Published March 3, 2017

RESEARCH

Spelt: Agronomy, Quality, and Flavor of Its Breads from 30 Varieties Tested across Multiple Environments

Matthias Rapp, Heinrich Beck, Hermann Gütler, Wendelin Heilig, Norbert Starck, Peter Römer, Catherine Cuendet, Friedrich Uhlig, Hannes Kurz, Tobias Würschum, and C. Friedrich H. Longin*

ABSTRACT

Spelt (*Triticum aestivum* L. ssp. *spelta*) is an old hulled wheat currently receiving renewed interest of consumers, bakers, millers, and farmers. Our objectives were (i) to assess the genetic variability and heritability of agronomic and quality traits together with the flavor and odor of breads, (ii) to investigate correlations among these traits, and (iii) to draw conclusions for spelt breeding targeting improved yield, quality, and flavor of end products. Therefore, we investigated 30 spelt varieties in up to six field locations and determined important agronomic parameters and numerous quality traits, as well as flavor and odor of breads made from all varieties. As for the closely related bread wheat (*T. aestivum* L. ssp. *aestivum*), protein and gluten content were tightly correlated in spelt, but both were only moderately correlated with protein quality. The correlation between the sedimentation volume determined either with sodium dodecyl sulfate (SDSS) or according to Zeleny was very high ($r = 0.94$, $p < 0.001$). However, SDSS differentiated the spelt varieties better and is thus proposed for future spelt breeding and evaluations. We determined a significant genetic variation for bread flavor with a heritability of 0.56. Furthermore, this flavor was not correlated with traits important for spelt breeding, i.e., protein quality and agronomy. Thus, future breeding can simultaneously target improved yield, bread-making quality, and a more aromatic bread flavor of new spelt varieties, which would also be of interest for bread and durum wheat (*Triticum durum* Desf.). These findings highlight the need of an intensified interdisciplinary research to develop faster methods for flavor and odor evaluation of breads.

M. Rapp, T. Würschum, C.F.H. Longin, State Plant Breeding Institute, Univ. of Hohenheim, 70599 Stuttgart, Germany; H. Beck, BeckaBeck, 72587 Römerstein, Germany; H. Gütler, Stelzenmühle, 88410 Bad Wurzach, Germany; W. Heilig, Kreislandwirtschaftsamt, 72525 Münsingen, Germany; N. Starck, Pflanzenzucht Oberlimpurg, 74523 Schwäbisch Hall, Germany; P. Römer, Südwestdeutsche Saatzucht GmbH & Co. KG, 76437 Rastatt, Germany; C. Cuendet, Getreidezüchtung Peter Kunz GmbH, 64287 Darmstadt, Germany; H. Kurz, State Institute of Agricultural Chemistry, Univ. of Hohenheim, 70599 Stuttgart, Germany; F. Uhlig, Saaten Zentrum Schöndorf, 99427 Weimar, Germany. Received 17 May 2016. Accepted 22 Nov. 2016. *Corresponding author (friedrich.longin@uni-hohenheim.de). Assigned to Associate Editor Esten Mason.

Abbreviations: Ale, Aschersleben; BLUE, best linear unbiased estimate; Das, Darmstadt; Hoh, Hohenheim; Ra, Rastatt; SDSS, sodium dodecyl sulfate micro-sedimentation test; Sha, Schwäbisch Hall; Wei, Weimar; Z-SDS, sedimentation volume according to Zeleny; σ^2_G, genotypic variance; $\sigma^2_{G \times E}$, variance due to genotype-by-environment interaction.

Spelt (*Triticum aestivum* L. ssp. *spelta*) is an old-world hulled wheat that currently attracts renewed interest as food and feed grain (Campbell, 1997; Longin et al., 2016). Spelt was one of the major cereals of the Alemannians in Southern Germany, Austria, and Switzerland between the 12th and 19th century (Miedaner and Longin, 2016) and was also grown in tens of thousands of hectares in the United States. In the last centuries, however, spelt was replaced by the higher-yielding bread wheat (*T. aestivum* L. ssp. *aestivum*). Beside its low yield, the grains from spelt are tightly enclosed by tough glumes, thus requiring a special dehulling treatment in the mill to separate the chaff from the grain, another reason for its replacement by bread wheat (Longin et al., 2016). However, in the last three decades, spelt enjoys growing demand from consumers, bakers, and farmers, having led to >100,000 ha grown especially in Germany and neighboring countries.

Published in Crop Sci. 57:739–747 (2017).
doi: 10.2135/cropsci2016.05.0331

© Crop Science Society of America | 5585 Guilford Rd., Madison, WI 53711 USA
All rights reserved.

This trend has several reasons. First, consumers orientate towards high-quality products with increased taste, serving as a carrier of pleasure, culture, and tradition (Barilla Center for Food & Nutrition, 2012; Puratos, 2012). An extreme example of this trend is the Slow Food movement. In addition to this, consumers pay increased attention to a wholesome diet and a more sustainable lifestyle. This manifests in the preference of healthy products coming from regional agriculture with decreased production intensity. To meet these criteria, spelt breads should ideally be produced with whole-grain flour, promising numerous health benefits (Aune et al., 2016), and should bear an intense, authentic, and pleasant bread flavor. Second, especially small- and medium-sized farmers search for alternative crops, which ideally make them independent of the global market of major crops and its volatility of prices, enabling a robust income while, in parallel, improving the farmers' image due to extensification and an increase in biodiversity. Third, artisan millers and bakers search for specialties, which enable them to compete against industrial bakeries with specialty segments not available in industrial bakery. Spelt fulfills these requirements because it is still regarded as a minor crop with current price margins far beyond bread wheat, and consumers associate tradition, good taste, and a healthier production with it. This increased demand across the whole product chain requires information about the agronomic performance, quality traits, and flavor of spelt and its products to maximize their acceptance and the profit along the product chain.

Considerable breeding efforts have been made in the last three decades in spelt, targeting especially a reduced plant height to minimize lodging while increasing yield. Consequently, a hulled yield, which includes chaff and grains, of up to 8 Mg ha^{-1} is possible nowadays (Longin and Würschum, 2014; Longin et al., 2016). After dehulling, about 70% of this hulled yield remains as final grain yield (Longin et al., 2016). Protein content of spelt is considerably higher than in bread wheat (Longin et al., 2016), but the protein quality is different, as indicated for instance by lower sedimentation volumes compared with bread wheat (Longin et al., 2016). Nevertheless, a significant genetic variation within different spelt varieties was identified for bread-making quality (Schober et al., 2002). Thereby, the loaf volume and the height-to-width ratio of the bread correlated quite well with the sedimentation volume. However, this study investigated only very old spelt varieties under low-input farming. Further, studies also investigating, beside agronomy and quality, the flavor and odor of the breads are, to the best of our knowledge, not available in the literature.

Consequently, our aim was to conduct a field-trial series at multiple locations in Germany with 30 current varieties and breeding lines of spelt, investigating agronomy and quality, as well as flavor and odor of the breads. In particular, our objectives were (i) to assess the genetic variability and heritability of agronomic and quality traits, as well as flavor and odor of the breads, (ii) to investigate the correlations among these traits, and (iii) to draw conclusions on breeding spelt cultivars with improved yield, quality, and flavor of their end products.

MATERIALS AND METHODS

Plant Material and Field Trials

Twenty-five recent spelt breeding lines from four breeding companies were evaluated together with five current spelt varieties. These five varieties were Oberkulmer Rotkorn (old and tall landrace), Franckenkorn (medium plant height, registered in 1996 in Germany), Zollernspelz (short plant height, registered in 2006 in Germany), Badenkrone (short plant height, registered in 2010 in Germany), and Cosmos (short plant height, registered in 2000 in Belgium). For simplicity of reading, we denote breeding lines in the following also as varieties.

The field trials were conducted as winter cropping (i.e., sowing in October 2014 and harvest in July 2015) at six diverse locations in Germany. Five locations were under conventional farming practices at Hohenheim (Hoh; 48°42′50″ N, 9°12′58″ E; 400 m asl, growing seasons mean temperature of 10.6°C, and mean precipitation of 670 mm), Aschersleben (Ale; 51°45′ N, 11°27′ E; 114 m asl, growing season mean temperature of 8.8°C), Rastatt (Ra; 48°51′ N, 8°12′ E; 114 m asl, growing season mean temperature of 10.3°C, and mean precipitation of 913 mm), Schwäbisch Hall (Sha; 49°12′ N, 9°30′ E; 270 m asl, growing season mean temperature of 10°C, and mean precipitation 765 mm), and Weimar (Wei; 51°1′3″ N, 11°21′27″ E; 268 m asl, growing seasons mean temperature of 8.6°C, and mean precipitation of 756 mm). One organic location was included in the study in Darmstadt (Das; 49°81′92″ N, 8°64′49″ E; 170 m asl, growing seasons mean temperature of 10.1°C and mean precipitation of 578 mm).

At each location, the material was evaluated in an α-lattice design with two replications (Schutz and Cockerham, 1962). Field plot size ranged between 5 and 9 m^2, sowing density from 280 to 350 dehulled grains m^{-2}, and number of rows from six to seven, with row spacing of 0.18 m, depending on the location. All plots were machine planted and combine harvested. The trials at Hoh, Ale, Ra, Wei, and Sha were conducted on conventional farming fields, treated once with herbicide (Ariane C, 1.5 L ha^{-1}), treated twice with growth regulator (once with Cycocel 720, 0.7 L ha^{-1}; once with Modus, 0.5 L ha^{-1}), and fertilized with 100 to 140 kg nitrogen ha^{-1} (ammonium nitrate). The organic trial in Das was grown after clover–grass mixture and was twice machine harrowed for weed control. No other treatments were applied.

Hulled yield was determined by the combine harvesters, adjusted to moisture content. Heading time was recorded as the day in the year, when 75% of the ears of the plot had fully emerged from the flag leaf. Plant height was measured in centimeters from the ground to the top of five representative ears of each plot, excluding awns. The different traits and their determination method, as well as their intensity of phenotyping described as number of locations and replications, are shown in Supplemental Table S1.

Laboratory Analyses

Subsequently, all samples of the locations Hoh and Sha were dehulled and cleaned using Mini-Petkus seed cleaner (Röber, Bad Oeynhausen, Germany) to separate hulls, straw, and damaged grains. The resulting samples were used for further quality analyses. Total protein content was determined with near infrared spectroscopy (ICC standard method 159, ICC, Vienna, Austria). Protein quality was determined with two sedimentation tests: as sedimentation volume in milliliters according to a sodium dodecyl sulfate micro-sedimentation test (SDSS, ICC standard method 151, ICC, Vienna, Austria) and as sedimentation value in milliliters according to Zeleny (Z-SDS, ICC standard method 116/1, ICC, Vienna, Austria).

To have enough seeds for bread-baking tests, the samples of both replications at the single locations Hoh and Sha were mixed. Baking tests were performed in an artisanal bakery. Grains were ground just before dough preparation by an industrial mill equipped with a 0.5-mm sieve (Mühlomat 100, Treffler, D-86554 Pöttmes-Echsheim) delivering fine whole-grain flour with low flour temperature. The mill has an automatic cleaning process based on high-pressure air, minimizing the intermixing of different samples. The whole-grain flour was used beside bread baking for further quality analyses comprising wet gluten content and gluten index, determined with a Glutomatic system from Perten (ICC standard method 155, ICC, Vienna, Austria). Furthermore, color measurement of the whole-grain flour (l, a, b-value) was performed with a Minolta chroma meter CR-400/410 (Konica Minolta, Osaka, Japan). Finally, mineral content of Ca, Cu, Fe, K, Mg, Mn, P, S, and Zn were determined by pressure digestion with nitric acid in quartz vessels in a microwave-heated, high-pressure digestion system (Ultra Clave II, MLS GmbH, Germany). Elements are measured by ICP–OES (inductively coupled plasma optical emission spectrometry) using a Varian Vista Pro instrument (radial plasma viewing) with a conical spray chamber, a Meinhardt type nebulizer, and external calibration against certified standards (various suppliers).

Breads were baked in loaf tins based on the following recipe: 2 kg whole-grain flour, 1.3 kg cold water, 40 g fresh yeast, and 40 g salt. The recipe corresponds to a standard spelt bread recipe of German bakeries. The formula was kept constant to satisfy "*ceteris paribus*" conditions for all samples. To avoid any further interference, the addition of further ingredients like ascorbic acid, enzymes, and others was omitted. All ingredients were put together and mixed for 4 min at slow speed and kneaded for 0.5 min at high speed in an Alpha mixer from Häussler (Häussler GmbH, D-88499 Heiligkreuztal). The rheological character of the gluten network as an estimation for dough quality was then judged by an experienced baker by stretching a sample of about 20 g dough in his fingers (Fig. 1). Dough being highly extensible without cracking was judged as very good, while a short dough rapidly cracking was judged as poor. We transformed the bakers' notation of very good, good, intermediate, and poor dough quality in the numbers 4, 3, 2, and 1, respectively, for statistical analysis. This manual method tries to roughly imitate Barbender's extensograph, which was unfortunately not available for our study. The complete doughs resulting from each whole-grain sample were then kneaded manually into two breads, put into loaf tins, and stored in a climate chamber with a 30°C temperature and 70% humidity. After 75 min of proving time, the loaf tins were transferred into a Thermo-Oel oven (HEUFT Thermo-Oel GmbH, D-56745 Bell) with an initial temperature of 250°C and then baked at 220°C for 50 min. After baking, the breads were left to cool down at room temperature before storing them under a cover until the next day.

Breads were judged according to the rules of the German Agricultural Society (DLG e.V.; DIN 10969:2001–05, DIN 10964:2014–11, and DIN 10975:2005–04), representing the standard method for testing bread-making quality in Germany, which almost all bakeries regularly perform for quality management. One day after bread making, the breads were judged, assuming that fresh bread most often tastes good, while a clearer differentiation among better and poorer quality is apparent the day after baking. The judgement was performed by a panel of four people, which had to give a jointly decided judgement on the bread quality. The odor was tested by closely smelling at the crumb, and the flavor was assessed by tasting the crumb. Thereby, a scale was used from 1 = no aroma to 5 = very intensive aroma. The quality of the crumb, the crust, and the loaf volume was estimated visually with a scale from 1 = poor to 5 = very good. However, differences for these three traits were low and difficult to judge visually, leading to heritability estimates of zero. Thus, these traits were disregarded in the following results, and we highly recommend running bread-baking tests for spelt, like for bread wheat, as bread rolls baked without loaf tins and judging them with photometric or other more sophisticated devices. However, the judgement of the crumb flavor and odor worked well with our method.

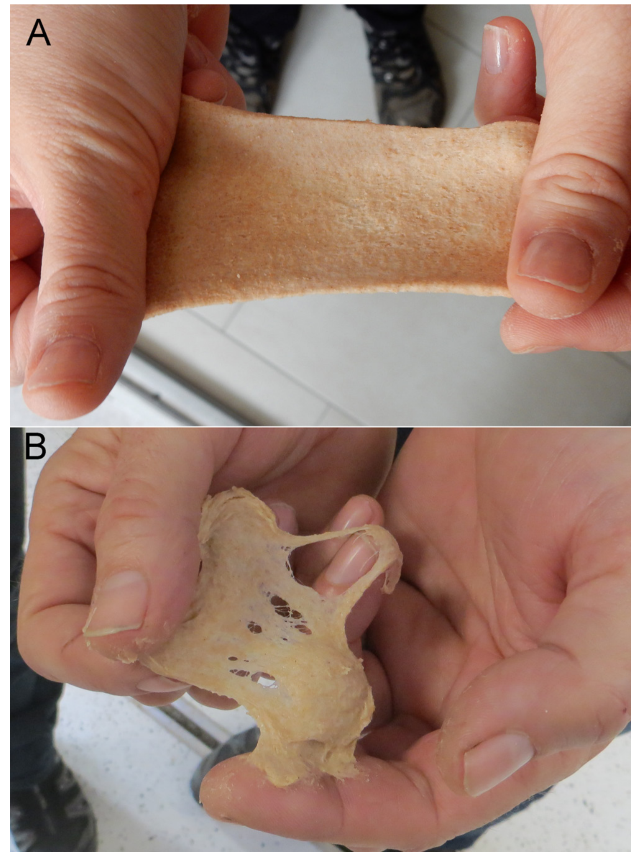

Fig. 1. Dough quality determined by manually stretching a sample of 20 g with (A) good extensible dough and (B) poor dough with short structure rapidly cracking.

Data Analyses

The data were analyzed according to the following statistical model:

$$y_{ikno} = \mu + g_i + \text{loc}_k + g_i{:}\text{loc}_k + r_{kn} + b_{kno} + e_{ikno},$$

where y_{ikno} was the phenotypic observation for the *i*th variety tested in the *k*th location in the *n*th replication in the *o*th incomplete block, μ was an intercept term, g_i was the genetic effect of the *i*th genotype, loc_k was the effect of the *k*th location, $g_i{:}loc_k$ was the genotype-by-location interaction, r_{kn} was the effect of the *n*th replication at the *k*th location, b_{kno} was the effect of the *o*th incomplete block in the *n*th replication at the *k*th location, and e_{ikno} was the residual. All effects were regarded as random.

Variance components were determined by the restricted maximum likelihood (REML) method assuming a random model and a heterogenic error for single locations using classical one-stage analyses (Cochran and Cox, 1957). Significance of variance component estimates was tested by model comparison with likelihood ratio tests in which halved *P*-values were used as approximation (Stram and Lee, 1994). Heritability was calculated as $h^2 = 1 - \vartheta/\sigma^2_G$, where ϑ is the mean variance of a difference of two best linear unbiased predictors (BLUP) and σ^2_G the genetic variance (Cullis et al., 2006; Piepho and Möhring, 2007). In addition, best linear unbiased estimates (BLUEs) across locations were estimated assuming fixed genetic effects.

As described above, for bread-baking tests, the samples of both replications at the investigated locations Hoh and Sha were mixed. Thus, the statistical model was reduced to $y_{ik} = \mu + g_i + loc_k + e_{ik}$. Phenotypic correlation coefficients (*r*) were estimated among BLUEs of the examined traits. All analyses were performed with the statistical software R (R Development Core Team, 2016) and the software ASReml 3.0 (Gilmour et al., 2009).

RESULTS

For almost all traits, a wide variation of the genotypic values was observed, resulting in σ^2_G significantly larger than zero ($p < 0.05$) (Tables 1 and 2). Estimates of the variances due to genotype-by-environment interaction ($\sigma^2_{G \times E}$) were considerably smaller than σ^2_G for heading time, plant height, SDSS, Z-SDS, SDSS/protein content, and Z-SDS/protein content but were approximately as high as σ^2_G for yield and protein content. Owing to the lack of replications, no $\sigma^2_{G \times E}$ could be determined for gluten content, gluten index, dough quality, different minerals, l-, a-, and b-values, flavor, or odor. A large range of heritability estimates for the different traits was observed. For instance, the lowest heritability of 0.36 was determined for the mineral content of P, while the highest heritability was estimated to be 0.96 for heading time. For hulled yield, a heritability of 0.77 was observed.

Phenotypic correlation coefficients varied largely between the different agronomic and quality traits (Table 3). For instance, highly significant negative correlations between hulled yield and protein content ($r = -0.60$, $p < 0.001$) or plant height ($r = -0.64$, $p < 0.001$) were observed. Both methods to determine the sedimentation volume (i.e., SDSS and Z-SDS) were very tightly correlated with each other ($r = 0.94$, $p < 0.001$), while both were only moderately correlated with protein content. No correlation was found between dough quality and protein or gluten content. The standardization of the sedimentation volume by the protein content yielded the highest correlation with dough quality ($r = 0.46$, $p < 0.05$ for SDSS and $r = 0.49$, $p < 0.01$ for Z-SDS).

Table 1. Minimum, mean, maximum values and the LSD of the phenotypic values, genetic variance (σ^2_G), variance due to genotype-by-environment interaction ($\sigma^2_{G \times E}$), pooled error variance (σ^2_E), and heritability (h^2) for agronomic and quality traits of 30 spelt varieties analyzed across locations and replications (SDSS, sedimentation volume; Z-SDS, sedimentation value according to Zeleny).

	Hulled yield	Heading time	Plant height	Hectoliter mass	Protein content	SDSS	Z-SDS	Gluten content	Gluten index	Dough quality	Flavor	Odor
	Mg ha^{-1}	d	cm	kg	% total solids	—— mL ——		g		—— score ——		
Min.	5.42	149.80	108.80	32.26	13.92	39.92	14.87	31.20	7.43	1.00	1.00	0.84
Mean	6.83	152.90	122.70	35.90	15.80	60.00	22.52	37.49	19.71	2.27	2.72	2.92
Max.	7.78	157.20	136.70	40.09	18.54	77.98	30.72	45.40	62.80	4.00	5.00	5.00
LSD 5%	0.62	1.14	7.12	4.25	1.51	9.26	3.46	5.83	25.51	1.71	1.94	2.11
σ^2_G	0.16***	4.09***	33.98***	2.09	0.57**	103.89***	13.23***	6.58*	100.79	0.47*	0.61*	0.46
$\sigma^2_{G \times E}$	0.17***	0.62***	19.99***	1.21	0.46***	18.61***	1.90**	–	–	–	–	–
σ^2_E	0.39	2.92	17.81	5.56	0.19	4.07	1.58	8.51	162.65	0.73	0.91	1.07
h^2	0.77	0.96	0.86	0.51	0.67	0.91	0.90	0.61	0.55	0.55	0.56	0.45

* Significant at the 0.05 probability level; ** significant at the 0.01 probability level; *** significant at the 0.001 probability level.

Table 2. Minimum, mean, maximum values and the LSD of the phenotypic values, genetic variance (σ^2_G), pooled error variance (σ^2_E), and heritability (h^2) for quality traits of 30 spelt varieties analyzed across locations.

	l-value	a-value	b-value	Ca	Cu	Fe	K	Mg	Mn	P	S	Zn
				—— mg kg^{-1} ——								
Min.	83.55	1.54	10.80	204.50	3.87	36.05	3680.00	1188.00	14.55	3943.00	1489.00	23.05
Mean	85.29	2.03	11.73	265.10	4.69	41.44	4173.00	1303.00	22.17	4361.00	1614.00	26.45
Max.	86.28	2.49	12.97	342.00	5.68	48.85	4493.00	1418.00	27.35	4626.00	1800.00	33.20
LSD 5%	0.51	0.17	0.39	29.81	0.52	4.91	254.08	110.07	3.62	319.95	130.05	4.18
σ^2_G	0.44***	4.00×10^{-4}***	0.24***	897.64***	0.14***	7.47***	3.51×10^{4}***	1.37×10^{3}	4.83***	7.16×10^{3}	2.67×10^{3}	3.03*
σ^2_G	0.07	1.00×10^{-4}	0.04	222.21	0.07	6.03	1.61×10^{4}	3.03×10^{3}	3.28	2.56×10^{3}	4.15×10^{3}	4.36
h^2	0.93	0.92	0.93	0.89	0.81	0.71	0.81	0.47	0.75	0.36	0.56	0.58

* Significant at the 0.05 probability level; ** significant at the 0.01 probability level; *** significant at the 0.001 probability level.

Table 3. Phenotypic correlation coefficients among agronomic and quality traits from 30 spelt varieties (SDSS, sedimentation volume; Z-SDS, sedimentation value according to Zeleny; SSM, sum of standardized mineral contents Ca, Cu, Fe, K, Mg, Mn, P, S, and Zn).

	Heading time	Plant height	hL mass	Protein content	SDSS	Z-SDS	l-value	a-value	b-value	Gluten content	Gluten index	Dough quality	Flavor	Odor	Ca	Cu	Fe	K	Mg	Mn	P	S	Zn	SSM
Hulled yield	−0.25	−0.64***	−0.05	−0.60***	−0.14	−0.23	0.24	−0.24	−0.39*	−0.51**	0.14	0.19	0.08	−0.02	0.06	−0.60***	−0.63***	0.62***	−0.33	−0.53**	−0.01	−0.50**	−0.55**	−0.39*
Heading time		0.23	−0.59***	−0.12	−0.36*	−0.36	0.01	0.28	−0.03	−0.06	−0.07	0.10	−0.12	−0.14	−0.18	0.12	−0.01	0.08	−0.07	0.15	0.00	0.03	0.02	−0.02
Plant height			0.03	0.39*	−0.05	0.07	−0.09	0.11	0.26	0.04	−0.06	−0.21	−0.01	0.18	−0.36*	0.47**	0.53**	−0.45*	0.50**	0.39*	0.11	0.34	0.32	0.32
hL mass				0.13	0.45*	0.47**	−0.02	−0.23	0.04	0.00	0.32	0.22	0.03	0.00	0.18	0.08	0.16	−0.12	0.18	0.07	0.05	0.05	-0.07	0.11
Protein content					0.41*	0.52**	−0.25	0.21	0.18	0.72***	−0.08	−0.05	−0.38*	−0.01	−0.22	0.54**	0.82***	−0.49**	0.47**	0.36	0.32	0.78***	0.68***	0.55**
SDSS						0.94***	0.12	−0.23	−0.17	0.50**	0.51**	0.39*	−0.28	−0.03	−0.01	0.15	0.14	−0.28	0.17	−0.03	0.06	0.27	-0.05	0.01
Z-SDS							−0.03	−0.05	−0.06	0.55**	0.47**	0.41*	−0.22	0.04	−0.14	0.19	0.31	−0.37	0.27	0.03	0.09	0.29	0.03	0.05
l-value								−0.77***	−0.67***	−0.27	0.33	0.09	0.06	0.00	−0.01	−0.09	−0.32	0.11	−0.08	−0.31	0.15	−0.06	-0.25	−0.11
a-value									0.28	0.14	−0.32	0.15	−0.08	0.02	−0.37*	0.05	0.32	−0.07	0.00	−0.02	−0.08	0.08	0.10	−0.08
b-value										0.27	−0.40*	−0.42*	0.13	0.02	0.15	0.26	0.33	−0.40*	0.06	0.57**	−0.25	0.00	0.33	0.18
Gluten content											−0.16	−0.09	−0.28	−0.07	0.04	0.43*	0.55**	−0.59***	0.20	0.41*	−0.05	0.58***	0.45*	0.32
Gluten index												0.43*	−0.24	−0.23	0.07	0.03	−0.13	0.34	0.18	−0.01	0.30	−0.09	-0.33	0.03
Dough quality													−0.22	−0.12	−0.22	−0.26	−0.11	0.05	−0.10	−0.38*	−0.05	−0.18	-0.44*	−0.40*
Flavor														0.65***	0.18	−0.20	−0.14	−0.03	0.05	−0.19	−0.05	−0.20	-0.08	−0.04
Odor															−0.02	−0.20	0.09	−0.07	0.31	−0.30	0.22	0.06	0.04	0.10
Ca																−0.07	−0.28	0.08	−0.11	0.20	−0.10	−0.16	0.04	0.21
Cu																	0.67***	−0.45**	0.49**	0.70***	0.37*	0.64***	0.61***	0.70***
Fe																		−0.53**	0.62***	0.51**	0.37*	0.70***	0.65***	0.66***
K																			−0.21	−0.39*	0.32	−0.33	-0.35	−0.10
Mg																				0.35	0.70***	0.57**	0.51**	0.75***
Mn																					0.06	0.34	0.52**	0.59***
P																						0.47**	0.48**	0.73***
S																							0.67***	0.68***
Zn																								0.78***

* Significant at the 0.05 probability level; ** significant at the 0.01 probability level; *** significant at the 0.001 probability level.

A positive and highly significant correlation of flavor and odor of the breads was observed ($r = 0.65$, $p < 0.001$). We further observed highly significant ($p < 0.01$) and negative correlation coefficients between hulled yield and the mineral salts Cu, Fe, Mn, S, and Zn. In contrast, hulled yield was positively correlated with the mineral salt K ($r = 0.62$, $p < 0.001$). Furthermore, a highly significant and positive correlation between protein content and the minerals Fe, S, and Zn was determined ($p < 0.001$).

The breads from the different varieties showed different colors and structures of the crust (Fig. 2). For instance, the bread of the breeding line GZPK_REPSA18 from location Hoh had a considerable lighter crust than the other three varieties shown in Fig. 2. However, the crust of the bread of this variety from the second location Sha was considerably darker (Fig. 2). Furthermore, heritability was zero for the judgment of the crust due to a very low σ^2_G but high error variance (σ^2_E) observed in our trial (data not shown). Thus, the crust color seems to be highly influenced by the environment and less by the variety. Although whole-grain flours differed largely in color, as indicated by high genetic variability for l-, a-, and b-value of the Minolta chroma meter (Table 2), these values did not correlate with the crust color of the breads (data not shown).

DISCUSSION

Spelt is an old-world crop that has received growing attention from consumers, bakers, breeders, and farmers in recent years. A sustainable reestablishment of spelt requires knowledge about best farming practices, quality for bread making, and parameters to measure it, as well as its potential to deliver high-quality products with favorable flavor and odor. For agronomic traits, two recent studies (Longin and Würschum, 2014; Longin et al., 2016) and, for bread-making traits, one older study are available in the literature (Schober et al., 2002), but odor and flavor of spelt breads and their inheritance have not been investigated yet. Furthermore, investigating agronomy and bread-making quality, as well as flavor and odor of the breads, in the same trial enables us to draw important conclusions about the correlations between all these traits affecting the complete product chain. However, such a study is yet completely lacking in the literature for spelt. This motivated us to investigate 30 different spelt varieties at six different locations in Germany for important agronomic and quality traits, as well as for odor and flavor of the final breads.

Although a considerably smaller number of varieties exists in spelt as compared with bread wheat, we observed a large range of genotypic values, resulting in highly significant σ^2_G for almost all traits (Tables 1 and 2). Heritability estimates were high and comparable with studies in spelt (Longin and Würschum, 2014; Longin et al., 2016), bread, and durum wheat (Longin et al., 2013a, 2013b). For gluten content, gluten index, dough quality, flavor,

Fig. 2. Different colors of breads made from different varieties grown at different locations. The upper row represents the location Schwäbisch Hall, the lower row the location Hohenheim, always with two breads from one variety.

and odor, heritability estimates ranged around 0.5 (Table 1), which especially for gluten index is lower than reports for other wheat species (Longin et al., 2013b). However, these traits were only measured at two locations without replications, and thus their heritability might be underestimated as compared, for instance, with yield measured at six locations and two replications in this study. Consequently, our dataset seems to be based on phenotypic data of high quality, allowing for meaningful conclusions for spelt in general and spelt breeding in particular.

Quality Traits for Bread-Making Potential

For protein and wet gluten content, we determined moderately high heritabilities of 0.67 and 0.61, respectively (Table 1). This is in accordance with other reports for spelt and wheat (Longin et al., 2016), showing that protein and gluten content are influenced not only by the variety, but also largely by the environment where the specific variety was grown. We further observed a high correlation between protein and gluten content of 0.72 ($p < 0.001$, Table 2), which is in accordance with results in spelt (Schober et al., 2006) and durum wheat (Longin et al., 2013b). As protein content can be determined rapidly and nondestructively by near-infrared spectroscopy, it could be used as a rough estimator of the gluten content, avoiding its more cumbersome determination. Interestingly, neither protein nor gluten content were correlated with dough quality. Baking studies in bread wheat with different flour blends showed that protein content and glutenin-to-gliadin ratios independently influence dough and baking characteristics (Uthayakumaran et al., 1999). Further studies in bread wheat revealed nonsignificant correlations between protein content and dough quality, and it was found that the protein quality is of higher importance than the relative protein content (Uhlen et al., 2004; Maphosa et al., 2015). In spelt, gluten content showed no significant correlation with bread volume, nor with the width-to-height ratio (Schober et al., 2002). Thus, other methods than protein or gluten content should be used to indirectly predict dough and bread-making quality in spelt.

In bread wheat, sedimentation volume is broadly used as an estimator for protein and bread-making quality. Thereby, two slightly different methods were applied, Z-SDS or SDSS. For both methods, very high heritabilities of 0.91 and 0.90 were determined for SDSS and Z-SDS, respectively (Table 1). Furthermore, a very tight positive correlation of $r = 0.94$ ($p < 0.001$) was observed between SDSS and Z-SDS (Table 3), which is in accordance with reports for bread wheat (Wang and Kovacs, 2002). In addition, both methods had a similarly high correlation with dough quality. Thus, both methods seem to be usable synonymously for spelt. However, the differentiation of the varieties was better with SDSS, with mean values ranging from 39.92 to 77.98 mL as compared with 14.87 to 30.72 mL for Z-SDS (Table 1). Therefore, the German Federal Office for the registration of varieties, as well as German spelt breeders, use the SDSS method for determining protein quality in spelt.

In contrast to breeders, millers often use the gluten index for estimating protein quality. We determined a moderately positive correlation of gluten index with SDSS ($r = 0.51$, $p <$

0.01) and Z-SDS ($r = 0.47$, $p < 0.01$), respectively (Table 3), which is lower than values reported for durum (Clarke et al., 2010; Longin et al., 2013b) and bread wheat (Gómez et al., 2009; Bonfil et al., 2015). However, the heritability for gluten index was 0.55, and thus much lower than the heritability of Z-SDS or SDSS of 0.90. This might be partly explained by the fact that we measured Z-SDS and SDSS at samples from two locations, each with two replications, while gluten index was determined based on the flour samples for baking, where we had already mixed the two replications from the specific locations to have enough grains and flour for baking. A further and more probable reason might be that a correct determination of the gluten index is tricky with spelt, due to its softer gluten as compared to durum or bread wheat, thereby increasing the error in its determination and thus reducing heritability. This is supported by the mean values of the varieties, where three varieties had an extraordinarily high gluten index, while the majority of varieties were quite close to each other (Supplemental Table S2). In contrast to these observations for gluten index, the mean values for Z-SDS, SDSS, and dough quality were roughly following a normal distribution. More research is required to verify whether these problems with gluten index are an artifact of our dataset or whether gluten index is generally an error-prone method for determining protein quality in spelt.

Dough quality was measured by stretching about 20 g of freshly prepared dough manually to observe the dough structure and extensibility. A significant and large σ^2_G was observed (Table 1, Fig. 1). Despite seeing these large differences, we unfortunately classified them in only four classes—very good, good, intermediate, and poor dough quality, which were later transformed into numbers 4, 3, 2, and 1, respectively, for statistical analysis. This might explain why we determined a heritability of only 0.55 for dough quality. We recommend for further research the use of an extensograph or, if not available, at least a broader scale for the manual extension, ranging from 1–9 to better capture the large differentiation between varieties.

We determined correlation coefficients of similar but only moderately high size (r ~ 0.40, $p < 0.05$) for dough quality and gluten index Z-SDS, or SDSS, respectively (Table 3; Bonfil et al., 2015). For Z-SDS and SDSS, this correlation could be increased to $r = 0.49$ ($p < 0.01$) and $r = 0.46$ ($p < 0.05$) by standardizing their value with the respective protein content. This standardization is based on the fact that values of Z-SDS and SDSS for specific varieties partly depend on the protein content of the respective variety, which can be seen in a moderate and significant correlation of Z-SDS and SDSS with protein content (Table 3). This is in accordance with findings in bread wheat (Graybosch et al., 1996; Nakamura et al., 2012) and is the reason why many bread wheat breeders use the ratio Z-SDS/protein content or SDSS/protein content as main selection criterion for protein and bread-making quality. A high correlation coefficient of SDSS and loaf volume and, in particular, the width-to-height ratio of breads was reported for bread-making tests with old spelt varieties (Schober et al., 2002), underlining the usefulness of determining protein quality for spelt.

As mentioned in the Materials and Methods, we also determined the loaf volume of the breads, but only visually on a scale from 1 to 5 due to the lack of image analyses or other techniques in a routine artisanal bakery, where we performed our study. Having made the breads in loaf tins, however, the difference between the varieties seen in dough quality was not any more visible on the final breads. Thus, we recommend conducting bread-making tests in spelt with bread rolls baked freely without loaf tins and measuring the bread volume and its width-to-height ratio by photo technique or similar methods.

Good Differentiation among Varieties Regarding Flavor and Odor of their Breads

We judged the flavor and odor of the different breads according to the rules of the German Agricultural Society (DLG e.V.; DIN 10969:2001–05, DIN 10964:2014–11, and DIN 10975:2005–04), representing the standard method used by German bakers to test bread-making quality. The breads were tested 1 d after baking by tasting and smelling the crumb and finding jointly, as a group of four people, an average grade on a scale from 1 = no flavor or odor to 5 = very intensive and aromatic flavor and odor. Despite using this simple method, we observed a large differentiation in flavor of the different breads, leading to a significant σ^2_G (Table 1). These findings underline results from other studies that also report significant differences in flavor between their few investigated varieties (Callejo et al., 2015; Starr et al., 2015).

We determined a heritability of 0.56 for flavor of the breads, which in our opinion is quite high, taking into account the investigation of samples from only two locations without replications and the complexity of flavor and odor (Pico et al., 2015; Heiniö et al., 2016). This underlines the feasibility to select for good flavor in breeding programs. However, these findings also show that flavor is not only influenced by the genetics, but also by the environmental conditions where the test varieties were grown, underlining the necessity of testing flavor on variety samples coming from multiple environments.

The differentiation between the varieties was lower for odor, with a higher environmental impact (Table 1). Nevertheless, a correlation coefficient of $r = 0.65$ ($p < 0.001$) was determined between flavor and odor, showing that these two traits are partly influenced by the same compounds and genetics. By tasting and smelling the breads from 30 different varieties grown at two different locations, we observed another surprising fact. Despite having used the identical recipe and preparation protocol, some of the breads had an intensive smell and taste of the yeast, while others did not (data not shown). Interestingly, neither the color measurement of the whole-grain flour by the l-, a-, and b-value of Minolta's chroma meter, nor the content of the nine measured minerals or their sum, were correlated with the flavor or odor of the bread. Previous studies in bread wheat have shown a correlation between the grain color, the phenolic acid content, and flavor (Challacombe et al., 2012; Starr et al., 2013). While these contrasting results might be due to a general difference between spelt and bread wheat, they might also be explained by the composition of the evaluated wheat sets. In these studies, a highly diverse set of genotypes ranging from dark red grains to light grains was used, whereas our spelt set consisted entirely of normal colored spelt varieties with little variation in grain color. Further research is required to investigate the flavor and odor of bread cereals in general, to compare the different standard methods (e.g., the DLG method we used with sensory panel evaluation), and to elaborate the environmental and genetic reasons influencing them and the potential to exploit these aroma profiles most efficiently in breeding and for preparation of better end products. An established method for odor evaluation of Basmati rice (*Oryza sativa* L.) is the use of 1.7% KOH solution, which allows one to conduct tests with small samples of only few grains (Hien et al., 2006; Jeng et al., 2015). This could also be a promising method for spelt breeding if the grain or flour odor and the odor of the end products are sufficiently correlated. This requires more interdisciplinary research from agronomy and cereal science, and especially enormous efforts to improve throughput in phenotyping odor and flavor rapidly for high sample numbers. As plant breeding is always dealing with high numbers of potential new varieties, the development of a rapid test method, ideally requiring only few grains or flour for estimating the flavor and odor, would be necessary to maximize the efficiency of breeding.

Next-Generation Wheat Breeding Involves Selection on Flavor and Odor

Many bakers complain that the selection of new varieties is mainly based on high yield and better protein quality, which has reduced the flavor of the flour of new varieties as compared with historical wheat varieties. Beside subjective observations, this speculation is driven by the argument that an increase in grain yield is mainly based on an increased amount of starch in the inner endosperm, and the "ingredients" delivering taste are only maintained or even reduced in relation to starch. In our investigations on spelt, however, we did not see any correlation between flavor nor odor of the breads with yield, protein, or dough quality (Table 3). Furthermore, the variety Badenkrone, known for its outstanding yield, was the variety identified with the most favorable flavor (Fig. 3). In contrast, the old landrace Oberkulmer Rotkorn had a rather poor flavor in our evaluation. Thus, it seems possible to select new varieties with high yield, good protein, and high dough quality coupled with an intensive flavor and odor, opening the door for "next-generation" wheat breeding.

As flavor and odor can only yet be determined in bread-making tests and both have only moderately high heritabilities, we propose to delay their testing to the later-breeding generations, where larger sample sizes coming from different field locations are available. In contrast, the high heritability, coupled with the need of only few grams of seeds and its good differentiation among varieties, makes the selection on SDSS in early generations appealing. Combined with the classical selection criteria in early generations (i.e., plant height, heading time, and disease resistances), only candidates delivering improved qualities and agronomic characteristics can then be intensively tested on yield, and the highest-yielding ones later on dough and bread-making quality coupled with flavor and odor. However, selection on flavor and odor is cumbersome and expensive. Thus, breeders and farmers might only promote varieties with better flavor as long as their advantage is honored by higher prices compared with standard wheat or spelt.

Conflict of Interest

The authors declare there to be no conflict of interest.

Supplemental Material Available

Supplemental material for this article is available online.

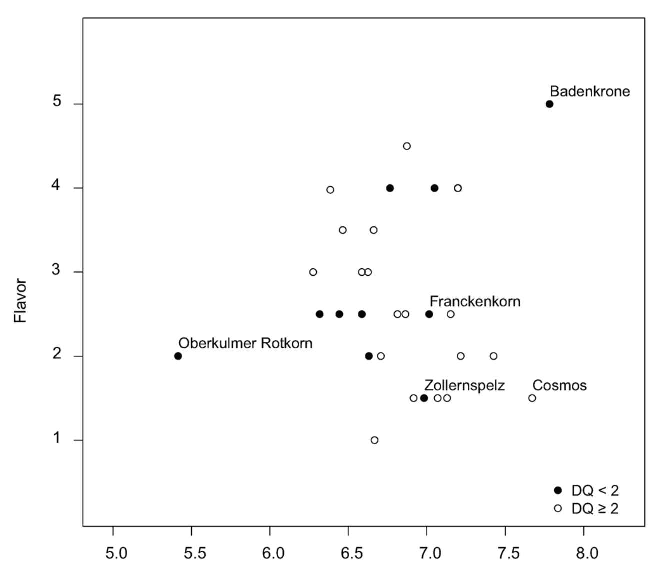

Fig. 3. Mean values of hulled yield plotted against flavor of 25 elite spelt breeding lines and five spelt varieties Cosmos, Badenkrone, Franckenkorn, Oberkulmer Rotkorn, and Zollernspelz; filled dots indicate a low dough quality, open dots a high dough quality.

References

Aune, D., N. Keum, E. Giovannucci, L.T. Fadnes, P. Boffetta, D.C. Greenwood et al. 2016. Whole grain consumption and risk of cardiovascular disease, cancer, and all cause and cause specific mortality: Systematic review and dose-response meta-analysis of prospective studies. BMJ 353:i2716. doi:10.1136/bmj.i2716

Barilla Center for Food & Nutrition. 2012. Eating in 2030: Trends and perspectives. Barilla Center for Food & Nutrition. https://www.barillacfn.com/en/publications/eating-in-2030-trends-and-perspectives (accessed 11 May 2016).

Bonfil, D.J., S. Abbo, and T. Svoray. 2015. Sowing date and wheat quality as determined by gluten index. Crop Sci. 55:2294–2306. doi:10.2135/cropsci2014.08.0562

Callejo, M.J., M.E. Vargas-Kostiuk, and M. Rodriguez-Quijano. 2015. Selection, training and validation process of a sensory panel for bread analysis: Influence of cultivar on the quality of breads made from common wheat and spelt wheat. J. Cereal Sci. 61:55–62. doi:10.1016/j.jcs.2014.09.008

Campbell, K.G. 1997. Spelt: Agronomy, genetics, and breeding. In: J. Janick, Plant breeding reviews. Vol. 15. John Wiley & Sons, New York. p. 187–213.

Challacombe, C.A., E.S.M. Abdel-Aal, K. Seetharaman, and L.M. Duizer. 2012. Influence of phenolic acid content on sensory perception of bread and crackers made from red or white wheat. J. Cereal Sci. 56:181–188. doi:10.1016/j.jcs.2012.03.006

Clarke, F.R., J.M. Clarke, N.A. Ames, R.E. Knox, and R.J. Ross. 2010. Gluten index compared with SDS-sedimentation volume for early generation selection for gluten strength in durum wheat. Can. J. Plant Sci. 90:1–11. doi:10.4141/CJPS09035

Cochran, W.G., and G.M. Cox. 1957. Experimental designs. 2nd ed. John Wiley & Sons, New York.

Cullis, B.R., B. Smith, and N.E. Coombes. 2006. On the design of early generation variety trials with correlated data. J. Agric. Biol. Environ. Stat. 11:381–393. doi:10.1198/108571106X154443

Gilmour, A., B. Gogel, B. Cullis, and R. Thompson. 2009. ASReml user's guide. Release 3.0. VSN International Ltd, Hemel Hempstead, UK.

Gómez, M., N. Aparicio, E. Ruiz-París, B. Oliete, and P.A. Caballero. 2009. Evolution of bread-making quality of Spanish bread-wheat genotypes. Span. J. Agric. Res. 7:585–595. doi:10.5424/sjar/2009073-443

Graybosch, R.A., C.J. Peterson, D.R. Shelton, and P.S. Baenziger. 1996. Genotypic and environmental modification of wheat flour protein composition in relation to end-use quality. Crop Sci. 36:296–300. doi:10.2135/cropsci1996.0011183X003600020014x

Heiniö, R.L., M.W.J. Noort, K. Katina, S.A. Alam, N. Sozer, H.L. de Kock et al. 2016. Sensory characteristics of wholegrain and bran-rich cereal foods—A review. Trends Food Sci. Technol. 47:25–38. doi:10.1016/j.tifs.2015.11.002

Hien, N.L., T. Yoshihashi, W.A. Sarhadi, V.C. Thanh, Y. Oikawa, and Y. Hirata. 2006. Evaluation of aroma in rice (*Oryza sativa* L.) using KOH method, molecular markers and measurement of 2-acetyl-1-pyrroline concentration. Jpn. J. Trop. Agric. 50:190–198. doi:10.11248/jsta1957.50.190

Jeng, T.L., C.S. Chan, S.Y. Lin, D.P. Shung, C.Z. Pan, Y.J. Shih, and J.M. Sung. 2015. Comparative analysis of eating quality and yield of selected non-waxy red-pericarp aromatic rice mutants. J. Appl. Bot. Food Qual. 88:145–151.

Longin, C.F.H., M. Gowda, J. Mühleisen, E. Ebmeyer, E. Kazman, R. Schachschneider et al. 2013a. Hybrid wheat: Quantitative genetic parameters and consequences for the design of breeding programs. Theor. Appl. Genet. 126:2791–2801. doi:10.1007/s00122-013-2172-z

Longin, C.F.H., A.N. Sieber, and J.C. Reif. 2013b. Combining frost tolerance, high grain yield and good pasta quality in durum wheat. Plant Breed. 132:353–358. doi:10.1111/pbr.12064

Longin, C.F.H., and T. Würschum. 2014. Genetic variability, heritability and correlation among agronomic and disease resistance traits in a diversity panel and elite breeding material of spelt wheat. Plant Breed. 133:459–464. doi:10.1111/pbr.12182

Longin, C.F.H., J. Ziegler, R. Schweiggert, P. Koehler, R. Carle, and T. Würschum. 2016. Comparative study of hulled (einkorn, emmer, and spelt) and naked wheats (durum and bread wheat): Agronomic performance and quality traits. Crop Sci. 56:302–311. doi:10.2135/cropsci2015.04.0242

Maphosa, L., P. Langridge, H. Taylor, L.C. Emebiri, and D.E. Mather. 2015. Genetic control of grain protein, dough rheology traits and loaf traits in a bread wheat population grown in three environments. J. Cereal Sci. 64:147–152. doi:10.1016/j.jcs.2015.05.010

Miedaner, T., and C.F.H. Longin 2016. Neglected cereals: From ancient grains to superfood. Agrimedia, Clenze, Germany.

Nakamura, K., Y. Taniguchi, M. Taira, and H. Ito. 2012. Investigation of soft wheat flour quality factors associated with sponge cake sensory tenderness. Cereal Chem. 89:79–83. doi:10.1094/CCHEM-07-11-0081

Pico, J., J. Bernal, and M. Gomez. 2015. Wheat bread aroma compounds in crumb and crust: A review. Food Res. Int. 75:200–215. doi:10.1016/j.foodres.2015.05.051

Piepho, H.-P., and J. Möhring. 2007. Computing heritability and selection response from unbalanced plant breeding trials. Genetics 177:1881–1888. doi:10.1534/genetics.107.074229

Puratos. 2012. TASTE/TOMORROW from consumer insights to innovation. Puratos, Groot-Bijgaarden, Belgium. http://www.insites-consulting.com/global-food-survey-identifies-10-future-proof-consumer-trends/ (accessed 11 May 2016).

R Development Core Team. 2016. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Schober, T.J., S.R. Bean, and M. Kuhn. 2006. Gluten proteins from spelt (*Triticum aestivum* ssp. *spelta*) cultivars: A rheological and size-exclusion high-performance liquid chromatography study. J. Cereal Sci. 44:161–173. doi:10.1016/j.jcs.2006.05.007

Schober, T.J., C.I. Clarke, and M. Kuhn. 2002. Characterization of functional properties of gluten proteins in spelt cultivars using rheological and quality factor measurements. Cereal Chem. 79:408–417. doi:10.1094/CCHEM.2002.79.3.408

Schutz, W.M., and C.C. Cockerham. 1962. The effect of field blocking on gain from selection. Institute of Statistics Mimeograph Ser. No. 328. North Carolina State Univ., Raleigh, North Carolina.

Starr, G., W.L.P. Bredie, and Å.S. Hansen. 2013. Sensory profiles of cooked grains from wheat species and varieties. J. Cereal Sci. 57:295–303. doi:10.1016/j.jcs.2012.11.014

Starr, G., T.S. Hansen, M.A. Petersen, and W.L.P. Bredie. 2015. Aroma of wheat porridge and bread-crumb is influenced by the wheat variety. LWT-. Food Sci. Technol. (Campinas) 63:590–598.

Stram, D.O., and J.W. Lee. 1994. Variance components testing in the longitudinal mixed effects model. Biometrics 50:1171–1177. doi:10.2307/2533455

Uhlen, A.K., S. Sahlstrøm, E.M. Magnus, E.M. Færgestad, J.A. Dieseth, and K. Ringlund. 2004. Influence of genotype and protein content on the baking quality of hearth bread. J. Sci. Food Agric. 84:887–894. doi:10.1002/jsfa.1797

Uthayakumaran, S., P.W. Gras, F.L. Stoddard, and F. Bekes. 1999. Effect of varying protein content and glutenin-to-gliadin ratio on the functional properties of wheat dough. Cereal Chem. 76:389–394. doi:10.1094/CCHEM.1999.76.3.389

Wang, C., and M.I.P. Kovacs. 2002. Swelling index of glutenin test. I. Method and comparison with sedimentation, gel-protein, and insoluble glutenin tests. Cereal Chem. 79:183–189. doi:10.1094/CCHEM.2002.79.2.183