaki, M., Sakamoto, T., Toshima, H., Koga, Y., Imaizumi, T., & Sasazuki, T. (1997) Mutations in the cardiac troponin I gene associated with hypertrophic cardiomyopathy. *Nat Genet*, 16(4), 379-382.
- Kleinsmith, L.J., & Kish, V.M. (1995) *Principles of Cell and Molecular Biology*. New York: HarperCollins College Publishers.
- Kloepfer, H.W., & Yalley, C. (1958) Autosomal recessive inheritance of Duchenne-type muscular dystrophy. *Ann Hum Genet*, 22, 138-143.
- Knapp, M., Seuchter, S.A., & Baur, M.P. (1994) Two-locus disease models with two marker loci: the power of affected-sib pair tests. *Am J Hum Genet*, 55(3), 482-498.
- Knight, S.J., Flannery, A.V., Hirst, M.C., Campbell, L., Christodoulou, Z., Phelps, S.R., Pointon, J., Middleton-Price, H.R., Barnicoat, A., Pembrey, M.E., Holland, J., Oostra, B.A., Bobrow, M., & Davies, K.E. (1993) Trinucleotide repeat amplification and hypermethylation of a CpG island in FRAXE mental retardation. *Cell*, 74(1), 127-134.
- Koenig, M., Hoffman, E., Bertelson, C., Monaco, A., Feener, C., & Knunkel, L. (1987) Complete cloning of the Duchenne muscular dystrophy (DMD) cDNA and preliminary genomic organisation of the DMD gene in normal and affected individuals. *Cell*, 51, 509-517.
- Koenig, M., Beggs, A.H., Moyer, M., Scherpf, S., Heindrich, K., Bettecken, T., Meng, G., Muller, C.R., Lindlof, M., & Kaariainen, H. (1989) The molecular basis for Duchenne versus Becker muscular dystrophy: correlation of severity with type of deletion. *Am J Hum Genet*, 45(4), 498-506.



- Korenberg, J.R., Yang-Feng, T., Schreck, R., & Chen, X.N. (1992) Using fluorescence *in situ* hybridization (FISH) in genome mapping. *Trends Biotechnol*, 10, 27-32.
- Kumar, S. & Bansal, M. (1996) Structural and sequence characteristics of long alpha helices in globular proteins. *Biophys J*, 71(3), 1574-1586.
- Kurnit, D.M. & Seed, B. (1990) Improved genetic selection for screening bacteriophage libraries by homologous recombination *in vivo*. *Proc Natl Acad Sci USA*, 87, 3166-3169.
- Krawczak, M., & Schmidtke, J. (1994) *DNA Fingerprinting*. Oxford, UK: BIOS Scientific Publishers.
- Krause, S., Fakan, S., Weis, K., & Wahle, E. (1994) Immunodetection of poly(A) binding protein II in the cell nucleus. *Exp Cell Res*, 214(1), 75-82.
- Kress, W., Halliger-Keller, B., Grimm, T., Porschke, H., Engelhardt, A., Goebel, H.H., & Muller-Mysok, B. (1998) No evidence for heterogeneity in oculopharyngeal muscular dystrophy. *J Med Genet*, 35(7), 613-614.
- Kronert, W.A., O'Donnell, P.T., Fieck, A., Lawn, A., Vigoreaux, J.O., Sparrow, J.C., & Bernstein, S.I. (1995) Defects in the *Drosophila* myosin rod permit sarcomere assembly but cause flight muscle degeneration. *J Mol Biol*, 249(1), 111-125.
- Kruglyak, L, Daly, M.J., Reeve-Daly, M.P., & Lander, E.S. (1996) Parametric and nonparametric linkage analysis: a unified multipoint approach. *Am J Hum Genet*, 58, 1347-1363.
- Kuhn, E., Fiehn, W., Schroder, J.M., Assmus, H., & Wagner, A. (1979) Early myocardial disease and cramping myalgia in Becker-type muscular dystrophy: a kindred. *Neurol*, 29(8), 1144-1149.
- Kuhn, E., & Schroder, J.M. (1981) A new type of distal myopathy in two brothers. *J Neurol*, 226(3), 181-185.
- Lacomis, D., Kupsky, W.J., Kuban, K.K., & Specht, L.A. (1991) Childhood onset oculopharyngeal muscular dystrophy. *Pediatr Neurol*, 7(5), 382-384.
- Lander, E.S. & Schork, N.J. (1994) Genetic dissection of complex traits. *Science*, 265, 2037-2048.

- Laing, N.G. (1993) Molecular genetics and genetic counselling for Duchenne/Becker muscular dystrophy. In T. Partridge (Ed.), *Molecular and Cell Biology of Muscular Dystrophy*, (pp. 37-84). London: Chapman & Hall.
- Laing, N.G. (1995a) Inherited disorders of contractile proteins in skeletal and cardiac muscle. *Curr Opin Neurol*, 8, 391-396.
- Laing, N.G., Laing, B.A., Meredith, C., Wilton, S.D., Robbins, P., Honeyman, K., Dorosz, S., Kozman, H.M., Mastaglia, F.L., & Kakulas, B.A. (1995b) Autosomal dominant distal myopathy: linkage to chromosome 14. *Am J Hum Genet*, 56, 422-427.
- Lankford, E.B., Epstein, N.D., Fananapazir, L., & Sweeney, H.L. (1995) Abnormal contractile properties of muscle fibers expressing beta-myosin heavy chain gene mutations in patients with hypertrophic cardiomyopathy. *J Clin Invest*, 95(3), 1409-1414.
- Laporte, J., Hu, L.J., Kretz, C., Mandel, J-L., Kioschis, P., Coy, J.F., Klauck, S.M., Poustka, A., & Dahl, N. (1996) A gene mutated in X-linked myotubular myopathy defines a new putative tyrosine phosphatase family conserved from yeast. *Nat Genet*, 13, 175-182.
- Lathrop, G.M., Lalouel, J.M., Julier, C., & Ott, J. (1985) Multilocus linkage analysis in humans: detection of linkage and estimation of recombination. *Am J Hum Genet*, 55(5), 354-361.
- Lee, E.Y-H., To, H., Shew, J-Y., Bookstein, R., Schull, P., & Lee, W-H. (1988) Inactivation of the retinoblastoma susceptibility gene in human breast cancers. *Science*, 241, 218-221.
- Lee, W-H., Bookstein, R., Hong, F., Young, L-J., Shew, J-Y., & Lee, E.Y-H. (1987) Human retinoblastoma susceptibility gene: cloning, identification and sequence. *Science*, 235, 1394-1399.
- Lemmers, R.J., van der Maarel, S.M., van Deutekom, J.C., van der Wielen, M.J., Deidda, G., Dauwerse, H.G., Hewitt, J., Hofker, M., Bakker, E., Padberg, G.W., & Frants, R.R. (1998) Inter- and intrachromosomal subtelomeric rearrangements on 4q35: implications for facioscapulohumeral muscular dystrophy (FSHD) aetiology and diagnosis. *Hum Mol Genet*, 7(8), 1207-1214.
- Liechti-Gallati, S., Muller, B., Grimm, T., Kress, W., Muller, C., Boltshauser, E., Moser, H., & Braga, S. (1991) X-linked centronuclear myopathy: mapping the gene to Xq28. *Neuromuscul Disord*, 1(4), 239-245.

- Li, T., Arranz, M., Aitchison, K.J., Bryant, C., Liu, X., Kerwin, R.W., Murray, R., Sham, P., & Collier, D.A. (1998) Case-control, haplotype relative risk and transmission disequilibrium analysis of a dopamine D2 receptor functional promoter polymorphism in schizophrenia. *Schiz Res*, 32(2), 87-92.
- Liang, P., & Pardee, A.B. (1992) Differential display of eukaryotic messenger RNA by means of the polymerase chain reaction. *Science*, 257, 967-971.
- Lim, L.E., Duclos, F., Broux, O., Bourg, N., Sunada, Y., Allamand, V., Meyer, J., Richard, I., Moomaw, C., Slaughter, C., Tome, F.M.S., Fardeau, M., Jackson, C.E., Beckmann, J.S., & Campbell, K.P. (1995) Beta-sarcoglycan: characterisation and role in limb-girdle muscular dystrophy linked to 4q12. *Nat Genet*, 11(3), 257-265.
- Linoli, G., Tomelleri, G., & Ghezzi, M. (1991) Oculopharyngeal muscular dystrophy. Description of a case with involvement of the central nervous system. *Pathologica*, 83(1085), 325-334.
- Liu, L., Forsell, C., Lilius, L., Axelman, K., Corder, E.H., & Lannfelt, L. (1996) Allelic association but only weak evidence for linkage to the apolipoprotein E locus in late-onset Swedish Alzheimer families. *Am J Hum Genet*, 67(3), 306-311.
- Liu, J., Aoki, M., Illa, I., Wu, C., Fardeau, M., Angelini, C., Serrano, C., Urtizberea, J.A., Hentati, F., Hamida, M.B., Bohlega, S., Culper, E.J., Amato, A.A., Bossie, K., Oeltjen, J., Bejaoui, K., McKenna-Yasek, D., Hosler, B.A., Schurr, E., Arahata, K., de Jong, P.J., & Brown, R.H. (Jr.) (1998) Dysferlin, a novel skeletal muscle gene, is mutated in Miyoshi myopathy and limb girdle muscular dystrophy. *Nat Genet*, 20(1), 31-36.
- Liu, Q., & Sommer, S. (1995) Restriction endonuclease fingerprinting (REF): a sensitive method for screening mutations in long contiguous segments of DNA. *BioTechniques*, 18, 470-477.
- Lohmann D.R., (1999) RB1 gene mutations in retinoblastoma. *Hum Mutat* 14(4), 283-288.
- Lonjou, C., Collins, A., & Morton, N.E. (1999) Allelic association between marker loci. *Proc Natl Acad Sci, USA*, 96(4), 1621-1626.
- Lovett, M., Kere, J., & Hinton, L.M. (1991) Direct selection: a method for the isolation of cDNAs encoded by large genomic regions. *Proc Natl Acad Sci, USA*, 88, 9628-9632.
- Lupas, A. (1996) Coiled coils: new structures and new functions. *Trends Biochem Sci*, 21(10), 375-382.

- MacArthur, M.W., & Thornton, J.M. Influence of proline residues on protein conformation. (1991) *J Mol Biol*, 218(2), 397-412.
- MacDonald, M.E., Lin, C., Srinidhi, L., Bates, G., Altherr, M., Whaley, W.L., Lehrach, H., Wasmuth, J., & Gusella, J.F. (1991) Complex patterns of linkage disequilibrium in the Huntington disease region. *Am J Hum Genet*, 49, 723-734.
- Magee, K.R., & DeJong, R.N. (1965) Hereditary distal myopathy with onset in infancy. *Arch Neurol*, 13(4), 387-390.
- Mahjneh, I., Vannelli, G., Bushby, K., & Marconi, G.P. (1992) A large inbred Palestinian family with two forms of muscular dystrophy. *Neuromuscul Disord*, 2(4), 277-283.
- Markesbery, W.R., Griggs, R.C., Leach, R.P., & Lapham, L.W. (1974) Late onset hereditary distal myopathy. *Neurol*, 23, 127-134.
- Markesbery, W.R., Griggs, R.C., & Herr, B. (1977) Distal myopathy: electron microscopic and histochemical studies. *Neurol*, 27(8), 727-735.
- Markesbery, W.R., & Griggs R.C. (1986) Distal myopathies. In A. K. Engel & B. Q. Banker (Eds.), *Myology* (pp. 1313-1325) New York: McGraw-Hill.
- Martinsson, T., Darin, N., Kyllerman, M., Oldfors, A., Hallberg, B., Wahlstrom, J. (1999) Dominant hereditary inclusion-body myopathy gene (*IBM3*) maps to chromosome region 17p13.1. *Am J Hum Genet.*, 64(5), 1420-1426.
- Martinsson, T., Oldfors, A., Darin, N., Berg, K., Tajsharghi, H., Kyllerman, M., & Wahlstrom, J. (2000) Autosomal dominant myopathy: missense mutation (Glu-706 --> Lys) in the myosin heavy chain IIa gene. *Proc Natl Acad Sci USA*, 97(26), 14614-14619.
- Mastaglia, F.L. (1991) *Genetic myopathies*. In M. Swash & J. Oxbury (Eds.) *Clinical Neurology*, (pp. 1300-1301). Edinburgh: Churchill Livingstone.
- Mastaglia, F.L., & Laing, N.G. (1996) Investigation of muscle disease. *J Neurol Neurosurg Psychiatry*, 60, 256-274.
- Matthews, P.B. (1981) Evolving views on the internal operation and functional role of the muscle spindle. *J Physiol*, 320, 1-30.
- Matsuoka, R., Chambers, A., Kimura, M., Kanda, N., Bruns, G., Yoshida, M., & Takao, A. (1989) Molecular cloning and chromosomal localization of a gene coding for human cardiac myosin heavy chain. *Am J Med Genet*, 29, 369-376.

- McKusick, V.A. (1973) Phenotypic diversity of human diseases resulting from allelic series. *Am J Hum Genet*, 25(4), 446-456.
- McKusick, V.A. (1990) Foreward. In T.D. Gelehrter & F.S. Collins, *Principles of Medical Genetics* (p. v). Baltimore: Williams & Wilkins.
- McKusick, V.A. (1992) *Mendelian inheritance in man: catalogues of autosomal dominant, autosomal recessive and X-linked phenotypes* (10th ed.). Baltimore: John Hopkins University Press.
- McLachlan, A.D., & Karn, J. (1982) Periodic charge distributions in the myosin rod amino acid sequence match cross-bridge spacings in muscle. *Nature*, 299(5880), 226-231.
- Meola, G., Sansone, V., Rotondo, G., Tome, F.M.S., & Bouchard J.P. (1997) Oculopharyngeal muscular dystrophy in Italy. *Neuromuscul Disord*, 7(Suppl. 1), S53-S56.
- Meredith, C., Laing, B.A., Wilton, S.D., Mastaglia, F.L., Robbins, P., Kozman, H., Honeyman, K., Dorosz, S., Kakulas, B.A., & Laing, N.G. (1994a) Autosomal dominant distal myopathy: linkage on chromosome 14. *Muscle & Nerve*, Suppl. 1, S89. (VIIth International Congress on Neuromuscular Diseases, Kyoto, Japan 10-15 July).
- Meredith, C., Laing, B.A., Wilton, S.D., Mastaglia, F.L., Robbins, P., Kozman, H., Honeyman, K., Dorosz, S., Kakulas, B.A., & Laing, N.G. (1994b) Autosomal dominant distal myopathy: linkage to chromosome 14. In *Human Genome Data Base* (GDB), GOO-378-277. (The second international chromosome 14 workshop; Oxford, September 1-3).
- Merriman, T., Twells, R., Merriman, M., Eaves, I., Cox, R., Cucca, F., McKinney, P., Shield, J., Baum, D., Bosi, E., Pozzilli, P., Nistico, L., Buzzetti, R., Joner, G., Ronningen, K., Thorsby, E., Undlien, D., Pociot, F., Nerup, J., Bain, S., Barnett, A., & Todd, J. (1997) Evidence by allelic association-dependent methods for a type 1 diabetes polygene (IDDM6) on chromosome 18q21. *Hum Mol Genet*, 6(7), 1003-1010.
- Messina, D.N., Speer, M.C., Pericak-Vance, M.A., & McNally, E.M. (1997) Linkage of familial dilated cardiomyopathy with conduction defect and muscular dystrophy to chromosome 6q23. *Am J Hum Genet*, 61(4), 909-917.
- Milhorat, A.T., & Wolff, H.G. (1943) Studies in diseases of muscle. XII. Progressive muscular dystrophy of atrophic distal type; report of autopsy. *Arch Neurol Psychiatry*, 49, 655-664.
- Miller, S., Dykes, D., & Polesky, H. (1988) A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Res*, 16, 1215.

- Minetti, C., Sotgia, F., Bruno, C., Scartezzini, P., Broda, P., Bado, M., Masetti, E., Mazzocco, M., Egeo, A., Donati, M.A., Volonte, D., Galbiati, F., Cordone, G., Bricarelli, F.D., Lisanti, M.P., & Zara, F. (1998) Mutations in the caveolin-3 gene cause autosomal dominant limb-girdle muscular dystrophy. *Nat Genet*, 18(4), 365-368.
- Mitrani-Rosenbaum, S., Argov, Z., Blumenfeld, A., Siedman, C.E., & Siedman, J.G. (1996) Hereditary inclusion body myopathy maps to chromosome 9p1-q1. *Hum Mol Genet*, 5, 159-163.
- Miyoshi, K., Kawai, H., Iwasa, M., Kuska, K., & Nishino, H. (1986) Autosomal recessive distal muscular dystrophy as a new type of progressive muscular dystrophy: seven cases in eight families, including an autopsied case. *Brain*, 109, 31-54.
- Mizuno, Y., Noguchi, S., Yamamoto, H., Yoshida, M., Suzuki, A., Hagiwara, Y., Hayashi, Y.K., Arahata, K., Nonaka, I., Hirai, S., & Ozawa, E. (1994) Selective defect of sarcoglycan complex in severe childhood autosomal recessive muscular dystrophy muscle. *Biochem Biophys Res Commun*, 203(2), 979-983.
- Mogensen J., Klausen, I.C., Pedersen, A.K., Egeblad, H., Bross, P., Kruse, T.A., Gregersen, N., Hansen, P.S., Baandrup, U., & Borglum, A.D. (1999) Alpha-cardiac actin is a novel disease gene in familial hypertrophic cardiomyopathy. *J Clin Invest*, 103(10), R39-43.
- Mohire, M.D., Tandan, R., Fries, T.J., Little, B.W., Pendlebury, W.W., & Bradley, W.G. (1998) Early-onset benign autosomal dominant limb-girdle myopathy with contractures (Bethlem myopathy). *Neurol* 38(4), 573-580.
- Monaco, A.P., Neve, R.L., Colletti-Feener, C., Bertelson, C.J., Kurnit, D.M., & Kunkel, L.M. (1986) Isolation of candidate cDNAs for portions of the Duchenne muscular dystrophy gene. *Nature*, 323, 646-650.
- Monaco, A.P., Bertelson, C., Liechti-Gallati, S., Moser, H., & Kunkel, L. (1988) An explanation for the phenotypic differences between patients bearing partial deletions of the DMD locus. *Genomics*, 2, 90-95.
- Moncman, C.L., & Wang, K. (1999) Functional dissection of nebulin demonstrates actin binding of nebulin-like repeats and Z-line targeting of SH3 and linker domains. *Cell Motil Cytoskeleton*, 44(1), 1-22.
- Montanaro, V., Casamassimi, A., D'Urso, M., Yoon, J.Y., Freije, W., Schlessinger, D., Muenke, M., Nussbaum, R.L., Saccone, S., Maugeri, S., Santoro, A.M., Motta, S., & Valle, G.D. (1991) *In situ* hybridization to cytogenetic bands of yeast artificial chromosomes covering 50% of human Xq24-Xq28 DNA. *Am J Hum Genet*, 48, 183-194.

- Moreira, E.S., Vainzof, M., Marie, S.K., Sertie, A.L., Zatz, M., & Passos-Bueno, M.R. (1997) The seventh form of autosomal recessive limb-girdle muscular dystrophy is mapped to 17q11-12. *Am J Hum Genet*, 61(1), 151-159.
- Morrison, N.A., Qi, J.C., Tokita, A., Kelly, P.J., Crofts, L., Nguyen, T.V., Sambrook, P.N., & Eisman, J.A. (1994) Prediction of bone density from vitamin D receptor alleles. *Nature*, 367(6460), 284-287.
- Morton, N.E. (1955) Sequential test for the detection of linkage. *Am J Hum Genet*, 7, 277-318.
- Morton, N.E. (1991) Parameters of the human genome. *Proc Natl Acad Sci USA*, 88, 7474-7476.
- Morton, N.E. (1995) LODs past and present. *Genetics*, 140(1), 7-12.
- Morton, N.E., & Lio, P. (1997) Oligogenic linkage and map integration. In I-H. Pawlowitzki, J.H. Edwards, & E.A. Thompson (Eds.), *Genetic Mapping of Disease Genes* (pp. 17-21). San Diego: Academic Press Inc.
- Morton, N.E. (1998) Significance levels in complex inheritance. *Am J Hum Genet*, 62(3), 690-697.
- Morton, N.E. & Collins, A. (1998) Tests and estimates of allelic association in complex inheritance. *Proc Natl Acad Sci, USA*, 95(19), 11389-11393.
- Muller-Seitz, M., Kaupmann, K., Labeit, S., & Jockusch, H. (1993) Chromosomal localization of the mouse titin gene and its relation to "muscular dystrophy with myositis" and nebulin genes on chromosome 2. *Genomics*, 18(3), 559-561.
- Mulley, J.C., Kozman, H.M., Phillips, H.A., Gedeon, A.K., McCure, J.A., Iles, D.E., Gregg, R.G., Hogan, K., Couch, F.J., MacLennan, D.H., & Haan, E.A. (1993) Refined genetic localization for central core disease. *Am J Hum Genet*, 52, 398-405.
- Mullis, K., Falcorna, F., Scharf, F., Snikl, R., Horn, G., & Erlich, H. (1986) Specific amplification of DNA in vitro: the polymerase chain reaction. *Cold Spring Harbor Symposium on Quantitative Biology*, 51, 260.
- Murakami, N., Ihara, Y., & Nonaka, I. (1995) Muscle fiber degeneration in distal myopathy with rimmed vacuole formation. *Acta Neuropathol (Berl)*, 89(1), 29-34.
- Murray, J.C., Duyk, G.M., Sheffield, V.C., Weber, J.L., Buetow, K.H., Weir, R.F., & Newkirk, N. (1993) CHLC REPORT, 1(1), 1-18. Available <http://lpg.nci.nih.gov/CHLC/> (1999, December 1).

- Murray, J.C., Duyk, G.M., Sheffield, V.C., Weber, J.L., Buetow, K.H., Weir, R.F., & Newkirk, N. (1994) CHLC Version 2.5 (V2.5) Maps: Special Edition. Available <http://lpg.nci.nih.gov/CHLC/> (1999, December 1).
- Myers, R.M., Fischer, S.G., Lerman, L.S., & Maniatis, T. (1985) Detection of single base substitutions by ribonuclease cleavage at mismatches in RNA:DNA heteroduplexes. *Science*, 230, 1242-1246.
- Nagel, R.L. (1999) Will the genetic individualization of disease force a new paradigm? *C R Acad Sci III*, 322(1), 1-4.
- Nakamura, Y., Leppert, M., O'Connell, P., Wolff, R., Holm, T., Culver, M., Martin, C., Fujimoto, E., Hoff, M., Kumlin, E., & White, R. (1987) Variable number tandem repeat (VNTR) markers for human gene mapping. *Science*, 235, 1616-1622.
- Neitzel, H. (1986) A routine method for the establishment of permanent growing lymphoblastoid cell lines. *Hum Genet*, 73, 320-326.
- Nelson, D.L., Ledbetter, S.A., Corbo, L., Victoria, M.F., Ramirez-Solis, R., Webster, D.H., Ledbetter, D.H., & Caskey, C.T. (1989) *Alu* polymerase chain reaction: a method for rapid isolation of human-specific sequences from complex DNA sources. *Proc Natl Acad Sci USA*, 86, 6686-6690.
- Nemeth, A., Krause, S., Blank, D., Jenny, A., Jeno, P., Lustig, A., & Wahle, E. (1995) Isolation of genomic and cDNA clones encoding bovine poly(A) binding protein II. *Nucleic Acids Res*, 23(20), 4034-4041.
- Newman, B., Austin, M.A., Lee, M., & King, M.C. (1988) Inheritance of human breast cancer: evidence for autosomal dominant transmission in high-risk families. *Proc Natl Acad Sci USA*, 85(9), 3044-3048.
- Nicholls, E.M., & Stark, A.E. (1971) Bayes' theorem. *Med J Aust*, 2(26), 1335-1339.
- Nigro, V., de Sa Moreira, E., Piluso, G., Vainzof, M., Belsito, A., Politano, L., Puca, A.A., Passos-Bueno, M.R., & Zatz, M. (1996) Autosomal recessive limb-girdle muscular dystrophy, LGMD2F, is caused by a mutation in the delta-sarcoglycan gene. *Nat Genet*, 14(2), 195-198.
- NIH/CEPH Collaborative Mapping Group. (1992) A comprehensive genetic linkage map of the human genome. *Science*, 258, 62-86.
- Nishimura, M., Miyamoto, K., Motoyoshi, Y., Sugie, H., & Tanabe, H. (1991) Autosomal recessive oculopharyngeal "muscular dystrophy" - clinical features and association with reduced activity of myophosphorylase. *Rinsho Shinkeigaku*, 31(4), 383-390. (English abstract)



- Nizetic D, Gellen L, Hamvas RM, Mott R, Grigoriev A, Vatcheva R, Zehetner G, Yaspo ML, Dutriaux A, & Lopes C. (1994) An integrated YAC-overlap and 'cosmid-pocket' map of the human chromosome 21. *Hum Mol Genet*, 3(5), 759-770.
- Noguchi, S., McNally, E.M., Ben Othmane, K.B., Hagiwara, Y., Mizuno, Y., Yoshida, M., Yamamoto, H., Bonnemann, C.G., Gussoni, E., Denton, P.H., Kyriakides, T., Middleton, L., Hentati, F., Ben Hamida, M., Nonaka, I., Vance, J.M., Kunkel, L.M., & Ozawa, E. (1995) Mutations in the dystrophin-associated protein gamma-sarcoglycan in chromosome 13 muscular dystrophy. *Science*, 270(5237), 819-822.
- Nonaka, I., Sunohara, N., Ishiura, S., & Satoyoshi, E. (1981) Familial distal myopathy with rimmed vacuole and lamellar (myeloid) body formation. *J Neurol Sci*, 51, 141-155.
- Nonaka, I., Sunohara, N., Satoyoshi, E., Terasawa, K., & Yonemoto, K. (1985) Autosomal recessive distal myopathy with rimmed vacuole formation. *Ann Neurol*, 17, 51-59.
- Nowak, K.J., Wattanasirichaigoon, D., Goebel, H.H., Wilce, M., Pelin, K., Donner, K., Jacob, R.L., Hubner, C., Oexle, K., Anderson, J.R., Verity, C.M., North, K.N., Iannaccone, S.T., Muller, C.R., Nurnberg, P., Muntoni, F., Sewry, C., Hughes, I., Sutphen, R., Lacson, A.G., Swoboda, K.J., Vigneron, J., Wallgren-Pettersson, C., Beggs, A.H., & Laing, N.G. (1999) Mutations in the skeletal muscle alpha-actin gene in patients with actin myopathy and nemaline myopathy. *Nat Genet*, 23(2), 208-212.
- O'Brien, S.J., & Graves, J.A.M. (1991) Report of the Committee on Comparative Gene Mapping. *Cytogenet Cell Genet*, 58, 1124-1151.
- O'Connell, J.R., & Weeks, D.E. (1995) The VITESSE algorithm for rapid exact multilocus linkage analysis via genotype set-recording and fuzzy inheritance. *Nat Genet*, 11(4), 402-408.
- Offer, G., & Knight, P. (1996) The structure of the head-tail junction of the myosin molecule. *J Mol Biol*, 256(3), 407-16.
- Ohya, K., Tachi, N., Kozuka, N., Kon, S., Kikuchi, K., & Chiba, S. (1997) Detection of the mutation in facioscapulohumeral muscular dystrophy patients. *Acta Paediatr Jpn*, 39(1), 92-96.
- Okagaki, T., Weber, F.E., Fischman, D.A., Vaughan, K.T., Mikawa, T., & Reinach, F.C. (1993) The major myosin-binding domain of skeletal muscle MyBP-C (C protein) resides in the COOH-terminal, immunoglobulin C2 motif. *J Cell Biol*, 123(3), 619-626.
- Okayama, H., & Berg, P. (1983) A cDNA cloning vector that permits expression of cDNA inserts in mammalian cells. *Mol Cell Biol*, 2, 151-170.

- Online Mendelian Inheritance in Man. (1998) 164300 Oculopharyngeal muscular dystrophy; OPMD. Available <http://www.ncbi.nlm.nih.gov/omim/> (1999, December 6).
- Ott, J. (1974) Estimation of the recombination fraction in human pedigrees: efficient computation of the likelihood for linkage studies. *Am J Hum Genet*, 26, 588-597.
- Ott J. (1997) Genetic mapping in complex disorders. In I-H., Pawlowitzki, J.H. Edwards & E.A. Thompson (Eds.), *Genetic Mapping of Disease Genes* (pp. 23-30). San Diego: Academic Press Inc.
- Olson, M.V., Hood, L., Cantor, C., & Botstein, D. (1989) A common language for physical mapping of the human genome. *Science*, 245, 1434-1435.
- Orita, M., Iwahana, H., Kanazawa, H., & Sekiya, T. (1989) Detection of polymorphisms of human DNA by gel electrophoresis as single-strand conformation polymorphisms. *Proc Natl Acad Sci USA*, 86, 2766-2770.
- Pratt, M.F., & Meyers, P.K. (1986) Oculopharyngeal muscular dystrophy: recent ultrastructural evidence for mitochondrial abnormalities. *Laryngoscope*, 96(4), 368-373.
- Ozawa, E., Noguchi, S., Mizuno, Y., Hagiwara, Y., & Yoshida, M. (1998) From dystrophinopathy to sarcoglycanopathy: evolution of a concept of muscular dystrophy. *Muscle & Nerve*, 21(4), 421-438.
- Pan, T.C., Zhang, R.Z., Pericak-Vance, M.A., Tandan, R., Fries, T., Stajich, J.M., Viles, K., Vance, J.M., Chu, M.L., & Speer, M.C. (1998) Missense mutation in a von Willebrand factor type A domain of the alpha 3(VI) collagen gene (COL6A3) in a family with Bethlem myopathy. *Hum Mol Genet* 7, (5), 807-812.
- Panegyres, P.K., Mastaglia, F.L., & Kakulas, B.A. (1990) Limb girdle syndromes. Clinical, morphological and electrophysiological studies. *J Neurol Sci*, 95(2), 201-218.
- Partridge, T. (1993) Preface. In T. Partridge (Ed.), *Molecular and Cell Biology of Muscular Dystrophy*, (p. xiii-xvi). London: Chapman & Hall.
- Pallavicini, A., Zimbello, R., Tiso, N., Muraro, T., Rampoldi, L., Bortoluzzi, S., Valle, G., Lafranchi, G., & Danieli, G.A. (1997) The preliminary transcript map of a human skeletal muscle. *Hum Mol Genet*, 6(9), 1445-1450.
- Passos-Bueno, M.R., Terwilliger, J., Ott, J., Vainzof, M., Love, D.R., Davies, K.E., & Zatz, M. (1991a) Linkage analysis in families with autosomal recessive limb-girdle muscular dystrophy (LGMD) and 6q probes flanking the dystrophin-related sequence. *Am J Med Genet*, 38(1), 140-146.

- Passos-Bueno, M.R., Vainzof, M., Pavanello, R de C., Pavanello-Filho, I., Lima, M.A., & Zatz, M. (1991b) Limb-girdle syndrome: a genetic study of 22 large Brazilian families. Comparison with X-linked Duchenne and Becker dystrophies. *J Neurol Sci*, 103(1), 65-75.
- Passos-Bueno, M.R., Moreira, E.S., Vainzof, M., Marie, S.K., & Zatz, M. (1996) Linkage analysis in autosomal recessive limb-girdle muscular dystrophy (AR-LGMD) maps a sixth form to 5q33-34 (LGMD2F) and indicates that there is at least one more subtype of AR LGMD. *Hum Mol Genet*, 5(6), 815-20.
- Passos-Bueno, M.R., Vainzof, M., Moreira, E.S., & Zatz, M. (1999) Seven autosomal recessive limb-girdle muscular dystrophies in the Brazilian population: from LGMD2A to LGMD2G. *Am J Med Genet*, 82(5), 392-398.
- Paterson, B.M., & Bishop, J.O. (1977) Changes in the mRNA population of chick myoblasts during myogenesis *in vitro*. *Cell*, 12, 751-765.
- Pauling, L., Itano, H.A., Singer, S.J., & Wells, I.C. (1949) Sickle cell anaemia, a molecular disease. *Science*, 110, 543-548.
- Pauzner, R., Blatt, I., Mouallem, M., Ben-David, E., Farfel, Z., & Sadeh, M. (1991) Mitochondrial abnormalities in oculopharyngeal muscular dystrophy. *Muscle & Nerve*, 14(10), 947-952.
- Pedrosa-Domellof, F., & Thornell, L.E. (1994) Expression of myosin heavy chain isoforms in developing human muscle spindles. *J Histochem Cytochem*. 42(1):77-88.
- Pelin, K., Ridanpaa, M., Donner, K., Wilton, S., Krishnarajah, J., Laing, N., Kolmerer, B., Millevoi, S., Labeit, S., de la Chapelle, A., & Wallgren-Pettersson, C. (1997) Refined localisation of the genes for nebulin and titin on chromosome 2q allows the assignment of nebulin as a candidate gene for autosomal recessive nemaline myopathy. *Eur J Hum Genet*, 5(4), 229-234.
- Pintado, E., de Diego, Y., Hmadcha, A., Carrasco, M., Sierra, J., & Lucas, M. (1995) Instability of the CGG repeat at the FRAXA locus and variable phenotypic expression in a large fragile X pedigree. *J Med Genet*, 32(11), 907-908.
- Plomin, R., Owen, M.J., & McGuffin, P. (1994) The genetic basis of complex human behaviours. *Science*, 264(5166), 1733-1739.

- Poetter, K., Jiang, H., Hassanzadeh, S., Master, S.R., Chang, A., Dalakas, M.C., Rayment, I., Sellers, J.R., Fananapazir, L., & Epstein, N.D. (1996) Mutations in either the essential or regulatory light chains of myosin are associated with a rare myopathy in human heart and skeletal muscle. *Nat Genet*, 13(1), 63-69.
- Porschke, H., Kress, W., Reichmann, H., Goebel, H.H., & Grimm, T. (1997) Oculopharyngeal muscular dystrophy in a northern German family linked to chromosome 14q, and presenting carnitine deficiency. *Neuromuscul Disord*, 7 (Suppl. 1), S57-S62.
- Poulsen, F.R., & Lowy, J. (1983) Small-angle X-ray scattering from myosin heads in relaxed and rigor frog skeletal muscles. *Nature*, 303(5913), 146-152.
- Preus, M., & Ayme, S. (1983) Formal analysis of dysmorphism: objective methods of syndrome definition. *Clin Genet*, 23, 1-16.
- Ramachandran, G.N., & Lakshminarayanan A.V. (1966) Conformation of side groups in amino acids and peptides. *Biopolymers*, 4(4), 495-497.
- Ranum, L.P., Rasmussen, P.F., Benzow, K.A., Koob, M.D., & Day, J.W. (1998) Genetic mapping of a second myotonic dystrophy locus. *Nat Genet* 19(2), 196-198.
- Redwood, C.S., Moolman-Smook, J.C., & Watkins, H. (1999) Properties of mutant contractile proteins that cause hypertrophic cardiomyopathy. *Cardiovasc Res*, 44(1), 20-36.
- Retzel, E.F., Collett, M.S., & Faras, A.J. (1980) Enzymatic synthesis of deoxyribonucleic acid by the avian retrovirus reverse transcriptase in vitro: optimum conditions required for transcription of large ribonucleic acid templates. *Biochemistry*, 19(3), 513-518.
- Richards, R.I., Holman, K., Friend, K., Kremer, E., Hillen, D., Staples, A., Brown, W.T., Goonewardena, P., Tarleton, J., Schwartz, C., & Sutherland, G. R. (1992) Evidence of founder chromosomes in fragile X syndrome. *Nat Genet*, 1(4), 257-260.
- Richard, I., Broux, O., Allamand, V., Fougereousse, F., Chiannikulchai, N., Bourg, N., Brenguier, L., Devaud, C., Pasturaud, P., Roudaut, C., Hillaire, D., Passos-Bueno, M.-R., Zatz, M., Tischfield, J. A., Fardeau, M., Jackson, C. E., Cohen, D., & Beckmann, J. S. (1995) Mutations in the proteolytic enzyme calpain 3 cause limb-girdle muscular dystrophy type 2A. *Cell*, 81(1), 27-40.
- Richard, I., & Beckmann, J.S. (1995) How neutral are synonymous codon mutations? *Nat Genet*, 10(3), 259.

- Richardson, J.S., & Richardson, D.C. (1988) Helix lap-joints as ion-binding sites: DNA-binding motifs and Ca-binding "EF hands" are related by charge and sequence reversal. *Proteins*, 4(4), 229-239.
- Ridley, M. (1999) *GENOME: The Autobiography of a Species in 23 Chapters*. London: Fourth Estate Limited.
- Riley, J.H., Butler, R., Ogilvie, D., Finniear, R., Jenner, D., Powell, S., Anand, R., Smith, J.C., & Markham, A.F. (1990) A novel, rapid method for the isolation of terminal sequences from yeast artificial chromosome (YAC) clones. *Nucleic Acids Res*, 18(10), 2887-2890.
- Riley, J.H., Ogilvie, D., & Anand, R. (1992) Construction, characterisation and screening of YAC libraries. In R. Anand (Ed.) *Techniques for the Analysis of Complex Genomes* (pp 59-79). London: Academic Press Inc.
- Riordan, J.R., Rommens, J.M., Kerem, B., Alon, N., Rozmahel, R., Grzelczak, Z., Zielanski, J., Lok, S., Plavsic, N., Chou, J.L., Drumm, M.L., Iannuzzi, M.C., Collins, F.S., & Tsui, L-C. (1989) Identification of the cystic fibrosis gene: cloning and characterisation of complementary DNA. *Science*, 245, 1066-1073.
- Risch, N., Spiker, D., Lotspeich, L., Nouri, N., Hinds, D., Hallmayer, J., Kalaydjieva, L., McCague, P., Dimiceli, S., Pitts, T., Nguyen, L., Yang, J., Harper, C., Thorpe, D., Vermeer, S., Young, H., Hebert, J., Lin, A., Ferguson, J., Chiotti, C., Wiese-Slater, S., Rogers, T., Salmon, B., Nicholas, P., & Myers, R.M. (1999) A genomic screen of autism: evidence for a multilocus etiology. *Am J Hum Genet*, 65(2), 493-507.
- Roberds, S., Leturcq, F., Allamand, V., Piccolo, F., Jeanpierre, M., Anderson, R., Lim, L., Lee, J., Tome, F., Romero, N., Fardeau, M., Beckmann, J., Kaplan, J-C., & Campbell, K. (1994) Missense mutations in the adhalin gene linked to autosomal recessive muscular dystrophy. *Cell*, 78, 625-633.
- Rogers, T., Kalaydjieva, L., Hallmayer, J., Petersen, P.B., Nicholas, P., Pingree, C., McMahon, W.M., Spiker, D., Lotspeich, L., Kraemer, H., McCague, P., Dimiceli, S., Nouri, N., Peachy, T., Yang, J., Hinds, D., Risch, N., & Myers, R.M. (1999) Exclusion of linkage to the HLA region in ninety multiplex sibships with autism. *J Autism Dev Disord*, 29(3), 195-201.
- Rosenthal, N. (1987) Identification of regulatory elements of cloned genes with functional assays. *Methods Enzymol*, 152, 704-720.
- Roz, L., Wu, C.L., Porter, S., Scully, C., Speight, P., Read, A., Sloan, P., & Thakker, N. (1996) Allelic imbalance on chromosome 3p in oral dysplastic lesions: an early event in oral carcinogenesis. *Cancer Res*, 56(6), 1228-1231.

- Russell, & Lupas, A.N. (1999) COILS program, version 2.2 with MTK matrix. Available [http://ulrec3.unli.ch/software/COILS\\_form.html](http://ulrec3.unli.ch/software/COILS_form.html) (2000, December 23).
- Sadeh, M., Gadoth, N., Hadar, H., & Ben-David, E. (1993) Vacuolar myopathy sparing the quadriceps. *Brain*, 116 (Pt 1), 217-232.
- Saiki, R.F., Scharf, S., Faloona, F., Mullis, K.B., Horn, G.T., Erlich, H.A., & Arnheim, N. (1985) Enzymatic amplification of beta-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. *Science*, 230, 1350-1354.
- Saiki, R.K. (1989) The Design and Optimisation of the PCR. In H. A. Erlich (Ed.), *PCR Technology: Principles and Applications for DNA Amplification* (pp. 7-16). New York: Stockton Press.
- Saiki, R. K. (1990) Amplification of Genomic DNA. In M. A. Innis, D. H. Gelfand, J. J. Sninsky & T. W. White (Eds.), *PCR Protocols: A Guide to Methods and Applications* (pp. 13-20). San Diego: Academic Press Inc.
- Salmikangas, P., Mykkanen, O.M., Gronholm, M., Heiska, L., Kere, J., & Carp, N.O. (1999) Myotilin, a novel sarcomeric protein with two Ig-like domains, is encoded by a candidate gene for limb-girdle muscular dystrophy. *Hum Mol Genet*, 8(7), 1329-1336.
- Salmon, B., Hallmayer, J., Rogers, T., Kalaydjieva, L., Petersen, P.B., Nicholas, P., Pingree, C., McMahon, W., Spiker, D., Lotspeich, L., Kraemer, H., McCague, P., Dimiceli, S., Nouri, N., Pitts, T., Yang, J., Hinds, D., Myers, R.M., & Risch, N. (1999) Absence of linkage and linkage disequilibrium to chromosome 15q11-q13 markers in 139 multiplex families with autism. *Am J Med Genet*, 88(5), 551-556.
- Sanger, F. (1981) Determination of nucleotide sequences in DNA. *Science*, 214(4526), 1205-1210.
- Sambrook, J., Fritsch, E.F., & Maniatis, T. (1989) *Molecular Cloning: a Laboratory Manual* (2nd Ed) Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.
- Santavuori, P., & Leisti, J. (1980) Muscle, eye and brain disease (MEB) In A.W. Eriksson, H.R. Forsius, H.R. Nevanlinna, P.L. Workman, R.K Norio (Eds.), *Population Structure and Genetic Disorders*, (pp. 647-651). New York: Academic Press.
- Sarfarazi, M., Upadhyaya, M., Padberg, G., Pericak-Vance, M., Siddique, T., Lucotte, G., & Lunt, P. (1989) An exclusion map for facioscapulohumeral (Landouzy-Dejerine) disease. *J Med Genet*, 26, 481-484.

- Sargent, C.A., Affara, N.A., Bentley, E., Pelmeur, A., Bailey, D.M.D., Davey, P., Dow, D., Leversha, M., Aplin, H., Besley, G.T.N., & Ferguson-Smith, M.A. (1993) Cloning of the X-linked glycerol kinase deficiency gene and its identification by sequence comparison to the *Bacillus subtilis* homologue. *Hum Mol Genet*, 2, 97-106.
- Sarkar, G., Yoon, H-S., & Sommer, S.S. (1992a) Screening for mutations by RNA single-strand conformation polymorphism (rSSCP): comparison with DNA-SSCP. *Nucleic Acids Res*, 20, 871-878.
- Sarkar G., Yoon, H-S., & Sommer, S.S. (1992b) Dideoxy fingerprinting (ddF): a rapid and efficient screen for the presence of mutations. *Genomics*, 13, 441-443.
- Satoh, M., Takahashi, M., Sakamoto, T., Hiroe, M., Marumo, F., & Kimura, A. (1999) Structural analysis of the titin gene in hypertrophic cardiomyopathy: identification of a novel disease gene. *Biochem Biophys Res Commun*, 262(2), 411-417.
- Satoyoshi, E., & Kinoshita, M. (1977) Oculopharyngodistal myopathy. *Arch Neurol*, 34, 89-92.
- Satoyoshi, E., Sunohara, N., Nonaka, I. (1998) Distal myopathy with rimmed vacuoles, inclusion-body myositis, and related disorders in Japan. In V. Askanas, G. Serratrice and W. King Engel (Eds.), *Inclusion-body myositis and myopathies* (pp. 244-251). Cambridge, UK: Cambridge University Press.
- Schaffer, A.A. (1996) Faster linkage analysis computations for pedigrees with loops or unused alleles. *Hum Hered*, 46(4), 266-235.
- Schalling, M., Hudson, T.J., Buetow, K.H., & Housman, D.E. (1993) Direct detection of novel expanded trinucleotide repeats in the human genome. *Nat Genet*, 4, 135-139.
- Schellenberg, G.D., Payami, H., Wijsman, E.M., Orr, H.T., Goddard, K.A., Anderson, L., Nemen, s E., White, J.A., Alonso, M.E., Ball, M.J., Kaye, J., Morris, J.C., Chui, H., Sadovnick, A.D., Heston, L.L., Martin, G.M., & Bird, T.D. (1993) Chromosome 14 and late-onset familial Alzheimer disease (FAD). *Am J Hum Genet*, 53(3), 619-628.
- Schlosstein, L., Terasaki, P.I., Bluestone, R., & Pearson, C.M. (1973) High association of an HL-A antigen, W27, with ankylosing spondylitis. *New Eng.J. Med*, 288(14), 704-706.
- Schena, M., Shalon, D., Heller, R., Chai, A., Brown, P.O., & Davis, R.W. (1996) Parallel human genome analysis: microarray-based expression monitoring of 1000 genes. *Proc Natl Acad Sci USA*, 93(20), 10614-1069.

- Schimmel, P.R., & Flory, P.J. (1968) Conformational energies and configurational statistics of copolypeptides containing L-proline. *J Mol Biol*, 34(1), 105-120.
- Schon, E.A., Bonilla, E., & DiMauro, S. (1997) Mitochondrial DNA mutations and pathogenesis. *J Bioenerg Biomembr*, 29(2), 131-149.
- Schotland, D.L., & Rowland, L.P. (1964) Muscular dystrophy. Features ocular myopathy, distal myopathy and myotonic dystrophy. *Arch Neurol*, 10, 433-445.
- Schroder, J.M., Krabbe, B., & Weis, J. (1995) Oculopharyngeal muscular dystrophy: clinical and morphological follow-up study reveals mitochondrial alterations and unique nuclear inclusions in a severe autosomal recessive type. *Neuropathol Appl Neurobiol*, 21(1), 68-73.
- Scoppetta, C., Vaccario, M.L., Casali, C., Di Trapani, G., & Mennuni, G. (1984) Distal muscular dystrophy with autosomal recessive inheritance. *Muscle & Nerve*, 7(6), 478-481.
- Scoppetta C., Casali, C., La Cesa, I., Sermoni, A., Mercuri, B., Pierelli, F., & Vaccario, M.L. (1995) Infantile autosomal dominant distal myopathy. *Acta Neurologica Scandinavica*, 92(2):122-126.
- Scrimgeour, E.,M., & Mastaglia, F.,L. (1984) Oculopharyngeal and distal myopathy: a case study from Papua New Guinea. *Am J Med Genet*, 17(4), 763-771.
- Sears, L.E., Moran, L.S., Kissinger, C., Creasey, T., Sutherland, E., Perry-O'Keefe, H., Roskey, M., & Slatko, B. (1992) Thermal cycle sequencing and alternative manual and automated DNA sequencing protocols using the highly thermostable Vent<sub>R</sub><sup>TM</sup> (exo-) DNA polymerase. *BioTechniques*, 13, 626-636.
- Sellers, J.R. (1999) *Myosins*. Oxford: Oxford University Press.
- Setlow, R.B. (1978) Repair deficient human disorders and cancer. *Nature*, 271(5647), 713-717.
- Serratrice, G. (1998) Evolving concepts of inclusion-body myositis. In V. Askanas, G. Serratrice and W. King Engel (Eds.), *Inclusion-body myositis and myopathies* (pp.81-104) Cambridge, UK: Cambridge University Press.
- Shalling, M., Hudson, T.J., Buetow, K.H., & Housman, D.E. (1993) Direct detection of novel expanded trinucleotide repeats in the human genome. *Nat Genet*, 4, 135-139.



- Shalon, D., Smith, S.J., & Brown, P.O. (1996) A DNA microarray system for analyzing complex DNA samples using two-color fluorescent probe hybridization. *Genome Res*, 6(7), 639-645.
- Sham, P.C., & Curtis, D. (1995) An extended transmission/disequilibrium test (TDT) for multi-allele marker loci. *Ann Hum Genet*, 55(5), 1093-1095.
- Sheffield, V.C., Beck, J.S., Kwitek, A.E., Sandstrom, D.W., & Stone, E.M. (1993) The sensitivity of single-strand conformation polymorphism analysis for the detection of single base substitutions. *Genomics*, 16(2), 325-332.
- Shizuya, H., Birren, B., Kim, U-J., Mancino, V., Slepak, T., Tachiiri, Y., & Simon, M. (1992) Cloning and stable maintenance of 300-kilobase-pair fragments of human DNA in *Escherichia coli* using an F-factor based vector. *Proc Natl Acad Sci USA*, 89(18), 8794-8797.
- Shokeir, M.H., & Kobrinsky, N.L. (1976) Autosomal recessive muscular dystrophy in Manitoba Hutterites. *Clin Genet*, 9(2), 197-202.
- Sindhwani, R., Ismail-Beigi, F., & Leinwand, L.A. (1994) Post-transcriptional regulation of rat alpha cardiac myosin heavy chain gene expression. *J Biol Chem*, 269(5), 3272-3276.
- Sinke, R.J., Carlton, V.E., Juijn, J.A., Delhaas, T., Bull, L., van Berge Henegouwen, G.P., van Hattum, J., Keller, K.M., Sinaasappel, M., Bijleveld, C.M., Knol, I.E., Ploos van Amstel, H.K., Pearson, P.L., Berger, R., Freimer, N.B., & Houwen, R.H. (1997) Benign recurrent intrahepatic cholestasis (BRIC): evidence of genetic heterogeneity and delimitation of the BRIC locus to a 7-cM interval between D18S69 and D18S64. *Hum Genet*, 100(3-4), 382-387.
- Smalley, S.L., Bailey, J.N., Palmer, C.G., Cantwell, D.P., McGough, J.J., Del'Homme, M.A., Asarnow, J.R., Woodward, J.A., Ramsey, C., & Nelson, S.F. (1998) Evidence that the dopamine D4 receptor is a susceptibility gene in attention deficit hyperactivity disorder. *Mol Psychiatry*, 3(5), 427-430.
- Smith, C.A.B. & Stephens, D.A. (1997) Simple likelihood and probability calculations for linkage analysis. In I-H Pawlowitzki, J.H. Edwards & E.A. Thompson (Eds.), *Genetic Mapping of Disease Genes*, (pp 73-76). San Diego: Academic Press Inc.
- Soeda, E., Hou, D-E., Osoegawa, K., Atsuchi, Y., Yamagata, T., Shimokawa, T., Kishida, H., Soeda, E., Okano, S., Chumakov, I., Cohen, D., Raff, M., Gairdner, K., Graw, S.L., Patterson, D., de Jong, P., Ashworth, L.K., Slezak, T., & Carrano, A.V. (1995) Cosmid assembly and anchoring to human chromosome 21. *Genomics*, 25, 73-84.

- Sohn, R.L., Vikstrom, K.L., Strauss, M., Cohen, C., Szent-Gyorgyi, A.G., Leinwand, L.A. (1997) A 29 residue region of the sarcomeric myosin rod is necessary for filament formation. *J Mol Biol*, 266(2), 317-330.
- Soukup, T., & Thornell, L.E. (1999) Unusual intrafusal fibres in human muscle spindles. *Physiol Res*, 48(6), 519-523.
- Somer, H. (1994) *Report of the 25th European Neuro Muscular Centre (ENMC) International Workshop: Distal Myopathies*. Baarn, The Netherlands: ENMC.
- Somer, H., Paetau, A., & Udd, B. (1998) Distal myopathies. In A.E.H. Emery (Ed.), *Neuromuscular Disorders: Clinical and Molecular Genetics* (pp. 181-201) West Sussex: John Wiley & Sons Ltd.
- Southern, E.M. (1975) Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J Mol Biol*, 98, 503-517.
- Speer, M.C., Yamaoka, L.H., Gilchrist, J.H., Gaskell, C.P., Stajich, J.M., Vance, J.M., Kazantsev, A., Lastra, A.A., Haynes, C.S., Beckmann, J.S., Cohen, D., Weber, J. L., Roses, A. D., Pericak-Vance, M. A. (1992) Confirmation of genetic heterogeneity in limb-girdle muscular dystrophy: linkage of an autosomal dominant form to chromosome 5q. *Am J Hum Genet*, 50(6), 1211-1217.
- Speer, M.C., Tandan, R., Rao, P.N., Fries, T., Stajich, J.M., Bolhuis, P.A., Jobsis, G.J., Vance, J.M., Viles, K.D., Sheffield, K., James, C., Kahler, S.G., Pettenati, M., Gilbert, J.R., Denton, P.H., Yamaoka, L.H., Pericak-& Vance, M.A. (1996) Evidence for locus heterogeneity in the Bethlem myopathy and linkage to 2q37. *Hum Mol Genet*, 5(7), 1043-1046.
- Speer, M.C., Vance, J.M., Grubber, J.M., Lennon, G.F., Stajich, J.M., Viles, K.D., Rogala, A., McMichael, R., Chutkow, J., Goldsmith, C., Tim, R.W., & Pericak-Vance, M.A. (1999) Identification of a new autosomal dominant limb-girdle muscular dystrophy locus on chromosome 7. *Am J Hum Genet*, 64(2), 556-562.
- Spielman, R.S., McGinnis, R.E., & Ewens, W.J. (1993) Transmission test for linkage disequilibrium: the insulin gene region and insulin dependent diabetes mellitus (IDDM) *Am J Hum Genet*, 52(3), 506-516.
- Spielman, R.S. & Ewens, W.J. (1998) A sibship test for linkage in the presence of association: the sib transmission/disequilibrium test. *Am J Hum Genet*, 62(2), 450-458.
- Strachan, T., & Read, A.P. (1996) *Human Molecular Genetics*. Oxford: BIOS Scientific Publishers Limited.

- Sternberg, N.L. (1992) Cloning high molecular weight DNA fragments by the bacteriophage P1 system. *Trends Genet*, 8(1), 11-16.
- Stanford Whole Human Genome Radiation Hybrid (RH) Panels. (1995) *The 83 hybrid G3 panel and the 90 hybrid TNG4 RH panel*. Available <http://www-shgc.stanford.edu/> (1998, December 16).
- Straub, R.E., Lehner, T., Luo, Y., Loth, J.E., Shao, W., Sharpe, L., Alexander, J.R., Das, K., Simon, R., Fieve, R.R., Lerer, B., Endicott, J., Ott, J., Gilliam, T.C., & Baron, M. (1994) A possible vulnerability locus for bipolar affective disorder on chromosome 21q22.3. *Nat Genet*, 8(3), 291-296.
- Summers, N.L., & Karplus, M. (1990) Modeling of globular proteins. A distance-based data search procedure for the construction of insertion/deletion regions and Pro-->non-Pro mutations. *J Mol Biol*, 216(4), 991-1016.
- Sunohara, N., Nonaka, I., Kamei, N., & Satoyoshi, E. (1989) Distal myopathy with rimmed vacuole formation. A follow-up study. *Brain*, 112 (Pt 1), 65-83.
- Sunohara, N, Arahata K., Hoffman EP, Yamada H, Nishimiya J, Arikawa E, Kaido M, Nonaka I, Sugita H. (1990) Quadriceps myopathy: forme fruste of Becker muscular dystrophy. *Annals of Neurology*, 28(5), 634-639.
- Strathdee, C.A., Duncan, A.M., & Buchwald, M. (1992a) Evidence for at least four Fanconi anaemia genes including FACC on chromosome 9. *Nat Genet*, 1(3), 196-198.
- Strathdee, C.A., Gavish, H., Shannon, W.R., & Buchwald, M. (1992b) Cloning of cDNAs for Fanconi's anaemia by functional complementation. *Nature*, 356, 763-767.
- Sturtevant, A.H. (1913) The linear arrangement of six sex-linked factors in *Drosophila* as shown by their mode of association. *J Exp Zool*, 14, 43-59.
- Suarez, B.K., Rice, J., & Reich, T. (1978) The generalized sib pair IBD distribution: its use in the detection of linkage. *Ann Hum Genet*, 42(1), 87-94.
- Sumner, D., Crawford, M.A., & Harriman, D.G. (1971) Distal muscular dystrophy in an English family. *Brain*, 94(1), 51-60.

- Swanson, J.M., Sunohara, G.A., Kennedy, J.L., Regino, R., Fineberg, E., Wigal, T., Lerner, M., Williams, L., LaHoste, G.J., & Wigal, S. (1998) Association of the dopamine receptor D4 (DRD4) gene with a refined phenotype of attention deficit hyperactivity disorder (ADHD): a family-based approach. *Mol Psychiatry*, 3(1), 8-41.
- Teh, B.T., Sullivan, A.A., Farnebo, F., Zander, C., Li, F.Y., Strachan, N., Schalling, M., Larsson, C., & Sandstrom, P. (1997) Oculopharyngeal muscular dystrophy (OPMD): report and genetic studies of an Australian kindred. *Clin Genet*, 51(1), 52-55.
- Thage, O. (1965) The "quadriceps syndrome". An electromyographic and histological evaluation. *Acta Neurol Scand*, (Suppl.13, Pt 1), 245-249.
- The Huntington's Disease Collaborative Research Group. (1993) A novel gene containing a tinuiceotide repeat that is expanded and unstable on Huntington's disease chromosomes. *Cell*, 72, 971-983.
- Thierfelder L, Watkins H, MacRae C, Lamas R, McKenna W, Vosberg HP, Seidman JG, & Seidman CE. (1994) Alpha-tropomyosin and cardiac troponin T mutations cause familial hypertrophic cardiomyopathy: a disease of the sarcomere. *Cell*, 77(5), 701-712.
- Thompson, E.A. (1997) Conditional gene identity in affected individuals. In I-H Pawlowitzki, J.H. Edwards & E.A. Thompson (Eds.), *Genetic Mapping of Disease Genes*, (pp 137-146). San Diego: Academic Press Inc.
- Thompson, E.A. (1997) Final Comments. In I-H Pawlowitzki, J.H. Edwards & E.A. Thompson (Eds.), *Genetic Mapping of Disease Genes*, ( pp. 275-277). San Diego: Academic Press Inc.
- Thompson, E.A., & Neel, J.V. (1997) Allelic association and allele frequency distribution as a function of social and demographic history. *Am J Hum Genet*, 60, 197-204.
- Thomson, G. (1995) Mapping disease genes: family-based association studies. *Am J Hum Genet*, 57(2), 487-498.
- Thornell, L-E., Edström, L., Billeter, R., Butler-Browne, G.S., Kjöll, U., & Whalen, R.G. (1984) Muscle fibre type composition in distal myopathy (Welder). An analysis with enzyme and immuno-histochemical, gel electrophoresis and ultrastructural techniques. *J Neurol Sci*, 65, 269-292.
- Toda, T., Kanazawa, I., Nakamura, Y. (1993) Localization of a gene responsible for Fukuyama type congenital muscular dystrophy to chromosome 9q31-33 by linkage analysis. *Human Genome Mapping Workshop 93*, 20. (abstract)

- Todd, J.A., Bell, J.I., & McDevitt, H.O. (1987) HLA-DQ beta gene contributes to susceptibility and resistance to insulin-dependent diabetes mellitus. *Nature*, 329(6140), 599-604.
- Todd, J.A., Bell, J.I., & McDevitt, H.O. (1988) HLA antigens and insulin-dependent diabetes. *Nature*, 333(6175), 710.
- Todd, J.A. & Farrall, M. (1996) Panning for gold: genome-wide scanning for linkage in type 1 diabetes. *Hum Mol Genet*, 5, 1443-1448.
- Tome, F.M.S., & Fardeau, M. (1980) Nuclear inclusions in oculopharyngeal dystrophy. *Arch Neuropath*, 49, 85-87.
- Tome, F., M., S., Leclerc, A., Lopez, N., Chateau, D., Alvarez, R.B., Richardet, J.M., Askanas, V., & Fardeau, M. (1993) Childhood-onset myopathy characterised by cytoplasmic and nuclear inclusions containing 16-18 nm tubulofilaments. *Neurology*, 43 (Suppl. 2), A201-202.
- Tome, F.M.S., Evangelista, T., Leclerc, A., Sunada, Y., Manole, E., Estournet, B., Barois, A., Campbell, K.P., & Fardeau, M. (1994) Congenital muscular dystrophy with merosin deficiency. *C R Acad Sci Paris, Life Sci*, 317(4), 351-357.
- Tome, F.M., Chateau, D., Helbling-Leclerc, A., Fardeau, M. (1997) Morphological changes in muscle fibers in oculopharyngeal muscular dystrophy. *Neuromuscul Disord*, 7 (Suppl 1), S63-S99.
- Thomasson, K.A., & Applequist, J. (1990), Bond-optimized ring closure for proline: comparison of conformations and semiempirical energies with small molecule X-ray structures. *Biopolymers*, 30(3-4), 437-450.
- Towner, P. (1995) Cloning the transcribed portion of the genome. In F. Farzaneh & D.N. Cooper (Eds.), *Functional Analysis of the Human Genome* (p. 69-85). Oxford: BIOS Scientific Publishers Limited.
- Tyler, F.H., & Wintrobe, M.M. (1950) Studies in disorders of muscle I. The problem of progressive muscular dystrophy. *Ann Internal Med*, 32, 72.
- Udd, B., Kaarianen, H., & Somer, H. (1991) Muscular dystrophy with separate clinical phenotype in a large family. *Muscle & Nerve*, 14, 1050-1058.

- Udd, B., Partanen, J., Halonen, P., Falck, B., Hakamies, L., Heikkila, H., Ingo, S., Kalimo, H., Kaariainen, H., Laulumaa, V., Paljarvi, L., Rapola, J., Reunanen, M., Sonninen, V., & Somer, H. (1993) Tibial muscular dystrophy. Late adult-onset distal myopathy in 66 Finnish patients. *Arch Neurol*, 50(6), 604-608.
- Udd, B., Kalimo, H., Nokelainen, P., & Somer, H. (1998) Tibial muscular dystrophy: clinical genetic and morphologic characteristics. In V. Askanas, G. Serratrice and W. King Engel (Eds.), *Inclusion-body myositis and myopathies* (pp.232-243). Cambridge, UK: Cambridge University Press.
- Uyeda, T.Q., Abramson, P.D., & Spudich, J.A. (1996) The neck region of the myosin motor domain acts as a lever arm to generate movement. *Proc Natl Acad Sci USA*, 93(9), 4459-4464.
- Valdes, J.M., Tagle, D.A., & Collins, F.S. (1994) Island rescue PCR: a rapid and efficient method for isolating transcribed sequences from yeast artificial chromosomes and cosmids. *Proc Natl Acad Sci USA*, 91, 5377-5381.
- van den Anker, J.N., van Vught, E.E., Zandwijken, G.R., Cohen-Overbeek, T.E., & Lindhout, D. (1993) Severe limb abnormalities: analysis of a cluster of five cases born during a period of 45 days. *Am J Med Genet*, 45(5), 659-667.
- van der Kooi, A.J., van Meegen, M., Ledderhof, T.M., McNally, E.M., de Visser, M., & Bolhuis, P.A. (1997) Genetic localization of a newly recognized autosomal dominant limb-girdle muscular dystrophy with cardiac involvement (LGMD1B) to chromosome 1q11-21. *Am J Hum Genet*, 60(4), 891-895.
- van der Meulen, M.A., & te Meerman, G.J. (1997) Association and haplotype sharing due to identity by descent, with an application to genetic mapping. In I-H Pawlowitzki, J.H. Edwards & E.A. Thompson (Eds.), *Genetic Mapping of Disease Genes*, (pp 115-135). San Diego: Academic Press Inc.
- van Ommen, G.J., Breuning, M.H., & Raap, A.K. (1995) FISH in genome research and molecular diagnostics. *Curr Opin Genet & Develop*, 5(3), 304-308.
- Verloes, A. (1995) Numerical Syndromology: a mathematical approach to the nosology of complex phenotypes. *Am J Med Genet*, 55, 433-443.
- Victor, M., Hayes, R., & Adams, R.D. (1962) Oculopharyngeal muscular dystrophy. A familial disease of late life characterized by dysphagia and progressive ptosis of the eyelids. *New Eng J Med*, 267, 1267-1272.
- Vikstrom, K.L., & Leinwand, L.A. (1996) Contractile protein mutations and heart disease. *Curr Opin Cell Biol*, 8(1), 97-105.

- Vita, G., Dattola, R., Santoro, M., & Messina, C. (1983) Familial oculopharyngeal muscular dystrophy with distal spread. *J Neurol*, 230(1), 57-64.
- Vollrath, D., Davis, R.W., Connelly, C., & Hieter, P. (1988) Physical mapping of large DNA by chromosome fragmentation. *Proc Natl Acad Sci USA*, 85, 6027-6031.
- von Melchner, H., & Ruley, H.E. (1989) Identification of cellular promoters by using a retrovirus promoter trap. *J Virol* 63, 3227-3233.
- von Melchner, H., Reddy, S., & Ruley H.E. (1990) Isolation of cellular promoters by using a retrovirus promoter trap. *Proc Natl Acad Sci USA*, 87, 3733-3737.
- von Melchner, H., & Ruley, H.E. (1995) Gene entrapment. In F. Farzaneh & D.N. Cooper (Eds.), *Functional Analysis of the Human Genome* (pp. 109-118). Oxford: BIOS Scientific Publishers Limited.
- Voit, T., Hermann, E., Nuen-Jacob., & Cohn, R.D. (1998) Autosomal dominant infantile onset tibial myopathy. *Neuromuscul Disord*, 8(3,4), 251. (abstract)
- Voit, T., Kutz, P., Leube, B., Neuen-Jacob, E., Schroder, J.M., Cavallotti, D., Vaccario, M.L., Schaper, J., Broich, P., Cohn, R., Baethmann, M., Gohlich-Ratmann, G., Scoppetta, C., & Herrmann R. (2001) Autosomal dominant distal myopathy: further evidence of a chromosome 14 locus. *Neuromuscul Disord*, 11(1), 11-19.
- Vosberg, H.P. (1989) The polymerase chain reaction: an improved method for the analysis of nucleic acids. *Hum Genet*. 83, 1-15.
- Wade, R., Eddy, R., Shows, T.B., & Kedes, L. (1990) cDNA sequence, tissue specific expression, and chromosomal mapping of the human slow-twitch skeletal muscle isoform of troponin I. *Genomics*, 7, 346-357.
- Wallace, M.R., Marchuk, D.A., Anderson, L.B., Letcher, R., Odeh, H.M., Saulino, A.M., Fountain, J.W., Brereton, A, Nicholson, J., Mitchell, A.L., Brownstein, B.H., & Collins, F.S. (1990) Type 1 neurofibromatosis gene: identification of a large transcript disrupted in three NF1 patients. *Science*, 249, 181-186.
- Walton, J.N., & Natrass, F.J. (1954) On the classification, natural history and treatment of myopathies. *Brain*, 77, 169-251.
- Walton, J.N., & Gardner-Medwin, D. (1981) Progressive muscular dystrophy and the Myotonic disorders. In J.N. Walton (Ed.), *Disorders of voluntary muscle* (4th ed., pp. 481-524). Edinburgh: Churchill Livingstone.

- Walro, J.M., & Kucera, J. (1999) Why adult mammalian intrafusal and extrafusal fibers contain different myosin heavy-chain isoforms. *Trends Neurosci*, 22(4), 180-184.
- Walter, M.A., Spillet, D.J., Thomas, P., & Weissenbach, J., & Goodfellow, P.N. (1994) A method for constructing radiation hybrid maps of whole genomes. *Nat Genet*, 7(1), 22-28.
- Wang, W., Wu, W., Desai, T., Ward, D.C., & Kaufman, S.J. (1995) Localization of the alpha 7 integrin gene (*ITGA7*) on human chromosome 12q13: clustering of integrin and Hox genes implies parallel evolution of these gene families. *Genomics*, 26(3), 568-570.
- Wang, S., Sun, C.E., Walczak, C.A., Ziegler, J.S., Kipps, B.R., Goldin, L.R., & Diehl, S.R. (1995) Evidence for a susceptibility locus for schizophrenia on chromosome 6pter-p22. *Nat Genet*, 10(1), 41-46.
- Watkins, H., Conner, D., Thierfelder, L., Jarcho, J.A., MacRae, C., McKenna, W.J., Maron, B.J., Seidman, J.G., & Seidman, C.E. (1995) Mutations in the cardiac myosin binding protein-C gene on chromosome 11 cause familial hypertrophic cardiomyopathy. *Nat Genet*, 11(4), 434-437.
- Weber, F., de Villiers, J., & Schaffner, W. (1984) An SV40 "enhancer trap" incorporates exogenous enhancers or generates enhancers from its own sequences. *Cell*, 36, 983-992.
- Weber, J.L., & May, P.E. (1989) Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction. *Am J Hum Genet*, 44, 388-396.
- Weber J.L. (1990) Informativeness of human (dC-dA)<sub>n</sub>. (dG-dT)<sub>n</sub> polymorphisms. *Genomics*, 7, 524-530.
- Weeks, D.E., & Lange, K. (1988) The affected-pedigree-member method of linkage analysis. *Am J Hum Genet*, 42(2), 315-326.
- Weeks, D.E., & Lange, K. (1992) A multilocus extension of the affected-pedigree-member method of linkage analysis. *Am J Hum Genet*, 50(4), 859-868.
- Weiss, A., Schiaffino, S., & Leinwand, L.A. (1999) Comparative sequence analysis of the complete human sarcomeric myosin heavy chain family: implications for functional diversity. *J Mol Biol*, 290(1), 61-75.
- Weiler, T., Greenberg, C.R., Nylén, E., Halliday, W., Morgan, K., Eggertson, D., & Wrogemann, K. (1996) Limb-girdle Muscular Dystrophy and Myoshi Myopathy in an aboriginal Canadian kindred to map to *LGMD2B* and segregate with the same haplotype. *Am J Hum Genet*, 59, 872-878.



- Weiler, T., Bashir, R., Anderson, L.V., Davison, K., Moss, J.A., Britton, S., Nylen, E., Keers, S., Vafiadaki, E., Bushby, C.R., & Wrogemann, K. (1999) Identical mutation in patients with limb girdle muscular dystrophy type 2B or Miyoshi myopathy suggests a role for modifier gene(s) *Hum Mol Genet*, 8(5), 871-877.
- Weissenbach, J., Gyapay, G., Dib, C., Vignal, A., Morrisette, J., Millasseau, P., Vaysseix, G., & Lathrop, M. (1992) A second-generation linkage map of the human genome. *Nature*, 359, 794-801.
- Welander, L. (1951) Myopathia distalis tarda hereditaria. *Acta Medica Scand*, 141(Suppl. 265), 1-124.
- Weiner, S.J., Kollman, P.A., Nguyen, D.T., & Case D.A. (1986) Rigid-geometry energy maps with dihedral angles fixed at ideal values. *J Comp Chem*, 7, 230-252.
- Wijmenga, C., Hewitt, J.E., Sandkuijl, L.A., Clark, L.N., Wright, T.J., Dauwerse, H.G., Gruter, A.M., Hofker, M.H., Moerer, P., Williamson, R., van Ommen, G.-J. B., Padberg, G. W., & Frants, R. R. (1992) Chromosome 4q DNA rearrangements associated with facioscapulohumeral muscular dystrophy. *Nat Genet*, 2(1), 26-30.
- Wolfe, S.L. (1993) *Molecular and Cellular Biology*. Belmont: Wadsworth Inc.
- Wong, K.T., Dick, D., & Anderson, J.R. (1996) Mitochondrial abnormalities in oculopharyngeal muscular dystrophy. *Neuromuscul Disord*, 6(3), 163-166.
- Worton, R.G. & Thompson, M.W. (1988) Genetics of Duchenne muscular dystrophy. *Ann Rev Genet*, 22, 601-629.
- Yokota, J., Wada, M., Shimosato, Y., Teraada, M., & Sugimura, T. (1987) Loss of heterozygosity on chromosomes 3, 13 and 17 in small-cell carcinoma and on chromosome 3 in adenocarcinoma of the lung. *Proc Natl Acad Sci USA*, 84, 9252-9257.
- Youil, R., Kemper, B.W., & Cotton, R.G.H. (1995) Screening for mutations by enzyme mismatch cleavage with T4 endonuclease VII. *Proc Natl Acad Sci USA*, 92, 87-91.
- Yunis, E.J., & Samaha, F.J. (1971) Inclusion body myositis. *Lab Invest*, 25, 240-248.
- Zimmerman, S.S., Pottle, M.S., Nemethy, G., & Scheraga, H.A. (1977) Conformational analysis of the 20 naturally occurring amino acid residues using ECEPP. *Macromolecules*, 10(1), 1-9.

## **APPENDIX A**

### **Thesis Publications**

### **Thesis Publications**

- Eyre, H.J., Akkari, P.A., Meredith, C., Wilton, S.D., Callen, D.C., Kedes, L., & Laing, N.G. (1993) Assignment of the human slow skeletal muscle troponin gene (*TNNI1*) to 1q32 by fluorescence *in situ* hybridisation. *Cytogenet Cell Genet*, 62, 181-182.
- Akkari, P.A., Eyre, H., Wilton, S.D., Callen, D.C., Meredith, C., Kedes, L., & Laing, N.G. (1994). Assignment of the human skeletal muscle alpha actin gene (*ACTA1*) to 1q42 by fluorescence *in situ* hybridisation. *Cytogenet Cell Genet*, 65, 265-267.
- Laing, N.G., Laing, B.A., Meredith, C., Wilton, S.D., Robbins, P., Honeyman, K., Dorosz, S., Kozman, H.M., Mastaglia, F.L., & Kakulas, B.A. (1995). Autosomal dominant distal myopathy: linkage to chromosome 14. *Am J Hum Genet*, 56, 422-427.
- Hunt, C.C.J, Eyre, H.J., Akkari, P.A., Meredith, C., Dorosz, S.M., Wilton, S.D., Callen, D.F., Laing, N.G., & Baker E. (1995). Assignment of the beta tropomyosin gene (*TPM2*) to band 9p13 by fluorescence *in situ* hybridisation. *Cytogenet Cell Genet*, 71, 94-95.
- Felice, K.J., Meredith, C., Binz, N., Butler, A., Jacob, R., Akkari, P., Hallmayer, J., & Laing, N. (1999). Autosomal distal myopathy not linked to the known distal myopathy loci. *Neuromuscul Disord*, 9, 59-65.

## Assignment of the human slow skeletal muscle troponin gene (TNNI1) to 1q32 by fluorescence in situ hybridisation

H.J. Eyre,<sup>1</sup> P.A. Akkari,<sup>2</sup> C. Meredith,<sup>3</sup> S.D. Wilton,<sup>2</sup> D.C. Callen,<sup>1</sup> L. Kedes,<sup>4</sup> and N.G. Laing<sup>2</sup>

<sup>1</sup>Department of Cytogenetics and Molecular Genetics, Adelaide Children's Hospital, North Adelaide (Australia); <sup>2</sup>Australian Neuromuscular Research Institute, QEII Medical Centre, Nedlands (Western Australia); <sup>3</sup>Edith Cowan University, Joondalup (Western Australia); and <sup>4</sup>University of Southern California School of Medicine, Institute for Genetic Medicine, Los Angeles, CA (USA)

**Abstract.** The human gene for slow-twitch skeletal muscle troponin I (TNNI1) has previously been mapped to 1q12→qter using somatic cell hybrids. The TNNI1 locus has now been fur-

ther localised to 1q32 using fluorescence in situ hybridization. This result confirms the previous assignment of this locus and maps the gene to a single chromosome band.

Each muscle-specific gene should be associated with an inherited muscle disease, unless mutations do not occur in that gene or all mutations of that gene are lethal. Precise mapping of muscle genes in the human genome therefore becomes important in relation to mapping muscle diseases.

Striated muscle contraction is regulated by  $Ca^{2+}$  through the troponin complex (Ebashi et al., 1967; Zot et al., 1987). The troponin complex is composed of three polypeptides: troponin C (TnC), troponin I (TnI), and troponin T (TnT). Each polypeptide of the troponin complex is a member of a multigene family that is muscle-fibre specific. Thus, the TnI family consists of three isotypes, TnI-fast, TnI-slow, and TnI-cardiac, which are components of fast and slow skeletal muscle fibres and cardiac muscle fibres, respectively (Cummins and Perry, 1977).

The isoform TnI-slow was previously mapped to 1q12→qter by analysis of a somatic cell hybrid panel, using an isoform-specific probe from the 3' untranslated region (3' UTR) of the slow troponin I mRNA (Wade et al., 1990). The chromosome 1 locus was designated TNNI1 (Wade et al., 1990). We describe the assignment of the TNNI1 locus to 1q32 by utilising the 3' UTR probe to initially isolate genomic clones of TNNI1, which were then used for in situ hybridization.

### Materials and methods

#### Isolation of genomic clones

The genomic library was constructed in  $\lambda$ Gem 11 (Promega) following partial *Sau*3A digestion. The library was plated out at six-fold redundancy, and duplicate lifts were made using Hybond N<sup>+</sup> colony/plaque screening membranes (Amersham). The probe, LK 603, derived from the 3' UTR region of TNNI1, was radiolabeled with [ $\alpha$ -<sup>32</sup>P]dCTP (Amersham) using a random-primer kit (Promega). The duplicate lifts were hybridised with the probe for 16 h at 42 °C in a hybridisation solution of 50% formamide, 10% dextran sulphate, 1 M NaCl, 50 mM Tris-HCl [pH 7.5], and 0.1 mg/ml herring sperm DNA. Lifts were washed in 0.1% SDS, 0.1 × SSC for 30 min at 65 °C. Autoradiography was performed overnight with Dupont Cronex 4 X-ray film and Quanta III autoradiography screens at -80 °C. Plaques that were duplicated in the autoradiographs were considered positive. The purity of the positive plaques was ensured by a second and third round of screening.

Pure clones were then amplified by the plate lysate method (Sambrook et al., 1989). Phage were eluted overnight in SM buffer (5.8 g NaCl, 2 g  $MgSO_4 \cdot 7H_2O$ , 50 ml of 1 M Tris [pH 7.5], 5 ml of 2% gelatin solution, and sufficient  $H_2O$  to make up 1 liter) at 4 °C. Eluates were collected and centrifuged at 9,000 rpm in an SS34 Dupont rotor for 15 min to pellet any bacteria that might be present. The supernatant was collected, and the phage were precipitated with 10% polyethylene glycol, 1 M NaCl on ice for 1 h and then centrifuged at 1,500 rpm in an SS34 Dupont rotor. Supernatant was poured off, and the pellet was resuspended in 0.5 ml of TE (pH 8.0). The phage suspension was digested with 0.24 mg/ml RNase A at 37 °C for 30 min, then digested with 0.4 mg/ml proteinase K, 20 mM EDTA, 0.5% SDS for 15 min at 65 °C. The sample was extracted three times with a mixture of phenol and chloroform and once with chloroform alone. DNA integrity was checked by running 9  $\mu$ l of each sample on a 0.8% agarose gel at 100 V for 30 min, followed by staining with ethidium bromide. Single, high-molecular-weight bands were seen on the gel. Whole clones were then used for in situ hybridisation.

#### In situ hybridisation

The probes were nick-translated with biotin-14-dATP and hybridised in situ at a total probe concentration of 15 ng/ $\mu$ l to metaphases from two normal males. The fluorescence in situ hybridisation (FISH) method was modified from that previously described (Callen et al., 1990) in that only two rounds of amplification with fluorescein-conjugated avidin and biotinylated goat anti-avidin were required, and the chromosomes were stained before analysis with both propidium iodide (as a counterstain) and 4',6-diamidino-2-phenylindole (DAPI) (for chromosome identification).

Supported by the Adelaide Children's Hospital Research Fund (H.E. and D.C.), the Neuromuscular Foundation of Western Australia (P.A., S.W., and N.L.), the Edith Cowan University Research Fund (C.M.), the National Institutes of Health (L.K.), and the Muscular Dystrophy Association (L.K.).

Received 14 September 1992; accepted 15 October 1992.

Request reprints from Dr N. G. Laing, Australian Neuromuscular Research Institute, Fourth Floor, "A" Block, QEII Medical Centre, Nedlands, Western Australia 6009 (Australia).

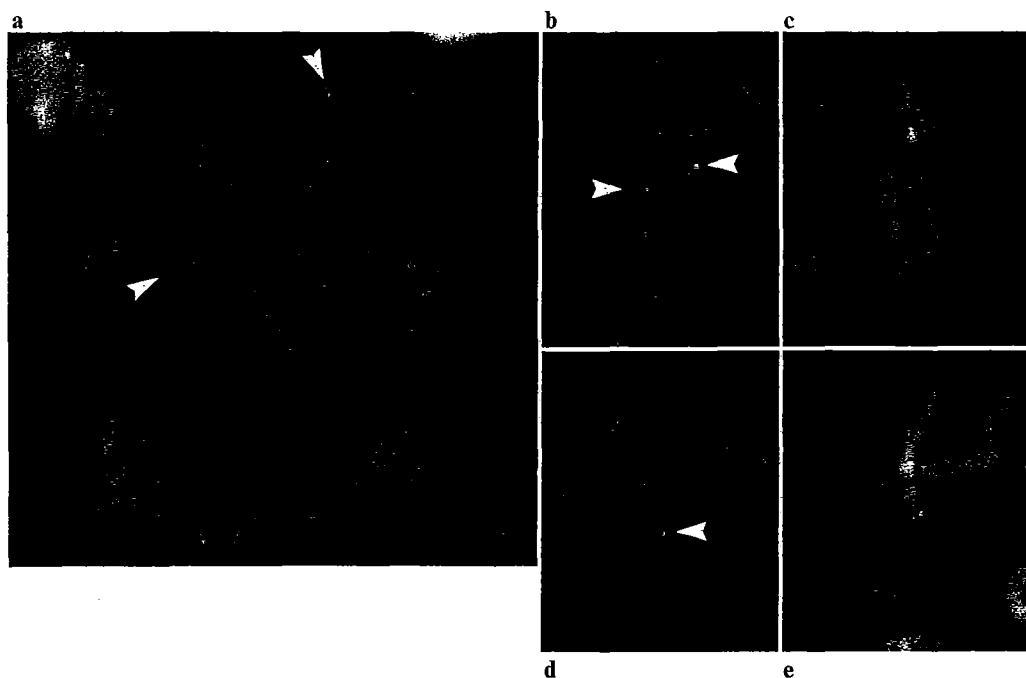


Fig. 1. Whole and partial metaphases showing in situ hybridization with the biotinylated probe (TNNI1). (a, b, d) Normal male chromosomes stained with propidium iodide. Hybridisation sites on chromosome 1 are indicated by arrows. (c, e) DAPI staining of the same metaphases for chromosome identification.

### Results and discussion

Twenty-five metaphases from the first normal male were examined for fluorescent signals. Twenty-three of these metaphases showed signal on one or both chromatids of chromosome 1 in the region of 1q31→q32. Ninety percent of this signal was at q32 (Fig. 1). A total of nine nonspecific background dots were observed in these 25 metaphases.

A similar result was obtained by hybridization of the 3' UTR probe of TnI-slow to metaphases derived from the second normal male (data not shown).

To localize TNNI1 to 1q32, it was necessary to use a probe that distinguished this TnI-slow isoform from the TnI-fast and TnI-cardiac isoforms. Although the 3' UTR probe of TnI-slow distinguishes it from the other two isoforms of the TnI gene

family, the size of the probe (250 bp) was not large enough for the in situ hybridization method used. The 3' UTR probe was therefore used to isolate genomic clones whose insert size was large enough for in situ hybridization.

The results of this study confirm the previous assignment of the human TNNI1 locus to 1q12→qter (Wade et al., 1990) and further localise the gene to 1q32. The localisation of TNNI1 to 1q32 should be of interest to laboratories searching for linkage of inherited human muscle disorders, especially those that have already linked diseases to this region.

### Acknowledgements

We thank Dr. L. Abrahams for the gift of the genomic library.

### References

- Callen DF, Baker E, Eyre HJ, Chernos JE, Bell JA, Sutherland GR: Reassessment of two apparent deletions of chromosome 16p to an ins(11;16) and at (1;16) by chromosome painting. *Annls Génét* 33:219-221 (1990).
- Cummins P, Perry V: Troponin I from human skeletal and cardiac muscles. *Biochem J* 171:251-259 (1977).
- Ebashi S, Ebashi F, Kodama A: Troponin as the Ca<sup>++</sup> receptive protein in the contractile system. *J Biochem, Tokyo* 62:137-138 (1967).
- Sambrook J, Fritsch EF, Maniatis T: *Molecular Cloning: A Laboratory Manual*, 2nd Ed (Cold Spring Harbor Laboratory Press, Cold Spring Harbor 1989).
- Wade R, Eddy R, Shows TB, Kedes L: cDNA sequence, tissue specific expression, and chromosomal mapping of the human slow-twitch skeletal muscle isoform of troponin I. *Genomics* 7:346-357 (1990).
- Zot AS, Potter JD: Structural aspects of troponin-tropomyosin regulation of skeletal muscle contraction. *A Rev biophys Chem* 16:535-539 (1987).

## Assignment of the human skeletal muscle alpha actin gene (ACTA1) to 1q42 by fluorescence in situ hybridisation

P. A. Akkari,<sup>1</sup> H. J. Eyre,<sup>2</sup> S. D. Wilton,<sup>1</sup> D.F. Callen,<sup>2</sup> S. A. Lane,<sup>2</sup>  
C. Meredith,<sup>3</sup> L. Kedes,<sup>4</sup> and N. G. Laing<sup>1</sup>

<sup>1</sup>Australian Neuromuscular Research Institute, QEII Medical Centre, Nedlands,

<sup>2</sup>Department of Cytogenetics and Molecular Genetics, Adelaide Children's Hospital, North Adelaide,

<sup>3</sup>Edith Cowan University, Joondalup (Australia) and

<sup>4</sup>University of Southern California School of Medicine, Institute for Genetic Medicine, Los Angeles CA (USA)

**Abstract.** The human skeletal muscle alpha actin gene (ACTA1) has previously been localized to 1p21→qter using somatic cell hybrids and a specific probe from the 3' untranslated

region of the gene. Using fluorescence in situ hybridization the localization has been confirmed and the ACTA1 gene precisely mapped to 1q42.

Alpha actin is one of the major components of the contractile unit in skeletal muscle. The human skeletal muscle alpha actin gene, ACTA1 has previously been localized to 1p21→qter by somatic cell hybrid analysis using an isoform specific probe derived from the 3' untranslated region of a full length cDNA clone of ACTA1 (Gunning et al., 1984).

We describe the mapping of the ACTA1 gene to 1q42 by initially using the ACTA1 sequence tagged site (STS) (Freeman and States, 1991) to isolate pure genomic clones, followed by fluorescence in situ hybridization (FISH) (Callen et al., 1990) of these clones to metaphase chromosome spreads.

### Materials and methods

#### Isolation of genomic clones

The genomic library was constructed in  $\lambda$  Gem 11 (Promega) following a partial *Sau* 3A digestion. The library was plated at 6-fold redundancy and duplicate lifts were made using Hybond N+ colony-plaque screening membranes (Amersham). The PCR amplified STS of the ACTA1 gene (Freeman

and States, 1991) used as a probe, was labeled with [ $\alpha$ -<sup>32</sup>P]dCTP (Amersham) using a random-primer kit (Promega). The duplicate lifts were hybridised with probe at 42 °C for 16 hours in a hybridisation solution of 50% formamide, 10% dextran-sulphate, 1M NaCl, 50 mM Tris-HCl pH 7.5, 0.1 mg/ml herring sperm DNA. Lifts were washed in 0.1% SDS and 0.1 × SSC at 65 °C for 30 min. Dupont Cronex 4 X-ray film with Quanta III autoradiography screens were used for autoradiography at -80 °C overnight. Plaques showing up in duplicate on the autoradiographs were considered positive. STS-PCR assay confirmed the presence of the ACTA1 sequence in the genomic clones.

Pure clones were then amplified by the plate lysate method (Sambrook et al., 1989). Phage were eluted in SM buffer (100 mM NaCl, 8 mM MgSO<sub>4</sub>·7H<sub>2</sub>O, 50 mM Tris pH 7.5, 0.01% gelatin solution) at 4 °C overnight. Eluates were collected, phage purified on caesium chloride gradient (Sambrook, et al., 1989) and DNA extracted. DNA integrity was checked by running each sample on a 0.8% agarose gel at 100 volts for 30 min then staining with ethidium bromide. Confirmation that the extracted DNA contained ACTA1 gene was obtained by STS-PCR and also by PCR amplification of the 5' and 3' ends of the gene (5' primer set: [TGA-GACTTCTGCGCTGATGCA, ACTCACGGTGGTCCTTCCGGA], 3' primer set: [GCACAGGCATCGTGCTGGACT, TATGTACACGTTA-TAAACACTG]). Whole clones were then used as probes for in situ hybridisation.

#### In situ hybridisation

The probes were nick-translated with biotin-14-dATP and each was hybridized in situ at a final concentration of 20 ng/ $\mu$ l to metaphases from two normal males. The FISH method was modified from that previously described (Callen, et al., 1990) in that chromosomes were stained before analysis with both propidium iodide (as counterstain) and DAPI (for chromosome identification).

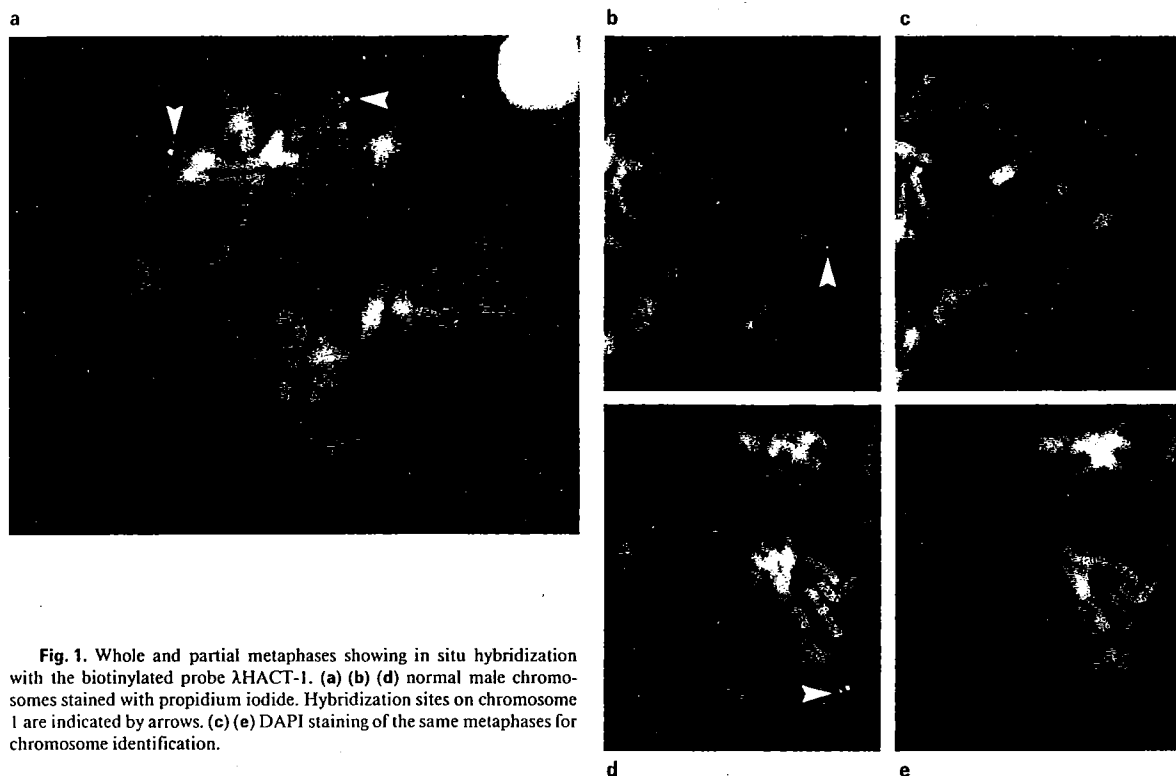
#### Cell lines

The construction of the mouse-human hybrid cell lines was described by Callen et al. (1992). Table I shows the human chromosome 1 content of the cell lines.

Supported by the Neuromuscular Foundation of Western Australia (P.A., S.W., N.L.), the Adelaide Children's Hospital Research Fund (H.E., S.L., D.C.), the Edith Cowan University Research Fund (C.M.) and National Institutes of Health USA (L.K.).

Received 2 July 1993; manuscript accepted 3 August 1993.

Request reprints from Dr N.G. Laing, Australian Neuromuscular Research Institute, 4th Floor 'A' Block, QEII Medical Centre, Nedlands, Western Australia, 6009, Australia; telephone: 61-9-389-2818; fax: 61-9-389-3487; Email: nlaing@niwa.uwa.edu.au



**Fig. 1.** Whole and partial metaphases showing in situ hybridization with the biotinylated probe  $\lambda$ HACT-1. (a) (b) (d) normal male chromosomes stained with propidium iodide. Hybridization sites on chromosome 1 are indicated by arrows. (c) (e) DAPI staining of the same metaphases for chromosome identification.

## Results and discussion

In order to map ACTA1 to 1q42 by FISH, probes greater than 1 kb in length are required. The isoform specific STS PCR product (188bp) (Freeman and States, 1991) was initially used as a hybridization probe to the genomic library which has an average insert size of 15 kb. Six positive plaques were picked and respread at lower density so that individual plaques could be isolated. The STS primer set was used to assay selected pure plaques. Two pure plaques which gave positive STS signals,  $\lambda$ HACT-1 and  $\lambda$ HACT-2, were then further propagated.

$\lambda$ HACT-1 gave positive results for both the 3' and 5' primers and is therefore likely to contain the whole ACTA1 gene.  $\lambda$ HACT-2 gave a positive signal for the 3' primer set but not the 5' primer set demonstrating that only part of the gene is present in this clone.

$\lambda$ HACT-1 and  $\lambda$ HACT-2 were then used for FISH to metaphase spreads from two normal males. Twenty-five metaphases from the first normal male were examined for fluorescent signal with  $\lambda$ HACT-1. Twenty-one of these metaphases showed signal on one or both chromatids of chromosome 1 in the region q42→q43; 73% of this signal was at q42 (Fig. 1). Seven non-specific background dots were observed in the 25 metaphases. Similar results were obtained in 20 metaphases from the second normal male hybridized with this probe and from hybridizations using  $\lambda$ HACT-2 as the probe. These results assign the ACTA1 locus to 1q42.

**Table I.** Assignment of ACTA1 by STS-PCR amplification of chromosome 1 hybrids

Hybrid cell line	Human chromosome content	Hybridization <sup>a</sup>
CY15	der(16)t(1;16)(p34;p12.2) <sup>b</sup>	-
CY129	der(16)t(1;16)(p13.1;p11.2) <sup>b</sup>	-
CY152	der(16)t(1;16)(p11;p11.2) <sup>b</sup>	-
CY128	der(1)(1;16)(q43;q22.1)	+
CY196	der(16)t(1;16)(q12;p13.3)	+
CY13	der(16)t(1;16)(q44;p12.3) <sup>b</sup>	-

<sup>a</sup> PCR amplification of the ACTA1 STS is indicated by (+).

<sup>b</sup> Although these cell lines contain other human chromosomes, the presence of chromosome 1 has been excluded.

PCR analysis of the hybrid cell panel showed that the hybrid cell lines CY128 and CY196 contained the ACTA1 sequence (Table I). The only region of chromosome 1 unique to these cell lines is q12→q43, which is consistent with the in situ hybridization result placing ACTA1 in q42.

These results confirm the assignment of the ACTA1 locus to 1p21→qter using a specific probe for the 3' untranslated region with somatic cell hybrids (Gunning et al., 1984) and further localise the gene to 1q42.

The skeletal muscle alpha-actin gene has been localized to chromosome 8 in the mouse (*Actsk-1*) (Alonso et al., 1993;

Abonia et al., 1993) and chromosome 1 in human (ACTA1) (Gunning et al., 1984; this report). Alonso et al. (1993) and Abonia et al. (1993) show that angiotensinogen (*Agt*) is closely linked to *Actsk-1* on chromosome 8 in the mouse while AGT has been localised to 1q42→q43 in human (Isa et al., 1990; Gaillard-Sanchez et al., 1990). The present FISH localisation of ACTA1 to 1q42 is further evidence of a conserved region of homology between mouse chromosome 8 and human chromosome 1 and confirms the inference that ACTA1 in human should map in the region of 1q42→q43 (Alonso et al., 1993; Abonia et al., 1993).

The localization of ACTA1 to 1q42 should be of interest to laboratories searching for linkage of inherited neuromuscular disorders, especially those that have already linked such diseases to this region.

#### Acknowledgements

We thank Dr. L. Abrahams for the gift of the genomic library. We thank Drs Josephine Peters and Tony Searle for information on the mouse-human homologies.

#### References

- Abonia PJ, Abel KJ, Eddy RL, Elliott RW, Chapman VM, Shows TB, Gross KW: Linkage of *Agt* and *Actsk-1* to distal mouse chromosome 8 loci: a new conserved linkage. *Mammal Genome* 4:25-32 (1993).
- Alonso S, Montagutelli X, Simon-Chazottes D, Guenet J-L, Buckingham M: Re-localization of *Actsk-1* to mouse chromosome 8, a new region of homology with human chromosome 1. *Mammal Genome* 4:15-20 (1993).
- Callen DF, Baker E, Eyre HJ, Chernos JE, Bell JA, Sutherland GR: Reassessment of two apparent deletions of chromosome 16p to an ins(11;16) and a t(1;16) by chromosome painting. *Ann Genet* 33:219-221 (1990).
- Callen DF, Doggett NA, Stallings RL, Chen LZ, Whitmore SA, Lane SA, Nancarrow JK, Apostolou S, Thompson AD, Lapsys NM, Eyre H, Baker E, Phillips H, Holman K, Shen Y, Richards RI, Weber JL, Sutherland GR: High resolution cytogenetic-based physical map of human chromosome 16. *Genomics* 13:1178-1185 (1992).
- Freeman BC, Slates JC: An STS in the human skeletal  $\alpha$ -actin gene. *Nucl Acids Res* 19:5086 (1991).
- Gaillard-Sanchez I, Mattei MG, Clauser E, Corvol P: Assignment by in situ hybridisation of the angiotensinogen gene to chromosome band 1q4, the same region as the human renin gene. *Hum Genet* 84:341-343 (1990).
- Gunning P, Mohun T, Ng S-Y, Ponte P, Kedes L: Evolution of the human sarcomeric actin genes: evidence for units of selection within the 3' untranslated regions of the mRNAs. *J molec Evol* 20:202-214 (1984).
- Isa MN, Boyd E, Morrison N, Harrap S, Clauser E, Connor JM: Assignment of the human angiotensin gene to chromosome 1q42→q43 by non-isotopic in situ hybridisation. *Genomics* 8:598-600 (1990).
- Sambrook J, Fritsch EF, Maniatis T: *Molecular Cloning: A Laboratory Manual* (2nd ed.). (Cold Spring Harbor Laboratory Press, Cold Spring Harbor 1989).



## Autosomal Dominant Distal Myopathy: Linkage to Chromosome 14

N. G. Laing,<sup>1</sup> B. A. Laing,<sup>1</sup> C. Meredith,<sup>3</sup> S. D. Wilton,<sup>1</sup> P. Robbins,<sup>2</sup> K. Honeyman,<sup>3</sup> S. Dorosz,<sup>1</sup> H. Kozman,<sup>4</sup> F. L. Mastaglia,<sup>1</sup> and B. A. Kakulas<sup>1</sup>

<sup>1</sup>Australian Neuromuscular Research Institute and <sup>2</sup>Department of Pathology, QEII Medical Centre, and <sup>3</sup>Edith Cowan University, Perth; and <sup>4</sup>Department of Cytogenetics and Molecular Genetics, Adelaide Children's Hospital, Adelaide

### Summary

We have studied a family segregating a form of autosomal dominant distal myopathy (MIM 160500) and containing nine living affected individuals. The myopathy in this family is closest in clinical phenotype to that first described by Gowers in 1902. A search for linkage was conducted using microsatellite, VNTR, and RFLP markers. In total, 92 markers on all 22 autosomes were run. Positive linkage was obtained with 14 of 15 markers tested on chromosome 14, with little indication of linkage elsewhere in the genome. Maximum two-point LOD scores of 2.60 at recombination fraction .00 were obtained for the markers MYH7 and D14S64—the family structure precludes a two-point LOD score  $\geq 3$ . Recombinations with D14S72 and D14S49 indicate that this distal myopathy locus, *MPD1*, should lie between these markers. A multipoint analysis assuming 100% penetrance and using the markers D14S72, D14S50, MYH7, D14S64, D14S54, and D14S49 gave a LOD score of exactly 3 at MYH7. Analysis at a penetrance of 80% gave a LOD score of 2.8 at this marker. This probable localization of a gene for distal myopathy, *MPD1*, on chromosome 14 should allow other investigators studying distal myopathy families to test this region for linkage in other types of the disease, to confirm linkage or to demonstrate the likely genetic heterogeneity.

### Introduction

The distal myopathies are a heterogeneous group of disorders showing both autosomal dominant and autosomal recessive inheritance, with diverse phenotypic features and pathological changes (Mastaglia 1991). Gowers (1902) described a patient with weakness and wasting of the hand, forearm, and anterior tibial muscles, which had developed by the age of 10 years; and by the age of 18 years there

was also severe wasting of the sternomastoids, wasting of the tongue, and weakness of the facial muscles. This is accepted as the first description of the hereditary distal myopathies (MIM 160500; McKusick 1992), although it has been suggested that Gowers's patient may have had myotonic dystrophy, which shows a similar pattern of muscle atrophy and weakness (see Walton and Gardner-Medwin 1981). The better-known of the distal myopathies are the dominantly transmitted forms, occurring particularly in Scandinavia, which may have a late onset and either a relatively benign course (Welander 1951) or a more rapid progression (Edström et al. 1980). Recessive forms have also been described, particularly in Japan, where a distinctive myopathy characterized histologically by the presence of rimmed vacuoles in muscle fibers occurs (Nonaka et al. 1981).

To date, genetic linkage has not been demonstrated in any of the forms of distal myopathy, nor have any candidate genes been identified (McKusick 1992). We report here positive linkage to the long arm of chromosome 14 in a family with dominantly inherited distal myopathy with phenotypic features closely resembling those described in Gowers's original report (Gowers 1902). These results have previously been reported in brief (Meredith et al. 1994).

### Subjects and Methods

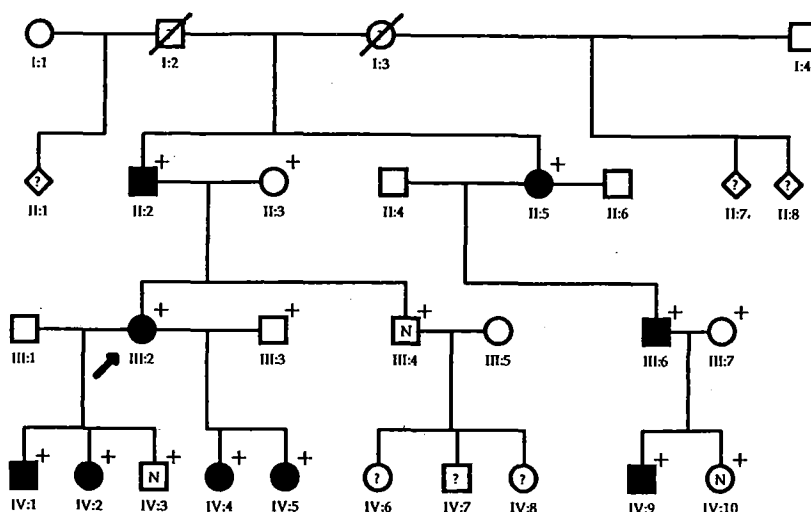
#### Family

The condition was dominantly transmitted with high penetrance in three generations of the family, which is of English/Welsh origin (fig. 1). Nine affected and three unaffected members from three generations were examined.

#### Clinical Features

Age at onset was 4–25 years, with selective weakness of the toe and ankle extensors and the neck flexors, followed, after several years, by progressive weakness of the finger extensors. The finger flexors and intrinsic hand muscles were relatively unaffected, but certain proximal muscle groups such as the hip abductors and external rotators and shoulder abductors were mildly affected. Although progression was gradual, there was eventually a moderate degree of incapacity, as in II:2, the oldest living affected family member, who was initially investigated in 1969 and who

Received October 22, 1993; accepted for publication October 19, 1994.  
Address for correspondence and reprints: Dr. Nigel G. Laing, Australian Neuromuscular Research Institute, 4th Floor 'A' Block, QEII Medical Centre, Perth, Western Australia 6009, Australia.  
© 1995 by The American Society of Human Genetics. All rights reserved.  
0002-9297/95/5602-0010\$02.00



**Figure 1** Pedigree of the family segregating distal myopathy. The proband (III:2) is indicated by the arrow. Each individual marked with a cross was examined clinically, and an immortalized cell line was established for him or her. Blackened symbols denote affected individuals; an "N" in the center of the symbol denotes that the individual is clinically unaffected, and a question mark (?) in the center of the symbol denotes that the individual's status is unknown.

was reexamined in 1992 when he was still walking but had difficulty maintaining an erect posture when standing. In all affected individuals the deep-tendon reflexes were preserved, and the plantar responses were flexor. There was no myotonia, sensory impairment, or other neurological abnormality. Serum creatine kinase levels were elevated (216–531 U/liter [normal <180 U/liter]) in three affected individuals.

#### Electrodiagnostic Studies

Electromyography and nerve conduction studies were performed in the proband (III:2; fig. 1) and in II:2 and IV:2. Muscle sampling with a concentric needle electrode showed striking myopathic motor-unit potential changes (i.e., low amplitude and brief duration units, with many polyphasic units) and a full, low-amplitude (<0.5 mV) interference pattern particularly in affected distal limb muscles and, to a lesser extent, in some proximal muscles. Occasional spontaneous fibrillation potentials and positive waves were present in some affected muscles in III:2. Myotonia was not found. Results of motor- and sensory-nerve conduction studies in the upper and lower limbs were normal.

#### Histological Changes

An open biopsy taken from the left vastus lateralis muscle in III:2 showed occasional necrotic and regenerating fibers, excessive variation in fiber size, increased numbers of fibers with internal nuclei, occasional angulated atrophic

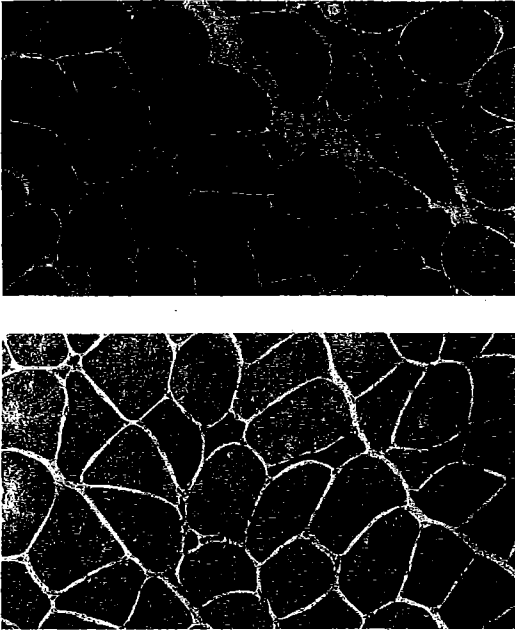
fibers, and nuclear clumps (fig. 2). There were no rimmed vacuoles or sarcoplasmic inclusions. Histochemistry showed only nonspecific changes with moth-eaten fibers and increased enzyme activity in angulated fibers in the NADH-TR preparations (fig. 2). There was no pathological fiber-type grouping. There were no sarcoplasmic inclusion bodies as found in the family described by Edström et al. (1980). Results of dystrophin and desmin immunohistochemistry were normal. A biopsy of the left tibialis anterior muscle from II:2, taken in 1969, showed evidence of end-stage disease with extreme muscle-fiber atrophy, many central nuclei and nuclear clumps, and condensation of endomysial connective tissue.

#### Blood Samples

Blood samples were taken, and immortalized cell lines were established for the 15 family members who volunteered to participate in the study (fig. 1). Venous blood samples (≤40 ml) from adults were split 50:50 into lithium-heparin and EDTA tubes for, respectively, lymphocyte isolation and immortalization with Epstein-Barr virus (Neitzel 1986) and immediate DNA isolation by either conventional phenol-chloroform or salt-precipitation (Miller et al. 1988) techniques.

#### Linkage Analysis

Ninety-two polymorphic markers spread over all autosomes were tested. These were D1S80, D1S185, D1S176, CRP, D1S104, AT3, and ACTN2; APOB, D2S44, HOX4,



**Figure 2** Top, Vastus lateralis biopsy from III:2, showing excessive variation in myofibril size and nuclear clumps: hematoxylin and eosin ( $\times 400$ ). Bottom, Vastus lateralis biopsy from III:2, showing darkly stained angulated atrophic type II fibers and moth-eaten type I fibers: NADH-TR ( $\times 400$ ).

and D2S102; D3S1307, D3S1304, D3S1285, D3S1287, D3S1279, D3S1282, D3S1262, and D3S1311; HD, D4S174, and D4S171; D5S268, D5S112, D5S39, APC, and D5S210; DMDL; D7S472, D7S435, D7S474, GCK, D7S440, and CF; D8S166, D8S84, D8S198, and D8S200; D9S104, D9S15, and D9S53; D10S28; D11S875, D11S873, CD3D, and D11S836; D12S43 and D12S60; RB1; D14S72, D14S50, MYH7, D14S64, D14S54, D14S49, D14S52, D14S76, D14S53, D14S74, D14S48, D14S81, D14S45, D14S51, and D14S31; D15S87; D16S291, D16S292, D16S287, D16S295, D16S298, D16S300, D16S308, D16S265, D16S186, D16S301, D16S260, D16S266, and D16S305; D17S30, D17S122, GX-Alu, HOXB, and D17S26; D18S40 and D18S51; D19S75, APOC2, and DM; D20S66; D21S210 and D21S213; and D22S258.

#### VNTR Analysis

Probes for genotyping the VNTR markers D2S44, D10S28, D14S13, and D17S26 were purchased from Promega. For each family member, 4  $\mu$ g of DNA isolated as above was digested overnight with *Hae*III (Amersham). Fragments separated in a 0.8% agarose gel electrophoresis were alkali-transferred to Hybond N<sup>+</sup> nylon membranes

(Amersham) and hybridized with inserts labeled by random priming (Amersham). Cronex 4 film (Du Pont) was exposed for 7–10 d at  $-80^{\circ}\text{C}$  with single intensifying screens.

#### Microsatellite and PCR Polymorphisms

A Cyclone Plus DNA Synthesiser (Milligen/Bioscience) was used to synthesize primers for microsatellite (Weber and May 1989) and PCR polymorphisms. Primer sequences were obtained from the Genome Data Base. Microsatellite PCR conditions for a reaction volume of 20  $\mu$ l were as follows: 4  $\mu$ l of 5  $\times$  buffer (335 mM Tris-HCl [pH 8.8] at  $25^{\circ}\text{C}$ , 83 mM  $(\text{NH}_4)_2\text{SO}_4$ , 100  $\mu\text{M}$  dNTP, 0.2  $\mu$ l of 25 mM  $\text{MgCl}_2$ , 50 ng of each primer, 0.1  $\mu$ l of  $^{32}\text{P}$ -dCTP (3,000 Ci/mmol; Amersham), 50 ng of target DNA, and 0.5 unit of *Tth* polymerase (Biotech International). The reactions were overlaid with mineral oil, and the cycling conditions were as follows:  $94^{\circ}\text{C}$  for 5 min to denature,  $58^{\circ}\text{C}$  for 6 min to anneal and elongate for one cycle, followed by 35 cycles of  $94^{\circ}\text{C}$  for 1 min and  $58^{\circ}\text{C}$  for 6 min. Where necessary, annealing temperatures were raised to improve primer specificity. Products were then refrigerated at  $4^{\circ}\text{C}$  until gel electrophoresis. Aliquots of the reaction product were mixed with an equal volume of formamide-loading buffer and, depending on expected allele size, were electrophoresed on standard 6% or 4% acrylamide denaturing sequencing gels (Sambrook et al. 1989, chap. 13). PCR reaction conditions for four nonmicrosatellite polymorphisms (APOB, AT3, HD, and D17S30) were similar, except that no radioactive nucleotide was included in the reaction mix and aliquots of the reaction products were electrophoresed in 2% agarose.

#### Linkage Analysis

Two-point linkage analysis was carried out using the method of maximum likelihood and LOD scores (Morton 1955), by the computer program LIPED (Ott 1974). Analysis of the combined data from all the markers was performed using the EXCLUDE program (Edwards 1987; Sarfarazi et al. 1989). The chromosomal positions of loci were estimated from the NIH/CEPH Collaborative Mapping Group (1992) linkage maps, the second-generation linkage map of the human genome (Weissenbach et al. 1992), and data from the Cooperative Human Linkage Center. Multipoint linkage analysis was performed using the CRIMAP version 2.4 package, as described elsewhere (Mulle et al. 1993), at 100% and 80% penetrance.

#### Results

##### Linkage

In total, 92 marker loci spread over all 22 autosomes were typed in the distal myopathy pedigree. Two markers on chromosome 8 gave positive two-point LOD scores  $>1.00$ . D8S84 gave a LOD score of 1.11 at recombination

Table 1

Results of Two-Point Linkage Analyses, for Linkage between MPD1 and Markers on Chromosome 14

Locus	Location*	LOD Score at $\theta$ =							
		0	.001	.01	.05	.10	.20	.30	.40
D14S72 .....	q11	-3.13	-1.56	-.59	.03	.23	.31	.26	.15
D14S50 .....	q11	2.31	2.31	2.27	2.12	1.93	1.51	1.06	.57
MYH7 .....	q11	2.60	2.59	2.55	2.36	2.12	1.61	1.08	.53
D14S64 .....	q11-q32	2.60	2.59	2.55	2.36	2.12	1.61	1.08	.53
D14S54 .....	q11-q32	.41	.41	.40	.35	.29	.17	.09	.02
D14S49 .....	q11-q32	-2.49	-1.25	-.30	.30	.47	.49	.36	.19
D14S52 .....	q11-q32	-.99	-3.41	-1.45	-.22	.17	.32	.23	.10
D14S76 .....	q11-q32	-.99	-6.39	-3.42	-1.46	-.72	-.17	.01	.05
D14S53 .....	q11-q32	-.99	-3.31	-1.38	-.14	.27	.49	.44	.27
D14S74 .....	q11-q32	-.99	-.99	-.02	.55	.69	.66	.52	.30
D14S48 .....	q11-q32	-.99	-.47	.49	1.00	1.06	.87	.56	.26
D14S81 .....	q11-q32	.20	.20	.19	.15	.10	.04	.01	.00
D14S45 .....	q32.1-qter	-.99	-2.20	-1.21	-.57	-.34	-.15	-.07	-.02
D14S51 .....	q32.1-qter	-.99	-3.77	-1.80	-.53	-.10	.15	.15	.08
D14S13 .....	q32.1-qter	-.99	-1.61	-.63	-.03	.15	.19	.13	.04

\* From NIH CEPH Collaborative Mapping Group (1992) and Weissenbach et al. (1992).

fraction ( $\theta$ ) .00, and D8S200 gave a LOD score of 1.09 at  $\theta$  = .00. These two markers were, however, only informative in one branch of the family, and other markers on chromosome 8—i.e., D8S166 and D8S198—which gave information for the other branch of the family, showed multiple recombinants. Thus, significantly negative LOD scores were obtained for these markers, and chromosome 8 is unlikely to be the site of the disease locus. More strongly positive two-point LOD scores were obtained with markers on chromosome 14. D14S50, MYH7, and D14S64 showed no recombinants (table 1) and two-point LOD scores of 2.31, 2.60, and 2.60, respectively. A multipoint LOD score of 3.00 was obtained for MYH7, and 2.99 for D14S64, when the information for D14S72, D14S50, MYH7, D14S64, D14S54, and D14S49 was used (fig. 3), under the assumption of 100% penetrance. A multipoint LOD score of 2.80 was obtained for MYH7 at a penetrance of 80%. Recombinants with D14S49 and D14S72 indicate that the disease locus, *MPD1*, should lie between these markers.

EXCLUDE

The EXCLUDE program (Edwards 1987) was used to estimate the likelihood that the distal myopathy gene in this family lies on chromosome 14 and to estimate the proportion of the genome excluded as the possible site for the disease gene. Running the EXCLUDE program for the entire data set gave a probability of 99.9% that the gene responsible for autosomal dominant distal myopathy in this family, *MPD1*, lies on chromosome 14. The same probability of 99.9% for the disease gene lying on chromo-

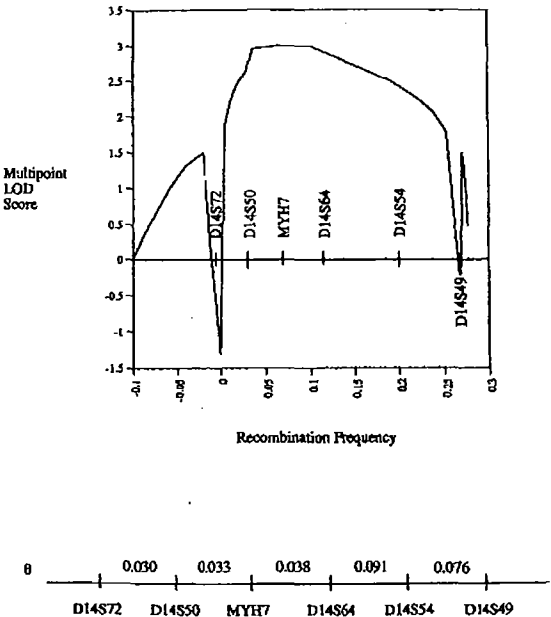


Figure 3 Multipoint analysis for the markers D14S72, D14S50, MYH7, D14S64, D14S54, and D14S49, assuming 100% penetrance. The graph shows multipoint LOD scores for chromosome 14; the LOD score for a penetrance of 80% was 2.8 at MYH7. Shown beneath the graph is the background map used to localize the disease genes.  $\theta$  is sex-averaged and is calculated by using Haldane's mapping function.

some 14 was obtained from EXCLUDE in an affecteds-only analysis. Running the EXCLUDE program for the data set minus the data for chromosome 14 showed that a considerable proportion,  $\geq 50\%$ , of the genome had been excluded as the possible site of the disease gene in this family.

## Discussion

The pedigree investigated displays a form of distal myopathy that is quite distinct from the better-known distal myopathies described in Scandinavia and Japan. Welander (1951) described a benign Scandinavian form with weakness first occurring in the fingers and wrists and later progressing to involve the anterior tibial and calf muscles. There was also early involvement of the thumb, wrist, and finger muscles in the severe Scandinavian form described by Edström et al. (1980), with weakness progressing to other muscles and leading to major disability within 15 years of onset of the disease. In comparison, the distal myopathy described here is very similar to that reported by Gowers (1902), with early involvement of the distal lower limb muscles and the sternomastoids and with relative sparing of the hand and forearm muscles.

The distal myopathy in this family shows obvious dominant inheritance. A considerable portion of the genome has been scanned with the 92 markers tested. The best evidence of linkage is seen with chromosome 14. The EXCLUDE program (Edwards 1987) gives a 99.9% likelihood of linkage to chromosome 14, both in an affecteds-only analysis and when data for all the family members are used, while multipoint analysis gives a LOD score of 3.00 at MYH7, under the assumption of 100% penetrance, and gives a LOD score of 2.8 at 80% penetrance. There is little evidence of linkage elsewhere. It is unlikely that a larger LOD score will be obtained from study of this family, since the Australian branch of the family has lost contact with any remaining relatives in England/Wales, and, despite attempts to identify such relatives, none have been found. It is also possible that there are no other affected relatives.

Both the linkage to MYH7 and D14S64 and the recombinations with D14S49 and D14S72 indicate that the mutated gene, *MPD1*, should lie in the region of these markers. Both the distal myopathy in this family and the other distal myopathies affect restricted groups of muscles in defined locations of the body; other muscle diseases affect other restricted muscle groups. Exactly what type of genes that, when mutated, could cause muscle diseases restricted to certain muscle groups is uncertain. They could perhaps be genes coding for muscle structural proteins, though this might seem unlikely because such proteins should be expressed in all muscle groups. The mutated genes might, on the other hand, be genes involved in delineating body

segments during development. Recently a repeated homeo-domain has been implicated in another restricted form of muscle disease, facioscapulohumeral muscular dystrophy (Wijmenga et al. 1992), although the exact relationship between the homeodomain and the disease has been brought into question by the identification of recombinations between the homeodomain marker and the disease (Weiffenbach et al. 1993).

Two muscle genes that are known to lie within the *MPD1* linkage region on chromosome 14 are the alpha and beta cardiac myosin genes MYH6 and MYH7 (Matsuo et al. 1989). Although beta cardiac myosin is expressed in skeletal muscle (Jandreski et al. 1987), it is unlikely that mutations in that gene are the basis of the distal myopathy in the family under investigation, since mutations in the beta cardiac myosin gene lead to hypertrophic cardiomyopathy, which is not present in this family. Markesbery et al. (1974) reported autopsy findings of cardiomyopathy in their late-onset hereditary distal myopathy patients, but the pattern of muscle weakness in their patients was different from that in the family described here. Nevertheless, the cardiac myosin genes must be considered as candidate genes for *MPD1* and should be investigated for mutations in affected family members.

Despite the phenotypic differences between the Gowers and other forms of distal myopathy, it is possible that some are allelic, with different mutations producing the different phenotypes. The suggested linkage to chromosome 14 should now allow other investigators researching distal myopathy families to confirm linkage or demonstrate the heterogeneity that might be expected from the variable clinical phenotypes seen in the condition.

## Acknowledgments

This work was supported by the Neuromuscular Foundation of Western Australia and the Edith Cowan University Internal Grants Scheme. We thank Drs. John Mulley, Ted Edkins, and Maija Cohonen-Corish for running polymorphic markers on various chromosomes. We thank the family members for agreeing to take part in the study.

## References

- Edström L, Thornell L, Eriksson A (1980) A new type of hereditary distal myopathy with characteristic sarcoplasmic bodies and intermediate (skeleton) filaments. *J Neurol Sci* 47:171-190
- Edwards JH (1987) Exclusion mapping. *J Med Genet* 24:539-543
- Gowers WR (1902) A lecture on myopathy and a distal form. *Br Med J* 2:89-92
- Jandreski MA, Sole MJ, Liew C-C (1987) Two different forms of beta myosin heavy chain are expressed in human striated muscle. *Hum Genet* 77:127-131

- Markesbery WR, Griggs RC, Leach RP, Lapham LW (1974) Late onset hereditary distal myopathy. *Neurology* 24:127-134
- Mastaglia FL (1991) Genetic myopathies. In: Swash M, Oxbury J (eds) *Clinical neurology*. Churchill Livingstone, Edinburgh, pp 1300-1301
- Matsuoka R, Chambers A, Kimura M, Kanda N, Bruns G, Yoshida M, Takao A (1989) Molecular cloning and chromosomal localization of a gene coding for human cardiac myosin heavy chain. *Am J Med Genet* 29:369-376
- McKusick VA (1992) Mendelian inheritance in man: catalogs of autosomal dominant, autosomal recessive and X-linked phenotypes, 10th ed. John Hopkins University Press, Baltimore and London
- Meredith C, Laing BA, Wilton SD, Mastaglia FL, Robbins P, Kozman H, Honeyman K, et al (1994) Autosomal dominant distal myopathy: linkage on chromosome 14. *Muscle Nerve Suppl* 1:S89
- Miller SA, Dykes DD, Polesky HF (1988) A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Res* 16:1215
- Morton NE (1955) Sequential tests for the detection of linkage. *Am J Hum Genet* 7:277-318
- Mulley JC, Kozman HM, Phillips HA, Gedeon AK, McCure JA, Iles DE, Gregg RG, et al (1993) Refined genetic localization for central core disease. *Am J Hum Genet* 52:398-405
- Neitzel H (1986) A routine method for the establishment of permanent growing lymphoblastoid cell lines. *Hum Genet* 73:320-326
- NIH/CEPH Collaborative Mapping Group (1992) A comprehensive genetic linkage map of the human genome. *Science* 258:62-86
- Nonaka I, Sunohara N, Ishiura S, Satoyoshi E (1981) Familial distal myopathy with rimmed vacuole and lamellar (myeloid) body formation. *J Neurol Sci* 51:141-155
- Ott J (1974) Estimation of the recombination fraction in human pedigrees: efficient computation of the likelihood for linkage studies. *Am J Hum Genet* 26:588-597
- Sambrook J, Fritsch EF, Maniatis T (1989) *Molecular cloning: a laboratory manual*, 2d ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
- Sarfrazi M, Upadhyaya M, Padberg G, Pericak-Vance M, Siddique T, Lucotte G, Lunt P (1989) An exclusion map for facioscapulohumeral (Landouzy-Dejerine) disease. *J Med Genet* 26:481-484
- Walton JN, Gardner-Medwin D (1981) Progressive muscular dystrophy and the myotonic disorders. In: Walton JN (ed) *Disorders of voluntary muscle*. Churchill Livingstone, Edinburgh, London, Melbourne, and New York, pp 481-524
- Weber JL, May PE (1989) Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction. *Am J Hum Genet* 44:388-396
- Weiffenbach B, Dubois J, Storvick D, Tawil R, Jacobsen SJ, Gilbert J, Wijmenga C, et al (1993) Mapping the facioscapulohumeral muscular dystrophy gene is complicated by chromosome 4q35 recombination events. *Nat Genet* 4:165-169
- Weissenbach J, Gyapay G, Dib C, Vignal A, Morissette J, Millasseau P, Vaysseix G, et al (1992) A second-generation linkage map of the human genome. *Nature* 359:794-801
- Welander L (1951) Myopathia distalis tarda hereditaria. *Acta Medica Scand* 141, Suppl 265:1-124
- Wijmenga C, Hewitt JE, Sandkuijl LA, Clark LN, Wright TJ, Dauwerse HG, Gruter A-M, et al (1992) Chromosome 4q DNA rearrangements associated with facioscapulohumeral muscular dystrophy. *Nat Genet* 2:26-30

## Assignment of the human beta tropomyosin gene (TPM2) to band 9p13 by fluorescence in situ hybridisation

C.C. J. Hunt,<sup>1,2</sup> H.J. Eyre,<sup>3</sup> P.A. Akkari,<sup>2</sup> C. Meredith,<sup>1</sup> S.M. Dorosz,<sup>2</sup>  
S.D. Wilton,<sup>2</sup> D.F. Callen,<sup>3</sup> N.G. Laing,<sup>2</sup> and E. Baker<sup>3</sup>

<sup>1</sup> Edith Cowan University, Joondalup, Western Australia;

<sup>2</sup> Australian Neuromuscular Research Institute, QEII Medical Centre, Nedlands, Western Australia;

<sup>3</sup> Centre for Medical Genetics, Department of Cytogenetics and Molecular Genetics, Women's and Children's Hospital, Adelaide (Australia)

**Abstract.** A sequence tagged site (STS) was developed for the human beta tropomyosin gene (TPM2). The STS was used to amplify DNA from somatic cell hybrids to localise TPM2 to human chromosome 9. Genomic clones isolated with the STS

product were in turn used in fluorescence in situ hybridisation to metaphase chromosome spreads to further localise TPM2 to 9p13.

Each muscle-specific gene should be associated with an inherited muscle disease, unless mutations do not occur in that gene, or all mutations of that gene are lethal. Precise mapping of muscle genes therefore becomes important in relation to mapping muscle diseases (Eyre et al., 1993; 1994; Wilton et al., 1994).

Tropomyosin is one of the components of the thin filaments of muscle, binding to actin, and, together with troponin, regulating contraction in a calcium-dependent manner (Clayton et al., 1988). There are at least four distinct tropomyosin genes in humans (TPM1–4) and each may encode different protein isoforms by alternate splicing (Clayton et al., 1988). The alpha tropomyosin gene TPM1 has recently been localised to 15q22 (Eyre et al., 1994) and has been shown to be mutated in some cases of familial hypertrophic cardiomyopathy (Thierfelder et al., 1994). The alpha tropomyosin gene TPM3 has been recently localised to 1q22→q23 (Wilton et al., 1994) and has been shown to be mutated in a family with autosomal dominant nemaline myopathy (Laing et al., 1995).

We describe the mapping of the human beta tropomyosin gene TPM2 (Widada et al., 1988) to 9p13 by development of a

sequence tagged site (STS) (Olson et al., 1989), testing of a somatic cell hybrid panel, and fluorescence in situ hybridization (FISH).

### Materials and methods

#### STS for the TPM2 gene

The cDNA sequence of the skeletal muscle transcript of the TPM2 gene (Widada, et al., 1988) was used, along with the known conserved exon pattern of the tropomyosin genes (Clayton et al., 1988), to design primers for amplification which would amplify across intronic sequences (Wilton et al., 1994). The primers were chosen such that at least one was located in a skeletal muscle specific exon (Clayton, et al., 1988; Widada, et al., 1988) to help avoid amplifying pseudogene sequences. The primers eventually used were in the muscle specific exon VIII<sub>sk</sub> (TPM2.8R: 5'-AGTTACTGTAGTG-GAGGGAG-3') and the common exon VI (TPM2.6F: 5'-ATTCCACCAAA-GAAGATAAA-3') (Widada, et al., 1988). These primers amplified an STS product approximately 1.5 kb long. The amplification conditions for a reaction volume of 25 µl were: 5 µl of 5 × buffer (335 mM Tris HCl, pH 8.8 at 25 °C; 83 mM (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 1 mM dNTPs, 10 mM MgCl<sub>2</sub>, 1 mg gelatin/ml, and 2.25% Triton X-100), 50 ng target DNA and 0.5 U of Tth polymerase (Biotech International). Thermal cycling conditions were: 35 cycles of 94 °C for 30 s, 58 °C for 6 min, with an initial denaturing period at 94 °C for 4.5 min. The 1.5-kb product was partially sequenced using an ABI 373A DNA sequencer to confirm that the STS product did indeed contain parts of the TPM2 gene.

#### Isolation of genomic clones

The genomic library was constructed in λEMBL3 SP6/T7 (Clontech) following a partial *Bam*HI digestion. The library was plated out at 6-fold redundancy and duplicate lifts made using Hybond N+ colony/plaque screening membranes (Amersham). The probe used was the 1.5-kb PCR amplified STS of the TPM2 gene which was radiolabelled with [ $\alpha$ -<sup>32</sup>P]dCTP (Bresatec) using a random-prime kit (Promega). The duplicate lifts were hybridised

Supported by the Neuromuscular Foundation of Western Australia (P.A., S.W., N.L.); the J.H. & J.D. Gunn Medical Research Foundation, the National Health & Medical Research Council of Australia (P.A., N.L., H.E., E.B., D.C.); and the Edith Cowan University Research Fund (C.H., C.M.).

Received 17 October 1994; revision accepted 27 January 1995.

Request reprints from Dr N.G. Laing, Australian Neuromuscular Research Institute, 4th Floor "A" Block, QEII Medical Centre, Nedlands, Western Australia, 6009 (Australia); telephone: 61-9-389-2818; fax: 61-9-389-3487.

with probe at 42 °C for 16 h in a hybridisation solution of 50% formamide, 10% dextran sulphate, 1 M NaCl, 50 mM Tris-HCl pH7.5, 0.1 mg/ml hering sperm DNA. Lifts were washed in 0.1% SDS and 0.1 × SSC at 65 °C for 30 min. Autoradiography was carried out with Dupont Cronex 4 X-ray film with Quanta III autoradiography screens at -80 °C overnight. Plaques showing up in duplicate on the autoradiographs were considered positive. Confirmation of the TPM2 sequence in the genomic clones was ensured by STS PCR assay and partial sequencing.

Pure clones were then amplified by the plate lysate method (Sambrook et al., 1989). Phage were eluted in lambda dilution buffer (100 mM NaCl, 10 mM MgSO<sub>4</sub>·7H<sub>2</sub>O, 100 mM Tris pH 7.5, 0.01% gelatin solution) at 4 °C overnight. Eluates were collected and phage were purified on caesium chloride gradients (Sambrook, et al., 1989) and DNA extracted. DNA integrity was checked by running each sample on a 0.8% agarose gel at 100 V for 30 min and staining with ethidium bromide. Extracted DNA was confirmed to contain the TPM2 gene by STS PCR and sequencing. Whole clones were then used for in situ hybridisation.

#### *In situ hybridisation*

The probes were nicktranslated with biotin-14-dATP and hybridized in situ at a final concentration of 20 ng/μl to metaphases from a total of three normal males. The fluorescence in situ hybridization (FISH) method was as described previously (Eyre et al., 1993; 1994; Wilton et al., 1994) with chromosomes stained before analysis with both propidium iodide (as counter-stain) and DAPI (for chromosome identification). Images of metaphase preparations were captured by a CCD camera and computer enhanced.

#### *Cell lines*

A panel of rodent-human somatic cell hybrids was obtained from the National Institute of General Medical Sciences Human Genetic Mutant Cell Repository [NIGMS/mapping panel #2]. The construction of the mouse-human hybrid cell lines was described by Dubois and Naylor (1992). PCR amplification of the STS was performed with 50 ng of hybrid DNA and the primers TPM2.6F and TPM2.8R as above.

## Results and discussion

### *Somatic cell hybrid panel*

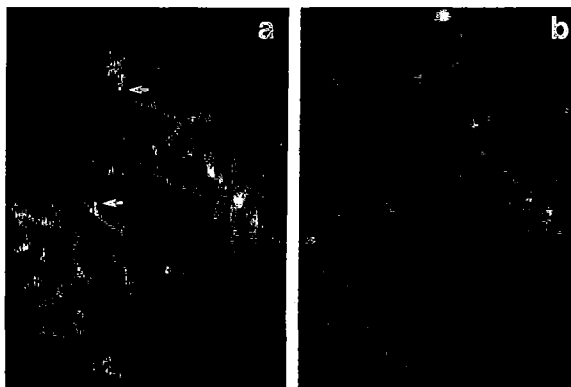
The TPM2 STS gave an amplification product from only the control human DNA and the somatic cell hybrid GM/NA10611 (Dubois et al., 1992) (data not shown). Only chromosome 9 was unique to this cell line indicating that TPM2 is on human chromosome 9.

### *FISH*

The STS PCR product was used as a hybridization probe to the genomic library which has an average insert size of 15 kb.

## References

- Clayton L, Reinach FC, Chumbley GM, MacLeod AR: Organization of the hTM<sub>α</sub> gene: Implications for the evolution of muscle and non-muscle tropomyosins. *J molec Biol* 201:507-515 (1988).
- Dubois BL, Naylor SL: Characterisation of NIGMS human/rodent somatic cell hybrid mapping panel 2 by PCR. *Genomics* 16:315-319 (1992).
- Eyre HJ, Akkari PA, Meredith C, Wilton SD, Callen DC, Kedes L, Laing NG: Assignment of the slow skeletal muscle troponin gene (TNNI1) to 1q32 by fluorescence in situ hybridisation. *Cytogenet Cell Genet* 62:181-182 (1993).
- Eyre H, Akkari PA, Wilton SD, Callen DC, Baker E, Laing NG: Assignment of the human skeletal muscle alpha tropomyosin gene TPM1 to 15q22 by fluorescence in situ hybridisation. *Cytogenet Cell Genet* 69:15-17 (1995).
- Laing NG, Wilton SD, Akkari PA, Dorosz S, Boundy K, Kneebone C, Blumbers P, White S, Watkins H, Love DR, Haan E: A mutation in the alpha tropomyosin gene TPM3 associated with autosomal dominant nemaline myopathy NEM1. *Nature Genetics* 9:75-79 (1995).
- Olson M, Hood L, Cantor C, Botstein D: A common language for physical mapping of the human genome. *Science* 245:1434-1435 (1989).
- Sambrook J, Fritsch EF, Maniatis T: *Molecular Cloning*. Second Edition (Cold Spring Harbour Laboratory, Cold Spring Harbour 1989).
- Thierfelder L, Watkins H, MacRae C, Lamas R, McKenna W, Vosberg H-P, Seidman JG, Seidman CE: α-Tropomyosin and cardiac troponin T mutations cause familial hypertrophic cardiomyopathy: a disease of the sarcomere. *Cell* 77:701-712 (1994).
- Widada JS, Ferraz C, Capony JP, Liautaud JP: Complete nucleotide sequence of the adult skeletal isoform of human skeletal muscle β-tropomyosin. *Nucl Acids Res* 16:3109 (1988).
- Wilton SD, Eyre H, Akkari PA, Watkins HC, MacRae C, Laing NG, Callen DC: Assignment of the human alpha tropomyosin gene TPM3 to 1q22 → q23 by fluorescence in situ hybridisation. *Cytogenet Cell Genet* 68:122-124 (1994).



**Fig.1.** Metaphase showing FISH with the probe TPM2.2. (a) Normal male chromosomes stained with propidium iodide. Hybridization sites on chromosome 9 are indicated by arrows. (b) The same metaphase as (a) stained with DAPI for chromosome identification.

Four positive plaques were picked and respread at lower density enabling individual plaques to be isolated. The STS primer set was used to assay selected pure plaques. Two of those which gave positive STS signals, TPM2.2 and TPM2.4, were further propagated.

TPM2.2 and TPM2.4 were then used for FISH to metaphase spreads from a total of 3 normal males. Twenty-five metaphases from the first normal male were examined for fluorescent signal with TPM2.2. Eighteen of these metaphases showed signal on one or both chromatids of chromosome 9 in the region 9p13 → p21; 95% of this signal was at 9p13.1 (Fig. 1). There was a total of 11 non-specific background dots observed in these 25 metaphases. A similar result was obtained from hybridization of this probe to 15 metaphases from the second normal male and from hybridization using the probe TPM2.4 to 20 metaphases from a third normal male (data not shown).

The localization of TPM2 to 9p13 should be of interest to laboratories interested in positional cloning of inherited muscle disorders, especially those that have been linked to this region.





PERGAMON

Neuromuscular Disorders 9 (1999) 59–65



## Autosomal dominant distal myopathy not linked to the known distal myopathy loci

K.J. Felice<sup>a,\*</sup>, C. Meredith<sup>b,c</sup>, N. Binz<sup>b</sup>, A. Butler<sup>b</sup>, R. Jacob<sup>b</sup>,  
P. Akkari<sup>c,d</sup>, J. Hallmayer<sup>e</sup>, N. Laing<sup>b</sup>

<sup>a</sup>Department of Neurology, University of Connecticut School of Medicine, 263 Farmington Avenue, Farmington, CT, 06030-1840, USA

<sup>b</sup>Australian Neuromuscular Research Institute, Department of Pathology, University of Western Australia,

QEH Medical Centre, Nedlands, Western Australia, 6009 Australia

<sup>c</sup>Centre for Human Genetics, Edith Cowan University, Joondalup, Western Australia 6027, Australia

<sup>d</sup>Division of Neurosciences, Duke University Medical Center, Durham, NC, USA

<sup>e</sup>Centre for Research in Psychiatry, Department of Psychiatry, University of Western Australia, Graylands Hospital, Mt. Claremont, Western Australia 6010, Australia

Received 9 July 1998; received in revised form 29 September 1998; accepted 15 October 1998

### Abstract

The distal myopathies are clinically, pathologically and genetically heterogeneous. Thus far, seven types of distal myopathy have been linked to four chromosome loci. We recently examined four affected members from three generations of an autosomal dominant distal myopathy kindred. A muscle biopsy was performed on the index case. Muscle histopathology showed non-specific myopathic findings including increased variation in fiber size and increased internalized nuclei. No abnormal inclusions or vacuoles were present. Microsatellite markers for the four distal myopathy loci on chromosomes 2, 9 and 14 were studied on affected and several unaffected family members. Affected patients developed distal weakness in anterior foreleg muscles followed by progressive distal upper and proximal lower extremity involvement. Chromosome 2, 9 and 14 regional markers were informative and demonstrated recombinations with affected individuals in the pedigree. The resulting LOD scores obtained from the multipoint analyses gave no evidence of positive linkage to any of the regions and positively excluded (LOD score less than  $-2$ ) all, or virtually all, of the candidate regions examined. This autosomal dominant distal myopathy family does not show evidence of linkage to any of the known distal myopathy loci, suggesting the existence of at least one more distal myopathy locus. Furthermore, the clinical and pathological features appear distinct from other previously described but genetically-undetermined autosomal dominant distal myopathies. © 1999 Elsevier Science B.V. All rights reserved

**Keywords:** Distal myopathy; Autosomal dominant; Chromosome

### 1. Introduction

The familial distal myopathies are a pathologically and genetically heterogeneous group of disorders which share the clinical feature of myopathic weakness beginning in the distal muscles of the upper or lower extremities, or both [1,2]. Previously, these disorders were classified by specific clinical and pathological features into several distinct disorders including Welander, Markesbery, Nonaka and Miyoshi myopathies. Recently, the genetic loci for seven

types of distal myopathy have been elucidated and, therefore, a revised genetic classification of these disorders is now emerging. In this report, we describe a new autosomal dominant distal myopathy family which does not show evidence of linkage to any of the known distal myopathy loci. Furthermore, the clinical and pathological features appear distinct from other previously described but genetically-undetermined autosomal dominant distal myopathies.

### 2. Case reports

The pedigree for this family is shown in Fig. 1. Cases 1–4

\* Corresponding author. Tel.: +1-860-679-3186; fax: +1-860-679-4446; e-mail: felice@nso.uchc.edu

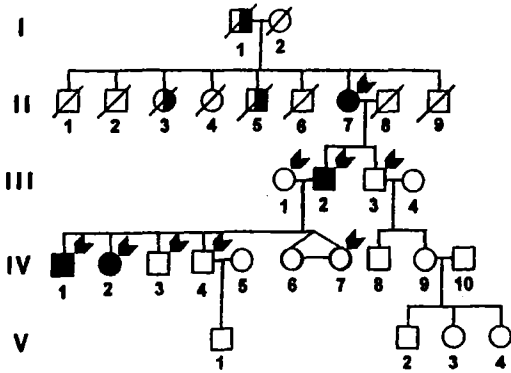


Fig. 1. Pedigree for family. The index case is III-2. Symbols represent the following: squares, males; circles, females; filled, definitely affected; half-filled, probably affected based on history; slashes, deceased; and arrow heads, patients examined.

were examined by one of the authors (KJF). Cases 5–7 were not examined but were probably affected based on the history.

#### 2.1. Case 1 (III-2)

This 57-year-old man first became aware of weakness while working on a volunteer fire rescue and ambulance service at age 35 years. Prior to this, he described himself as strong and athletic. The initial problems included frequent tripping due to foot drop and difficulty with stair climbing. The leg weakness progressed slowly since that time. For the past 2 years, he has had occasional falls due to gait imbalance. Recently, he started using a cane and bilateral ankle-foot orthoses. Other problems included progressive exertional dyspnea, bilateral cataract surgery (capsular cataracts) at age 44 years, gouty arthritis, and hypertension. He denied double vision, swallowing problems, weight loss, upper extremity weakness, muscle cramps, sensory loss or paresthesias, passage of dark-colored urine, exercise- or cold-induced weakness, or voiding difficulties. On examination, his height and weight were 73 inches and 254 lbs. The heart and lung sounds were normal. The forced vital capacity (FVC) was 1.7 l (30% of predicted). There were no dysmorphic features, scapular winging, scoliosis, or foot deformities. The cranial nerve examination was normal with no weakness of eyelid, ocular, facial, pharyngeal, lingual or neck muscles. There was no sternocleidomastoid (SCM) muscle atrophy. Muscle bulk, tone and strength were normal in the upper extremities except for weakness (Medical Council Research grade; MRC) of the arm extensors (3), finger extensors (4), and the flexor digitorum profundus of digits 5 (2). There was marked foreleg and foot muscle atrophy (Fig. 2). He needed to push-off with both hands to stand from a chair. The gait was unsteady with a combined waddle and steppage quality.

The MRC-graded lower limb muscle strength was as follows: hip flexors (3), knee flexors (4), knee extensors (4), foot dorsiflexors (1), planter flexors (2), toe flexors (1), and toe extensors (1). There was no myotonia. Myotatic reflexes were hypoactive or absent. No pathological reflexes were present. The sensory examination was normal. The creatine kinase (CK) was 534 U/l ( $n < 269$ ). Other laboratory studies including thyroid stimulating hormone, thyroxine, carnitine, anti-nuclear antibody titer, lactate and aldolase were normal. Leukocyte DNA studies for the myotonic dystrophy CTG trinucleotide repeat expansion on 19q13 were negative (Athena Laboratories, Worcester, MA). The electrocardiogram (ECG) and chest X-ray were normal. Sensory and motor nerve conduction, and repetitive motor nerve stimulation studies were normal. Concentric needle electromyography (CNE) revealed slightly increased insertional activity (complex repetitive discharges), rare fibrillations and positive waves, and increased recruitment of short duration, low amplitude, polyphasic motor unit action potentials (i.e. myopathic recruitment). A left vastus lateralis muscle biopsy revealed non-specific myopathic abnormalities including increased fiber size variation, increased internalized nuclei, and mildly increased endomysial fat and connective tissue (Fig. 3). No rimmed vacuoles, inclusions, cores or rods were present. There was no abnormal storage of lipid or glycogen. Electron microscopy revealed normal-appearing mitochondria with no abnormal inclusions or filaments. These results were similar to a first muscle biopsy done on this patient at another Center several years prior.

#### 2.2. Case 2 (IV-1)

This 35-year-old son of case 1 began to notice weakness while in the army at age 20 years. He recalled having problems walking on his heels, performing sit-ups and push-ups, and running. The leg weakness slowly progressed. More recently, he began to suffer occasional falls due to leg and foot weakness. He denied double vision, swallowing problems, breathing problems, arm weakness, sensory loss or cramps. He also has small bilateral capsular cataracts. On examination, the height and weight were 63 inches and 287 lbs. The FVC was 4.5 l (82% of predicted). There were no dysmorphic features or scapular winging. The abdomen was protuberant and the lumbar lordosis was exaggerated. There was no ocular, facial, lingual, pharyngeal or neck muscle weakness. The gait had both a waddling and steppage quality. There was moderate foreleg and foot muscle atrophy. Upper extremity muscle strength was normal except for mild finger extensor weakness. The MRC-graded lower limb muscle strength was as follows: hip flexors (3), knee flexors (5), knee extensors (5), foot dorsiflexors (3), planter flexors (4), toe flexors (2), and toe extensors (2). He was able to stand on his toes but not his heels. There was no myotonia. The myotatic reflexes and sensory examinations were normal. Myotonic dystrophy DNA studies were negative. The CK was 1701 U/l ( $n < 269$ ). The ECG was nor-

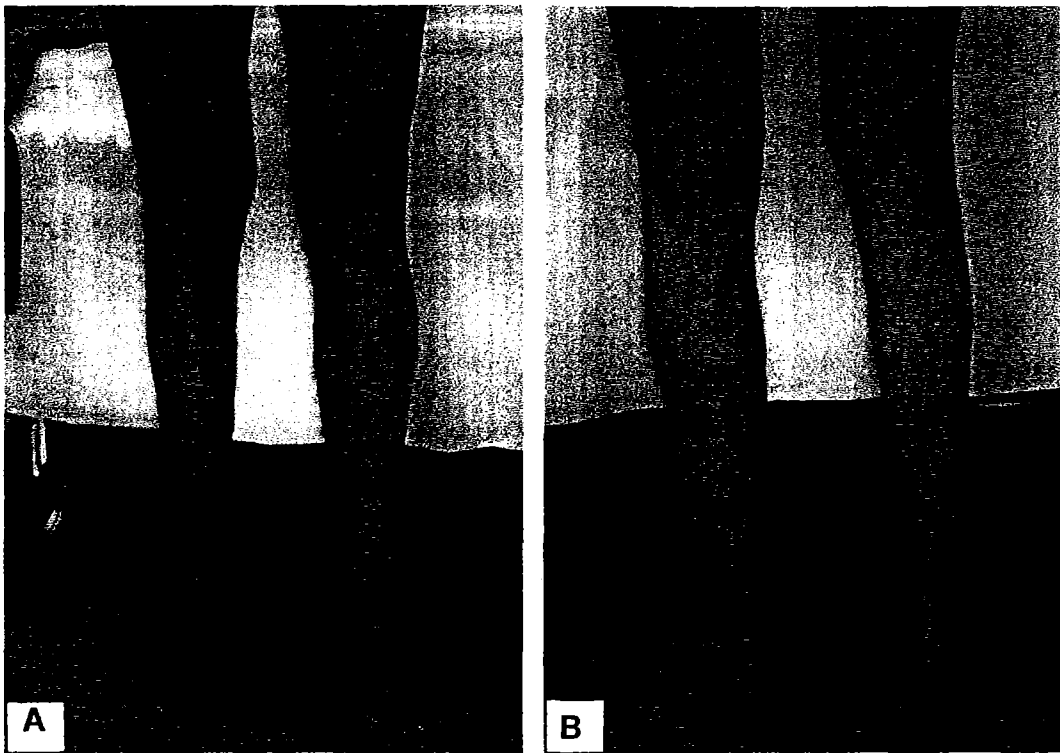


Fig. 2. Photograph of patient 1 showing marked atrophy of anterior (a) and posterior (b) foreleg muscles.

mal. Nerve conduction studies revealed absent peroneal compound muscle action potentials with otherwise normal sensory and motor studies. CNE revealed rare fibrillations, positive waves, and myopathic-type recruitment in proximal and distal limb muscles.

#### 2.3. Case 3 (IV-2)

This 32-year-old daughter of case 1 denied any weakness or myalgias. Her past medical history was remarkable for mild hypothyroidism which was treated with Synthroid. A recent ophthalmologic examination was negative for cataracts. On examination, her weight and height were 151 lbs and 67 inches. The cranial nerve examination was normal. There was no SCM atrophy. The gait was essentially normal except that she could not walk on her heels. Her foreleg and foot muscles were thin. She had minimal difficulty standing from a full squat. Upper extremity muscle strength was normal. There was no scapular winging. There was no myotonia. In the lower extremities, there was mild weakness of the hip flexors, and foot and toe extensors. The sensory and reflex examinations were normal. Nerve conduction studies and CNE were normal. The CK was 406 U/l ( $n < 230$ ).

#### 2.4. Case 4 (II-7)

This 83-year-old mother of case 1 admitted to mild difficulty climbing stairs for the past 2–3 years. The past medical history was remarkable for rheumatoid arthritis and mild late-onset cataracts. The examination was remarkable for moderate difficulty standing from a chair, inability to heel-walk, and mild to moderate weakness of the hip flexors and foot dorsiflexors.

#### 2.5. Case 5 (I-1)

This is the maternal grandfather of case 1. He and his wife were born in Poland. There is no family history of consanguinity. He suffered from coronary artery disease and died at age 83 years. The family recalls that he walked with a cane and had difficulty climbing stairs, both disabilities were believed to be due to leg weakness. Specific details of his medical history were unknown.

#### 2.6. Case 6 (II-3)

This is a maternal aunt of case 1. She died at age 93 years.

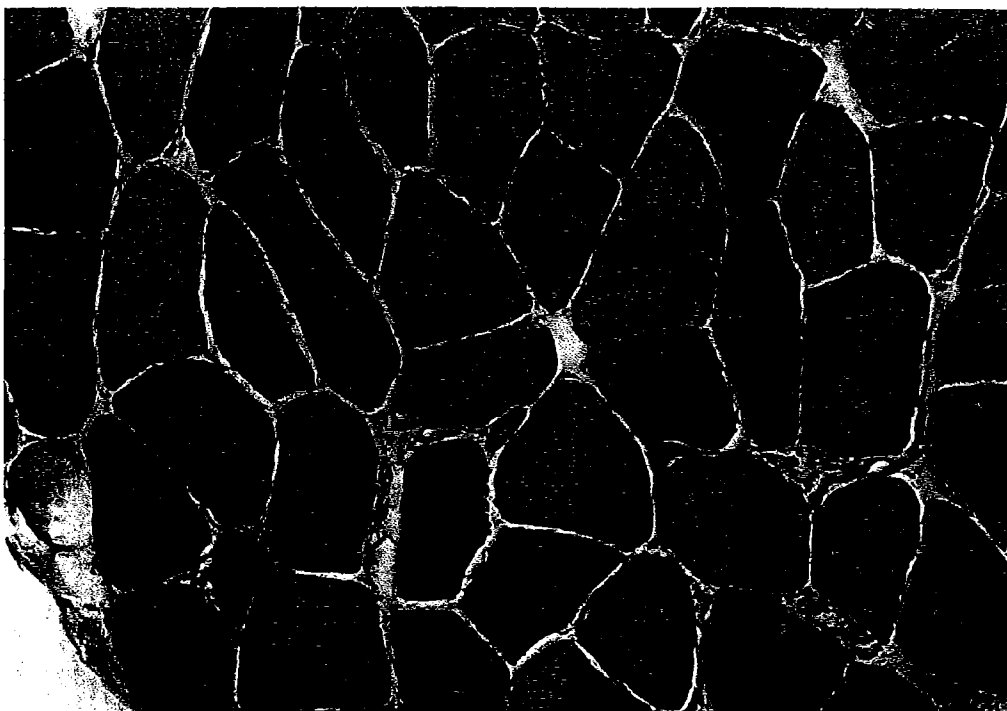


Fig. 3. Left vastus lateralis muscle biopsy from patient 1 showing non-specific myopathic changes including increased fiber size variation and increased internalized nuclei (hematoxylin and eosin,  $\times 100$ ).

From her mid-adult life on, she had problems walking up and down stairs, leg weakness, and difficulty rising from the floor. Specific details of her medical history are unknown.

### 2.7. Case 7 (II-5)

This is a maternal uncle of case 1. He died at age 72 years. He was diagnosed with a progressive form of 'muscular dystrophy' which began in his early 40's and primarily affected his legs. He required a wheelchair in his 50's. Specific details of his medical history are unknown.

The following family members were found to have normal clinical examinations by one of the authors (KJF): III-1, III-3, IV-3, IV-4, and IV-7. The following family members were not examined but, by history, appear to have not been affected or show no signs of weakness at this time: I-2, II-1, II-2, II-4, II-6, II-8, II-9, IV-6, IV-8, IV-9, V-1, V-2, V-3, V-4.

## 3. DNA analyses

### 3.1. Methods

#### 3.1.1. Microsatellite analysis

The following microsatellite markers for the four known

distal myopathy loci were analyzed: (1) Miyoshi recessive distal myopathy locus on chromosome 2: D2S303, D2S291, D2S292, D2S2111, D2S2109, D2S145, D2S286; (2) autosomal dominant tibial muscular dystrophy locus on chromosome 2: D2S300, D2S148, D2S364, D2S2273, D2S350, D2S152, D2S389; (3) Nonaka recessive distal myopathy locus on chromosome 9: D9S319, D9S43, D9S248 and D9S276; and (4) autosomal dominant distal myopathy MPD1 locus on chromosome 14: D14S72, D14S283, MYH7.1, D14S264, D14S80, D14S275, D14S262, D14S252, D14S257, D14S49. The microsatellites were amplified and analyzed using standard conditions [3]. The annealing temperature was 55°C for all markers except D14S257 at 52°C; D2S2281, D2S2273, D2S2366, D2S389 at 57°C; D14S49, D14S80, D14S252, D14S253, D14S262, D14S275, D14S283 at 58°C; D14S264 at 60°C; D14S72 at 61°C and MYH7.1 at 64°C; extension temperature was 72°C for all markers. The PCR reactions were performed using Tth Plus enzyme (Biotech International) in a reaction volume of 20  $\mu$ l with 50 ng genomic DNA and final concentrations of 1 mM  $MgCl_2$ , 20  $\mu$ M dNTPs and 1  $\mu$ Ci dCTP (Amersham), for 35 cycles. The amplified samples were electrophoresed on 6% 19:1 acrylamide/bisacrylamide gels at 1300 V for 3 h. Gels were exposed to Dupont Cronex X-ray film at room temperature overnight.

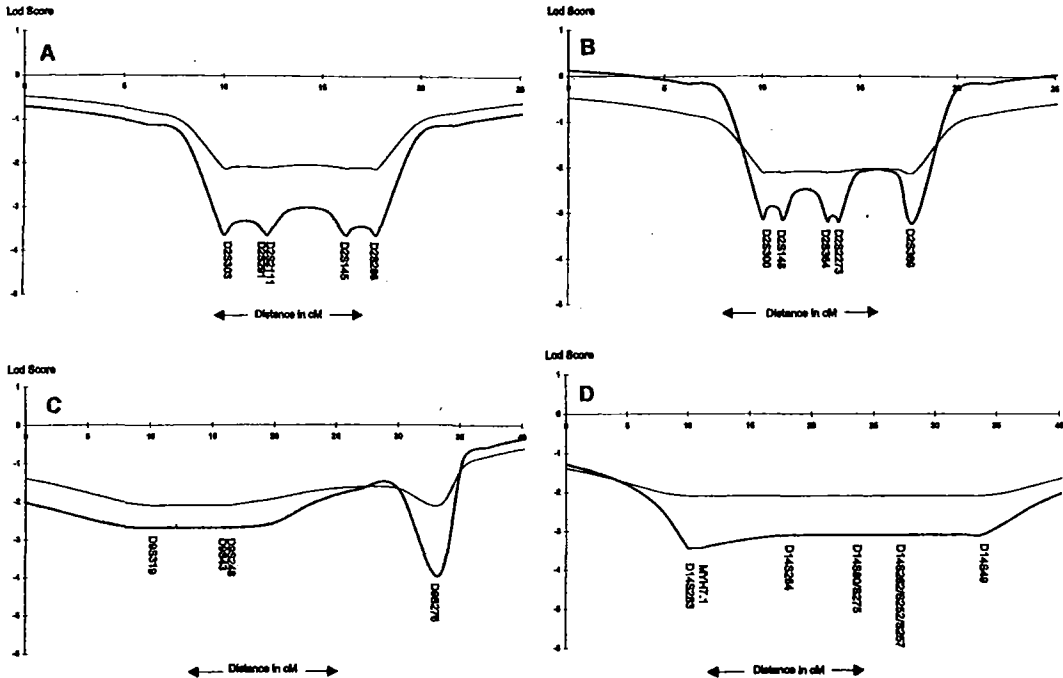


Fig. 4. Multipoint linkage analysis of distal myopathy in a single family from Connecticut against fixed maps of markers for the Miyoshi distal myopathy (a) and tibial muscular dystrophy (b) regions on chromosome 2, the Nonaka distal myopathy region on chromosome 9 (c), and the autosomal dominant MPD1 region on chromosome 14 (d) obtained as described in Section 3.1. For all regions, distal myopathy was treated as a dominant disease with a penetrance of 70% and a gene frequency of 0.001. Two sets of calculations were performed for all regions: either the phenotypic status of unaffected individuals was considered to be unknown (thin line), or phenotypic information of unaffected individuals was included (thick line). At each locus, one marker was arbitrarily set at 10 cM and the other markers were then set according to the chromosome maps.

### 3.1.2. Multipoint linkage analysis

Multipoint LOD scores were calculated between a putative disease locus and a marker locus using the program GENEHUNTER [4]. Distances between markers on chromosome 2 were fixed according to the maps given by Gyapay et al. [5], the maps at the Genethon web site (<http://www.genethon.fr/>), and those of Bashir et al. [6]. Distances on chromosomes 9 and 14 were obtained from the maps in Gyapay et al. [5], and the maps at the Genethon (<http://www.genethon.fr/>) and Co-operative Human Linkage Centre (<http://www.chlc.oorg/>) web sites. In the linkage calculations the gene frequency was assumed to be 0.001. The number of affected family members is not large enough to draw any firm conclusions in regard to the penetrance. The penetrance was therefore set to the conservative value 70%. In addition, calculations were carried out ignoring the phenotypic status of the unaffected individuals ('penetrance free model').

### 3.2. Results

Markers at all four distal myopathy loci were informative and demonstrated recombinations with affected individuals

in the pedigree. The resulting LOD scores obtained from the multipoint analyses were less than  $-2$  for virtually the entire regions tested under either model (Fig. 4). All four distal myopathy regions can therefore be excluded for linkage and the disease is therefore not linked to any of the known distal myopathy loci.

### 4. Discussion

The clinicopathological features of the above described family included: (1) slowly progressive autosomal dominant distal myopathy with known affected members spanning four generations, (2) onset of anterior foreleg muscle weakness in the second to third decades with progression to other distal and proximal muscles over time, (3) no cranial or neck muscle involvement, (4) respiratory muscle involvement in the index case, (5) early-onset cortical cataracts in patients 1 and 2, (6) no clinical or EMG evidence of myotonia, (7) mild to moderate elevations (1.8–6.3 times the upper limit of normal) in CK values, (8) non-specific muscle histopathology with no evidence of vacuoles or inclusions, and (9) no genetic linkage to any of the known distal myopathy loci.

The distal myopathies have been classified into autosomal dominant and recessive forms [1,2]. The dominant conditions include Welander [7–9] and Markesbery [10,11] myopathies, a Finnish form of tibial muscular dystrophy [11–14] and autosomal dominant distal myopathy (MPD1) [3]. The recessive forms include Miyoshi [15–19] and Nonaka [20,21] myopathies. In addition, less common forms of hereditary distal myopathy have been described including several forms of infantile- and childhood-onset distal myopathies [1,2,22] distal myopathy with abnormal accumulation of desmin [23], and hereditary inclusion body myositis [24].

Recently, the genetic loci for seven types of distal myopathy have been assigned to four distinct loci and, therefore, a revised genetic classification of these disorders is now emerging (Table 1). In 1995, Laing et al. localized an autosomal dominant distal myopathy kindred to chromosome 14 (MPD1) [3]. In the same year, Bejaoui et al. found linkage of Miyoshi myopathy, an autosomal recessive disorder, to chromosome 2p12–14 [15]. Miyoshi myopathy appears to be allelic to two other autosomal recessive disorders, limb girdle muscular dystrophy 2B (LGMD2B) and a newly-described distal anterior compartment myopathy [6,25,26]. Moreover, the two disease phenotypes, Miyoshi myopathy and LGMD2B, have both been seen in one Canadian aboriginal kindred and in a family from Russia [16,17,19]. In a recent study by Liu et al., the gene product for Miyoshi myopathy, LGMD2B, and anterior compartment myopathy has been discovered and named dysferlin [18]. Tibial muscular dystrophy was recently localized to chromosome 2q31–33 and the late-onset distal myopathy of Markesbery links to the same region [11,13]. Interestingly, a proximal limb-girdle myopathy rather than a distal myopathy may develop in patients who are homozygous for the tibial muscular dystrophy gene defect on 2q31–33 [14]. In 1996, Mitrani-Rosenbaum et al. found linkage of hereditary inclusion

body myopathy, a disorder with quadriceps-sparing and prominent distal weakness, to chromosome 9p1–q1 [24]. Recently, Ikeuchi et al. have found similar linkage to chromosome 9 for Japanese families with autosomal recessive distal myopathy (Nonaka myopathy) suggesting that this disorder and hereditary inclusion body myopathy are allelic disorders [21]. Genetic localization of Welander myopathy is not yet published but recent work by Ahlberg et al. has excluded linkage to the Miyoshi or MPD1 distal myopathy loci [9].

In reference to the present family, five of the six major forms of distal myopathy are excluded based on DNA analyses (Markesbery, Miyoshi, MPD1, Nonaka, tibial muscular dystrophy). Moreover, several clinicopathological features seem to distinguish the present family from the dominantly-inherited distal myopathies. First, vacuolar changes on muscle histopathology are often observed in Welander, Markesbery and the Finnish tibial muscular dystrophies. Vacuoles were not found in two muscle specimens on patient 1 from this family. Second, weakness is usually begins after the age of 35 years in Welander, Markesbery and Finnish tibial muscular dystrophies. In the current family, the symptoms began in the second to third decade. Third, the pattern of weakness is also a distinguishing feature in at least two of the autosomal dominant distal myopathies. In Welander distal myopathy, weakness begins in the hand muscles and progresses over many years to involve the distal lower extremity muscles. Proximal weakness is uncommon in this condition. Finnish tibial muscular dystrophy patients usually present with weakness of the anterior tibial muscles and, later, the long toe extensors. Upper extremity and proximal lower extremity involvement is quite rare. The pattern of weakness in Markesbery myopathy, which includes onset in the anterior tibial muscles followed by progressive weakness of intrinsic hand and proximal limb muscles, more closely follows that of the current family. Some differences are noteworthy, however.

Table 1  
Tentative classification of distal myopathies based gene location<sup>a</sup>

Myopathy	Mode of inheritance	Gene location	References
Autosomal recessive distal myopathy (Miyoshi myopathy) <sup>b</sup>	AR	2p12–14	[4,8]
Distal anterior compartment myopathy	AR	2p12–14	[7]
Tibial muscular dystrophy (Udd myopathy)	AD	2q31–33	[11–13,22]
Late-onset distal myopathy (Markesbery myopathy)	AD	2q31–33	[13,21]
Nonaka myopathy	AR	9p1–q1	[15,23]
Hereditary inclusion body myopathy	AR	9p1–q1	[14]
Autosomal dominant distal myopathy	AD	14	[3]
Welander myopathy	AD	?	[16,19,20]

<sup>a</sup>Classification excludes rare and genetically-undetermined disorders including the kindred in this report (see text), and other myopathies (e.g. myotonic dystrophy) in which distal weakness may be a prominent feature (see Ref. [2]).

<sup>b</sup>Limb girdle muscular dystrophy type 2B also localizes to 2p12–14.

Cases 1 and 2 in this series had mild weakness of finger extensors, finger flexors, or both but neither had involvement of intrinsic hand muscles which is in contrast to most of the patients described by Marksbery et al. [10]. Moreover, cardiomyopathy was not present in this family but was noted in one of the seven cases reported by Marksbery et al. [10]. Fourth, other inconsistent features in this family including respiratory muscle involvement and early cataracts may be unique to this form of distal myopathy. Finally, of the dominantly-inherited distal myopathies, Welander myopathy could not be excluded based on DNA studies as the locus for this myopathy has not been elucidated [9]. However, in addition to the distinctive clinical features as noted above, Welander myopathy is usually restricted geographically to Sweden and parts of Finland, thereby, further distinguishing this myopathy from the family described in this report.

Other dominantly-inherited muscular dystrophies in which distal limb muscle weakness may be prominent including myotonic dystrophy and facioscapulohumeral muscular dystrophy warrant further mention. Despite the findings of distal weakness and early cataracts in this family, myotonic dystrophy was excluded based on the absence of clinical and EMG myotonia, and the negative DNA studies. Facioscapulohumeral muscular dystrophy was excluded based on the absence of characteristic clinical features including facial weakness and scapular winging.

In conclusion, this autosomal dominant distal myopathy family does not show evidence of linkage to any of the known distal myopathy loci, suggesting the existence of at least one more distal myopathy locus. Furthermore, the clinical and pathological features appear distinct from other previously described but genetically-undetermined autosomal dominant distal myopathies. The identification of larger families with this phenotype will be needed in order to provide further genetic elucidation of this seemingly distinct distal myopathy.

## References

- [1] Griggs RC, Marksbery WR. Distal myopathies. In: Engel AG, Franzini-Armstrong C, editors. *Myology: basic and clinical*. New York: McGraw-Hill, 1994:1246–1257.
- [2] Barohn RJ. Distal myopathies and dystrophies. *Semin Neurol* 1993;13:247–255.
- [3] Laing NG, Laing BA, Meredith C. et al. Autosomal dominant distal myopathy: linkage to chromosome 14. *Am J Genet* 1995;56:422–427.
- [4] Kruglyak L, Daly MJ, Reeve-Daly MP, Lander ES. Parametric and nonparametric linkage analysis: a unified multipoint approach. *Am J Hum Genet* 1996;58:1347–1363.
- [5] Gyapay G, Morissette J, Vignal A. et al. The 1993–94 Genethon human genetic linkage map. *Nat Genet* 1994;7:246–339.
- [6] Bashir R, Strachan T, Keers S. et al. A gene for autosomal recessive limb-girdle muscular dystrophy maps to chromosome 2p. *Hum Mol Genet* 1994;3:455–457.
- [7] Welander L. Myopathia distalis tarda hereditaria. *Acta Med Scand* 1951;141 (S265):1–124.
- [8] Borg K, Ahlberg G, Borg J, Edstrom L. Welander's distal myopathy: clinical, neurophysiological and muscle biopsy observations in young and middle aged adults with early symptoms. *J Neurol Neurosurg Psychiatry* 1991;54:494–498.
- [9] Ahlberg G, Borg K, Edstrom L, Anvret M. Welander distal myopathy is not linked to other defined distal myopathy loci. *Neuromusc Disord* 1997;7:256–260.
- [10] Marksbery WR, Griggs RC, Leach RP, Lapham LW. Late onset hereditary distal myopathy. *Neurology* 1974;23:127–134.
- [11] Haravuori H, Makela-Benga P, Figlewicz DA. et al. Tibial muscular dystrophy and late-onset distal myopathy are linked to the same locus on chromosome 2q. *Neurology* 1998;50:A186.
- [12] Udd B, Partanen J, Halonen P. et al. Tibial muscular dystrophy: late adult-onset distal myopathy in 66 Finnish patients. *Arch Neurol* 1993;50:604–608.
- [13] Haravuori H, Makela-Benga P, Udd B. et al. Assignment of the tibial muscular dystrophy locus to chromosome 2q31. *Am J Hum Genet* 1998;62:620–626.
- [14] Udd B. Limb-girdle type muscular dystrophy in a large family with distal myopathy: homozygous manifestation of a dominant gene. *J Med Genet* 1992;29:383–389.
- [15] Bejaoui K, Hirabayashi K, Hentati F. et al. Linkage of Miyoshi myopathy (distal autosomal recessive muscular dystrophy) locus to chromosome 2p12–14. *Neurology* 1995;45:768–772.
- [16] Bulayeva K, Tsuji S. Clinical and molecular analysis of a large family with three distinct phenotypes of progressive muscular dystrophy. *Brain* 1996;119:1895–1909.
- [17] Illarioshkin S, Ivanovasmolenskaya I, Tanaka H. et al. Refined genetic location of the chromosome 2p-linked progressive muscular dystrophy gene. *Genomics* 1997;42:345–348.
- [18] Liu J, Aoki M, Illa I. et al. Dysferlin, a novel skeletal muscle gene, is mutated in Miyoshi myopathy and limb girdle muscular dystrophy. *Nat Genet* 1998;20:31–36.
- [19] Weiler T, Greenberg CR, Nylen E. et al. Limb-girdle muscular dystrophy and Miyoshi myopathy in an aboriginal Canadian kindred map to LGMD2B and segregate with the same haplotype. *Am J Hum Genet* 1996;59:872–878.
- [20] Nonaka I, Sunohara N, Ishiura S, Satoyoshi E. Familial distal myopathy with rimmed vacuole and lamellar (myeloid) body formation. *J Neurol Sci* 1981;51:141–155.
- [21] Ikeuchi T, Asaka T, Saito M. et al. Gene locus for autosomal recessive distal myopathy with rimmed vacuoles maps to chromosome 9. *Ann Neurol* 1997;41:432–437.
- [22] Magee KR, DeJong RN. Hereditary distal myopathy with onset in infancy. *Arch Neurol* 1965;13:387–390.
- [23] Horowitz SH, Schmalbruch H. Autosomal dominant distal myopathy with desmin storage: a clinicopathologic and electrophysiologic study of a large kinship. *Muscle Nerve* 1994;17:151–160.
- [24] Mitran-Rosenbaum S, Argov Z, Blumenfeld A. et al. Hereditary inclusion body myopathy maps to chromosome 9p1-q1. *Hum Mol Genet* 1996;5:159–163.
- [25] Bashir R, Keers S, Strachan T. et al. Genetic and physical mapping at the limb-girdle muscular dystrophy locus (LGMD2B) on chromosome 2p. *Genomics* 1996;33:46–52.
- [26] Illa I, Serrano C, Gallardo E. et al. Distal anterior compartment myopathy: a new severe dystrophic phenotype linked to chromosome 2p13. *Neurology* 1998;50:A186.

## **APPENDIX B**

### **Lod Scores for the Segregation of the WAMPD and QMPD Disease Loci with Marker Loci**

**Table B4.1    *Lod Scores for Linkage between the MPD1  
Locus and Marker Loci.***

**Table B5.1    *Lod Scores for Linkage between the QMPD  
Locus and Marker Loci.***



**Table B4.1     *Lod Scores for Linkage between the MPD1 Locus and Marker Loci.***

Analysis includes data from all family members. Chromosome position originally determined as % of pter-qter from NIH/CEPH Collaborative Mapping Group (1992), then progressively updated from CHLC Report (Murray *et al.*, 1993 and 1994), Cox *et al.* (1994) and "The National Centre For Biotechnology Information" GB4 GeneMap'98 (<http://www.ncbi.nlm.nih.gov/>).

D-numbers indicate anonymous DNA fragments (D-segments), other markers are associated with the genes: CRP - C-reactive protein; AT3 -antithrombin II; ACTN2 - alpha-2 actinin; APOB - apolipoprotein B; HOXD4 - homeobox D4; APC - adenomatosis polyposis coli; SCA1 - spino cerebellar ataxia; DMDL/Duchenne muscular dystrophy like (utrophin); GCK - Glucokinase; CFTR/cystic fibrosis transmembrane conductance regulator; CD3D - delta polypeptide antigen in TYT3 complex; RB1 - retinoblastoma 1; MYH7 - myosin beta chain (heavy polypeptide); TPM1 - tropomyosin 1 (alpha); GX-Alu - Alu polymorphism intragenic in NF1 gene; HOXB6 - homeobox B6; APOC2 - apolipoprotein C-II; DM - dystrophia myotonica protein kinase.

Two of the markers associated with genes, APOB and AT3, are non-microsatellite polymorphisms. The remainder are microsatellites.

**Table B4.1** *Lod Scores for Linkage between the MPD1 Locus and Marker Loci on Chromosomes 1, 2, 3, 4 and 5.*

Chromosome %pter-qter	Locus	0.00	0.05	0.10	0.20	( $\theta$ ) 0.30	0.40	Equivalent No. Recombinations	Equivalent No. Informative Meioses
1.05	D1S80	-99.99	-2.75	-1.59	-0.59	-0.15	-0.02	3.3	6.5
1.19	D1S164	-99.99	-2.88	-1.52	-0.59	-0.35	0.14	1.7	5.8
1.44	D1S223	-99.99	-1.27	-0.49	0.08	0.22	0.23	2.1	6.5
1.55	D1S185	-99.99	-4.20	-2.79	-1.48	-0.78	-0.32	6.3	12.6
1.57	D1S176	-99.99	-0.09	0.09	0.14	0.09	0.02	0.3	1.7
1.59	CRP	-99.99	-1.31	-0.81	-0.41	-0.24	-0.14	1.8	3.6
1.62	D1S104	-99.99	-4.34	-2.91	-1.54	-0.80	-0.32	6.6	13.1
1.67	AT3	0.13	0.10	0.08	0.05	0.02	0.01	0.0	0.4
1.89	ACTN2	-99.99	-1.67	-0.90	-0.26	-0.03	0.03	1.5	3.8
2.07	APOB	-99.99	-2.99	-1.89	-0.89	-0.41	-0.14	4.3	8.5
2.55	D2S44	-99.99	-3.73	-2.35	-1.10	-0.49	-0.16	5.3	10.6
2.66	HOXD4	-99.99	-3.04	-1.92	-0.91	-0.41	-0.14	4.0	9.0
2.94	D2S102	-99.99	-2.80	-1.74	-0.81	-0.38	-0.13	3.9	7.8
3.00	D3S1307	-99.99	-2.64	-1.79	-0.98	-0.52	-0.22	4.0	8.1
3.07	D3S1304	-99.99	-3.61	-2.22	-0.97	-0.38	-0.09	5.0	10.0
3.38	D3S1285	-99.99	-0.77	-0.22	0.19	0.28	0.20	2.4	7.9
3.52	D3S1278	-99.99	-3.22	-2.04	-0.96	-0.42	-0.13	4.6	9.2
3.73	D3S1279	-99.99	-2.24	-1.20	-0.37	-0.05	0.04	1.6	4.1
3.81	D3S1282	-99.99	-1.96	-1.14	-0.41	-0.10	0.02	1.1	2.6
3.89	D3S1262	-99.99	-2.52	-1.65	-0.82	-0.39	-0.14	3.7	7.4
3.99	D3S1311	-99.99	-2.54	-1.62	-0.78	-0.37	-0.13	3.6	7.3
4.01	D4S45	-99.99	-3.75	-2.39	-1.19	-0.62	-0.27	5.4	10.8
4.28	D4S174	-99.99	-0.43	0.02	0.31	0.37	0.12	1.7	6.6
4.98	D4S171	-99.99	-3.61	-2.22	-0.97	-0.38	-0.09	5.0	10.0
5.18	D5S268	0.82	0.73	0.64	0.42	0.27	0.14	0.0	2.6
5.38	D5S112	-99.99	-0.34	0.11	0.38	0.37	0.23	0.9	4.6
5.39	D5S39	-99.99	-1.20	-0.46	0.10	0.25	0.20	2.1	7.0

Table B4.1 *Lod Scores for Linkage between the MPD1 Locus and Marker Loci on Chromosomes 5 (cont), 6, 7, 8, 9, 10, 11 and 12.*

Chromosome Position	Locus	0.00	0.05	Recombination 0.10	Fraction 0.20	( $\theta$ ) 0.30	0.40	Equivalent No. Recombinations	Equivalent No. Informative Meioses
5.60	APC	-99.99	-3.62	-2.24	-1.02	-0.44	-0.14	5.1	10.1
5.80	D5S210	-99.99	-3.24	-2.09	-1.01	-0.46	-0.15	4.7	9.4
6.13	SCA1	-99.99	-0.73	-0.21	0.15	0.24	0.19	0.7	2.0
6.85	DMDL	-99.99	-0.09	0.09	0.14	0.09	0.02	0.3	1.7
7.05	D7S472	-99.99	-4.29	-2.86	-1.51	-0.78	-0.32	6.4	12.9
7.18	D7S435	-99.99	-1.45	-0.69	-0.09	0.10	0.12	5.3	13.2
7.20	D7S474	-99.99	-0.71	-0.24	0.11	0.18	0.13	1.5	5.1
7.26	GCK	-99.99	-1.99	-1.19	-0.50	-0.19	-0.04	2.7	5.4
7.47	D7S440	-99.99	-3.88	-2.47	-1.17	-0.53	-0.17	5.6	11.1
7.74	CFTR	-99.99	-1.56	-0.97	-0.43	-0.17	-0.04	2.2	4.4
8.41	D8S166	-99.99	-4.97	-3.27	-1.68	-0.85	-0.33	7.4	14.6
8.53	D8S84	1.11	0.99	0.86	0.61	0.38	0.18	0.0	3.7
8.65	D8S200	1.09	0.97	0.85	0.61	0.38	0.18	0.0	3.6
8.70	D8S198	-99.99	-3.73	-2.35	-1.10	-0.49	-0.16	5.3	10.6
9.20	D9S104	-99.99	-0.41	0.04	0.32	0.33	0.21	2.8	9.3
9.29	D9S15	-99.99	-0.25	0.05	0.28	0.31	0.21	2.6	8.5
9.55	D9S53	-99.99	-2.88	-1.77	-0.78	-0.30	-0.07	4.0	8.0
10.18	D10S28	-99.99	-3.81	-2.46	-1.25	-0.66	-0.29	5.5	11.1
11.09	D11S875	-99.99	-2.46	-1.61	-0.83	-0.44	-0.19	2.6	5.2
11.52	D11S873	-99.99	-3.69	-2.31	-1.08	-0.48	-0.16	5.2	10.4
11.70	CD3D	1.39	1.25	1.10	0.80	0.49	0.22	0.0	4.6
11.76	D11S836	-99.99	-2.77	-1.70	-0.76	-0.34	-0.12	3.8	7.6
12.47	D12S43	-99.99	-2.53	-1.58	-0.71	-0.27	-0.06	3.9	7.8
12.95	D12S60	-99.99	-2.75	-1.75	-0.94	-0.46	-0.06	4.1	8.3

**Table B4.1** *Lod Scores for Linkage between the MPD1 Locus and Marker Loci on Chromosomes 13, 14, 15 and 16.*

Chromosome	Position	Locus	0.00	0.05	0.10	0.20	0.30	0.40	Equivalent No. Recombinations	Equivalent No. Informative Meioses
			Recombination Fraction					( $\theta$ )		
13.41	RB1		0.13	0.10	0.08	0.05	0.02	0.01	0.0	0.4
13.74	D13S71		-99.99	-1.47	-0.75	-0.19	0.03	0.10	1.4	3.4
14.08	D14S72		-3.13	0.03	0.23	0.31	0.26	0.15	0.9	4.1
14.10	D14S283		-1.93	1.14	1.24	1.11	0.81	0.43	1.1	8.8
14.10	D14S50		2.31	2.12	1.93	1.51	1.06	0.57	0.0	7.7
14.11	MYH7		2.60	2.36	2.12	1.61	1.08	0.53	0.0	8.6
14.15	D14S64		2.60	2.36	2.12	1.61	1.08	0.53	0.0	8.6
14.15	D14S264		2.60	2.36	2.12	1.61	1.08	0.53	0.0	8.6
14.17	D14S80		1.71	1.57	1.42	1.11	0.77	0.41	0.0	5.7
14.18	D14S275		2.31	2.12	1.93	1.51	1.06	0.57	0.0	7.7
14.18	D14S262		2.90	2.64	2.38	1.82	1.22	0.61	0.0	9.6
14.18	D14S252		1.05	0.95	0.86	0.65	0.45	0.23	0.0	3.5
14.20	D14S54		0.49	0.41	0.34	0.21	0.10	0.03	0.0	1.6
14.20	D14S257		1.99	1.81	1.61	1.21	0.79	0.37	0.0	6.6
14.22	D14S49		-2.50	0.30	0.47	0.49	0.36	0.19	0.8	4.6
14.37	D14S52		-99.99	-0.22	0.17	0.32	0.23	0.10	0.9	4.2
14.50	D14S76		-99.99	-1.46	-0.72	-0.17	0.01	0.05	2.1	5.1
14.52	D14S53		-99.99	-0.14	0.27	0.49	0.44	0.27	1.8	7.6
14.54	D14S74		-99.99	0.55	0.69	0.66	0.52	0.30	0.9	6.0
14.58	D14S48		-99.99	1.00	1.06	0.87	0.56	0.26	0.7	6.8
14.63	D14S81		0.20	0.15	0.11	0.04	0.01	0.00	0.0	0.7
14.69	D14S45		-99.99	-0.57	-0.34	-0.15	-0.07	-0.02	0.8	1.5
14.70	D14S51		-99.99	-0.54	-0.10	0.15	0.16	0.08	0.8	3.0
14.78	D14S13		-99.99	-0.03	0.15	0.20	0.13	0.04	0.4	2.2
15.49	TPM1		-99.99	-1.30	-0.78	-0.32	-0.10	-0.03	1.5	3.0
15.98	D15S87		-99.99	-1.63	-0.86	-0.24	-0.02	0.04	2.0	4.9
16.04	D16S291		-99.99	-1.20	-0.46	0.10	0.25	0.20	3.1	9.4

MPD1 Linkage Region

Table B4.1 *Lod Scores for Linkage between the MPD1 Locus and Marker Loci on Chromosomes 16(cont), 17, 18, 19, 20, 21 & 22.*

Chromosome Position	Locus	Recombination Fraction ( $\theta$ )						Equivalent No. Recombinations	Equivalent No. Informative Meioses
		0.00	0.05	0.10	0.20	0.30	0.40		
16.13	D16S292	-99.99	-0.80	-0.50	-0.23	-0.09	-0.02	1.1	2.3
16.17	D16S287	-99.99	-0.94	-0.36	0.08	0.19	0.14	2.1	6.7
16.20	D16S295	-3.40	-0.59	-0.34	-0.12	-0.04	-0.01	0.7	1.5
16.28	D16S298	-99.99	0.18	0.34	0.37	0.27	0.14	0.6	3.6
16.58	D16S308	-99.99	-0.83	-0.31	0.10	0.20	0.16	0.7	1.4
16.68	D16S265	-99.99	-3.99	-2.53	-1.18	-0.52	-0.16	5.7	11.4
16.74	D16S186	-99.99	-3.57	-2.43	-1.35	-0.73	-0.31	5.5	11.0
16.74	D16S301	0.51	0.45	0.40	0.29	0.19	0.09	0.0	1.7
16.80	D16S260	-3.43	-0.25	-0.04	0.09	0.08	0.03	0.4	1.7
16.87	D16S266	-99.99	-1.27	-0.75	-0.30	-0.11	-0.03	1.7	3.4
16.99	D16S305	-99.99	-5.55	-3.80	-2.10	-1.15	-0.49	8.6	17.1
17.04	D17S30	-2.19	0.63	0.75	0.67	0.44	0.18	0.5	4.7
17.29	D17S122	-99.99	-1.36	-0.63	-0.07	0.11	0.11	5.2	12.9
17.37	GX-Alu	-99.99	-3.62	-2.25	-1.02	-0.44	-0.14	5.1	10.1
17.54	HOXB6	-99.99	-1.19	-0.61	-0.13	0.04	0.06	2.2	5.5
17.85	D17S26	-99.99	-1.24	-0.49	0.08	0.24	0.19	2.0	6.6
18.11	D18S40	-99.99	-2.45	-1.39	-0.49	-0.12	0.02	3.1	6.3
18.61	D18S51	-99.99	-2.35	-1.30	-0.44	-0.09	0.03	2.9	5.9
19.50	D19S75	-99.99	-3.55	-2.40	-1.30	-0.70	-0.30	5.4	10.8
19.76	APOC2	0.20	0.17	0.13	0.07	0.03	0.01	0.0	0.7
19.77	DM	-99.99	-5.44	-3.68	-1.98	-1.04	-0.42	8.3	16.6
20.25	D20S66	-99.99	-2.52	-1.47	-0.56	-0.16	0.01	3.9	9.0
21.32	D21S210	-99.99	-2.48	-1.41	-0.50	-0.12	0.02	0.9	2.3
21.45	D21S213	-99.99	-3.04	-1.92	-0.91	-0.41	-0.14	4.3	8.7
22.35	D22S258	-99.99	-2.78	-1.71	-0.77	-0.35	-0.12	3.8	7.7

**Table B5.1     *Lod Scores for Linkage between the QMPD Locus and Marker Loci.***

Analysis includes data from all family members. Chromosome position originally determined as % of pter-qter from NIH/CEPH Collaborative Mapping Group (1992), then progressively updated from CHLC Report (Murray *et al.*, 1993 and 1994), Cox *et al.* (1994) and "The National Centre For Biotechnology Information" GB4 GeneMap'98 (<http://www.ncbi.nlm.nih.gov/>).

D-numbers indicate anonymous DNA fragments (D-segments), other markers are associated with the genes: ACTN2 - alpha-2 actinin; APOB - apolipoprotein B; SCA1 - spino cerebellar ataxia; DMDL/Duchenne muscular dystrophy like (utrophin); MYH7 - myosin beta chain (heavy polypeptide); TPM1 - tropomyosin 1 (alpha); GX-Alu - Alu polymorphism intragenic in NF1 gene; DM - dystrophin myotonia protein kinase.

One of the markers associated with genes, APOB is a non-microsatellite polymorphisms. The remainder are microsatellites.

**Table B5.1** *Lod Scores for Linkage between the QMPD Locus and Marker Loci on Chromosomes 1, 2, 4, 6, 7, 8 & 13.*

Chromosome %pter-qter	Locus	0.00	0.05	0.10	0.20	0.30	0.40	Equivalent No. Recombinations	Equivalent No. Informative Meioses
1.19	D1S164	-3.89	-0.72	-0.46	-0.23	-0.13	-0.06	1.0	2.1
1.44	D1S223	-99.99	-1.13	-0.63	-0.27	-0.03	0.04	2.4	4.7
1.45	D1S187	1.35	1.21	1.05	0.74	0.43	0.16	0.0	4.5
1.55	D1S185	0.13	0.11	0.09	0.05	0.02	0.01	0.0	0.4
1.64	D1S242	-99.99	-3.44	-2.29	-1.18	-0.56	-0.23	5.1	10.3
1.89	ACTN2	-99.99	-1.49	-0.94	-0.46	-0.22	-0.09	2.1	4.3
2.07	APOB	-99.99	-1.08	-0.54	-0.06	0.07	0.11	2.3	4.6
2.88	D2S102	-99.99	-2.44	-1.59	-0.76	-0.37	-0.13	3.6	7.1
4.28	D4S174	-99.99	-2.17	-1.65	-0.73	-0.42	-0.16	3.0	6.0
4.98	D4S171	-99.99	-0.08	0.13	0.25	0.23	0.15	1.0	4.1
5.18	D5S268	-99.99	-2.44	-1.59	-0.76	-0.37	-0.13	3.6	7.1
5.38	D5S112	-2.50	0.12	0.32	0.42	0.36	0.22	1.2	5.5
5.39	D5S127	-99.99	-0.88	-0.38	0.02	0.14	0.12	2.2	6.4
5.65	D5S399	-3.40	-0.72	-0.44	-0.19	-0.08	0.02	1.0	2.0
6.13	SCA1*	-99.99	-2.47	-1.63	-0.88	-0.50	-0.24	3.7	7.4
6.85	DMDL*	-2.80	-0.18	0.27	0.14	0.14	0.08	0.6	2.6
7.00	D7S531	-99.99	-4.72	-3.24	-1.79	-0.96	-0.41	7.3	14.6
7.05	D7S472	0.90	0.79	0.68	0.43	0.21	0.05	0.0	3.0
7.06	D7S517	-99.99	-4.72	-3.24	-1.79	-0.96	-0.41	7.3	14.6
7.95	D7S483	-3.59	-0.33	-0.12	0.01	0.02	0.01	0.8	1.9
8.70	D8S198	-99.99	-2.44	-1.59	-0.76	-0.37	-0.13	3.6	7.1
13.74	D13S71	-99.99	-0.94	-0.32	0.05	0.12	0.13	1.7	4.6





**APPENDIX C**

***MYH6***

***MYH7***

***α-MyHC***

***β-MyHC***

<b>Table C7.1</b>	<b><i>Sequence Comparisons of MYH6 and MYH7 cDNA, and of α-MyHC and β-MyHC Amino Acids.</i></b>
<b>Table C7.2</b>	<b><i>Genetic and Amino Acids Codes.</i></b>
<b>Table C7.3</b>	<b><i>Sequence of MYH7 and its Primers.</i></b>

**Table C7.1      *Sequence Comparisons of MYH6 and MYH7 cDNA, and of  $\alpha$ -MyHC and  $\beta$ -MyHC Amino Acids.***

The application program used to make these comparisons was *SeqEd*, Version 1.0.3 (Myers, 1992: ABI). The sense-strand exonic sequences for both genes were imported from a BLAST search (<http://www.ncbi.nlm.nih.gov/BLAST/>) which identified the full published genomic sequences, *MYH6* (Epp et al., 1993) and *MYH7* (Jaenicke et al., 1990).

Row one shows the approximate base number of each gene. 'Misnumbering' bases is an unavoidable consequence of aligning these two sequences, because their 5' untranslated regions are of different lengths, so *MYH6* and *MYH7* have different starting points. In addition, *MYH7* has four deleted codons compared to *MYH6*.

Row two and three show the sense-strand exonic sequences of *MYH6* and *MYH7*. Row four compares these *MYH6* and *MYH7* exonic sequences. Dashes indicate identical bases, asterisks indicate non-identical bases.

Row five and six show the derived amino acid sequences for the polypeptide encoded by *MYH6*, which is  $\alpha$ -MyHC, and by *MYH7*, which is  $\beta$ -MyHC. Epp et al., (1993) and Jaenicke et al. (1990) derived the  $\alpha$ -MyHC and  $\beta$ -MyHC amino acid sequences from the genomic sequences for these genes. Row seven compares these amino acid sequences. In this row, dashes indicate identical amino acids, asterisks indicate non-identical amino acids.

Table C7.2 in this Appendix aligns genetic and amino acid codes.

<p>MYH6 cDNA MYH7 cDNA cDNA:MYH7vsMYH6 alpha-MyHC beta-MyHC alpha vs beta</p>	<p>15 30 45 60 75 90 105</p> <p>atagagagactctgctgagccagattcttcaggattctcgtgaagggataaccaggggaagcaccagATGACCGATGCCAGATGGCTGACTTTGGGG -tgtctttccctgctgctctcaggtccctgagggccttgccctttctcatctgtagacacacttgagtagccagggcacagccATGGGAGATTCGGAGATGGCAGTCTTTGGGG *****</p> <p>M T D A Q M A D F G M G D S E M A V F G</p>
<p>MYH6 cDNA MYH7 cDNA cDNA:MYH7vsMYH6 alpha-MyHC beta-MyHC alpha vs beta</p>	<p>120 135 150 165 180 195 210 225</p> <p>CAGCGGCCCCAGTACCTCCGCAAGTCAGAGAAGGAGCGCTCTAGAGGCCAGACCCGGCCCTTTGACATTTCGCACTGAGTGCTTCGTGCCCGATGACAAGGAAGAGTTTGTCAAAGCCAA CTGCCGCCCCCTACCTGCGCAAGTCAGAGAAGGAGCGGCTAGAGAAGCGCAGACAGGCCCTTTGACCTCAAGAAGGATGCTTCGTGCCTTGATGACAACAGGAGTTTGTCAAAGGCCAA *****</p> <p>A A A Q Y L R K S E K E R L E A Q T R P F D I R T E C F V P D D K E E F V K A K A A A P Y L R K S E K E R L E A Q T R P F D L K K D V F V P D D K Q E F V K A K</p>
<p>MYH6 cDNA MYH7 cDNA cDNA:MYH7vsMYH6 alpha-MyHC beta-MyHC alpha vs beta</p>	<p>240 255 270 285 300 315 330 345</p> <p>GATTTTGTCCCGGAGGGAGGCAAGGTCATGTCTGAAACCGAGAATGGAAGACGGTGACTGTGAAGGAGGACCAGGTGTTGTCAGCAGAACCCACCCAAAGTTTCGACAAGATTCAGGAC GATCGTGTCTCGAGAGGGGTGGCAAAGTCACGTGCCGAGACTGAGTATGGCAAGACAGTGACCGTGAAGGAGGACCAGGTGATGTCAGCAGAACCCACCCAAAGTTTCGACAAGATTCAGGAC *****</p> <p>I L S R E G G K V I A E T E N G K T V T V K E D Q V L Q Q N P P K F D K I Q D I V S R E G G K V T A E T E Y G K T V T V K E D Q V M Q Q N P P K F D K I E D</p>
<p>MYH6 cDNA MYH7 cDNA cDNA:MYH7vsMYH6 alpha-MyHC beta-MyHC alpha vs beta</p>	<p>360 375 390 405 420 435 450 465</p> <p>ATGGCCATGCTGACCTTCTCTGACGAGCCCGCGGTGCTTTTCAACCTCAAGGAGCGCTACGCGGCTGGATGATATATACCTACTCGGGGCTCTTCTGTGTCACTGTCAACCCCTACA ATGGCCATGCTGACCTTCTCTGACGAGCCCGCGGTGCTTACAACTCAAGGATGCTACGCGCTCTGGATGATCTACACCTACTCGGGGCTCTTCTGTGTCACTGTCAACCCCTACA *****</p> <p>M A M L T F L H E P A V L F N L K E R Y A A W H I Y T Y S G L F C V T V N P Y M A M L T F L H E P A V L Y N L K D R Y G A S W M I Y T Y S G L F C V T V N P Y</p>
<p>MYH6 cDNA MYH7 cDNA cDNA:MYH7vsMYH6 alpha-MyHC beta-MyHC alpha vs beta</p>	<p>480 495 510 525 540 555 570 585</p> <p>AGTGGCTGCCGGTGTACAAATGCCGAGGTGGTGGCCGCTACCGGGGCAAGAAGAGGAGTGAGGCCCCGCCCCACATCTCTCCATCTCCGACAACGCCCTATCAGTACATGCTGACAGA AGTGGCTGCCGGTGTACAAATGCCGAGGTGGTGGCTGCTACCGGGGCAAGAAGAGGAGGAGGCCCCGCCCCACATCTCTCCATCTCCGACAACGCCCTATCAGTACATGCTGACAGA *****</p> <p>K W L P V Y N A E V V A A Y R G K K R S E A P P H I F S I S D N A Y Q Y M L T D K W L P V Y T P E V V A A Y R G K K R S E A P P H I F S I S D N A Y Q Y M L T D</p>
<p>MYH6 cDNA MYH7 cDNA cDNA:MYH7vsMYH6 alpha-MyHC beta-MyHC alpha vs beta</p>	<p>600 615 630 645 660 675 690 705</p> <p>TCGGGAGAACCAGTCCATCTCTATCACGGGAGAATCCGGGGCGGGGAAGACTGTGAACACCAAGCGTGTCACTCCAGTACTTTGCCAGCATTCGAGCCATAGGTGACCGTGGCAAGAAG CAGAGAAAACCAGTCCATCTCTATCACCGGAGAATCCGGAGCAGGGAAGACAGTCAACACCAAGAGGGTCACTCCAGTACTTTGCTGTTATTCGAGCCATTCGGGACCGCAGCAAGAAG *****</p> <p>R E N Q S I L I T G E S G A G K T V N T K R V I Q Y F A S I A A I G D R G K K R E N Q S I L I T G E S G A G K T V N T K R V I Q Y F A V I A A I G D R S K K</p>

	720735750765780795810825
MYH6 cDNA MYH7 cDNA cDNA:MYH7vsMYH6 alpha-MyHC beta-MyHC alpha vs beta	GACAATGCCAATGCGAACAAAGGGCACCCCTGGAGGACCAGATCATCCAGGCCAACCCGCTCTGGAGGCCTTCGGCAATGCCAAGACTGTCCGGAACGACAACTCCTCCCGCTTTGGGA GACCAGAGCCCGGGC++AAGGGCACCCCTGGAGGACCAGATCATCCAGGCCAACCCGTCTGGAGGCCTTTGGCAATGCCAAGACCGTCCGGAACGACAACTCCTCCCGCTTCGGGA ***** D N A N A N K G T L E D Q I I Q A N P A L E A F G N A K T V R N D N S S R F G D Q S P G ? K G T L E D Q I I Q A N P A L E A F G N A K T V R N D N S S R F G *****
	840855870885900915930
MYH6 cDNA MYH7 cDNA cDNA:MYH7vsMYH6 alpha-MyHC beta-MyHC alpha vs beta	AATTCATTAGGATCCACTTTGGGGCCACTGGAAAGCTGGCTTCTGCAGACATAGAGACCTACCTGCTGGAGAAGTCCCGGGTGATCTTCCAGCTGAAAGCTGAGAGAAACTACCACAT AATTCATTTCGAATTCATTTTGGGGCAACAGGAAGTTGGCATCTGCAGACATAGAGACCTATCTTCTGGAAAATCCAGAGTTATTTTCCAGCTGAAAGCAGAGAGAGATTATCACAT ***** K F I R I H F G A T G K L A S A D I E T Y L L E K S R V I F Q L K A E R N Y H I K F I R I H F G A T G K L A S A D I E T Y L L E K S R V I F Q L K A E R D Y H I *****
	9459609759901005102010351050
MYH6 cDNA MYH7 cDNA cDNA:MYH7vsMYH6 alpha-MyHC beta-MyHC alpha vs beta	CTTCTACCAGATTCTGTCCAACAAGAAGCCGGAGTTGCTGGACATGCTGCTGGTCACCAACAATCCCTACGACTACGCCTTCGTGTCTCAGGGAGAGGTGTCGGTGGCCTCCATTGAT TTTCTACCAAAATCCTGTCTAACAAGAGCCTGAGCTGCTGGACATGCTGCTGATCACCACCAACCCCTACGATTATGCATTTCATCTCCCAAGGAGAGACCACCGTGGCCTCCATTGAT ***** F Y Q I L S N K K P E L L D M L L V T N N P Y D Y A F V S Q G E V S V A S I D F Y Q I L S N K K P E L L D M L L I T N N P Y D Y A F I S Q G E T T V A S I D *****
	10651080109511101125114011551170
MYH6 cDNA MYH7 cDNA cDNA:MYH7vsMYH6 alpha-MyHC beta-MyHC alpha vs beta	GACTCCGAGGAGCTCATGGCCACCGATAGTGCTTTGACGTGCTGGGCTTCACCTTCAGAGGAGAAAGCTGGCGTCTACAAGCTGACGGGAGCCATCATGCACTACGGGAACATGAAGT GACGCTGAGGAGCTCATGGCCACTGATAACGCTTTTGATGTGCTGGGCTTCACCTTCAGAGGAGAAAACTCCATGTATAAGCTGACAGGCGCCATCATGCACTTTGGAAACATGAAGT ***** D S E E L M A T D S A F D V L G F T S E E K A G V Y K L T G A I M H Y G N M K D A E E L M A T D N A F D V L G F T S E E K N S M Y K L T G A I M H F G N M K *****
	11851200121512301245126012751290
MYH6 cDNA MYH7 cDNA cDNA:MYH7vsMYH6 alpha-MyHC beta-MyHC alpha vs beta	TCAAGCAGAAGCAGCGGGAGGAGCAGGCGGAGCCAGACGGCACCGAAGATGCTGACAAGTCTGGGCTACCTCATGGGGCTGAACTCAGCTGACCTGCTCAAGGGGCTGTGCCACCCCTCG TCAAGCTGAAGCAGCGGGAGGAGCAGGCGGAGCCAGACGGCACCTGAAGAGGCTGACAAGTCTGGCTACCTCATGGGGCTGAACTCAGCGGACCTGCTCAAGGGGCTGTGCCACCCCTCG ***** F K Q K Q R E E Q A E P D G T E D A D K S A Y L M G L N S A D L L K G L C H P R F K L K Q R E E Q A E P D G T E E A D K S A Y L M G L N S A D L L K G L C H P R *****
	13051320133513501365138013951410
MYH6 cDNA MYH7 cDNA cDNA:MYH7vsMYH6 alpha-MyHC beta-MyHC alpha vs beta	GGTGAAAGTGGGCAACGAGTATGTACCAAGGGGAGAGCGTGCAGCAGGTGTACTACTCCATCGGGGCTCTGGCCAAAGGCAGTGTATGAGAAGATGTTCAACTGGATGGTGACGCGC GGTGAAAGTGGGCAATGAGTACGTACCAAGGGGAGAAATGTCCAGCAGGTGTATATGCACTGGGGCACTGGCCAAAGGCAGTGTATGAGAGGATGTTCAACTGGATGGTGACGCGC ***** V K V G N E Y V T K G Q S V Q Q V Y Y S I G A L A K A V Y E K M F N W M V T R V K V G N E Y V T K G Q N V Q Q V I Y A T G A L A K A V Y E R M F N W M V T R *****

	1425	1440	1455	1470	1485	1500	1515	1530
MYH6 cDNA MYH7 cDNA cDNA: MYH7 vs MYH6 alpha-MyHC beta-MyHC alpha vs beta	ATCAACGCCACCCTGGAGACCAAGCAGCCACGCCAGTACTTCATAGGAGTCTCTGGACATCGCTGGCTTCGAGATCTTCGACTTCAACAGCTTTGAGCAGCTCTGCATCAACTTCACCA ATCAATGCCACCCTGGAGACCAAGCAGCCACGCCAGTACTTCATAGGAGTCTCTGGACATCGCTGGCTTCGAGATCTTCGACTTCAACAGCTTTGAGCAGCTCTGCATCAACTTCACCA * I N A T L E T K Q P R Q Y F I G V L D I A G F E I F D F N S F E Q L C I N F T I N A T L E T K Q P R Q Y F I G V L D I A G F E I F D F N S F E Q L C I N F T							
	1545	1560	1575	1590	1605	1620	1635	1650
MYH6 cDNA MYH7 cDNA cDNA: MYH7 vs MYH6 alpha-MyHC beta-MyHC alpha vs beta	ACGAGAAGCTGCAGCAGTTCTTCAACCACCACATGTTCTGCTGCTGGAGCAGGAGGAGTACAAGAAGGAGGGCATTGAGTGGACATTCAITGACTTTGGCATGGACCTGCAGGCTGCAT ACGAGAAGCTGCAGCAGTTCTTCAACCACCACATGTTTGTGCTGCTGGAGCAGGAGGAGTACAAGAAGGAGGGCATTGAGTGGACATTCAITGACTTTGGCATGGACCTGCAGGCTGCAT * N E K L Q Q F F N H H M F V L E Q E E Y K K E G I E W T F I D F G M D L Q A C I N E K L Q Q F F N H H M F V L E Q E E Y K K E G I E W T F I D F G M D L Q A C I							
	1665	1680	1695	1710	1725	1740	1755	1770
MYH6 cDNA MYH7 cDNA cDNA: MYH7 vs MYH6 alpha-MyHC beta-MyHC alpha vs beta	TGACCTCATCGAGAAGCCCATGGGCATCATGTCCATCTCTGGAGGAGGAGTGCATGTTCCCAAGGCCACTGACATGACCTTCAAGGCCAAGCTGTACGACAACCACCTGGGCAAGTCC TGACCTCATCGAGAAGCCCATGGGCATCATGTCCATCTCTGGAGGAGGAGTGCATGTTCCCAAGGCCACTGACATGACCTTCAAGGCCAAGCTGTGTGACAACCACCTGGGCAAGTCC * D L I E K P M G I M S I L E E E C M F P K A T D M T F K A K L Y D N H L G K S D L I E K P M G I M S I L E E E C M F P K A T D M T F K A K L F D N H L G K S							
	1785	1800	1815	1830	1845	1860	1875	
MYH6 cDNA MYH7 cDNA cDNA: MYH7 vs MYH6 alpha-MyHC beta-MyHC alpha vs beta	AACAAATTTCCAGAAGCCACGCAACATCAAGGGGAAGCAGGAAGCCCACTTCTCCTTGATCCACTACGCCGGCACTGTGGACTACAACATCTCTGGGCTGGCTGGAAAAAACAAGGATC GCCAACTTTCCAGAAGCCACGAAATATCAAGGGGAAGCCTGAAGCCCACTTCTCCTTGATCCACTATGCCGGCACTGTGGACTACAACATCAITGGCTGGCTGCAGAAAGAACAAGGATC * N N F Q K P R N I K G K Q E A H F S L I H Y A G T V D Y N I L G W L E K N K D A N F Q K P R N I K G K P E A H F S L I H Y A G I V D Y N I I G W L Q K N K D							
	1890	1905	1920	1935	1950	1965	1980	1995
MYH6 cDNA MYH7 cDNA cDNA: MYH7 vs MYH6 alpha-MyHC beta-MyHC alpha vs beta	CTCTCAACGAGACTGTGTGGCCCTGTACCAGAAGTCTCCTCCTCAAGCTCATGGCCACTCTCTTCTCCTCTACGCAACTGCCGATAGTGGGACAGTGGTAAAGCAAAGGAGGCAA CTCTCAATGAGACTGTCTGTGGGCTGTATCAGAAGTCTCCTCCTCAAGTGTGCTCAGCACCTGTCTTGGCAACTATGCTGGGGCTGATGCGCCT++ATTGAGAAGGGCAAAGGCAAGGC * P L N E T V V A L Y Q K S S L K L M A T L F S S S Y A T A D T G D S G K S K G G K P L N E T V V G L Y Q K S S L K L L S T L F A N Y A G A D A P ? I E K G K G K A							
	2010	2025	2040	2055	2070	2085	2100	2115
MYH6 cDNA MYH7 cDNA cDNA: MYH7 vs MYH6 alpha-MyHC beta-MyHC alpha vs beta	GAAAAAGGGCTCATCTTCCAGACGGTGTCTGGCTCTCCACCAGGAAAAATCTCAACAAGCTAATGACCAACCTGAGGACCACCCATCTCTCACTTTGTGCGTTGCATCATCCCAATGAG CAAGAAAGGCTCGTCTTTTCAAGCTGTGTACGCTCTGCACAGGAAAAATCTGAACAAGCTGATGACCAACTTGGCGCTCCACCCATCCCACTTTGTACGTTGTATCATCCCTAATGAG * K K G S S S F Q T V S A L H R E N L N K L M T N L R T T H P H F V R C I I P N E K K G S S S F Q T V S A L H R E N L N K L M T N L R S T H P H F V R C I I P N E							

	2130	2145	2160	2175	2190	2205	2220	2235
MYH6 cDNA MYH7 cDNA cDNA:MYH7vsMYH6 alpha-MyHC beta-MyHC alpha vs beta	CGGAAGGCTCCAGGGGTGATGGACAACCCCTGGTCATGCACCAAGCTGCGCTGCAATGGCGTGTCTGGAGGGCATCCGCATCTGCAGGAAGGGCTTCCCCAACCCGCATCCTCTACGGGG ACAAAGTCTCCAGGGCTGATGGACAACCCCTGGTCATGCACCAAGCTGCGCTGCAATGGTGTGTCTGGAGGGCATCCGCATCTGCAGGAAGGGCTTCCCCAACCCGCATCCTCTACGGGG ***** R K A P G V M D N P L V M H Q L R C N G V L E G I R I C R K G F P N R I L Y G T K S P G V M D N P L V M H Q L R C N G V L E G I R I C R K G F P N R I L Y G *****							
	2250	2265	2280	2295	2310	2325	2340	2355
MYH6 cDNA MYH7 cDNA cDNA:MYH7vsMYH6 alpha-MyHC beta-MyHC alpha vs beta	ACTTCCGGCAGAGGTATCGCATCCTGAACCCAGTGGCCATCCCTGAGGGACAGTTCATTGATAGCAGGAAGGGGGACAGAGAAGCTGCTCAGCTCTCTGGACATTGATCACAAACAGTA ACTTCCGGCAGAGGTATCGCATCCTGAACCCAGCGGCCATCCCTGAGGGACAGTTCATTGATAGCAGGAAGGGGGCAGAGAAGCTGCTCAGCTCCCTGGACATTGATCACAAACAGTA ***** D F R Q R Y R I L N P V A I P E G Q F I D S R K G T E K L L S S L D I D H N Q Y D F R Q R Y R I L N P A A I P E G Q F I D S R K G A E K L L S S L D I D H N Q Y *****							
	2370	2385	2400	2415	2430	2445	2460	2475
MYH6 cDNA MYH7 cDNA cDNA:MYH7vsMYH6 alpha-MyHC beta-MyHC alpha vs beta	CAAGTTTGGCCACACCAAGGTGTTCTTCAAGGCAGGGCTGCTTGGGCTGCTGGAGGAGATGCGGGATGAGAGGCTGAGCCGCATCATCAGCGCATGCAAGGCCAAGCCCGGGGCGCAG CAAGTTTGGCCACACCAAGGTGTTCTTCAAGGCAGGGCTGCTTGGGCTGCTGGAGGAAATGAGGGACGAGAGGCTGAGCCGCATCATCAGCGCATGCAAGGCCAAGCCCGGGGCTGTG ***** K F G H T K V F F K A G L L G L L E E M R D E R L S R I I T R M Q A Q A R G Q K F G H T K V F F K A G L L G L L E E M R D E R L S R I I T R M Q A Q A R G V *****							
	2490	2505	2520	2535	2550	2565	2580	2595
MYH6 cDNA MYH7 cDNA cDNA:MYH7vsMYH6 alpha-MyHC beta-MyHC alpha vs beta	CTCATGCGCATTGAGTTCAAGAAGATAGTGGAAACGCGAGGGATGCCCTGCTGGTAAATCCAGTGGAAACATTCCGGGCCCTTCATGGGGGTCAAGAATTGGCCCTGGATGAAGCTCTACTTCA CTCGCCAGAATGGAGTACAAAAGCTGCTGGAAACGTAGAGACTCCCTGCTGGTAAATCCAGTGGAAACATTCCGGGCCCTTCATGGGGGTCAAGAATTGGCCCTGGATGAAGCTCTACTTCA ***** L M R I E F K K I V E R R D A L L V I Q W N I R A F M G V K N W P W M K L Y F L A R M E Y K K L L E R R D S L L V I Q W N I R A F M G V K N W P W M K L Y F *****							
	2610	2625	2640	2655	2670	2685	2700	
MYH6 cDNA MYH7 cDNA cDNA:MYH7vsMYH6 alpha-MyHC beta-MyHC alpha vs beta	AGATCAAGCCGCTGCTGAAGAGCGCAGAGACGGAGAAGGAGATGGCCACCATGAAGGAAGAGTTCTGGGCGCATCAAGAGAGCGCTGGAGAAGTCCGAGGCTCGCCGCAAGGAGCTGGA AGATCAAGCCGCTGCTGAAGAGTGCAGAAAGAGAGAAGGAGATGGCCATCATGAAGGAGGAGTTACACACGCTCAAGAGAGCGCTAGAGAAGTCCGAGGCTCGCCGCAAGGAGCTGGA ***** K I K P L L K S A E T E K E M A T M K E E F G R I K E T L E K S E A R R R K E L E K I K P L L K S A E R E K E M A S M K E E F T R L K E A L E K S E A R R R K E L E *****							
	2715	2730	2745	2760	2775	2790	2805	2820
MYH6 cDNA MYH7 cDNA cDNA:MYH7vsMYH6 alpha-MyHC beta-MyHC alpha vs beta	GGAGAAGATGGTGTCCCTGCTGCAGGAGAAGAATGACCTGCAGCTCCAAGTGCAGGCGGAACAAAGACAACCTCAATGATGCTGAGGAGCGCTGCGACCAAGCTGATCAAAAACAAGATT GGAGAAGATGGTGTCCCTGCTGCAGGAGAAGAATGACCTGCAGCTCCAAGTGCAGGCGGAACAAAGACAACCTGCGAGATGCTGAGGAGCGCTGTGATCAGCTGATCAAAAACAAGATT ***** E K M V S L L Q E K N D L Q L Q V Q A E Q D N L N D A E E R C D Q L I K N K I E K M V S L L Q E K N D L Q L Q V Q A E Q D N L A D A E E R C D Q L I K N K I *****							



	3555 3570 3585 3600 3615 3630 3645
MYH6 cDNA MYH7 cDNA cDNA:MYH7vsMYH6 alpha-MyHC beta-MyHC alpha vs beta	CTGGAAGAGGCCGGCGGGGCCACGTCCTCGTGAGATCGAGATGAACAAGAGCGCGAGGCCGAGTTCCAGAAGATGCGCGGGGACCTGGAGGAGGCCACGTCGACGACGAGGCCACTG CTGGAAGAGGCCGGCGGGGCCACGTCCTCGTGAGATCGAGATGAACAAGAGCGCGAGGCCGAGTTCCAGAAGATGCGCGGGGACCTGGAGGAGGCCACGTCGACGACGAGGCCACTG L E E A G G A T S V Q I E M N K K R E A E F Q K M R R D L E E A T L Q H E A T L E E A G G A T S V Q I E M N K K R E A E F Q K M R R D L E E A T L Q H E A T
	3660 3675 3690 3705 3720 3735 3750 3765
MYH6 cDNA MYH7 cDNA cDNA:MYH7vsMYH6 alpha-MyHC beta-MyHC alpha vs beta	CCGCGGGCCCTGCGCAAGAAGCAGCGCGACAGCGTGGCGGAGCTGGGGCGAGCAGATCGACAACCTGCGAGCGGTGAAGCAGAAGCTGGAGAAGGAGAAGAGCGAGTTCAAGCTGGAGCT CCGCGGGCCCTGCGCAAGAAGCAGCGCGACAGCGTGGCGGAGCTGGGGCGAGCAGATCGACAACCTGCGAGCGGTGAAGCAGAAGCTGGAGAAGGAGAAGAGCGAGTTCAAGCTGGAGCT A A A L R K K H A D S V A E L G E Q I D N L Q R V K Q K L E K E K S E F K L E L A A A L R K K H A D S V A E L G E Q I D N L Q R V K Q K L E K E K S E F K L E L
	3780 3795 3810 3825 3840 3855 3870 3885
MYH6 cDNA MYH7 cDNA cDNA:MYH7vsMYH6 alpha-MyHC beta-MyHC alpha vs beta	GGATGACGTCACCTCCAACATGGAGCAGATCATCAAGGCCAAGGCCAAACCTGGAGAAAGTGTCTCGGACGCTGGAGGACCAGGCCAATGAGTACCGCGTGAAGCTAGAAGAGGCCCAA GGATGACGTCACCTCCAACATGGAGCAGATCATCAAGGCCAAGGCCAAACCTGGAGAAAGTGTCTCGGACGCTGGAGGACCAGGCCAATGAGTACCGCGTGAAGCTAGAAGAGGCCCAA D D V T S N M E Q I I K A K A N L E K V S R T L E D Q A N E Y R V K L E E A Q D D V T S N M E Q I I K A K A N L E K M C R T L E D Q M N E H R S K A E E T Q
	3900 3915 3930 3945 3960 3975 3990 4005
MYH6 cDNA MYH7 cDNA cDNA:MYH7vsMYH6 alpha-MyHC beta-MyHC alpha vs beta	CGCTCCCTCAATGATTTCAACACCCAGCGAGCCAAGCTGCAGACCCAGAGATGGAGAGTTGGCCCGGCAGCTAGAGGAAAAGGAGGCGCTAATCTCGCAGCTGACCCGGGGGAAGCTCT CGTTCTGTCAACGACCTCACCAGCCAGCGGGCCAAAGTTGCAACCGGAGATGGTGAGCTGTCCCGGCAGCTGGATGAGAAGGAGGCACTGATCTCCAGCTGACCCGAGGCAAGCTCA R S L N D F T T S Q R A K L Q T E N G E L A R Q L E E K E A L I S Q L T R G K L R S V N D L T S Q R A K L Q T E N G E L S R Q L D E K E A L I S Q L T R G K L
	4020 4035 4050 4065 4080 4095 4110 4125
MYH6 cDNA MYH7 cDNA cDNA:MYH7vsMYH6 alpha-MyHC beta-MyHC alpha vs beta	CTTATACCCAGCAAATGGAGGACCTCAAAGGCAGCTGGAGGAGGAGGGCAAGGCCGAAGAACGCCCTGGCCCCATGCACTGCAGTCAGCCCGGCATGACTGCGACCTGCTGCGGGAGCA CCTACACCCAGCAGCTGGAGGACCTCAAAGGCAGCTGGAGGAGGAGGTTAAGGCCGAAGAACGCCCTGGCCCCACGCACTGCAGTCGGCCCGGCATGACTGCGACCTGCTGCGGGAGCA S Y T Q Q M E D L K R Q L E E E G K A K N A L A H A L Q S A R H D C D L L R E Q T Y T Q Q L E D L K R Q L E E E V K A K N A L A H A L Q S A R H D C D L L R E Q
	4140 4155 4170 4185 4200 4215 4230 4245
MYH6 cDNA MYH7 cDNA cDNA:MYH7vsMYH6 alpha-MyHC beta-MyHC alpha vs beta	GTACGAGGAGGAGACAGAGGCCAAGGCCGAGCTGCAGCGCGTCTGTGTCGAAGGCCAACTCGGAGGTGGGCCAGTGGAGGACCAAGTATGAGACGGACGCCATTTCAGCGGACTGAGGAG GTACGAGGAGGAGACAGAGGCCAAGGCCGAGCTGCAGCGCGTCTGTGTCGAAGGCCAACTCGGAGGTGGGCCAGTGGAGGACCAAGTATGAGACGGACGCCATTTCAGCGGACTGAGGAG Y E E E T E A K A E L Q R V L S K A N S E V A Q W R T K Y E T D A I Q R T E E Y E E E T E A K A E L Q R V L S K A N S E V A Q W R T K Y E T D A I Q R T E E



	4260 4275 4290 4305 4320 4335 4350 4365
MYH6 cDNA MYH7 cDNA cDNA:MYH7vsMYH6 alpha-MyHC beta-MyHC alpha vs beta	CTCGAAGAGGCCAAAAGAAGCTGGCCACGGCTGCAGGATGCCGAGGAGGCCGTGGAGGCTGTTAATGCCAAGTGCTCCTCACTGGAGAAGACCAAGCACCGGCTACAGAATGAGA CTCGAGGAGGCCAAGAAGAAGCTGGCCACGGCTGCAGGAAGCTGAGGAGGCCGTGGAGGCTGTTAATGCCAAGTGCTCCTCGCTGGAGAAGACCAAGCACCGGCTACAGAATGAGA *-----* L E E A K K K L A Q R L Q D A E E A V E A V N A K C S S L E K T K H R L Q N E L E E A K K K L A Q R L Q E A E E A V E A V N A K C S S L E K T K H R L Q N E *-----*
	4380 4395 4410 4425 4440 4455 4470
MYH6 cDNA MYH7 cDNA cDNA:MYH7vsMYH6 alpha-MyHC beta-MyHC alpha vs beta	TAGAGGACTTGATGGTGGACGTAGAGCGCTCCAACGCCGCTGCTGCAGCCCTGGACAAGAAGCAGAGAACTTTGACAAGATCCTGGCCGAGTGGAGCAGAGATATGAGGAGTCGCA TCGAGGACTTGATGGTGGACGTAGAGCGCTCCAATGCTGCTGCTGCAGCCCTGGACAAGAAGCAGAGGAACCTTCGACAAGATCCTGGCCGAGTGGAGCAGAGATATGAGGAGTCGCA *-----* I E D L M V D V E R S N A A A A A L D K K Q R N F D K I L A E W K Q K Y E E S Q I E D L M V D V E R S N A A A A A L D K K Q R N F D K I L A E W K Q K Y E E S Q *-----*
	4485 4500 4515 4530 4545 4560 4575 4590
MYH6 cDNA MYH7 cDNA cDNA:MYH7vsMYH6 alpha-MyHC beta-MyHC alpha vs beta	GTCTGAGCTGGAGTCCCTACAGAAGGAGGCTCGCTCCCTCAGCACAGAGCTCTTCAAGCTCAAGAAGCCCTACGAGGAGTCCCTGGAGCACCTAGAGACCTTCAAGCGGGAGAACAAAG GTCCGAGCTGGAGTCCCTCGAGAAGGAGGCTCGCTCCCTCAGCACAGAGCTCTTCAAACTCAAGAAGCCCTATGAGGAGTCCCTGGAACATCTGGAGACCTTCAAGCGGGAGAACAAA *-----* S E L E S S Q K E A R S L S T E L F K L K N A Y E E S L E H L E T F K R E N K S E L E S S Q K E A R S L S T E L F K L K N A Y E E S L E H L E T F K R E N K *-----*
	4605 4620 4635 4650 4665 4680 4695 4710
MYH6 cDNA MYH7 cDNA cDNA:MYH7vsMYH6 alpha-MyHC beta-MyHC alpha vs beta	AACCTTCAGGAGGAAATCTCGACCTTACTGAGCAGCTAGGAGAAGGAGGAAAGATGTGCATGAGCTGGAGAAGGTCCGCAACAGCTGGAGGTGGAGAAGCTGGAGCTGCAGTCAG AACCTGCAGGAGGAGATCTCCGACTTGACTGAGCAGTTGGGTTCCAGCGGAAAGACTATCCATGAGCTGGAGAAGGTCCGAAAGCAGCTGGAGGCCGAGAAGATGGAGCTGCAGTCAG *-----* N L Q E E I S D L T E Q L G E G G K N V H E L E K V R K Q L E V E K L E L Q S N L Q E E I S D L T E Q L G S S G K T I H E L E K V R K Q L E A E K M E L Q S *-----*
	4725 4740 4755 4770 4785 4800 4815 4830
MYH6 cDNA MYH7 cDNA cDNA:MYH7vsMYH6 alpha-MyHC beta-MyHC alpha vs beta	CCCTGGAGGAGGCAGAGGCCCTCCCTGGAGCACGAGGAGGGCAAGATCCTCCGGGCCAGCTAGAGTTCAACCAGATCAAGGCAGAGATCGAGCGGAAGCTGGCAGAGAAGGACGAGGA CCCTGGAGGAGGCCGAGGCTCCCTGGAGCACGAGGAGGGCAAGATCCTCCGGGCCAGCTGGAGTTCAACCAGATCAAGGCAGAGATCGAGCGGAAGCTGGCAGAGAAGGACGAGGA *-----* A L E E A E A S L E H E E G K I L R A Q L E F N Q I K A E I E R K L A E K D E E A L E E A E A S L E H E E G K I L R A Q L E F N Q I K A E I E R K L A E K D E E *-----*
	4845 4860 4875 4890 4905 4920 4935 4950
MYH6 cDNA MYH7 cDNA cDNA:MYH7vsMYH6 alpha-MyHC beta-MyHC alpha vs beta	GATGGAACAGGCCAAGCGCAACCACCTGCGGGTGGTGGACTCGCTGCAGACCTCCCTGGATGCAGAGACACGCAGCCGCAACGAGGCTCTGAGGGTGAAGAAGAAGATGGAAGGAGAC GATGGAACAGGCCAAGCGCAACCACCTGCGGGTGGTGGACTCGCTGCAGACCTCCCTGGACGCAGAGACACGCAGCCGCAACGAGGCCCTGAGGGTGAAGAAGAAGATGGAAGGAGAC *-----* M E Q A K R N H Q R V V D S L Q T S L D A E T R S R N E V L R V K K K M E G D M E Q A K R N H L R V V D S L Q T S L D A E T R S R N E A L R V K K K M E G D *-----*



	56705685570057155730574557605775
MYH6 cDNA MYH7 cDNA cDNA:MYH7vsMYH6 alpha-MyHC beta-MyHC alpha vs beta	AAAAAAGAACCTGCTGCGGCTACAGGACCTGGTGGACAAGCTGCAACTGAAGGTCAAGGCCTACAAGCGCCAGGCCGAGGAGGCGGAGGAGCAAGCCAAACCAACCTGTCCAAGTTCC AGGAAAAAACCTGCTGCGGCTGACAGGACCTGGTAGACAAGCTGCAGCTAAAGGTCAAGGCCTACAAGCGCCAGGCCGAGGAGGCGGAGGAGCAAGCCAAACCAACCTGTCCAAGTTCC ***-----* K K N L L R L Q D L V D K L Q L K V K A Y K R Q A E E A E E Q A N T N L S K F R K N L L R L Q D L V D K L Q L K V K A Y K R Q A E E A E E Q A N T N L S K F *-----*
	57905805582058355850586558805895
MYH6 cDNA MYH7 cDNA cDNA:MYH7vsMYH6 alpha-MyHC beta-MyHC alpha vs beta	GCAAGGTGCAGCATGAGCTGGATGAGGCAGAGGAGCGGGCGGACATCGCTGAGTCCCAGGTCAACAAGCTTCGAGCCAAAGAGCCGTGACATTGGTGCCAAGCAAAAAATGCACGATGA GCAAGGTGCAGCACGAGCTGGATGAGGCAGAGGAGCGGGCGGACATCGCCGAGTCCCAGGTCAACAAGCTGCGGGCCAAAGAGCCGTGACATTGGCACGAAG++++GGCTTGAATGA *-----* R K V Q H E L D E A E E R A D I A E S Q V N K L R A K S R D I G A K Q K M H D E R K V Q H E L D E A E E R A D I A E S Q V N K L R A K S R D I G T K ? ? G L N E *-----*
	59105925594059555970598560006015
MYH6 cDNA MYH7 cDNA cDNA:MYH7vsMYH6 alpha-MyHC beta-MyHC alpha vs beta	GGAGTGAcactgcctcggaacctcactcttgccaaacctgtaataaatatgagtgccca GGAGTAGctttgccacatcttgatctgctcagccctggaggtgccagcaaaagcccatgctggagcctgtgtaacagctccttgggaggaagcagaataaaagcaattttccttgaagc -----* E E -----
	60306045606060756090610561206135
MYH6 cDNA MYH7 cDNA cDNA:MYH7vsMYH6 alpha-MyHC beta-MyHC alpha vs beta	<== cga <== <== <== <==

**Table C7.2 Genetic and Amino Acids Codes.**

	AGA									UUA					AGC					
	AGG									UUG					AGU					
GCA	CGA						GGA			CUA				CCA	UCA	ACA			GUA	
GOC	CGC						GGC		AUA	CUC				CCC	UCC	AOC			GUC	UAA
GOG	CGG	GAC	AAC	UGC	GAA	CAA	GGG	CAC	AUC	CUG	AAA			UUC	CCG	UCG	ACG		GUG	UAG
GCU	CGU	GAU	AAU	UGU	GAG	CAG	GGU	CAU	AUU	CUU	AAG	AUG	UUU	CCU	UCU	ACU	UGG	UAU	GUU	UGA
Ala	Arg	Asp	Asn	Cys	Glu	Gln	Gly	His	Ile	Leu	Lys	Met	Phe	Pro	Ser	Thr	Trp	Tyr	Val	Stop
A	R	D	N	C	E	Q	G	H	I	L	K	M	F	P	S	U	W	Y	V	
GCT	CGT	GAT	AAT	TGT	GAG	CAG	GGT	CAT	ATT	CTT	AAG	ATG	TTT	CCT	TCT	ACT	TGG	TAT	GTT	TGA
GCG	CGG	GAC	AAC	TGC	GAA	CAA	GGG	CAC	ATC	CTG	AAA		TTC	CCG	TCG	ACG		TAC	GTG	TAG
GOC	CGC						GGC		ATA	CTC				CCC	TOC	ACC			GTC	TAA
GCA	CGA						GGA			CTA				CCA	TCA	ACA			GTA	
	AGG									TTG					AGT					
	AGA									TTA					AGC					

RNA codons above, amino acid codes in the middle, and DNA 'codons' below.

**Table C7.3      *Sequence of MYH7 and Its Primers***

*SeqEd*, Version 1.0.3 (Myers, 1992: ABI) was used to convert the published single stranded 'sense strand' *MYH7* cDNA sequence (Jaenicke et al., 1990).

The reading frame can be deduced at any point in the *MYH7* sequence as the numbers on the top ruler identify the last base of every fifth codon. Table C7.2 in this Appendix is the genetic code. Primers were designed as described in Section 7.3.1.3 (pp. 183-187).















**APPENDIX D**

**Abbreviations and Symbols**

**Tables**

**Figures**

Abbreviations and Symbols	400
Tables	411
Figures	414

## **Abbreviations and Symbols**

A	adenine
Å	Angstrom
ABI	Applied Biosystems International
ACD	Assembly Conserved Domain (at the C-terminal of the myosin heavy chain)
<i>ACTA1</i>	skeletal muscle alpha actin gene
<i>ACTC</i>	Actin, $\alpha$ , cardiac muscle gene
AD	autosomal dominant
AD-VCPMD	AD-distal myopathy with vocal cord and pharyngeal weakness MPD2
ADHD	attention deficit hyperactivity disorder
$\alpha$ -MyHC	alpha myosin heavy chain polypeptide
$\alpha$ -SG	alpha-sarcoglycan (adhalin)
$\alpha$ - <sup>32</sup> P	alpha-phosphate of a deoxynucleoside triphosphate radiolabelled with phosphorous isotope <sup>32</sup> P
ANRI	Australian Neuromuscular Research Institute
APM	affected pedigree member
APOE	apolipoprotein E
AR	autosomal recessive
ASP	affected sib pair
B	blank
BAC	bacterial artificial chromosome
$\beta$ -SG	beta-sarcoglycan
BMD	Becker muscular dystrophy
$\beta$ -MyHC	beta myosin heavy chain polypeptide
$\beta$ -MyHC A1663P	substitution of alanine by proline at residue 1663 in beta-MyHC

BRIC	benign recurrent intrahepatic cholestasis
bp	base pairs
C	cardiac
C	cytosine
<i>C. elegans</i>	<i>Caenorhabditis elegans</i>
CANP3	calpain-3 gene
CAV3	caveolin-3 gene
CCD	central core disease
cDNA	complementary DNA
CEPH	Centre d'Etude du Polymorphisme Humain
CENP-E	Human centromeric protein-E
$\chi^2$	chi-square statistical test
CHLC	Cooperative Human Linkage Center
Ci	Curie
CK	creatine kinase
cM	centiMorgan
CMD	congenital muscular dystrophy
CMH	cardiomyopathy hypertrophic/hypertrophic cardiomyopathy
CMH1	cardiomyopathy hypertrophic 1, caused by mutations in <i>MYH7</i> , the myosin heavy polypeptide 7 cardiac muscle, beta gene
CMH2	cardiomyopathy hypertrophic 2, caused by mutations in <i>TNNT2</i> , the troponin T2, cardiac gene
CMH3	cardiomyopathy hypertrophic 3, caused by mutations in <i>TPM1</i> , the tropomyosin 1 alpha gene
CMH4	cardiomyopathy hypertrophic 4, caused by mutations in <i>MYBPC3</i> , the myosin binding protein, cardiac C3
CMH5	a 'wastebasket' CMH category for forms of CMH not linked to any known loci
CMH6	hypertrophic cardiomyopathy associated with Wolff- Parkinson-White Syndrome

CMH7	cardiomyopathy hypertrophic 7, caused by mutations in <i>TNNI3</i> , the troponin I, cardiac muscle isoform
CMH9	cardiomyopathy hypertrophic 7, caused by mutations in <i>TTN</i> , the titin gene
CMPD	Connecticut Myopathy Distal (family)
CNE	concentric needle electromyography
CNS	central nervous system
<i>COL6A1</i>	Collagen 6A1 gene
<i>COL6A2</i>	Collagen 6A2 gene
<i>COL6A3</i>	Collagen 6A3 gene
CpG islands	short stretches of DNA, often <1 kb, containing CpG dinucleotides which are unmethylated. Associated with genes at a higher frequency than they are found elsewhere in the genome
cpm	counts per minute
C-terminal	towards the carboxyl end of a polypeptide
d	map distance between the markers
D-segments	anonymous DNA fragments without gene association
DAPI	4', 6 diamidino-2-phenylindole
dATP	deoxyadenosine triphosphate
dCTP	deoxycytidine triphosphate
°C	degrees Celsius
dGTP	deoxyguanosine triphosphate
DMAT	distal anterior compartment (tibial) myopathy
DM1	myotonic dystrophy 1
DM2	myotonic dystrophy 2
DMD	Duchenne muscular dystrophy
<i>DMPK</i>	protein kinase gene with an expanded CTG repeat in its 3' untranslated region which is associated with myotonic dystrophy
DNA	deoxyribonucleic acid
dNTPs	deoxynucleoside triphosphates

D-number	anonymous number, eg patient or sample number
D-segment	anonymous DNA fragment
DOE	US Department of Energy
DRPLA	dentatorubral pallidoluysian atrophy
ds	double-stranded (DNA)
δ-SG	delta-sarcoglycan
dTTP	deoxythymidine triphosphate
<i>DYSF</i>	dysferlin gene
e	base of natural logarithms
ECG	electrocardiogram
<i>E. coli</i>	<i>Escherichia coli</i>
EDB	extensor digitorum brevis
EDMD	Emery-Dreifuss muscular dystrophy
EDTA	disodium ethylene diamine tetra-acetate
ELC	essential light chain (of the myosin II hexamer)
EM	electron microscope
<i>Emb</i>	embryonic muscle
ENMC	European Neuromuscular Centre
<i>EMD</i>	Emerin gene
EMG	electromyography
EST(s)	expressed sequence tags
F	familial (as applied to disease)
F	forward primer (in PCR or sequencing)
Φ	Yule coefficient
FA	Fanconi anaemia
<i>FACC</i>	Fanconi anaemia complementation group-C gene
FDC-CDM	AD familial dilated cardiomyopathy (FDC), cardiac conduction-system disease (CD) and myopathy (M) with adult-onset limb-girdle muscular dystrophy [FDC-CDM]



F-factors	'fertility' plasmids in bacteria
<i>F8</i>	factor VIII, haemophilia A, disease gene
FISH	fluorescent <i>in situ</i> hybridisation
5'	five prime
<i>FMR1</i>	fragile-X gene
FRAXA/FRAXE	fragile-X syndromes
FSHD	facioscapulo-humeral muscular dystrophy
FVC	forced vital capacity
G	guanine
g	grams
$\gamma$ -SG	gamma-sarcoglycan
GDB	Genome Data Base
gDNA	genomic DNA
$^3\text{H}$	tritium
$H_0$	disease and marker locus are unlinked
$H_1$	disease and marker locus are linked
h-IBM	hereditary inclusion body myopathy
HD	Huntington disease
HGP	Human Genome Project
HIV-1	human immunodeficiency virus 1
HMM	heavy meromyosin fragment of the myosin heavy chain
HUGO	Human Genome Organisation
HLA	histocompatibility complex of loci
<i>HLA-DR2</i>	HLA allele within the major histocompatibility complex
HRR	haplotype relative risk
IBD	identical by descent
IBM	inclusion body myopathy
IBM3	hereditary inclusion-body myopathy associated with a mutation in a chromosome 17 MYH gene

IBS	identical by state
IDDM	insulin-dependent diabetes mellitus
IEC	International Equipment Company
INI	intranuclear tubular filament inclusions
IRP	CpG Island Rescue PCR
<i>in situ</i>	in location
<i>in vitro</i>	in glass/in the test tube
<i>in vivo</i>	in a living system
IRE-PCR	interspersed repeat element PCR
<i>ITGA7</i>	integrin- $\alpha$ 7 gene
J	Joule
kb	kilobase
l	litre
lac-z	beta-galactosidase gene
<i>LAMA2</i>	alpha-2 laminin gene
$\lambda$	lambda bacteriophage
$\lambda$ EMBL3 SP6/T7	human genomic library (CLONTECH)
$\lambda$ Gem 11	human genomic library (Promega)
LMM	light meromyosin proteolytic fragment of the myosin heavy chain molecule which comprises the C-terminal two-thirds of the rod. Responsible for MyHC self-association and forms the backbone of the thick filament
lod (score)	the decadic logarithm of the odds that loci are linked
LGMD	limb-girdle muscular dystrophy
LIPED	linkage analysis computer program
LODM	AD-Markesbery distal myopathy
LOH	loss of heterozygosity
<i>m</i>	connecting meioses
M	molar
Mb	megabase

MCA	multiple congenital anomaly (syndromes)
MD(s)	muscular dystrophy (dystrophies)
MEI	muscle, eye and brain disease
MHC	major histocompatibility complex
MIM	Mendelian inheritance in man
$\mu$ l	microlitre
$\mu$ g	micrograms
ml	millilitre
mM	millimolar
MM	Miyoshi myopathy
MMLV-RT	Maloney murine Leukaemia virus-Reverse Transcriptase
MPD1	myopathy, distal 1
MPD2	myopathy, distal 2
MRC	Medical Research Council
mRNA	messenger ribonucleic acid
MS	multiple sclerosis
MTM	myotubular myopathy
<i>MTM1</i>	myotubularin 1 gene
MyBP-C	myosin binding protein-C
<i>MYBPC3</i>	Myosin binding protein, cardiac C3 gene
MYH	myosin heavy chain genes
<i>MYH6</i>	myosin heavy polypeptide 6, cardiac muscle, alpha gene
<i>MYH7</i>	myosin heavy polypeptide 7, cardiac muscle, beta gene
MyHC	myosin heavy chain polypeptide (of the myosin II hexamer)
MyHC-A	fruit fly <i>Drosophila melanogaster</i> embryonic muscle myosin heavy chain
<i>MYL3</i>	Myosin, light chain, alkali; ventricular and skeletal slow gene
<i>MYL2</i>	Myosin, light chain 2; regulatory ventricular gene

N	unaffected (within pedigree symbol)
NCS	nerve conduction studies
NIH	National Institutes of Health
NCBI	National Centre for Biotechnology Information
ng	nanogram
NM	non-muscle
N-terminal	towards the amide end of a polypeptide
NT-3	neurotrophin-3-deficient (mice)
OAT	ornithine- <i>d</i> -aminotransferase
OMIM	online Mendelian inheritance in man
OPMD	oculopharyngeal muscular dystrophy
ORF(s)	open reading frame(s)
<i>p</i>	probability/threshold of significance
p	short arm (petite) of human chromosomes
P1	bacteriophage plasmid vector
<i>PABP2</i>	poly(A) binding protein 2 gene
<i>PAH</i>	phenylalanine hydroxylase gene/PKU disease gene
PAC	P1-derived artificial chromosomes
<i>parA/parB</i>	F-factor genes which limit copy number to 1 or 2 per cell
PCR	polymerase chain reaction
PEI 'paper'	chromatographic acetate strip
pfu	plaque forming units
<i>pg</i>	picograms
pH	decadic logarithm of the hydrogen ion concentration
phage	bacteriophage
PKU	phenylketonuria
poly-A	RNA sequence with multiple repeating adenine bases
pREP4	cDNA expression shuttle vector
pter	p terminus of a chromosome

PTS(s)	promoter-tagged site(s)
q	long arm of human chromosomes
QMPD	Queensland Myopathy Distal (family)
?	unknown status (within pedigree symbol) OR haplotype/recombination uncertainty (next to haplotype)
qter	q terminus of a chromosome
R	reverse primer (in PCR or sequencing)
<i>RB1</i>	retinoblastoma 1 gene
RED	repeat expansion detection
REF	restriction endonuclease fingerprinting
RFLP(s)	restriction fragment length polymorphisms(s)
RH panel	radiation hybrid panel
RLC	regulatory light chain (of the myosin II hexamer)
rSSCP	RNA single-strand conformation polymorphism
RV(s)	rimmed vacuole(s)
RNA	ribonucleic acid
RNase H	enzyme which will produce nicks and gaps in cDNA: mRNA hybrids, thus creating a series of primers for subsequent amplification
RT	reverse transcriptase
S	skeletal muscle
S	sporadic
S1	myosin heavy chain proteolytic subfragment comprising the globular head, the convertor domain and an $\alpha$ -helical neck region
S2	myosin heavy chain proteolytic subfragment comprising the N-terminal one-third of the MyHC rod region which contains a myosin binding protein-C binding site at its N- terminal end, and a hinge-like region at its C-terminal end
SBMA	spinal and bulbar muscular atrophy
SCA	spinocerebellar ataxia
SDS	sodium dodecyl sulphate
SCM	sternocleido-mastoid

SCARMMD	severe childhood autosomal recessive muscular dystrophy
<i>SeqEd</i>	sequence editor (used for DNA and polypeptide sequences)
SG	sarcoglycan
SGP	sarcoglycanopathy
<i>Sm</i>	smooth muscle
SM buffer	sodium chloride, magnesium sulphate, Tris Cl buffer
SMMMyHC	human smooth muscle myosin heavy chain
SNP(s)	single nucleotide polymorphism(s)
SOD1	copper zinc superoxide dismutase
SSC buffer	sodium chloride, sodium citrate buffer
SSCP	DNA single-strand conformation polymorphism
ss	single-stranded (applied to DNA and RNA)
<i>St</i>	striated muscle
STE buffer	sodium chloride, Tris, EDTA buffer
STS	single-copy, unique 100-250 bp sequence of DNA
<i>t</i>	original symbol for recombination fraction/ probability of recombination
T	thymine
<i>Taq</i>	acronym prefix for thermostable ' <i>Taq</i> ' DNA polymerase extracted from <i>Thermus aquaticus</i> . Also used as a prefix to describe the buffer for <i>Taq</i> DNA polymerase
<i>Taq</i> polymerase	<i>Taq</i> DNA polymerase
TDT	transmission disequilibrium test
$\theta$	theta (recombination fraction)
3'	three prime
TMD	tibial muscular dystrophy (Udd myopathy)
<i>TNNI1</i>	slow skeletal muscle troponin gene
<i>TNNI3</i>	troponin I, cardiac muscle isoform gene
<i>TNNT2</i>	troponin T2, cardiac gene

<i>TPM1</i>	tropomyosin 1 alpha gene
<i>TPM2</i>	beta-tropomyosin gene
<i>TTN</i>	titin gene
Tris Cl	chloride stabilised Tris
tRNA	transfer RNA
TTE	transthoracic echocardiogram
<i>Tth</i>	acronym prefix for thermostable ' <i>Tth</i> ' DNA polymerase extracted from <i>Thermus thermus</i> . Also used as a prefix to describe the buffer for <i>Tth</i> DNA polymerase
<i>Tth</i> polymerase	<i>Tth</i> DNA polymerase
TBE	Tris-borate/EDTA electrophoresis buffer
U	unit (physiological)
u	unit (enzymatic)
UTR	untranslated region
VDR	Vitamin D Receptor
VNTR(s)	variable number tandem repeat(s)
V	volts
WA	Western Australian
WAMPD	Western Australian Myopathy Distal (family)
WAOPMD	a Western Australian family segregating a putative form of oculopharyngeal muscular dystrophy
WDM	Welander distal myopathy
wG	weeks of gestation
X	'times' or multiplied by
XL-PCR	extra-long PCR
XL-r polymerase	recombinant 'extra-long' DNA polymerase with editing/proof reading capacity
YAC	yeast artificial chromosomes

## **Tables**

	<b>Page</b>
Table 1.1 <i>Muscular Dystrophy Nosology, 1993.</i>	9
Table 2.1 <i>Physical Mapping.</i>	48
Table 2.2 <i>Cloning Vectors.</i>	50
Table 2.3 <i>Novel Genomic DNA Transcript Identification.</i>	60
Table 2.4 <i>Common PCR Mutation Detection Methods.</i>	65
Table 4.1 <i>Nosology of the Hereditary Distal Myopathies.</i>	94
Table 4.2 <i>Distribution of muscle weakness in 9 distal myopathy individuals in the WAMPD family.</i>	99
Table 4.3 <i>LIPED generated two-point lod scores between MPD1 disease locus and loci lying within &amp; flanking the chromosome 14 MPD1 candidate region.</i>	121
Table 5.1 <i>Affecteds-only LIPED generated two-point lod scores between QMPD disease locus and loci lying within and flanking the MPD1 candidate region on chromosome 14.</i>	134
Table 5.2 <i>Affecteds-only LIPED generated two-point lod scores between CMPD disease locus and loci lying within and flanking the MPD1 candidate region on chromosome 14.</i>	146
Table 5.3 <i>Laing, Silburn and Felice myopathies.</i>	151
Table 6.1 <i>Affecteds-only LIPED generated two-point lod scores between WAOPMD disease locus and loci lying within and flanking the MPD1 candidate region on chromosome 14.</i>	170
Table 7.1 <i>MYH6 cDNA/Exonic Primers.</i>	185
Table 7.2 <i>MYH7 cDNA/Exonic Primers.</i>	186



		Page
Table 7.3	MYH7 gDNA Intronic Primers.	187
Table 7.4	<i>MYH6 amplicon 12 sequences. The sequence of the WAMPD proband, III.2, and 8 unaffected individuals compared to the equivalent published sequences for MYH6 and MYH7.</i>	196
Table 7.5	<i>MYH7 Amplicon Sequencing.</i>	201
Table 7.6	<i>MYH7 cDNA Sequence Differences in Nine Western Australian Residents.</i>	202
Table 7.7	<i>MYH7 Sequence Differences in Proband III.2.</i>	206
Table 8.1	<i>Cardiomyopathy Hypertrophic, Familial (CMH) and Associated Disease Genes.</i>	225
Table 8.2	<i>Derived <math>\beta</math>-MyHC Rod Region Amino Acid Sequence for WAMPD Laing Myopathy Proband, III.2.</i>	234
Table 8.3	<i>Comparison of Number 30 Human <math>\beta</math>-MyHC 28-Residue Repeat to Equivalent Regions in other MyHCs.</i>	237
Table 8.4	<i>COILS Analysis of Polypeptides with Varying Degrees of Homology to Human <math>\beta</math>-MyHC.</i>	263
Table 9.1	<i>Molecular Genetics and Autosomal Dominant Muscular Dystrophy Nosology.</i>	278
Table 9.2	<i>Molecular Genetics and X-linked Recessive and Autosomal Recessive Muscular Dystrophy Nosology.</i>	279
Table 9.3	<i>Nosology of the Hereditary Distal Myopathies.</i>	282
Table 9.4	<i>Classification of Distal Myopathies based on Gene Location.</i>	285
Table 9.5	<i>Distal Myopathies Excluded from the Distal Myopathy Nosology.</i>	296

		<b>Page</b>
Table B4.1	<i>Lod Scores for Linkage between the MPD1 Locus and Marker Loci.</i>	380
Table B5.1	<i>Lod Scores for Linkage between the QMPD Locus and Marker Loci.</i>	384
Table C7.1	<i>Sequence Comparisons of MYH6 and MYH7 cDNA, and of <math>\alpha</math>-MyHC and <math>\beta</math>-MyHC Amino Acids.</i>	387
Table C7.2	<i>Genetic and Amino Acids Codes.</i>	396
Table C7.3	<i>Sequence of MYH7 and its Primers.</i>	397

## **Figures**

		<b>Page</b>
Figure 3.1	Autoradiographs of a $\lambda$ Gem 11 genomic library quaternary screen for positive <i>TNNI1</i> clones.	88
Figure 3.2	A whole metaphase showing <i>in situ</i> hybridisation of the biotinylated <i>TNNI1</i> probe to 1q32.	89
Figure 4.1	Pedigree of a Western Australian family segregating distal myopathy (WAMPD family).	96
Figure 4.2	Vastus lateralis biopsy from III.2.	101
Figure 4.3	Microsatellite D22S258 and VNTR D17S26 segregation in the WAMPD pedigree.	107
Figure 4.4	Microsatellite D4S174 and VNTR D2S44 segregation in the WAMPD pedigree.	108
Figure 4.5	Identification of the MPD1 candidate region.	110
Figure 4.6	Multipoint analysis for the markers D14S72, D14S50, MYH7, D14S64, D14S54 & D14S49 assuming 100%, 80% and 50% penetrance.	112
Figure 4.7	WAMPD family 'affecteds-only' exclusion map.	113
Figure 4.8	WAMPD family 'affecteds-only' exclusion map with chromosome 14 data removed.	114
Figure 4.9	Chromosome 14 sex-averaged genetic linkage maps.	116
Figure 4.10	Microsatellites D14S257 and D14S262 segregation in the WAMPD pedigree.	118
Figure 4.11	Microsatellite markers defining the MPD1 candidate region.	119
Figure 4.12	Refinement of the MPD1 candidate region.	120

		<b>Page</b>
Figure 5.1	Pedigree of a Queensland family (QMPD) segregating distal myopathy.	128
Figure 5.2	Segregation of the MPD1 candidate region microsatellite, D14S64 in the QMPD pedigree.	133
Figure 5.3	Genotypes of participating QMPD pedigree members for marker loci spanning the MPD1 candidate region.	136
Figure 5.4	Segregation of chromosome 17 microsatellite GX-Alu in QMPD pedigree.	138
Figure 5.5	Affecteds-only exclusion map for QMPD family.	139
Figure 5.6	Pedigree of a Connecticut family (CMPD) segregating distal myopathy.	141
Figure 5.7	Genotypes of participating CMPD pedigree members for marker loci spanning the MPD1 candidate region.	148
Figure 5.8	Multipoint linkage analysis of distal myopathy at 70% penetrance in a single family from Connecticut against fixed maps of markers on chromosome 14.	149
Figure 6.1	Pedigree of a Western Australian family (WAOPMD) segregating putative-OPMD.	164
Figure 6.2	Segregation of the microsatellite markers, D14S283 and D14S49, in the WAOPMD pedigree.	167
Figure 6.3	Segregation of the microsatellite marker D14S252 in the WAOPMD pedigree.	168
Figure 6.4	Genotypes of participating WAOPMD pedigree members for marker loci spanning the MPD1 candidate region and the refined French-Canadian OPMD candidate region.	169
Figure 6.5	'Affecteds-only' exclusion map for the WAOPMD family.	172

		Page
Figure 7.1	RNA extractions from human muscle using the TRI-REAGENT method.	193
Figure 7.2	The copper zinc superoxide dismutase (SOD1) test to ascertain the success of cDNA manufacture.	194
Figure 7.3	Chromatogram for the <i>MYH6</i> bases, 1947-1982.	197
Figure 7.4	The original six <i>MYH7</i> amplicons.	198
Figure 7.5	The reverse sequence chromatogram of the illegitimate 42F<->1041R amplicon over its last twenty-two 3' bases.	199
Figure 7.6	Differences to the published <i>MYH7</i> cDNA sequence detected in the <i>MYH7</i> cDNA sequences of the WAMPD Laing myopathy proband, III.2, and in nine unaffected, unrelated WA residents.	203
Figure 7.7	Differences to the published <i>MYH7</i> cDNA sequence detected only in the <i>MYH7</i> cDNA sequences of the WAMPD Laing myopathy proband, III.2.	206
Figure 7.8	The segregation of the gDNA, <i>MYH7</i> 23623-23633 base sequence in the WAMPD pedigree.	208
Figure 7.9	The chromatograms of <i>MYH7</i> gDNA sequence 23623<->23633 in selected members of the WAMPD family.	209
Figure 7.10	The segregation of <i>MYH7</i> 23623-23633 in the WAMPD pedigree, from III.2 to IV:1 and IV:3.	210
Figure 7.11	<i>MYH7</i> cDNA amplicon 10 and <i>MYH7</i> gDNA amplicon 11 sequences for two European families segregating early onset AD-MPD.	213
Figure 7.12	Screening the <i>MYH7</i> gDNA amplicon 11 (23531F<->23825R) for the G23628C SSCP.	214

		<b>Page</b>
Figure 8.1	A schematic synopsis of the myosin II hexamer.	219
Figure 8.2	Superimposition of CMH1 mutations on a cartoon of the <i>MYH7</i> gene.	221
Figure 8.3	A dipeptide showing two peptide bonds, two C $\alpha$ -C $\alpha$ vectors and the conventional labelling of carbon atoms in amino acids.	232
Figure 8.4	Section of a right-handed polypeptide $\alpha$ -helix.	240
Figure 8.5	Molecular model of proline in a polypeptide $\alpha$ -helix.	242
Figure 8.6	Steric clashes between the proline C $\beta$ H <sub>2</sub> group and preceding C $\delta$ H <sub>2</sub> and NH groups.	243
Figure 8.7	A curved helix involving residues 14 to 32 in avian pancreatic polypeptide.	247
Figure 8.8	A schematic diagram showing how the $\alpha$ -helical kink angle ( $\theta$ ) is defined.	251
Figure 8.9	The pattern formed by the side chains of two superimposed $\alpha$ -helices.	254
Figure 8.10	Two right-handed $\alpha$ -helices forming a left-handed, coiled-coil superhelix.	255
Figure 8.11	The global parameters of a two stranded coiled-coil.	256
Figure 8.12	Two heptad wheels showing the sequence motif of an $\alpha$ -helical, coiled-coil dimer.	257
Figure 8.13	Crick's $\alpha$ -helical, coiled-coil dimer with heptad repeats.	258
Figure 8.14	Helix-net representation of dimer coiled-coil interactions.	259

		<b>Page</b>
Figure 8.15	Cross-sectional view of amino acid side chain interactions in the hydrophobic core of a coiled-coil dimer.	260
Figure 8.16	Model of assembly of myosin molecules to form the thick filament.	266
Figure 8.17	Schematic diagram of cross-bridge and helical periodicity in thick filaments as determined from X-ray diffraction patterns.	268
Figure 8.18	A single skip residue causing a 100° azimuthal shift of MyHC hydrophobic core residues.	270