

Article

# Dissecting the Genetic Basis of Lateral and Central Spikelet Development and Grain Traits in *Intermedium-Spike Barley (Hordeum vulgare Convar. Intermedium)*

Helmy M. Youssef <sup>1,2,3,\*</sup> , Mohamed Allam <sup>1,4</sup>, Faiza Boussora <sup>1,5</sup>, Axel Himmelbach <sup>1</sup>, Sara G. Milner <sup>1</sup>, Martin Mascher <sup>1</sup>  and Thorsten Schnurbusch <sup>1,3,\*</sup> 

<sup>1</sup> Leibniz Institute of Plant Genetics and Crop Plant Research (IPK), Corrensstr. 3, OT Gatersleben, 06466 Seeland, Germany; mohamed.allam@studenti.unitus.it (M.A.); boussora@ipk-gatersleben.de (F.B.); himmelba@ipk-gatersleben.de (A.H.); milner@ipk-gatersleben.de (S.G.M.); mascher@ipk-gatersleben.de (M.M.)

<sup>2</sup> Faculty of Agriculture, Cairo University, Giza 12613, Egypt

<sup>3</sup> Institute of Agricultural and Nutritional Sciences, Faculty of Natural Sciences III, Martin Luther University Halle-Wittenberg, 06120 Halle, Germany

<sup>4</sup> Faculty of Agriculture, Assuit University, Assuit 71526, Egypt

<sup>5</sup> Institute of Arid Lands of Medenine, Route du Djorf km 22.5, Médénine 4100, Tunisia

\* Correspondence: helmy.mohamed-youssef-ibrahim@landw.uni-halle.de (H.M.Y.); schnurbusch@ipk-gatersleben.de (T.S.); Tel.: 49-3455522683 (H.M.Y.)

Received: 19 October 2020; Accepted: 23 November 2020; Published: 26 November 2020



**Abstract:** Barley (*Hordeum vulgare* L.) is one of the major grain crops worldwide and considered as a model plant for temperate cereals. One of the barley row-type groups, named *intermedium*-barley, was used in our previous study where we reported that other genetic loci rather than *vrs1* and *Int-c* could play a role in lateral spikelet development and even in setting grains. To continue this work, we used phenotypic and genotypic data of 254 *intermedium*-spike barley accessions aimed at dissecting the genetic basis of development and grain traits of lateral and central spikelet using genome wide association (GWAS) analysis. After genotypic data filtering, 8,653 single-nucleotide polymorphism (SNPs) were used for GWAS analysis. A total of 169 significant associations were identified and we focused only on the subset of associations that exceeded the  $p < 10^{-4}$  threshold. Thirty-three highly significant marker-trait-associations (MTAs), represented in 28 different SNPs on all seven chromosomes for the central and/or lateral spikelet traits; such as kernel length, width, area, weight, unfilled spikelet and 1000-kernel weight, were detected. Highly significant associated markers were anchored physically using barley genome sequencing to identify candidate genes to either contain the SNPs or the closest gene to the SNP position. The results showed that 12 MTAs were specific for lateral spikelet traits, nine MTAs were specific for central spikelet traits and seven MTAs for both central and lateral traits. All together, the GWAS and candidate gene results support our hypothesis that lateral spikelet development could be regulated by loci different from those regulating central spikelet development.

**Keywords:** *intermedium* barley; spikelet development; fertility; central spikelet; lateral spikelet; GWAS

## 1. Introduction

Barley (*Hordeum vulgare* L.) is one of the most highly produced and consumed grains in the world, classified as the fourth major cereal crop globally [1]. It is an excellent experimental model for other cereals, due to its simple inbreeding and diploid genome. The barley inflorescence (called spike)

consists of three single-flowered spikelets—one central and two lateral spikelets per rachis node—that confer unique row-type identity to barley spikes. The main distinction among cultivars is made based on the number of kernel rows, i.e., two-rowed and six-rowed barley spikes. In two-rowed barley only the central spikelet is fertile, resulting in a single grain; whereas there are three grains at each node in six-rowed barleys [2]. Domestication evolution has influenced morphological and physiological features of crops [3,4] constituting the domestication syndrome, for which crops depend on humans for their growth and reproduction. Disabling sterility in the lateral florets seems a straightforward way to triple the number of grains at each rachis node, that was the route of six-rowed barley to arise from the ancestral two-rowed barley. The *intermedium*-barley might be a transition form from two- to six-rowed barley.

There are more than 400,000 accessions of barley in worldwide gene banks [5] that are considered as an excellent genetic resource to be efficiently utilized in barley breeding programs. These accessions differ in their genetic constitution. The genetic regulation of crop plant fitness (the mutant  $\times$  genotype interaction) across diverse environments [6] is under the focus of recent research. In addition, the high-quality barley reference genome sequence assembly became available in 2017 [7] and increased the interest of barley as an important model crop plant that now can be used to understand the genetic basis of barley (root, shoot, spike and spikelet) development. One of these genetic resources is known as *intermedium*-spike barley (*Hordeum vulgare* convar. *intermedium* (Körn.) Mansf. (syn. *H. intermedium* Körn.) that exhibits various degrees of lateral spikelet fertility and grain development, an intermediate form between two- and six-rowed types [8]. It might therefore provide an excellent opportunity to study the genetic basis controlling central and lateral spikelet fertility separately.

There are five row-type genes identified and characterised in barley, *Six-rowed spike 1* (*Vrs1*) encodes a homeodomain leucine zipper class I transcriptional activator [9,10] that inhibits the fertility of lateral spikelets. Functional mutations in *VRS1* are sufficient to generate a full six-rowed spike phenotype. Meanwhile *VRS2*, *VRS3*, *VRS4* and *VRS5/INTERMEDIUM-SPIKE C* (*INT-C*) encode homologues of SHORT INTERNODES (SHI), putative Jumonji C-type H3K9me2/me3 demethylase, LATERAL ORGAN BOUNDARY (LOB) and TEOSINTE BRANCHED1/CYCLOIDEA/PROLIFERATING CELL NUCLEAR ANTIGEN FACTOR 1 (TB1), respectively [11–16]. The *int-c* locus, located on chromosome 4HS [12,17], has been reported to control the fertility of lateral spikelets and accordingly the row-type identity.

In our previous work [18], we re-sequenced the entire *Vrs1*, *Vrs4* and *Int-c* genes in the *intermedium* barley collection. The results showed that some of the *intermedium* accessions showed a complete six-rowed phenotype while carrying the wildtype two-rowed allele of *Vrs1* [9] in combination with the six-rowed allele of *int-c* that is considered as a modifier of lateral spikelet development [12]. Additionally, we found that the *Int-c.a* allele was not able to overpower the suppression of lateral florets in the genomic background of wild barleys and two-rowed-like forms of *intermedium* barley accessions that carried known functional *Vrs4* alleles [13] and the *Vrs1.b* allele [18]. These results suggested that other genetic loci or allelic combinations could play a role in lateral spikelet development and even in setting grains.

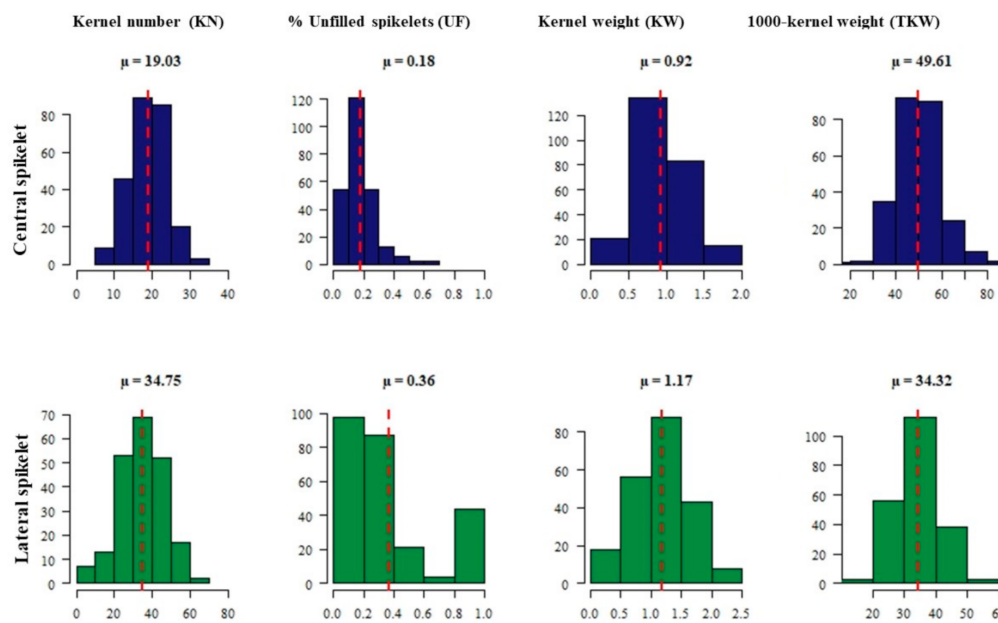
The rapid development of the genomic tools in DNA sequencing revolutionized the way to genetically improve important traits (e.g., yield traits) to bridge the gap between the world production and demands. One of the recent applications of next generation sequencing (NGS) is the genotyping by sequencing (GBS) approach that provides a huge amount of high-quality genome-wide genetic markers (single nucleotide polymorphisms, SNPs). These SNPs are suitable for studying diversity and marker-trait associations [19,20] and accelerate genetic breeding programs in crops [20]. Many powerful statistical genetics methods were also proposed to identify alleles controlling target traits. The associations between genotype and phenotype are dependent on linkage disequilibrium (LD), the non-random association of alleles, being broken down by many generations of recombination. A genome-wide association study (GWAS) is one of these useful methods that can be successfully applied to identify candidate genes for many important traits in barley as it tests the association between SNP markers and the phenotype of a target trait [21–23].

Here, we aimed to (1) better understand the genetic basis controlling spikelet development and growth. Moreover, we sought to (2) identify novel gene(s)/loci that could play a role in lateral spikelet development and even setting grains in the *intermedium* barley collection. Association mapping analyses have been used to identify SNP markers associated with heading and flowering time in barley (see, e.g., [24,25]). However, the genetic variation of spike morphology and its components is less studied. In this study, we are the first in barley to dissect the genetic basis of development and grain traits of lateral and central spikelet separately. Fourteen spikelet morphology and fertility-related traits were scored for central and lateral spikelets distinctly, and then phenotypic-genotypic association analysis was performed. Further quantitative trait loci (QTLs) analysis identified candidate genes in this study that will help to propose a genetic network (among novel and already known genes) underlying spikelet fertility traits in barley.

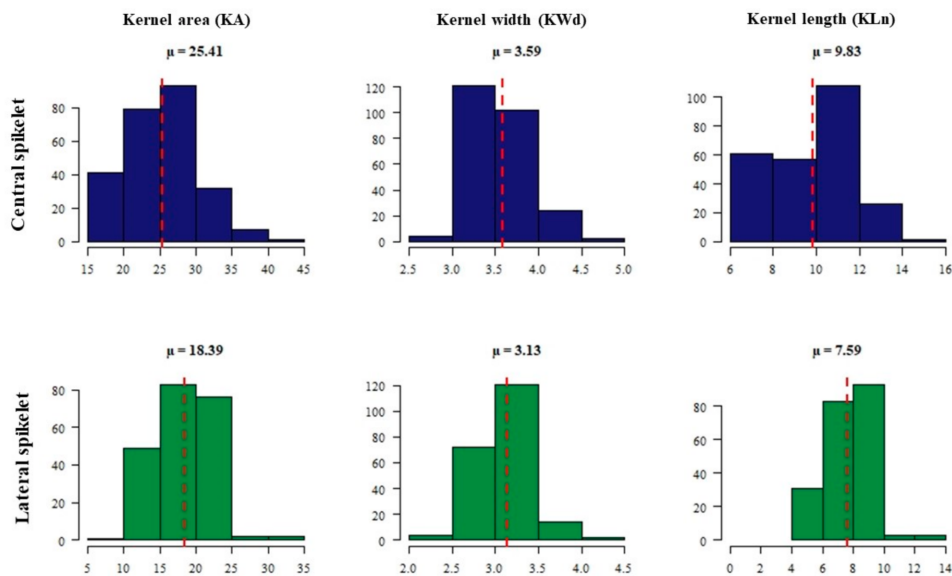
## 2. Results

### 2.1. Phenotypic Data Analysis

Phenotypic variation was observed based on the average of the four biological replicates of each accession among the 254 *intermedium*-spike barley accessions for all traits. Frequency distributions of all accessions for the investigated traits are presented in Figures 1 and 2.



**Figure 1.** The phenotype's frequency distribution histogram of kernel number (KN), percentage (%) of unfilled spikelets (UF), kernel weight (KW) and 1000-kernel weight (TKW) for central (blue color) and lateral (green color) spikelets among 254 *intermedium*-spike barley accessions.



**Figure 2.** The phenotype's frequency distribution histogram of kernel area (KA), kernel width (KWd), kernel length (KLn) and spikelet awn length (SAL) for central (blue color) and lateral (green color) spikelets among 254 *intermedium*-spike barley accessions.

The *intermedium*-spike barley collection showed different fertility degrees for lateral spikelets, with a range from 0%, i.e., completely infertile (two-rowed like), to 100% fertile (six-rowed like), with a mean value of 36%. However, only an average of 18% infertility was noticed for central spikelets. As expected, a higher number of kernels (KN) and lower 1000-kernel weight (TKW) of lateral spikelets were observed. The mean value for KN per spike of central spikelets was 19.03, while it was 34.75 for lateral spikelets. The mean value for TKW of central spikelets was about 50 g, while for lateral spikelets it was 34.32 g (Figure 1).

The average of kernel area (KA) of central spikelets was 25.41 mm<sup>2</sup>, with a range from 17 to 42 mm<sup>2</sup>. While, KA of lateral spikelets (18.39 mm<sup>2</sup>) was smaller by about 27.6%, with a range from 9.7 to 33.78 mm<sup>2</sup>. The mean values of kernel width (KWd) and length (KLn) were 3.59 and 9.83 mm<sup>2</sup>, respectively, of central spikelets. However, KWd and KLn of lateral spikelets were smaller by about 12.8% and 22.4%, respectively (Figure 2).

## 2.2. SNP Data and Population Structure Analysis

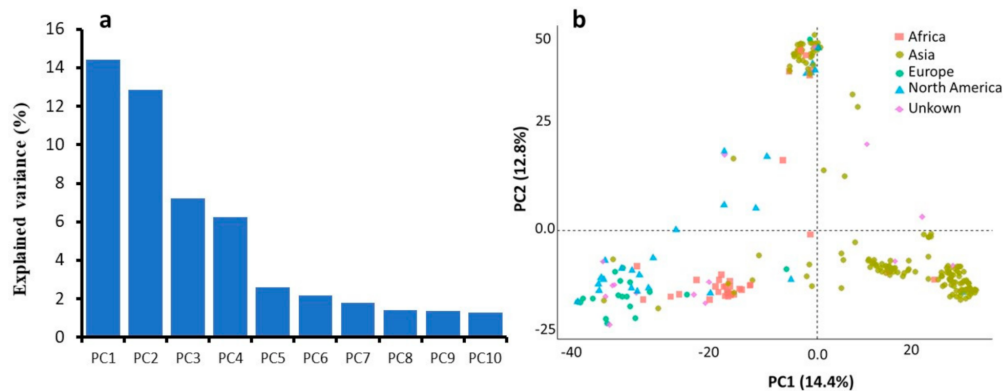
A total of 8,653 high quality SNPs distributed on the seven chromosomes (Supplementary Table S2) were selected for further analyses. The highest number of SNPs (17%, 1,458 of 8,653) was located on chromosome 2H and covered from nucleotide position 30,562 bp to 765,158,691 bp of the chromosome, while chromosome 4H contained the least number of SNPs (10%, 905 of 8653) and covered from 54,127 bp to 646,290,210 bp of the chromosome. The average number of SNPs/Mb was the lowest (7.4) on chromosome 1H and the highest (10.2) on chromosome 7H.

The population structure in the panel of *intermedium* barley was analysed using 8653 SNPs and a model-based approach in STRUCTURE program. The number of subpopulation (K) was plotted against the  $\Delta K$  calculated and the peak of the broken line graph was observed at K = 6 (Supplementary Materials), which indicates that the most appropriate number of clusters are represented by six sub-groups in this population.

A scree plot generated to visualize the fraction of variance explained by each of the 10 principle components, showed that PC1 and PC2 represented the highest proportion of the total variance, that together described about 27% of the population stratification (Figure 3a). PC1 explained about 14.4%, and PC2 explained about 12.8% of the total variance. The results of the PCA highlighted



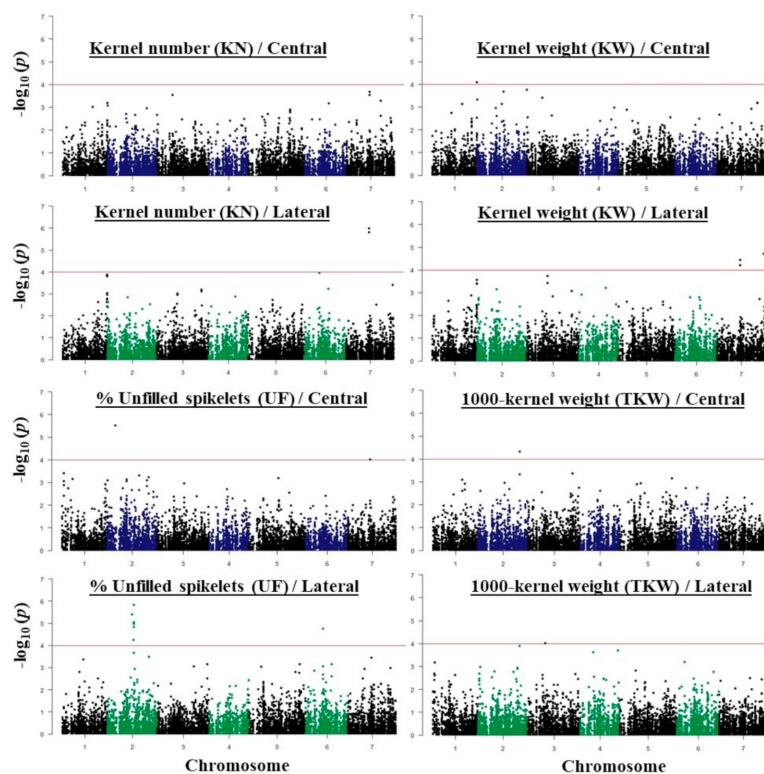
different sub-populations could be explained by geographical differences between accessions with few exceptions (Figure 3b).



**Figure 3.** Plots for the principal components (PCs). (a) The scree plot shows the variance explained by the first 10 PCs. (b) The distribution of the 254 *intermediuim*-spike barley lines in PC1 vs. PC2 ('continents' used as covariate).

### 2.3. GWAS Analysis and Candidate Genes

The results of a GWAS conducted for fertility and kernel traits of the *intermediuim*-spike barley are presented in Figures 4 and 5. The highest number of associations was found on chromosome 2H (41 associations) and the lowest was on chromosome 4H and 6H (13 associations). In this study, we focused only on the subset of 33 highly significant associations represented in 28 different SNPs that exceeded the  $p < 10^{-4}$  threshold (Tables 1 and 2).



**Figure 4.** Manhattan plots of genome wide association (GWAS) conducted on kernel number (KN), percentage (%) of unfilled spikelets (UF), kernel weight (KW), and 1000-kernel weight (TKW) for central spikelets and lateral spikelets.

**Table 1.** Significant single-nucleotide polymorphisms (SNPs) showing the highest marker-trait associations for the central spikelet.

Traits	SNP ID	Chr	SNP Physical Position	Ref Allele	Alt Allele	Gene	HC/LC	Annotation
Kernel number "KN"	m1_556903593	1H	556903593	T	C	HORVU1Hr1G094980	HC_G	Early flowering 3 (ELF3)
	m7_168762979	7H	168762979	G	T	HORVU7Hr1G048880	HC_G	Sec14p-like phosphatidylinositol transfer family protein
	m7_168763009	7H	168763009	C	T	HORVU7Hr1G048880	HC_G	Sec14p-like phosphatidylinositol transfer family protein
% Unfilled spikelets "UF"	m2_35441939	2H	35441939	C	T	HORVU2Hr1G001640	HC_G	Mediator of RNA polymerase II transcription subunit 15a
	m7_109046266	7H	109046266	C	T	HORVU7Hr1G040320	LC_u	undescribed protein
Kernel weight "KW"	m1_556903593	1H	556903593	T	C	HORVU1Hr1G094980	HC_G	Early flowering 3
	m2_647258179	2H	647258179	A	G	HORVU2Hr1G091170	HC_G	Expansin B3
	m2_653986096		653986096	C	A	HORVU2Hr1G092530	HC_G	Aldehyde dehydrogenase family 3 member F1
Thousand kernel weight "TKW"	m2_727927744	2H	727927744	T	C	HORVU2Hr1G113260	HC_G	N-terminal nucleophile aminohydrolases (Ntn hydrolases) superfamily protein
	m2_727927750	2H	727927750	G	T	HORVU2Hr1G113260	HC_G	N-terminal nucleophile aminohydrolases (Ntn hydrolases) superfamily protein
	m2_727927781	2H	727927781	G	A	HORVU2Hr1G113260	HC_G	N-terminal nucleophile aminohydrolases (Ntn hydrolases) superfamily protein
	m3_661788536	3H	661788539	A	C	HORVU3Hr1G099530	HC_G	UDP-Glycosyltransferase superfamily protein
	m3_661788539	3H	661788539	T	G	HORVU3Hr1G099530	HC_G	UDP-Glycosyltransferase superfamily protein
	m3_661788540	3H	661788539	A	G	HORVU3Hr1G099530	HC_G	UDP-Glycosyltransferase superfamily protein
	m3_661788542	3H	661788539	A	T	HORVU3Hr1G099530	HC_G	UDP-Glycosyltransferase superfamily protein

Table 1. Cont.

Traits	SNP ID	Chr	SNP Physical Position	Ref Allele	Alt Allele	Gene	HC/LC	Annotation
Kernel area "KA"	m3_661788536	3H	661788539	A	C	HORVU3Hr1G099530	HC_G	UDP-Glycosyltransferase superfamily protein
	m3_661788539	3H	661788539	T	G	HORVU3Hr1G099530	HC_G	UDP-Glycosyltransferase superfamily protein
	m3_661788540	3H	661788539	A	G	HORVU3Hr1G099530	HC_G	UDP-Glycosyltransferase superfamily protein
	m3_661788542	3H	661788539	A	T	HORVU3Hr1G099530	HC_G	UDP-Glycosyltransferase superfamily protein
	m4_557589531	4H	557589531	C	T	HORVU4Hr1G067140	HC_G	Protein kinase superfamily protein
	m7_548419008	7H	548419008	G	T	HORVU7Hr1G090180	LC_u	undescribed protein
Kernel width "KWd"	m2_727927744	2H	727927744	T	C	HORVU2Hr1G113260	HC_G	N-terminal nucleophile aminohydrolases (Ntn hydrolases) superfamily protein
	m2_727927750	2H	727927750	G	T	HORVU2Hr1G113260	HC_G	N-terminal nucleophile aminohydrolases (Ntn hydrolases) superfamily protein
	m2_727927781	2H	727927781	G	A	HORVU2Hr1G113260	HC_G	N-terminal nucleophile aminohydrolases (Ntn hydrolases) superfamily protein
Kernel Length "KLn"	m7_548419008	7H	548419008	G	T	HORVU7Hr1G090180	LC_u	undescribed protein

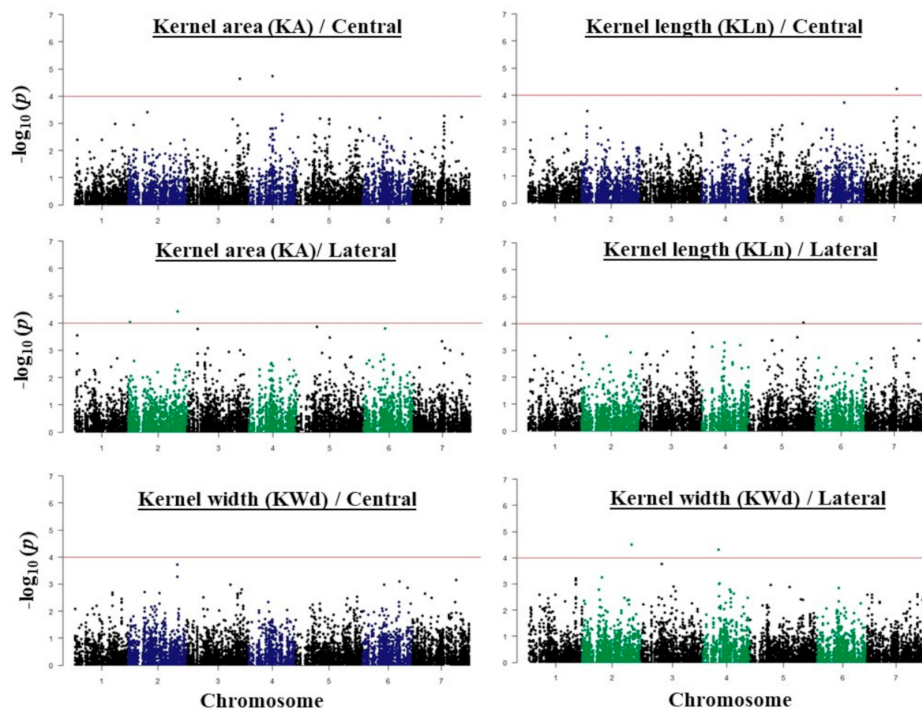
**Table 2.** Significant SNPs showing the highest marker-trait associations for the lateral spikelet.

Traits	SNP ID	Chr	SNP Physical Position	Ref Allele	Alt Allele	Gene	HC/LC	Annotation
Kernel number "KN"	m7_168762979	7H	168762979	G	T	HORVU7Hr1G048880	HC_G	Sec14p-like phosphatidylinositol transfer family protein
	m7_168763009	7H	168763009	C	T	HORVU7Hr1G048880	HC_G	Sec14p-like phosphatidylinositol transfer family protein
	m7_651215909	7H	651215909	C	T	HORVU7Hr1G120030	HC_G	Delta(24)-sterol reductase
% Unfilled spikelets "UF"	m2_647258179	2H	647258179	A	G	HORVU2Hr1G091170	HC_G	Expansin B3
	m2_649489024	2H	649489024	G	A	HORVU2Hr1G091760	LC_TE	Retrotransposon protein, putative, unclassified
	m2_649557855	2H	649557855	T	C	HORVU2Hr1G091830	HC_G	undescribed protein
	m2_651372029	2H	651372029	G	A	HORVU2Hr1G092180	HC_G	Elongation factor G
	m2_653986096	2H	653986096	C	A	HORVU2Hr1G092530	HC_G	Aldehyde dehydrogenase family 3 member F1
	m2_654165703	2H	654165703	A	G	HORVU2Hr1G092600	HC_G	multiprotein bridging factor 1A
	m2_654165718	2H	654165718	C	G	HORVU2Hr1G092600	HC_G	multiprotein bridging factor 1A
	m2_654165739	2H	654165739	C	A	HORVU2Hr1G092600	HC_G	multiprotein bridging factor 1A
	m6_115087828	6H	115087828	G	T	HORVU6Hr1G028720	HC_G	V-type ATP synthase subunit D
Kernel weight "KW"	m7_168762979	7H	168762979	G	T	HORVU7Hr1G048880	HC_G	Sec14p-like phosphatidylinositol transfer family protein
	m7_168763009	7H	168763009	C	T	HORVU7Hr1G048880	HC_G	Sec14p-like phosphatidylinositol transfer family protein
	m7_651215909	7H	651215909	C	T	HORVU7Hr1G120030	HC_G	Delta(24)-sterol reductase
Thousand kernel weight "TKW"	m2_727927744	2H	727927744	T	C	HORVU2Hr1G113260	HC_G	N-terminal nucleophile aminohydrolases (Ntn hydrolases) superfamily protein
	m2_727927750	2H	727927750	G	T	HORVU2Hr1G113260	HC_G	N-terminal nucleophile aminohydrolases (Ntn hydrolases) superfamily protein
	m3_208358229	3H	208358229	C	A	HORVU3Hr1G036930	HC_G	Mediator of RNA polymerase II transcription subunit 27
	m4_26351423	4H	26351423	G	A	HORVU4Hr1G009300	HC_G	HXXXD-type acyl-transferase family protein

Table 2. Cont.

Traits	SNP ID	Chr	SNP Physical Position	Ref Allele	Alt Allele	Gene	HC/LC	Annotation
Kernel area "KA"	m2_8906731	2H	8906731	T	G	HORVU2Hr1G003720	HC_G	GDSL esterase/lipase
	m2_727927744	2H	727927744	T	C	HORVU2Hr1G113260	HC_G	N-terminal nucleophile aminohydrolases (Ntn hydrolases) superfamily protein
	m2_727927750	2H	727927750	G	T	HORVU2Hr1G113260	HC_G	N-terminal nucleophile aminohydrolases (Ntn hydrolases) superfamily protein
	m3_208358229	3H	208358229	C	A	HORVU3Hr1G036930	HC_G	Mediator of RNA polymerase II transcription subunit 27
	m7_548419008	7H	548419008	G	T	HORVU7Hr1G090180	LC_u	undescribed protein
Kernel width "KWd"	m2_727927744	2H	727927744	T	C	HORVU2Hr1G113260	HC_G	N-terminal nucleophile aminohydrolases (Ntn hydrolases) superfamily protein
	m2_727927750	2H	727927750	G	T	HORVU2Hr1G113260	HC_G	N-terminal nucleophile aminohydrolases (Ntn hydrolases) superfamily protein
	m3_208358229	3H	208358229	C	A	HORVU3Hr1G036930	HC_G	Mediator of RNA polymerase II transcription subunit 27
	m4_26351423	4H	26351423	G	A	HORVU4Hr1G009300	HC_G	HXXXD-type acyl-transferase family protein
Kernel Length "KLn"	m5_619972700	5H	619972700	A	G	HORVU5Hr1G104750	HC_G	Bifunctional inhibitor/lipid-transfer protein/seed storage 2S albumin superfamily protein none





**Figure 5.** Manhattan plots of GWAS conducted on KA, KWd, and KLn, for central spikelets and lateral spikelets.

The significant marker-trait association (MTA) was detected on chromosome 1H, at 556,903,593 pb (m1\_556903593) for KW and KN of central spikelets only. Interestingly we found that this SNP is located in the Early flowering 3 (ELF3) gene, in the exon at 45 bp from the start codon (Table 1). Another two highly significant MTAs were detected on chromosome 7H (m7\_168762979 and m7\_168763009) for KN of both central and lateral spikelets. These two markers were also found in KW of lateral spikelets only. At this position we found that these two SNPs are located in the Sec14p-like gene (Tables 1 and 2). Another highly significant association was detected for KW of lateral spikelets and also located on chromosome 7H (m7\_651215909). However, this marker was only identified for KN of lateral spikelets (Table 2). For UF, ten significant associations were detected on chromosome 2H: for central spikelets at 35,441,939 bp, while located in the range of 647,258,179 to 654,165,739 bp for lateral spikelets. From the candidate genes at latter position expansin B3, elongation factor G, aldehyde dehydrogenase family 3 member F1 and multiprotein bridging factor 1A were found only for lateral spikelets (Table 2). Interestingly, in this region the Vrs1 gene is co-located (chr2H:652,031,295 to 652,032,562) (Figure 4). Two other MTAs were found for UF, on chromosome 7H at 109,046,266 bp for central spikelets, and on chromosome 6H at 115,087,828 bp for lateral spikelets.

On chromosome 2H, the SNPs m2\_727927744, m2\_727927750 and m2\_727927781 representing the same genomic region were detected for TKW and KWd of central and lateral spikelets and KA of lateral spikelets. However, another region was found on chromosome 3H (m3\_661788536, m3\_661788539, m3\_661788540 and m3\_661788542), associated with TKW and KA, but only for central spikelets, while m3\_208358229 was detected for TKW, KA, and KWd, but only for lateral spikelets. On chromosome 4H at 26,351,423 bp, another MTA was detected for TKW and KWd for lateral spikelets. On chromosome 7H, marker (m7\_548419008) was detected for KA of both central and lateral spikelets, and KLn of lateral spikelets only.

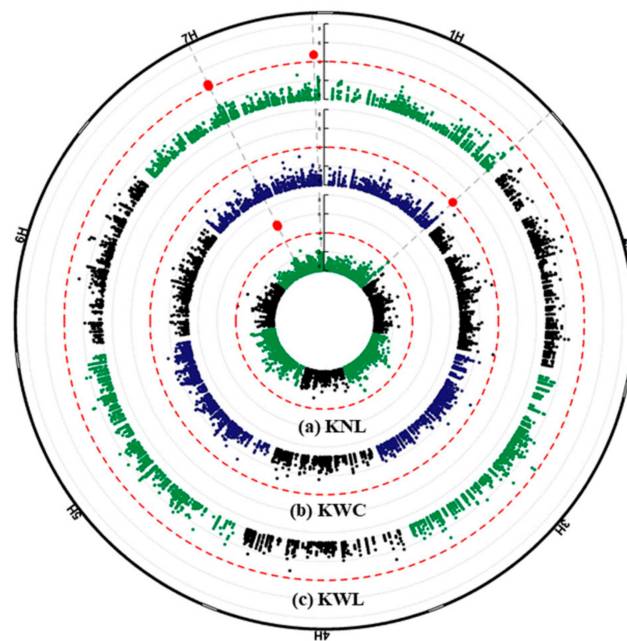
### 3. Discussion

In this study, we focused on a better understanding of the spikelet development and its effects on barley grain production. Yield is a complex agronomic trait, kernel number per spike and kernel weight

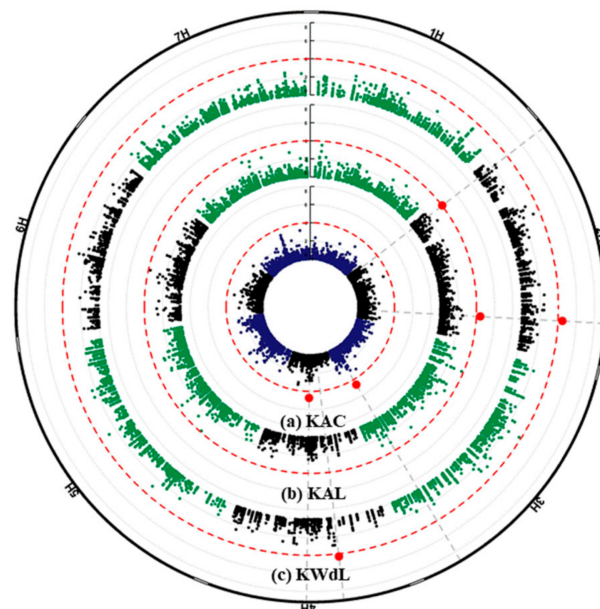
and size are key factors determining final barley yield. Although some QTLs related to yield traits have been identified, such as grain length [26–30], TKW [29–33], grain width [26,27,29,30,32,34,35], grain area [26,29,34] on almost all seven chromosomes, not many genes affecting yield and spike architecture have been identified so far [2,9,12,36].

The current study is the first report exploring the genetic basis of natural variation in separated central and lateral spikelets of *intermedium*-spike barleys for morphological and fertility-related traits along the barley inflorescence using GWAS. We phenotyped mature spikes from 254 *intermedium*-spike barley accessions and found 169 significant MTAs position on all the seven chromosomes that showed associations with lateral and/or central spikelets traits. Here we focused our attention only on the 28 highest significant ones (Tables 1 and 2) out of which 12 MTAs positions were detected for the lateral spikelets traits, nine MTAs for the central spikelets traits and seven MTAs positions for both central and lateral traits. From all of these SNP positions, only in a single case the well-known row-type gene *Vrs1* is co-located on Chr. 2H at position chr2H:652,031,295–652,032,562. Provided that most *intermedium* barley carry the functional two-rowed allele at *vrs1* and the six-rowed allele at *Int-c* [18], we assume that allelic variation is mainly fixed for both genes in our panel. Given that variation might be present but in low frequency at both genes, obtaining significant peaks from those genes appears less likely. We therefore assume that the peak on 2 HL, close to the *Vrs1* gene, is most likely related to other genes in LD. Nevertheless, these results support our previous finding that different loci, i.e., other than *vrs1* and *Int-c*, may control lateral and/or central spikelet development in this *intermedium* barley collection [18].

The thirteen MTAs detected for the lateral spikelets are located on chromosomes 2H to 7H. One of the interesting ones is located on chromosome 2H between 649,489,024 and 654,165,739 nucleotide positions for the %UF containing approx. ~160 HC genes. Interestingly, we found in this region *multiprotein bridging factor 1A* gene that carried three of the SNPs (m2\_654165703, m2\_654165718 and m2\_654165739). Tsuda et al. (2004) [37] reported that *multiprotein bridging factor 1 (MBF1)* in *Arabidopsis* functions as a transcriptional co-activator that mediates transcriptional activation by bridging between an activator and a TATA-box binding protein (TBP). Annotations of genes located on the other positions on chromosome 2H encode three different proteins; *Expansin B3* at m2\_647258179, *Elongation factor G (EF-G)* at m2\_651372029 and *Aldehyde dehydrogenase family 3member F1* at m2\_653986096. All of these genes have unknown effects on the lateral spikelet development of barley. For lateral spikelets, an MTA at nucleotide position (208,358,229) on chromosome 3H was detected for KWd, KA and TKW, and another MTAs at nucleotide position (26,351,423) on chromosome 4H was detected but only for the KWd and TKW. Another MTAs on chromosome 2H (positions: m2\_727927744 and m2\_727927750) were detected for the KWd and KA for lateral spikelets (Table 2, Figure 6) as well as MTAs on chromosome 7H at positions (m7\_168762979 and m7\_168763009) for the KN and KW of lateral spikelets (Table 2, Figure 7). These results suggested a high correlation and effect of KWd KA and of KN, KW on TKW. The two main known row-type genes affecting barley grain size or weight that have been characterized so far are *Vrs1* located on chromosome 2 HL [9,38] and *Int-c* located on chromosome 4 HS [12] at a position 17,599,033–17,600,737 which is about 8.8 Mb distant from the detected SNP on 4 HS (m4\_26351423). These results re-confirmed our previous findings that *Vrs1* and *Int-c* are not the only contributors affecting KWd, KA and TKW in this collection [18]. However, future work is required for validating the detected candidate regions by using either other GWAS panels or biparental mapping populations.



**Figure 6.** Circular Manhattan plots of genome-wide association scans for central (blue) and lateral (green) spikelets for (a) Kernel number (KN), (b) and (c) kernel weight (KW). The significance level of marker-trait associations ( $-\log p$  values) is represented by the vertical scale bar. Individual chromosomes are represented on the outer circle and separated from each other by white borders.  $-\log p$  values ( $p < 10^{-4}$ ) are indicated by dashed circles. Genomic regions of detected QTL on the respective chromosomes are colored in red (outer circle).

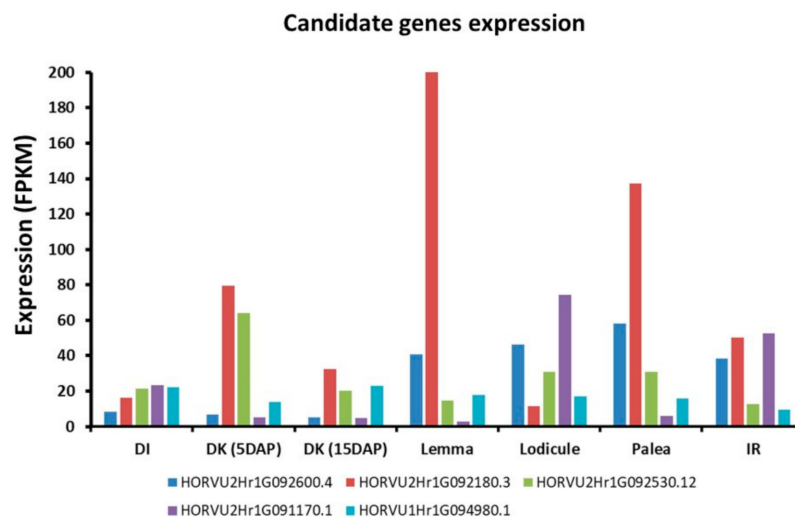


**Figure 7.** Circular Manhattan plots of genome-wide association scans for central (blue) and lateral (green) spikelets for (a) and (b) kernel area (KA), and (c) kernel width (KWd). The significance level of marker-trait associations ( $-\log p$  values) is represented by the vertical scale bar. Individual chromosomes are represented on the outer circle and separated from each other by white borders.  $-\log p$  values ( $p < 10^{-4}$ ) are indicated by dashed circles. Genomic regions of detected QTL on the respective chromosomes are colored in red (outer circle).

On chromosome 3H, we detected highly significant MTAs for KA and TKW but only for the central spikelet, suggesting that there are different loci in central spikelets playing a role in controlling

these traits. Interestingly, the four detected SNPs (m3\_661788536, m3\_661788539, m3\_661788540 and m3\_661788542) are located in a single gene, a uridine-diphosphate (UDP)-glycosyltransferase superfamily protein. We found that three of these SNPs are causing changes in the amino acid sequence of this gene; m3\_661788539 caused a serine (S) to arginine (R) change and m3\_661788540 and m3\_661788542 caused arginine (R) to glycine (G) changes at amino acid positions three and four, respectively. UDP-glycosyltransferases are enzymes that work on a large and diverse group of highly complex and diverse substrates such as different phytohormones (such as auxin, cytokinin and salicylic acid) and flavonoids as well as many others substrates to regulate plant growth and development, including biotic and abiotic resistance [39,40]. Their assigned function supports a putative role as a candidate gene in central spikelet development in barley; however, also in this case follow-up experiments are urgently required to verify this candidate.

Interestingly, we detected an MTA for central kernel number and kernel weight at 556,903,593 bp position on chromosome 1H. The annotation data showed that this SNP is located at *ELF3*. The *ELF3* gene is responsible for suppressing flowering under non-inductive photoperiods by regulating gibberellin production and *FLOWERING LOCUS T 1 (FT1)* expression [41]. These results support our previous findings that phytohormones may have a key role during spikelet development in barley [14]. We also found MTA for two SNPs (m2\_727927744, m2\_727927750) on chromosome 2H for central and lateral spikelets associated with traits TKW, KA and KWd as well as MTA at a position 548,419,008 bp on chromosome 7H associated with KLn and KA for central and lateral spikelets. These results supported our hypothesis that KWd has an impact on KW, while KLn has an impact on KA. Notably, we found that all the candidate genes presented here are expressed in the developing inflorescence tissues (Figure 8).



**Figure 8.** Expression analysis of the most important candidate genes: HORVU1Hr1G094980—*early flowering 3 (ELF3)*; HORVU2Hr1G092600—*multiprotein bridging factor 1A*; HORVU2Hr1G092180—*elongation factor G*; HORVU2Hr1G092530—*aldehyde dehydrogenase family 3 member F1*; HORVU2Hr1G091170—*expansin B3* in the developing inflorescence tissues; DI—developing inflorescences (1–1.5 cm); DK (5 DAP)—developing kernels (5 DAP); DK (15 DAP)—developing kernels (15 DAP); IR—inflorescences rachis (35 DAP); FPKM—fragments per kilobase of exon model per million reads mapped. (Source: Colmsee et al. 2015 [42]; <https://apex.ipk-gatersleben.de/apex/f?p=284:10:::>).

All together, the GWAS and candidate gene results support our hypothesis that lateral spikelet development could be regulated by loci different from those regulating central spikelet development. Nevertheless, future experiments may have to reveal the nature of these insights in more detail.



## 4. Materials and Methods

### 4.1. Genetic Materials and Phenotyping

A set of 254 *intermedium*-spike barley accessions, classified based on lateral spikelet size as *intermedium*-spike barley [8] (*Hordeum vulgare* L. convar. *intermedium* (Körn.) Mansf.) from the German Federal ex situ Gene bank hosted at IPK-Gatersleben (Seeland, Germany), were used in this study (Supplementary Table S1). The association mapping panel is composed of spring and winter barley lines from different regions worldwide, originating from four continents (Asia, Africa, North America and Europe; Supplementary Materials) and 18 different countries (Supplementary Table S1 in Youssef et al. 2017a [18]).

Four plants from each accession of the *intermedium*-spike barleys were germinated in 96 wells trays and kept for 10–15 days till they reached the two- to three-leaf stage. At the three leaves stage, plantlets were vernalized for six weeks at 4 °C followed by hardening for one week under 12/12 h light conditions (day/night) and temperature 12 ± 2 °C day and night. Afterwards, each plant was transplanted into 14 cm diameter pot and grown in potting substrate (peatmoss fertilised with N|:P|:K, 14|:16|:18) in the greenhouse under long-day conditions (16 h/8 h) and temperature 20 ± 2 °C during the day and 16 ± 2 °C during the night, and labelled as the main culm. The pots were randomly arranged and periodically moved around to avoid positional effects. The plants were watered when needed.

Briefly, the data were scored from the spike of all labelled main culm plants after harvest for central/lateral spikelet traits, separately as follows: (1) the percentage of unfilled spikelets “UF” (%) were evaluated. Then central and lateral kernels were separated and the following grain traits were scored using a MARVIN grain analyser (GTA Sensorik GmbH, Neubrandenburg, Germany): (2) kernel number “KN”, (3) kernel length “KLn”, (4) kernel width “KWd”, (5) kernel area “KA” (6) kernel weight “KW” per main spike in gram (g), and (7) thousand kernel weight “TKW”. Four labelled spikes from the main culm of each accession were checked to obtain the average for all traits used for GWAS analysis.

### 4.2. Phenotypic Data Analysis

The experiment was arranged as a completely randomized design with 4 biological replications (plants) per accession. The frequency distribution of all traits was visualized using a histogram. Analysis of variance (ANOVA) of the phenotypic data was performed, which included 254 *intermedium*-spike barley accessions. All statistical analyses in this study were conducted using R 3.5.3 (R Core Team 2018 [43], Vienna, Austria). Data transformations were performed where necessary to normalise the trait distributions.

### 4.3. GBS-SNP Genotyping

All *intermedium* barley accessions were genotyped using genotyping-by-sequencing (GBS). DNA was extracted according to Doyle and Doyle protocol [44]. GBS was performed following (Mascher et al. 2013) [45]. Adapters were trimmed from reads with cutadapt [46] (Mascher 2011). Trimmed reads were mapped to the barley reference genome [7] with BWA-MEM [47], converted to BAM format with SAMtools [48] and sorted with Novosort (Novocraft Technologies Sdn Bhd, Petaling Jaya, Selangor, Malaysia, <http://www.novocraft.com/>). Multi-sample variant calling was done with SAMtools [49] using the command “mpileup” with the parameters “-DV”. The VCF output file was filtered with an AWK script available from [https://bitbucket.org/ipk\\_dg\\_public/vcf\\_filtering](https://bitbucket.org/ipk_dg_public/vcf_filtering). Only biallelic SNPs were considered. SNPs with minor allele frequency (MAF) of less than 5% and missing data of more than 20% were filtered out. Finally, we used 8653 SNPs for GWAS analysis of 254 *intermedium*-spike barley accessions.



#### 4.4. Population Structure and Kinship Analyses

The population structure of the panel was assessed using STRUCTURE software 2.3.4, Stanford, California, USA [50]. Numbers of hypothetical subpopulation ranging  $k = 1$  to 10 were assessed in an admixture model using 10,000 burn-in followed by 10,000 recorded Markov chain iterations. To estimate the sampling variance (robustness) of population structure inference, 10 independent runs were carried out for each  $k$ . The output from STRUCTURE was analyzed in STRUCTURE HARVESTER [51]. The  $\Delta K$  statistics based on the rate of change in the logarithm of the probability of likelihood [LnP(D)] value between successive  $k$  values [52] was used to predict the optimum number of subpopulations. Population structure was also investigated using principle component analysis (PCA) using the Genome Association and Prediction Integrated Tool (GAPIT) R package [53]. To determine the number of PCs to use in clustering and GWAS analysis, a scree plot was generated by plotting the percentage of variance explained by the first 10 PCs against the number of PCs. The resulting PC loading for each line were exported for the creation of the graph using the ggplot function of R [43]. A kinship matrix (K) was constructed using GAPIT by setting the 'group. from' and 'group. to' parameters equal to the population size and the 'group. by' parameter equal to one.

#### 4.5. Genome-Wide Association Mapping Analysis and Candidate Gene Identification

Genome-wide association mapping was performed with the Genome Association and Prediction Integrated Tool (GAPIT) [53]. Several models were tested for the identification of marker-trait associations (MTAs). A naive model utilizing only genotypic and phenotypic data using the general linear model (GLM). The principle component (P) model using the first (PC1) and second (PC2) components that were selected based on generated scree plots, were included in the model as fixed-effect covariate to correct for population structure. The kinship matrix (K) model, the relatedness between individuals was calculated and included in the model as random-effect covariate. Finally, the model (P + K), both the population structure PCs and kinship matrix K were included in mixed liner model (MLM) as fixed and random-effect covariates, respectively [54].

The optimal model (P + K) was selected based on the quantile-quantile (Q-Q) plots of  $p$ -values comparing the uniform distribution of the expected  $-\log_{10}(p)$  to the observed  $-\log_{10}(p)$  of all the evaluated traits (Supplementary Materials) and genome-wide association mapping was performed with the identified model. SNP sites with the lowest  $p$  value in the peak region ( $p < 10^{-4}$ ) were considered highly significant SNPs for phenotypic variation. We calculated the false discovery rate (FDR) at  $p$  value  $< 0.05$  to determine the significance level of the SNP and exclude false-positive MTAs. Loci with a physical distance less than 100 bp were considered the same locus. Visualization of the significant SNPs was reported using Manhattan plots generated using the R package *qqman* [55]. The significant (FDR ( $p < 0.05$ ) or highly significant (FDR ( $p < 0.01$ )) identified associated markers were anchored physically using the recently published barley genome sequence [7]. Based on the SNPs physical position we isolate the candidate genes to be either contain the SNPs or the closest gene to the SNP position using the BARLEX IPK server (<http://apex.ipk-gatersleben.de/apex/f?p=284:10>).

**Supplementary Materials:** The following are available online at <http://www.mdpi.com/2223-7747/9/12/1655/s1>, Figure S1: Model-based subdivision of population structure. The plot shows the estimation of the number of populations by delta K value ( $\Delta K$ )., Figure S2 The quantile-quantile (Q-Q) plots of  $p$ -values comparing the uniform distribution of the expected  $-\log_{10}(p)$  to the observed  $-\log_{10}(p)$  of all the evaluated traits., Table S1: Total number of *intermedium*-spike barley accessions and their origin., Table S2: The average number of SNPs/Mb for all chromosomes.

**Author Contributions:** T.S. and H.M.Y. conceived the study; H.M.Y., M.A., F.B., S.G.M., M.M. and A.H. performed the experiments; H.M.Y., M.A. and F.B. analysed the data; H.M.Y., M.A. and T.S. wrote the manuscript with input from all co-authors. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was supported by grants from the HEISENBERG Program of the German Research Foundation (DFG), grant No. SCHN 768/8-1 and SCHN 768/15-1 to TS.

**Acknowledgments:** The authors gratefully acknowledge the careful maintenance and supply of the *intermedium* barley accessions from the Federal ex situ Genebank, IPK, Germany. We are grateful to Michael Grau for supporting

the phenotypic characterization of the *intermedium* collection. We thank Ahmed Alqudah for his general advices on the analysis and critically reading a previous version of the manuscript; and Hend Ibrahim, Mechthild Pürschel and Corinna Trautewig for excellent technical assistance.

**Conflicts of Interest:** The authors declare that there is no conflict of interest.

## References

1. FAOSTAT. *Food and Agriculture Organization of the United Nations (FAO)*; FAOSTAT Database; FAO: Rome, Italy, 2016.
2. Koppolu, R.; Schnurbusch, T. Developmental pathways for shaping spike inflorescence architecture in barley and wheat. *J. Integr. Plant Biol.* **2019**, *61*, 278–295. [[CrossRef](#)]
3. Abbo, S.; van-Oss, R.P.; Gopher, A.; Saranga, Y.; Ofner, I.; Peleg, Z. Plant domestication versus crop evolution: A conceptual framework for cereals and grain legumes. *Trends Plant Sci.* **2014**, *19*, 351–360. [[CrossRef](#)] [[PubMed](#)]
4. Haas, M.; Schreiber, M.; Mascher, M. Domestication and crop evolution of wheat and barley: Genes, genomics, and future directions. *J. Integr. Plant Biol.* **2019**, *61*, 204–225. [[CrossRef](#)] [[PubMed](#)]
5. Knüpffer, H. Triticeae genetic resources in ex situ genebank collections. In *Genetics and Genomics of the Triticeae, Plant Genetics and Genomics: Crops and Models 7*; Muehlbauer, G.J., Feuillet, C., Eds.; Springer: Berlin/Heidelberg, Germany, 2009; pp. 31–79.
6. Allam, M.; Ordás, B.; Djemel, A.; Tracy, W.F.; Revilla, P. Linkage disequilibrium between fitness QTLs and the *sugary1* allele of maize. *Mol. Breed.* **2019**, *39*, 3. [[CrossRef](#)]
7. Mascher, M.; Gundlach, H.; Himmelbach, A.; Beier, S.; Twardziok, S.O.; Wicker, T.; Bayer, M. A chromosome conformation capture ordered sequence of the barley genome. *Nature* **2017**, *544*, 427–433. [[CrossRef](#)] [[PubMed](#)]
8. Mansfeld, R. Das morphologische System der Saatgerste, *Hordeum vulgare* L. *Theor. Appl. Genet.* **1950**, *20*, 8–24. [[CrossRef](#)]
9. Komatsuda, T.; Pourkheirandish, M.; He, C.; Azhaguvel, P.; Kanamori, H.; Perovic, D.; Lundqvist, U. Six-rowed barley originated from a mutation in a homeodomain-leucine zipper I-class homeobox gene. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 1424–1429. [[CrossRef](#)]
10. Sakuma, S.; Pourkheirandish, M.; Hensel, G.; Kumlehn, J.; Stein, N.; Tagiri, A.; Yamaji, N.; Ma, J.F.; Sassa, H.; Koba, T.; et al. Divergence of expression pattern contributed to neofunctionalization of duplicated HD-Zip I transcription factor in barley. *New Phytol.* **2013**, *197*, 939–948. [[CrossRef](#)]
11. Shuai, B.; Reynaga-Peña, C.; Springer, P. The lateral organ boundaries gene defines a novel, plant-specific gene family. *Plant Physiol.* **2002**, *129*, 747–761. [[CrossRef](#)]
12. Ramsay, L.; Comadran, J.; Druka, A.; Marshall, D.F.; Thomas, W.T.; Macaulay, M.; Hayes, P.M. INTERMEDIUM-C, a modifier of lateral spikelet fertility in barley, is an ortholog of the maize domestication gene TEOSINTE BRANCHED 1. *Nat. Genet.* **2011**, *43*, 169–172. [[CrossRef](#)]
13. Koppolu, R.; Anwar, N.; Sakuma, S.; Tagiri, A.; Lundqvist, U.; Pourkheirandish, M.; Rutten, T.; Seiler, C.; Himmelbach, A.; Ariyadasa, R.; et al. *Six-rowed spike4 (Vrs4)* controls spikelet determinacy and row-type in barley. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 13198–13203. [[CrossRef](#)] [[PubMed](#)]
14. Youssef, H.M.; Eggert, K.; Koppolu, R.; Alqudah, A.M.; Poursarebani, N.; Fazeli, A.; Lundqvist, U. VRS2 regulates hormone-mediated inflorescence patterning in barley. *Nat. Genet.* **2017**, *49*, 157–161. [[CrossRef](#)] [[PubMed](#)]
15. Bull, H.; Casao, M.C.; Zwirek, M.; Flavell, A.J.; Thomas, W.T.; Guo, W.; Druka, A. Barley *SIX-ROWED SPIKE3* encodes a putative Jumonji C-type H3K9me2/me3 demethylase that represses lateral spikelet fertility. *Nat. Commun.* **2017**, *8*, 936. [[CrossRef](#)] [[PubMed](#)]
16. van Esse, G.W.; Walla, A.; Finke, A.; Koornneef, M.; Pecinka, A.; von Korff, M. *Six-Rowed Spike3 (VRS3)* is a histone demethylase that controls lateral spikelet development in barley. *Plant Physiol.* **2017**, *174*, 2397–2408. [[CrossRef](#)]
17. Pourkheirandish, M.; Komatsuda, T. The importance of barley genetics and domestication in a global perspective. *Ann. Bot.* **2007**, *100*, 999–1008. [[CrossRef](#)]

18. Youssef, H.M.; Mascher, M.; Ayoub, M.A.; Stein, N.; Kilian, B.; Schnurbusch, T. Natural diversity of inflorescence architecture traces cryptic domestication genes in barley (*Hordeum vulgare* L.). *Genet. Resour. Crop Evol.* **2017**, *64*, 843–853. [[CrossRef](#)]
19. Wenzl, P.; Raman, H.; Wang, J.; Zhou, M.; Huttner, E.; Kilian, A. A DArT platform for quantitative bulked segregant analysis. *BMC Genom.* **2007**, *8*, 196. [[CrossRef](#)]
20. Raman, H.; Raman, R.; Nelson, M.N.; Aslam, M.N.; Rajasekaran, R.; Wratten, N.; Schondelmaier, J. Diversity Array Technology markers: Genetic diversity analyses and linkage map construction in rapeseed (*Brassica napus* L.). *DNA Res.* **2012**, *19*, 51–65. [[CrossRef](#)]
21. Waugh, R.; Jannink, J.; Muehlbauer, G.J.; Ramsay, L. The emergence of whole genome association scans in barley. *Curr. Opin. Plant Biol.* **2009**, *12*, 218–222. [[CrossRef](#)]
22. George, A.W.; Cavanagh, C. Genome-wide association mapping in plants. *Theor. Appl. Genet.* **2015**, *128*, 1163–1174. [[CrossRef](#)]
23. Liu, H.J.; Yan, J. Crop genome-wide association study: A harvest of biological relevance. *Plant J.* **2019**, *97*, 8–18. [[CrossRef](#)] [[PubMed](#)]
24. Maurer, A.; Draba, V.; Jiang, Y.; Schnaithmann, F.; Sharma, R.; Schumann, E.; Pillen, K. Modelling the genetic architecture of flowering time control in barley through nested association mapping. *BMC Genom.* **2015**, *16*, 290. [[CrossRef](#)] [[PubMed](#)]
25. Sannemann, W.; Huang, B.E.; Mathew, B.; Léon, J. Multi-parent advanced generation inter-cross in barley: High-resolution quantitative trait locus mapping for flowering time as a proof of concept. *Mol. Breed.* **2015**, *35*, 86. [[CrossRef](#)]
26. Ayoub, M.; Symons, S.; Edney, M.; Mather, D. QTLs affecting kernel size and shape in a two-rowed by six-rowed barley cross. *Theor. Appl. Genet.* **2002**, *105*, 237–247. [[CrossRef](#)]
27. Walker, C.K.; Ford, R.; Muñoz-Amatriaín, M.; Panozzo, J.F. The detection of QTLs in barley associated with endosperm hardness, grain density, grain size and malting quality using rapid phenotyping tools. *Theor. Appl. Genet.* **2013**, *126*, 2533–2551. [[CrossRef](#)]
28. Zhou, Q.; Dong, Y.; Shi, Q.; Zhang, L.; Chen, H.; Hu, C.; Li, Y. Verification and fine mapping of qGW1.05, a major QTL for grain weight in maize (*Zea mays* L.). *Mol. Genet. Genom.* **2017**, *292*, 871–881. [[CrossRef](#)]
29. Xu, X.; Sharma, R.; Tondelli, A.; Russell, J.; Comadran, J.; Schnaithmann, F.; Flavell, A.J. Genome-wide association analysis of grain yield-associated traits in a pan-European barley cultivar collection. *Plant Genome* **2018**, *11*, 170073. [[CrossRef](#)]
30. Wang, Q.; Sun, G.; Ren, X.; Du, B.; Cheng, Y.; Wang, Y.; Sun, D. Dissecting the genetic basis of grain size and weight in barley (*Hordeum vulgare* L.) by QTL and comparative genetic analyses. *Front. Plant Sci.* **2019**, *10*, 469. [[CrossRef](#)]
31. Comadran, J.; Russell, J.R.; Booth, A.; Pswarayi, A.; Ceccarelli, S.; Grando, S.; Benbelkacem, A. Mixed model association scans of multi-environmental trial data reveal major loci controlling yield and yield related traits in *Hordeum vulgare* in Mediterranean environments. *Theor. Appl. Genet.* **2011**, *122*, 1363–1373. [[CrossRef](#)]
32. Kalladan, R.; Worch, S.; Rolletschek, H.; Harshavardhan, V.T.; Kuntze, L.; Seiler, C.; Röder, M.S. Identification of quantitative trait loci contributing to yield and seed quality parameters under terminal drought in barley advanced backcross lines. *Mol. Breed.* **2013**, *32*, 71–90. [[CrossRef](#)]
33. Maurer, A.; Draba, V.; Pillen, K. Genomic dissection of plant development and its impact on thousand grain weight in barley through nested association mapping. *J. Exp. Bot.* **2016**, *67*, 2507–2518. [[CrossRef](#)] [[PubMed](#)]
34. Sameri, M.; Komatsuda, T. Localization of quantitative trait loci for yield components in a cross oriental × occidental barley cultivar (*Hordeum vulgare* L.). *Jpn. Agric. Res. Q.* **2007**, *41*, 195–199. [[CrossRef](#)]
35. Cu, S.T.; March, T.J.; Stewart, S.; Degner, S.; Coventry, S.; Box, A.; Eglinton, J. Genetic analysis of grain and malt quality in an elite barley population. *Mol. Breed.* **2016**, *36*, 129. [[CrossRef](#)]
36. Zwirek, M.; Waugh, R.; McKim, S.M. Interaction between row-type genes in barley controls meristem determinacy and reveals novel routes to improved grain. *New Phytol.* **2019**, *221*, 1950–1965. [[CrossRef](#)]
37. Tsuda, K.; Tsuji, T.; Hirose, S.; Yamazaki, K. Three Arabidopsis MBF1 homologs with distinct expression profiles play roles as transcriptional co-activators. *Plant Cell Physiol.* **2004**, *45*, 225–231. [[CrossRef](#)]
38. Sakuma, S.; Lundqvist, U.; Kakei, Y.; Thirulogachandar, V.; Suzuki, T.; Hori, K.; Shimada, Y. Extreme suppression of lateral floret development by a single amino acid change in the VRS1 transcription factor. *Plant Physiol.* **2017**, *175*, 1720–1731. [[CrossRef](#)]

39. Huang, J.; Pang, C.; Fan, S.; Song, M.; Yu, J.; Wei, H.; Yu, S. Genome-wide analysis of the family 1 glycosyltransferases in cotton. *Mol. Genet. Genom.* **2015**, *290*, 1805–1818. [[CrossRef](#)]
40. Xing, L.P.; He, L.Q.; Xiao, J.; Chen, Q.G.; Li, M.H.; Shang, Y.; Wang, X.E. An UDP-glucosyltransferase gene from barley confers disease resistance to Fusarium head blight. *Plant Mol. Biol. Rep.* **2017**, *35*, 224–236. [[CrossRef](#)]
41. Boden, S.A.; Weiss, D.; Ross, J.J.; Davies, N.W.; Trevaskis, B.; Chandler, P.M.; Swain, S.M. EARLY FLOWERING 3 regulates flowering in spring barley by mediating gibberellin production and FLOWERING LOCUS T expression. *Plant Cell* **2014**, *26*, 1557–1569. [[CrossRef](#)]
42. Colmsee, C.; Beier, S.; Himmelbach, A.; Schmutzer, T.; Stein, N.; Scholz, U.; Mascher, M. BARLEX—The Barley Draft Genome Explorer. *Mol. Plant* **2015**, 964–966. [[CrossRef](#)]
43. R Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2018.
44. Doyle, J.J.; Doyle, J.L. Isolation of plant DNA from fresh tissue. *Focus* **1990**, *12*, 13–15.
45. Mascher, M.; Wu, S.; Amand, P.S.; Stein, N.; Poland, J. Application of genotyping-by-sequencing on semiconductor sequencing platforms: A comparison of genetic and reference-based marker ordering in barley. *PLoS ONE* **2013**, *8*, e76925. [[CrossRef](#)] [[PubMed](#)]
46. Mascher, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet. J.* **2011**, *17*, 10–12.
47. Li, H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv* **2013**, arXiv:1303.3997.
48. Li, H.; Handsaker, B.; Wysoker, A.; Fennell, T.; Ruan, J.; Homer, N.; Durbin, R. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **2009**, *25*, 2078–2079. [[CrossRef](#)]
49. Li, H. A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics* **2011**, *27*, 2987–2993. [[CrossRef](#)]
50. Pritchard, J.K.; Stephens, M.; Donnelly, P. Inference of population structure using multi-locus genotype data. *Genetics* **2000**, *155*, 945–959.
51. Earl, D.A.; von Holdt, B.M. Structure harvester: A website and program for visualizing structure output and implementing the Evanno method. *Conserv. Genet. Resour.* **2012**, *4*, 359–361. [[CrossRef](#)]
52. Evanno, G.; Regnaut, S.; Goudet, J. Detecting the number of clusters of individuals using the software structure: A simulation study. *Mol. Ecol.* **2005**, *14*, 2611–2620. [[CrossRef](#)]
53. Lipka, A.E.; Tian, F.; Wang, Q.; Peiffer, J.; Li, M.; Bradbury, P.J.; Zhang, Z. GAPIT: Genome association and prediction integrated tool. *J. Bioinform.* **2012**, *28*, 2397–2399. [[CrossRef](#)]
54. Yu, J.; Pressoir, G.; Briggs, W.H.; Bi, I.V.; Yamasaki, M.; Doebley, J.F.; Kresovich, S. A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. *Nat. Genet.* **2006**, *38*, 203–208. [[CrossRef](#)] [[PubMed](#)]
55. Turner, S.D. qqman: An R package for visualizing GWAS results using Q-Q and Manhattan plots. *J. Open Source Softw.* **2018**, *3*, 731. [[CrossRef](#)]

**Publisher’s Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).