Insight into the genetic contribution of maximum yield potential, spikelet development and abortion in barley

Alqudah, Ahmad M.; Sharma, Rajiv; Börner, Andreas

Published in:
Plants, People, Planet

DOI:
10.1002/ppp3.10203

Print publication: 01/11/2021

Document Version
Publisher's PDF, also known as Version of record

Received: 27 July 2020 | Revised: 2 April 2021 | Accepted: 23 April 2021

DOI: 10.1002/ppp3.10203

Plants People Planet PPP Open Access

RESEARCH ARTICLE

Insight into the genetic contribution of maximum yield potential, spikelet development and abortion in barley

Ahmad M. Alqudah[1,2] | Rajiv Sharma[3] | Andreas Börner[1,2]

[1]Research Group Resources Genetics and Reproduction, Department Genebank, Leibniz Institute of Plant Genetics and Crop Plant Research (IPK), Stadt Seeland, Germany

[2]Institute of Agricultural and Nutritional Sciences, Martin Luther University Halle-Wittenberg, Halle (Saale), Germany

[3]Scotland's Rural College (SRUC), Edinburgh, UK

Correspondence
Ahmad M. Alqudah, Institute of Agricultural and Nutritional Sciences, Martin Luther University Halle-Wittenberg, Betty-Heimann-Str. 3, 06120, Halle (Saale), Germany.
Email: ahmad.alqudah@landw.uni-halle.de; ahqudah@gmail.com

Societal Impact Statement

To feed the world's ever-increasing population, new genetic approaches are required. Increasing the number of living spikelets is one promising way to improve grain yield. This, in turn, increases the number of spikelets per plant, thereby increasing the total yield. We present the first evidence for genetic control of alive spikelets in barley. Discovering natural variation as well as genomic regions associated with these traits will serve as a benchmark in future breeding for improving grain yield.

Summary

- The primary goal of most breeding programmes is to increase grain yield. However, one of the many methods for raising yield that is yet to be fully investigated is increasing the number of spikelets by minimising spikelet abortion. Spikelet abortion dramatically increases during the late reproductive phase, but the molecular and genetic mechanisms remain unknown. Here, we employed a phenotyping approach in which developed and undeveloped spikelets were detected and counted during spike development and their maximum yield potential (MYP) was investigated.
- We studied 20 agronomic and spikelet-related traits using a set of 184 diverse spring barley accessions under field conditions. By employing a set of >125K SNPs, GWAS was conducted.
- Our analysis revealed 26 genetic clusters associated with MYP and the number of developed and undeveloped spikelets. Most of the significant associated genomic regions were co-located near the candidate genes of phytohormones such as ABA, auxin, and cytokinin suggesting the importance of phytohormones in keeping spikelets alive, their development, and MYP.
- Our findings point to a potential link between jasmonic acid and the MYP, development and abortion of spikelets. We further provide genetic evidence that sugar-related genes and sucrose have the potential to regulate MYP, spikelet development and spikelet survival. Our findings can be used for marker-assisted breeding and as a resource for future molecular and genetic validation. Collectively, we propose a new genetic network linking spikelet-related traits to grain yield determinants.

KEYWORDS
alive spikelet, barley, candidate genes, GWAS, *Hordeum vulgare* L, maximum yield potential, phytohormones, spikelet development, sugar-related genes

1 | INTRODUCTION

Barley (*Hordeum vulgare* L.) is one of the most important cereal crops in the world that grows under a wide range of conditions. Barley production has increased in the last decades mainly due to the improvement in yield production of new barley cultivars and the deployment of new cultivation practices. Nevertheless, further improvements in yield production are required to meet the predicted increase in population growth and the plant's ability to cope with environmental stress.

The natural form of barley spike architecture has an indeterminate number of rachis nodes, which are the physical positions of the basic reproductive unit of grain yield known as the spikelet. The number of spikelets at each node is varied and controlled by numerous loci/genes; three single spikelets in *SIX-ROWED SPIKE 1* (*VRS1*), or supernumerary spikelets per node, for example *VRS2* and *VRS4* (McKim et al., 2018). The first sign of spikelet development appears in the form of a double ridge (leaf primordia ridges at the lower position and spikelet primordia ridges at the upper position) during the spikelet initiation phase (Kirby & Appleyard, 1987). Spikelet ridge development crosses several biological processes, including the formation of floral organs (glume, lemma, palea, stamen, pistil and carpel; McKim et al., 2018) until it reaches awn primordium (AP). The maximum number of spikelet primordia per spike is determined during the AP developmental stage, which is described as the maximum yield potential (MYP; Alqudah & Schnurbusch, 2014).

The spikelets develop further, and new distinct organs emerge, extending the filaments upwards, lengthening the anthers and increasing the size of the lodicules. According to Alqudah and Schnurbusch (2014), AP to awn tipping (TIP) is the most decisive developmental phase for keeping spikelets alive and determining final grain yield. This phase includes anther developmental stages, which consist of the white anther (WA), green anther (GA) and yellow anther (YA) stages (Kirby & Appleyard, 1987). In addition to their importance in fertilisation, anther developmental stages are also crucial in determining the number of alive spikelets (AS) or spikelet survival. Principally, the MYP includes undeveloped spikelets (incomplete spikelet structure and organs), which are mostly located at the top of the spike and are ultimately aborted at maturity. Therefore, counting the number of developed and undeveloped spikelets (DS and UDS) at the WA or GA stage is crucial to determine how many spikelets are alive (survival) and to understand the mechanism of spikelet development, its survival (alive) and abortion.

The biological meaning of AS is that it allows the survival of developed spikelets with full structure and all spikelet organs, resulting in increased grains. Spikelet abortion, which occurs at a specific developmental phase, reduces grain yield. The competition for assimilates among spikelets, the spikelet positions on the spike (basal, central and apical) and within the spikelet (lateral and central), and distortion of the vascular tissue and cell structure are a few of the factors that contribute to UDS (Alqudah & Schnurbusch, 2014; Arisnabarreta & Miralles, 2004, 2006; Kernich et al., 1997). Spikelet abortion appeared to be higher in six-rowed barley than in two-rowed barley. In the former, more number of spikes and awn primordia are formed resulting in greater competition among spikelets (Alqudah & Schnurbusch, 2014; Arisnabarreta & Miralles, 2004; del Moral et al., 2002; Miralles et al., 2000).

Alive spikelet is a complex trait influenced by genetic and environmental factors. In barley plants, spikelet survival was found to be an inherited trait (Alqudah & Schnurbusch, 2014). Environmental factors, such as temperature, photoperiod, light intensity, and soil characters, are crucial contributors to spikelet alive/survival (Alqudah & Schnurbusch, 2014; Arisnabarreta & Miralles, 2006; Kernich et al., 1997). Endogenous biological factors, including phytohormones (Marzec & Alqudah, 2018) and metabolites (sugar; Ghiglione et al., 2008), are also involved in the regulation of spikelet/floret development, which is essential for maintaining source–sink relationships (Lunn et al., 2014), meiosis or abnormality of sexual organs (Arisnabarreta & Miralles, 2006; Higgins, 2013) and autophagy with phytohormone crosstalk (Gou et al., 2019).

An increase in the number of AS and a decrease in spikelet abortion are important for improving yield. However, there is limited knowledge on the general factors influencing the number of AS, abortion, and the underlying genetic and molecular mechanisms. In addition, no genetic research has been conducted to identify the underlying genetic or molecular causes of MYP, developed or undeveloped spikelets, and AS or abortion in barley, as well as their contribution to the number of grains per spike.

Therefore, the aim of this study is to detect the genetic factors underlying the number of MYP, DS, UDS and AS. To this end, we phenotyped hundreds of spikes at three major developmental stages (MYP, GA and TIP) under field conditions. We demonstrate that there is ample natural variation in the studied traits, which leads to detection of previously unknown genomic regions and candidate genes. We propose that abortion candidate genes including phytohormone- and sugar-related genes have the potential to regulate MYP, spikelet development and spikelet survival.

2 | MATERIALS AND METHODS

2.1 | Plant materials, SNP genotyping and population stratification analyses

The number of EcoSeed spring barley collection obtained from a collection of 21,405 accessions stored at the Genebank, IPK-Gatersleben, Germany, was 184. This number, as described by Nagel

et al. (2019), represents 23 countries and includes 116 two-rowed and 68 six-rowed spike accessions, consisting of 105 cultivars, 11 breeding lines, three hybrid lines and 65 landraces. Accession names and details about their origin as well as phenotype information of the accessions are provided in Dataset S1.

The single nucleotide polymorphism (SNP) of the EcoSeed collection was discovered by applying the genotyping-by-sequencing (GBS) and Illumina 9K marker chip technologies. The SNP chip underwent a quality control and filtration process, applying a minor allele frequency of ≥5% (Alqudah, Sallam, et al., 2020). After conducting quality checks, 4809 SNPs markers out of 7791 SNPS and 122,213 SNPs out of 233,095 SNPs corresponded to EcoSeed collection were finally obtained. A total of 127,022 SNPs (4809 Illumina 9K + 122,213 GBS) were included in the GWAS analysis as they physically anchored to barley physical positions "Morex version 1" (Mascher et al., 2017; Milner et al., 2019). To calculate the population structure, principal component analysis was used which was computed in GAPIT (Lipka et al., 2012). To display the population structure, graphics were generated in R version (RStudio-Team, 2015) using the R package GGPLOT2 (Wickham, 2016). The kinship matrix along with three principal components was used to control false positives using the mixed linear model method implemented in Genome Association and Prediction Integrated Tool, GAPIT (Lipka et al., 2012).

The genome-wide pairwise estimates of linkage disequilibrium (LD) were calculated as a squared correlation between pairs of polymorphic SNPs (r^2) for the whole genome using GenStat 18 (VSN International, 2016). Finally, LD decay patterns were visualised by plotting the LD estimated as r^2 versus the distance between pairs of polymorphic SNPs (Mbp).

2.2 | Novel phenotyping of MYP and AS-related traits

The collection was grown under field conditions at the IPK, Germany campus during the spring 2018 growing season. Two hundred seeds of each accession were planted on 1 April 2018 directly into clay loam soil and were grown in a randomised block design, with three 1 m^2 plots of each accession spaced by two rows within each plot. Standard local agronomic practices were applied to the field plots. In total, six to nine main spikes of each accession (two to three from each plot) were randomly selected from the middle of the plots for phenotyping. These were dissected manually starting from the stem elongation stage (Alqudah & Schnurbusch, 2014) to measure and count the MYP, DS and UDS at GA and TIP developmental stages using portable USB Digital Microscope Magnification (Banik et al., 2020; Rather et al., 2019). Each genotype was dissected three to four times between the stem elongation stage and the heading stage (HD) to define MYP, GA and TIP developmental stages precisely. MYP is the total number of spikelets per spike, including DS and UDS as shown in Figure 1. The developmental signs of the GA and TIP stages were defined according to previous research (Alqudah & Schnurbusch, 2014; Kirby & Appleyard, 1987; Waddington et al., 1983). Phenotyping of MYP, AS, DS, UDS, GDD, spike length and density traits was measured as described in Table 1. The growing degree-days (GDD) for each developmental stage were collected when most of the plants in three plots of each accession reached a particular stage (Alqudah et al., 2014). The number of developed spikelets at HD and the number of grains per spike at harvest stages (Hrv) were counted (Figure 1). Spike length and density (the ratio between the developed spikelet or grain to spike length) were calculated. In addition, 15 main spikes per accession from each plot were randomly selected to determine grain number per spike (GPS). The AS percentage among the stages was calculated according to Alqudah and Schnurbusch (2014). For example, AS% of developed spikelets at the GA and Hrv (AS_Hrv_GA) is the ratio between them multiplied by 100.

FIGURE 1 Phenotypic variation in spikelet abortion among barley accessions at spike heading (HD) developmental stage and harvest (Hrv). Spikelet abortion is indicated by the red line

2.3 | Phenotypic data analysis and heritability estimation

Significant differences among accessions were determined using analysis of variance at $p \leq .05$ in Genstat v18 (VSN International, 2016). The Student's *t* test was used to compare trait differences attributed to row types (i.e. two-rowed vs. six-rowed), biological status (i.e. cultivars, breeding lines and landraces) and among geographical origin. Trait relationship was calculated using Pearson correlation analysis applying the corrplot function in R package corrplot (Wei & Simko, 2017). Residual maximum likelihood in mixed linear models was used to analyse the phenotypic data, and the best linear unbiased estimates (BLUEs) were used to determine the phenotypic

TABLE 1 Name and description of measured traits in spring barley

Name	Description of trait measurement
Alive spikelet	The developed spikelet with full structure and all spikelet organs which can survive
Percentage of alive spikelets between stage *X* and *Y*	AS% of developed spikelets/grains at the *Y* and *X* stages as the ratio between them multiplied by 100, where *X* is a specific stage and *Y* is stage prior to *X*:AS − Hrv − GA = DS(Hrv)/DS(GA) × 100%
Number of developed spikelets	The number of spikelets with full structure and all spikelet organs
Percentage of developed spikelets	The number of developed spikelets at stage *X* divided by the total number of spikelets at the same stage multiplied by 100
Green anther	The second stage of anther development when the anther changes to green
Growing degree-days	The accumulated thermal time (°C.day^{-1}) to reach specific development stage
Broad-sense heritability	The ratio of genetic variance and phenotypic variance (H^2)
Maximum yield potential	The total number of spikelets per spike, including developed and undeveloped spikelet
Spike density	The ratio between the developed spikelet or grain and spike length (grain per cm)
Spike length	The length of spike from base of spike till tip without awn (cm)
Awn tipping	The first awns visible from leaf
Number of undeveloped spikelets	The number of incomplete spikelets with missing structures or organs and mostly at the tip of the spike
Percentage of undeveloped spikelets	The number of undeveloped spikelets at stage *X* divided by the total number of spikelets at the same stage multiplied by 100

means of each trait for each accession using the lme4 package (Bates et al., 2015) as described in Alqudah, Sallam, et al. (2020). Finally, broad-sense heritability (H^2) for each trait was calculated as a ratio of the genetic variance and the phenotypic variance that includes the error variance.

$$H^2 = \frac{\sigma^2 G}{\sigma^2 G + \frac{\sigma^2 e}{Br}}$$

where σ^2G is the variance of genotypes (accessions), σ^2e is the variance of error, *B* is the number of blocks and *r* is the number of replicates.

2.4 | Genome-wide models and assessments of statistical significance

GWAS analysis was performed using the GAPIT R package (Lipka et al., 2012) as described above using the mixed linear model method, which evaluated the 127,022 SNPs and BLUE values of studied traits of the accessions. Mixed linear model was used for GWAS analysis as a powerful model for association detection and controlling the population structure (Alqudah, Sallam, et al., 2020). More details about the mixed linear model can be found in Lipka et al. (2012).

False discovery rate (FDR) was calculated for each trait separately, and the association signals that passed the threshold of FDR at 0.001 (-$\log_{10}$ *p*-values ≥FDR) were used for further analyses (Alqudah, Sallam, et al., 2020). For each measured trait, Manhattan plots were generated to visualise the marker–trait associations across the seven barley chromosomes using the R package ShinyAIM (Hussain et al., 2018).

2.5 | Genotype–phenotype network

SNP interactions across the traits were calculated using the Network-Based Genome-Wide Association Studies (NETGWAS) R package (Behrouzi et al., 2017). SNPs passing the FDR (threshold at 0.001) were used in this analysis. The complex genotype–phenotype network analysis evaluates the genetic data of multi-SNPs with multi-phenotypic traits, and the interaction between them; this helps to better understand the genetic basis of such complex traits. Intra- and inter-chromosomal interactions were also analysed using the NETGWAS R package (Behrouzi et al., 2017). The most significant SNPs (key SNPs) directly connected with the studied traits (key traits) were used in locating candidate genes. The corresponding positions of these SNPs along with traits were finally plotted using Synthesis-View as a data visualisation application (http://visualization.ritchielab.org/).

2.6 | Candidate gene detection and validation

Key SNPs were used to define the closest candidate gene within the LD interval based on their physical position using barley database BARLEX (Colmsee et al., 2015) that hosts barley genome assembly. BARLEX is a web-based platform that contains gene annotations, gene ontology (GO), and description of barley gene space and gene annotations (https://apex.ipk-gatersleben.de/apex/f?p=284:10::::::). Only high confidence candidate genes with associated key SNPs within the physical length of LD were used for further evaluation (Alqudah, Sallam, et al., 2020). Furthermore, gene expression profile for high confidence candidate genes was measured in FPKM (fragments per kilobase of transcript per million mapped reads) from the

RNA-seq expression database for spikelets and grain organs, in addition to spike-related traits at different developmental stages in the barley cultivar Morex (Mascher et al., 2017). In the current study, we only considered the gene expression in INF2: Inflorescence (1–1.5 cm), CAR5, 15: caryopses at 5 and 15 DPA (days post-anthesis), LEM: Lemma (6 weeks PA [post-anthesis]), LOD: Lodicule (6 weeks PA), PAL: Palea (6 weeks PA), EPI: Epidermis (4 weeks), RAC: Rachis (5 weeks PA). More details on the RNA expression data analysis are available elsewhere (Mascher et al., 2017; Rapazote-Flores et al., 2019).

3 | RESULTS

3.1 | SNP genotyping and population stratification

A total number of 127,022 SNPs were used for the analysis after filtering and quality control check (Figure S1a). The number of SNPs dramatically increased (25X) when we combined the SNP information from the 9K chip and GBS in the present study compared to around 4800 SNP markers published earlier from these accessions (Alqudah et al., 2014; Nagel et al., 2019). The population structure analysis using 127,022 SNPs (principal component analysis) shows more distinct groupings among the barley accessions. It demonstrates that when a high number of sequenced-based markers are employed, the grouping becomes clearer compared to chip-based 9K SNPs. It is not surprising given that 9K SNPs suffer from ascertainment bias, particularly when revealing diversity across Ethiopian barleys (Figure S1b–d). Using 127,000 SNPs, the average genome-wide LD decay at r^2 = 0.15 is approximately 2.5 Mbp (Figure S2). Thus, we expect to discover most of the quantitative trait loci (QTLs) in the present study with a reasonable distance from the candidate genes.

3.2 | Extensive natural variation in MYP and AS%-related traits

A wide variation among the measured traits was observed (Figure S3). Further analysis of the phenotypic variation showed significant (p < .05) variation among the accessions when classified into row type, biostatus and origins (Dataset S1).

A wide range of variation was detected in MYP, DS% and UDS% at the GA and TIP stages among the accessions (Figure 2a). Six-rowed accessions developed significantly higher MYP than two-rowed accessions, which leads to a higher UDS% in the former, whereas the latter have a higher DS% in both stages (Figure 2b). This trend is seen in the biostatus of the accessions where landraces exhibit highly significant natural variation in MYP and an increase in UDS% instead of DS% compared with other groups (Figure 2c). The accessions from Africa and America had significantly higher MYP, but the accessions from Africa had more UDS% at the GA and TIP stages, while accessions from Middle and Western Europe produced significantly more DS%, especially at the TIP stage (Figure 2d).

Highly significant differences (p < .0001) in AS% were found among the accessions at all developmental stages (Figure 3a). Approximately 75% of spikelets at MYP were alive at the GA stage (~25% aborted) and this number reduced to ~65% at the TIP stage (~35% abortion from MYP), indicating that the duration between MYP and GA is the most critical phase for AS in barley (Figure 3a). About 50% reduction in AS% between Hrv_MYP indicates that ~15%–20% of spikelets were aborted after the TIP stage (Figure 3a). Two-rowed accessions have significantly higher AS% at all developmental stages than six-rowed accessions, except between Hrv_TIP stages (Figure 3b). The biostatus and origin of the accessions confirm that the landrace group shows the lowest AS%, whereas breeding and hybrid lines show the highest AS% during the critical phase (Figure 3c,d).

Analysis of thermal time as GDD for phase transition shows significant differences (p < .0001) among the accessions (Figure S4a). From sowing, the required thermal time to reach each developmental stage is significantly shorter in six-rowed accessions (Figure S4b). There are no significant variation for GDD between accessions based on their biostatus in the MYP stage, whereas accessions that are categorised as cultivars have a longer phase transition to the TIP stages (Figure S4c). Clear significant differences among the accessions from different origins in GDD were detected especially at TIP. The accessions from Middle, Northern and Western Europe have a significantly longer phase transition, whereas those from America, Asia and Eastern Europe have a significantly shorter phase transition (Figure S4d).

For the spike architecture traits, there is a clear significant difference (p < .001) in all measured traits. The spike length at the TIP stage and awn length at Hrv stage are shorter in the breeding lines, whereas the final spike length at the Hrv stage is shorter in the landraces and hybrids (Figure S5a). The shortest spike at the TIP stage appeared in the accessions from Africa, whereas American accessions have the shortest spike and awn lengths compared to other regions (Figure S5b). Two-rowed accessions have significantly (p < .001) longer spikes than six-rowed accessions (Figure S5c). There are no clear results for spike density at the TIP stage, although six-rowed landraces from Africa carry a denser spike (Figure S5d–f).

During the field trial of 2018 weather conditions were dry and probably mimicked the drought conditions of Europe. Weather data show that between April and August, when plants exhibited vigorous growth and development, dry and high temperature and low amount of precipitation were observed (Figure S6). The average temperature and precipitation during April and May 2018 were 14 ± 2.8 and 17 ± 4.1°C and 30 ± 3.1 and 20 ± 0.8 mm respectively. According to the weather data at IPK and Deutschlandwetter, the dry condition during the 2018 growing season made it an exceptional season when compared with previous decades (https://www.ipk-gatersleben.de/en/databases/weather-data/) with the lowest rainfall since 1881 (https://www.dwd.de/DE/presse/pressemitteilungen/DE/2018/20181228_deutschlandwetter_jahr2018_news.html).

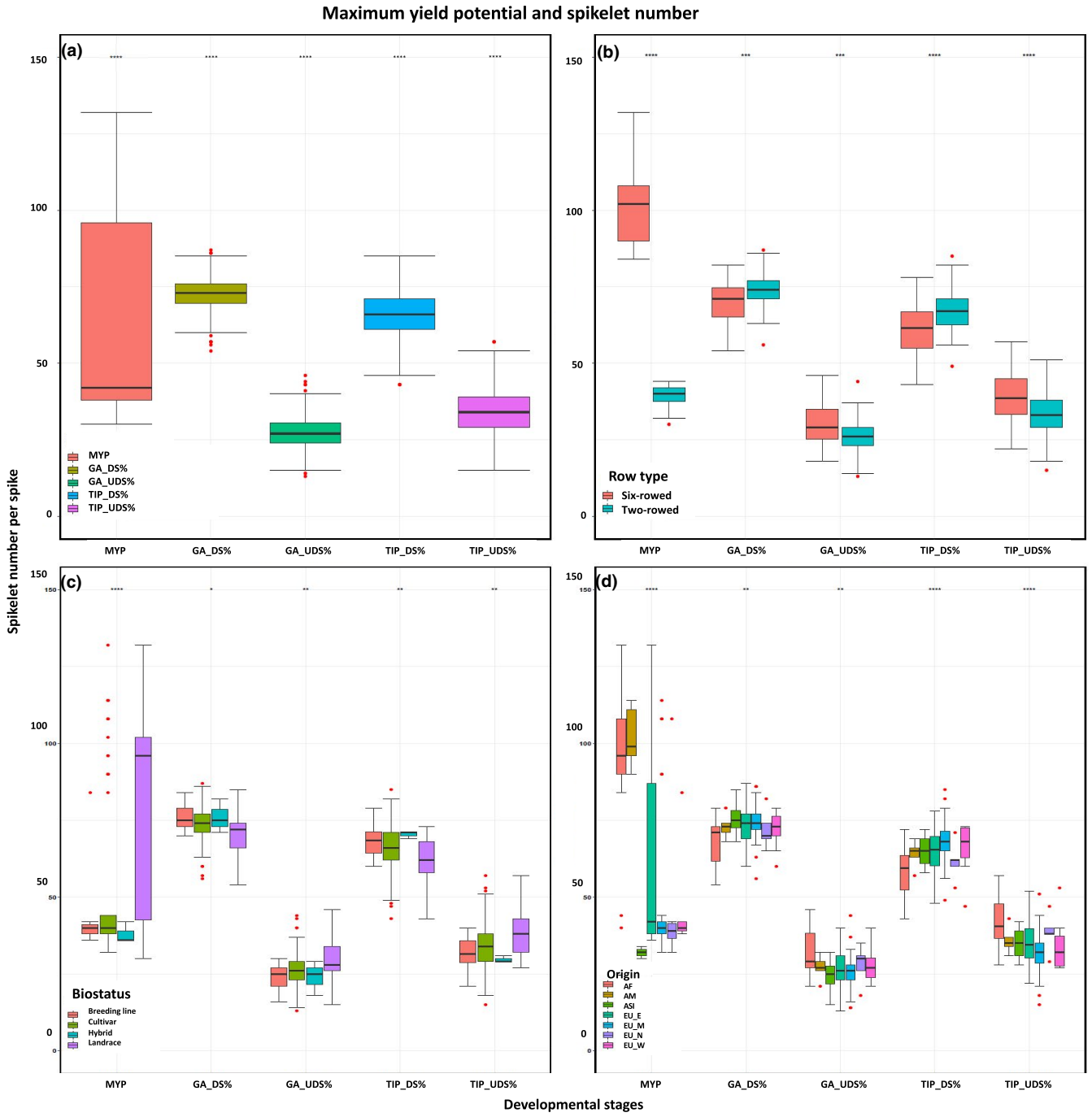

FIGURE 2 Phenotypic variation in MYP, DS% and UDS% at GA and TIP stages between (a) all accessions, (b) row type, (c) biostatus and (d) origin. MYP: maximum yield potential, DS%: number of developed spikelets, UDS%: number of undeveloped spikelets, GA: green anther, TIP: awn tipping, AF: Africa, AM: America, ASI: Asia, EU_: Europe (E: Eastern, M: Middle, N: Northern, W: Western). The asterisk denotes significant difference at $p \leq .05$ according to the *t* test among the accessions of each group at the same developmental stage. The degree of significance is indicated as *p, 0.05; **p, 0.01; ***p, 0.001; ****p, 0.0001. ns, not significant

3.3 | Correlation, clustering and heritability of phenotypic traits

In general, four major clusters are visualised according to the correlation among traits (Figure 4a). The first cluster includes three traits, MYP with UDS at the GA and TIP stages, with a strong correlation between UDS at both stages, which are moderately positively correlated with MYP (Figure 4a). The second cluster consists of DS% and AS% at the MYP, GA and TIP stages. A high correlation between DS% at the GA and TIP stages was detected, which reflects the high correlation between AS%_MYP at the GA and TIP stages (Figure 4a). Particularly, MYP (cluster 1) has a

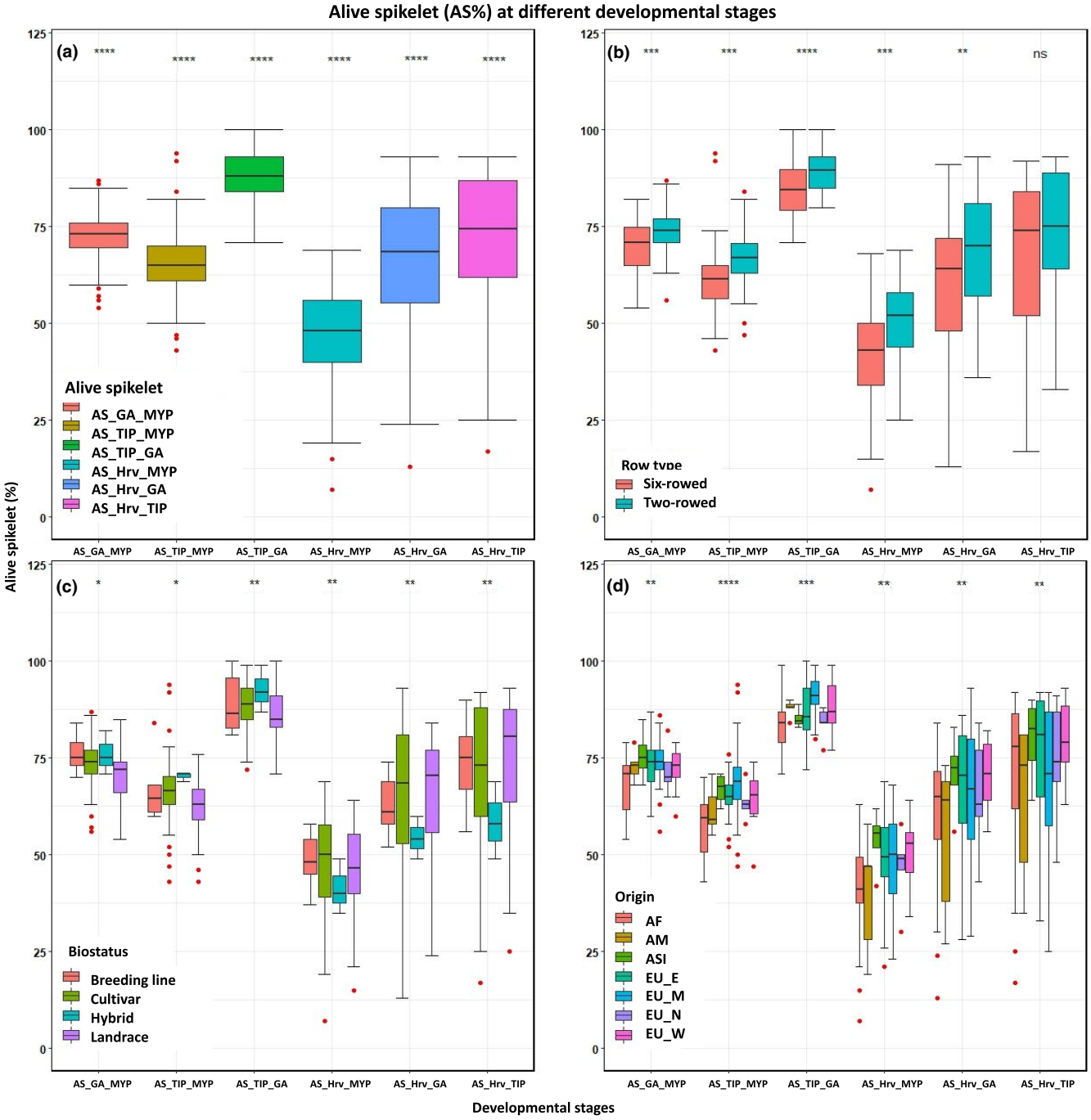

FIGURE 3 Phenotypic variation in AS% among the developmental stages between (a) all accessions, (b) row type, (c) biostatus and (d) origin. AS%: alive spikelet percentage, MYP: maximum yield potential, GA: green anther, TIP: awn tipping, Hrv: Harvest, AF: Africa, AM: America, ASI: Asia, EU_: Europe (E: Eastern, M: Middle, N: Northern, W: Western). The asterisk denotes significant difference at $p \leq .05$ according to the *t* test among the accessions of each group at the same developmental stage. The degree of significance is indicated as *p, 0.05; **p, 0.01; ***p, 0.001; ****p, 0.0001. ns, not significant

moderately negative correlation (~−0.40) with all the traits in cluster 2 (Figure 4a). The correlation among the developmental stage transition (MYP, GA and TIP) by GDD had a moderate negative correlation (~−0.30) with MYP and no clear correlation with AS% at the GA or TIP stages (Figure 4a). However, stage transition had a moderate positive correlation (0.20–0.30) with AS% between Hrv and the MYP, GA or TIP stages (cluster 4). In addition, there was a moderate negative correlation (−0.25, −0.40) between MYP and stage transition (Figure 4a).

Separately, in the two- and six-rowed accessions, the same trend of clusters was detected. MYP clustered with stage transition in two-rowed accessions (Figure 4b), whereas the cluster is with the DS% and AS% groups at the MYP, GA and TIP stages in six-rowed accessions (Figure 4c). MYP had a moderate positive correlation with

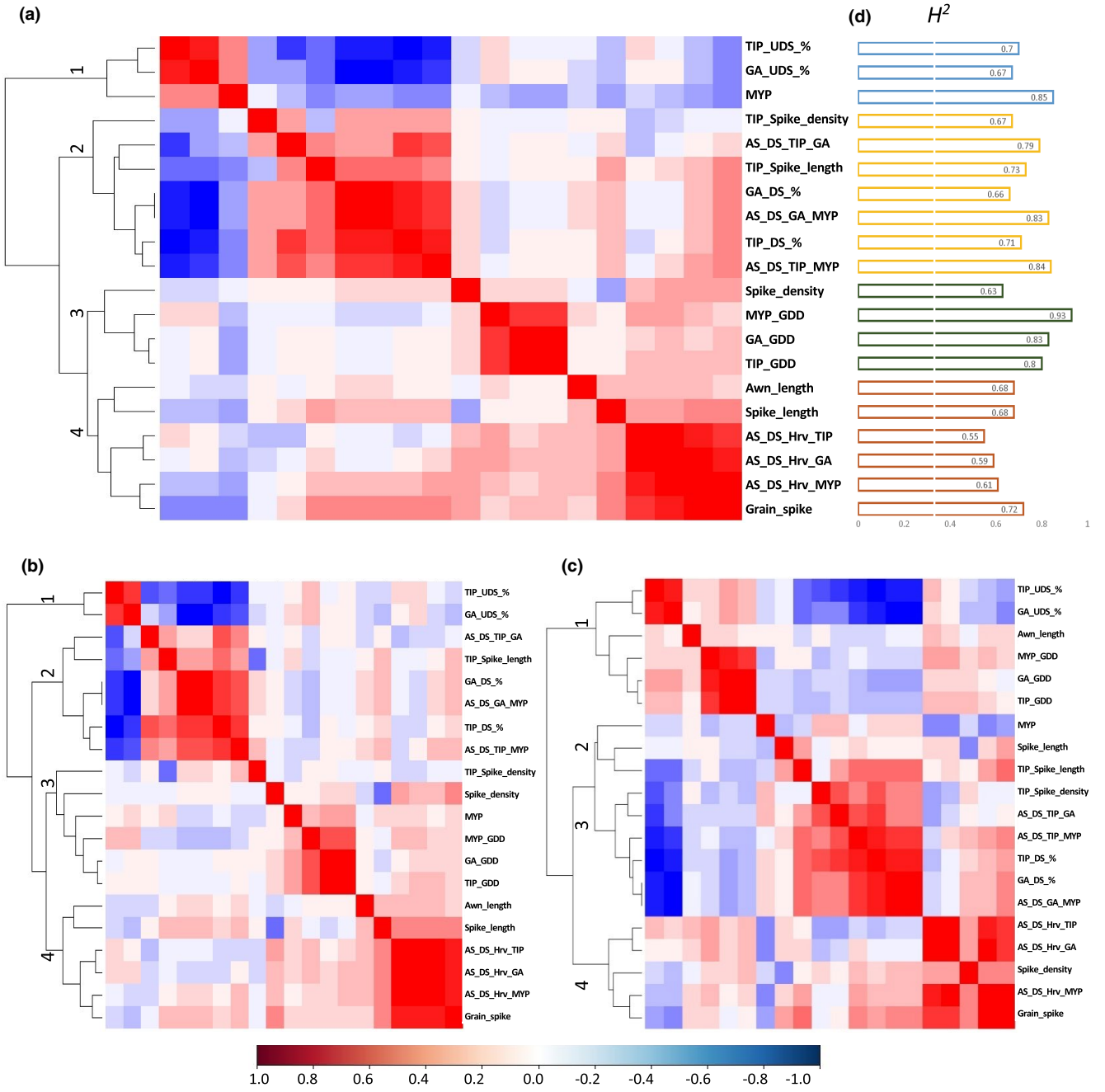

FIGURE 4 Heat map representing the cluster of the Pearson correlation coefficient between each pair of 20 traits in (a) all collections, (b) two-rowed accessions and (c) six-rowed accessions. The strength of the Pearson's correlation coefficients is represented with the blue-white-red colour scheme. (d) Broad-sense heritabilities (H^2) for the traits in the whole collection

stage transition (especially GA and TIP) in two-rowed accessions; in contrast, a negative correlation was found in six-rowed accessions. In six-rowed accessions, AS%_Hrv_MYP negatively correlated with MYP and UDS% at the GA and TIP stages but positively correlated with DS% at these stages (Figure 4c) compared with two-rowed accessions (Figure 4b). DS% at the GA and TIP stages, and AS% were positively correlated with spike length and density at the TIP stage, final spike, and awn length in addition to GPS (Figure 4a–c). These correlations are much stronger in six-rowed accessions.

Broad-sense heritability values (H^2) ranging from 0.55 to 0.93 are obtained for the studied traits. High heritability for MYP is also found when expressed in thermal time (GDD) with $H^2 \geq 0.93$ (Figure 4d). The DS% and UDS% at the GA and TIP stages reduced the moderate H^2 values ranging from 0.66 to 0.71 (Figure 4d). AS% among early stages, for example AS%_GA_MYP, exhibited high H^2 values (0.79–0.84), whereas the moderate H^2 reflected the proportion of environmental effects on the AS% at Hrv with early-stage, for example AS%_Hrv_MYP ranging from 0.55 to 0.61. Other

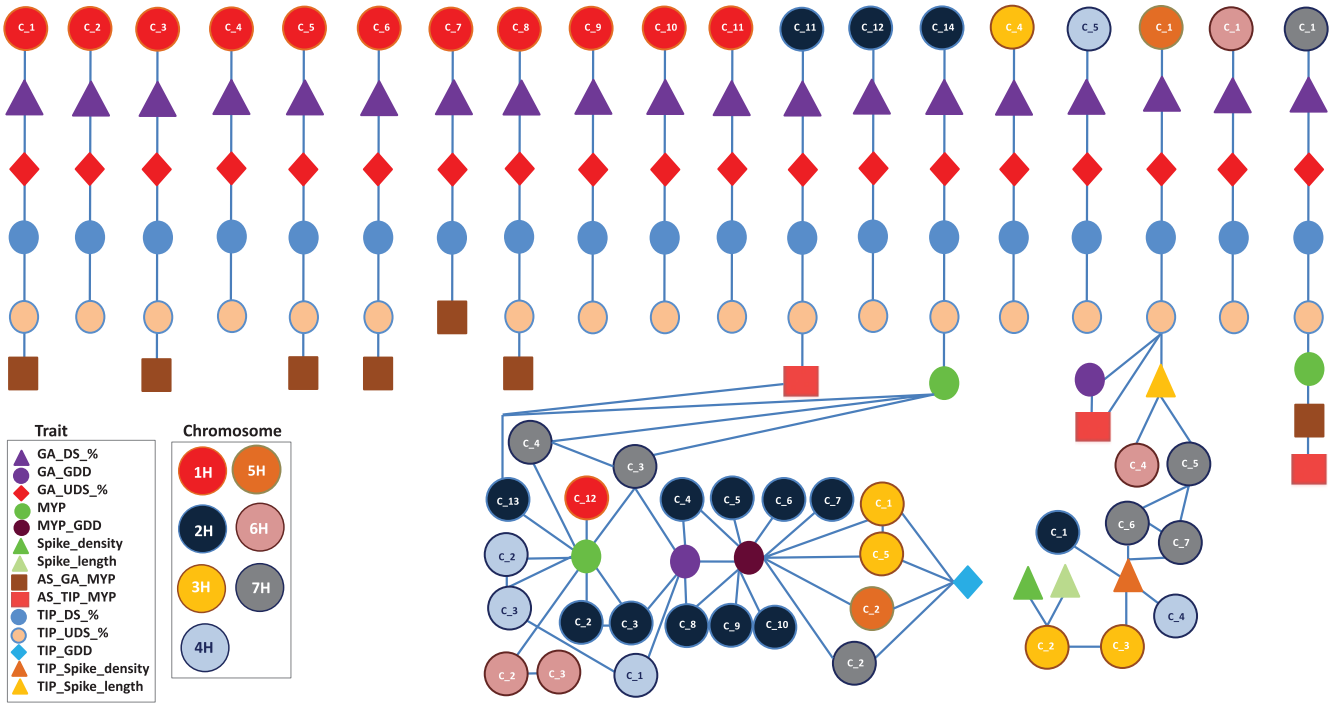

FIGURE 5 Genotype–phenotype network analysis for defining the key genetic cluster of markers associated with the studied traits

agronomic traits, such as grain yield, had moderate heritability values (0.63, 0.72; Figure 4d).

3.4 | Investigating relationship of MYP and AS%-related traits using networks

In total, 574 SNP marker–trait association combinations over the whole genome were identified in GWAS that passed the FDR test ($-\log_{10}(p)$ of SNP ≥ FDR at 0.001; Dataset S2; Figure S7). These were subsequently investigated in the network analysis (Dataset S2). Significant SNPs are fall into genetic clusters on each chromosome based on their physical positions, using a defined LD interval of LD ± 1.25 Mbp. Network analysis provided a visual summary of the relationships that allowed the identification of associated genetic clusters with multiple phenotypes (Figure 5). In total, 299 SNPs (~52%) made 50 genetic clusters that show a significant network with multiple traits (>2 traits), suggesting linkage or pleiotropic effects (Dataset S3). Chromosome 2H has the largest number of clusters (14_C, 28%), followed by chromosome 1H with 12 clusters (24%; Dataset S3). C_2 on chromosome 5H has the highest number of associated SNPs (21 SNPs), followed by C_2 and C_5 on chromosome 1H with 15 SNPs each (Dataset S3).

There are 26 genetic clusters in the whole genome connected with MYP, DS% and UDS% at the GA and TIP developmental stages, of which 11 clusters belong to chromosome 1H (Figure 5). Six genetic clusters controlling DS% and UDS% at the GA and TIP stages on 1H are also connected with AS_GA_MYP% (Figure 5). SNPs within cluster number 5 on 1H (SNPs: m_10811 and m_10812 at 666,503,44 and 666,733,71 respectively) show contrasting effects (~−4% to ~+4%) on AS_GA_MYP that can be explained by allelic variation in this genomic region (Figure 6; Dataset S3).

Furthermore, clusters C_11 (AS_TIP_MYP%) and C_14 (MYP) connects with DS% and UDS% during the GA and TIP phases, respectively. This region is located on chromsome 2H (Figure 5; Dataset S3). Allelic variation at SNPs within cluster number 14 on 2H had contrasting effects on MYP (~−10 to ~+10 spikelet) with phenotypic variation explained by the marker being ~5.5% (Figure 6).

These clusters are also connected with three genetic clusters (C_13 2H, C_3 and C_4 7H), which connects to numerous clusters controlling MYP (Figure 5; Dataset S3). The genetic clusters have inter- and intra-specific interactions with each other, of which C_3 2H, C_1 4H and C_3 7H directly connect with MYP_GDD. Also, GDD at the MYP and GA stages share genetic clusters from 2H, while GDD at the GA and TIP stages have other genetic clusters from 3H, 5H and 7H (Figure 5; Dataset S3). Genetic cluster C_1 5H had a genetic connection with DS% and UDS% at the GA and TIP stages in addition to GA_GDD and AS_TIP_MYP. The same genetic cluster is also connected with spike architecture traits (spike length and density) at the TIP and Hrv stages through a network connection with other genetic clusters in 3H, 4H, 6H and 7H (Figure 5; Dataset S3). Finally, the most important genetic cluster in the network is C_1 on 7H that controls DS% and UDS% at the GA, TIP and MYP stages, that also shared with AS_GA_MYP% and AS_TIP_MYP%, indicating the importance of this major genetic region in barley yield improvement.

3.5 | Candidate genes for MYP and AS%-related traits

On all chromosomes, 296 candidate genes were found to be potentially associated with MYP and AS%-related traits within the ±1.25 Mbp interval (Dataset S4). The list includes 48 candidate genes that might be involved in spikelet development and abortion (Table 2; Dataset S4). For example, the list includes many gene annotations as 6-phosphogluconate dehydrogenase, decarboxylating 1, ATP-dependent, APETALA2 (AP2), sucrose and cellulose synthase, fatty acid, glycerol-3-phosphate acyltransferase 8 (GPAT), phytohormones (such as abscisic acid [ABA], auxin, gibberellin and ethylene), cytochrome P450, DNA-binding one zinc finger, heavy metal transport, leucine-rich repeat receptor-like protein kinase and NAC domain proteins (Table 2; Dataset S5).

Strikingly, 7H (7,906,852–9,914,430 bp) is an important genetic region that includes numerous interesting putative candidate genes that potentially control spikelet development, MYP and spikelet abortion (Table 2; Figure 5). These genes are involved in phosphogluconate dehydrogenase, sucrose synthase, GPAT, lipoxygenase 2, GDSL esterase/lipase, and ozone-responsive stress-related protein (Table 2; Figure 5), which are essential for plant developmental mechanisms such as shoot apical meristem.

The SNP associated with 2H is from the *PHOTOPERIOD RESPONSE LOCUS1* (*Ppd-H1*) gene (29,127,021 bp) that shows the highest phenotypic variance (~9.1%) and associated with the MYP and showing a negative effect by reducing spikelets (~−10) (Dataset S3). Early flowering in the accessions means longer time for spikelet abortion and thus finding the association at this region is not surprising. In addition, the region close to the major row type *SIX-ROWED SPIKE 1* (*VRS1*) gene (2H, 646,237,676 bp) was found to be associated with MYP, accounting for ~7.3% of the variation and reducing ~20 spikelets (Dataset S3). As evident from the correlations of the traits, six-rowed barleys have higher spikelet-bearing potential but several spikelets were aborted due to competition for nutrients among the spikelets.

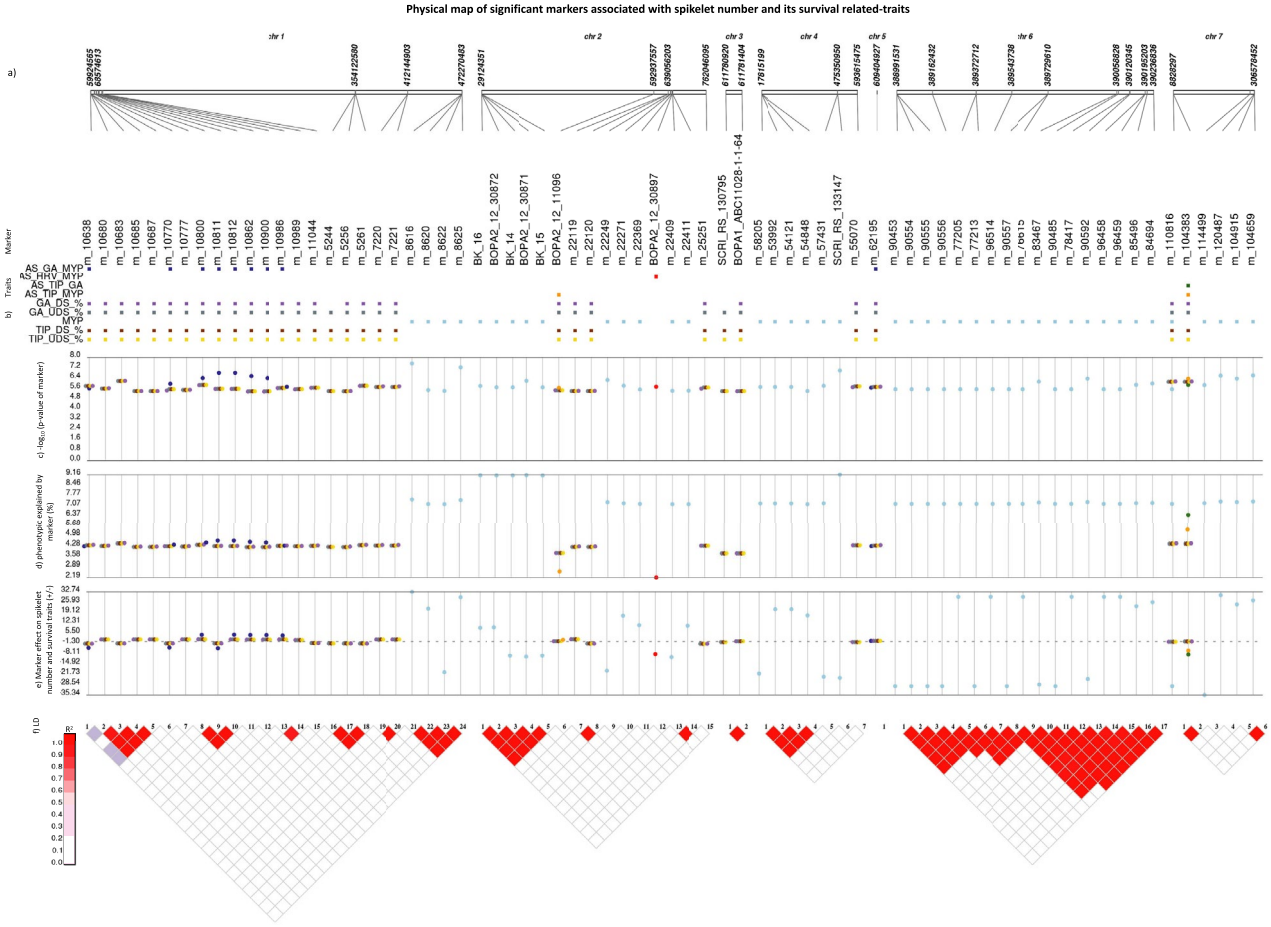

FIGURE 6 Physical map of significant markers associated with spikelet number and its survival-related traits. (a) The physical position of associated markers, (b) the name of markers and associated traits, (c) plot of minus $\log_{10}$ (*p*-value) of associated markers, (d) phenotypic variance explained by marker (%), (e) marker effect on spikelet number and its survival traits (+/−), (f) linkage disequilibrium (LD) among the significant markers presented as R^2

3.6 | Expression analysis of candidate genes

Candidate genes in different barley spike and grain organs at different developmental stages show a wide range of gene expression (Dataset S5; Figure S8). Although several potential candidate genes exist, three genes (14-3-3-like protein GF14-E, 6-phosphogluconate dehydrogenase, decarboxylating 1 and ozone-responsive stress-related protein) and their annotations (*HORVU3Hr1G010530.2*, *HORVU7Hr1G006160.4* and *HORVU7Hr1G007610.1*) discovered on chromosome 3H and 7H, respectively, displayed the highest expression in INF2, which is around the stage of awn primordia (MYP). These genes are also highly expressed in other grain and spike organs, for example, CAR 5, 15 DPA, LEM, LOD, PAL, EPI and RAC. These stages of expression levels demonstrate that the genes play important biological roles in spike and spikelet development as well as in other important organs involved in grain and spikelet traits (Dataset S5; Figure S8). Another interesting gene discovered was alanine aminotransferase 2 on chromosome 1H *HORVU1Hr1G018540.17* that is highly expressed in grain and spikelet organs but not in INF2. Notably, *HORVU7Hr1G007220.1* is the most highly expressed gene across the organs, especially in specific ones like LEM, PAL and RAC (Dataset S5; Figure S8).

4 | DISCUSSION

In this study, we discovered the natural variation of MYP by dissecting the traits that contribute to spikelet survival in barley (Figure 1). Due to the labour-intensive measurements of these traits, particularly under field conditions, genetic studies related to these traits are difficult. In the light of the prevalent constraints of measuring these traits, we performed this study under large field plots and under replication. We show that these traits have a direct effect on spikelet number maintenance and, as a result, yield. We propose a novel phenotyping strategy for accurate measurement of MYP, DS, UDS, and AS-related traits (Figure 1), which result in robust and reliable identification of QTLs in GWAS analysis. An increase in the total number of spikelets was proposed previously as a key factor for yield improvements (Alqudah & Schnurbusch, 2014); however, dissecting and separating the DS and UDS from the total spikelets maximises the ability to detect genetic regions of studied traits. Therefore, the present study demonstrates that nested phenotypes can be further dissected and undetected genomic regions can be discovered. In our study, six-rowed accessions, which are mostly landraces from Africa and America, produced more spikelets at MYP and thus reduced DS% compared with two-rowed accessions, which are predominantly European cultivars. Low AS% in six-rowed accessions can be attributed to high competition among spikelets at the MYP stage, combined with a faster phase transition (GA and TIP) compared to European accessions (Alqudah & Schnurbusch, 2014; Arisnabarreta & Miralles, 2004; del Moral et al., 2002; Miralles et al., 2000). The results demonstrate that the MYP stage can be measured between AP and WA, which is genotype dependent, and its response to environmental cues.

Ppd-H1 (*PSEUDO-RESPONSE REGULATOR 7*, *PRR7*) contributes to early heading/flowering under the long day European growing season, suggesting that the reduction in AS% could be influenced by the natural variation in the *Ppd-H1* gene. The *Ppd-H1* sensitive allele originated mostly from African and American accessions in our study. We do not expect wide differences in European barley accessions as they mostly harbour the photoperiod insensitive late heading allele (*ppd-H1*). This is not surprising as evident in recently surveyed sequences of a large barley collection in this gene (Sharma et al., 2020).

A huge peak near the *Vrs1* gene highlights the major differences in spikelet arrangement at rachis nodes despite controlling for population structure. These results demonstrate that row-type genes indeed have a huge influence on MYP and related traits. Whether these differences are directly due to major morphological changes in spike arrangements or whether there are independent genetic players influencing the traits is under investigation. These results also highlight the limitation of six-rowed landraces in bearing more spikelets than their MYP. Of note is that the genetic map position of another row-type gene *labile* (Youssef et al., 2020) has recently been corrected, which also overlapped in the region of the peak mentioned above on chromosome 2H. Interestingly the *labile* barleys are Ethiopian in origin.

The phase transition between GA and TIP is critical as this is the time which determines the AS%. In this study, most of the two-rowed accessions (~70% of two-rowed *VRS1*) carried the less sensitive allele of *Ppd-H1*, which are late heading/flowering and bear high spikelet numbers, but several of them aborted before maturity. This suggests that delayed transition (GA and TIP) is not always effective in improving final grain yield because it also likely increases UDS despite an increase in total spikelets. Therefore, achieving a balance in transition time between developmental stages is very crucial in increasing yield by improving the number of DS while retaining a large number of spikelets that have already formed. Heritability analysis demonstrated that MYP, DS% and UDS% at the GA and TIP stages are largely controlled by genetic factors and less influenced by the environment as reflected in their high heritability values. As phase transition and spikelet development are directly influenced by environmental factors like temperature, photoperiod and stress (Alqudah & Schnurbusch, 2014; Arisnabarreta & Miralles, 2006; Kernich et al., 1997), high temperature with low precipitation amounts could be the main environmental factor for increasing UDS% and the abortion rate in our study.

In this research, we observed a considerable number of genetic clusters (QTLs) that are important for improving MYP, DS and AS%. For example, on chromosome 1H, six genetic clusters controlling DS% and UDS% at the GA and TIP stages were shared with AS_GA_MYP%, suggesting that improving DS% during the critical phase (MYP_GA) can increase AS% and thus improve grain yield. Notably, the genetic cluster C_1 at 5H shows the genetic connection among AS% at TIP and GA stages, AS% TIP_MYP with phase

TABLE 2 List of high confidence candidate genes with their physical position based on Morex (v1) genome pseudomolecule sequence and Morex (v2) and annotation with gene ontology (GO) terms

	Gene name	Chr.	Cluster	Start (v1)	End (v1)	Start (v2)	End (v2)	Gene annotation	GO terms
1	*HORVU1Hr1G017770.2*	1H	C_2	61,271,383	61,272,571	56,910,766	56,911,954	Auxin-responsive protein IAA4	GO:0005515, GO:0005634, GO:0006355
2	*HORVU1Hr1G018140.5*	1H	C_5	65,592,292	65,593,858	60,974,240	60,975,806	Dehydrogenase/reductase SDR family member 4	GO:0008152, GO:0016491
3	*HORVU1Hr1G018540.17*	1H	C_6	68,365,660	68,370,235	63,678,515	63,683,090	Alanine aminotransferase 2	GO:0003824, GO:0009058, GO:0030170
4	*HORVU1Hr1G018700.1*	1H	C_7	69,599,927	69,605,808	64,830,766	64,836,374	S-adenosyl-L-methionine-dependent methyltransferases	None
5	*HORVU1Hr1G018930.2*	1H	C_8	71,393,292	71,400,691	67,107,128	67,114,527	Mitochondrial transcription termination factor family protein	GO:0003690, GO:0005739, GO:0006355
6	*HORVU1Hr1G019000.2*	1H	C_8	71,846,595	71,859,008	67,770,302	67,782,722	ATP-dependent RNA helicase, putative	GO:0003676, GO:0005524
7	*HORVU1Hr1G019290.1*	1H	C_8	72,311,829	72,319,723	68,111,524	68,119,418	Transcription regulatory protein SNF2	None
8	*HORVU1Hr1G019330.7*	1H	C_8	72,525,628	72,535,676	68,394,352	68,404,400	Serine/threonine protein phosphatase family protein	GO:0004721, GO:0005515, GO:0009742, GO:0016787
9	*HORVU1Hr1G067160.1*	1H	C_12	476,758,262	476,759,481	448,036,529	448,037,748	AP2/Ethylene-responsive transcription factor	GO:0003677, GO:0003700, GO:0006355
10	*HORVU2Hr1G013400.32*	2H	C_2	29,123,785	29,127,889	23,720,341	23,724,445	Pseudo-response regulator 7	GO:0000160, GO:0005515
11	*HORVU2Hr1G081570.1*	2H	C_11	591,656,893	591,657,870	528,539,430	528,540,407	Heavy metal transport/ detoxification superfamily protein	GO:0030001, GO:0046872
12	*HORVU2Hr1G081650.10*	2H	C_11	592,334,638	592,338,706	529,114,148	529,118,216	Cytochrome P450 superfamily protein	GO:0005506, GO:0016705, GO:0020037, GO:0055114
13	*HORVU2Hr1G081670.4*	2H	C_11	592,342,208	592,343,991	529,108,665	529,110,639	ATP-dependent 6-phosphofructokinase 2	GO:0003872, GO:0006002, GO:0006096
14	*HORVU2Hr1G081780.2*	2H	C_11	593,165,621	593,169,320	529,848,477	529,852,168	Protein FATTY ACID EXPORT 6	GO:0016020
15	*HORVU3Hr1G010530.2*	3H	C_1	23,184,058	23,187,544	18,313,375	18,316,861	14-3-3-like protein GF14-E	GO:0019904
16	*HORVU3Hr1G072810.1*	3H	C_3	548,753,336	548,756,420	498,496,462	498,499,545	Gibberellin 2-oxidase	GO:0005506, GO:0016491, GO:0055114
17	*HORVU3Hr1G085720.3*	3H	C_4	613,897,980	613,903,093	555,623,119	555,628,232	BEL1-like homeodomain 8	GO:0003677, GO:0006355
18	*HORVU3Hr1G114970.1*	3H	C_5	692,190,558	692,192,500	622,380,719	622,382,661	Two-component response regulator ARR1	GO:0003677

(Continues)

TABLE 2 (Continued)

	Gene name	Chr.	Cluster	Start (v1)	End (v1)	Start (v2)	End (v2)	Gene annotation	GO terms
19	*HORVU4Hr1G079860.14*	4H	C_5	613523682	613527868	594,559,030	594,563,201	Cytokinin riboside 5'-monophosphate phosphoribohydrolase	GO:0009691, GO:0016787
20	*HORVU5Hr1G097530.9*	5H	C_1	604,619,681	604,628,095	541,640,209	541,648,578	Protein kinase superfamily protein	GO:0004672, GO:0005524, GO:0006468
21	*HORVU5Hr1G097620.1*	5H	C_1	605,046,252	605,048,334	541,969,425	541,971,507	DOF zinc finger protein 1	GO:0003677, GO:0006355
22	*HORVU5Hr1G097630.1*	5H	C_1	605,102,179	605,108,556	541,902,068	541,908,445	Serine/threonine, ABA	GO:0004672, GO:0005524, GO:0006468
23	*HORVU5Hr1G097730.1*	5H	C_1	605,404,198	605,406,017	542,257,633	542,259,452	maternal effect embryo arrest 14	None
24	*HORVU5Hr1G097870.1*	5H	C_1	606,030,788	606,033,849	542,700,019	542,703,080	Cytochrome b6-f complex subunit 7	GO:0009512
25	*HORVU5Hr1G098000.2*	5H	C_1	606,265,773	606,270,239	543,130,751	543,135,250	Thioredoxin superfamily protein	GO:0045454
26	*HORVU5Hr1G098040.1*	5H	C_1	606,306,734	606,312,889	542,976,469	542,982,570	Proteasome assembly chaperone 2	None
27	*HORVU5Hr1G098450.2*	5H	C_1	607,811,609	607,815,014	544,349,551	544,352,956	AP2-like ethylene-responsive transcription factor	GO:0003677, GO:0003700, GO:0006355
28	*HORVU5Hr1G098490.11*	5H	C_1	607,885,609	607,888,626	544,434,092	544,437,175	Cellulose-synthase-like C4	None
29	*HORVU5Hr1G098570.6*	5H	C_1	608,243,814	608,250,710	544,667,511	544,674,395	Homeobox protein knotted-1-like 12	GO:0003677, GO:0005634, GO:0006355, GO:0043565
30	*HORVU5Hr1G098670.1*	5H	C_1	608,528,936	608,532,693	545,021,870	545,025,578	MLO-like protein 5	GO:0006952, GO:0016021
31	*HORVU5Hr1G098890.10*	5H	C_1	608,994,464	608,997,728	545,300,924	545,304,190	Malate dehydrogenase	GO:0003824, GO:0005975, GO:0006108, GO:0016491
32	*HORVU5Hr1G099380.1*	5H	C_1	610,154,782	610,157,638	546,532,937	546,535,718	Aspartyl/glutamyl-tRNA(Asn/Gln) amidotransferase subunit B	None
33	*HORVU5Hr1G099470.2*	5H	C_1	610,287,150	610,289,276	546,458,717	546,460,843	NAC domain protein,	GO:0003677, GO:0006355
34	*HORVU5Hr1G099490.2*	5H	C_1	610,332,314	610,338,919	546,605,726	546,612,331	Phosphatidylinositol N-acetylglucosaminyltransferase subunit GPI1	GO:0006506, GO:0016021, GO:0017176
35	*HORVU5Hr1G099700.2*	5H	C_1	610,623,383	610,626,183	546,844,350	546,847,147	Hexosyltransferase	GO:0016757
36	*HORVU6Hr1G059810.2*	6H	C_3	396138647	396147162	394,178,091	394,186,606	Pyridoxine/pyridoxamine 5'-phosphate oxidase	GO:0004733, GO:0008615, GO:0010181, GO:0016491
37	*HORVU6Hr1G059940.4*	6H	C_3	397015876	397017242	394,840,267	394,841,633	GRAM domain-containing protein / ABA-responsive protein-related	None
38	*HORVU7Hr1G006070.1*	7H	C_1	7,906,852	7,908,220	7,604,123	7,605,491	6-phosphogluconate dehydrogenase, decarboxylating 1	GO:0004616, GO:0006098, GO:0016491, GO:0055114

(Continues)

TABLE 2 (Continued)

	Gene name	Chr.	Cluster	Start (v1)	End (v1)	Start (v2)	End (v2)	Gene annotation	GO terms
39	*HORVU7Hr1G006180.4*	7H	C_1	8,015,847	8,020,157	7,664,696	7,669,128	Guanine nucleotide-binding protein alpha-2 subunit	GO:0003924, GO:0004871, GO:0007165, GO:0007186
40	*HORVU7Hr1G006240.1*	7H	C_1	8,168,380	8,169,150	8,053,906	8,054,676	Avenin-like a4	GO:0045735
41	*HORVU7Hr1G006320.4*	7H	C_1	8,220,885	8,226,058	7,996,206	8,001,398	Lipoxygenase 2, chloroplastic	GO:0005515, GO:0016491, GO:0016702, GO:0046872
42	*HORVU7Hr1G006450.1*	7H	C_1	8,514,242	8,514,744	849,187	849,689	6-phosphogluconate dehydrogenase, decarboxylating 1	GO:0004616, GO:0006098, GO:0016491, GO:0055114
43	*HORVU7Hr1G007160.1*	7H	C_1	9,561,365	9,562,490	9,271,297	9,272,422	Sucrose synthase 2	None
44	*HORVU7Hr1G007220.1*	7H	C_1	9,611,885	9,618,894	9,221,890	9,228,899	Sucrose synthase 1	None
45	*HORVU7Hr1G007570.1*	7H	C_1	9,844,143	9,845,752	9,435,160	9,436,769	Glycerol-3-phosphate acyltransferase 8	GO:0008152, GO:0016746
46	*HORVU7Hr1G007580.1*	7H	C_1	9,879,963	9,884,227	9,474,373	9,478,637	Cytochrome P450 superfamily protein	GO:0005506, GO:0016705, GO:0020037, GO:0055114
47	*HORVU7Hr1G007590.1*	7H	C_1	9,889,081	9,894,380	9,483,977	9,489,276	GDSL esterase/lipase	GO:0016788
48	*HORVU7Hr1G007610.1*	7H	C_1	9,914,430	9,916,941	9,509,120	9,511,437	Ozone-responsive stress-related protein	None

transition (GA_GDD) and spike length, demonstrating the importance of this cluster in controlling multiple traits. It also highlights the role of phase transition in keeping the spikelet alive. The genetic cluster C_1 on 7H controls DS% and UDS% at the GA and TIP stages; MYP is connected with AS_GA_MYP% and AS_TIP_MYP%, indicating the importance of this genetic region in grain yield improvement. Network analysis showed the degree of complexity of the studied traits; for example C_11 and C_14 on 2H are connected with MYP, DS% and UDS% at the GA and TIP stages, and AS_TIP_MYP% is linked with C_13 on 2H, and C_3 and C_4 on 7H, which have inter- and intra-connection with many other clusters controlling MYP and phase transition. Such traits cannot be detected without partitioning the total number of spikelets into DS and UDS, besides using a high-density SNP number and post-genome-wide state-of-the-art analysis (e.g. networking).

Using networking and a GWAS approach, we have identified hundreds of putative candidate genes, of which 48 are potentially involved in spikelet development and abortion. Based on the associated candidate genes, we have highlighted the major candidate genes and the underlying pathways in Figure 7 and Table 2. The association signals in several cases either come from the candidate genes or quite near the candidate genes involved in phytohormones such as ABA on chromosome 5H, auxin-responsive protein IAA4 on chromosome 1H and cytokinin on chromosome 4H. ABA is suggested to be a crucial determinant of spikelet development and key player in spikelet abortion (Marzec & Alqudah, 2018). Furthermore, high ABA concentrations at the tip of the spike also led to higher spikelet abortion (Boussora et al., 2019).

FIGURE 7 Proposed model of the regulation of spikelet development and abortion. Green box and arrow indicate the promotion of spikelet development and survival. Red box and arrow indicate the promotion of spikelet abortion. Ambiguous regulation is indicated by a question mark

Cai et al. (2014) reported a critical role of jasmonic acid (JA) in regulating the determinacy of rice floral meristem and spikelet morphogenesis (*Extra Glume 1* [*EG1*] and *EG2*). However, we did not detect any candidate genes for JA, brassinosteroids and strigolactones near the associated regions, which could be attributed to either rare allele segregation of these genes in our study or the developmental differences in these species.

The associated genomic region at 5H for DS, UDS, spike length and developmental stages measured in GDD coincides with cellulose synthase gene *HvCslF7* that regulates polysaccharides and β-glucan synthases. The spatiotemporal role of *HvCsl* genes in controlling grain development and how polysaccharides are used by the plant during spikelet and spike development remain unknown. In this study, two genes encoding sucrose synthase (SuSy), which is an essential glycosyltransferase enzyme for sugar metabolism, were detected on chromosome 7H. SuSy is a member of the glycosyltransferase-4 subfamily that includes trehalose synthase and trehalose phosphorylase (Granot & Stein, 2019). SuSy plays a crucial role in the shoot apical meristem and in leaf development in transgenic tomato by altering the expression of auxin-related genes (Goren et al., 2017). Plants overexpressing SuSy have shown vigorous growth by modulating the meristem.

Chromosome 5H genomic region harbors NAC transcription factor gene, which is considered to be one of the largest plant transcription factor families regulate a wide range of different developmental processes (Nadolska-Orczyk et al., 2017). NAC showed strong expression in the developing caryopsis and grain filling stages by modulating programmed cell death in the endosperm during grain maturation (van Doorn & Woltering, 2005). NAC has also been associated with spikelet/floret sterility in modern wheat cultivars (Alqudah, Haile, et al., 2020) suggesting its potential biological function in controlling spikelet development. Finally, chromosome 2H peaks reside near the gene encoding *APETALA2* (AP2), which is known to modify spike density by controlling internode length along the axis of barley spike (Houston et al., 2013). Recent studies revealed that *HvAP2* acts in the peduncle internode by modulating the JA pathway. It suppresses internode cell proliferation and expansion, which are thought to be involved in spikelet abortion (Patil et al., 2019).

This study provided valuable insights into the importance of spikelet development, abortion and grain yield. The results showed that the MYP stage is genotype dependent, in addition to its response to environmental conditions. MYP determined grain number in two-rowed barley while spikelet abortion was the key player in final grain number in six-rowed barley. Novel genomic regions were found in the study, for example, on chromosome 7H harbouring essential candidate genes for sugar metabolism. We conclude that the presence of sucrose-related genes is crucial for spikelet development and survival at the tip of the spike, which is where most of the spikelet abortions occurred. In addition, peaks near the genes related to phytohormone genes, for example auxin, ABA, cytokinin and JA, were observed. The peaks signify the involvement of plant hormones in spikelet abortion and development. Finally, the present study provides new routes to improve grain yield as we showed that genetic dissection of spikelet development identifies both novel and known genes.

ACKNOWLEDGEMENTS

We thank Claudia Krebes, Anette Marlow and Stefanie Sehmisch for excellent technical assistance. We also thank Gudrun Schütze for help with images of spikes and IPK gardeners for help during this work. This work was funded by the Leibniz Institute of Plant Genetics and Crop Plant Research (IPK), Germany.

CONFLICT OF INTEREST

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

AUTHOR CONTRIBUTIONS

Conceived and designed the experiments: Ahmad M. Alqudah and Andreas Börner. Performed the experiments: Ahmad M. Alqudah. Conducted phenotyping measurement in the field: Ahmad M. Alqudah. Analysed the data: Ahmad M. Alqudah and Rajiv Sharma. Wrote the paper: Ahmad M. Alqudah, Rajiv Sharma and Andreas Börner.

ORCID

Ahmad M. Alqudah https://orcid.org/0000-0002-0436-9724
Rajiv Sharma https://orcid.org/0000-0003-3488-8688

REFERENCES

Alqudah, A. M., Haile, J. K., Alomari, D. Z., Pozniak, C. J., Kobiljski, B., & Börner, A. (2020). Genome-wide and SNP network analyses reveal genetic control of spikelet sterility and yield-related traits in wheat. *Scientific Reports*, *10*(1), 2098.

Alqudah, A. M., Sallam, A., Stephen Baenziger, P., & Börner, A. (2020). GWAS: Fast-forwarding gene identification and characterization in temperate Cereals: Lessons from Barley – A review. *Journal of Advanced Research*, *22*, 119–135.

Alqudah, A. M., & Schnurbusch, T. (2014). Awn primordium to tipping is the most decisive developmental phase for spikelet survival in barley. *Functional Plant Biology*, *41*(4), 424–436.

Alqudah, A. M., Sharma, R., Pasam, R. K., Graner, A., Kilian, B., & Schnurbusch, T. (2014). Genetic dissection of photoperiod response based on GWAS of pre-anthesis phase duration in spring barley. *PLoS One*, *9*(11), e113120.

Arisnabarreta, S., & Miralles, D. J. (2004). The influence of fertiliser nitrogen application on development and number of reproductive primordia in field-grown two- and six-rowed barleys. *Australian Journal of Agricultural Research*, *55*(3), 357–366.

Arisnabarreta, S., & Miralles, D. J. (2006). Floret development and grain setting in near isogenic two- and six-rowed barley lines (*Hordeum vulgare* L.). *Field Crops Research*, *96*(2–3), 466–476.

Banik, S., Mahato, K. K., Antonini, A., & Mazumder, N. (2020). Development and characterization of portable smartphone-based imaging device. *Microscopy Research and Technique*, *83*(11), 1336–1344.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1).

Behrouzi, P., Arends, D., & Wit, E. C. (2017). netgwas: An R package for network-based genome-wide association studies. arXiv preprint arXiv:1710.01236.

Boussora, F., Allam, M., Guasmi, F., Ferchichi, A., Rutten, T., Hansson, M., Youssef, H. M., & Borner, A. (2019). Spike developmental stages and

ABA role in spikelet primordia abortion contribute to the final yield in barley (*Hordeum vulgare* L.). *Botanical Studies*, *60*(1), 13.

Cai, Q., Yuan, Z., Chen, M., Yin, C., Luo, Z., Zhao, X., Liang, W., Hu, J., & Zhang, D. (2014). Jasmonic acid regulates spikelet development in rice. *Nature Communications*, *5*(1), 3476.

Colmsee, C., Beier, S., Himmelbach, A., Schmutzer, T., Stein, N., Scholz, U., & Mascher, M. (2015). BARLEX – The barley draft genome explorer. *Molecular Plant*, *8*(6), 964–966.

del Moral, L. G., Miralles, D. J., & Slafer, G. (2002). Initiation and appearance of vegetative and reproductive structures throughout barley development. In G. A. Slafer, J. L. Molina-Cano, R. Savin, J. Araus, & I. Romagosa (Eds.), *Barley science: Recent advances from molecular biology to agronomy of yield and quality* (pp. 243–267). Food Product Press.

Ghiglione, H. O., Gonzalez, F. G., Serrago, R., Maldonado, S. B., Chilcott, C., Cura, J. A., Miralles, D. J., Zhu, T., & Casal, J. J. (2008). Autophagy regulated by day length determines the number of fertile florets in wheat. *The Plant Journal*, *55*(6), 1010–1024.

Goren, S., Lugassi, N., Stein, O., Yeselson, Y., Schaffer, A. A., David-Schwartz, R., & Granot, D. (2017). Suppression of sucrose synthase affects auxin signaling and leaf morphology in tomato. *PLoS One*, *12*(8), e0182334.

Gou, W., Li, X., Guo, S., Liu, Y., Li, F., & Xie, Q. (2019). Autophagy in plant: A new orchestrator in the regulation of the phytohormones homeostasis. *International Journal of Molecular Sciences*, *20*(12), 2900.

Granot, D., & Stein, O. (2019). An overview of sucrose synthases in plants. *Frontiers in Plant Science*, *10*, 95.

Higgins, J. D. (2013). Analyzing meiosis in barley. In W. P. Pawlowski, M. Grelon, & S. Armstrong (Eds.), *Plant meiosis: Methods and protocols* (pp. 135–144). Humana Press.

Houston, K., McKim, S. M., Comadran, J., Bonar, N., Druka, I., Uzrek, N., Cirillo, E., Guzy-Wrobelska, J., Collins, N. C., Halpin, C., Hansson, M., Dockter, C., Druka, A., & Waugh, R. (2013). Variation in the interaction between alleles of HvAPETALA2 and microRNA172 determines the density of grains on the barley inflorescence. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(41), 16675–16680.

Hussain, W., Campbell, M., Walia, H., & Morota, G. (2018). ShinyAIM: Shiny-based application of interactive Manhattan plots for longitudinal genome-wide association studies. *Plant Direct*, *2*(10), e00091.

Kernich, G. C., Halloran, G. M., & Flood, R. G. (1997). Variation in duration of pre-anthesis phases of development in barley (*Hordeum vulgare*). *Australian Journal of Agricultural Research*, *48*(1), 59–66.

Kirby, E., & Appleyard, M. (1987). *Cereal development guide*. NAC Cereal Unit.

Lipka, A. E., Tian, F., Wang, Q., Peiffer, J., Li, M., Bradbury, P. J., Gore, M. A., Buckler, E. S., & Zhang, Z. (2012). GAPIT: Genome association and prediction integrated tool. *Bioinformatics*, *28*(18), 2397–2399.

Lunn, J. E., Delorge, I., Figueroa, C. M., Van Dijck, P., & Stitt, M. (2014). Trehalose metabolism in plants. *The Plant Journal*, *79*(4), 544–567.

Marzec, M., & Alqudah, A. M. (2018). Key hormonal components regulate agronomically important traits in barley. *International Journal of Molecular Sciences*, *19*(3).

Mascher, M., Gundlach, H., Himmelbach, A., Beier, S., Twardziok, S. O., Wicker, T., Radchuk, V., Dockter, C., Hedley, P. E., Russell, J., Bayer, M., Ramsay, L., Liu, H., Haberer, G., Zhang, X.-Q., Zhang, Q., Barrero, R. A., Li, L., Taudien, S., … Stein, N. (2017). A chromosome conformation capture ordered sequence of the barley genome. *Nature*, *544*(7651), 427–433.

McKim, S. M., Koppolu, R., & Schnurbusch, T. (2018). Barley inflorescence architecture. In N. Stein & G. J. Muehlbauer (Eds.), *The barley genome* (pp. 171–208). Springer International Publishing.

Milner, S. G., Jost, M., Taketa, S., Mazón, E. R., Himmelbach, A., Oppermann, M., Weise, S., Knüpffer, H., Basterrechea, M., König, P., Schüler, D., Sharma, R., Pasam, R. K., Rutten, T., Guo, G., Xu, D., Zhang, J., Herren, G., Müller, T., … Stein, N. (2019). Genebank genomics highlights the diversity of a global barley collection. *Nature Genetics*, *51*(2), 319–326.

Miralles, D. J., Richards, R. A., & Slafer, G. A. (2000). Duration of the stem elongation period influences the number of fertile florets in wheat and barley. *Functional Plant Biology*, *27*(10), 931–940.

Nadolska-Orczyk, A., Rajchel, I. K., Orczyk, W., & Gasparis, S. (2017). Major genes determining yield-related traits in wheat and barley. *TAG. Theoretical and Applied Genetics.*, *130*(6), 1081–1098.

Nagel, M., Alqudah, A. M., Bailly, M., Rajjou, L., Pistrick, S., Matzig, G., Börner, A., & Kranner, I. (2019). Novel loci and a role for nitric oxide for seed dormancy and preharvest sprouting in barley. *Plant, Cell and Environment*, *42*(4), 1318–1327.

Patil, V., McDermott, H. I., McAllister, T., Cummins, M., Silva, J. C., Mollison, E., Meikle, R., Morris, J., Hedley, P. E., Waugh, R., Dockter, C., Hansson, M., & McKim, S. M. (2019). APETALA2 control of barley internode elongation. *Development*, *146*(11). https://doi.org/10.1242/dev.170373

Rapazote-Flores, P., Bayer, M., Milne, L., Mayer, C.-D., Fuller, J., Guo, W., Hedley, P. E., Morris, J., Halpin, C., Kam, J., McKim, S. M., Zwirek, M., Casao, M. C., Barakate, A., Schreiber, M., Stephen, G., Zhang, R., Brown, J. W. S., Waugh, R., & Simpson, C. G. (2019). BaRTv1.0: An improved barley reference transcript dataset to determine accurate changes in the barley transcriptome using RNA-seq. *BMC Genomics*, *20*(1), 968.

Rather, Z. A., Khuroo, A. A., Dar, A. R., & Dar, T. U. H. (2019). Smartphone-integrated field microscopy (SPFM): A low-cost and portable tool to study live biological specimens in the wild. *Plant Biosystems – An International Journal Dealing with all Aspects of Plant Biology*, *154*(5), 757–765.

RStudio-Team. (2015). *RStudio: Integrated development for R. R: A language and environment for statistical computing*. R Foundation for Statistical Computing: RStudio, Inc. http://www.rstudio.com/

Sharma, R., Shaaf, S., Neumann, K., Go, Y., Mascher, M., David, M., Al-Yassin, A., Özkan, H., Blake, T., Hübner, S., Castañeda-Álvarez, N., Grando, S., Ceccarelli, S., Baum, M., Graner, A., Coupland, G., Pillen, K., Weiss, E., Mackay, I., … Kilian, B. (2020). On the origin of photoperiod non-responsiveness in barley. *bioRxiv*: 2020.2007.2002.185488.

van Doorn, W. G., & Woltering, E. J. (2005). Many ways to exit? Cell death categories in plants. *Trends in Plant Science*, *10*(3), 117–122.

VSN International. (2016). *VSN. GenStat for Windows 18th edition*. VSN International. GenStat.co.uk

Waddington, S. R., Cartwright, P. M., & Wall, P. C. (1983). A quantitative scale of spike initial and pistil development in barley and wheat. *Annals of Botany*, *51*(1), 119–130.

Wei, T., & Simko, V. (2017). corrplot: Visualization of a correlation matrix. *R Package Version*, *0.84 230(231)*, 11.

Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis*. Springer.

Youssef, H. M., Koppolu, R., Rutten, T., Korzun, V., Schweizer, P., & Schnurbusch, T. (2020). Correction to: Genetic mapping of the labile (lab) gene: A recessive locus causing irregular spikelet fertility in labile-barley (*Hordeum vulgare* convar. labile). *Theoretical and Applied Genetics*, *133*(9), 2759.

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the Supporting Information section.

How to cite this article: Alqudah AM, Sharma R, Börner A. Insight into the genetic contribution of maximum yield potential, spikelet development and abortion in barley. *Plants, People, Planet*. 2021;00:1–16. https://doi.org/10.1002/ppp3.10203