Theoretical and Applied Genetics
https://doi.org/10.1007/s00122-018-3163-x

ORIGINAL ARTICLE

Assessing the variation and genetic architecture of asparagine content in wheat: What can plant breeding contribute to a reduction in the acrylamide precursor?

Matthias Rapp[1] · Klaus Schwadorf[2] · Willmar L. Leiser[1] · Tobias Würschum[1] · C. Friedrich H. Longin[1]

Received: 16 May 2018 / Accepted: 6 August 2018

Abstract

Key message **A large genetic variation, moderately high heritability, and promising prediction ability for genomic selection show that wheat breeding can substantially reduce the acrylamide forming potential in bread wheat by a reduction in its precursor asparagine.**

Abstract Acrylamide is a potentially carcinogenic substance that is formed in baked products of wheat via the Maillard reaction from carbonyl sources and asparagine. In bread, the acrylamide content increases almost linearly with the asparagine content of the wheat grains. Our objective was, therefore, to investigate the potential of wheat breeding to contribute to a reduction in acrylamide by decreasing the asparagine content in wheat grains. To this end, we evaluated 149 wheat varieties from Central Europe at three locations for asparagine content, as well as for sulfur content, and five important quality traits regularly assessed in bread wheat breeding. The mean asparagine content ranged from 143.25 to 392.75 mg/kg for the different wheat varieties, thus underlining the possibility to reduce the acrylamide content of baked wheat products considerably by selecting appropriate varieties. Furthermore, a moderately high heritability of 0.65 and no negative correlations with quality traits like protein content, sedimentation volume and falling number show that breeding of quality wheat with low asparagine content is feasible. Genome-wide association mapping identified few QTL for asparagine content, the largest explaining 18% of the genotypic variance. Combining these QTL with a genome-wide prediction approach yielded a mean cross-validated prediction ability of 0.62. As we observed a high genotype-by-environment interaction for asparagine content, we recommend the costly and slow laboratory analysis only for late breeding generations, while selection in early generations could be based on marker-assisted or genomic selection.

Introduction

Health aspects associated with the consumption of bread wheat (*Triticum aestivum* L.) are of major relevance, as wheat is one of the most important staple crops worldwide (FAOSTAT 2013). In general, bread wheat plays an important role for a healthy diet, as supported by medical studies and recognized by several international organizations like WHO, FAO, and EFSA (cf. Huang et al. 2015). However, acrylamide is formed in potentially harmful concentrations when cereals such as wheat are treated for a long time with heat during the processing to food products (Claus et al. 2008a). Acrylamide can be found in most cereal products such as bread and breakfast cereals, but in particularly high concentrations in crispbread, cookies, and gingerbread (Becalski et al. 2003; Svensson et al. 2003; Sadd and Hamlet 2005). Consequently, a range of methods was developed to reduce the formation of acrylamide during the production process. First of all, the reduction in heat, but also a prolonged fermentation time or the addition of specific ingredients, for example cysteine or asparaginase, has been proposed (Claus et al. 2008a).

Communicated by Xianchun Xia.

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s00122-018-3163-x) contains supplementary material, which is available to authorized users.

✉ C. Friedrich H. Longin
friedrich.longin@uni-hohenheim.de

[1] State Plant Breeding Institute, University of Hohenheim, 70599 Stuttgart, Germany

[2] Core Facility Hohenheim, Module Analytical Chemistry, University of Hohenheim, 70599 Stuttgart, Germany

Published online: 20 August 2018

Another promising approach is to decrease the precursors in the raw material and thereby reduce the potential for the formation of acrylamide. Acrylamide forms mainly from carbonyl sources and free asparagine as part of the Maillard reaction (Mottram et al. 2002; Stadler et al. 2002; Zyzak et al. 2003). As during the production process of cereal products the carbonyl sources are available in excess, asparagine is the limiting precursor for the formation of acrylamide in cereal products (Surdyk et al. 2004). For this reason, studies aiming to reduce the precursors in the raw material of cereals mainly focused on the asparagine content. For wheat, it was shown by experimental field trials that the asparagine content in the kernels drastically increased under sulfur-deficient conditions (Shewry et al. 1983; Muttucumaru et al. 2006). Conversely, however, there was no reduction in the asparagine content in the wheat kernels by an additional application of sulfur fertilizer in non-sulfur-deprived environments (Claus et al. 2006). Studies performed with few wheat varieties grown at several locations reported an effect of the variety on asparagine content, but often also a high effect of the environment (Claus et al. 2006; Corol et al. 2016; Curtis et al. 2018). These studies indicated that variety selection and plant breeding might contribute to lower asparagine concentrations in the raw material, but the number of varieties tested was low and information about the heritability and the genetic architecture of asparagine content in wheat grown under field conditions was lacking. Only one study based on 92 Australian wheat genotypes systematically assessed the heritability and the genetic architecture of asparagine content in wheat and reported a heritability of 0.32 (Emebiri 2014). In addition, this study detected putative QTL on chromosome 5A, which explained between 14 and 24% of the observed variation. However, this study was performed under greenhouse conditions and it is unclear whether the results are transferable to wheat grown under field conditions.

We therefore evaluated 149 bread wheat varieties grown at three locations for asparagine content and investigated the phenotypic variation as well as the heritability of this trait. We further analyzed the association between asparagine and sulfur content, as well as the correlation of asparagine content with other important quality traits to assess possible negative effects of breeding for low asparagine content in bread wheat. Moreover, we explored the genetic architecture of asparagine content in order to evaluate the potential of genomic-assisted breeding.

Materials and methods

Plant material and field trials

The plant material consisted of 149 different European wheat varieties, mainly originating from Germany, the UK, and France. Plants were grown in a winter cropping system, with sowing in October and harvest in July of the following year. The field experiments were performed at three locations in the season 2015/2016. The genotypes were grown in observation plots with two rows of 1 m length, arranged as a partially replicated design with a replication factor of 1.125 (Williams et al. 2011). The locations were Hohenheim (HOH, 48°43′07.3″N 9°11′08.7″E, altitude 403 m, Germany), Oberer Lindenhof (OLI, 48°28′19.0″N 9°18′29.3″E, altitude 700 m, Germany), and Eckartsweier (48°32′52.4″N 7°52′32.5″E, altitude 140 m, Germany).

Laboratory analysis

For the analysis of asparagine and sulfur content, finely ground whole-grain flour samples were produced using a laboratory mill equipped with a 1-mm sieve (Cyclotec 1093, FOSS, Hilleroed, Denmark). The asparagine content was determined according to European Commission Regulation (EC No 152/2009, Annex III, Method F) with moderate modifications. When extracting free natural asparagine from plant matrices using mixtures of hydrochloric acid, there is a risk that under these conditions, a difficult to estimate amount is converted into aspartic acid. The conversion into aspartic acid with elimination of ammonia leads to a non-quantified proportion of asparagine during the measurement procedure. Therefore, as an extracting agent for asparagine the so-called sample dilution buffer, pH 2.20, was used, which is composed of water, methanol, phosphoric acid, formic acid, glacial acetic acid, and caprylic/octanoic acid. The determination of asparagine in the wheat samples was performed as follows: 2 g of the prepared sample material was weighed into a PE screw-cap bottle, 50 mL of sample dilution buffer was added, and the bottle then tightly closed. The mixture was shaken vigorously for 105 min on a laboratory shaker. Thereafter, the content of the bottle was quantitatively transferred in a 100-mL volumetric flask, which was filled up to mark with sample dilution buffer and treated for 15 min in an ultrasonic bath. Following a centrifugation step (10 min with 3100 rpm), the supernatant was decanted into a PE tube and membrane-filtered by syringe filter (cellulose acetate, 0.2 µM). This filtrate was used for chromatography. Asparagine was separated and determined by ion exchange chromatography and ninhydrin post-column derivatization using an amino acid analyzer (Biochrom 30; Biochrom Ltd., Cambridge, UK). For the analysis of asparagine with Biochrom 30, the buffer separation program had to be specially adapted. The recovery for the whole procedure was > 98%, and the lower limit of quantification was about 50 mg/kg asparagine.

Sulfur content was determined by quantitative elemental analysis resp. CHNS analysis using a vario EL cube (Elementar Analysensysteme GmbH, Langenselbold, Germany).

Thirty milligrams of the prepared sample material was inserted into the combustion tube of a furnace (1150 °C) and burned in an excess of oxygen. The elements bound in the sample burn to form gaseous reaction products, e.g., SO_2 and SO_3 for sulfur. In various traps, the individual measuring components in the reaction gas mixture are collected and in further steps (with adsorption columns/tubes, reduction tube), separated from each other. The operational procedure to measure the SO_3 component includes a reduction of SO_3 to SO_2 in the reduction tube followed by heating the adsorption column to SO_2 desorption temperature. The SO_2 is then desorbed, and it enters the infrared detector with the carrier gas (Helium). The resulting measuring signal is finally evaluated via PC. The recovery for the whole procedure was > 90% (corrected by an every working day newly determined factor) and the lower limit of quantification was about 700 mg/kg sulfur.

Total protein content (PC) was determined with near-infrared spectroscopy (NIRS) on the grains (ICC standard method 159, ICC, Vienna, Austria). In a further attempt, we additionally utilized the obtained NIR spectra in order to establish a NIRS calibration for asparagine and sulfur content. We used the ratio of performance to deviation (RPD) to evaluate the goodness of fit of the NIRS models (Williams 1987). For a good prediction of trait values by NIRS models, a high RPD is required, with a value of at least 3 for sample screening (Batten 1998). Protein quality was measured by the sedimentation value in mL according to Zeleny (Z-SDS, ICC standard method 116/1, ICC, Vienna, Austria). The falling number as estimator for the alpha-amylase activity was determined with a 'Perten Falling Number 1800' device (Perten Instruments, Hägersten, Sweden) based on whole-grain flour (ICC standard method 107/1, Vienna, Austria). Thousand kernel weight in g (TKW) was measured using a Marvin seed analyzer (GTA Sensorik, Neubrandenburg, Germany). The hectoliter weight was estimated by the weight of grains fitting in a cylindrical can with a volume of 26 mL as, $\text{HLW} = \frac{\text{grain weight (kg)}}{\text{volume (l)}} \times 100\,\text{l}$.

Molecular data analysis

Dominant silico-DArTs and SNP marker information was obtained for the whole plant material by genotyping-by-sequencing at Diversity Arrays Technology (Yarralumla, Australia) using the Wheat GBS 1.0 assay. The dominant silico-DArTs and codominant single-nucleotide polymorphism (SNP) markers are in the following denoted by their clone ID and the marker-type prefix 'D' or 'S,' respectively. Markers showing more than 25% missing values or a minor allele frequency lower than 5% were removed from the initial marker set. Remaining missing values were imputed by the software package LinkImpute that utilizes LD-kNNi, a k-nearest neighbor genotype imputation (Money et al. 2015). The imputation reached an estimated accuracy of 0.95, which is the proportion of masked known genotypes that were correctly imputed. After the imputation, markers with a minor allele frequency lower than 5% were again removed, resulting in 41,604 markers of which 22,122 markers had a known map position (Li et al. 2015). In a first run, markers with known and unknown map position were included in the association mapping. For unmapped markers showing significant marker–trait association, we assigned a map position, when they were in high linkage disequilibrium with many mapped markers at a similar chromosomal position. Thus, the estimated map position of a marker with unknown position refers to the position of the marker that was in highest linkage disequilibrium and had a known map position. In the final association mapping, we then used 22,158 markers consisting of markers with known map position (22,122) or a map position assigned based on linkage disequilibrium with mapped markers (36). Of the 22,158 markers, 8339 had unique positions on the genetic map that has a total length of 3302.5 cM (Li et al. 2015).

Phenotypic data analysis

The asparagine content data showed a left-side skewed distribution. We applied a logarithmic transformation to the data that led to normally distributed residuals and a considerably increased homogeneity of the residuals. The transformed asparagine data were used in the phenotypic analyses, the association mapping, and the genomic prediction. In case of computations of correlations among traits and estimations of marker effects, we used the back-transformed data. The phenotypic data were analyzed according to the following statistical model:

$$y_{ijk} = \mu + g_i + \text{env}_j + g_i : \text{env}_j + b_{jk} + e_{ijk},$$

where y_{ijk} was the phenotypic observation of the ith genotype tested at the jth environment in the kth incomplete block, μ was the general mean, g_i the genotypic effect of the ith genotype, env_j the effect of the jth environment, $g_i : \text{env}_j$ was the genotype-by-environment interaction, b_{jk} was the effect of the kth block at the jth environment, and e_{ijk} was the residual error. Estimates for variance components were computed by applying the restricted maximum likelihood (REML) method assuming a fully random model in a classical one-stage analysis (Cochran and Cox 1957). A likelihood ratio test with model comparisons was performed to test for significance of variance components (Stram and Lee 1994). We computed best linear unbiased estimates (BLUEs) across all environments assuming fixed genetic effects. The heritability (h^2) was calculated according to the formula proposed by Piepho and Möhring (2007) as $h^2 = 1 - \frac{\vartheta}{2\sigma_G^2}$, where ϑ is the mean variance of a difference of two best linear

unbiased predictors and σ_G^2 the genetic variance. All analyses were based on the statistical software R (R Development Core Team 2018) and the software ASReml 3.0 (Gilmour et al. 2009).

Association mapping and genomic prediction

We performed genome-wide association mapping applying a mixed linear model approach, including a kinship matrix to correct for population structure (Yu et al. 2006). We utilized the function 'polygenic' of the R package GenABEL to obtain *P* values for marker–trait associations. An explorative significance threshold of $P < 0.001$ was used to identify significant marker–trait associations. To correct for collinearity among markers, i.e., several significant markers identifying the same QTL, we fitted all significant markers simultaneously in a linear model in the order of the strength of the marker–trait association, beginning with the marker with the lowest *P* value (Würschum et al. 2015). For the applied additive genetic model, marker information was coded as 0, 1, 2, with 0 and 2 being the two homozygous allelic classes, and the obtained R^2_{adj} was used to calculate the total proportion of explained genetic variance as $p_G = \frac{R^2_{\text{adj}}}{h^2} \times 100\%$ (Würschum et al. 2016). To estimate the p_G of individual markers, the sums of squares from the ANOVA of the linear model including the significant markers were used in the ratio $p_G = \frac{\frac{SS_m}{SS_{\text{total}}}}{h^2} \times 100\%$, where SS_m are the sums of squares of the individual marker and SS_{total} corresponds to the total sums of squares of all markers together. Only the markers explaining more than 0.5% of the genotypic variance in this approach were finally reported as putative QTL.

A genomic prediction approach with the R package 'RR-BLUP' (Endelman 2011; Endelman and Jannink 2012) was utilized to capture the effects of additional additive genetic effects that did not pass the significance threshold of the association mapping. This genomic prediction approach was based on a ridge regression BLUP (RR-BLUP) and a fivefold cross-validation (Würschum et al. 2013, 2014). Additionally, a weighted ridge regression best linear unbiased prediction (wRR-BLUP) approach was applied, including the putative QTL which individually explained more than 10% of the genotypic variance. In contrast to the general marker set, the identified markers explaining more than 10% were modeled as fixed effects (Boeven et al. 2016; Zhao et al. 2014; Spindel et al. 2016). In addition, we used the R package 'BGLR' in order to compare the prediction ability of ridge regression BLUP with further genomic prediction models including Bayesian and reproducing kernel Hilbert space (RKHS) approaches.

Results

In a panel of 149 bread wheat varieties grown at three locations, we assessed the classical wheat quality parameters protein content, sedimentation volume (Z-SDS), falling number (FN), thousand kernel weight (TKW), and hectoliter weight (HLW), and in addition, the asparagine and sulfur content. For all traits, we observed significant genotypic variances and except for sulfur content also a significant genotype-by-environment interaction. The heritability estimates ranged between 0.65 for protein content and 0.90 for TKW (Table 1). Asparagine content showed no strong correlation with any of the other investigated traits, the strongest being the negative correlation with Z-SDS ($r = -0.29$, $P < 0.001$) (Fig. 1). There was no significant correlation between asparagine and sulfur content. Sulfur content, by contrast, showed a significant and positive correlation with protein content ($r = 0.67$, $P < 0.001$).

A wide range of phenotypic values was found for asparagine and sulfur content, ranging from 143.25 to

Table 1 Summary statistics for asparagine content, sulfur content, protein content, sedimentation volume according to Zeleny (Z-SDS), falling number, thousand kernel weight (TKW), and hectoliter weight (HLW)

Parameter	Asparagine[a] (mg/kg)	Sulfur (mg/kg)	Protein content (%)	Z-SDS (mL)	Falling number (s)	TKW (g)	HLW (kg)
Min	143.25	1131.00	10.82	10.18	207.42	28.41	67.87
Mean	241.69	1281.93	12.76	24.09	383.19	44.49	75.20
Max	392.75	1506.09	15.18	45.10	475.92	58.42	79.98
σ_G^2	4.42×10^{-3}***	6433.68***	0.33***	51.64***	1140.38***	18.36***	4.27***
$\sigma_{G\times E}^2$	3.94×10^{-3}***	161.61	0.17*	10.06***	465.95*	4.31***	0.65***
σ_e^2	2.62×10^{-3}	4628.20	0.26	9.86	1037.13	1.40	1.05
h^2	0.65	0.79	0.65	0.88	0.68	0.90	0.88

σ_G^2 genotypic variance, $\sigma_{G\times E}^2$ genotype-by-environment interaction variance, σ_e^2 error variance, h^2 heritability

*, **, ***Significant at the 0.05, 0.01, and 0.001 probability level, respectively

[a]Min, mean, max, values are given in back-transformed unit, variance components in log10-transformed form

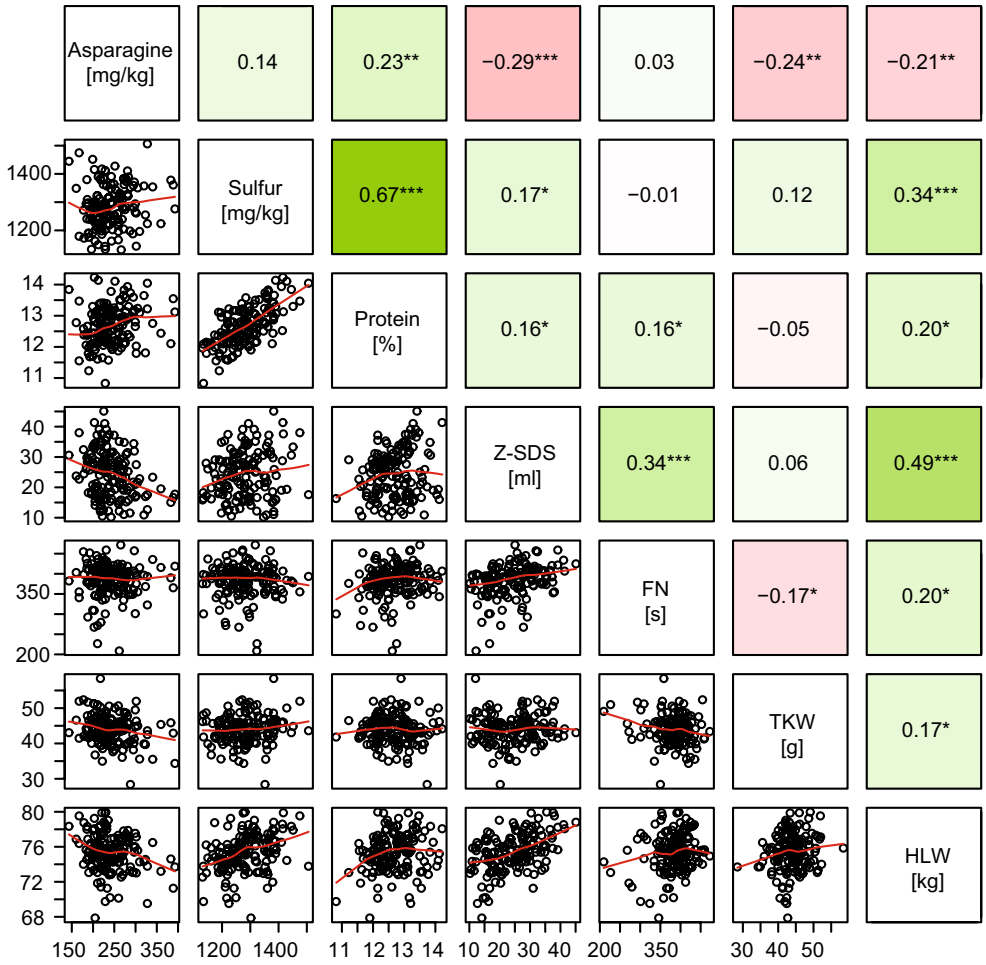

Fig. 1 Trait correlations between asparagine content, sulfur content, protein content, sedimentation volume (Z-SDS), falling number (FN), thousand kernel weight (TKW), and hectoliter weight (HLW) assessed in a diverse bread wheat population. Below the diagonal, the bivariate scatter plots with a locally weighted regression line are shown, and above the diagonal the correlation coefficients (*, **, *** significantly different from zero at $P<0.05$, $P<0.01$, and $P<0.001$, respectively)

392.75 mg/kg and from 1131.00 to 1506.09 mg/kg, respectively (Fig. 2a). The wheat varieties of the diversity panel stem from different decades, ranging from registration dates in the late 1920s until 2013, thus allowing to assess the traits for temporal trends. For asparagine content, we observed no large change across the different decades of variety registration. Accordingly, no trend was observable for the quality traits protein content, falling number, thousand kernel weight, and hectoliter weight (Fig. S1). A slight decrease in sulfur content over the decades (Fig. 2b, Fig. S1) and a slight increase in Z-SDS (Fig. S1) were observable in modern wheat varieties.

We evaluated the potential to predict asparagine and sulfur content by NIRS. To assess the goodness of fit of NIRS models, Williams (1987) suggested using the ratio of performance to deviation (RPD), which is the ratio of the standard error of the prediction to the standard deviation of the samples. In our study, the NIRS calibration for asparagine and sulfur content resulted in a RPD of 2.6 for sulfur content and a RPD of 1.4 for asparagine content.

The genotyping-by-sequencing approach and the subsequent quality checks resulted in 22,122 markers with known map position (Li et al. 2015). A principal coordinate analysis indicated no clear clusters or population structure according to the date of variety registration (Fig. S2). Based on the marker information and the adjusted trait values of each variety, the genome-wide scan identified eight putative QTL each for asparagine and sulfur content (Table 2, Fig. S3, Fig. S4). Except for the QTL on chromosome 6B, which was identified at a similar map position for both traits, there appeared to be no overlapping QTL for asparagine and sulfur content. For asparagine content, the detected QTL jointly explained 78.5% of the genotypic variance, with the QTL on chromosome 7B explaining the largest proportion with 18.0%. The eight QTL detected for sulfur content jointly explained 60.8% genotypic variance and individually up

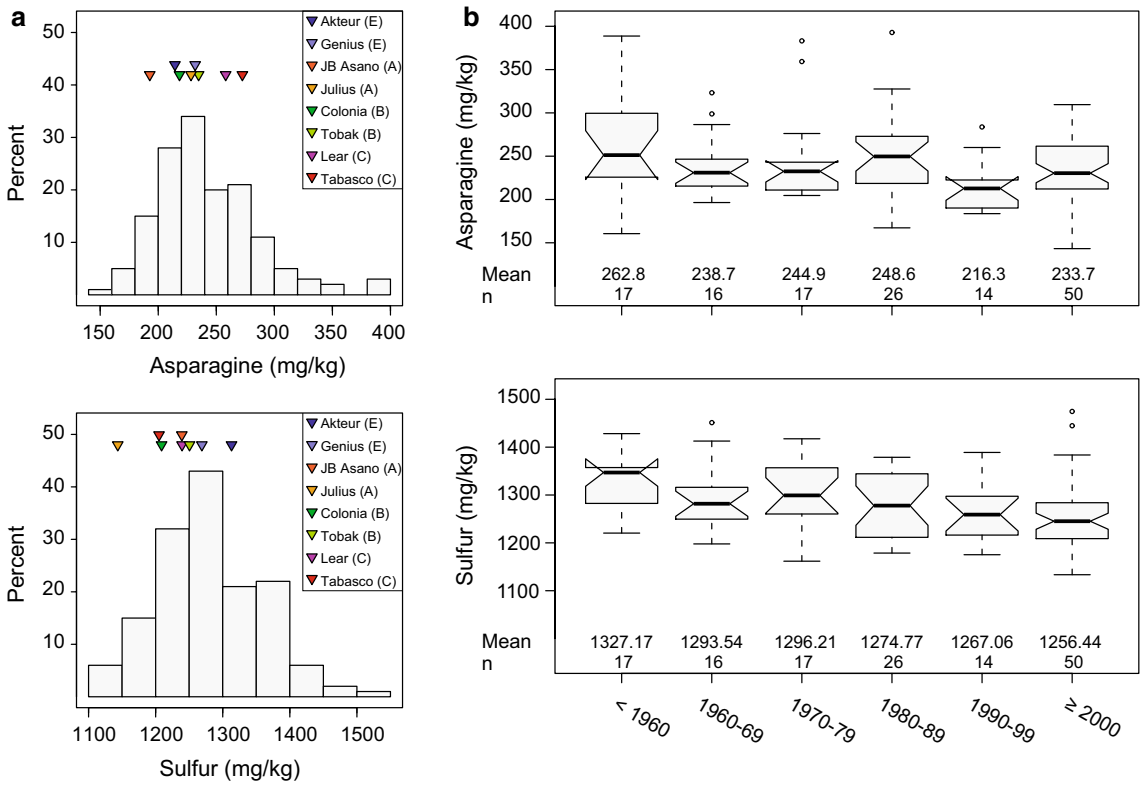

Fig. 2 **a** Histogram of the BLUEs of 149 wheat varieties for asparagine and sulfur content. **b** Box plots showing asparagine and sulfur content dependent on the period of variety registration. Numbers in the plot indicate the mean values and the number of genotypes in each group. Genotypes with lacking registration information were omitted

to 15.9% for a QTL on chromosome 6D. Regarding QTL that individually explained more than 10% of the genotypic variance of asparagine or sulfur content, genotypes carrying the favorable allele at all three QTL showed considerably improved trait values compared to genotypes with the negative alleles at these loci (Fig. 3, Fig. S5).

We next explored the potential of genomic-assisted breeding by utilizing RR-BLUP and wRR-BLUP. For both approaches, this resulted in a higher cross-validated prediction ability for sulfur content than for asparagine content (Fig. 4). The prediction ability of the weighted genome-wide approach that incorporates identified QTL as fixed effects resulted in a substantially higher prediction ability for both asparagine and sulfur content. The prediction ability obtained by RR-BLUP was in a similar range as that obtained by Bayesian or RKHS approaches (Fig. S6).

Discussion

For a large proportion of the world population, bread wheat is an integral part of the daily diet and delivers besides carbohydrates and protein important health-promoting ingredients, especially when consumed as whole grain (cf. Awika 2011; Huang et al. 2015). However, the heat-induced formation of acrylamide from asparagine and carbonyl sources in wheat products is assumed to be a health risk (Claus et al. 2008a). During the bread-making process, the carbonyl sources are available in high amounts and, consequently, asparagine is the precursor limiting the formation of acrylamide in cereal products (Surdyk et al. 2004). Our aim was to explore the potential of wheat breeding to reduce acrylamide by decreasing the content of its precursor asparagine in wheat grains. For this purpose, we investigated the phenotypic variation and the genetic architecture of asparagine content in wheat and evaluated the potential of genomic-assisted breeding of wheat with low asparagine content. Furthermore, we determined the correlation of asparagine content with important quality traits to evaluate the potential to combine low asparagine content with high bread-making quality.

The relation of asparagine and sulfur content in wheat

In our experiment, we observed no significant correlation between asparagine and sulfur content. Previous studies reported a considerably higher accumulation of asparagine

Table 2 Marker–trait associations indicating putative QTL for asparagine and sulfur content

Trait	Marker	Chr.	Pos. (cM)	Phys. Pos. (bp)	MAF	p_G	α-Effect
Asparagine	D1109543	1A	155.92	307,408,774	0.22 (+)	10.92	13.51
	D1097684	4A	250.70	743,752,702[a]	0.14 (+)	16.61	20.25
	D1037340	7A	80.01	54,898,194	0.22 (+)	8.98	− 12.77
	D1202736	1B	104.00	40,806,058	0.09 (+)	9.63	− 24.97
	S1218159	6B	60.43	164,035,370	0.12 (+)	6.39	19.75
	D2322503[b]	7B	249.93	750,604,024	0.44 (−)	17.96	12.58
	D1249605	2D	40.78	19,847,184	0.36 (+)	6.11	13.59
	D1133231	6D	187.31	466,977,892	0.40 (−)	5.89	− 13.23
						78.41	
Sulfur	D1008850	3A	17.38	19,181,539	0.16 (+)	4.66	− 38.5
	D1034244	5A	114.89	536,372,070	0.06 (+)	2.94	− 53.9
	D1233578	6A	24.53	11,008,363	0.09 (+)	4.28	41.6
	D3960474[b]	2B	75.87	58,748,715	0.45 (−)	16.22	− 27.2
	D993029	6B	62.93	470,811,416	0.23 (−)	5.09	26.5
	D1139179	6B	154.09	713,336,444	0.21 (+)	1.30	22.8
	D2290582	7B	10.34	5,513,108	0.27 (+)	14.03	− 31.9
	D3533237[b]	6D	1.136	4,627,044	0.21 (+)	15.86	37.8
						60.83	

Chr. chromosome, *Pos.* chromosome position in cM, Phys. Pos. sequence start position in base pairs according to the bread wheat reference genome (IWGSC RefSeq v1.0), minor allele frequency (MAF) for which the (±) sign indicates an increasing or decreasing effect of the minor allele on the trait, p_G proportion of genotypic variance explained by the QTL in percent, and allele substitution (α) effect in trait unit. The total proportion of explained genotypic variance of all detected QTL is given underneath the line

[a]Marker D1097684 did not produce any hits in the BLASTn; the physical position was estimated with that of the highly collinear marker D1228636

[b]Unmapped marker for which a genetic map position was assigned based on the position of mapped markers in high linkage disequilibrium with it

when wheat was grown under sulfur deficiency (Muttucumaru et al. 2006; Shewry et al. 2009). However, under non-sulfur-deprived growing conditions, an additional application of sulfur had no effect on asparagine content (Claus et al. 2006). The excess of free amino acids, such as asparagine, under sulfur deficiency might be explained by a lower synthesis rate of gluten proteins. The lower protein synthesis is due to a lack of required sulfur-containing amino acids, leading to an excess of the non-sulfur-containing asparagine, which cannot be fully built into proteins (Shewry et al. 1983, 2009). Sulfur content had no effect on asparagine content in our trial performed at non-sulfur-deprived locations, illustrating that sulfur content is not generally associated with asparagine content, but may only affect it under certain environmental conditions. Nevertheless, it is possible that varieties with a high sulfur content bound in proteins are more responsive to sulfur deficiency and show a stronger increase in asparagine content under such conditions. This might explain the change of the variety ranking for asparagine content under sulfur deficiency reported by Curtis et al. (2018) and requires further research.

Selection for low asparagine content is possible

We observed a wide range of phenotypic values leading to a significant genotypic variance, which is the basis for reducing the asparagine content in wheat grains by plant breeding. The heritability estimate of asparagine content was at the level of that of protein content, so that a response to selection comparable to protein content can be expected (Table 1). The ratio between the genotypic variance and the genotype-by-environment interaction variance was 1.12 for asparagine content and thus much lower than for other quality traits. This is in line with previous studies that also reported a significant genotype-by-environment interaction for asparagine content and substantiates the strong effect of the environment on this trait (Curtis et al. 2009; Corol et al. 2016). Consequently, evaluation of wheat varieties and selection on asparagine content in breeding programs should be based on multi-location field trials in the target environments.

Asparagine content showed no strong correlation with any of the other evaluated traits, which included the most important quality traits in bread wheat breeding. The sedimentation volume (Z-SDS), as an estimator for baking

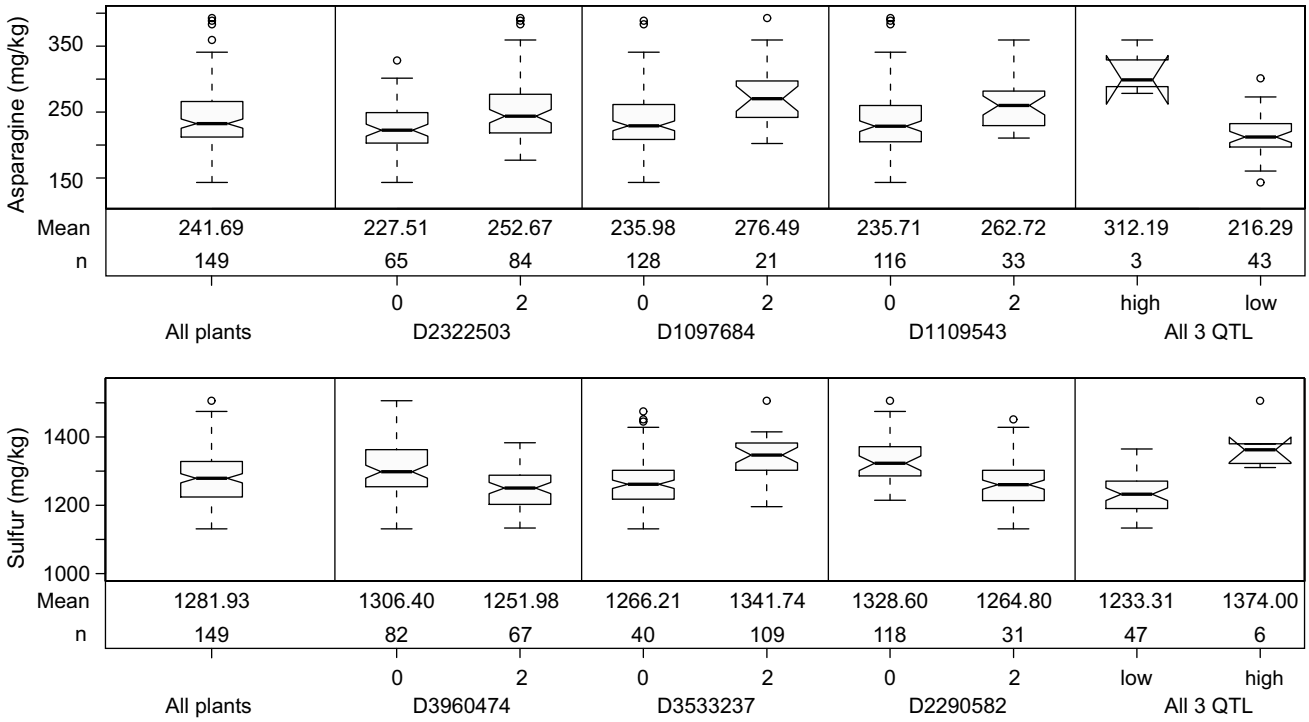

Fig. 3 Box plots showing asparagine and sulfur content of genotypes carrying different alleles for the detected putative QTL explaining more than 10% of the genotypic variance. The two rightmost box plots show trait values for genotypes which carry the high or low allele at all three QTL. The numbers underneath the box plots indicate the mean value and the number (n) of genotypes in each group

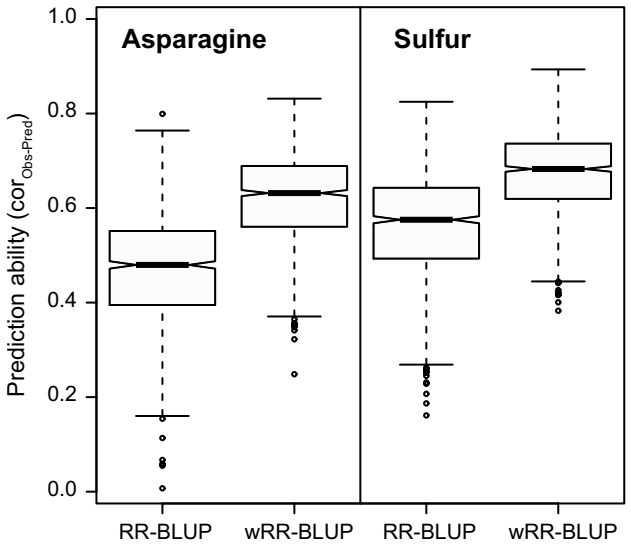

Fig. 4 Prediction ability from 1000 runs of a fivefold cross-validation in the bread wheat panel ($n = 149$) for asparagine and sulfur content for genome-wide prediction with ridge regression BLUP (RR-BLUP), or weighted ridged regression BLUP (wRR-BLUP) including the detected QTL explaining more than 10% of the genotypic variance as fixed effects in the model

quality, showed the highest correlation with asparagine content, which, however, was still weak ($r = -0.29$, Fig. 1). An experiment with 150 bread wheat varieties grown at one Hungarian location also reported only a weak correlation between asparagine content and Z-SDS (Corol et al. 2016). Thus, both results contradict conclusions of previous studies that higher quality in bread wheat would lead to increased levels of asparagine (Claus et al. 2006, 2008a). In this context, it is important to mention that these previous reports considered the protein content as quality criterion, but even for protein and asparagine content we did not find a strong correlation. In addition, a recent study by Curtis et al. (2018), which analyzed different quality groups of bread wheat, also observed no significant correlation between wheat quality and asparagine content. Thus, the absence of strong correlations with any of the other investigated traits suggests that it is possible to integrate a selection on low asparagine content in breeding programs, without negative effects on the commonly selected quality traits.

The large range of asparagine content observed among different varieties illustrates that the choice of variety can have a tremendous effect on asparagine content, and thus, on the potential to form acrylamide in baked products (Table 1, Fig. 2a). When the minimum and maximum values are considered, a reduction of 64% in asparagine is achievable just by choosing an appropriate wheat variety. This is more effective than the addition of reducing additives such as cysteine and comparable to reductions achieved by a prolonged fermentation process during dough preparation (Claus et al. 2008b).

We hypothesize that a further reduction is possible by targeted crosses among lines with low asparagine content and pyramiding of QTL alleles for low asparagine. Thus, breeding can contribute substantially to reduce acrylamide and its potentially negative health aspects in wheat end products.

Asparagine and sulfur content in old and modern varieties

Comparing varieties from different registration periods revealed no change in their asparagine content (Fig. 2b). This is in accordance with observations by Corol et al. (2016) who reported only a weak correlation between asparagine content and the year of variety registration. Thus, breeding efforts over the last decades have not substantially altered the asparagine content of bread wheat varieties. As it has not been directly selected for this trait, this shows that asparagine content has also not been the target of indirect selection, which substantiates the conclusion of it not being correlated with other agronomic traits selected for in breeding programs.

For sulfur content, a slight tendency toward lower levels in the more recently registered varieties was observed (Fig. 2b). The slight tendency toward lower sulfur content in the more recently registered varieties requires validation, but may be associated with selection for quality or yield, as previous studies reported a weak-to-moderate negative correlation between sulfur content and grain yield in bread wheat and spelt (Zhao et al. 2009; Rapp et al. 2017).

How to breed for low asparagine content?

Our results illustrate that the breeding progress of the past decades has not altered the average asparagine content of wheat varieties. However, our results also underline that plant breeding could contribute substantially to a lower acrylamide content in end products by developing varieties with reduced asparagine content. A particularly high acrylamide content is formed during the production of crisp bread, breakfast cereals, cookies, and gingerbread (Becalski et al. 2003; Svensson et al. 2003; Sadd and Hamlet 2005), and these products generally do not require the same quality properties demanded for normal bread making. For most of these products, manufacturers usually use the soft wheat classes according to the UK and US quality classifications (Nabim 2018; U.S. Wheat Associates 2018), which approximately correspond to the Ck quality group or sometimes to the B quality group of the German system. An initial breeding program for low asparagine varieties should therefore focus on those quality groups, but the extension to all quality groups is recommended.

The moderately high heritability and the large genotype-by-environment interaction imply that breeding decisions should only be based on multi-site testing. Furthermore, analysis of asparagine is much more complex than for the quality traits routinely assessed in bread wheat, leading to higher costs and a considerably longer time required per analyzed sample. Consequently, we recommend the laboratory analysis of asparagine content for the few candidates in late breeding generations. This implies, however, the risk that the most promising candidates for low asparagine content are already rejected in early generations. A rapid method such as NIRS, which allows a rough testing of many samples in a relatively short time, would therefore be extremely valuable. However, in our case the NIRS calibration yielded a too low RPD of only 1.4 for asparagine content. This value might be increased with a broader base of samples, but the recommended threshold of a RPD of at least 3 for a sample screening does not appear to be readily achievable (Batten 1998).

Alternatively, knowledge about the genetic architecture might allow the utilization of genomic tools to assist selection on asparagine content in early generations. In our genome-wide association mapping, we detected eight putative QTL for asparagine content, which jointly explained 78.5% of the genetic variance (Table 2, Fig S2). Among these, a putative QTL detected on chromosome 7B explained the highest proportion of genotypic variance with 18.4%. The detected QTL for asparagine content were different from those reported in an Australian association mapping study, which was based on a greenhouse experiment (Emebiri 2014). This is in line with the strong effect of the environment on this trait and suggests that identification of QTL to assist breeding should be based on phenotypic data obtained in the target environments. Marker technologies that enable a high throughput can be utilized to screen a high number of lines at low costs. The mean value of genotypes carrying the favorable low asparagine content allele at the three putative QTL explaining more than 10% of the genotypic variance was 216.3 mg/kg as compared to 277.7 mg/kg for genotypes carrying at least two unfavorable high asparagine content alleles at these loci (Fig. 3, Fig. S5). Thus, if these putative QTL for asparagine content can be confirmed, they might be used for selection of genotypes with low asparagine content in early generations.

In general, the absence of major QTL and the phenotypic distribution point toward a quantitative inheritance of asparagine content with a complex genetic architecture. We therefore also evaluated the potential of genomic selection to reduce asparagine content, which resulted in prediction abilities comparable to other quantitative traits in bread wheat (Boeven et al. 2016; Würschum et al. 2016). In our study, the combination of identified QTL with genomic selection in a weighted RR-BLUP minimized the variation of the prediction ability and maximized the mean prediction ability to 0.62 (Fig. 4). This is a level for which simulation studies on optimum breeding schemes showed a significant

increase in the annual response to selection when genomic prediction was included in the breeding scheme (Marulanda et al. 2016). In addition, for breeding programs already using genomic prediction tools it would mean a negligible extra effort to consider asparagine content as an additional trait. Thus, if the monetary and logistic resources of the breeding scheme allow the implementation of genomic selection, it might significantly enhance the selection gain for low asparagine content in wheat.

In summary, a promising strategy to breed for low asparagine content would start by determining the asparagine content of a range of varieties and breeding lines in multi-location field trials, and subsequent crossing of genotypes with low asparagine content followed by recurrent selection. Marker-assisted or genomic selection appears promising to assist selection in early generations. In later breeding generations, the asparagine content can be assessed by laboratory analysis as a final selection criterion, as well as to improve the prediction accuracy of genomic selection. It is also conceivable that institutions for variety registration might be persuaded to include the analysis of asparagine content as an additional trait in their registration trials. Such trials could deliver reliable data from a high number of locations, and in case of a successful variety registration, these data would be of high value to guide the food industry in their choice of variety. Finally, the demand for wheat varieties with low asparagine content will play a crucial role for the possible contribution of plant breeding. Breeding companies are only able to run costly programs for developing low asparagine wheat varieties when there is the prospect of an economic benefit from such varieties, either through higher achievable prices or through a gain in market share.

Author contribution statement CFHL and TW designed the study; KS and WLL collected the data; MR performed the analyses; MR, KS, CFHL, and TW wrote the paper; all coauthors approved the final manuscript.

Acknowledgements We thank the International Wheat Genome Sequencing Consortium (http://www.wheatgenome.org) for providing pre-publication access to IWGSC RefSeq v1.0 to substantiate our results from the genome-wide association mapping with estimated physical map positions. The financial support of Deutsche Forschungsgemeinschaft is highly acknowledged (DFG LO 1816 – 4/1).

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Ethical standard The authors declare that the experiments comply with the current laws of Germany.

References

Awika JM (2011) Major cereal grains production and use around the world. In: Awika JMA, Piironen V, Bean S (eds) Advances in cereal science: implications to food processing and health promotion. Soil & Crop Science Dept./Nutrition and Food Science Dept., Texas A&M University, 2474 TAMU, TX

Batten GD (1998) Plant analysis using near infrared reflectance spectroscopy: the potential and the limitations. Aust J Exp Agric 38:697–706

Becalski A, Lau BPY, Lewis D, Seaman SW (2003) Acrylamide in foods: occurrence, sources, and modeling. J Agric Food Chem 51:802–808

Boeven PHG, Longin CFH, Leiser WL, Kollers S, Ebmeyer E, Würschum T (2016) Genetic architecture of male floral traits required for hybrid wheat breeding. Theor Appl Genet 129:2343–2357

Claus A, Schreiter P, Weber A, Graeff S, Herrmann W, Claupein W et al (2006) Influence of agronomic factors and extraction rate on the acrylamide contents in yeast-leavened breads. J Agric Food Chem 54:8968–8976

Claus A, Carle R, Schieber A (2008a) Acrylamide in cereal products: a review. J Cereal Sci 47:118–133

Claus A, Mongili M, Weisz G, Schieber A, Carle R (2008b) Impact of formulation and technological factors on the acrylamide content of wheat bread and bread rolls. J Cereal Sci 47:546–554

Cochran WG, Cox GM (1957) Experimental designs, 2nd edn. Wiley, New York

Corol DI, Ravel C, Rakszegi M, Charmet G, Bedo Z, Beale MH et al (2016) 1H-NMR screening for the high-throughput determination of genotype and environmental effects on the content of asparagine in wheat grain. Plant Biotechnol J 14:128–139

Curtis TY, Muttucumaru N, Shewry PR, Parry MAJ, Powers SJ, Elmore JS et al (2009) Effects of genotype and environment on free amino acid levels in wheat grain: implications for acrylamide formation during processing. J Agric Food Chem 57:1013–1021

Curtis TY, Powers SJ, Wang R, Halford NG (2018) Effects of variety, year of cultivation and sulphur supply on the accumulation of free asparagine in the grain of commercial wheat varieties. Food Chem 239:304–313

Emebiri LC (2014) 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 94:1422–1429

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

Endelman JB, Jannink J-L (2012) Shrinkage estimation of the realized relationship matrix. G3 Genes Genomes Genet 2:1405–1413

FAOSTAT (2013) Food and agriculture organization of the united nations. Crop. Natl. Prod. http://www.fao.org/faostat/en/#data/FBS. Accessed 27 Apr 2018

Gilmour AR, Gogel BJ, Cullis BR, Thompson R (2009) ASReml user guide release 3.0. VSN International Ltd, Hemel Hempstead, UK

Huang T, Xu M, Lee A, Cho S, Qi L (2015) Consumption of whole grains and cereal fiber and total and cause-specific mortality: prospective analysis of 367,442 individuals. BMC Med 13:1–9

Li H, Vikram P, Singh RP et al (2015) A high density GBS map of bread wheat and its application for dissecting complex disease resistance traits. BMC Genom 16:216

Marulanda JJ, Mi X, Melchinger AE, Xu J-L, Würschum T, Longin CFH (2016) Optimum breeding strategies using genomic selection for hybrid breeding in wheat, maize, rye, barley, rice and triticale. Theor Appl Genet 129:1901–1913

Money D, Gardner K, Migicovsky Z, Schwaninger H, Zhong G-Y, Myles S (2015) LinkImpute: fast and accurate genotype

imputation for non-model organisms. G3 Genes Genomes Genet 5:2383–2390
Mottram DS, Wedzicha BL, Dodson AT (2002) Acrylamide is formed in the Maillard reaction. Nature 419:448–449
Muttucumaru N, Halford NG, Elmore JS, Dodson AT, Parry M, Shewry PR et al (2006) Formation of high levels of acrylamide during the processing of flour derived from sulfate-deprived wheat. J Agric Food Chem 54:8951–8955
National Association of British & Irish Millers (2018) Wheat guide 2018. Nabim, London, pp 1–3
Piepho H-P, Möhring J (2007) Computing heritability and selection response from unbalanced plant breeding trials. Genetics 177:1881–1888
R Development Core Team (2018) R: a language and environment for statistical computing. http://www.r-project.org. Accessed 27 Apr 2018
Rapp M, Beck H, Gütler H, Heilig W, Starck N, Römer P et al (2017) Spelt: agronomy, quality, and flavor of its breads from 30 varieties tested across multiple environments. Crop Sci 57:739–747
Sadd P, Hamlet C (2005) The formation of acrylamide in UK cereal products. In: Friedman M, Mottram D (eds) Chemistry and safety of acrylamide in food. Advances in experimental medicine and biology. Springer, Boston, pp 415–429
Shewry PR, Franklin J, Parmar S, Smith SJ, Miflin BJ (1983) The effects of sulphur starvation on the amino acid and protein compositions of barley grain. J Cereal Sci 1:21–31
Shewry PR, Zhao FJ, Gowa GB, Hawkins ND, Ward JL, Beale MH et al (2009) Sulphur nutrition differentially affects the distribution of asparagine in wheat grain. J Cereal Sci 50:407–409
Spindel JE, Begum H, Akdemir D, Collard B, Redoña E, Jannink J, McCouch S (2016) Genome-wide prediction models that incorporate de novo GWAS are a powerful new tool for tropical rice improvement. Heredity 116:395–408
Stadler RH, Blank I, Varga N, Robert F, Hau J, Guy PA et al (2002) Acrylamide from Maillard reaction products. Nature 419:449–450
Stram DO, Lee JW (1994) Variance components testing in the longitudinal mixed effects model. Biometrics 50:1171–1177
Surdyk N, Rosén J, Andersson R, Åman P (2004) Effects of asparagine, fructose, and baking conditions on acrylamide content in yeast-leavened wheat bread. J Agric Food Chem 52:2047–2051
Svensson K, Abramsson L, Becker W, Glynn A, Hellenäs KE, Lind Y, Rosén J (2003) Dietary intake of acrylamide in Sweden. Food Chem Toxicol 41:1581–1586
U.S. Wheat Associates (2018) uswheat.org. In: Wheat classes. http://www.uswheat.org/wheatClasses. Accessed 23 Apr 2018
Williams PC (1987) Variables affecting near-infrared reflectance spectroscopic analysis. In: Near-infrared technology in the agricultural and food industries. American Association of Cereal Chemists Inc., Saint Paul
Williams E, Piepho HP, Whitaker D (2011) Augmented p-rep designs. Biometrical J 53:19–27
Würschum T, Reif JC, Kraft T, Janssen G, Zhao Y (2013) Genomic selection in sugar beet breeding populations. BMC Genet 14:85
Würschum T, Abel S, Zhao Y (2014) Potential of genomic selection in rapeseed (*Brassica napus* L.) breeding. Plant Breed 133:45–51
Würschum T, Langer SM, Longin CFH (2015) Genetic control of plant height in European winter wheat cultivars. Theor Appl Genet 128:865–874
Würschum T, Leiser WL, Kazman E, Longin CFH (2016) Genetic control of protein content and sedimentation volume in European winter wheat cultivars. Theor Appl Genet 129:1685–1696
Yu J, Pressoir G, Briggs WH et al (2006) A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. Nat Genet 38:203–208
Zhao FJ, Su YH, Dunham SJ, Rakszegi M, Bedo Z, McGrath SP, Shewry PR (2009) Variation in mineral micronutrient concentrations in grain of wheat lines of diverse origin. J Cereal Sci 49:290–295
Zhao Y, Mette MF, Gowda M, Longin CFH, Reif JC (2014) Bridging the gap between marker-assisted and genomic selection of heading time and plant height in hybrid wheat. Heredity 112:638–645
Zyzak DV, Sanders RA, Stojanovic M, Tallmadge DH, Eberhart BL, Ewald DK et al (2003) Acrylamide formation mechanism in heated foods. J Agric Food Chem 51:4782–4787