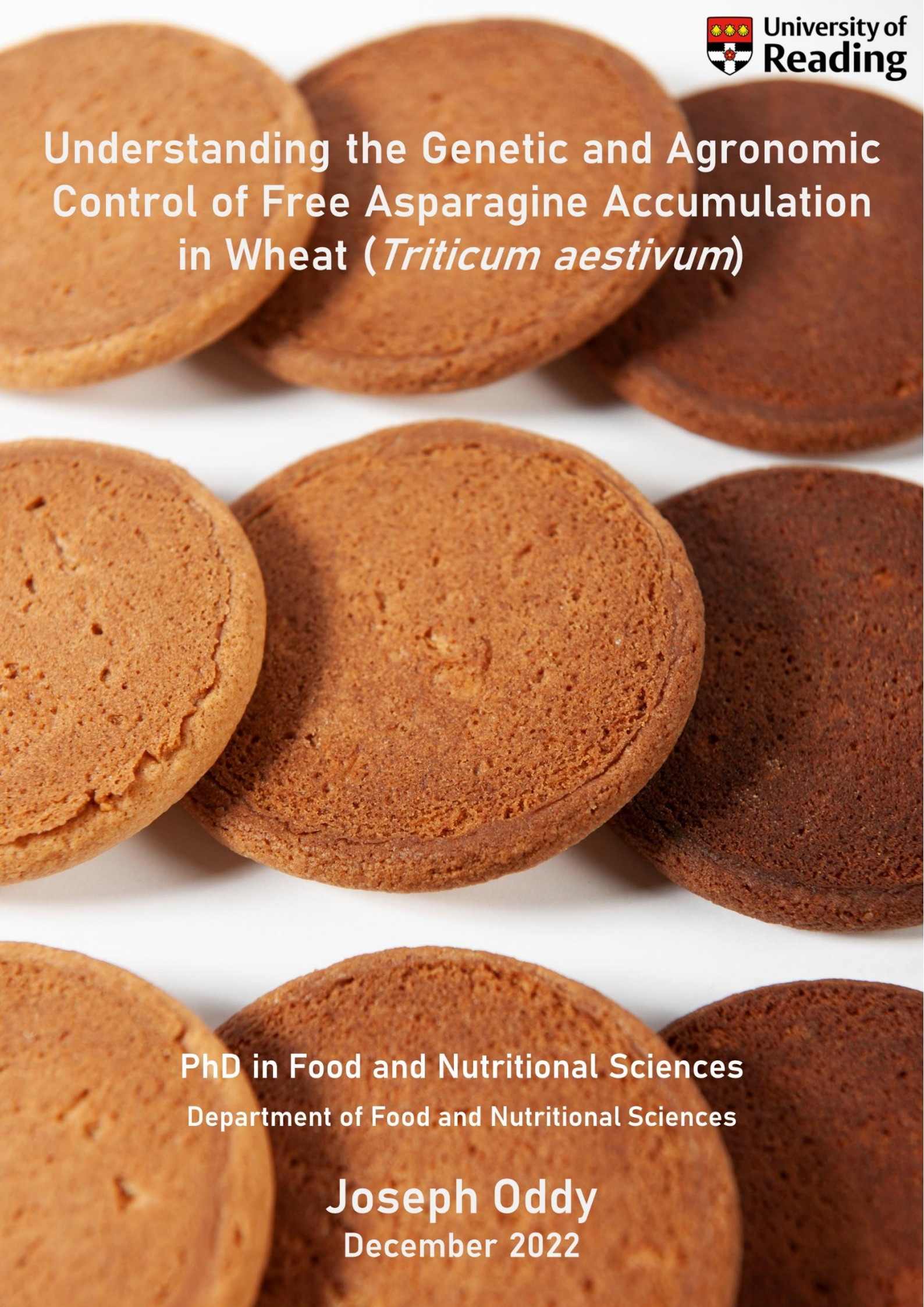


University of Reading

Understanding the Genetic and Agronomic Control of Free Asparagine Accumulation in Wheat (*Triticum aestivum*)

PhD in Food and Nutritional Sciences

Department of Food and Nutritional Sciences

Joseph Oddy

December 2022

Abstract

Acrylamide is a processing contaminant found in many different foods. It is formed from free (soluble, non-protein) asparagine and reducing sugars during the Maillard reaction, which occurs during low-moisture, high-temperature processing. In wheat-based food products, free asparagine is the key determinant of how much acrylamide forms, so the amount of acrylamide in wheat-based food products can be decreased by minimising the amount of free asparagine that accumulates in wheat grain. Agronomic and genetic strategies have previously been explored to reduce free asparagine accumulation in wheat grain, but a small proportion of products still continue to exceed recommendations set by the European Union, likely due to fluctuations in free asparagine content. Consequently, there is a need to develop more strategies to limit free asparagine accumulation in wheat grain.

In this thesis, new findings relevant to the genetic and agronomic control of grain asparagine accumulation are discussed. Firstly, a presence/absence variation of one of the *asparagine synthetase 2* homoeologs was analysed and found to reduce grain free asparagine content in the field. Secondly, variation in grain asparagine content and other traits was explored in a soft wheat mapping population, enabling identification of a QTL controlling free asparagine. A soft wheat population was used because this project sought to reduce the acrylamide content of soft wheat products, most notably biscuits. Finally, the impact of different fertilisers on grain asparagine accumulation was investigated in the field, facilitating further analysis of how this translates to differences in biscuit acrylamide content and the potential for imaging of plants and seeds to predict grain asparagine content. This showed that a nitrogen to sulphur application ratio of 10:1 (kg/ha) was sufficient to prevent large increases in free asparagine and that imaging could accurately predict free asparagine content. Overall, this body of work highlights interesting effects of genetic and environmental factors on grain asparagine content that can be used to inform future strategies for grain asparagine reduction.

Declaration of Original Authorship

I confirm that this is my own work and the use of all material from other sources has been properly and fully acknowledged.

Joseph Oddy

Personal Acknowledgements

Firstly, I would like to thank my supervisors. I had many great experiences and opportunities throughout my PhD and I found it all very enjoyable! A lot of this is due to Nigel Halford's excellent supervision and mentorship, and I am grateful to have had him as my supervisor. I would also like to thank my second supervisor at Rothamsted, Mark Wilkinson, for always making the time to ensure that I feel supported and for his guidance in the lab. Many thanks also to my industrial supervisors from Mondelēz, Isabel Moreira de Almeida and Nicholas Cryer, for giving me the opportunity to see what the applications of my work could be and for their help and advice throughout. Finally, I would like to thank my academic supervisor at Reading, Stephen Elmore, for his continued support and input.

Secondly, I would like to thank all the people at Rothamsted who have helped me along the way. Thanks to Sarah Raffan for her advice and for providing a good example to follow on how to do a PhD at Rothamsted! Many thanks to Chris Hall, who was like another supervisor to me in the Jenkinson, for teaching me how to deal with wheat, weevils, and everything else. Thanks to the Rothamsted Farm team and Chris Mackay in particular for their work managing our trials, without which the project could not have occurred. I would also like to thank the Rothamsted Statistics team, particularly John Addy, Andrew Mead, and Suzanne Clark, for providing me with great advice and training. I would also like to thank Carole Elleman, Daniel Clark, Jason Beasley, and Daniela Bedoya Orozco for their help during my placement at Reading Scientific Service Ltd. and Donna Fellowes for her support throughout the studentship. Many thanks to Graham Shephard for taking the photo of the biscuits I made for the front cover of the thesis.

Finally, I would like to thank my parents, my wife, and everyone else in my family for their support throughout.

Formal Acknowledgements and Contribution to Papers

I was supported by a Biotechnology and Biological Sciences Research Council (BBSRC) Collaborative Training Partnership Studentship whilst undertaking my PhD at Rothamsted Research in partnership with University of Reading and Mondelēz UK R&D Ltd. I received additional support in the form of a scholarship from the Society of Chemical Industry.

Chapter 1: All text by Joseph Oddy.

Chapter 2: Text, tables, and figures by Joseph Oddy with edits from Sarah Raffan, Mark Wilkinson, Stephen Elmore, and Nigel Halford.

Chapter 3: Text, tables, and figures by Joseph Oddy with edits from Sarah Raffan, Mark Wilkinson, Stephen Elmore, and Nigel Halford.

Chapter 4: This work was a collaborative project that combined one of the sub-projects from this PhD with work carried out by a team at Colorado State University (Rocío Alarcón-Reverte, Karl Ravet, and Stephen Pearce). Text, tables, and figures by Joseph Oddy in Results sections "Characterisation of the *TaASN-B2* deletion", "Wider screening for the presence/absence of ASN-B2", and "Contribution of *TaASN-B2* to free asparagine concentration in the grain" with edits from all co-authors. This relates to Figures 1 – 4 and 7 – 8 in the main text. The analyses presented in sections "Natural diversity in the asparagine synthetase gene family in wheat" and "Expression profiles of wheat *ASN* genes during development" (as shown in Table 1 and Figures 5 - 6) were carried out by the team at Colorado State University with the exception of the RT-qPCR analysis presented in Figure 7, which was performed by Joseph Oddy. Text in Introduction and Discussion sections was written by Joseph Oddy, Nigel Halford, and Stephen Pearce, with edits from all co-authors.

Work discussed in methodology sections "Genomic analyses" and "PCR assays to detect *TaASN-B2*" was performed by both Joseph Oddy and the team in Colorado. Work described in methodology sections "Plant materials and germination", "DNA extraction", "RT-

qPCR", and "Effect of *TaASN-B2* on free asparagine in the grain" was performed by Joseph Oddy under statistical guidance from Andrew Mead and Andrea Minter. Work described in section "RNA-seq data analysis" was performed by the team in Colorado. Supplementary Figures 1 and 2 and Supplementary Tables 1, 2b, and 4 were made by the Colorado team whilst remaining supplementary materials were made by Joseph Oddy.

Chapter 5: Text, tables, and figures by Joseph Oddy with edits from all co-authors. The mapping population was developed and managed in field trials by Monika Chhetry and Rajani Awal at the John Innes Centre. Genotyping of the population was subsequently performed at Limagrain (UK) and amino acid analysis was performed at Curtis Analytics (UK). All other practical and computational analyses were performed by Joseph Oddy, with assistance from Eve Murray and Rebecca Testa with the collection of grain morphology measurements and advice from John Addy and Luzie Wingen on statistical analysis. Dan Smith and Rob King also provided guidance on bioinformatic analysis of candidate genes.

I would like to thank Steve Hanley, Steve Thomas, Andy Phillips, Friedrich Longin, Matthias Rapp, and Keywan Hassani-Pak for their discussions with me concerning this study.

Chapter 6: Text, tables, and figures by Joseph Oddy with edits from all co-authors. The field trials were drilled and maintained by the Rothamsted Farm team and harvested grain was processed with assistance from the Sample Processing Facility team at Rothamsted. Amino acid analysis was performed by Curtis Analytics (UK) and statistical design and analysis of trials was performed in consultation with the Rothamsted Statistics group. Acrylamide analysis was performed by Daniela Bedoya Orozco at Reading Scientific Services Ltd (UK). All other practical and computational analysis was performed by Joseph Oddy.

I would also like to thank March Castle and Andrew Riche for guidance on collecting multispectral measurements from the field; Carole Elleman, Daniel Clark, Ananai Garcia, Jason Beasley, and Daniela Bedoya Orozco for their assistance at Reading Scientific Services Ltd; the Rothamsted farm team for managing the field trials; Graham Shephard for photography; and Gavin McDiarmid for drone photography.

Chapter 7: All text by Joseph Oddy.

Table of Contents

Abbreviations

AHDB: Agriculture and Horticulture Development Board

ANOVA: Analysis of Variance

BBSRC: Biotechnology and Biological Sciences Research Council

BLAST: Basic Local Alignment Search Tool

CIM: Composite Interval Mapping

CRISPR: Clustered Regularly Interspersed Short Palindromic Repeats

CTAB: Cetyl Trimethyl Ammonium Bromide

DPA: Days Post Anthesis

EFSA: European Food Safety Administration

EMS: Ethyl Methane Sulphonate

EU: European Union

FAO: Food and Agriculture Organisation of the United Nations

FDA: Food and Drug Administration

FSANZ: Food Standards Agency of Australia and New Zealand

G × E: Genotype-by-Environment Interaction

GA: Gibberellin

GATase: Glutamine amidotransferase

GM: Genetic Modification/Genetically Modified

GWAS: Genome-wide Association Study

GY: Grain Yield

Hd: Heading Date

HFN: Hagberg Falling Number

HLW: Hectolitre Weight

HMW: High Molecular Weight

HPLC: High-Performance Liquid Chromatography

IWGSC: International Wheat Genome Sequencing Consortium

K: Potassium

kb: kilo base pair

KASP: Kompetitive Allele Specific PCR

KHI: Kernel Hardness Index

KW: Kernel Weight

LMW: Low Molecular Weight

LSD: Least Significant Difference

LTR: Long Terminal Repeat

Mbp: Mega base pair

N: Nitrogen

NDVI: Normalised Difference Vegetation Index

NIRS: Near Infrared Spectroscopy

NUE: Nitrogen Use Efficiency

P: Phosphorus

PAV: Presence/Absence Variation

PCA: Principal Component Analysis

PCR: Polymerase Chain Reaction

PH: Plant Height

PHS: Pre-Harvest Sprouting

PLSR: Partial Least Squares Regression

PSRI: Plant Senescence Reflectance Index

QTL: Quantitative Trait Locus/Loci

REML: Restricted Maximum Likelihood

RNA-seq: RNA sequencing

RT-qPCR: Reverse Transcription quantitative PCR

S: Sulphur

SIM: Simple Interval Mapping

TKW: Thousand Kernel Weight

TPM: Transcripts Per Million

USDA: United States Department of Agriculture

YDT: Yield Gap-Based Drought Tolerance

Z-SDS: Zeleny Sedimentation Index

1

Introduction

1.1.Project background

The aims of this project were initially developed by Professor Nigel Halford (Rothamsted Research), Dr. Isabel Moreira de Almeida (Mondelēz International), Dr. Nicholas Cryer (Mondelēz International), and Dr. Stephen Elmore (University of Reading), as part of the development of a BBSRC Collaborative Training Partnership studentship (UK Research and Innovation, 2019). Mondelēz is a global snack food company, selling a range of products that are impacted by acrylamide regulations, such as biscuits. Professor Nigel Halford and Dr. Stephen Elmore have both worked extensively on the subject of acrylamide in food, so this partnership was set up in order to develop practical strategies to reduce the acrylamide-forming potential of wheat.

1.2.The acrylamide issue

Acrylamide is an organic compound formed during the Maillard reaction from free (soluble, non-protein) asparagine and reducing sugars (e.g. glucose, fructose) (Figure 1.1), a reaction that is also responsible for the formation of many desirable colour and flavour compounds (Mottram et al., 2002; Stadler et al., 2002). It is present in many different foods; but the largest food groups contributing to dietary acrylamide intake in the European Union are estimated to be potato and wheat-based food products, and coffee (CONTAM Panel, 2015). Dietary acrylamide consumption is very variable, depending on the types of food consumed, the processes those foods have been through, and the raw materials used to manufacture those foods.

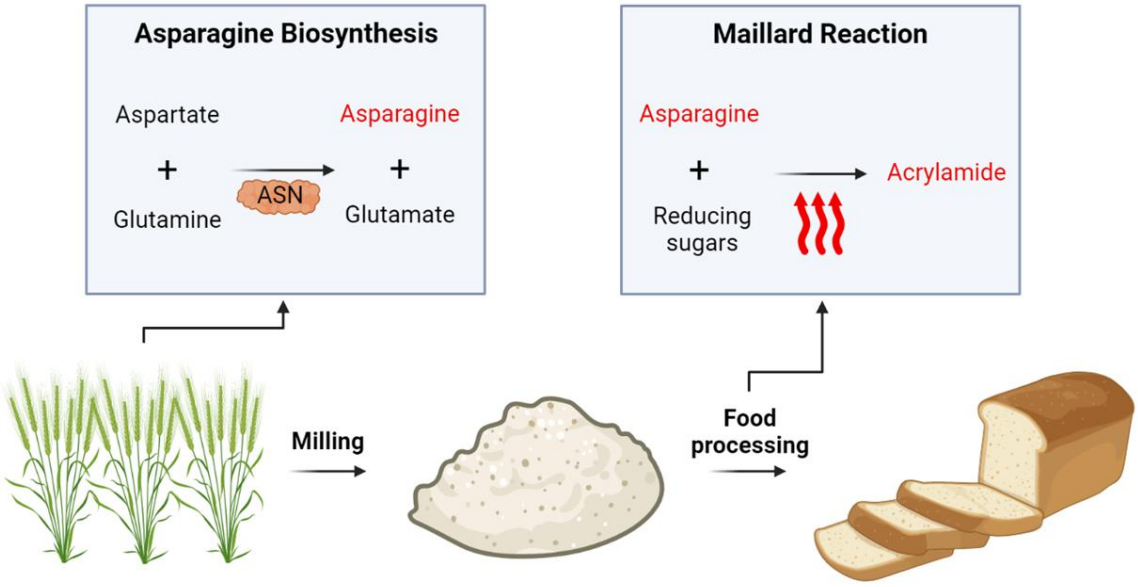


Figure 1.1. Simplified representation of asparagine biosynthesis in wheat (catalysed by asparagine synthetase) and the formation of acrylamide in the maillard reaction during the manufacture of wheat-based food products (catalysed by heat). ASN (asparagine synthetase).

The main issue concerning dietary consumption of acrylamide is its carcinogenicity, as its other negative health effects (neurological, reproductive, and developmental) are not predicted to be an issue at the concentrations consumed in the diet (CONTAM Panel, 2015). Acrylamide is known to cause cancers in rodents when given at high doses (National Toxicology Program, 2012) and mechanistically there is a feasible route through which it can cause cancer in humans. In humans, acrylamide is metabolised to glycidamide (a genotoxic compound) so it is possible that acrylamide could cause cancer in humans through this mechanism (Koszucka et al., 2020). Zhivagui et al. (2019) showed that a mutational signature from glycidamide was present in various human cancers, suggesting that this could be contributing to oncogenesis (although the study did not show that these mutations were oncogenic). On the other hand, epidemiological studies of dietary acrylamide intake have not found consistent associations between acrylamide and cancer in either food frequency questionnaire studies or biomarker studies (the latter being the more reliable methodology of the two) (Lipworth et al., 2012; Virk-Baker et al., 2014). Consequently, it has not been possible to conclude with certainty that acrylamide causes cancer in humans at the concentrations found in food. Nor has it been possible to conclude that all concentrations of acrylamide in food are

safe, due to the evidence from mechanistic and animal studies suggesting that it could cause cancer.

As a result, some regulatory bodies have taken a precautionary approach and have established regulations for acrylamide in food since negative health effects from dietary acrylamide consumption cannot be excluded (see Raffan & Halford (2019) for review). Of all regulatory bodies, the European Commission has the strongest regulations, with recommended "Benchmark Levels" indicating concentrations of acrylamide in food that they think industry can achieve (European Commission, 2017). However, since toxicologically-safe limits for dietary acrylamide consumption have not been determined, these levels are instead based on what is reasonably achievable in different foods. Consequently, the Benchmark Level for wheat-based bread is relatively low (50 micrograms per kilogram) when compared to the Benchmark Level for potato crisps (750 micrograms per kilogram). In most regulatory regions, however, no such Benchmark or equivalent levels for acrylamide exist.

Despite the lack of legal limits across all regulatory regions for acrylamide in food, some manufacturers have made substantial progress in reducing the acrylamide content of their products, using the many changes to product processing and formulation that are available (Powers et al., 2021). For example, some products can have asparaginase added during processing, which vastly reduces the amount of free asparagine in the product, although at an additional cost (Xu et al., 2016). Replacing ammonium bicarbonate and high-fructose corn syrup with alternatives has also been shown to be effective at reducing acrylamide concentrations in food (Amrein et al., 2006). Without changing product formulation, quality control checks based on colour can also be effective at removing high-acrylamide products, due to the association between acrylamide and colour (Gökmen et al., 2006; Gökmen et al., 2008).

Another possibility is to use raw materials that do not favour the formation of acrylamide. For wheat-based food products, free asparagine is the precursor that determines the concentration of acrylamide in the final food product (reviewed in Raffan & Halford, 2019). Consequently, the use of flour that is low in free asparagine would help to keep food products low in acrylamide. Asparagine is an amino acid that is present naturally in plants, formed from the ATP dependent reaction of aspartate with glutamine through the action of the enzyme asparagine synthetase (ASN) (Figure 1.1) (Lea et al., 2007). As an amino acid, asparagine is an important component of proteins, but the role of free (soluble, non-protein) asparagine in

wheat is less clear. Numerous stressors have been shown to cause large increases in the free asparagine content of wheat grain (e.g. sulphur deficiency (Curtis et al., 2014; Curtis et al., 2018) and disease (Curtis et al., 2016)), but it is unclear why this occurs. It is known that the ASNs have an important role in asparagine accumulation, as knocking out some of these genes (the ASN2s, gene accessions provided in Table 1.1) has proved to be an effective method of reducing grain free asparagine content (Raffan et al., 2021; Raffan et al., 2023).

Table 1.1. Ensembl accessions for the asparagine synthetase (*ASN2*) homeologs in varieties Chinese Spring and Landmark (Howe et al., 2021).

Gene name	Chinese Spring accession	Landmark accession
TaASN-A2	TraesCS3A02G077100	TraesLDM3A03G01333300
TaASN-B2	Not present	TraesLDM3B03G01566640
TaASN-D2	TraesCS3D02G077300	TraesLDM3D03G01816420

Many previous research projects have been undertaken to develop strategies to reduce the free asparagine content of wheat grain, including research into wheat agronomy and genetics (covered in Chapters 2 and 3). These have been successful and led to practical recommendations (FoodDrinkEurope, 2019); however, industry continues to be troubled by food products that exceed recommended acrylamide concentrations. For example, approximately 15% of breakfast cereals in Spain sampled between 2006 and 2018 were found to exceed Benchmark Levels in one study (Mesías et al., 2019). As the European Union looks set to introduce Maximum Levels (legal limits) at current Benchmark Levels (European Commission, 2022), further strategies to help manufacturers keep their acrylamide concentrations low need to be developed to reduce the possibility of exceeding these limits.

1.3. Project objectives

This project sought to develop new strategies to reduce the acrylamide content of food products by finding new ways of keeping free asparagine concentrations low in wheat grain. As Mondelēz are a snack food company, this project was also designed with a particular focus on soft wheat, as this is the wheat group used in the baking of products such as biscuits, wafers,

and other snack foods (see Chapter 6 Introduction for a more detailed discussion of soft wheat in the UK).

At the beginning of my PhD, a literature review (Oddy et al., 2020) was written to summarise what was known about free asparagine accumulation in wheat grain (Chapter 2). This review explored the known functions of asparagine in wheat and other species, as well as strategies that existed at the time to control grain free asparagine content. Later on in the project, a second literature review (Oddy et al., 2022) was written focusing on the relationship between grain free asparagine content and other traits, with a greater focus on wheat in particular, and how breeding for low free asparagine wheat may be achieved (Chapter 3). Together, these literature reviews provide a comprehensive background to this project and what is known about free asparagine accumulation in wheat grain.

The first sub-project of this PhD (Oddy et al., 2021) was to characterise a deletion of the B genome homeolog of the asparagine synthetase 2 gene (*TaASN-B2*; *Ta* stands for *Triticum aestivum*, *ASN* stands for asparagine synthetase, B denotes the B genome homoeolog and the number (2) following the subgenome denotes the gene number) (Chapter 4). A number of asparagine synthetase genes had been shown to exist in the wheat genome prior to this project, but it was noticed that the Chinese Spring reference genome did not possess *TaASN-B2* (Table 1.1) whereas unpublished data suggested that other varieties did have a *TaASN-B2* gene (Xu et al., 2018). When I started this project, other colleagues in my group were using CRISPR/Cas9 technology to edit the *TaASN2* genes (without any subgenome identifier, *TaASN2* refers to all A, B, and D subgenome genes) to reduce grain free asparagine content (Raffan et al., 2021). As a result, a cDNA sequence for *TaASN-B2* in the wheat variety Cadenza had been identified. Beyond this data, nothing else was known about *TaASN-B2* and no analyses had been performed to locate the site and size of the deletion in Chinese Spring or to investigate differences between homeologs. Consequently, the pursuit of a detailed investigation into the variable presence and absence of this gene and its impact on grain free asparagine content was of significant interest.

The second sub-project of this PhD was to investigate a soft wheat mapping population formed from the soft wheat parental varieties, Claire and Robigus (Chapter 5). The genetic control of grain free asparagine content was not well known during the development of this project, with QTL identified for grain free asparagine content from only three previous studies (Emebiri, 2014; Peng et al., 2018; Rapp et al., 2018). In order to breed a new commercial

variety with low grain free asparagine content, ideally QTL would be identified in parents commonly used in breeding programs, such as Claire and Robigus. This mapping population had already been developed between Rothamsted Research and the John Innes Centre before the planning stages of this project, so could be used as soon as the project began.

The third and final sub-project of this PhD was to investigate different agronomic treatments of soft wheat varieties to see how these impacted grain free asparagine content (Chapter 6). Although sulphur deficiency was known to cause increases in grain free asparagine content (Curtis et al., 2014; Curtis et al., 2018), no field experiments had been performed testing the importance of the ratio of nitrogen to sulphur application. Furthermore, no field trials had been performed to investigate the effects of phosphorus and potassium deficiencies on grain free asparagine content, although these are common components of fertilisers. Consequently, the treatments used in the field trials were a combination of different nitrogen, sulphur, potassium, and phosphorus application rates. Twelve different soft wheat varieties were used in these trials to ensure that my conclusions about the effect of treatments were not variety-specific (i.e. the conclusions drawn about the effect of treatment could be applicable to many soft wheat varieties). As part of this sub-project, a three-month industrial placement was planned to investigate acrylamide formation and other quality traits in biscuits baked from the field trial samples. This placement was originally intended to take place at the Mondelēz Biscuit R&D site in Saclay (France), but due to difficulties caused by the COVID-19 pandemic, this placement instead took place at Mondelēz's Reading Scientific Services Ltd.

1.4. References

Amrein, T. M., Andres, L., Manzardo, G. G., & Amadò, R. (2006). Investigations on the promoting effect of ammonium hydrogencarbonate on the formation of acrylamide in model systems. *Journal of Agricultural and Food Chemistry*, *54*(26), 10253-10261.

Curtis, T., Halford, N. G., Powers, S. J., McGrath, S. P. and Zazzeroni, R. (2014). Effect of sulphur fertilisation on the acrylamide-forming potential of wheat (HGCA Project Report No. 525). *Home Grown Cereals Authority (HGCA) Stoneleigh*. https://ahdb.org.uk/effect-of-sulphur-fertilisation-on-the-acrylamide-forming-potential-of-wheat

Curtis, T. Y., Powers, S. J., & Halford, N. G. (2016). Effects of fungicide treatment on free amino acid concentration and acrylamide-forming potential in wheat. *Journal of Agricultural and Food Chemistry*, *64*(51), 9689-9696.

Curtis, T. Y., Powers, S. J., Wang, R., & Halford, N. G. (2018). Effects of variety, year of cultivation and sulphur supply on the accumulation of free asparagine in the grain of commercial wheat varieties. *Food Chemistry*, *239*, 304-313.

EFSA Panel on Contaminants in the Food Chain (CONTAM). (2015). Scientific opinion on acrylamide in food. *EFSA Journal*, *13*(6), 4104.

Emebiri, L. C. (2014). Genetic variation and possible SNP markers for breeding wheat with low-grain asparagine, the major precursor for acrylamide formation in heat-processed products. *Journal of the Science of Food and Agriculture*, *94*(7), 1422-1429.

European Commission. (2017). Commission regulation (EU) 2017/2158 of 20 November 2017 establishing mitigation measures and Benchmark levels for the reduction of the presence of acrylamide in food. Brussels, Belgium: European Commission.

European Commission. (2022). Food Safety: Acrylamide. https://ec.europa.eu/food/safety/chemical-safety/contaminants/catalogue/acrylamide_en (accessed October 11, 2022).

FoodDrinkEurope. (2019). Acrylamide toolbox 2019. Brussels: Food Drink Europe; 2019.

Gökmen, V., & Şenyuva, H. Z. (2006). Study of colour and acrylamide formation in coffee, wheat flour and potato chips during heating. *Food Chemistry, 99*(2), 238-243.

Gökmen, V., Açar, Ö. Ç., Arribas-Lorenzo, G., & Morales, F. J. (2008). Investigating the correlation between acrylamide content and browning ratio of model cookies. *Journal of Food Engineering*, *87*(3), 380-385.

Howe, K.L., Achuthan, P., Allen, J., Allen, J., Alvarez-Jarreta, J., Amode, M.R., Armean, I.M., Azov, A.G., Bennett, R., & Bhai, J. (2021). Ensembl 2021. *Nucleic Acids Research*, *49*, D884–D891

Koszucka, A., Nowak, A., Nowak, I., & Motyl, I. (2020). Acrylamide in human diet, its metabolism, toxicity, inactivation and the associated European Union legal regulations in food industry. *Critical Reviews in Food Science and Nutrition*, *60* (10), 1677-1692.

Lea, P. J., Sodek, L., Parry, M. A., Shewry, P. R., & Halford, N. G. (2007). Asparagine in plants. *Annals of applied biology*, *150*(1), 1-26.

Lipworth, L., Sonderman, J. S., Tarone, R. E., & McLaughlin, J. K. (2012). Review of epidemiologic studies of dietary acrylamide intake and the risk of cancer. *European Journal of Cancer Prevention*, *21*(4), 375-386.

Mesías, M., Sáez-Escudero, L., Morales, F. J., & Delgado-Andrade, C. (2019). Reassessment of acrylamide content in breakfast cereals. Evolution of the Spanish market from 2006 to 2018. *Food Control*, *105*, 94-101.

Mottram, D. S., Wedzicha, B. L., & Dodson, A. T. (2002). Acrylamide is formed in the Maillard reaction. *Nature*, *419*(6906), 448-449.

National Toxicology Program. (2012). Toxicology and carcinogenesis studies of acrylamide (CASRN 79-06-1) in F344/N rats and B6C3F1 mice (feed and drinking water studies). *National Toxicology Program Technical Report Series*, (575), 1-234.

Oddy, J., Raffan, S., Wilkinson, M. D., Elmore, J. S., & Halford, N. G. (2020). Stress, nutrients and genotype: understanding and managing asparagine accumulation in wheat grain. *CABI Agriculture and Bioscience*, *1*(1), 1-14.

Oddy, J., Alarcón-Reverte, R., Wilkinson, M., Ravet, K., Raffan, S., Minter, A., ... & Pearce, S. (2021). Reduced free asparagine in wheat grain resulting from a natural deletion of TaASN-B2: investigating and exploiting diversity in the asparagine synthetase gene family to improve wheat quality. *BMC Plant Biology*, *21*(1), 1-17.

Oddy, J., Raffan, S., Wilkinson, M. D., Elmore, J. S., & Halford, N. G. (2022). Understanding the relationships between free asparagine in grain and other traits to breed low-asparagine wheat. *Plants*, *11*(5), 669.

Peng, Y., Liu, H., Chen, J., Shi, T., Zhang, C., Sun, D., He, Z., Hao, Y., & Chen, W. (2018). Genome-wide association studies of free amino acid levels by six multi-locus models in bread wheat. *Frontiers in Plant Science, 9*, 1196

Powers, S. J., Mottram, D. S., Curtis, A., & Halford, N. G. (2021). Progress on reducing acrylamide levels in potato crisps in Europe, 2002 to 2019. *Food Additives & Contaminants: Part A, 38*(5), 782-806.

Raffan, S., & Halford, N. G. (2019). Acrylamide in food: Progress in and prospects for genetic and agronomic solutions. *Annals of Applied Biology, 175*(3), 259-281.

Raffan, S., Sparks, C., Huttly, A., Hyde, L., Martignago, D., Mead, A., ... & Halford, N. G. (2021). Wheat with greatly reduced accumulation of free asparagine in the grain, produced by CRISPR/Cas9 editing of asparagine synthetase gene TaASN2. *Plant Biotechnology Journal, 19*(8), 1602-1613.

Raffan, S., Oddy, J., Mead, A., Barker, G., Curtis, T. Y., Usher, S. L., ... & Halford, N. G. (2023). Field assessment of genome edited, low asparagine wheat: Europe's first CRISPR wheat field trial. *Plant Biotechnology Journal*. Advance online publication. https://doi.org/10.1111/pbi.14026

Rapp, M., Schwadorf, K., Leiser, W. L., Würschum, T., & Longin, C. F. H. (2018). Assessing the variation and genetic architecture of asparagine content in wheat: What can plant breeding contribute to a reduction in the acrylamide precursor? *Theoretical and Applied Genetics, 131*(11), 2427-2437.

Stadler, R. H., Blank, I., Varga, N., Robert, F., Hau, J., Guy, P. A., ... & Riediker, S. (2002). Acrylamide from Maillard reaction products. *Nature, 419* (6906), 449-450.

UK Research and Innovation. (2019). Crop management strategies for low asparagine grains to limit acrylamide-forming potential. https://gtr.ukri.org/projects?ref=studentship-2288997#/tabOverview

Virk-Baker, M. K., Nagy, T. R., Barnes, S., & Groopman, J. (2014). Dietary acrylamide and human cancer: a systematic review of literature. *Nutrition and Cancer, 66*(5), 774-790.

Xu, F., Oruna-Concha, M. J., & Elmore, J. S. (2016). The use of asparaginase to reduce acrylamide levels in cooked food. *Food Chemistry*, *210*, 163-171.

Xu, H., Curtis, T. Y., Powers, S. J., Raffan, S., Gao, R., Huang, J., ... & Halford, N. G. (2018). Genomic, biochemical, and modeling analyses of asparagine synthetases from wheat. *Frontiers in Plant Science*, *8*, 2237

Zhivagui, M., Ng, A. W., Ardin, M., Churchwell, M. I., Pandey, M., Renard, C., ... & Zavadil, J. (2019). Experimental and pan-cancer genome analyses reveal widespread contribution of acrylamide exposure to carcinogenesis in humans. *Genome Research*, *29*(4), 521-531.

2

Stress, nutrients and genotype: understanding and managing asparagine accumulation in wheat grain

Joseph Oddy[1], Sarah Raffan[1], Mark D. Wilkinson[1], J. Stephen Elmore[2] and Nigel G. Halford[1]

1. *Plant Sciences Department, Rothamsted Research, Harpenden AL5 2JQ, UK.*
2. *Department of Food and Nutritional Sciences, University of Reading, Whiteknights, PO Box 226, Reading RG6 6AP, UK.*

CABI Agriculture and Bioscience, 1, 10 (2020).

This chapter has been modified from its published form to allow for edits and corrections within this Thesis.

2.1. Introduction to paper

At the outset of this PhD, a comprehensive review of free asparagine in terms of its functions and in particular its relationship with stress was lacking from the literature. The relationship between free asparagine accumulation and stress had been highlighted in a number of studies but had not been discussed in depth. This relationship is highly relevant to the formation of acrylamide, since it is under stressful conditions that the highest concentrations of free asparagine are often observed. Consequently, in this review I aimed to cover three main topics: (1) the function of asparagine and the asparagine synthetases in wheat and other species, (2) the functional role of asparagine accumulation during stress and the molecular basis of this response, and (3) what strategies are available to prevent the accumulation of free asparagine in wheat.

2.2. Background

The amino acid asparagine has long been of interest to plant biologists because of its role in nitrogen transport and stress responses (Lea et al., 2007). In the early 2000s, however, asparagine in crops gained new significance because of the discovery that free (soluble, non-protein) asparagine is a precursor for the food processing by-product and contaminant, acrylamide. Acrylamide forms from free asparagine and reducing sugars, mainly glucose, fructose, and maltose, during a non-enzymatic reaction called the Maillard reaction. This reaction occurs when food is heated above 120 °C under low moisture content in processes such as frying, roasting, baking, and toasting (Mottram et al., 2002; Stadler et al., 2002). In wheat products, free asparagine concentration determines the potential for acrylamide formation (reviewed in Raffan & Halford (2019)).

Dietary acrylamide intake is concerning because of its links to cancer (reviewed in Raffan & Halford (2019)) and authorities such as the European Commission have been prompted to introduce regulations on acrylamide concentrations in food (European Commission, 2017). Food manufacturers have adapted their processes and applied more effective quality control measures to reduce the concentrations of acrylamide in their products, but there is a limit to what can be achieved with that approach without affecting product quality. In order to make further improvements they need raw materials with consistently-low potential for acrylamide formation. Consequently, wheat growers need effective crop-management strategies to prevent excess free asparagine accumulation in the grain while not compromising desirable traits, such

as crop yield, protein content, disease resistance and stress tolerance. These strategies can only be developed through a greater understanding of the mechanisms and functions underlying asparagine accumulation.

2.3. Asparagine in normal development

As a proteinogenic amino acid, asparagine is a structural component of proteins, but its role in plant biology extends beyond this. Asparagine has an essential role in nitrogen storage and transport in many plant species. It acts as the main transport molecule of reduced nitrogen in the vasculature; likely because it is the amino acid with the highest nitrogen to carbon ratio (Lea et al., 2007; Gaufichon et al., 2010). Nitrogen is initially taken up from the soil as ammonium or nitrate, the latter being reduced by nitrate reductase (NR) to form nitrite, which is then reduced by nitrite reductase (NiR) to form ammonium (Masclaux-Daubresse et al., 2010). Assimilation then occurs via the glutamine synthetase (GS)–glutamate synthetase (GOGAT) cycle, in which GS incorporates ammonia into glutamate to form glutamine. Finally, asparagine can then be synthesised from glutamine by asparagine synthetase.

Asparagine is synthesised similarly across eukaryotes by the asparagine synthetases, which transfer an amine group from glutamine to aspartate (Dang et al., 1996; Lea et al., 2007; Lomelino et al., 2017). This reaction requires ATP to occur and results in the formation of asparagine and glutamate (Fig. 2.1). The breakdown of asparagine is controlled by different proteins called asparaginases, which are also present across eukaryotes and catalyse the breakdown of asparagine back into aspartate in a hydrolysis reaction that releases ammonia (Fig. 2.1) (Lea et al., 2007; Batool et al., 2016). In bread wheat, there are five distinct asparagine synthetase genes: *TaASN1*, *TaASN2*, *TaASN3.1*, *TaASN3.2*, and *TaASN4*, which show different patterns of spatial and temporal expression during development (Gao et al., 2016; Xu et al., 2018). *TaASN1* is the most highly expressed during early development in vegetative tissues, whereas *TaASN2* is the most highly expressed in the embryo and endosperm during late development; both are upregulated in response to sulphur deficiency (Gao et al., 2016; Xu et al., 2018; Curtis et al., 2019). These studies also highlighted that *TaASN3.1*, *TaASN3.2*, and *TaASN4* are all expressed at much lower levels than *TaASN1* and *TaASN2* and these expression patterns are reflected in publicly-available RNA-seq datasets (Ramírez-González et al., 2018). At the homoeologue level, *TaASN1* is differentially expressed across the A, B, and D genomes and *TaASN2* is expressed differently in the A and D genomes

(Ramírez-González et al., 2018; Curtis et al., 2019). The absence of a B genome *TaASN2* homoeologue in some varieties and its presence in others has been noted previously (Xu et al., 2018), but the origin and effect of this deletion currently remains unclear.

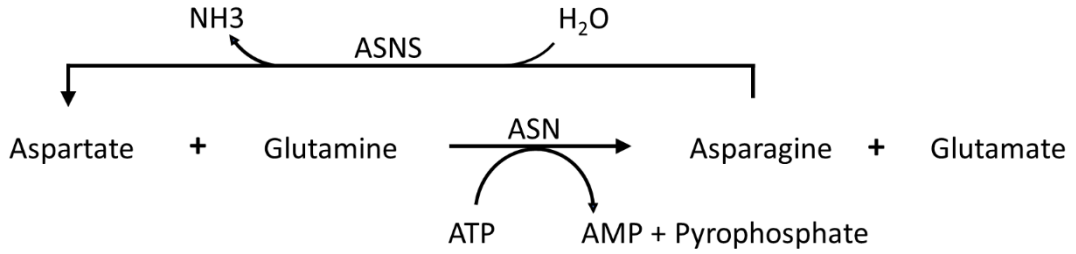


Figure 2.1. **The reactions of asparagine synthetase (ASN) and asparaginase (ASNS).** Asparagine and glutamate are formed from aspartate and glutamine in an ATP-dependent reaction catalysed by ASN. Asparagine is broken down into aspartate in a hydrolysis reaction catalysed by ASNS, releasing ammonia.

The wheat asparaginase genes are less well characterised, but there is some information regarding their genetics and expression. Curtis et al. (2019), for example, identified seven putative asparaginase genes in the wheat genome and showed that the embryo and endosperm specific expression of one of these asparaginases was responsive to sulphur deficiency. In tobacco and lupin, asparaginase gene expression is highest in young developing tissues and in tissues of the developing seed, both of which are nitrogen sinks (Grant & Bevan, 1994). Potassium-dependent and potassium-independent forms of asparaginase are also present (Sieciechowicz et al., 1988a), but it is not known if equivalent forms exist in wheat. Expression of both potassium-dependent and potassium-independent asparaginases during development in Arabidopsis is similarly localised to sink tissues, but the potassium-dependent asparaginase is expressed at lower levels (Bruneau et al., 2006) and asparaginase function is not required for normal development (Ivanov et al. 2012). In contrast, the potassium-dependent asparaginase of *Lotus japonicus* is crucial for plant growth and seed production (Credali et al., 2013). Further analysis of asparaginase gene expression is required to fully understand the function of asparaginase in wheat and its impact on asparagine concentrations and nitrogen mobilisation.

Across different plant species, asparagine transport and accumulation dynamics differ widely. In bread wheat (*Triticum aestivum* L.), the dynamics and magnitude of asparagine transport

within the plant are unclear. In particular, it is not known if asparagine is imported into the grain or how much this import, if occurring, contributes to grain free asparagine concentrations relative to *in situ* asparagine synthesis. Although this has not been studied directly, expression analyses of asparagine synthetase genes highlight in situ synthesis as the foremost determinant, as discussed below. In some species, asparagine transport enables nitrogen mobilisation from source to sink organs in the xylem and enables nitrogen remobilisation in the phloem, in processes such as seed-filling in Arabidopsis (Lam et al., 2003) and leaf senescence (Herrera-Rodríguez et al., 2006; Lea et al., 2007; de Michele et al., 2009). In many, but not all species, free asparagine accumulates to higher concentrations than any other amino acid during germination in many, but not all, species (Lea et al., 2007). This mobilisation and remobilisation of nitrogen in the form of asparagine allows some plant species to transport nitrogen safely, whereas accumulation of excess ammonium, for example, is toxic (Britto et al., 2001). These differences in asparagine transport dynamics have important consequences for crop plants: most asparagine in chicory roots, for example, is transported from leaves (Soares, 2020), whereas asparagine is synthesised *in situ* in potato tubers and is not imported from the leaves at all (Chawla et al., 2012; Muttucumaru et al., 2014). Such differences will influence any strategies used to reduce asparagine concentrations in crop plants. The function of asparagine as a means of nitrogen mobilisation in some plant species is supported by a number of functional studies on asparagine synthetase genes, which also highlight the different roles of these genes and the enzymes they encode (Table 2.1). Such studies have also suggested that the overexpression of certain endogenous asparagine synthetase genes may be a viable strategy for improving nitrogen use efficiency for some species (reviewed in McAllister et al., (2012)), but the outcome is variable depending on the species and the gene.

Table 2.1. Functional studies of asparagine synthetase genes (*ASN*) and the conclusions made from those studies. N (nitrogen).

Species	Modulation	Reported conclusions	Study
Arabidopsis thaliana	*ASN1* overexpression	Tolerance to N deprivation in germination. Enhanced seed protein content.	Lam et al., (2003)
		Slightly more N and protein content. Higher dry weight.	Gaufichon et al., (2017)
	CaAS1 overexpression	Enhanced disease resistance.	Hwang et al., (2011)
	ASN1 knockout	Some disruption of seed formation. Slightly less N and more carbon. Aberrant cell patterns in the embryo.	Gaufichon et al., (2017)
	ASN1 silencing	Negligible effect on virus replication.	Fernández-Calvino et al., (2016)
	ASN2 overexpression	Increased asparagine concentrations.	Igarashi et al., (2009)
	ASN2 knockout	Reduced salt tolerance.	Maaroufi-Dgumi et al., (2011)
		Ammonium accumulation and defective growth.	Gaufichon et al., (2013)
	ASN3 knockout	No visible phenotype in development. No difference in seed carbon or N.	Gaufichon et al., (2016)
Brassica napus (Oilseed rape)	*Asna (E. coli)* overexpression	Poorer performance at low N application than WT. Better performance at high N.	Seiffert et al., (2004)
Capsicum annuum (Capsicum pepper)	*CaAS1* silencing	Increased susceptibility to disease.	Hwang et al., (2011)
Lactuca sativa (Garden lettuce)	*Asna (E. coli)* overexpression	Faster vegetative growth and greater dry weight. Improved N status.	Giannino et al., (2008)
Nicotiana benthamiana	*ASN* silencing	Morphological defects upon infection. Reduced virus accumulation.	Fernández-Calvino et al., (2016)
Oryza sativa (Rice)	*ASN1* knockout	Reduced stature and fewer tillers.	Luo et al., (2016)

2.4. Why does asparagine accumulate during stress?

Alongside its role in normal development, asparagine has been shown to accumulate in response to diverse types of abiotic and biotic stressors in many different species. These stressors include disease, salt and water stress, and nutrient deficiencies, the latter in particular when nitrogen is plentiful but other minerals are lacking (see Lea et al., (2007), Stewart & Larher, (1980) for reviews). Functional studies have indicated that this stress-induced asparagine accumulation may be adaptive against some stressors such as disease and mineral limitation (Table 2.1). One possible explanation for this is that asparagine accumulates during stress as part of the nitrogen remobilisation process or to store nitrogen, similarly to its function in normal development (Fig. 2.2). This has been proposed to happen during infection, in order to divert nitrogen away from the pathogen and to sequester it elsewhere, based on analyses of tomato leaves in response to *Pseudomonas syringae* (Olea et al., 2004). In this model, ammonia released from stress-induced proteolysis is assimilated into glutamine in mesophyll cells. Glutamine is subsequently exported from these cells into the phloem, where asparagine is then synthesised. Asparagine transport then allows mobilisation of nitrogen to healthy tissue and the pathogen is deprived of a source of nitrogen from the host. Hwang et al. (2011) also show that infection-induced asparagine synthetase expression leads to reactive oxygen species (ROS) bursts and nitric oxide (NO) production, both of which have major roles in defence, so asparagine accumulation may also increase disease resistance this way. Consequently, overexpression of asparagine synthetase genes can confer greater disease resistance, whereas asparagine synthetase gene silencing can confer greater sensitivity (Hwang et al., 2011). However, Fernández-Calvino et al. (2016) show that silencing of asparagine synthetase gene expression in *Nicotiana benthamiana* can cause a reduction in viral replication. They suggest that this is because asparagine accumulation detoxifies the cell of ammonium, allowing the cell to remain healthy and consequently allowing the virus to replicate to a greater extent. Asparagine accumulation has also been recorded in wheat grain in response to disease, specifically in response to the withdrawal of fungicide application (Curtis et al., 2016), indicating that these processes may also occur in wheat.

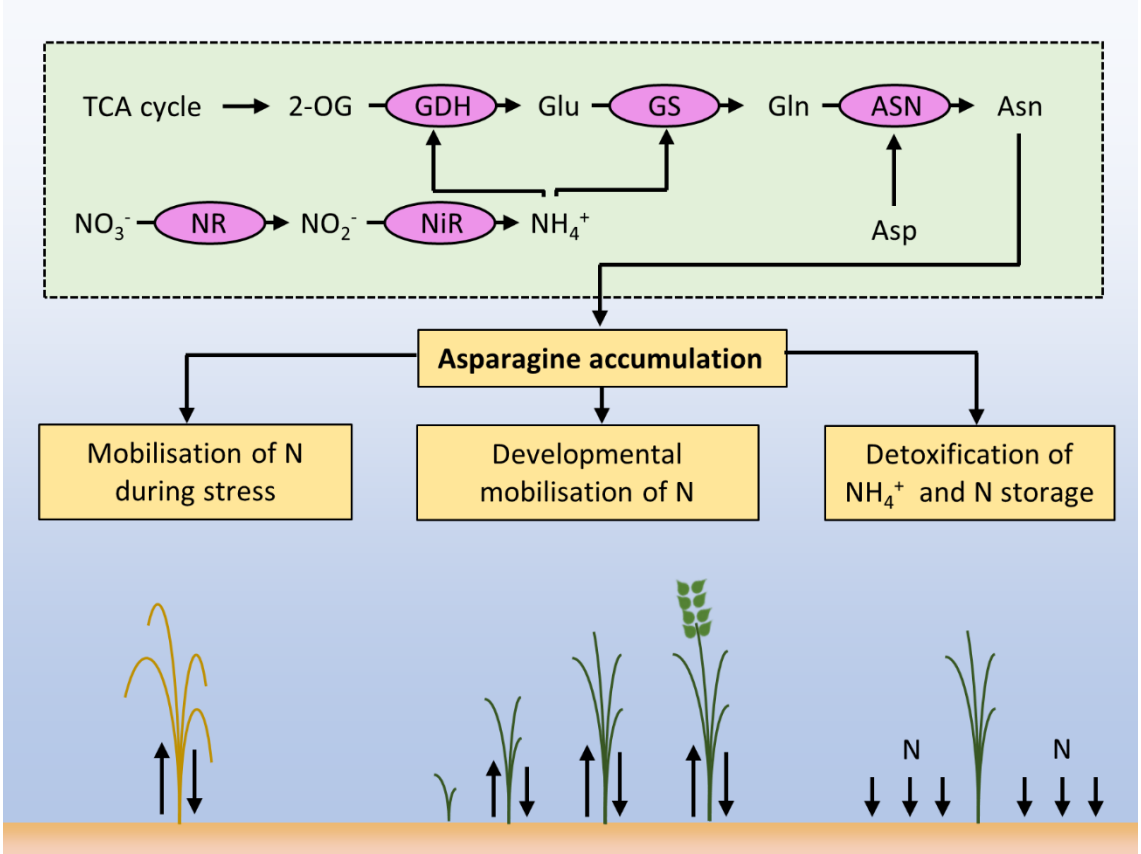


Figure 2.2. Synthesis and functions of asparagine. Nitrate (NO_3^-) is taken up from the soil and reduced to ammonium (NH_4^+) through nitrate reductase (NR) and nitrite reductase (NiR). Ammonium is assimilated via glutamate dehydrogenase (GDH) and glutamine synthetase (GS) to form glutamine (Glu), which is then used to form asparagine (Asn) alongside aspartate (Asp). Asparagine mobilises nitrogen from source to sink tissues during germination, vegetative growth, senescence, and seed filling, as well as during stress. Detoxification of ammonium may also be an important function of asparagine accumulation when nitrogen (N) is abundant and during stress when ammonium accumulates.

Asparagine accumulation may also play a role in nitrogen remobilisation and ammonium detoxification during abiotic stress. Beato et al. (2014) suggest that asparagine accumulation occurs in these situations as a result of a high ammonium to hexose ratio in the cell. Many different abiotic stressors can induce general energy stress (Lastdrager et al., 2014) and proteolysis (Hildebrandt et al., 2015), causing decreases in cellular hexose concentrations and increases in ammonium. Excess ammonium produced by proteolysis can be recycled into amino acids and asparagine in particular by the GS/GOGAT cycle and asparagine synthetase, respectively, thereby preventing the toxic build-up of ammonium. This function of asparagine

accumulation remains experimentally unverified, but Lam et al. (2003) have shown that asparagine accumulation is adaptive against nitrogen limitation during germination, and Maaroufi-Dguimi et al. (2011) have also suggested a role in salt tolerance.

The proposed role of asparagine accumulation in salt and drought tolerance is comparable to the role of the amino acid proline. Proline is considered to play a role in stress tolerance against drought and salt stress: its accumulation allows it to function as an osmolyte and prevent ROS bursts, although its adaptive value across species is uncertain (Szabados & Savouré, 2010). Rashmi et al. (2019) suggest that asparagine may have a similar role as an osmolyte, because of the upregulation of asparagine synthetase gene expression during salt stress in *Pandanus odorifer*. However, Yadav et al. (2019) demonstrate that asparagine correlates negatively with yield-gap based drought tolerance in wheat. Based on this, they suggest that asparagine instead accumulates as a result of drought-induced senescence and that asparagine accumulation is indicative of poor drought tolerance. Varieties of wheat less tolerant to drought will show more drought-induced senescence, which increases asparagine concentrations and asparagine synthetase gene expression. Therefore, the function of asparagine accumulation during abiotic stress in wheat may be principally one of nitrogen remobilisation and ammonium detoxification instead.

2.5. How does asparagine accumulate during stress?

The accumulation of asparagine during plant stress is reflected in the upregulation of asparagine synthetase genes in response to diverse stressors in a range of plant species. Such upregulation in response to diverse types of stress has been observed, for example, in sunflower (*Helianthus annuus*) (Herrera-Rodriguez et al., 2007), Arabidopsis (*Arabidopsis thaliana*) (Lam et al., 1998; Baena-González et al., 2007), maize (*Zea mays*) (Chevalier et al., 1996), and wheat (*Triticum aestivum*) (Wang et al., 2005; Curtis et al., 2019). Upregulation has also been observed in soybean (*Glycine max*) (Antunes et al., 2008), sunflower (*H.annuus*) (Herrera-Rodríguez et al., 2004), barley (*Hordeum vulgare*) (Avila-Ospina et al., 2015), poplar (*Populus simonii x Populus nigra*) (Qu et al., 2019), and common bean (*Phaseolus vulgaris*) (Osuna et al., 2001) in response to nitrogen. This implies that the asparagine that accumulates during stress is synthesised predominantly *de novo* and not just as a result of proteolysis and amino acid catabolism. Such studies showing the upregulation of asparagine synthetase gene expression, alongside a comprehensive metabolic network (Curtis et al., 2018a) detailing its

regulation in plants, suggest that asparagine synthetase is primarily regulated at the transcriptional level, and therefore that asparagine accumulation is determined in the main by the level of asparagine synthetase gene expression. In the proposed regulatory pathways for the asparagine synthetase genes in Arabidopsis, there are two distinct pathways of upregulation for *AtASN1* and *AtASN2* (Curtis et al., 2018a). For *AtASN1*, the signalling pathway (Fig. 2.3) starts off with activation of SnRK1.1. and SnRK1.2 (sucrose nonfermenting-1-related protein kinases) in response to low glucose availability and/or the activity of SnAKs (SnRK1-activating kinases). The SnRK1s are protein kinases that are active during periods of low cellular energy status and are involved in many diverse processes, including autophagy, metabolism, and stress responses (Wurzinger et al., 2018; Rodriguez et al., 2019). The SnAKs are mostly known for their ability to regulate the SnRK1s and the downstream processes they control (Glab et al., 2017). The signal is then relayed through basic leucine zipper (bZIP) transcription factors. These are known to regulate a range of developmental and stress responses, often through specific dimerisations with one another (Dröge-Laser & Weiste, 2018). Upon activation by the SnRK1s, the basic region/leucine zipper (bZIP) proteins upregulate *AtASN1* to drive asparagine accumulation.

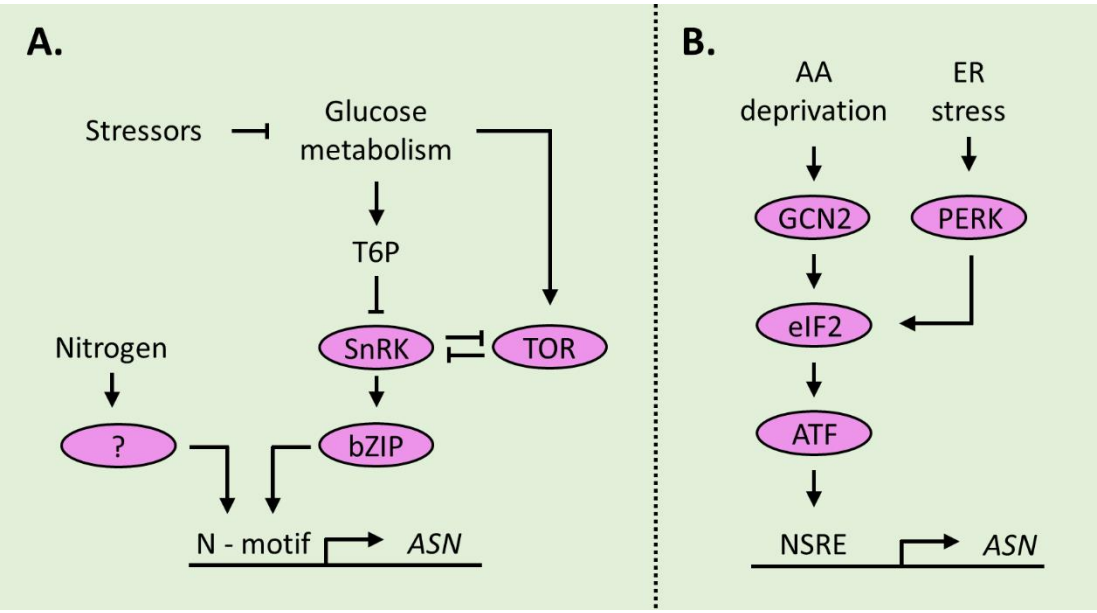


Figure 2.3. Models of asparagine synthetase regulation. a. Regulation of asparagine synthetase (*ASN*) gene expression in plants by stress and nitrogen. **b.** Regulation of *ASN* in mammals by amino acid (AA) deprivation and endoplasmic reticulum (ER) stress. T6P (trehalose-6-phosphate), SnRK (sucrose non-fermenting related kinase), TOR (target of rapamycin), bZIP (basic region/leucine zipper), N (nitrogen), GCN2 (general control

nonderepressible 2), PERK (protein kinase R-like endoplasmic reticulum kinase), eIF2 (eukaryotic initiation factor 2), ATF (activating transcription factor), NSRE (nutrient sensing response element).

Studies in Arabidopsis have indicated that regulation of *AtASN1* expression by SnRK1 and bZIP transcription factors is highly responsive to nutrients (Baena-González et al., 2007). This nutrient control is mostly exerted by trehalose-6-phosphate (T6P), which acts as an inhibitor of SnRK1 (Zhang et al., 2009). T6P abundance closely traces the concentrations of cellular sucrose, so acts as a sugar signal. Consequently, SnRK1 is inhibited under conditions of high sucrose, such as during the dark and under stress (Baena-González et al., 2007). SnRK1 has also been shown to be impacted by stressors that downregulate glucose metabolism. For example, Dong et al. (2017) showed that sulphur deprivation was able to activate SnRK1 through downregulated glucose metabolism and through the reduction in activity of target of rapamycin (TOR). TOR is a protein kinase that, like SnRK1, responds to many different environmental cues to regulate growth and metabolism, but mostly in the opposite direction to SnRK1 (Shi et al., 2018). It is worth noting at this point that the relationship between sucrose and hexoses in different tissues will depend on the activities of the enzymes that interconvert them, such as sucrose synthase, invertase, sucrose phosphate synthase and sucrose phosphate phosphatase. The *Fusarium* toxin deoxynivalenol has also recently been shown to upregulate *TaSnRK1* (Perochon et al., 2019), further suggesting that the responsiveness of SnRK1 to diverse stressors may partially explain the patterns of expression shown by the *ASNs*.

Multiple members of the bZIP family have been shown to activate *AtASN1* (Baena-González et al., 2007; Hanson et al., 2008; Dietrich et al., 2011) and some of these bZIPs possess amino acid motifs that match SnRK1 phosphorylation target sites (Zhang et al., 2008a). Recent experimental evidence has also demonstrated that SnRK1 phosphorylates bZIP63 *in vivo* (Mair et al., 2015), changing its dimerisation preferences and consequently its activity, confirming direct regulation of bZIPs by SnRK1. Therefore, the asparagine accumulation that occurs in response to sugar deprivation is likely an outcome of signalling through T6P, SnRK1, and bZIP transcription factors, eventually resulting in upregulated *AtASN1* expression.

The promoter region of wheat *TaASN1* also contains a potential regulatory element that is known to be targeted by bZIPs. This element is called the nitrogen (N)—motif, or alternatively a general control nonderepressible-4 (GCN4)—like motif (Albani et al., 1997). bZIPS have

also been implicated in the responses observed in wheat grain under sulphur deficiency (Curtis et al., 2019; Raffan et al., 2020), further strengthening the link between stress and asparagine accumulation.

Asparagine accumulation is also likely impacted by the activity and regulation of the asparaginases, although this is less well characterised than asparagine synthetase regulation. Firstly, the activity of asparaginase is responsive to environmental cues. Potassium is necessary in order for some asparaginases to function (Sodek et al., 1980) and the activity of the potassium-dependent asparaginase is negatively regulated by glutamine (Tonin & Sodek, 1990). Asparaginase activity is also diurnally regulated, with greater activity in the light (during ATP abundancy, the glutamate synthase cycle, and protein biosynthesis), and less activity in the dark (during proteolysis) (Sieciechowicz et al., 1988b, c; Sieciechowicz & Ireland, 1989). Secondly, the expression of asparaginase genes in the grain is responsive to the environment, being downregulated by high rates of nitrogen fertiliser (Zheng et al., 2018) or sulphur deficiency (Curtis et al., 2019). The impact of these different environmental cues on asparaginase activity and expression is consistent with the asparagine accumulation that is seen in response to different stressors, suggesting that changes in asparaginase activity could be at least partly responsible for this asparagine accumulation. This is further suggested by functional studies of asparaginase genes, which demonstrate that the absence of asparaginase activity can greatly increase asparagine accumulation. In *Lotus japonicus*, knockout of asparaginase genes increases asparagine concentrations significantly (Credali et al., 2013), while Arabidopsis mutants lacking asparaginases accumulate more free asparagine in mature seed (Ivanov et al., 2012). Consequently, the asparaginases may also play an important role in controlling stress-induced asparagine accumulation.

2.6. Insights on asparagine synthetase regulation from other eukaryotes

The regulation of asparagine synthetase in mammals is, like plants, principally mediated at the transcriptional level. Deprivation of amino acids and glucose has been shown to cause an upregulation of mammalian asparagine synthetase gene expression in many studies (Gong et al., 1991; Guerrini et al., 1993; Hutson and Kilberg, 1994; Hutson et al., 1997; Barbosa-Tessmann et al., 1999a, b). The first regulators implicated in this nutritional control of transcription were found to be elements in the promoter of the asparagine synthetase gene, subsequently called nutrient sensing response elements (NSREs) and unfolded protein response

(UPR) elements (Guerrini et al., 1993; Barbosa-Tessmann et al., 1999a, 2000) (Fig. 3). These elements were found to be targeted by a combination of upregulating transcription factors, such as activating transcription factors 4 (ATF4) and 5 (ATF5) (Siu et al., 2002; Al Sarraj et al., 2005), alongside downregulating transcription factors, such as activating transcription factor 3 (ATF3), CCAAT-enhancer-binding protein β (C/EBPβ), and C/EB homology protein (CHOP) (Chen et al., 2004; Su & Kilberg, 2008; Thiaville et al., 2008). The ATFs belong to the mammalian bZIP family and have regulatory roles during development and stress (Hai, 2007). The transcription factors CHOP and C/EBPβ belong to the CCAAT-enhancer-binding protein (C/EBP) family of transcription factors and also regulate genes involved in growth and metabolism (Nerlov, 2007).

ATF4 and ATF5 are regulated to a large extent by eukaryotic translation initiation factor 2 (eIF2), which is in turn regulated by the protein kinases general control nonderepressible (GCN)-2 and protein kinase-like endoplasmic reticulum kinase (PERK) (Balasubramanian et al., 2013). eIF2 is also regulated by RNA-dependent protein kinase (PKR) and haem-regulated inhibitor (HRI), which respond to viral infection and haem deprivation respectively (Wek et al., 2006), but these have not been discussed as much as GCN2 and PERK in the context of asparagine synthetase regulation. GCN2 is activated by uncharged tRNA, which is abundant at low amino acid concentrations, whereas PERK is activated by endoplasmic reticulum stress (Balasubramanian et al., 2013). Consequently, the signalling pathways controlled by GCN2 and PERK are called the amino acid response (AAR) and unfolded protein response (UPR), respectively, and they show how nutrient status and stress are linked to asparagine synthetase gene regulation. Interestingly, the SnRK1 homologue, AMP-activated protein kinase (AMPK), does not seem to play a large role in mammalian asparagine synthetase gene regulation: AMPK does not affect asparagine synthetase gene expression during glucose deprivation (Cui et al., 2007), for example, in contrast to the role of SnRK1 in plants described above.

In fungi, GCN2 is activated by diverse stressors such as amino acid limitation, oxidative stress, and glucose starvation, and relays this signal by phosphorylating the alpha subunit of eIF2 (eIF2α). In budding yeast (*Saccharomyces cerevisiae*) this leads to an increase in translation of GCN4, a homologue of ATF4, whilst general protein synthesis is reduced (Hinnebusch, 1994; Yang et al., 2000; Mascarenhas et al., 2008). GCN4 binds a "GCN4 box" regulatory motif to induce asparagine synthetase gene expression (Dang et al., 1996; Natarajan et al., 2001). The GCN4 box is analogous to the NSREs in mammals and identical to the putative regulatory motif identified in wheat *TaASN1* described above (Gao et al., 2016). Although plants do not

possess a GCN4/ATF homologue (Halford, 2005), the nitrogen element may be targeted by other bZIPs instead (Albani et al., 1997), as discussed above. Plants also do not possess *PERK* (Ruberti & Brandizzi, 2014), *HRI*, or *PRK* homologues (Halford et al., 2004), but do possess a *GCN2* homologue (Zhang et al., 2003). Like its mammalian and yeast counterparts, plant GCN2 responds to amino acid deprivation (Lageix et al., 2008; Zhang et al., 2008b) and a range of other stressors, including UV-radiation, cold shock, wounding, treatment with methyl jasmonate, salicylate, and cadmium salts (Lageix et al., 2008; Sormani et al., 2011). Plant GCN2 is also activated by interacting with uncharged tRNA, in the same way as the yeast enzyme (Li et al., 2013). However, when *GCN2* is overexpressed in wheat it reduces the expression of *TaASN1* (Byrne et al., 2012), suggesting that it is a negative regulator, so more research is required to elucidate the role of GCN2 in controlling plant asparagine synthetase gene expression. Interestingly, asparagine synthetase forms filaments in yeast during nutrient stress, whereas in mammals the enzymes localise to the mitotic spindle, suggesting "moonlighting" functions distinct from their better known function (Noree et al., 2018). It is not known whether plant asparagine synthetases also show distinct cellular localisations and moonlighting functions.

In recent years, asparagine synthetase regulation in mammals has been of particular interest because of its relevance to cancer biology. The mammalian asparagine synthetase catalyses an identical reaction to the plant asparagine synthetase (Lomelino et al., 2017) and its basal expression is low in all organs except for the pancreas, where it is important for protecting against nutrient stress (Mukherjee et al., 2020). Asparagine synthetase gene expression in human tumour cells increases in response to glucose deprivation, glutamine deprivation, cisplatin treatment, and hypoxia (Cui et al., 2007; Ameri et al., 2010; Zhang et al., 2014) and appears to play a role in cell proliferation (Gong & Basilico, 1990). Notably, Knott et al. 2018 showed that depleting asparagine by dietary restriction or by using asparaginases in a murine model of breast cancer could reduce the number of metastases, the development of which is associated with poorer survival. Asparagine may achieve these effects by acting as an amino acid exchange factor: intracellular asparagine could be exchanged with extracellular amino acids to regulate amino acid homeostasis and metabolism (Krall et al., 2016). Asparagine could also achieve these effects by promoting an epithelial-mesenchymal transition (Knott et al., 2018), a cellular transition associated with increased drug resistance and survivability. The increased understanding of the role of asparagine in cancer (Chiu et al., 2020; Kanarek et al., 2020) has highlighted a need to further assess the role of dietary asparagine in future studies.

There is also a link between GCN2 activity, asparagine synthetase gene expression and sulphur metabolism in mammalian systems, which is intriguing given the response of asparagine synthetase to sulphur deficiency in wheat: both phosphorylation of eIF2α and expression of asparagine synthetase have been shown to be higher in liver cells of rats fed a diet deficient in sulphur-containing amino acids compared to well-nourished rats (Sikalidis & Stipanuk, 2010).

2.7. Can asparagine accumulation in wheat be reduced?

Some strategies to reduce asparagine accumulation in wheat cultivation have been published and included in the Acrylamide Toolbox produced by FoodDrinkEurope (2019). The Acrylamide toolbox provides a range of mitigation techniques for all actors in the food supply chain, including the use of asparaginase during processing, changing cooking temperature, excluding overly-browned products after cooking and processing, and the use of crop varieties with low acrylamide-forming potential. In the USA, low asparagine GM potato varieties (Innate® and Innate2®) have been available for several years (USDA-APHIS, 2014), but regulations in the European Union make commercialisation of new GM crops in Europe just about impossible (Halford, 2019). No similar GM varieties of wheat or other cereals have been produced, anyway, but there is considerable variation in the basal concentrations of free asparagine in the grain of conventional wheat varieties (Taeymans et al., 2004; Claus et al., 2006; Curtis et al., 2009, 2018b), suggesting that asparagine concentrations could be lowered through selection and breeding. Heritability of free asparagine concentrations in the grain ranges from low to moderate and, as we have discussed already, asparagine accumulation is highly influenced by environmental factors. Consequently, the SNPs and QTL associated with the trait explain only a small proportion of the variance and no common QTL have been identified (Emebiri, 2014; Rapp et al., 2018). Asparagine concentrations for new varieties entering the market are also not required to be published, so it is impossible for the food industry to tell farmers which varieties they would like them to grow. In addition, food businesses generally purchase wheat grain from the world market rather than local suppliers. This ensures that they get the best price available but means that they have little control over the varieties that are grown. As a result, recommended practices for asparagine reduction in wheat currently emphasise crop-management strategies (FoodDrinkEurope, 2019), specifically the application of nitrogen fertiliser at the minimum levels to ensure that the optimum yield

and required protein content is achieved, the application of sulphur fertiliser to ensure that sulphur deficiency is avoided, and the use of fungicides to control disease.

Application of nitrogen fertiliser is known to increase grain asparagine concentrations across wheat varieties (Claus et al., 2006; Weber et al., 2008; Martinek et al., 2009), reflecting the upregulation of asparagine synthetase gene expression in response to nitrogen seen in many plant species. Application of excess nitrogen fertiliser compromises food safety this way, as well as negatively impacting many aspects of environmental health (Cameron et al., 2013), but is also necessary to achieve good yields to ensure food security and to produce protein concentrations necessary for breadmaking (Weber et al., 2008). The application of inorganic and organic nitrogen in conventional and organic farming systems, respectively, is associated with higher concentrations of free asparagine in the grain (Stockmann et al., 2018) but organic methods have been shown to achieve lower asparagine concentrations than conventional farming (Stockmann et al., 2019), likely due to the slower release of nitrogen in that system. Organic systems are also associated with a reduced yield and lower protein content (Stockmann et al., 2018), but the association between protein content and asparagine concentrations was not strong, indicating that asparagine concentrations could be reduced through decreased nitrogen application without compromising protein concentrations too much, at least in the varieties used in this study. Martinek et al. (2009) similarly found correlations ranging from weak to strong between protein content and asparagine concentrations in varieties tested in their study, further emphasising the untapped potential of variety selection. Nitrogen application rates of 180 kg per hectare have been described as enabling accumulation of protein to levels sufficient for breadmaking (Weber et al., 2008) whilst minimising asparagine concentrations (Stockmann et al., 2018), so this application rate may be optimal for some varieties. Nevertheless, varietal differences in the response to nitrogen application in terms of protein and asparagine accumulation point to a need for asparagine screening in plant breeding programmes in order to develop reliable and useful recommendations.

Adequate sulphur application is another important part of a fertiliser regime in order to reduce free asparagine concentrations in wheat, as increasing application rates are known to reduce asparagine accumulation (Granvogl et al., 2007; Curtis et al., 2018b). In fact, sulphur deficiency can cause a many-fold increase in free asparagine concentration. A large increase in free asparagine concentrations is one of the issues most likely to result in the acrylamide concentrations of a food product becoming unacceptably high (reviewed in Raffan et al., (2020)). As a result, the application of sulphur at a rate of 20 kg per hectare has previously

been recommended for all wheat destined for human consumption (Raffan et al., 2020). Sulphur deficiency may increase asparagine concentrations as a result of increased proteolysis and energy stress, in order to mobilise nitrogen and detoxify ammonium as outlined above (Beato et al., 2014; Dong et al., 2017), or to store nitrogen for synthesis of sulphur-rich proteins when sulphur availability increases (Zhao et al., 1999). Similarly to the relationship between nitrogen and asparagine, the effect of sulphur on asparagine concentrations also varies greatly depending on the wheat genotype (Curtis et al., 2018b).

The final recommendation is the application of fungicide and good phytosanitary practices, i.e. disease control (FoodDrinkEurope, 2019). Application of fungicides has been shown to reduce asparagine accumulation in wheat (Martinek et al., 2009; Curtis et al., 2016), reflecting the effect of disease in general on asparagine accumulation. The response of asparagine concentrations to fungicide application differs greatly between genotypes (Curtis et al., 2016), likely as a result of different levels of disease resistance, indicating again the potential for variety selection to help control asparagine concentrations.

These crop-management strategies help to address some of the agronomic causes of increases in free asparagine concentrations, but, as discussed above, research indicates that free asparagine accumulation in wheat grain probably occurs in response to many other stressors. Drought, for example, can cause an increase in free asparagine in potato, but the responses are, again, genotype-dependent and complex, with moderate drought stress in potato actually decreasing acrylamide-forming potential in some varieties (Muttucumaru et al., 2015). In wheat, drought stress is known to increase asparagine concentrations (Carillo et al., 2005), which Yadav et al. (2019) attribute to senescence during severe stress.

Deficient or excessive levels of macro- and micro-nutrients can lead to asparagine accumulation in many plant species (see Lea et al., (2007), Stewart & Larher (1980) for review) but the effects of precise application rates on asparagine concentrations in wheat have not been tested. Nutrient availability is significantly altered by arbuscular mycorrhizal fungi (AMF), leading to increases in asparagine concentrations with greater AMF colonisation in some experiments (Salvioli et al., 2012; Whiteside et al., 2012; Gaude et al., 2015), whilst leading to decreases in others (Saia et al., 2015). Provision of nutrients such as phosphorus by AMF reduces asparagine concentrations under nitrogen deficiency (Saia et al., 2015) but, under conditions of nitrogen sufficiency, this is likely to be counterbalanced by the effect of increased nitrogen availability. Consequently, different AMF soil inoculation strategies may be able to

modulate asparagine concentrations, and the effectiveness of such an approach could be enhanced if it were used alongside the other crop-management strategies described above. Overall, crop-management strategies are crucial in ensuring that free asparagine concentrations are kept as low as possible (Fig. 2.4).

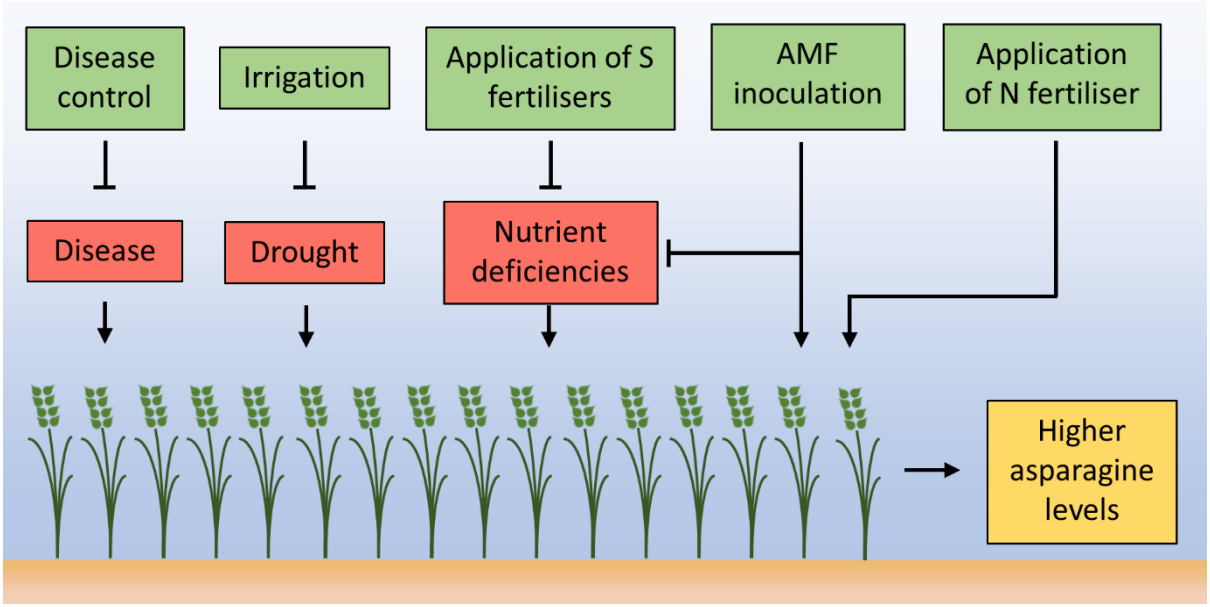


Figure 2.4. Overview of crop-management strategies and their effects on asparagine concentrations. Crop-management strategies (green boxes) help to mitigate plant stressors (red boxes) and influence plant growth, with variable impacts on asparagine concentrations. Disease control, irrigation, and application of nutrients like sulphur (S) help to combat stressors which otherwise increase asparagine accumulation. Arbuscular mycorrhizal fungi (AMF) inoculation has mixed effects on asparagine accumulation due to its provision of nitrogen and other nutrients. Application of nitrogen (N) fertiliser increases asparagine concentrations as a result of increased nitrogen storage and its mobilisation as asparagine.

2.8. Conclusions

Attempts to reduce free asparagine concentrations in wheat grain, and thereby dietary acrylamide intake, must be balanced against other aspects of wheat cultivation and strategies to improve public health. For example, free asparagine concentration could be greatly reduced by withholding all nitrogen application, but this would devastate yields and grain protein composition (Hawkesford, 2014). Dietary acrylamide intake could also be reduced by less consumption of wholegrain foods (Raffan & Halford, 2019), but these foods are generally high

in fibre, which is known to decrease chronic inflammation, thereby reducing the risk of cancers, heart disease, and many other health issues (Swann et al., 2019). With regards to flavour, the Maillard reaction is responsible for the formation of many desirable flavour compounds alongside acrylamide (Raffan & Halford, 2019), so total reduction of the Maillard reaction would negatively impact flavour. Consequently, strategies to reduce asparagine in wheat and acrylamide formation in baked wheat products must account for such outcomes.

Fortunately, the reduction of asparagine in wheat is likely to go hand-in-hand with the reduction of plant stress, which is desirable for all in the food supply chain, and is unlikely to impact flavour formation as much as other acrylamide mitigation strategies (Xu et al., 2016). While the molecular mechanisms linking free asparagine accumulation in wheat grain and nutrient availability have been elucidated in part, the effects of other abiotic stresses remain mostly unknown, although research from other plants and other eukaryotes have provided insights, which we have summarised in this review. Similarly, the function of free asparagine accumulation in wheat has mostly been inferred from its function in other plants. A greater understanding of the functions and mechanisms of asparagine accumulation in wheat will therefore enable the development of new strategies for asparagine reduction, and perhaps the breeding of reliably low asparagine varieties.

2.9. References

al Sarraj, J., Vinson, C., & Thiel, G. (2005). Regulation of asparagine synthetase gene transcription by the basic region leucine zipper transcription factors ATF5 and CHOP. *Biological Chemistry*, *386*(9), 873-879.

Albani, D., Hammond-Kosack, M. C., Smith, C., Conlan, S., Colot, V., Holdsworth, M., & Bevan, M. W. (1997). The wheat transcriptional activator SPA: a seed-specific bZIP protein that recognizes the GCN4-like motif in the bifactorial endosperm box of prolamin genes. *The Plant Cell*, *9*(2), 171-184.

Ameri, K., Luong, R., Zhang, H., Powell, A. A., Montgomery, K. D., Espinosa, I., ... & Jeffrey, S. S. (2010). Circulating tumour cells demonstrate an altered response to hypoxia and an aggressive phenotype. *British Journal of Cancer*, *102*(3), 561-569.

Antunes, F., Aguilar, M., Pineda, M., & Sodek, L. (2008). Nitrogen stress and the expression of asparagine synthetase in roots and nodules of soybean (*Glycine max*). *Physiologia Plantarum*, *133*(4), 736–743.

Avila-Ospina, L., Marmagne, A., Talbotec, J., Krupinska, K., & Masclaux-Daubresse, C. (2015). The identification of new cytosolic glutamine synthetase and asparagine synthetase genes in barley (*Hordeum vulgare* L.), and their expression during leaf senescence. *Journal of Experimental Botany*, *66*(7), 2013–2026.

Baena-González, E., Rolland, F., Thevelein, J. M., & Sheen, J. (2007). A central integrator of transcription networks in plant stress and energy signalling. *Nature*, *448*(7156), 938–942.

Balasubramanian, M. N., Butterworth, E. A., & Kilberg, M. S. (2013). Asparagine synthetase: regulation by cell stress and involvement in tumor biology. *American Journal of Physiology-Endocrinology and Metabolism*, *304*(8), 789-799.

Barbosa-Tessmann, I. P., Chen, C., Zhong, C., Schuster, S. M., Nick, H. S., & Kilberg, M. S. (1999). Activation of the unfolded protein response pathway induces human asparagine synthetase gene expression. *Journal of Biological Chemistry*, *274*(44), 31139-31144.

Barbosa-Tessmann, I. P., Chen, C., Zhong, C., Siu, F., Schuster, S. M., Nick, H. S., & Kilberg, M. S. (2000). Activation of the human asparagine synthetase gene by the amino acid response and the endoplasmic reticulum stress response pathways occurs by common genomic elements. *Journal of Biological Chemistry*, *275*(35), 26976-26985.

Barbosa-Tessmann, I. P., Pineda, V. L., Nick, H. S., Schuster, S. M., & Kilberg, M. S. (1999). Transcriptional regulation of the human asparagine synthetase gene by carbohydrate availability. *Biochemical Journal*, *339*(1), 151-158.

Batool, T., Makky, E. A., Jalal, M., & Yusoff, M. M. (2016). A comprehensive review on L-asparaginase and its applications. *Applied Biochemistry and Biotechnology*, *178*, 900-923.

Beato, V. M., Rexch, J., Navarro-Gochicoa, M. T., Camacho-Cristóbal, J. J., Herrera-Rodríguez, M. B., & González-Fontes, A. (2014). Boron deficiency increases expressions of asparagine synthetase, glutamate dehydrogenase and glutamine synthetase genes in tobacco roots irrespective of the nitrogen source. *Soil Science and Plant Nutrition*, *60*(3), 314–324.

Britto, D. T., Siddiqi, M. Y., Glass, A. D., & Kronzucker, H. J. (2001). Futile transmembrane NH4+ cycling: a cellular hypothesis to explain ammonium toxicity in plants. *Proceedings of the National Academy of Sciences*, *98*(7), 4255-4258.

Bruneau, L., Chapman, R., & Marsolais, F. (2006). Co-occurrence of both L-asparaginase subtypes in Arabidopsis: At3g16150 encodes a K+-dependent L-asparaginase. *Planta*, *224*, 668-679.

Byrne, E. H., Prosser, I., Muttucumaru, N., Curtis, T. Y., Wingler, A., Powers, S., & Halford, N. G. (2012). Overexpression of GCN2-type protein kinase in wheat has profound effects on free amino acid concentration and gene expression. *Plant Biotechnology Journal*, *10*(3), 328-340.

Cameron, K. C., Di, H. J., & Moir, J. L. (2013). Nitrogen losses from the soil/plant system: a review. *Annals of Applied Biology*, *162*(2), 145-173.

Carillo, P., Mastrolonardo, G., Nacca, F., & Fuggi, A. (2005). Nitrate reductase in durum wheat seedlings as affected by nitrate nutrition and salinity. *Functional Plant Biology*, *32*(3), 209–219.

Chawla, R., Shakya, R., & Rommens, C. M. (2012). Tuber-specific silencing of asparagine synthetase-1 reduces the acrylamide-forming potential of potatoes grown in the field without affecting tuber shape and yield. *Plant Biotechnology Journal*, *10*(8), 913-924.

Chen, H., Pan, Y. X., Dudenhausen, E. E., & Kilberg, M. S. (2004). Amino acid deprivation induces the transcription rate of the human asparagine synthetase gene through a timed program of expression and promoter binding of nutrient-responsive basic region/leucine zipper transcription factors as well as localized histone acetylation. *Journal of Biological Chemistry*, *279*(49), 50829-50839.

Chevalier, C., Bourgeois, E., Just, D., & Raymond, P. (1996). Metabolic regulation of asparagine synthetase gene expression in maize (*Zea mays* L.) root tips. *The Plant Journal*, *9*(1), 1–11.

Chiu, M., Taurino, G., Bianchi, M. G., Kilberg, M. S., & Bussolati, O. (2020). Asparagine synthetase in cancer: beyond acute lymphoblastic leukemia. *Frontiers in Oncology*, *9*, 1480

Claus, A., Schreiter, P., Weber, A., Graeff, S., Herrmann, W., Claupein, W., ... & Carle, R. (2006). Influence of agronomic factors and extraction rate on the acrylamide contents in yeast-leavened breads. *Journal of Agricultural and Food Chemistry*, *54*(23), 8968-8976.

Credali, A., García-Calderón, M., Dam, S., Perry, J., Díaz-Quintana, A., Parniske, M., ... & Márquez, A. J. (2013). The K+-dependent asparaginase, NSE1, is crucial for plant growth and seed production in *Lotus japonicus*. *Plant and Cell Physiology*, *54*(1), 107-118.

Cui, H., Darmanin, S., Natsuisaka, M., Kondo, T., Asaka, M., Shindoh, M., Higashino, F., Hamuro, J., Okada, F., Kobayashi, M., Nakagawa, K., Koide, H., & Kobayashi, M. (2007). Enhanced expression of asparagine synthetase under glucose-deprived conditions protects pancreatic cancer cells from apoptosis induced by glucose deprivation and cisplatin. *Cancer Research*. *67*(7), 3345-3355

Curtis, T. Y., Bo, V., Tucker, A., & Halford, N. G. (2018). Construction of a network describing asparagine metabolism in plants and its application to the identification of genes affecting asparagine metabolism in wheat under drought and nutritional stress. *Food and Energy Security*, *7*(1), e00126.

Curtis, T. Y., Muttucumaru, N., Shewry, P. R., Parry, M. A. J., Powers, S. J., Elmore, J. S., Mottram, D. S., Hook, S., & Halford, N. G. (2009). Effects of genotype and environment on free amino acid levels in wheat grain: Implications for acrylamide formation during processing. *Journal of Agricultural and Food Chemistry, 57(3), 1013-1021*

Curtis, T. Y., Powers, S. J., & Halford, N. G. (2016). Effects of fungicide treatment on free amino acid concentration and acrylamide-forming potential in wheat. *Journal of Agricultural and Food Chemistry, 64*(51), 9689–9696.

Curtis, T. Y., Powers, S. J., Wang, R., & Halford, N. G. (2018). Effects of variety, year of cultivation and sulphur supply on the accumulation of free asparagine in the grain of commercial wheat varieties. *Food Chemistry, 239, 304-313.*

Curtis, T. Y., Raffan, S., Wan, Y., King, R., Gonzalez-Uriarte, A., & Halford, N. G. (2019). Contrasting gene expression patterns in grain of high and low asparagine wheat genotypes in response to sulphur supply. *BMC Genomics*, *20*(1), 628.

Dang, V. D., Valens, M., Bolotin-Fukuhara, M., & Daignan-Fornier, B. (1996). Cloning of the ASN1 and ASN2 genes encoding asparagine synthetases in *Saccharomyces cerevisiae*:

differential regulation by the CCAAT-box-binding factor. *Molecular Microbiology*, *22*(4), 681-692.

de Michele, R., Formentin, E., Todesco, M., Toppo, S., Carimi, F., Zottini, M., Barizza, E., Ferrarini, A., Delledonne, M., Fontana, P., & lo Schiavo, F. (2009). Transcriptome analysis of *Medicago truncatula* leaf senescence: Similarities and differences in metabolic and transcriptional regulations as compared with Arabidopsis, nodule senescence and nitric oxide signalling. *New Phytologist*, *181*(3), 563-575.

Dietrich, K., Weltmeier, F., Ehlert, A., Weiste, C., Stahl, M., Harter, K., & Dröge-Lasera, W. (2011). Heterodimers of the Arabidopsis transcription factors bZIP1 and bZIP53 reprogram amino acid metabolism during low energy stress. *The Plant Cell*, *23*(1), 381-395.

Dong, Y., Silbermann, M., Speiser, A., Forieri, I., Linster, E., Poschet, G., Allboje Samami, A., Wanatabe, M., Sticht, C., Teleman, A. A., Deragon, J. M., Saito, K., Hell, R., & Wirtz, M. (2017). Sulfur availability regulates plant growth via glucose-TOR signaling. *Nature Communications*, *8*(1), 1174.

Dröge-Laser, W., & Weiste, C. (2018). The C/S1 bZIP Network: A Regulatory Hub Orchestrating Plant Energy Homeostasis. *Trends in Plant Science*, *23*(5), 422-433.

Emebiri, L. C. (2014). Genetic variation and possible SNP markers for breeding wheat with low-grain asparagine, the major precursor for acrylamide formation in heat-processed products. *Journal of the Science of Food and Agriculture*, *94*(7), 1422-1429.

European Commission. (2017). Commission Regulation (EU) 2017/2158: establishing mitigation measures and benchmark levels for the reduction of the presence of acrylamide in food. *Official Journal of the European Union*.

Fernández-Calvino, L., Guzmán-Benito, I., del Toro, F. J., Donaire, L., Castro-Sanz, A. B., Ruíz-Ferrer, V., & Llave, C. (2016). Activation of senescence-associated Dark-inducible (DIN) genes during infection contributes to enhanced susceptibility to plant viruses. *Molecular Plant Pathology*, *17*(1), 3-15.

FoodDrinkEurope. (2019). Acrylamide toolbox 2019. *Food Drink Europe, Brussels*.

Gao, R., Curtis, T. Y., Powers, S. J., Xu, H., Huang, J., & Halford, N. G. (2016). Food safety: Structure and expression of the asparagine synthetase gene family of wheat. *Journal of Cereal Science*, *68*, 122-131.

Gaude, N., Bortfeld, S., Erban, A., Kopka, J., & Krajinski, F. (2015). Symbiosis dependent accumulation of primary metabolites in arbuscule-containing cells. *BMC Plant Biology*, *15*(1), 1-9.

Gaufichon, L., Marmagne, A., Belcram, K., Yoneyama, T., Sakakibara, Y., Hase, T., Grandjean, O., Clément, G., Citerne, S., & Boutet-Mercey, S. (2017). ASN 1-encoded asparagine synthetase in floral organs contributes to nitrogen filling in Arabidopsis seeds. *The Plant Journal*, *91*(3), 371–393.

Gaufichon, L., Marmagne, A., Yoneyama, T., Hase, T., Clément, G., Trassaert, M., Xu, X., Shakibaei, M., Najihi, A., & Suzuki, A. (2016). Impact of the disruption of ASN3-encoding asparagine synthetase on Arabidopsis development. *Agronomy*, *6*(1), 12.

Gaufichon, L., Masclaux-Daubresse, C., Tcherkez, G., Reisdorf-Cren, M., Sakakibara, Y., Hase, T., Clement, G., Avice, J., Grandjean, O., & Marmagne, A. (2013). *Arabidopsis thaliana* ASN2 encoding asparagine synthetase is involved in the control of nitrogen assimilation and export during vegetative growth. *Plant, Cell & Environment*, *36*(2), 328–342.

Gaufichon, L., Reisdorf-Cren, M., Rothstein, S. J., Chardon, F., & Suzuki, A. (2010). Biological functions of asparagine synthetase in plants. *Plant Science*, *179*(3), 141-153.

Giannino, D., Nicolodi, C., Testone, G., Frugis, G., Pace, E., Santamaria, P., Guardasole, M., & Mariotti, D. (2008). The overexpression of asparagine synthetase a from E. coli affects the nitrogen status in leaves of lettuce (*Lactuca sativa* L.) and enhances vegetative growth. *Euphytica*, *162*, 11-22.

Glab, N., Oury, C., Guérinier, T., Domenichini, S., Crozet, P., Thomas, M., Vidal, J., & Hodges, M. (2017). The impact of *Arabidopsis thaliana* SNF1-related-kinase 1 (SnRK1)-activating kinase 1 (SnAK1) and SnAK2 on SnRK1 phosphorylation status: characterization of a SnAK double mutant. *The Plant Journal*, *89*(5), 1031-1041.

Gong, S. S., & Basilico, C. (1990). A mammalian temperature-sensitive mutation affecting G1 progression results from a single amino acid substitution in asparagine synthetase. *Nucleic Acids Research*, *18*(12), 3509-3509.

Gong, S. S., Guerrini, L., & Basilico, C. (1991). Regulation of asparagine synthetase gene expression by amino acid starvation. *Molecular and Cellular Biology*, *11*(12), 6059-6066.

Grant, M., & Bevan, M. W. (1994). Asparaginase gene expression is regulated in a complex spatial and temporal pattern in nitrogen-sink tissues. *The Plant Journal*, *5*(5), 695-704.

Granvogl, M., Wieser, H., Koehler, P., von Tucher, S., & Schieberle, P. (2007). Influence of sulfur fertilization on the amounts of free amino acids in wheat. Correlation with baking properties as well as with 3-aminopropionamide and acrylamide generation during baking. *Journal of Agricultural and Food Chemistry*, *55*(10), 4271–4277.

Guerrini, L., Gong, S. S., Mangasarian, K., & Basilico, C. (1993). Cis- and trans-acting elements involved in amino acid regulation of asparagine synthetase gene expression. *Molecular and Cellular Biology*, *13*(6), 3202-3212.

Hai, T. (2007). The ATF Transcription Factors in Cellular Adaptive Responses. In *Gene Expression and Regulation* (pp. 329-340). Springer, New York, NY.

Halford, N. G. (2005). Regulation of carbon and amino acid metabolism: roles of *sucrose nonfermenting-1-related protein kinase-1* and *general control nonderepressible-2-related protein kinase*. *Advances in Botanical Research*, *43*, 93-142.

Halford, N. G. (2019). Legislation governing genetically modified and genome-edited crops in Europe: the need for change. *Journal of the Science of Food and Agriculture*, *99*(1), 8-12.

Halford, N. G., Hey, S., Jhurreea, D., Laurie, S., McKibbin, R. S., Zhang, Y., & Paul, M. J. (2004). Highly conserved protein kinases involved in the regulation of carbon and amino acid metabolism. *Journal of Experimental Botany*, *55*(394), 35-42.

Hanson, J., Hanssen, M., Wiese, A., Hendriks, M. M. W. B., & Smeekens, S. (2008). The sucrose regulated transcription factor bZIP11 affects amino acid metabolism by regulating the expression of *asparagine synthetase 1* and *proline dehydrogenase 2*. *The Plant Journal*, *53*(6), 935-949.

Hawkesford, M. J. (2014). Reducing the reliance on nitrogen fertilizer for wheat production. *Journal of Cereal Science*, *59*(3), 276-283.

Herrera-Rodríguez, M. B., Maldonado, J. M., & Pérez-Vicente, R. (2004). Light and metabolic regulation of HAS1, HAS1. 1 and HAS2, three asparagine synthetase genes in *Helianthus annuus*. *Plant Physiology and Biochemistry*, *42*(6), 511–518.

Herrera-Rodríguez, M. B., Maldonado, J. M., & Pérez-Vicente, R. (2006). Role of asparagine and asparagine synthetase genes in sunflower (*Helianthus annuus*) germination and natural senescence. *Journal of Plant Physiology*, *163*(10), 1061-1070.

Herrera-Rodriguez, M. B., Perez-Vicente, R., & Maldonado, J. M. (2007). Expression of asparagine synthetase genes in sunflower (*Helianthus annuus*) under various environmental stresses. *Plant Physiology and Biochemistry*, *45*(1), 33–38.

Hildebrandt, T. M., Nunes Nesi, A., Araújo, W. L., & Braun, H. P. (2015). Amino Acid Catabolism in Plants. *Molecular Plant*, *8*(11), 1563-1579.

Hinnebusch, A. G. (1994). Translational control of GCN4: an in vivo barometer of initiation-factor activity. *Trends in Biochemical Sciences*, *19*(10), 409-414.

Hutson, R. G., & Kilberg, M. S. (1994). Cloning of rat asparagine synthetase and specificity of the amino acid-dependent control of its mRNA content. *Biochemical Journal*, *304*(3), 745-750.

Hutson, R. G., Kitoh, T., Moraga Amador, D. A., Cosic, S., Schuster, S. M., & Kilberg, M. S. (1997). Amino acid control of asparagine synthetase: Relation to asparaginase resistance in human leukemia cells. *American Journal of Physiology-Cell Physiology*, *272*(5), C1691-C1699.

Hwang, I. S., An, S. H., & Hwang, B. K. (2011). Pepper asparagine synthetase 1 (CaAS1) is required for plant nitrogen assimilation and defense responses to microbial pathogens. *The Plant Journal*, *67*(5), 749–762.

Igarashi, D., Ishizaki, T., Totsuka, K., & Ohsumi, C. (2009). ASN2 is a key enzyme in asparagine biosynthesis under ammonium sufficient conditions. *Plant Biotechnology*, *26*(1), 153-159.

Ivanov, A., Kameka, A., Pajak, A., Bruneau, L., Beyaert, R., Hernández-Sebastià, C., & Marsolais, F. (2012). Arabidopsis mutants lacking asparaginases develop normally but exhibit enhanced root inhibition by exogenous asparagine. *Amino Acids*, *42*, 2307-2318.

Kanarek, N., Petrova, B., & Sabatini, D. M. (2020). Dietary modifications for enhanced cancer therapy. *Nature*, *579*(7800), 507–517.

Knott, S. R. V., Wagenblast, E., Khan, S., Kim, S. Y., Soto, M., Wagner, M., Turgeon, M. O., Fish, L., Erard, N., Gable, A. L., MacEli, A. R., DIckopf, S., Papachristou, E. K., D'Santos, C.

S., Carey, L. A., Wilkinson, J. E., Harrell, J. C., Perou, C. M., Goodarzi, H., … Hannon, G. J. (2018). Asparagine bioavailability governs metastasis in a model of breast cancer. *Nature*, *554*(7692), 378-381.

Krall, A. S., Xu, S., Graeber, T. G., Braas, D., & Christofk, H. R. (2016). Asparagine promotes cancer cell proliferation through use as an amino acid exchange factor. *Nature Communications*, *7*(1), 11457.

Lageix, S., Lanet, E., Pouch-Pélissier, M. N., Espagnol, M. C., Robaglia, C., Deragon, J. M., & Pélissier, T. (2008). Arabidopsis eIF2α kinase GCN2 is essential for growth in stress conditions and is activated by wounding. *BMC Plant Biology*, *8*(1), 1-9.

Lam, H., Hsieh, M., & Coruzzi, G. (1998). Reciprocal regulation of distinct asparagine synthetase genes by light and metabolites in *Arabidopsis thaliana*. *The Plant Journal*, *16*(3), 345–353.

Lam, H. M., Wong, P., Chan, H. K., Yam, K. M., Chen, L., Chow, C. M., & Coruzzi, G. M. (2003). Overexpression of the ASN1 gene enhances nitrogen status in seeds of Arabidopsis. *Plant Physiology*, *132*(2), 926–935.

Lastdrager, J., Hanson, J., & Smeekens, S. (2014). Sugar signals and the control of plant growth and development. *Journal of Experimental Botany*, *65*(3), 799-807.

Lea, P. J., Sodek, L., Parry, M. A. J., Shewry, R., & Halford, N. G. (2007). Asparagine in plants. *Annals of Applied Biology*, *150*(1), 1–26.

Li, M. W., Auyeung, W. K., & Lam, H. M. (2013). The GCN2 homologue in *Arabidopsis thaliana* interacts with uncharged tRNA and uses Arabidopsis eIF2α molecules as direct substrates. *Plant Biology*, *15*(1), 13-18.

Lomelino, C. L., Andring, J. T., McKenna, R., & Kilberg, M. S. (2017). Asparagine synthetase: Function, structure, and role in disease. *Journal of Biological Chemistry*, *292*(49), 19952-19958.

Luo, L., Qin, R. Y., Liu, T., Yu, M., Yang, T. W., & Xu, G. H. (2019). OsASN1 Plays a Critical Role in Asparagine-Dependent Rice Development. *International Journal of Molecular Sciences*, *20*(1).

Maaroufi-Dguimi, H., Debouba, M., Gaufichon, L., Clément, G., Gouia, H., Hajjaji, A., & Suzuki, A. (2011). An Arabidopsis mutant disrupted in ASN2 encoding asparagine synthetase 2 exhibits low salt stress tolerance. *Plant Physiology and Biochemistry*, *49*(6), 623–628.

Mair, A., Pedrotti, L., Wurzinger, B., Anrather, D., Simeunovic, A., Weiste, C., Valerio, C., Dietrich, K., Kirchler, T., Nägele, T., Carbajosa, J. V., Hanson, J., Baena-González, E., Chaban, C., Weckwerth, W., Dröge-Laser, W., & Teige, M. (2015). SnRK1-triggered switch of bZIP63 dimerization mediates the low-energy response in plants. *Elife*, *4*, e05828.

Martinek, P., Klem, K., Váňová, M., Bartáčková, V., Večerková, L., Bucher, P., & Hajšlová, J. (2009). Effects of nitrogen nutrition, fungicide treatment and wheat genotype on free asparagine and reducing sugars content as precursors of acrylamide formation in bread. *Plant, Soil and Environment*, *55*(5), 187-195.

Mascarenhas, C., Edwards-Ingram, L. C., Zeef, L., Shenton, D., Ashe, M. P., & Grant, C. M. (2008). Gcn4 is required for the response to peroxide stress in the yeast *Saccharomyces cerevisiae*. *Molecular Biology of the Cell*, *19*(7), 2995-3007.

Masclaux-Daubresse, C., Daniel-Vedele, F., Dechorgnat, J., Chardon, F., Gaufichon, L., & Suzuki, A. (2010). Nitrogen uptake, assimilation and remobilization in plants: Challenges for sustainable and productive agriculture. *Annals of Botany*, *105*(7), 1141-1157.

Mcallister, C. H., Beatty, P. H., & Good, A. G. (2012). Engineering nitrogen use efficient crop plants: The current status. *Plant Biotechnology Journal*, *10*(9), 1011-1025.

Mottram, D. S., Wedzicha, B. L., & Dodson, A. T. (2002). Food chemistry: Acrylamide is formed in the Maillard reaction. *Nature*, *419*(6906), 448-449.

Mukherjee, A., Ahmed, N., Rose, F. T., Ahmad, A. N., Javed, T. A., Wen, L., Bottino, R., Xiao, X., Kilberg, M. S., & Husain, S. Z. (2020). Asparagine synthetase is highly expressed at baseline in the pancreas through heightened PERK signaling. *Cellular and Molecular Gastroenterology and Hepatology*, *9*(1), 1-13.

Muttucumaru, N., Keys, A. J., Parry, M. A. J., Powers, S. J., & Halford, N. G. (2014). Photosynthetic assimilation of 14C into amino acids in potato (*Solanum tuberosum*) and asparagine in the tubers. *Planta*, *239*, 161-170.

Muttucumaru, N., Powers, S. J., Elmore, J. S., Mottram, D. S., & Halford, N. G. (2015). Effects of water availability on free amino acids, sugars, and acrylamide-forming potential in potato. *Journal of Agricultural and Food Chemistry*, *63*(9), 2566–2575.

Natarajan, K., Meyer, M. R., Jackson, B. M., Slade, D., Roberts, C., Hinnebusch, A. G., & Marton, M. J. (2001). Transcriptional profiling shows that GCN4p is a master regulator of gene expression during amino acid starvation in yeast. *Molecular and Cellular Biology*, *21*(13), 4347-4368.

Nerlov, C. (2007). The C/EBP family of transcription factors: a paradigm for interaction between gene expression and proliferation control. *Trends in Cell Biology*, *17*(7), 318-324.

Noree, C., Monfort, E., & Shotelersuk, V. (2018). Human asparagine synthetase associates with the mitotic spindle. *Biology Open*, *7*(12), bio038307.

Olea, F., Pérez-García, A., Cantón, F. R., Rivera, M. E., Cañas, R., Ávila, C., Cazorla, F. M., Cánovas, F. M., & de Vicente, A. (2004). Up-regulation and localization of asparagine synthetase in tomato leaves infected by the bacterial pathogen *Pseudomonas syringae*. *Plant and Cell Physiology*, *45*(6), 770-780.

Osuna, D., Gálvez-Valdivieso, G., Piedras, P., Pineda, M., & Aguilar, M. (2001). Cloning, characterization and mRNA expression analysis of PVAS1, a type I asparagine synthetase gene from *Phaseolus vulgaris*. *Planta*, *213*(3).

Perochon, A., Váry, Z., Malla, K. B., Halford, N. G., Paul, M. J., & Doohan, F. M. (2019). The wheat SnRK1α family and its contribution to Fusarium toxin tolerance. *Plant Science*, *288*, 110217.

Qu, C., Hao, B., Xu, X., Wang, Y., Yang, C., Xu, Z., & Liu, G. (2019). Functional research on three presumed asparagine synthetase family members in poplar. *Genes*, *10*(5), 326.

Raffan, S., & Halford, N. G. (2019). Acrylamide in food: Progress in and prospects for genetic and agronomic solutions. *Annals of Applied Biology*, *175*(3), 259-281.

Ramírez-González, R. H., Borrill, P., Lang, D., Harrington, S. A., Brinton, J., Venturini, L., Davey, M., Jacobs, J., van Ex, F., Pasha, A., Khedikar, Y., Robinson, S. J., Cory, A. T., Florio, T., Concia, L., Juery, C., Schoonbeek, H., Steuernagel, B., Xiang, D., … Uauy, C. (2018). The transcriptional landscape of polyploid wheat. *Science*, *361*(6403), eaar6089.

Rapp, M., Schwadorf, K., Leiser, W. L., Würschum, T., & Longin, C. F. H. (2018). Assessing the variation and genetic architecture of asparagine content in wheat: What can plant breeding contribute to a reduction in the acrylamide precursor? *Theoretical and Applied Genetics, 131*, 2427-2437.

Rashmi, D., Barvkar, V. T., Nadaf, A., Mundhe, S., & Kadoo, N. Y. (2019). Integrative omics analysis in *Pandanus odorifer* (Forssk.) Kuntze reveals the role of asparagine synthetase in salinity tolerance. *Scientific Reports*, *9*(1), 932.

Rodriguez, M., Parola, R., Andreola, S., Pereyra, C., & Martínez-Noël, G. (2019). TOR and SnRK1 signaling pathways in plant response to abiotic stresses: Do they always act according to the "yin-yang" model? *Plant Science*, *288*, 110220.

Ruberti, C., & Brandizzi, F. (2014). Conserved and plant-unique strategies for overcoming endoplasmic reticulum stress. *Frontiers in Plant Science*, 69.

Saia, S., Ruisi, P., Fileccia, V., di Miceli, G., Amato, G., & Martinelli, F. (2015). Metabolomics suggests that soil inoculation with arbuscular mycorrhizal fungi decreased free amino acid content in roots of durum wheat grown under N-limited, P-rich field conditions. *PLoS One*, *10*(6), e0129591.

Salvioli, A., Zouari, I., Chalot, M., & Bonfante, P. (2012). The arbuscular mycorrhizal status has an impact on the transcriptome profile and amino acid composition of tomato fruit. *BMC Plant Biology*, *12*, 1-12.

Seiffert, B., Zhou, Z., Wallbraun, M., Lohaus, G., & Möllers, C. (2004). Expression of a bacterial asparagine synthetase gene in oilseed rape (*Brassica napus*) and its effect on traits related to nitrogen efficiency. *Physiologia Plantarum*, *121*(4), 656-665.

Shi, L., Twary, S. N., Yoshioka, H., Gregerson, R. G., Miller, S. S., Samac, D. A., Gantt, J. S., Unkefer, P. J., & Vance, C. P. (1997). Nitrogen assimilation in alfalfa: Isolation and characterization of an asparagine synthetase gene showing enhanced expression in root nodules and dark-adapted leaves. *The Plant Cell*, *9*(8), 1339-1356.

Shi, L., Wu, Y., & Sheen, J. (2018). TOR signaling in plants: conservation and innovation. *Development*, *145*(13), dev160887.

Sieciechowicz, K. A., & Ireland, R. J. (1989). Isolation and properties of an asparaginase from leaves of *Pisum sativum*. *Phytochemistry*, *28*(9), 2275-2279.

Sieciechowicz, K. A., Ireland, R. J., & Joy, K. W. (1988). Diurnal changes in asparaginase activity in pea leaves: II. Regulation of activity. *Journal of Experimental Botany*, *39*(6), 707-721.

Sieciechowicz, K. A., Joy, K. W., & Ireland, R. J. (1988a). Diurnal changes in asparaginase activity in pea leaves: I. The requirement for light for increased activity. *Journal of Experimental Botany*, *39*(6), 695-706.

Sieciechowicz, K. A., Joy, K. W., & Ireland, R. J. (1988b). The metabolism of asparagine in plants. *Phytochemistry*, *27*(3), 663-671.

Sikalidis, A. K., & Stipanuk, M. H. (2010). Growing rats respond to a sulfur amino acid–deficient diet by phosphorylation of the α subunit of eukaryotic initiation factor 2 heterotrimeric complex and induction of adaptive components of the integrated stress response. *The Journal of Nutrition*, *140*(6), 1080-1085.

Siu, F., Bain, P. J., Leblanc-Chaffin, R., Chen, H., & Kilberg, M. S. (2002). ATF4 is a mediator of the nutrient-sensing response pathway that activates the human asparagine synthetase gene. *Journal of Biological Chemistry*, *277*(27), 24120-24127.

Soares, E., Shumbe, L., Dauchot, N., Notte, C., Prouin, C., Maudoux, O., & Vanderschuren, H. (2020). Asparagine accumulation in chicory storage roots is controlled by translocation and feedback regulation of asparagine biosynthesis in leaves. *New Phytologist*, *228*(3), 922-931.

Sodek, L., Lea, P. J., & Miflin, B. J. (1980). Distribution and properties of a potassium-dependent asparaginase isolated from developing seeds of *Pisum sativum* and other plants. *Plant Physiology*, *65*(1), 22-26.

Sormani, R., Delannoy, E., Lageix, S., Bitton, F., Lanet, E., Saez-Vasquez, J., Deragon, J. M., Renou, J. P., & Robaglia, C. (2011). Sublethal cadmium intoxication in *Arabidopsis thaliana* impacts translation at multiple levels. *Plant and Cell Physiology*, *52*(2), 436-447.

Stadler, R. H., Blank, I., Varga, N., Robert, F., Hau, J., Guy, P. A., Robert, M. C., & Riediker, S. (2002). Acrylamide from Maillard reaction products. *Nature*, *419*(6906), 449-450.

Stewart, G. R., & Larher, F. (1980). Accumulation of amino acids and related compounds in relation to environmental stress. In *Amino Acids and Derivatives* (pp. 609-635). Academic Press.

Stockmann, F., Weber, E. A., Mast, B., Schreiter, P., Merkt, N., Claupein, W., & Graeff-Hönninger, S. (2019). Acrylamide-formation potential of cereals: What role does the agronomic management system play? *Agronomy*, *9*(10), 584.

Stockmann, F., Weber, E. A., Schreiter, P., Merkt, N., Claupein, W., & Graeff-Hönninger, S. (2018). Impact of nitrogen and sulfur supply on the potential of acrylamide formation in organically and conventionally grown winter wheat. *Agronomy*, *8*(12), 284.

Su, N., & Kilberg, M. S. (2008). C/EBP homology protein (CHOP) interacts with activating transcription factor 4 (ATF4) and negatively regulates the stress-dependent induction of the asparagine synthetase gene. *Journal of Biological Chemistry*, *283*(50), 35106-35117.

Swann, O. G., Kilpatrick, M., Breslin, M., & Oddy, W. H. (2019). Dietary fiber and its associations with depression and inflammation. *Nutrition Reviews*, *78*(5), 394-411.

Szabados, L., & Savouré, A. (2010). Proline: a multifunctional amino acid. *Trends in Plant Science*, *15*(2), 89-97.

Taeymans, D., Wood, J., Ashby, P., Blank, I., Studer, A., Stadler, R. H., Gondé, P., van Eijck, P., Lalljie, S., Lingnert, H., Lindblom, M., Matissek, R., Müller, D., Tallmadge, D., O'Brien, J., Thompson, S., Silvani, D., & Whitmore, T. (2004). A review of acrylamide: An industry perspective on research, analysis, formation, and control. *Critical Reviews in Food Science and Nutrition*, *44*(5), 323-347.

Thiaville, M. M., Dudenhausen, E. E., Zhong, C., Pan, Y. X., & Kilberg, M. S. (2008). Deprivation of protein or amino acid induces C/EBPβ synthesis and binding to amino acid response elements, but its action is not an absolute requirement for enhanced transcription. *Biochemical Journal*, *410*(3), 473-484.

Tonin, G. S., & Sodek, L. (1990). Asparaginase, allantoinase and glutamine synthetase activities in soybean cotyledons grown *in vitro*. *Phytochemistry*, *29*(9), 2829-2831.

USDA-APHIS. (2014). *JR Simplot company petition (14–093-01p) for determination of nonregulated status for Innate^{TM} potatoes with late blight resistance, low acrylamide potential, reduced black spot and lowered reducing sugars: Russet Burbank event W8*.

Wang, H. B., Liu, D. C., Sun, J. Z., & Zhang, A. M. (2005). Asparagine synthetase gene TaASN1 from wheat is up-regulated by salt stress, osmotic stress and ABA. *Journal of Plant Physiology*, *162*(1), 81–89.

Waterhouse, R. N., Smyth, A., Massonneau, A., Presser, L. M., & Clarkson, D. T. (1996). Molecular cloning and characterisation of asparagine synthetase from *Lotus japonicus*: Dynamics of asparagine synthesis in N-sufficient conditions. *Plant Molecular Biology, 30,* 883-897.

Weber, E. A., Graeff, S., Koller, W. D., Hermann, W., Merkt, N., & Claupein, W. (2008). Impact of nitrogen amount and timing on the potential of acrylamide formation in winter wheat (*Triticum aestivum* L.). *Field Crops Research*, *106*(1), 44-52.

Wek, R. C., Jiang, H. Y., & Anthony, T. G. (2006). Coping with stress: EIF2 kinases and translational control. *Biochemical Society Transactions*, *34*(1), 7-11.

Whiteside, M. D., Garcia, M. O., & Treseder, K. K. (2012). Amino acid uptake in arbuscular mycorrhizal plants. *PLoS ONE, 7*(10), e47643.

Wurzinger, B., Nukarinen, E., Nägele, T., Weckwerth, W., & Teige, M. (2018). The snrk1 kinase as central mediator of energy signaling between different organelles. *Plant Physiology*, *176*(2), 1085-1094.

Xu, H. W., Curtis, T. Y., Powers, S. J., Raffan, S., Gao, R. H., Huang, J. H., Heiner, M., Gilbert, D. R., & Halford, N. G. (2018). Genomic, biochemical, and modeling analyses of asparagine synthetases from wheat. *Frontiers in Plant Science*, *8*.

Yadav, A. K., Carroll, A. J., Estavillo, G. M., Rebetzke, G. J., & Pogson, B. J. (2019). Wheat drought tolerance in the field is predicted by amino acid responses to glasshouse-imposed drought. *Journal of Experimental Botany*, *70*(18), 4931-4948.

Yang, R., Wek, S. A., & Wek, R. C. (2000). Glucose limitation induces GCN4 translation by activation of GCN2 protein kinase. *Molecular and Cellular Biology, 20*(8), 2706-2717.

Zhang, J., Fan, J., Venneti, S., Cross, J. R., Takagi, T., Bhinder, B., Djaballah, H., Kanai, M., Cheng, E. H., Judkins, A. R., Pawel, B., Baggs, J., Cherry, S., Rabinowitz, J. D., & Thompson, C. B. (2014). Asparagine plays a critical role in regulating cellular adaptation to glutamine depletion. *Molecular Cell*, *56*(2), 205-218.

Zhang, Y., Andralojc, P. J., Hey, S. J., Primavesi, L. F., Specht, M., Koehler, J., Parry, M. A. J., & Halford, N. G. (2008). Arabidopsis sucrose non-fermenting-1-related protein kinase-1 and calcium-dependent protein kinase phosphorylate conserved target sites in ABA response element binding proteins. *Annals of Applied Biology*, *153*(3), 401-409.

Zhang, Y., Dickinson, J. R., Paul, M. J., & Halford, N. G. (2003). Molecular cloning of an arabidopsis homologue of GCN2, a protein kinase involved in co-ordinated response to amino acid starvation. *Planta, 217*, 668-675.

Zhang, Y., Primavesi, L. F., Jhurreea, D., Andralojc, P. J., Mitchell, R. A. C., Powers, S. J., Schluepmann, H., Delatte, T., Wingler, A., & Paul, M. J. (2009). Inhibition of SNF1-related protein kinasel activity and regulation of metabolic pathways by trehalose-6-phosphate1. *Plant Physiology*, *149*(4), 1860-1871.

Zhang, Y., Wang, Y., Kanyuka, K., Parry, M. A. J., Powers, S. J., & Halford, N. G. (2008). GCN2-dependent phosphorylation of eukaryotic translation initiation factor-2α in Arabidopsis. *Journal of Experimental Botany*, *59*(11), 3131-3141.

Zhao, F. J., Hawkesford, M. J., & McGrath, S. P. (1999). Sulphur assimilation and effects on yield and quality of wheat. *Journal of Cereal Science*, *30*(1), 1-17.

Zheng, T., Qi, P. F., Cao, Y. L., Han, Y. N., Ma, H. L., Guo, Z. R., Wang, Y., Qiao, Y. Y., Hua, S. Y., Yu, H. Y., Wang, J. P., Zhu, J., Zhou, C. Y., Zhang, Y. Z., Chen, Q., Kong, L., Wang, J. R., Jiang, Q. T., Yan, Z. H., … Zheng, Y. L. (2018). Mechanisms of wheat (*Triticum aestivum*) grain storage proteins in response to nitrogen application and its impacts on processing quality. *Scientific Reports*, *8*(1), 11928.

3

Understanding the relationships between free asparagine in grain and other traits to breed low-asparagine wheat

Joseph Oddy[1], Sarah Raffan[1], Mark D. Wilkinson[1], J. Stephen Elmore[2] and Nigel G. Halford[1]

1. *Plant Sciences Department, Rothamsted Research, Harpenden AL5 2JQ, UK.*
2. *Department of Food and Nutritional Sciences, University of Reading, Whiteknights, PO Box 226, Reading RG6 6AP, UK.*

Plants, 11(5), 669 (2022).

This chapter has been modified from its published form to allow for edits and corrections within this Thesis.

3.1. Introduction to paper

Following on from the review of free asparagine and its relationship with stress, I aimed to summarise the relationships between free asparagine and other traits. In order for low free asparagine to become a breeding target, it needs to be shown that it does not impact negatively on any desirable characteristics. Many studies have measured free asparagine in wheat and how it correlates with other traits, but it was not clear whether the patterns observed in these studies were consistent or environmentally dependent. Consequently, the aims of this review were threefold: (1) assess the relationships between free asparagine and quality traits, (2) assess the relationships between free asparagine and agronomic traits, and finally (3) evaluate how breeding for low free asparagine wheat can be achieved.

3.2. Introduction

Wheat is one of the world's most important crops, contributing an estimated 18.6% to global daily calorie intake and 19.8% to global daily protein intake in 2018 (FAO, 2021). The contribution of wheat to daily calorie and protein intake varies substantially by region, with certain regions having greater dependence on wheat than others. For example, the contribution of wheat to daily calorie and protein intake was approximately double the global average at 39.1% and 38.4%, respectively, in Central Asia in 2018 (Kazakhstan, Kyrgyzstan, Tajikistan, Turkmenistan, and Uzbekistan (as described by the Food and Agriculture Organisation of the United Nations) (FAO, 2021)). Consequently, it is essential to ensure that the supply and quality of wheat is safeguarded against emerging challenges. This can be achieved by the development of new crop-management strategies and crop protection products, or through the breeding of new varieties.

Wheat breeding and research has been greatly facilitated in recent years by the sequencing of multiple wheat genomes (Walkowiak et al., 2020), the development of numerous marker technologies (Wilkinson et al., 2020), and the use of new gene editing technologies (Smedley et al., 2021). With the development of these technologies, we can begin to investigate and improve traits that may have been prohibitively costly or time consuming to improve in the past, and the free amino acid composition of wheat grain is one such trait.

The free (soluble, non-protein) amino acid content of wheat grain has been of most interest to wheat geneticists in recent years because of the food safety issues associated with free

asparagine, the precursor to the 'probably carcinogenic' processing contaminant, acrylamide (Tareke et al., 2002). Free asparagine reacts with reducing sugars to form acrylamide (Mottram et al., 2002; Stadler et al., 2002), but free asparagine concentration has been shown to be the major determinant of acrylamide concentration in wheat products in several studies (see Raffan & Halford (2019) for review). Halford et al. (2007), for example, used data generated by Muttucumaru et al. (2006), to plot free asparagine concentration against acrylamide formation in wheat flour heated for 20 min at 160 or 180 °C, and obtained coefficients of determination (R^2) of 0.956 and 0.998 for pot- and field-grown plants, respectively. In contrast, there was no relationship between the concentration of reducing sugars and the amount of acrylamide that formed. Muttucumaru et al. (2006) had shown sulphur deficiency to cause very high concentrations of free asparagine to accumulate in wheat grain, and Granvogl et al. (2007) obtained very similar results, with acrylamide formation closely related to free asparagine concentration, except in flours from extremely sulphur-deprived plants, in which free asparagine concentration was so high that it was no longer limiting.

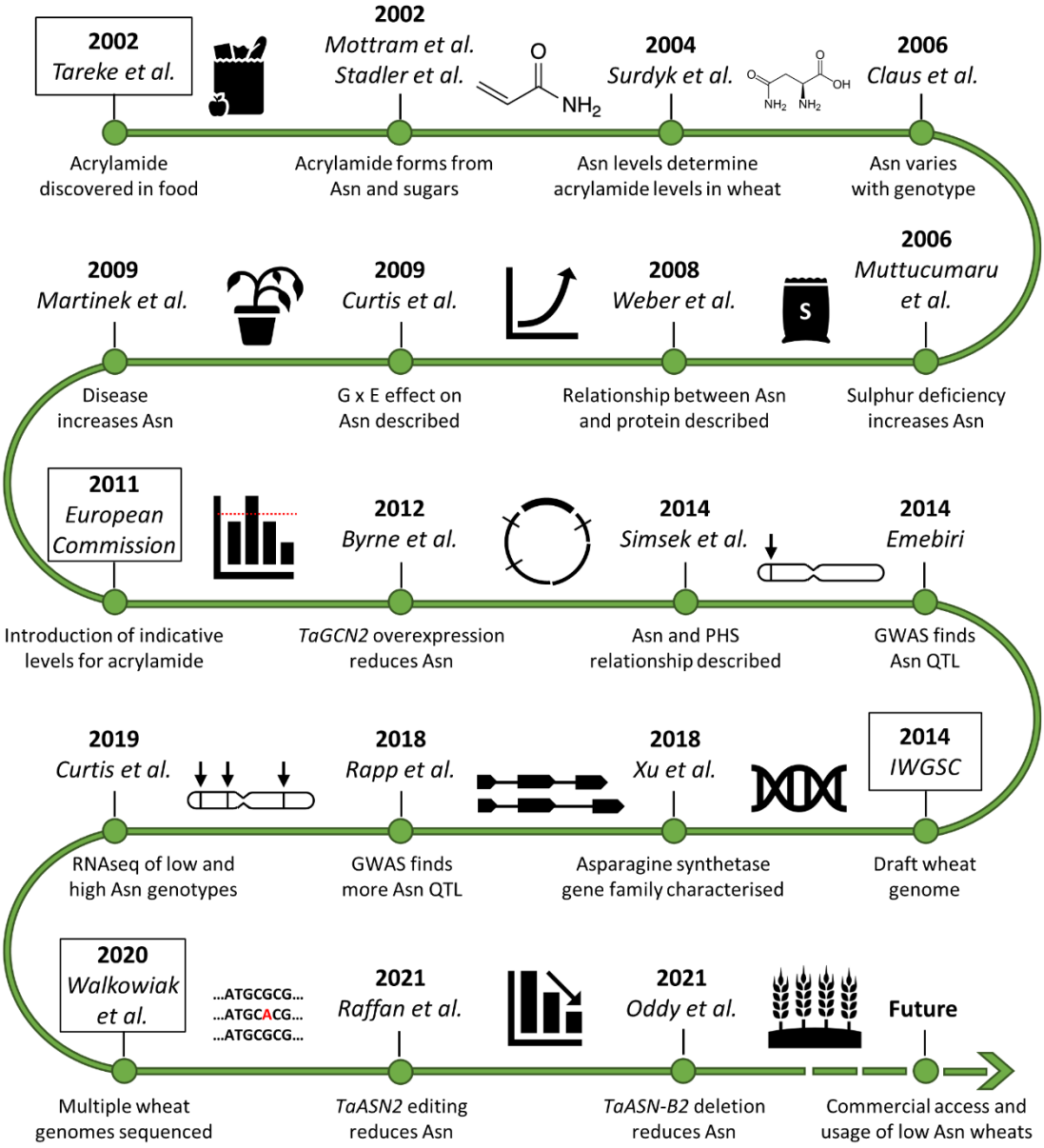


Figure 3.1. Timeline of asparagine research in wheat since the discovery of acrylamide in food. Asn (asparagine), IWGSC (International Wheat Genome Sequencing Consortium), G × E (genotype-by-environment interaction), GWAS (genome-wide association study), PHS (pre-harvest sprouting), QTL (quantitative trait locus/loci), RNA-seq (RNA sequencing), *TaASN2* (Asparagine synthetase 2), *TaGCN2* (general control nonderepressible 2).

Curtis et al. (2009) also studied the effect of sulphur deficiency and showed acrylamide formation to rise with free asparagine concentration (R^2 = 0.9945), up to free asparagine

concentrations of approximately 25 mmol/kg or higher, something only seen in flour from extremely sulphur-deprived plants.

Acrylamide also forms in potato products and the relationship between free asparagine and reducing sugar concentration and acrylamide formation for potato is very different. Potato tubers have higher concentrations of free asparagine than cereal grains and reducing sugar concentration is usually the limiting factor for acrylamide formation, although free asparagine concentration does contribute to the variance in some datasets (Muttucumaru et al., 2014). Nevertheless, the clear relationship between free asparagine concentration and acrylamide formation in wheat products means that strategies to control acrylamide formation in wheat-based foods over the last 20 years have targeted free asparagine (Figure 3.1).

Although we now better understand the environmental and genetic factors that influence grain free asparagine content, there are still many unanswered questions around how these factors interact and how they relate to other traits. Here, we summarise some of the research regarding the relationship between free asparagine concentration in the grain and other traits, and how genetic improvements might be made using this information.

3.3. Relationships between free asparagine, quality and agronomic traits

3.3.1. Free asparagine concentration and quality traits

Quality traits in wheat are those that impact the functionality of the end product (i.e., the baking and nutritional quality of the grain), so encompass traits such as pre-harvest sprouting (PHS), protein content and hardness. Grain free-asparagine content has sometimes been found to correlate with some of these quality traits, but this differs greatly between studies (Table 3.1 and Table 3.2). Few quality traits have been tested for a relationship with free asparagine in more than one study, and those that have often show different relationships across studies (Table 3.1), implying that free asparagine concentration is unlikely to correlate strongly with quality traits.

Malunga et al. (2019) undertook the largest study of free asparagine in relation to quality traits, screening 42 quality traits and assessing their relationship with free asparagine, in both wholemeal and white flours. This analysis revealed that free asparagine in wholemeal samples did not correlate with any quality parameters, except for a weak correlation ($r = -0.389$, $p = 0.0339$) with the extensograph A parameter. Similarly, free asparagine in white flour only

correlated weakly with extensograph R_{max} ($r = -0.370$, $p = 0.0444$), extensograph A ($r = -0.378$, $p = 0.0394$) and water dough colour b^* parameters ($r = 0.373$, $p = 0.0426$). Corol et al. (2016) also performed correlation analyses of free asparagine with quality traits and did find some weak associations, but these have not been corroborated by further studies (Table 3.1).

One potentially interesting relationship is that between free asparagine and PHS, because of the potential for protein hydrolysis during PHS to release free asparagine. PHS negatively impacts wheat quality in a range of ways, reducing flour yield, the quality of baked products, and nutrient content (Simsek et al., 2014b). Simsek et al. (2014a) reported a moderately strong ($r = 0.6$–0.7) positive correlation between free asparagine, sprouting score, and endoprotease activity in samples of sprouted wheat grain, suggesting that there was a relationship between asparagine and PHS at high levels of sprouting. Additionally, in a study designed to render the asparagine synthetase 2 genes (*TaASN2*) non-functional through gene editing, Raffan et al. (2021) observed a poor germination phenotype that could be rescued through exogenous application of asparagine to the soil, implying that low grain free asparagine content may inhibit germination and could perhaps also affect PHS. Further research is required to confirm the germination phenotype, but asparagine synthetases are known to play important roles in germination in other species (Herrera-Rodríguez et al., 2006; Canales et al., 2012). No correlation has been observed to date between free asparagine content and Hagberg falling number (HFN) (Table 3.1), which is indicative of α-amylase activity and, therefore, PHS. However, it is possible that a relationship between grain free asparagine content, germination and PHS could exist when free asparagine concentration is very low (e.g., in *TaASN2* edited lines) or very high (e.g., in artificially sprouted wheat samples).

Table 3.1. Association between free asparagine and selected quality traits. Asn (asparagine), HFN (Hagberg falling number), Z-SDS (Zeleny sedimentation index).

Asn Measurement	Sample type	Trait	*r*	*p*	Reference
Log_e transformation	White flour	Farinograph absorption	0.94	<0.001	Liu et al. (2011)
		Nitrogen to sulphur ratio	0.73	<0.01	
		Nitrogen content	0.62	<0.05	
Log_e transformation	Wholemeal flour	Sprouting score	0.68	<0.001	Simsek et al. (2014a)
		Endoprotease activity (sprouted)	0.69	<0.001	
		Endoprotease activity (ΔD)	0.60	<0.01	
Untransformed	Wholemeal flour	HFN	0.07	0.39	Corol et al. (2016)
		Z-SDS	0.37	<0.001	
		Gluten content	0.44	<0.001	
		Starch content	−0.32	<0.001	
		Water absorption	0.35	<0.001	
		Hardness index	0.03	0.68	
Log_e transformation	Wholemeal flour	Absorption	−0.03	>0.05	Ohm et al. (2017)
Untransformed	Wholemeal flour	Hardness index	0.15	>0.05	Navrotskyi et al. (2018)
Log_{10} back-transformed	Wholemeal flour	Sulphur grain content	0.14	>0.05	Rapp et al. (2018)
		HFN	0.03	>0.05	
		Z-SDS	−0.29	<0.001	
Untransformed	Wholemeal flour	HFN	−0.17	0.36	Malunga et al. (2019)
		Gluten index	−0.36	<0.05	
		Flour starch damage	−0.18	0.33	
		Farinograph absorption	−0.12	0.5436	

In contrast to other quality traits, the relationship between grain free asparagine content and protein content has been tested numerous times and the results suggest that there is a positive correlation between the two traits, varying from weak to strong, under different conditions (Table 3.2). The protein content of wheat is important both for breadmaking functionality and for its nutritional quality, especially as the global agricultural system shifts towards the cultivation of more plant protein for sustainability reasons. A more detailed analysis of the relationship between protein content and free asparagine content was undertaken by Simsek et al. (2014a), who found significant positive associations between free asparagine content and extractable F4 (albumin/globulin), F5, and F6 (hydrolysed polymeric/non-gluten protein) HPLC protein fractions. This is consistent with the release of free asparagine from the hydrolysis of proteins under PHS. Simsek et al. (2014a) also found significant negative associations between asparagine and unextractable F1 (HMW glutenin polymers) and F2 (LMW glutenin polymers) protein fractions. This was further supported by Ohm et al. (2018), where significant negative ($p < 0.05$) genotypic and phenotypic correlations were found between free asparagine and unextractable F1 protein fractions, but not between free asparagine and extractable F1 fractions.

The contrasting relationships between free asparagine and the different protein fractions has interesting implications for quality, because the ratio of unextractable HMW polymeric proteins to extractable LMW polymeric proteins is a better determinant of quality than total protein measurements (Ohm et al., 2009; Tsilo et al., 2010). Consequently, lower free asparagine content in the grain may be associated with higher bread-making quality. This conclusion was drawn by Ohm et al. (2018), who further suggested that measurements of unextractable polymeric protein may allow for selection of varieties that simultaneously have high-quality bread-making potential and are low in free asparagine content. Such a correlation between free asparagine and bread-making quality has not been consistently observed across studies (see Table 3.1), so the relationship is probably more complex than this. Higher protein content may also be desirable, independent of its effect on bread-making quality.

The complexity of the factors determining protein content and the free asparagine content of grain can be illustrated by looking at soft wheat varieties. These varieties typically have lower protein content than hard wheats, making them unsuitable for bread-making but suitable for biscuits, breakfast cereals, pastries and other baked goods. Based on this, the grain of soft wheats might be expected to have lower free asparagine content than hard wheats, due to the positive correlations often found with protein (Table 3.2). Curtis et al. (2018) did show that

varieties with consistently-low free-asparagine concentration were often soft wheats, but the difference between hard and soft variety groups was not significant, with high and low-asparagine varieties in both groups. It is possible that the association of some soft wheat varieties with consistently low free asparagine content was due to the deletion of one of the asparagine synthetase 2 homeologues, *TaASN-B2*, which has been shown to be associated with lower grain free asparagine content and was more common in the soft wheats used in the trial (Oddy et al., 2021). The effect of this deletion is only apparent when the plants have adequate sulphur, with the effect being overwhelmed by the huge increase in free asparagine concentration that occurs under sulphur deficiency, adding more complexity to the control of grain free asparagine content.

Table 3.2. Associations between free asparagine content and protein content. * These values refer to R^2 values, not *r* values. r_p (phenotypic correlation), r_g (genotypic correlation), UN (untreated control), T (intensive fungicide treatment), N (agronomic intensity experiments).

Asparagine Measure	**Sample type**	**Protein Measure**	**R^2/*r***	***p***	**Reference**
Untransformed	Wholemeal flour	Crude protein	0.86 *	<0.001	Weber et al. (2008)
Untransformed	Flour (undetermined)	Protein content (2006 UN)	0.93	<0.01	Martinek et al. (2009)
		Protein content (2006 T)	0.63	<0.05	
		Protein content (2007 UN)	0.75	>0.05	
		Protein content (2007 T)	0.27	>0.05	
		Protein content (2006 N)	0.73	<0.01	
		Protein content (2007 N)	0.89	<0.01	
Log_e transformation	Wholemeal flour	Protein content (non-sprouted)	NA	>0.05	Simsek et al. (2014a)
		Protein content (sprouted)	NA	>0.05	
		Protein content (ΔD)	NA	>0.05	
Untransformed	Grain	Protein content	0.45	<0.001	Corol et al. (2016)

	Wholemeal flour	Protein content	0.51	<0.001	
	White flour	Protein content	0.38	<0.001	
Log_e transformation	Wholemeal flour	Protein content	0.43	<0.001	Ohm et al. (2017)
Log_e transformation	Wholemeal flour	Protein content (r_p)	−0.03	>0.05	Ohm et al. (2018)
		Protein content (r_g)	−0.37	>0.05	
Untransformed	Wholemeal flour	Protein content	0.52	<0.01	Navrotskyi et al. (2018)
Log_{10} back transformed	Wholemeal flour	Protein content	0.23	<0.01	Rapp et al. (2018)
Untransformed	White flour	Crude protein	0.36 *	NA	Stockmann et al. (2018)
Untransformed	White flour	Crude protein	0.04 *	NA	Stockmann et al. (2019)
Untransformed	Wholemeal flour	Protein content	−0.08	0.66	Malunga et al. (2019)
	White flour	Protein content	−0.14	0.46	

Although the relationship between free asparagine content and the protein composition of grain is complex, there are two factors that are well known to affect both: nitrogen and sulphur fertilisers. Nitrogen application increases both the free asparagine content and protein content of grain, whereas sulphur application decreases free asparagine content and improves protein composition (see Oddy et al. (2020) for review). This is reflected in the correlations between free asparagine and nitrogen, and between free asparagine and the nitrogen to sulphur ratio in wheat grain (Liu et al., 2011) (Table 3.1), and implies that wheat uses free asparagine as a nitrogen store in the grain when sulphur is limiting (reviewed in Raffan & Halford (2019)). Application of more sulphur is, therefore, desirable for both traits, except for its environmental pollution effects (Hinckley et al., 2020), whereas a balance between higher protein/higher free asparagine and lower protein/lower free asparagine must be struck when it comes to nitrogen application. Similar trade-offs arise because of the association of nitrogen with desirable agronomic traits, but there may be solutions in breeding, as discussed below. In the meantime, our advice is that nitrogen application should be accompanied with sufficient sulphur (typically 20 kg sulphur per hectare) to prevent the nitrogen ending up as free asparagine instead of protein.

3.3.2. Free asparagine and agronomic traits

As a result of the positive association between free asparagine content and nitrogen application, it might be expected that there would be a similar association between free asparagine content and traits related to growth because of the positive relationship between plant growth and nitrogen. Positive correlations between free asparagine content and yield have indeed been found (Table 3.3) but, perhaps surprisingly, these correlations have not been consistent across studies. Xie et al. (2021), for example, found that free asparagine content (measured in milligrams per gram of protein) was negatively correlated with grain yield in one year when the yield was low (between two and four tonnes per hectare), but positively associated in another year, when the yield was higher (between four and eight tonnes per hectare), suggesting a non-linear relationship. A reduction in plant stress could explain the negative correlation observed over lower yield values, whilst the positive correlation could be due to greater nitrogen availability in the soil. However, the authors note that the relationship between absolute free asparagine content (measured without normalisation to protein) and yield was not as strong as the relationship when the normalisation of free asparagine to protein was performed. The lack of comprehensive yield/free asparagine studies does not provide strong support for hypotheses linking the two traits, but it could be worthwhile investigating the nature of the relationship between free asparagine content and yield in more detail in future studies.

Another interesting correlation shown in Table 3.3 is that between the asparagine response (measured as the ratio of asparagine in treated vs. asparagine in untreated plants) and the yield gap-based measure of drought tolerance (YDT), as studied by Yadav et al. (2019). YDT provides a measurement of how well a variety performs under drought stress relative to unstressed conditions. The negative correlation between the asparagine response and YDT in the study indicated that plants that were less tolerant to drought tended to accumulate more asparagine. This relationship is consistent with the general observation that free asparagine accumulates under stress (reviewed in Oddy et al. (2020)), and Yadav et al. (2019) suggested that the relationship could be caused by the remobilisation of nitrogen during stress-induced senescence. Curtis et al. (2018) also showed that asparagine metabolism is affected by drought stress in wheat, by constructing a detailed network describing the genes and other factors involved, using a Unique Network Identification Pipeline to show the inter-relationships between genes that changed in expression in response to drought stress, in both leaves and roots.

The relationship between free asparagine and senescence in wheat is not well understood, but Emebiri (2014) did find a negative correlation between asparagine and flowering time (Table 3), which may reflect an association between senescence and asparagine. Senescence is known to cause the remobilisation of nitrogen via asparagine and the activation of asparagine synthetases in other species, including sunflower, tobacco, and barley (Herrera-Rodríguez et al., 2006; Masclaux-Daubresse et al., 2006; Avila-Ospina et al., 2015; Bovet et al., 2019), and early-senescing barley lines show greater expression of asparagine synthetase in senescing tissues relative to later senescing lines (Jukanti et al., 2008). Navrotskyi et al. (2018) also found a positive correlation between free asparagine and the number of days until harvest (Table 3.3), again implying that longer periods of senescence might be responsible for this association.

Table 3.3. Associations between free asparagine content and agronomic measurements. TKW (thousand kernel weight), HH (as measured from heading to harvest date), HLW (hectolitre weight), YDT (yield gap-based drought tolerance).

Asparagine Measure	**Agronomic Measure**	***r***	***p***	**Reference**
Log_e back-transformed	Flowering time	−0.67	<0.001	Emebiri (2014)
Untransformed	Plant height	0.41	<0.001	Corol et al. (2016)
	TKW	0.03	0.75	
	Mean kernel diameter	0.13	0.11	
	Mean kernel weight	0.06	0.45	
	Yield	−0.14	0.09	
	Precipitation (HH)	−0.85	<0.05	
	Temperature (HH)	0.74	0.10	
Log_e transformation	HLW	−0.40	<0.001	Ohm et al. (2017)
Untransformed	Mean kernel diameter	0.37	<0.05	Navrotskyi et al. (2018)
	Mean kernel weight	0.37	<0.05	
	Yield	−0.32	>0.05	
	Days from flowering to harvest	0.61	<0.001	

Log_{10} back transformed	TKW	−0.24	<0.01	Rapp et al. (2018)
	HLW	−0.21	<0.01	
Untransformed	Nitrogen application	0.63	NA	Stockmann et al. (2018)
Untransformed	TKW	−0.27	0.15	Malunga et al. (2019)
	HLW	−0.07	0.71	
Log_e transformed responses	YDT	−0.73	<0.05	Yadav et al. (2019)
Per unit protein	Yield (2018)	0.74	NA	Xie et al. (2021)
	Yield (2019)	−0.56	NA	
Untransformed	Yield	0.75	<0.001	Malunga et al. (2021)

Further research in this area could be greatly facilitated by investigating free asparagine accumulation in stay-green varieties of wheat. These varieties show delayed senescence, leading to a prolonged green phenotype, and generally have higher yields and better stress tolerance, although this leads to a trade-off with protein and micronutrient content due to the yield dilution effect (Distelfeld et al., 2014). Heyneke et al. (2019) undertook an experiment comparing the leaf metabolome of early and late senescing wheat lines and found that free asparagine content in the leaf decreased as senescence progressed, but this was not statistically significant. The ratio of asparagine to aspartic acid (as well as the ratio of glutamine to glutamic acid) did increase significantly, in both early- and late- senescing lines, as senescence progressed. The authors of this study interpret the increase in nitrogen-rich amino acids (asparagine and glutamine) relative to their precursors (aspartic acid and glutamic acid, respectively) as being indicative of nitrogen remobilisation to other active organs. The remobilisation of free asparagine from senescing leaves to developing grain may, therefore, be a mechanism which connects senescence and grain free-asparagine content. Further investigation of early- and late- senescing lines should be undertaken to shed more light on the relationship between free asparagine and senescence.

The development of stay-green varieties is part of a larger effort to develop varieties with better nitrogen-use efficiency (NUE), in order to reduce agricultural inputs, since nitrogen fertilisers are a major source of environmental pollution (Cameron et al., 2013). NUE can be defined in

many different ways but is commonly described as a productivity index measuring yield per unit of nitrogen (see Hawkesford & Riche (2020) for review). Strategies to enhance NUE in wheat may impact grain free asparagine content due to effects on nitrogen uptake and partitioning within the plant, especially those methods that modulate genes involved in amino acid synthesis and transport.

For example, Wan et al. (2021) reported that overexpression of the starchy endosperm amino acid transporter, *TaAAP13*, in endosperm tissue increased grain size and grain weight, but decreased grain yield and seed number per plant overall, as well as increasing free asparagine content. In another study, Tiong et al. (2021) transformed rice, wheat, and barley plants with a stress-inducible barley alanine aminotransferase, OsAnt1:HvAlaAT, resulting in increased grain yield for some of the resulting lines. The authors also showed that free asparagine content in the roots and shoots of the mutant rice plants was lower relative to wild-type plants, but these measurements were not repeated in the mutant wheat plants and grain free asparagine content was not measured. Hu et al. (2018) also improved NUE using modulated amino acid transporters, but this time by overexpressing an isoform of glutamine synthetase 2. This increased grain yield under conditions of both high and low nitrogen, but total amino acid and glutamine content (with which asparagine is often strongly positively correlated) only increased significantly under high nitrogen conditions. These examples show that the many different strategies for improving NUE are likely to have different effects on grain free asparagine content, based on their individual mechanisms, and in combination with different environments and management practices. It is, therefore, important that effects on free asparagine concentration (and, therefore, acrylamide-forming potential) are assessed in plants in which NUE has been improved.

The relationship between free asparagine content, quality, and agronomic traits, as described above, is summarised in Figure 3.2, below.

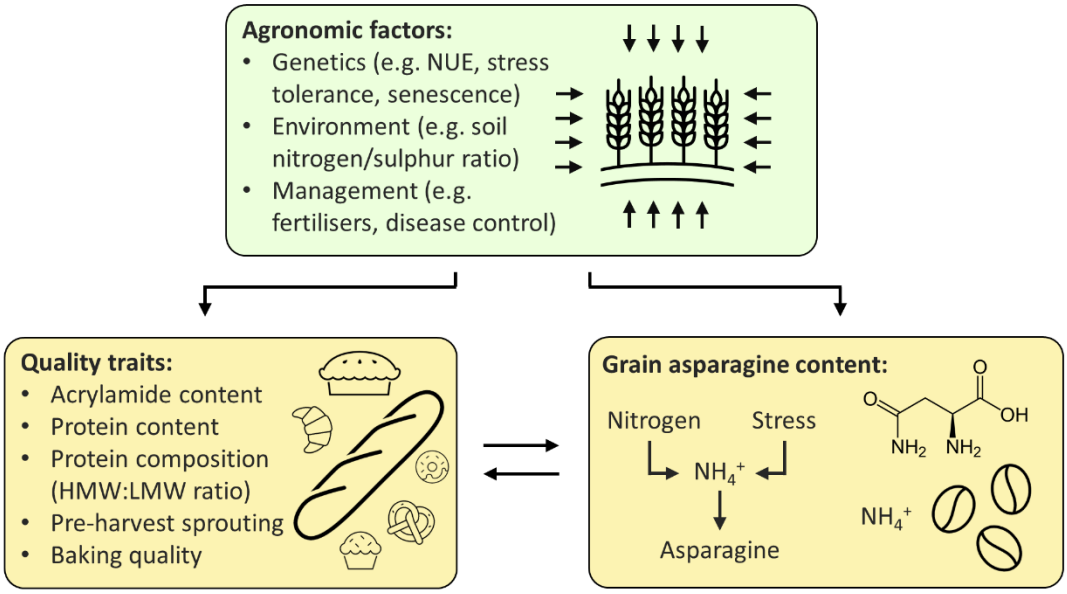


Figure 3.2. Proposed relationship between agronomic factors, quality traits, and grain free asparagine content. Agronomic factors influence both quality traits and grain free asparagine content, whilst quality traits and grain asparagine are linked to one another. For example, pre-harvest sprouting can influence free asparagine content, whilst free asparagine content determines acrylamide content and colour. NUE (nitrogen-use efficiency), HMW (high molecular weight), LMW (low molecular weight).

3.4. Breeding wheat with low free asparagine content

Selection for desirable traits in wheat (e.g., disease resistance and increased yield) has occurred since humans first started cultivating diploid and tetraploid wheats, approximately 10,000 years ago (Giles & Brown, 2006; Zohary et al., 2013). A historical analysis of varieties, registered from the late 1800s to the present day, indicates that commercial plant breeding has altered the amino acid composition of wheat grain, along with many other agronomic and quality traits (Poudel et al., 2021). However, free asparagine showed no discernible change across the measured period and another study by Rapp et al. (2018) also did not find any temporal trend in free asparagine content across the varieties screened in that study. This is in contrast with Corol et al. (2016), who detected a weak negative correlation between variety release year and grain free asparagine content ($r = -0.255$, $p = 0.0019$). This slight negative correlation may be due to the decreasing protein content of varieties, as a result of selection for increasing yields and the yield-dilution effect.

The lack of any strong correlation between variety release year and grain free asparagine content reflects that free asparagine concentration is not strongly linked to any other traits that have been selected for, over the course of commercial wheat breeding history. However, free asparagine concentration does display a moderate heritability in some studies (Table 3.4), with the study that used one of the more robust estimates of heritability (Piepho and Möhring method), estimating heritability at 0.65, similar to the heritability estimates obtained for protein content and falling number from the same study (Rapp et al., 2018). Further accurate measurements of grain asparagine heritability are required to corroborate this, as well as measurements taken across multiple environments, but this indicates that there is scope for reducing the free asparagine content of wheat grain through breeding. However, there is undoubtedly a substantial environmental (E), as well as genetic (G), effect on free asparagine concentration, together with a G × E interaction, which may have discouraged breeders from attempting to develop low-asparagine varieties to date.

Table 3.4. Heritability estimates of asparagine in wheat (given to 2 significant figures).

Heritability Method	h^2	**Reference**
Broad-sense	0.31	[20]
Surrogate method	0.13	[29]
Piepho and Möhring	0.65	[23]
Broad-sense	0.41	[61]

Breeding low-asparagine wheat could potentially be achieved in three main ways: directly, by using either existing or induced variation, or indirectly, through selection for related traits (Figure 3.3). New wheat varieties are commonly developed using existing variation. However, the only multi-environment quantitative trait locus (QTL) for low-asparagine known at present is the one in which the *TaASN-B2* gene is either present or deleted. This has been shown to affect the free asparagine content of grain in two different field trials (Oddy et al., 2021). Selection for the *TaASN-B2* deletion represents an easy gain for breeders, but further trials testing the effect of the deletion should be performed to confirm the stability of the effect across more environments. Other QTL controlling grain free asparagine content have also been identified, but these have not yet been verified across more than one environment (Emebiri, 2014; Rapp et al., 2018). Identification of multi-environment QTL, in combination with

genomic and marker assisted selection (Rapp et al., 2018), could enable low-asparagine wheat to be developed, without the time-consuming or expensive need to screen large numbers of plants for free asparagine concentration.

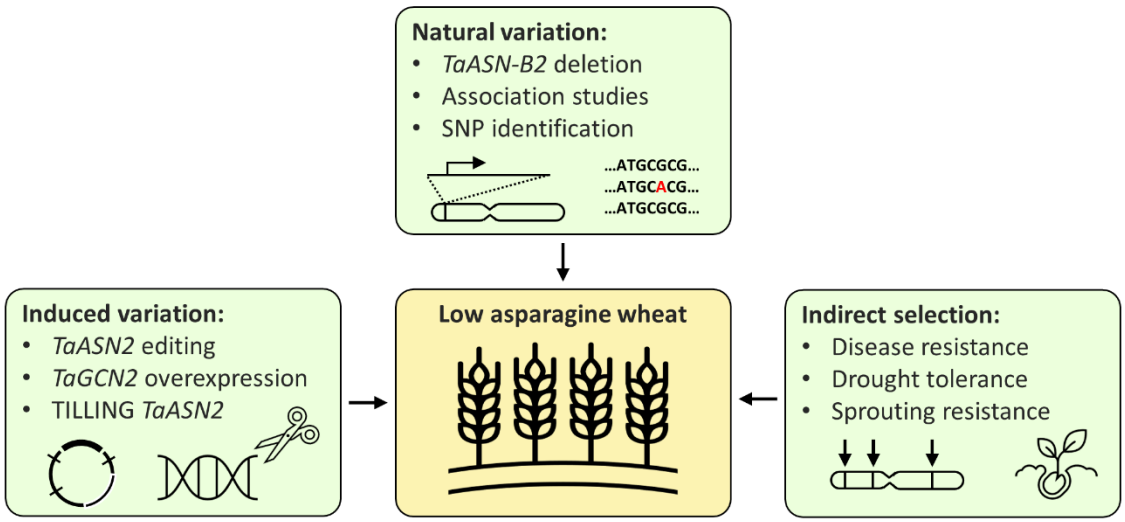


Figure 3.3. Strategies for the breeding of low-asparagine wheat. *TaASN2* (asparagine synthetase 2), *TaGCN2* (general control nonderepressible 2), TILLING (targeted induced local lesions in genomes).

Relying on natural variation is limited by the availability of existing variation, whereas techniques that induce or increase variation in the wheat genome could generate new variants with free asparagine content below the normal range. This has been demonstrated by the use of CRISPR/Cas9 technology to 'knock out' the *TaASN2* genes, reducing grain free asparagine content by up to 90% in glasshouse experiments (Raffan et al., 2021). The edited lines still need to undergo trials to confirm the stability of this phenotype in the field, but the stability of the 'natural' *TaASN-B2* deletion phenotype under field conditions (Oddy et al., 2021) is encouraging, suggesting that the *TaASN2*-edited phenotypes may be similarly stable. However, the interaction between the *TaASN-B2* deletion and sulphur deficiency implies that *TaASN2* variants may not be sufficient to control grain free asparagine content during sulphur deficiency or other stresses, again highlighting the effects of E and G × E. On the other hand, the varieties carrying the *TaASN-B2* deletion have intact *TaASN-A2* and *TaASN-D2* genes, whereas the edited lines lack any functional *TaASN2* genes, so the edited lines will

be valuable for investigating whether this prevents free asparagine accumulation under conditions of sulphur deficiency or other stresses.

The benefits of inducing variation in candidate genes was also recently demonstrated by Alarcón-Reverte et al. (2022), in which wheat plants possessing EMS-induced null *TaASN-A2* alleles were grown in the field and tested for grain free-asparagine content. Reductions of between 9% and 34% were achieved, without any negative side effects on quality traits, demonstrating again the utility of induced variation and the lack of strong associations between free asparagine and quality traits (except for acrylamide and colour).

As a result of the potential loss or partial loss of the low-asparagine phenotype of *TaASN2* knockouts under stress, a third, complementary option for controlling grain free asparagine content can also be adopted: breeding for stress tolerance. As discussed above, stress and grain free asparagine content are closely linked, and it is often during stress that the highest grain free asparagine contents are observed (Curtis et al., 2016; Curtis et al., 2018). Breeding for stress tolerance could, therefore, ensure that a low-asparagine phenotype would be retained under stress. The relationship between stress and asparagine requires further investigation to confirm this however. Selection for other related traits, such as those discussed above (e.g., PHS resistance, delayed senescence), could also provide indirect selection for lower-grain asparagine, but these traits are not as clearly linked with asparagine as asparagine is with stress.

3.5. Conclusions

Asparagine is, of course, an important plant metabolite, and since the discovery that it can be converted to acrylamide during the cooking and processing of food, there has been debate over how much its concentration could be reduced before effects were seen on other important traits. It was also recognized by the food industry that the production of fried, roasted, toasted and baked coffee, potato and cereal products containing no acrylamide at all was not possible, and that they should aim to reduce acrylamide to concentrations ‘as low as reasonably achievable’ (FoodDrinkEurope, 2019). More recently, at least in the European Union, it has become a regulatory compliance issue, with manufacturers striving to keep the acrylamide concentrations in their products below benchmark levels set by the European Commission (European Commission, 2017). The European Commission is currently considering replacing benchmark levels (described as ‘performance indicators’) for some products with maximum levels; i.e.,

concentrations of acrylamide above which it would be illegal to sell a product (European Commission, 2021). If maximum levels were set at or lower than the current benchmark levels, it would have serious implications for the food industry. There is a paucity of data in the public domain on acrylamide concentrations in cereal products, but a recent study in Spain found that 15% of breakfast cereals contained acrylamide above the benchmark level, which for wheat-based breakfast cereals is 300 parts per billion (ppb) (Mesías et al., 2019). Manufacturers could, therefore, face the prospect of product recalls and even prosecution if a maximum level of 300 ppb was imposed. This makes it more important than ever that wheat breeders engage with the acrylamide issue, especially as many strategies involving agronomy and food-processing technology have already been implemented (FoodDrinkEurope, 2019) and the opportunities for further gains involving those approaches may be limited.

The strategies outlined here show that the breeding of low-asparagine wheat, using natural and induced variation, is feasible and unlikely to negatively impact other traits, with the exception of germination, which may be affected, but only if free asparagine concentration is reduced to very low concentrations (Raffan et al., 2021). Furthermore, breeding solutions stand to be more sustainable, cost-effective, and less impactful on flavour than the solutions provided by agronomic and food sciences, and could make additional agronomic or food industry modifications unnecessary. Consequently, development of low-asparagine phenotypes in elite wheat varieties should be considered in future wheat breeding programs.

3.6. References

Alarcón-Reverte, R., Xie, Y., Stromberger, J., Cotter, J. D., Mason, R. E., & Pearce, S. (2022). Induced mutations in ASPARAGINE SYNTHETASE-A2 reduce free asparagine concentration in the wheat grain. *Crop Science*, *62*(4), 1484-1496.

Avila-Ospina, L., Marmagne, A., Talbotec, J., Krupinska, K., & Masclaux-Daubresse, C. (2015). The identification of new cytosolic glutamine synthetase and asparagine synthetase genes in barley (*Hordeum vulgare* L.), and their expression during leaf senescence. *Journal of Experimental Botany*, *66*(7), 2013-2026.

Bovet, L., Cheval, C., Hilfiker, A., Battey, J., Langlet, D., Broye, H., ... & Goepfert, S. (2019). Asparagine synthesis during tobacco leaf curing. *Plants*, *8*(11), 492.

Byrne, E. H., Prosser, I., Muttucumaru, N., Curtis, T. Y., Wingler, A., Powers, S., & Halford, N. G. (2012). Overexpression of GCN2-type protein kinase in wheat has profound effects on free amino acid concentration and gene expression. *Plant Biotechnology Journal*, *10*(3), 328-340.

Cameron, K. C., Di, H. J., & Moir, J. L. (2013). Nitrogen losses from the soil/plant system: a review. *Annals of Applied Biology*, *162*(2), 145-173.

Canales, J., Rueda-López, M., Craven-Bartle, B., Avila, C., & Cánovas, F. M. (2012). Novel insights into regulation of asparagine synthetase in conifers. *Frontiers in Plant Science*, *3*, 100.

Claus, A., Schreiter, P., Weber, A., Graeff, S., Herrmann, W., Claupein, W., ... & Carle, R. (2006). Influence of agronomic factors and extraction rate on the acrylamide contents in yeast-leavened breads. *Journal of Agricultural and Food Chemistry*, *54*(23), 8968-8976.

Corol, D. I., Ravel, C., Rakszegi, M., Charmet, G., Bedo, Z., Beale, M. H., ... & Ward, J. L. (2016). ^{1}H-NMR screening for the high-throughput determination of genotype and environmental effects on the content of asparagine in wheat grain. *Plant Biotechnology Journal*, *14*(1), 128-139.

Curtis, T. Y., Bo, V., Tucker, A., & Halford, N. G. (2018). Construction of a network describing asparagine metabolism in plants and its application to the identification of genes affecting asparagine metabolism in wheat under drought and nutritional stress. *Food and Energy Security*, *7*(1), e00126.

Curtis, T. Y., Muttucumaru, N., Shewry, P. R., Parry, M. A., Powers, S. J., Elmore, J. S., ... & Halford, N. G. (2009). Effects of genotype and environment on free amino acid levels in wheat grain: implications for acrylamide formation during processing. *Journal of Agricultural and Food Chemistry*, *57*(3), 1013-1021.

Curtis, T. Y., Powers, S. J., & Halford, N. G. (2016). Effects of fungicide treatment on free amino acid concentration and acrylamide-forming potential in wheat. *Journal of Agricultural and Food Chemistry*, *64*(51), 9689-9696.

Curtis, T. Y., Powers, S. J., Wang, R., & Halford, N. G. (2018). Effects of variety, year of cultivation and sulphur supply on the accumulation of free asparagine in the grain of commercial wheat varieties. *Food Chemistry*, *239*, 304-313.

Curtis, T. Y., Raffan, S., Wan, Y., King, R., Gonzalez-Uriarte, A., & Halford, N. G. (2019). Contrasting gene expression patterns in grain of high and low asparagine wheat genotypes in response to sulphur supply. *BMC Genomics*, *20*(1), 1-18.

Distelfeld, A., Avni, R., & Fischer, A. M. (2014). Senescence, nutrient remobilization, and yield in wheat and barley. *Journal of Experimental Botany*, *65*(14), 3783-3798.

Emebiri, L. C. (2014). Genetic variation and possible SNP markers for breeding wheat with low-grain asparagine, the major precursor for acrylamide formation in heat-processed products. *Journal of the Science of Food and Agriculture*, *94*(7), 1422-1429.

European Commission. (2021). Food Safety: Acrylamide. Available online: https://ec.europa.eu/food/safety/chemical-safety/contaminants/catalogue/acrylamide_en (accessed on 22 February 2022).

European Commission. (2011). *Commission Recommendation of 10.1.2011 on Investigations into the Levels of Acrylamide in Food*; European Commission: Brussels, Belgium.

European Commission. (2017). *Commission Regulation (EU) 2017/2158 of 20 November 2017 Establishing Mitigation Measures and Benchmark Levels for the Reduction of the Presence of Acrylamide in Food*; European Commission: Brussels, Belgium, 2017.

FAO. (2021). FAOSTAT: Food Balances (2014-). Available online: http://www.fao.org/faostat/en/ (accessed on 16 November 2021).

Food Drink Europe. (2019). *Acrylamide Toolbox 2019*; Food Drink Eur.: Brussels, Belgium, 2019.

Giles, R. J., & Brown, T. A. (2006). GluDy allele variations in *Aegilops tauschii* and *Triticum aestivum*: implications for the origins of hexaploid wheats. *Theoretical and Applied Genetics*, *112*(8), 1563-1572.

Granvogl, M., Wieser, H., Koehler, P., Von Tucher, S., & Schieberle, P. (2007). Influence of sulfur fertilization on the amounts of free amino acids in wheat. Correlation with baking properties as well as with 3-aminopropionamide and acrylamide generation during baking. *Journal of Agricultural and Food Chemistry*, *55*(10), 4271-4277.

Halford, N. G., Muttucumaru, N., Curtis, T. Y., & Parry, M. A. (2007). Genetic and agronomic approaches to decreasing acrylamide precursors in crop plants. *Food Additives and Contaminants*, *24*(sup1), 26-36.

Hawkesford, M. J., & Riche, A. B. (2020). Impacts of G x E x M on nitrogen use efficiency in wheat and future prospects. *Frontiers in Plant Science*, *11*, 1157.

Herrera-Rodríguez, M. B., Maldonado, J. M., & Pérez-Vicente, R. (2006). Role of asparagine and asparagine synthetase genes in sunflower (*Helianthus annuus*) germination and natural senescence. *Journal of Plant Physiology*, *163*(10), 1061-1070.

Heyneke, E., Watanabe, M., Erban, A., Duan, G., Buchner, P., Walther, D., ... & Hoefgen, R. (2019). Effect of senescence phenotypes and nitrate availability on wheat leaf metabolome during grain filling. *Agronomy*, *9*(6), 305.

Hinckley, E. L. S., Crawford, J. T., Fakhraei, H., & Driscoll, C. T. (2020). A shift in sulfur-cycle manipulation from atmospheric emissions to agricultural additions. *Nature Geoscience*, *13*(9), 597-604.

Hu, M., Zhao, X., Liu, Q., Hong, X., Zhang, W., Zhang, Y., ... & Tong, Y. (2018). Transgenic expression of plastidic glutamine synthetase increases nitrogen uptake and yield in wheat. *Plant Biotechnology Journal*, *16*(11), 1858-1867.

Jukanti, A. K., Heidlebaugh, N. M., Parrott, D. L., Fischer, I. A., McInnerney, K., & Fischer, A. M. (2008). Comparative transcriptome profiling of near-isogenic barley (*Hordeum vulgare*) lines differing in the allelic state of a major grain protein content locus identifies genes with possible roles in leaf senescence and nitrogen reallocation. *New Phytologist*, *177*(2), 333-349.

Liu, Y., Ohm, J. B., Hareland, G., Wiersma, J., & Kaiser, D. (2011). Sulfur, protein size distribution, and free amino acids in flour mill streams and their relationship to dough rheology and breadmaking traits. *Cereal Chemistry*, *88*(2), 109-116.

Lukaszewski, A. J., Alberti, A., Sharpe, A., Kilian, A., Stanca, A. M., Keller, B., ... & Dubská, Z. (2014). A chromosome-based draft sequence of the hexaploid bread wheat (*Triticum aestivum*) genome. *Science*, *345*(6194), 1251788.

Malunga, L. N., Ames, N., Khorshidi, A. S., Thandapilly, S. J., Yan, W., Dyck, A., ... & Scanlon, M. G. (2021). Association of asparagine concentration in wheat with cultivar, location, fertilizer, and their interaction. *Food Chemistry*, *344*, 128630.

Malunga, L. N., Ames, N. P., Masatcioglu, M. T., Khorshidi, A. S., Thandapilly, S. J., Cuthbert, R. D., ... & Scanlon, M. G. (2019). Free asparagine concentrations in Canadian hard red spring wheat cultivars. *Canadian Journal of Plant Science*, *99*(3), 338-347.

Martinek, P., Klem, K., Vanova, M., Bartackova, V., Vecerkova, L., Bucher, P., & Hajslova, J. (2009). Effects of nitrogen nutrition, fungicide treatment and wheat genotype on free asparagine and reducing sugars content as precursors of acrylamide formation in bread. *Plant, Soil and Environment*, *55*(5), 187-195.

Masclaux-Daubresse, C., Reisdorf-Cren, M., Pageau, K., Lelandais, M., Grandjean, O., Kronenberger, J., ... & Suzuki, A. (2006). Glutamine synthetase-glutamate synthase pathway and glutamate dehydrogenase play distinct roles in the sink-source nitrogen cycle in tobacco. *Plant Physiology*, *140*(2), 444-456.

Mesías, M., Sáez-Escudero, L., Morales, F. J., & Delgado-Andrade, C. (2019). Reassessment of acrylamide content in breakfast cereals. Evolution of the Spanish market from 2006 to 2018. *Food Control*, *105*, 94-101.

Mottram, D. S., Wedzicha, B. L., & Dodson, A. T. (2002). Acrylamide is formed in the Maillard reaction. *Nature*, *419*(6906), 448-449.

Muttucumaru, N., Halford, N. G., Elmore, J. S., Dodson, A. T., Parry, M., Shewry, P. R., & Mottram, D. S. (2006). Formation of high levels of acrylamide during the processing of flour derived from sulfate-deprived wheat. *Journal of Agricultural and Food Chemistry*, *54*(23), 8951-8955.

Muttucumaru, N., Powers, S. J., Elmore, J. S., Briddon, A., Mottram, D. S., & Halford, N. G. (2014). Evidence for the complex relationship between free amino acid and sugar concentrations and acrylamide-forming potential in potato. *Annals of Applied Biology*, *164*(2), 286-300.

Navrotskyi, S., Baenziger, P. S., Regassa, T., Guttieri, M. J., & Rose, D. J. (2018). Variation in asparagine concentration in Nebraska wheat. *Cereal Chemistry*, *95*(2), 264-273.

Oddy, J., Alarcón-Reverte, R., Wilkinson, M., Ravet, K., Raffan, S., Minter, A., ... & Pearce, S. (2021). Reduced free asparagine in wheat grain resulting from a natural deletion of TaASN-B2: investigating and exploiting diversity in the asparagine synthetase gene family to improve wheat quality. *BMC Plant Biology*, *21*(1), 1-17.

Oddy, J., Raffan, S., Wilkinson, M. D., Elmore, J. S., & Halford, N. G. (2020). Stress, nutrients and genotype: understanding and managing asparagine accumulation in wheat grain. *CABI Agriculture and Bioscience*, *1*(1), 1-14.

Ohm, J. B., Hareland, G., Simsek, S., & Seabourn, B. (2009). Size-exclusion HPLC of protein using a narrow-bore column for evaluation of breadmaking quality of hard spring wheat flours. *Cereal Chemistry*, *86*(4), 463-469.

Ohm, J. B., Mergoum, M., & Simsek, S. (2017). Variation of free asparagine concentration and association with quality parameters for hard red spring wheat grown in North Dakota. *Cereal Chemistry*, *94*(4), 712-716.

Ohm, J. B., Simsek, S., & Mergoum, M. (2018). Variation of protein MWD parameters and their associations with free asparagine concentration and quality characteristics in hard red spring wheat. *Journal of Cereal Science*, *79*, 154-159.

Poudel, R., Bhinderwala, F., Morton, M., Powers, R., & Rose, D. J. (2021). Metabolic profiling of historical and modern wheat cultivars using proton nuclear magnetic resonance spectroscopy. *Scientific Reports*, *11*(1), 1-10.

Raffan, S., & Halford, N. G. (2019). Acrylamide in food: Progress in and prospects for genetic and agronomic solutions. *Annals of Applied Biology*, *175*(3), 259-281.

Raffan, S., Sparks, C., Huttly, A., Hyde, L., Martignago, D., Mead, A., ... & Halford, N. G. (2021). Wheat with greatly reduced accumulation of free asparagine in the grain, produced by CRISPR/Cas9 editing of asparagine synthetase gene TaASN2. *Plant Biotechnology Journal*, *19*(8), 1602-1613.

Rapp, M., Schwadorf, K., Leiser, W. L., Würschum, T., & Longin, C. F. H. (2018). Assessing the variation and genetic architecture of asparagine content in wheat: What can plant breeding contribute to a reduction in the acrylamide precursor? *Theoretical and Applied Genetics*, *131*(11), 2427-2437.

Simsek, S., Ohm, J. B., Lu, H., Rugg, M., Berzonsky, W., Alamri, M. S., & Mergoum, M. (2014). Effect of pre-harvest sprouting on physicochemical changes of proteins in wheat. *Journal of the Science of Food and Agriculture*, *94*(2), 205-212.

Simsek, S., Ohm, J. B., Lu, H., Rugg, M., Berzonsky, W., Alamri, M. S., & Mergoum, M. (2014). Effect of pre-harvest sprouting on physicochemical properties of starch in wheat. *Foods*, *3*(2), 194-207.

Smedley, M. A., Hayta, S., Clarke, M., & Harwood, W. A. (2021). CRISPR-Cas9 based genome editing in wheat. *Current Protocols*, *1*(3), e65.

Stadler, R. H., Blank, I., Varga, N., Robert, F., Hau, J., Guy, P. A., ... & Riediker, S. (2002). Acrylamide from Maillard reaction products. *Nature*, *419*(6906), 449-450.

Stockmann, F., Weber, E. A., Merkt, N., Schreiter, P., Claupein, W., & Graeff-Hönninger, S. (2019). Impact of row distance and seed density on grain yield, quality traits, and free asparagine of organically grown wheat. *Agronomy*, *9*(11), 713.

Stockmann, F., Weber, E. A., Schreiter, P., Merkt, N., Claupein, W., & Graeff-Hönninger, S. (2018). Impact of nitrogen and sulfur supply on the potential of acrylamide formation in organically and conventionally grown winter wheat. *Agronomy*, *8*(12), 284.

Surdyk, N., Rosén, J., Andersson, R., & Åman, P. (2004). Effects of asparagine, fructose, and baking conditions on acrylamide content in yeast-leavened wheat bread. *Journal of Agricultural and Food Chemistry*, *52*(7), 2047-2051.

Tareke, E., Rydberg, P., Karlsson, P., Eriksson, S., & Törnqvist, M. (2002). Analysis of acrylamide, a carcinogen formed in heated foodstuffs. *Journal of Agricultural and Food Chemistry*, *50*(17), 4998-5006.

Tiong, J., Sharma, N., Sampath, R., MacKenzie, N., Watanabe, S., Metot, C., ... & Okamoto, M. (2021). Improving nitrogen use efficiency through overexpression of alanine aminotransferase in rice, wheat, and barley. *Frontiers in Plant Science*, *12*, 628521.

Tsilo, T. J., Ohm, J. B., Hareland, G. A., & Anderson, J. A. (2010). Association of size-exclusion HPLC of endosperm proteins with dough mixing and breadmaking characteristics in a recombinant inbred population of hard red spring wheat. *Cereal Chemistry*, *87*(2), 104-111.

Walkowiak, S., Gao, L., Monat, C., Haberer, G., Kassa, M. T., Brinton, J., ... & Pozniak, C. J. (2020). Multiple wheat genomes reveal global variation in modern breeding. *Nature*, *588*(7837), 277-283.

Wan, Y., Wang, Y., Shi, Z., Rentsch, D., Ward, J. L., Hassall, K., ... & Hawkesford, M. J. (2021). Wheat amino acid transporters highly expressed in grain cells regulate amino acid accumulation in grain. *PloS One*, *16*(2), e0246763.

Weber, E. A., Graeff, S., Koller, W. D., Hermann, W., Merkt, N., & Claupein, W. (2008). Impact of nitrogen amount and timing on the potential of acrylamide formation in winter wheat (*Triticum aestivum* L.). *Field Crops Research*, *106*(1), 44-52.

Wilkinson, P. A., Allen, A. M., Tyrrell, S., Wingen, L. U., Bian, X., Winfield, M. O., ... & Barker, G. L. (2020). CerealsDB—new tools for the analysis of the wheat genome: update 2020. *Database*, *2020*, baaa060.

Xie, Y., Malunga, L. N., Ames, N. P., Waterer, J., Khorshidi, A. S., & Scanlon, M. G. (2021). Effects of growing environment, genotype, and commercial fertilization levels on free asparagine concentration in Western Canadian wheat. *Cereal Chemistry*, *98*(1), 89-99.

Xu, H., Curtis, T. Y., Powers, S. J., Raffan, S., Gao, R., Huang, J., ... & Halford, N. G. (2018). Genomic, biochemical, and modeling analyses of asparagine synthetases from wheat. *Frontiers in Plant Science*, *8*, 2237.

Yadav, A. K., Carroll, A. J., Estavillo, G. M., Rebetzke, G. J., & Pogson, B. J. (2019). Wheat drought tolerance in the field is predicted by amino acid responses to glasshouse-imposed drought. *Journal of Experimental Botany*, *70*(18), 4931-4948.

Zohary, D., Hopf, M., & Weiss, E. (2012). *Domestication of Plants in the Old World: The origin and spread of domesticated plants in Southwest Asia, Europe, and the Mediterranean Basin*. Oxford University Press.

4

Reduced free asparagine in wheat grain resulting from a natural deletion of *TaASN-B2*: investigating and exploiting diversity in the asparagine synthetase gene family to improve wheat quality

Joseph Oddy[1], Rocío Alarcón-Reverte[2], Mark Wilkinson[1], Karl Ravet[2], Sarah Raffan[1], Andrea Minter[3], Andrew Mead[3], J. Stephen Elmore[4], Isabel Moreira de Almeida[5], Nicholas C. Cryer[6], Nigel G. Halford[1], Stephen Pearce[2]

1. Plant Sciences Department, Rothamsted Research, Harpenden, Hertfordshire, AL5 2JQ, United Kingdom

2. Department of Soil and Crop Sciences, Colorado State University, Fort Collins, CO 80523, USA

3. Computational and Analytical Sciences Department, Rothamsted Research, Harpenden, Hertfordshire, AL5 2JQ, United Kingdom

4. Department of Food & Nutritional Sciences, University of Reading, Whiteknights, Reading, RG6 6DZ, United Kingdom

5. Mondelēz R&D International, Paris Saclay Tech Center, 6 Rue René Razel, 91400 Saclay, France

6. Mondelēz UK R&D Ltd, Bournville Lane, Bournville, Birmingham, B30 2LU, United Kingdom

BMC Plant Biology, 21, 302 (2021).

This chapter has been modified from its published form to allow for edits and corrections within this Thesis.

4.1. Introduction to paper

The first sub-project of my PhD was to characterise the natural presence/absence variation (PAV) of *TaASN-B2* and assess its impact on grain free asparagine content. I started this sub-project by performing bioinformatic analysis to locate where this presence/absence variation occurred in the wheat genome and to compare the DNA and protein structures of the *TaASN2* homeologs. Using these sequence data, I developed a PCR test to screen for the presence/absence of *TaASN-B2* and began screening varieties from previous field trials. Shortly afterwards, the COVID-19 pandemic began so some of this screening work was postponed until work on site started again. It was during this time that our group at Rothamsted was contacted by another group at Colorado State University who were also investigating the *TaASN-B2* PAV. This led to our collaboration on this sub-project, as they had RNA-seq data and wheat pangenomic analyses that complemented our molecular and field trial data (see Formal Acknowledgements and Contribution to Papers section for more detail). This allowed us to perform a more comprehensive analysis of *TaASN-B2* and corroborated the analyses I had performed.

4.2. Background

Asparagine in its free (soluble, non-protein) form is an important nitrogen transport and storage molecule in plants (see Lea et al. (2007) for review). It also accumulates during abiotic and biotic stress and has potential roles in ammonium detoxification and reactive oxygen species/nitrous oxide production (see Oddy et al. (2020) for review). However, free asparagine is also the precursor for acrylamide (C_3H_5NO), a carcinogenic contaminant that forms during the frying, roasting, baking, toasting and high-temperature processing of grains, tubers, beans and storage roots (reviewed in Raffan & Halford (2019)). Acrylamide is classified as an extremely hazardous substance in the United States (USA), a serious health hazard with acute toxicity in the European Union (EU), and a Group 2A carcinogen (probably carcinogenic to humans) by the International Agency for Research on Cancer (IARC, 1994).

The European Commission has led the way in developing a regulatory system for acrylamide concentrations in food (see Raffan & Halford (2019) for a comprehensive review). The current EU regulation on acrylamide in food (Commission Regulation (EU) 2017/2158 (European Commission, 2017)) states that acrylamide in food 'potentially increases the risk of developing cancer for consumers in all age groups'. It also sets Benchmark Levels for acrylamide in

different food types, and carries an explicit threat to set Maximum Levels (i.e. concentrations above which it would be illegal to sell a product) in the future.

In the USA, the federal government has not introduced equivalent regulations, although the Food and Drug Administration (FDA) has issued an acrylamide 'action plan' (Food and Drug Administration, 2016). However, as long ago as 2005 the Attorney General of the State of California filed a lawsuit against five food companies and four restaurant chains for failing to label their products with a warning to alert consumers to the presence of acrylamide (reviewed in Raffan & Halford (2019)). California has also seen private lawsuits brought against the coffee industry over the lack of warning notices. Regulators in other countries that have taken a position on dietary acrylamide include Health Canada, Food Standards Australia New Zealand (FSANZ) and authorities in Japan and Hong Kong (reviewed in Raffan & Halford (2019)). Meanwhile, a recent study identified a unique mutational 'signature' associated with acrylamide and its metabolite, glycidamide, that was found at high frequency in multiple human tumour types (Zhivagui et al., 2019). This represents the strongest evidence yet of a link between dietary acrylamide intake and cancer in humans.

Reduced free asparagine concentrations in crop raw materials would greatly assist the food industry in complying with regulations on acrylamide in food products. In wheat, free asparagine concentrations in the grain are typically higher in plants grown in sulphur-deficient soils, or in plants infected by pathogens. Therefore, crop-management strategies, including ensuring that wheat is supplied with sufficient sulphur during cultivation (Raffan et al., 2020), and protected from pathogen infection (Martinek et al., 2009; Curtis et al., 2016), are the most common strategies to reduce free asparagine concentrations. Nevertheless, wheat breeders are under pressure from food businesses to develop varieties with reduced concentrations of free asparagine in the grain. Although free asparagine concentrations do vary across genotypes and exhibit reasonably high heritability (Emebiri, 2014; Curtis et al., 2018; Rapp et al., 2018), the large effect of crop management and other environmental (E) factors, both per se and in combination with genetic factors (G × E), means that breeding for low free asparagine concentrations will not be a simple task (reviewed in Raffan & Halford (2019)). Through association mapping, several quantitative trait loci (QTL) for free asparagine content have been identified, but no common QTL have been identified between studies (Emebiri, 2014; Rapp et al., 2018).

Asparagine is synthesised by the transfer of an amino group from glutamine to aspartate to make asparagine and glutamate in a reaction catalysed by asparagine synthetase. The cereal asparagine synthetase gene family comprises between two and five genes per diploid genome (Raffan & Halford, 2021), with members of the Triticeae tribe all having five genes per genome, assigned to four groups: 1, 2, 3 (subdivided into 3.1 and 3.2) and 4. The one documented exception is that some hexaploid common wheat (*Triticum aestivum* L.; genomes AABBDD) and tetraploid emmer wheat (*T. turgidum*; genomes AABB) genotypes lack a group 2 gene in the B genome (*TaASN-B2*/*TdASN-B2*) (Raffan & Halford, 2021; Xu et al., 2018). This gene is absent from the IWGSC RefSeq v1.1 genome assembly of the common wheat landrace Chinese Spring but is present in the cultivar Cadenza (Raffan & Halford et al., 2021). In tetraploid wheats, *ASN-B2* is absent in wild emmer wheat (*T. turgidum* ssp. *dicoccoides*) genotype Zavitan, but present in domesticated durum wheat cultivar Svevo (*T. turgidum* L. ssp. *durum* (Desf.) Husn.) (Raffan & Halford et al., 2021). Since wild emmer wheat is believed to be the B genome donor for both tetraploid durum and hexaploid common wheat (Dubcovsky et al., 2007), one possible explanation for the presence of *ASN-B2* in some cultivars but not others is that the hybridisation event that produced hexaploid wheat occurred more than once, involving emmer wheats with and without *ASN-B2*. This would be consistent with evidence of multiple hybridisations found from wider analyses of genome data (Marcussen et al., 2014).

The extent of the presence/absence of *ASN-B2* in different genotypes is particularly interesting because the *TaASN2* genes are the most highly expressed members of the asparagine synthetase family in the grain (Gao et al., 2016; Curtis et al., 2019). Wheat plants carrying CRISPR/Cas9 induced edits in all six *TaASN2* alleles (homozygous edits in each subgenome) exhibit greatly reduced free asparagine concentration in their grains (Raffan et al., 2021). It follows that the natural *ASN-B2* deletion could represent a valuable genetic variant for wheat breeders to exploit in order to reduce free asparagine content in the grain. In the present study, therefore, natural genetic variation in the asparagine synthetase gene family in wheat was characterised. The presence/absence of *ASN-B2* was screened in a panel of UK and global common wheat varieties, as well as wheat progenitor genomes and wild species. The deletion of *TaASN-B2* is associated with an overall reduction in *TaASN2* transcript levels and grain free asparagine concentrations and may be a useful allele for wheat breeding programmes to develop varieties with lower concentrations of free asparagine.

4.3. Results

4.3.1 Natural diversity in the asparagine synthetase gene family in wheat

Full-length coding sequences of *ASN* genes from the wheat landrace Chinese Spring were used as queries in BLASTn searches against the genome assemblies of 14 common wheat varieties and spelt wheat (*T. aestivum* ssp. *spelta*) to characterise natural allelic variation in the wheat asparagine synthetase gene family. The results are shown in Table 4.1, ordered by gene name (Xu et al., 2018; Raffan & Halford, 2021) and the corresponding annotated gene model ID from the Chinese Spring RefSeq v1.1 genome assembly (IWGSC, 2018). For each orthologous gene, Sorting Intolerant From Tolerant (SIFT) analysis was performed on the translated protein to predict whether the variation in amino acid sequences was likely to disrupt protein function (highlighted in yellow in Table 4.1) or to be tolerated (highlighted in green). Full details of specific amino acid changes for all wheat varieties are provided in Table S4.1.

Table 4.1. Natural variation in ASN proteins in 14 wheat varieties. Each gene is annotated by name (Xu et al., 2018; Raffan & Halford, 2021) and genome (hence *TaASN-A1*, for example, is the A genome *TaASN1* gene), and the gene ID from the Chinese Spring RefSeq v1.1 gene models (note that *TaASN-B2* is absent in Chinese Spring). Green shading indicates no predicted impact of variety-specific mutations; yellow shading indicates one or more mutations that are predicted to disrupt protein function; red shading indicates that the gene is not present in that genome.

Protein	IWGSC RefSeq v1.1 ID	Robigus	CDC Landmark	Julius	Claire	Jagger	Cadenza	Paragon	Arina	Norin 61	CDC Stanley	SY Mattis	Lancer	Mace	Spelt
TaASN-A1	*TraesCS5A02G153900*														
TaASN-B1	*TraesCS5B02G152600*														
TaASN-D1	*TraesCS5D02G159100*														
TaASN-A2	*TraesCS3A02G077100*														
TaASN-B2	Not present														
TaASN-D2	*TraesCS3D02G077300*														
TaASN-A3.1	*TraesCS1A02G382800*														
TaASN-B3.1	*TraesCS1B02G408200*														
TaASN-D3.1	*TraesCS1D02G390500*														
TaASN-A3.2	*TraesCS1A02G422100*														
TaASN-B3.2	*TraesCS1B02G453600*														
TaASN-D3.2	*TraesCS1D02G430300*														
TaASN-A4	*TraesCS4A02G109900*														
TaASN-B4	*TraesCS4B02G194400*														
TaASN-D4	*TraesCS4D02G195100*														

There were deletions, polymorphisms and presence/absence variation in several wheat *ASN* genes. For example, *TaASN-B1* was deleted in SY Mattis, while eight other varieties carried an allelic variant with a 16 bp deletion in exon seven, introducing a frame shift

and bringing a premature stop codon into frame. The presence of this deletion means that the gene is predicted to encode a 375 amino acid protein with a C-terminal truncation of 209 amino acids, including part of the asparagine synthetase domain, indicating that this protein is likely to be non-functional (Fig. S4.1). *TaASN-B3.2* was deleted in Norin 61, CDC Stanley and Lancer, whereas *TaASN-A3.2* and *TaASN-D3.2* were present in all analysed varieties, and showed no polymorphisms predicted to impact protein function (Table 4.1). In contrast, 12 wheat varieties carried *TaASN-A3.1* alleles with polymorphisms predicted to disrupt protein function (Table 4.1). Some varieties carry a combination of alleles predicted to disrupt the function of multiple asparagine synthetase proteins. For example, CDC Stanley carries alleles predicted to affect the function of the enzymes encoded by *TaASN-A1*, *TaASN-B1* and *TaASN-A3.1*, in addition to a deletion of *TaASN-B3.2*, while SY Mattis carries deletions of *TaASN-B1* and *TaASN-B2*, and disruptive alleles of *TaASN-A3.1* and *TaASN-D2* (Table 4.1). The most common presence/absence variation was of *TaASN-B2*, which was deleted in eight of the 15 genotypes assayed, including Chinese Spring (Table 4.1).

4.3.2. Characterisation of the TaASN-B2 deletion

The deletion containing *TaASN-B2* was mapped to chromosome arm 3BS in the Chinese Spring RefSeq v1.1 genome assembly (Fig. 4.1a). Alignment of the surrounding region was performed between the annotated Chinese Spring genome and the corresponding region of the Svevo and Jagger genomes, both of which contain the *ASN-B2* gene, to evaluate other features of this locus (Fig. 4.1a and b). The deletion in Chinese Spring was 12,752 bp with respect to the Svevo genome and 12,770 bp with respect to the Jagger genome. A putative open reading frame predicted to encode an F-box protein was detected upstream of *TaASN-B2* in the deleted region (Fig. 4.1a), but further analysis of this gene was not undertaken and this gene is not predicted in the latest gene models of the the 10+ Wheat Genomes Project (https://webblast.ipk-gatersleben.de/wheat_ten_genomes/). Directly downstream of the deletion in Chinese Spring and the corresponding region in Svevo and Jagger there is a large, long terminal repeat (LTR) retrotransposon, *Inga*, belonging to the Ty1-*copia* family (Wicker & Keller, 2007) (Fig. 4.1a), the identity of which was confirmed using the TREP database (Wicker et al., 2002). Analysis of the other genome assemblies for the genotypes shown in Table 4.1 revealed that all eight varieties lacking *TaASN-B2* had identical breakpoints.

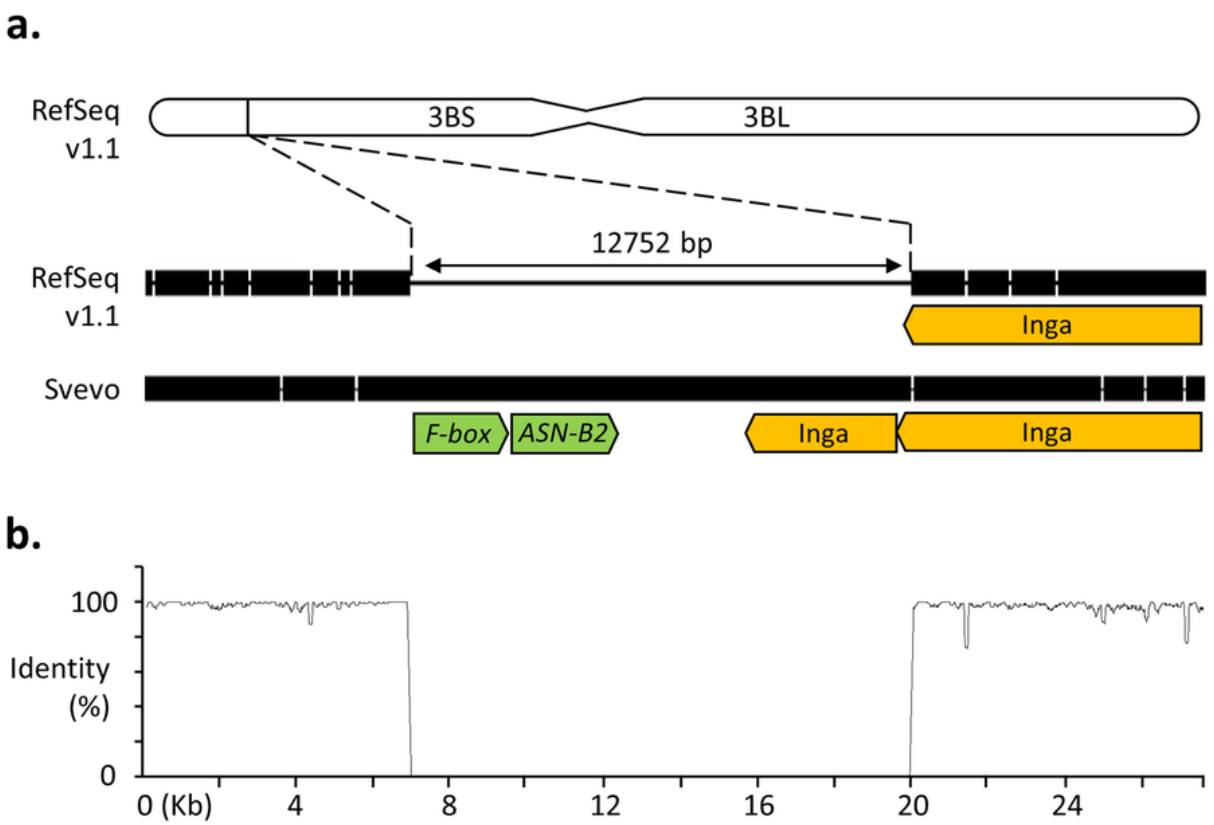


Figure 4.1. Variation at the *TaASN-B2* locus a. Diagram showing the location of the deletion in Chinese Spring and alignment with the corresponding region in variety Svevo. The deletion occurs at position 60,301,515 bp on chromosome 3B in the RefSeq v1.1 genome assembly. Notable gene and transposon annotations are shown. Breaks in RefSeq v1.1 and Svevo show indel differences between the alignment. b. Plot of the nucleotide sequence identity (%) between Chinese Spring and Svevo in regions flanking the deletion, from approximately 7 kb upstream to 8 kb downstream. Sliding window average of 100 bp

4.3.3. Wider screening for the presence/absence of ASN-B2

Because of their potential role in determining free asparagine concentrations in the wheat grain, allelic variation in *ASN2* genes was explored in a broader set of wheat germplasm. Comparison of the three *TaASN2* homeologues in the Cadenza genome revealed they share a common gene structure, each containing 11 exons (Fig. 4.2a). The encoded proteins shared > 99% identity at the amino acid level, with only eight polymorphic residues between homeologues (Fig. 4.2b). Although four of these polymorphisms fell in the glutamine amidotransferase (GATase)

domain (Fig. 4.2b), none were predicted to affect protein function according to the SIFT analysis (Table 4.1).

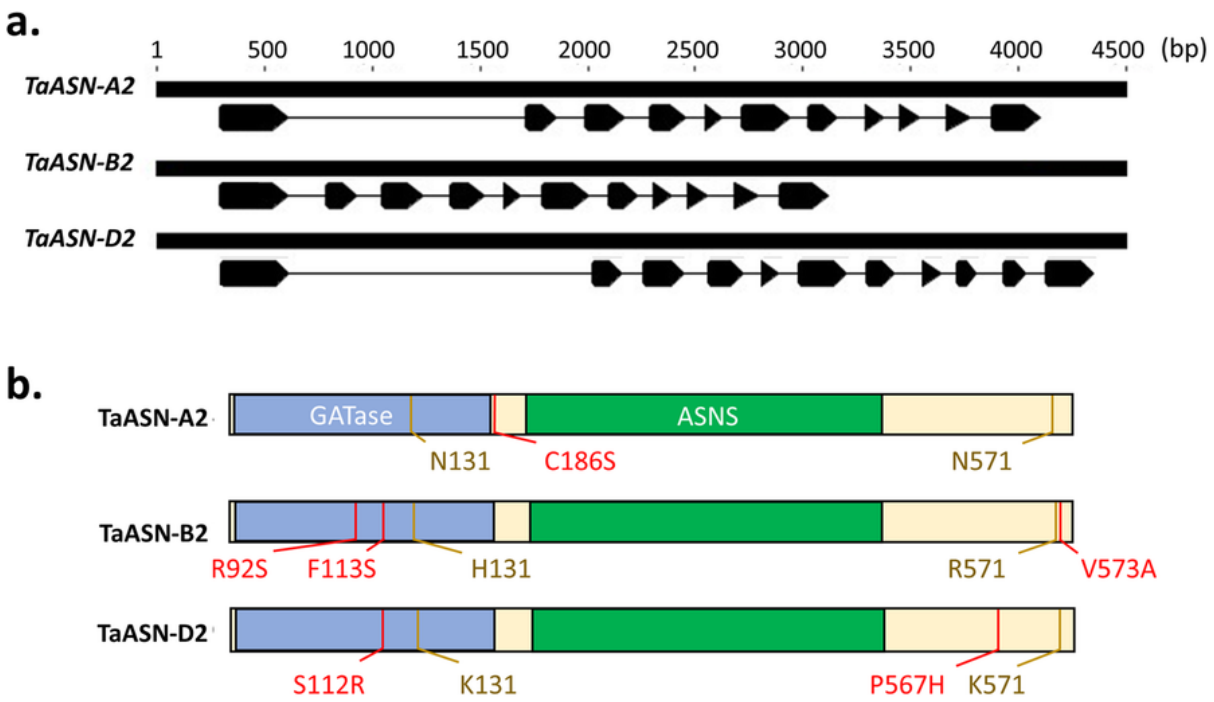


Figure 4.2. Structural characteristics of the *TaASN2* genes and proteins from Cadenza. **a**. Diagrammatic representation of the exon/intron structures of the *TaASN-A2*, *TaASN-B2* and *TaASN-D2* homeologues. Open arrows indicate exons and lines indicate introns. **b**. Diagram representing the structure and similarity of the proteins encoded by each *TaASN2* homeologue, showing the glutamine amidotransferase (GATase) domain (amino acids approx. 2–185) and asparagine synthetase (ASNS) domain (amino acids approx. 210–450). Residues that differ across all three homeologues are highlighted in yellow, whereas residues that differ in a single homeologue are highlighted in red

The length of intron 1 varied between homeologues and was 1104 bp in *TaASN-A2*, 1411 bp in *TaASN-D2*, but only 175 bp in *TaASN-B2* (Fig. 4.2a). A pair of redundant primers were designed to amplify a DNA fragment from the first intron of all three homeologues, allowing for the reaction products to be readily distinguished based on size and to detect the presence of *TaASN-B2*. A second pair of homeologue-specific primers were designed to anneal upstream and downstream of the deleted region containing *TaASN-B2*. These primers amplify a DNA fragment only in genotypes carrying this deletion as the amplicon is too large to amplify in

genotypes that possess *TaASN-B2*. The presence of *TaASN-B2* was, therefore, demonstrated by the amplification of a 434 bp product with the first primer pair and failure to amplify a PCR product using the second primer pair. The results of the analysis are shown in Fig. 4.3a-d and summarised in Table S2a. Overall, *TaASN-B2* was deleted in 52 of 63 UK winter wheat varieties assayed (82.5%) (Fig. 4.3e). The deletion was most common in the biscuit (G3) class (93.3%) and least common in the breadmaking (G1) (70%) class (Fig. 4.3e). An additional set of 24 global wheat varieties were analysed using a similar PCR assay (Fig. S4.2a) and the results are shown in Fig. S4.2b, and summarised in Table S4.2b. The *TaASN-B2* deletion was less common among these wheats than in the UK varieties, being present in just 50% of the genotypes (Table S4.2b).

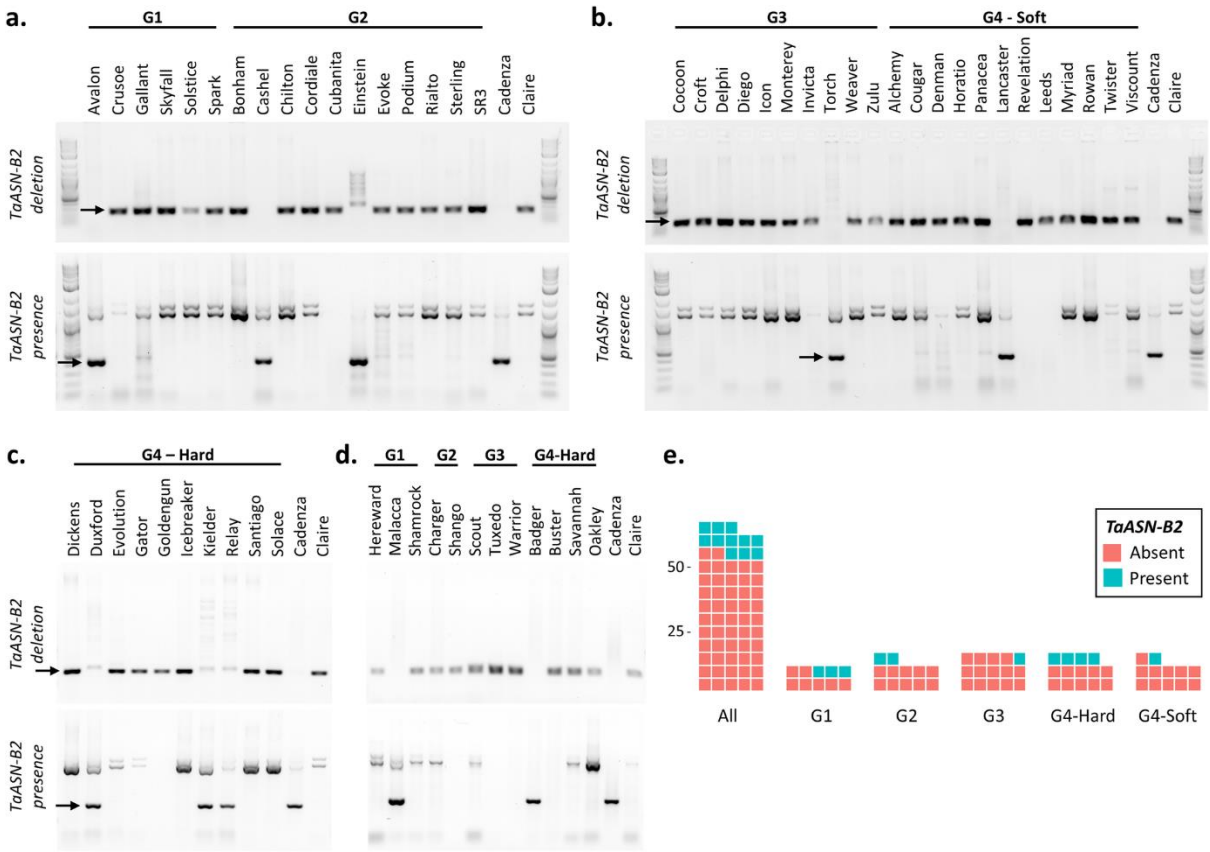


Figure 4.3. Presence/absence of *TaASN-B2* in UK wheat varieties. **a-d:** Electrophoresis gels of PCR products from assays to distinguish the presence and absence of *TaASN-B2* in a collection of UK wheat varieties grown across two years (Curtis et al., 2018). **a.** UK Flour Millers milling group 1 and 2 hard bread wheats (grown in 2012–2013). **b.** UK Flour Millers group 3 and 4 soft wheats (grown in 2012–2013). **c.** UK Flour Millers group 4 hard wheats (grown in 2012–2013). **d.** Remaining varieties grown only in 2011–2012. Varieties Cadenza

and Claire were used as controls for *TaASN-B2* presence and absence, respectively. The distinguishing PCR products are indicated with arrows. **e.** Diagram showing the frequency of the *TaASN-B2* deletion in 63 UK wheat varieties, separated into UK Flour milling groups: G1 (breadmaking), G2 (breadmaking potential), G3 (soft/biscuit), G4 (feed/other)

A selection of other wheat species was also screened for the presence of an *ASN-B2* gene (Fig. 4.4). An *ASN2* gene was identified in *Aegilops speltoides* (genome BB); however, while an *ASN-B2* gene was present in some tetraploid wheat genotypes (genomes AABB) it was absent in others (Fig. 4.4). Both pasta wheat (*T. turgidum* ssp. *durum)* varieties assayed in the study, Svevo and Kronos, were shown to have an *ASN-B2* gene, as was Polish wheat (*T. turgidum* ssp. *Polonicum*), but the gene was absent in rivet wheat (*T. turgidum* ssp. *turgidum*). There was some ambiguity in the result for makha wheat (*T. macha*) in that there was a clear positive result for the presence of the *ASN-B2* gene but a faint band amplified in the assay for the deletion (Fig. 4.4). Further investigation is required to establish the presence/absence of *TaASN-B2* in *T. macha*.

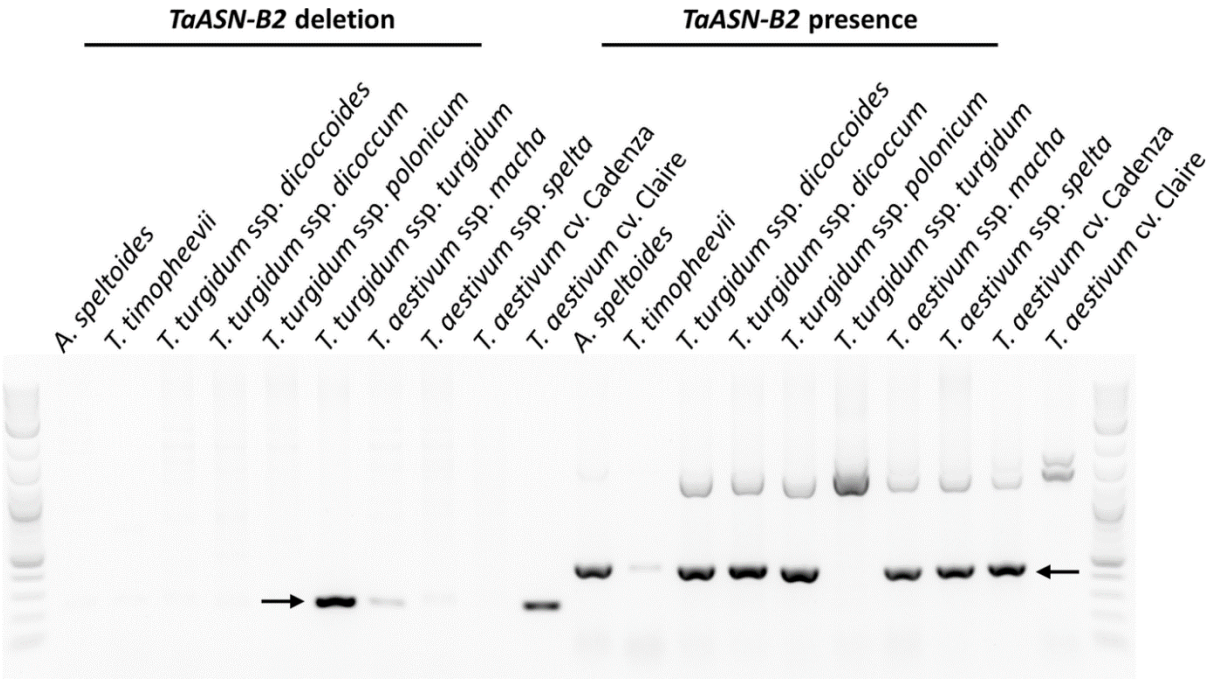


Figure 4.4. Presence/absence of *ASN-B2* in different wheat species. Electrophoresis gel of PCR products from assays to distinguish the presence and absence of *ASN-B2* in a selection of wheat species. *T. aestivum* varieties Cadenza and Claire were used as controls for *ASN-*

B2 presence and absence, respectively. The distinguishing PCR products are indicated with arrows.

4.3.4. Expression profiles of wheat ASN genes during development

Raw sequencing reads from public RNA-seq datasets were mapped to the IWGSC RefSeq v1.1 genome assembly to provide a comprehensive overview of the expression profile of each *TaASN* gene. The expression values for each dataset are provided in supplementary file 4.1 (available at github.com/JosephOddy/ThesisSupplementaryFiles) as mean Transcripts Per Million (TPM) values. In a dataset encompassing roots, leaves, stems, spike and grain, each sampled at three developmental stages (Ramírez-González et al., 2018), *TaASN1* transcript levels were highest in young roots and leaves, whereas the three homeologues of *TaASN3.1* and *TaASN3.2* showed a broader expression profile, with transcripts detected in all assayed tissue types across different stages of development (Fig. 4.5a). *TaASN4* homeologues were also broadly expressed, with *TaASN-A4* transcript levels highest in root and spike tissues, and *TaASN-B4* and *TaASN-D4* more highly expressed during stem development (Fig. 4.5a). As shown previously (Gao et al., 2016; Curtis et al., 2019), *TaASN2* showed a grain-specific expression profile, with transcript levels highest at Zadoks stage 85 (Z85), which corresponds to the soft dough stage (Zadoks et al., 1974) (Fig. 4.5a). Furthermore, *TaASN-A2* accounted for 83.3% of all *TaASN* transcripts in grain tissues at Z85, while *TaASN-D2* contributed just 3.0%, consistent with previous results (Curtis et al., 2019).

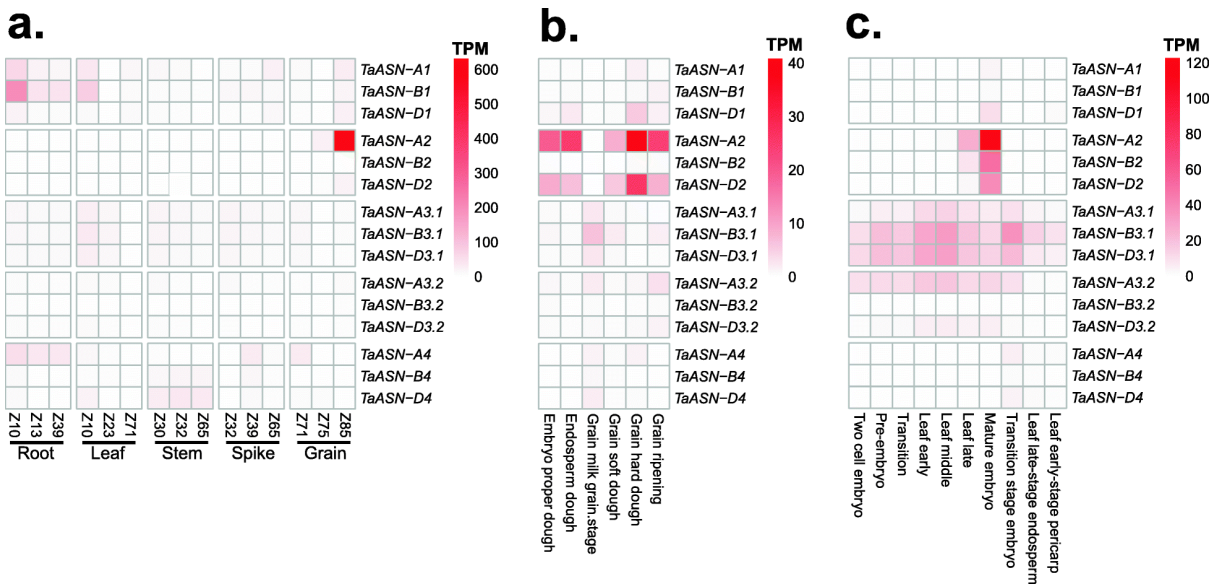


Figure 4.5. Expression profiles of *TaASN* genes in wheat derived from RNA-seq mapping data. **a.** Expression of *TaASN* genes in Chinese Spring across different tissues and developmental stages. **b.** Expression of *TaASN* genes in grain tissues of the variety Azhurnaya across six stages of development. **c.** Expression of *TaASN* genes in variety AC Barrie across an embryo development timecourse.

To further explore the expression of *TaASN* genes in the grain, an expression dataset from six stages of grain development in the variety Azhurnaya was analysed (Ramírez-González et al., 2018). As expected, high transcript levels of *TaASN-A2* and *TaASN-D2* were found in five developmental stages which, when combined, accounted for between 69 and 86% of all *TaASN* transcripts in these tissues (Fig. 4.5b). The exception was the grain milk stage, where *TaASN3.1* transcript levels were higher than *TaASN2* (Fig. 4.5b), as shown previously (Gao et al., 2016). *TaASN-B2* transcripts were detected at negligible levels in this dataset, suggesting its deletion in the Azhurnaya genome.

Analysis of expression data from an embryo development timecourse in the common wheat variety AC Barrie (Nirmal et al., 2016) revealed that *TaASN2* transcript levels were highest in the mature embryo stage (Fig. 4.5c). Among *TaASN2* homeologues, *TaASN-A2* was again the most highly expressed gene, while *TaASN-B2* and *TaASN-D2* were expressed at similar levels (Fig. 4.5c). However, *TaASN2* transcripts were detected only at negligible levels at all other developmental timepoints, including earlier stages of embryo development and in endosperm and pericarp tissues, where *TaASN3.1* transcripts were more abundant (Fig. 4.5c). Taken

together, these data confirm the specific expression of *TaASN2* in grain tissues and the mature embryo, and indicate a broader role for *TaASN3.1* genes across development, including the early stages of embryo development.

4.3.5. Inter-varietal variation in TaASN expression profiles during grain development

To analyse variation in *ASN* transcript levels in wheat grain, RNA-seq reads were mapped from grain samples taken at 14 days post anthesis (DPA) and 30 DPA from 27 worldwide wheat varieties (Nirmal et al., 2016). At 14 DPA, total *TaASN3.1* transcript levels ranged from 2 to 32 TPM and were greater than *TaASN2* in 22 of the 27 varieties assayed (Fig. 4.6a), consistent with previous results (Curtis et al., 2019). At 30 DPA, *TaASN2* homeologues were the most highly expressed asparagine synthetase genes in all varieties assayed (Fig. 4.6b). At this later timepoint, total *TaASN2* transcript levels showed large variation between genotypes, ranging from 28 to 242 TPM (Supplementary file 4.1). Several lines exhibited very low *TaASN-B2* transcript levels and the deletion of this gene was confirmed in five of these lines using the PCR assay (Fig. S4.2b; Table S4.2b). There were also lines with readily detectable *TaASN-B2* transcripts (>five TPM), and the presence of this gene was confirmed for four of these lines (Fig. S4.2b; Table S4.2b).

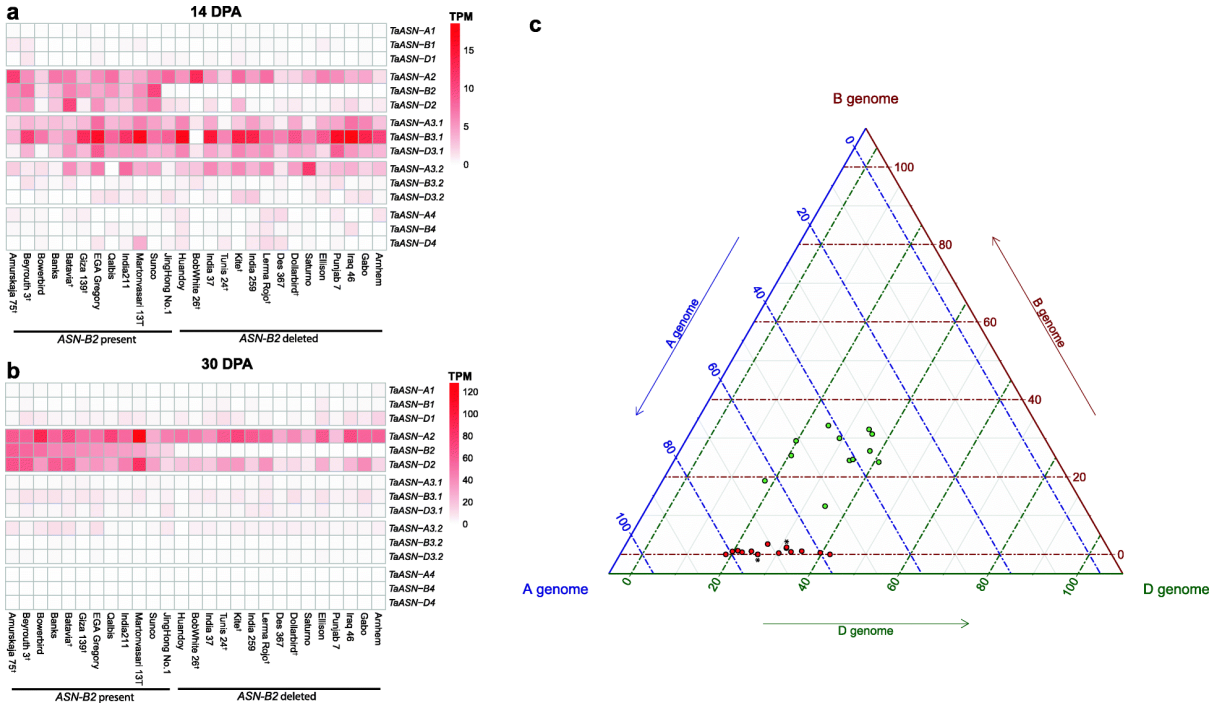


Figure 4.6. Expression of *TaASN* genes in 27 worldwide wheat varieties from RNA-seq mapping data (Nirmal et al., 2016). a. Grain tissues 14 DPA. b. Grain tissues 30 DPA. † indicates varieties for which the presence or absence of *TaASN-B2* was confirmed by PCR c. Ternary plot illustrating the relative contribution of A, B and D genome *TaASN2* homeologues to total *TaASN2* gene expression in grain tissues at 30 DPA in each variety. Varieties with *TaASN-B2* present are indicated with green circles, varieties with *TaASN-B2* deleted are indicated with red circles

A ternary plot showing the relative contributions of each homeologue to overall *TaASN2* transcript levels in the grain at 30 DPA in different wheat varieties (Fig. 4.6c) revealed that *TaASN-A2* transcript levels were generally greater than *TaASN-D2* in varieties with very low *TaASN-B2* transcript levels (likely associated with the deletion of this gene in these varieties). By contrast, in varieties with relatively high *TaASN-B2* expression, *TaASN-A2* and *TaASN-D2* were generally more evenly expressed, and in four varieties, *TaASN-D2* transcript levels were higher than *TaASN-A2* (Fig. 4.6c). Nevertheless, overall TPM values for *TaASN-A2* and *TaASN-D2* were not higher in varieties that lacked *TaASN-B2* compared with those in which *TaASN-B2* was present (Supplementary file 4.1), so there was no evidence of increased expression of these genes to compensate for the lack of *TaASN-B2* transcripts.

To investigate the expression dynamics further, the expression of the *TaASN2* homeologues was also analysed by RT-qPCR in two wheat varieties possessing *TaASN-B2* (Cadenza and Duxford) and two varieties lacking it (Spark and Claire) (Fig. 4.7). The results of the analysis of variance for this experiment are shown in Table S3a, revealing significant effects ($p < 0.001$) of variety, timepoint, and homeologue, and the interactions between these factors, on relative expression levels. In Cadenza and Duxford, mean *TaASN-A2* expression was the highest of the three homeologues across all timepoints (14, 21 and 28 DPA), whereas mean *TaASN-D2* expression was the lowest (Fig. 4.7a and b). In both varieties, mean *TaASN-B2* expression was greater than *TaASN-D2* expression at all timepoints, and in Cadenza at 21 DPA, matched the levels of *TaASN-A2* expression (Fig. 4.7a). In Claire and Spark, mean *TaASN-A2* expression was greater than *TaASN-D2* expression in all samples, and showed similar expression dynamics across timepoints to Cadenza and Duxford (Fig. 4.7c and d). Notably, there was no evidence of higher expression of *TaASN-A2* or *TaASN-D2* in Claire and Spark compared with Cadenza and Duxford.

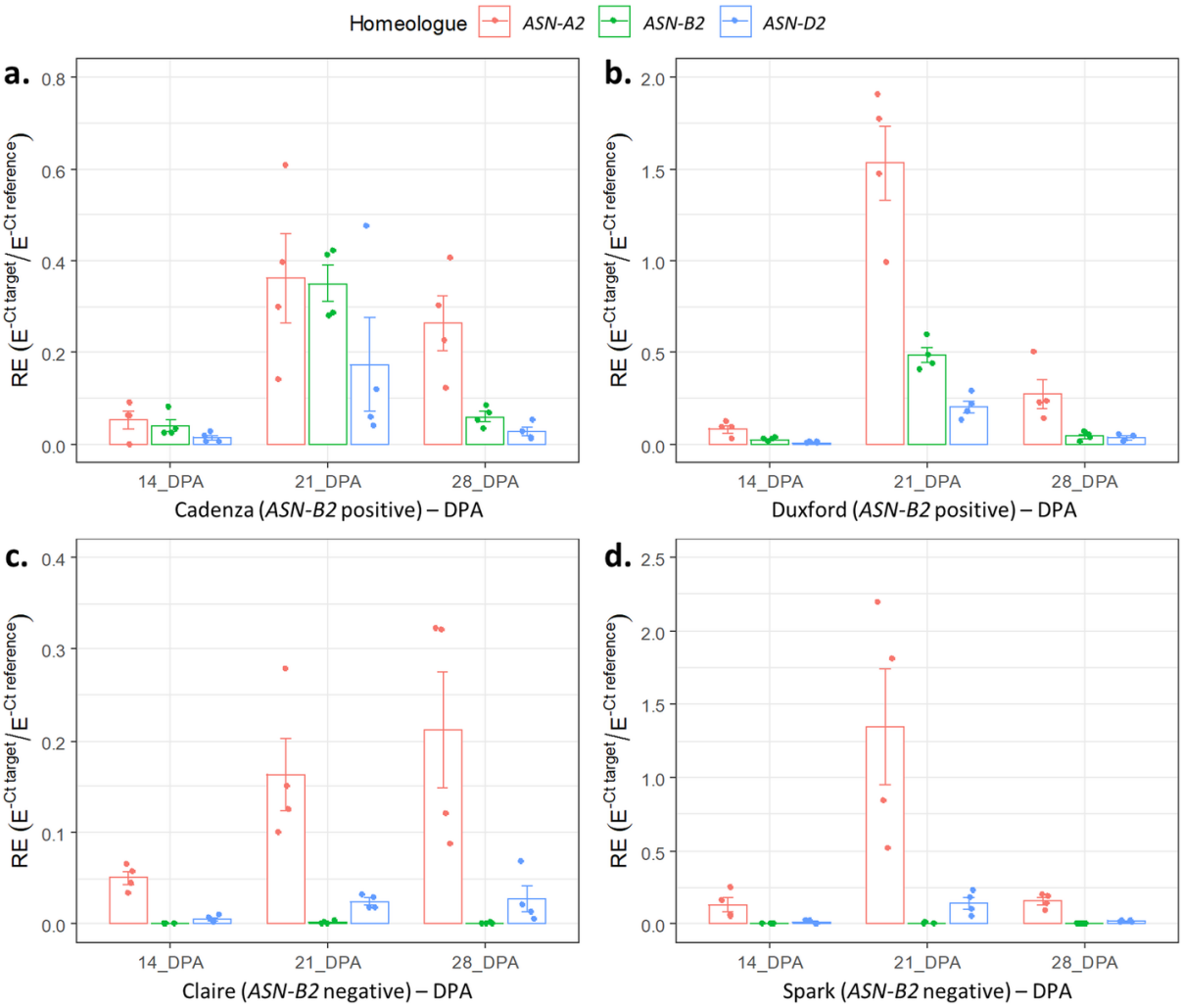


Figure 4.7. Expression analysis by RT-qPCR of the *TaASN2* homeologues in the embryo of four varieties of wheat either possessing the *TaASN-B2* gene (Cadenza (**a.**) and Duxford (**b.**)) or lacking one (Claire (**c.**) and Spark (**d.**)) at 14-, 21-, and 28-days post anthesis (DPA). Expression levels are relative to three reference genes (*GAPDH*, *PROSM*, and *SDH*). RE (relative expression), E (PCR efficiency calculated by LinRegPCR), Ct (threshold cycle). Error bars show standard error of the mean.

Contribution of TaASN-B2 to free asparagine concentration in the grain

The screen of varieties for the presence or absence of *TaASN-B2* (Fig. 4.3, Table S4.2) included 63 UK varieties for which free asparagine concentration in the grain had been determined in field trials grown in the UK over two growing seasons (2011–2012 and 2012–2013) (Curtis et al., 2018). This meant that an assessment could be made of the effect of the *TaASN-B2* deletion

on free asparagine concentrations in the grain. Of the 63 varieties in the field trials, eleven possessed *TaASN-B2* while 52 did not.

The grain from these field trials had been produced in plots in which sulphur was either supplied or withheld (Curtis et al., 2018). We analysed the effect of *TaASN-B2* alongside the other variables in these trials by ANOVA (Table S4.3b), which revealed a significant ($p < 0.001$) effect of the *TaASN-B2* deletion in the 2011–2012 field trial: varieties without *TaASN-B2* had 13.18% less free asparagine relative to those with *TaASN-B2* (Fig. 4.8a). There was no significant effect ($p > 0.05$) of the deletion by itself in the 2012–2013 field trial, but there was a significant ($p < 0.001$) interaction between *TaASN-B2* presence/absence and sulphur treatment: there was no significant difference ($p > 0.05$) in free asparagine concentrations between varieties with and without *TaASN-B2* under sulphur deficiency, but varieties without *TaASN-B2* had 32.60% less free asparagine ($p < 0.01$) than those with *TaASN-B2* under sulphur sufficiency (Fig. 4.8c).

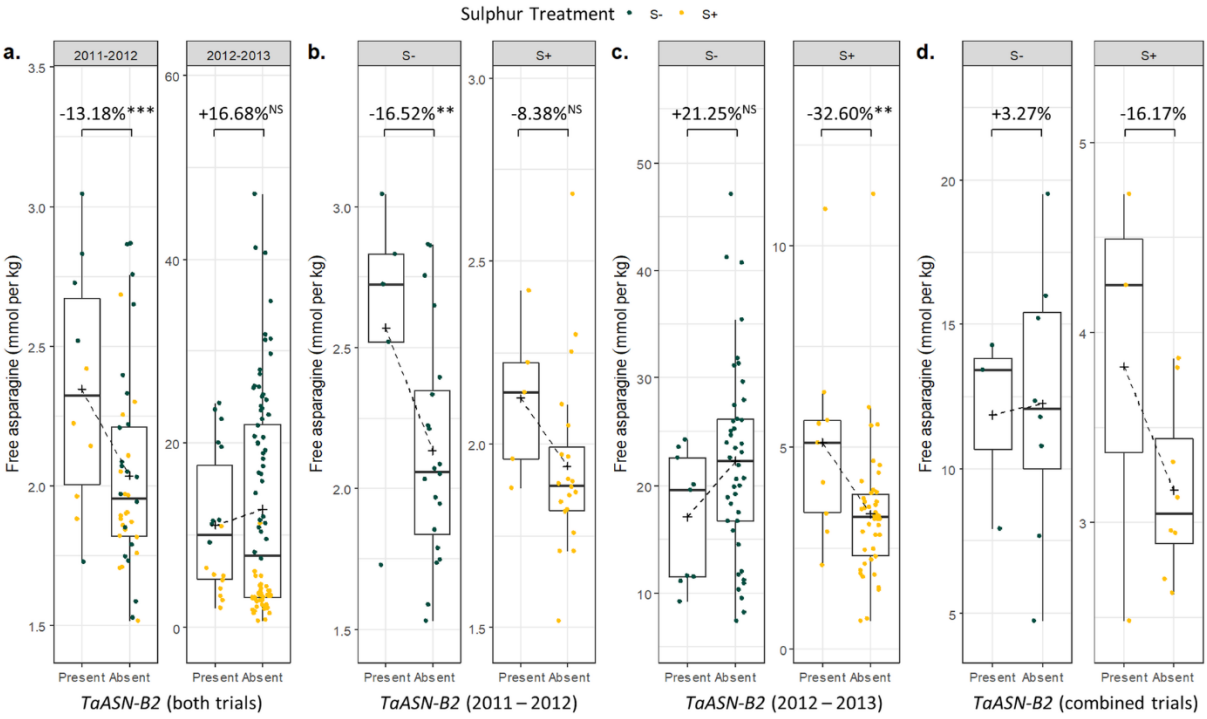


Figure 4.8. Concentrations of free asparagine in the grain of different varieties of wheat in which the *TaASN-B2* gene was either present or absent in two field trials, 2011–2012 and 2012–2013, showing the effect of the *TaASN-B2* deletion under different sulphur treatments (S- and S+) and the interaction between the presence/absence of the gene and treatment. (a) Back-transformed data from the 2011–2012 and 2012–2013 trials with concentrations from the

S- and S+ treatments plotted together. (b) Back-transformed data from the 2011–2012 season with concentrations from the S- and S+ treatments plotted separately. (c) Back-transformed data from the 2012–2013 season with concentrations from the S- and S+ treatments plotted separately. (d) Back-transformed predicted means from the REML analysis for the 11 varieties grown in both trials. Crosses indicate the mean of each group. Boxes show the interquartile range and median. Whiskers show the smallest and largest value within 1.5 times of the interquartile range above and below the 75th and 25th percentile, respectively. Significance scores for the main effect of *TaASN-B2* in (a) were taken from separate ANOVA analyses performed on $\log_e$ transformed data. Multiple comparisons with Bonferroni correction were performed following the ANOVA analyses for *TaASN-B2* group comparisons in 2011–2012 (b) and 2012–2013 (c). NS = $p > 0.05$, ** = $p < 0.01$, *** = $p < 0.001$.

The REML analysis based on the complete combined datasets also identified a significant effect of *TaASN-B2* as well as an interaction between *TaASN-B2* and sulphur treatment (Table S4.3b). This analysis also further demonstrated that the interaction between *TaASN-B2* and sulphur treatment changed by year, as is suggested by Fig. 4.8b and c, and that there is a significant interaction between year and sulphur treatment. Prediction and analysis of means from the REML analysis for the 11 varieties common to both trials (of the 63 that were analysed in total) suggests that the *TaASN-B2* deletion effect is greatest under sulphur sufficiency across different growing seasons (Fig. 4.8d). The effect of sulphur deficiency differed greatly between the trials, so it is difficult to reach a definitive conclusion regarding the effect of *TaASN-B2* under sulphur deficiency based on the predicted means of the REML analysis.

4.4. Discussion

4.4.1. A natural deletion of ASN-B2 in wheat

In this study, wheat genomic resources were utilised to characterise a natural deletion that includes the complete *TaASN-B2* gene and to design a molecular assay to trace its frequency in a diverse set of wheat germplasm. The break points for this deletion event were identical in all eight common wheat genome assemblies that were analysed, which would be consistent with a single, common origin for this allele in hexaploid wheats. The deleted region is in close proximity to a Ty1-*copia* transposable element (Fig. 4.1a), one of the most abundant classes of

LTR retrotransposons in the wheat genome (Wicker et al., 2018), and a potential causative agent for the deletion. The *ASN-B2* gene was intact in the B genome of *Ae. speltoides* (genome BB) (Fig. 4.4) and this genome is related to the B genome progenitor of domesticated wheat species. However, the direct ancestor of the hybridisation event originating emmer wheat (genomes AABB) has yet to be identified (Dubcovsky et al., 2007), so firm conclusions of the origins of the *ASN-B2* deletion in diploid wheat progenitors cannot be drawn. Some wild and domesticated tetraploid wheats carried an intact *ASN-B2* gene whereas in others it was absent, which, as discussed elsewhere (Marcussen et al., 2014; Raffan & Halford, 2021), is consistent with the occurrence of independent hybridisation events in wheat's evolutionary history. Studying haplotype variation at this locus in more diverse wheat germplasm collections, including more wild and domesticated emmer wheats from different subpopulations within the fertile crescent (Dubcovsky et al., 2007), could shed light on the origins of the *ASN-B2* deletion.

Among different classes of UK winter common wheat varieties, this deletion was found at high frequencies (Fig. 4.3e). The unbalanced distribution of this allele could result from the lack of genetic diversity during early selections in UK wheat variety development, be an artefact of the selection of varieties analysed in this study, or simply have occurred by chance, indicating that this allele is under neutral selection. However, it is also possible that this allele is subject to direct or indirect selection due to a positive impact on plant fitness or performance. If so, it is unlikely to have been selected due to its association with free asparagine content in the grain described in the current study, because selections based on free asparagine content have been made only very recently, if at all. Based on the grain-specific expression profile of *TaASN-B2*, other possible traits that might account for positive selection of the deletion are pre-harvest sprouting resistance or other quality traits shown in some studies to be correlated with free asparagine concentrations (Simsek et al., 2014). However, other studies have found no association between free asparagine content and a host of baking quality traits (Rapp et al., 2018), so this association requires continued analysis. A further possibility is that this deletion has been selected indirectly due to a beneficial genetic variant in linkage disequilibrium with this locus. Evidence of marker-trait associations for other agronomic traits in this region would support this hypothesis, although it would be difficult to test directly. A more detailed functional characterisation of the *ASN2* genes focused on grain development and quality traits may reveal additional, previously unidentified roles of this gene; similar to the moonlighting

roles found for asparagine synthetase in yeast and mammalian cells (Noree et al., 2018; Noree et al., 2019).

4.4.2. Expression profiles of wheat asparagine synthetase genes

Our expression data confirmed previous findings (Gao et al., 2016; Curtis et al., 2019) that, in many varieties, *TaASN-A2* is the most highly expressed *TaASN2* homeologue during grain development (Figs. 4.5, 4.6 and 4.7). In five of the 12 varieties possessing *TaASN-B2* that have been assayed by RNA-seq, *TaASN-B2* contributes a higher proportion of *ASN2* transcripts than *TaASN-D2*, whereas *TaASN-D2* expression is greater in the other seven varieties (Fig. 4.6c). Higher *TaASN-B2* expression over *TaASN-D2* was also observed in varieties Cadenza and Duxford via RT-qPCR, throughout the sampling time course from 14 to 28 DPA (Fig. 4.7). In every variety lacking *TaASN-B2*, *TaASN-A2* contributed a greater proportion of *ASN2* transcripts than *TaASN-D2* (Fig. 4.6c). By contrast, in varieties with *TaASN-B2* present the three homeologues were more evenly expressed, although *TaASN-A2* still contributed the greatest proportion of *ASN2* transcripts in most lines (Fig. 4.6c). However, based on our combined RNA-seq and RT-qPCR data, the loss of *TaASN-B2* is associated with an overall reduction of *ASN2* transcript levels and no compensatory upregulation of the A or D homeologues, which is consistent with the effect of this deletion on grain free asparagine concentration. This is in contrast to other biosynthetic pathways, for example in GA signalling (Middleton et al., 2012), in which feedback mechanisms modulate the transcript levels of biosynthetic genes to compensate for loss of expression.

Our expression analyses also provide additional insight into the potential role of *TaASN3.1* during embryo and grain development. Transcript levels of this gene were most abundant at earlier stages of grain development, while *TaASN2* transcripts predominated later in grain development (Figs. 4.5 and 4.6). In most, but not all, varieties, expression of *TaASN3.1* remained higher than *TaASN2* even by 14 DPA, in contrast to a previous study in which *TaASN2* was found to be the most highly expressed asparagine synthetase gene in the embryo and endosperm of the common wheat variety Spark and doubled haploid line SR3 at this developmental stage (Gao et al., 2016; Curtis et al., 2019). This is most likely explained by differences in growth conditions and developmental rates; indeed, differential rates of development and asparagine synthetase gene expression were observed between Spark and SR3 even under identical growth conditions (Curtis et al., 2019). Therefore, it would be interesting to characterise the function of *TaASN3.1* in wheat to help understand the extent to

which this gene contributes to asparagine biosynthesis in the grain, and its importance for grain development. Allelic variation in this gene could also contribute to reduced free asparagine content in the grain, and it would be possible to test this hypothesis either by characterizing the effect of potentially disruptive natural *TaASN-A3.1* alleles carried by some varieties (Table 4.1), or by targeted mutagenesis. This approach could also be applied to characterise the function of other wheat asparagine synthetase genes, to help assess the possibility of integrating genetic variation in this family into wheat breeding programmes.

4.4.3. Breeding for reduced grain free asparagine content

The *TaASN-B2* deletion was associated with a 13.18% reduction in grain free asparagine concentration in the 2011–2012 field trial (Fig. 4.8a) and with a 32.60% reduction under sulphur sufficiency in the 2012–2013 field trial (Fig. 4.8c), showing that this variant can contribute to reduced grain free asparagine concentrations in field conditions. The milder effect of sulphur deficiency on free asparagine concentrations in 2011–2012 was previously noted (Curtis et al., 2018) and is likely why the effect of *TaASN-B2* was seen across both sulphur treatments. In contrast, sulphur deficiency caused a much greater increase in free asparagine concentrations in the 2012–2013 trial, bringing about an interaction between sulphur treatment and *TaASN-B2* (Table S4.3b) and demonstrating that the deletion does not have an effect under severe sulphur deficiency when free asparagine concentrations are higher (Fig. 4.8c). This is consistent with earlier studies showing strong environmental (E) as well as genetic (G) impacts on free asparagine content, as well as complex G × E interactions (Raffan & Halford, 2019). It implies that the effect of the *TaASN-B2* deletion can be maximised by employing appropriate crop-management strategies to reduce stress.

This retrospective analysis of the *TaASN-B2* deletion on grain free asparagine concentrations was limited because varieties in the trials were not originally selected based on the presence or absence of *TaASN-B2*, so there were many more varieties without *TaASN-B2* than those with it in this analysis. Consequently, it will be important to characterise the impact of the *TaASN-B2* deletion in replicated field trials in additional and more diverse environments, including phenotyping a broader set of traits to confirm that there are no detrimental pleiotropic effects associated with this allele. To minimise the impact of other genetic variation, these trials could be performed in a common genetic background, either by developing near isogenic lines, or by directly inducing genetic variation by mutagenesis.

Measuring free asparagine concentrations in grain directly is expensive, requires specialist analytical equipment and is often impractical for breeders, but the use of our simple PCR screen could enable wheat breeders to exclude genotypes that are more likely to have high free asparagine concentrations in the grain when grown in sulphur-sufficient conditions (Fig. 4.8b). Although the effect is mild, this could contribute to reducing the public health risk associated with dietary acrylamide and help food manufacturers to comply with the difficult and evolving regulations on the presence of acrylamide in their products. This deletion can be fully exploited in all regulatory environments due to its natural origins, in the same way that other quality traits have been exploited in different crop species. For example, some barley genotypes possess reduced cadmium accumulation due to the natural insertion of a transposable element upstream of a cadmium transporter, and the allele carrying this insertion can be used without restriction (Lei et al., 2020).

Although the impact of this specific variant may be limited for UK wheat breeders because of its high frequency in UK winter wheat varieties (Fig. 4.3e, Table S4.2a), there may be greater opportunity to apply this allele in other regions of the world, since just 50% of wheat varieties from a selection of global varieties carried the deletion (Table S4.2b). This panel included varieties from Australia, Africa and Europe, so it would be worthwhile exploring the frequency of this deletion in broader collections of wheat germplasm. Although only two durum wheat varieties were included in the current study, both carried the *ASN-B2* gene, possibly indicating that the historic hybridisation events giving rise to durum wheats may have included the *ASN-B2* gene at higher frequencies than for common wheat. Therefore, durum wheat breeders may have an opportunity to reduce free asparagine concentration in the grain by identifying and selecting genotypes carrying this deletion. Furthermore, because the durum wheat genome is tetraploid, those varieties lacking the *ASN-B2* homeologue may show a proportionally greater reduction in grain asparagine than that found in hexaploid common wheat. A major use of durum wheat is for pasta production and although acrylamide is present in pasta, it is at relatively low concentrations (EFSA, 2015). However, durum wheat grain is also incorporated into grists for making products in which acrylamide concentrations are likely to be higher, such as pizza bases, pitta bread and other flatbreads.

Although we detected only two different *TaASN-D2* alleles (one only found in the variety SY Mattis) and no variation in *TaASN-A2* (Table 4.1), it is possible that broader screens of more diverse germplasm may yield additional natural variants that could be integrated into breeding programmes to select for reduced free asparagine concentration. Previous association mapping

studies have identified QTL controlling free asparagine content (Emebiri, 2014; Rapp et al., 2018), but these QTL did not map to regions of the genome containing asparagine synthetase genes. However, reverse genetics tools, such as EMS- or CRISPR/Cas9-induced mutagenesis, provide the potential to engineer allelic diversity that does not exist among wheat germplasm, including combinations of recessive mutations that are unlikely to be selected due to functional redundancy in polyploid genomes (Krasileve et al., 2017; Zhang et al., 2020). The power of this approach was demonstrated in a recent study in which plants exhibiting reductions in free asparagine concentrations of up to 90% were developed by editing all six *ASN2* alleles using CRISPR/Cas9 (Raffan et al., 2021). Furthermore, the presence of three homeologues of this gene allows for selection of combinations of allelic knockouts that may allow breeders to balance reduced free asparagine content with other grain development traits. Although this would be a powerful and rapid approach to engineer and characterise potentially valuable genetic variation, it is important to note the complex and dynamic regulatory landscape that currently restricts applications of CRISPR-Cas9 in plant breeding in some regions of the world (Schmidt et al., 2020). Such restrictions do not apply to the use of EMS in plant breeding, so crossing genotypes that naturally lack *TaASN-B2* with genotypes containing EMS induced mutations in *TaASN-A2* and *TaASN-D2* would be a feasible alternative.

4.5. Conclusions

Characterisation of natural allelic variation in the wheat asparagine synthetase gene family identified a deletion of just under 13 kb encompassing *TaASN-B2* that is present at high frequencies among UK winter wheat varieties. The deletion was also present in some wild emmer wheats, suggesting its ancient origins and retention during domestication and modern breeding. The allele carrying the deletion was associated with a reduction in free asparagine content in field experiments and could be selected using an inexpensive PCR assay to help breeders develop low-asparagine wheat varieties.

4.6. Methods

4.6.1. Genomic analyses

Nucleotide sequence data for the wheat *ASN* genes from different wheat genotypes were obtained using the BLAST tools of the 10+ Wheat Genomes Project (https://webblast.ipk-

gatersleben.de/wheat_ten_genomes/), the Grassroots Genomics Project (https://wheatis.tgac.ac.uk/grassroots-portal/blast) (Bian et al., 2017), and the Graingenes database (https://wheat.pw.usda.gov/cgi-bin/seqserve/blast_wheat.cgi) (Blake et al., 2019). Some *ASN* genes lacked complete sequence information, and these exceptions are described in Table S4). Geneious Prime 2020.1.2 was used for alignments and sequence identity analyses between genes. The annotated genome from the pasta wheat (*T. durum*) variety Svevo (https://www.interomics.eu/durum-wheat-genome) (Maccaferri et al., 2019) was used to compare the genomic region containing *TaASN-B2* with the corresponding region in the RefSeq v1.1 genome from the common wheat (*T. aestivum* L.) landrace Chinese Spring (from Ensembl plants https://plants.ensembl.org/wheat) (IWGSC, 2018). The softberry-FGENESH tool (Solovyev et al., 2006) was used to identify putative genes from the 12,770 bp deleted region encompassing *TaASN-B2* from the variety Jagger. This analysis identified five putative ORFs. Each was analysed with HMMScan (Potter et al., 2018) using an e-value cutoff of 0.05, which confirmed the presence of *TaASN-B2* and a second gene encoding a protein containing an F-box PFam domain (PF00646). Transposon annotations were confirmed using the BLAST tool in TREP (TRansposable Elements Platform) (http://botserv2.uzh.ch/kelldata/trep-db/index.html) (Wicker et al., 2002). SIFT analysis (Sim et al., 2012) was performed by comparing each ASN protein from Chinese Spring with the protein encoded by the orthologous gene from other wheat varieties. For *TaASN-B2*, the protein from Jagger was used as a reference.

4.6.2. Plant materials and germination

Seeds of UK cultivars were either maintained at Rothamsted Research or obtained from stocks produced in the field trials studied here (Curtis et al., 2018). Other wheat varieties were obtained from USDA-ARS National Small Grains Collection (https://www.ars.usda.gov/) and the Germplasm Resource Unit at the John Innes Centre (www.seedstor.ac.uk). The names of all varieties included in this study are listed in Table S2. Seed surface sterilisation was performed by incubating seeds in 70% ethanol for 10 min and then in 20% (v/v) sodium hypochlorite solution for 60 min with gentle agitation to ensure homogenous sterilisation of the seeds. Seeds were subsequently washed four times with sterile distilled water and left to germinate under continuous light at room temperature in sterile 90 mm Petri dishes on wet filter paper, sealed with Parafilm (Fisher Scientific Ltd., Loughborough, UK). For older seeds, plates

were wrapped in foil and incubated at 4 °C for two to seven days to break dormancy, before transferring to continuous light and room temperature for germination. For seeds unable to germinate using either method, the embryo was dissected and placed in 90 mm Petri dishes containing MS media (4.4 g/L MS salts (Murashige & Skoog, 1962), 3% sucrose (30 g/L), pH 5.8, 7 g/L agar). These plates were then sealed with Parafilm and left to germinate at room temperature and continuous light.

4.6.3. DNA extraction

DNA was extracted from leaf material of seedlings using the Wizard® Genomic DNA Purification Kit according to the manufacturer's instructions (Promega (UK) Ltd., Southampton, UK). For seeds that failed to germinate using the above methods, embryos were dissected from multiple seeds and ground together into a fine powder for DNA extraction by the CTAB method (Sambrook et al., 1989). DNA quality and abundance were assayed using a NanoDrop™ 1000 Spectrophotometer (Thermo Fisher Scientific).

4.6.4. PCR assays to detect TaASN-B2

Homeologue-specific primers ASN-B2-Deletion-F and ASN-B2-Deletion-R (Table S4.5) were designed to anneal upstream and downstream of the deletion site so that the absence of *TaASN-B2* could be demonstrated as a positive result with the amplification of a 232 bp DNA fragment. Another pair of primers to amplify the first intron of all *TaASN2* homeologues, ASN-2-Universal-F and ASN-2-Universal-R (Table S4.5), was designed to test for and distinguish the presence or absence of all three *TaASN2* homeologues based on size. These forward and reverse primers amplified DNA fragments of sizes 1363 bp, 434 bp and 1670 bp, corresponding to *TaASN-A2*, *TaASN-B2* and *TaASN-D2*, respectively. These two primer sets were used in combination to verify the presence or absence of *ASN-B2.*

These primers were used in PCR reactions in volumes of 25 µL using 1 × DreamTaq™ PCR Master Mix (1.5 mM $MgCl_2$) (Thermo Fisher Scientific, Epsom, UK) and including 1 µM of each primer and 50–150 ng of genomic DNA. Cycling conditions were identical for both primer sets: 5 min denaturation at 96 °C; 32 cycles of 30 s denaturation at 96 °C, 30 s annealing at 60.5 °C, 1.5 min extension at 72 °C; 10 min final extension at 72 °C. Reactions were analysed by electrophoresis on an agarose gel (1.0% w/v, ethidium bromide staining) with 1 kb Plus

DNA Ladder (NEB, UK) and visualised using UV light in the Geldoc imaging system (BioRad, USA).

A separate PCR assay was used to confirm the presence or absence of *ASN-B2* in a set of 24 global wheat varieties (Fig. S4.2). This assay used two sets of primers in a single PCR to amplify different amplicons depending on the presence or absence of *ASN-B2*. The primers used were ASN-B2_qF1 (P3) and ASN-B2_qR1 (P4), which amplify a DNA fragment of 125 bp in varieties carrying *TaASN-B2*, along with ASN-B2_CS_F3 (P1) and ASN-B2_CS_R1 (P2), which amplify a DNA fragment of 189 bp in varieties lacking the *TaASN-B2* gene (Fig. S4.2). The PCR mixture included, in a total volume of 25 µL, 0.2 µM of primers P1 and P2 and 0.24 µM of primers P3 and P4, 1 × Standard *Taq* buffer, 250 ng template DNA and 0.125 µL *Taq* polymerase (New England Biolabs, Ipswich, MA, USA). Amplification was carried out using the following conditions: 95 °C for 30 s; 35 cycles of: 95 °C for 15 s, 59 °C for 30 s, 68 °C for 30 s; 68 °C for 5 min. Amplified DNA fragments were separated by electrophoresis on a 3% agarose gel stained with SYBR Safe (ApexBio, Houston, TX, USA). A single amplified DNA fragment of either 189 bp or 125 bp was expected from each reaction (Fig. S4.2b). Full, uncropped images of all electrophoresis gels are provided in Supplementary file 4.2.

4.6.5. RT-qPCR

Two common wheat varieties carrying *TaASN-B2* (Cadenza and Duxford) and two lacking the gene (Claire and Spark) were grown in a randomised block design in a glasshouse. Plants were grown in individual pots for destructive sampling and four replicates were taken at each timepoint. RNA was extracted from embryo tissue at three timepoints (14-, 21-, and 28-days post anthesis) using a standardised RNA extraction protocol (Chang et al., 1993). The RNA was cleaned further using the ReliaPrep™ RNA Clean-Up and Concentration System (Promega) according to the manufacturer's instructions. DNA was then removed from these samples using RQ1 RNase-Free DNase (Promega) according to the manufacturer's instructions, and the RNA was quantified using a NanoDrop™ 1000 Spectrophotometer (Thermo Fisher Scientific).

cDNA was synthesised by mixing 2 µg RNA in nuclease-free water with oligo-dT primer, dNTPs, and SuperScript™ III Reverse Transcriptase kit components (Invitrogen), according to

the manufacturer's instructions, and placing in a thermocycler using the following programme: 95 °C for 5 min; 60 °C for 60 min; 72 °C for 15 min.

RT-qPCR was performed using an Applied Biosystems™ 7500 Real-Time PCR System set to ddCt (relative quantitation) mode. Each reaction contained 10 µL SYBR Green Master Mix (Applied Biosystems), 5 µL primer mastermix (containing 0.04 mM of each primer and ROX reference dye in nuclease-free water), and 5 µL cDNA (diluted to 6 ng/µL). The expression of each target gene was measured relative to three reference genes; *GAPDH*, *PROSM*, and *SDH*. Details of the primers used are found in Table S4.5.

Relative expression values were calculated as described by Rieu and Powers (2009). Applied Biosystems 7500 Real-Time PCR Software version 2.0.5 was used to calculate Ct values and exported Rn data were converted to PCR efficiency data using LinRegPCR (Ramakers et al., 2003). Statistical tests were performed using GenStat (VSN, 2020) to account for the blocking structure of the experiment and graphs were plotted in R (R Core Team, 2020) using the package ggpubr (Kassambara, 2020).

4.6.6. RNA-seq data analysis

Raw RNA-seq reads from a developmental timecourse in the landrace Chinese Spring (Choulet et al., 2014), grain development samples from the variety Azhurnaya (Ramírez-González et al., 2018), an embryo development timecourse from the variety AC Barrie (Xiang et al., 2019) and grain expression at 14 DPA and 30 DPA from a set of worldwide wheat varieties (Nirmal et al., 2016), were downloaded from the NCBI GEO database (https://www.ncbi.nlm.nih.gov/geo/) processed and mapped to the IWGSC RefSeq v1.1 genome following the approach described previously (Pearce et al., 2015). An additional contig corresponding to the *TaASN-B2* coding sequence and including 1 kb of sequence upstream and downstream of the protein coding region was added to the reference genome to assay transcript levels of this gene. Raw counts were converted into TPM using a custom python script. Heatmaps were generated in R (v1.12.5019) using the command 'heatmap' within the gplots package (Warnes et al., 2013). The ternary plot was created using the ggtern package (Hamilton & Ferry, 2018) within ggplot2 (Wickham, 2016). All expression data in TPM are presented in Supplementary file 4.1.

4.6.7. Effect of TaASN-B2 on free asparagine in the grain

Data from field trials performed in 2011–2012 and 2012–2013 (Curtis et al., 2018) were used to investigate the effect of the *TaASN-B2* deletion on grain asparagine concentrations with sulphur either supplied (S+) or withheld (S-). Varieties from this field trial were screened as described above for the presence or absence of *TaASN-B2*. Data were $\log_e$ transformed to account for heterogeneity of variance, as identified in the previous study, before performing the ANOVA and REML analyses. Analyses and plotting were performed in R (R Core Team, 2020) with the package ggpubr (Kassambara, 2020). ANOVA and REML analyses were performed in GenStat (VSN, 2020) to account for the split-plot blocking structure of the field trials.

The re-analysis of the field trial data included an additional factor identifying the presence/absence of *TaASN-B2*, with the effect of variety nested within the *TaASN-B2* factor. The trials in the individual years were analysed according to the designs indicated before (Curtis et al., 2018) using ANOVA, with the addition of the *TaASN-B2* factor. The data combined across the two trials was analysed as a linear mixed model using the REML algorithm, allowing for the different design structures and sets of varieties included in the two trials, providing an overall comparison of both the presence/absence of *TaASN-B2*, and the differences between the varieties included across the two trials (allowing comparisons of varieties included in different years).

4.7. References

Bian, X., Tyrrell, S., & Davey, R. P. (2017). The Grassroots life science data infrastructure. https://grassroots.tools.

Blake, V. C., Woodhouse, M. R., Lazo, G. R., Odell, S. G., Wight, C. P., Tinker, N. A., ... & Sen, T. Z. (2019). GrainGenes: centralized small grain resources and digital platform for geneticists and breeders. *Database, 2019*, baz065.

Chang, S., Puryear, J., & Cairney, J. (1993). A simple and efficient method for isolating RNA from pine trees. *Plant Molecular Biology Reporter, 11*(2), 113-116.

Choulet, F., Alberti, A., Theil, S., Glover, N., Barbe, V., Daron, J., ... & Feuillet, C. (2014). Structural and functional partitioning of bread wheat chromosome 3B. *Science*, *345*(6194), 1249721.

Curtis, T. Y., Powers, S. J., & Halford, N. G. (2016). Effects of fungicide treatment on free amino acid concentration and acrylamide-forming potential in wheat. *Journal of Agricultural and Food Chemistry*, *64*(51), 9689-9696.

Curtis, T. Y., Powers, S. J., Wang, R., & Halford, N. G. (2018). Effects of variety, year of cultivation and sulphur supply on the accumulation of free asparagine in the grain of commercial wheat varieties. *Food Chemistry*, *239*, 304-313.

Curtis, T. Y., Raffan, S., Wan, Y., King, R., Gonzalez-Uriarte, A., & Halford, N. G. (2019). Contrasting gene expression patterns in grain of high and low asparagine wheat genotypes in response to sulphur supply. *BMC Genomics*, *20*(1), 1-18.

Dubcovsky, J., & Dvorak, J. (2007). Genome plasticity a key factor in the success of polyploid wheat under domestication. *Science*, *316*(5833), 1862-1866.

Emebiri, L. C. (2014). Genetic variation and possible SNP markers for breeding wheat with low-grain asparagine, the major precursor for acrylamide formation in heat-processed products. *Journal of the Science of Food and Agriculture*, *94*(7), 1422-1429.

European Commission. Commission regulation EU. (2017). 2017/2158 establishing mitigation measures and benchmark levels for the reduction of the presence of acrylamide in food. Brussels: European Commission; 2017.

European Food Safety Authority Panel on Contaminants in the Food Chain (CONTAM Panel). (2015). Scientific opinion on acrylamide in food. *EFSA Journal*, *13*, 4104.

Food and Drug Administration. (2016). Guidance for Industry, Acrylamide in Foods. College Park MD: Food and Drug Administration.

Gao, R., Curtis, T. Y., Powers, S. J., Xu, H., Huang, J., & Halford, N. G. (2016). Food safety: structure and expression of the asparagine synthetase gene family of wheat. *Journal of Cereal Science*, *68*, 122-131.

Genstat for Windows 21st Edition. (2020). VSN International, Hemel Hempstead, UK. https://Genstat.co.uk.

Hamilton, N. E., & Ferry, M. (2018). ggtern: Ternary diagrams using ggplot2. *Journal of Statistical Software*, *87*, 1-17.

IARC. (1994). Some industrial chemicals: IARC monographs on the evaluation of carcinogenic risks to humans. Vol 60. *Occupational and Environmental Medicine*, *52*(5), 360.

Kassambara A. (2020). ggpubr: 'ggplot2' Based Publication Ready Plots. R package version 0.4.0. https://CRAN.R-project.org/package=ggpubr

Krasileva, K. V., Vasquez-Gross, H. A., Howell, T., Bailey, P., Paraiso, F., Clissold, L., ... & Dubcovsky, J. (2017). Uncovering hidden variation in polyploid wheat. *Proceedings of the National Academy of Sciences*, *114*(6), E913-E921.

Lea, P. J., Sodek, L., Parry, M. A., Shewry, P. R., & Halford, N. G. (2007). Asparagine in plants. *Annals of Applied Biology*, *150*(1), 1-26.

Lei, G. J., Fujii-Kashino, M., Wu, D. Z., Hisano, H., Saisho, D., Deng, F., ... & Ma, J. F. (2020). Breeding for low cadmium barley by introgression of a Sukkula-like transposable element. *Nature Food*, *1*(8), 489-499.

Maccaferri, M., Harris, N. S., Twardziok, S. O., Pasam, R. K., Gundlach, H., Spannagl, M., ... & Cattivelli, L. (2019). Durum wheat genome highlights past domestication signatures and future improvement targets. *Nature Genetics*, *51*(5), 885-895.

Marcussen, T., Sandve, S. R., Heier, L., Spannagl, M., Pfeifer, M., International Wheat Genome Sequencing Consortium,, ... & Praud, S. (2014). Ancient hybridizations among the ancestral genomes of bread wheat. *Science*, *345*(6194), 1250092.

Martinek, P., Klem, K., Vanova, M., Bartackova, V., Vecerkova, L., Bucher, P., & Hajslova, J. (2009). Effects of nitrogen nutrition, fungicide treatment and wheat genotype on free asparagine and reducing sugars content as precursors of acrylamide formation in bread. *Plant, Soil and Environment*, *55*(5), 187-195.

Middleton, A. M., Úbeda-Tomás, S., Griffiths, J., Holman, T., Hedden, P., Thomas, S. G., ... & Owen, M. R. (2012). Mathematical modeling elucidates the role of transcriptional feedback in gibberellin signaling. *Proceedings of the National Academy of Sciences*, *109*(19), 7571-7576.

Murashige, T., & Skoog, F. (1962). A revised medium for rapid growth and bio assays with tobacco tissue cultures. *Physiologia Plantarum*, *15*(3), 473-497.

Nirmal, R. C., Furtado, A., Wrigley, C., & Henry, R. J. (2016). Influence of gene expression on hardness in wheat. *PLOS One*, *11*(10), e0164746.

Noree, C., Monfort, E., & Shotelersuk, V. (2018). Human asparagine synthetase associates with the mitotic spindle. *Biology Open*, *7*(12), bio038307.

Noree, C., Sirinonthanawech, N., & Wilhelm, J. E. (2019). *Saccharomyces cerevisiae ASN1* and *ASN2* are asparagine synthetase paralogs that have diverged in their ability to polymerize in response to nutrient stress. *Scientific Reports*, *9*(1), 1-8.

Oddy, J., Raffan, S., Wilkinson, M. D., Elmore, J. S., & Halford, N. G. (2020). Stress, nutrients and genotype: understanding and managing asparagine accumulation in wheat grain. *CABI Agriculture and Bioscience*, *1*(1), 1-14.

Pearce, S., Vazquez-Gross, H., Herin, S. Y., Hane, D., Wang, Y., Gu, Y. Q., & Dubcovsky, J. (2015). WheatExp: an RNA-seq expression database for polyploid wheat. *BMC Plant Biology*, *15*(1), 1-8.

Potter, S. C., Luciani, A., Eddy, S. R., Park, Y., Lopez, R., & Finn, R. D. (2018). HMMER web server: 2018 update. *Nucleic Acids Research*, *46*(W1), W200-W204.

R Core Team. (2013). R: A language and environment for statistical computing. https://www.R-project.org/

Raffan, S., & Halford, N. G. (2019). Acrylamide in food: Progress in and prospects for genetic and agronomic solutions. *Annals of Applied Biology*, *175*(3), 259-281.

Raffan, S., & Halford, N. G. (2021). Cereal asparagine synthetase genes. *Annals of Applied Biology*, *178*(1), 6-22.

Raffan, S., Oddy, J., & Halford, N. G. (2020). The sulphur response in wheat grain and its implications for acrylamide formation and food safety. *International Journal of Molecular Sciences*, *21*(11), 3876.

Raffan, S., Sparks, C., Huttly, A., Hyde, L., Martignago, D., Mead, A., ... & Halford, N. G. (2021). Wheat with greatly reduced accumulation of free asparagine in the grain, produced by CRISPR/Cas9 editing of asparagine synthetase gene *TaASN2*. *Plant Biotechnology Journal*, *19*(8), 1602-1613.

Ramakers, C., Ruijter, J. M., Deprez, R. H. L., & Moorman, A. F. (2003). Assumption-free analysis of quantitative real-time polymerase chain reaction (PCR) data. *Neuroscience Letters*, *339*(1), 62-66.

Ramírez-González, R. H., Borrill, P., Lang, D., Harrington, S. A., Brinton, J., Venturini, L., ... & Uauy, C. (2018). The transcriptional landscape of polyploid wheat. *Science*, *361*(6403), eaar6089.

Rapp, M., Schwadorf, K., Leiser, W. L., Würschum, T., & Longin, C. F. H. (2018). Assessing the variation and genetic architecture of asparagine content in wheat: What can plant breeding contribute to a reduction in the acrylamide precursor? *Theoretical and Applied Genetics*, *131*(11), 2427-2437.

Rieu, I., & Powers, S. J. (2009). Real-time quantitative RT-PCR: design, calculations, and statistics. *The Plant Cell*, *21*(4), 1031-1033.

Sambrook J, Fritsch EF, Maniatis T. (1989). Molecular cloning: a laboratory manual. *CSHL Press*.

Schmidt, S. M., Belisle, M., & Frommer, W. B. (2020). The evolving landscape around genome editing in agriculture: Many countries have exempted or move to exempt forms of genome editing from GMO regulation of crop plants. *EMBO Reports*, *21*(6), e50680.

Sim, N. L., Kumar, P., Hu, J., Henikoff, S., Schneider, G., & Ng, P. C. (2012). SIFT web server: predicting effects of amino acid substitutions on proteins. *Nucleic Acids Research*, *40*(W1), W452-W457.

Simsek, S., Ohm, J. B., Lu, H., Rugg, M., Berzonsky, W., Alamri, M. S., & Mergoum, M. (2014). Effect of pre-harvest sprouting on physicochemical changes of proteins in wheat. *Journal of the Science of Food and Agriculture*, *94*(2), 205-212.

Solovyev, V., Kosarev, P., Seledsov, I., & Vorobyev, D. (2006). Automatic annotation of eukaryotic genes, pseudogenes and promoters. *Genome Biology*, *7*(1), 1-12.

International Wheat Genome Sequencing Consortium (IWGSC), Appels, R., Eversole, K., Stein, N., Feuillet, C., Keller, B., ... & Singh, N. K. (2018). Shifting the limits in wheat research and breeding using a fully annotated reference genome. *Science*, *361*(6403), eaar7191.

Warnes, G. R., Bolker, B., Bonebakker, L., Gentleman, R., Huber, W., Liaw, A., ... & Venables, B. (2009). gplots: Various R programming tools for plotting data. *R package version*, *2*(4), 1. http://CRAN.R-project.org/package=gplots.

Wicker, T., Gundlach, H., Spannagl, M., Uauy, C., Borrill, P., Ramírez-González, R. H., ... & Choulet, F. (2018). Impact of transposable elements on genome structure and evolution in bread wheat. *Genome Biology*, *19*(1), 1-18.

Wicker, T., & Keller, B. (2007). Genome-wide comparative analysis of copia retrotransposons in Triticeae, rice, and Arabidopsis reveals conserved ancient evolutionary lineages and distinct dynamics of individual copia families. *Genome Research*, *17*(7), 1072-1081.

Wicker, T., Matthews, D. E., & Keller, B. (2002). TREP: a database for Triticeae repetitive elements. *Trends in plant science*, *7*(12), 561-562.

Wickham, H. (2016). Package 'ggplot2': elegant graphics for data analysis. *Springer-Verlag New York*, *10*, 978-0.

Xiang, D., Quilichini, T. D., Liu, Z., Gao, P., Pan, Y., Li, Q., ... & Datla, R. (2019). The transcriptional landscape of polyploid wheats and their diploid ancestors during embryogenesis and grain development. *The Plant Cell*, *31*(12), 2888-2911.

Xu, H., Curtis, T. Y., Powers, S. J., Raffan, S., Gao, R., Huang, J., ... & Halford, N. G. (2018). Genomic, biochemical, and modeling analyses of asparagine synthetases from wheat. *Frontiers in Plant Science*, *8*, 2237.

Zadoks, J. C., Chang, T. T., & Konzak, C. F. (1974). A decimal code for the growth stages of cereals. *Weed Research*, *14*(6), 415-421.

Zhang, Y., Pribil, M., Palmgren, M., & Gao, C. (2020). A CRISPR way for accelerating improvement of food crops. *Nature Food*, *1*(4), 200-205.

Zhivagui, M., Ng, A. W., Ardin, M., Churchwell, M. I., Pandey, M., Renard, C., ... & Zavadil, J. (2019). Experimental and pan-cancer genome analyses reveal widespread contribution of acrylamide exposure to carcinogenesis in humans. *Genome Research*, *29*(4), 521-531.

4.8. Supplementary figures and tables

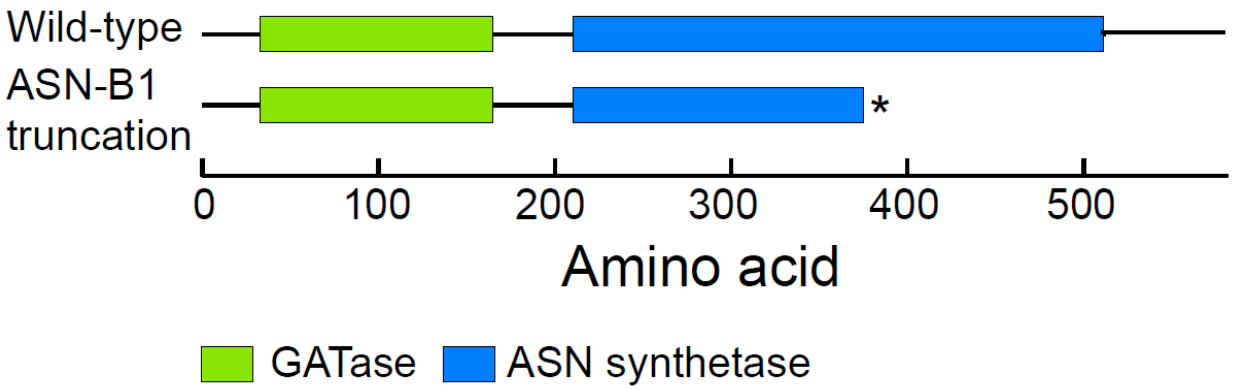


Fig. S4.1. Allelic diversity in *TaASN1*. Predicted amino acid sequence of the full-length ASN-B1 protein encoded by varieties Robigus, Julius, Norin 61, Mace and Spelt wheat (wild-type), compared to the truncated protein encoded by varieties CDC Landmark, Claire, Jagger, Cadenza, Paragon, Arina, CDC Stanley and Lancer (ASN-B1 truncation). In the latter varieties, a 16 bp deletion in exon 7 is predicted to introduce a premature stop codon at amino acid residue 375, indicated by *. The conserved GATase and ASN synthetase domains are highlighted.

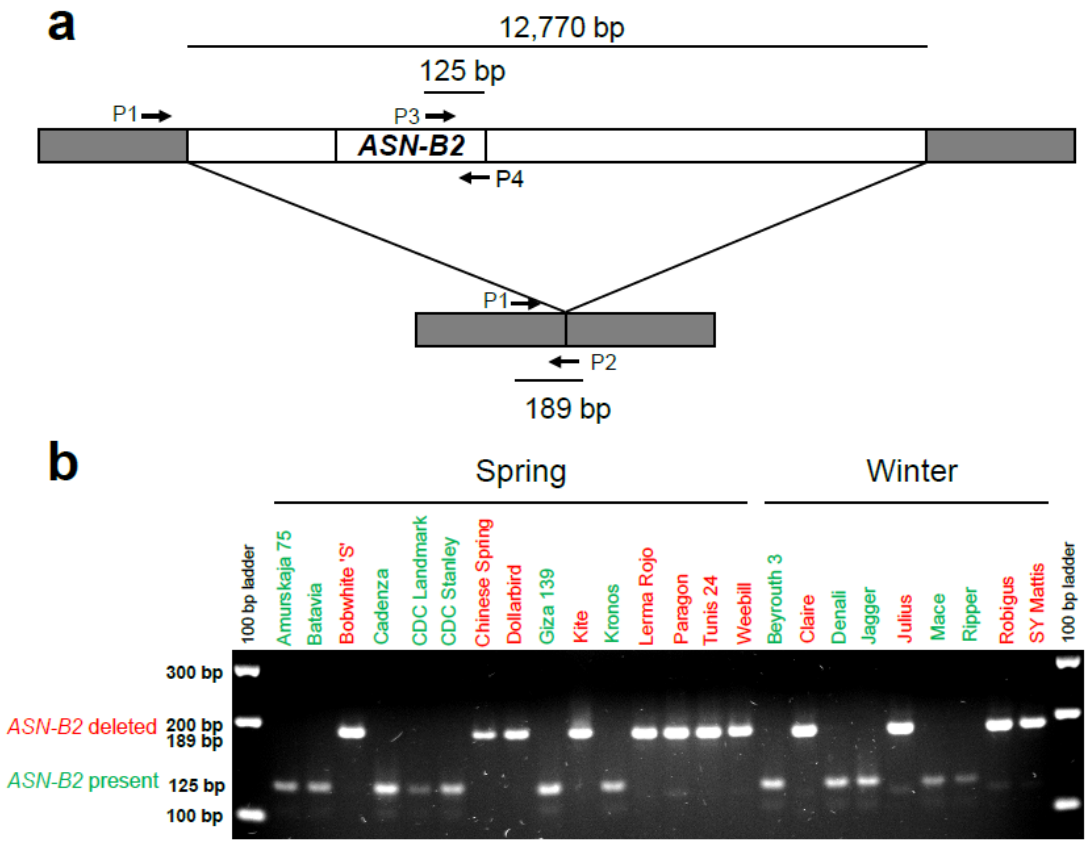


Fig. S4.2. PCR assay to distinguish presence and absence of *TaASN-B2* in a collection of 24 global wheat varieties. **a.** Schematic diagram of the assay to show primer positions and expected amplicon sizes. Amplification of a 189 bp product with primers P1 and P2 indicates that *TaASN-B2* is deleted, while amplification of a 125 bp product with primers P3 and P4 indicates that *TaASN-B2* is present. One amplified fragment is expected in each reaction. **b.** Agarose gel electrophoresis of PCR products from the assay. Varieties with *TaASN-B2* deleted are highlighted in red, while varieties with the gene present are highlighted in green. A 100 bp ladder is shown in the first and last well of the gel for size comparison. Among the varieties are five carrying the *TaASN-B2* deletion and four with *TaASN-B2* present that were used to assay *ASN* expression in the grain. Full details of each variety are given in Table 2.

Table S4.1. Natural variation in ASN proteins in wheat. For each protein, shades of green indicate that all amino acid substitutions are predicted to be tolerated and the encoded protein is predicted to be functional. Shades of yellow/orange indicate that at least one polymorphism is predicted to be disruptive for protein function based on SIFT analysis. Red indicates the gene is deleted in that variety. Full details of each protein type are provided in the key below the main table, where (T) indicates the amino acid substitution at that position is predicted to be tolerated, and (APF) indicates the change is predicted to affect protein function.

Protein	IWGSC RefSeq v1.1 Gene ID	Robigus	CDC Landmark	Julius	Claire	Jagger	Cadenza	Paragon	Arina	Norin 61	CDC Stanley	SY Mattis	Lancer	Mace	Spelt
TaASN-A1	*TraesCS5A02G153900*														
TaASN-B1	*TraesCS5B02G152600*														
TaASN-D1	*TraesCS5D02G159100*														
TaASN-A2	*TraesCS3A02G077100*														
TaASN-B2	Not present														
TaASN-D2	*TraesCS3D02G077300*														
TaASN-A3.1	*TraesCS1A02G382800*														
TaASN-B3.1	*TraesCS1B02G408200*														
TaASN-D3.1	*TraesCS1D02G390500*														
TaASN-A3.2	*TraesCS1A02G422100*														
TaASN-B3.2	*TraesCS1B02G453600*														
TaASN-D3.2	*TraesCS1D02G430300*														
TaASN-A4	*TraesCS4A02G109900*														
TaASN-B4	*TraesCS4B02G194400*														
TaASN-D4	*TraesCS4D02G195100*														

Table S4.1. (cont.).

Protein	RefSeqv1.0 gene ID*	Type 2	Type 3	Type 4
TaASN-A1	*TraesCS5A02G153900*	V19A (T) R329S (APF)	R329S (APF)	
TaASN-B1	*TraesCS5B02G152600*	Y113F (T) I529V (T)	STOP codon on 3rd exon	Not present
TaASN-D1	*TraesCS5D02G159100*	G268R (APF)		
TaASN-A2	*TraesCS3A02G077100*			
TaASN-B2	Not present	*		
TaASN-D2	*TraesCS3D02G077300*	R112S (T) K131N (T) P175L (T) R418W (APF)		
TaASN-A3.1	*TraesCS1A02G382800*	G110R (APF) S474P (T) A559S (T)	R26_R31del G110R (APF) S474P (T) A559S (T)	
TaASN-B3.1	*TraesCS1B02G408200*	K527E (T)	H565P (T)	
TaASN-D3.1	*TraesCS1D02G390500*			
TaASN-A3.2	*TraesCS1A02G422100*	A533T (T) E537K (T)		
TaASN-B3.2	*TraesCS1B02G453600*	Not present		
TaASN-D3.2	*TraesCS1D02G430300*			
TaASN-A4	*TraesCS4A02G109900*	E439_P440Ins		
TaASN-B4	*TraesCS4B02G194400*			
TaASN-D4	*TraesCS4D02G195100*			

**ASN-B2* sequence based on Jagger reference genome.

Table S4.2. a. List of UK winter wheat (*Triticum aestivum*) varieties with *TaASN-B2* present or absent, separated by market class. **b.** List of common wheat varieties with *TaASN-B2* present or deleted among a panel of 24 global wheat varieties. ID, accession numbers and country of origin are provided.

a.

UK winter wheat group	***TaASN-B2* present**	***TaASN-B2* deleted**	
Group 1: Bread making	Avalon	Crusoe	Skyfall
	Cadenza	Gallant	Solstice
	Malacca	Hereward	Spark
		Shamrock	
Group 2: Bread making potential	Cashel	Bonham	Evoke
	Einstein	Charger	Podium
		Chilton	Rialto
		Cordiale	Shango
		Cubanita	Sterling
Group 3: Biscuit	Torch	Claire	Monterey
		Cocoon	Robigus
		Croft	Scout
		Delphi	Tuxedo
		Diego	Warrior
		Icon	Weaver
		Invicta	Zulu
Group 4: Soft	Lancaster	Alchemy	Leeds
		Cougar	Myriad
		Denman	Rowan
		Horatio	Twister
		Panacea	Viscount
		Revelation	
Group 4: Hard	Badger	Buster	Icebreaker
	Duxford	Dickens	Oakley
	Kielder	Evolution	Santiago
	Relay	Gator	Savannah
		Goldengun	Solace

b.

Growth habit	*TaASN-B2* present	Accession number	Country of origin	*TaASN-B2* deleted	Accession number	Country of origin
Spring	Amurskaja 75	PI 372145	Russia	Bobwhite 'S'		Mexico
	Batavia	PI-572700	Australia	Chinese Spring	CItr 14108	China
	Cadenza	id#39740	UK	Dollarbird	PI-525198	Australia
	CDC Landmark	id#39741	Canada	Kite	PI-386162	Australia
	CDC Stanley	id#39742	Canada	Lerma Rojo	CItr 13651	Mexico
	Giza 139	PI-185612	Egypt	Paragon	id#39749	UK
	Kronos	PI-576168	USA	Tunis 24	PI-278561	Tunisia
				Weebill	id#39754	Mexico
Winter	Beyrouth	PI 278533	Lebanon	Claire	id#39743	UK
	Denali	PI 664256	USA	Julius	id#39745	Germany
	Jagger	PI-593688	USA	Robigus	id#39751	UK
	Mace	id#39746	Australia	SY Mattis	id#39753	UK
	Ripper	PI 644222	USA			

Table S4.3. Significance values for RT-qPCR and field analysis. **a.** ANOVA Analysis was performed using Timepoint*Variety*Homeologue as the treatment structure and Block/Subblock/Plot as the blocking structure. **b.** Significance values for factors in the ANOVA and REML analyses of field trial data. All analyses were performed on $\log_e$ transformed data. ANOVA analyses were performed using Block/MainPlot/SplitPlot as the random model and (*TaASN-B2*/Variety)*Sulphur Treatment as the treatment model. REML analysis was performed using Year/Block/MainPlot/SplitPlot as the random model and Year*(*TaASN-B2*/Variety)*Sulphur Treatment as the treatment model.

a

Factor	**P-value**
Timepoint	<.001
Variety	<.001
Timepoint*Variety	<.001
Homeologue	<.001
Timepoint*Homeologue	<.001
Variety*Homeologue	<.001
Timepoint*Variety*Homeologue	<.001

b

	2011-2012 ANOVA	2012-2013 ANOVA	Combined REML
Treatment	0.037	0.010	<.001
TaASN-B2	<.001	0.236	0.027
*TaASN-B2**Variety	<.001	0.007	0.007
*TaASN-B2**Treatment	0.221	<.001	0.006
*TaASN-B2**Variety*Treatment	0.074	0.007	0.063
Year			<.001
Year**TaASN-B2*			0.951
Year*Treatment			<.001
Year**TaASN-B2**Treatment			0.004
Year**TaASN-B2**Variety			0.300
Year**TaASN-B2**Variety*Treatment			0.192

Table S4.4: Details of missing sequence data in *ASN* genes in some genome assemblies. Similarity among sequences was determined using all available sequence but some varieties had regions of 'Ns' within *ASN* genes, as indicated in the table below.

Gene	IWGSC RefSeq v1.1 Gene ID	Information on sequence availability
TaASN-A1	*TraesCS5A02G153900*	Full-length sequence in all varieties.
TaASN-B1	*TraesCS5B02G152600*	Full-length sequence in all varieties.
TaASN-D1	*TraesCS5D02G159100*	Full-length sequence in all varieties.
TaASN-A2	*TraesCS3A02G077100*	Full-length sequence in all varieties.
TaASN-B2	*Not present*	Full-length sequence in all varieties.
TaASN-D2	*TraesCS3D02G077300*	Cadenza - missing sequences in exons 10 and 11.
TaASN-A3.1	*TraesCS1A02G382800*	Full-length sequence in all varieties.
TaASN-B3.1	*TraesCS1B02G408200*	Full-length sequence in all varieties.
TaASN-D3.1	*TraesCS1D02G390500*	Full-length sequence in all varieties.
TaASN-A3.2	*TraesCS1A02G422100*	Robigus and Claire – missing sequence from exons 1-3. Julius, Jagger, Arina, Lancer, Mace and Spelt – missing sequence from exon 1. Cadenza - missing sequence at the end of exon 1. Paragon - missing sequence in intron 3*.
TaASN-B3.2	*TraesCS1B02G453600*	Robigus, Cadenza and Paragon - missing sequence in exon 4. Spelt - missing sequence in exon 1.
TaASN-D3.2	*TraesCS1D02G430300*	Robigus, Claire and Paragon - missing sequence in intron 3*
TaASN-A4	*TraesCS4A02G109900*	Paragon and Norin 61 - missing sequence in exon 1.
TaASN-B4	*TraesCS4B02G194400*	Full-length sequence in all varieties.
TaASN-D4	*TraesCS4D02G195100*	Full-length sequence in all varieties.

* Intron 3 contains some sequence similarity to exon 4, although in each variety, missing sequence means it is not possible to determine whether this is a complete copy.

Table S4.5. Primers used in this study.

Primer name	Sequence (5'-3')
ASN-B2-Deletion-F	CGTATAGACCCCGACTCATTGG
ASN-B2-Deletion-R	GCGAGTTAAGGCATGAGCTAAATATC
ASN-2-Universal-F	CGCTCTACAACGAGGACAAG
ASN-2-Universal-R	CCAATGTAGAGAGGCGTGAC
ASN-B2_CS_F3 (P1*)	AGCAAGCCTTCACCATCATT
ASN-B2_CS_R1 (P2*)	GATGTAGGCATGTCAACGAGA
ASN-B2_qF1 (P3*)	AACAAGCCTGGGGTGATGAG
ASN-B2_qR1 (P4*)	TTGTCTCAAAAAGAAAAAGAACTTG
ASN-A2-F	TCAACGGGGAGGTCTACAAC
ASN-A2-R	GCAATGAAGCTGTTATCTCGTG
ASN-B2-F	GTCAACGGGGAGATCTACAACC
ASN-B2-R	GCGATGAAGCTGTGATCTCTTG
ASN-D2-F	GTGAACGGGGAGATTTACAAC
ASN-D2-R	GCAATGAAGCTCTTATCTCGTG
GAPDH-F	ACTTCCAGGGTGACAACAGG
GAPDH-R	GTGCTGTATCCCCCACTCGTT
PROSM-F	CGAGATCGACCAAGAATGG
PROSM-R	TGAGTGTGGCCTCCCTCC
SDH-F	GCTGCCATCATATCCATTCC
SDH-R	AGCAATGTTACCCCTCATCG

*P1-P4 denotes the primer names used to illustrate the assay in Fig. S2.

5

Genetic control of grain amino acid composition in a UK soft wheat mapping population

Joseph Oddy[1], Monika Chhetry[2], Rajani Awal[2], John Addy[1], Mark Wilkinson[1], Dan Smith[1], Robert King[1], Chris Hall[1], Rebecca Testa[1], Eve Murray[1], Sarah Raffan[1], Tanya Y. Curtis[3], Luzie Wingen[2], Simon Griffiths[2], Simon Berry[4], J. Stephen Elmore[5], Nicholas Cryer[6], Isabel Moreira de Almeida[7], and Nigel G. Halford[1*]

1. Rothamsted Research, Harpenden, Hertfordshire AL5 2JQ, UK

2. John Innes Centre, Norwich Research Park, Colney Lane, Norwich, UK

3. Curtis Analytics Limited, Discovery Park, Sandwich, UK

4. Limagrain UK Ltd, Market Rasen LN7 6DT, UK

5. Department of Food & Nutritional Sciences, University of Reading, RG6 6DZ, UK

6. Mondelēz UK R&D Ltd, Bournville Lane, Bournville, Birmingham, B30 2LU, UK

7. Mondelēz R&D International, 6 Rue René Razel, 91400 Saclay, France

As submitted to The Plant Genome after one round of peer review.

This chapter has been modified from its submitted form to allow for edits and corrections within this Thesis.

5.1. Introduction to paper

Whilst work was underway on the sub-project investigating *TaASN-B2*, I had started looking into the work required for the sub-project investigating the Claire × Robigus mapping population. This mapping population was selected for this sub-project as both parents were soft wheat varieties that exhibited differences in free asparagine content (Robigus being higher than Claire in grain free asparagine content). Consequently, any free asparagine QTL identified from this study would be directly relevant for soft wheat breeding. Another benefit of using this mapping population was that scaffold-level genome assemblies of Claire and Robigus were released near the time that the sub-project began, creating an opportunity to investigate any QTL at candidate gene level. I demonstrated this by investigating a plausible candidate gene underlying the lysine QTL identified in this study, due to the nutritional importance of this amino acid (as discussed in greater detail below).

In order to address some of the questions raised by Chapter 3 regarding the relationship between free asparagine and quality traits, I also investigated some grain morphology traits (e.g. grain length, weight) and quality traits (kernel hardness, Hagberg Falling Number). This enabled a comparison of QTL controlling different traits to see which (if any) overlapped with grain free asparagine QTL and whether there were comparable levels of heritability between any of the amino acids and grain size/quality traits. Building on these comparisons, I also performed genomic prediction for each trait to compare this approach with the conventional QTL mapping employed in this study.

5.2. Introduction

The nutritional quality of wheat has profound impacts on human health. As one of the largest sources of average daily calorie intake in the world (18.2% in 2019) (FAOSTAT, 2021), wheat is an essential source of macro and micronutrients. In 2019, 19.5% of average daily global protein intake was estimated to be provided by wheat-based foods (FAOSTAT, 2021). Similarly, between 2008 and 2017 in the UK, over 25% of average daily fibre intake was provided by wheat-based foods (Gressier & Frost, 2021). Wheat flour is often fortified to increase its nutrient content: in the UK, for example, wheat is fortified with calcium, iron, thiamine, niacin and, most recently, folate (DEFRA, 1998; DHSC, 2021). The quantities of different macro and micronutrients in wheat can have large impacts on population health because of the scale at which wheat products are consumed. For example, it is estimated that

the addition of folate to UK flour will lead to a 20% decrease in neural tube defects in babies (DHSC, 2021). Consequently, it is essential to ensure that the nutritional profile of wheat is as beneficial as it can be for human health.

One way in which the nutritional profile of wheat can be improved is *via* optimisation of its amino acid composition, with the concentrations of lysine and asparagine most important. Free (soluble, non-protein) asparagine can be converted to the processing contaminant, acrylamide, during high-temperature cooking and processing, and this has led to ongoing efforts to reduce free asparagine concentration (Oddy et al., 2022). Lysine, on the other hand, is not produced endogenously by humans or other monogastric animals, making it an essential amino acid in the diet, but it is present in only small quantities in wheat and other cereal grain and populations reliant on cereals for their nutrition may suffer from lysine deficiency (Galili & Amir, 2013). Indeed, fortifying wheat flour by adding lysine has been shown to improve indices of nutritional status in clinical trials in Pakistan, northern China, and Syria (Hussein et al., 2004; Zhao et al., 2004; Ghosh et al., 2008). Flour fortification is unlikely to be a sustainable solution in developing countries and it would be much cheaper and more efficient to increase the intrinsic lysine content of wheat grain. Therefore, the amino acid composition of wheat grain could be optimised both by decreasing grain free asparagine content and increasing lysine content.

In recent years, studies have investigated genetic strategies for the reduction of free asparagine content in wheat grain. Induced and natural variation in the asparagine synthetase 2 genes, for example, has been found to impact significantly on free asparagine content (Oddy et al., 2021; Raffan et al., 2021; Alarcon-Reverte et al., 2022) and quantitative trait loci (QTL) for grain free asparagine content have been identified from previous GWAS studies (Emebiri, 2014; Peng et al., 2018; Rapp et al., 2018). However, the small number of stable QTL available to breeders limits the progress that can be made to reduce grain free asparagine content in breeding programmes and no genetic strategies for soft (biscuit) wheat specifically have been investigated. Similarly, there are limited strategies currently available for increasing lysine content in wheat grain. Lysine biofortification *via* QTL identification and marker-assisted breeding has been studied extensively in both rice (Wang et al., 2008; Zhong et al., 2011; Yoo, 2017; Jang et al., 2020) and maize (Prasanna et al., 2020), but only two studies have previously investigated lysine biofortification in wheat through association studies. Peng et al. (2018) successfully identified QTL controlling free lysine and Jiang et al. (2013) identified QTL for total lysine.

Consequently, the aim of this study was to investigate QTL, genomic prediction accuracy, and candidate genes controlling the free amino acid composition of wheat grain in a soft wheat mapping population developed from the varieties Claire and Robigus. Like many UK varieties, these parents both lack the B genome homeologue of the asparagine synthetase-2 gene, *TaASN-B2* (TraesLDM3B03G01566640 in variety Landmark), the presence/absence of which is a known source of grain free asparagine content variation (Oddy et al., 2021). This mapping population, therefore, represents a useful resource for identifying additional variation. Claire and Robigus are also represented by scaffold-level genome assemblies in the wheat pangenome, facilitating candidate gene analysis. Furthermore, we investigated other traits, such as grain size, hardness, and Hagberg falling number (HFN), to determine whether QTL controlling nutritional traits overlapped with those controlling other traits of interest.

5.3. Materials and Methods

5.3.1. Production of Doubled Haploid lines

171 Doubled Haploid lines of Robigus x Claire were produced using a modified Knox et al. (2000) method. Wheat spikes were emasculated between growth stages GS55 and GS59. Once the stigma was receptive it was fertilised with freshly shed donor maize pollen. After one day, wheat florets were treated with Dicamba (20mgL^{-1}) (Sigma-Aldrich, D5417) and injected into the plant stem (100mgL^{-1}). Developing embryos were excised between 14 and 21 days. Under aseptic conditions, seeds were removed from the spikelets, surface sterilised with 70% (v/v) ethanol (EtOH) for 1 min, rinsed with sterile distilled water, and immersed in 20% (v/v) commercial bleach solution with a few drops of Tween® 20 for 20 mins. They were then rinsed with sterile distilled water three times.

Haploid embryos were excised and grown on 90mm Petri dishes in the dark on Gamborg's B5 media with minimal organics (Gamborg et al., 1968), 2% (w/v) sucrose, pH 5.8, 9gL^{-1} Difco bactoagar at 20^{o}C. When showing signs of germination, embryos were transferred to a light incubator at 20^{o}C. Any non-germinated 1-month old embryos were given cold shock treatment at 4^{o}C for 7 days to promote germination. Germinated plantlets were vernalised for 4 weeks and were grown in the glasshouse until the 4-tiller stage. Plants were then given colchicine (Sigma-Aldrich, C9754) treatment for 5 to 6 hours in the light at room temperature, washed and transplanted to soil, acclimatised and grown in a glasshouse. The mapping population was genotyped by Limagrain using a proprietary SNP array. The genetic map comprising 872 loci was constructed using MSTMap Online (http://mstmap.org/).

The mapping population was grown in field trials at the John Innes Centre Morley Mill Hill field site (52°33′15.1″N 1°01′59.2″E) in 2017 to 2018 (abbreviated as H18) and at the Church Farm field site (52°38′N 1°10′E) in 2018 to 2019 (abbreviated as H19). All 171 Doubled Haploid lines of the mapping population were grown in each trial. Within each trial, one replicate of each line was drilled in 6 m^2 plots in a completely randomised design. The H18 field trial was drilled on the 21st September 2017 and harvested on the 1st August 2018. The H19 field trial was drilled on the 14th September 2018 and harvested on the 12th August 2019. Growth habit, heading date, plant height, and yield traits were scored in the field.

DoubleTop fertiliser (27N $30SO_3$) was applied at a rate of 150 kg/ha on the 20th March 2018 for H18 and the 23rd February 2019 for H19. In both trials, slug control pellets were applied at a rate of 7 kg/ha after drilling to control slug pests (Gusto (metaldehyde) pellets on 27th September 2017 for H18 and Sluxx (Ferric phosphate) pellets on 24th September 2018 for H19). Herbicide mixtures were applied in autumn (21st November 2017 for H18 and 24th September 2018 for H19) and spring (23rd May 2018 for H18 and 19th March 2019 for H19) for both trials to control weeds.

5.3.2. Phenotyping

Grain diameter, kernel hardness index (KHI), and grain weight measurements were recorded for 300 kernels from each line in the population using a Perten Single Kernel Classification System (SKCS) 4100 (Calibre Control International Ltd., Warrington, UK). Grain length (mm), width (mm), and area (mm^2) measurements were recorded in triplicate for each sample using a MARVIN Seed Analyser and software Marvin 4.0 (MARViTECH GmbH, Wittenburg, Germany). Grain samples were milled to wholemeal flour in a coffee grinder and flour moisture content was recorded using a Minispec nuclear magnetic resonance (NMR) analyser (Minispec Mq10, Bruker Inc., Germany). Hagberg falling number measurements were recorded using an FN 1000 as the average of two technical replicates (Perten, Sweden), adjusting for flour moisture content as required according to manufacturer's instructions. Amino acid analysis was performed on wholemeal flour samples by HPLC as described previously (Raffan et al., 2021) by Curtis Analytics (Sandwich, UK). Briefly, free amino acids were extracted from 0.5g of wholemeal flour and underwent precolumn derivatisation (Curtis et al., 2018). Samples were then run on an HPLC system identically to previously described (Raffan et al., 2021). Three technical replicates were taken for each sample for amino acid measurement.

5.3.3. Phenotypic data analysis

Skewness and kurtosis were measured for all variables in each environment and normal plots visually inspected in Genstat (VSN International, 2021) to determine if variables required transformation. The data were appropriately transformed according to their distribution if necessary (see Tables S1 and S2 for details of transformations). Subsequent analyses were performed on transformed variables unless otherwise stated. Plotting was performed in R (R Core Team, 2021) with the packages ggplot2 (Wickham, 2016), tidyverse (Wickham et al., 2019), and cowplot (Wilke, 2020).

Broad-sense heritability for each trait was estimated as described in Covarrubias-Pazaran (2019) using the packages dplyr (Wickham et al., 2022) and lme4 (Bates et al., 2015). Kendall rank correlation coefficients were performed on non-transformed data and adjusted *p* values (Bonferroni correction) were calculated for plotting using R (R Core Team, 2021) and the package corrplot (Wei & Simko, 2021). Principal component analysis was performed on untransformed, scaled variables using the package factoextra (Kassambara and Mundt, 2020). Correlation network analysis was performed and plotted by filtering for significant correlations where $p < 0.001$ using Kendall correlation with Bonferroni correction using the packages corrr (Kuhn, Jackson and Cimentada, 2020), igraph (Csardi and Nepusz, 2006), and ggraph (Pedersen, 2021).

Bayesian modelling was performed on untransformed variables in R using the package rstanarm (Goodrich et al., 2020). Variables were scaled before modelling and individual linear models for each predictor variable were created to guide the selection of informative priors. Simulations of the posterior distribution were subsequently performed to check model fit and intervals were plotted using the package bayesplot (Gabry and Mahr, 2022). R^2 estimates were obtained by taking the median of leave-one-out cross validation adjusted estimates.

5.3.4. Linkage analysis

Single-environment linkage analysis was performed in R using packages qtl (Broman et al., 2003) and qtl2 (Broman et al., 2018). Single-environment linkage analysis was made into an interactive app using the packages shiny (Chang et al., 2021), plyr (Wickham, 2011), and rsconnect (Atkins, McPherson & Allaire, 2021), accessible at https://t9onwp-wheatworker.shinyapps.io/QTL_Browser/ and in supplementary data file 5.1 (available at github.com/JosephOddy/ThesisSupplementaryFiles). As before, SIM was performed first to identify covariates for use in CIM. Identified QTL from CIM were then used to create single

QTL models as well as additive QTL models. Upper and lower 95% confidence intervals for QTL location were calculated using the Bayesian credible interval method in R/qtl and expanded to the closest markers. Pseudomarkers were generated every 2 cM in the map and the minimum marker covariate proximity was set at 20 cM. A logarithm of the odds (LOD) score of 3 was used as the significance threshold.

Multi-environment single trait linkage analysis was performed in Genstat for each trait to detect QTL present in both environments, following selection of the most appropriate variance-covariance model according to the Bayesian information criterion. Simple interval mapping (SIM) was initially performed to identify putative QTL. These QTL were then used as covariates in composite interval mapping (CIM). QTL identified from CIM were then used to construct the final QTL models. Pseudo-markers were generated every 2 cM in the map. The minimum cofactor proximity was set at 30 cM and the minimum separation for selected QTL at 20 cM. Significance thresholds were determined by the Li and Ji method (Li & Ji, 2005) with a genome-wide significance level of 0.05.

5.3.5. Genomic prediction

Genomic prediction was performed for each trait *via* five-fold cross validation with 10,000 permutations using the R package rrBLUP (Endelman, 2011). The "mixed.solve" function within this package was used to estimate marker effects for each trait, with the identity matrix being left unspecified. Pearson correlation coefficients were calculated for the results from the training and testing datasets to estimate genomic prediction accuracy. For within year prediction estimates, training and testing datasets came from the same trial. For between year prediction estimates, training and testing datasets were from different trials. Further detail is available as R markdown in supplementary data file 5.2. Scripts were submitted to the high-performance computing cluster at Rothamsted Research via SLURM for execution.

5.3.6. Candidate gene analysis

The gene content of the lysine QTL was determined for all wheat pangenome varieties at chromosome scale assembly by identifying the location of the markers in these varieties and extracting genes from Ensembl Biomart (Howe et al., 2021). Genes residing within the region in variety Chinese Spring v1.0 were submitted to KnetMiner (https://knetminer.com/Triticum_aestivum/) (Hassani-Pak et al., 2021) for ranking on relevant keywords ("Lysine", "Storage proteins"). Expression of the top hits was then investigated in expVIP (Borrill, Ramirez-Gonzalez, & Uauy, 2016) to further narrow down plausible

candidate genes. Transcript per million (TPM) data for the Azhurnaya developmental time-course experiment were extracted from expVIP for plotting in R using the package pheatmap (Kolde, 2019). Corresponding Claire and Robigus genes were then identified from these Chinese Spring candidate genes in Ensembl and pairwise aligned via BLAST using Geneious Prime 2020.1.2 to identify variation.

5.4. Results

5.4.1. Phenotypic analysis

We measured free amino acid concentrations and other traits of interest in the Robigus × Claire mapping population from field trials grown in 2017–2018 (H18) and 2018–2019 (H19) (Figure 5.1; Figure S5.1). Transgressive segregation was observed for asparagine in H19, indicating that this population was suitable for QTL detection of asparagine (Figure S5.2). Aspartic acid, asparagine, and glutamic acid were the most abundant of the free amino acids measured, with concentrations of free amino acids consistently higher in H19 than in H18 (Figure 5.1a). Principal component analysis revealed harvest year to be a key driver of variation in this dataset (Figure 5.1b) and, notably, the second harvest year (H19) also showed lower yield alongside the increased free amino acid content of the grain (Figure 5.1b). PCA and correlation network analysis revealed that most of the other traits measured here were uncorrelated with the amino acids (Figure 5.1b; Figure 5.1c; Figure S5.3; Figure S5.4), except for grain yield which showed negative correlations with a subset of amino acids (Figure 5.1b; Figure 5.1c; Figure 5.2a).

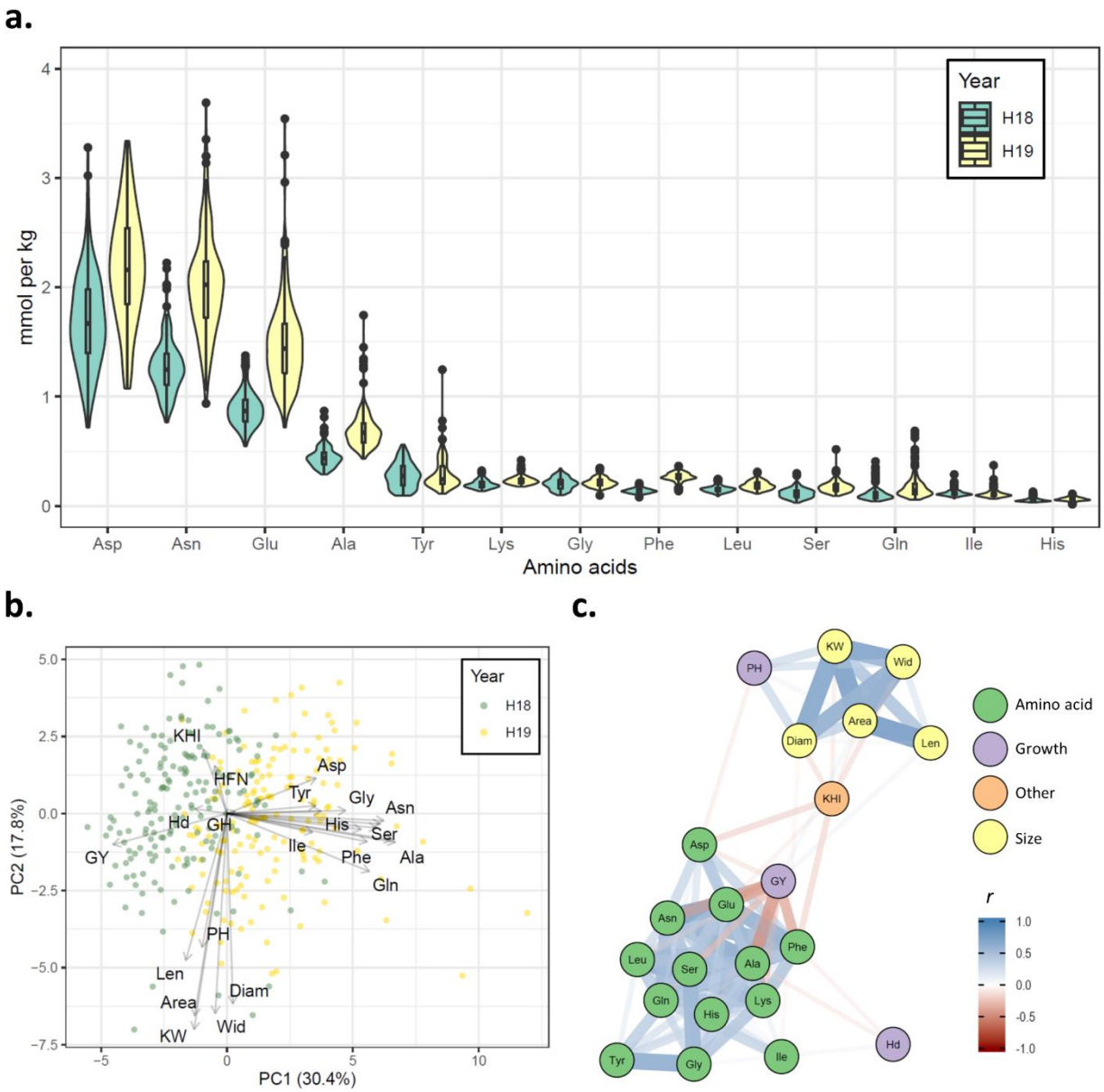


Figure 5.1. Characterisation of the Robigus × Claire mapping population. a. Measurements of amino acids in the 2017–2018 (H18) and 2018–2019 (H19) harvest years. **b.** Principal component analysis of all traits in both years along the first two principal components. **c.** Correlation network analysis of all traits across both years (GH omitted, Kendall correlation, only links with significance <0.001 shown). Diam (grain diameter), GH (growth habit), GY (grain yield), H18 (2017 – 2018 trial), H19 (2018 – 2019 trial), Hd (heading date), HFN (Hagberg falling number), KHI (kernel hardness index), KW (kernel weight), Len (grain length), PH (plant height), Wid (grain width).

To understand whether any of the traits we measured could predict free asparagine or lysine content in the grain, we constructed Bayesian linear models with the traits and harvest year as

explanatory variables (Figure 5.2b; Figure 5.2c). In both the free asparagine (Figure 5.2b) and lysine (Figure 5.2c) models, environment had the greatest effect whereas other variables had little explanatory power. Nevertheless, the variance explained in the models was still reasonable for asparagine at 56.5%, but only 22.2% for lysine.

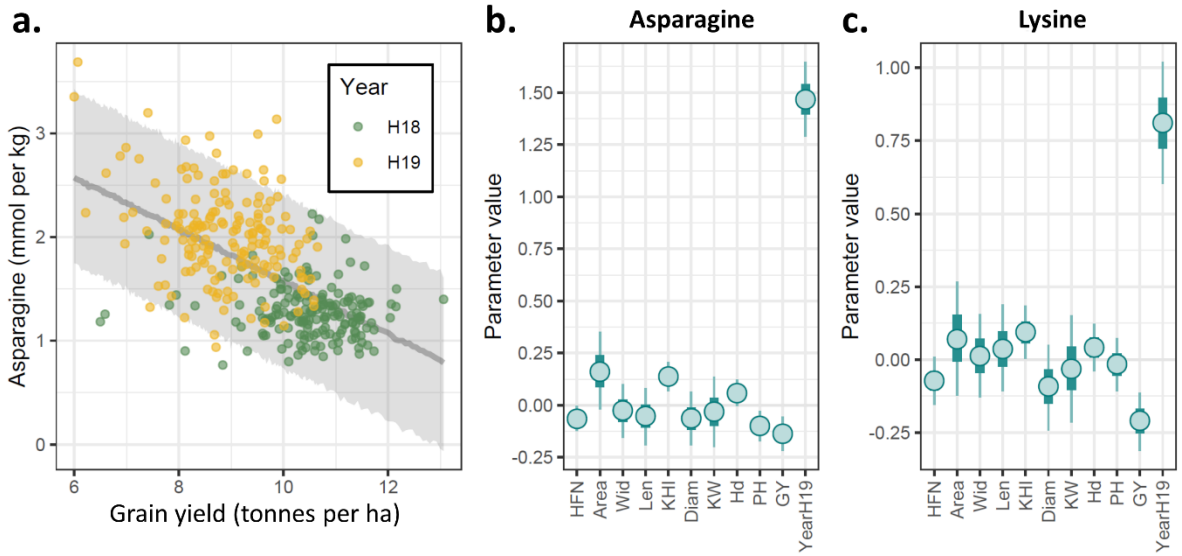


Figure 5.2. Relationships between free asparagine/lysine and other agronomic measurements. a. Linear modelling of free asparagine content against grain yield. The grey shaded ribbon shows 95% prediction intervals sampled from the posterior distribution. **b.** and **c.** Parameter values from multiple linear modelling of asparagine **(b.)** and lysine **(c.)** as explained by other traits measured in this population. The thin lines show 90% posterior uncertainty intervals whereas the thicker lines show 50% posterior uncertainty intervals. The point shows the posterior median. Diam (grain diameter), GY (grain yield), H18 (2017 – 2018 trial), H19 (2018 – 2019 trial), Hd (heading date), HFN (Hagberg falling number), KHI (kernel hardness index), KW (kernel weight), Len (grain length), PH (plant height), Wid (grain width).

5.4.2. QTL analysis

Broad-sense heritability estimates varied substantially between the different amino acids, with free asparagine and lysine showing heritability estimates of 0.60 and 0.45, respectively (Table S5.1). Aspartic acid showed the highest heritability of the amino acids measured here, with an estimate of 0.82. Heritability estimates for the size traits were generally very high, as expected, and correlation of these values between years was also stronger than the correlation of amino acids between years (Table S5.1).

We identified QTL for grain free asparagine content and lysine content on chromosomes 4B and 1A, respectively (Figure 5.3a; Figure 5.3b; Table 5.1), which had significant effects across both environments but were also affected by QTL by environment effects (Figure 5.3c; Figure 5.3d; Table 5.1; Table S5.2). The asparagine QTL on 4B explained 2.6% of the variance in H18, when free asparagine concentrations were lower overall, whereas it explained 14.8% of the variance in H19, when free asparagine concentrations were elevated (Table 5.1). In both years, the Robigus allele was associated with the higher free asparagine concentrations. In contrast, the lysine QTL on 1A explained 12.1% of the variance in H18, when free lysine was lower overall, and only 2.6% of the variance in H19, when free lysine concentrations were elevated. The Claire allele was associated with higher free lysine concentrations in both years in this case. Multi-environment linkage analysis of amino acid and grain measurements revealed many QTL controlling the other amino acids and traits (Table 5.1; Table S5.2; Table S5.3).

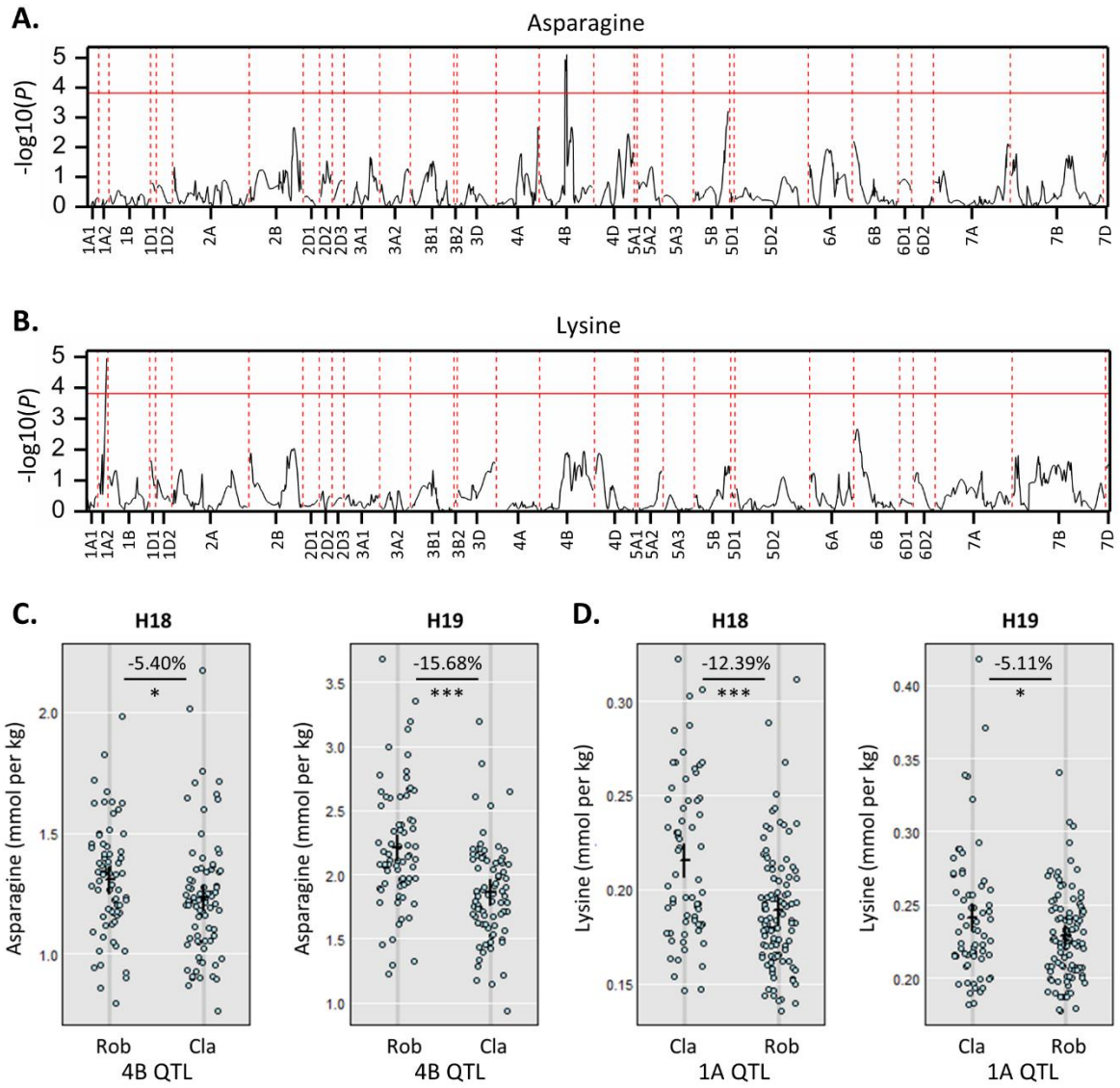


Figure 5.3. Identification of QTL controlling free asparagine and free lysine. a. Multi – environment genome scan plot for asparagine. **b.** Multi – environment genome scan plot for lysine. **c.** Impact of the asparagine QTL on free asparagine concentrations in both field trials. **d.** Impact of the lysine QTL on free lysine concentrations in both field trials. Error bars show plus and minus two times standard error of the mean. Significance values are taken from the corresponding years of the multi-environment linkage analysis. H18 (2017 – 2018 trial), H19 (2018 – 2019 trial).

Table 5.1. Multi-environment QTL for measured amino acids. Chr. (Chromosome), cM (centimorgan), H18 (2017 – 2018 trial), H19 (2018 – 2019 trial), H18 (%)/H19 (%) (percentage variation explained by QTL in each environment), High val. (high value parent (i.e. the parent which posseses the allele conferring the higher value)), Marker (name of the closest marker to the QTL), Mbp (megabase pair location of closest marker in Chinese Spring v1.0).

Trait	Multi-environment single trait linkage analysis (H18 and H19)							
	Marker	Chr.	cM	Mbp	$-\log_{10}(p)$	H18 (%)	H19 (%)	High val.
Ala	WC.0223839	7B	211.2	719	5.03	7.1	5.7	Robigus
Asn	WC.0221262	4B	114.47	601	5.96	2.6	14.8	Robigus
Asp	WC.0218489	1B	54.4	530	5.4	8	5.9	Claire
	WC.0214359	3A2	2.3	738	7.95	7.3	15.3	Robigus
	WC.0221037	4A	148.8	703	8.08	12.6	9.3	Claire
	WC.0227146	4D	48.8	16	3.7	5.5	4.1	Claire
Gln	WC.0221302	4B	103.7	547	3.5	5.4	4.5	Robigus
	WC.0228471	6B	19.7	25	5.09	8.2	6.7	Claire
Glu	WC.0221329	4B	100.8	518	4.27	3.7	10.1	Robigus
Gly	WC.0226796	4B	155.2	327	4.26	3.2	5.3	Robigus
Iso	WC.0223785	7B	211.2	717	3.6	6.8	3.7	Robigus
Lys	WC.0218011	1A2	27.3	593	4.95	12.1	2.6	Claire
Phe	WC.0220622	3B1	78.1	116	3.83	6.2	5.6	Robigus

The QTL controlling asparagine on chromosome 4B appeared to overlap with QTL for glutamine, glutamic acid, and glycine (Table 5.1). Each of these QTL had a greater effect in H18 than in H19 and the Robigus allele was associated with the higher value in each case, suggesting that these QTL are caused by the same variant. These QTL were also located near to a QTL for KHI and more distantly to QTL for grain diameter, plant height, and grain weight (Table 5.2, Figure S5.5), which are likely caused by the *Rht-B1* polymorphism. QTL for aspartic acid also appeared to overlap with QTL for other traits (Table 5.2). For aspartic acid on 4A and 4D, there are co-locating HFN QTL, suggesting that these two traits are under the control of the same locus. The location of the QTL on 4D matches the *Rht-D1* polymorphism between Claire and Robigus found at 18.78 Mbp in Chinese Spring. Of all the amino acids measured in this study, we identified the most QTL controlling aspartic acid (Table 5.1). Other potential sources of variation underpinning the QTL in this study are presented in Table S5.4.

Table 5.2. Multi-environment QTL impacting both amino acids and other traits on chromosomes 4A, 4B, and 4D. Chr. (Chromosome), cM (centimorgan), H18 (2017 – 2018 trial), H19 (2018 – 2019 trial), H18 (%)/H19 (%) (percentage variation explained by QTL in each environment), High val. (high value parent (i.e. the parent which posseses the allele conferring the higher value)), Marker (name of the closest marker to the QTL), Mbp (megabase pair location of closest marker in Chinese Spring v1.0).

Chr.	Multi-environment single trait linkage analysis (H18 and H19)							
	Trait	**Marker**	**cM**	**Mbp**	**-log$_{10}$(*p*)**	**H18 (%)**	**H19 (%)**	**High val.**
4A	Asp	WC.0221037	148.8	703	8.08	12.6	9.3	Claire
	KHI	WC.0221037	148.8	703	8.26	14.8	14.7	Robigus
	HFN	WC.0188904	147.1	733	8.24	11.5	10.3	Robigus
	Area	WC.0220938	149.7	709	2.22	3	3.2	Claire
	Length	WC.0221119	149.7	702	7.12	1.8	6.5	Claire
4B	Asn	WC.0221262	114.47	601	5.96	2.6	14.8	Robigus
	KHI	WC.0226741	110.8	594	4.30	4.2	8.6	Robigus
	Diam	WC.0226868	82.5	32	18.53	17.7	18.7	Claire
	Height	WC.0226868	82.5	32	17.98	19.8	20.7	Claire
	Weight	WC.0226868	82.5	32	9.25	12.2	14.1	Claire
4D	Asp	WC.0227146	48.8	16	3.7	5.5	4.1	Claire
	Width	WC.0227146	48.8	16	5.98	8	8.9	Robigus
	Diam	WC.0227146	48.8	16	7.84	7.6	8.1	Robigus
	HFN	WC.0227149	56.9	17	10.92	23.5	5.6	Robigus
	Height	WC.0213051	56.9	17	28.97	27.8	38.9	Robigus

5.4.3. Genomic prediction

Following our modelling of asparagine and lysine using agronomic measurements and QTL models, we calculated the accuracy of genomic prediction (GP) for within and between year prediction of traits (Figure 5.4a; Figure S5.7; Table S5.1). Prediction accuracy was more consistent when performed across years rather than within years (Figure S5.7), so these were used for further interpretation. Prediction accuracy for lysine was the lowest of all traits at a mean accuracy of 0.10, whereas accuracy for asparagine was around 0.34. Of all amino acids, aspartic acid had the greatest prediction accuracy results. Prediction accuracies for the other functional traits were generally higher than the accuracies for amino acids, as expected from

the higher heritability of these traits. Comparing the amount of variation explained by genomic prediction methods and additive QTL models, we can see that the GP models explain more variance than the additive QTL models for all traits (Figure 5.4b).

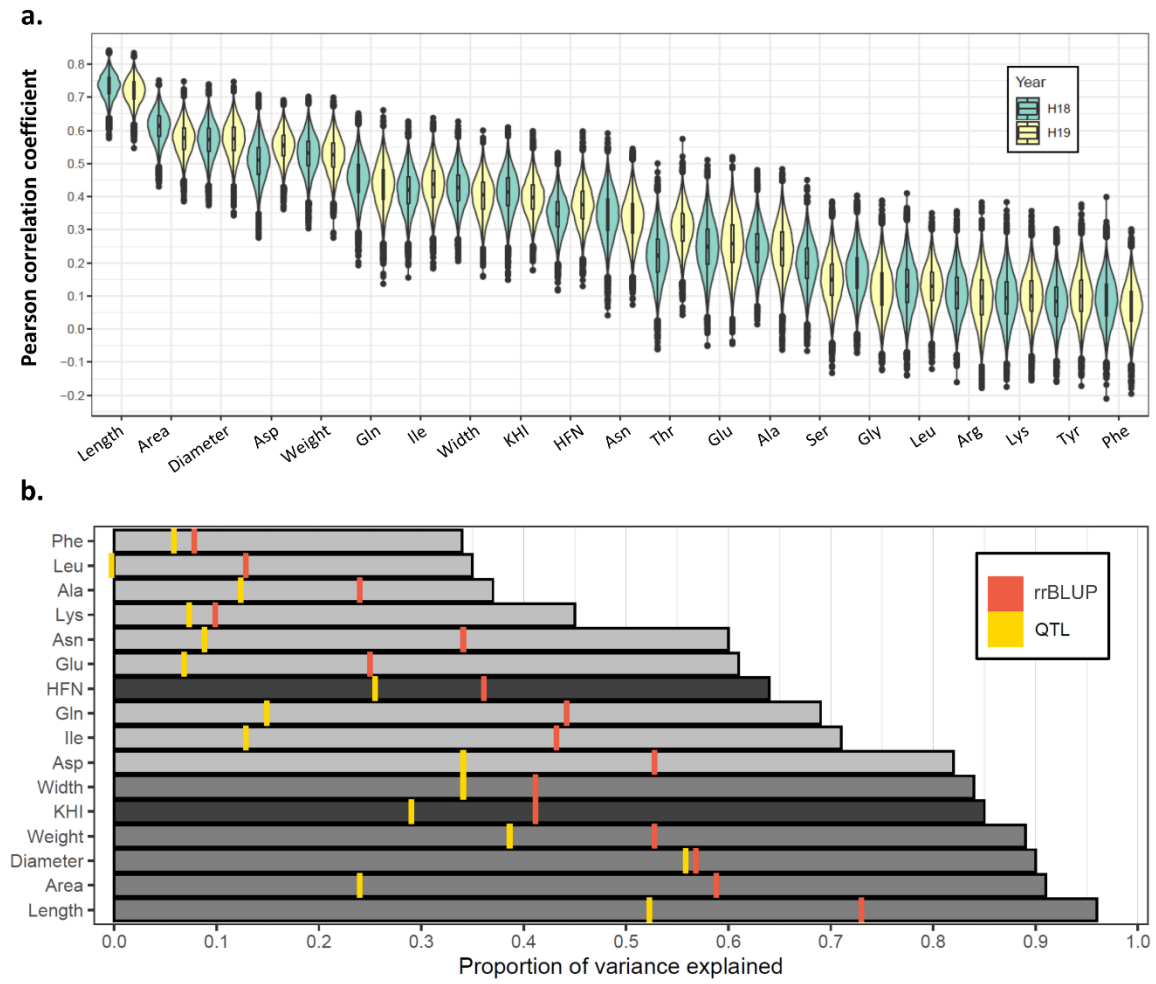


Figure 5.4. Variation explained by heritability, genomic prediction, and QTL. a. Genomic prediction accuracy between years. **b.** Additive QTL effects and genomic prediction (rrBLUP) accuracy (yellow and red marks, respectively) plotted alongside broad-sense heritability (shown as bars). Bars are shaded according to the trait group that they belong to (amino acid, size, or other). H18 (2017 – 2018 trial), H19 (2018 – 2019 trial).

5.4.4. Lysine QTL candidate gene analysis

The gene content and QTL size of the lysine QTL on 1A, the HFN/aspartic acid/KHI QTL on 4A, and the asparagine QTL on 4B differed substantially (Table S5.5). Due to the size of the 4A and 4B QTL, we were unable to plausibly narrow down candidate genes, whereas the lysine QTL on 1A was much smaller so amenable to further analysis. We investigated the gene

content of the lysine QTL for all genomes assembled to chromosome scale in the wheat pangenome and gene content varied to a small extent between the different varieties (Table S5.6). Most notably, the QTL did not match any locations in variety Julius and matched to an unanchored scaffold in Stanley.

KnetMiner analysis of the genes residing in Chinese Spring in the lysine QTL was undertaken with relevant keywords to highlight possible candidate genes (Table S5.7), and these genes were subsequently investigated for their expression patterns from expVIP (Figure S5.8). Pairwise analysis of the top KnetMiner hits in the lysine QTL showed that the top hit from expression analysis (TRAESCS1A02G445700) differed between Claire and Robigus. TRAESCS1A02G445700, or *TaHDT-A1*, has been identified as a member of the histone deacetylase family in wheat. A deletion within the CDS of the gene in Robigus means that the most highly expressed transcript cannot be expressed (Figure 5.5) and the two missing exons from this most highly expressed transcript form a zinc finger/C2H2 DNA binding domain, which is important for transcriptional regulation.

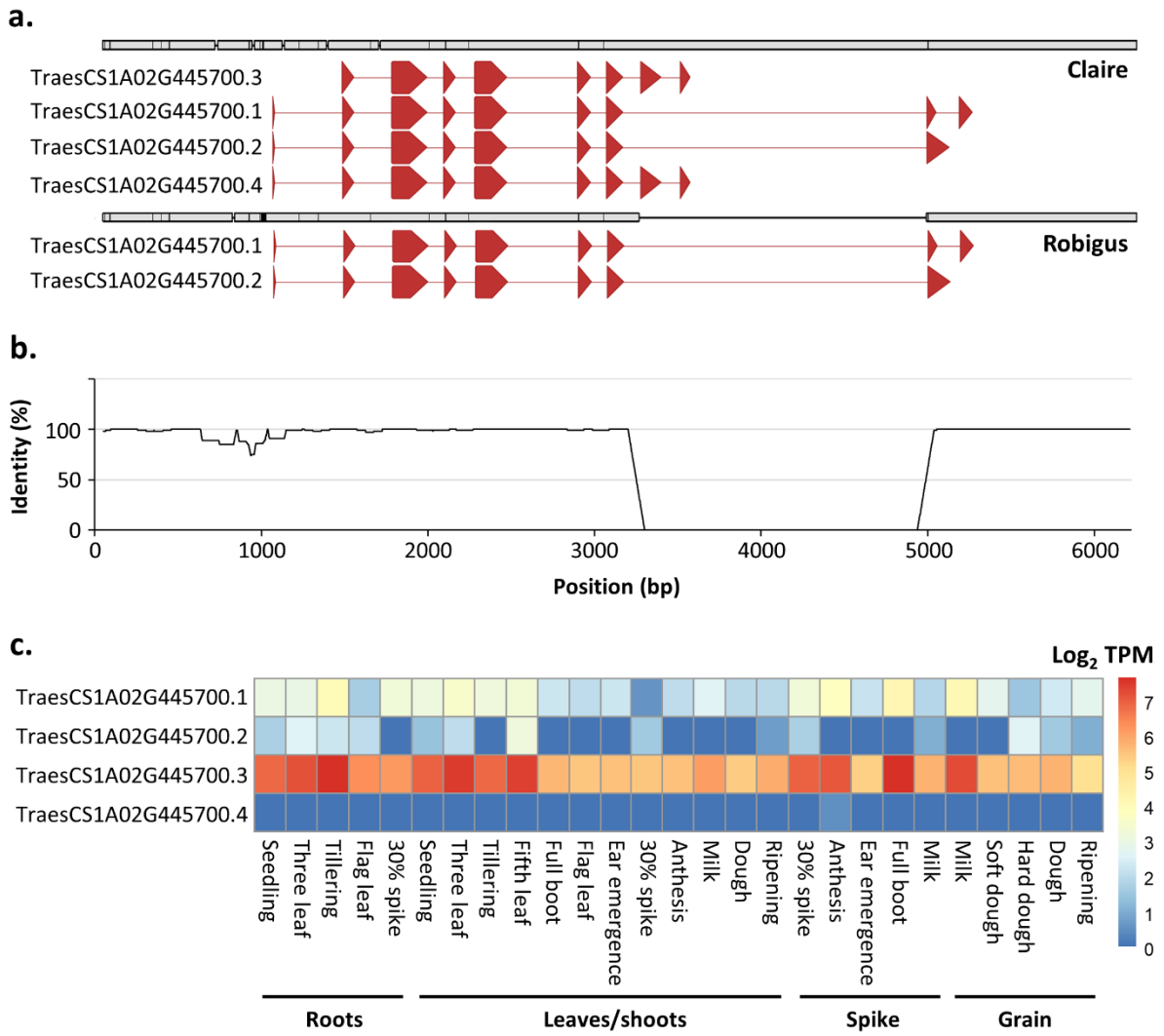


Figure 5.5. Analysis of the *TaHDT-A1* candidate gene for lysine QTL between parents Claire and Robigus. a. Pairwise alignment of the two genes. **b.** Percentage identity calculated as a sliding window average of 100 bp. **c.** Expression of the four transcripts throughout development in variety Azhurnaya. Bp (basepair), TPM (transcripts per million).

5.5. Discussion

5.5.1. Limited variation in Claire and Robigus for asparagine and lysine improvement

Soft wheat breeding in the UK has relied heavily upon Claire and Robigus as parents since their development in 1999 and 2005, respectively. A recent study found that UK winter wheats developed between 2002 and 2017 could be clustered into four distinct populations, and two of these populations were characterised by their Claire or Robigus heritage (Shorinola et al, 2022). The varieties within these population groups characterised by Claire and Robigus heritage are also almost entirely soft wheat varieties, further emphasising the importance of these two varieties in UK soft wheat breeding. This large contribution of Claire and Robigus

as parents to soft wheat breeding means that opportunities for nutritional improvement have often been limited to variation between these two parents.

Our analysis found that there is variation between Claire and Robigus and that this does impact asparagine and lysine content to a small extent. Asparagine had a moderate heritability (0.60) across both field trials in the study, whereas the heritability for lysine was lower (0.45). One QTL was found for asparagine and lysine each, both explaining less than 10% of the variance on average. The asparagine QTL identified here (peak at 601.4 Mbp in Chinese Spring) lies around 60 Mbp from another QTL (peak at 660.7 Mbp in Chinese Spring) identified by Peng et al. (2018), suggesting that these may coincide, whereas the lysine QTL does not overlap with previously identified QTL. Genomic selection had a predictive ability of 0.34 on average for asparagine, indicating that this method may be better suited for breeding because of the genetic architecture of this trait (many small-effect QTL). Rapp et al. (2018) also found that GS had a predictive ability of around 0.5 on average for asparagine, the higher estimate in this study likely due to within environment prediction and analysis of a more diverse mapping population. GS only achieved a predictive ability of 0.10 for lysine, indicating that only incremental advances in lysine content are possible using Claire and Robigus. Previous GWAS studies using more diverse panels have found more, larger-effect QTL controlling asparagine and lysine content (Peng et al., 2018; Rapp et al., 2018), indicating that there may be beneficial alleles in more diverse germplasm. Consequently, UK soft wheat germplasm will require diversity beyond Claire and Robigus to make changes to asparagine and lysine content beyond the incremental improvements found here.

5.5.2. Trade-offs between amino acid content and other traits

Another aspect we wanted to investigate in this population was whether there were any relationships between amino acids and other traits. Amino acids tended to correlate positively with one another and were mostly unrelated to the other measured traits, with the exception of grain yield and kernel hardness index. A negative correlation between grain yield and free asparagine has previously been documented (Xie et al., 2021), but in other experiments the association has been positive (Malunga et al., 2021; Xie et al., 2021). In our analysis, this association mostly arose because of the effect of environment on both yield and asparagine. Environmental stress can lead to decreases in yield whilst increasing free asparagine, whilst other variables (e.g., nitrogen fertiliser) can lead to increases in both yield and free asparagine (see Oddy et al. (2022) for review). Our modelling of asparagine through these variables mostly indicated environment as the driving force in our study, but there was still a slight negative

association with yield and plant height as well as a slight positive association with kernel hardness. Kernel hardness, like grain free asparagine content, is known to increase with nitrogen application, which may underly this small association with asparagine.

A strong environmental effect on free asparagine concentration has been observed in response to many different stressors (see Oddy et al., 2020 for review) and it is under stressful conditions that the highest asparagine concentrations are often observed. These increases in grain free asparagine concentration vary massively, causing unexpected blips in acrylamide content in food products. These environmentally induced increases pose the greatest threat to food safety and regulatory compliance, so elimination of this environmental response would be of great interest. A weak environmental effect was seen in this study: during the 2018–2019 season the average amino acid concentrations rose whilst the yields dropped. The lack of environmental data for this trial prevents us from drawing conclusions on why this environmental effect was seen. Interestingly, the asparagine QTL we identified here had greater effect in this season, enabling reductions of 15.68% in free asparagine concentrations in those lines possessing the Claire allele over those possessing the Robigus allele. This suggests that this QTL may be more effective under more stressful conditions, so selection of the Claire allele at this locus may prove beneficial for reducing the large free asparagine increases observed following environmental stress. This is in contrast to the effect of the *TaASN-B2* deletion, which has a greater effect when grain free asparagine concentrations are lower (Oddy et al., 2021), when plants are not suffering from sulphur deficiency. Future work would therefore benefit from identification of similar QTL that are associated with lowering free asparagine content from the high concentrations seen during stress. This would enable the stacking of alleles that are beneficial under both stress and non-stress conditions, to ensure that free asparagine concentrations are minimised in all environments.

We also wanted to understand whether any QTL controlling amino acid content had pleiotropic effects on other traits. The asparagine QTL we identified on chromosome 4B appeared to overlap with QTL for plant height in the first year, suggesting that there might be an impact of the *Rht-B1b* allele on asparagine. The *Rht* genes are dwarfing genes used during the green revolution that have many impacts on crop traits beyond height (Casebow et al., 2016) and Claire and Robigus both possess different *Rht* genes on 4B and 4D (Wilkinson et al., 2020). However, this QTL overlap was not present in the second year of analysis when the asparagine QTL had a greater effect, suggesting that the QTL controlling height and asparagine may be distinct. However, a more detailed analysis is required to comprehensively assess the impact of *Rht-B1* alleles (and dwarfing genes in general) on grain free asparagine content. The

QTL controlling asparagine did overlap consistently with a QTL for KHI, with the 'increasing allele' belonging to Robigus for both traits. Kernel hardness and free asparagine content are both known to correlate under certain conditions with nitrogen content (Oddy et al., 2022), so this QTL may be linked to nitrogen use efficiency/uptake. The KHI QTL on 4B also exhibited a similar genotype by environment effect pattern to the asparagine QTL, with a greater effect of the QTL observed in the second trial year. Selection for the Claire allele at this QTL would therefore be suitable in the context of soft wheat breeding, where both softer textures and lower free asparagine content are desirable.

Interestingly, we found much more genetic control of free aspartic acid concentration in this population compared to the other amino acids. Heritability was high (> 0.8), genomic prediction accuracy was moderate (> 0.5, same as grain weight), and there were four multi-environment QTL controlling the trait. Two of the QTL controlling aspartic acid also overlapped with QTL controlling HFN. One of these QTL was situated on 4D and overlapped with traits for plant height and grain size, indicating that this may be due to *Rht-D1* allele status, which is known to impact HFN as well as plant height (Fradgley et al., 2022). The second QTL controlling both aspartic acid and HFN was situated on 4A and also overlapped with traits for grain size and KHI. Previous work has identified a major QTL underlying pre-harvesting sprout (PHS) variation on 4A, but both Claire and Robigus share the same *MKK3-A* allele which underlies this QTL (Shorinola et al., 2017). Li et al. (2021) also identified a PHS QTL in a similar region on 4A but this does not overlap with the region identified here. One possible source of variation underlying the QTL controlling aspartic acid and HFN on 4A is the *Triticum dicoccoides* introgression in Robigus, which matches the region this QTL is found in (Przewieslik-Allen et al., 2021). The antagonistic relationship between HFN and asparagine at this QTL could be a result of increased HFN reducing proteolysis, and thereby preventing accumulation of free amino acids.

5.5.3. Lysine candidate genes

Scaffold-level genome assemblies of Claire and Robigus (Walkowiak et al., 2020) enabled us to investigate the lysine QTL in greater depth, identifying the candidate gene *TaHDT-A1*, encoding a histone deacetylase. The wheat histone deacetylase family is very large, encompassing approximately 50 genes (Jin et al., 2020, Li et al., 2022). Histone deacetylases function mainly to inhibit gene expression because histone deacetylation causes chromatin condensation, with roles in many different developmental processes and environmental responses. The link between this histone deacetylase and lysine content will require greater

investigation as it is not clear how the two are related. In wheat, it is known that differences in grain lysine content can be caused by differential expression of lysine-poor storage proteins (prolamins). Gill-Humanes et al. (2014), for example, identified downregulation of gliadins (a class of prolamins) as a method of increasing lysine content in wheat, and Moehs et al. (2019) showed that mutation of wheat prolamin binding factor (*WPBF*), a DOF-class transcription factor, increased lysine concentration. Lower prolamin protein content is also associated with increased lysine content in barley (Rustgi et al., 2019). However, the prolamins confer the viscoelastic properties of wheat dough that are required for the manufacture of many products, including bread, so this must also be considered when trying to breed for higher lysine content.

In maize, grain lysine content is similarly affected by the abundance of lysine-poor proteins in the prolamin family called zeins. The expression of particular zein genes is determined by a bZIP transcription factor called *Opaque2* (Gavazzi et al., 2007), and the mutant line lacking a functional *Opaque2* gene is characterised by higher kernel lysine content (Mertz et al., 1964). Interestingly, the lysine QTL identified in this study is situated upstream of an *Opaque2* orthologue on chromosome 1A: TraesCS1A02G329900, otherwise known as *SPA* (storage protein activator), which is known to activate storage protein synthesis in wheat (Albani et al., 1997). The A genome homeologue of *SPA* does not differ in sequence between Claire and Robigus, but differential expression of *SPA* (through differences in HDT1 regulation) is a possible mechanism by which this QTL could affect lysine content.

Future work investigating *HDT1*, *SPA*, and other regulatory genes of storage proteins in wheat would help to elucidate their effects on grain lysine content and would be useful for expanding the germplasm available to increase lysine content, given the limited QTL and small effect of GS we found. Chromosome-level assemblies of Claire and Robigus would also enable further analysis of this mapping population in the future. Combining both increased diversity and pangenomes, sequencing of the Watkins collection and construction of genome assemblies will enable novel diversity to be identified that can be introgressed into elite soft wheat germplasm (Shewry et al., 2022).

5.6. Conclusions

The nutritional quality of UK soft wheat can be improved incrementally using diversity from Claire and Robigus, but greater diversity is required to make larger gains. The genetic architecture of different amino acids differs considerably, and they are often controlled by QTL that impact other traits. Future soft wheat breeding in the UK should therefore consider use of more genetic diversity and using pleiotropic QTL to the benefit of farmers and consumers.

5.7. References

Alarcón-Reverte, R., Xie, Y., Stromberger, J., Cotter, J. D., Mason, R. E., & Pearce, S. (2022). Induced mutations in ASPARAGINE SYNTHETASE-A2 reduce free asparagine concentration in the wheat grain. *Crop Science, 62*(4), 1484-1496.

Albani, D., Hammond-Kosack, M.C.U., Smith, C., Conlan, S., Colot, V., Holdsworth, M., & Bevan, M.W. (1997). The wheat transcriptional activator SPA: A seed-specific bZIP protein that recognizes the GCN4-like motif in the bifactorial endosperm box of prolamin genes. *The Plant Cell, 9*(2), 171-184.

Atkins, A., McPherson, J., & Allaire, J.J. (2021). rsconnect: Deployment Interface for R Markdown Documents and Shiny Applications. R package version 0.8.25. https://CRAN.R-project.org/package=rsconnect

Bates, D., Maechler, M., Bolker, B., and Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software, 67*(1), 1-48.

Borrill, P., Ramirez-Gonzalez, R., & Uauy, C. (2016). expVIP: a customizable RNA-seq data analysis and visualization platform. *Plant Physiology, 170,* 2172–2186.

Broman, K.W., Gatti, D.M., Simecek, P., Furlotte, N.A., Prins, P., Sen, Ś., Yandell, B.S., & Churchill, G.A. (2019). R/qtl2: software for mapping quantitative trait loci with high-dimensional data and multiparent populations. *Genetics, 211,* 495–502.

Broman, K.W., Wu, H., Sen, Ś., & Churchill, G.A. (2003). R/qtl: QTL mapping in experimental crosses. *Bioinformatics, 19,* 889–890.

Casebow, R., Hadley, C., Uppal, R., Addisu, M., Loddo, S., Kowalski, A., Griffiths, S., & Gooding, M. (2016). Reduced height (Rht) alleles affect wheat grain quality. *PloS one, 11,* e0156056.

Chang, W., Cheng, J., Allaire, JJ., Sievert, C., Schloerke, B., Xie, Y., Allen, J., McPherson, J., Dipert, A.,& Borges, B. (2021). shiny:Web Application Framework for R. R package version 1.7.1. https://CRAN.R-project.org/package=shiny

Covarrubias-Pazaran, G. (2019). Heritability: Meaning and Computation. https://excellenceinbreeding.org/sites/default/files/manual/EiB-M2_Heritability_18-02-20.pdf (accessed March 1, 2022).

Csardi, G., Nepusz, T. (2006) The igraph software package for complex network research, InterJournal, Complex Systems 1695. https://igraph.org

Curtis, T.Y., Powers, S.J., Wang, R., & Halford, N.G. (2018). Effects of variety, year of cultivation and sulphur supply on the accumulation of free asparagine in the grain of commercial wheat varieties. *Food chemistry*, *239*, 304-313.

DEFRA (1998). The Bread and Flour Regulations 1998. https://www.legislation.gov.uk/uksi/1998/141/contents/made (accessed September 16, 2022)

Department of Health and Social Care UK Government. (2021). Folic Acid Added to Flour to Prevent Spinal Conditions in Babies. https://www.gov.uk/government/news/folic-acid-added-to-flour-to-prevent-spinal-conditions-in-babies (accessed February 24, 2022).

Emebiri, L.C. (2014). Genetic variation and possible SNP markers for breeding wheat with low-grain asparagine, the major precursor for acrylamide formation in heat-processed products. *Journal of the Science of Food and Agriculture, 94(7), 1422-1429.*

Endelman, J.B. (2011). Ridge regression and other kernels for genomic selection with R package rrBLUP. *The Plant Genome, 4*, 250–255.

FAO. (2021). FAOSTAT: Food Balances (2014-). http://www.fao.org/faostat/en/ (accessed February 1, 2022).

Fradgley, N.S., Gardner, K., Kerton, M., Swarbreck, S.M., & Bentley, A.R. (2022). Trade-offs in the genetic control of functional and nutritional quality traits in UK winter wheat. *Heredity*, *128,* 420–433.

Gabry J, Mahr T (2022). "bayesplot: Plotting for Bayesian Models." R package version 1.9.0. https://mc-stan.org/bayesplot/

Galili, G., & Amir, R. (2013). Fortifying plants with the essential amino acids lysine and methionine to improve nutritional quality. *Plant Biotechnology Journal, 11*, 211–222.

Gamborg, O.L., Miller, R., Ojima, K. (1968). Nutrient requirements of suspension cultures of soybean root cells. *Experimental Cell Research, 50,* 151–158.

Gavazzi, F., Lazzari, B., Ciceri, P., Gianazza, E., & Viotti, A. (2007). Wild-type opaque2 and defective opaque2 polypeptides form complexes in maize endosperm cells and bind the opaque2-zein target site. *Plant Physiology, 145*, 933–945.

Gil-Humanes, J., Pistón, F., Altamirano-Fortoul, R., Real, A., Comino, I., Sousa, C., Rosell, C.M., & Barro, F. (2014). Reduced-gliadin wheat bread: an alternative to the gluten-free diet for consumers suffering gluten-related pathologies. *PloS one, 9,* e90898.

Goodrich, B., Gabry, J., Ali, I., & Brilleman, S. (2020). rstanarm: Bayesian applied regression modeling via Stan. R package version 2.21.1 https://mc-stan.org/rstanarm.

Gressier, M., & Frost, G. (2022). Minor changes in fibre intake in the UK population between 2008/2009 and 2016/2017. *European Journal of Clinical Nutrition, 76,* 322–327.

Hassani-Pak, K., Singh, A., Brandizi, M., Hearnshaw, J., Parsons, J.D., Amberkar, S., Phillips, A.L., Doonan, J.H., & Rawlings, C. (2021). KnetMiner: a comprehensive approach for supporting evidence-based gene discovery and complex trait analysis across species. *Plant Biotechnology Journal, 19,* 1670–1678.

Howe, K.L., Achuthan, P., Allen, J., Allen, J., Alvarez-Jarreta, J., Amode, M.R., Armean, I.M., Azov, A.G., Bennett, R., & Bhai, J. (2021). Ensembl 2021. *Nucleic Acids Research, 49,* D884–D891

Jang, S., Han, J.-H., Lee, Y.K., Shin, N.-H., Kang, Y.J., Kim, C.-K., & Chin, J.H. (2020). Mapping and validation of QTLs for the amino acid and total protein content in brown rice. *Frontiers in Genetics, 11,* 240.

Jiang, X., Deng, Z., Ru, Z., Wu, P., & Tian, J. (2013). Quantitative trait loci controlling amino acid contents in wheat (*Triticum aestivum* L.). *Australian Journal of Crop Science, 7,* 820–829.

Jin, P., Gao, S., He, L., Xu, M., Zhang, T., Zhang, F., Jiang, Y., Liu, T., Yang, J., & Yang, J. (2020). Genome-wide identification and expression analysis of the histone deacetylase gene family in wheat (*Triticum aestivum* L.). *Plants, 10,* 19.

Kassambara, A., & Mundt, F. (2020). factoextra: Extract and Visualize the Results of Multivariate Data Analyses. R package version 1.0.7. https://CRAN.R-project.org/package=factoextra

Knox, R.E., Clarke, J.M., & DePauw, R.M. (2000). Dicamba and growth condition effects on doubled haploid production in durum wheat crossed with maize. *Plant Breeding, 119,* 289–298.

Kolde, R. (2019). pheatmap: Pretty Heatmaps. R package version 1.0.12. https://CRAN.R-project.org/package=pheatmap

Kuhn, M., Jackson, S., and Cimentada, J. (2020). corrr: Correlations in R. R package version 0.4.3. https://CRAN.R-project.org/package=corrr

Li, H., Liu, H., Pei, X., Chen, H., Li, X., Wang, J., & Wang, C. (2021). Comparative genome-wide analysis and expression profiling of histone acetyltransferases and histone deacetylases involved in the response to drought in wheat. *Journal of Plant Growth Regulation, 41,* 1065-1078.

Li, L., Zhang, Y., Zhang, Y., Li, M., Tian, X., Song, J., Luo, X., Xie, L., Wang, D., & He, Z. (2021). Genome-wide linkage mapping for preharvest sprouting resistance in wheat using 15K single-nucleotide polymorphism arrays. *Frontiers in Plant Science, 12,* 749206.

Malunga, L.N., Ames, N., Khorshidi, A.S., Thandapilly, S.J., Yan, W., Dyck, A., Waterer, J., Malcolmson, L., Cuthbert, R., Sopiwnyk, E., & Scanlon, M.G. (2021). Association of asparagine concentration in wheat with cultivar, location, fertilizer, and their interaction. *Food Chemistry, 344.*

Mertz, E.T., Bates, L.S., & Nelson, O.E. (1964). Mutant gene that changes protein composition and increases lysine content of maize endosperm. *Science, 145,* 279–280.

Moehs, C.P., Austill, W.J., Holm, A., Large, T.A.G., Loeffler, D., Mullenberg, J., Schnable, P.S., Skinner, W., van Boxtel, J., & Wu, L. (2019). Development of decreased-gluten wheat enabled by determination of the genetic basis of lys3a barley. *Plant Physiology, 179*, 1692–1703.

Oddy, J., Alarcón-Reverte, R., Wilkinson, M., Ravet, K., Raffan, S., Minter, A., Mead, A., Elmore, J.S., de Almeida, I.M., Cryer, N.C., Halford, N.G., & Pearce, S. (2021). Reduced free asparagine in wheat grain resulting from a natural deletion of TaASN-B2: investigating and exploiting diversity in the asparagine synthetase gene family to improve wheat quality. *BMC plant biology, 21*, 1-17.

Oddy, J., Raffan, S., Wilkinson, M.D., Elmore, J.S., & Halford, N.G. (2020). Stress, nutrients and genotype: understanding and managing asparagine accumulation in wheat grain. *CABI Agriculture and Bioscience, 1*, 1-14.

Oddy, J., Raffan, S., Wilkinson, M.D., Elmore, J.S., & Halford, N.G. (2022). Understanding the relationships between free asparagine in grain and other traits to breed low-asparagine wheat. *Plants, 11*, 669.

Pedersen, TL. (2021). ggraph: An Implementation of Grammar of Graphics for Graphs and Networks. R package version 2.0.5. https://CRAN.R-project.org/package=ggraph

Peng, Y., Liu, H., Chen, J., Shi, T., Zhang, C., Sun, D., He, Z., Hao, Y., & Chen, W. (2018). Genome-wide association studies of free amino acid levels by six multi-locus models in bread wheat. *Frontiers in Plant Science, 9*, 1196.

Prasanna, B.M., Palacios-Rojas, N., Hossain, F., Muthusamy, V., Menkir, A., Dhliwayo, T., Ndhlela, T., San Vicente, F., Nair, S.K., & Vivek, B.S. (2020). Molecular breeding for nutritionally enriched maize: status and prospects. *Frontiers in Genetics, 1392.*

Przewieslik-Allen, A.M., Wilkinson, P.A., Burridge, A.J., Winfield, M.O., Dai, X., Beaumont, M., King, J., Yang, C., Griffiths, S., & Wingen, L.U. (2021). The role of gene flow and chromosomal instability in shaping the bread wheat genome. *Nature Plants, 7*, 172–183.

R Core Team (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing,Vienna, Austria. https://www.R-project.org/

Raffan, S., Sparks, C., Huttly, A., Hyde, L., Martignago, D., Mead, A., Hanley, S.J., Wilkinson, P.A., Barker, G., Edwards, K.J., Curtis, T.Y., Usher, S., Kosik, O., & Halford, N.G. (2021). Wheat with greatly reduced accumulation of free asparagine in the grain, produced by CRISPR/Cas9 editing of asparagine synthetase gene TaASN2. *Plant Biotechnology Journal, 19*. https://doi.org/10.1111/pbi.13573

Rapp, M., Schwadorf, K., Leiser, W.L., Würschum, T., & Longin, C.F.H. (2018). Assessing the variation and genetic architecture of asparagine content in wheat: What can plant breeding contribute to a reduction in the acrylamide precursor? *Theoretical and Applied Genetics*. https://doi.org/10.1007/s00122-018-3163-x

Rustgi, S., Shewry, P., Brouns, F., Deleu, L.J., & Delcour, J.A. (2019). Wheat seed proteins: factors influencing their content, composition, and technological properties, and strategies to reduce adverse reactions. *Comprehensive Reviews in Food Science and Food Safety, 18*, 1751–1769.

Shewry, P.R., Lovegrove, A., Wingen, L.U., & Griffiths, S. (2022). Opinion exploiting genomics to improve the benefits of wheat: Prospects and limitations. *Journal of Cereal Science, 105*, 10344.

Shorinola, O., Balcárková, B., Hyles, J., Tibbits, J.F.G., Hayden, M.J., Holušova, K., Valárik, M., Distelfeld, A., Torada, A., & Barrero, J.M. (2017). Haplotype analysis of the pre-harvest sprouting resistance locus Phs-A1 reveals a causal role of TaMKK3-A in global germplasm. *Frontiers in Plant Science, 8*, 1555.

Shorinola, O., Simmonds, J., Wingen, L.U., & Uauy, C. (2022). Trend, population structure, and trait mapping from 15 years of national varietal trials of UK winter wheat. *G3*, *12*, jkab415

Walkowiak, S., Gao, L., Monat, C., Haberer, G., Kassa, M. T., Brinton, J., ... & Pozniak, C. J. (2020). Multiple wheat genomes reveal global variation in modern breeding. *Nature, 588*(7837), 277-283.

Wang, L., Zhong, M., Li, X., Yuan, D., Xu, Y., Liu, H., He, Y., Luo, L., & Zhang, Q. (2008). The QTL controlling amino acid content in grains of rice (*Oryza sativa*) are co-localized with the regions involved in the amino acid metabolism pathway. *Molecular Breeding, 21*, 127–137.

Wei, T., and Simko, V. (2021). R package 'corrplot': Visualization of a Correlation Matrix (Version 0.92). https://github.com/taiyun/corrplot

Wickham, H. (2011). The split-apply-combine strategy for data analysis. *Journal of Statistical Software, 40*, 1–29.

Wickham, H. (2016) ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York.

Wickham et al., (2019). Welcome to the tidyverse. *Journal of Open Source Software, 4*(43), 1686, https://doi.org/10.21105/joss.01686

Wickham, H., François, R., Henry, L., and Müller, K. (2022). dplyr: A Grammar of Data Manipulation. R package version 1.0.8. https://CRAN.R-project.org/package=dplyr

Wilke, CO. (2020). cowplot: Streamlined Plot Theme and Plot Annotations for 'ggplot2'. R package version 1.1.1. https://CRAN.R-project.org/package=cowplot

Wilkinson, P.A., Allen, A.M., Tyrrell, S., Wingen, L.U., Bian, X., Winfield, M.O., Burridge, A., Shaw, D.S., Zaucha, J., Griffiths, S., Davey, R.P., Edwards, K.J., & Barker, G.L.A. (2020). CerealsDB - New tools for the analysis of the wheat genome: Update 2020. *Database, 2020,* baaa060.

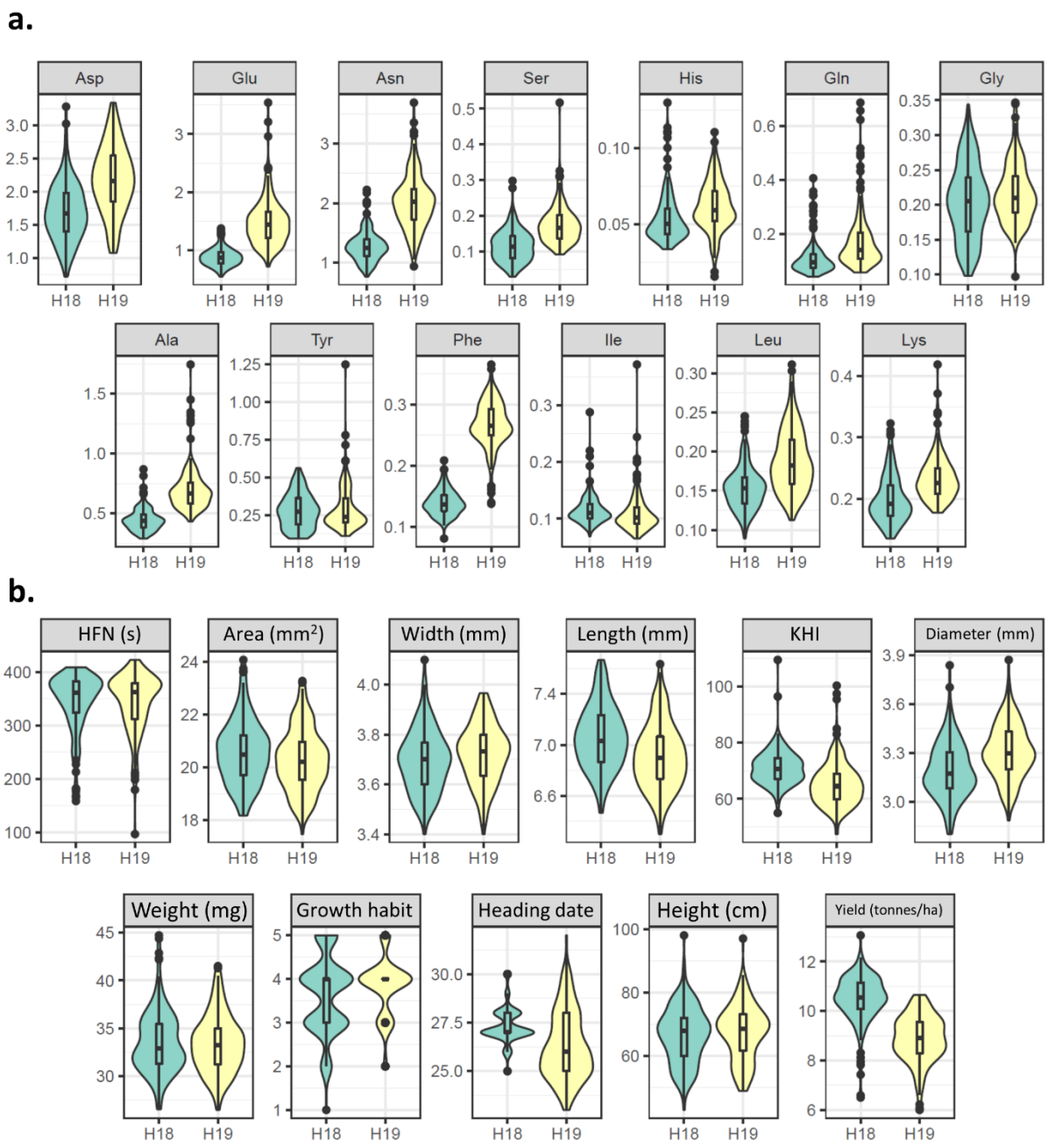
Wilkinson, P.A., Winfield, M.O., Barker, G.L.A., Allen, A.M., Burridge, A., Coghill, J.A., & Edwards, K.J. (2012). CerealsDB 2.0: an integrated resource for plant breeders and scientists. *BMC Bioinformatics, 13,* 1–6.

Xie, Y., Malunga, L.N., Ames, N.P., Waterer, J., Khorshidi, A.S., & Scanlon, M.G. (2021). Effects of growing environment, genotype, and commercial fertilization levels on free asparagine concentration in Western Canadian wheat. *Cereal Chemistry, 98*(1), 89-99.

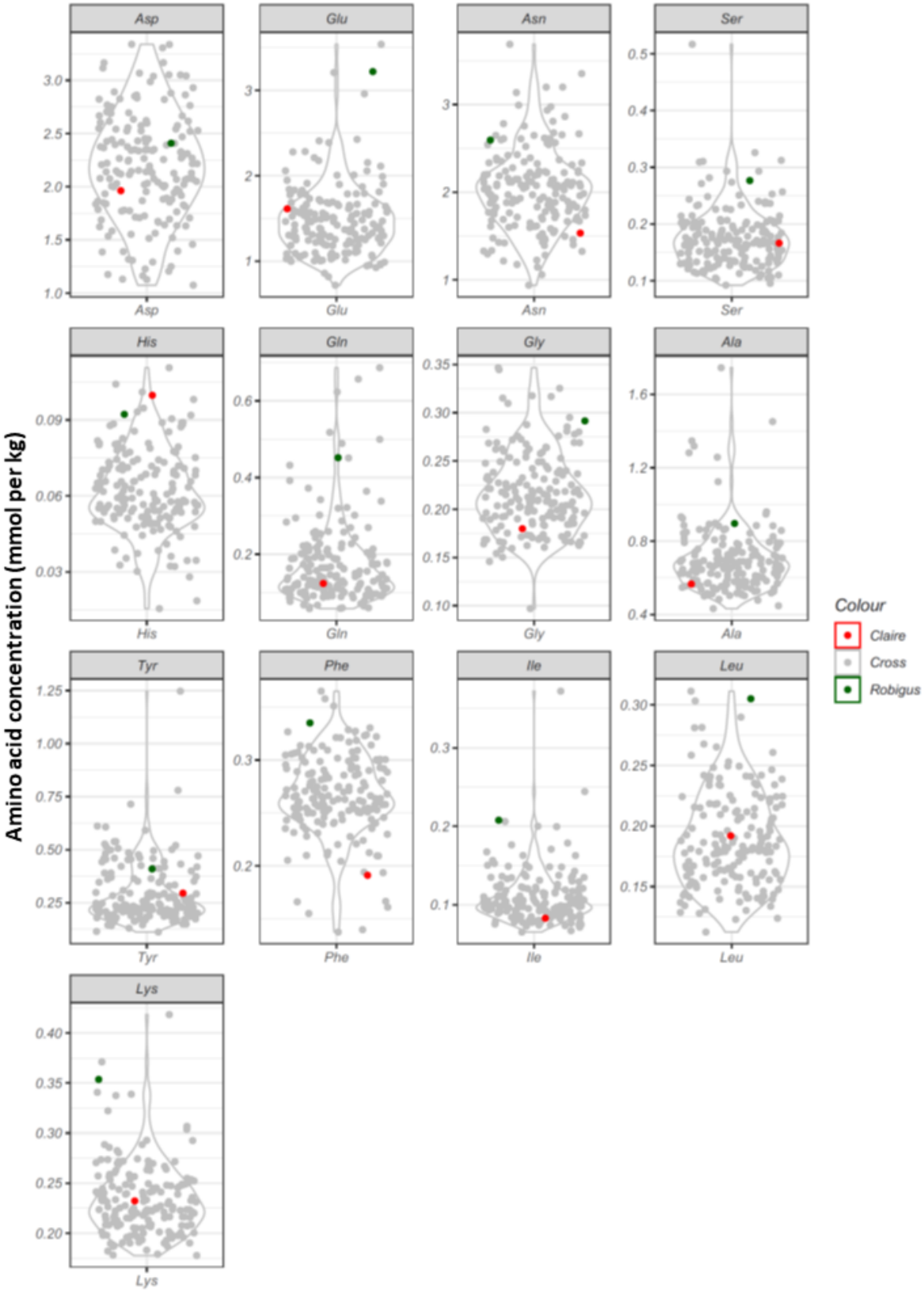
Yoo, S.-C. (2017). Quantitative trait loci controlling the amino acid content in rice (Oryza sativa L.). *Journal of Plant Biotechnology*, *44*, 349–355.

Zhong, M., Wang, L., Yuan, D., Luo, L., Xu, C., & He, Y. (2011). Identification of QTL affecting protein and amino acid contents in rice. *Rice Science*, *18*, 187–195.

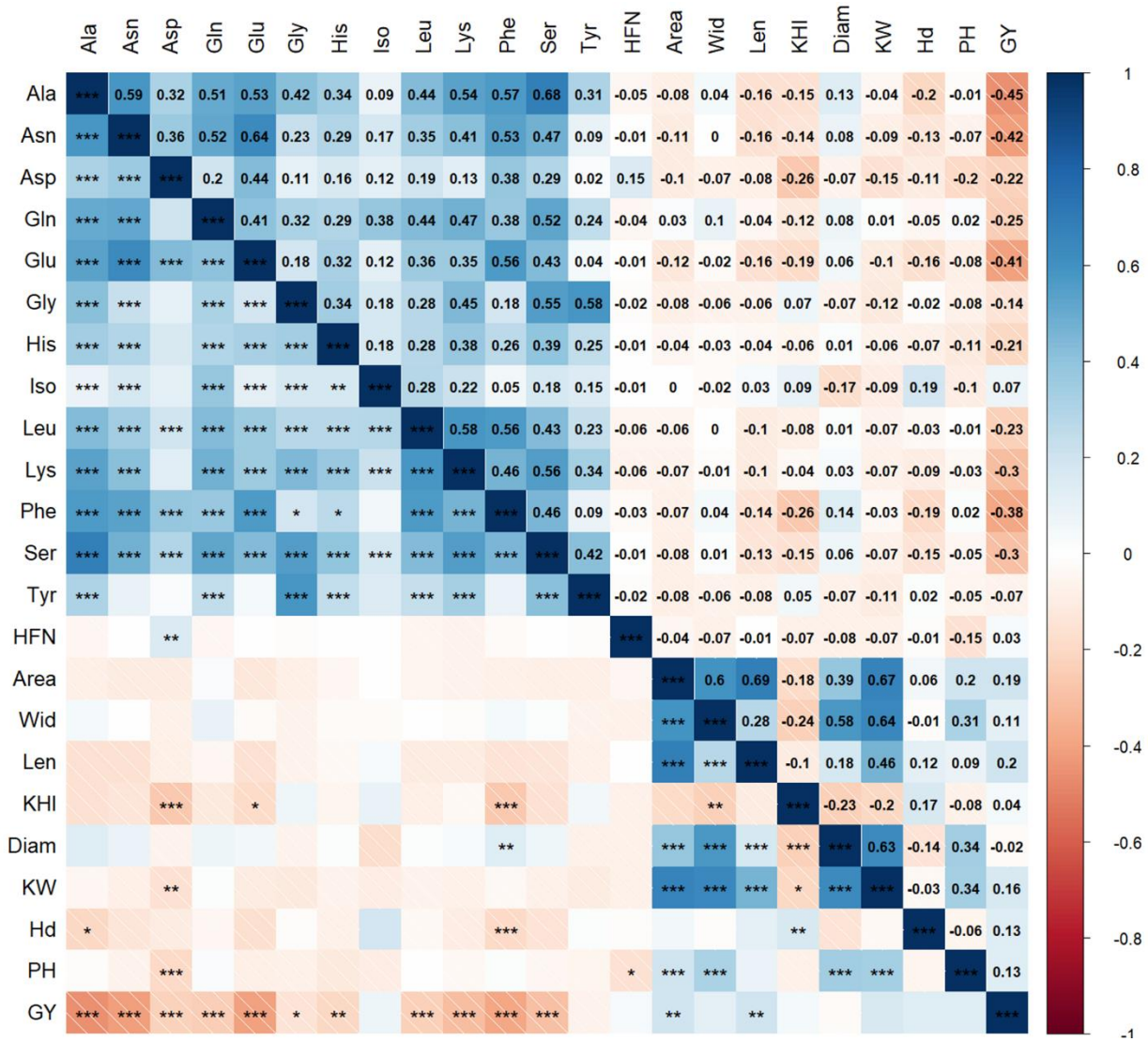
5.8. Supplementary figures and tables



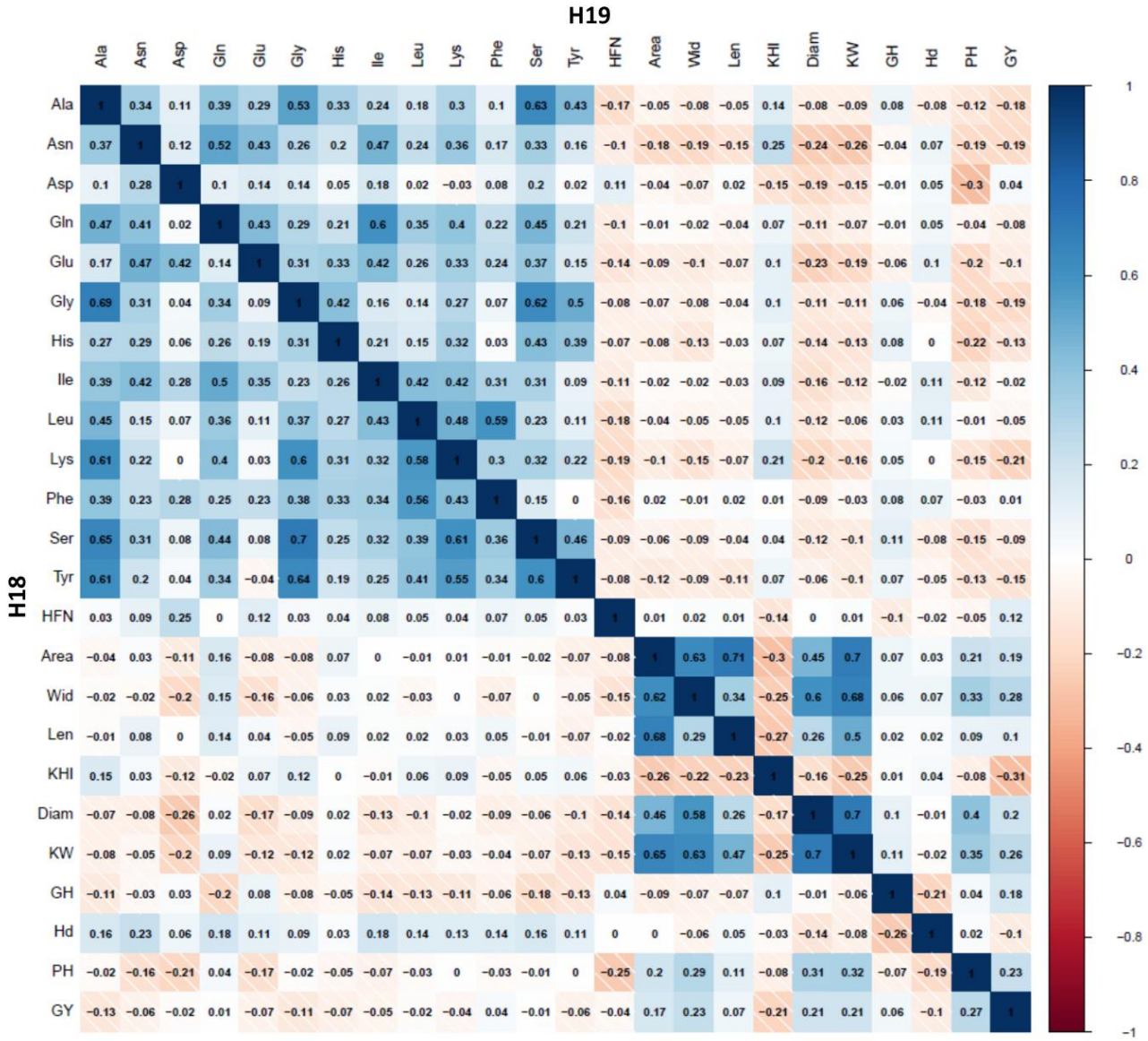
Supplementary figure 5.1. Trait measurements from the Claire x Robigus mapping population grown across two years (H18 and H19). a. Concentration of amino acids (mmol per kg) in wheat grain in both environments. **b.** Measurements of other quality and agronomic traits across both environments.



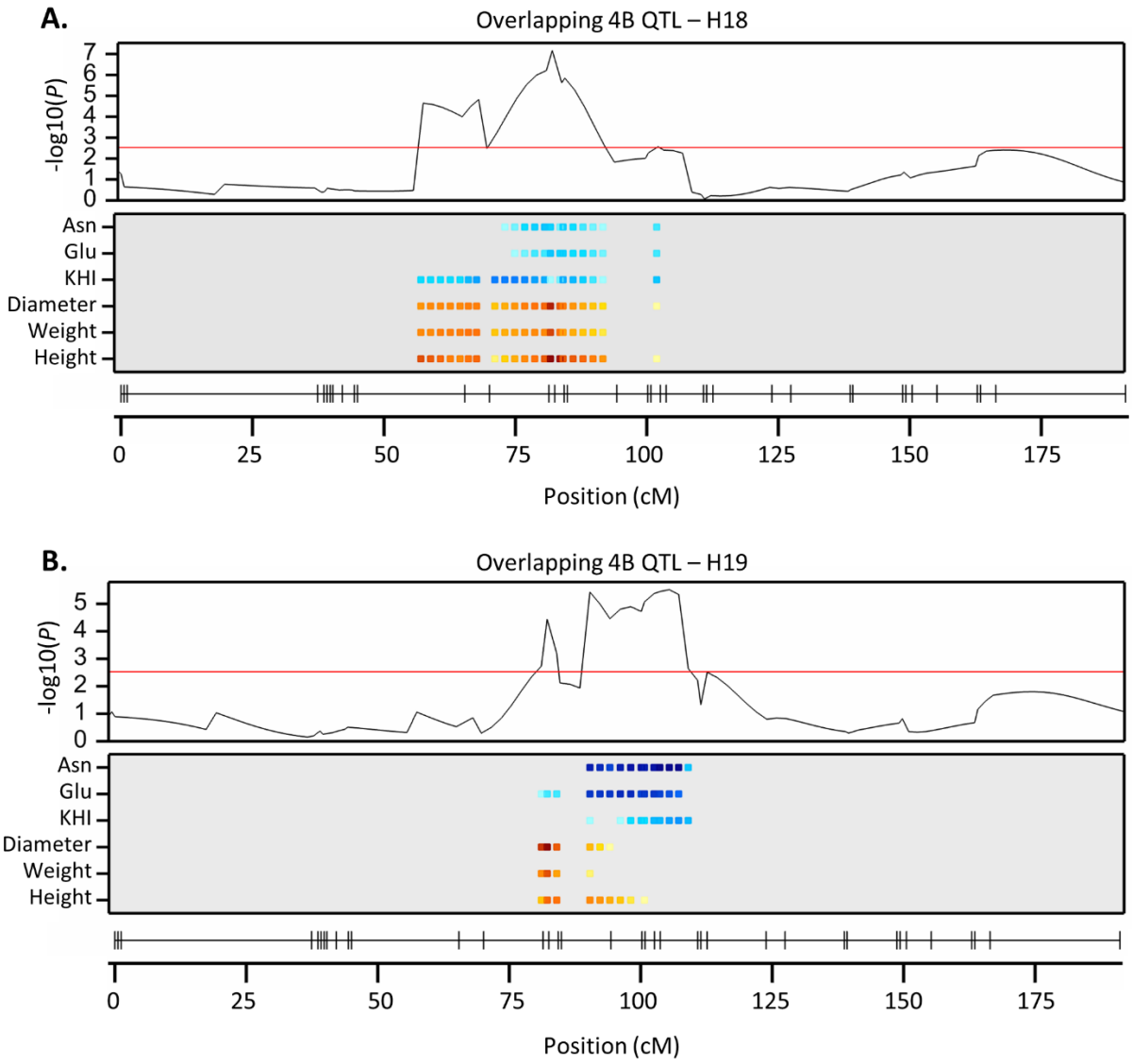
Supplementary figure 5.2. Grain amino acid concentrations in the Claire × Robigus mapping population in the 2018 – 2019 trial, relative to parental lines Claire and Robigus. Data points are jittered on the x axis to improve visualisation of the data distribution.



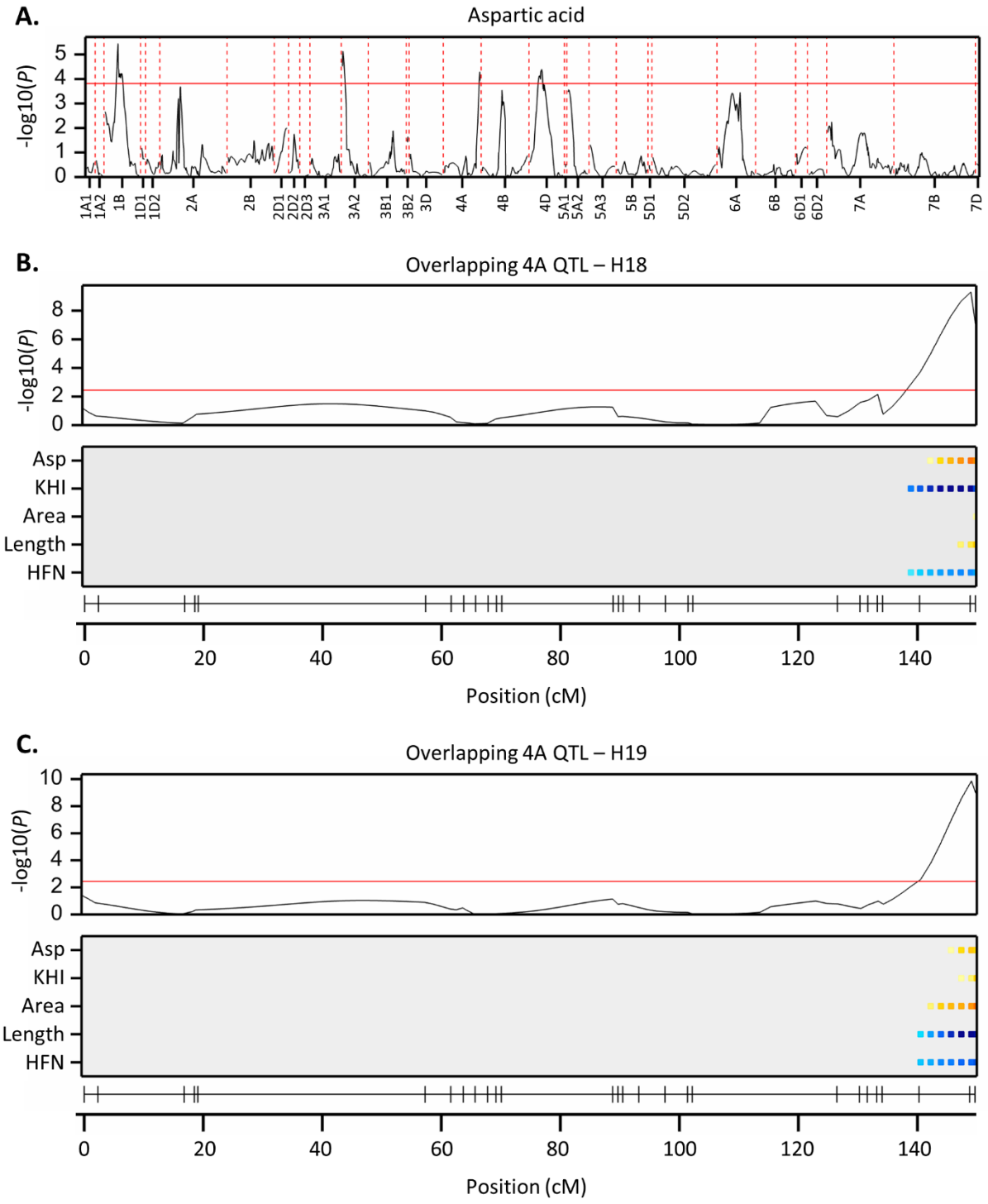
Supplementary figure 5.3. Correlation of traits taken from both H18 and H19 environments. Kendall correlation coefficients shown in upper right triangle and significance asterisks from adjusted p values (Bonferroni correction) shown in lower left triangle. HFN (Hagberg falling number), KHI (kernel hardness index), Diam (diameter), KW (kernel weight), Hd (heading date), PH (plant height), GY (grain yield).



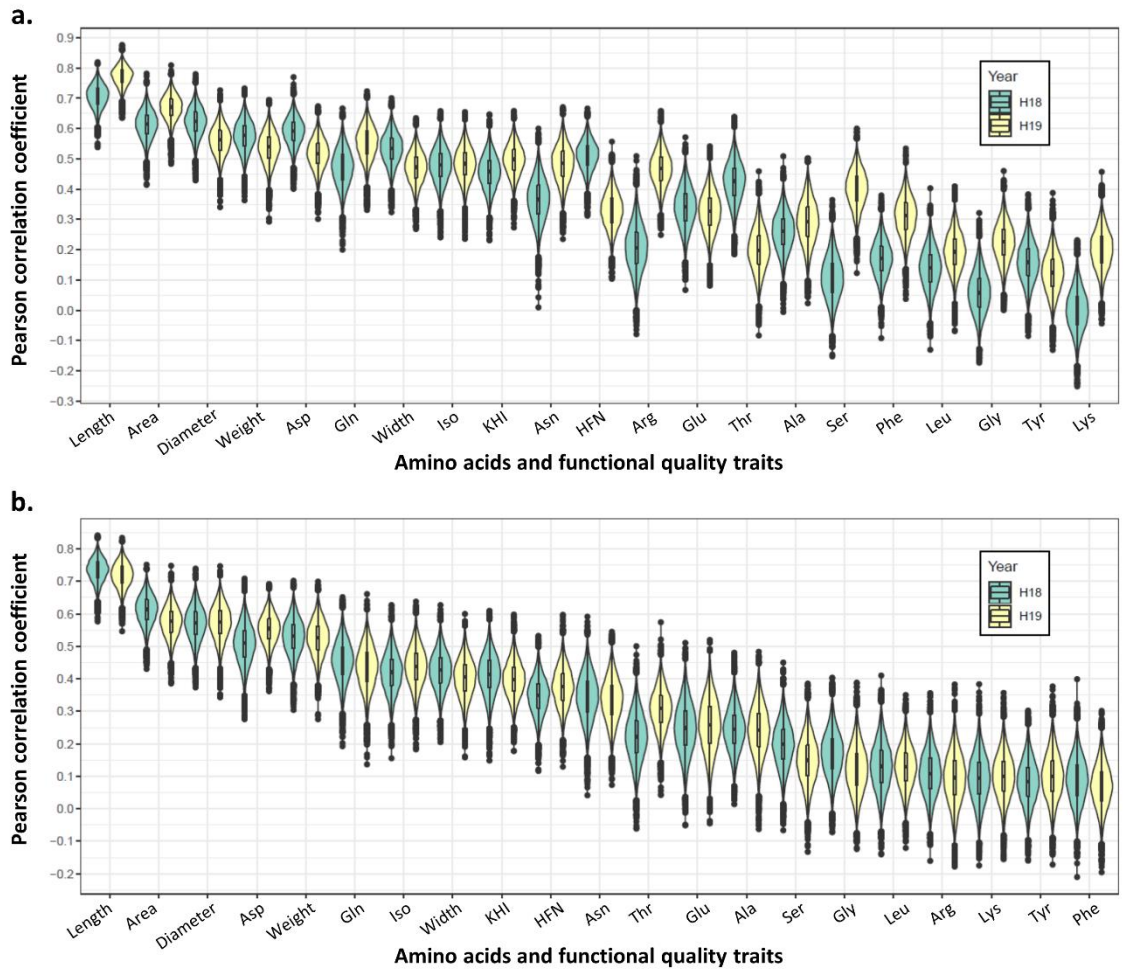
Supplementary figure 5.4. Kendall correlation coefficients between traits within each environment. H18 is the bottom left triangle, H19 is the upper right triangle. HFN (Hagberg falling number), KHI (kernel hardness index), Diam (diameter), KW (kernel weight), Hd (heading date), PH (plant height), GY (grain yield).



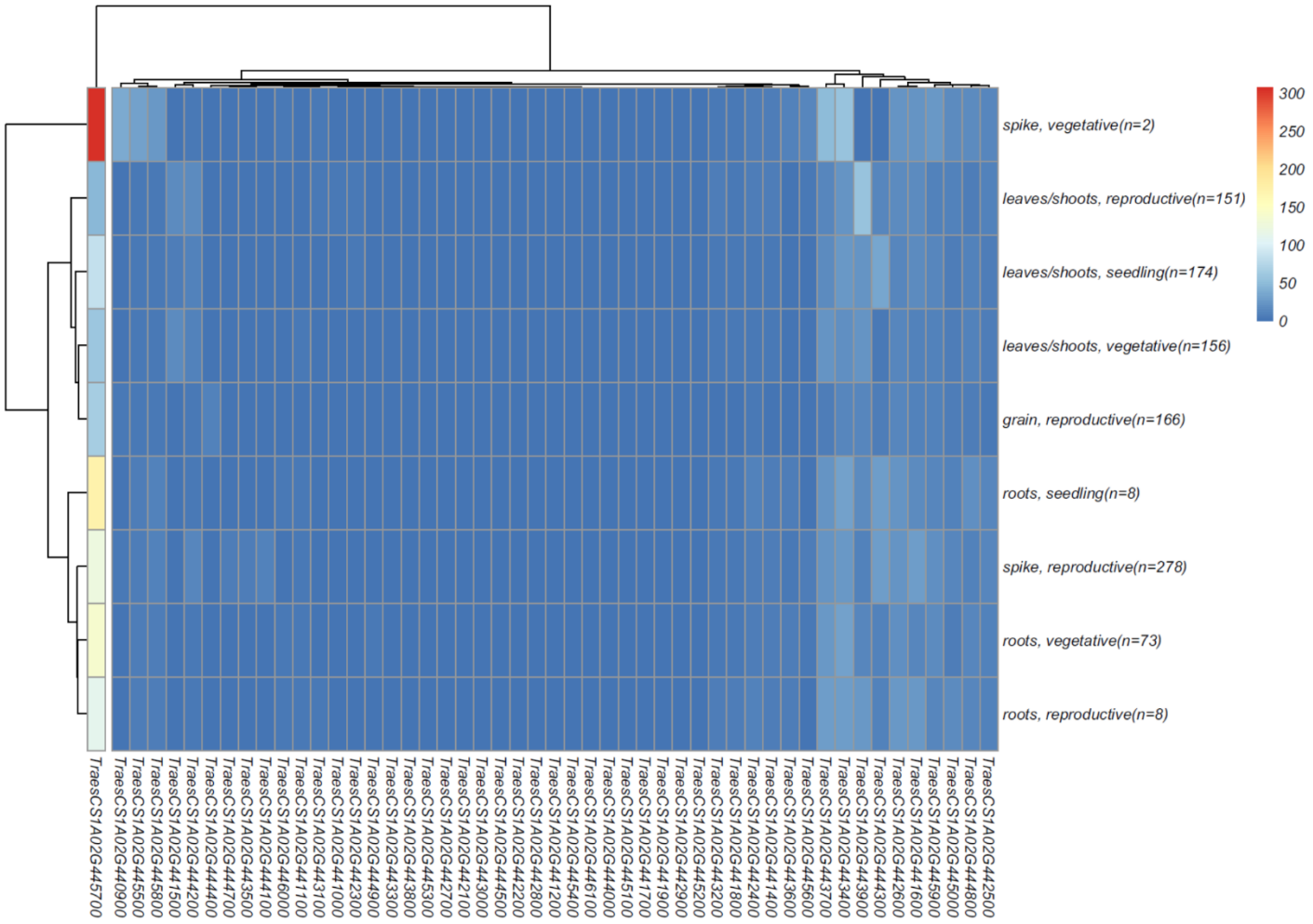
Supplementary figure 5.5. Multi-trait analysis of the asparagine QTL on 4B in the Claire x Robigus mapping population. Blue indicates Robigus additive allele whilst red indicates Claire additive allele. The darkness of colour corresponds to the magnitude of the effect. KHI (kernel hardness index).



Supplementary figure 5.6. Multi-environment and multi-trait linkage analysis of aspartic acid QTL in the Claire x Robigus mapping population. KHI (kernel hardness index), HFN (Hagberg falling number).



Supplementary figure 5.7. Accuracy of genomic selection for each trait measured in the mapping population using within (a.) and between (b.) year prediction.



Supplementary figure 5.8. Expression (transcripts per million) of candidate genes residing in the QTL for lysine in different tissues and developmental stages. The tissue is written first (spike, leaves/shoots, grain, roots) followed by the developmental stage (seedling, vegetative, reproductive). Candidate genes are given as their accession in Chinese Spring. Data obtained from expVIP.

Supplementary Table 5.1. Broad sense heritability estimates, Kendall correlation coefficients for traits across both environments, and within and between environments average genomic prediction accuracies (as Pearson correlation coefficients). Heritability estimates and genomic prediction performed on transformed data, Kendall correlation coefficients performed on non-transformed data. KHI (Kernel Hardness Index), HFN (Hagberg Falling Number), h^2 (heritability), *r* (Kendall correlation coefficient), GP (genomic prediction), H18 (first field trial), H19 (second field trial).

			Within environment GP *r*			**Between environments GP *r***		
Traits	h^2	***r***	**H18**	**H19**	**Mean**	**H18 train**	**H19 train**	**Mean**
Amino acids								
Alanine	0.37	0.00	0.26	0.29	0.27	0.24	0.24	0.24
Asparagine	0.60	0.34	0.36	0.48	0.42	0.35	0.33	0.34
Aspartic acid	0.82	0.51	0.59	0.52	0.55	0.51	0.55	0.53
Glutamic acid	0.61	0.33	0.34	0.33	0.33	0.25	0.26	0.25
Glutamine	0.69	0.30	0.47	0.55	0.51	0.45	0.43	0.44
Glycine	0.00	-0.12	0.06	0.22	0.14	0.17	0.12	0.15
Isoleucine	0.71	0.39	0.48	0.48	0.48	0.42	0.44	0.43
Leucine	0.35	0.18	0.14	0.19	0.17	0.13	0.13	0.13
Lysine	0.45	0.15	0.00	0.20	0.10	0.09	0.10	0.10
Phenylalanine	0.34	0.17	0.17	0.31	0.24	0.09	0.07	0.08
Serine	0.00	-0.14	0.11	0.40	0.25	0.20	0.15	0.17
Tyrosine	0.00	-0.23	0.16	0.12	0.14	0.08	0.10	0.09
Functional traits								
Area	0.91	0.63	0.61	0.67	0.64	0.61	0.58	0.59
Length	0.96	0.77	0.71	0.77	0.74	0.73	0.72	0.73
Width	0.84	0.55	0.53	0.47	0.50	0.43	0.40	0.41
Diameter	0.90	0.60	0.62	0.56	0.59	0.57	0.57	0.57
KHI	0.85	0.58	0.46	0.50	0.48	0.41	0.40	0.41
Weight	0.89	0.59	0.58	0.54	0.56	0.53	0.53	0.53
HFN	0.64	0.34	0.51	0.33	0.42	0.35	0.37	0.36

Supplementary Table 5.2. Multi-environment QTL for measured amino acids. Chr (chromosome), cM (centimorgan), Mbp (megabase pair location in Chinese Spring), QTL x E (QTL by environment interaction), ratio (ratio of effects between years), S.E. (standard error of the mean), % Expl. (percentage of variance explained).

	Single trait linkage analysis (H18 and H19)							H18					H19				
Trait	**Marker**	**Chr.**	**cM**	**Mbp**	**-log$_{10}$(*p*)**	**QTL x E**	**Ratio**	**Effect**	**S.E.**	**High value**	***p***	**% Expl.**	**Effect**	**S.E.**	**High value**	***p***	**% Expl.**
Asparagine	WC.0221262	4B	114.47	601	5.96	yes	4.34	0.041	0.020	Robigus	0.040	2.6	0.178	0.034	Robigus	<0.001	14.8
Log$_e$ alanine	WC.0218011	1A2	27.3	593	5.10	yes	0.34	0.065	0.016	Claire	<0.001	10.9	0.022	0.015	Robigus	0.159	1
	WC.0223839	7B	211.2	719	5.03	no	1.00	0.052	0.012	Robigus	<0.001	7.1	0.052	0.012	Robigus	<0.001	5.7
Aspartic acid	WC.0218489	1B	54.4	530	5.40	no	1.00	0.123	0.027	Claire	<0.001	8	0.123	0.027	Claire	<0.001	5.9
	WC.0214359	3A2	2.3	738	7.95	yes	1.68	0.118	0.029	Robigus	<0.001	7.3	0.198	0.033	Robigus	<0.001	15.3
	WC.0221037	4A	148.8	703	8.08	no	1.00	0.154	0.027	Claire	<0.001	12.6	0.154	0.027	Claire	<0.001	9.3
	WC.0227146	4D	48.8	16	3.70	no	1.00	0.102	0.028	Claire	<0.001	5.5	0.102	0.028	Claire	<0.001	4.1
Log$_e$ glutamate	WC.0221329	4B	100.8	518	4.27	yes	2.41	0.034	0.013	Robigus	0.011	3.7	0.082	0.019	Robigus	<0.001	10.1
Log$_e$ glutamine	WC.0218486	1B	117.93	660	2.16	yes	9.91	0.011	0.038	Robigus	0.760	0.1	0.109	0.038	Robigus	0.004	4.9
	WC.0221302	4B	103.7	547	3.50	no	1.00	0.105	0.029	Robigus	<0.001	5.4	0.105	0.029	Robigus	<0.001	4.5
	WC.0228471	6B	19.7	25	5.09	no	1.00	0.129	0.029	Claire	<0.001	8.2	0.129	0.029	Claire	<0.001	6.7
Glycine	WC.0218011	1A2	27.3	593	5.37	yes	0.53	0.017	0.004	Claire	<0.001	9.6	0.009	0.003	Robigus	0.005	4.3
	WC.0226796	4B	155.2	327	4.26	no	1.00	0.010	0.002	Robigus	<0.001	3.2	0.010	0.002	Robigus	<0.001	5.3
-1/Isoleucine	WC.0221386	4B	94.3	172	7.31	yes	5.10	0.161	0.120	Robigus	0.180	1	0.821	0.151	Robigus	<0.001	14.1
	WC.0223785	7B	211.2	717	3.60	no	1.00	0.421	0.115	Robigus	<0.001	6.8	0.421	0.115	Robigus	<0.001	3.7
Log$_e$ lysine	WC.0218011	1A2	27.3	593	4.95	yes	0.38	0.064	0.014	Claire	<0.001	12.1	0.024	0.011	Claire	0.038	2.6
Phenylalanine	WC.0220622	3B1	78.1	116	3.83	yes	1.80	0.005	0.002	Robigus	0.001	6.2	0.009	0.003	Robigus	0.002	5.6
Log$_e$ serine	WC.0226730	4B	123.8	632	3.39	yes	2.33	0.033	0.031	Claire	0.279	0.7	0.077	0.020	Robigus	<0.001	7.9

Supplementary Table 5.3. Multi-environment QTL for selected grain traits. Chr (chromosome), cM (centimorgan), Mbp (megabase pair location in Chinese Spring), QTL x E (QTL by environment interaction), ratio (ratio of effects between years), S.E. (standard error of the mean), % Expl. (percentage of variance explained), KHI (Kernel Hardness Index), HFN (Hagberg Falling Number).

Trait	Single trait linkage analysis (H18 and H19)							H18					H19				
	Marker	Chr.	cM	Mbp	-log$_{10}$(*p*)	QTLxE	Ratio	Effect	High value	S.E.	*p*	% Expl.	Effect	High value	S.E.	*p*	% Expl.
Log$_e$ KHI	WC.0221037	4A	148.8	703	8.26	yes	1.41	0.032	Robigus	0.006	<0.001	14.8	0.045	Robigus	0.008	<0.001	14.7
	WC.0226741	4B	110.8	594	4.30	yes	2.06	0.017	Robigus	0.006	0.002	4.2	0.035	Robigus	0.008	<0.001	8.6
	WC.0222754	6A	83.22	108	6.09	yes	4.40	0.005	Claire	0.006	0.377	0.4	0.022	Robigus	0.008	0.007	3.4
	WC.0228678	7A	152.17	539	4.45	no	1.00	0.024	Claire	0.006	<0.001	8.1	0.024	Claire	0.006	<0.001	4.1
Area	WC.0217441	2A	70.8	176	7.48	yes	1.43	0.325	Robigus	0.080	<0.001	7.4	0.464	Robigus	0.080	<0.001	15.8
	WC.0220938	4A	149.7	709	2.22	no	1.00	0.208	Claire	0.076	0.006	3	0.208	Claire	0.076	0.006	3.2
	WC.0193228	7A	158.09	610	4.94	no	1.00	0.360	Robigus	0.082	<0.001	9.1	0.360	Robigus	0.082	<0.001	9.5
Length	WC.0225130	2A	67.2	166	8.15	no	1.00	0.077	Robigus	0.013	<0.001	9.6	0.077	Robigus	0.013	<0.001	9.1
	WC.0212864	3A1	92.1	63	4.86	no	1.00	0.058	Robigus	0.013	<0.001	5.5	0.058	Robigus	0.013	<0.001	5.1
	WC.0225952	3B1	128.8	739	7.18	no	1.00	0.072	Claire	0.013	<0.001	8.5	0.072	Claire	0.013	<0.001	8
	WC.0221119	4A	149.7	702	7.12	yes	1.91	0.034	Claire	0.014	0.019	1.8	0.065	Claire	0.014	<0.001	6.5
	WC.0221932	5A1	0.6	1	9.23	no	1.00	0.084	Robigus	0.014	<0.001	11.4	0.084	Robigus	0.014	<0.001	10.7
	WC.0223413	7A	171.6	641	7.23	no	1.00	0.072	Robigus	0.013	<0.001	8.4	0.072	Robigus	0.013	<0.001	7.9
	WC.0222406	7B	40.6	34	7.35	yes	1.54	0.048	Claire	0.014	0.001	3.7	0.074	Claire	0.014	<0.001	8.4
Width	WC.0217441	2A	72.25	176	4.88	no	1.00	0.031	Robigus	0.007	<0.001	6.1	0.031	Robigus	0.007	<0.001	6.9
	WC.0095497	2B	120.78	545	4.04	yes	3.44	0.009	Claire	0.008	0.257	0.5	0.031	Claire	0.008	<0.001	7.1
	WC.0227146	4D	48.8	16	5.98	no	1.00	0.035	Robigus	0.007	<0.001	8	0.035	Robigus	0.007	<0.001	8.9
	WC.0228194	6A	86.1	127	7.35	yes	0.59	0.044	Robigus	0.008	<0.001	12.9	0.026	Robigus	0.008	0.001	4.9
	WC.0228849	7A	174.8	641	3.04	no	1.00	0.024	Robigus	0.007	0.001	3.8	0.024	Robigus	0.007	0.001	4.2
	WC.0212528	7B	107.2	641	4.60	yes	0.11	0.028	Robigus	0.008	<0.001	5.2	0.003	Robigus	0.008	0.732	0.1
Diameter	WC.0217441	2A	72.25	176	3.62	no	1.00	0.030	Robigus	0.008	<0.001	3.1	0.030	Robigus	0.008	<0.001	3.2
	WC.0226154	3B1	81.6	237	9.72	no	1.00	0.051	Claire	0.008	<0.001	8.9	0.051	Claire	0.008	<0.001	9.4
	WC.0226868	4B	82.5	32	18.53	no	1.00	0.072	Claire	0.008	<0.001	17.7	0.072	Claire	0.008	<0.001	18.7
	WC.0227146	4D	48.8	16	7.84	no	1.00	0.047	Robigus	0.008	<0.001	7.6	0.047	Robigus	0.008	<0.001	8.1
	WC.0221859	5A2	2.3	30	5.62	no	1.00	0.038	Claire	0.008	<0.001	5	0.038	Claire	0.008	<0.001	5.2
	WC.0212957	6A	99.37	499	8.73	no	1.00	0.050	Robigus	0.008	<0.001	8.4	0.050	Robigus	0.008	<0.001	8.8
	WC.0222406	7B	40.6	34	4.43	no	1.00	0.034	Claire	0.008	<0.001	3.9	0.034	Claire	0.008	<0.001	4.1
Weight	WC.0217441	2A	72.25	176	6.45	no	1.00	0.936	Robigus	0.184	<0.001	8.4	0.936	Robigus	0.184	<0.001	9.7
	WC.0229571	2B	9.37	12	3.18	no	1.00	0.629	Claire	0.185	0.001	3.8	0.629	Claire	0.185	0.001	4.4
	WC.0226868	4B	82.5	32	9.25	no	1.00	1.127	Claire	0.182	<0.001	12.2	1.127	Claire	0.182	<0.001	14.1
	WC.0212957	6A	99.37	499	8.28	no	1.00	1.085	Robigus	0.186	<0.001	11.3	1.085	Robigus	0.186	<0.001	13.1

-1/(500-HFN)	WC.0188904	4A	147.1	733	8.24	no	1.00	0.001	Robigus	0.000	<0.001	11.5	0.001	Robigus	0.000	<0.001	10.3
	WC.0227149	4D	56.9	17	10.92	yes	0.00	0.001	Robigus	0.000	<0.001	23.5	0.000	Robigus	0.000	<0.001	5.6
Heading date	WC.0226616	4A	130.3	632	4.14	no	1.00	0.258	Claire	0.065	<0.001	8.6	0.258	Claire	0.065	<0.001	2.2
	WC.0221774	5A3	104.9	706	4.51	yes	3.88	0.146	Claire	0.065	0.024	2.8	0.566	Claire	0.124	<0.001	10.5
Plant height	WC.0226868	4B	82.5	32	17.98	no	1.00	3.921	Claire	0.444	<0.001	19.8	3.921	Claire	0.444	<0.001	20.7
	WC.0213051	4D	56.9	17	28.97	yes	1.16	4.647	Robigus	0.471	<0.001	27.8	5.376	Robigus	0.471	<0.001	38.9
Grain yield	WC.0227571	7B	16.5	328	5.62	yes	0.66	0.354	Robigus	0.070	<0.001	13.9	0.235	Robigus	0.070	0.001	6.2

Supplementary Table 5.4. Sources of variation related to asparagine and falling number screened in this study in the Claire x Robigus mapping population. Chr. (chromosome).

Source of variation	Claire	Robigus	Chr.	Reference
ASN-B2 PAV	Absent	Absent	3B	Oddy et al., 2021
ASN-B1	Non-functional	Functional	5B	Oddy et al., 2021
ASN-A3.1	Non-functional	Non-functional	1A	Oddy et al., 2021
Rht-B1	*Rht-B1a* (WT)	*Rht-B1b* (Dwarf)	4B	Wilkinson et al., 2020
Rht-D1	*Rht-D1b* (Dwarf)	*Rht-D1a* (WT)	4D	Wilkinson et al., 2020
T. dicoccoides introgression	Absent	Present	4A	Przewieslik-Allen et al., 2021
TaMKK3A	A	A	4A	Shorinola et al., 2016
PM19-A1 promoter InDel	Deletion	Deletion	4A	Shorinola et al., 2016

Supplementary Table 5.5. Physical locations of the Hagberg Falling Number (HFN), asparagine (Asn), and lysine (Lys) QTL in Chinese Spring. Chr. (chromosome), Mbp (megabase pair location), CI (confidence interval), bp (base pair).

Trait	Chr.	Peak (Mbp)	Lower CI (Mbp)	Upper CI (Mbp)	QTL size (bp)	No. of genes
HFN	4A	733	691	745	54,058,906	824
Asn	4B	601	533	632	96,765,195	754
Lys	1A	593	590	594	4,471,109	50

Supplementary Table 5.6. Physical location of the lysine QTL in the wheat pangenome (chromosome level assemblies) and gene content.

Genome	Chr	Lower CI	Upper CI	QTL size	No. of genes
IWGSC	1A	590	594	4,471,109	50
ArinaLrFor	1A	598	603	4,559,393	55
Jagger	1A	592	596	4,481,576	60
Julius	NA	NA	NA	NA	NA
Lancer	1A	591	595	4,548,456	59
Landmark	1A	593	595	2,101,319	38
Mace	1A	586	591	4,477,846	55
SY Mattis	1A	596	601	4,560,453	66
Norin61	1A	589	594	4,645,175	58
Stanley	scaffold_v3_2071	5	NA	NA	NA

Supplementary Table 5.7. Knetminer analaysis of candidate genes in the lysine QTL on chromosome 1A. Gene accessions are from Chinese Spring. Knetscore was obtained from Knetminer using keywords "lysine" and "grain protein". The total Knetscore represents the Knetscore without any keyword filtering. The normalised Knetscore is the Knetscore divided by the total Knetscore. Genes are ranked from highest normalised Knetscore to lowest.

Accession	Gene name	Knetscore	Total Knetscore	Normalised Knetscore
TRAESCS1A02G441200	ABI2	925.05	2441.35	0.38
TRAESCS1A02G443700	TRAESCS1A02G443700	21.07	95.44	0.22
TRAESCS1A02G443600	NLE1	37.66	205.90	0.18
TRAESCS1A02G445700	HDT3	141.74	797.61	0.18
TRAESCS1A02G441000	TRAESCS1A02G441000	24.25	152.19	0.16
TRAESCS1A02G443900	UGE2	81.68	586.08	0.14
TRAESCS1A02G441500	TRAESCS1A02G441500	14.41	104.80	0.14
TRAESCS1A02G442600	TRAESCS1A02G442600	7.66	57.80	0.13
TRAESCS1A02G443500	ACA1	13.42	114.24	0.12
TRAESCS1A02G445000	SMC5	19.57	172.24	0.11
TRAESCS1A02G442200	CYP75B1	27.84	272.94	0.10
TRAESCS1A02G440900	TRAESCS1A02G440900	5.32	52.82	0.10
TRAESCS1A02G442800	CKL7	31.99	344.35	0.09
TRAESCS1A02G444700	NANA	9.67	104.95	0.09
TRAESCS1A02G443200	TRAESCS1A02G443200	44.52	493.42	0.09
TRAESCS1A02G445800	TRAESCS1A02G445800	16.70	188.32	0.09
TRAESCS1A02G441100	TRAESCS1A02G441100	36.82	419.12	0.09
TRAESCS1A02G441600	TRAESCS1A02G441600	8.49	104.28	0.08
TRAESCS1A02G442400	RGD3	11.67	150.50	0.08
TRAESCS1A02G443100	GRIK2	91.61	1190.50	0.08
TRAESCS1A02G444800	TRAESCS1A02G444800	17.22	226.55	0.08
TRAESCS1A02G446000	TRAESCS1A02G446000	23.23	322.73	0.07
TRAESCS1A02G443400	RH15	48.49	677.60	0.07
TRAESCS1A02G444900	ATL42	31.03	435.91	0.07
TRAESCS1A02G441400	TRAESCS1A02G441400	25.45	361.70	0.07
TRAESCS1A02G442100	CYP81D1	126.83	1835.06	0.07
TRAESCS1A02G442500	FTSH4	27.64	408.35	0.07
TRAESCS1A02G444300	NDK4	37.80	574.85	0.07
TRAESCS1A02G441800	TRAESCS1A02G441800	22.40	344.37	0.07
TRAESCS1A02G442300	CYP75B1	13.43	207.77	0.06
TRAESCS1A02G444500	BAM3	68.80	1162.49	0.06
TRAESCS1A02G445500	AGO4	19.24	457.74	0.04
TRAESCS1A02G445900	TRAESCS1A02G445900	2.27	62.34	0.04
TRAESCS1A02G445600	TRAESCS1A02G445600	16.16	565.10	0.03
TRAESCS1A02G441700	TRAESCS1A02G441700	0.00	4.69	0.00
TRAESCS1A02G441900	TRAESCS1A02G441900	0.00	51.12	0.00

TRAESCS1A02G442700	TRAESCS1A02G442700	0.00	21.09	0.00
TRAESCS1A02G442900	TRAESCS1A02G442900	0.00	52.06	0.00
TRAESCS1A02G443000	TRAESCS1A02G443000	0.00	78.46	0.00
TRAESCS1A02G443300	TRAESCS1A02G443300	0.00	35.47	0.00
TRAESCS1A02G443800	TRAESCS1A02G443800	0.00	39.49	0.00
TRAESCS1A02G444000	TRAESCS1A02G444000	0.00	83.65	0.00
TRAESCS1A02G444100	TRAESCS1A02G444100	0.00	49.88	0.00
TRAESCS1A02G444200	TRAESCS1A02G444200	0.00	7.82	0.00
TRAESCS1A02G444400	TRAESCS1A02G444400	0.00	76.16	0.00
TRAESCS1A02G445100	TRAESCS1A02G445100	0.00	40.24	0.00
TRAESCS1A02G445200	TRAESCS1A02G445200	0.00	41.99	0.00
TRAESCS1A02G445300	TRAESCS1A02G445300	0.00	48.70	0.00
TRAESCS1A02G445400	TRAESCS1A02G445400	0.00	30.87	0.00
TRAESCS1A02G446100	TRAESCS1A02G446100	0.00	55.12	0.00

6

Potential for reducing dietary-acrylamide exposure from wheat products through crop management and imaging

Joseph Oddy[1], John Addy[1], Andrew Mead[1], Chris Hall[1], Chris Mackay[1], Tom Ashfield[1,2], Faye McDiarmid[2], Tanya Y. Curtis[3], Sarah Raffan[1], Mark Wilkinson[1], J. Stephen Elmore[4], Nicholas Cryer[5], Isabel Moreira de Almeida[6], and Nigel G. Halford[1*]

1. *Rothamsted Research, Harpenden, Hertfordshire AL5 2JQ, United Kingdom*
2. Crop Health and Protection (CHAP), Rothamsted, Harpenden AL5 2JQ, United Kingdom
3. Curtis Analytics Limited, Discovery Park, Sandwich, United Kingdom
4. Department of Food & Nutritional Sciences, University of Reading, RG6 6DZ, UK
5. Mondelēz UK R&D Ltd, Bournville Lane, Bournville, Birmingham, B30 2LU, UK
6. Mondelēz R&D International, 6 Rue René Razel, 91400 Saclay, France

As submitted to Journal of Agricultural and Food Chemistry.
This chapter has been modified from its submitted form to allow for edits and corrections within this Thesis.

6.1. Introduction to paper

The two field trials described in this chapter were designed as the main component of the collaboration between Rothamsted Research, University of Reading, and Mondelēz International. In the original plan, field trials consisting of twelve varieties and eleven treatments across three years were intended to take place. Of most interest in these planned field trials was the best application ratio of nitrogen and sulphur to prevent free asparagine accumulation. The presence/absence of phosphorus and potassium was also investigated, but sulphur deficiency was already known to be one of the greatest influencers of grain free asparagine content. However, due to unforeseen challenges arising within the project, only two trials were analysed (with one trial having four varieties only analysed). Despite these difficulties, our findings from this study still provided useful insight into which agronomic treatments are best to control free asparagine accumulation in wheat grain.

In addition to the analysis of different agronomic treatments in this study, this PhD project was also designed with the intention to bake biscuits from the grains from these field trials at the Mondelēz Biscuit Research and Development site in Saclay to test biscuit quality and acrylamide content. However due to the COVID-19 pandemic, this analysis was performed instead at Reading Scientific Services Ltd. Nevertheless, the facilities available there enabled an analysis of biscuits baked from different flours and yielded useful information regarding how agronomy impacts biscuit quality.

The last aspect of this study to be designed was the multispectral imaging of plants and seeds. After discussions regarding how best to measure fertiliser uptake to ensure that plants in the second trial were adequately fertilised , I talked with other researchers at Rothamsted who recommended using a multispectral imaging device. This enabled me to calculate vegetation indices that could be used to estimate whether the trial was responding as expected to different fertiliser treatments. The usefulness of these measurements led me to continue measuring our trial throughout the growing season, collecting sufficient data to construct the statistical model used in this study. Following on from this, I then developed a collaboration with Crop Health and Protection (CHAP) to see whether similar multispectral measurements from seeds could also be used to construct models to predict grain free asparagine content. Although this investigation used material from one trial only, it still gave useful data on the effectiveness of imaging for predicting grain free asparagine content.

6.2. Introduction

Since large-scale industrial manufacturing of biscuits and cakes started in the 1800s, these foods have become staple items in the food culture of many parts of the world (Manley, 2011). In 2020, participants of the UK National Diet and Nutrition Survey recorded a consumption of 20 g of biscuits and 16.75 g of buns, cakes, and pastries each day (averaged across all age groups) (Public Health England, 2021). This was reflected by the 2.96 billion GBP in UK biscuit sales in 2020 and by the estimate that 99.5 % of all households purchased biscuits in 2020 (Pladis, 2021). Consequently, there is a large market in the UK for soft milling wheat flour, with UK flour millers producing an average of 541,000 tonnes of biscuit flour and 81,000 tonnes of cake flour annually from 1991 to 2020 (AHDB, 2022). In the USA, approximately 9.82 million tonnes of soft red winter wheat production are forecast for the year 2021/2022 (Economic Research Service USDA, 2022), providing flour for a biscuit market worth approximately 11.7 billion USD in 2021 (Mordor Intelligence, 2022).

Soft milling wheats (UK Flour Millers Group 3) are the primary crop used in the baking of biscuits, cakes, breakfast cereals and fine bakery products because they have lower protein content than hard, breadmaking wheats (11-11.5 % protein content requirement for soft wheats vs. 13 % requirement for breadmaking) and have a soft endosperm texture (UK Flour Millers, 2022). Soft wheat grains are easily fractured, so exhibit less starch damage and less water absorption during milling and processing than hard wheat grains (Pasha, Anjum & Morris, 2010). Due to their lower protein requirement, Group 3 wheats do not require as much nitrogen fertilizer as Group 1 (breadmaking) or 2 (breadmaking potential) wheats (AHDB, 2022). At the time of writing, there are ten Group 3 varieties on the 2022/2023 UK winter wheat recommended list but only four Group 1 and four Group 2 varieties (AHDB, 2022). These factors may drive an increase in soft wheat cultivation, but increasing prices obtainable for breadmaking wheat, the soaring cost of nitrogen fertilizer and many other factors will also affect farmers' decision-making (Sylvester-Bradley & Clarke, 2022).

Another factor that farmers are increasingly having to manage is the acrylamide-forming potential of their wheat. Acrylamide is a 'probably carcinogenic', neurotoxic, and reproductively toxic contaminant that forms from free asparagine and reducing sugars (principally glucose, fructose and maltose) in the Maillard reaction during frying, baking, roasting, toasting and high temperature processing (Mottram et al., 2002; Stadler et al., 2002). Other amino acids also participate in the Maillard reaction, giving rise to the color and flavor

compounds that impart fried, baked, roasted and toasted products with their signature characteristics. Dietary exposure to acrylamide is considered to be a public health risk by the European Food Safety Authority (EFSA, 2015; EFSA, 2022), prompting the European Commission to introduce a series of risk management measures, most recently Commission Regulation (EU) 2017/2158 (European Commission, 2017), which came into force in 2018. Commission Regulation (EU) 2017/2158 set Benchmark Levels (described by the Commission as performance indicators) for acrylamide in different food categories. These included 50 µg/kg for soft bread, 350 µg/kg for biscuits (150 µg/kg if they are for infants), 400 µg/kg for crackers, 300 µg/kg for wheat-based breakfast cereals, 150 µg/kg for breakfast cereals made with other grains, and 40 µg/kg for cereal-based baby foods.

Based on estimates of dietary acrylamide intake from EFSA (EFSA Panel on Contaminants in the Food Chain, 2015; Raffan and Halford, 2019), soft wheat products (biscuits, crackers, breakfast cereals, crispbreads, cakes and pastries) are major contributors to dietary acrylamide intake, even more so than bread if taken together. Consequently, biscuit, breakfast cereal, crispbread, cake and pastries manufacturers must minimize the concentrations of acrylamide in their products as much as possible, and various strategies for doing so have been compiled in FoodDrinkEurope's "Acrylamide Toolbox" (FoodDrinkEurope, 2019). However, while reducing acrylamide formation, food businesses must avoid impacting flavor, aroma, texture and color, and ending up with a bland, insipid product that consumers reject. The lower the concentration of free asparagine in the grain they use as raw material, the easier it would be for them to achieve that.

A factor that makes it more difficult for food businesses to keep the acrylamide concentrations in their products consistently below the Benchmark Level is the highly variable and unpredictable concentrations of free asparagine and reducing sugars in the raw materials they use. For example, average potato chip (UK crisp) acrylamide concentrations in Europe have declined substantially since acrylamide was discovered in food in 2002 and mitigation strategies began to be introduced, with European Snacks Association data showing a reduction of 54 % between 2002 and 2019 (Powers et al., 2021). Nevertheless, in the three-year period from 2017 to 2019, 7.75 % of potato chip samples still failed the 750 µg/kg Benchmark Level, and seasonal and geographical factors exacerbated the problem, with almost 18 % of samples in southern Europe in January and above 10 % in every region for some of the year exceeding the Benchmark Level. Similarly, a recent study in Spain (Mesias et al., 2019) found that 15 % of breakfast cereals contained acrylamide above the Benchmark Level. It is, therefore, alarming

that the European Commission is considering replacing Benchmark Levels with Maximum Levels (i.e., concentrations above which it would be illegal to sell a product), and look likely to set Maximum Levels at or below the current Benchmark Levels (European Commission, 2022).

For soft milling wheat products, it is the environmental impact on free asparagine concentration in the grain that poses the largest risk for acrylamide formation, as free asparagine is the key determinant of acrylamide formation in wheat products (Raffan & Halford, 2019). Unfortunately, free asparagine is difficult and expensive to measure, usually requiring HPLC or GC/LC-MS methods for quantification (Raffan & Halford, 2019). Enzymatic methods for free asparagine detection remove some of this complexity and cost (Lecart et al., 2018), but they still require multi-step sample preparation. Measurement of grain protein content can be achieved rapidly and non-destructively using near infra-red spectroscopy (NIRS) (Caporaso, Whitworth & Fisk, 2018), but similar attempts to use NIRS to measure free asparagine in wheat grain have found low predictive ability (Rapp et al., 2018).

Both abiotic and biotic stressors are known to increase free asparagine in the grain (Oddy et al., 2020), so certain crop-management strategies are included in the compulsory mitigation measures set out in Commission Regulation (EU) 2017/2158. These include avoiding excessive nitrogen (N) application while ensuring adequate sulphur (S) supply. However, there is still uncertainty about the optimal levels of N and S *per se* that should be applied and the effect of the N:S ratio. Additionally, the impact of other minerals (phosphorus and potassium) on free asparagine concentrations is not known. Consequently, this study aimed to investigate these questions, encompassing not only the effects of fertilization on free asparagine concentration, but also the impacts on biscuit quality and acrylamide concentration after baking. We also investigated whether free asparagine could be predicted from multispectral measurements of plants growing in the field and of seeds.

6.3. Materials and methods

6.3.1. Screening and selection of soft wheat varieties

DNA was extracted from a selection of soft wheat varieties and screened for the presence of the *ASN-B2* gene (TraesLDM3B03G01566640 in variety Landmark; Ensembl, 2021) as described previously (Oddy et al., 2021). Varieties lacking *ASN-B2* were then used in this

study, comprising Arkeos (2010, Limagrain), Barrel (2014, KWS), Basset (2015, KWS), Claire (1997, Limagrain), Croft (2012, KWS), Elicit (2017, Elsoms Wheat), Firefly (2017, KWS), Horatio (2011, Limagrain), Invicta (2008, Limagrain), Leeds (2011, KWS), Myriad (2011, Limagrain), and Zulu (2012, Limagrain). Data on variety registration dates and breeding companies were obtained from the EU plant variety database (European Commission, 2022) and UK national lists (APHA, 2022). Claire was used as a negative control and Cadenza as a positive control when screening for the presence of *ASN-B2* due to the availability of these genomes in Ensembl Plants (Howe et al., 2021). The results of this screen are displayed in Supplementary Figure 6.1.

6.3.2. Field trials

Field trials were undertaken at two different locations across the Rothamsted Research experimental farm site at Woburn: Stackyard (51° 59' 53.3832" N 0° 37' 1.3008" W) in 2019/2020 and Butt Clong (52° 0' 43.7184" N 0° 35' 45.5388" W) in 2020/2021 (Supplementary Figure 6.2). Key dates for these trials are given in Supplementary Table 6.1. The trial at Stackyard was undertaken using treatments A to K listed in Table 6.1, whereas the trial at Butt Clong used treatments A to L. The size of experimental plots in each trial was 9 × 1.8 m. Stackyard has a loamy sand to sandy loam soil, whereas Butt Clong has a sand to loamy sand soil (Watts, 2015). Unfortunately, detailed nutrient analysis of these soils prior to the trials were performed is not available. See Supplementary Figure 6.3 for the layout of each trial.

Nitrogen (N) and sulphur (S) were applied as DoubleTop (CF fertilizers) (ammonium sulphate and ammonium nitrate mixture, 27N ($30SO_3$)). Nitram (CF fertilizers) (ammonium nitrate, NH_4NO_3) was applied for sulphur deficient plots and to supplement the DoubleTop application where necessary to reach the required nitrogen treatment rates. Phosphorus was applied as Triple Super Phosphate (TSP) (Diamond Fertilizers) and potassium was applied as muriate of potash (MOP) (Diamond Fertilizers) (K_2O). See Supplementary File 6.1 for further details of fertilizer treatments (available at github.com/JosephOddy/ThesisSupplementaryFiles).

At Stackyard, varieties were drilled at a rate of 350 seeds/m^2 except for Croft (343 seeds/m^2), Invicta (376 seeds/m^2), and Leeds (289 seeds/m^2) due to differences in germination, seed damage, and seed availability. Due to a problem with drilling, plot 11 was smaller than the rest and reliable yield measurements could not be taken. Sulphur, potassium, and

phosphorus were applied at the same time as the first nitrogen split in this trial (10/03/2020). Herbicide was sprayed as a mixture of Palio (Corteva), Cogent (Intracrop), and Sprinter (Nufarm) at a rate of 0.265 kg/ha on 24/03/2020 to control weeds.

Table 6.1. Fertilizer treatments applied in this study. Application rates given in kilograms per hectare.

Treatment	Nitrogen	Sulphur	Phosphorus	Potassium
A	200	40	35	62
B	200	20	35	62
C	200	10	35	62
D	200	0	35	62
E	100	40	35	62
F	100	20	35	62
G	100	10	35	62
H	100	0	35	62
I	100	20	35	0
J	100	20	0	62
K	100	20	0	0
L	200	0	0	0

At Butt Clong, all varieties were drilled at a rate of 350 seeds/m^2. Only Nitram was applied during the first split (23/02/2021), with the other fertilizers being applied on 11/04/2021. Pesticide was sprayed as a mixture of Samurai (Bayer CropScience) (3 L/ha) and Buffalo Elite (Intracrop) (1 L/ha) on 24/06/2021. Plots 242 and 295 were mixed during harvest, preventing further analysis of these plots, but this did not affect yield measurements.

Weather measurement data were retrieved from the Rothamsted electronic archive resource (Rothamsted Research, 2022), which contains daily data from a weather station at the Woburn experimental field site. Daily temperature, rainfall, and solar radiation measurements over the periods that both trials were grown are shown in Supplementary Figure 6.4.

6.3.3. Phenotyping

After harvest, a sub-sample of each plot was weighed to calculate fresh weight. This sub-sample was then oven-dried at 80 °C for 24 hours and then re-weighed to measure dry weight. The percentage reduction in weight from lost moisture content was then used to adjust yield estimates taken at harvest on the combine and to calculate grain yield at 85 % dry matter. For long-term storage, grain was oven-dried to reduce moisture content to between 8 and 10 %. Thousand grain weight (TGW) measurements were subsequently taken by counting 500 seeds using a seed counter (Elmor model C1, Switzerland), drying overnight at 80 °C, and then weighing. This was repeated twice to give TGW measurements.

For samples from the first trial at Stackyard, approximately 80 g from each plot were milled to fine wholemeal flour using a Retsch 400 ultra-centrifugal mill (Retsch GmbH, Germany). Samples of grain from Butt Clong were milled to wholemeal flour in a coffee grinder. Flour moisture content was determined using a Minispec nuclear magnetic resonance (NMR) analyzer (Minispec Mq10, Bruker Inc., Germany). Following determination of moisture content, Hagberg Falling Number was measured using an FN 1000 (Perten, Sweden). Free asparagine analysis was performed on wholemeal flour samples by HPLC as described previously (Raffan et al., 2021) by Curtis Analytics (Sandwich, UK). The white flour used in biscuit baking was also measured for amino acids according to the same protocol.

Multispectral measurements were taken in the field at Butt Clong using a Tec5 HandySpec Field spectrometer (Oberursel, Germany) as described previously (Holman et al., 2019). Measurements were taken for all 432 plots on six different dates from the 17th May 2021 to the 6th August 2021. Reflectance values were obtained for 65 wavelengths at 10 nm intervals between 360 nm and 1000 nm. $NDVI_{680}$ (normalized difference vegetation index) and PSRI (plant senescence reflectance index) (Merzlyak et al., 1999) were calculated as previously described from wavelengths as shown below:

1. $\mathrm{NDVI} = \frac{780-680}{780+680}$

2. $\mathrm{PSRI} = \frac{680-500}{750}$

Grain samples from 72 plots at Butt Clong (all three replicates of all twelve varieties for the treatments N100 S10 +P +K and N200 S0 -P -K) were further analyzed using the Videometer SeedLab system (Videometer, DK) available in the Crop Health and Protection

(UK)'s Digital Phenotyping Laboratory. This automated grain/seed imaging system can be used to determine reflectance values for 19 wavelengths ranging from 365nm to 970nm and fluorescence values by the optional use of 4 long-pass filters (400nm, 500nm, 600nm and 700nm cut-offs). In total, 70 features were calculated from the image data; 19 reflectance values, 31 fluorescence bands, and 20 morphological and colour-based features (see Supplementary File 6.2). Seed and chaff were separated from background pixels and one another using custom classifiers developed using the Videometer SeedLab system. Only data from pixels classified as seeds were used in further analysis. See Supplementary File 6.2 for the data obtained from this analysis.

6.3.4. Baking tests and acrylamide analyses

Grain samples of variety Basset from the second trial at Butt Clong were milled to white flour using a Bühler mill by Campden BRI (Chipping Campden, UK), with an average extraction rate of 75.93 % ± 0.42 % (95 % confidence intervals given by ± symbol). This flour was used to bake biscuits according to a modified AACC 10-53.01 protocol (Baking Quality of Cookie Flour: Macro Wire-Cut Formulation). Flour moisture content was measured using a HG63 Halogen Moisture Analyser (Mettler Toledo) to adjust water and flour volumes according to flour moisture content. Mean flour moisture content was 13.51 % ± 0.23 %.

To form the creamed mass, non-fat dry milk, salt, sodium bicarbonate, sugar, and palm oil were mixed in a Hobart N50 mixer for 3 minutes at speed 1 (136 rpm), stopping and scraping the contents every minute. The creamed mass was subsequently mixed with a solution of ammonium bicarbonate and high-fructose corn syrup (42 %) in distilled water at speed 1 for 1 minute and at speed 2 (281 rpm) for a further minute. Finally, the flour was mixed in to form the dough at speed 1 for 2 minutes, stopping and scraping contents every 30 s.

Dough was then rolled and cut into four × 5 cm portions on a single aluminum tray. Dough water activity was measured using a 4TE water activity meter (Aqualab), with mean water activity of 0.80 ± 0.01, and dough weight was measured before baking, with a mean weight of 101.0 g ± 1.2 g. Biscuits were baked for 11 minutes in a five-chamber Polin Elettrodrago oven at 205°C, left to cool for 5 minutes on the tray outside the oven, followed by a 45-minute cooling period on a wire rack before storing in air-tight containers. This protocol was repeated twice for each flour sample.

Biscuit diameter, stack height, and weight measurements were taken as the mean of all four biscuits from each half-batch. For color analysis, images of both the top and bottom of the biscuits were taken inside an LED light box (SAMTIAN) with a color temperature of 5500 Kelvin and a FinePix S8000fd digital camera (Fujifilm). Images were captured at a shutter speed of 1/250 seconds, an aperture size of f/4, and an ISO of 100. Biscuit pixel area was segmented from background pixels using the Simple Interactive Object Extraction plugin in Fiji. RGB images were then converted to Lab Stack images and mean values for CIELAB color space parameters (L*, a*, and b*) were taken from the total area of all biscuits in an image. One biscuit was taken from each half-batch for acrylamide analysis at Reading Scientific Services Ltd (UK), with one technical replicate being taken from each sample.

6.3.5. Statistical design and analyses

Both field trials were designed as split-plot designs with 3 replicate blocks, with nutrient treatments applied to main plots (each comprising a linear array of 12 sub-plots) and varieties applied to sub-plots. Given the overall size of each experiment and the potential impact of farm operations, the allocation of varieties followed an incomplete (first trial) 11-by-12 Latin Square design or complete (second trial) 12-by-12 Latin Square design for each of the 3 sets of 11 or 12 main plots arranged down the length of the trial: this was to ensure that varieties were spread evenly across the width of each trial, and therefore reduce any bias due to spatial variability (Supplementary Figure 6.3). However, this additional blocking structure was not incorporated into the analysis model, with sub-plots just assumed to be nested within main plots. For free asparagine, three technical replicates were collected from each sub-plot. A row of discard plots was incorporated into the design of the second trial to account for where an old hedgerow used to be in the field. Both field trials were designed using GenStat (VSN International, 2022).

Grain free asparagine content was $\log_e$ transformed to account for non-normality and improve heteroscedasticity because it was positively skewed. Analysis of Variance (ANOVA) was used to investigate the effects of different experimental factors whilst accounting for the structure of the trials. For visualization of results, least significant differences were calculated at 5% to plot alongside means from each model. Graphs were plotted in R (R Core Team, 2021) with data manipulation using package “gdata” (Warnes et al., 2014) and in Genstat. Details of each model (treatment structure, blocking structure, and full ANOVA tables) are available in

Supplementary Files 6.3 and 6.4, along with a description of the data filtering used in each analysis. Data used in these analyses are available in Supplementary File 6.5. Some of the analyses considered data from just a subset of the varieties whereas other analyses considered the combined data from both trials. For the combined analysis, we included terms to test for the consistency of treatments between the two trials (the trial-by-treatment interactions).

For analysis of multispectral field data via Partial Least Squares Regression (PLSR), all wavelength measurements across all timepoints were combined (including NDVI and PDRI measurements) to form the predictor variables. Log_e transformed grain free asparagine content and non-transformed yield measurements were used as the response variables. Based on the mean square error and R-squared plots investigating the optimal number of components to include for each trait, three components were retained for the yield PLSR model and ten components were retained for the asparagine PLSR model. Five-fold cross validation repeated 1000 times was used to test each model and collect R-squared estimates. PLSR and plotting of multispectral data was performed using python and python packages NumPy (Harris et al., 2020), pandas (Reback et al., 2020), Scikit-learn (Pedregosa et al., 2011), and plotnine. Data used for analysis are available in Supplementary File 6.6.

For analysis of data obtained from the Videometer SeedLab, measurements were obtained for a minimum of 200 seeds for each sample, and the mean calculated for each variable to obtain a single measurement for each variable from each sample. Principal component analysis and linear discriminant analysis were performed and visualized in R with the packages factoextra (Kassambara & Mundt, 2020), MASS (Venables & Ripley, 2013), ggplot2 (Wickham, 2016) and cowplot (Wilke, 2020). Correlation matrices were used for both principal component analyses performed in this study using the function "prcomp" and option "scale = TRUE" to account for different scales of measurement between variables. Gaussian naïve Bayes classification was performed and visualized using python and the same packages as described for PLSR above. Five-fold cross validation repeated 1000 times was also used to test the balanced accuracy of this model.

Analysis and visualization of biscuit data was performed in R with the packages factoextra (Kassambara & Mundt, 2020), ggplot2 (Wickham, 2016) and cowplot (Wilke, 2020). Hue angle (a color appearance parameter) was calculated using the below formula:

$$h° = \frac{\left(\tan^{-1}\frac{b*}{a*}\right) \times 360}{2\pi}$$

6.4. Results

6.4.1. Impact of environment, fertilizers, and variety on the free asparagine content of wheat grain

In order to analyze the impact of environment, fertilizer treatment, variety, and the interaction between these factors on the free asparagine content of wholemeal flour, we constructed ANOVA models investigating the overall impact of treatment (Table 6.2), the N:S ratio (Table 6.3), and the application of P and/or K (Table 6.4). Full details of the data used for each analysis and the modelling terms are provided in Supplementary Files 6.2 and 6.3, in addition to an analysis of nitrogen and sulphur as interacting terms (Supplementary Table 6.2, Supplementary File 6.3) and our analyses of yield. These analyses showed that treatment significantly impacted the free asparagine content of wholemeal flour across both trials (Table 6.2, Figure 6.1A, Supplementary Figure 6.6C/D) and that this was principally due to the N:S ratio (Table 6.3, Figure 6.1C, Supplementary Figure 6.7A), whereas potassium and phosphorus did not significantly impact wholemeal flour free asparagine content in either trial (Table 6.4, Figure 6.1D/E, Supplementary Figure 6.7B/C). Variety did significantly impact the free asparagine content of wholemeal flour across both trials; however, the differences between varieties were not as great as those between trials or treatments (Figure 6.1B, Supplementary Figure 6.6A/B). Non-transformed free asparagine data are shown in the form of heatmaps in Supplementary Figure 6.5.

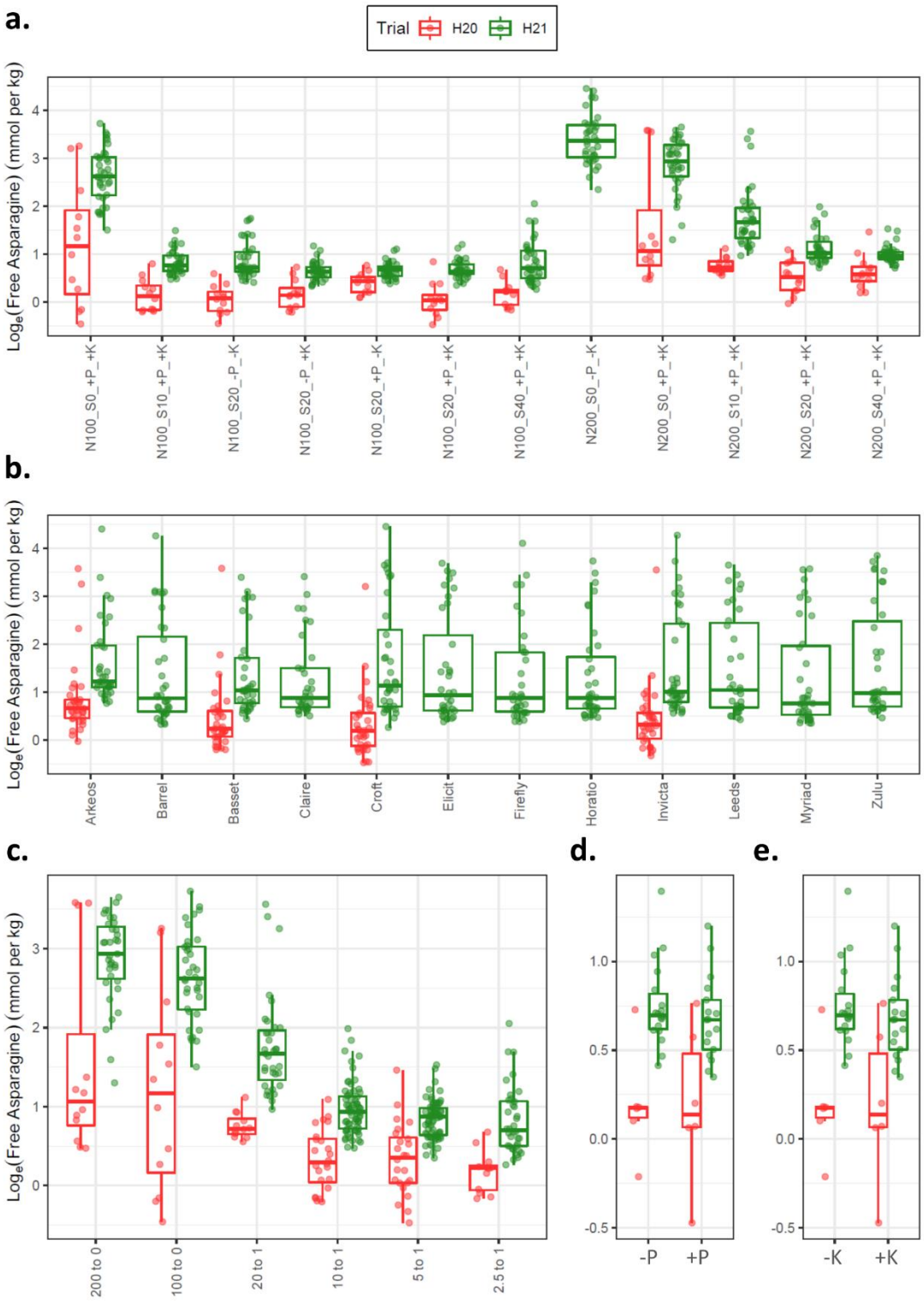


Figure 6.1. Free asparagine measurements in wholemeal flour from both field trials. **a.** Free asparagine measurements separated by agronomic treatment. **b.** Free asparagine measurements separated by variety. **c.** Free asparagine measurements separated by nitrogen to sulphur ratio. **d.** Free asparagine measurements separated by phosphorus treatment. **e.** Free asparagine measurements separated by potassium treatment. Boxes show first and third quartiles alongside

the median. Whiskers extend to the largest data points within 1.5 times interquartile range. H20 (2019 – 2020 trial), H21 (2020 – 2021 trial).

Table 6.2. Significance values (F probabilities) of terms in ANOVA models for analysis of Log_e transformed free asparagine content in wholemeal flour. H20 (2019 – 2020 trial), H21 (2020 – 2021 trial).

Source of variation	H20	H21	Both (nested)	Both (full)
Treatment	0.030	<.001	<.001	<.001
Variety	<.001	<.001	<.001	<.001
Treatment*Variety	0.183	0.005	0.004	0.119
Trial	NA	NA	NA	0.006
Trial*Treatment	NA	NA	NA	0.524
Trial*Variety	NA	NA	NA	<.001
Trial*Treatment*Variety	NA	NA	NA	0.056

Table 6.3. Significance values (F probabilities) of terms in N:S ratio ANOVA models for analysis of Log_e transformed free asparagine content in wholemeal flour. H20 (2019 – 2020 trial), H21 (2020 – 2021 trial).

Source of variation	H20	H21	Both
N:S ratio	0.050	<.001	<.001
Variety	<.001	<.001	<.001
N:S ratio*Variety	0.024	<.001	0.034
Trial	NA	NA	0.012
Trial*N:S ratio	NA	NA	0.471
Trial*Variety	NA	NA	<.001
Trial*N:S ratio*Variety	NA	NA	0.009

Table 6.4. Significance values (F probabilities) of terms in phosphorus/potassium ANOVA models for analysis of Log_e transformed free asparagine content in wholemeal flour. H20 (2019 – 2020 trial), H21 (2020 – 2021 trial).

Source of variation	H20	H21	Both
Phosphorus	0.353	0.413	0.700
Potassium	0.321	0.353	0.195
Phosphorus*Potassium	0.087	0.284	0.381
Variety	<.001	<.001	<.001
Phosphorus*Variety	0.898	0.295	0.400
Potassium*Variety	0.498	0.723	0.572
Phosphorus*Potassium*Variety	0.540	0.688	0.335
Trial	NA	NA	<.001
Trial*Phosphorus	NA	NA	0.244
Trial*Potassium	NA	NA	0.719
Trial*Phosphorus*Potassium	NA	NA	0.035
Trial*Variety	NA	NA	0.094
Trial*Phosphorus*Variety	NA	NA	0.583
Trial*Potassium*Variety	NA	NA	0.810
Trial*Phosphorus*Potassium*Variety	NA	NA	0.872

There was a significant interaction between variety and trial (Tables 6.2 and 6.3), indicating that the free asparagine content in wholemeal flour was not consistent for the varieties across environments. For example, Croft grain had the lowest mean free asparagine content in the 2020 harvest (H20), but the highest in the 2021 harvest (H21) (Figure 6.1B, Supplementary Figure 6.6A/B). In contrast, there was no interaction between the N:S ratio and trial (Table 6.3), suggesting that the N:S ratio had the same effect across environments. This can be seen in Figure 6.1C, where free asparagine content of wholemeal flour decreases in both environments as the N:S ratio decreases from 200:0 to 10:1, and then remains fairly constant at lower ratios.

6.4.2. Predicting free asparagine content from plant and grain measurements

In trial H21, we collected multispectral measurements in the field for all 432 plots at six different timepoints until harvest (Figure 6.2A). Normalized difference vegetation index (NDVI) values calculated from these data showed differences between treatments, with plots lacking sulphur generally having lower NDVI values (Supplementary Figure 6.9). In order to test whether these data could be used to predict the free asparagine content of wholemeal flour, we constructed partial least squares regression (PLSR) models using data from all six timepoints and tested the accuracy of this method for predicting free asparagine content and yield (Figure 6.2B). The model for free asparagine had an average R^2 value of 71.26 %, whereas the model for yield had an average R^2 value of 81.75 %. We also tested a classification model using Gaussian naïve Bayes to see whether these measurements could distinguish between sulphur-deficient (S0) and sulphur-fed (S10, S20, or S40) plots (Supplementary Figure 6.10). This model had a mean balanced accuracy of 0.76 (improvement of 0.26 over random classification).

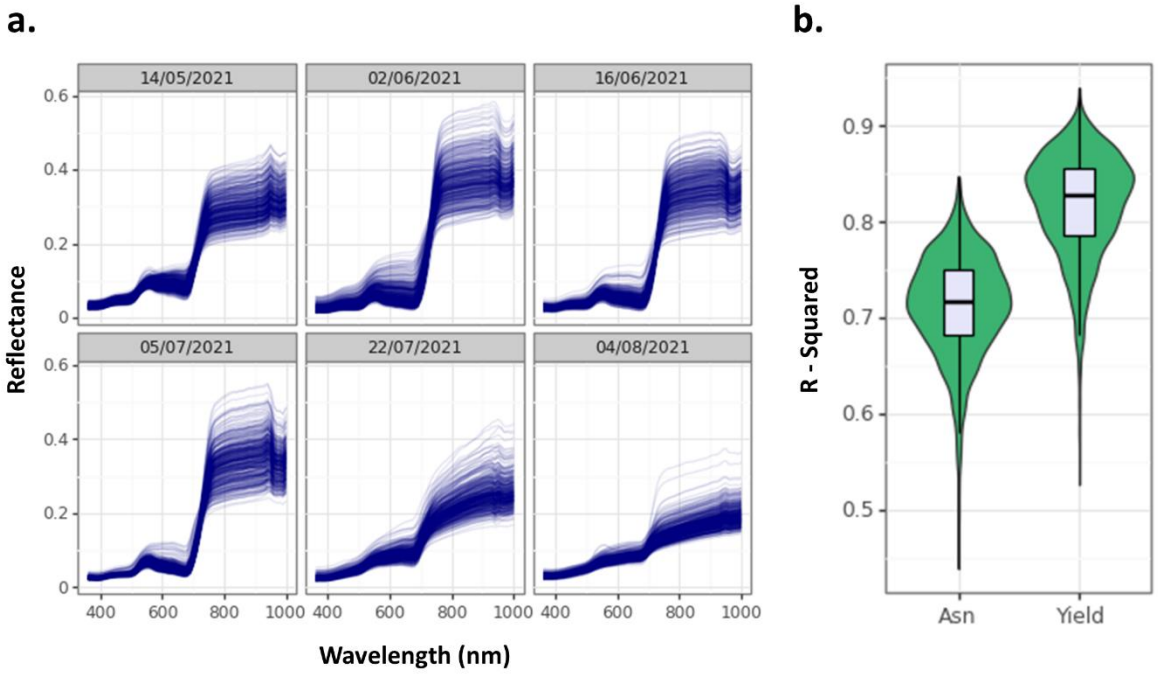


Figure 6.2. A. Multispectral measurements taken from field plots for trial H21. **B**. R^2 values from Partial Least Squares Regression analysis of the data for free asparagine (Asn) and yield.

Following on from this experiment, we investigated whether multispectral, fluorescence, and morphology measurements from the seed itself could be used to distinguish high asparagine

seeds (> 10 mmol per kg, from treatment N200 S0 -P -K) from low asparagine seeds (< 5 mmol per kg, from treatment N100 S10 +P +K). Seeds from different agronomic treatments did tend to separate out along the second principal component from our PCA analysis (Figure 6.3A), indicating that a classification model may be effective. We then performed a linear discriminant analysis which showed good separation of treatments (Figure 6.3B) and tested the accuracy of a Gaussian naïve Bayes classifier using 1000 repeated five-fold cross validation (Figure 6.3C). This model was able to classify samples to the correct agronomic treatment group with a balanced accuracy of 0.86 (improvement of 0.36 over random classification), and was able to classify samples to the correct variety group with a balanced accuracy of 0.42 (improvement of 0.34 over random classification).

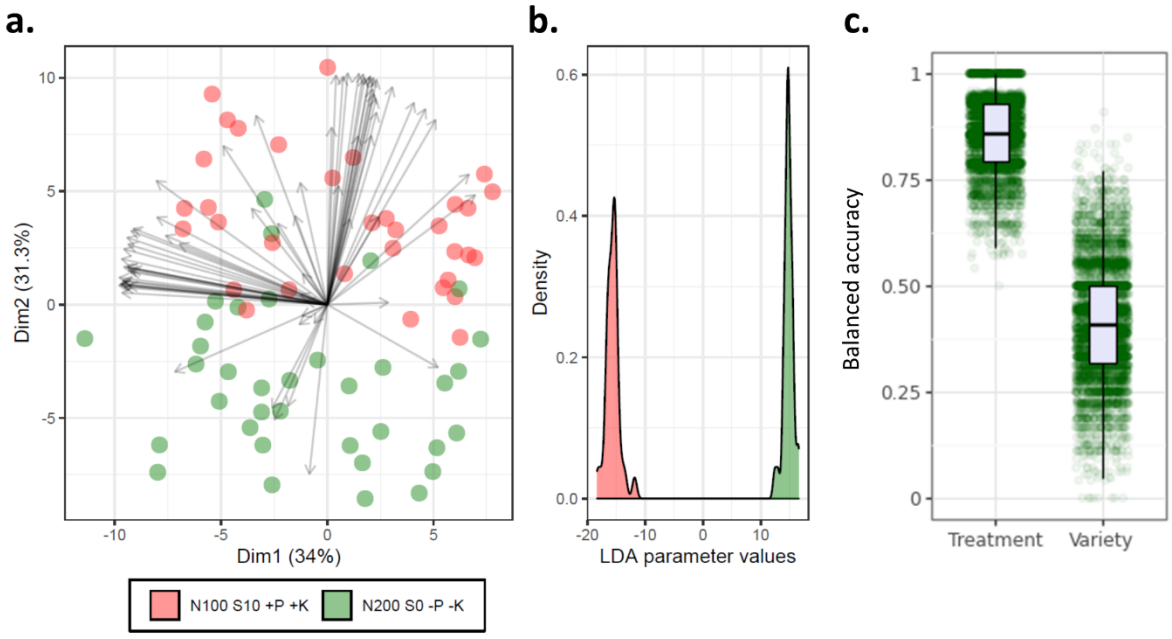


Figure 6.3. Measurements of selected seeds from trial H21. **A.** Principal component analysis of all measured traits from Videometer SeedLab and wholemeal flour free asparagine content. **B.** Linear discriminant analysis of seeds separated by agronomic treatment. **C.** Balanced accuracy scores from gaussian naïve Bayes classification predicting sample treatment and variety.

6.4.3. Impact of different agronomic treatments on biscuit quality and acrylamide formation

In order to assess the impact of different agronomic treatments on end products, we baked biscuits from selected flours in the H21 trial (Figure 6.4). We chose to bake biscuits using flour from variety Basset in three different agronomic treatments (N100 S10 +P +K, N100 S0 +P +K, and N200 S0 -P -K) because of the range in wholemeal flour free asparagine content shown by these samples. Biscuits baked from these different agronomic treatments tended to separate along the first principal component in the principal component analysis (Figure 6.4B), largely due to differences in acrylamide content, white flour free asparagine content, and color. The groups did not separate out along the second principal component, which mostly highlighted differences in moisture content and diameter.

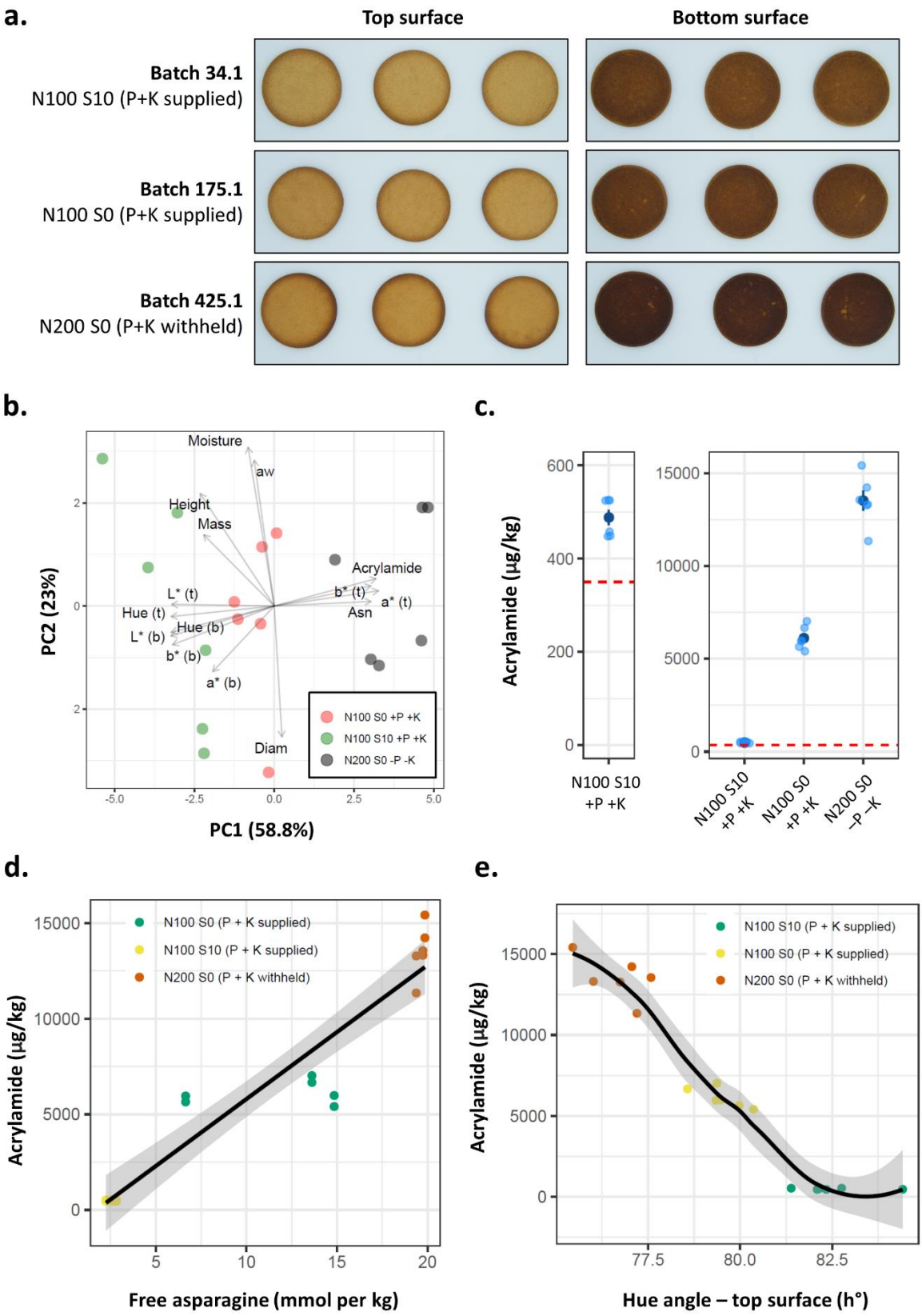


Figure 6.4. Acrylamide measurements in biscuits and associations with other variables. **A.** Representative images of biscuits baked in this study. **B.** PCA of all measurements taken from

biscuits. **C.** Acrylamide concentration of biscuits produced from grain from different agronomic treatments. The EU Benchmark Value of 350 μg/kg is given by the dashed red line. Dark blue points and bars show mean and standard error of the means. This plot is split to better visualize the lowest acrylamide concentrations of the S10 treatment. **D.** Association between the free asparagine content of white flour from plots selected for baking and biscuit acrylamide content. Linear model line fitted. **E.** Association between biscuit color (as measured by hue angle of the top surface) and acrylamide content. Lowess smoothing curve fitted. Grey shaded region indicates the standard error of the means.

Acrylamide content varied widely between the three agronomic groups, with all samples exceeding the EU Benchmark Level of 350 μg/kg for biscuits (Figure 6.4C). Whereas the biscuits from the N100 S10 +P +K treatment group contained 488 μg/kg mean acrylamide, the N100 S0 +P +K and N200 S0 -P -K treatment groups contained 6114 μg/kg and 13,523 μg/kg mean acrylamide, respectively. These differences were significant between all treatment groups (one-way ANOVA and Tukey tests, $p < 0.001$). Biscuit acrylamide content did correlate with free asparagine content of white flour (Figure 6.4D) and top surface hue angle (Figure 6.4E), with Kendall correlations of 0.79 and – 0.79 and R^2 values from linear models of 89 % and 91 %, respectively ($p < 0.001$).

6.5. Discussion

6.5.1. Optimizing fertilizer application to reduce free asparagine content of wheat grain

Our results indicate that the application of nitrogen (N) and sulphur (S) at a ratio of 10:1 (kg/ha) is sufficient to prevent large increases in asparagine accumulation in the grain of wheat grown on loamy sand/sandy loam soils. Previous studies investigating a range of S application rates have recommended that 20 kg/ha S should be applied (equivalent to 50 kg/ha SO_3) (Curtis et al., 2014). Our findings agree that at higher application rates of N (200 kg/ha), 20 kg/ha S application is required to minimize asparagine accumulation. However, at lower rates of N application (100 kg/ha), 10 kg/ha S application is sufficient. Application rates greater than a 10:1 N to S ratio did not contribute to any meaningful further reduction in free asparagine content in wholemeal flour, so application above this rate should be carefully considered due to the potential negative effects of S over-application on the environment (Hinckley et al.,

2020). The average field rate for N application on winter wheat in the UK in 2021 was 188 kg/ha, whilst the average S application rate was 20.8 kg/ha (equivalent to 52 kg/ha SO_3) (Department for Environment, Food and Rural Affairs, 2022), equivalent to a mean N:S ratio of approximately 9:1. However, while N was applied to 99 % of the winter wheat area, S was applied to only 73 %, meaning that 27 % of the winter wheat area in 2021 received high levels of N without any S. Different soil types will require different application rates of N and S due to differences in endogenous nutrient content, so further testing on sites with different soil profiles should be undertaken, and some of the wheat not receiving S may be used for feed or bioenergy. However, this does suggest that persuading farmers who are currently not applying S to their winter wheat to do so would facilitate regulatory compliance on acrylamide for food businesses and reduce the exposure of consumers to acrylamide from wheat products. S application also improves a number of desirable traits, which we did not measure in our study (Zhao et al., 1999).

Interestingly, we found that withholding potassium and phosphorus application did not cause increases in free asparagine content in wholemeal flour from either trial when S was applied at 20 kg/ha, but we did observe an increase in the sulphur-deficient plots in the second trial when phosphorus and potassium were absent. Previous research has shown that potassium and phosphorus deficiencies cause increases in asparagine in the root, stem, and leaves of a range of plant species (see Stewart and Lahrer (1980), and Lea et al. (2007) for reviews), so we thought we might observe a similar increase in wholemeal flour regardless of S application, but this was not the case. Our study did not detect an overall effect of phosphorus and/or potassium application on yield either, suggesting that phosphorus and potassium may already have been present in sufficient concentrations in the soil at both trial sites, despite the soil type. This further emphasizes the need to test soil nutrient content and tailor fertilizer application so that only the required amount is applied (AHDB, 2022). Nevertheless, our observation that phosphorus and potassium deficiencies may cause increases in wholemeal flour free asparagine content during S deficiency warrants further study.

In addition to different N:S ratios, free asparagine content of wholemeal flour also differed between the different varieties used in this study. Varietal differences in free asparagine in wholemeal flour have been observed in many studies, but they are often much smaller than the differences associated with environmental factors (Curtis et al., 2016; Curtis et al., 2018). In this study, the differences between varieties were also much smaller than the differences between fertilizer treatments. The same pattern has been observed between

varieties that do and do not possess the *TaASN-B2* gene: while there were differences in free asparagine content in wholemeal flour between such varieties, much larger differences were again caused by S deficiency (Oddy et al., 2021). Consequently, the use of varieties that are lower in free asparagine content in the grain will only be effective if a low N:S ratio is maintained. Sulphur deficiency also impacts on desirable traits (Zhao et al., 1999), and our overall recommendation is for farmers to focus on applying nitrogen and sulphur at a ratio of 10:1 kg/ha.

6.5.2. Predicting the free asparagine content of grain using imaging technology

Burnett et al. (2021a) demonstrated for the first time that hyperspectral imaging of plants can be used to effectively predict certain metabolites produced during stress (in their study abscisic acid and proline were analyzed during drought stress) and outlined how such analyses can be performed (Burnett et al., 2021b). Similarly, we found that multispectral measurements of wheat plants growing in the field were able to predict grain free asparagine content with an average accuracy of 71% when used in our PLSR model. Much of the predictive ability of our model is likely due to the dual impact of sulphur deficiency on wheat canopy color (sulphur deficiency caused a yellowing of the canopy, shown by our NDVI and PSRI measurements) and grain free asparagine content. Few studies have investigated sulphur deficiency through multispectral imaging, as most studies of this sort have focused on nitrogen (Berger et al., 2020), but Mahajan et al. (2014) found that certain vegetation indices could predict sulphur content in wheat with an accuracy of 0.46. Our results indicate that multispectral measurements from the field could be used to accurately predict grain free asparagine content and sulphur deficiency, enabling farmers with access to such technology to monitor the acrylamide forming potential of their crop.

Our model for classification of seeds was also able to distinguish sulphur-deficient from sulphur-fed samples. Models such as these, if developed appropriately, could be useful for millers to quickly screen grain samples to determine grain quality. Classification of wheat seeds using spectroscopy can be used for traits such as protein content, Hagberg Falling Number, and pathogen damage (Caporaso et al., 2018), so prediction of free asparagine content could be integrated into such models to give an overall measurement of grain quality. However, for models on both plants and seeds to have broad applicability, they would need to be trained using samples from many more diverse environments and under diverse stressors. Many other

stressors are associated with grain asparagine accumulation (Oddy et al., 2020), and spectroscopy can be used to measure such stressors (Lowe, Harrison & French, 2017), so future experiments should investigate the accuracy of these models under more stressors.

6.5.3. Impact of different treatments on biscuit quality

Flours from different agronomic treatments differed greatly in terms of acrylamide content in this study and there was a strong correlation between asparagine and acrylamide, showing that agronomic strategies to control grain free asparagine content can effectively control the acrylamide content of biscuits. There was also a strong correlation between color (hue angle) and acrylamide content, as observed previously in biscuits (Schouten et al., 2022), indicating that biscuit color can also be used to predict acrylamide concentration. In-line color sorting could therefore be implemented on biscuit production lines to eliminate high-acrylamide products, as has been recommended for other products (FoodDrinkEurope, 2019). Acrylamide formation will differ in biscuits depending on differences in ingredients and processing, so it would be advisable to check the correlation between the free asparagine content of the flour used and the end-product acrylamide concentration to ensure that strategies to reduce grain free asparagine content will be effective to control acrylamide.

Interestingly, we found acrylamide concentrations exceeding the Benchmark Level for biscuits (350 μg/kg) even in those samples where free asparagine concentration was low (2 – 3 mmol/kg). This is likely due to the baking methodology and oven used in this study. For example, our recipe included ammonium bicarbonate and high-fructose corn syrup, the combination of which is known to greatly elevate acrylamide formation compared to sodium bicarbonate and glucose (Amrein et al., 2006). Consequently, it is important to implement baking processes that do not favor acrylamide formation even when using low asparagine flours, as unfavorable processing conditions can create products exceeding Benchmark Levels from flours which are fairly low risk.

6.6. References

AHDB. (2022). UK human and industrial cereal usage. https://ahdb.org.uk/cereals-oilseeds/uk-human-industrial-cereal-usage (accessed May 30, 2022).

AHDB. (2022). Winter wheat recommended list 2022/2023. https://ahdb.org.uk/knowledge-library/recommended-lists-for-cereals-and-oilseeds-rl#h21 (accessed June 1, 2022).

AHDB. (2022). RB209 Section 4 Arable Crops. https://ahdb.org.uk/knowledge-library/rb209-section-4-arable-crops (accessed June 1, 2022).

Amrein, T. M., Andres, L., Manzardo, G. G., & Amadò, R. (2006). Investigations on the promoting effect of ammonium hydrogencarbonate on the formation of acrylamide in model systems. *Journal of Agricultural and Food Chemistry*, *54*(26), 10253-10261.

APHA. (2022). Plant Varieties and Seeds Gazette. Special edition April 2022. https://www.gov.uk/government/publications/plant-varieties-and-seeds-gazette-2020 (accessed May 31, 2022).

Berger, K., Verrelst, J., Feret, J. B., Wang, Z., Wocher, M., Strathmann, M., ... & Hank, T. (2020). Crop nitrogen monitoring: Recent progress and principal developments in the context of imaging spectroscopy missions. *Remote Sensing of Environment*, *242*, 111758.

Burnett, A. C., Serbin, S. P., Davidson, K. J., Ely, K. S., & Rogers, A. (2021). Detection of the metabolic response to drought stress using hyperspectral reflectance. *Journal of Experimental Botany*, *72*(18), 6474-6489.

Burnett, A. C., Anderson, J., Davidson, K. J., Ely, K. S., Lamour, J., Li, Q., ... & Serbin, S. P. (2021). A best-practice guide to predicting plant traits from leaf-level hyperspectral data using partial least squares regression. *Journal of Experimental Botany*, *72*(18), 6175-6189.

Caporaso, N., Whitworth, M. B., & Fisk, I. D. (2018). Near-Infrared spectroscopy and hyperspectral imaging for non-destructive quality assessment of cereal grains. *Applied Spectroscopy Reviews*, *53*(8), 667-687.

Curtis, T., Halford, N. G., Powers, S. J., McGrath, S. P. and Zazzeroni, R. (2014). Effect of sulphur fertilisation on the acrylamide-forming potential of wheat (HGCA Project Report No. 525). *Home Grown Cereals Authority (HGCA) Stoneleigh*. https://ahdb.org.uk/effect-of-sulphur-fertilisation-on-the-acrylamide-forming-potential-of-wheat

Curtis, T. Y., Powers, S. J., & Halford, N. G. (2016). Effects of fungicide treatment on free amino acid concentration and acrylamide-forming potential in wheat. *Journal of Agricultural and Food Chemistry, 64*(51), 9689-9696.

Curtis, T. Y., Powers, S. J., Wang, R., & Halford, N. G. (2018). Effects of variety, year of cultivation and sulphur supply on the accumulation of free asparagine in the grain of commercial wheat varieties. *Food Chemistry, 239*, 304-313.

Department for Environment, Food and Rural Affairs. (2022). British Survey of Fertilizer Practice 2021. https://www.gov.uk/government/statistics/british-survey-of-fertilizer-practice-2021 (accessed 8th September 2022).

Economic Research Service USDA. (2022). Wheat Data: Yearbook Tables. https://www.ers.usda.gov/data-products/wheat-data/ (accessed 3rd August 2022).

EFSA Panel on Contaminants in the Food Chain (CONTAM). (2015). Scientific opinion on acrylamide in food. *EFSA Journal, 13*(6), 4104.

European Commission. (2006). Commission Regulation (EC) No 401/2006 of 23 February 2006 laying down the methods of sampling and analysis for the official control of the levels of mycotoxins in foodstuffs. *Official Journal of the European Union, 70*, 12-34.

European Commission. (2017). Commission regulation (EU) 2017/2158 of 20 November 2017 establishing mitigation measures and Benchmark levels for the reduction of the presence of acrylamide in food. Brussels, Belgium: European Commission.

European Commission. (2022). Food Safety: Acrylamide. https://ec.europa.eu/food/safety/chemical-safety/contaminants/catalogue/acrylamide_en (accessed June 2, 2022).

European Commission (2022). EU Plant variety database (v.3.4). https://ec.europa.eu/food/plant/plant_propagation_material/plant_variety_catalogues_databases/search/public/index.cfm?event=SearchForm&ctl_type=A (accessed May 31, 2022)

European Food Safety Authority (EFSA), Benford, D., Bignami, M., Chipman, J. K., & Ramos Bordajandi, L. (2022). Assessment of the genotoxicity of acrylamide. *EFSA Journal, 20*(5), e07293.

FoodDrinkEurope. (2019). Acrylamide toolbox 2019. Brussels, Belgium: FoodDrinkEurope.

Harris, C. R., Millman, K. J., Van Der Walt, S. J., Gommers, R., Virtanen, P., Cournapeau, D., ... & Oliphant, T. E. (2020). Array programming with NumPy. *Nature*, *585*(7825), 357-362.

Hinckley, E. L. S., Crawford, J. T., Fakhraei, H., & Driscoll, C. T. (2020). A shift in sulfur-cycle manipulation from atmospheric emissions to agricultural additions. *Nature Geoscience*, *13*(9), 597-604.

Holman, F. H., Riche, A. B., Castle, M., Wooster, M. J., & Hawkesford, M. J. (2019). Radiometric calibration of 'commercial off the shelf'cameras for UAV-based high-resolution temporal crop phenotyping of reflectance and NDVI. *Remote Sensing*, *11*(14), 1657.

Howe, K.L., Achuthan, P., Allen, J., Allen, J., Alvarez-Jarreta, J., Amode, M.R., Armean, I.M., Azov, A.G., Bennett, R., & Bhai, J. (2021). Ensembl 2021. *Nucleic Acids Research*, *49*, D884–D891

Kassambara, A., Mundt, F. (2020). factoextra: Extract and Visualize the Results of Multivariate Data Analyses. R package version 1.0.7. https://CRAN.R-project.org/package=factoextra

Lecart, B., Jacquet, N., Anseeuw, L., Renier, M., Njeumen, P., Bodson, B., ... & Richel, A. (2018). Nonconventional enzymatic method to determine free asparagine level in whole-grain wheat. *Food Chemistry*, *251*, 64-68.

Lowe, A., Harrison, N., & French, A. P. (2017). Hyperspectral image analysis techniques for the detection and classification of the early onset of plant disease and stress. *Plant Methods*, *13*(1), 1-12.

Mahajan, G. R., Sahoo, R. N., Pandey, R. N., Gupta, V. K., & Kumar, D. (2014). Using hyperspectral remote sensing techniques to monitor nitrogen, phosphorus, sulphur and potassium in wheat (*Triticum aestivum* L.). *Precision Agriculture*, *15*(5), 499-522.

Mänd, P., Hallik, L., Peñuelas, J., Nilson, T., Duce, P., Emmett, B. A., ... & Kull, O. (2010). Responses of the reflectance indices PRI and NDVI to experimental warming and drought in European shrublands along a north–south climatic gradient. *Remote Sensing of Environment*, *114*(3), 626-636.

Manley, D. (2011). Setting the scene: A history and the position of biscuits. In *Manley's Technology of Biscuits, Crackers and Cookies* (pp. 1-9). Woodhead Publishing.

Merzlyak, M. N., Gitelson, A. A., Chivkunova, O. B., & Rakitin, V. Y. (1999). Non-destructive optical detection of pigment changes during leaf senescence and fruit ripening. *Physiologia Plantarum, 106*(1), 135-141.

Mordor Intelligence. (2022). BISCUITS MARKET - GROWTH, TRENDS, COVID-19 IMPACT, AND FORECASTS (2022 - 2027). https://www.mordorintelligence.com/industry-reports/biscuits-market (accessed 3rd August 2022).

Mottram, D. S., Wedzicha, B. L., & Dodson, A. T. (2002). Acrylamide is formed in the Maillard reaction. *Nature, 419*(6906), 448-449.

Oddy, J., Raffan, S., Wilkinson, M. D., Elmore, J. S., & Halford, N. G. (2020). Stress, nutrients and genotype: understanding and managing asparagine accumulation in wheat grain. *CABI Agriculture and Bioscience, 1*(1), 1-14.

Oddy, J., Alarcón-Reverte, R., Wilkinson, M., Ravet, K., Raffan, S., Minter, A., ... & Pearce, S. (2021). Reduced free asparagine in wheat grain resulting from a natural deletion of TaASN-B2: investigating and exploiting diversity in the asparagine synthetase gene family to improve wheat quality. *BMC plant biology, 21*(1), 1-17.

Pasha, I., Anjum, F. M., & Morris, C. F. (2010). Grain hardness: a major determinant of wheat quality. *Food Science and Technology International, 16*(6), 511-522.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., ... & Duchesnay, E. (2011). Scikit-learn: Machine learning in Python. *The Journal of Machine Learning Research, 12*, 2825-2830.

Pladis. (2021). Winning with Biscuits. Annual Biscuit Review 2020. https://www.readkong.com/page/winning-with-biscuits-annual-biscuit-review-2020-pladis-1454169 (accessed May 31, 2022).

Powers, S. J., Mottram, D. S., Curtis, A., & Halford, N. G. (2021). Progress on reducing acrylamide levels in potato crisps in Europe, 2002 to 2019. *Food Additives & Contaminants: Part A, 38*(5), 782-806.

Public Health England. (2021). NDNS: Diet and physical activity – a follow-up study during COVID-19. https://www.gov.uk/government/statistics/ndns-diet-and-physical-activity-a-follow-up-study-during-covid-19 (accessed May 31, 2022).

R Core Team (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

Raffan, S., & Halford, N. G. (2019). Acrylamide in food: Progress in and prospects for genetic and agronomic solutions. *Annals of Applied Biology, 175*(3), 259-281.

Raffan, S., Sparks, C., Huttly, A., Hyde, L., Martignago, D., Mead, A., ... & Halford, N. G. (2021). Wheat with greatly reduced accumulation of free asparagine in the grain, produced by CRISPR/Cas9 editing of asparagine synthetase gene TaASN2. *Plant Biotechnology Journal, 19*(8), 1602-1613.

Rapp, M., Schwadorf, K., Leiser, W. L., Würschum, T., & Longin, C. F. H. (2018). Assessing the variation and genetic architecture of asparagine content in wheat: What can plant breeding contribute to a reduction in the acrylamide precursor? *Theoretical and Applied Genetics, 131*(11), 2427-2437.

Reback, J., McKinney, W., Van Den Bossche, J., Augspurger, T., Cloud, P., Klein, A., ... & Seabold, S. (2020). pandas-dev/pandas: Pandas 1.0. 5. *Zenodo.*

Rothamsted Research. (2022). Woburn Meteorological Data. Electronic Rothamsted Archive, Rothamsted Research. http://www.era.rothamsted.ac.uk/index.php (accessed 3rd August 2022).

Schouten, M. A., Tappi, S., Glicerina, V., Rocculi, P., Angeloni, S., Cortese, M., ... & Romani, S. (2022). Formation of acrylamide in biscuits during baking under different heat transfer conditions. *LWT, 153*, 112541.

Stadler, R. H., Blank, I., Varga, N., Robert, F., Hau, J., Guy, P. A., ... & Riediker, S. (2002). Acrylamide from Maillard reaction products. *Nature, 419*(6906), 449-450.

Sylvester-Bradley, R. & Clarke, S. (2022). Review of how best to respond to expensive fertilizer nitrogen for use in 2022. Part four: Late N for Milling Wheat. https://ahdb.org.uk/how-best-to-respond-to-costly-fertilizer-nitrogen-for-use-in-2022 (accessed June 1, 2022).

UKFM. (2022). Wheat. https://www.ukflourmillers.org/wheat (accessed May 31, 2022).

Venables, W. N., & Ripley, B. D. (2013). Modern applied statistics with S-PLUS. Springer Science & Business Media.

VSN International (2022). Genstat for Windows 22nd Edition. VSN International, Hemel Hempstead, UK.

Watts, C. (2015). Woburn Soil Maps and Legends: Woburn Soils Texture. http://www.era.rothamsted.ac.uk/eradoc/article/SoilMaps-2-2 (accessed June 6, 2022)

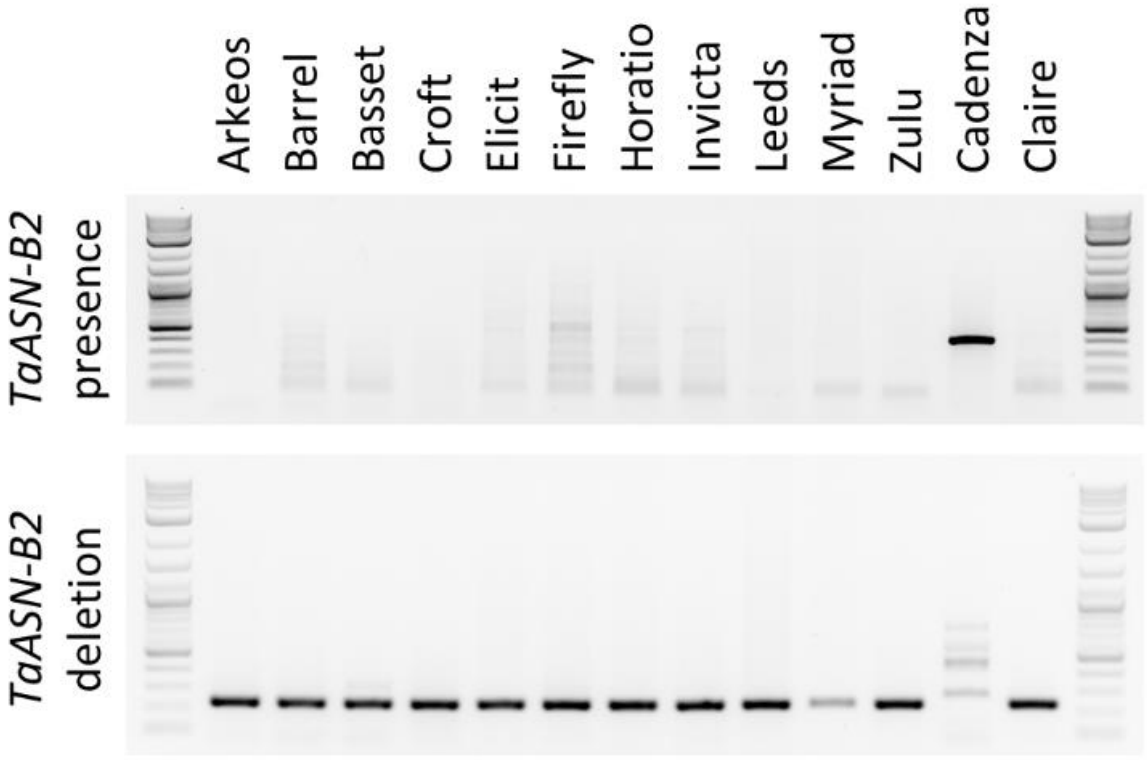
Warnes, G. R., Bolker, B., Gorjanc, G., Grothendieck, G., Korosec, A., Lumley, T., ... & Rogers, J. (2014). gdata: Various R programming tools for data manipulation. R package version, 2(3), 35.

Wickham, H. (2016). ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York.

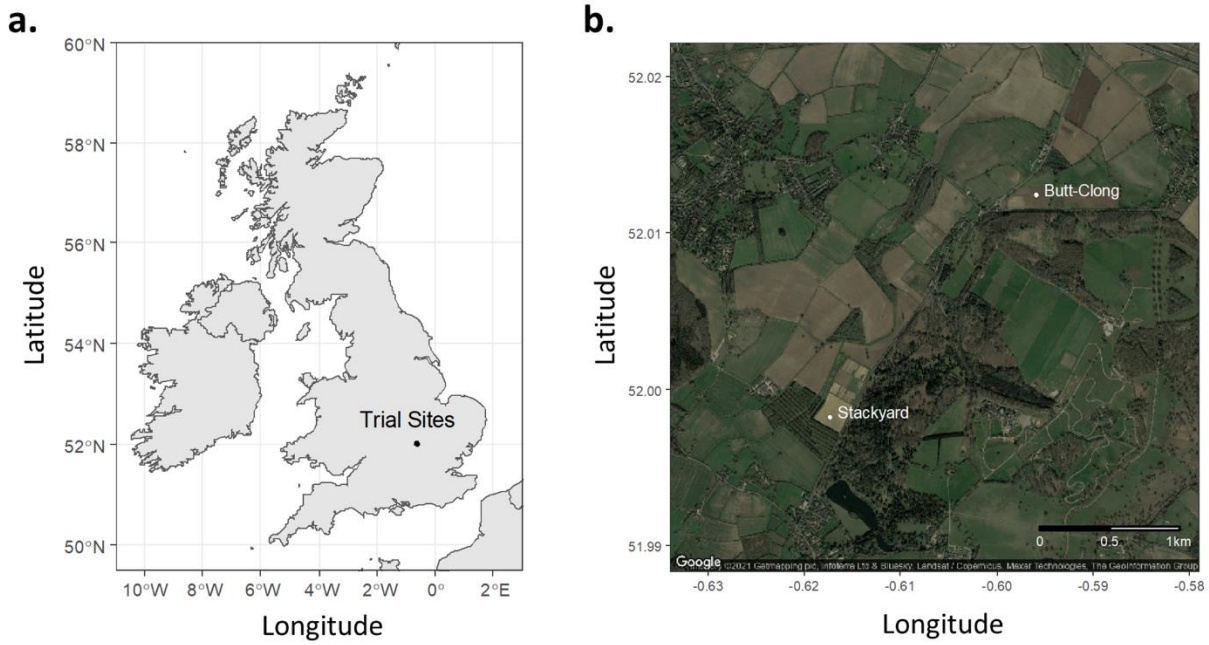
Wilke, CO. (2020). cowplot: Streamlined Plot Theme and Plot Annotations for 'ggplot2'. R package version 1.1.1. https://CRAN.R-project.org/package=cowplot

Zhao, F. J., Salmon, S. E., Withers, P. J. A., Monaghan, J. M., Evans, E. J., Shewry, P. R., & McGrath, S. P. (1999). Variation in the breadmaking quality and rheological properties of wheat in relation to sulphur nutrition under field conditions. *Journal of Cereal Science*, *30*(1), 19-31.

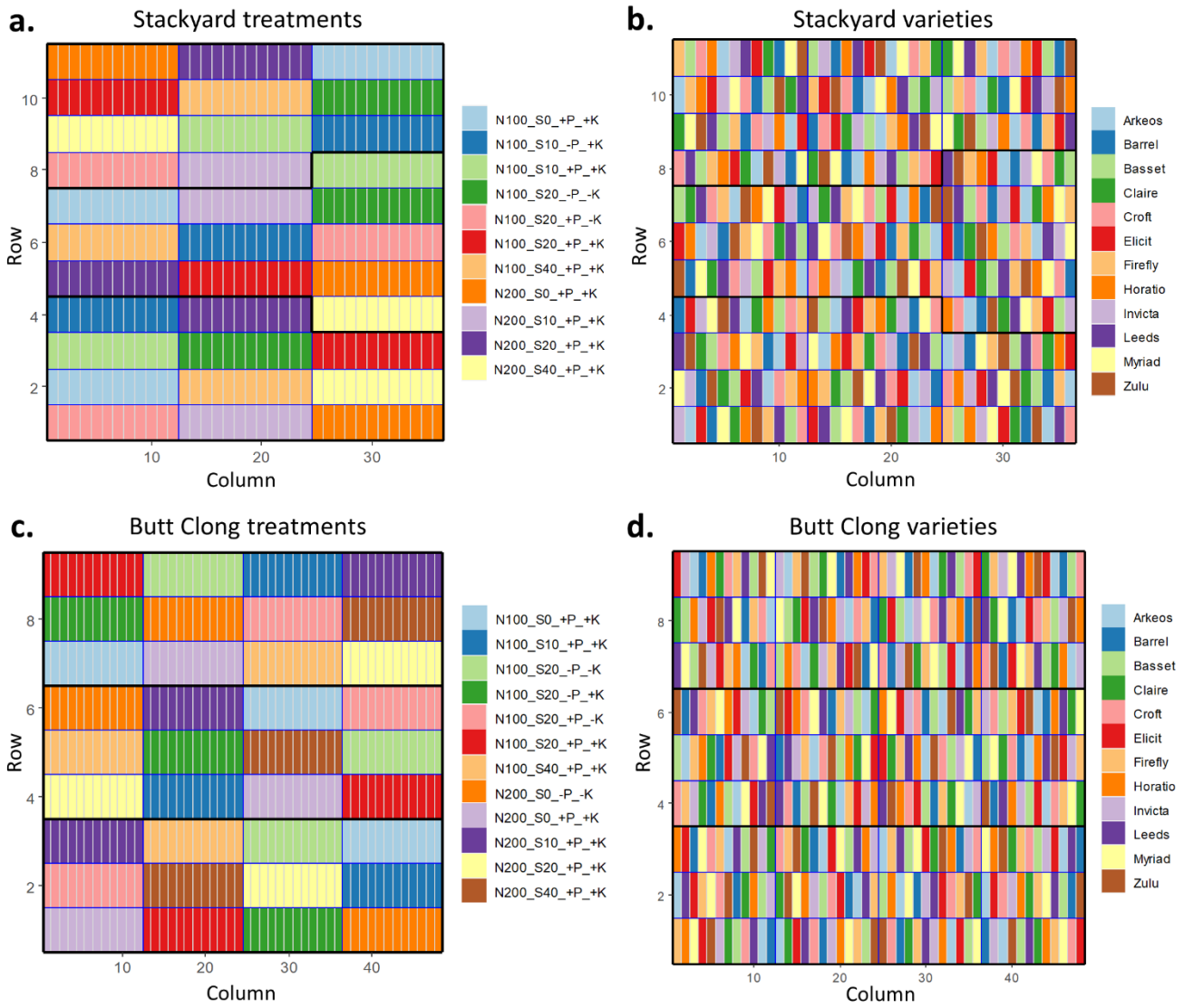
6.7. Supplementary figures and tables



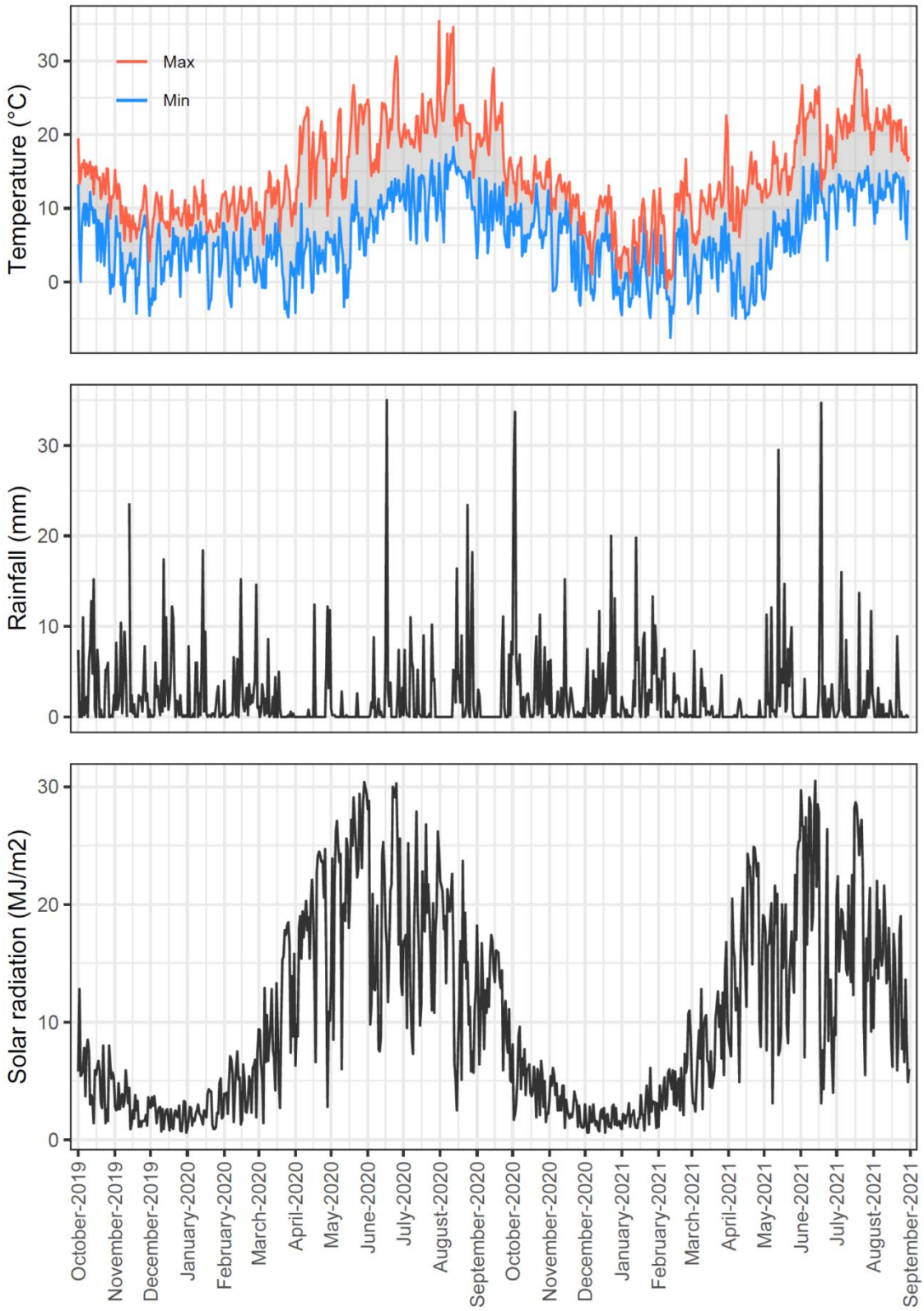
Supplementary Figure 6.1. Screening for the presence/absence of the *TaASN-B2* gene in the soft wheat varieties analysed in this study. Cadenza was used as a positive control and Claire as a negative control.



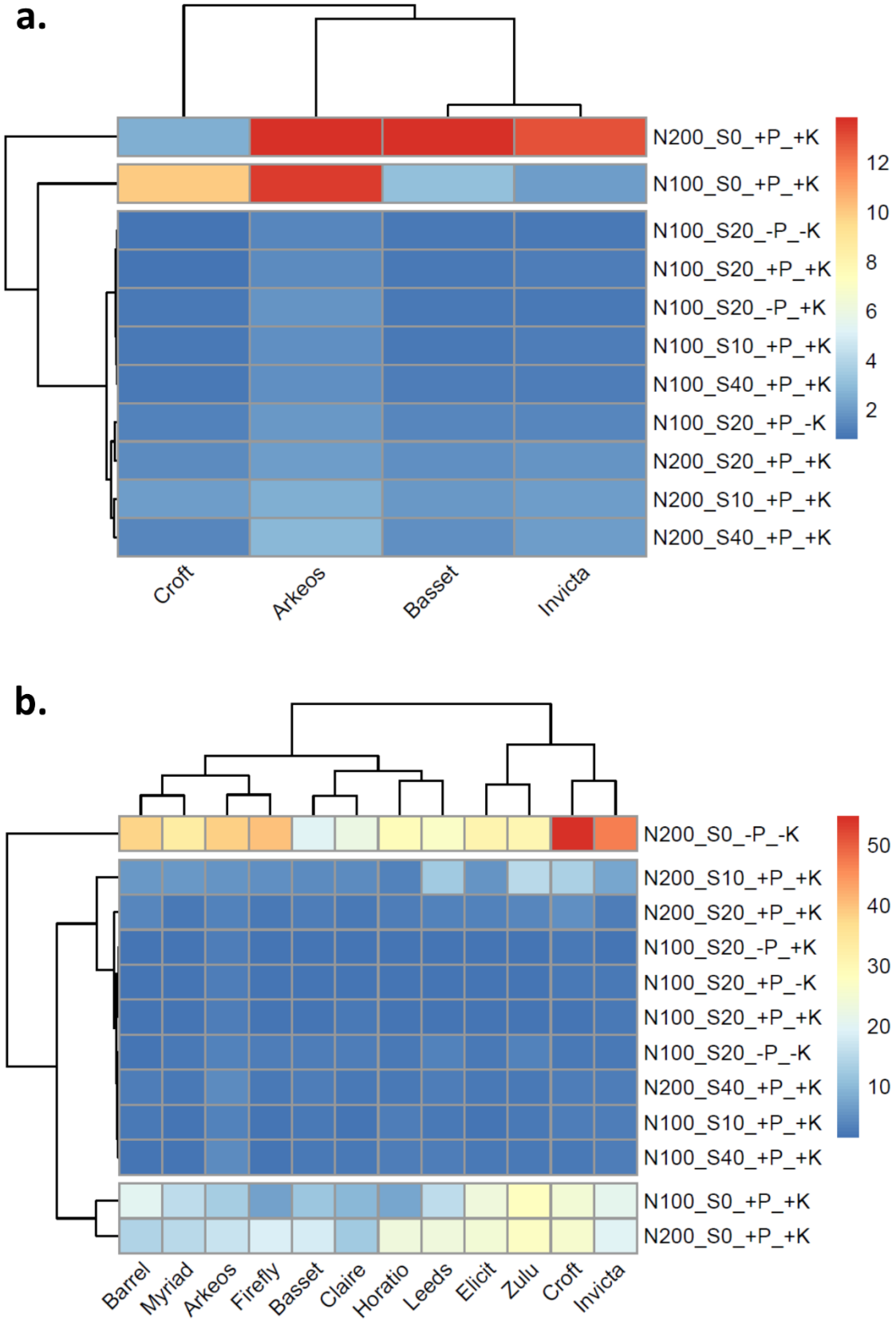
Supplementary Figure 6.2. Location of the trial sites relative to the British Isles (**a.**) and at finer scale (**b.**).



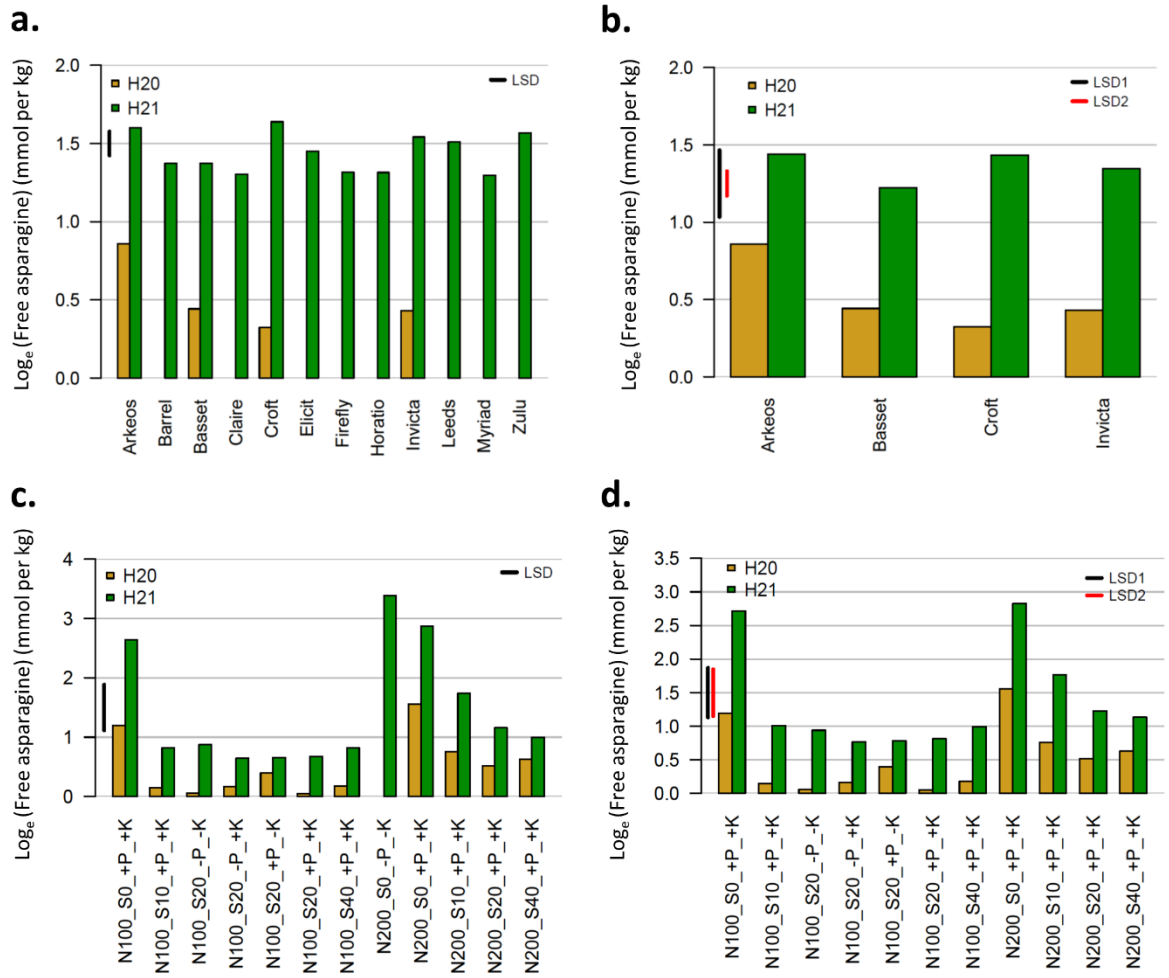
Supplementary Figure 6.3. Layout of the treatments and varieties in Stackyard in 2019 – 2020 (**a.** and **b.**) and in Butt Clong in 2020 – 2021 (**c.** and **d.**).



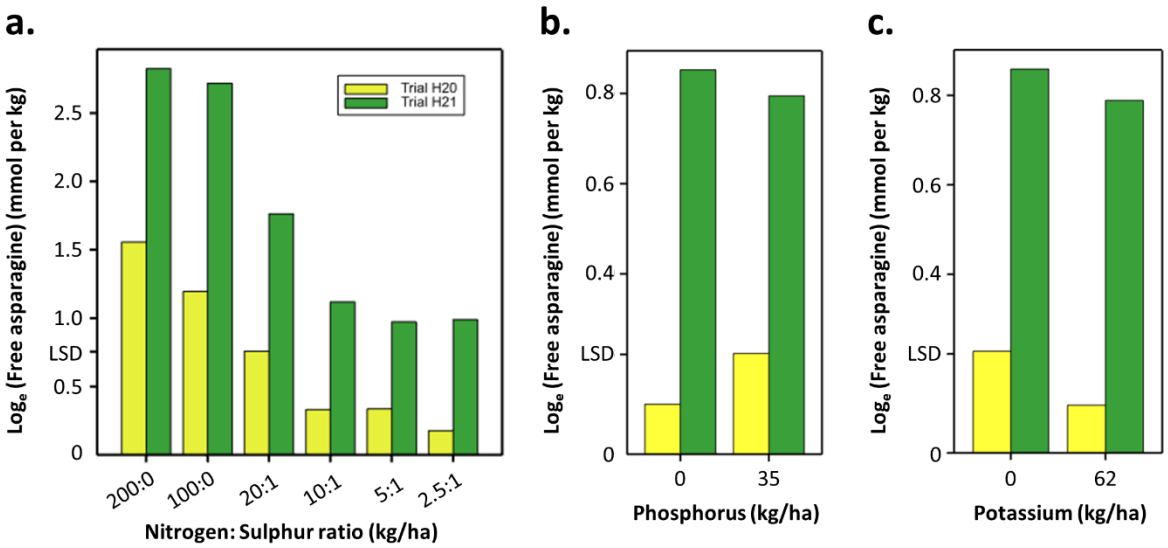
Supplementary Figure 6.4. Temperature, rainfall, and solar radiation measurements during both field trials measured at the Woburn weather station.



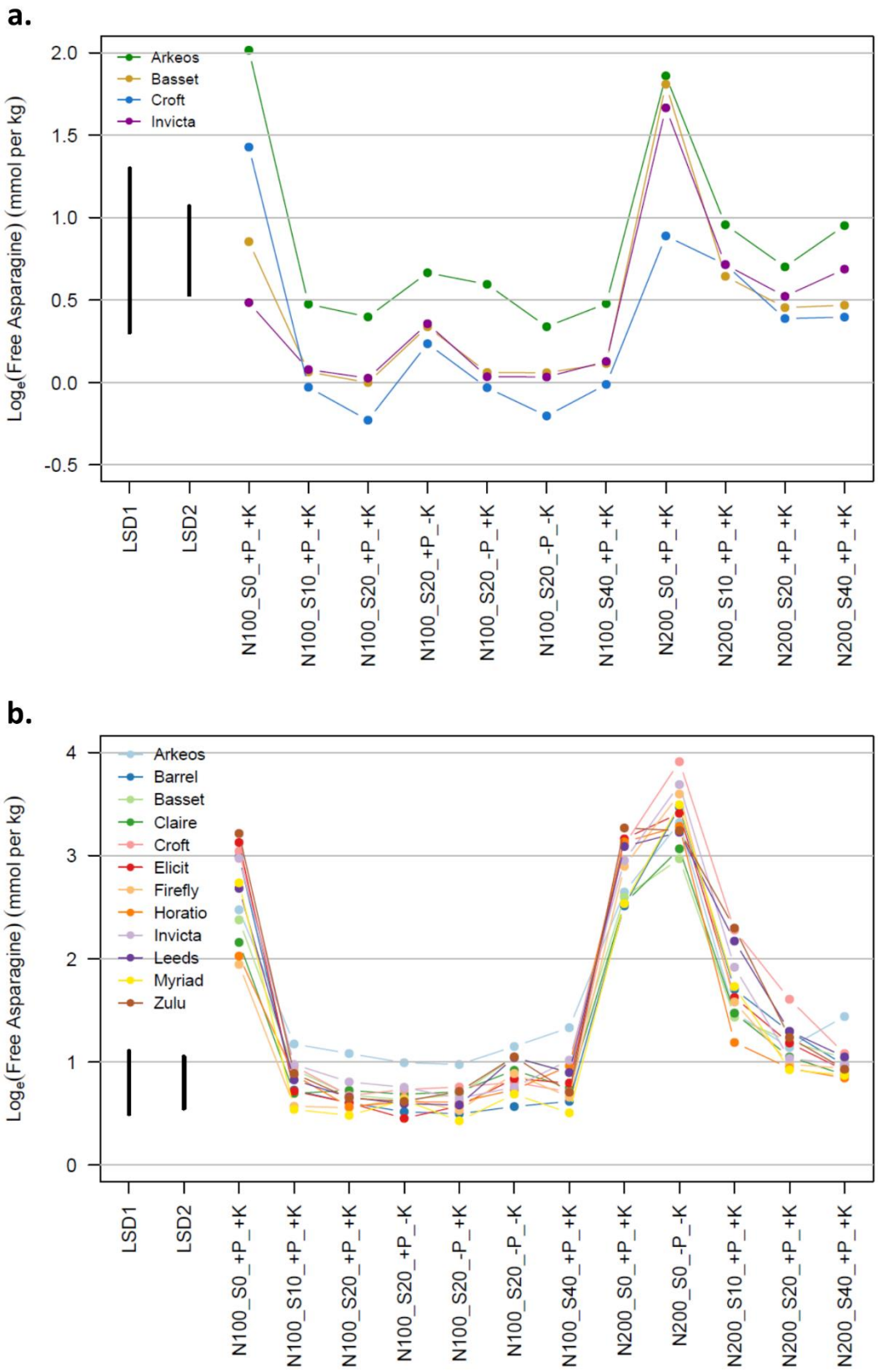
Supplementary Figure 6.5. Non-transformed mean asparagine content in trial H20 (**a.**) and H21 (**b.**) for each variety and treatment combination.



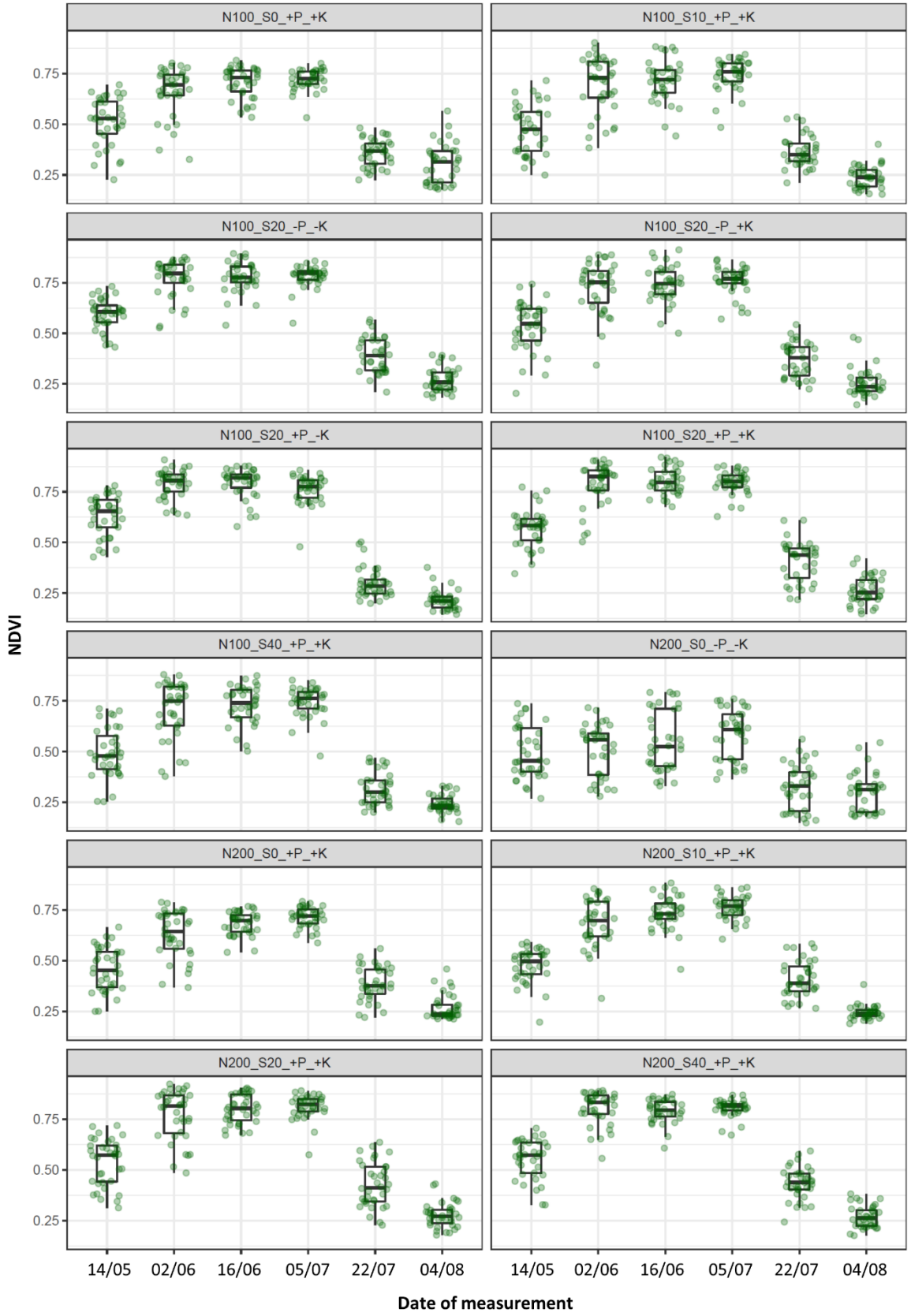
Supplementary Figure 6.6. Mean grain free asparagine content for each variety and fertiliser treatment in both trials and in different analyses. **a.** Variety means from nested model. **b.** Variety means from full model. **c.** Treatment means from nested model. **d.** Treatment means from full model. N (nitrogen), S (sulphur), P (phosphorus), K (potassium). Fertiliser application rates for N and S given in kg per hectare. Nested model LSDs (least significant differences) set at 5% and only valid for comparisons within each trial. Full model LSDs (least significant differences) set at 5%. LSD1s valid for comparison between trials and treatments/varieties. LSD2s only valid when comparing same level of trial.



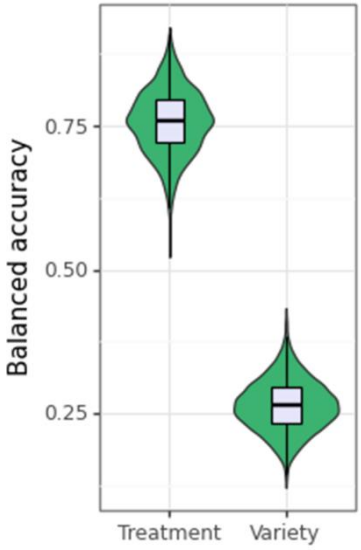
Supplementary Figure 6.7. Mean grain free asparagine content at different nitrogen: sulphur ratios (**a.**) and in the presence or absence of phosphorus (**b.**) and potassium (**c.**) fertiliser. LSDs valid for comparison between trials and treatments.



Supplementary Figure 6.8. Mean log$_e$ transformed free asparagine content in trials H20 (**a.**) and H21 (**b.**) for each variety and treatment combination.



Supplementary Figure 6.9. Normalised difference vegetation index measurements for all plots over all six measurement dates separated by agronomic treatment.



Supplementary Figure 6.10. Balanced accuracy values from gaussian naïve bayes classification models distinguishing sulphur deficient vs. sulphur fed plots (Treatment) and varieties using multispectral data from the field of trial H21.

Supplementary Table 6.1. Key dates in the two trials undertaken in this study.

Trial	Drilling	First nitrogen application	Second nitrogen application	Harvest
Stackyard	04/12/2019	10/03/2020	01/06/2020	10-11/09/2020
Butt Clong	26/11/2020	23/02/2021	17/05/2021	20/08/2021

Supplementary Table 6.2. Significance values (F probabilities) of terms in Nitrogen/Sulphur ANOVA models for analysis of Log_e transformed grain free asparagine content.

Source of variation	H20	H21	Both
Nitrogen	0.078	<.001	0.005
Sulphur	0.028	<.001	<.001
Nitrogen*Sulphur	0.987	0.015	0.664
Variety	<.001	<.001	<.001
Nitrogen*Variety	0.133	0.166	0.178
Sulphur*Variety	0.773	0.026	0.684
Nitrogen*Sulphur*Variety	0.069	0.050	0.067
Trial	NA	NA	0.012
Trial*Nitrogen	NA	NA	0.673
Trial*Sulphur	NA	NA	0.245
Trial*Nitrogen*Sulphur	NA	NA	0.935
Trial*Variety	NA	NA	<.001
Trial*Nitrogen*Variety	NA	NA	0.052
Trial*Sulphur*Variety	NA	NA	0.086
Trial*Nitrogen*Sulphur*Variety	NA	NA	0.462

7

Conclusions

7.1. Discussion of sub-projects

At the start of this PhD, there were three sub-projects to address. The ultimate goal of each sub-project was to develop strategies that would result in low acrylamide wheat-based food products (biscuits in particular). Each sub-project is restated and discussed in the context of our results below.

7.1.1. Characterisation of the TaASN-B2 *deletion at the molecular level and the impact that it has in the field*

The deletion of *TaASN-B2* described in this study (Oddy et al., 2021) is the first naturally occurring genetic variant controlling grain free asparagine content to be precisely identified and verified in field trials. The study provided the location and structure of *TaASN-B2* (Figure 4.1), its expression relative to the other *TaASN2* homeologues (Figure 4.7), its presence/absence in wheat relatives (Figure 4.4), and its impact on grain free asparagine content in two field trials of commercial UK wheat cultivars (Figure 4.8). The interaction between the *TaASN-B2* PAV (presence/absence variation) and sulphur fertiliser also demonstrated how genotype by environment interactions (G × E) need to be considered when testing varieties for grain free asparagine content. Although this variant appears to be effective under conditions of sulphur sufficiency only (when grain free asparagine content is low), it can be identified easily and used further in plant breeding programs to contribute to incremental changes in grain free asparagine content.

Although the old dataset of grain free asparagine content used in this study to investigate the impact of *TaASN-B2* (Curtis et al., 2018) gave us insight into how this PAV impacts grain free asparagine content in commercial wheat varieties, the background genotypic differences between the varieties also precluded the drawing of any definitive conclusions about the effect of *TaASN-B2*. Other studies have compared wheat plants in the same background with altered *TaASN2* genes (using TILLING and CRISPR), showing that the differences observed in grain free asparagine content between plants can be ascribed to *TaASN2* differences (Raffan et al., 2021; Alarcón-Reverte et al., 2022). In order for *TaASN-B2* screening to be more accessible for plant breeders, a higher throughput assay than the one used in our study would need to be developed (e.g., KASP). It would also be interesting to investigate further the origin of the *TaASN-B2* PAV. We identified a retrotransposon adjacent to the PAV and screened some wild wheat relatives to better understand when the PAV may have

originated, but a deeper investigation would be required to understand fully the origin of this variant.

7.1.2. Investigation of the soft wheat Claire × Robigus mapping population for genetic variation underpinning grain free asparagine content and other quality traits

Similar to the *TaASN-B2* study, my analysis of the Claire × Robigus mapping population also showed that incremental changes in grain free asparagine content could be achieved using existing soft wheat germplasm. The genetic control of amino acids, grain size traits, grain quality traits, and other developmental measurements was investigated using both conventional association mapping analyses (Figure 5.3; Table 5.1) and genomic prediction. This showed that genomic prediction outperformed the conventional mapping analysis for all traits and revealed interesting differences between the heritability and genetic control of these different traits (Figure 5.4). Using the recently released sequences of the 10+ wheat pangenome project and other available wheat genomic data (Walkowiak et al., 2020), potential candidate genes underpinning QTL controlling lysine content were also explored (Figure 5.5).

As mentioned briefly in Chapter 6, identification of candidate genes was made more difficult in this study because the genomes for Claire and Robigus were only scaffold-level (not assembled into chromosome level assemblies) (Walkowiak et al., 2020). This meant that direct alignments could not be performed easily between large chromosomal regions of interest; for example at the telomeric end of chromosome 4A in Robigus, which contains an introgression from *Triticum diccocoides* and which matches the location of some of the QTL identified in this study (Przewieslik-Allen et al., 2021). The low-density nature of the genetic map (872 markers) also limited the precision of the QTL mapping. Compared to other studies (Peng et al., 2018; Rapp et al., 2018), fewer QTL controlling grain free asparagine content were identified in this study. This may be expected because of the use of less diversity (a cross between two elite UK soft wheat cultivars) compared to other studies that used many more and more distantly-related wheat genotypes. Finally, the grain free asparagine contents measured in the two trials of this study were already quite low (all samples were under 4 mmol/kg) (Figure 5.1). Future association mapping studies should therefore focus on growing genotypes under stressed conditions (when grain free asparagine content is higher) in order to identify QTL that can reduce grain free asparagine content under stress; for example, in fields where

sulphur is not supplied or in glasshouses where plants can be infected with fungal pathogens to simulate disease in the field.

7.1.3. Investigation of how different agronomic strategies impact on soft wheat grain free asparagine content and acrylamide formation in biscuits

The final sub-project of my PhD refined the existing fertiliser recommendations, to minimise grain free asparagine content, and explored how multispectral imaging could be used to monitor free asparagine content. The fertiliser recommendations established in this project provide greater detail on the application rates of nitrogen and sulphur (in particular, the optimal ratio (10:1 kg/ha) at which they should be applied to minimise free asparagine accumulation) (Figure 6.1c). The study also provided the first investigation into how phosphorus and potassium deficiencies impact grain free asparagine content in the field (Figure 6.1d; Figure 6.1e). To demonstrate that progress made in reducing grain free asparagine content would be translated into lower acrylamide food products, biscuits were baked using grain samples from these trials (Figure 6.4). During the latter half of this sub-project, I also became interested in how imaging technology can contribute to the monitoring of grain free asparagine content. This led to the investigations into the relationship between multispectral measurements of plants/seeds and grain free asparagine content, showing that models developed using such technology can provide good estimates (Figure 6.2; Figure 6.3).

One difficulty that arose during this study was that I did not screen grain free asparagine content for all varieties in the first trial due to poor Hagberg Falling Numbers (HFN). Only those varieties that showed HFNs greater than the minimum (62) were taken forward for grain free asparagine content analysis since pre-harvest sprouting (PHS) had badly damaged the grain quality of all varieties. Such damaged grains would not usually be used to bake biscuits or any other products for human consumption, so only the four best (but still badly PHS damaged) varieties were selected for measurement of grain free asparagine. As a result, conclusions regarding the effect of environment (trial year and location) were limited to those four varieties that were present and analysed in both field trials. Furthermore, a third field trial using the same design structure as the second was supposed to run at the Rothamsted Broom's Barn site, but this trial could not be analysed due to a mistake in fertiliser application which prevented analysis using the original design structure. Consequently, the interpretation of the effect of environment on grain free asparagine content is limited to only four varieties across

two environments. Another aspect of this study that should be improved upon in future studies was the soil analysis. Differences in soil probably impact on grain free asparagine content to a large extent due to the accessibility of different nutrients. Although both trials were performed on similar soil types (loamy sand/sandy loam), greater detail on soil nutrient composition would have enabled a more thorough investigation into how this fundamental part of agronomy impacts grain free asparagine content.

7.2. Conclusions and future work

The results obtained in this project have a common theme underlying them: the differences between genotypes were smaller than differences caused by environmental factors. In the *TaASN-B2* study, the presence/absence variation did have an effect in sulphur-sufficient conditions when grain free asparagine content was already fairly low, but this difference was not observed under sulphur deficiency, which caused much larger increases in grain free asparagine content. In the mapping population study, a QTL which appears to account for a small amount of the variation in grain free asparagine content was identified, but again the effect of this QTL was less than the difference in grain free asparagine content between different trial years. Finally, in the fertiliser study, differences between varieties were again eclipsed by the much larger differences between fertiliser treatments and trial environment. In terms of practical recommendations then, agronomic management should be emphasised as one of the key strategies for keeping grain free asparagine content low.

As discussed in Chapter 6, grain free asparagine content is most problematic when it is at high concentrations caused by environmental stressors such as sulphur deficiency and disease. Whilst stressors can sometimes be managed through fertiliser application, disease control, and irrigation, management strategies cannot always protect a plant from stress. Consequently, agronomic management is unlikely to be sufficient to eliminate high asparagine grain from the field, due to many uncontrollable environmental factors. As agronomic management may not always be capable of eliminating high free asparagine content crops, monitoring may therefore be a partial solution to this problem.

As demonstrated in Chapter 6, imaging technologies appear to have some use in predicting the free asparagine content of plants and seeds. If agronomic strategies employed by a farmer fail to sufficiently control grain free asparagine content, this could be detected and the farmer could therefore sell this crop for industrial purposes other than consumption, such as

bioethanol or starch production (AHDB, 2022). Evidence for low free asparagine may also be requested from food manufacturers if stricter regulations are put in place. Our results also indicated that monitoring could be put in place further down the supply chain to detect high asparagine seeds; for example, at a mill to ensure that only low grain free asparagine content wheat were used. Finally, the study also shows that high acrylamide biscuits could be eliminated through colour sorting as a final quality control check. So, although agronomic management cannot always prevent the accumulation of high grain free asparagine content, monitoring from the farm through to the production line could eliminate those products with high asparagine or high acrylamide. In order to build on this work and develop further strategies for reducing acrylamide in food, three possible research directions are discussed below.

7.2.1. Comprehensive studies to dissect the environmental variation in free asparagine content

Environmental variation in grain free asparagine content (the variation due to trial year/site) remains mostly unaccounted for. Many possible causes for this environmental variation have been proposed previously, such as differences in weather, disease pressure, and soil type of the trial site. For example, Curtis et al. (2018) did not observe an effect of sulphur deficiency on grain free asparagine content in one year of field trials, which was speculated to have been an effect of the weather. Malunga et al. (2020) found that in their study of grain free asparagine content across two trial sites, 80% of the observed variation in asparagine was due to location. Xie et al. (2020) similarly found that 44.3% of observed variation in grain free asparagine content across four different field trials was attributed to environmental effects (trial year and/or location). The same study also noted that sulphur fertilisation did not have any impact at trial sites where sulphur was already plentiful in the soil.

To determine what causes this environmental variation, further field trials should be undertaken measuring all relevant factors that could impact grain free asparagine content. This would require measurements of weather, disease pressure, soil nutrients, and all other relevant environmental variables and agronomic management practices over several sites and years. By breaking the effect of environment down into these constituent parts, the amount of variation attributable to each component could be measured. Alternatively, glasshouse studies that specifically introduce individual stressors could be undertaken to isolate and compare the effect each stressor has on grain free asparagine content in a controlled environment. This experiment could be more easily performed due to its smaller scale, but interactions between stressors

would not be considered so it would be less relevant to cultivation of wheat in the field. Consequently, a series of field trials across many sites and years may be the best way of dissecting the effects of environmental stressors on grain free asparagine content.

7.2.2. Development of imaging strategies to predict grain free asparagine content

A more detailed understanding of the environmental factors that control grain free asparagine content (as outlined above) could enable farmers to effectively predict the grain free asparagine content of their crop. However, such detailed data are normally unavailable to farmers. A potentially easier solution would be to develop accessible imaging tools that could predict grain free asparagine content from measurements of plants and seeds. The findings of this study suggested that such imaging tools could feasibly be developed; however, a number of challenges would need to be addressed.

Firstly, these predictive models need to be built from many more measurements of plants and seeds. By training models on a more diverse dataset, the predictions of these models would be likely to be more accurate when used in different environments. This could be achieved by collecting imaging data from a series of field trials in different locations and years, similar to those described above for the study of environmental variation. This simultaneous collection of imaging and environmental variation data would greatly enhance our ability to predict grain free asparagine content. The ability to build such models also relies on the assumption that stressors that cause free asparagine to accumulate in the grain also leave a "signature" that can be detected by imaging technology (in the case of our study, a yellowing of the wheat canopy caused by sulphur deficiency). Many stressors do cause such signatures (Lowe et al., 2017), but the relationship between these signatures of stress and grain free asparagine content are not clear. Consequently the development of such models should assess the conditions under which predictions are most accurate, and those in which the models under- and over-estimate.

Another challenge facing the deployment of such a monitoring strategy is the accessibility of imaging tools. This study relied on a labour-intensive form of monitoring, which required the imaging device to be held above each plot to take a measurement. Similar multispectral measurements can be taken with unmanned aerial vehicles (UAVs) (Feng et al., 2021), and this would save time and labour if automated. Although UAVs may not be accessible to every farmer, this may be something contract agronomists could help with.

Alternatively, it may be possible to use satellite data to estimate grain free asparagine content. The spatial resolution of satellite data collection continues to improve (e.g., resolution of 0.25 m^2 in some satellites) and some satellites are able to collect multispectral data (Li et al., 2021), so it is feasible that satellite data could be used to build models to predict grain free asparagine content. Access to and interpretation of these data would still require specialist knowledge (as with the other monitoring methods), but the use of satellite data may require less input from the farmer in terms of time, labour, and money.

7.2.3. Testing of stress resistant wheat varieties

Whilst the results in this thesis showed that genotypic differences were not sufficient to control grain free asparagine content under diverse environments, this may not be indicative of what plant breeding could achieve to control grain free asparagine content. Many plant varieties have been bred to resist or tolerate the environmental stresses that cause free asparagine to accumulate (Hussain et al., 2021), so it is feasible that such varieties would show reduced free asparagine accumulation. The plant varieties tested in these studies were from a fairly narrow genetic base, mostly being elite UK cultivars. Consequently, more diverse germplasm (e.g., the Watkins collection of landraces (Wingen et al., 2017)) is likely to harbour alleles to deal with stress. One way of finding useful genetic variants would be to screen a diverse collection of stress-resistant wheat varieties under diverse stressors for grain free asparagine content. This would enable us to test the hypothesis that stress resistant wheat varieties accumulate less grain asparagine and also enable us to perform an association mapping analysis to find the genetic variants responsible. An analysis of stay-green varieties (Distelfeld et al., 2014) in such a study would also allow for an investigation into the impact of senescence on grain free asparagine content. If stress resistance and stay-green traits were found to control grain free asparagine content effectively, then these traits could be stacked to create a durable low grain free asparagine content wheat variety.

Induced variation (via TILLING or CRISPR) is another promising source of variation for grain free asparagine content. Wheat lines with non-functional *TaASN1* and *TaASN2* alleles accumulate less grain free asparagine (Raffan et al., 2021; Alarcón-Reverte et al., 2022), but these lines still need to be tested to see whether they can resist the large increases in grain free asparagine content caused by environmental stressors. If these lines perform well in field trials, then the non-functional *TaASN1* and *TaASN2* alleles, if generated via TILLING, could be used

in breeding programs. If generated by CRISPR, breeders would have to comply with the evolving regulations that restrict the commercialisation and cultivation of gene edited crops (Turnbull et al., 2021), but this may soon be easier in the UK as legislation to facilitate the use of gene edited crops and animals is currently going through parliament (UK Parliament, 2022).

7.3. Last word

Strategies such as those outlined in this thesis can provide effective control over acrylamide formation in wheat products, so long as they are used in combination with other processing strategies. As shown in our study, even when a crop is grown under good environmental conditions and has a low concentration of free asparagine in the grain, unfavourable processing conditions can cause acrylamide concentration to accumulate beyond current Benchmark Levels. Those involved at all stages of the food supply chain, from plant breeders and farmers through to millers and food manufacturers, can exert control over the acrylamide content of food (Figure 7.1), but must coordinate their efforts to ensure that the end product is low in acrylamide.

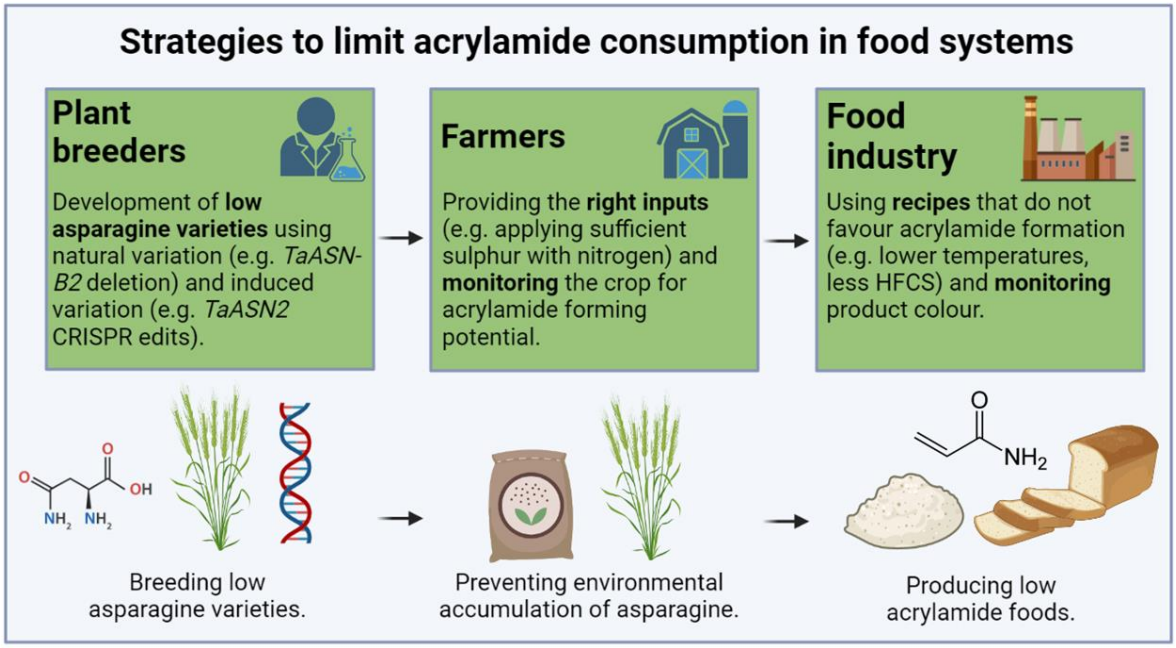


Figure 7.1. Summary of strategies through the food supply chain that can help to reduce the intake of dietary acrylamide. HFCS (high-fructose corn syrup), *TaASN2* (asparagine synthetase 2).

7.4. References

AHDB. (2022). UK human and industrial cereal usage. https://ahdb.org.uk/cereals-oilseeds/uk-human-industrial-cereal-usage (accessed October 11, 2022).

Alarcón-Reverte, R., Xie, Y., Stromberger, J., Cotter, J. D., Mason, R. E., & Pearce, S. (2022). Induced mutations in ASPARAGINE SYNTHETASE-A2 reduce free asparagine concentration in the wheat grain. *Crop Science, 62*(4), 1484-1496.

Curtis, T. Y., Powers, S. J., Wang, R., & Halford, N. G. (2018). Effects of variety, year of cultivation and sulphur supply on the accumulation of free asparagine in the grain of commercial wheat varieties. *Food Chemistry, 239,* 304-313.

Distelfeld, A., Avni, R., & Fischer, A. M. (2014). Senescence, nutrient remobilization, and yield in wheat and barley. *Journal of Experimental Botany, 65*(14), 3783-3798.

Feng, L., Chen, S., Zhang, C., Zhang, Y., & He, Y. (2021). A comprehensive review on recent applications of unmanned aerial vehicle remote sensing with various sensors for high-throughput plant phenotyping. *Computers and Electronics in Agriculture, 182,* 106033.

Hussain, B., Akpınar, B. A., Alaux, M., Algharib, A. M., Sehgal, D., Ali, Z., ... & Budak, H. (2021). Wheat genomics and breeding: bridging the gap. *AgriRxiv, (2021),* 20210081805.

Li, D., Quan, C., Song, Z., Li, X., Yu, G., Li, C., & Muhammad, A. (2021). High-throughput plant phenotyping platform (HT3P) as a novel tool for estimating agronomic traits from the lab to the field. *Frontiers in Bioengineering and Biotechnology, 8,* 623705.

Lowe, A., Harrison, N., & French, A. P. (2017). Hyperspectral image analysis techniques for the detection and classification of the early onset of plant disease and stress. *Plant Methods, 13*(1), 1-12.

Malunga, L. N., Ames, N., Khorshidi, A. S., Thandapilly, S. J., Yan, W., Dyck, A., ... & Scanlon, M. G. (2021). Association of asparagine concentration in wheat with cultivar, location, fertilizer, and their interaction. *Food Chemistry, 344,* 128630.

Oddy, J., Alarcón-Reverte, R., Wilkinson, M., Ravet, K., Raffan, S., Minter, A., ... & Pearce, S. (2021). Reduced free asparagine in wheat grain resulting from a natural deletion of TaASN-

B2: investigating and exploiting diversity in the asparagine synthetase gene family to improve wheat quality. *BMC Plant Biology*, *21*(1), 1-17.

Peng, Y., Liu, H., Chen, J., Shi, T., Zhang, C., Sun, D., He, Z., Hao, Y., & Chen, W. (2018). Genome-wide association studies of free amino acid levels by six multi-locus models in bread wheat. *Frontiers in Plant Science*, *9*, 1196

Przewieslik-Allen, A. M., Wilkinson, P. A., Burridge, A. J., Winfield, M. O., Dai, X., Beaumont, M., ... & Edwards, K. J. (2021). The role of gene flow and chromosomal instability in shaping the bread wheat genome. *Nature Plants*, *7*(2), 172-183.

Raffan, S., Sparks, C., Huttly, A., Hyde, L., Martignago, D., Mead, A., ... & Halford, N. G. (2021). Wheat with greatly reduced accumulation of free asparagine in the grain, produced by CRISPR/Cas9 editing of asparagine synthetase gene TaASN2. *Plant Biotechnology Journal*, *19*(8), 1602-1613.

Rapp, M., Schwadorf, K., Leiser, W. L., Würschum, T., & Longin, C. F. H. (2018). Assessing the variation and genetic architecture of asparagine content in wheat: What can plant breeding contribute to a reduction in the acrylamide precursor? *Theoretical and Applied Genetics*, *131*(11), 2427-2437.

Turnbull, C., Lillemo, M., & Hvoslef-Eide, T. A. (2021). Global regulation of genetically modified crops amid the gene edited crop boom–a review. *Frontiers in Plant Science*, *12*, 630396.

UK Parliament. (2022). Genetic Technology (Precision Breeding) Bill. https://bills.parliament.uk/bills/3167 (accessed October 27, 2022).

Walkowiak, S., Gao, L., Monat, C., Haberer, G., Kassa, M. T., Brinton, J., ... & Pozniak, C. J. (2020). Multiple wheat genomes reveal global variation in modern breeding. *Nature*, *588*(7837), 277-283.

Wingen, L. U., West, C., Leverington-Waite, M., Collier, S., Orford, S., Goram, R., ... & Griffiths, S. (2017). Wheat landrace genome diversity. *Genetics*, *205*(4), 1657-1676.

Xie, Y., Malunga, L. N., Ames, N. P., Waterer, J., Khorshidi, A. S., & Scanlon, M. G. (2021). Effects of growing environment, genotype, and commercial fertilization levels on free asparagine concentration in Western Canadian wheat. *Cereal Chemistry, 98*(1), 89-99.