

2021

Perennial ryegrass contains gluten-like proteins that could contaminate cereal crops

Sophia Escobar Correias
Edith Cowan University

James A. Broadbent

Alicja Andraszek

Sally Stockwell

Crispin A. Howitt

See next page for additional authors

Follow this and additional works at: <https://ro.ecu.edu.au/ecuworkspost2013>



Part of the [Nutrition Commons](#)

[10.3389/fnut.2021.708122](https://doi.org/10.3389/fnut.2021.708122)

Escobar-Correias, S., Broadbent, J. A., Andraszek, A., Stockwell, S., Howitt, C. A., Juhász, A., & Colgrave, M. L. (2021). Perennial ryegrass contains gluten-like proteins that could contaminate cereal crops. *Frontiers in Nutrition*, 8, article 708122. <https://doi.org/10.3389/fnut.2021.708122>

This Journal Article is posted at Research Online.
<https://ro.ecu.edu.au/ecuworkspost2013/10904>

Authors

Sophia Escobar Correas, James A. Broadbent, Alicja Andraszek, Sally Stockwell, Crispin A. Howitt, Angela Juhasz, and Michael Colgrave



Perennial Ryegrass Contains Gluten-Like Proteins That Could Contaminate Cereal Crops

Sophia Escobar-Correas^{1,2†}, James A. Broadbent^{1†}, Alicja Andraszek¹, Sally Stockwell¹, Crispin A. Howitt³, Angéla Juhász² and Michelle L. Colgrave^{1,2*}

¹ CSIRO Agriculture and Food, St. Lucia, QLD, Australia, ² Australian Research Council Centre of Excellence for Innovations in Peptide and Protein Science, School of Science, Edith Cowan University, Joondalup, WA, Australia, ³ CSIRO Agriculture and Food, Canberra, ACT, Australia

OPEN ACCESS

Edited by:

Anil K. Verma,
Marche Polytechnic University, Italy

Reviewed by:

Mahir Turhan,
Mersin University, Turkey
Priscila Farage,
Universidade Federal de Goiás, Brazil
Giorgia Vici,
University of Camerino, Italy

*Correspondence:

Michelle L. Colgrave
michelle.colgrave@csiro.au

[†] These authors have contributed
equally to this work and share first
authorship

Specialty section:

This article was submitted to
Food Chemistry,
a section of the journal
Frontiers in Nutrition

Received: 11 May 2021

Accepted: 23 June 2021

Published: 28 July 2021

Citation:

Escobar-Correas S, Broadbent JA,
Andraszek A, Stockwell S, Howitt CA,
Juhász A and Colgrave ML (2021)
Perennial Ryegrass Contains
Gluten-Like Proteins That Could
Contaminate Cereal Crops.
Front. Nutr. 8:708122.
doi: 10.3389/fnut.2021.708122

Background: To ensure safe consumption of gluten-free products, there is a need to understand all sources of unintentional contamination with gluten in the food chain. In this study, ryegrass (*Lolium perenne*), a common weed infesting cereal crop, is analysed as a potential source of gluten-like peptide contamination.

Materials and Methods: Ten ryegrass cultivars were analysed using shotgun proteomics for the presence of proteins from the prolamin superfamily. A relative quantitative assay was developed to detect ryegrass gluten-like peptides in comparison with those found in 10 common wheat cultivars.

Results: A total of 19 protein accessions were found across 10 cultivars of ryegrass for the protein families of PF00234-Tryp_alpha_amyl, PF13016-Gliadin, and PF03157-Glutinin_HMW. Protein and peptide homology searches revealed that gliadin-like peptides were similar to avenin and gamma-gliadin peptides. A total of 20 peptides, characteristic of prolamin superfamily proteins, were selected for liquid chromatography mass spectrometry (LC-MS) with multiple reaction monitoring (MRM). Only two of the monitored peptides were detected with high abundance in wheat, and all others were detected in ryegrass. Glutenin and alpha-amylase/trypsin inhibitor peptides were reported for the first time in ryegrass and were noted to be conserved across the Poaceae family.

Conclusion: A suite of gluten-like peptides were identified using proteomics that showed consistent abundance across ryegrass cultivars but were not detected in wheat cultivars. These peptides will be useful for differentiating wheat gluten contamination from ryegrass gluten contamination.

Keywords: wild grass, cereal, ryegrass, gluten, wheat, proteomics, LC-MS/MS

INTRODUCTION

Gluten proteins are the most abundant proteins found in commercial cereal grains, including wheat, barley, rye and oats (1). Consumption of these proteins will trigger gluten-related disorders (GRD) in ~100 million people globally (2, 3). At least six GRDs (4) have been described; these may be autoimmune, allergic, or neither and are caused by a mix of environmental and genetic factors

(4, 5). Coeliac disease (CD) is the most recognised GRD, which is currently diagnosed based on serology and small intestinal biopsies and is estimated in 0.7–1.4% of the global population (2). Affected individuals are genetically susceptible to generate an autoimmune inflammatory response in the small intestine when exposed to gluten proteins. Long-term exposure results in chronic intestinal inflammation and villi degradation for these individuals (6).

The only treatment for GRDs is to avoid the intake of trigger proteins, i.e., commit to a life-long gluten-free diet (GF-diet), although this is far from simple or easy. Complications arise through the difficulty in avoiding gluten-containing additives, inadvertent gluten intake, and food contamination. Gluten is frequently included in nongrain-based items such as sausages, salad dressing, imitation meats, and even some medications (7, 8). Therefore, the GF-diet mainly consists of fruits, vegetables, legumes, meat, seafood, nuts, dairy and bakery products, which include GF cereals or pseudocereals, such as rice, corn, quinoa and millet (9, 10). Nevertheless, these products may contain hidden gluten due to unintentional contamination through the food supply chain.

Cross-contamination within the food supply chain can happen at different stages, such as production, milling, export and retail. There are multiple possibilities from the moment the grain is grown until the GF-flour is packed, including agricultural co-mingling through crop rotation, storage and transport. Some studies provide examples in the contamination of GF-oats with gluten-containing cereals during harvest (11–13). One further aspect in this regard concerns the contamination of GF products due to weeds growing in the field, a topic that is under-researched. Weed management in crops is a challenging task. Farmers spend thousands of dollars each year in an effort to control weed invasion; however, this outcome is not always completely accomplished, and farmers must deal with weed seed contamination, which can become a serious problem for GF-cereals and other crops that are supposed to be free from gluten, such as pulses (14).

One of the most common weed seeds found in cereal samples is ryegrass (genus *Lolium*) (14). This grass is the most common weed infesting cereal grain fields in Australia and has small dense seeds that are difficult to eliminate during automated grain cleaning (14, 15). Ryegrass belongs to the same Poaceae grass family as wheat and other gluten-containing crops, wherein the storage proteins primarily comprise gluten-like proteins. Ryegrass has been subjected to Western blotting followed by liquid chromatography mass spectrometry/mass spectrometry (LC-MS/MS) analysis, demonstrating that it contains proteins with structural similarity with gluten proteins (15).

Herein, a combination of discovery and targeted proteomic analyses were undertaken to confirm the presence of gluten-like proteins in ryegrass cultivars. A targeted LC-MS/MS assay was developed and used to assess relative levels of target peptides across 10 ryegrass cultivars and 10 wheat cultivars to identify the differences in peptide abundance patterns and to identify possible ryegrass-specific peptide markers.

This study investigates whether gluten contamination can potentially originate from sources other than traditional cereal grains, such as field contaminants. Understanding the possible

origins of gluten contamination and establishing identification and quantification methods for these new protein species could help to provide a more accurate characterisation of food and assure food safety for the population affected by GRDs.

MATERIALS AND METHODS

Sample Material

Grain samples from 10 ryegrass and 10 wheat cultivars were obtained (Supplementary Table 1) from the Australian pasture collection and Australian Winter Cereals Collection (Tamworth, Australia). All samples were manually inspected to exclude foreign seed contamination. Flour samples were obtained by milling with a Metefem Hungarian Mill (model FQD2000, Hungary).

Protein Extraction and Digestion

Methods were performed according to the study of Bose et al., with minor changes (16). Flour (20 mg) was weighed (four replicates of wheat and three replicates of ryegrass) and mixed with 200 μ L of 55% isopropanol and 2% dithiothreitol (IPA/DTT). Samples were vortexed and sonicated for 5 min and then incubated on a thermo heating mixer block (Thermo Scientific) at 300 rpm at 65°C for 45 min. The mixtures were centrifuged for 15 min at 20,800 \times g, and the supernatants were transferred to Protein LoBind Tubes (Eppendorf). The protein content was quantified *via* Bradford assay, and 200 μ g of protein was loaded onto 10-kDa molecular weight cut-off (MWCO) filters (Millipore, Sydney, Australia). Aliquots were adjusted to 200 μ L with UA buffer (8 M urea in 0.1 M Tris-HCl, pH 8). The filter content was washed two times with 200 μ L UA buffer followed by centrifugation at 20,800 \times g for 15 min at room temperature (RT).

Alkylation of cysteines was performed by the addition of 100 μ L of 50 mM iodoacetamide (in UA buffer) and incubation in the dark at RT, with 300 rpm shaking for 20 min. Samples were centrifuged for 15 min at 20,800 \times g and then washed with 200 μ L UA buffer and centrifuged again, as previously described. Flowthrough was discarded and the buffer was changed to 50 mM AmBic (ammonium bicarbonate, pH 8.0) by washing the filters two times with 200 μ L of this digestion buffer followed by centrifugation. Filters were transferred to fresh collection tubes. Protein digestion was achieved by adding 200 μ L of trypsin (Promega, Wisconsin, USA) prepared as 200 μ g/mL in 50 mM AmBic (pH 8.0) containing 1 mM CaCl₂ and incubated overnight in the dark at 37°C with 300 rpm shaking. Sample filters were centrifuged at 20,800 \times g for 15 min, followed by two washes with 200 μ L of AmBic (pH 8.0) and centrifugation. Filtrates were evaporated to dryness in a Savant SpeedVac concentrator (Thermo Scientific, USA). Peptides were reconstituted in 0.1% formic acid to a protein concentration of 1 μ g/ μ L for LC-MS/MS analysis.

Mass Spectrometry and Protein Identification

Liquid chromatography mass spectrometry was performed using an Ekspt nanoLC415 (Eksigent, Dublin, CA) coupled to a TripleTOF® 6600 mass spectrometer (SCIEX, Redwood City,

CA, USA). The specifications of the acquisition parameters have been previously described (15, 16).

Discovery data was generated for 10 ryegrass cultivars, and the spectra were searched against the Poaceae grass family subset of the UniProt database (version 2021/01) appended with the common repository of adventitious proteins. The database from a higher taxonomic group was used due to poor representation of the *Lolium perenne* proteome (748 nonredundant protein sequences). The UniProt UniRef 100 redundancy reduction was applied to remove Poaceae proteins with 100% sequence identity, leading to a total of 1,953,474 protein sequences.

ProteinPilot v5.0.3 software (SCIEX) encompassing the Paragon and ProGroup algorithms (17) was used to identify peptides, infer proteins, and generate false discovery rate (FDR) reports. Results from discovery analysis were curated using an in-house script (git-hub/Sophia-006¹) (18). To ensure quality in the identification of gluten proteins and peptides, the following curation parameters were applied: 1% FDR or 99% peptide confidence, requisite tryptic and semi-tryptic peptides, up to two missed cleavages, and variable modifications of glutamine to pyro-glutamic acid, carbamidomethyl cysteine, and oxidation of methionine.

Protein summaries (Supplementary Data 1) were analysed to identify proteins of interest, i.e., gluten proteins. In this regard, a protein family (Pfam) search was performed that encompassed searching for three specific domains: Gliadin (PF13016), Glutenin_HMW (PF03157), and Tryp_alpha_amyl (PF00234). The Pfam search was performed using profile hidden Markov models 3 (HMMER3) (19). Protein and peptide homology searches were performed using the BLAST algorithm and Peptide Search tool available at UniProt², respectively.

Proteins that contained gliadin, glutenin_HMW, and Tryp_alpha_amyl domains were selected for targeted assay development. These protein sequences were gathered into a FASTA file that was used to construct a table of peptide multiple reaction monitoring (MRM) transitions using Skyline Software (MacCoss Lab Software, Washington, USA) (20). MRM transitions were determined for each peptide. Peptides were selected for MRM based on the following criteria: (1) tryptic or semi-tryptic; (2) identified with 95% confidence; (3) unmodified or common modifications only; and (4) high peak signal intensity. Transitions for each selected peptide were prioritised from the acquired discovery data, including precursor ion (Q1) and fragment ion (Q3) m/z, and rolling collision energy (Supplementary Data 2). In total, 20 peptides were selected for MRM experiments. Four semi-tryptic peptides were included; their fully tryptic versions were not included in the final MRM due to insufficient evidence in the discovery data. Three MRM transitions were monitored per peptide based on the intensity and lack of interferences.

Digested peptides were separated with a Shimadzu Nexera UHPLC (Rydalmere, Australia) and analysed with a 6500 QTRAP mass spectrometer (SCIEX), as described previously (15, 16). Relative quantitation was achieved using scheduled

MRM scanning experiments with a 60-s detection window for each MRM transition with retention time as determined in preliminary MRM experiments and a 0.6-s cycle time. Peaks were integrated using Skyline, wherein all three transitions were required to coelute at the same retention time (min) with a signal-to-noise ratio (S/N) > 3. The peak areas for the three monitored MRM transitions were summed. The mean, SD, and co-efficient of variation (CV) were calculated for technical replicates for each peptide (Supplementary Data 3). Batch and injection order effects were removed from the data by monitoring external standards interspersed with the unknown samples while retaining the differences between ryegrass and wheat samples.

Graphs were generated in the R statistical computing environment using the ggplot package (21) and Morpheus³ (Broad Institute, Cambridge MA, USA).

RESULTS

Gluten Protein Identification Yield From Shotgun Proteomics

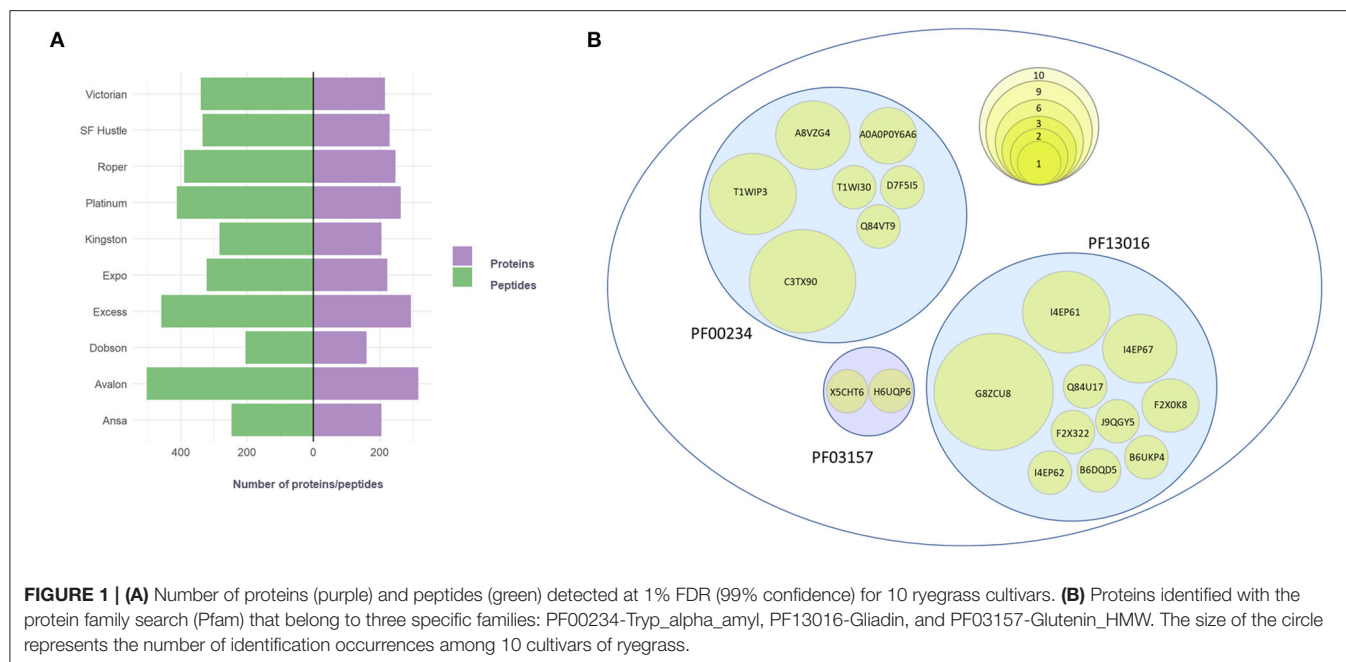
Ten varieties of *L. perenne* were studied (Supplementary Table 1) to confirm the suspected presence of gluten in this species. For the identification of gluten proteins, high-resolution data acquisition and database searching was performed using a UniProt database from a higher taxonomic group belonging to the family Poaceae. The total number of proteins identified with 99% confidence between the 10 varieties of ryegrass varied between 160 and 316 (Figure 1A). Additionally, peptides discovered with tryptic digestion varied between 205 to 503 (Figure 1A).

Next, a protein family domain search was performed for each of the cultivars with the aim to find proteins containing domains characteristic of gliadin, glutenin_HMW, and tryp_alpha_amyl domain families. The results revealed that between three to eight candidate proteins were found for each of the cultivars (Supplementary Data 4). A total of 19 protein accessions were found among the 10 varieties of ryegrass. A table showing positive identification of proteins representing each domain family across the varieties is provided in Supplementary Data 4. For the tryp_alpha_amyl-domain containing protein family (Figure 1B), the candidate protein with the highest number of identification occurrences was UniProt accession C3TX90 (present in nine varieties), followed by proteins T1WIP3 (present in six varieties) and A8VZG4 (present in three varieties). The lowest number of identification occurrences were for A0A0P0Y6A6 (present in two varieties) and for proteins D7F5I5, Q84VT9, and T1WI30 (present in one variety). The gliadin-domain containing protein family (Figure 1B) was represented by the proteins G8ZCU8 (present in all the varieties) and I4EP61 (present in six varieties). Further gliadin-domain containing members were detected between one and three varieties (I4EP57, F2X0K8, I4EP62, J9QGY5, B6DQD5, Q84U17, F2X322, and B6UKP4). The glutenin_HMW-domain containing protein family (Figure 1B) was the family with the lowest representation

¹https://github.com/Sophia-006/FDR_proteins_peptides.git

²Uniprot.org

³<https://software.broadinstitute.org/morpheus/>



with two candidate proteins, H6UQP6 and X5CHT6, detected in one ryegrass variety.

Of note, these domains could represent multiple protein types due to the mixed nature of gluten-type proteins, i.e., families which comprise a range of proteins that have the Tryp_alpha_amyl domain or the gliadin-domain are not necessary α -amylase/trypsin inhibitor proteins or gliadins. As such, additional protein and peptide homology searches were performed to determine the full-length homology of these proteins.

Gluten proteins tend to have conserved regions, and further bioinformatic analysis is necessary to determine which peptides could function as peptide markers to differentiate between grass species. As a result, candidate proteins and peptides were subject to searches to identify non-ryegrass orthologues, i.e., to determine the potential for these peptide sequences to be observed in additional species. In consideration of the limitations of mass spectrometric detection, search settings included leucine and isoleucine equivalence (isobaric amino acids) and a requisite C-terminal arginine or lysine (trypsin cleavage site).

Results of the searches are shown in **Table 1**, and a detailed table specifying the species within the tribe identity match can be found in **Supplementary Table 2**. Peptides SQILQQSSCQVMR (G8ZCU8), CPAHSVVQAILQK (I4EP61), QFLVQQCSPVAEVPFLR (I4EP61), and QQAQFEGMR (I4EP57) were found to belong exclusively to avenins present in oats, while peptides QQCCQQLAQIPQQLR (F2X0K8) and APFASIVASIGGQE (F2X322) belong to gamma-gliadins that are present in wheat. An additional gamma-gliadin peptide, APFASIVAGIGGQYR (B6DQD5), was found in species of the Triticeae tribe.

Two peptides, QQCCQQLAQIPEQSR and SQMLQQSSCHVIR, from the protein J9QGY5, which is a

low molecular weight glutenin (LMW-GS), belong exclusively to the wild grass species *Dasyphyrum villosum* of the Triticeae tribe. HMW glutenin (HMW-GS) peptides DVSAKCRPVAVSQVAR, ELQESSLEACRQVVDQQLAGR, and QLQCERELQESSLEACR (X5CHT6) were moderately conserved and found across species of the Triticeae tribe. The peptide DGSFYPPGEATPPQQLQQR (H6UQP6) was exclusively found in *Elymus libanoticus*, a species that belongs to the Triticeae tribe.

Peptide EGMEVFPGCR (T1WIP3) was found exclusively in the species *Elymus grandis* (Triticeae); peptide LTAASVPAVCK (T1WI30) was found in several species of the Triticeae tribe. The peptide LLQQQLNPCR (A8VZG4) was found in an α -gliadin sequence of species *Dasyphyrum hordeaceum* (Triticeae). For A0A0P0Y6A6, which presented protein homology to a lipid-transfer protein, the conserved peptide TACNCLK was found in species of families Poaceae, Fabaceae, Asteraceae, among others; however, the peptide CGVSIPYTISPISIDCSR was exclusively found in species of rice of tribe Oryzeae. The two peptides DPYYEQCPMRK and SDLYGPNLQGEVTMLMER (C3TX90), from a puroindoline-like protein were found exclusively in *Brachypodium sylvaticum* of tribe Brachypodieae. Peptide QLSQIAPQCR (D7FSI5), characteristic of puroindolines and hordoidolines, was present in several species of the Triticeae tribe.

Gluten-Like Peptide Quantification in Wheat and Ryegrass

Peptides characteristic of gluten-like proteins were selected for LC-MRM-MS analysis. Based on the results obtained from homology searches, 20 peptides from 13 proteins were measured with a relative quantitative analysis (**Supplementary Data 5**). The selected peptides were measured in 10 cultivars of perennial ryegrass (*L. perenne*) to determine the abundance of gluten

TABLE 1 | Protein and peptide homology search results.

Peptide sequence	Protein accession	Protein family domain	Protein homology (from BLAST)	Peptide search 100% identity match (Tribe)
SQILQQSSCQVMR	G8ZCU8	PF13016	Avenin	Aveneae
CPAIHSVVQAILQK	I4EP61	PF13016	Avenin	Aveneae
QLFVQQCSPVAEVPFLR	I4EP61	PF13016	Avenin	Aveneae
AFALQALPAMCDVYVPPHCSVA	I4EP61	PF13016	Avenin	Aveneae
QQAQFEGMR	I4EP57	PF13016	Avenin	Aveneae
QCCQQLAQIPQQLR	F2X0K8	PF13016	Gamma-gliadin	Triticeae
APFASIVASIGGQE	F2X322	PF13016	Gamma-gliadin	Triticeae
APFASIVAGIGGQYR	B6DQD5	PF13016	Gamma-gliadin	Triticeae
QCCQQLAQIPEQSR	J9QGY5	PF13016	LMW-glutenin	Triticeae
SQMLQQSSCHVIR	J9QGY5	PF13016	LMW-glutenin	Triticeae
DVSAKCRPVAVSQVAR	X5CHT6	PF03157	HMW-glutenin	Triticeae
ELQESSLEACRQVVDQQLAGR	X5CHT6	PF03157	HMW-glutenin	Triticeae
ELQESSLEACR	X5CHT6	PF03157	HMW-glutenin	Triticeae
QLQCERELQESSLEACR	X5CHT6	PF03157	HMW-glutenin	Triticeae
DGSFYFGEATPPQQLQQR	H6UQP6	PF03157	HMW-glutenin	Triticeae
RCCDELSAIPAYCR	Q84VT9	PF00234	Trypsin inhibitor	Triticeae
LQCVGSQVPEAVLR	T1WIP3	PF00234	Dimeric alpha-amylase inhibitor	Triticeae
EGMEVFPGCR	T1WIP3	PF00234	Dimeric alpha-amylase inhibitor	Triticeae
LLQQQLNPCR	A8VZG4	PF00234	Alpha-gliadin	Triticeae
LTAASVPAVCK	T1WI30	PF00234	Dimeric alpha-amylase inhibitor	Triticeae
TACNCLK	A0A0P0Y6A6	PF00234	Nonspecific lipid-transfer protein	Highly conserved. Families: Poaceae, Fabaceae, Asteraceae
CGVSIPYITSPSIDCSR	A0A0P0Y6A6	PF00234	Nonspecific lipid-transfer protein	Oryzeae
DPYEQCPMRK	C3TX90	PF00234	Puroindoline-like protein	Brachypodieae
SDLYGPNLQGEVTMLMER	C3TX90	PF00234	Puroindoline-like protein	Brachypodieae
QLSQIAPQCR	D7FSI5	PF00234	Puroindoline, Hordoindoline	Triticeae

Peptide sequence, protein accession and protein family are specified. Protein homology refers to all protein types where the peptide was found. Peptide search results refer to all tribes where the peptide was found with a 100% identity match.

candidates. At the same time, these same peptides were measured in 10 wheat varieties to compare the abundance between these two species, aiming to discover peptides that can differentiate ryegrass from wheat.

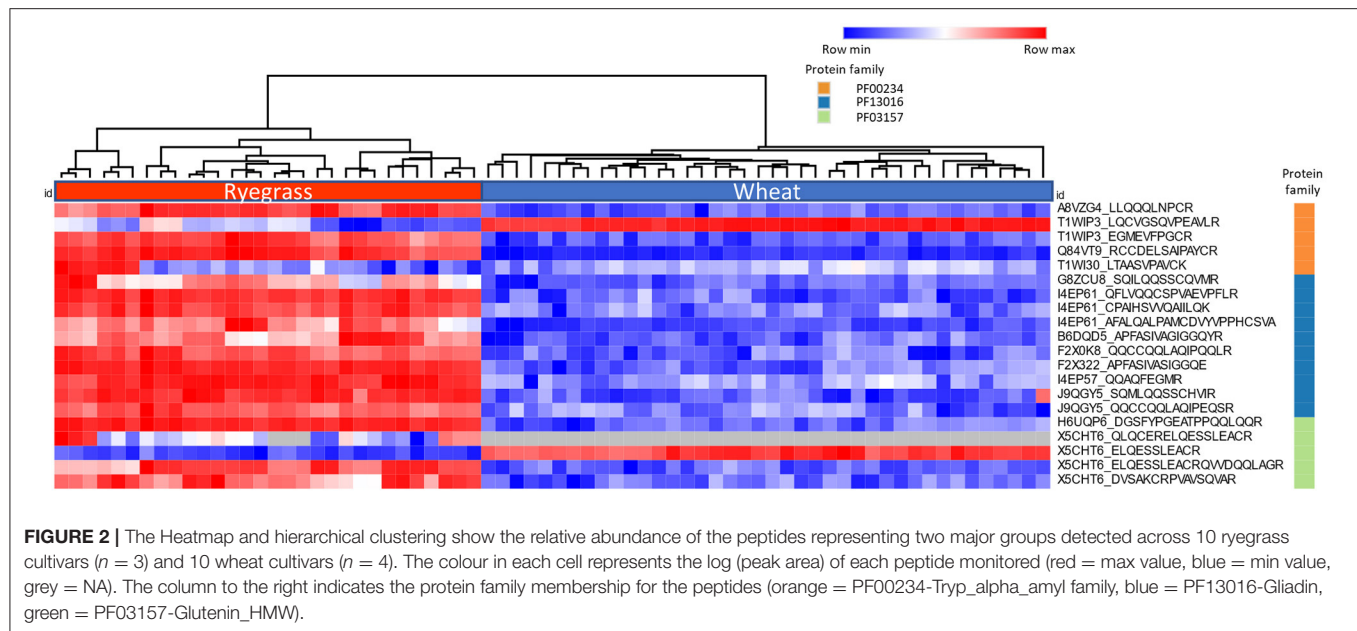
The heatmap (**Figure 2**) shows the logarithmic relative abundance of the measured peptides across 10 ryegrass cultivars ($n = 3$) and 10 wheat cultivars ($n = 4$). The complete Euclidean linkage method was used for hierarchical clustering, which shows clear stratification of ryegrass and wheat. Protein family domain membership for each protein is also specified within the graph. The analysis revealed the peptides measured by LC-MRM-MS in ryegrass cultivars had low abundance or were not detected in wheat.

Relative Quantitation of Gluten Proteins Across Ten Cultivars of Ryegrass

Except for the peptides mentioned in the section Gluten-like peptide quantification in wheat and ryegrass above, the remaining peptides were only quantified in ryegrass. Their summary MRM peak area results are shown in **Figures 3–5** according to their protein family domain membership.

The biological variation between cultivars is presented in **Supplementary Data 6**.

Results for the gluten-like peptides of family PF13016 Gliadin domain-containing proteins (akin to avenin-like proteins and gamma-gliadins) are shown in **Figure 3**. Peptide SQILQQSSCQVMR (**Figure 3A**) from protein G8ZCU8 showed high variability across ryegrass cultivars with the highest abundance in the cultivar Victorian. Peptide QQAQFEGMR (**Figure 3B**) from protein I4EP57 showed variability among all cultivars of ryegrass with a biological co-efficient of variation (CV) of 32%. Peptide AFALQALPAMCDVYVPPHCSVA (I4EP61) varied across the ryegrass cultivars and showed high values; however, CPAIHSVVQAILQK and QFLVQQCSPVAEVPFLR (I4EP61) (**Figure 3C**) showed lower values but were consistently found among all ryegrass cultivars, with a biological CV of <20%. The peptide QCCQQLAQIPQQLR (**Figure 3D**) from protein F2X0K8 shows moderate variance across the cultivars, showing higher values for Dobson and Victorian. Peptide APFASIVASIGGQE from F2X322 (**Figure 3E**) was consistent across all ryegrass cultivars showing a CV of 20%. Peptide APFASIVAGIGGQYR from protein B6DQD5 was variable among ryegrass cultivars and was higher in cultivars Expo



and SF Hustle. From protein J9QGY5, two peptides were measured (**Figure 3G**): QCCCQLAQIPEQSR was not only more abundant but also more variable across cultivars and was higher in the Dobson cultivar and SQMLQSSCHVIR was in lower abundance but was consistent ($CV < 25\%$) across ryegrass cultivars.

Figure 4 shows the quantitation of peptides from proteins belonging to the family PF03157 HMW-glutenin. Peptides belonging to two proteins, namely X5CHT6 and H6UQP6, were measured. Three peptides were quantified from protein X5CHT6, each showing high variability across the cultivars. Peptide DVSAKCRPVAVSQVAR showed higher abundance in Kingston, Platinum, Roper and SF Hustle; peptide ELQESSLEACRQVVDQQLAGR was highest in abundance in Dobson, Excess and SF Hustle, while peptide QLQCERELQESSLEACR was the highest in Victorian (**Figure 4A**). The second HMW glutenin, protein H6UQP6, was quantified using peptide DGSFYPGPEATPPQQLQQR with low biological variance across cultivars with a $CV < 15\%$ (**Figure 4B**).

The detection of peptides from PF00234 Tryp_alpha_amyl domain-containing proteins are shown in **Figure 5**; these proteins were also shown to have homology to gluten-like proteins (**Table 1**). One peptide EGMEVFPQCR was observed to be variable across ryegrass cultivars. This peptide is found in protein T1WIP3, a dimeric alpha-amylase inhibitor (**Figure 5A**). Peptide LLQQQLNPCR was variable across ryegrass cultivars and is present in protein A8VZG4, which is an α -gliadin (**Figure 5B**). Another peptide LTAASVPAVCK showed variable abundance with high levels in cultivars Roper and Victorian. This peptide is found in protein T1WI30, a dimeric alpha-amylase inhibitor (**Figure 5C**). Peptide RCCDELSAIPAYCR (detected as protein Q84VT9, a trypsin inhibitor) showed

good signal intensity but varied across the ryegrass cultivars (**Figure 5D**).

DISCUSSION

Ryegrass (genus *Lolium*) has been identified as one of the most challenging weeds for cropping systems due to its ability to resist herbicides, consequently affecting different farming practises (22). Furthermore, preliminary investigation of ryegrass proteins has revealed its potential as a source of gluten contamination (15, 23, 24). In this regard, an isolation and purification method for the prolamin fraction of ryegrass grains, which was named loliin, was established as early as the 1930s (23). This study preceded the study of Shewry in 1986, who characterised prolamins from different grasses, including *L. perenne*, and established homology of ryegrass γ -prolamins to those from wheat, barley and rye. More recently, Colgrave et al. revealed the reactivity of ryegrass prolamins to the anti-gliadin antibody and identified possible antigenic proteins through LC-MS analysis of gel-separated proteins (15).

Herein, we report the identification of prolamin super-family peptides in ryegrass cultivars and their quantitation when compared with wheat, with the aim to measure differences in peptide abundance and identify potential peptide markers of ryegrass contamination. To this end, grain proteins from 10 cultivars of diploid perennial ryegrass (*L. perenne*) were processed to enrich for seed storage proteins and measured by LC-MS/MS in the search for gluten-like sequences.

Shotgun proteome measurement and database searching were used for the initial discovery of gluten-like protein sequences. The protein search database comprised the public sequences of the Poaceae family proteins from UniProt, a higher taxonomic group for wheat and ryegrass representing diverse genetic

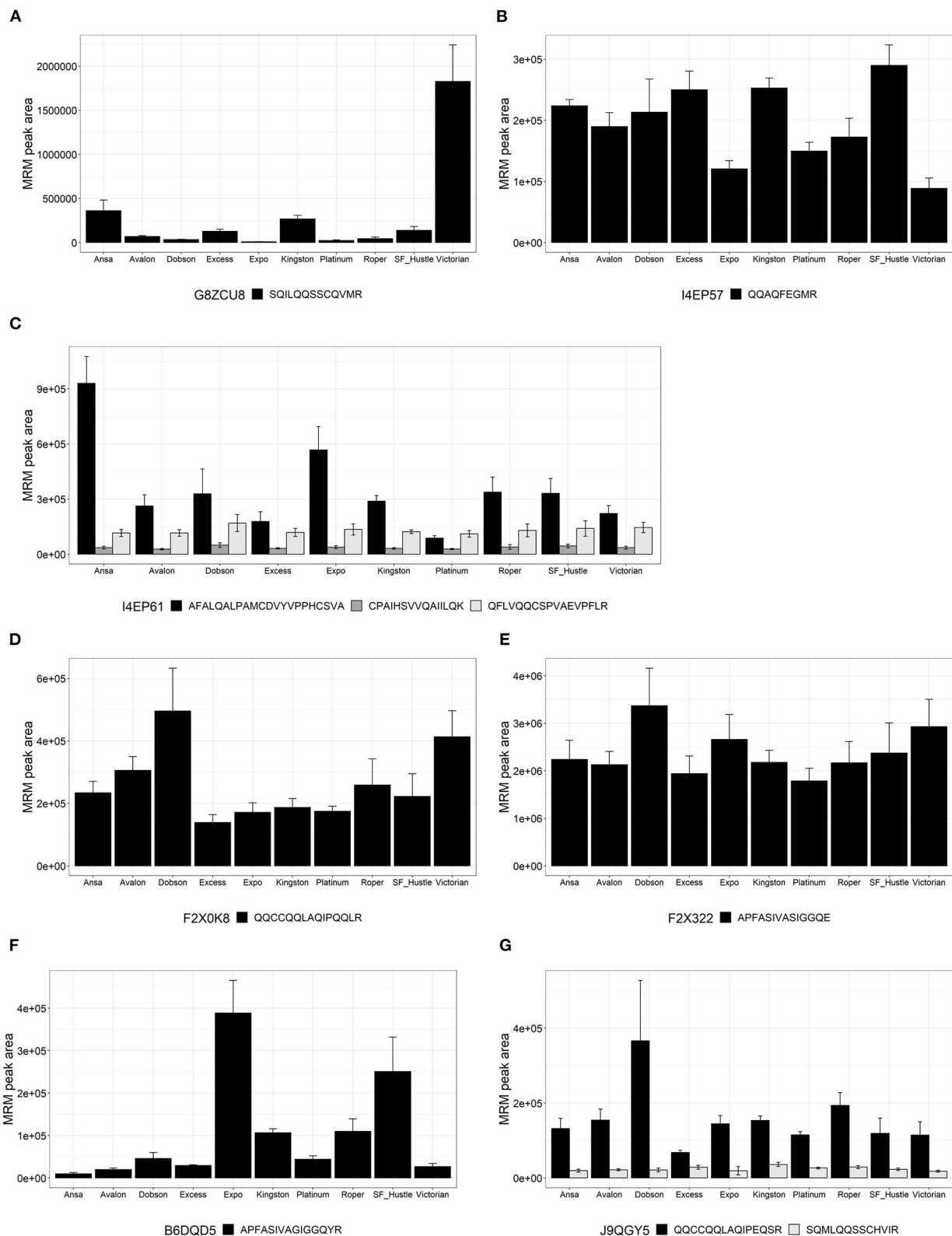
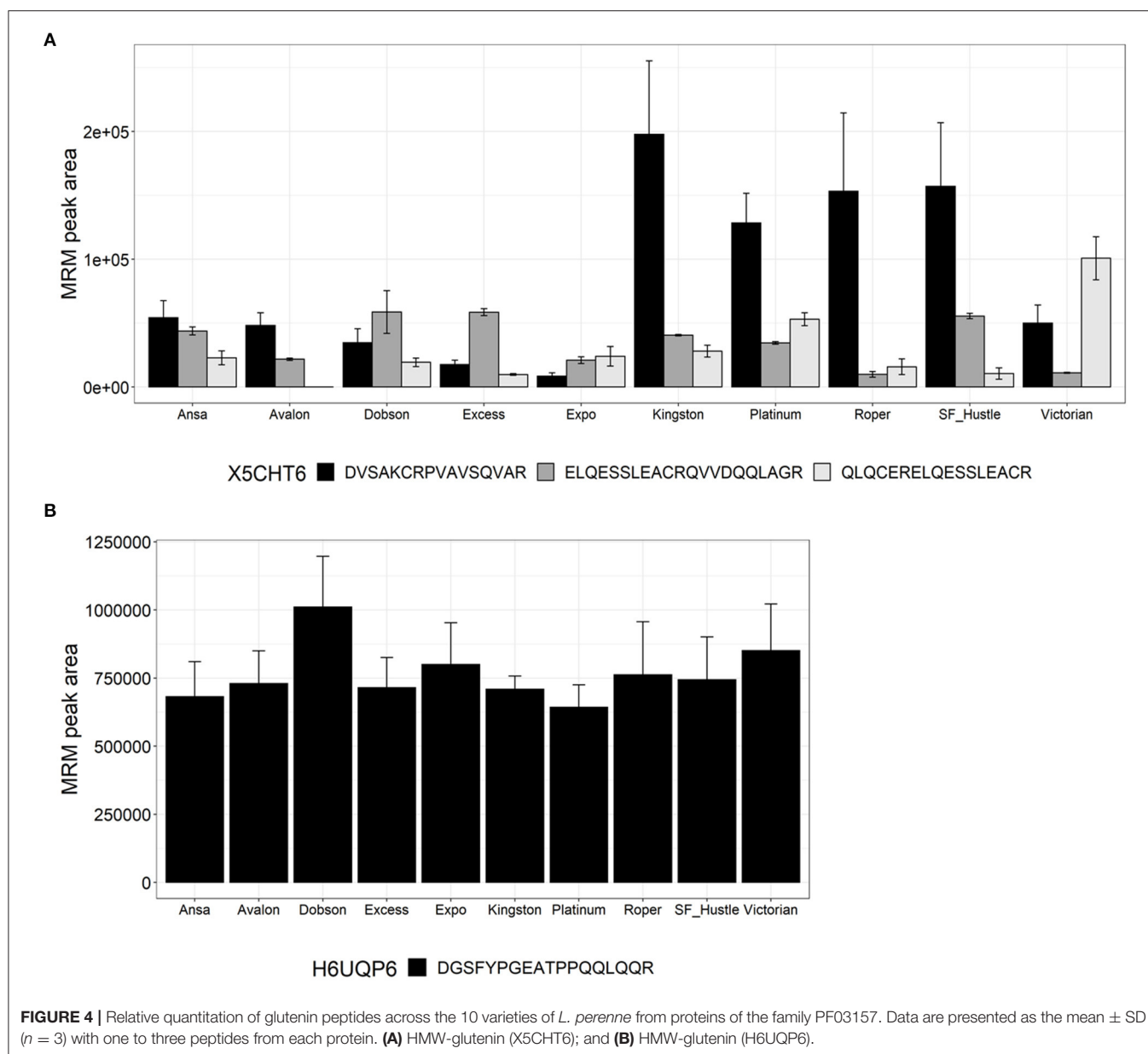


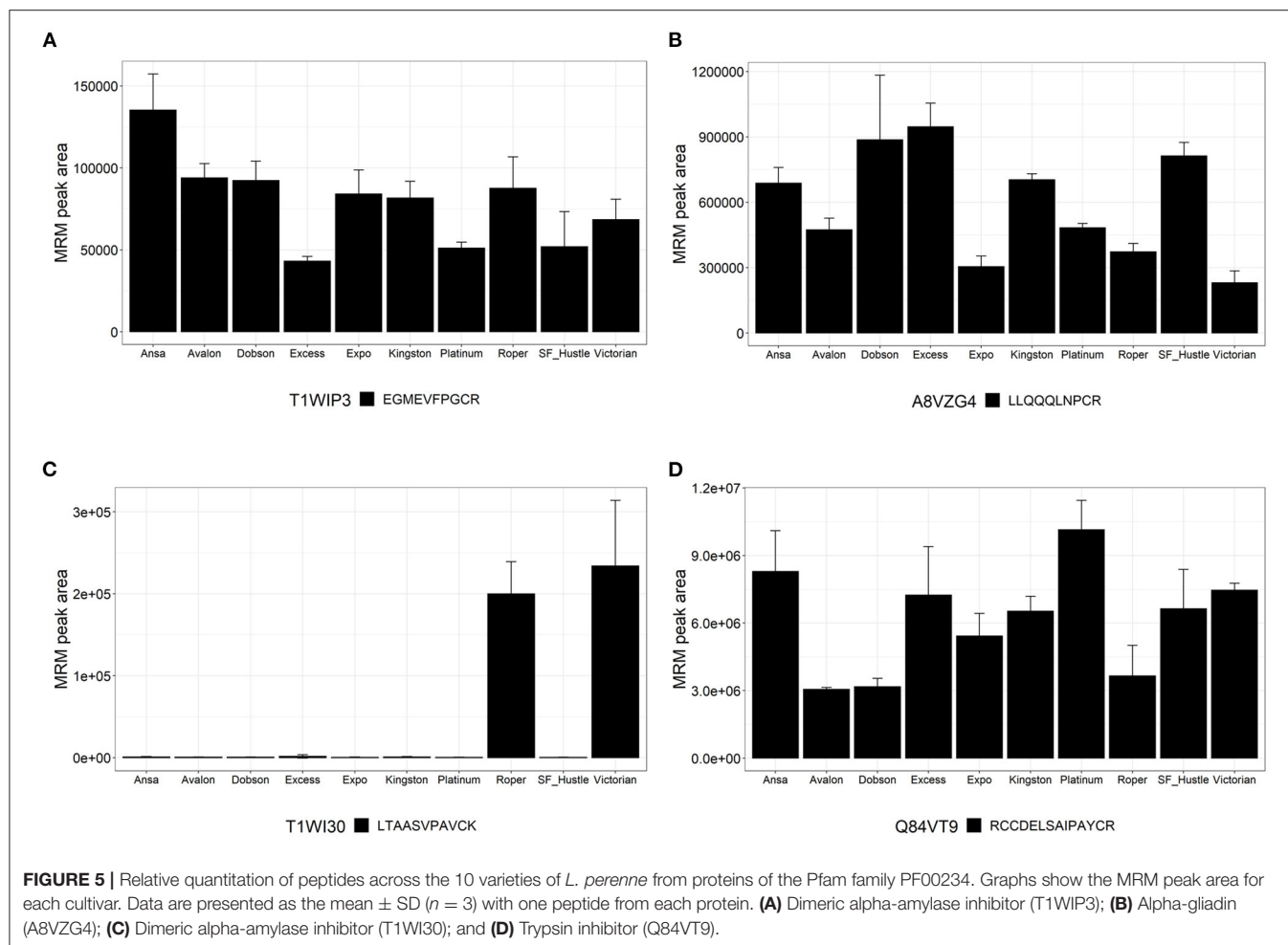
FIGURE 3 | Relative quantitation expressed as multiple reaction monitoring (MRM) peak area for peptides from proteins of the family PF13016 across 10 cultivars of *Lolium perenne*. Data are presented as the mean \pm SD ($n = 3$) with one to three peptides from each protein. **(A)** Avenin protein (G8ZCU8); **(B)** Avenin (I4EP57); **(C)** Avenin (I4EP61); **(D)** Gamma-gliadin (F2X0K8); **(E)** Gamma-gliadin (F2X322); **(F)** Gamma-gliadin (B6DQD5); and **(G)** LMW-glutenin (J9QGY5).



variations. Previous proteomics studies in *L. perenne* have used a different database approach where *Brachypodium distachyon* sequences were searched to successfully characterise drought response in the leaves from this grass (25); because of the focus on this compartment, no gluten-like proteins were identified in this study. No other studies besides Colgrave et.al. reported proteome measurement in ryegrass seeds (15).

The discovery strategy herein was complemented by a protein family domain search identifying three to eight gluten-like proteins in each ryegrass cultivar (**Supplementary Data 4**). This number is lower when compared with the frequency of gluten proteins identified when searching traditional cereal grain data (e.g., wheat, rye, barley) against the same Poaceae subset of the UniProt-KB database, which typically varies

between 5 and 47 gluten proteins (26, 27). However, the gluten-like protein sequences in ryegrass may differ from those in wheat and other Poaceae members. Therefore, further investigation is needed to precisely characterise the prolamins in ryegrass. Nevertheless, the strategy to determine gluten-like proteins through detecting the protein family domains for Gliadin, Glutenin_HMW, and Tryp_alpha_amyl revealed 19 protein accessions representing these families, including: 10 gliadin-domain containing proteins, 2 glutenin_HMW domain containing proteins, and 7 Tryp_alpha_amyl domain-containing proteins (**Figure 1**). Due to a paucity of data (protein sequences) in the *Lolium* subset of the UniProt database (748 nonredundant protein sequences), the identification of gluten-like peptides is likely not exhaustive and has led to the detection of proteins from



other species, i.e., orthologous proteins. The implementation of genomic and/or transcriptomic sequencing efforts would allow more ryegrass gluten-like proteins to be discovered.

A targeted MRM assay was developed for specifically measuring peptides found in proteins with homology to gliadins, glutenins, and ATIs. Prior to measuring these peptides, we expected that some targets would be detected at high levels in ryegrass and not in wheat, since target peptides were identified from ryegrass discovery proteomics. The results of this study revealed clear differences between the peptide content in ryegrass and wheat, with peptides predominant in ryegrass, regardless of belonging to the same family of Poaceae.

Gluten Relative Quantitation and Potential Markers

Gliadin and HMW Glutenin Family

In this study, nine peptides were identified from seven protein sequences, characteristic of prolamin proteins, and were measured by LC-MRM-MS. Six of these proteins belong to the Gliadin domain containing protein family and one to the Tryp_alpha_amyl family. Peptide search analysis revealed that five peptides (SQILQQSSCQVMR, CPAIHSVVQAILQK, QFLVQQCSPVAEVPFLR, AFALQALPAMCDVYVPPHCSVA,

and QQAQFEGMR) were primarily detected in avenin proteins (Table 1, Supplementary Table 2). In ryegrass, the peptides detected in all cultivars with consistent levels were QFLVQQCSPVAEVPFLR and CPAIHSVVQAILQK (Figure 3C). These peptides are not found in wheat, barley, or rye; therefore, these peptides are possible markers for differentiating ryegrass contamination from traditional gluten-containing grains. However, they are also common to *Avena sp.* and thus will not discriminate ryegrass from oats (Supplementary Table 2). Another peptide, LLQQQLNPCR (Figure 5B), from protein A8VZG4, a Tryp_alpha_amyl domain containing protein family member, was present in an α -gliadin of the species *D. hordeaceum* of the Triticeae tribe. However, this peptide may not be an ideal peptide marker due to its lack of uniform signal. Three peptides (QQCCQQLAQIPQQLR, APFASIVASIGGQE, and APFASIVAGIGGQYR, Figures 3D–F) were matched back to γ -gliadin sequences from species of the Triticeae tribe; interestingly, though these specific peptides were not detected in the wheat cultivars tested. These peptides showed variable abundance across the ryegrass cultivars, and although not detected in the wheat cultivars examined herein, these peptides will not make ideal markers for ryegrass contamination due to their presence in known wheat γ -gliadin sequences.

The nomenclature of prolamin super-family proteins is diverse depending on the cereal of origin. In wheat, these proteins are gliadins; in rye, they are secalins; in barley, hordeins, and oats, they are avenins. The phylogenetic relationship between prolamin proteins from different species has been demonstrated with homology comparisons between avenin sequences and α - and γ -gliadins from wheat, B-hordeins from barley, and γ -secalins from rye (28–30). Herein, we showed that ryegrass has gliadin-like peptides that share a certain level of similarity to avenins and γ -gliadins. Moreover, this agrees with a comparative genomic study that revealed conserved genetic maps in terms of orthology and collinearity in the lineage of the Triticeae (wheat), Aveneae (oat), and Poeae tribes (ryegrass) (31–33), although phylogeny studies place *Lolium* closer to *Avena* than to *Triticum* and *Hordeum* (34–37).

The other fraction of prolamins is constituted by glutenins, which are present as high molecular weight (HMW) and low molecular weight (LMW) subunits that join to make multimeric proteins held by disulphide bonds (1). Immune reactive epitopes have also been reported for HMW and LMW glutenins (38–43).

In this study, seven peptides were measured from three glutenin-like protein sequences. None of these peptides have been reported previously in ryegrass (15). Two peptides QQCCQQLAQIPEQSR and SQMLQQSSCHVIR (**Figure 3G**) from protein J9QGY5 (Gliadin-domain containing protein family) were found exclusively in LMW-GS of the species *D. villosum* of the Triticeae tribe. Both peptides were experimentally detected in ryegrass but not in wheat; the peptide SQMLQQSSCHVIR showed consistent levels across the ryegrass cultivars. Five peptides were characteristic of HMW-GS. These included four peptides (ELQESSLEACR, DVSAKCRPVAVSQVAR, ELQESSLEACRQVVDQQLAGR, and QLQCERELQESSLEACR, **Supplementary Figures 1C,D, Figure 4**) matching to protein X5CHT6. These were determined to be conserved within the Triticeae tribe and had variable abundance within the analysed ryegrass cultivars. The peptide ELQESSLEACR (**Supplementary Figures 1C,D**) was highly abundant in the wheat cultivars. Peptide DGSFYPPGEATPPQQLQQR was found consistently in ryegrass samples and was not detected in the wheat extracts; it was exclusively found in a protein sequence from *E. libanoticus*, which is a wild species of the Triticeae tribe.

Two peptides characteristic of LMW glutenin (QQCCQQLAQIPEQSR and SQMLQQSSCHVIR) and one peptide characteristic of HMW glutenin (DGSFYPPGEATPPQQLQQR) are found within Triticeae tribe; however, they were not present in wheat protein sequences. As such, these sequences may be candidate peptides to differentiate wheat and ryegrass. Nevertheless, this study reports candidate glutenin peptides for the first time in ryegrass cultivars.

Tryp_Alpha_Amyl Family

Alpha-amylase/trypsin inhibitor proteins are not considered as gluten proteins; however, recent research suggests that some GRDs are not triggered by gluten proteins but by other types of proteins with similar structures, including ATI proteins (44, 45). These ATI proteins are involved in plant defence; however,

they also activate the innate immune system and trigger pro-inflammatory cytokines (46–49). It is believed that ATIs can affect individuals with a sensitive type of GRD (46, 47).

ATI proteins were identified in this study, and diagnostic peptides were measured by LC-MS. The strategic approach consisted of searching peptides characteristic of proteins with the protein family domain PF00234, which lead to the initial recognition of several proteins and peptides; however, a deeper analysis using protein and peptide homology searches revealed that some of these peptides were found in other members of the PF00234-domain containing proteins, namely nsLTPs and puroindolines. Ultimately, four peptides characteristic of ATI proteins were measured from three target proteins. Two peptides (LQCVGSQVPEAVLR and EGMEVFPGCR, **Supplementary Figures 1A,B, Figure 5A**) from protein T1WIP3 were conserved for species of the Triticeae tribe, which coincides with the quantitative measurements showing a higher abundance in the wheat cultivars. Peptide LTAASVPAVCK (**Figure 5C**) from the protein T1WI30 had variable abundance across the ryegrass cultivars and was found in sequences from several species of the Triticeae tribe. Consequently, these peptides are unlikely candidates for markers of ryegrass presence. The last peptide, RCCDELSAIPAYCR (**Figure 5D**), from protein Q84VT9 showed high abundance in ryegrass and was not detected in wheat. This peptide was found exclusively in alpha-amylase/trypsin inhibitor proteins of species *Hordeum vulgare*; therefore, it is not present in wheat databases and is a candidate peptide to differentiate ryegrass from wheat but not barley.

The results described herein for the identification of ATI proteins provide evidence for peptide sequence similarity when comparing ryegrass to protein members of the Triticeae tribe ATIs, but further analysis, supplemented by genomic/transcriptomic sequencing efforts, would be needed to ascertain the extent of protein sequence homology. Principally, the ATIs targeted herein from ryegrass cultivars may not necessarily have the same immunoinflammatory properties as the wheat ATIs that are implicated in GRD. ATIs often differ in secondary structure, i.e., the number and position of intrachain disulphide bonds, and may contain different arrangements of α -helices that can influence the ability to activate the innate immune response (47).

CONCLUSION

Ryegrass is a common field contaminant with dense seeds and herbicide-resistant properties. These factors render ryegrass with the potential to enter the supply chain and be inadvertently consumed by the general population. This study provides evidence that gluten-like peptides are present in perennial ryegrass, thereby highlighting the potential risk of unintentional consumption of gluten through the cross-contamination of traditional cereal grains. Indeed, this food safety risk is underscored by a lack of knowledge around the potential for the industry standard ELISA assessment to produce inaccurate information. If these new gluten protein species trigger a Coeliac response but fail to produce an ELISA signal, then a food

safety issue may remain unchecked. Conversely, should ryegrass produce no Coeliac response but a strong ELISA signal, then foods may not meet safety requirements in error, thereby presenting an unnecessary challenge for food manufacturers. Future studies are warranted to explore the immunogenic potential of these new gluten-like proteins to determine their presence in nontraditional cereal grains, determine their natural variation, and deploy methods that can be used to differentiate ryegrass from wheat.

Through targeted measurement of prolamin super-family proteins, a suite of peptides were identified that showed consistent abundance across ryegrass cultivars but were not detected in wheat. These peptides could potentially be used in an assay for detecting ryegrass contamination in food products and differentiating ryegrass from wheat contamination in other cereal grains or processed foods. The methodology developed herein could also be applied to determine the extent of ryegrass presence in commodity grain or after processing into food ingredients. There are however no studies that have reported on whether the gluten-like proteins from ryegrass can trigger CD or other GRDs. Nevertheless, one study revealed cross-reactivity between ryegrass pollen and wheat endosperm proteins (50). Future investigation is required to measure the immune reactive potential of ryegrass and continue this important body of work that now spans over 90 years. Importantly, further studies are also required to enhance the genomic resources available for ryegrass so that species-specific proteins are readily identifiable rather than relying on sequence variation from related taxa to identify peptides and proteins. A combination of clinical studies supplemented by analytical workflows to understand the risk associated with agricultural co-mingling and dietary exposure are needed to ensure food safety and avoid the inadvertent failure of the GF-diet.

REFERENCES

- Wieser H. Chemistry of gluten proteins. *Food Microbiol.* (2007) 24:115–9. doi: 10.1016/j.fm.2006.07.004
- Singh P, Arora A, Strand TA, Leffler DA, Catassi C, Green PH, et al. Global prevalence of celiac disease: systematic review and meta-analysis. *Clin Gastroenterol Hepatol.* (2018) 16:823–36.e2. doi: 10.1007/978-3-319-91316-2
- King JA, Jeong J, Underwood FE, Quan J, Panaccione N, Windsor JW, et al. Incidence of celiac disease is increasing over time: a systematic review and meta-analysis. *Am J Gastroenterol.* (2020) 115:507–25. doi: 10.14309/ajg.0000000000000523
- Sapone A, Bai JC, Ciacci C, Dolinsek J, Green PHR, Hadjivassiliou M, et al. Spectrum of gluten-related disorders: consensus on new nomenclature and classification. *BMC Med.* (2012) 10:13. doi: 10.1186/1741-7015-10-13
- Gujral N, Freeman HJ, Thomson ABR. Celiac disease: prevalence, diagnosis, pathogenesis and treatment. *World J Gastroenterol.* (2012) 18:6036–59. doi: 10.3748/wjg.v18.i42.6036
- Walker MM, Murray JA. An update in the diagnosis of coeliac disease. *Histopathology.* (2011) 59:166–79. doi: 10.1111/j.1365-2559.2010.03680.x
- Silvester JA, Weiten D, Graff LA, Walker JR, Duerksen DR. Is it gluten-free? Relationship between self-reported gluten-free diet adherence and knowledge of gluten content of foods. *Nutrition.* (2016) 32:777–83. doi: 10.1016/j.nut.2016.01.021
- Cruz JE, Cocchio C, Lai PT, Hermes-DeSantis E. Gluten content of medications. *Am J Health Sys Phar.* (2015) 72:54–60. doi: 10.2146/ajhp140153
- Bascuñán KA, Vespa MC, Araya M. Celiac disease: understanding the gluten-free diet. *Eur J Nutr.* (2017) 56:449–59. doi: 10.1007/s00394-016-1238-5
- El Khoury D, Balfour-Ducharme S, Joye JJ. A review on the gluten-free diet: technological and nutritional challenges. *Nutrients.* (2018) 10:1410. doi: 10.3390/nu10101410
- Falcomer AL, Santos Araújo L, Farage P, Santos Monteiro J, Yoshio Nakano E, Puppini Zandonadi R. Gluten contamination in food services and industry: a systematic review. *Crit Rev Food Sci Nutr.* (2020) 60:479–93. doi: 10.1080/10408398.2018.1541864
- Koerner TB, Clérout C, Poirier C, Cantin I, Alimkulov A, Elamparo H. Gluten contamination in the Canadian commercial oat supply. *Food Addit Contam Part A Chem Anal Control Expo Risk Assess.* (2011) 28:705–10. doi: 10.1080/19440049.2011.579626
- Storsrud S, Malmheden Yman I, Lenner RA. Gluten contamination in oat products and products naturally free from gluten. *Eur Food Res Technol.* (2003) 217:481–5. doi: 10.1007/s00217-003-0786-0
- Niknam SR, Moerkerk M, Cousens R. Weed seed contamination in cereal and pulse crops. In: *Proceedings of the 13th Australian Weeds Conference.* (2002). p. 59–62.
- Colgrave ML, Goswami H, Byrne K, Blundell M, Howitt CA, Tanner GJ. Proteomic profiling of 16 cereal grains and the application of targeted proteomics to detect wheat contamination. *J Proteome Res.* (2015) 14:2659–68. doi: 10.1021/acs.jproteome.5b00187
- Bose U, Broadbent JA, Byrne K, Hasan S, Howitt CA, Colgrave ML. Optimisation of protein extraction for in-depth profiling of the cereal grain proteome. *J Prot.* (2019) 197:23–33. doi: 10.1016/j.jprot.2019.02.009

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession numbers can be found below: <https://doi.org/10.25919/8ehe-yr54>.

AUTHOR CONTRIBUTIONS

SE-C, JB, MC, and CH designed the study. SS and AA carried out sample preparation and extraction. SE-C and JB carried out acquisition, analysis, interpretation of the data and created the scripts for data curation, and visualisation. AJ provided guidance for building the database, data analysis, data visualisation, and data interpretation. SE-C, JB, AJ, CH, and MC drafted the manuscript. All authors contributed to the article and approved the submitted version.

ACKNOWLEDGMENTS

The authors thank Malcolm Blundell for milling the grains to be used for proteomic analysis. Sophia Escobar-Correas was supported by a Research Plus CERC Postdoctoral Fellowship from CSIRO, Australia. Alicja Andraszek was supported by the CSIRO Agriculture and Food summer vacation scholarship.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fnut.2021.708122/full#supplementary-material>

17. Shilov IV, Seymour SL, Patel AA, Loboda A, Tang WH, Keating SP, et al. The paragon algorithm, a next generation search engine that uses sequence temperature values and feature probabilities to identify peptides from tandem mass spectra. *Mol Cell Prot.* (2007) 6:1638–55. doi: 10.1074/mcp.T600050-MCP200
18. Escobar CS. R code to filter FDR reports from discovery proteomics generated in ProteinPilot Software. v2. CSIRO. *Softw Coll.* (2021) doi: 10.25919/e2aw-vd18
19. Stanke M, Waack S. Gene prediction with a hidden Markov model and a new intron submodel. *Bioinformatics.* (2003) 19(Suppl. 2):ii215–25. doi: 10.1093/bioinformatics/btg1080
20. MacLean B, Tomazela DM, Shulman N, Chambers M, Finney GL, Frewen B, et al. Skyline: an open source document editor for creating and analyzing targeted proteomics experiments. *Bioinformatics.* (2010) 26:966–8. doi: 10.1093/bioinformatics/btq054
21. Wickham H. An introduction to ggplot: an implementation of the grammar of graphics in R. *Statistics* (2006).
22. Thom ER, Waugh CD, McCabe RJ. Growth and persistence of perennial and hybrid ryegrasses when grazed by dairy cows in the central Waikato region of New Zealand. *New Zealand J Agricult Res.* (1998) 41:477–86. doi: 10.1080/00288233.1998.9513331
23. Jodidi SL. Isolation and purification of the alcohol-soluble protein (prolamin) occurring in English ryegrass (*Lolium perenne*). *J Agricult Res.* (1930) 40:361–70.
24. Shewry PR, Smith SJ, Lew EJ-L, Kasarda DD. Characterization of prolamins from meadow grasses: homology with those of wheat, barley and rye. *J Exp Botany.* (1986) 37:633–9. doi: 10.1093/jxb/37.5.633
25. Vanani FR, Shabani L, Sabzalian MR, Dehghanian F, Winner L. Comparative physiological and proteomic analysis indicates lower shock response to drought stress conditions in a self-pollinating perennial ryegrass. *PLoS ONE.* (2020) 15:e0234317. doi: 10.1371/journal.pone.0234317
26. Fallahbaghery A, Zou W, Byrne K, Howitt CA, Colgrave ML. Comparison of gluten extraction protocols assessed by LC-MS/MS analysis. *J Agric Food Chem.* (2017) 65:2857–66. doi: 10.1021/acs.jafc.7b00063
27. Lexhaller B, Colgrave ML, Scherf KA. Characterization and relative quantitation of wheat, rye, and barley gluten protein types by liquid chromatography–tandem mass spectrometry. *Front Plant Sci.* (2019) 10:1530. doi: 10.3389/fpls.2019.01530
28. Juhász A, Haraszi R, Maulis C. ProPepper: a curated database for identification and analysis of peptide and immune-responsive epitope composition of cereal grain protein families. *Database.* (2015) bav100. doi: 10.1093/database/bav100
29. Juhász A, Belova T, Florides CG, Maulis C, Fischer I, Gell G, et al. Genome mapping of seed-borne allergens and immunoresponsive proteins in wheat. *Sci Adv.* (2018) 4:eaar8602. doi: 10.1126/sciadv.aar8602
30. Tanner G, Juhász A, Florides CG, Nye-Wood M, Békés F, Colgrave ML, et al. Preparation and characterization of avenin-enriched oat protein by chill precipitation for feeding trials in Celiac disease. *Front Nutr.* (2019) 6:162. doi: 10.3389/fnut.2019.00162
31. Pfeifer M, Martis M, Asp T, Mayer KFX, Lübberstedt T, Byrne S, et al. The perennial ryegrass GenomeZipper: targeted use of genome resources for comparative grass genomics. *Plant Physiol.* (2013) 161:571–82. doi: 10.1104/pp.112.207282
32. Jones ES, Mahoney NL, Hayward MD, Armstead IP, Jones JG, Humphreys MO, et al. An enhanced molecular marker based genetic map of perennial ryegrass (*Lolium perenne*) reveals comparative relationships with other Poaceae genomes. *Genome.* (2002) 45:282–95. doi: 10.1139/g01-144
33. Sim S, Chang T, Curley J, Warnke SE, Barker RE, Jung G. Chromosomal rearrangements differentiating the ryegrass genome from the Triticeae, oat, and rice genomes using common heterologous RFLP probes. *Theor Appl Genet.* (2005) 110:1011–9. doi: 10.1007/s00122-004-1916-1
34. Soreng RJ, Peterson PM, Romaschenko K, Davidse G, Zuloaga FO, Judziewicz EJ, et al. A worldwide phylogenetic classification of the Poaceae (Gramineae). *J Syst Evol.* (2015) 53:117–37. doi: 10.1111/jse.12150
35. Kellogg EA. Relationships of cereal crops and other grasses. *PNAS.* (1998) 95:2005–10. doi: 10.1073/pnas.95.5.2005
36. Bouchenak-Khelladi Y, Salamin N, Savolainen V, Forest F, Bank M van der, Chase MW, et al. Large multi-gene phylogenetic trees of the grasses (Poaceae): progress towards complete tribal and generic level sampling. *Mol Phylog Evol.* (2008) 47:488–505. doi: 10.1016/j.jmpev.2008.01.035
37. Saarela JM, Burke SV, Wysocki WP, Barrett MD, Clark LG, Craine JM, et al. A 250 plastome phylogeny of the grass family (Poaceae): topological support under different data partitions. *PeerJ.* (2018) 6:e4299. doi: 10.7717/peerj.4299
38. Vader W, Kooy Y, Van Veelen P, De Ru A, Harris D, Benckhuijsen W, et al. The gluten response in children with celiac disease is directed toward multiple gliadin and glutenin peptides. *Gastroenterology.* (2002) 122:1729–37. doi: 10.1053/gast.2002.33606
39. Stepniak D, Vader LW, Kooy Y, van Veelen PA, Moustakas A, Papandreou NA, et al. T-cell recognition of HLA-DQ2-bound gluten peptides can be influenced by an N-terminal proline at p-1. *Immunogenetics.* (2005) 57:8–15. doi: 10.1007/s00251-005-0780-8
40. Bodd M, Kim C, Lundin KEA, Sollid LM. T-Cell response to gluten in patients with HLA-DQ2.2 reveals requirement of peptide-MHC stability in Celiac disease. *Gastroenterology.* (2012) 142:552–61. doi: 10.1053/j.gastro.2011.11.021
41. Van de wal Y, Kooy YMC, Veelen P van, Vader W, August SA, Drijfhout JW, et al. Glutenin is involved in the gluten-driven mucosal T cell response. *Eur J Immunol.* (1999) 29:3133–9. doi: 10.1002/(SICI)1521-4141(199910)29:10<3133::AID-IMMU3133>3.0.CO;2-G
42. Kooy-Winkelaar Y, van Lummel M, Moustakas AK, Schweizer J, Mearin ML, Mulder CJ, et al. Gluten-specific T cells cross-react between HLA-DQ8 and the HLA-DQ2α/DQ8β transdimer. *J Immunol.* (2011) 187:5123–9. doi: 10.4049/jimmunol.1101179
43. Sollid LM, Qiao S-W, Anderson RP, Gianfrani C, Koning F. Nomenclature and listing of Celiac disease relevant gluten T-cell epitopes restricted by HLA-DQ molecules. *Immunogenetics.* (2012) 64:455–60. doi: 10.1007/s00251-012-0599-z
44. Salcedo G, Quirce S, Diaz-perales A. Wheat allergens associated with Baker's asthma. *J Invest Allergol Clin Immunol.* (2011) 21:81.
45. Verbeke K. Nonceliac gluten sensitivity: what is the culprit? *Gastroenterology.* (2018) 154:471–3. doi: 10.1053/j.gastro.2018.01.013
46. Junker Y, Zeissig S, Kim S-J, Barisani D, Wieser H, Leffler DA, et al. Wheat amylase trypsin inhibitors drive intestinal inflammation via activation of toll-like receptor 4. *J Exp Med.* (2012) 209:2395–408. doi: 10.1084/jem.20102660
47. Zevallos VF, Raker V, Tenzer S, Jimenez-Calvente C, Ashfaq-Khan M, Rüssel N, et al. Nutritional wheat amylase-trypsin inhibitors promote intestinal inflammation via activation of myeloid cells. *Gastroenterology.* (2017) 152:1100–13. doi: 10.1053/j.gastro.2016.12.006
48. Bellinghausen I, Weigmann B, Zevallos V, Maxeiner J, Reißig S, Waisman A, et al. Wheat amylase-trypsin inhibitors exacerbate intestinal and airway allergic immune responses in humanized mice. *J Allerg Clin Immunol.* (2019) 143:201–12. doi: 10.1016/j.jaci.2018.02.041
49. Ziegler K, Neumann J, Liu F, Fröhlich-Nowoisky J, Cremer C, Saloga J, et al. Nitration of wheat amylase trypsin inhibitors increases their innate and adaptive immunostimulatory potential *in vitro*. *Front Immunol.* (2019) 9:3174. doi: 10.3389/fimmu.2018.03174
50. Donovan GR, Baldo BA. Crossreactivity of IgE antibodies from sera of subjects allergic to both ryegrass pollen and wheat endosperm proteins: evidence for common allergenic determinants. *Clin Exp Allergy.* (1990) 20:501–9. doi: 10.1111/j.1365-2222.1990.tb03142.x

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2021 Escobar-Correas, Broadbent, Andraszek, Stockwell, Howitt, Juhász and Colgrave. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.