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
Rothamsted Research
Rocio Alárcon-Reverte
Colorado State University
Mark Wilkinson
Rothamsted Research
Karl Ravet
Colorado State University
Sarah Raffan
Rothamsted Research
Andrea Minter
Rothamsted Research
Andrew Mead
Rothamsted Research
J. Stephen Elmore
University of Reading
Isabel Moreira Almeida
Mondelēz International (United Kingdom)
Nicholas C. Cryer
Mondelēz International (United Kingdom)
Nigel G. Halford
Rothamsted Research
Stephen Pearce (Stephen.Pearce@colostate.edu)
Colorado State University

Research Article

Keywords: wheat, Triticum aestivum, asparagine metabolism, acrylamide, asparagine synthetase, food safety, wheat breeding

Posted Date: March 11th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-276186/v1

Abstract

Background: Understanding the determinants of free asparagine concentration in wheat grain is necessary to reduce levels of the processing contaminant acrylamide in baked and toasted wheat products. Although crop management strategies can help reduce asparagine levels, breeders have limited options to select for genetic variation underlying this trait. Asparagine synthetase enzymes catalyse a critical step in asparagine biosynthesis in plants and, in wheat, are encoded by five homeologous gene triads that exhibit distinct expression profiles. Within this family, *TaASN2* genes are highly expressed during grain development but *TaASN-B2* is absent in some varieties.

Results: Natural genetic diversity in the asparagine synthetase gene family was assessed in different wheat varieties revealing instances of presence/absence variation and other polymorphisms, including some predicted to affect the function of the encoded protein. The presence and absence of *TaASN-B2* was determined across a range of UK and global common wheat varieties and related species, showing that the deletion encompassing this gene was already present in some wild emmer wheat genotypes. Expression profiling confirmed that *TaASN2* transcripts were only detectable in the grain, while *TaASN3.1* genes were highly expressed during the early stages of grain development. *TaASN-A2* was the most highly expressed *TaASN2* homeologue in most assayed wheat varieties. *TaASN-B2* and *TaASN-D2* were expressed at similar, lower levels in varieties possessing *TaASN-B2*. Expression of *TaASN-A2* and *TaASN-D2* did not increase to compensate for the absence of *TaASN-B2,* so total *TaASN2* expression was lower in varieties lacking *TaASN-B2*. Consequently, free asparagine levels in field-produced grain were, on average, lower in varieties lacking *TaASN-B2*, although the effect was lost when free asparagine accumulated to very high levels as a result of sulphur deficiency.

Conclusions: Selecting wheat genotypes lacking the *TaASN-B2* gene may be a simple and rapid way for breeders to reduce free asparagine levels in commercial wheat grain.

Background

Asparagine in its free (soluble, non-protein) form is an important nitrogen transport and storage molecule in plants (see [1] for review). It also accumulates during abiotic and biotic stress and has potential roles in ammonia detoxification and reactive oxygen species/nitrous oxide production (see [2] 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 [3]). 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 [4].

The European Commission has led the way in developing a regulatory system for acrylamide levels in food (see [3] for a comprehensive review). The current EU regulation on acrylamide in food (Commission Regulation (EU) 2017/2158 [5]) 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 levels in different food types, and carries an explicit threat to set Maximum Levels (i.e. levels 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' [6]. 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 [3]). 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 [3]). 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 [7]. 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, reductions in free asparagine concentration in the grain have mostly been achieved through crop management strategies, including ensuring that wheat is supplied with sufficient sulphur during cultivation [8] and is protected from pathogen infection [9, 10]. Wheat breeders are under pressure from food businesses to develop varieties with reduced concentrations of free asparagine in the grain. Although free asparagine levels do vary across genotypes and exhibit reasonably high heritability [11–13], 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 asparagine levels will not be a simple task (reviewed in [3]). Through association mapping, several quantitative trait loci (QTL) for asparagine content have been identified, but they explain only a small proportion of the variance and no common QTL have been identified between studies [12, 13].

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 [14], 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*) [14, 15]. 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 [14]. 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.) [14]. Since wild emmer wheat is believed to be the B genome donor for both tetraploid durum and hexaploid common wheat [16], the most likely 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* [14]. This would be consistent with evidence of multiple hybridisations found from wider analyses of genome data [17].

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 [18, 19], and their expression in the embryo is likely a key factor determining free asparagine concentrations in the grain as a whole [19]. 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 asparagine concentrations and may be a useful allele for wheat breeding programmes to develop varieties with lower levels of free asparagine.

Results

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 1, ordered by gene name [14, 15] and the corresponding annotated gene model ID from the Chinese Spring RefSeq v1.1 genome assembly [20]. 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 1) or to be tolerated (highlighted in green). Full details of specific amino acid changes for all wheat varieties are provided in Additional File 1, Table S1.

Table 1. Natural variation in ASN proteins in 14 wheat varieties. Each gene is annotated by name [14, 15] 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 [18] (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.

		Robigus	CDC Landmark	Julius	Claire	Jagger	Cadenza	Paragon	Arina	Norin 61	CDC Stanley	SY Mattis Mattis	Lan
Protein	IWGSC RefSeq v1.1 ID												
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 (Additional File 1, Fig. S1). *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 1). In contrast, twelve wheat varieties carried *TaASN-A3.1* alleles with polymorphisms predicted to disrupt protein function (Table 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 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 1).

Characterisation of the *TaASN-B2* deletion

The deletion containing *TaASN-B2* mapped to chromosome arm 3BS in the Chinese Spring RefSeq v1.1 genome assembly (Fig. 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. 1a and 1b). 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. 1a). 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 [21] (Fig. 1a), the identity of which was confirmed using the TREP database [22]. Analysis of the other genome assemblies for the genotypes shown in Table 1 revealed that all eight varieties lacking *TaASN-B2* had identical breakpoints.

Wider screening for the presence/absence of *ASN-B2*

Because of their potential role in determining free asparagine levels 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. 2a). The encoded proteins shared >99 % identity at the amino acid level, with only eight polymorphic residues between homeologues (Fig. 2b). Although four of these polymorphisms fell in the glutamine amidotransferase (GATase) domain (Fig. 2b), none were predicted to affect protein function according to the SIFT analysis (Table 1).

The length of intron 1 varied between homeologues and was 1,104 bp in *TaASN-A2*, 1,411 bp in *TaASN-D2*, but only 175 bp in *TaASN-B2* (Fig. 2a). A pair of redundant primers was 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 was designed to anneal upstream and downstream of the deleted region containing *TaASN-B2* to amplify a DNA fragment only in genotypes carrying this deletion. 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. 3a-d and summarised in Table 2a. Overall, *TaASN-B2* was deleted in 52 of 63 UK winter wheat varieties assayed (82.5 %) (Fig. 3e). The deletion was most common in the biscuit (G3) class (93.3 %) and least common in the breadmaking (G1) (70 %) class (Fig. 3e). An additional set of 24 global wheat varieties were analysed using a similar PCR assay (Additional File 1, Fig. S2a) and the results are shown in Additional File 1, Fig. S2b, and summarised in Table 2b. The *TaASN-B2* deletion was less frequent among these wheats than in the UK varieties, being present in just 50 % of the genotypes (Table 2b).

Table 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.

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

a.
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			

A selection of other wheat species was also screened for the presence of an *ASN-B2* gene (Fig. 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). 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). This band was still present when the experiment was repeated (data not shown) and is likely due to genetic heterogeneity in the sample.

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 Additional File 2 as mean Transcripts Per Million (TPM) values. In a dataset encompassing roots, leaves, stems, spike and grain, each sampled at three developmental stages [23], *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. 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. 5a). As shown previously [18, 19], *TaASN2* showed a grain-specific expression profile, with transcript levels highest at Zadoks stage 85 (Z85), which corresponds to the soft dough stage [24] (Fig. 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 [19].

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 [25]. As expected [18, 19], 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. 5b). The exception was the grain milk stage, where *TaASN3.1* transcript levels were higher than *TaASN2* (Fig. 5b), as shown previously [18]. *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 [26] revealed that *TaASN2* transcript levels were highest in the mature embryo stage (Fig. 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. 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. 5c). Taken together, these data confirm the specific activity 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.

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 at 14 days post anthesis (DPA) and 30 DPA taken from 27 worldwide wheat varieties [26]. 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. 6a), consistent with previous results [19]. At 30 DPA, *TaASN2* homeologues were the most highly expressed asparagine synthetase genes in all varieties assayed (Fig. 6b). At this latter timepoint, total *TaASN2* transcript levels showed large variation between genotypes, ranging from 28 to 242 TPM (Additional File 2). 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 (Additional File 1, Fig. S2b; Table 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 (Additional File 1, Fig. S2b; Table 2b).

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. 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. 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 (Additional File 2), 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. 7). The results of the analysis of variance for this experiment are shown in Table 3a, 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. 7a and 7b). 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. 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. 7c and 7d). Notably, there was no evidence of higher expression of *TaASN-A2* or *TaASN-D2* in Claire and Spark compared with Cadenza and Duxford.

Table 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. *p*-values are displayed from the ANOVA analyses and Chi probability values are displayed from the REML analysis.

a

Factor	P-value
Timepoint	<0.001
Variety	<0.001
Timepoint*Variety	<0.001
Homeologue	<0.001
Timepoint*Homeologue	<0.001
Variety*Homeologue	<0.001
Timepoint*Variety*Homeologue	<0.001

b

	2011-2012 ANOVA	2012-2013 ANOVA	Both years' REML
Treatment	0.04	0.009	<0.001
ASN-B2	<0.001	0.27	0.040
*ASN-B2**Variety	<0.001	0.002	0.003
*ASN-B2**Treatment	0.195	<0.001	0.012
*ASN-B2**Variety*Treatment	0.068	0.008	0.100
Year			<0.001
Year**ASN-B2*			0.948
Year*Treatment			<0.001
Year**ASN-B2**Variety			0.439
Year**ASN-B2**Treatment			0.009
Year**ASN-B2**Treatment*Variety			0.326

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

The screen of varieties for the presence or absence of *TaASN-B2* (Fig. 3, Table 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) [11]. This meant that an assessment could be made of the effect of the *TaASN-B2* deletion on free asparagine levels 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 [11]. Box and scatter plots showing free asparagine levels in each variety, grouped by sulphur treatment and *TaASN-B2* presence/absence, are shown for the 2011 – 2012 field trial varieties (Fig. 8a), 2012 – 2013 field trial varieties (Fig. 8b), and combined field trial varieties (Fig. 8c). Analysis of variance revealed a significant ($p < 0.05$) effect of *TaASN-B2* in the 2011 – 2012 field trial, with higher asparagine levels in the varieties possessing *TaASN-B2*, but not in the 2012 – 2013 field trial (Table 3b). In the 2012 – 2013 field trial, however, there was a significant interaction between *TaASN-B2* and treatment, with a *post-hoc* test revealing significantly ($p < 0.05$) higher free asparagine levels in varieties possessing *TaASN-B2* under sulphur sufficiency (Fig. 8b). Restricted Maximum Likelihood (REML) analysis performed on the combined years' dataset also revealed a significant increase in asparagine levels for varieties possessing *TaASN-B2* and an interaction between *TaASN-B2* and treatment (Table 3b).

Discussion

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, strongly suggesting a single, common origin for this allele. The deleted region is in close proximity to a Ty1-*copia* transposable element (Fig. 1a), one of the most abundant classes of LTR retrotransposons in the wheat genome [27], 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) 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 [16], 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 [14, 17], suggests 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 [16], 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. 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 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 asparagine levels [28]. However, other studies have found no association between asparagine content and a host of baking quality traits [29], 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 as well; similar to the moonlighting roles found for asparagine synthetase in yeast and mammalian cells [30, 31].

Expression profiles of wheat asparagine synthetase genes

Our expression data confirmed previous findings [18, 19] that, in many varieties, *TaASN-A2* is the most highly expressed *TaASN2* homeologue during grain development (Figs. 5, 6 and 7). In five of the twelve 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. 6c, Additional file 2). 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. 7). In varieties lacking *TaASN-B2*, *TaASN-A2* contributed a greater proportion of *ASN2* transcripts than *TaASN-D2*, while in varieties with *TaASN-B2* present, the three homeologues were more evenly expressed (Fig. 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 asparagine levels. This is in contrast to other biosynthetic pathways, for example in GA signalling [32], 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. 5 and 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 [18, 19]. 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 [19]. 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 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.

Breeding for reduced grain asparagine content

The *TaASN-B2* deletion was associated with a reduction in free asparagine in the grain from two field trials and showed a significant interaction with sulphur treatment, with proportionally greater reductions in free asparagine when sulphur was plentiful (Fig. 8). This interaction suggests that the deletion has less of an effect under sulphur deficiency, when free asparagine levels are higher, but also suggested that it may reduce free asparagine levels under sulphur deficiency in some varieties.

Looking at the two field trials separately, the effect of the deletion under different sulphur treatments becomes more complex. In the 2011 – 2012 field trial, the *TaASN-B2* deletion itself was significant, whereas there was no interaction between sulphur treatment and the deletion. This may be a result of the milder effect of sulphur deficiency on free asparagine levels observed in this trial, with free asparagine increasing to a lesser extent under sulphur deficiency, as noted in the previous study [11]. In contrast, the *TaASN-B2* deletion itself was not significantly associated with free asparagine levels across both sulphur treatments in the 2012 – 2013 field trial. Instead, there was a significant interaction between the deletion and sulphur treatment, with the deletion conferring lower free asparagine levels only under sulphur sufficiency. This may be a result of the more severe effects of sulphur deficiency observed during this growing season: asparagine levels increased much more in response to sulphur deficiency than in the 2011 – 2012 trial, likely masking any effect of the deletion. Consequently, the effect of the deletion was most obvious when sulphur deficiency was minimised, consistent with earlier studies showing a strong environmental impact on free asparagine content [3].

Measuring free asparagine levels 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 levels of free asparagine in the grain from their

breeding programmes. 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 [33].

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.

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. 3e, Table 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 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 levels [34]. However, durum wheat grain is also incorporated into grists for making products in which acrylamide levels 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 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. However, previous association mapping studies for asparagine content detected only a small number of minor-effect QTL [12, 13], none of which mapped to regions of the genome containing asparagine synthetase genes. This highlights the limited options for breeders aiming to develop low asparagine varieties using natural genetic variation. 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 [35, 36]. An obvious application of these tools is to induce gene knockouts in all three *TaASN2* homeologues. 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 [37].

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.

Methods

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) [38], and the Graingenes database (https://wheat.pw.usda.gov/cgi-bin/seqserve/blast_wheat.cgi) [39]. Some *ASN* genes lacked complete sequence information, and these exceptions are described in Additional File 1, Table S2). 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) [40] 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) [20]. The softberry-FGENESH tool [41] 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 [42] 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) [22]. SIFT analysis [43] 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.

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 [11]. 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). 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 [44], 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.

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 [45]. DNA quality and abundance were assayed using a NanoDrop™ 1000 Spectrophotometer (Thermo Fisher Scientific).

PCR assays to detect *TaASN-B2*

Homeologue-specific primers ASN-B2-Deletion-F and ASN-B2-Deletion-R (Additional File 1, Table S3) 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 (Additional File 1, Table S3), 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 1,363 bp, 434 bp and 1,670 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 (1.5mM $MgCl_2$) (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 (Additional File 1, Fig. S2). 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 (Additional File 1, Fig. S2). 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 (Additional File 1, Fig. S2b). Full, uncropped images of all electrophoresis gels are provided in Additional File 3.

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 [46]. 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 Additional File 1, Table S3.

Relative expression values were calculated as described by Rieu and Powers [47]. 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 [48]. Statistical tests were performed using GenStat [49] to account for the blocking structure of the experiment and graphs were plotted in R [50] using the package ggpubr [51].

RNA-seq data analysis

Raw RNA-seq reads from a developmental timecourse in the landrace Chinese Spring [23], grain development samples from the variety Azhurnaya [25], an embryo development timecourse from the variety AC Barrie [52] and grain expression at 14 DPA and 30 DPA from a set of worldwide wheat varieties [26], 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 [53]. 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 [54]. The ternary plot was created using the ggtern package [55] within ggplot2 [56]. All expression data in TPM are presented in Additional File 2.

Effect of *TaASN-B2* on free asparagine in the grain

Data from the 2011 – 2012 and 2012 - 2013 field trials [11] were used to investigate the effect of the *TaASN-B2* deletion on grain asparagine levels with and without sulphur deficiency. 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 [11]. Analyses and plotting were performed in R [50] with the package ggpubr [51]. ANOVA and REML analyses were performed in GenStat [49] to account for the split-plot blocking structure of the field trials and *post-hoc* Bonferroni tests were performed to allow for multiple comparisons between varieties.

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 [11] 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).

Declarations

Ethics approval and consent to participate

Not applicable

Consent for publication

Not applicable

Availability of data and material

Expression data was generated from public RNA-seq datasets that were downloaded from https://www.ncbi.nlm.nih.gov/geo/ and processed as described in the results section. All expression values used to generate expression graphs are presented in Additional File 2. Amino acid data for protein sequences in the landrace Chinese Spring were described previously [11]. DNA and protein sequences from other wheat varieties were derived from 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) [38], and the Graingenes database (https://wheat.pw.usda.gov/cgi-bin/seqserve/blast_wheat.cgi).

Funding

JO is supported by a BBSRC Collaborative Training Partnership Studentship with partners: University of Reading and Mondelez UK R&D Ltd. SR was supported by a Biotechnology and Biological Sciences Research Council (BBSRC) SWBio iCASE DTP Studentship, with partners: University of Bristol, AHDB, KWS UK Ltd, Saaten Union UK Ltd, RAGT Seeds Ltd, Syngenta UK Ltd, and Limagrain UK Ltd. NGH is supported at Rothamsted Research by the BBSRC via the Designing Future Wheat Programme (BB/P016855/1). BBSRC is part of UK Research and Innovation. Research at Colorado State University was partially funded by the Colorado Wheat Research Foundation and Colorado Wheat Administrative Committee.

Competing interests

The funders listed above had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish.

Authors' contributions

JO performed sequence analysis and annotation, designed and ran molecular assays, performed statistical analysis, and wrote the manuscript. RAR performed sequence analysis and annotation, and designed and ran molecular assays. MW designed and ran molecular assays. KR performed expression analyses and wrote the manuscript. SR performed sequence analysis and annotation. AM1 and AM2 advised and directed the experimental design and statistical analysis. JSE, IMdA and NCC advised on experimental design and data analysis; SP and NGH wrote the manuscript. All authors read and approved the manuscript.

Acknowledgements

Not applicable

Authors' information

Not applicable

Abbreviations

ANOVA, Analysis of Variance; BLAST, Basic Local Alignment Search Tool; CRISPR, Clustered Regularly Interspersed Short Palindromic Repeats; CTAB, Cetyl Trimethyl Ammonium Bromide; DPA, Days Post Anthesis; EMS, Ethyl Methane Sulphonate; EU, European Union; FDA, Food and Drug Administration; FSANZ, Food Standards Agency of Australia and New Zealand; GA, Gibberellin; GATase, Glutamine amidotransferase; IWGSC, International Wheat Genome Sequencing Consortium; kb, kilo base pair; LTR, Long Terminal Repeat; PCR, Polymerase Chain Reaction; QTL, Quantitative Trait Locus; REML, Restricted Maximum Likelihood; TPM, Transcripts Per Million.

References

1. Lea PJ, Sodek L, Parry MA, Shewry PR, Halford NG Asparagine in plants. Ann Appl Biol 2007;150:1–26.
2. Oddy, J., Raffan, S., Wilkinson, M.D. et al. Stress, nutrients and genotype: understanding and managing asparagine accumulation in wheat grain. CABI Agric Biosci. 2020;1:10.
3. Raffan S, Halford NG. Acrylamide in food: progress in and prospects for genetic and agronomic solutions. Ann Appl Biol 2019;175:259–281.
4. IARC International Agency for Research on Cancer. Some Industrial Chemicals; IARC Monographs on the Evaluation of Carcinogenic Risks to Humans. Vol. 60. Lyon: International Agency for Research on Cancer; 1994.
5. European Commission. Commission Regulation EU. 2017/2158 establishing mitigation measures and benchmark levels for the reduction of the presence of acrylamide in food. Brussels: European Commission; 2017.
6. Food and Drug Administration. Guidance for Industry, Acrylamide in Foods. College Park MD, Food and Drug Administration; 2016.
7. Zhivagui M, Ng AWT, Ardin M, Churchwell MI, Pandey M, Renard C, Villar S, Cahais V, Robitaille A, Bouaoun L, Heguy A, Guyton KZ, Stampfer MR, McKay J, Hollstein M, Olivier M, Rozen SG, Beland FA, Korenjak M, Zavadil J. Experimental and pan-cancer genome analyses reveal widespread contribution of acrylamide exposure to carcinogenesis in humans. Genome Res. 2019;29:521–531.
8. Raffan S, Oddy J, Halford NG. The sulphur response in wheat and its implications for acrylamide formation and food safety. Int J Molec Sci. 2020;21:3876.
9. Martinek P, Klem K, Vanova M, Bartackova V, Vecerkova L, Bucher P, Hajslova J. 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 Env. 2009;55:187–95.
10. Curtis TY, Powers SJ, Halford NG. Effects of fungicide treatment on free amino acid concentration and acrylamide-forming potential in wheat. J Agric Food Chem. 2016;64:9689–96.
11. Curtis TY, Powers SJ, Wang R, Halford, NG. Effects of variety, year of cultivation and sulphur supply on the accumulation of free asparagine in the grain of commercial wheat varieties. Food Chem. 2018;239:304–13.
12. Rapp M, Schwadorf K, Leiser WL, Würschum T, Longin CFH. Assessing the variation and genetic architecture of asparagine content in wheat: What can plant breeding contribute to a reduction in the acrylamide precursor? Theor Appl Genet. 2018;131:2427–37.
13. Emebiri LC. Genetic variation and possible SNP markers for breeding wheat with low-grain asparagine, the major precursor for acrylamide formation in heat-processed products. J Sci Food Agric 2014;94:1422–29.
14. Raffan S, Halford NG. Phylogenetic analysis of cereal asparagine synthetase genes. Ann Appl Biol 2021;178:6–22.
15. Xu H, Curtis TY, Powers SJ, Raffan S, Gao R, Huang J, Heiner M, Gilbert D, Halford NG. Genomic, biochemical and modelling analyses of asparagine synthetases from wheat. Frontiers Plant Sci. 2018;8:2237.
16. Dubcovsky J, Dvorak J. Genome plasticity a key factor in the success of polyploid wheat under domestication [published correction appears in Science. 2007 Oct 19;318(5849):393]. Science. 2007;316:1862–1866.
17. Marcussen T, Sandve SR, Heier L, Spannag M, Pfeifer M, The International Wheat Genome Sequencing Consortium, Jakobsen KS, Wulff BBH, Steuernage B, Mayer KFX, Olsen O-A. Ancient hybridizations among the ancestral genomes of bread wheat. Science 2014;345:1250092.
18. Gao R, Curtis TY, Powers SJ, Xu H, Huang J, Halford NG. Food safety: Structure and expression of the asparagine synthetase gene family of wheat. J Cer Sci 2016;68:122–31.
19. Curtis TY, Raffan S. Wan Y, King R, Gonzalez-Uriarte A, Halford NG. Contrasting gene expression patterns in grain of high and low asparagine wheat genotypes in response to sulphur supply. BMC Genomics 2019;20:628.
20. The International Wheat Genome Sequencing Consortium (IWGSC), Appels R, Eversole K, Feuillet C, Keller B, Rogers J, Stein N, et al. Shifting the limits in wheat research and breeding using a fully annotated reference genome. Science. 2018;361:eaar7191.
21. Wicker T, Keller B. 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 Res. 2007;17:1072–81.
22. Wicker T, Matthews DE, Keller B. TREP: A database for Triticeae repetitive elements. Trends Plant Sci. 2002;7:561–62.
23. Choulet F, Alberti A, Theil S, Glover N, Barbe V, Daron J, Pingault L, Sourdille P, Couloux A, Paux E, Leroy P, Mangenot S, Guilhot N, Le Gouis J, Balfourier F, Alaux M, Jamilloux V, Poulain J, Durand C, Bellec A, Gaspin C, Safar J, Dolezel J, Rogers J, Vandepoele K, Aury JM, Mayer K, Berges H, Quesneville H, Wincker P, Feuillet C. Structural and functional partitioning of bread wheat chromosome 3B. Science. 2014;345:1249721.
24. Zadoks JC, Chang TT, Konzak CF. A decimal code for the growth stages of cereals. Weed Res. 1974;14:415–21.

25. Ramírez-González RH, Borrill P, Lang D, Harrington SA, Brinton J, Venturini L, Davey M, Jacobs J, van Ex F, Pasha A, Khedikar Y, Robinson SJ, Cory AT, Florio T, Concia L, Juery C, Schoonbeek H, Steuernagel B, Xiang D, Ridout CJ, Chalhoub B, Mayer KFX, Benhamed M, Latrasse D, Bendahmane A; International Wheat Genome Sequencing Consortium, Wulff BBH, Appels R, Tiwari V, Datla R, Choulet F, Pozniak CJ, Provart NJ, Sharpe AG, Paux E, Spannagl M, Bräutigam A, Uauy C. The transcriptional landscape of polyploid wheat. Science. 2018;361:eaar6089.
26. Nirmal RC, Furtado A, Wrigley C, Henry RJ. Influence of gene expression on hardness in wheat. PLoS ONE. 2016;11:e0164746.
27. Wicker T, Gundlach H, Spannagl M, Uauy C, Borrill P, Ramírez-González RH, De Oliveira R, International Wheat Genome Sequencing Consortium, Mayer K, Paux E, Choulet F. Impact of transposable elements on genome structure and evolution in bread wheat. Genome Biol. 2018;19:103.
28. Simsek S, Ohm JB, Lu H, Rugg M, Berzonsky W, Alamri MS, Mergoum M. Effect of pre-harvest sprouting on physicochemical changes of proteins in wheat. J Sci Food Agric. 2014;94:205–12.
29. Rapp M, Schwadorf K, Leiser WL, Würschum T, Longin CF. Assessing the variation and genetic architecture of asparagine content in wheat: What can plant breeding contribute to a reduction in the acrylamide precursor? Theor Appl Genet. 2018;131:2427–37.
30. Noree C, Sirinonthanawech N, Wilhelm JE. Saccharomyces cerevisiae *ASN1* and *ASN2* are asparagine synthetase paralogs that have diverged in their ability to polymerize in response to nutrient stress. Sci. Rep. 2019;9:278.
31. Noree C, Monfort E, Shotelersuk V. Human asparagine synthetase associates with the mitotic spindle. Biol. Open. 2018;15;7(12).
32. Middleton AM, Úbeda-Tomás S, Griffiths J, Holman T, Hedden P, Thomas SG, Phillips AL, Holdsworth MJ, Bennett MJ, King JR, Owen MR. Mathematical modeling elucidates the role of transcriptional feedback in gibberellin signaling. Proc Natl Acad Sci USA. 2012;109:7571–6.
33. Lei GJ, Fujii-Kashino M, Hisano H, Saisho D, Deng F, Yamaji N, Sato K, Zhao FJ, Ma JF. Breeding for low cadmium barley by introgression of a Sukkula-like transposable element. Nat. Food. 2020;1:489–99.
34. European Food Safety Authority Panel on Contaminants in the Food Chain (CONTAM Panel). Scientific opinion on acrylamide in food. EFSA J. 2015;13:4104.
35. Krasileva KV, Vasquez-Gross HA, Howell T, Bailey P, Paraiso F, Clissold L, Simmonds J, Ramirez-Gonzalez RH, Wang X, Borrill P, Fosker C. Uncovering hidden variation in polyploid wheat. Proc Natl Acad Sci USA. 2017;114:E913-21.
36. Zhang Y, Pribil M, Palmgren M, Gao C. A CRISPR way for accelerating improvement of food crops. Nat. Food. 2020;1:200–5.
37. Schmidt SM, Belisle M, Frommer WB. 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 Rep. 2020;19:e50680.
38. Bian X, Tyrrell S, Davey RP. The Grassroots life science data infrastructure. 2017. https://grassroots.tools.
39. Blake VC, Woodhouse MR, Lazo GR, Odell SG, Wight CP, Tinker NA, Wang Y, Gu YQ, Birkett CL, Jannink JL, Matthews DE, Hane DL, Michel SL, Yao E, Sen TZ. GrainGenes: centralized small grain resources and digital platform for geneticists and breeders. Database (Oxford). 2019;baz065.
40. Maccaferri M, Harris NS, Twardziok SO, et al. Durum wheat genome highlights past domestication signatures and future improvement targets. Nature Genet. 2019;51:885–95.
41. Solovyev V, Kosarev P, Seledsov I, Vorobyev D. Automatic annotation of eukaryotic genes, pseudogenes and promoters. Genome Biol. 2006;7(S1):S10.
42. Potter SC, Luciani A, Eddy SR, Park Y, Lopez R, Finn RD. HMMER web server: 2018 update. Nucleic Acids Res. 2018;46:W200-4.
43. Sim NL, Kumar P, Hu J, Henikoff S, Schneider G, Ng PC. SIFT web server: predicting effects of amino acid substitutions on proteins. Nucleic Acids Res. 2012;40:W452-7.
44. Murashige T, Skoog F. A revised medium for rapid growth and bio assays with tobacco tissue cultures. Physiol Plantarum. 1962;15:473–97.
45. Sambrook J, Fritsch EF, Maniatis T. Molecular cloning: a laboratory manual. CSHL Press. 1989.
46. Chang S, Puryear J, Cairney J. A simple and efficient method for isolating RNA from pine trees. Plant Molec Biol Rep. 1993;11:113–6.
47. Rieu I, Powers SJ. Real-time quantitative RT-PCR: design, calculations, and statistics. Plant Cell 2009;21:1031–3.
48. Ramakers C, Ruijter JM, Deprez RH, Moorman AF. Assumption-free analysis of quantitative real-time polymerase chain reaction (PCR) data. Neurosci. Lett. 2003;339:62–6.
49. Genstat for Windows 21st Edition. VSN International, Hemel Hempstead, UK. 2020. https://Genstat.co.uk.
50. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. 2020. https://www.R-project.org/.
51. Kassambara A. ggpubr: 'ggplot2' Based Publication Ready Plots. R package version 0.4.0. 2020. https://CRAN.R-project.org/package=ggpubr
52. Xiang D, Quilichini TD, Liu Z, Gao P, Pan Y, Li Q, Nilsen KT, Venglat P, Esteban E, Pasha A, Wang Y, Wen R, Zhang Z, Hao Z, Wang E, Wei Y, Cuthbert R, Kochian LV, Sharpe A, Provart N, Weijers D, Gillmor CS, Pozniak C, Datla R. The transcriptional landscape of polyploid wheats and their diploid ancestors during embryogenesis and grain development. Plant Cell. 2019;31:2888–2911.
53. Pearce S, Vazquez-Gross H, Herin SY, Hane D, Wang Y, Gu YQ, Dubcovsky J. WheatExp: an RNA-seq expression database for polyploid wheat. BMC Plant Biol. 2015;15:299.
54. Warnes GR, Bolker B, Bonebakker L, Gentleman R, Huber W, Liaw A, Lumley T, Maechler M, Magnusson A, Moeller S, Schwartz M, Venables B. gplots: Various R programming tools for plotting data2013. R package version 2.12.1. 2013. http://CRAN.R-project.org/package=gplots.
55. Hamilton NE, Ferry M. ggtern: Ternary diagrams using ggplot2. Journal of Statistical Software. 2018;8:3863.
56. H. Wickham. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York. 2016.

Figures

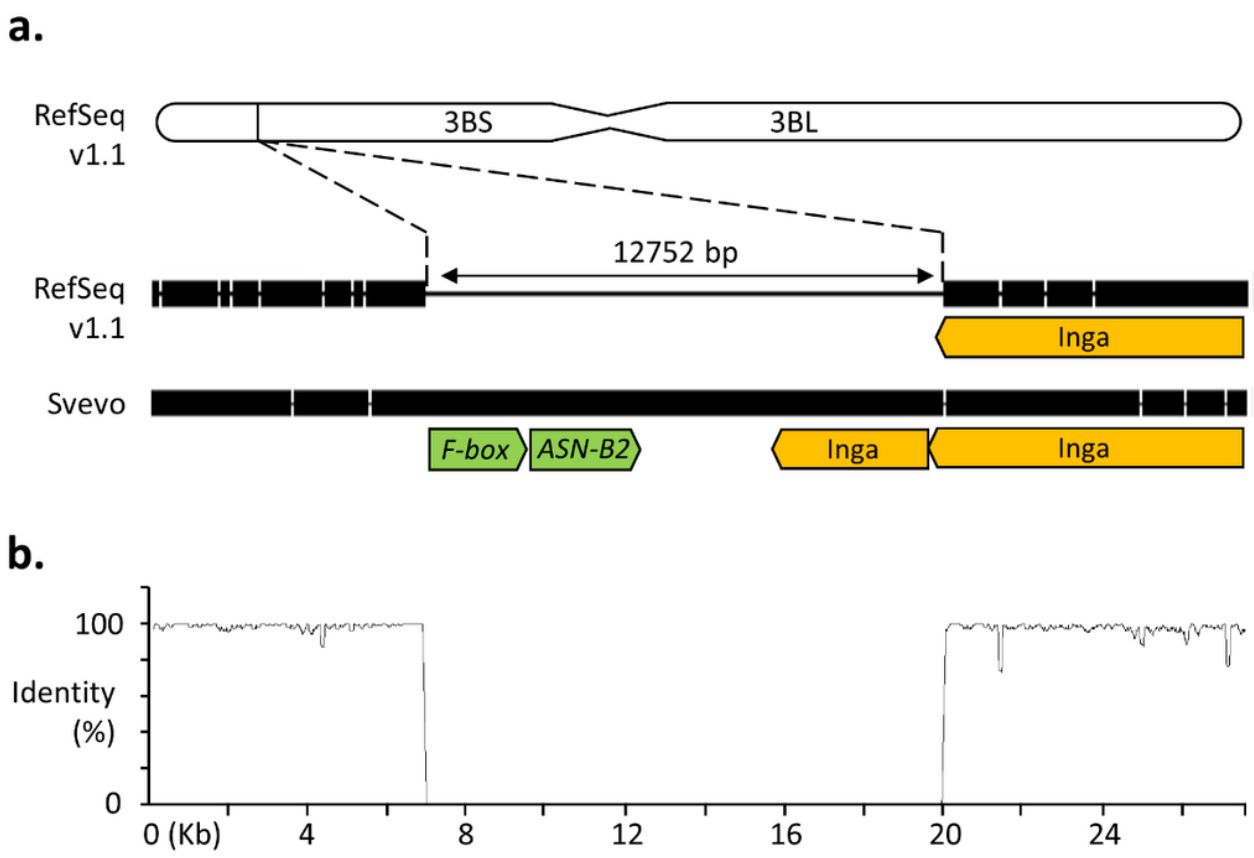

Figure 1

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. 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.

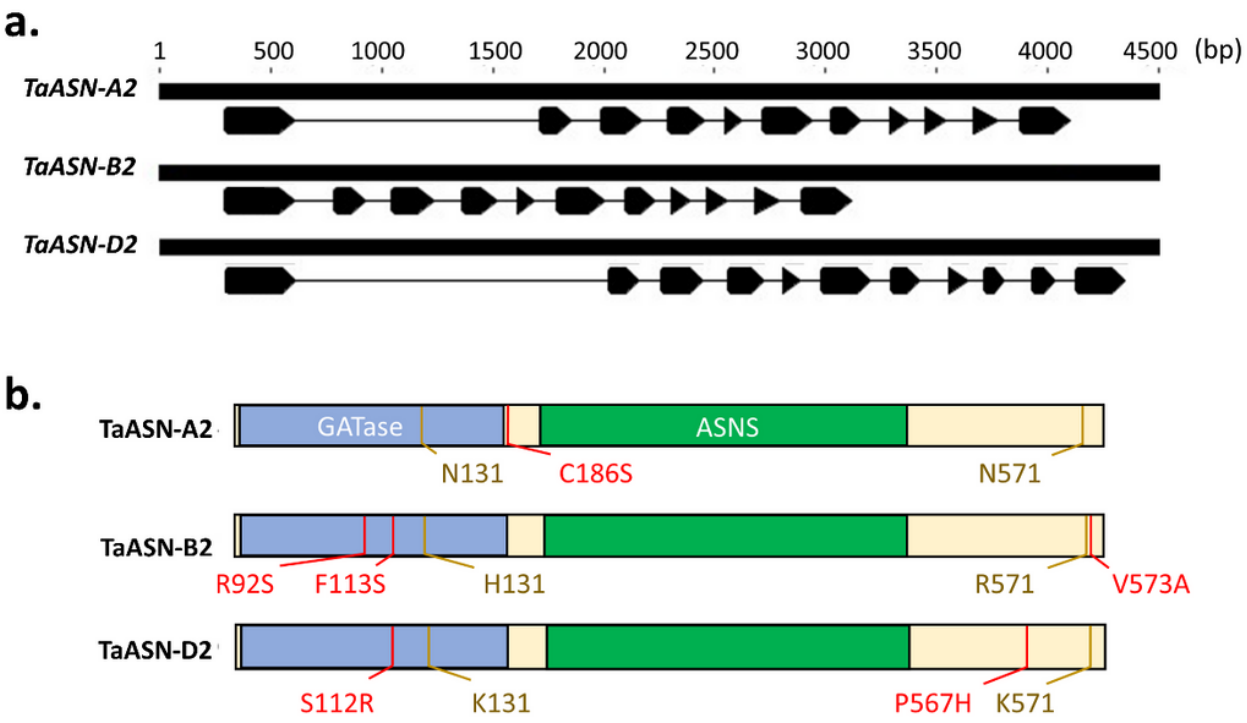

Figure 2

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 [15]. 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 amidotransferases (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.

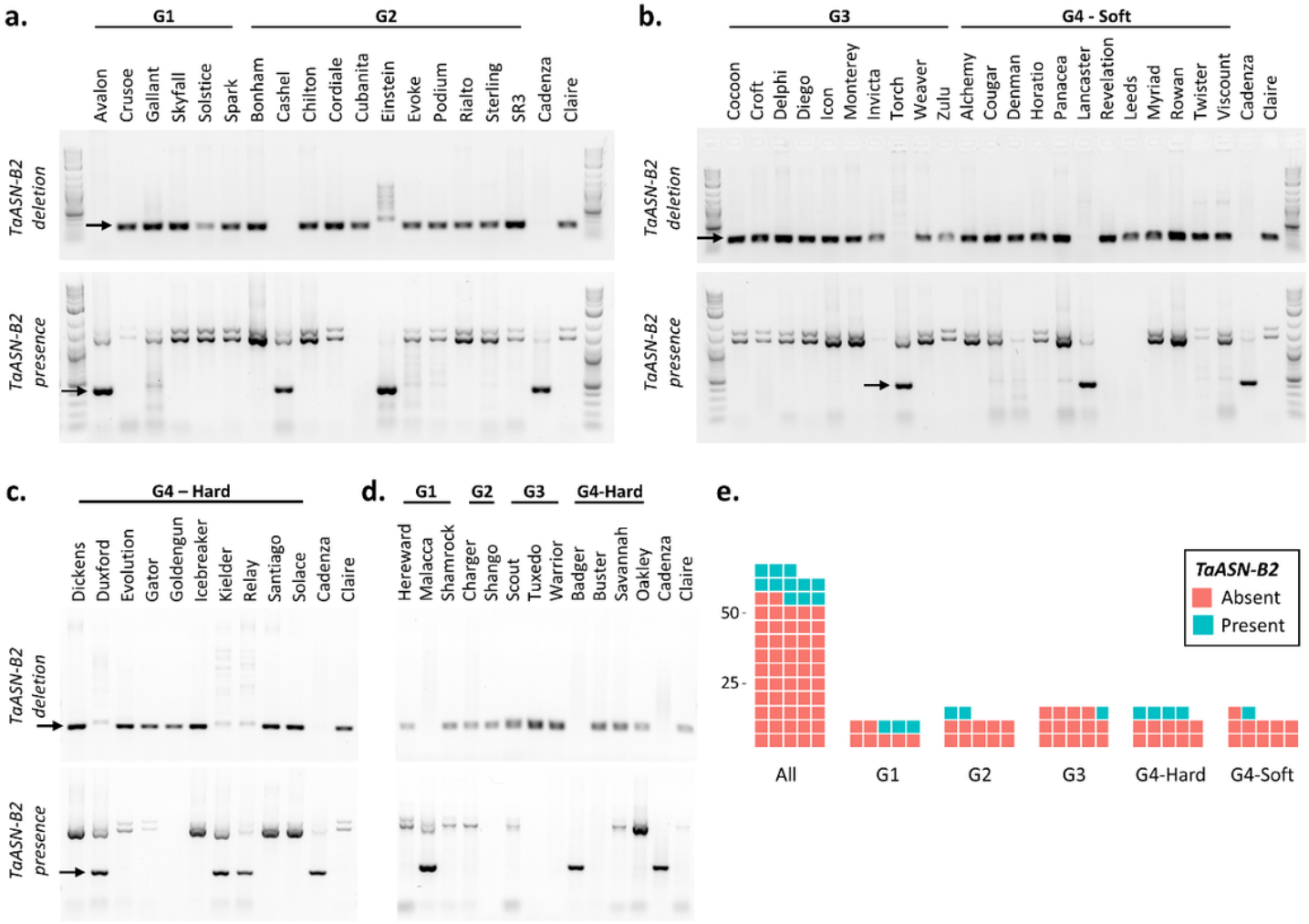

Figure 3

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 [11]. 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).

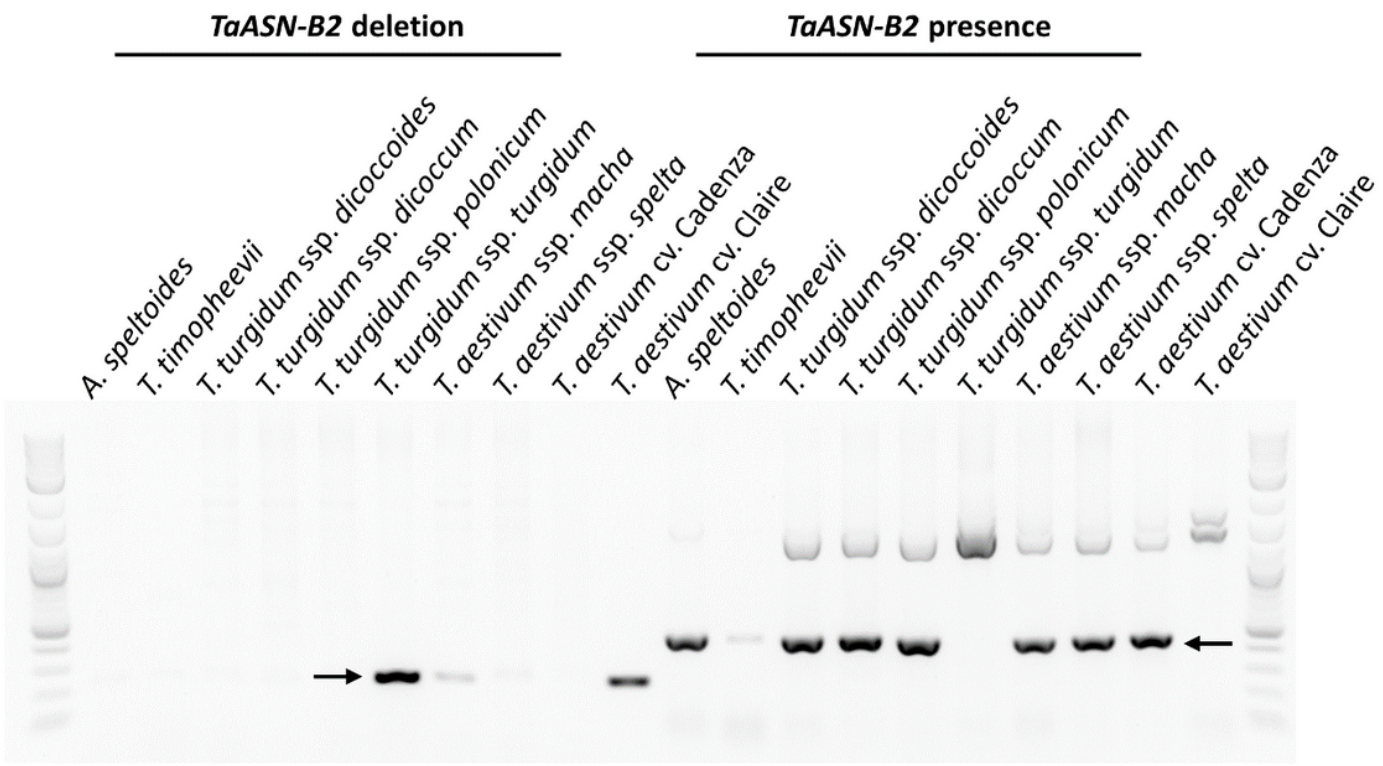

Figure 4

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.

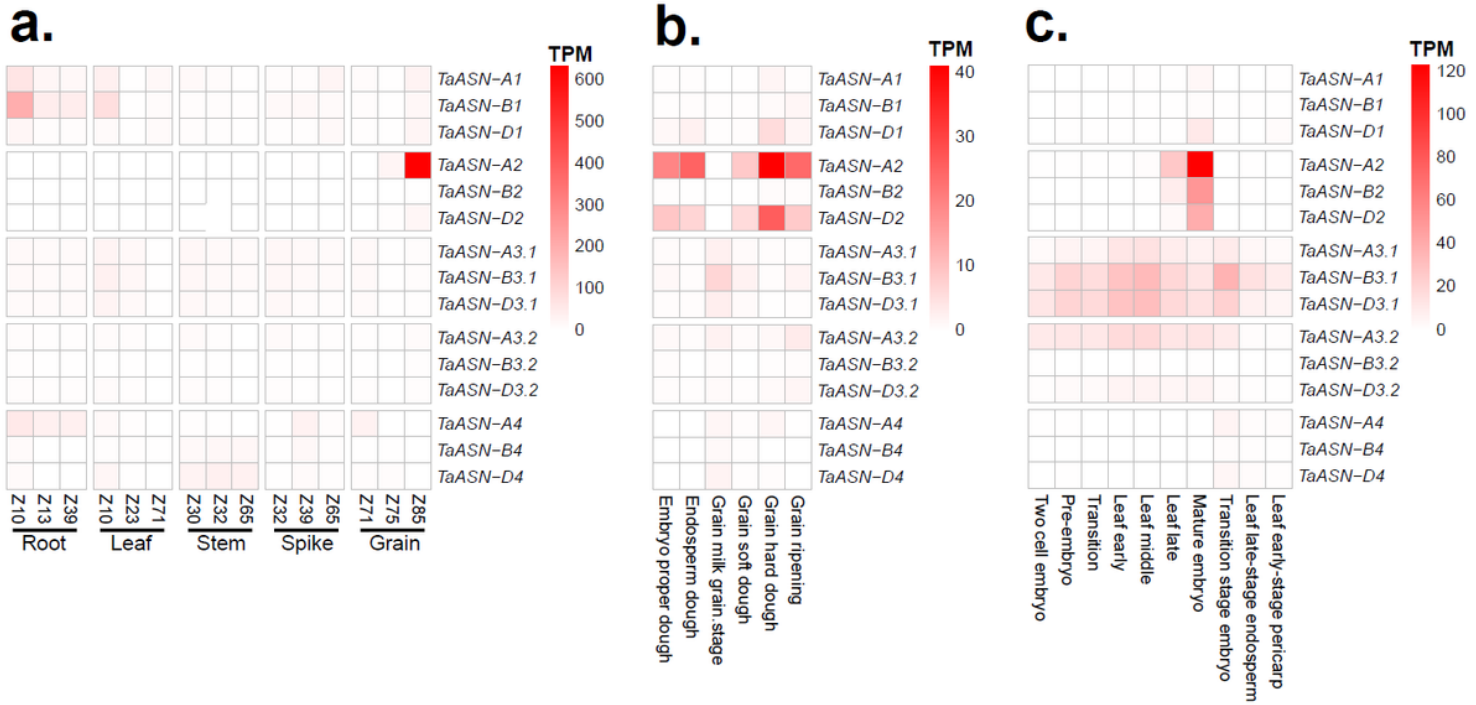

Figure 5

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 [23]. b. Expression of TaASN genes in grain tissues of the variety Azhurnaya across six stages of development [25]. c. Expression of TaASN genes in variety AC Barrie across an embryo development timecourse [55].

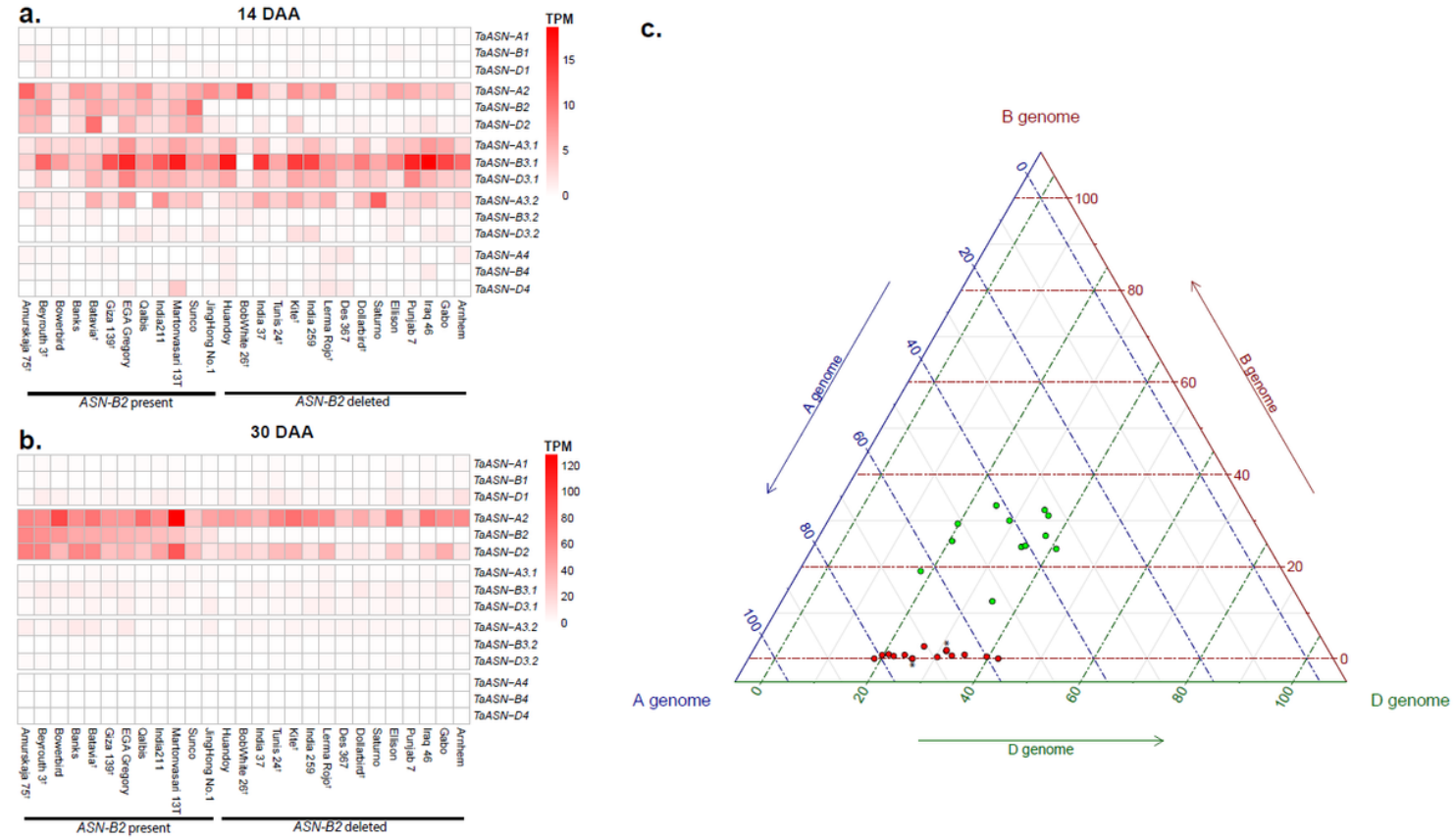

Figure 6

Expression of TaASN genes in 27 worldwide wheat varieties from RNA-seq mapping data [26]. 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.

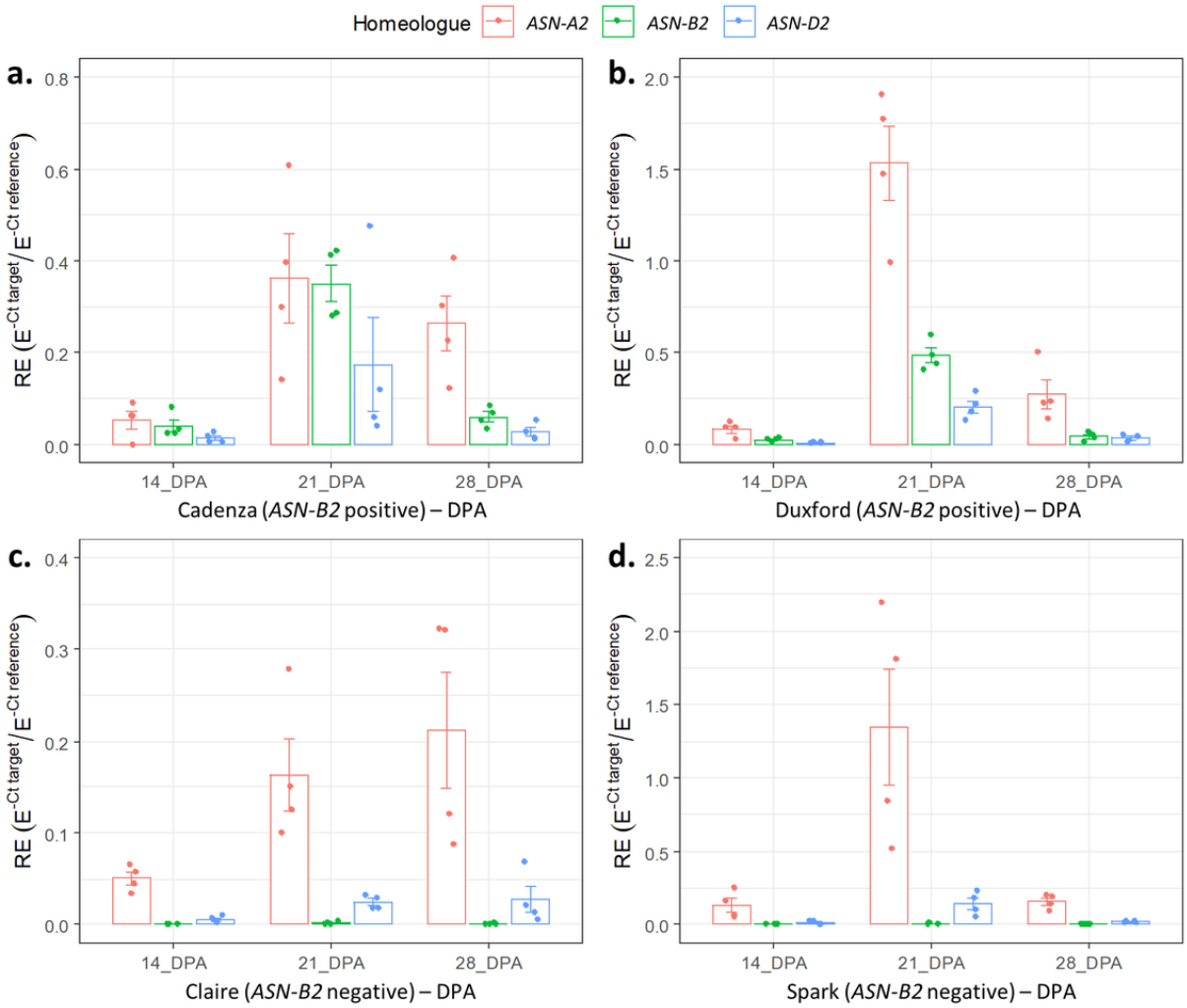

Figure 7

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.

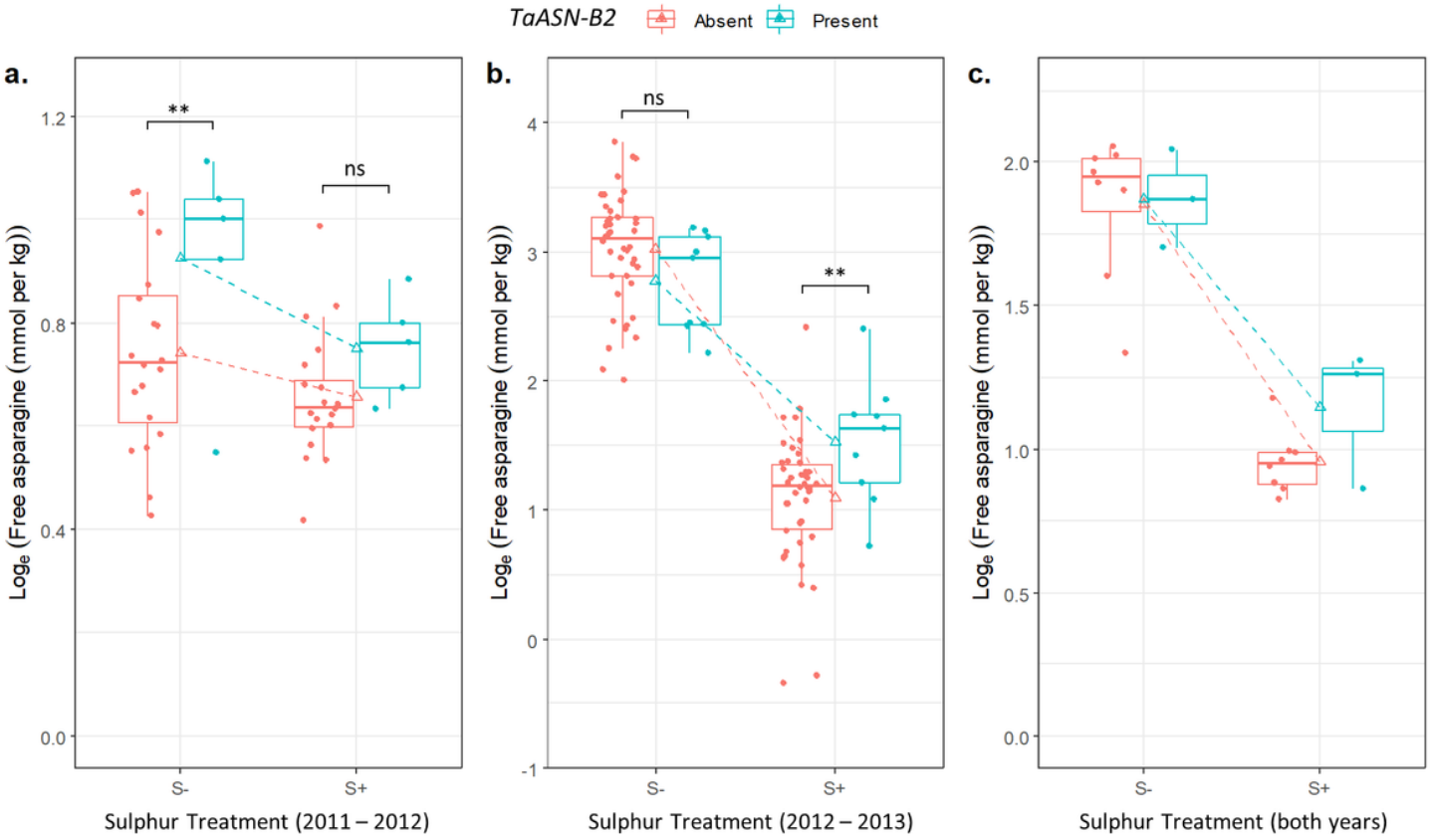

Figure 8

Effect of the TaASN-B2 deletion on grain asparagine levels under different field sulphur conditions. In 2011-2012 (a.), 2012-2013 (b.), and for varieties grown in both years (c.). P-values calculated by post-hoc Bonferroni tests following ANOVA analysis. ns = $p > 0.05$, * = $p < 0.05$, ** = $p < 0.01$, *** = $p < 0.001$. Triangles show the means of each group.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Additionalfile1.pdf
- Additionalfile2.xlsx
- Additionalfile3.pdf