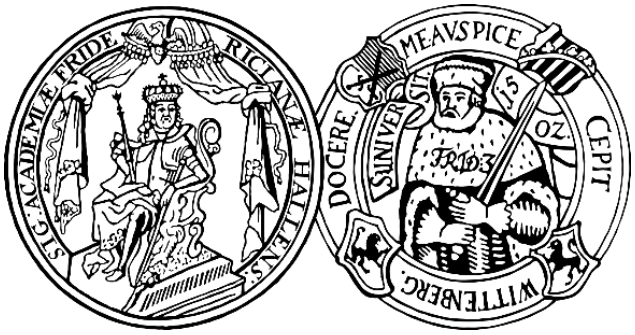
Insights into the genetics of pre-anthesis tip degeneration and related traits in barley (*Hordeum vulgare* L.)



Kumulative Dissertation

zur Erlangung des Doktorgrades der Agrarwissenschaften

Doctor agriculturae (Dr. agr.)

der Naturwissenschaftlichen Fakultät III Agrar- und Ernährungswissenschaften, Geowissenschaften und Informatik

der Martin-Luther-Universität Halle-Wittenberg

vorgelegt von

Frau Roop Kamal Muqaddasi

Gutachter

1. Prof. Dr. Thorsten Schnurbusch (IPK Gatersleben, Universität Halle-Wittenberg
2. Prof. Dr. Kevin P. Smith (University of Minnesota, USA)

verteidigt am 28.04.2025 in Haale (Saale)

ACKNOWLEDGMENTS

I express my deepest appreciation to all those who supported me throughout my doctoral journey. First, I express my sincere gratitude to Prof. Thorsten Schnurbusch for the opportunity to be part of his laboratory and to work on an interesting project. Prof. Schnurbusch's guidance and feedback encouraged me to grow as a scientist, and his helpful and friendly nature created a productive and positive work atmosphere that helped me achieve my targets.

I deeply appreciate all the present and former Plant Architecture (PBP) group members for their support during the field trials. Especially, I thank Dr. Yongyu Huang for his suggestions and advice during the research. I also thank my friends, Nandhakumar Shanmuguraj and Guojing Jiang, for the friendly office atmosphere and for making the past years very enjoyable. I very much thank the efforts of the technical staff of PBP group including Corinna Trautewig, Kerstin Wolf, Ellen Weiβ, and former members, Angelika Pueschel, and Franziska Backhaus for their outstanding support. I appreciate the hard work of IPK gardeners in caring for the plants.

I extend my thanks to Dr. Andreas Börner, Dr. Martin Mascher, and Dr. Murukarthick Jayakodi for providing the plant material and genotypic data, which were an indispensable part of my research. A special thanks to Dr. Britt Leps for her assistance in administrative work and personal affairs.

I am very grateful to my loving parents, Mr. Ranjeet Singh and Mrs. Jasvir Kaur, and my sweet brother Ekam Singh for believing in me, encouraging me to follow my dreams, and continuously supporting me with their unconditional love. I remember and pay my heartfelt thanks to my late grandparents, Mr. Karnail Singh and Mrs. Surjeet Kaur, for their blessings.

In closing, I thank my beloved husband, Quddoos Muqaddasi, who encouraged me at every step and greatly helped me—both personally and professionally. I appreciate his scientific contributions to my career, and all the valuable discussions and fun we had along the road.

Roop Kamal Muqaddasi
IPK Gatersleben

TABLE OF CONTENTS

CHAPTER 1. GENERAL INTRODUCTION 1

1.1 Origin and domestication 1

1.2 Barley production and uses 2

1.3 Barley classification 3

1.3.1 Classification based on growth habits 3
1.3.2 Classification based on gene pools 3
1.3.3 Classification based on lateral spikelet fertility 3

1.4 Barley growth and development stages 4

1.4.1 Vegetative phase 5
1.4.2 Reproductive phase 6
1.4.3 Grain-filling phase 7
1.4.4 Waddington scale 7
1.4.5 Regulation of phase transition in barley 8

1.5 Progression of inflorescence development in grasses 9

1.6 Spikelet/floret degeneration in cereals 10

1.6.1 Pre-anthesis tip degeneration in barley 12
1.6.2 Hypotheses for spikelet or floret degeneration 13
1.6.3 Barley pre-anthesis tip degeneration in comparison to other cereals 14

1.7 Association mapping in plants 15

1.7.1 Selection of barley association panel for this dissertation 17

1.8 Objectives of the dissertation 17

CHAPTER 2. Influence of potential and final spikelet number on pre-anthesis tip degeneration 18

2.1 Abstract 18

2.2 Introduction 18

2.3 Material and methods 20

2.3.1 Panel selection, preliminary evaluation, and field trials 20
2.3.2 Investigated traits and their phenotyping 22
2.3.3 Phenotypic data collection via microscopic dissection 23
2.3.4 Phenotypic data collection on the field 23
2.3.5 Pre-anthesis tip degeneration calculation 23
2.3.6 Within-year data analyses 23
2.3.7 Across-years data analyses 24
2.3.8 Correlations and path analysis among the investigated traits 25
2.3.9 Calculation of direct, indirect, and total effects via best-fit path analysis model 25

2.4 Results 26

2.4.1 Potential spikelet number bears a large and significant genotypic variation and significant positive correlation with heading date and plant height in six-rowed barleys 26
2.4.2 Final spikelet number is highly heritable and shows a significant positive correlation with potential spikelet number, heading date, and plant height 29

2.4.3 PTD's genotypic diversity highlights the potential for high selection gain and is negatively associated with final spikelet number 30
2.4.4 Wide variation for heading date and plant height in six-rowed barley accessions 31
2.4.5 The geographical origin of six-rowed barley accessions revealed distinct patterns for the investigated traits 31
2.4.6 Path analysis highlights the direct and indirect relationships among the investigated traits 33
2.5 Discussion 37
2.5.1 Path analysis revealed concealed impact of potential spikelet number on pre-anthesis tip degeneration 37
2.5.2 The geographical origin affects pre-anthesis tip degeneration in six-rowed barleys 38
2.5.3 Effect of within-year environmental fluctuations on the variance and correlation analyses 39
CHAPTER 3. Relationship of pre-anthesis tip degeneration with spike, grain, and shoot traits 40
3.1 Abstract 40
3.2 Introduction 41
3.3 Material and Methods 43
3.3.1 Plant materials 43
3.3.2 Phenotyping 43
3.3.3 Variance component analyses and calculation of BLUEs 44
3.3.4 Principal component and correlation analyses 46
3.3.5 Regression analyses: linear, quadratic, multiple, and generalized additive models 46
3.4 Results 47
3.4.1 Variance component analyses show large genotypic variance and high heritability estimates 47
3.4.2 Principal component analysis revealed the opposite genetic nature of spike and grain morphological traits 50
3.4.3 Correlation analyses show stronger interdependencies among the grain traits as compared to spike and shoot traits 51
3.4.4 Hierarchical clustering highlights six distinct clusters separating spike and shoot traits from grain traits 52
3.4.5 Geographical origin of the accessions significantly impact the variation for all the traits 54
3.5 Discussion 56
3.5.1 Plausible relationship between awn length, pre-anthesis tip degeneration and grain development. 56
3.5.2 Genetic interactions of PTD with other spike- and grain-traits is generally better revealed by generalized additive models as compared to linear and quadratic regressions 58
3.5.3 Increased days to heading increases PSN and FSN but decreases grain set and thousand-grain weight 59
CHAPTER 4. Genome-wide association analyses for pre-anthesis tip degeneration and other investigated traits 60
4.1 Abstract 60
4.2 Introduction 60
4.3 Material and Methods 62
4.3.1 Barley association mapping panel: selection criteria, field trials, and phenotyping 62
4.3.2 Whole genome shotgun sequencing, SNP calling, population structure, and phylogenetic analyses 63

4.3.3 Genome-wide association studies: selection of best-fit models, QTL identification, and reference phenotype-genotype map 63
4.3.4 LD clumping, unique and shared QTL, LD-based candidate gene identification, and expression analyses 64
4.3.5 TILLING for HvRAP gene 64
4.3.6 Chlorophyll measurements 65
4.3.7 Chlorophyll fluorescence analyses 65
4.4 Results 66
4.4.1 A diverse and representative panel of six-rowed spring barleys is a resource for high-resolution genetic studies 66
4.4.2 QTL for spikelet number traits, awn length, heading date, grain number, and grain set 70
4.4.3 QTL network analysis unfolds shared genetic architecture among barley traits 74
4.4.4 GWAS reveals unique genetic architecture of barley spike PTD 76
4.4.5 HvRAP affects chloroplast structure, photosystem II efficiency, and spikelet and grain number in barely 83
4.5 Discussion 87
CHAPTER 5. GENERAL DISCUSSION 90
5.1 Evolutionary reason behind spikelet/floret degeneration 90
5.2 Proxy traits for PTD and importance of panel selection to investigate PTD 91
5.3 Image analyses pipelines and genomic selection may help exploit PTD in breeding and genetics studies 91
5.4 *RFL* genes in relation to pre-anthesis tip degeneration 93
5.5 Potential association of *HvRAN2* with barley pre-anthesis tip degeneration 96
5.6 Putative candidate genes for final spikelet number, plant height, and culm dry weight 98
CHAPTER 6. SUMMARY 100
CHAPTER 7. Zusammenfassung 102
REFERENCES 104

FIGURE LEGENDS

Figure 1.1 Harvestable area and production for major cereals. 1
Figure 1.2 The top ten barley (a) growing and (b) producing countries in 2021. 2
Figure 1.3 Different spike images based on the lateral spikelet fertility. 4
Figure 1.4 Barley growth stages based on the Zadok Scale (Zadoks et al., 1974). 5
Figure 1.5 Vegetative or leaf initiation phase in barley. 6
Figure 1.6 Different stages in the early reproductive phase in barley. 7
Figure 1.7 Floral degeneration in cereals. 10
Figure 1.8 Progression of spikelet initiation and development and pre-anthesis tip degeneration in barley. 13
Figure 2.1 Progression of pre-anthesis tip degeneration (PTD) in barley. 20
Figure 2.2 Geographical distribution of the six-rowed barley panel. 21
Figure 2.3 Panel distribution based on the Gene material and Continents. 22
Figure 2.4 Variance component and heritability analyses. 27
Figure 2.5 Phenotypic distribution and correlation of the investigated traits in a panel of 417 six-rowed spring barley accessions. 28
Figure 2.6 Environment (year) specific phenotypic distribution of the investigated traits in a panel of 417 six-rowed spring barley accessions. 28
Figure 2.7 Comparison of accessions according to their geographical origin. 32
Figure 2.8 Path diagram elucidating the relationship between the traits. 34
Figure 3.1 Proportion of the different variance components and heritability for each investigated trait in a panel of 417 six-rowed spring barley accessions. 47
Figure 3.2 Phenotypic distribution of the investigated traits in a panel of 417 six-rowed spring barley accessions. 49
Figure 3.3 Scree and PCA plots for the investigated traits. 50
Figure 3.4 Hierarchical clustering based on the Pearson-product moment correlation analysis among the BLUEs of the investigated traits. 53
Figure 3.5 Comparison of accessions according to their geographical origin for the spike traits. 54
Figure 3.6 Comparison of accessions according to their geographical origin for the grain traits. 55
Figure 3.7 Comparison of accessions according to their geographical origin for the shoot traits. 55
Figure 4.1 Diversity of the six-rowed barley panel used for GWAS. 67
Figure 4.2 Summary of genome-wide association studies depicted as Manhattan plots where x-axis shows seven barley chromosomes and y-axis shows the marker significances as –log10 (P). 68
Figure 4.3 Summary of genome-wide association studies depicted as Manhattan plots where x-axis shows seven barley chromosomes and y-axis shows the marker significances as –log10 (P). 69
Figure 4.4 Shared QTL among the investigated traits. 72
Figure 4.5 QTL based network among the spike, grain and shoot traits. 73
Figure 4.6 A reference genotype-phenotype map highlighting the most significant trait-associated markers 75
Figure 4.7 Allelic distribution, expression and sequencing analyses reveal association of *HvRAP* with pre-anthesis tip degeneration (PTD) in barley. 81
Figure 4.8 Trait comparison and chlorophyll concentration with respect to major haplotypes for pre-anthesis tip degeneration. 82
Figure 4.9 FluorCam results for Sebastian (WT) and *hvrap.g* TILLING lines. 84
Figure 4.10 Light microscopy (LM) and transmission electronic microscopic (TEM) analysis of chloroplasts of the 3rd leaf of cv. Sebastian and *hvrap.g* TILLING lines. 85
Figure 4.11 Spikelet and grain number analyses of plants from the segregating family of the *hvrap.g* line. 86
Figure 5.1 Expression pattern of known barley *HvRFL* genes. 94
Figure 5.2 Expression pattern of unknown barley *RFL* gene on chromosome 1H. 95
Figure 5.3 Expression pattern for barley *RAN* genes. 98
Figure 5.4 Comparison for potential spikelet number (PSN), final spikelet number (FSN) and pre-anthesis tip degeneration (PTD) with respect to major haplotypes (Hap1: $n = 201$; Hap2: $n = 27$) for the gene *HvRAN2*. 98

TABLE LEGENDS

Table 1.1 Barley growth stages based on the Waddington scale 8
Table 2.1 Path analysis summary of the three investigated models. 35
Table 3.1 List of spike, grain and shoot traits evaluated on a panel of 417 six-rowed barley accessions collected from across the globe. 43
Table 3.2 Multiple linear regression for pre-anthesis tip degeneration (PTD) 58
Table 4.1 List of high confidence genes on chromosome 1H, 3H and 7H for pre-anthesis tip degeneration in barley 77
Table 4.2 Annotation results for the Pentatricopeptoide repeat gene cluster on chromosome 1H for pre-anthesis tip degeneration. 79
Table 5.1 BART IDs along with tissue expression for the PPR gene cluster on chromosome 1H. 96

SUPPLEMENTARY TABLE LEGENDS

Table S2.1 ANOVA for the investigated traits in 2018 120
Table S2.2 ANOVA for the investigated traits in 2019 121
Table S2.3 ANOVA for the investigated traits in 2020 122
Table S2.4 Across years ANOVA for the investigated traits 123
Table S2.5 Weather data during the 2018 growing season at IPK, Gatersleben. (Submitted in a separate Excel file) 124
Table S2.6 Weather data during the 2019 growing season at IPK, Gatersleben. (Submitted in a separate Excel file) 124
Table S2.7 Weather data during the 2020 growing season at IPK, Gatersleben. (Submitted in a separate Excel file) 124
Table S2.8 Path analysis results for potential spikelet number (PSN) as an independent variable for pre-anthesis tip degeneration (PTD) 124
Table S2.9 Path analysis results for final spikelet number (PSN) as an independent variable for pre-anthesis tip degeneration (PTD) 125

Table S3.1 ANOVA for the spike, grain and shoot traits in 2018 126
Table S3.2 ANOVA for the spike, grain and shoot traits in 2019 128
Table S3.3 ANOVA for the spike, grain and shoot traits in 2020 131
Table S3.4 Across year ANOVA for the spike, grain and shoot traits 134
Table S3.5 Multiple regression results for all the traits underlining the trait displaying the highest regression coefficient for a selected trait 137
Table S3. 6 Comparison between linear, quadratic and generalised additive model (GAM) to study the relationship between each investigated trait and pre-anthesis tip degeneration 139

Table 4.1 Complete information regarding barley association panel consisting of 416 six-rowed barley accessions (Submitted in a separate excel file)
Table 4.2 Phenotypic data for the investigated traits (Submitted in a separate excel file)
Table 4.3 Bonferroni Correction (BC) and False discovery rate (FDR) values for each investigated trait (Submitted in a separate excel file)
Table 4.4 List of genes present within the most significant region/SNP on each chromosome for potential spikelet number (PSN) (Submitted in a separate excel file)
Table 4.5 List of genes present within the most significant region/SNP on each chromosome for final spikelet number (FSN) (Submitted in a separate excel file)
Table 4.6 List of genes present within the most significant region/SNP on each chromosome for awn length (AL) (Submitted in a separate excel file)
Table 4.7 List of genes present within the most significant region/SNP on each chromosome for heading date (HD). (Submitted in a separate excel file)
Table 4.8 List of genes present within the most significant region/SNP on each chromosome for grain number per spike (GNS) (Submitted in a separate excel file)
Table 4.9 List of genes present within the most significant region/SNP on each chromosome for grain set (GS) (Submitted in a separate excel file)
Table 4.10 List of genes present within the most significant region/SNP on each chromosome for spike length (SL) (Submitted in a separate excel file)
Table 4.11 List of genes present within the most significant region/SNP on each chromosome for spike weight (SW) (Submitted in a separate excel file)
Table 4.12 List of genes present within the most significant region/SNP on each chromosome for spike density (SD) (Submitted in a separate excel file)
Table 4.13 List of genes present within the most significant region/SNP on each chromosome for grain length (GL) (Submitted in a separate excel file)
Table 4.14 List of genes present within the most significant region/SNP on each chromosome for grain width (GWi) (Submitted in a separate excel file)
Table 4.15 List of genes present within the most significant region/SNP on each chromosome for grain area (GA) (Submitted in a separate excel file)

Table 4.16 List of genes present within the most significant region/SNP on each chromosome for thousand grain weight (TGW) (Submitted in a separate excel file)
Table 4.17 List of genes present within the most significant region/SNP on each chromosome for grain weight (GWe) (Submitted in a separate excel file)
Table 4.18 List of genes present within the most significant region/SNP on each chromosome for plant height (PH) (Submitted in a separate excel file)
Table 4.19 List of genes present within the most significant region/SNP on each chromosome for culm dry weight (CDW) (Submitted in a separate excel file)
Table 4.20 Protein sequence alignment of HvRAP (HORVU.MOREX.r2.3HG0193400) in difference plant species (Submitted in a separate excel file)
Table 4.21 Raw phenotypic data for two PTD haplotypes (Submitted in a separate excel file)
Table 4.22 Information regarding the TILLING lines for barley HvRAP gene (Submitted in a separate excel file)

SUPPLEMENTARY FIGURE LEGENDS

Figure S2. 1 Representative immature barley spike at maximum yield potential (MYP) stage. 141
Figure S2.2 Variance component and repeatability analyses in 2018. 142
Figure S2.3 Variance component and repeatability analyses in 2019. 143
Figure S2.4 Variance component and repeatability analyses in 2020. 144
Figure S2.5 Phenotypic distribution and correlation analysis of the investigated traits in a panel of ~288 six-rowed spring barley accessions in 2018..... 145
Figure S2.6 Phenotypic distribution and correlation analysis of the investigated traits in a panel of 417 six-rowed spring barley accessions in 2019. 146
Figure S2. 7 Phenotypic distribution and correlation analysis of the investigated traits in a panel of 417 six-rowed spring barley accessions in 2020..... 147
Figure S2.8 Comparison of weather data in three consecutive growing seasons from 2018-2020..... 148
Figure S2. 9 Comparison of accessions with respect to accession nature, viz., cultivar and landraces..... 149
Figure S2.10 Path analysis models used to study the relationship between the investigated traits. 150
Figure S2.11 Within year Pearson's product-moment correlation (*r*) among the traits by dividing the panel into two sub-groups viz., group 1 and group 2. 151

Figure S3.1 Representative samples for awn length in six-rowed barley panel. 152
Figure S3.2 Proportion of the different variance components and heritability for each investigated trait in a panel of 288 six-rowed spring barley accessions in the year 2018..... 153
Figure S3.3 Proportion of the different variance components and heritability for each investigated trait in a panel of 417 six-rowed spring barley accessions in the year 2019..... 154
Figure S3.4 Proportion of the different variance components and heritability for each investigated trait in a panel of 417 six-rowed spring barley accessions in the year 2020..... 155
Figure S3.5 Phenotypic distribution of the investigated traits in a panel of 288 six-rowed spring barley accessions in 2018. 156
Figure S3.6 Phenotypic distribution of the investigated traits in a panel of 417 six-rowed spring barley accessions in 2019. 157
Figure S3.7 Phenotypic distribution of the investigated traits in a panel of 417 six-rowed spring barley accessions in 2020. 158
Figure S3.8 Environment (year) specific phenotypic distribution of the investigated traits..... 159
Figure S3.9 Contribution of the traits to the individual principal components (PCs)..... 160
Figure S3.10 Within- and across-year Pearson-product moment correlation (*r*) analyses for spike traits..... 161
Figure S3.11 Within- and across-year Pearson-product moment correlation (*r*) analyses for grain traits..... 162
Figure S3.12 Within- and across-year Pearson-product moment correlation (*r*) analyses for shoot traits. 163
Figure S3.13 Comparison of accessions with respect to accession nature, viz., cultivar and landraces..... 164
Figure S3.14 Distribution (grouping) of the 417 barley accessions based on their awn length (AL) on 1 to 6 ordinal scale..... 165
Figure S3.15 Effect of awn length (AL) of the grain traits..... 166
Figure S3.16 Comparison of different regression models explaining the relationship between pre-anthesis tip degeneration (PTD) and all the investigated spike traits..... 167
Figure S3.17 Comparison of different regression models explaining the relationship between pre-anthesis tip degeneration (PTD) and all the investigated grain traits..... 168
Figure S3.18 Comparison of different regression models explaining the relationship between pre-anthesis tip degeneration (PTD) and all the investigated shoot traits. 169

Figure S4.1 Quantile-Quantile (QQ) plots for all the investigated traits. 170
Figure S4. 2 Unique QTL for spike traits distributed on all seven chromosomes of barley..... 171
Figure S4. 3 Linkage disequilibrium and expression pattern for *HORVU.MOREX.r2.7HG0543420.* 172
Figure S4.4 Linkage disequilibrium plots for final spikelet number on chromosome 3HL. 173
Figure S4.5 Expression pattern for (a) *HORVU.MOREX.r2.3HG0263320*, and (b) *HORVU.MOREX.r2.7HG0602460* in different tissues..... 174
Figure S4.6 Linkage disequilibrium plots for awn length 175
Figure S4.7 Shoot traits' unique QTL distribution on seven barley chromosomes. 176

Figure S4.8 Linkage disequilibrium plots for heading date.....177
Figure S4.9 Grain traits' unique QTL distribution on seven barley chromosomes. 178
Figure S4.10 Linkage disequilibrium plots for grain number spike 179
Figure S4.11 Linkage disequilibrium plots for grain set 180
Figure S4.12 Linkage disequilibrium plots for spike length 181
Figure S4.13 Linkage disequilibrium plots for spike weight.....182
Figure S4.14 Linkage disequilibrium plots for spike density.....183
Figure S4.15 Linkage disequilibrium (LD) plots for grain morphometric traits. 184
Figure S4.16 Linkage disequilibrium plots for thousand grain weight.....185
Figure S4.17 Linkage disequilibrium plots for grain weight.....186
Figure S4.18 Linkage disequilibrium plots for plant height.....187
Figure S4.19 Linkage disequilibrium plots for culm dry weight.....188
Figure S4.20 Expression pattern of (a) *HORVU.MOREX.r2.1HG0078540* annotated as zinc finger protein CONSTANS, (b) *HORVU.MOREX.r2.1HG0078570* annotated as zinc finger protein CONSTANS, and (c) *HORVU.MOREX.r2.1HG0078680* annotated as zinc finger CCCH domain protein in different tissues. .189
Figure S4.21 Linkage disequilibrium plots for spike pre-anthesis tip degeneration (PTD).....190
Figure S4.22 Gene expression and phylogenetic analysis of *HvRAN2* (*HORVU.MOREX.r2.7HG0592180*). ...191
Figure S4.23 Frequency distribution for pre-anthesis tip degeneration (PTD).....192
Figure S4.24 GWAS results for spike pre-anthesis tip degeneration (PTD) in 409 six-rowed spring barley accessions. 193
Figure S4.25 Comparison of GWAS results for PSN, FSN, PTD and SL between panel with 416 accessions and panel with 409 accessions.....194
Figure S4.26 Comparison of GWAS results for SW, SD, AL and GNS between panel with 416 accessions and panel with 409 accessions.....195
Figure S4.27 Comparison of GWAS results for GL, GWi, GA and GWe between panel with 416 accessions and panel with 409 accessions.....196
Figure S4.28 Comparison of GWAS results for GS, TGW, HD, PH and CDW between panel with 416 accessions and panel with 409 accessions. 197
Figure S4.29 PCA plot distinguishing the major haplotypes for PTD and accessions based on the acquisition year 198

LIST OF ABBREVIATIONS

Abbreviation	Description
AL:	Awn length
AP:	Awn primordium
CDW:	Culm dry weight
DR:	Double ridge
FDR:	False discovery rate
FNN:	Final (rachis) node number
FSN:	Final spikelet number
GA:	Grain area
GEMMA:	Genome-wide efficient mixed-model association
GL:	Grain length
GNS:	Grain number per spike
GS:	Grain set
GWAS:	Genome-wide association studies
GWe:	Grain weight
GWi:	Grain width
GY:	Grain yield
H^2**:**	Broad-sense Heritability
HD:	Heading date
IM:	Inflorescence meristem
LD:	Linkage disequilibrium
LMM:	Linear mixed model
LP:	Lemma primordium
MAF:	Minor allele frequency
MTA:	Marker trait association
MYP:	Maximum yield potential
PCA:	Principal component analysis
PH:	Plant height
PNN:	Potential (rachis) node number
QQ:	Quantile-Quantile
QTL:	Quantitative trait locus/loci
SA:	Spikelet abortion
SD:	Spike density
SL:	Spike length
SNP:	Single nucleotide polymorphism
SP:	Stamen primordium
SS:	Spikelet survival
SW:	Spike weight
TGW:	Thousand grain weight
TM:	Triple mound
W:	Waddington scale
WA:	Winter anther
Z:	Zadoks scale

A note on the organization of dissertation

The present dissertation is written in the form of a cumulative dissertation that includes a general introduction (chapter 1), three scientific articles (chapter 2-4), and a general discussion (chapter 5). At the end, the scientific articles are summarized in English and German languages (chapter 6-7).

Chapter 1 introduces barley, its production, uses, growth and developmental stages, association mapping and its tools, pre-anthesis tip degeneration (PTD) that is also known as spikelet abortion, and details of the association panel selected to study the spike, grain, and shoot traits in relation to PTD.

Chapter 2-3 holds two already published scientific articles, mainly from traits' phenotypic analyses standpoint—the web links are provided for both articles. The chapter 4 discusses the GWAS results for all the investigated traits and is currently under review. One important point to consider is the *terminology of the degeneration process*. In the two published articles, the degeneration process is referred to as *spikelet abortion*; however, in the third manuscript draft and subsequent dissertation, the degeneration process is referred to as *pre-anthesis tip degeneration* (PTD).

Chapter 5 discusses significant findings from all three scientific articles to provide a general outlook and better understanding of the genetic nature of PTD and other investigated traits.

The chapter 4 and this dissertation bear my complete name as Roop Kamal Muqaddasi.

Chapter 1. GENERAL INTRODUCTION

Barley (*Hordeum vulgare* L.) belongs to the *Triticeae* tribe of the *Poaceae* family. The *Triticeae* tribe consist of various economically important cereals, forages, as well as 350 wild species. As illustrated below in Figure 1.1, barley ranks fourth for harvestable area after wheat (*Triticum aestivum* L.), maize (*Zea mays* L.), and rice (*Oryza sativa* L.) in the world. Barley also ranks fourth in world cereal production after maize, rice, and wheat (FAO, 2021). In 2021, ~146 million tons of barley was produced on almost 49 million hectares. The genus *Hordeum* contains both perennial species, such as *Hordeum bulbosum,* and annual species, such as *Hordeum vulgare* and *Hordeum marinum* (Bothmer, 1992) and includes 32 species and 45 taxa. *Hordeum* species have a basic chromosome number of seven with different ploidy levels, such as diploid (2n = 2x = 14), tetraploid (2n = 4x = 28), and hexaploid (2n = 6x = 42) (Komatsuda et al., 1999). Cultivated barley (*Hordeum vulgare* ssp. *vulgare*) is known to be originated from wild barley (*Hordeum spontaneum* (C. Koch) Thell or *Hordeum vulgare* ssp. *spontaneum*) (Badr et al., 2000; Kilian et al., 2009). Both cultivated and wild barley are diploid. The diploid nature of barley, its adaptability along longitudes and latitudes, self-pollinating mating system, and availability of the genome sequence and genomic resources make it a model organism for all other cereal crops in the *Triticeae* tribe (Terzi et al., 2017).

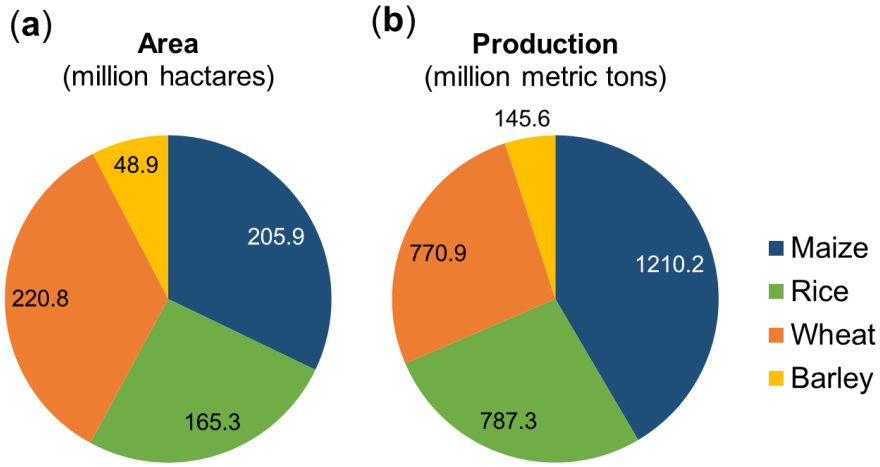


Figure 1.1 Harvestable area and production for major cereals. (a) Worldwide harvestable area, and (b) worldwide production of four major cereal crops: wheat (*Triticum aestivum* L.), maize (*Zea mays* L.), rice (*Oryza sativa* L.), and barley (*Hordeum vulgare* L.). Courtesy: http://faostat.fao.org (FAO, 2021).

1.1 Origin and domestication

Barley is one of the earliest domesticated crops and is considered one of the founder crops of "old-world" agriculture (Zohary & Hopf, 2000). Archaeological evidence shows that the transition from hunter-gathering to agriculture occurred in the Fertile Crescent ~12,000–9,500 years ago, and agriculture spread throughout Asia, Europe and Africa (Pourkheirandish et al., 2015). Wild barley naturally grows in southwest Asia, ranging from the eastern Mediterranean regions to the semi-deserts of Afghanistan (Harlan & Zohary, 1966). It is believed that barley was domesticated more than once. For instance, more than 10,000 years ago, barley was first

cultivated in the Fertile Crescent (now Israel and Jordan) and, therefore, Fertile Crescent was regarded as the first center of domestication for barley. The east of Fertile Crescent was considered the second domesticated area and contributed to Central and East Asian barleys. The third domesticated area includes the Himalayas, Eritrea, and Ethiopia. (Nevo et al., 1986; Badr et al., 2000; Morrell & Clegg, 2007; Orabi et al., 2007). Tibetan Plateau and surrounding areas is also considered as another center of barley domestication and cultivation (Dai et al., 2012; Ren et al., 2013). In conclusion, barley domestication regions extend from Central Asia to the Far East and Tibet (Wang et al., 2019).

1.2 **Barley production and uses**

Among the first domesticated crops, barley was a staple food for early farmers and was considered the "poor man's bread." The ten biggest barley-growing countries are shown in Figure 1.2a, with the Russian Federation having the highest harvestable area and production. Although Germany has the least harvestable area, its production is more than other countries (Figure 1.2b).

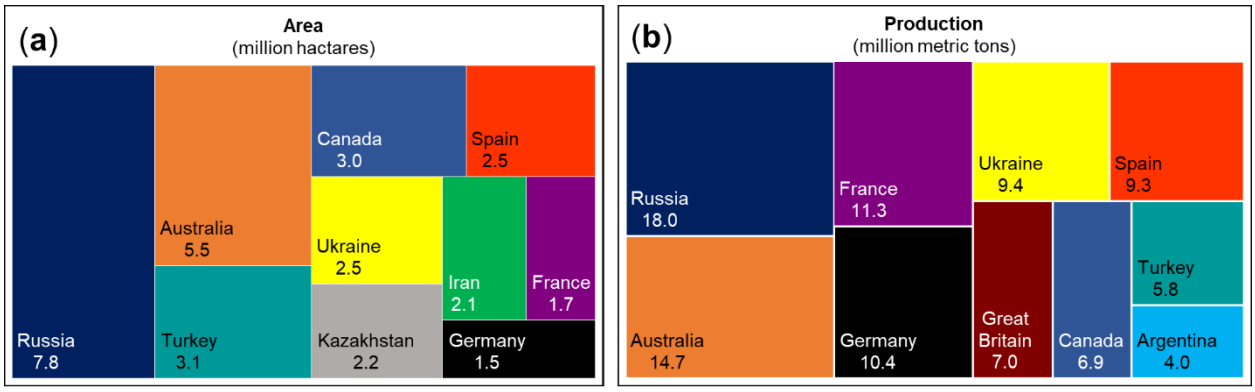


Figure 1.2 The top ten barley (a) growing and (b) producing countries in 2021. Raw data from http://faostat.fao.org.

Because of its adaptability to a wide range of environments, barley is cultivated worldwide and used in various economic segments such as stock feed, beverage industry, food sector, and biofuel production. The major portion of cultivated barley is used for stock feed as a source of crude proteins, essential amino acids, potassium, vitamin A, and β-glucan. Generally, >70% of barley is used for stock feed, ~20% in malting, brewing and distilling industries, and ~5% for human feed (Baik & E. Ullrich, 2008; Griffey et al., 2010; Tricase et al., 2018). Barley cultivars and their growing and harvesting practices help determine their use for stock feed. For example, good starch and protein in the barley grains make it a major energy source for ruminant, non-ruminant livestock, and poultry.

The second significant use of barley is malting. It is the primary cereal used in the brewing industry—both for alcoholic and non-alcoholic drinks production. Abundance of barley in brewing may be ascribed to its historic availability compared to other cereals. Apart from this, there are several reasons for its use in malting. For example, barley has tightly cemented lemma and palea physiologically, which protect the embryo during grain handling. In addition, the lemma and palea (hulls) aid in filtering the brewing mash. Finally, the steeped barley kernel has a stronger consistency than wheat and rye and can be handled at high moisture

levels with less risk of damage (Burger & LaBerge, 1985). Several non-alcoholic drinks such as barley infusions (as a coffee substitute), barley water, barley tea, and malted beverages such as malted milk are also made from barley (Tricase et al., 2018).

Barley is still used as a major staple food in several regions—this includes North Africa and Near East, highlands of Central Asia, Horn of Africa, Andean countries, and Baltic States. Barley grain has a unique chemical composition such as low-fat content, balanced protein content, and good concentration of complex carbohydrates, vitamins, and minerals—all these provide health benefits such as diabetes prevention and control, reduced cholesterol and heart disease, and weight control. Barley grain consists of about 65–68% starch, 10–17% protein, 4–9% β-glucan, 2–3% free lipids and 1.5–2.5% minerals (Quinde et al., 2004). In humans, β-glucan lowers blood cholesterol levels and glycemic index (Pins & Kaur, 2006) and controls glucose levels which in turn influence the cardiovascular health and diabetes (Baik & E. Ullrich, 2008). Barley flour is used to make bread, cakes, cookies, noodles, and snacks, and can also be blended into various food products to enhance the texture, aroma, and flavor.

1.3 **Barley classification**

1.3.1 Classification based on growth habits

Based on growth habits in Germany, barley can be divided into three types:

1. Winter barley is vernalization-sensitive, planted in late fall and harvested in the following summer. It does not flower or flower too late if sown in the spring season.

2. Spring barley is vernalization-insensitive, planted in spring and harvested in the same summer. If planted in the fall, it dies due to the low-temperature-induced shoot apex injury.

3. Facultative barleys are vernalization-insensitive and can be planted in either spring or fall, as these are cold-tolerant.

1.3.2 Classification based on gene pools

Barley germplasm can be divided into three gene pools (Harlan & de Wet, 1971):

1. Primary gene pool comprises both wild (*Hordeum vulgare* ssp. *spontaneum*) and cultivated barley.

2. Secondary gene pool contains only *Hordeum bulbosum* L.

3. Tertiary gene pool includes all other remaining species of the genus *Hordeum*.

1.3.3 Classification based on lateral spikelet fertility

Based on the lateral spikelet fertility, the cultivated barley *Hordeum vulgare* has several convarieties (Mansfeld, 1950) and the respective spike images are shown in Figure 1.3:

1. *Hordeum vulgare* convar. *distichon* is two-rowed barley where central spikelets are fully fertile, but two lateral spikelets are sterile.

2. *Hordeum vulgare* convar. *hexastichon* is six-rowed barley where both central and lateral spikelets are fertile.

3. *Hordeum vulgare* convar. *labile*—also known as "*irregular*" barley—is found in Ethiopia, Eritrea, and some North India-Pakistan districts. The labile phenotype can vary from spike to spike or even within a single plant. For example, one spike can be completely reduced to a *deficiens* phenotype with other spikes expressing various degrees of lateral spikelet fertility within an individual plant.

4. *Hordeum vulgare* convar. *deficiens* are Ethiopian two-rowed barleys that have no or extremely reduced lateral spikelets.

5. *Hordeum vulgare* convar. *intermedium* exhibits varying degrees of lateral spikelet fertility, i.e., an intermediate form between two- and six-rowed types.

Figure 1.3 Different spike images based on the lateral spikelet fertility. Left-right: *Hordeum vulgare* convar. *Distichon* (two-rowed), *Hordeum vulgare* convar. *Hexastichon* (six-rowed), *Hordeum vulgare* convar. *Labile*, *Hordeum vulgare* convar. *Deficiens*, and *Hordeum vulgare* convar. *intermedium* (Spike images from Youssef (2016)).

1.4 Barley growth and development stages

Several scales have been developed to study the barley and wheat growth stages. These scales include Zadok scale (Zadoks et al., 1974), Kirby scale (Kirby, E & Appleyard, M, 1987), Waddington scale (Waddington et al., 1983), Haun scale (Haun, 1973), and Feekes' scale (Feekes, 1941). Based on Zadok's scale, cereal growth and development is divided into ten stages, from germination to ripening (Figure 1.4). The breeders commonly use this scale to score the developmental phases in wheat and barley.

Broadly, barley growth and development is divided into three phases: vegetative, reproductive, and grain filling and maturity (Kirby, E & Appleyard, M, 1987; Slafer et al., 2002). The duration of the vegetative and early reproductive phases (together known as the pre-anthesis phase) determines the number of spikelet primordia initiated on an immature spike, while the late reproductive phase determines the number of spikelet primordia that could develop into fertile floret (Alqudah & Schnurbusch, 2014; Digel et al., 2015b).

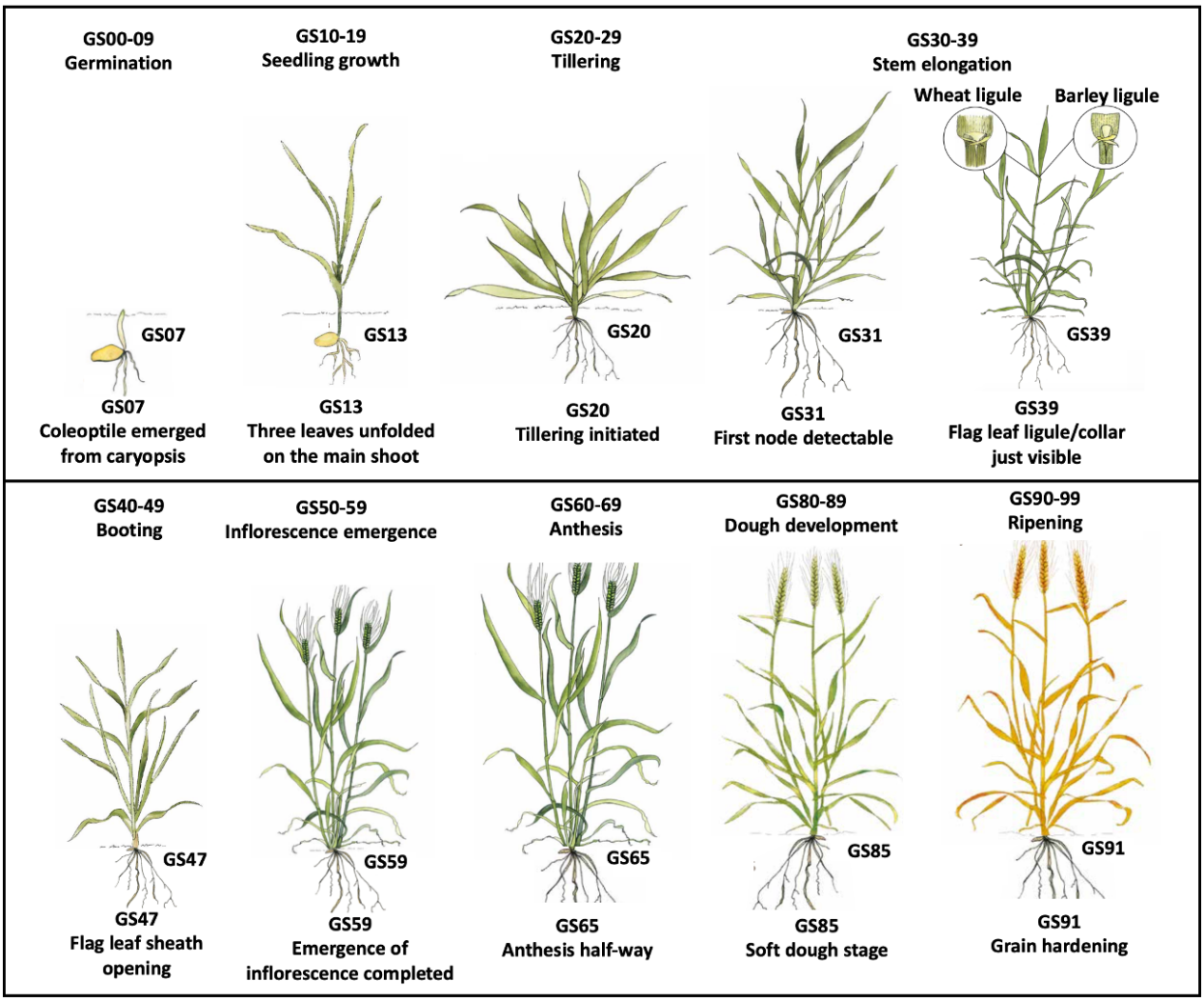


Figure 1.4 Barley growth stages based on the Zadok Scale (Zadoks et al., 1974) and the respective plant images are from AHDB et al. (2023). GS = growth stage.

1.4.1 Vegetative phase

The vegetative or leaf initiation phase is the first phase in barley development. The appearance of the leaf primordium as a visible dome marks the beginning of this phase (Figure 1.5). Two-rowed barley possesses a longer leaf initiation phase than six-rowed barley (Kirby & Riggs, 1978). This phase continues until collar formation (Kirby, E & Appleyard, M, 1987; Sreenivasulu & Schnurbusch, 2012).

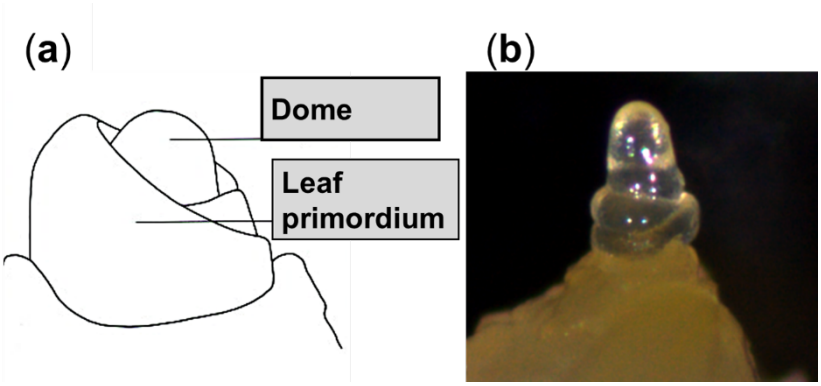


Figure 1.5 Vegetative or leaf initiation phase in barley. Left: Vegetative apex in barley, as described in Kirby, E and Appleyard, M (1987), and, Right: dissected vegetative apex in cv. Bowman.

1.4.2 Reproductive phase

It is the most extended phase in barley development, and collar formation is considered the transition point from vegetative to reproductive phase (Kirby, E & Appleyard, M, 1987; Sreenivasulu & Schnurbusch, 2012). This phase is divided into two sub-phases: early and late reproductive phases.

Early reproductive or spikelet initiation phase consists of several distinct stages: double ridge, triple mound, glume primordium, lemma primordium, stamen primordium, and awn primordium. In double ridge stage, both leaf (lower) and spikelet (upper) primordial ridges develop as a single unit. In triple mound, the upper spikelet primordial ridge develops into three mounds, one central spike meristem and two lateral spike meristems. Based on the fertility of these three spikelet meristems, barley is classified into two and six-rowed types. Glume primordium and the subsequent stages are marked by the appearance of various structures on the central spikelet. Here, the first structures to differentiate are glumes. This stage merges into the lemma primordium, in which a crescent shape structure is developed under the central and between lateral spikelets. Development of three stamen primordia is marked as stamen primordium stage, and lastly, in the awn primordium stage, lemma grows to form awns (Figure 1.6).

Late reproductive or spike growth and development phase includes tipping, heading, and anther extrusion. Awn primordium stage is considered a transition point to late reproductive stage. All stages of the anther development, e.g., white, green, and yellow stages, occur between awn primordium and tipping. Fertilization or anthesis, later, occurs between tipping and heading. The respective stage images for the late reproductive phase are illustrated in Alqudah et al. (2014).

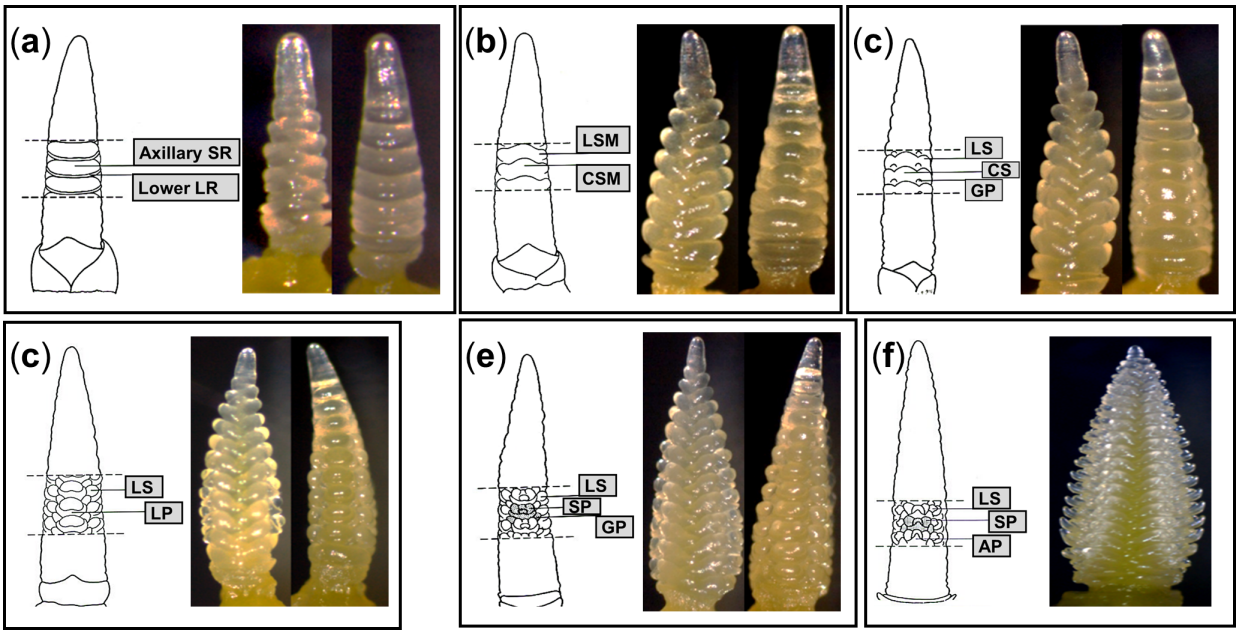


Figure 1.6 Different stages in the early reproductive phase in barley. (a) double ridge, (b) triple mound, (c) glume primordium, (d) lemma primordium, (e) stamen primordium, and (f) awn primordium, as given in Kirby, E and Appleyard, M (1987).
SR = Spikelet ridge; LR = leaf ridge; LSM = lateral spikelet mound; CSM = central spikelet mound; LS = lateral spikelet; CS = central spikelet; GP = glume primordium; LP = lemma primordium; SP = stamen primordia and AP = awn primordium.

1.4.3 Grain-filling phase

This is the last phase in barley development and is essential from the yield perspective. During this phase, the fertile spikelets that reach anthesis grow into caryopsis and ultimately to grain. This phase usually starts around 10 days post-anthesis, i.e., between heading and anther extrusion (Zadoks et al., 1974).

Each developmental phase has its importance; for example, the vegetative phase is important for leaf initiation, the early reproductive phase is vital for spikelet initiation and development, the late reproductive phase is essential for spike development, and finally, the grain-filling phase is necessary for dry matter accumulation of the caryopsis. Therefore, manipulating any of these phases remains the primary target of any barley improvement program to eventually improve the grain yield (GY).

1.4.4 Waddington scale

Waddington scale is a quantitative scale of development from seedling emergence (0) to pollination (10) and was proposed using spring barley cv. Koru and spring wheat cv. Highbury (Waddington et al., 1983). This scale is firstly based on the morphogenesis of spike initial, commencing at the "transition apex," then on floret morphogenesis, and finally on the pistil development. Different stages in the Waddington scale are listed in Table 1.1. In the present dissertation, Waddington scale was used to correctly mark the important stages.

Table 1.1 Barley growth stages based on the Waddington scale as described in Waddington et al. (1983).

Stages	Description
W1	Transition apex
W1.5	Early double ridge
W2	Double ridge
W2.25	Triple mound
W2.5	Glume primordium
W3	Lemma primordium
W3.5	Stamen primordium
W4	Pistil primordium / Awn primordium
W4.5	Carpel primordium/ Development of awns
W5	Carpel extending round three sides of ovule
W5.5	Closing of stylar canal
W6	Remain of stylar canal is visible as narrow opening
W6.5	Elongation of styles begins
W7	At the tip of styles, stigmatic branches differentiate as swollen cells
W7.5	Elongation of stigmatic branches
W8–8.5	Elongation of stigmatic branches and hair on ovary
W8.75	Elongation of stigmatic branches and hair on ovary and stigmatic branches form a tangles mass
W9	Style and stigmatic branches become erect and beginning of differentiation of stigmatic hair
W9.5	Well-developed stigmatic hairs are visible and stigmatic branches spread outwards
W10	Pollination, stigmatic branches spread wide and pollen grains are visible on the stigmatic hair

1.4.5 Regulation of phase transition in barley

In barley and wheat, floret development follows a two-phase system that starts with the initiation of spikelet primordia on the apex and (if the internal and external conditions are favorable) is followed by floral morphogenesis (Aspinall, 1966; Gol, Leonard et al., 2017). Barley is a facultative long-day plant controlled by vernalization and photoperiod response-related genes. In winter barley, the transition from vegetative to reproductive phase is controlled by vernalization response genes, i.e., *vernalization1* (*VRN1*) and *vernalization2* (*VRN2*), where *VRN1* is up-regulated during vernalization and represses the expression of *VRN2* in the leaf (von Zitzewitz et al., 2005; Sasani et al., 2009). In spring barley, vegetative to reproductive phase transition is independent of the cold temperatures as insertions and deletions in the first intron of *VRN1* cause its up-regulation (Fu et al., 2005; Szűcs et al., 2007). Moreover, spring barleys lack the functional copy of *VRN2*—either due to naturally occurring deletions of the

entire *VRN2* locus or loss-of-function mutations in the coding sequence of *VRN2* gene (Dubcovsky et al., 2005).

Another key regulator of cereal inflorescence development is *Photoperiod1* (*PPD1*) gene that, in its dominant form, accelerates flowering under long-day conditions (Turner et al., 2005). *PPD1* (encoding a *Pseudo Response Regulator*; *PRR* protein) is homologous to *Arabidopsis PRR3/PRR7* of the circadian clock and is characterized by a pseudoreceiver and a CCT domain. *PPD1* also induces the expression of *VRN3*, a homolog of *Arabidopsis Flowering locus T* (*FT*). FT protein translocate through the phloem channel from leaves to the shoot apical meristem, which induces the transition from vegetative to reproductive phase (Corbesier et al., 2007). Barley carries several *FT*-like genes, which act as central regulators of the transition from vegetative to reproductive growth phases. Natural variations in the *Ppd-H1* and *HvFT1* expression were reported to impact inflorescence development and floret fertility (Digel et al., 2015b). *FT2* is the paralog of *FT1,* and the overexpression of *FT2* in barley is associated with precocious flowering and reduced spikelet number. The *ft2* mutants in wheat also showed a prolonged spike development phase leading to a significant increase in the spikelet number. The expression analysis of wheat leaves showed that *FT2* was expressed later than *FT1* and *FT2* transcripts were detected in the shoot apical meristem and subsequently increased during early spike development. Therefore, it was suggested that in contrast to *FT1*, *FT2* plays an essential role in spike development and fertility and a limited role in the timing of the transition between the vegetative and reproductive phases (Shaw et al., 2018). Another *FT*-like gene, *HvFT3*, modified the expression of barley row-type genes such as *six-rowed-spike1* (*VRS1*), *six-rowed-spike4* (*VRS4*), *intermedium-spike c*, and several other floral homeotic genes. The overexpression *HvFT3* accelerate the initiation of spikelet primordia and the early reproductive development but not floral development (Mulki et al., 2018).

In a nutshell, the transition from vegetative to reproductive development phases is controlled by vernalization genes, whereas *PPD1* and associated *FT*-like genes are involved in inflorescence development.

1.5 **Progression of inflorescence development in grasses**

The members of the grass family show varying degrees of inflorescence complexity, and the variation in the inflorescence diversity is primarily controlled by two activities: meristem maintenance and meristem determinacy (Wang et al., 2021). All the aerial parts of the plants are generated by a group of proliferating, undifferentiated stem cells known as shoot apical meristem (SAM) (Traas & Vernoux, 2002). SAM functions both in vegetative and reproductive phases. Vegetative SAM produces leaf whorls, branches, and stems and after perceiving internal and external stimuli, vegetative SAM transitions into the reproductive SAM. The reproductive SAM, also known the inflorescence meristem (IM), produces either branch meristem (BM), spikelet meristem (SM), or floret meristems (FM). IM diversity and development is complex in grasses, leading to different types of inflorescences such as highly branched or compound inflorescence in rice (panicle) and maize tassel (racemose) to simplified spike-type inflorescence in wheat and barley (Koppolu et al., 2022).

In rice inflorescence, the determinate IM produces several primary branch meristems (PBMs) which further initiate secondary branch meristems (SBMs). Both PBMs and SBMs

elongate to produce the typical panicle architecture and later terminate into SM that differentiates into a single FM. In maize tassels, the IM first initiates several indeterminate branch meristems on which several spikelet pair meristems (SPMs) are distichously generated. These SPMs produce a pair of SMs, which develop into one pedicellate spikelet and one sessile spikelet. In case of wheat and barley, BMs are absent leading to comparatively simpler meristem differentiation and organization. Three levels of meristem organization form the spike inflorescence: (1) IM producing rachis and SMs, (2) SMs producing rachilla, and (3) FMs producing florets (Koppolu & Schnurbusch, 2019; Sakuma & Koppolu, 2023). The wheat IM is determinate and terminates into a terminal spikelet meristem. Each SM is indeterminate and produces 10–12 FMs in a distichous manner. In contrast to wheat, barley shows a peculiar meristem differentiation and organization where IM differentiates into triple spikelet meristem (TSM). The TSM further differentiates into two lateral spikelet meristem (LSM) and one central spikelet meristem (CSM), each producing a single FM. All functional SMs lead to six-rowed barley, whereas only functional CSM produce two-rowed barley. The FM ultimately converts into florets that bear grains (Koppolu & Schnurbusch, 2019). However, not all the florets survive to develop into grains. The details of the spikelet/floret degeneration are mentioned in the next section.

1.6 **Spikelet/floret degeneration in cereals**

Spikelet or floret degeneration in cereals is considered a severe constraint in improving GY as it decreases the floret number and final grain number in cereals. Hence, one of the promising avenues for enhancing GY in cereals involves reducing spikelet mortality. For example, in wheat, multiple florets degenerate within a spikelet (Kirby, 1988; Bancal, 2009; Ferrante et al., 2013b; Guo et al., 2017; Sakuma et al., 2019); in maize, the degeneration is known as ear tip barrenness (Qilin et al., 1999; Meng et al., 2007; Li et al., 2008); and, in rice panicle degeneration occurs (Wang et al., 2018; Ali et al., 2019) (Figure 1.7).

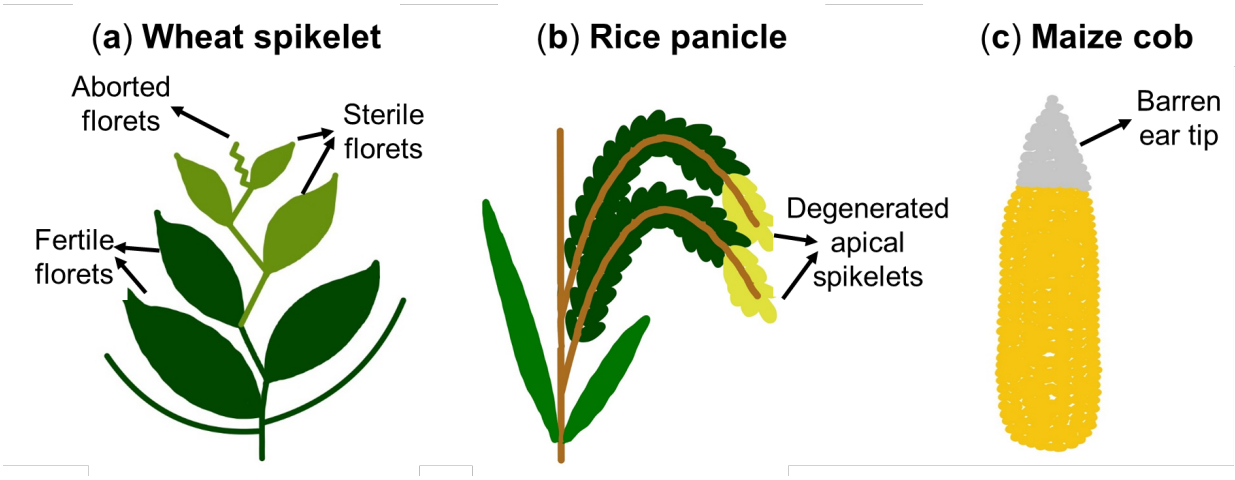


Figure 1.7 Floral degeneration in cereals. (a) Floret abortion in wheat spikelet, (b) panicle degeneration in rice and (c) ear tip barrenness in maize. Wheat spikelet figure adapted from Sakuma and Schnurbusch (2019).

Panicle degeneration in rice is well studied and based on the degeneration time and position on the panicle, it can be divided into two types: spikelet degeneration (pre-flowering

floret abortion) and spikelet barrenness (Wang et al., 2018). Two classical theories, namely, "resource limitation" and "self-organization" were proposed to explain the degeneration process (Wang et al., 2018). Based on "resource limitation" theory, low occurrence of non-structural carbohydrates causes spikelets to starve under stress leading to a strong competition for food, and as a result, inferior spikelets are forced to degenerate. However, this theory failed to explain why degeneration occurs under favorable conditions. According to the "self-organization" theory, panicle degeneration is considered a self-organized process where random migration of the initial resources leads to an imbalance of resource allocation among the developmental sinks. This imbalance leads to more resource allocation to superior sinks, whereas the inferior sink that could not get enough resources degenerate (Ganeshaiah & Uma Shaanker, 1994).

In recent years, several studies were conducted in rice to elucidate the genetic and molecular mechanism behind panicle degeneration. Several abortion mutants such as *panicle apical abortion* (*paab1-1*; Heng et al. (2018)), *degenerated panicle and partial sterility 1* (*dps1*; Zafar et al. (2020)), *apical spikelet abortion* (*asa*; Zhou et al. (2021)), *apical panicle abortion1331* (*apa1331*; Ali et al. (2022)), *panicle apical abortion 7* (*paa7*; Dai et al. (2022)), *tutou1* (Bai et al., 2015), and *tutou2* (Zhu et al., 2022) have been identified. Various players involved in panicle degeneration were identified in all these studies—for instance, deficiency of malate in the apical part of the panicle, changes in anther cuticle morphology, increase in reactive oxygen species (ROS) leading to enhanced programmed cell death, alterations in boron distribution within the panicle, increased salicylic acid levels and hydrogen peroxide accumulations, and reduction in vascular bundle number and downregulations of vital biological process in the apical part. Also, reviews such as Wang et al. (2018) and Ali et al. (2019) highlight the positive and negative regulation of various phytohormones in the degeneration process. Overall, apical panicle degeneration is a complex process attributed to changes in various essential aspects of plant development.

In maize, degeneration of female inflorescence is known as "ear tip barrenness" or "ear apical degeneration." Using microscopy and anatomical analyses, Pei et al. (2022) characterized the shortened ear mutant in maize as *ear apical degeneration 1* (*ead1*). They found that degeneration occurred due to the accumulation of ROS, leading to programmed cell death at the apex of the female inflorescence. The causal gene identified for this mutant was *EAD1,* which encodes an aluminum (Al)-activated malate transporter (AMLT). The malate concentrations were significantly reduced in the middle and apical parts of immature *ead1* ears compared to wild type, suggesting that *EAD1* plays a role in malate transport in these tissues. The short ear phenotype was rescued with the exogenous application of malate into *ead1* immature ears.

In wheat, numerous studies have been conducted to understand physiological and genetic mechanisms behind floret development and fertility (Whingwiri, EE & Stern, WR, 1982; Miralles et al., 1998; González et al., 2003c; González et al., 2003a; González et al., 2005a; González et al., 2005b; González et al., 2006; Bancal, 2008; Bancal, 2009; Ferrante et al., 2010; Ferrante et al., 2013b; Ferrante et al., 2013a; Ferrante et al., 2015; Guo et al., 2016; Guo et al., 2017). Since grain number is related to the number of fertile florets—a result of floret initiation and degeneration dynamics—these studies investigated the effect of duration of

stem elongation phase, the effect of photoperiod and vernalization genes on the stem elongation phase, resource availability, date of seeding, stem-spike competition, spike dry matter at anthesis, and stages of floret development to understand the dynamics of floret development. Based on the findings in these studies, two contrasting models were proposed to underline the floret degeneration process. These models are "trophic" and "pure development" models (Ferrante et al., 2013b). Trophic model favored that floret death is triggered by the dynamics of spike dry weight between terminal spikelet initiation and anthesis and floret death starts at the onset of rapid spike growth. Pure developmental model, on the other hand, suggests that the floret death is triggered by the developmental stage of the most advanced floret primordium of the central spikelets. Sakuma et al. (2019) identified the locus *Grain Number Increase 1* (*GNI1*)—a wheat ortholog of *Vrs1* in barley—that encodes a homeodomain leucine zipper class I transcription factor and affects the floret fertility and grain number in wheat. *GNI1* is expressed in parts of the rachilla as well as most apical florets suggesting its role in rachilla growth and development. Knockdown of *GNI1* led to an increase in the number of fertile florets and consequently the grains in hexaploid wheat. In short, *GNI1* functions as a suppressor of apical florets development within the spikelets and provides an essential finding toward understanding the floret fertility and abortion mechanism. Furthermore, *GNI1* reveals the potential of improved floret fertility by decreasing the apical degeneration as a successful approach to increase the grain number in cereals.

1.6.1 Pre-anthesis tip degeneration in barley

In barley, the number of spikelets per spike determines GY. In the spikelet primordia initiation phase, spikelets initiate as axillary meristems on the flanks of the immature spike. After passing successive developmental stages, namely, double ridge, triple mound, glume primordium, lemma primordium, stamen primordium, and awn primordium, the immature barley spikes reach the maximum yield potential (MYP) stage. At MYP stage, the maximum number of spikelet primordia are produced on the immature spike; after this, the plateau phase arrives where no new spikelet primordia are initiated. Hence, the maximum number of nodes or spikelets at the MYP stage represents the potential of the spike to produce spikelet primordia and is, therefore, designated as potential node number (PNN) or potential spikelet number (PSN). The MYP stage is retained for a few developmental stages that usually span Waddington stages 5–7 (Waddington et al., 1983; Thirulogachandar & Schnurbusch, 2021).

Not all spikelets produced by the end of the MYP stage survive until heading—this is due to pre-anthesis tip degeneration (hereafter, PTD; Figure 1.8). Barley PTD starts from the tip of IM, i.e., the activity of IM ceases and it begins to collapse, followed by the degeneration of the subjacent spikelet primordia and rachis up to a pre-destined position on the spike. During spike growth, the aborted spikelets appear desiccated as they lose their turgidity. Only those nodes or spikelets that are retained on the spike after degeneration—termed as final node number (FNN) or final spikelet number (FSN)—reach the grain filling phase (Figure 1.8). Thus, only a fraction of the PSN remains due to this floral initiation and degeneration process (Kirby, EJM & Appleyard, M, 1987; del Moral et al., 2003; González et al., 2003a). The surviving floral organs contribute to the final grain number per unit area that is reflected in the GY (Baethgen et al., 1995; Boonchoo et al., 1998; Del Moral et al., 1999; Arisnabarreta & Miralles, 2004). Nonetheless, the spikelet degeneration process is not only restricted to the apical part. A

few (usually 2–3) spikelet nodes also develop poorly and do not set grains at the base of the spike (Appleyard et al., 1982).

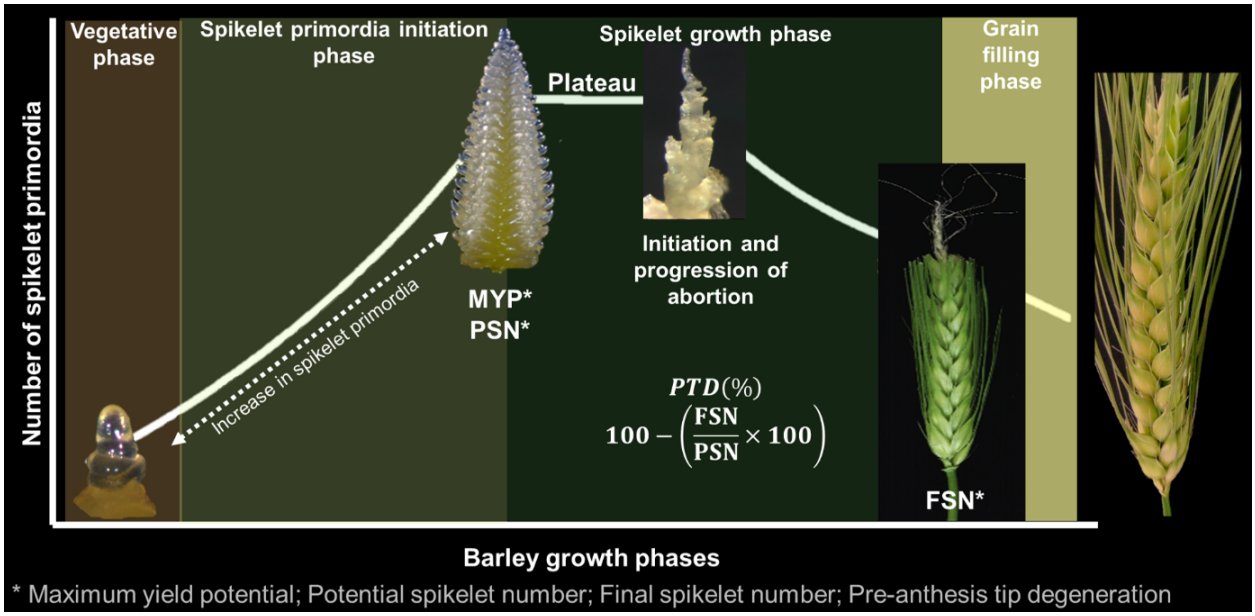


Figure 1.8 Progression of spikelet initiation and development and pre-anthesis tip degeneration in barley.

1.6.2 Hypotheses for spikelet or floret degeneration

Various causes for spikelet/floret degeneration have been hypothesized: the most favored ones being the competition for assimilates or stem-spike competition (Kirby, 1988; Arisnabarreta & Miralles, 2004; González et al., 2005a; Gonzalez et al., 2011; Ferrante et al., 2013b), competition between spikelets (Appleyard et al., 1982), the position of the spikelets within the spike (Arisnabarreta, Sebastián & Miralles, Daniel J, 2006), abortion of distal florets in wheat spikelets due to poor connection with main vascular (Wolde & Schnurbusch, 2019), abortion activation by the developmental stage of the most advanced floret in the spikelet (Whingwiri, EE & Stern, WR, 1982; Bancal, 2008; Bancal, 2009), and hormonal distribution and dynamics within the spike (Cotterell et al., 1981b; Youssef et al., 2017; Boussora et al., 2019). Some of these hypotheses for spikelet degeneration are explained below.

Competition for assimilates: In wheat, it was observed that floret death coincides with vigorous stem growth. During stem growth, the photosynthetic surface does not increase and root growth also ceases (Brooking & Kirby, 1981). At this stage, all the leaves were fully expanded and had reached maximum dry mass, i.e., there was no further increase in the supply of resources. This results in competition for resources between the stem and the florets, which leads to the death of the latter. The role of the peduncle in floret death was also emphasized as the peduncle length showed considerable variation and, hence, could be a determining factor for floret survival.

Competition between the spikelets: Appleyard et al. (1982) studied spike growth and development in the progeny of crosses between six-rowed and two two-rowed spring barley varieties and observed no differences in the rates of primordium initiation amongst the

progenies and parents. However, the proportion of spikelet primordium that ultimately survived to form grains was less in spikes with most spikelet primordia. Therefore, it was hypothesized that spikelets die due to "within" spike resource competition.

The position of spikelet within the spike: The dynamics of spikelet growth and development (Arisnabarreta, Sebastián & Miralles, Daniel J, 2006) shed light on the position of spikelets within the spike that could explain spikelet degeneration. In both barley row types, the floret primordium mortality starts at the beginning of the spike active growth phase and continues when the stem and spike growth are at their maximum rate. However, the grain setting is lower in six-rowed than in two-rowed barley. This might be due to the smaller carpels in the distal spikelet position that reduce their chance to set grains. It was, therefore, concluded that the lower grain setting in six-rowed barley was due to the fact that fewer florets survived owning to the lower degree of floret development in distal and lateral spikelet positions.

Recently, Huang et al. (2023) characterized *HvCMF4* (encoding a *CCT* [CONSTANS (CO), CO-like, and TIMING OF CAB1 (TOC1)] domain-containing *MOTIF FAMILY 4* protein) gene that is responsible for a PTD mutant known as *tip-sterile* (*tst2*) in barley. It was reported that spikelet development is altered by a vascular-specific circadian clock that regulates floral initiation and growth along with changes in chlorophyll biosynthesis and chloroplast development. Also, Shanmugaraj et al. (2023)—using transcriptome analysis—showed that senescence and defense-responsive transcription factor families, such as NAC [contains three transcription factors: NAM (no apical meristem), ATAF1-2 (*Arabidopsis thaliana* activating factor) and CUC2 (cup-shaped cotyledon)], HD-ZIPs (homeodomain-leucine zipper), bZIPs (basic leucine-zipper), and MYBs (myeloblastosis viral oncogene homolog), are amongst the putative candidate genes responsible for apical degeneration. Moreover, it was shown that barley PTD is associated with sugar depletion, amino acid degradation, and late abscisic acid (ABA) biosynthesis and signaling. Because of the genetic complexity of PTD, efforts are ongoing to discover more genetic factors underlying spike PTD in barley.

1.6.3 Barley pre-anthesis tip degeneration in comparison to other cereals

As spikelet/floret degeneration also occurs in other cereals, it is important to understand their similarities and differences with barley PTD. In short, spikelet primordia initiate until the maximum yield potential (MYP) stage in barley and elongation of the IM dome marks the completion of MYP. Afterward, the elongated IM loses its turgidity and starts to disintegrate which is followed by the degeneration of the youngest subjacent spikelet primordium. The spikelet primordia degeneration occurs in a basipetal manner and spikelet primordia at the apical part of the spike degenerate, decreasing the overall spikelet number and eventually grain number.

As per available literature, varying degrees of panicle degeneration happened in mutant plants during the later stages of plant growth in rice. For instance, in *tutou1*, *paab1-1*, *asa*, and *dps1* mutants, apical degeneration starts only after heading or when the panicle has attained its final size (12–15 cm) (Bai et al., 2015; Heng et al., 2018; Zafar et al., 2020; Zhou et al., 2021). In ear degeneration mutant (*ead1)* of maize, IM starts to shrink, and the SMs around the apex

also collapse. The apical IM and SMs degeneration extend and further deteriorate in 20 mm or bigger inflorescences, ultimately leading to abortion of SMs and FMs in *ead1* (Pei et al., 2022). The indeterminate wheat spikelet produces 10–12 floret primordia until the "green anther" stage. However, due to floret abortion, only a small fraction of the initiated floret primordia survives. Thus, the grain number per spikelet is much smaller than the actual floret primordia initiated (Guo et al., 2016). The visible floret degradation occurs over several floret development stages, especially from the green anther stage to anthesis (Guo, Z. & Schnurbusch, T., 2015). Spikelets at green anther stage harbor maximum floret primordia; however, after anthesis, only three to four floret primordia survive. It is worth mentioning that, sometimes, even the first and second florets inhibit the grain formation in the third and fourth florets that usually bear grains (Rawson & Evans, 1970).

In all the mentioned rice panicle degeneration mutants, degeneration of the apical part happened just before or after the heading or during late panicle development and also affects the middle spikelets in some mutants—this is contradictory to barley PTD where degeneration happens only before heading. Hence, though spikelet degeneration occurs both in rice and barley, the intensity and timing of degeneration are different in both crops. Furthermore, the start and the direction of degeneration are unclear in rice mutants. As mentioned above, two theories, i.e., resource limitation and self-organization, were proposed to explain panicle degeneration; however, if we consider spikelet degeneration in other cereals, it does not occur randomly. For instance, spikelet degeneration happens at the base and the top of a rice panicle, floret degeneration in wheat occurs at the distal position in a spikelet or both at base and top of a spike, and kernel abortion mainly happens at the top of maize ear. The occurrence of degeneration at a specific part of the inflorescence does not support the theory of random migration of initial resources favoring the self-organization theory, finally leading to degeneration of the inferior sink.

As wheat and barley belong to the *Triticeae* tribe, it is expected to find similarities in the degeneration process. However, changes in the inflorescence architectures make the degeneration process slightly different between these crops. For instance, in barley, IM first starts to degenerate, but in wheat, development of IM terminates into terminal spikelet meristem, i.e., there is no degradation of IM. Because of the indeterminate nature of IM in barley, numerous spikelet primordia are initiated on a young spike leading to dying of spikelets during the degeneration process. In wheat, on the other hand, spikelets are indeterminate and produce several floret primordia within the spikelet. Therefore, during degeneration, floret primordia die instead of spikelet primordia. Thus, the degeneration process of barley and wheat is similar, i.e., excess primordia (be it spikelet or floret) aborts to sustain the development of a few but well-developed middle and basal primordia. The process is different in terms of developmental timings of the degeneration (in barley after MYP and in wheat after green anther) and degenerated primordia identity.

1.7 **Association mapping in plants**

Advances in genomics, development of robust analyses tools, and improved computational facilities have helped researchers rapidly test thousands of markers for their association with important traits. Trait-linked makers representing quantitative trait loci are usually identified

via linkag mapping (usually involving bi-parental populations) and association mapping (lines with diverse genetic makeup). Association mapping is considered a powerful tool to disect genetic architecutre of traits as it offers high mapping resolution, reduced research time, and greater allelic diversity (Yu & Buckler, 2006).

According to Risch and Merikangas (1996), association mapping falls into two broad categories: candidate gene association mapping and genome-wide association mapping. The extent of linkage disequilibrium is the main factor deciding whether to opt for a genome-wide association studies (GWAS) or candidate gene-based approach. (Nordborg & Tavaré, 2002; K. & R., 2011). GWAS analyses the genetic variation at whole genome level to identify association signals for numerous complex traits and is adopted by researchers interested in conducting comprehensive genome-wide analyses of traits by testing thousands of molecular markers distributed across the genome. Candidate-gene association mapping, on the other hand, is more of a hypothesis-driven approach because mapping is restricted to genes that are good candidates for the trait of interest (Neale & Savolainen, 2004; Hall et al., 2010) and was first conducted for flowering time in maize (Remington et al., 2001). This approach is useful when no significant associations are found for a trait, even after multiple testing and correction of false discovery rate. To increase the power and precision of QTL detection, this approach can be used in parallel to GWAS, as shown in maize (Lipka et al., 2013; Gupta et al., 2014).

The factors such as phenotypic variation, population size, linkage disequilibrium, population structure, cryptic familial relatedness and allele frequency affects the precision of the GWAS. Using raw phenotypic data is not ideal for GWAS and usually, best linear unbiased estimations and best linear unbiased predictions can adjust the raw phenotypic data collected across several locations and years. Trait heritability is an important estimate as it reveals the percentage of genetic variance contributed to the phenotype. Only traits with moderate to high heritability values should be considered for GWAS. Another point to consider is the non-random association of allele or linkage disequilibrium (LD) at different loci—ignoring LD may lead to incorporation of both casual and non-casual alleles for further analyses that likely give false results. Population structure—present due to the relatedness among the individuals of a population— must be considered carefully while performing the analyses and interpreting the results. Failing to correct for population structure may lead to spurious marker-trait associations. Understanding and correcting the above-mentioned factors aid in successful GWAS analyses.

Over the recent years, GWAS was successful used to elucidate the genetic architecture of several important traits. For instance, to study spike architecture, spikelet number, grain yield, photoperiod, plant height, grain traits and nitrogen use efficiency in barley (Ramsay et al., 2011; Pasam et al., 2012; Alqudah et al., 2014; Karunarathne et al., 2020; Thabet et al., 2020). GWAS was used to study plant architecture, spikelet number, heading date, grain morphometric traits, abitotic stress in rice (Huang et al., 2010; Yano et al., 2019; Kang et al., 2020; Yuan et al., 2020). In wheat, genes or loci associated with spike morphology as well as floret fertility traits, total spikelet number, plant height, tiller number, root architecture. etc were identified via GWAS (Guo et al., 2017; Beyer et al., 2019; Muqaddasi et al., 2019; Ali et al., 2021).

1.7.1 Selection of barley association panel for this dissertation

In the present dissertation, a panel of 417 diverse six-rowed spring barleys was selected from the barley collection of the German federal *ex-situ* genebank hosted at the IPK, Gatersleben. The panel selection criterion was primarily based on the genotypes covering the maximum genotypic diversity space of domesticated six-rowed barley accessions. The selected accessions were fingerprinted with genotyping-by-sequencing (Milner et al., 2019). The selected accessions originate from different continents, i.e., from Africa (n = 73), Asia (n = 183), Europe (n = 80), North America (n = 28), and South America (n = 12). A total of 41 accessions were of unknown geographical origin. Although most accessions (n = 350; 84%) were landraces, we selected a decent proportion of recent cultivars (n = 67; 16%) to cover recent allelic diversity.

Only six-rowed accessions were selected to avoid population structure associated with the major row-type gene *Six-rowed spike 1* (*Vrs1*) (Komatsuda et al., 2007; Sakuma et al., 2013; Zwirek et al., 2019). Also, six-rowed barley occupies a significant market share in central Europe. In addition, accessions with *Ppd-H1* sensitive allele were chosen: accessions carrying *ppd-H1* insensitive allele often flower late and exhibit longer and slower growth that might affect spikelet degeneration or survival (i.e., pre-anthesis tip degeneration; PTD) by longer and slower growth of the plants (Turner et al., 2005). Besides, it was reported that six-rowed barleys generally show higher spikelet primordia degeneration (Whingwiri, EE & Stern, WR, 1982; Kernich et al., 1997; Miralles & Richards, 2000; Garcıa del Moral et al., 2002; Arisnabarreta & Miralles, 2004; Arisnabarreta, Sebastián & Miralles, Daniel J, 2006). Owing to the above-mentioned reasons, six-rowed accessions proved ideal to studying barley PTD and associated traits. The details regarding the field design, sowing conditions, and sample collection are mentioned in Kamal et al. (2022b) and the investigated traits are potential spikelet number (PSN), final spikelet number (FSN), pre-anthesis tip degeneration (PTD), spike length (SL), spike weight (SW), spike density (SD), awn length (AL), grain number per spike (GNS), grain length (GL), grain width (GWi), grain area (GA), grain weight (GWe), grain set (GS), thousand-grain weight (TGW), heading date (HD), plant height (PH), and culm dry weight (CDW).

1.8 **Objectives of the dissertation**

As degeneration of spikelet primordia is a crucial phenomenon affecting final grain number per spike, the general aim of this dissertation was to study pre-anthesis tip degeneration (PTD) in barley in depth. The specific objectives were to:

1. Analyze the natural genotypic variation for PTD in a population of six-rowed barley.
2. Analyze the PTD in relation to other spike, grain, and shoot traits.
3. Perform genome-wide association studies for PTD and related traits using SNPs developed from whole-genome shotgun sequencing.
4. Identification of genomic target regions and genes for PTD and related traits.

Chapter 2. **Influence of potential and final spikelet number on pre-anthesis tip degeneration**

Spikelet abortion in six-rowed barley is mainly influenced by final spikelet number, with potential spikelet number acting as a suppressor trait

Published in: Journal of Experimental Botany (2022) 73:2005-2020.

DOI: https://doi.org/10.1093/jxb/erab529

Authors: Roop Kamal, Quddoos H. Muqaddasi, Yusheng Zhao, Thorsten Schnurbusch

The original publication is available online at:

https://academic.oup.com/jxb/article/73/7/2005/6449485

Note: In the original publication, pre-anthesis tip degeneration (PTD) was described as spikelet abortion (SA). However, to maintain the consistency in the thesis, pre-anthesis tip degeneration was used in the following text.

2.1 Abstract

The potential to increase barley grain yield lies in the indeterminate nature of its inflorescence meristem, which produces spikelets, the basic reproductive unit in grasses that are linked to reproductive success. During early reproductive growth, barley spikes pass through the maximum yield potential—a stage after which no new spikelet ridges are produced. Subsequently, pre-anthesis tip degeneration (PTD), a phenomenon in which spikelets degenerate during spike growth, imposes a bottleneck for increasing the grain yield potential. Here, we studied the potential of main culm spikes by counting potential spikelet number (PSN) and final spikelet number (FSN), and computed the corresponding PTD (%) in a panel of 417 six-rowed spring barleys. Our phenotypic data analyses showed a significantly large within- and across-years genotypic variation with high broad-sense heritability estimates for all the investigated traits, including PTD. Asian accessions displayed the lowest PTD, indicating the presence of favourable alleles that may be exploited in breeding programs. A significantly negative Pearson's product–moment correlation was observed between FSN and PTD. Our path analysis revealed that PSN and FSN explain 93% of the observed phenotypic variability for PTD, with PSN behaving as a suppressor trait that magnifies the effect of FSN. Based on a large set of diverse barley accessions, our results provide a deeper phenotypic understanding of the quantitative genetic nature of PTD, its association with traits of high agronomic importance, and a resource for further genetic analyses.

Keywords: final spikelet number, heading date, plant height, potential spikelet number, pre-anthesis degeneration, spikelet primordia.

2.2 Introduction

Barley (*Hordeum vulgare* L.), like wheat, belongs to the Triticeae tribe; and mainly due to its diploid nature and relatively smaller well-studied genome, it is a model organism for genetic

studies. Both barley and wheat produce an unbranched inflorescence known as a spike. The barley spike is indeterminate in nature with determinate spikelets that are produced in a distichous manner with a single floret. The wheat spike, on the other hand, is determinate but with indeterminate spikelets, that is, each spikelet contains multiple florets (Bonnett, 1966; Kirby & Appleyard, 1984; Koppolu & Schnurbusch, 2019). The number of spikelets per spike is one of the major traits that determines the final grain yield (GY) at three developmental phases, namely, vegetative, reproductive, and grain filling. During the early reproductive phase, spikelets start to initiate as axillary meristems on the flanks of the spike from the spikelet ridges (double ridge stage; formation of the lower leaf and upper spikelet ridge on the spike). After passing successive developmental stages, namely, triple mound, glume, lemma and stamen primordium, and awn primordia, barley spikes reach their maximum yield potential (MYP). At the MYP stage, the maximum number of spikelet primordia are produced, and after this stage, no further primordia are initiated on the immature spikes. Consequently, the maximum number of nodes or spikelets at the MYP stage represents the spike's potential to set grains and is, therefore, designated as potential node number or potential spikelet number (PSN). The spikes retain their MYP for a few developmental stages that usually span Waddington stages 5–7 (Waddington et al., 1983; Thirulogachandar & Schnurbusch, 2021). The retaining of the MYP is, hereafter, indicated as the MYP phase. However, not all spikelets produced by the end of the MYP phase survive until the heading stage resulting in a phenomenon called apical pre-anthesis tip degeneration (PTD; Figure. 2.1). The process of pre-anthesis tip degeneration is not only restricted to the apical part; a few (usually 2–3) spikelet nodes at the base of the spike also develop poorly and do not set grains (Appleyard et al., 1982). The basal abortion was attributed to the intervention of leaf development in the spikelet developmental program, and depending upon genetic background, ~30–40% of the maximum number of spikelet primordia abort (Kirby & Faris, 1972; Gallagher & Biscoe, 1978).

The degeneration process starts from the tip of the inflorescence meristem, and proceeds downwards (Figure. 2.1b-e). During spike growth, the degeneration spikelets appear desiccated as they lose their turgidity. It should be noted that spikelet/floral degeneration is not only confined to barley, but also occurs in other cereals, for example, in wheat, where multiple florets degenerate in a spikelet (Kirby, 1988; Bancal, 2009; Ferrante et al., 2013b; Guo et al., 2017; Sakuma et al., 2019): in maize, it is known as ear tip barrenness (Qilin et al., 1999; Meng et al., 2007; Li et al., 2008); and in rice as panicle degeneration (Wang et al., 2018; Ali et al., 2019). Various causes for spikelet/floret degeneration have been hypothesized, the most favored ones being the competition for assimilates or stem-spike competition (Kirby, 1988; Arisnabarreta & Miralles, 2004; González et al., 2005a; Gonzalez et al., 2011; Ferrante et al., 2013b), competition between spikelets (Appleyard et al., 1982), the position of the spikelets within the spike (Arisnabarreta, Sebastián & Miralles, Daniel J, 2006), abortion of distal florets in wheat spikelets as they are not well connected with the main vascular bundles of the rachilla (Wolde & Schnurbusch, 2019), abortion activation by the developmental stage of the most advanced floret in the spikelet (Whingwiri, E & Stern, W, 1982; Bancal, 2008; Bancal, 2009), and hormonal distribution and dynamics within the spike (Cotterell et al., 1981a; Youssef et al., 2017; Boussora et al., 2019). Nevertheless, a conclusive representation of morphological or genetic mechanism(s) operating behind PTD remains elusive to date.

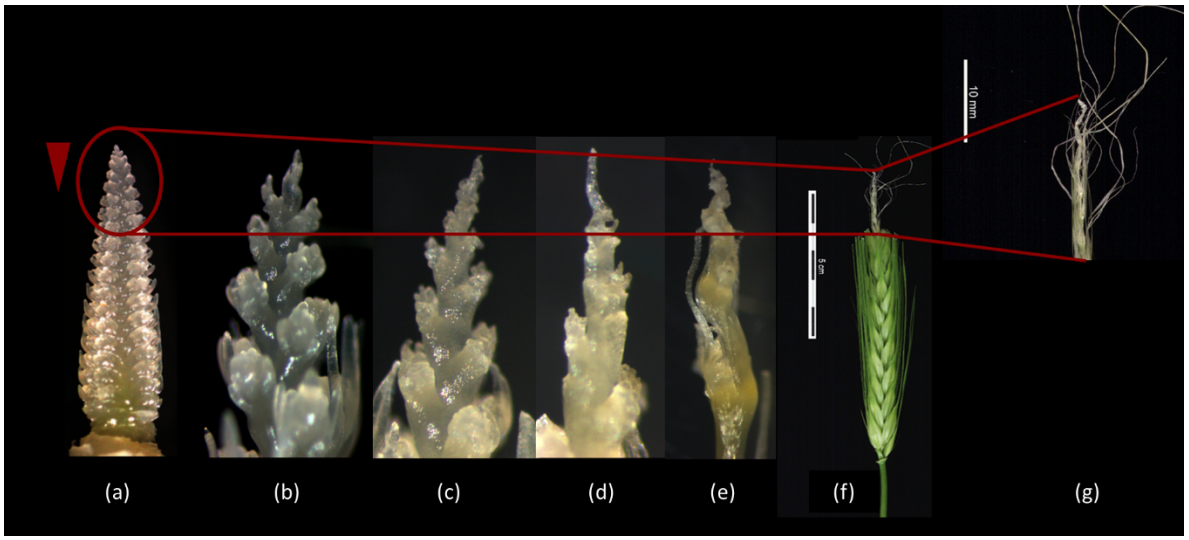


Figure 2.1 Progression of pre-anthesis tip degeneration (PTD) in barley. (a) Immature spike with maximum number of spikelet primordia; that is, potential spikelet number (PSN) during the maximum yield potential (MYP) phase, (b–e) Advancement in PTD from the tip of the immature spike in a downwards direction, (f) Spike after heading; the tip of the spike (g) shows the degenerated/aborted spikelets as a papery structure. The red circle represents the approximate number of apical spikelets that will degenerate after the MYP phase and the triangle shows the direction of PTD.

Only those nodes or spikelets that are retained on the barley spike after the degeneration—termed as final node number (FNN) or final spikelet number (FSN)—reach the grain-filling phase. The grain number resulting from this floral initiation and degeneration process represents only a fraction of the PSN (Kirby, EJM & Appleyard, M, 1987; del Moral et al., 2003; González et al., 2003b). Consequently, the surviving floral organs contribute to the final grain number per unit area that is directly reflected in the GY (Baethgen et al., 1995; Boonchoo et al., 1998; Del Moral et al., 1999; Arisnabarreta & Miralles, 2004). Hence, a reduction in spikelet number per spike from PSN to FSN in barley warrants an in-depth study of PTD. Since PTD is highly laborious to phenotype, previous studies in cereals, at large, were confined to a small number of genotypes which did not represent the breadth of genotypic variation either for PSN or for PTD.

Here, we investigated PTD in a diverse worldwide six-rowed spring barley panel comprising 417 accessions in replicated multi-year field trials. Our analyses revealed a large and significant genotypic variation for PSN, FSN, and PTD that resulted in high broad-sense heritability estimates. We also observed significant correlations of the investigated traits with PTD. We report that PSN shows a direct positive relationship while FSN shows a direct negative relationship with PTD.

2.3 **Material and methods**

2.3.1 Panel selection, preliminary evaluation, and field trials

An initial panel of 325 diverse six-rowed spring barleys was selected from the barley collection of the German federal *ex situ* genebank hosted at the IPK, Gatersleben. As shown in Figure. 2.2a, the panel selection criterion was primarily based on the genotypes covering the maximum

genotypic diversity space of 5038 domesticated six-rowed barley accessions fingerprinted with genotyping-by-sequencing (Milner et al., 2019). Six-rowed barley accessions were selected because six-rowed types occupy a significant market share in Central Europe, and to avoid phenotypic effects associated with the major row-type gene *Six-rowed spike 1* (*Vrs1*) that inhibits lateral spikelet development (Komatsuda et al., 2007; Sakuma et al., 2013; Zwirek et al., 2019). In addition, we chose lines harboring only the *Ppd-H1* sensitive allele; genotypes carrying the *ppd-H1* insensitive allele often flower late, and thus exhibit longer and slower growth that might affect PTD (Turner et al., 2005). Besides, it was reported that six-rowed barleys generally show higher PTD as–due to the fertile lateral spikelets—they possess a higher number of fertile spikelet primordia per spike compared with the two-rowed types (Whingwiri, E & Stern, W, 1982; Kernich et al., 1997; Miralles & Richards, 2000; Garcıa del Moral et al., 2002; Arisnabarreta & Miralles, 2004; Arisnabarreta, Sebastián & Miralles, Daniel J., 2006). Owing to the above-mentioned reasons, six-rowed barley accessions proved to be ideal for studying PTD and related traits.

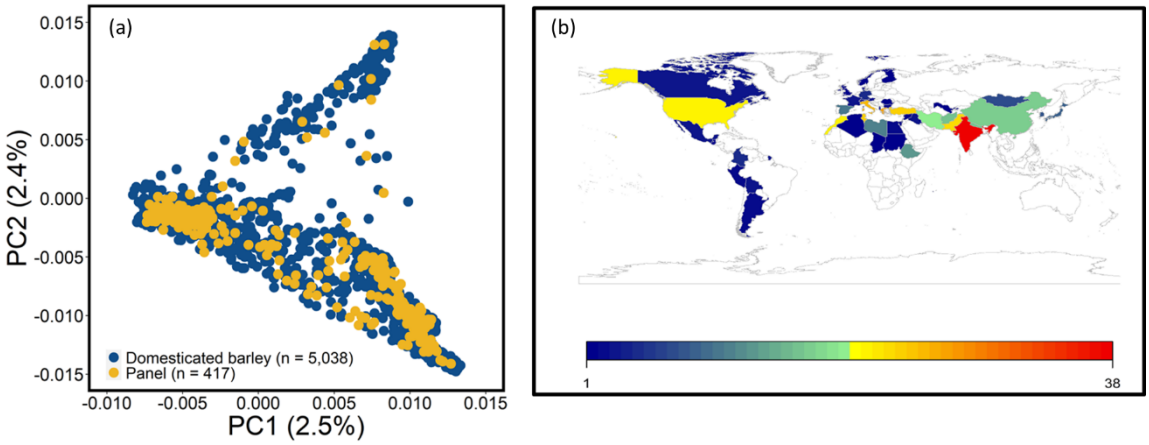


Figure 2.2 Geographical distribution of the six-rowed barley panel. (a) PCA plot highlighting the panel in the background of all the domesticated, six-rowed and spring-type barley accessions present in the German federal *ex-situ* gene bank, the *x*- and *y*-axes represents the first and second principal components namely, PC1 and PC2, respectively. (b) World map featuring all the accessions with respect to their country of origin, 1 represents the least number of accessions in a country, and 38 represents the maximum number of accessions that are from India.

It was reported that genebanks might harbor duplicated or wrongly passported genotypes (Lund et al., 2003). Consistent with this, when grown in the field in 2018, a total of 288 (89%) out of 325 accessions were found to be truly spring types and were, therefore, phenotyped for the investigated traits (see below). The remaining accessions (11%) either did not survive the field conditions or did not reach flowering. To increase the panel size, we added 129 new spring-type barley accessions to the panel, thereby reaching a total of 417 accessions to be evaluated in the following years—2019 and 2020. The complete panel ($n = 417$) represents the worldwide six-rowed spring barley diversity and comprises accessions from five different continents. (Figures. 2.2b and 2.3b). Although the majority of the accessions ($n = 350$; 84%) were landraces, we selected a decent proportion of recent cultivars ($n = 67$; 16%) to observe the trends for the selection of investigated traits (Figure. 2.3a).

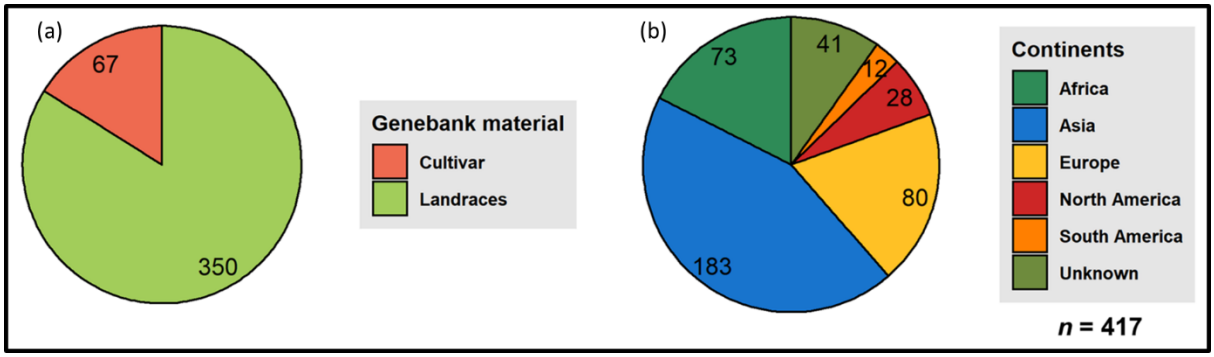


Figure 2.3 Panel distribution based on the Gene material and Continents. (a) Pie-chart highlighting the distribution of the panel with respect to the gene bank material, namely the cultivar and landraces and (b) pie-chart indicating the number of accessions belonging to each continent with minimum and maximum number of accessions from South America (12) and Asia (183), respectively. The keys describe the details for the gene bank material and the continents. '*n*' denotes the total number of accessions used in the panel.

The whole panel was evaluated at the IPK's field facilities (51°49′23″N, 11°17′13″E, 112 m altitude) in replicated trials (three replications per year) for three consecutive years, namely, 2018, 2019, and 2020. In 2018, the accessions were grown in a completely randomized design (CRD), while in the years 2019 and 2020, the panel was grown in an α-lattice design; the latter is known to have higher efficiency to reduce the residual variance and to estimate genetic variance more accurately (Abd El-Mohsen & Abo-Hegazy, 2013; Masood et al., 2018). The individual plot size was ~1.5 m^2, with each plot divided into six rows spaced 0.2 m apart. The sowing density was kept constant across the years with 20 kernels per row. Except for the plant growth regulators (PGRs), we applied standard agronomic practices in all years: PGRs are known to affect the investigated traits. Every year, in each replication, three main culms from the center of the plot were evaluated. The main culm was selected as it is less influenced by environmental perturbations and growth conditions, shows different spikelet primordia initiation and survival rates compared with tillers, and results in higher and consistent final grain number per spike (Cottrell et al., 1985).

2.3.2 *Investigated traits and their phenotyping*

We studied PTD by calculating spikelet survival (%) at two developmental events, namely, the MYP phase—determined based on the 'spikelet stop' method described in (Thirulogachandar & Schnurbusch, 2021)—and, at HD—calculated as the number of days from 1 January until 50% of the spikes emerged out from the flag leaf sheath in a plot at the growth stage of BBCH-59–61 (Zadoks et al., 1974). It is important to mention that, at the heading stage, the spikelets that survive the degeneration process develop and reach the grain-filling phase (Fig. 1F). Since data collection for PTD was difficult as different accessions start degenerating spikelets at different Waddington stages, we choose two developmental events to study the in-between ongoing PTD process. The first was PSN—the result for the spikelet initiation process (collected during the MYP phase), and the second was FSN—the consequence of PTD process (collected at the HD).

2.3.3 Phenotypic data collection via microscopic dissection

The main culms of three individual plants per replication and accession (in total, nine spikes per accession per year) were selected for microscopic analyses every year. The plants were dissected and individual rachis nodes (both differentiated and undifferentiated) were counted on the immature spikes under a stereomicroscope (Stemi 2000-c, Carl Zeiss Micro-Imaging, GmbH, Göttingen, Germany), as shown in Supplementary Figure. S2.1. In barley, each rachis node has the potential to produce three spikelets due to the formation of the triple spikelet meristem (Bonnett, 1966). Therefore, to achieve the maximum spikelet number, the total rachis node number was multiplied by three. In total, ~2600 spikes were individually dissected under the microscope in 2018, whereas ~3800 were dissected each in 2019 and 2020.

2.3.4 Phenotypic data collection on the field

Three main culm plants per replication and accession were selected from the center of the plot at HD stage. The total rachis node number was counted in the field and later multiplied by three to obtain the FSN. The same number of spikes were analyzed as for the microscopic dissection. In addition to PSN, FSN, and HD, the plant height (PH) was measured on the same three plants used for counting FSN in 2019 and 2020 as a distance (cm) from the soil surface to the base of the spike.

2.3.5 Pre-anthesis tip degeneration calculation

The percentage of the spikelets that are degenerated during PTD was calculated as:

$$\text{Pre} - \text{anthesis tip degeneration}(\%) = 100 - \left(\frac{\text{Final spikelet number}}{\text{Potential spikelet number}} \times 100\right) (1)$$

2.3.6 Within-year data analyses

Since we employed different field designs, namely CBD and α-lattice, for phenotypic data collection, different linear mixed-effect models were used to perform within-year phenotypic data analyses. For 2018, to compute individual variance components of the genotypes, replications, and residuals, the following mixed-effect model was used by assuming all effects except the intercept as random:

$$y_{ij} = \mu + g_i + r_j + \varepsilon_{ij} \ (2)$$

where, y_{ij} is the phenotypic record of the i^{th} genotype in j^{th} replication, μ is the common intercept term, g_i is the effect of the i^{th} genotype, r_j is the effect of the j^{th} replication and ε_{ij} denotes the corresponding residual term.

For the α-lattice field design in 2019 and 2020, we used the following model by assuming all effects, except the intercept, as random for individual variance component calculation:

$$y_{ijk} = \mu + g_i + r_j + \beta_{(j|k)} + \varepsilon_{ijk} \ (3)$$

where, y_{ijk} is the phenotypic record of the i^{th} genotype in the j^{th} replication and k^{th} block, μ is the common intercept term, g_i is the effect of i^{th} genotype, r_j is the effect of the j^{th} replication, $\beta_{j|k}$ is the block effect of the k^{th} block nested in the j^{th} replication and ε_{ijk} is the corresponding residual term. Within-year repeatability ($\widehat{H}^2$) was calculated as:

$$\widehat{H}^2 = \frac{\sigma_g^2}{\sigma_g^2 + \left(\frac{\sigma_\varepsilon^2}{n_R}\right)} (4)$$

where, σ_g^2 and σ_ε^2 represent the genotypic and residual variances, respectively; n_R denote the within-year number of replications.

Within-year calculations of the genotypic values (i.e., best linear unbiased estimations; BLUEs) for each investigated trait were performed using the Eqs. 2 and 3, assuming the effects of intercept and genotypes as fixed and all other effects as random.

2.3.7 Across-years data analyses

We calculated the across-years individual variance components of the genotype, genotype-by-year, and the residuals using a linear mixed-effect model by assuming all effects except the intercept as random as:

$$y_{ijkl} = \mu + g_i + y_j + (g \times y)_{(ij)} + (y \times r \times \beta)_{(j|k|l)} + \varepsilon_{ijkl} (5)$$

where, y_{ijkm} is the phenotypic record of the i^{th} genotype in the j^{th} year and k^{th} replication nested in l^{th} block, μ is the common intercept term, g_i is the effect of i^{th} genotype, y_j is the effect of the j^{th} year, $(g \times y)_{(ij)}$ is the genotype-by-year interaction effect of the i^{th} genotype and j^{th} year, $(y \times r \times \beta)_{(j|k|l)}$ is the l^{th} block nested in k^{th} replication in j^{th} year, and ε_{ijkl} is the corresponding residual term. The across-years heritability (H^2) was calculated as:

$$H^2 = \frac{\sigma_g^2}{\sigma_g^2 + \left(\frac{\sigma_{g\times y}^2}{n_y}\right) + \left(\frac{\sigma_\varepsilon^2}{n_y \times n_R}\right)} (6)$$

*w*here σ_g^2, $\sigma_{g\times y}^2$, and σ_ε^2 denote the genotypic, genotype-by-year, and the residual variance, respectively; n_y and n_R represent the average number of years and number of replications, respectively.

Since PTD (%) is a derived trait from PSN and FSN, we used the following model to compute the variance components of genotype and years:

$$y_{ij} = \mu + g_i + y_j + \varepsilon_{ij} (7)$$

where, y_{ij} is the phenotypic record of the i^{th} genotype in the j^{th} year, μ is the common intercept term, g_i is the effect of the i^{th} genotype, y_j is the effect of the j^{th} year and ε_{ij} is the corresponding residual term. Accordingly, the H^2 was calculated based on eq.4, except that n_R is replaced with n_y.

To calculate the across-years BLUEs, the intercept, and the genotypic effects were assumed to be fixed while all other effects were considered random in Eqs. 5 and 7. To check if the within- and across-years BLUEs were normally distributed, we applied the Shapiro-Wilk test at $P < 0.001$.

2.3.8 Correlations and path analysis among the investigated traits

To examine if genetic correlations exist among the investigated traits, we calculated the Pearson's product-moment correlation (r) and corresponding significance (P) values among the BLUEs calculated across-years in Eqs. 5 and 7. Moreover, to check the across-years general performance of a singular given trait, we calculated the average correlation ($\bar{r}$) by performing Fisher's z transformation, as described in Muqaddasi et al. (2020). This transformation is shown to provide a less biased estimate of the average correlation compared to the average correlation estimated based only on the mean of correlation values (Corey et al., 1998).

Path analysis—an extension of the multiple regression—allows comparing different models to determine the best fit hypothesis to explain the relationship between the endogenous (dependent) trait and two or more exogenous (independent) traits (Streiner, 2005). We performed the path analysis by employing three different models to study the relationship among the traits by setting PSN, FSN and PTD as dependent traits. In model 1, HD and PH (independent traits) were assumed to predict the variation in PSN; HD, PH and PSN were assumed to predict the variation in FSN; and PSN, FSN, HD and PH were assumed to predict PTD. In model 2, the independent traits for PSN and FSN remain the same as in model 1, but PTD was only assumed to be predicted by PSN and FSN. In model 3, the independent trait variables for PSN and PTD were kept the same as in model 2, with FSN being affected by PSN and HD.

The goodness of fit indices, viz., χ^2 test, comparative fit index (CFI), Tucker-Lewis index (TLI), and standard root mean square residual (SRMR), were used to analyze the goodness of fit of the model. The model was assumed acceptable if all of the following criteria were fulfilled: χ^2 value was non-significant ($P > 0.05$), CFI value as ≥ 0.90, TLI value as ≥ 0.95 and SRMR value as ≤ 0.08 (Suhr, 2008).

2.3.9 Calculation of direct, indirect, and total effects via best-fit path analysis model

Based on the best-fit indices, model 3 was selected to calculate the direct, indirect, and total effect of each independent trait on the dependent trait as mentioned below (for more details see the Results section; Supplementary Figure. S2.10). Let p be the path from one trait to another and T the trait number. In the first step, PSN was assumed to be affected by HD and PH, and, therefore, the direct (or total) effect for PSN was calculated as:

$$PSN = HD_{[p_{14}T_4]} + PH_{[p_{15}T_5]} (8)$$

In the second step, FSN was assumed to be directly affected by PSN and HD, whereas, HD and PH could also indirectly affect FSN via PSN. The direct effect for FSN was calculated as:

$$FSN = PSN_{[p_{21}T_1]} + HD_{[p_{24}T_4]} (9)$$

whereas, the indirect effects of HD and PH on FSN were calculated as:

$$HD = HD_{[p_{14}T_4]} \times PSN_{[p_{21}T_1]} (10)$$

$$PH = PH_{[p_{15}T_5]} \times PSN_{[p_{21}T_1]} (11)$$

The total effect for FSN was calculated by adding Eqs. 9–11 as;

$$FSN = \left(PSN_{[p_{21}T_1]} + HD_{[p_{24}T_4]}\right) + \left(HD_{[p_{14}T_4]} \times PSN_{[p_{21}T_1]}\right) + \left(PH_{[p_{15}T_5]} \times PSN_{[p_{21}T_1]}\right) (12)$$

In the third step, PSN and FSN directly explain the variation for PTD and also, indirectly via HD and PH. The direct effect of PSN and FSN on PTD was calculated as;

$$PTD = PSN_{[p_{31}T_1]} + FSN_{[p_{32}T_2]} (13)$$

and the indirect effects of HD and PH were calculated as:

$$HD = \left(HD_{[p_{14}T_4]} \times PSN_{[p_{31}T_1]}\right) + \left(HD_{[p_{24}T_4]} \times FSN_{[p_{32}T_2]}\right) (14)$$

$$PH = \left(PH_{[p_{15}T_5]} \times PSN_{[p_{31}T_1]}\right) + \left(PH_{[p_{15}T_5]} \times PSN_{[p_{21}T_1]} \times FSN_{[p_{32}T_2]}\right) (15)$$

The total effect for PTD was calculated by adding equations 13–15 as:

$$PTD = \left[PSN_{[p_{31}T_1]} + FSN_{[p_{32}T_2]}\right] + \left[\left(HD_{[p_{14}T_4]} \times PSN_{[p_{31}T_1]}\right) + \left(HD_{[p_{24}T_4]} \times FSN_{[p_{32}T_2]}\right)\right] + \left[\left(PH_{[p_{15}T_5]} \times PSN_{[p_{31}T_1]}\right) + \left(PH_{[p_{15}T_5]} \times PSN_{[p_{21}T_1]} \times FSN_{[p_{32}T_2]}\right)\right] (16)$$

Unless stated otherwise, all calculations were performed in software R (Team, 2013) by using the packages lme4 (Bates et al., 2015) and lavaan (Rosseel, 2012).

2.4 **Results**

2.4.1 Potential spikelet number bears a large and significant genotypic variation and significant positive correlation with heading date and plant height in six-rowed barleys

The PSN reflects a spike's maximum capacity to produce spikelets and, therefore, is a crucial GY component. We assessed 417 six-rowed spring barleys for PSN in replicated field trials across three years. Within-year variance component analyses attributed a significant proportion of the total variance for PSN to the genetic variance (Supplementary Figures. S2.2a, S2.4a, and S2.6a). For example, the within-year ANOVA for PSN revealed that the genotypic variance was significantly ($P < 0.001$) larger than zero (Supplementary Tables S2.1a, S2.2a, S2.3a and S2.4a), signifying a large genotypic diversity in the investigated panel. Within-year repeatability for PSN was high and ranged from 0.82 to 0.97 (Supplementary Figures. S2.2b, S2.4b, and S2.6b). Similarly, our across-years analyses reflected the within-year analyses showing that the genotypic variance was the principal contributor for PSN variation ($\sigma_g^2 = 72.68\%$) whereas only a small proportion of the total variance was attributed to residual components ($\sigma_\varepsilon^2 = 14.79\%$) (Figure. 2.4a). The large genotypic variance resulted in high broad-sense heritability ($H^2 = 0.93$) estimates for PSN, indicating a high-quality phenotypic data-set (Figure. 2.4b).

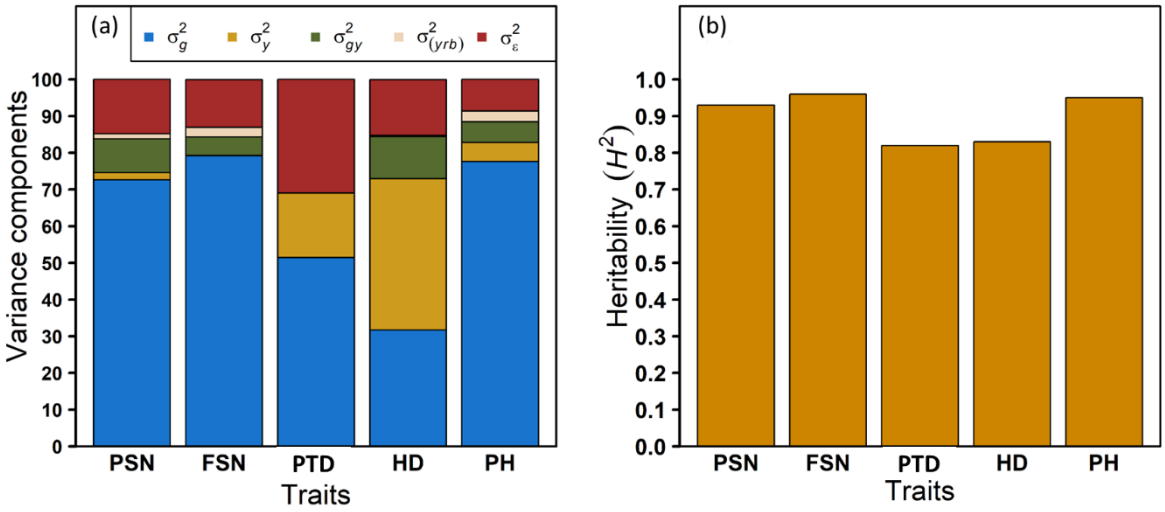


Figure 2.4 Variance component and heritability analyses. (a) Proportion of the different variance components for each investigated trait where σ_g^2 is the genotypic variance, σ_y^2 is the year variance, σ_{gy}^2 is the (genotype × year) interaction variance, $\sigma_{(yrb)}^2$ is the year and replication variance with replication nested into the blocks and σ_ε^2 is the error or residual variance. (b) Broad-sense heritability (H^2) for each trait. PSN, FSN, PTD, HD and PH represent potential spikelet number, final spikelet number, pre-anthesis tip degeneration (%), heading date (days from 1 January) and plant height (cm), respectively.

The BLUEs calculated in individual years for PSN exhibited a wide variation and approximated normal distribution (Shapiro-Wilk $P < 0.001$). For example, in 2018, PSN ranged from 36 to 130 with a mean of 89.9 (Supplementary Figure. S2.5a); in 2019, from 55 to 135 with a mean of 98.0 (Supplementary Figure. S2.6a), and in 2020, from 71 to 9135 with a mean of 100.6 (Supplementary Figure. S2.7a). The across-years BLUEs for PSN mirrored the within-year normal distributions (Shapiro-Wilk $P < 0.001$) and showed broad variation ranging from 70 to 132 with a mean of 99.3 (Figure. 2.5a). In addition, the performance of PSN analyzed via Fisher's z transformation showed a high average correlation ($\overline{r} = 0.83$) across three years (Figure. 2.6a). This indicates that the phenotypic data can be reliably used for further studies to make meaningful conclusions.

Besides PSN, we analyzed heading date (HD) and plant height (PH). Based on the hypothesis that HD and PH affect PSN, we performed Pearson's product-moment correlation (r) on within- and across-years BLUEs. PSN was significantly ($P < 0.001$) and positively correlated with both HD and PH. The correlation coefficient between PSN and HD was 0.53, 0.54, and 0.64 in 2018, 2019, and 2020, respectively (Supplementary Figures. S2.5f, S2.6f, S2.7f); the across-years correlation coefficient amounted to 0.65 (Figure. 2.5f). The association between PSN and PH remained consistent, with a correlation coefficient of 0.49. These analyses suggest that, on average, the taller and later heading plants show higher PSN.

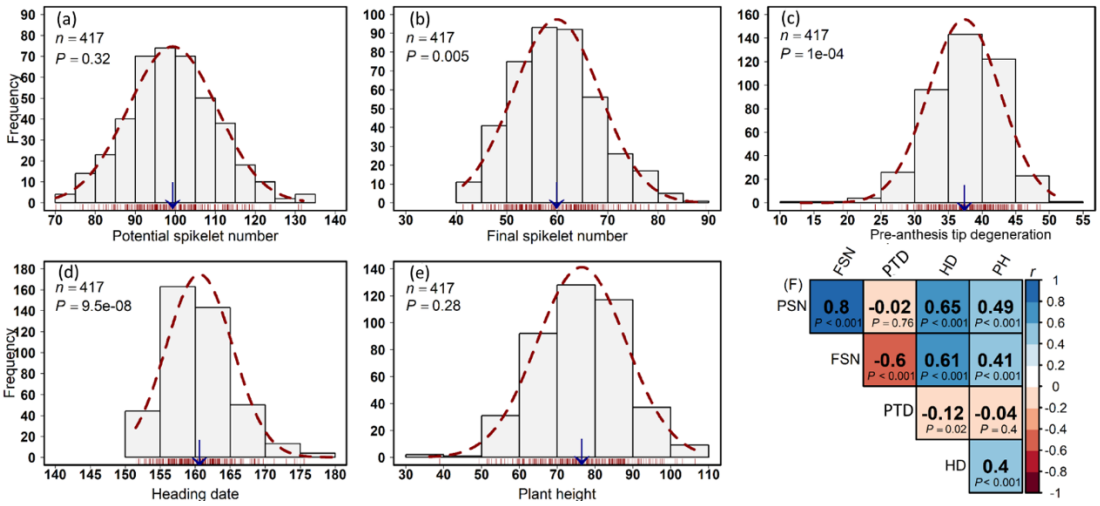


Figure 2.5 Phenotypic distribution and correlation of the investigated traits in a panel of 417 six-rowed spring barley accessions. Frequency distribution of (a) potential spikelet number (PSN), (b) final spikelet number (FSN), (c) pre-anthesis tip degeneration (PTD in %), (d) heading date (HD as days from 1 January) and (e) plant height (PH in cm). The *x*- and *y*-axes of each histogram denotes the individual trait and number of accessions (frequency), respectively. '*n*' is the number of accessions and '*P*' represents the Shapiro-Wilk's test result. The downward blue arrow indicates the mean value for the respective trait and the dashed curve shows the fit of the normal distribution curve. (f) Pearson's product-moment correlation (*r*) among the traits. *P*-value in the plot (f) denotes the significance level of the respective correlation.

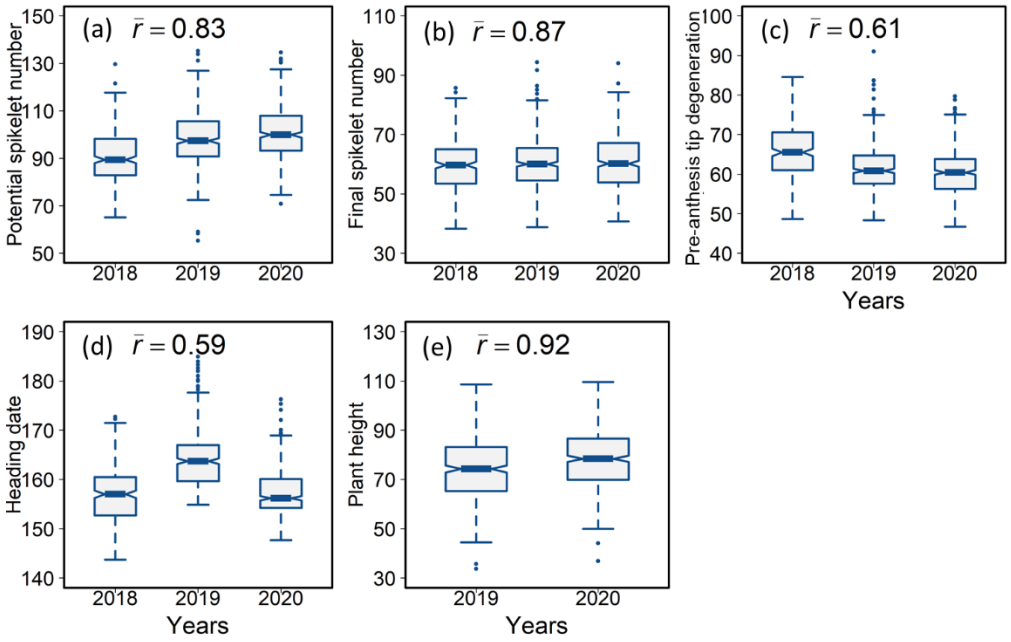


Figure 2.6 Environment (year) specific phenotypic distribution of the investigated traits in a panel of 417 six-rowed spring barley accessions with between-years average trait correlation ($\bar{r}$) calculated by performing the Fisher's *z* transformation. (a) Potential spikelet number (PSN), (b) final spikelet number (FSN), (c) pre-anthesis tip degeneration (PTD in %), (d)

heading date (HD as days from 1 January) and (e) plant height (PH in cm). The *x*- and *y*-axes of each plot indicate the years and the particular studied trait, respectively.

2.4.2 Final spikelet number is highly heritable and shows a significant positive correlation with potential spikelet number, heading date, and plant height

We studied FSN—the spikelets borne on a spike at HD after PTD has taken place. Our within-year ANOVA showed a large and significant (P<0.001) genotypic variance for FSN (Supplementary Tables S2.1b, S2.2b, S2.3b) and the proportion of genotypic variation in the years 2018, 2019, and 2020 amounted to 64, 78, and 90%, respectively. It should be noted that the residual variation of FSN was highest in the year 2018 (σ^2_ε=27.47%) compared with 2019 (σ^2_ε=17.17%) and 2020 (σ^2_ε=8.18%) (Supplementary Figures S2.2a, S2.3a, S2.4a) the most likely reasons for this could be a relatively smaller sample size (n=288) and higher environmental perturbations: the highest temperature and global solar radiation plus the lowest humidity were recorded in 2018 (Supplementary Figure. S2.8). Nevertheless, in the across-years ANOVA, we observed that the proportion of the genotypic variance was large (79%, Figure. 2.4a) and significantly (P<0.001) greater than zero (Supplementary Table S2.4b). The across-years broad-sense heritability for FSN amounted to 0.96. (Figure. 2.4b).

The BLUEs calculated for FSN within each year followed a statistically normal distribution and showed a wide variation. For instance, in the year 2018, FSN ranged from 38 to 86 with a mean of 59.3; in 2019, from 39 to 94 with a mean of 60.4; and in 2020, from 41 to 94 FSN with a mean of 60.8 (Supplementary Figures S2.5b, S2.6b, S2.7b). Our across-years BLUEs for FSN exhibited a similar pattern, with a statistically normal distribution, and showed a wide variation that ranged from 41 to 90 with a mean of 59.8 FSN per spike (Figure. 2.5b). As observed for PSN, Fischer's z transformation revealed a high average correlation ($\bar{r} = 0.87$)for FSN across three years (Figure. 2.6b). The high values for both broad-sense heritability and average correlation indicate the suitability of the phenotypic data to draw reliable conclusions.

We performed Pearson's product–moment correlation analysis among the across-year genotypic values (BLUEs) of PSN and FSN per spike plus HD and PH, which revealed that FSN was significantly and positively correlated with PSN (r = 0.80, P<0.001; Figure. 2.5f). This suggests that the more spikelet primordia are produced on an immature spike (i.e. PSN), the more the plant will retain after PTD (i.e. FSN). Within each year, FSN showed a strong association with PSN ($r_{[2018]}$=0.71; $r_{[2019]}$=0.75; $r_{[2020]}$=0.80; Supplementary Figures S2.5e, S2.6f, S2.7f). Here, it is important to note that with every year, the correlation between FSN and PSN improved; this may be attributed to the use of different field designs (2018, CBD; 2019 and 2020, α-lattice) and the improved accuracy for MYP sample collection and growth conditions. Significant differences among years concerning the average daily temperature, humidity, and global solar radiations during the three cropping seasons were observed (Supplementary Tables S2.5, S2.6, S2.7; Supplementary Figure. S2.8). There was no difference with respect to the soil type as the soil composition was homogeneous at the IPK fields with a clayey loam texture. In addition to PSN, FSN was significantly and positively correlated with both HD (r=0.61, P<0.001) and PH (r=0.41, P<0.001) (Figure. 2.5f). In 2020, we observed the

highest association of FSN with both HD (r=0.44, P<0.001) and PH (r=0.44, P<0.001) (Supplementary Figure. S2.7f).

2.4.3 *PTD's genotypic diversity highlights the potential for high selection gain and is negatively associated with final spikelet number*

We calculated PTD (%) using the BLUEs of PSN and FSN, as described in Equation 1. Since within-year PTD is a derived trait, the variance component and repeatability calculations were not possible. Nevertheless, the across-years ANOVA for PTD2. showed that genotype and year variances were significantly (P<0.001) larger than zero (Supplementary Table S2.4c). As previously observed for PSN and FSN, the variation in PTD was mainly due to genetic factors (σ^2_g=51.52%); however, environments also influenced PTD significantly (σ^2_ε=30.94%; P<0.001) (Figure. 2.4a). Within-year variation for PTD ranged from 2% to 57% with a mean of 34% (Supplementary Figure. S2.5c) in 2018, from 9% to ~52% with a mean of 38.3% in 2019 (Supplementary Figure. S2.6c), and from ~20% to 53% with a mean of 39.5% in 2020 (Supplementary Figure. S2.7c). In across-years, PTD varied from 13% to 51%, with a mean value of 37.3% (Figure. 2.5c). Nonetheless, the large and significant genotypic variance for PTD translated into high broad-sense heritability estimates that amounted to 0.82 (Figure. 2.4b). Furthermore, Fischer's z transformation revealed a high average correlation ($\overline{r} = 0.61$) for PTD (Figure. 2.6c). Coupled with large genotypic variance, a high broad-sense heritability estimate for PTD suggests that a high selection response could be expected in breeding programs.

To study the relationship between PTD and all other measured traits, we performed Pearson's product-moment correlation. PTD—except in 2018 ($r = -0.22$; Supplementary Figure. S2.5e)—showed an insignificant correlation with PSN (Figure. 2.5f, Supplementary Figures S2.6f, and S2.7f). Notably, the negative association observed in 2018 explained only ~4% of the total variation in PTD that could be attributed to PSN. In 2018, 2019, and across-years analysis, PTD showed a non-significant correlation with HD (Figure. 2.5f; Supplementary Figures S2.5e, S2.6f) and a weak negative but significant correlation (r= –0.25; P<0.001); Supplementary Figure. S2.7f) in 2020. Even though the correlation in 2020 was significant, it explained only ~6.3% of the total variation in PTD which is due to HD. With PH, PTD showed non-significant correlations within and across years, suggesting that PH does not directly influence PTD. An exception, however, was the trait FSN, which showed a significant negative correlation with PTD in all three years and their corresponding BLUEs (Figure. 2.5f; Supplementary Figures S2.5e, S2.6f, S2.7f). The across-year correlation coefficient between FSN and PTD was r= –0.60 (P<0.001) which explained ~36% of the total variation present for PTD. This shows that the higher the FSN, lower will be the PTD, and vice versa.

Since our investigated panel was comprised of both historic and recent landraces and cultivars, we explored the phenotypic differences of PTD associated therewith. We observed that the significant differences occurred only for FSN and PTD, while PSN, HD, and PH were not changed. As expected, compared with the landraces, the recent cultivars showed a significantly higher FSN, most probably due to lower (–5%) PTD (Supplementary Figure. S2.9).

2.4.4 *Wide variation for heading date and plant height in six-rowed barley accessions*

We calculated HD as the number of days from 1 January until 50% of the spikes were visible in a plot and employed the same three main culms used for FSN data collection to measure PH. The within-year variance components revealed that most of the variance for HD and PH was due to the genetic variance (Supplementary Figures S2.2a, S2.3a, S2.4a) that was reflected in the high repeatability values for both the traits (Supplementary Figures S2.2b, S2.3b, S2.4b). Also, the broad-sense heritability estimates were found to be high (H^2_{HD}= 0.83 and H^2_{PH}= 0.95; Figure. 2.4b).The across-years BLUEs also showed a wide variation ranging from 152 d to180 d with a mean value of 160.6 d (Figure. 2.5d), and this trend was also observed in the individual year data analysis (Supplementary Figures S2.5d, S2.6d, S2.7d). A high average correlation was observed between the BLUEs of each year, and however, this average correlation was lowest among all the traits (Figure. 2.6d). The within- and across-year BLUEs for PH approximated normal distribution (P<0.001) with a wide variation. In 2019, PH ranged from 34 cm to 109 cm with an average of 74 cm (Supplementary Figure. S2.6e), while in 2020, PH varied from 37 cm to 110 cm with a mean of 78 cm (Supplementary Figure. S7e). The across-year BLUEs for PH showed variation from 36 cm to 108 cm with an average of 76 cm (Figure. 2.5e). The average correlation for PH was the highest ($\bar{r} = 0.92$) among all the investigated traits (Figure. 2.6e). PH had a significant correlation with PSN and FSN as explained previously, and it also showed a low to moderate significant correlation with HD, with coefficients of 0.41 in 2019, 0.33 in 2020, and 0.40 in across-year analyses (Figure. 2.5f; Supplementary Figures S2.6f, S2.7f), but a non-significant correlation with PTD under all the scenarios.

2.4.5 *The geographical origin of six-rowed barley accessions revealed distinct patterns for the investigated traits*

To observe whether there exist any significant (P<0.001) differences among the spring barleys collected from different geographical origins, we investigated the accessions from all continents except Australia and Antarctica. The accessions (n=41) with unknown geographic origin were not used to interpret the results.

For PSN, the accessions from Africa showed significantly lower values and the European accessions, on the other hand, showed significantly higher PSN. However, the Asian, North-, and South-American accessions did not significantly differ (Figure. 2.7a). The FSN followed a similar trend to PSN—the African accessions showed significantly lower FSN values. Also, consistent with the correlation analyses (Figure. 2.5f), the African accessions that generally harbored lower PSN showed significantly lower FSN (Figure. 2.7b). Interestingly for PTD, the European accessions showed significantly higher values even though these accessions did not differ in their FSN, whereas the values from North- and South-American accessions did not differ significantly. (Figure. 2.7c). Despite being clear from Figure. 2.7c, African accessions did not show a significantly higher PTD. The African accessions had lower PSN and FSN; therefore, they abort fewer spikelets than the European accessions which produce more PSN and FSN.

The HD trend was similar to that of PSN; that is, the accessions belonging to the African and European continents showed significant differences: the African accessions took significantly fewer days to head whereas the European accessions took longer (Figure. 2.7d). Accessions from all other continents did not show any significant differences for heading days. For PH, the accessions belonging to the African, Asian, European, and South-American continents had significant differences, with European accessions showing the largest significant difference for PH followed by the Asian, African, and South-American accessions (Figure. 2.7e). The observed positive trend between HD and PSN was expected because the European accessions were generally later heading, thus explaining the higher number of potential spikelet primordia during the spikelet initiation and growth phases. The African accessions, in contrast, headed earlier: this means that they had less time in the spikelet initiation phase that was reflected in the lower PSN. Overall, the African accessions had shorter plant stature, earlier heading, and lower PSN and FSN. On the other hand, the European accessions, on average, were taller, later flowering, harbored more PSN, and consequently more PTD. The Asian accessions—even though there were no significant differences for PSN and FSN—showed the lowest PTD, suggesting that they may carry superior allelic combinations for high spikelet fertility. Since one crucial objective of any cereal breeding program is to increase spikelet fertility, the Asian accessions, in this regard, may possess favorable alleles that may be further exploited in breeding programs.

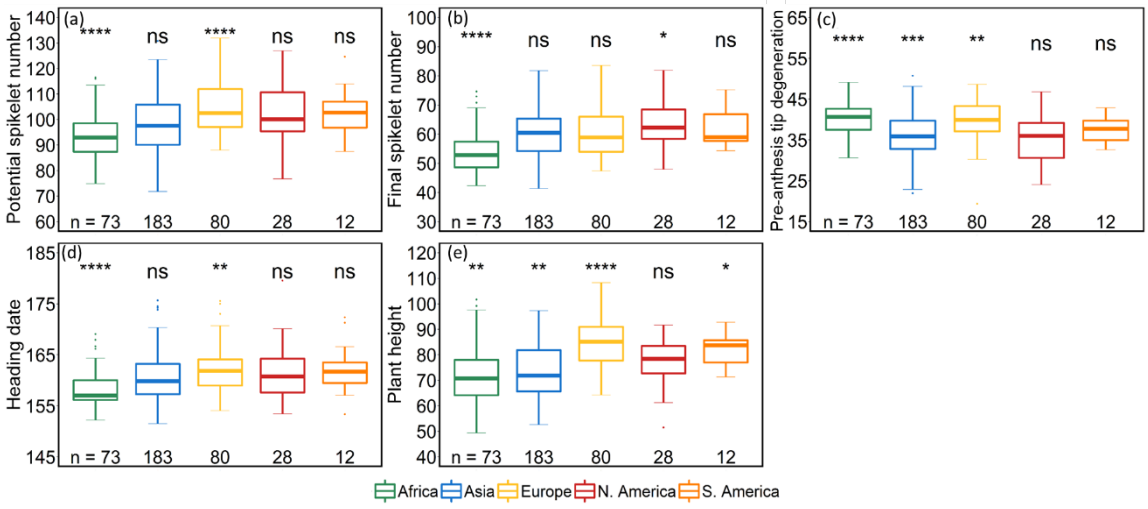


Figure 2.7 Comparison of accessions according to their geographical origin. (a) Potential spikelet number (PSN), (b) final spikelet number (FSN), (c) pre-anthesis tip degeneration (PTD in %), (d) heading date (HD as days from January 1st), and (e) plant height (PH in cm). '*n*' denotes the number of accessions in this panel and the number of the accessions belonging to each continent; N. America and S. America stands for North America and South America; *P*-values are significant at ****0.001, ***0.001, **0.01, and *0.5, respectively and 'ns' represents non-significant differences.

2.4.6 Path analysis highlights the direct and indirect relationships among the investigated traits

We performed the path analysis to study the directional and non-directional relationships between the independent and dependent traits. Here, the main objectives were to understand the correlation or covariance between the independent and dependent traits and explain the maximum variance of a dependent trait with the specified independent traits using itemized models.

Based on the prior understanding of the traits to explain the relationship among the investigated traits, we specified three different models (Supplementary Figure. S2.10). Model 1 exploited all the possible relationships among the investigated traits where PSN was assumed to be affected by HD and PH; FSN by PSN, HD, and PH; and PTD by PSN, FSN, HD, and PH. However, model 1 was 'just identified' (i.e. the degree of freedom was zero) and, even though the estimates of each independent trait on the dependent trait were calculated, they could not be taken into consideration as the corresponding fit indices were not computable (Table 2.1). In model 2, PSN was assumed to be affected by HD and PH; FSN by PSN, HD, and PH; and PTD by PSN and FSN. The $\chi 2$ test's *P*-value was significant ($P<0.001$) rendering model 2 unqualified. However, the $\chi 2$ test may not be very useful to check the goodness of fit of any specified model because of its sensitivity to the sample size; that is, the larger the sample size, the greater are the chances of obtaining a statistically significant $\chi 2$ value. Since path analysis is usually carried out with large sample size, the $\chi 2$ test is usually guaranteed to be significant and, therefore, does not provide useful information regarding the model fit. The CFI, TLI, RMSEA, and SRMR values for model 2 were, nevertheless, 0.99, 0.97, 0.07, and 0.01, respectively (Table 2.1). These indices were within the acceptable limits; however, the estimate of PH for FSN was non-significant. Model 3 was studied by keeping all the effects same as in model 2 except the effect of PH on FSN. The best fit indices for the model 3 were all within the acceptable range, namely CFI (0.99), TLI (0.98), RMSEA (0.09), and SRMR (0.01), with none found to be non-significant (Table 1). Therefore, model 3 was used to make a path diagram highlighting the relationship between the traits.

In the path diagram (Figure. 2.8), PH and HD were assumed to be the independent traits and, hence, straight arrowheads emerged from them. On the other hand, PSN and FSN behaved both as independent and dependent traits; thus, at least one straight arrow points towards them (PSN being the independent trait for FSN and PTD while being the dependent trait for PH and HD; whereas FSN is the independent trait for PTD and the dependent trait for PSN and HD). Finally, PTD behaved as a dependent trait for both PSN and FSN. Also, as described in the Materials and methods, the phenotypic data to study PTD were collected at the end of two developmental events, namely during the MYP phase (PSN) and at heading (FSN); unidirectional arrows go from PSN and FSN to PTD.

Here, we observed a direct significant positive effect of PSN on FSN (0.53) and PTD (0.60), suggesting that each unit increase in PSN resulted in a 0.53 and 0.60 unit increase in FSN and PTD, respectively. The indirect effect of PSN on PTD via FSN was calculated as –0.52 [$0.53 \times (-0.99)$] while the total effect of PSN on PTD was –0.08, [$0.6 + (-0.52)$]. FSN only directly affects PTD with a path coefficient of –0.99, indicating a rather substantial impact of FSN on PTD. A discrepancy was observed for PSN and PTD relationship in correlation and

path analyses. In correlation analysis, the association between PSN and PTD was non-significant, while it became significant in the path analysis. This is because PSN behaves as a suppressor trait in the path analysis; that is, PSN *per se* does not explain the variation present for PTD (Supplementary Table S2.8) but, when used in conjugation with FSN, it enhances the effect of FSN on PTD (Table 2.1; Supplementary Table S2.9). The R^2 values indicated that PSN and FSN together explained 93% of the variability present for PTD.

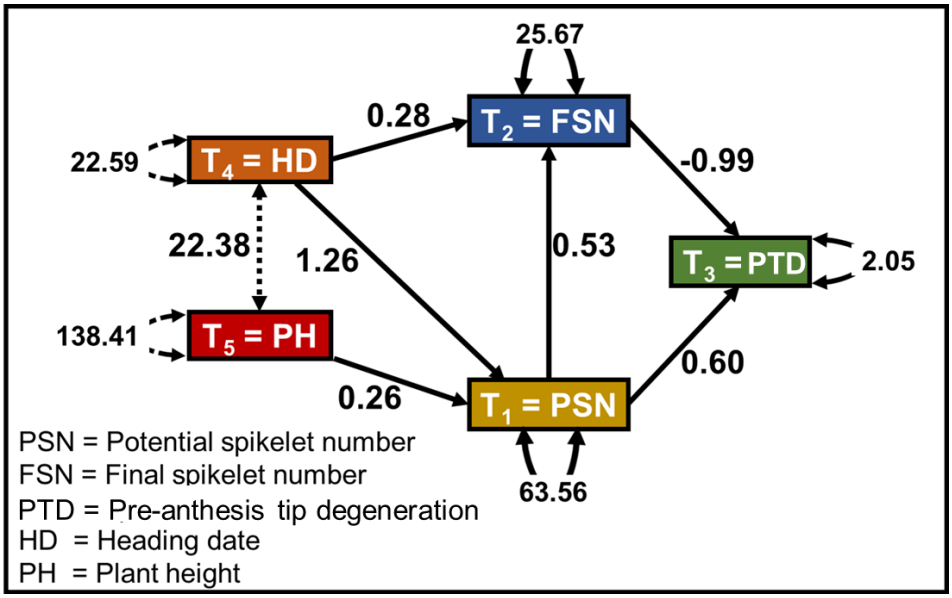


Figure 2.8 Path diagram elucidating the relationship between the traits. The single-headed arrow denotes that one trait directly affects another, and the dashed double-headed arrow implies a covariance between heading date (HD) and plant height (PH). The circular curved arrows represent the variance of a trait and the dashed circular curved arrows represent the variance of a trait that was not specified in the model. T_1–T_5 denotes the trait number from 1 to 5. The number on the unidirectional arrows represents the path coefficients and positive and negative values represent the direct positive and negative effect of one trait on another.

A strong direct positive effect of HD on PSN was observed compared with FSN, with path coefficients of 1.26 and 0.28, respectively: this indicates a decisive influence of HD on PSN. The indirect effect of HD on FSN via PSN was observed to be 0.67 as [1.26 × (0.53)], while the total effect of HD on FSN was 0.95 as [0.67 + (0.28)], which was still lower than the direct effect of HD on PSN (i.e. 1.26). We observed no direct effect, but an indirect effect of HD on PTD via PSN and FSN as 0.76 and –0.27, respectively; the total indirect effect was 0.49. In both within- and across-years correlation analyses, we observed non-significant results between PH and PTD. However, the indirect effect of PH on PTD was calculated. The indirect effect of PH on PTD via PSN was estimated to be 0.16. It should be noted that we did not calculate the effect of PH on PTD via FSN because of the non-significant effect of PH on FSN. Nonetheless, HD and PH explained 49% of the variability in PSN whereas PSN and HD explained 65% of the variability for FSN. These results, on the one hand, suggest the quantitative genetic architecture of the investigated traits because not all the variance could be explained by the explanatory traits. On the other hand—although we used large amounts of high-quality phenotypic data—they point to the existence of other non-investigated traits affecting PSN and FSN.

Table 2.1 Path analysis summary of the three investigated models.

	Model 1	Model 2	Model 3
Number of observations	417	417	417
Model test user model			
Test statistics	0.000	13.13	13.25
Degree of freedom	0.000	2	3
P- value (chi-square)	–	0.001	0.004
Model test baseline model			
Test statistics	1823.99	1823.99	1823.99
Degree of freedom	9	9	9
P- value (chi-square)	0.00	0.00	0.00
User model versus baseline model			
Comparative Fit Index (CFI)	1.00	0.99	0.99
Tucker-Lewis Index (TLI)	1.00	0.97	0.98
Root Mean Square Error of Approximation (RMSEA)	0.00	0.12	0.09
Standardised Root Mean Square Residual (SRMR)	0.00	0.009	0.01

Regression	Model 1				Model 2				Model 3			
	Estimate	Std. error	*z*-value	P (> |z|)	Estimate	Std. error	*z*-value	P (> |z|)	Estimate	Std. error	*z*-value	P (> |z|)
PSN ~												
HD	1.26	0.09	14.07	0.00	1.26	0.09	14.07	0.00	1.26	0.09	14.07	0.00
PH	0.26	0.04	7.28	0.00	0.26	0.04	7.28	0.00	0.26	0.04	7.28	0.00
FSN ~												
PSN	0.53	0.03	16.99	0.00	0.53	0.03	16.99	0.00	0.53	0.029	18.15	0.00
HD	0.28	0.07	4.01	0.00	0.28	0.07	4.01	0.00	0.28	0.07	4.08	0.00
PH	0.01	0.02	0.34	0.73	0.01	0.02	0.34	0.73	–	–	–	–
PTD ~												
PSN	0.59	0.01	52.14	0.00	0.59	0.01	57.65	0.00	0.59	0.01	57.65	0.00
FSN	-1.00	0.01	-73.37	0.00	-0.99	0.01	-73.32	0.00	-0.99	0.0	-73.32	0.00
HD	-0.07	0.02	-3.65	0.00	–	–	–	–	–	–	–	–
PH	0.00	0.01	0.49	0.63	–	–	–	–	–	–	–	–
R-square												
PSN	0.49				0.49				0.49			
FSN	0.65				0.65				0.65			
PTD	0.93				0.93				0.93			

Std. error, PSN, FSN, PTD, HD and PH denotes standard error, potential spikelet number, final spikelet number, pre-anthesis tip degeneration (%), heading date (days from January 1st) and plant height, respectively

2.5 Discussion

2.5.1 Path analysis revealed concealed impact of potential spikelet number on pre-anthesis tip degeneration

The significance of the early reproductive phase as a determinant of GY was studied in both wheat and barley previously (Appleyard et al., 1982; Kitchen & Rasmusson, 1983). The early reproductive phase was shown to be critical for final GY as it governs the maximum number of spikelet primordia that could potentially develop into grains. A study conducted in two-rowed barley proposed the importance of Waddington stage (W) 3.5 (i.e., stamen primordia stage—a part of an early reproductive phase) and concluded that the spikelet primordia that were initiated before the Waddington stage 3.5 corresponded to the final grain number per spike (Digel et al., 2015a). In contrast, the second school of thought focused on the late reproductive phase of stem elongation for PTD and described it as the most critical phase for GY determination (Miralles & Richards, 2000; González et al., 2003b; Ghiglione et al., 2008; Gonzalez et al., 2011; Alqudah & Schnurbusch, 2014; Guo, Zifeng & Schnurbusch, Thorsten, 2015). It was concluded that the late reproductive phase is marked by increased competition between the spike and the stem, and a limited quantity of assimilates allocated to the spikes induced PTD. Recently, Thirulogachandar et al. (2020), based on a panel of 27 genotypes, studied the GY components in both two- (n = 17) and six-rowed (n = 10) barley and concluded that, in two-rowed barley, the grain number was mainly determined by PSN (influenced by early reproductive phase), whereas, in six-rowed barley, it mainly follows the survival ability of the spikelets (influenced by the late reproductive phase).

In the present study, with an aim to explain the conundrum between two above-mentioned viewpoints, namely, the influence of early or late reproductive phases on the final GY, we evaluated a large and diverse set of six-rowed barleys in multi-year field trials. The correlation analysis—with an exception of 2018, where a mild correlation was observed—in both within- and across-year analyses revealed an insignificant relationship between PSN and PTD. However, the relationship between FSN and PTD was high and statistically significant in all the scenarios. Based exclusively on our correlation results, it becomes evident that variation in PSN does not influence the variation for PTD. However, our path analysis revealed a significant positive relationship between PSN and PTD that contradicts with the previously mentioned correlation results that FSN (governed by late reproductive phase) exclusively explains the variation for PTD.

The observed discrepancy between the correlation and path analysis results could be attributed to one of the traits being a suppressor variable in the path analysis (Thompson & Levine, 1997). A suppressor variable has three inherent properties: (i) it is uncorrelated with the dependent variable, (ii) by in and of itself, it is a non-significant predictor of the dependent variable, and (iii) it enhances the predictive ability of the other independent variables. In the path analysis, when PSN was used as the sole predictor for variation in PTD, their relationship was insignificant (Supplementary Table S2.8; $P > 0.05$) and, alone, it did not explain any variation for PTD ($R^2 = 0$). Furthermore, when FSN was used as the sole predictor for PTD variability, it only explained 35% of the variation present for PTD (Supplementary Table S2.9). However, when both PSN and FSN were used as predictors for PTD, the R^2equaled 0.93 (Table

2.1); that is, the effects of both traits were significant and these collectively explained 93% of the variation in PTD. Therefore, based on our large dataset and in line with the theoretical properties, it can be safely concluded that PSN behaves as a suppressor variable with being apparently uncorrelated with PTD (Figure. 2.5f, Supplementary Figures S2.6f and S2.7f), and, by itself not explaining the variation for PTD (Supplementary Table S2.8). Nevertheless, it enhances the predictive ability of FSN from 0.37 to 0.99 (neglecting the negative sign), that is, improving the R^2 from 35% to 93% (Table 2.1 and Supplementary Figure S2.9).

Hence, our analyses prove that in six-rowed barleys both the early and late reproductive phases influence PTD with PSN (governed during early reproductive phase) being a suppressor variable. In addition, our exploratory analyses suggest that on one hand it is paramount to use large and diverse data-sets to make meaningful conclusions, and, on the other hand, investigating the data from different logical standpoints and with diverse analyses helps to uncover hidden relationships among the variables/traits.

2.5.2 *The geographical origin affects pre-anthesis tip degeneration in six-rowed barleys*

In the past, PTD was reported to vary between 30–40% (Kirby & Faris, 1972). Kitchen and Rasmusson (1983), while studying 16 barley genotypes, reported ~60% of PTD. In our study, based on 417 diverse six-rowed barely accessions collected from across the globe, we also observed a large genotypic variation for PTD that ranged from 14 to 51%. This highlights that barley spikelet initiation and degeneration dynamics may lead to a decrease in the yield of up to ~50%. Since our panel represented the worldwide genetic diversity, the impact of geographies on PTD was comprehensively studied. The geographical impact revealed that, from all the continents, Asian accessions exhibit significantly ($P < 0.05$) lower PTD. In other words, the highest spikelet survival was seen in Asian accessions.

Most of the barley grown in northern and central Europe carry an insensitive photoperiod allele (*ppd-H1*) that confers delayed flowering under long day conditions; whereas, most barley from southwest Asia and the Mediterranean basin carries a sensitive photoperiod allele (*PPD-H1*) that induces early flowering under long day conditions in spring (Turner et al., 2005; Jones et al., 2008). The sensitive allele is favored in the Mediterranean region to escape the terminal drought conditions and the insensitive allele is favored in Europe to help plants increase their yield by increasing the days to heading. However, the long growth habitats of barley in Northern Europe might change towards Mediterranean conditions as a consequence of climate change, and, therefore, the environmental advantages of insensitive *ppd-H1* allele might thus dissipate in some European regions (Herzig et al., 2018). This, in turn, extends the possibility to explore the yield potential of *PPD-H1* sensitive barley under European conditions.

As we observed high across-years broad-sense heritability estimates (0.80) for PTD, and given a decent number and large genotypic variance among the accessions collected from each continent, it can be safely extrapolated that high broad-sense heritability estimates exist in separate populations. As stated elsewhere, large genotypic variance and high broad-sense heritability estimates promise high selection gain that can be pursued in six-rowed barley breeding programs. In addition, a wide genotypic variance points to a strong genetic basis of

PTD that can be exploited in future mapping and predictive breeding activities with high accuracy.

2.5.3 Effect of within-year environmental fluctuations on the variance and correlation analyses

We observed the lowest genetic variance, repeatability, and correlation among the traits in 2018 (Supplementary Figures S2.2, S2.3e). This could be attributed to the harsh environmental conditions in 2018 which resulted in a significant effect of these perturbations on the behavior of the investigated accessions. We observed significant differences in the weather conditions between three growth seasons, i.e., 2018 was an exceptionally dry year with the highest record-breaking temperature (Zscheischler & Fischer, 2020) and lowest humidity compared to 2019 and 2020 (Supplementary Figure. S2.9). We thus hypothesized that there may exist an effect of ambient environmental stimuli on the life cycle of the plants. To capture this effect, we divided the whole panel into two sub-groups, group-1 harbored accessions that head early or in a short time; whereas, group-2 contained accessions that take longer to head.

Here, we specifically analyzed the relationship between FSN and HD, as FSN counted at HD was at a stage relatively closer towards the end of the plant's life cycle. In 2018, the correlation coefficient between FSN and HD ($r = 0.47$) was higher in the group-1 accessions than in group-2 ($r = 0.36$; Supplementary Figure. S11a-b). In 2019 and 2020, since the environmental perturbations were lower, not much difference between the correlation coefficients of FSN and HD in the two sub-groups (Supplementary Figure S11c-f) was observed. The lower association in 2018 clearly shows the strong environmental effect during the growth cycle of the accessions. From these results, it can be deduced that the improved associations between FSN and HD in group-1 accessions resulted because they could—albeit harsher conditions—complete their life cycle. However, the accessions with longer HD genetic constitution (i.e., group-2) had the disadvantage of running out of water and nutrients and could not complete their developmental processes which led to a lower association between FSN and HD. Nonetheless, in 2019 and 2020, late-flowering accessions did not experience any water and nutrient deficiency which is reflected in the close correlation between FSN and HD in both the subgroups (Supplementary Figure S11c-f). These results show that—as expected—hotter and drier conditions hasten the growth cycle of plants, thereby lowering the chance for spikelets to fully develop. Therefore, with the changing climatic conditions, efforts should be made to select plants either with earlier HD genetic constitution in the hot and/or arid climatic regions, or select plants resistant to the changing environmental conditions. These efforts will aid in the better development of the spikelets that will consequently be reflected in higher GY.

Chapter 3. **Relationship of pre-anthesis tip degeneration with spike, grain, and shoot traits**

Genetic association of spikelet abortion with spike, grain, and shoot traits in highly-diverse six-rowed barley

Published in:

Frontiers in Plant Science (2022):13

DOI: https://doi.org/10.3389/fpls.2022.1015609

Authors: Roop Kamal, Quddoos H. Muqaddasi, Thorsten Schnurbusch

The original publication is available online at:

https://www.frontiersin.org/articles/10.3389/fpls.2022.1015609/full

Note: In the original publication, pre-anthesis tip degeneration (PTD) was described as spikelet abortion (SA). However, to maintain the consistency in the thesis, pre-anthesis tip degeneration was used in the following text.

3.1 **Abstract**

Pre-anthesis tip degeneration (PTD) is a phenomenon where apical spikelet primordia on an immature spike degenerate. Regardless of the row-type, both apical and basal spikelet degeneration occurs, and their extent decides the number of grain-bearing spikelets retained on the spike—thus, affecting the yield potential of barley. Reducing PTD, therefore, represents an opportunity to increase barley yields. Here, we investigated the variation for apical spikelet abortion along with 16 major spike, shoot, and grain traits in a panel of 417 six-rowed spring barleys. Our analyses showed a significantly large genotypic variation resulting in high heritability estimates for all the traits. PTD varies from 13 to 51% depending on the genotype and its geographical origin. Among the seven spike traits, PTD was negatively correlated with final spikelet number, spike length and density, while positively with awn length. This positive correlation suggests a plausible role of the rapidly growing awns during the spikelet abortion process especially after Waddington stage 5. In addition, PTD also showed a moderate positive correlation with grain length, grain area and thousand-grain weight. Our hierarchical clustering revealed distinct genetic underpinning of grain traits from the spike and shoot traits. Trait associations showed a geographical bias whereby European accessions displayed higher PTD and grain and shoot trait values, whereas the trend was opposite for the Asian accessions. To study the observed phenotypic variation of PTD explained by 16 other individual traits, we applied linear, quadratic, and generalized additive regression models (GAM). Our analyses of PTD revealed that the GAM generally performed superior in comparison to the other models.

The genetic interactions among traits suggest novel breeding targets and easy-to-phenotype "proxy-traits" for high throughput on-field selection for grain yield, especially in early generations of barley breeding programs.

Keywords: Final spikelet number, grain traits, grain morphometry traits, potential spikelet number, pre-anthesis tip degeneration, maximum yield potential, shoot traits, spike traits.

3.2 **Introduction**

Increasing grain yield (GY) remains one of the major goals of barley (*Hordeum vulgare* L.) breeders and geneticists. GY is a complex trait mainly influenced by the concerted action of several agronomic/morphological traits such as plants per m^2, spikes per m^2, spikelets per spike, spike length, grain number per spike, grain weight, thousand-grain weight, and plant height. Therefore, it is of utmost importance to study the variation and relationship among these traits in order to better understand the contribution of these traits towards GY. One of the promising areas to improve barley GY is to increase the spikelet number per spike, which could be achieved by understanding the spikelet abortion process and its association with other spike, grain and shoot traits.

During the early reproductive phase of a barley plant, the inflorescence meristem keeps producing spikelet primordia until the maximum yield potential (MYP) stage is reached; after this stage, the number of spikelet primordia plateaus (Thirulogachandar & Schnurbusch, 2021). The total spikelet primordia number obtainable at or after the MYP stage is known as the potential spikelet number (PSN) for a given immature spike. However, not all spikelet primordia survive until the grain filling phase. Due to pre-anthesis tip degeneration (PTD), spikelet primordia degenerate from the tip, thereby decreasing the actual harvestable GY of barley. This spikelet "initiation-and-degeneration" dynamics results in a proportion of the total spikelet primordia (Kirby, EJM & Appleyard, M, 1987; del Moral et al., 2003; González et al., 2003b). Spikelet primordia that survive until heading and reach the grain filling phase are the final spikelet number (FSN). Since GY is also determined by the FSN harboring grains, it is crucial to study the causal relationship of the extent of PTD with FSN and other major agronomically important traits.

Previously, Alqudah and Schnurbusch (2014) showed that spikelet survival was highly genetically controlled as the heritability was ~0.80. Kamal et al. (2021) reported similar results, where the broad-sense heritability for percentage of PTD was 0.81. Although, PTD has high heritability that governs its genetic nature, it is still a plastic trait in barley as the number of degenerated spikelets varies among the main culm and secondary tillers, genotypes and between row-types. Six-rowed barleys reported to display higher variation for the degenerated spikelets as the lateral spikelets are fertile due to the loss of function of *Vrs1* gene (Miralles & Richards, 2000; Garcıa del Moral et al., 2002; Arisnabarreta & Miralles, 2004; Komatsuda et al., 2007; Alqudah & Schnurbusch, 2014). In addition, the extent of PTD is also affected by the ambient environmental conditions, e.g., temperature, seasons, and locations. It was speculated that lower temperatures between MYP stage and spike emergence stages lead to more survived spikelets (Ellis & Kirby, 1980). Survival of the spikelets was also reported to be different between the main culm and tillers where the former bear more spikelets than the latter (Cottrell et al. (1985).

Similar to barley PTD, floret abortion occurs in wheat. Due to the indeterminate nature of the wheat spikelet, multiple florets are produced within each wheat spikelet but only 3−4 florets survive to develop into grains. Therefore, the number of surviving fertile florets in wheat was also the result of the dynamic "floret initiation and abortion" process. Several practices, such as shortening the photoperiod and increasing the nitrogen fertilizer, improved the number of fertile florets during anthesis primarily by increasing floret survival (Reynolds et al., 2012; González-Navarro et al., 2015; Zhang et al., 2021). One of the hypotheses for floret or spikelet abortion is the competition between the stem and spike (Kirby, 1988; Arisnabarreta & Miralles, 2004). In wheat, numerous studies (Miralles et al., 1998; Gonzalez et al., 2011; Zhu et al., 2019; Zhang et al., 2021) focused on reducing the assimilates partitioning to the stem and diverting the assimilates to the spike. This is because spike dry matter weight is known to have a significant positive correlation with number of fertile florets, which in turn, shows a positive correlation with grain number per spike (GNS). Based on the available wheat literature, two models, namely the trophic model and the developmental model, were proposed to explain the mechanism governing number of fertile florets and grain numbers (Ferrante et al., 2013b). The trophic model describes that floret death is triggered by the dynamics of the spike dry weight between the initiation of terminal spikelet and anthesis. On the other hand, the developmental model proposes that floret abortion is triggered by a fixed developmental stage of the most advanced floret primordium (Bancal, 2008; Bancal, 2009). This implies that the grain number has a stronger correlation with the time of floret death. Afterwards, Thirulogachandar et al. (2021) studied 27 barley accessions (17 two- and ten six-rowed barleys) to explain grain number determination and interactions between maximum spikelet number with other yield component traits and proposed two models, namely, the survival model and the developmental model. To understand the impact on grain number, they studied its association with PSN and spikelet survival. Their results showed that, in six-rowed barleys, GNS is associated with spikelet survival whereas, in two-rowed barley, GNS is strongly associated with PSN.

Since PTD is highly laborious to phenotype and requires the tracking of the MYP stage in each accession under study, earlier studies in wheat and barley were confined to a few accessions. Here, we calculated the variation for PTD in a diverse worldwide panel of 417 six-rowed barley accessions and studied its relationship with 16 other spike, grain and shoot traits. Previously, Kamal et al. (2021) explored that relationship between PSN, FSN, PTD along with heading date and plant height. They concluded that in six-rowed barley, FSN is mainly determined by the extent of PTD while PSN behaves as a suppressor trait and FSN and PSN together explained 93% of the variation for PTD. They also reported that heading date does not directly affect the extent of PTD; however, it indirectly affects PTD by altering number of spikelet primordia formed during early reproductive stage. Nevertheless, apart from PSN and FSN, several other morphological traits such as spike length, spike weight, awn length, grain morphometric traits, grain set, thousand grain weight, culm dry weight were not included in the previous studies. However, these traits could directly affect or interact with PTD.

Here we present analyses of 16 spike, grain, and shoot traits collected on a diverse worldwide six-rowed barley panel to study (1) their genotypic variation, (2) their interaction, and (3) their trait associations with respect to the geographical origin—all in relation to PTD. We report that among the spike traits, only awn length positively correlates with PTD and that PTD, in turn, influences grain morphometric traits—thus, representing a possible sink

competition for GY. Our hierarchical clustering points to a distinct genetical underpinning of grain traits from spike traits. We have also analyzed the data with different regression methods to check other non-linear associations among the traits. Our results provide a deeper understanding of the genetic interaction between PTD and the major spike, grain, and shoot traits.

3.3 **Material and Methods**

3.3.1 Plant materials

A panel of 417 six-rowed spring barleys was evaluated at the fields of Leibniz Institute of Plant Genetics and Crop Plant Research, Gatersleben, Germany (51°49'23''N, 11°17'13''E, 112 m altitude) for three consecutive years (2018–2020). Each year, three replications were planted, and three main culms were selected and tagged in the center of each plot (replication) for data collection. All the accessions were selected based on their spring growth habit, *Ppd-H1* sensitive allele and accessions imitating the genetic diversity harbored in the Federal ex-situ German genebank. The comprehensive details for the panel selection criteria trails are described in Kamal et al. (2021). The majority of the panel consist of landraces (n = 350; 84%) along with a decent proportion of recent cultivars (n = 67; 16%). Standard agronomic practices were applied except for the plant growth regulators.

3.3.2 Phenotyping

We studied 17 traits broadly divided into three categories, namely, spike (n = 7), grain (n = 7) and shoot (n = 3) traits (Table 3.1).

Table 3.1 List of spike, grain and shoot traits evaluated on a panel of 417 six-rowed barley accessions collected from across the globe.

Sr. No.	Spike traits	Grain traits	Shoot traits
1.	Potential spikelet number (PSN)	Grain number per spike (GNS)	Heading date (HD, days from 1st January)
2.	Final spikelet number (FSN)	Grain length (GL, mm)	Plant height (PH, cm)
3.	Pre-anthesis tip degeneration* (PTD, %)	Grain width (GWi, mm)	Culm dry weight (CDW, g)
4.	Spike length (SL, cm)	Grain area (GA, mm^2)	
5.	Spike weight (SW, g)	Grain weight per spike (GWe, g)	
6.	Spike density (SD)	Grain set (GS)	
7.	Awn length (AL)	Thousand-grain weight (TGW, g)	

* In the text, the abbreviation "PTD" represent the trait calculated from PSN and FSN using equation 1 whereas spikelet abortion represents the "in-between ongoing" abortion process during the early reproductive stages.

In 2018, the completely randomized design (CRD) was used, while in the years 2019 and 2020, the panel was grown in an α-lattice design. The individual plot size was ~1.5 m^2, with each plot divided into six rows spaced 0.2 m apart and the sowing density was kept constant across the years with 20 kernels row^{-1}. Comprehensive details of the experiment design are mentioned in Kamal et al. (2021). For PSN data collection, the MYP stage (Thirulogachandar & Schnurbusch, 2021) was tracked and three main culms per replication and accession (in total, nine spikes per accession per year) were selected for microscopic analyses. Not all the accessions reached the MYP stage simultaneously; therefore, main culms for each accession had collected on a routinely basis starting from the stem elongation phase (Anderson et al., 1995). Upon dissection under the stereomicroscope (Stemi 2000-c, Carl Zeiss Micro-Imaging, GmbH, Göttingen, Germany), individual rachis nodes (both differentiated and undifferentiated) were counted on the immature spikes. The total rachis node number was multiplied by three to obtain PSN. Later after the heading, three main culms were selected to calculate the number of spikelet per spike after the degeneration process. The retained spikelets represented FSN. The same culms were later harvested and spike length (SL, cm), spike weight (SW, g), spike density (SD, %, calculated as the ratio of FSN and SL) and awn length (AL) were measured. AL was measured on an ordinal scale ranging from 1–6 (Supplementary Figure. S3.1). PTD occurs over few Waddington stages, therefore, it is difficult to track the PTD process in such a big panel. Therefore, we selected two development stages namely, MYP stage and heading date to calculate PSN and FSN, respectively. PTD was calculated as

$$\mathrm{PTD}(\%) = 100 - \left(\frac{\text{Final spikelet number (FSN)}}{\text{Potential spikelet number (PSN)}} \times 100\right) (1)$$

In the text, the abbreviation "PTD" represent the trait calculated from PSN and FSN using equation 1 but developmentally, pre-anthesis tip degeneration represents the "in-between ongoing" degeneration process during the early reproductive stages. The grain traits namely grain number spike (GNS), grain length (GL, mm), grain width (GWi, mm), grain area (GA, mm^2), grain weight per spike (GWe, g) and thousand-grain weight (TGW, g) all were measured using a digital seed analyzer "Marvin" (GTA Sensorik GmBH, Neubrandenburg, Germany). Marvin analyzer takes into account GNS to calculate TGW as $[(\mathrm{GWe}/\mathrm{GNS}) \times 1000]$. Lastly, grain-set (GS; %) was calculated from the ratio of GNS to FSN. Among the shoot traits, heading date (HD, days) was calculated as the number of days from 1 January until 50% of the spikes are out from the flag leaf sheath. Plant height (PH, cm) was measured on the same three main culm selected earlier for FSN calculation as a distance from the soil surface to the base of the spike and culm dry weight (CDW, g) was measured after harvest.

3.3.3 Variance component analyses and calculation of BLUEs

We employed two field designs, namely, CRD) in 2018 and α-lattice in 2019 and 2020 for phenotypic data collection. The field design was changed from CRD to α-lattice as the latter is known to capture the genetic effects more accurately. Consequently, two different linear mixed-effect models were used to compute individual variance components based on the field designs. In 2018, within-year data analysis was performed by assuming all effects except the intercept as random in eq. 1:

$$y_{ij} = \mu + g_i + r_j + \varepsilon_{ij} (2)$$

where, y_{ij} is the phenotypic record of the i^{th} genotype in j^{th} replication, μ is the common intercept term, g_i is the effect of the i^{th} genotype, r_j is the effect of the j^{th} replication and ε_{ij} denotes the corresponding residual term.

For 2019 and 2020 within-year data analyses, we used the eq. 3 by assuming all effects except the intercept as random as:

$$y_{ijk} = \mu + g_i + r_j + \beta_{(j|k)} + \varepsilon_{ijk} \quad (3)$$

where, y_{ijk} is the phenotypic record of the i^{th} genotype in the j^{th} replication and k^{th} block, μ is the common intercept term, g_i is the effect of i^{th} genotype, r_j is the effect of the j^{th} replication, $\beta_{(j|k)}$ is the block effect of the k^{th} block nested in the j^{th} replication and ε_{ijk} is the corresponding residual term. Within-year repeatability ($\widehat{H}^2$) was calculated as:

$$\widehat{H}^2 = \frac{\sigma_g^2}{\sigma_g^2 + \left(\frac{\sigma_\varepsilon^2}{n_R}\right)} \quad (4)$$

where, σ_g^2 and σ_ε^2 represent the genotypic and residual variances, respectively; n_R denotes the within-year number of replications.Except for PTD, SD, and GS, the across-years variance component analyses were performed by assuming all effects except the intercept as random in eq. 5 as:

$$y_{ijkl} = \mu + g_i + y_j + (g \times y)_{(ij)} + (y \times r \times \beta)_{(j|k|l)} + \varepsilon_{ijkl} \quad (5)$$

where, y_{ijkl} is the phenotypic record of the i^{th} genotype in the j^{th} year and k^{th} replication nested in l^{th} block, μ is the common intercept term, g_i is the effect of i^{th} genotype, y_j is the effect of the j^{th} year, $(g \times y)_{(ij)}$ is the genotype-by-year interaction effect of the i^{th} genotype and j^{th} year, $(y \times r \times \beta)_{(j|k|l)}$ is the l^{th} block nested in k^{th} replication in j^{th} year, and ε_{ijkl} is the corresponding residual term. The across-years broad-sense heritability (H^2) was calculated as:

$$H^2 = \frac{\sigma_g^2}{\sigma_g^2 + \left(\frac{\sigma_{g\times y}^2}{n_y}\right) + \left(\frac{\sigma_\varepsilon^2}{n_y \times n_R}\right)} \quad (6)$$

*w*here σ_g^2, $\sigma_{g\times y}^2$, and σ_ε^2 denote the genotypic, genotype-by-year, and the residual variance, respectively; n_y and n_R represent the average number of years and number of replications, respectively.Since PTD, SD and GS were derived traits, we used the following model to compute the variance components of genotype and years:

$$y_{ij} = \mu + g_i + y_j + \varepsilon_{ij} (6)$$

where, y_{ij} is the phenotypic record of the i^{th} genotype in the j^{th} year, μ is the common intercept term, g_i is the effect of the i^{th} genotype, y_j is the effect of the j^{th} year and ε_{ij} is the corresponding residual term. Accordingly, the H^2 was calculated based on eq.4, except that n_R is replaced with n_y.

3.3.4 *Principal component and correlation analyses*

We drew a scree-plot to describe the percentage of variation accounted for by each principal component (PC) and the principal component analysis (PCA) plot to explain the relationships among the traits. A projection from the origin represented each trait. The length of the projections of a given trait from its origin measured the quality of the trait on the plot. In PCA plot, traits that are away from the origin are the major contributors of the corresponding PC. We also calculated the major trait contributors for the first five PCs.

Pearson's product-moment correlation (r) was computed to examine the relationship among the traits. Moreover, to check the across-years general performance of a given trait, we computed average correlation $(\bar{r})$ by performing Fisher's z transformation, as described in (Muqaddasi et al., 2020). Since the panel consists of accessions collected from five different continents, we used the student *t*-test to check if significant (P <0.05) differences exist for a trait with respect to geographical origin of the accessions.

3.3.5 *Regression analyses: linear, quadratic, multiple, and generalized additive models*

We implemented and compared linear (Su et al., 2012), quadratic (Bobbitt, 2020) and generalized additive model (GAM) (Wood, 2006) regression analyses to check the relationship between PTD and all other traits. For the linear regression, the following model was used:

$$y = \beta_0 + \beta_p x_p + \varepsilon \ (7)$$

where y is the response trait i.e., PTD, β_0 is the intercept, β_p is the regression coefficient, and x_p is a given predictor trait. To check the quadratic relationship among the traits, we used the following model:

$$y = \beta_0 + \beta_1 x_p + \beta_2 x_p^2 + \varepsilon \ (8)$$

where y is the response trait, β_0 is the intercept, β_1 , β_2 are the regression coefficients, x_p is the predictor trait, and x_p^2 is the quadratic function of that trait. In situations, where linear and quadratic regression failed to explain the relationship between the traits, we used the generalized additive model (GAM) as follows:

$$y = \beta_0 + s_1(x_1) + s_2(x_2) + \ldots + s_p(x_p) + \varepsilon \ (9)$$

where y is the response trait, β_0 is the intercept, s_1 , s_2, ... , s_p are the smooth function in GAM, and x_1 , x_2, ... , x_p are the predictor traits. Since GAMs combine the generalized linear model and additive model, they are not restricted to the normal distributions. Instead, these models used smoothening (splines or LOESS) functions that separate the data into "smooth + rough" parts to maximize the smooth and minimize the rough part. In addition to implementing different models, we performed ANOVA to examine the model fit, i.e., to check which model significantly better capture the variation in PTD with respect to a given independent trait. Unless stated otherwise, all calculations were performed in software R (Team, 2013) using the packages lme4 (Bates et al., 2015), corrplot, factoextra, factoMineR (Lê et al., 2008) and mgcv (Wood & Wood, 2015).

3.4 Results

3.4.1 Variance component analyses show large genotypic variance and high heritability estimates

We evaluated 417 six-rowed spring barleys for variation in seven spike-related traits (PNS, FSN, PTD, SL, SW, SD, and AL), seven grain traits (GNS, GL, GWi, GA, GWe, GS, and TGW) and three shoot (HD, PH, and CDW) traits to study the relationship between three classes of morphological traits. Restricted maximum likelihood based variance component analyses showed that both within- and across-years genotypic variance was significantly (P <0.001) greater than zero and was the principal contributor for variance in these traits (Supplementary Tables S3.1−S3.4).

The genotypic variance trend for spike traits was also reflected in broad-sense heritability (H^2) estimates with highest value for SD ($H^2 = 0.96$) and lowest for SW ($H^2 = 0.81$) (Figure. 3.1). Among grain traits, GNS, GWe and GS showed phenotypic plasticity either because of greater year variance or residual variance, whereas, GA, GL, and TGW were generally stable. The high genotypic variance for GA and GL resulted in the highest heritability estimates for both the traits ($H^2 = 0.95$), while the least heritability was observed for GNS ($H^2 = 0.75$) (Figure.3.1). For shoot traits, across-year analyses showed that PH has the highest genotypic variance, followed by CDW and HD (Figure.3.1). Because of temperature fluctuations and slightly different sowing dates over three years, we observed a higher yearly effect for HD. Nonetheless, the large genotypic variance also resulted in high heritability estimates for all the shoot traits (Figure.3.1 and Supplementary Figure S3.2).

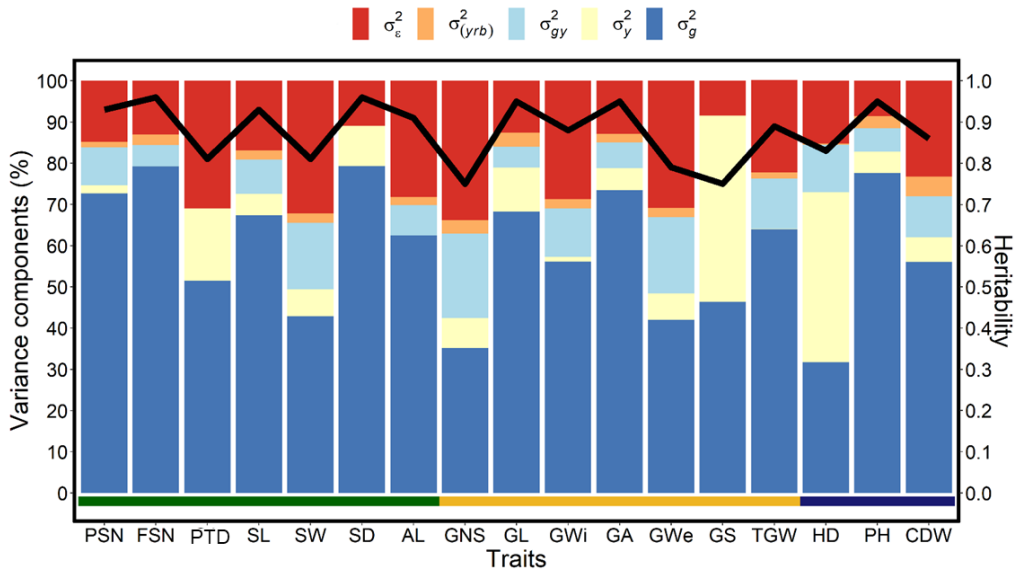


Figure 3.1 Proportion of the different variance components and heritability for each investigated trait in a panel of 417 six-rowed spring barley accessions. The x-axis represents all investigated traits, the left y-axis denotes the proportion of the variance components in percent, and the right y-axis represents the heritability scores. The black line represents the heritability value for the respective trait, σ_g^2 is the genotypic variance, σ_y^2 is the year variance, σ_{gy}^2 is the (genotype × year) interaction variance, $\sigma_{(yrb)}^2$is the year and replication variance with replication nested into the blocks and σ_ε^2 is the error or residual variance. The spike traits

indicated by the green horizontal line includes potential spikelet number (PSN), final spikelet number (FSN), pre-anthesis tip degeneration (PTD in %), spike length (SL in cm), spike weight (SW in g), spike density (SD) and awn length (AL). Grain traits represented by the yellow horizontal line includes grain number per spike (GNS), grain length (GL in cm), grain width (GWi in cm), grain area (GA in cm^2), grain weight per spike (GWe in g), grain set (GS in %), thousand-grain weight (TGW in g) and the shoot traits represented by blue horizontal line includes heading date (HD in days from 1 January), plant height (PH in cm) and culm dry weight (CDW in g).

The best linear unbiased estimations (BLUEs)—calculated within- and across-years—exhibited large genotypic variation (Figure. 3.2 and Supplementary Figures S3.3–3.5). Among spike traits, average correlations—calculated to observe the consistency of the phenotypic data across three growing years—was highest for SD ($\overline{r} = 0.89$) and lowest but reasonably good for PTD ($\overline{r} = 0.61$; Supplementary Figure S3.6a-g). In general, average correlation was high for all the spike traits, suggesting the suitability of the data to draw accurate conclusions. Furthermore, we also observed high average correlation values for all the grain traits with highest for GL ($\overline{r} = 0.91$; Supplementary Figure S3.6h-n). The average correlation ($\overline{r}$) amounted to 0.59 for HD, 0.80 for CDW and 0.92 for PH (Supplementary Figure S3.6o-q)—this reveals a relatively more significant environmental impact on HD than PH and CDW.

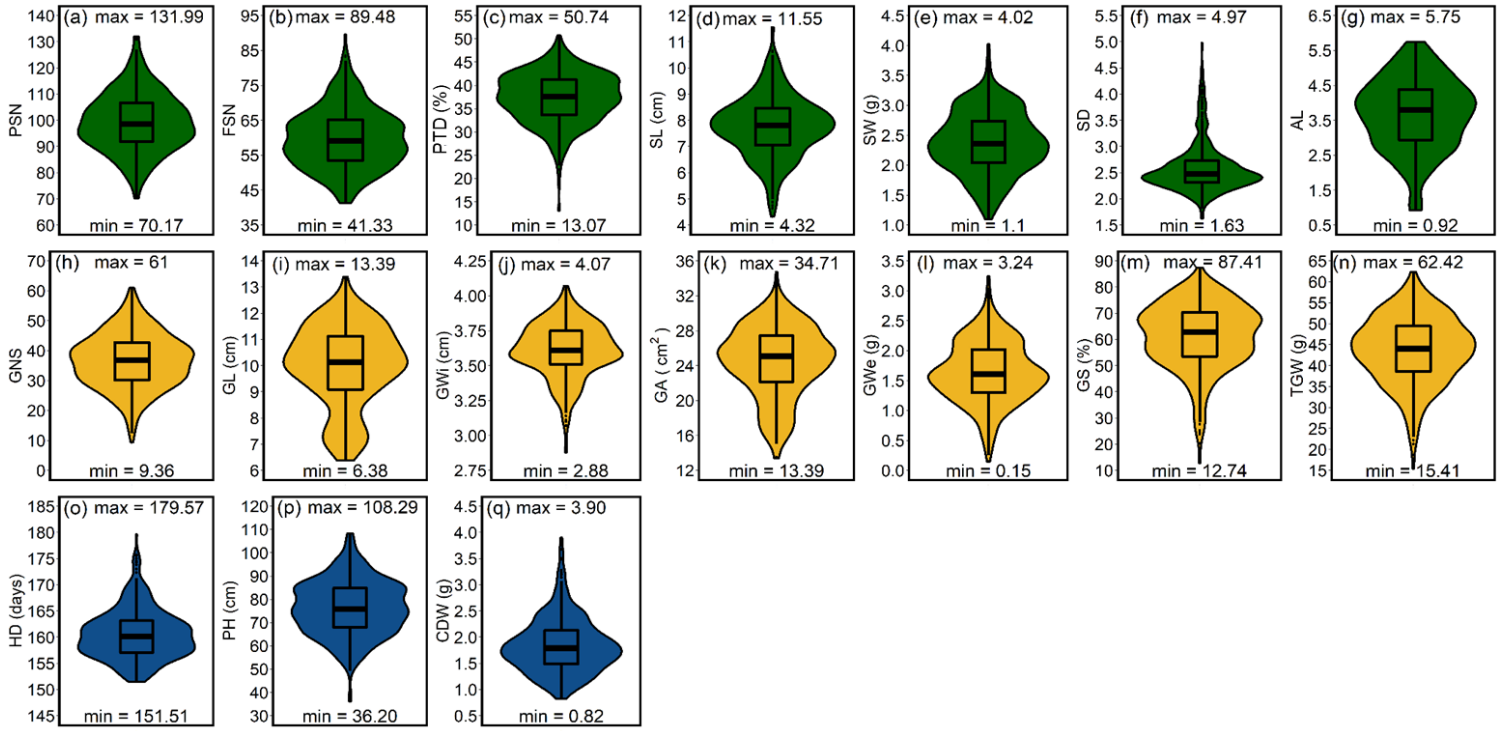


Figure 3.2 Phenotypic distribution of the investigated traits in a panel of 417 six-rowed spring barley accessions. (a-g) frequency distribution for the spike traits, (h-n) frequency distribution for grain traits and (o-q) frequency distribution for shoot traits. "max" and "min" represents the maximum and minimum value for each investigated trait and the box plot within the violin plots represents the lower quartile, median and upper quartile for each trait. PSN = potential spikelet number; FSN = final spikelet number; PTD = pre-anthesis tip degeneration (in %); SL = spike length (in cm); SW = spike weight (in g); SD = spike density; AL = awn length; GNS = grain number per spike; GL = grain length (in cm); GWi = grain width (in cm); GA = grain area (in cm^2); GWe = grain weight per spike (in g); GS = grain set (in %); TGW = thousand-grain weight (in g); HD = heading date (in days from January 1st); PH = plant height (in cm) and CDW = culm dry weight (in g).

3.4.2 Principal component analysis revealed the opposite genetic nature of spike and grain morphological traits

We performed principal component analysis (PCA) on all 17 traits to check the variation and major contributing traits for first five PCs. The scree-plot showed that first five PCs together explained 86.7% of the total variation for the data set (Figure. 3.3a). Interestingly, PC_1, explaining 31.6% variance, showed that all the grain traits and SW vary together (Figure. 3.3b, Supplementary Figure S3.7). On the other hand, PSN, FSN and shoot traits act differently—with all being in PC_2. SD and SL were the major contributor for PC_4 and PTD and AL were the major contributor for PC_5 (Supplementary Figure S3.7).

In PCA plot (Figure. 3.3b), the angle between the arrows/projections illustrates the relationship between the traits: an acute angle depicts a positive association, a 90° angle shows no association, and an obtuse angle describes a negative association. Likewise, arrow length for a given trait explains its impact on a particular PC. For instance, the arrow length for TGW and FSN was longer than other traits indicating that TGW and FSN were the major contributors for PC_1 and PC_2, respectively (Figure. 3.3b). The acute angle between PTD and grain traits such as GL, GWi, GA, GS, TGW and one spike trait AL depicted a positive association among these traits.

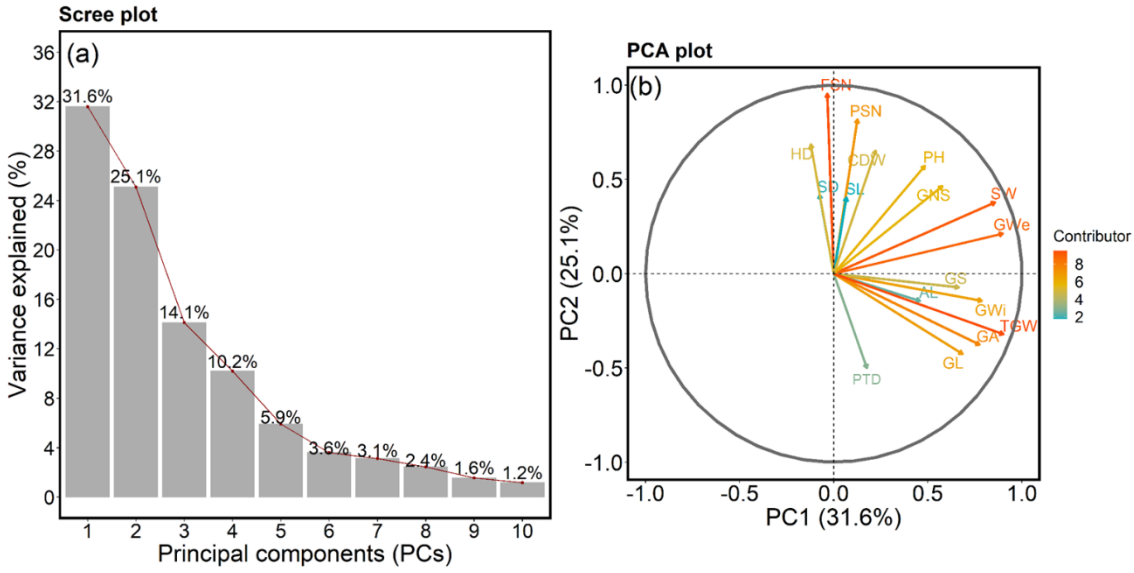


Figure 3.3 Scree and PCA plots for the investigated traits. **(a)** Scree plot highlighting the first ten principal components (PCs) describing the total variation present for all the studied traits. The *x*- and *y*-axis represent the first 10 PCs and variance explained in percent, respectively. The red line shows the trend of the percentage of variance explained starting from PC1 to PC10. **(b)** PCA plot showing the PC1 on the *x*-axis and PC2 on the *y*-axis. The contribution of each trait towards PC1 and PC2 is indicated by the color-coded arrow with maximum contributors indicated by orange to red arrows and minimum contributors by cyan-colored arrows. The legend describes the strength of the contribution. PSN = potential spikelet number; FSN = final spikelet number; PTD = pre-anthesis tip degeneration (in %); SL = spike length (in cm); SW = spike weight (in g); SD = spike density; AL = awn length; GNS = grain number per spike; GL = grain length (in cm); GWi = grain width (in cm); GA = grain area (in cm^2); GWe = grain weight per spike (in g); GS = grain set (in %); TGW = thousand-grain weight (in g); HD

= heading date (in days from January 1st); PH = plant height (in cm) and CDW = culm dry weight (in g).

We observed acute angles between shoot traits, PSN and FSN, suggesting alterations in HD directly affects the spikelet number and plant biomass accumulation. Similarly, an obtuse angles between PTD and most of the spike and shoot traits points towards the opposite nature of these traits, i.e., alternations in HD and consequently in PSN affects FSN, SL and SD further leading to changes in PTD extent. We also checked the contribution of each trait towards the first five PCs (Supplementary Figure. S3.7). As mentioned earlier, the first five PCs were selected as these explained most of the total variation. TGW, FSN, GS, SD and PTD were the major contributors and FSN, GS, FSN, GL and SW were the least contributors for PC_1, PC_2, PC_3, PC_4 and PC_5, respectively (Supplementary Figure. S3.7). It is worth mentioning that the contribution to PC_5 highlights the association between PTD and AL—both the traits tend to vary together.

3.4.3 Correlation analyses show stronger interdependencies among the grain traits as compared to spike and shoot traits

We performed the Pearson's product-moment correlation (r) on within- and across-years BLUEs and observed a strongest positive association between PSN and FSN (Supplementary Figure S3.8a), i.e., higher PSN leads to higher FSN. Both PSN and FSN were positively correlated with SL, SW and SD. The analysis showed that PSN was insignificantly correlated with PTD whereas FSN showed a negative correlation with PTD Across-year analyses showed a negative correlation of PTD with SD, i.e., denser spikes show more PTD ($r = -0.32; P < 0.001$); PTD was, however, positively correlated with AL ($r = 0.34; P < 0.001$; Supplementary Figure S3.8a). As expected, SL and SD showed a negative correlation. We also observed inconsistent significant levels between year 2018 and the remaining years especially for the correlation between PTD and PSN, SL and AL. This could be attributed to the harsh weather condition in 2018 affecting the overall plant growth and the number of spikelet primordia produced on an immature spike.

In contrary to the spike traits, higher correlation coefficients were observed among the grain traits (Supplementary Figure S3.8b). A high positive correlation was noticed among GSN, GWe, and GS ($r > 0.8$). Usually with increase in GNS, TGW tends to decreases, however, a positive correlation was observed between GNS and TGW both within- and across-year data analyses. GA shows higher correlation values with GL ($r = 0.96$) than GWi ($r = 0.76$). But, this trend was reversed for TGW, where GWi ($r = 0.84$) shows a higher correlation with TGW than GL ($r = 0.78$). Except for GA, GWi showed higher correlation coefficients for GNS, GWe, GS and TGW—as compared to GL—indicating that GWi could be an important factor deciding the grain parameters in barley For shoot traits, HD was more significantly correlated with CDW than PH. PH and CDW also show a positive and significant correlation (Supplementary Figure S3.8c).

3.4.4 Hierarchical clustering highlights six distinct clusters separating spike and shoot traits from grain traits

We performed hierarchical clustering based on the correlation matrix that divided all the 17 traits into two main clusters. Broadly, grain traits were entirely separated from shoot traits, whereas the spike traits were clustered with both grain and shoot traits (Figure. 3.4). Further dissection of these two main clusters revealed six distinct sub-clusters.

Cluster-1 contained two shoot (HD and CDW) and two spikes (PSN and FSN) traits. Both HD and CDW were positively correlated with PSN and FSN suggesting the positive influence of HD and CDW on the spikelet primordia initiation. Across sub-clusters, CDW was positively correlated with all the spike traits except PTD. In contrast, HD showed an insignificant correlation with PTD, AL and SW, suggesting that spike characteristics and eventually grain number varies with both variation in days to heading and culm biomass.

Two spike traits, AL and PTD were in cluster 2. Interestingly, AL was the only trait that showed a positive correlation with PTD ($r = 0.34$) suggesting a role for awn growth during spikelet degeneration process (further explained in the discussion section). Across clusters, all grain traits except GNS and GS also showed a positive correlation with AL. Hence, it could be speculated that simultaneous awn development and spikelet primordia growth during juvenile spike growth increases the number of aborted spikelets leading to lesser but bigger grains on the spike. PTD has a negative correlation with GNS and a positive correlation with TGW, but GNS has a positive correlation with TGW. One possible reason for the positive correlation between GNS and TGW could be that GNS is determined based on the PTD but more grains are filled because of the extra assimilates from the developed awns, thereby increasing the number of properly filled grain, grain parameter and finally the TGW. Cluster-3 also harbored two traits, namely, SL and PH. The hierarchical clustering highlighted a higher association of these traits with other shoot and spike traits than grain traits. Cluster-4 composed of GNS, GWe, GS and one spike trait, SW. The highest correlation was observed between GWe and SW ($r = 0.92$). Furthermore, GNS and GS showed a positive correlation with SW, suggesting that more fertile grains spikes are heavier and wider—an important indirect selection criterion in high-throughput on-field phenotyping, especially early generations of breeding programs. We observed an insignificant correlation for GS with most of the spike traits except FSN and SW. Cluster-5 contained only SD which was closely placed with other spike and shoot traits. SD was positively related to GNS and HD but negatively with PTD, GL, and GA.

Grain morphometric traits (GL, GWi and GA) and TGW were in cluster 6 (Figure. 3.4), highlighting the close association of underlying genetic mechanisms for the grain morphometric traits with TGW. Both GL and GA were negatively correlated with PSN and FSN indicating that higher PSN or FSN leads to smaller grains. Similarly, TGW showed a negative correlation with FSN, i.e., higher the spikelet number after degeneration, lower the TGW. Except PSN and FSN, cluster-6 traits positively correlated with PTD, SW and AL. Therefore, from the positive correlation of grain traits, especially with PTD, it can be concluded that an increase in PTD leads to fewer spikelets and hence larger GL, GA and TGW. This happens through proper development of the grains that further increases SW.

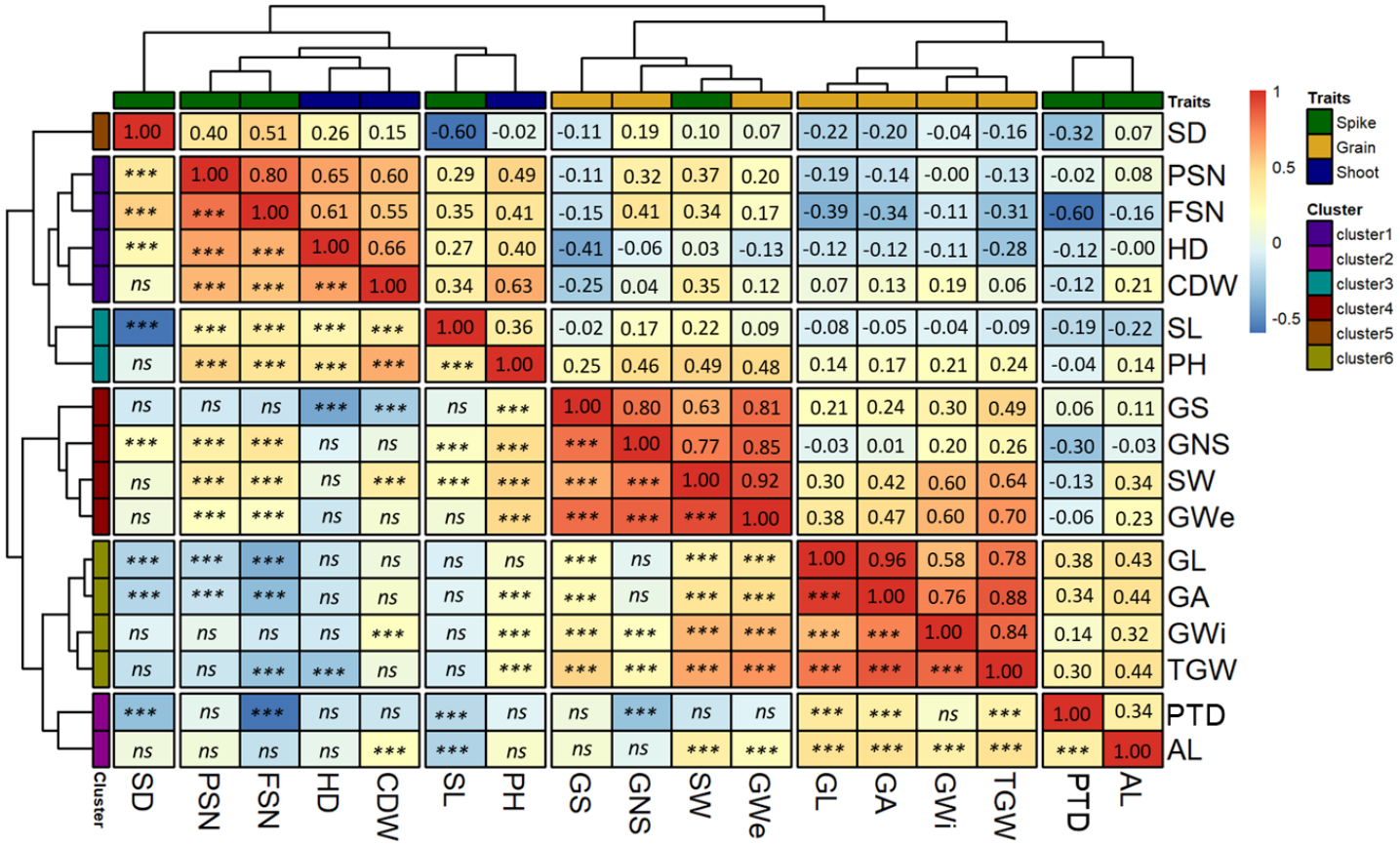


Figure 3.4 Hierarchical clustering based on the Pearson-product moment correlation analysis among the BLUEs of the investigated traits. The *x*- and *y*-axis both represent the investigated traits. The spike, grain and agronomic traits are color-coded as shown in the legend "Traits". The second legend shows the scale for Pearson-product moment correlation. The spike traits highlighted by the green color includes potential spikelet number (PSN), final spikelet number (FSN), pre-anthesis tip degeneration (PTD in %), spike length (SL in cm), spike weight (SW in g), spike density (SD) and awn length (AL). Grain traits highlighted by the yellow color includes grain number per spike (GNS), grain length (GL in cm), grain width (GWi in cm), grain area (GA in cm^2), grain weight per spike (GWe in g), grain set (GS in %), thousand-grain weight (TGW in g) and the shoot traits represented by blue color includes in heading date (HD in days from January 1^{st}), plant height (PH in cm) and culm dry weight (CDW in g).

3.4.5 Geographical origin of the accessions significantly impact the variation for all the traits

We studied the influence of geographical origin of the accessions on each trait by dividing the whole panel into six groups according to the continents, viz., African (*n* = 73), Asian (*n* = 183), European (*n* = 80), North American (*n* = 28), South American (*n* = 12) and accessions with unknown origin (*n* = 41). For comparison, we excluded the accessions with an unknown origin. For most of the traits, North- and South-American accessions did not show any significant differences from other continents (Figures 3.5-3.7). One reason for less variability within both the American accessions could be the low sample number as compared to other continents.

African accessions showed significant ($P<0.001$) lower values for PSN, FSN, SL, SD, GNS and HD. In the nutshell, African accessions, took fewer days to head, therefore, immature spikes stay in spikelet initiation phase for a shorter period leading to lower PSN, consequently lower FSN and ultimately lower GNS at harvesting. Contradictorily, European accessions displayed higher values for all the shoot traits (Figure 3.7). As the HD increases in European accessions, spikelet initiation phase is also expected to be longer (higher PSN; Figure 3.5).

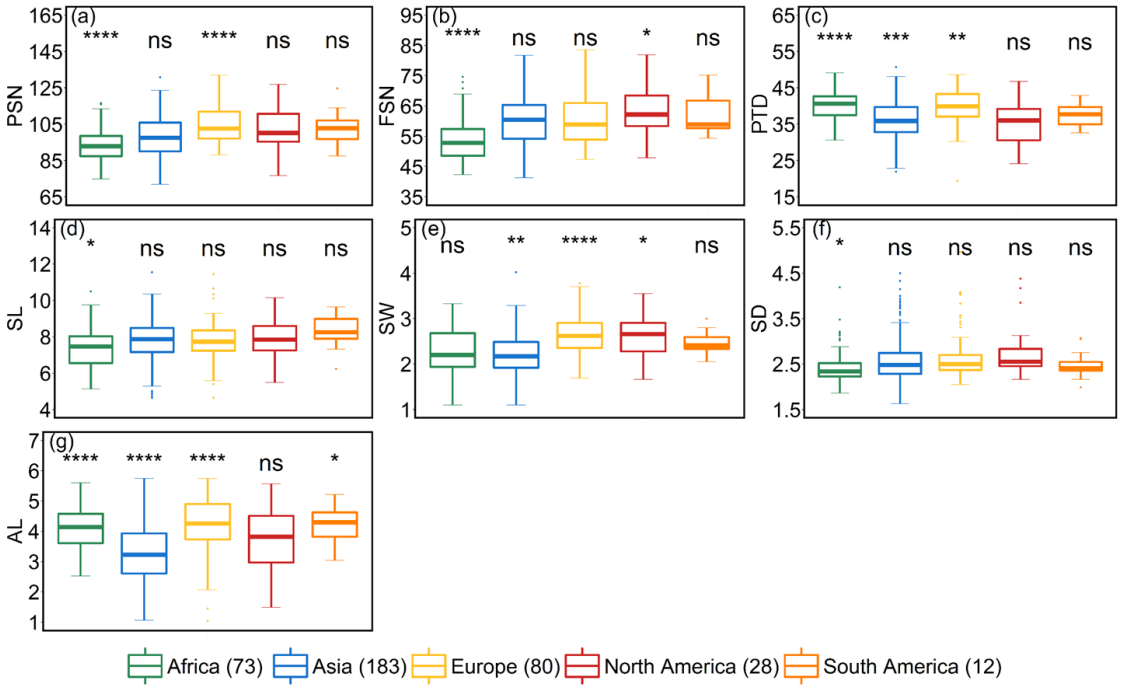


Figure 3.5 Comparison of accessions according to their geographical origin for the spike traits. (a) potential spikelet number (PSN), (b) final spikelet number (FSN), (c) pre-anthesis tip degeneration (PTD in %), (d) spike length (SL in cm), (e) spike weight (SW in g), (f) spike density (SD) and (g) awn length (AL). '*n*' denotes the number of accessions in this panel and the number of the accessions belonging to each continent; ****, ***, **, and * = significance level at the 0.0001, 0.001, 0.01, and 0.05 probability level, respectively and 'ns' represents insignificant differences and the legend represents different continents.

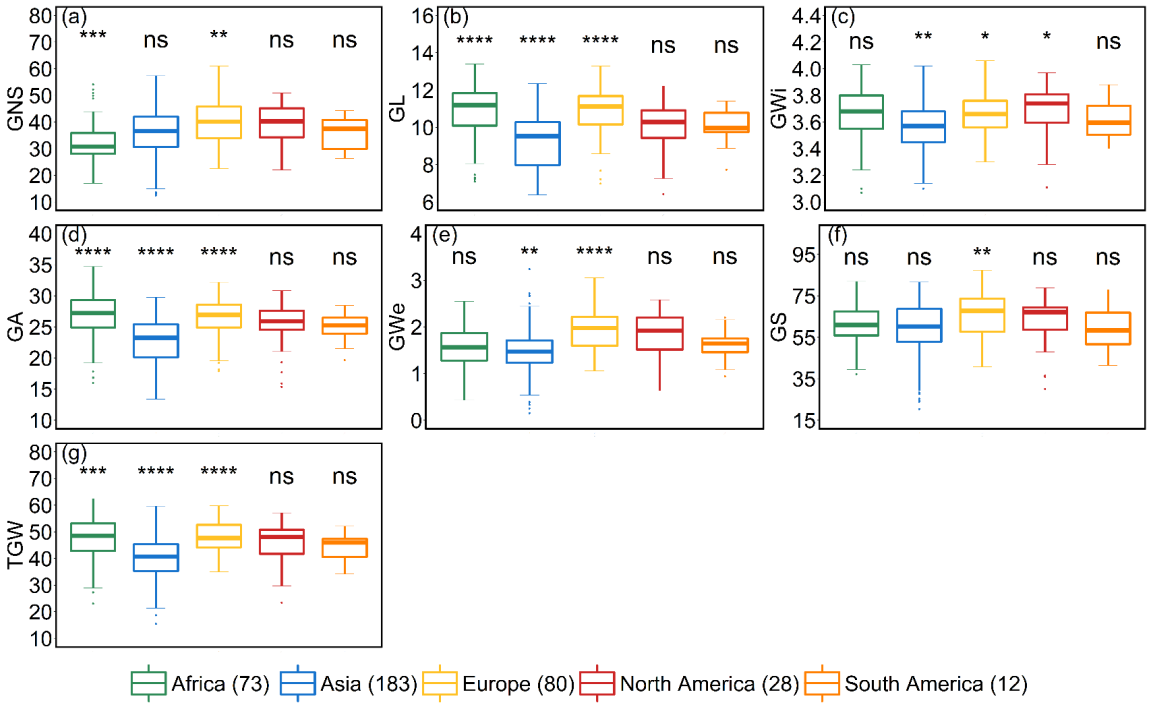


Figure 3.6 Comparison of accessions according to their geographical origin for the grain traits. (a) grain number per spike (GNS), (b) grain length (GL in cm), (c) grain width (GWi in cm), (d) grain area (GA in cm^2), (e) grain weight per spike (GWe in g), (f) grain set (GS in %) and (g) thousand-grain weight (TGW in g). '*n*' denotes the number of accessions in this panel and the number of the accessions belonging to each continent; ****, ***, **, and * = significance level at the 0.0001, 0.001, 0.01, and 0.05 probability level, respectively and 'ns' represents insignificant differences and the legend represents different continents.

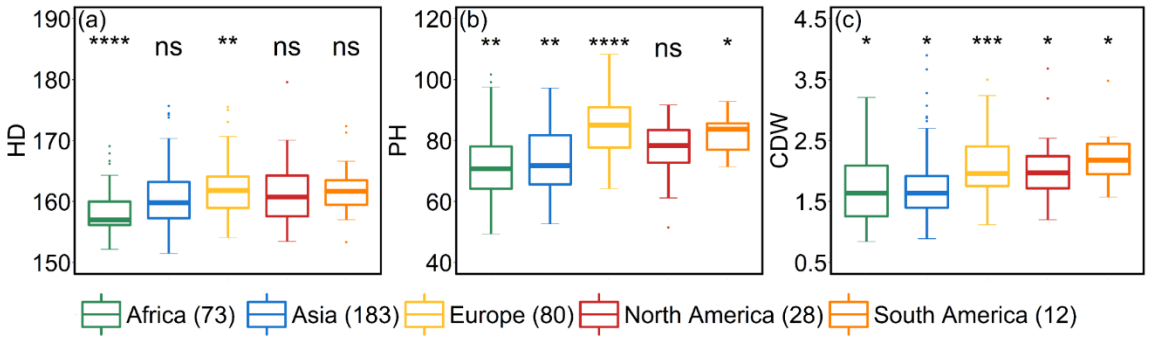


Figure 3.7 Comparison of accessions according to their geographical origin for the shoot traits. (a) heading date (HD in days from January 1st), (b) plant height (PH in cm) and (c) culm dry weight (CDW in g). '*n*' denotes the number of accessions in this panel and the number of the accessions belonging to each continent; ****, ***, **, and * = significance level at the 0.0001, 0.001, 0.01, and 0.05 probability level, respectively and 'ns' represents insignificant differences and the legend represents different continents.

As stated above, a positive correlation was observed between PTD and AL. We observed the same trend for African, Asian and European accessions. Asian accessions with smaller AL showed lower PTD. Whereas, African and European accessions with larger AL have high PTD. This further points to a potential role of awn development in affecting the extent

of PTD. However, the lower PTD in Asian accessions leads to smaller grains—all grain morphometric traits and TGW were reduced (Figure. 3.6). One possibility that could be exploited under European conditions is to pyramid the lower PTD alleles from Asian accessions and higher grain size alleles from the European accessions in crosses and further test the relationship between the PTD extent and grain size. Since accessions belonging to these continents can also be differentiated as landraces (n = 350) and cultivars (n = 67), we performed an analysis to check whether there exist any differences between landraces and cultivars for all the investigated traits. The significant differences were observed only for FSN, PTD SW and GNS. PTD was higher for landraces than cultivars resulting in lower FSN and consequently lower GNS and SW (Supplementary Figure S3.9). This hints towards the breeding progress for higher grain number and yield and consequently lower PTD.

3.5 **Discussion**

Early and late reproductive phases govern important agronomic traits in cereals. From a practical standpoint, early vigor and later stand strength correspond to better GY in barley and wheat. In early generations (e.g., F_4, F_5, DH_1), breeders select the genotypes or plots based on the plant and spike ideotypes to further test in yield plots. Since GY is an integral parameter of many component traits, such as number of spikes per unit area, number of spikelets per spike, number of grain per spike, spikelet fertility, spike density, spike length, grain weight per spike and many more, the extent of PTD directly or indirectly effect these traits. It is thus vital to study spikelet abortion from early to late reproductive stages and its association with other agronomically important traits.

3.5.1 Plausible relationship between awn length, pre-anthesis tip degeneration and grain development

PTD is a laborious trait to phenotype as it entails the phenotyping of a given accession at both early (for PSN) and late reproductive (for FSN) stages. Previously, in barley, the relationship of aborted or degenerated spikelets with other traits was elucidated but in only a handful of genotypes. This is the first study where we report the genotypic association of PTD with other spike, shoot, and grain traits at a large scale. As mentioned in the result section, AL was the only spike trait positively (moderately) correlated with PTD. Depending on the growing conditions and genetic background of an accession, awns act as a boon and bane for a plant. (Vervelde, 1953; McKenzie, 1972; Knott, 1986; Weyhrich et al., 1994; Motzo & Giunta, 2002; Martin et al., 2003; Ali et al., 2010; Guo & Schnurbusch, 2016). The positive effect of the awns has been attributed to their role in photosynthesis i.e., awns can function as an important source organ because, of their short nutrient transport route to grains, especially when flag leaves start to senesce (Li et al., 2006). While, the negative effect of awns arises as they act as a significant sink competitor of the growing spikelets for the available assimilates in an immature spike. Guo and Schnurbusch (2016) hypothesized that GY is influenced by the redistribution of assimilates within the spike associated with vigorous awn development. In wheat, it has been shown that the growing awns need more assimilates for their development as they represent 40% of the total spikelet biomass prior to spike emergence. (Rebetzke et al., 2016) This results in competition for assimilates between rapidly developing awns and florets, leading to fewer fertile spikelets, reduced floret fertility, and abortion of the distal florets.

In the present study, within- and across-year analyses revealed—except for 2018, where the correlation was non-significant but positive—a positive association between AL and PTD. One possible reason for the insignificant correlation in 2018 could be the extreme weather conditions. Kamal et al. (2021) showed that among the three growing seasons, highest temperature, global solar radiation, and least humidity levels were observed in 2018. The positive correlation between AL and PTD, nevertheless, showed that the proportion of the degenerated spikelets also increased with rapid awn growth from Waddington stages 4.5 and more. This is in line with a previously proposed hypothesis that competition between awns and floret development during the juvenile spikes might affect the number of fertile grains during harvest (Schaller & Qualset, 1975). Based on the AL ordinal scale, we also divided the whole panel into six groups and observed a general positive trend of the impact of AL on PTD (Supplementary Figure S10a). We further observed the association of geographical origin for each AL group with PTD. The extent of PTD varied according to the proportion of Asian, European and African accession in a given AL group, i.e., with an increase in the European and African accessions in a AL group, both AL and PTD increased (Supplementary Figure S10a). We observed that as AL increases, the grain morphometric traits and TGW also increases (Supplementary Figure S10b), pointing towards the "source" role of the developed awns.

Interestingly, while performing the multiple linear regression, AL was not one of the predictor variables for PTD. Both PSN and FSN were the major predictor of PTD (Kamal et al., 2021) but when one of the major predictor variables (PSN or FSN) was removed from the model, the effect of AL became significant (Table 3.2). Replacement of FSN with AL in the model leads to non-significant and significant results of PSN and AL, respectively (Table 3.2a). The non-significant results for PSN highlight that alone PSN cannot predict the variation for PTD. However, FSN and AL together predict ~41% of the variation for PTD (Table 3.2b) and AL alone explained ~11% of the variation in PTD (Table 3.2a). Nonetheless, since our whole panel comprises only awned accessions, evaluation and comparisons of PTD and AL and their effect on grain-traits in awnless or anwletted six-rowed barley would further shed light on this relationship. For example, isogenic lines of barley and wheat differing w.r.t. to AL as well as presence and absence of awn could be useful to study their effect on PTD and consequently GY (Schaller et al., 1972; Motzo & Giunta, 2002; Sanchez-Bragado et al., 2020); these lines appear similarly valuable to study the relationship PTD and AL using isogenic lines.

Table 3.2 Multiple linear regression for pre-anthesis tip degeneration (PTD) using potential spikelet number (PSN), final spikelet number (FSN) and awn length (AL) as predictor traits

2a. Multiple linear regression using PSN and AL as predictor traits					
PTD <- lm(PTD ~ PSN + AL)					
Summary (PTD)					
	Estimate	Standard error	t-value	Pr(>|t|)	Significance
(Intercept)	32.89982	2.31678	14.201	< 2e-16	*** (P < 0.001)
PSN	-0.02113	0.02218	-0.953	0.341	*ns*
AL	1.78447	0.24012	7.432	6.20E-13	*** (P < 0.001)

Residual standard error	5.026 on 414 degrees of freedom (DF)
Multiple R-squared	0.1179
Adjusted R-squared	**0.1136**
F-statistic	27.67 on 2 and 414 DF
P-value	5.27E-12

2b. Multiple linear regression using FSN and AL as predictor traits					
PTD <- lm(PTD ~ FSN + AL)					
Summary (PTD)					
	Estimate	Standard error	t-value	Pr(>|t|)	Significance
(Intercept)	53.31362	1.70494	31.27	< 2e-16	*** (P < 0.001)
FSN	-0.34718	0.02374	-14.624	< 2e-16	*** (P < 0.001)
AL	1.31447	0.19693	6.675	7.97E-11	*** (P < 0.001)

Residual standard error	4.086 on 414 degrees of freedom (DF)
Multiple R-squared	0.4171
Adjusted R-squared	**0.4143**
F-statistic	148.1 on 2 and 414 DF
P-value	<2.22-16

3.5.2 Genetic interactions of PTD with other spike- and grain-traits is generally better revealed by generalized additive models as compared to linear and quadratic regressions

In the generalized additive model (GAM), the linear predictors predict some unknown smooth monotonic function (*s*) of the expected value of the response where the response has a known mean-variance relationship (Wood et al., 2016; Hastie, 2017). Bera et al. (2021) used linear regression, polynomial regression, and GAM to study the canopy cover estimation in the dry deciduous forest of West Bengal, where they reported that GAM performed better than the other two regression models. Recently, GAM was used to predict soybean maturity under African environments (Marcillo et al., 2021). Here, we implemented linear, quadratic and GAM models to study the relationship between PTD and all other 16 traits (Supplementary Figure S11-13). PTD was used as a response trait whose variation was explained by the individual predictor trait.

Our ANOVA to examine the model efficiency revealed that GAM outperformed both linear and quadratic regression for most of the spike traits, namely, PSN, SL, SD, and AL (Supplementary Table S5, Supplementary Figure S11). For FSN, however, no significant differences were observed—linear model best described the relationship between FSN and PTD (Supplementary Figure S11b). For SW, both linear and quadratic models predicted the same but higher variation in PTD than GAM. Hence, either of them can be used to explain the relationship between PTD and SW (Supplementary Figure S11d).

For the grain-traits, namely, GNS, GL, GA and GS, GAM performed better than linear and quadratic regressions (Supplementary Table S5, Supplementary Figure S12). For GWe, however, no significant differences were observed among the three models. For GWi and TGW, we observed significant differences only between linear and GAM, where GAM explained higher variance for PTD than linear regression. These results point to the complex genetics of traits per se and their complex interaction. We show that since two traits are not always related to each other linearly or quadratically, it is essential to explore the other non-linear relationship between the traits in-depth. Nevertheless, a multiple-year, multiple-location large data set gathered from a robust breeding design is essential to determine the true genetic interactions among traits.

3.5.3 Increased days to heading increases PSN and FSN but decreases grain set and thousand-grain weight

Previously it was shown that HD affects spikelet initiation and growth phases in barley and wheat (Gol, L. et al., 2017; Mulki et al., 2018; Ochagavia et al., 2018; Prieto et al., 2018; Prieto et al., 2020). Our analyses revealed that all three shoot traits, viz., HD, PH, and CDW, showed a significant positive correlation with PSN and FSN. HD showed a slightly higher correlation with PSN ($r = 0.65; P < 0.001$) than FSN ($r = 0.61; P < 0.001$), suggesting that with an increase in HD, the length of the spikelet initiation phase might also increase, resulting in more spikelet primordia on the spike. With an increase in PSN, the chances of retaining more spikelets (FSN) also increases. Our results showed that PSN, FSN, and HD negatively impact grain morphometric traits. Also, variation in the PC_1 was mainly due to grain traits, and variation for PC_2 (25.1%) could be ascribed to spike and shoot traits (Figure 3.3b, Supplementary Figure S3.9a-b); thus, confirming correlation results.

From the geographical distribution analyses, we observed that Asian accessions had lower PTD, thus affecting all the grain morphometric traits. For example, in Asian accessions, we observed decreased GL, GWi, GA, GWe and TGW values (Figures 5, 6). In contrast, European accessions exhibited later HD, leading to higher PSN and ultimately higher PTD. Since PTD was higher compared to Asian accessions, this led to an increased GL, GWi, GA, GWe and ultimately TGW (Figures 5-7). Since one primary objective of any breeding program is to increase the GY by increasing the GNS or increasing the grain number per unit area while keeping grain size within a certain acceptable range, both European and Asian accessions can be exploited in the breeding schemes. The mining of favorable alleles for lower PTD from Asian accessions and alleles for larger grain size and TGW from European accessions may aid in developing ideotypic spike and grain architecture, thereby altering the harvestable GY.

Chapter 4. Genome-wide association analyses for pre-anthesis tip degeneration and other investigated traits

Insight into the genetics of pre-anthesis tip degeneration and related traits in barley

Submitted for peer review:

Authors: Roop Kamal Muqaddasi, Quddoos H. Muqaddasi, Corinna Trautewig, Yongyu Huang, Nandhakumar Shanmugaraj, Michael Melzer, Henning Tschiersch, Beata Chmielewska, Miriam Szurman-Zubrzycka, Iwona Szarejko, Axel Himmelbach, Murukarthick Jayakodi, Nils Stein, Martin Mascher, Thorsten Schnurbusch

4.1 Abstract

Identifying and improving key grain yield (GY) components, such as the number of spikelets per inflorescence, is an important goal of barley research. Pre-anthesis tip degeneration (PTD) is a common phenomenon where apical spikelet primordia on immature spike-type inflorescences degenerate prematurely in a basipetal pattern. The magnitude of spike PTD, thus, affects final grain-bearing spikelets and eventually GY. In this study, we conducted genome-wide association scans using ~22 million high-quality SNPs for spike PTD and 16 related traits in a diverse six-rowed spring barley panel of 416 accessions. In total, 132 QTL were detected, including three main-effect QTL for PTD. One PTD-QTL on chromosome 3H explained 14% of phenotypic variance and harbored a candidate gene—an octotricopeptide repeat protein with α-helical RNA-binding (RAP) domain. Natural sequence variations associated with this single copy barley RAP (*HvRAP*) gene showed a substitution in a conserved amino acid within the RAP domain, leading to enhanced PTD. Moreover, phenotypic analyses of *TILLING*-derived *hvrap.g* mutant family members revealed reduced photosynthetic efficiency, chloroplast structural defects, higher PTD, and decreased grain number per spike. Based on shared QTL, we propose a QTL network and putative candidate genes for important traits for future in-depth functional analyses.

Keywords: GWAS, pre-anthesis tip degeneration, *TILLING*, *HvRAP*, QTL-based network

4.2 Introduction

Barley (*Hordeum vulgare* L.) belongs to the *Triticeae* tribe of the *Poaceae* family. It is the fourth important cereal crop in the world after rice (*Oryza sativa* L.), maize (*Zea mays* L.), and wheat (*Triticum* sp. L.) (FAO, 2021) and is considered the genetic model for the cereals in *Triticeae* (Ariyadasa et al., 2014). Barley possesses an indeterminate spike-type inflorescence with single-flowered spikelets that develop grains. Barley plants pass through three development phases: vegetative (leaf initiation), reproductive, and grain-filling. The reproductive phase is further divided into early (spikelet initiation phase: double ridge stage to awn primordium stage) and late reproductive (spikelet growth and development; from awn primordium to anthesis) phases. Early reproductive phase determines the number of spikelet

primordia initiated while the late reproductive phase determines the number of spikelets that develop into the fertile floret (Kirby & Appleyard, 1987; Alqudah & Schnurbusch, 2014). After the main culm spike has reached its maximum yield potential stage (Thirulogachandar & Schnurbusch, 2021), the inflorescence meristem dome ceases its growth and starts to degenerate, followed by the degeneration of the subjacent spikelet primordia. Hence, due to apical spikelet primordia degeneration (henceforth named pre-anthesis tip degeneration; PTD), the full potential of a barley spike is not realized (it is important to note that in the previous publications, i.e., Kamal et al. (2022a); Kamal et al. (2022b) spike PTD was described as spikelet abortion, SA). Barley spike PTD is considered as a multilayered and developmentally programmed process that decreases the number of developing and maturing spikelets (Shanmugaraj et al., 2023); thus, it translates into lowered grain number per spike that is one of the major determinants of harvestable grain yields in barley. Although large genotypic variance exists for spike PTD in barley, it is also influenced by environmental fluctuations such as temperature, light conditions, and nutrition (Kamal et al., 2022b).

Improving the grain number per spike (GNS) remains one of the major goals of cereal breeding programs (Würschum et al., 2018; Sierra-Gonzalez et al., 2021). In cereals, GNS is a consequence of the spikelet primordia initiation and their subsequent development into fertile florets that bear grains. Consequently, it becomes crucial to discover the underlying genetic factors responsible for spikelet primordia number, PTD, and the number of spikelets retained after the degeneration. However, little is known about the genetic factors related to the degeneration process. Recently, Huang et al. (2023) characterized a *tip sterile 2* (*tst2*) mutant in barley. Mutations in the responsible *HvCMF4* gene, encoding a *CCT* [CONSTANS (**C**O), **C**O-like, and TIMING OF CAB1 (**T**OC1)] domain-containing *MOTIF FAMILY 4* protein, lead to excessive apical spikelet degeneration and pollination failure. It was proposed that the fate of spikelet development is timed by an altered vascular-specific circadian clock that coordinates floral initiation and growth. Moreover, it was shown that barley spike PTD is associated with sugar depletion, amino acid degradation, late abscisic acid (ABA) biosynthesis and signaling, as well as light regulation, chloroplast development, and chlorophyll biosynthesis (Huang et al., 2023; Shanmugaraj et al., 2023). Still, efforts are ongoing to understand and discover other genes related to PTD.

High-density genotyping platforms have helped reveal the genetic architecture of various agronomically important traits. In the recent past, barley researchers have mainly used the 9k (Comadran et al., 2012) and 50k (Bayer et al., 2017) SNP arrays to identify trait-linked loci. Thanks to cost-efficient sequencing and computational platforms (Muir et al., 2016; Levy & Boone, 2019), millions of SNPs can now be analyzed. Recently, Li et al. (2022) used about nine million SNPs to identify marker associations for drought-resistance traits in hulless barley. In addition to abundant SNPs, the statistical power of genome-wide association studies (GWAS) is affected by the population size and structure, linkage disequilibrium (LD), phenotypic variation, and heritability of a trait. Usually, a population comprising a few hundred genotypes with traits displaying wide phenotypic variations and high heritability is suitable for GWAS. Since the existence of population stratification within a population often leads to spurious associations, its correction based on structure and/or genomic relatedness helps detect true associations (Yu et al., 2006).

In this study, we analyzed 17 barley traits to identify the underlying QTL and candidate genes. By utilizing the whole genome shotgun approach, we identified approximately 22 million high-quality SNPs covering all barley chromosomes. We report novel QTL and highlight high-confidence genomic intervals containing genes for the investigated traits. In particular, we report a candidate gene underlying a major QTL for spike PTD that shows differences in two major haplotypes associated with differences in spike PTD, spikelet number, and chlorophyll levels of leaves and spikes. A mutation in the candidate gene annotated as an octotricopeptide repeat protein with α-helical RNA-binding (RAP) domain leads to loss of chlorophyll and decreased grain number in TILLING mutant lines. We report another spike PTD candidate gene, i.e., *HvRAN2* (Ras-related nuclear protein, belonging to the family of small GTP binding protein), that shows mRNA expression within the inflorescence meristem—a tissue where spike PTD typically occurs. Our analyses and reported candidate genes further aid towards a deeper understanding of the mechanism underlying spike PTD in barley and may help improve the grain set in this important cereal crop.

4.3 **Material and Methods**

4.3.1 Barley association mapping panel: selection criteria, field trials, and phenotyping

A spring barley association mapping panel consisting of 416 six-rowed accessions representing the genotypic diversity of accessions held at the German Federal *ex-situ* gene bank was selected for GWAS analyses (in previous chapters, a panel of 417 is mentioned. However, to perform the GWAS analyses, we had a panel of 416 accession as for one accession the WGS data was not available). To avoid phenotypic effects and severe population structure associated with the barley row-type gene *Six-rowed spike 1* (*Vrs1*), only six-rowed accessions were used to study spike PTD and its relationship with other traits. Moreover, all accessions did not possess the *ppd-H1* insensitive allele to avoid late flowering, as well as longer and slower growth, which often strongly impacts the production of spikelet primordia in barley (Turner et al., 2005).

Experiments were conducted at field facilities of the IPK (51°49’23’’N, 11°17’13’’E, 112 m altitude) in replicated completely randomized (CRD) and α-lattice experimental designs (Fisher, 1937; Le Clerg, 1966). CRD experimental design was implemented in 2018, whereas α-lattice in 2019 and 2020. Main culms were tagged to avoid tiller selection for the sample collection. Three main culms were selected from the center of each plot (i.e., replication) at three developmental stages: maximum yield potential (MYP) stage, heading date, and harvesting. PTD was calculated as:

$$\mathrm{PTD}(\%) = 100 - \left(\frac{\text{Final spikelet number}}{\text{Potential spikelet number}} \times 100\right)$$

The details regarding traits, sample collection at various stages and corresponding protocols, and phenotypic data analyses including heritability estimates for all 17 traits are described in Kamal et al. (2022a); Kamal et al. (2022b). Here, we used interquartile range method (Vinutha et al., 2018) for outlier removal for PTD.

4.3.2 *Whole genome shotgun sequencing, SNP calling, population structure, and phylogenetic analyses*

The whole genome shotgun sequencing was performed as described in Huang et al. (2024). To achieve the targeted depth coverage of ~3x, raw reads were trimmed with Cutadapt after aligning to the Morex genome (*v*2) (Monat et al., 2019) using minimap2 (Li, 2018). The alignments were sorted using Novosort (http://www.novocraft.com) and the bcftools/SAM Tools was used to call variants (SNPs and InDels). The resulting VCF was converted into Genomic Data Structure (GDS) format using SeqArray package (Zheng & Gogarten, 2017) in R to obtain a SNP matrix. The SNP mining yielded 49,450,236 SNPs. The quality criteria of removing markers with minor allele frequency < 0.05 and maximum of 10% missing data using *PLINK* *v*1.90b6.9 (Purcell et al., 2007) yielded 21,980,167 SNPs. The processed SNPs were further imputed using *BEAGLE (v5.1)* (Ayres et al., 2012) to generate a SNP matrix without missing data. The default parameters were used for the mentioned softwares.

The population structure within the panel was visualized using *PLINK*. The eigenvectors and eigenvalues were calculated based on the variance-standardized relationship matrix using the parameter *--pca*. The eigenvectors and eigenvalues were later used to calculate the percentage of variance explained, and to construct the two-dimensional PCA plot. Phylogenetic relationship among the selected accessions was studied via a neighbor-joining tree constructed using *PHYLIP* software (Baum, 1989) and visualized using *iTOL* (Letunic & Bork, 2021).

4.3.3 *Genome-wide association studies: selection of best-fit models, QTL identification, and reference phenotype-genotype map*

The GWAS for the traits was performed using two different models: *GEMMA* (Zhou & Stephens, 2012; Zhou & Stephens, 2014) and *BOLT-LMM* (Loh et al., 2015). For *GEMMA*, GWAS was performed using a univariate mixed linear model with the default parameters. The genomic-related matrix was generated using the *--gk* parameter, and then GWAS was performed using *--lmm* function using first three PCs as covariates. *BOLT-LMM* uses a Bayesian approach to model SNP effects with non-Gaussian prior distributions that better accommodates loci of both small and large effects. GWAS was performed using parameters *--bfile --phenoFile --phenoCol --covarFile --qCovarCol --LDscoresFile --maxModelSnps --lmmForceNonInf --numThreads --statsFile*. The association results were visualized as Manhattan and Quantile-Quantile (QQ) plots using *qqman* (Turner, 2018) in *R (v4.2.2)*. QQ plots were used to select the appropriate GWAS model for a given trait. For multiple statistical testing, false discovery rate and Bonferroni's correction at α-level of 0.05 were used to control false positives (Benjamini & Hochberg, 1995).

Manhattan plots for individual chromosomes were drawn to delineate significant regions of a trait. LD values for a significant region were calculated in *PLINK* using the most significant SNP (with highest $-\log_{10}(P)$ value) as a reference SNP with parameters *--file --r2 --ld-snp --ld-window-kb --ld-window --ldwindow-r2*. Significant peaks belonging to different LD blocks were considered as different QTL, whereas peaks within the same LD block were selected as one major QTL. The trait-associated markers and previously reported genes or QTL near the significant markers were anchored onto a reference genotype-phenotype map which

was aligned to the Morex *v*2 and visualized using Phenogram (Wolfe et al., 2013)—the URL is http://visualization.ritchielab.org/phenograms/plot.

4.3.4 LD clumping, unique and shared QTL, LD-based candidate gene identification, and expression analyses

To identify the most significant genetic variants related to a region, we used LD-based clumping using *PLINK*. LD clumping divides the most significant SNPs on different chromosomes into a small number of "clumps", which facilitates assessing the number of independent loci associated with a given trait. Clumping was performed using the default parameters *--bfile --clump --clump-verbose*. After identification of significant SNPs or peaks, unique and shared QTL were identified for each trait, and subsequently a QTL network was constructed.

Based on Morex *v*2, the genes associated directly either with significant SNPs or flanking to the significant SNPs were identified using the function *--clump-range* in *PLINK.* High confidence (HC) genes were selected in a LD block for all the traits. To identify the candidate genes among the HC genes, the expression pattern for each gene was observed in publicly available transcriptome databases e.g., *BARLEX* (Colmsee et al., 2015) using the IPK BRIDGE portal (https://bridge.ipk-gatersleben.de/#snpbrowser) and ePlant browser (http://bar.utoronto.ca/eplant_barley/) (Thiel et al., 2021).

4.3.5 TILLING for HvRAP gene

HvRAP (*HORVU.MOREX.r2.3HG0193400*)—barley homolog of an *Arabidopsis* gene (*AT2G31890*)—consists of three exons and is of 2,815 bp. The RAP domain of this gene extends from 582aa to 639aa. The primers used for TILLING covering the RAP domain were RAP.F: GCCGTGTCCATGATACAGAC and RAP.R: GCAGTTGCTAATTTCCTTTCTG that resulted in the amplicon length of 1,726 bp (817–2,093 bp of the genomic sequence).

The *Hor*TILLUS population was used for the *TILLING* analysis (Szarejko et al., 2017; Szurman-Zubrzycka et al., 2018). The mutational screening was performed on 6,144 M_2 plants. The eight-fold pools of DNA isolated from individual M_2 plants served as templates for PCR reactions. To check the efficiency of product amplification, randomly chosen samples were visualized on 1% agarose gel dyed with Simply Safe (Cat. No. E4600-01, EURx). The further steps included formation of homo- and hetero-duplexes, followed by the digestion of hetero-duplexes with Celery Juice Extract (CJE) containing Cel I endonuclease (Szurman-Zubrzycka et al., 2017). The capillary electrophoresis was carried out to visualize the product in Fragment Analyzer (5200 Fragment Analyzer System, Agilent) with the use of CRISPR Discovery Gel Kit (cat. No. DNF-910-K1000CP, Agilent). The products were analyzed with ProSize *v*4.0.1.4 software (Advanced Analytical Technologies, Inc., Ames, IA, USA). After identifying a potential mutation in a pool, the DNA of individual M_2 plants forming the pool was used as a template to perform a PCR reaction. The PCR products were visualized on 1% agarose gel, followed by their sequencing by Genomed SA, Poland. Lastly, mutations were analyzed using CodonCode Aligner software (https://www.codoncode.com).

4.3.6 Chlorophyll measurements

We measured the chlorophyll in leaf and spike tissues using methanol-based protocol. Five accessions were selected from major PTD haplotypes. Three biological replicates were used for each measurement. Briefly, leaf discs and spike samples at MYP stage (*Thirulogachandar & Schnurbusch, 2021*) were collected and grounded. In the grounded samples, 1 ml methanol was added, followed by centrifugation at 1,300 rpm for two minutes. Supernatant was collected in clean Eppendorf tubes labeled as set 1. In the remaining supernatant, 1 ml of methanol was added again, followed by centrifugation and supernatant collection in Eppendorf tubes labelled as set 2. The final sample was prepared by taking 600 µl of supernatant from each set, making the final volume 1200 µl. From the final volume, 800 µl was used for the downstream spectrophotometer measurements. The measurements were taken at two wavelengths, i.e., 652 nm and 665.2 nm, and, following *Porra et al. (1989)*, the chlorophyll concentration was calculated as:

$$\text{Chl. conc.} = 24.23A^{652} + 3.23A^{665.2}$$

4.3.7 *Chlorophyll fluorescence analyses*

Chlorophyll fluorescence parameters were measured using Pulse-amplitude modulated technique using FluorCam device (Photon Systems Instruments, Brno, Czech Republic) installed in an automated phenotyping platform (Tschiersch et al., 2017). The chlorophyll fluorescence parameters were analyzed at 36 DAS. Maximum quantum yield of photosystem (PS) II was measured in dark-adapted plants (2 hours). Minimal fluorescence level (F_0) was determined by applying weak, pulsed measuring light (PAR ≤ 0.2 µmol photons $m^{-2}s^{-1}$) and a saturating light pulse (800msec; PAR: 4100 µmol photons $m^{-2}s^{-1}$) was applied to induce transiently maximal fluorescence level (F_m). Variable fluorescence (F_v) was calculated as $F_v = F_m - F_0$ and maximum quantum yield of PSII as F_v/F_m.

PSII operating efficiency (Φ_{PSII}) was measured for fully light-adapted plants. Plants were acclimated to a light intensity of 600 µmol photons $m^{-2}s^{-1}$ for at least 10 min in the acclimation tunnel followed by 60s illumination (PAR: 600 µmol photons $m^{-2}s^{-1}$) after moving into the chlorophyll fluorescence imaging chamber. Following determination of steady-state fluorescence level under actinic illumination (F_s), maximum fluorescence yield in the light-adapted state (F_m') was measured during 800 ms exposure to a saturating light flash (PAR: 4100 µmol photons m^{-2} s^{-1}). Φ_{PSII} was determined as:

$$\Phi_{PSII} = (F_m' - F_s) / F_m'$$

The absorption (Abs) of the leaves was recorded as Normalized Difference Vegetation Index (NDVI) using the reflection differences in the near-infrared (NIR) and red (RED) spectral region.

4.4 **Results**

4.4.1 A diverse and representative panel of six-rowed spring barleys is a resource for high-resolution genetic studies

The selected panel tapped the worldwide allelic diversity of barley and comprises landraces, cultivars and breeding material. The accessions were collected from five continents and 45 countries (Supplementary Table S4.1). We sequenced these accessions with the whole-genome shotgun sequencing approach that yielded ~50 million SNPs. After filtering, we obtained ~22 million high-quality SNPs distributed across the chromosomes (Figure 4.1a). The data set provides a high-resolution mapping resource for reliable downstream genotype-phenotype analyses.

To observe if accessions were grouped according to their origin (e.g., countries or continents) or genetic nature (e.g., landraces, breeding lines, etc.), we performed principal component analysis (PCA): the first two PCs jointly explained 82.36% whereas the first five PCs explained > 95% of the variation in the data set (Figure 4.1b). The PCA plot differentiated Asian accessions from accessions originated from other continents (Figure 4.1c). This shows that, although only six-rowed barley accessions harboring the *Ppd-H1* allele were selected, a population structure among the accession nevertheless exists. Also, our phylogenetic analysis—performed to check the evolutionary relationship among the investigated accessions—broadly categorized the accessions into Eastern, Western, and Ethiopian clades (Figure 4.1d). Our constructed phylogenetic tree depicted that Eastern and Western accessions were genetically different from each other, with a few Asian accessions being categorized into the Western clade. Generally, the Eastern clade contained most of the Asian accessions along with a few accessions

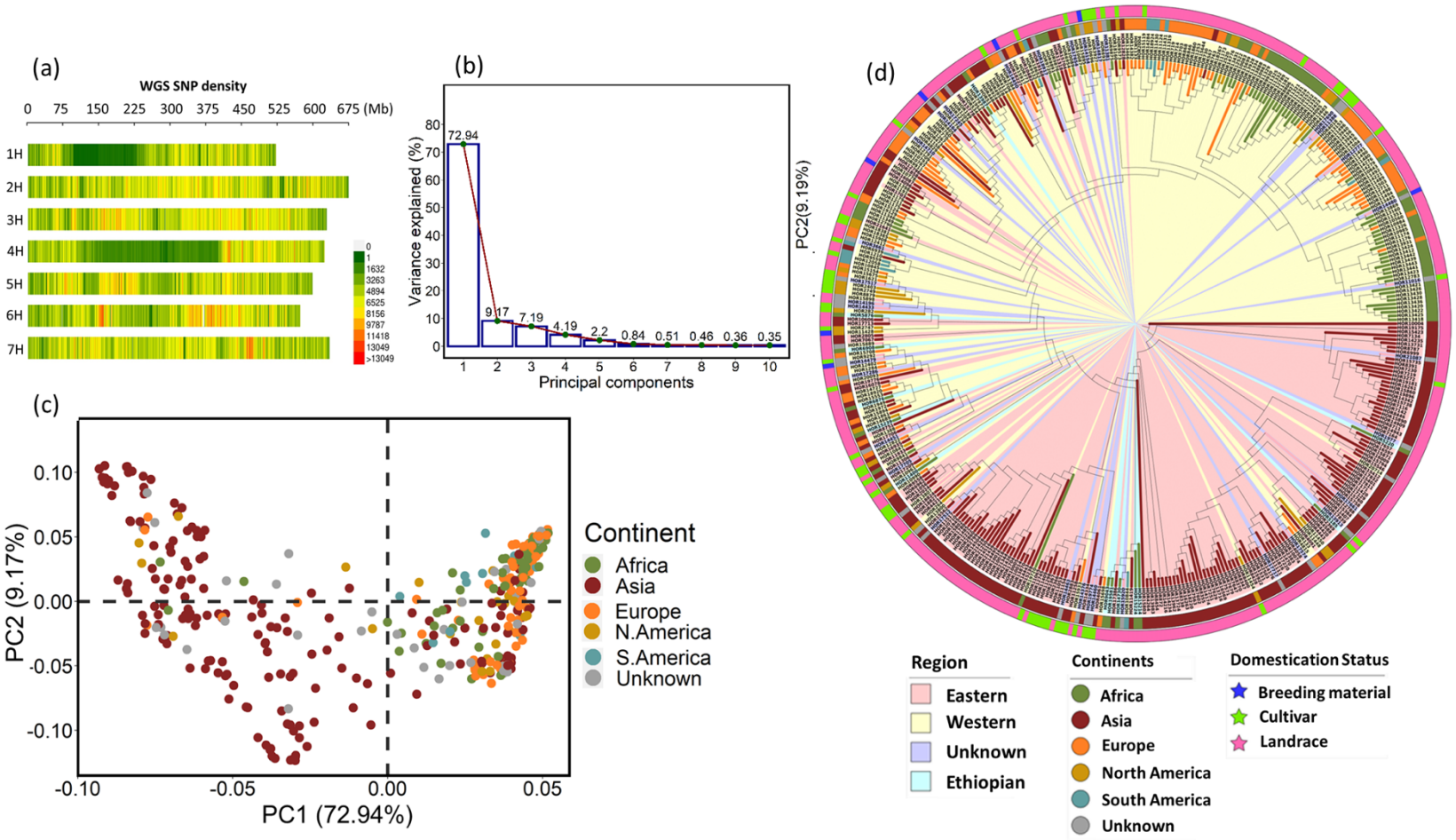


Figure 4.1 Diversity of the six-rowed barley panel used for GWAS. (a) Whole genome shotgun sequencing (WGS) SNP marker distribution across all seven barley chromosomes, (b) scree plot highlighting the first ten principal components, (c) PCA plot differentiating the accessions based on the continents, and (d) phylogenetic tree for the barley association panel.

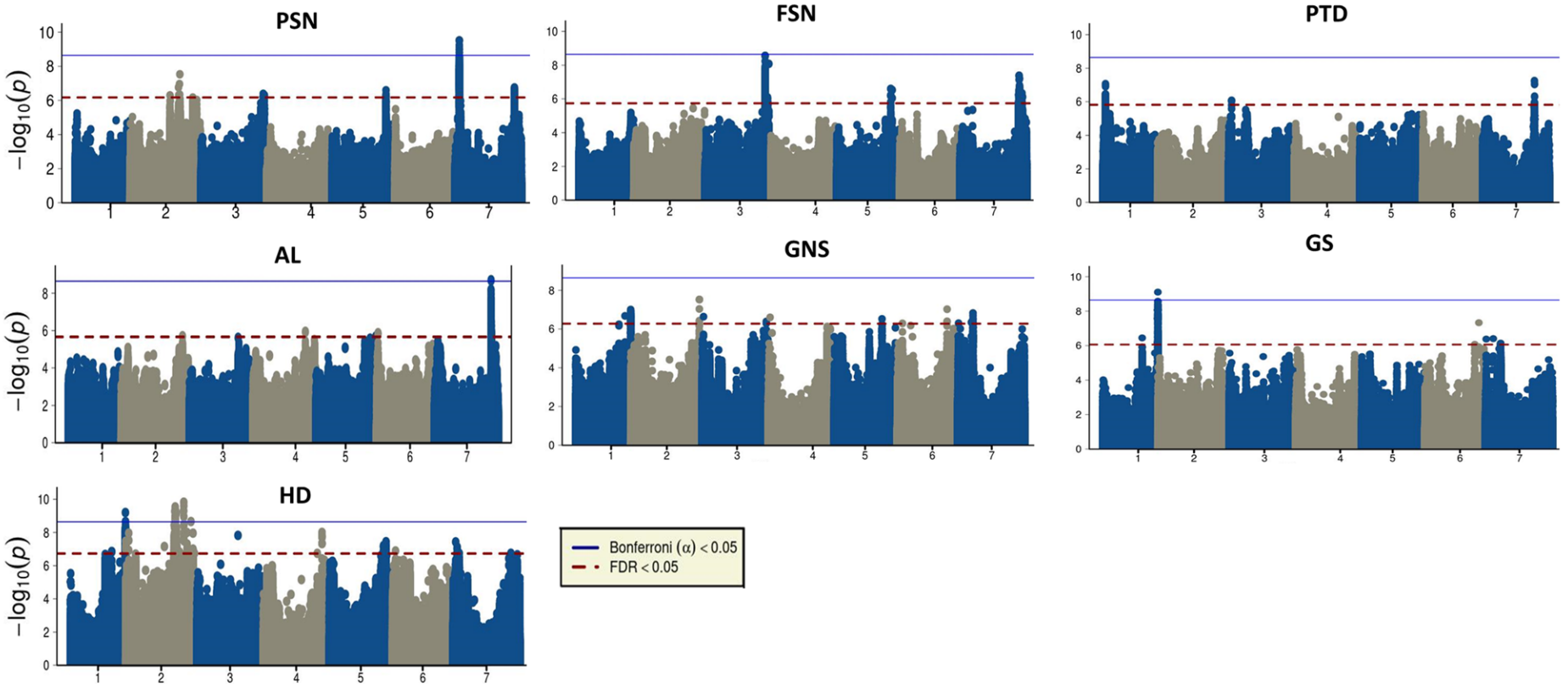


Figure 4.2 Summary of genome-wide association studies depicted as Manhattan plots where x-axis shows seven barley chromosomes and y-axis shows the marker significances as –log10 (P). A Bonferroni α level of 0.05 and false discovery rate (FDR) level of 0.05 was used to correct for multiple testing and identify significant markers. PSN = potential spikelet number, FSN = final spikelet number, PTD = pre-anthesis tip degeneration, AL = awn length, GNS = grain number per spike, GS = grain set, and HD = heading date.

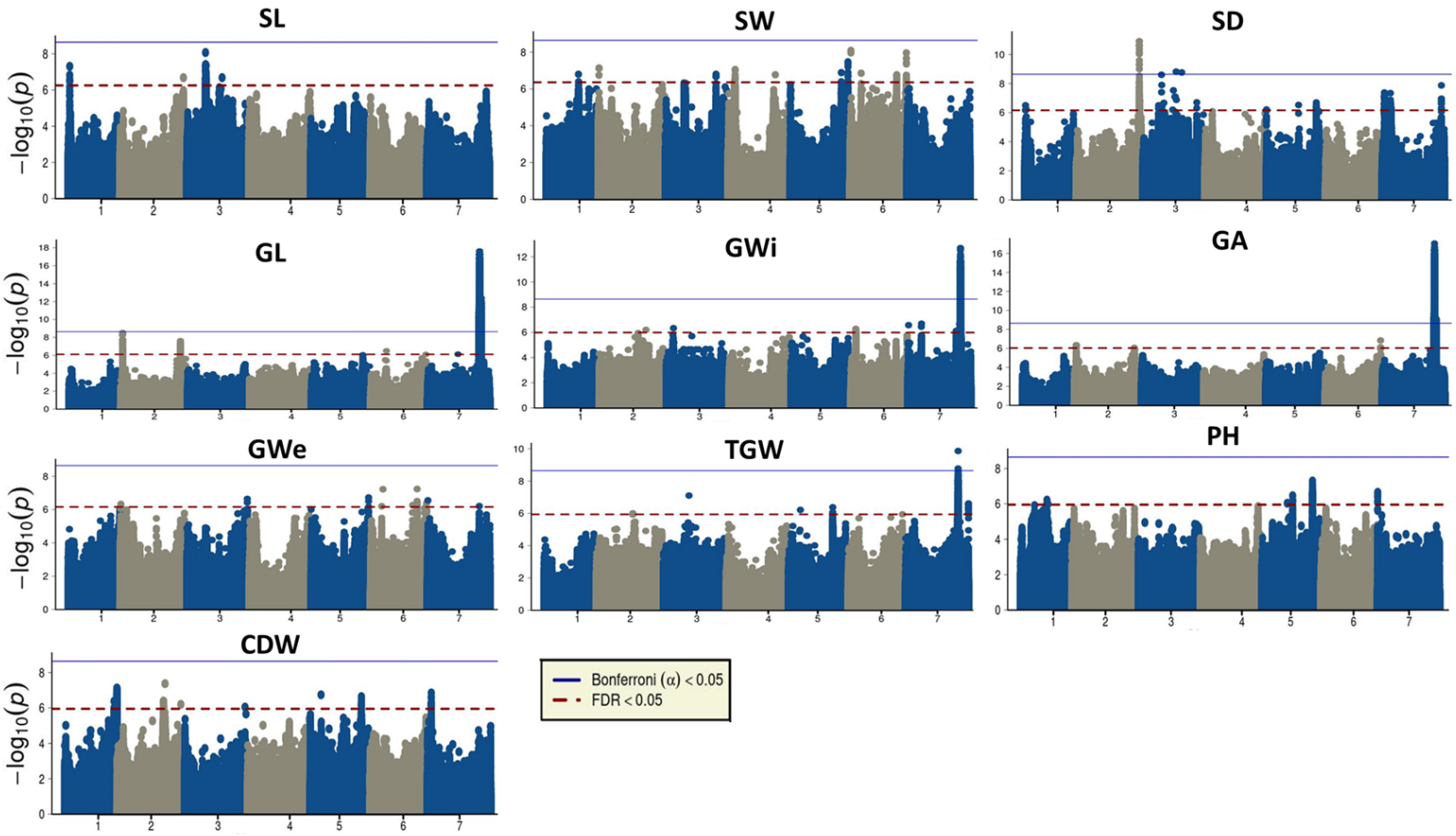


Figure 4.3 Summary of genome-wide association studies depicted as Manhattan plots where x-axis shows seven barley chromosomes and y-axis shows the marker significances as –log10 (P). A Bonferroni α level of 0.05 and false discovery rate (FDR) level of 0.05 was used to correct for multiple testing and identify significant markers. SL = spike length, SW = spike weight, SD = spike density, GL = grain length, GWi = grain width, GA = grain area, GWe = grain weight, TGW = thousand grain weight, PH = plant height, and CDW = culm dry weight.

We performed GWAS for 17 traits including spike PTD. The details about the phenotypic data collection and analyses were provided previously (Kamal et al., 2022a; Kamal et al., 2022b). Our GWAS detected 132 QTL for the 17 investigated traits—the marker significance values were plotted in the form of Manhattan plots (Figures 4.2-4.3), and the Bonferroni correction and FDR values for each trait are mentioned in Supplementary Table S4.3. The corresponding QQ plots (Supplementary Figure. S4.1) showed that the statistical models used to perform GWAS sufficiently accounted for spurious marker-trait associations. Thus, the significant QTL regions could be further dissected to uncover the underlying candidate genes.

4.4.2 QTL for spikelet number traits, awn length, heading date, grain number, and grain set

For PSN, we identified seven different significant loci with highest association in a 0.5kb interval on short arm of chr7H (peak: 40.85–41.37Mb; Figure 4.2, Supplementary Figure S4.2). All significant markers were in high LD (Supplementary Figure S4.3a) and harbored eight high-confidence (HC) genes (Supplementary Table S4.4). The most significant SNP was an intronic SNP within gene *HORVU.MOREX.r2.7HG0543420* (annotated as *MADS*-box transcription factor 5; ortholog of rice *OsMADS5*). For FSN, 12 QTL were detected with the most significant peak on chr3HL, which was in a tight LD interval of 442.6kb (peak: 573.05–573.49Mb; Figure 2, Supplementary Figures S4.2 and S4.4). All HC genes within significant regions for FSN are mentioned in Supplementary Table S4.5. Previously, a high correlation was observed between PSN and FSN, i.e., higher PSN formed during the early reproductive stages resulted in higher FSN (Kamal et al., 2022b). The relationship was also seen in the GWAS as three QTL on chr3HL, chr5HL, and chr7HL were shared between PSN and FSN (Figure 4.4a). The HC genes within the overlapping regions are mentioned in Supplementary Tables S4.4 and S4.5. Among these genes, *HORVU.MOREX.r2.3HG0263320* (annotated as agenet domain containing protein) was highly expressed in the young and developing inflorescence tissues and *HORVU.MOREX.r2.7HG0602460* (FBD-associated F-box protein) showed highest mRNA expression during inflorescence and grain development (Supplementary Figure S4.5). These listed HC genes need further in-depth studies to better relate them to the genetics of spikelet number formation in barley.

Eight QTL were detected for AL and (Figure 4.2, Supplementary Figure S4.2), of which the most significant association was identified on chr7HL. Further LD analysis, however, revealed two different QTL on chr7HL (Supplementary Figure S4.6). The first QTL region ranged from 559.93 to 562Mb (~2Mb) and harbored 16 HC genes, whereas the second spanned from 565.05 to 566.22Mb (~1.17Mb) and harbored 13 HC genes (Supplementary Table S4.6). A known gene for awn development, *short awn 2* (*lks2; HORVU.MOREX.r2.7HG0602340* annotated as Glycosyltransferase) was present in this interval (Yuo et al., 2012; Huang et al., 2021). We also observed that the chr7H QTL was shared between PSN, FSN and AL. Similarly, another overlapping region between FSN and AL (peak: 561.73–562Mb) was detected (Figure 4.4a). We previously conjectured the consequences of rapid awn elongation along the spikelet development Kamal et al. (2022a). This overlapping QTL region, hence, further points toward the common underlying genes during spikelet and awn development.

We detected 19 different QTL for HD (Figure 4.2, Supplementary Figure S4.7), of which three QTL were identified either closer to or within the genes *HvELF3* on chr1H (peak: 519.74–522.33Mb), *HvCEN* on chr2H (peak: 581.27–581.64Mb) and *VRN-H3/HvFT1* on chr7H (peak: 39.63–39.65Mb). We additionally detected another HD peak close to *Vrs1* on chr2H (581.26–581.64Mb). The remaining novel 16 QTL for HD are shown in Supplementary Figure S4.7, and corresponding LD region and HC genes are shown and listed in Supplementary Figure S4.8 and Supplementary Table S4.7, respectively. Among these novel QTL, seven genes were orthologous to *Arabidopsis* flowering time genes (Zhang, 2004; Ito et al., 2012; Li et al., 2014; Lee et al., 2019; Chu et al., 2022)—the details of the orthologous genes are mentioned in Supplementary Table S4.7. For GNS, the most significant association was detected on chr2HL (Figure 4.2). The high-LD 290 kb region harbored nine HC genes (Supplementary Figure S4.9; Supplementary Table S4.8) and the QTL on other chromosomes are highlighted in Figure S4.10 and the genes are listed in Supplementary Table S4.8. Four QTL—on chr1H, 4H, 6H, and 7H—were detected for GS with strongest association on chr1HL, harboring 20 HC genes (Figure 4.2, Supplementary Figure S4.11; Supplementary Table S4.9). The chr1HL QTL (521.66–522.02Mb) was shared between GS and GNS with same common peak SNP (Figure 4.4b). Furthermore, the same chr1H QTL region was also detected for HD (Figure 4.5). We conclude from these results that the duration of vegetative and reproductive growth phase (dictated by the flowering time genes) largely decides the number of spikelet primordia produced on an immature spike. For all remaining traits, i.e., SL, SW, SD, GL, GWi, GA, GWe, TGW, PH, and CDW, significant QTL regions are depicted in Figure 4.3, Supplementary Figure S4.7, and S4.9. The genomic regions in LD for these traits are highlighted in Supplementary Figures S4.12-S4.19 and the corresponding HC genes in Supplementary Tables S4.10-S4.19.

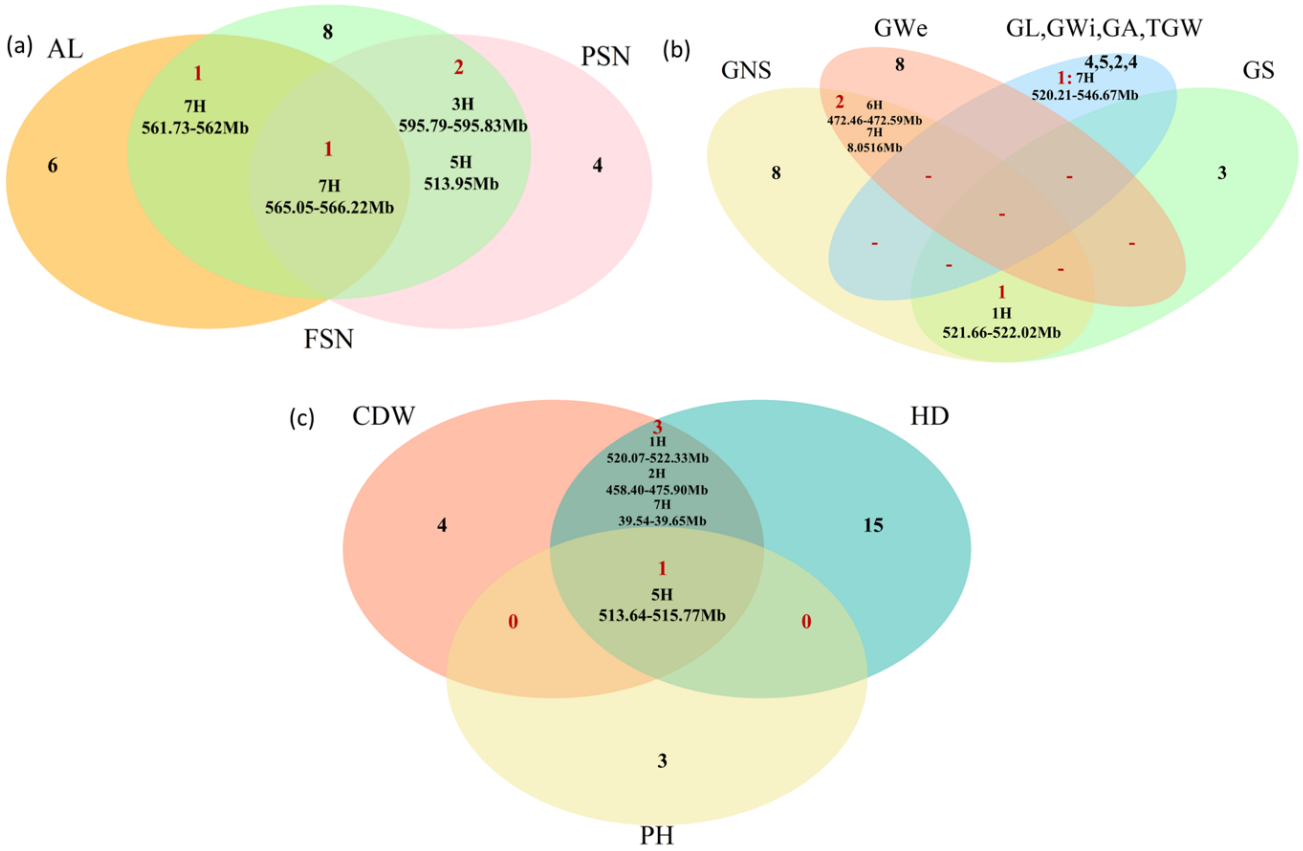


Figure 4.4 Shared QTL among the investigated traits. (a) venn diagram for shared QTL between PSN, FSN and AL, (b) venn diagram for shared QTL between GNS, GL, GWi, GA, GWe, GS, and TGW. The numbers 4,5,2,4 in the blue ellipse are the corresponding unique QTL for GL, GWi, GA and TGW, respectively. (c) venn diagram for shared QTL between HD, PH and CDW. The unique QTL for a particular trait are mentioned in black and the red numbers highlights the shared QTL. PSN = potential spikelet number, FSN = final spikelet number, AL = awn length, GNS = grain number per spike, GL = grain length, GWi = grain width, GA = grain area, GWe = grain weight, GS = grain set, TGW = thousand grain weight. HD = heading date, PH = plant height and CDW = culm dry weight.

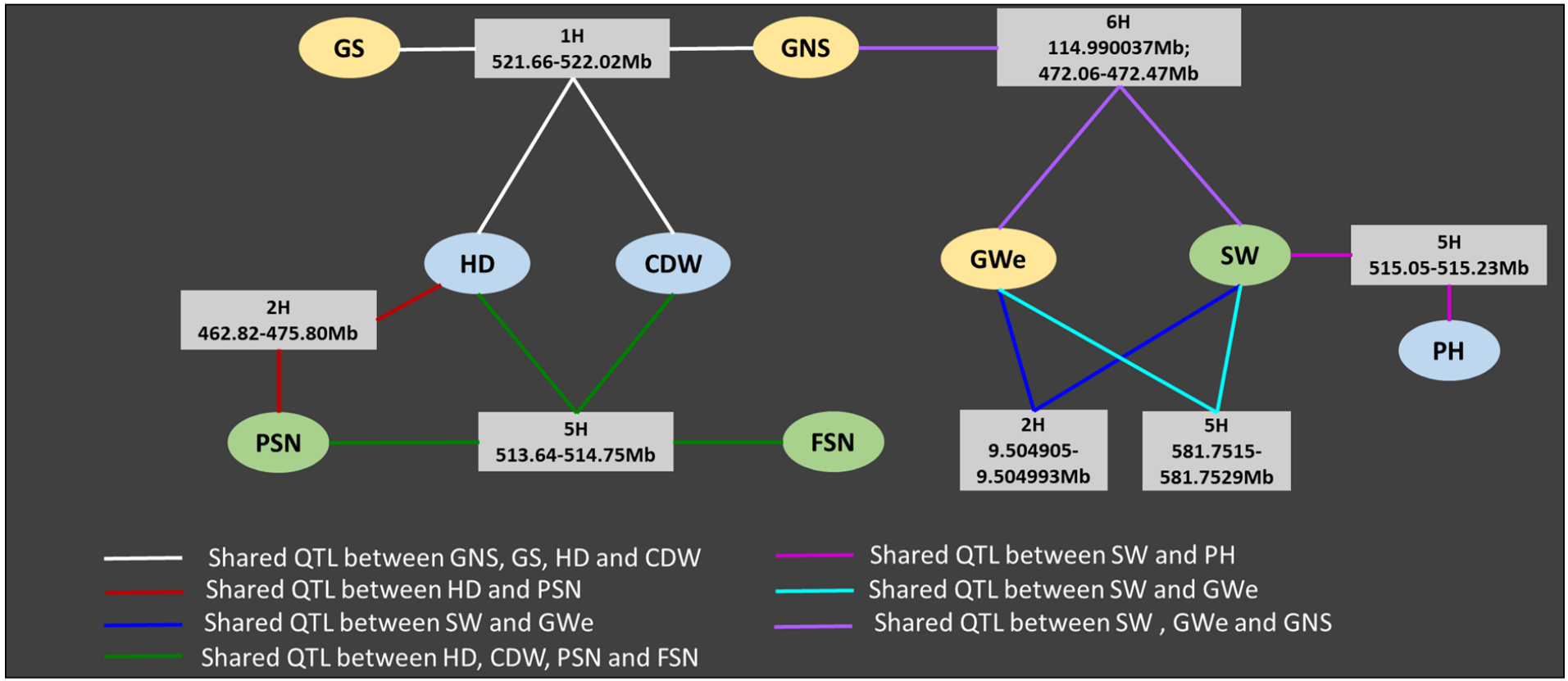


Figure 4.5 QTL based network among the spike, grain and shoot traits. PSN = potential spikelet number, FSN = final spikelet number, SW = spike weight, GNS = grain number per spike, GWe = grain weight, GS = grain set, HD = heading date, PH = plant height, and CDW = culm dry weight.

4.4.3 QTL network analysis unfolds shared genetic architecture among barley traits

We constructed a network plot based on QTL associated with more than one class of trait (Figure 4.5). Path analysis among spikelet number and HD traits showed that HD has a positive impact on both PSN and FSN (Kamal et al., 2022b). In this study, we identified two overlapping QTL between PSN, FSN, and HD (Figure 4.5). The first QTL overlapping between HD and PSN was on chr2H and extended from 462.82 to 475.80Mb. Although the interval was large, the whole locus was in tight LD. The second overlapping QTL on chr5H was among PSN, FSN and HD as well as CDW (peak: 513.64–514.75Mb), suggesting the impact of this locus on HD and, eventually, on the duration of vegetative and reproductive growth phases and biomass accumulation (Figure 4.5). HD and CDW also held a common QTL with GNS and GS. These results point towards an influence of heading date on the spikelet production, their development, as well as grain formation. We detected two overlapping QTL between SW and GWe and two overlapping QTL on chr6H between GNS, SW and GWe (Figure 4.5). In addition to the shared QTL analyses, we performed a systematic review of literature to determine the relative positions of the QTL identified in this study. Most of the QTL identified in our study are novel except for the regions around the known genes *HvELF3, VRN-H2, VRN-H3, AP2,* and *NUD*. The detection of known genes provided positive analytical proof for the strength of our GWAS analyses. The most significant markers for each trait are presented in the form of a reference genotype-phenotype map along with previously reported genes on each barley chromosome (Figure 4.6).

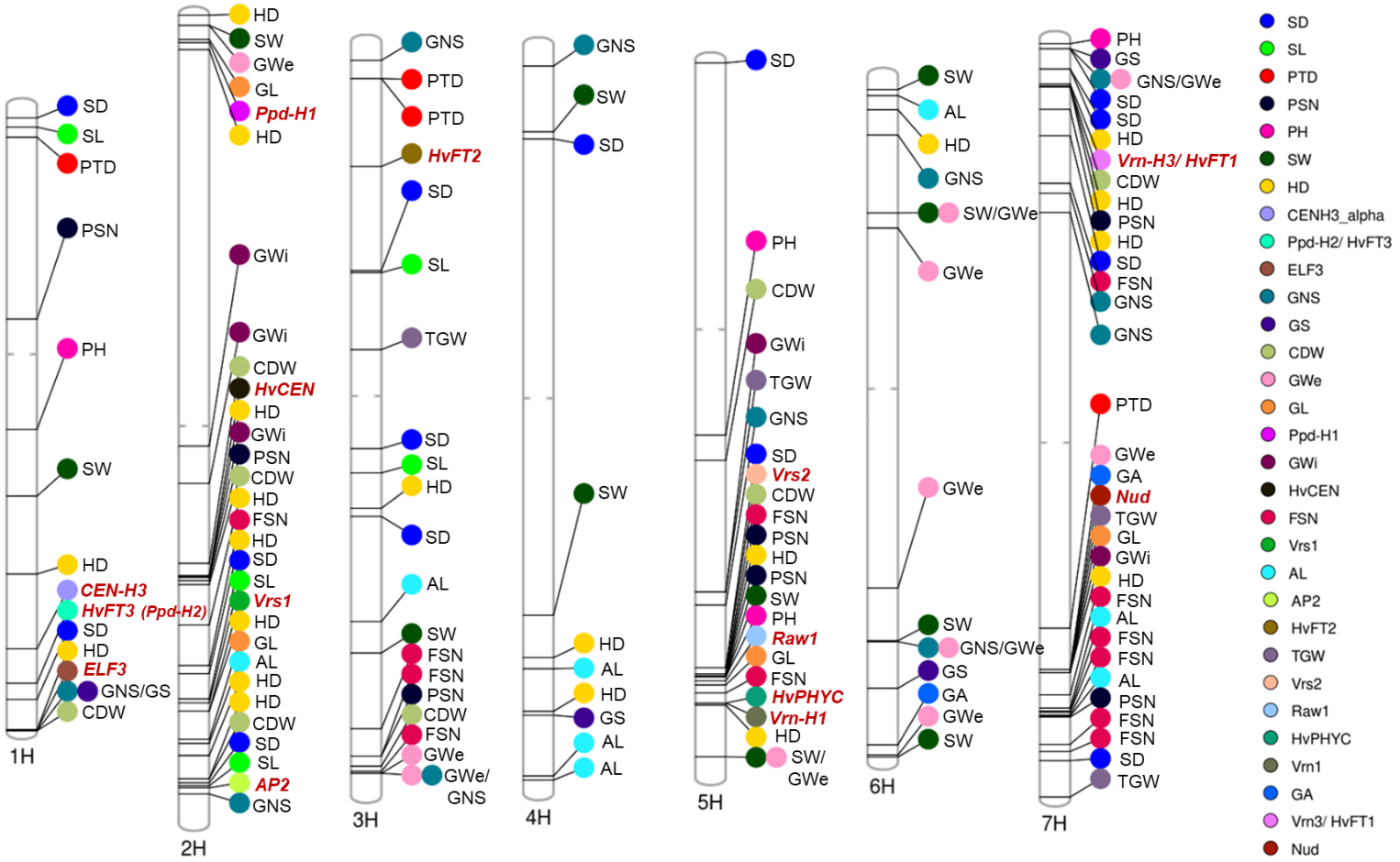


Figure 4.6 A reference genotype-phenotype map highlighting the most significant trait-associated markers (represent by the respective trait) on each chromosome aligned to the reference sequence of cv. Morex (RefSeq *v2*) along with previously reported genes (marked in bold red color).

4.4.4 *GWAS reveals unique genetic architecture of barley spike PTD*

Our GWAS analyses for spike PTD revealed three major-effect QTL on chromosomes 1H, 3H and 7H (Figure 4.2, Supplementary Figure S4.2), of which the most significant spike PTD-associated SNP explained 18%, 14.32%, and 16.19% genotypic variance, respectively. Based on LD, the chromosome 1H (chr1H) QTL spanned around 1.51Mb (Supplementary Figure S4.21) and harbored 20 HC genes (Table 4.1)—the majority of which were annotated either as Glucan endo-1, 3-beta-glucosidase or Pentatricopeptide repeats (PPR). PPR proteins participate in various RNA-associated processes, both in mitochondria and chloroplasts (Barkan & Small, 2014). Our annotation analyses showed that the PPR gene cluster on chr1H is the restorer of fertility-like genes (RFL or Rf, as described in *Triticeae*) or RNA processing factors (RPFs, as described in *Arabidopsis*) (Table 4.2). Barley has three high-density RFL gene clusters located on chromosomes 1H, 2H and 6H and the region on chromosome 1H is divided into three sub-clusters, with largest sub-cluster2 containing eight genes. The PTD chr1H QTL region belongs to sub-cluster2 and from the 10 detected PPR genes, five were previously annotated as *HvRFL4*, *HvRFL24*, *HvRFL8*, *HvRFL9* and *HvRFL10* (Table 4.2) (Melonek et al., 2019).

The 7H spike PTD-QTL spanned ~2Mb (491.71–493.76Mb) and contained 10 HC genes (Table 4.1): of these, only two, viz., *HORVU.MOREX.r2.7HG0592130* (transmembrane protein) and *HORVU.MOREX.r2.7HG0592180* (GTP-binding protein family) were expressed in the inflorescence meristem (Thiel et al., 2021). The transcript level of the gene annotated as GTP-binding nuclear protein showed a strong decrease starting from the lemma primordium stage and was also homologous to *Arabidopsis RAN* genes ((Pham et al., 2022; Qin et al., 2022); Supplementary Figure 4.22a). No orthologs were found for the transmembrane protein.

The spike PTD QTL on 3H spanned only 136.8kb (30.42–30.56Mb) and harbored three HC genes (Table 4.1). It should be noted that our initial GWAS analyses for all traits, including spike PTD, involved 416 accessions. Another round of GWAS was performed for spike PTD using only 409 accessions—because seven accessions were removed based on the outlier analyses (Supplementary Figure S4.23). Interestingly, the second round of GWAS unveiled the marker significance peak on 3H as being the strongest ($-\log_{10}(P) = 11.01$; Supplementary Figure S4.24a) of all three spike PTD-QTL. However, the GWAS results for all other traits remained unchanged (Supplementary Figures S4.25-4.28). Delineating the significant 3H region further showed that all peak SNPs were in tight LD (Supplementary Figure S4.24b) and were located in an intergenic region between *HORVU.MOREX.r2.3HG0193400* (AT2G31890-like protein) and *HORVU.MOREX.r2.3HG0193410* (Transposon Ty3-G Gag-Pol polyprotein

Table 4.1 List of high confidence genes on chromosome 1H, 3H and 7H for pre-anthesis tip degeneration in barley

Gene IDs		Chr.	Start position	End position	Gene length	Annotation
HORVU.MOREX.r3.1HG0011380	*HORVU.MOREX.r2.1HG0008430*	1H	24,531,408	24,534,253	2846	Zinc finger family protein
HORVU.MOREX.r3.1HG0011400	*HORVU.MOREX.r2.1HG0008450*	1H	24,554,200	24,556,680	2481	Pentatricopeptide repeat-containing protein
HORVU.MOREX.r3.1HG0011440	*HORVU.MOREX.r2.1HG0008490*	1H	24,581,624	24,583,389	1766	Glucan endo-1,3-beta-glucosidase, putative
HORVU.MOREX.r3.1HG0011450	*HORVU.MOREX.r2.1HG0008510*	1H	24,745,383	24,748,605	3223	Pentatricopeptide repeat-containing protein
HORVU.MOREX.r3.1HG0011570	*HORVU.MOREX.r2.1HG0008520*	1H	24,763,847	24,765,602	1756	Glucan endo-1,3-beta-glucosidase, putative
HORVU.MOREX.r3.1HG0011510	*HORVU.MOREX.r2.1HG0008580*	1H	24,785,334	24,788,247	2914	Pentatricopeptide repeat-containing protein
HORVU.MOREX.r3.1HG0011540	*HORVU.MOREX.r2.1HG0008620*	1H	24,846,745	24,847,836	1092	Glucan endo-1,3-beta-glucosidase, putative
HORVU.MOREX.r3.1HG0011560	*HORVU.MOREX.r2.1HG0008630*	1H	25,048,788	25,051,358	2571	Pentatricopeptide repeat-containing protein
HORVU.MOREX.r3.1HG0011590	*HORVU.MOREX.r2.1HG0008650*	1H	25,189,794	25,192,376	2583	Pentatricopeptide repeat-containing protein
HORVU.MOREX.r3.3HG0221850	*HORVU.MOREX.r2.1HG0008680*	1H	25,258,609	25,259,334	726	Tyrosine N-monooxygenase
HORVU.MOREX.3.1HG0011660	*HORVU.MOREX.r2.1HG0008690*	1H	25,262,307	25,263,542	1236	Glucan endo-1,3-beta-glucosidase, putative
NA	*HORVU.MOREX.r2.1HG0008700*	1H	25,263,705	25,264,106	402	Glucan endo-1,3-beta-glucosidase, putative
HORVU.MOREX.r3.1HG0011680	*HORVU.MOREX.r2.1HG0008710*	1H	25,492,390	25,494,972	2583	Pentatricopeptide repeat-containing protein
NA	*HORVU.MOREX.r2.1HG0008720*	1H	25,499,738	25,500,982	1245	Glucan endo-1,3-beta-glucosidase, putative
HORVU.MOREX.r3.1HG0011700	*HORVU.MOREX.r2.1HG0008730*	1H	25,506,911	25,509,364	2454	Pentatricopeptide repeat-containing protein
NA	*HORVU.MOREX.r2.1HG0008740*	1H	25,511,133	25,514,319	3187	Transposon Ty3-I Gag-Pol polyprotein
NA	*HORVU.MOREX.r2.1HG0008770*	1H	25,581,705	25,582,259	555	Pentatricopeptide repeat superfamily protein
HORVU.MOREX.r3.1HG0011730	*HORVU.MOREX.r2.1HG0008810*	1H	25,681,782	25,684,364	2583	Pentatricopeptide repeat-containing protein
HORVU.MOREX.r3.1HG0011790	*HORVU.MOREX.r2.1HG0008840*	1H	26,045,146	26,046,936	1791	Pentatricopeptide repeat-containing protein
HORVU.MOREX.r3.1HG0011820	*HORVU.MOREX.r2.1HG0008870*	1H	26,050,330	26,052,105	1776	Protein ABIL1
HORVU.MOREX.r3.3HG0232920	*HORVU.MOREX.r2.3HG0193390*	3H	30,418,454	30,420,684	2231	Pyrimidine-specific ribonucleoside hydrolase
HORVU.MOREX.r3.3HG0232930	*HORVU.MOREX.r2.3HG0193400*	3H	30,421,279	30,424,093	2815	AT2G31890-like protein
HORVU.MOREX.r3.3HG0232960	*HORVU.MOREX.r2.3HG0193420*	3H	30,569,270	30,570,529	1260	GDSL esterase/lipase
HORVU.MOREX.r3.7HG0713840	*HORVU.MOREX.r2.7HG0592090*	7H	491,667330	491,668928	1599	Cytochrome P450
HORVU.MOREX.r3.7HG0713850	*HORVU.MOREX.r2.7HG0592110*	7H	491,854408	491,854839	432	multidrug resistance-associated protein 2

HORVU.MOREX.r3.7HG0713850	*HORVU.MOREX.r2.7HG0592110*	7H	491,854408	491,854839	432	multidrug resistance-associated protein 2
HORVU.MOREX.r3.7HG0713870	*HORVU.MOREX.r2.7HG0592120*	7H	491,856949	491,858556	1608	Zinc finger homeobox protein 3
NA	*HORVU.MOREX.r2.7HG0592130*	7H	491,862,463	491,863,940	1478	transmembrane protein
HORVU.MOREX.r3.7HG0713920	*HORVU.MOREX.r2.7HG0592170*	7H	492,400,780	492,402,531	1752	Subtilisin-like protease
HORVU.MOREX.r3.7HG0713940	*HORVU.MOREX.r2.7HG0592180*	7H	492,465,956	492,467,995	2040	GTP-binding nuclear protein
HORVU.MOREX.r3.7HG0713960	*HORVU.MOREX.r2.7HG0592200*	7H	492,810,093	492,812,188	2096	Cytochrome P450
HORVU.MOREX.r3.7HG0714040	*HORVU.MOREX.r2.7HG0592230*	7H	493,295,517	493,297,642	2126	Cytochrome P450
HORVU.MOREX.r3.7HG0714100	*HORVU.MOREX.r2.7HG0592250*	7H	493,507,000	493,510,233	3234	26S protease regulatory subunit, putative
HORVU.MOREX.r3.7HG0714120	*HORVU.MOREX.r2.7HG0592260*	7H	493,813,281	493,820,291	7011	Eukaryotic translation initiation factor 3 subunit I

2

Table 4.2 Annotation results for the Pentatricopeptoide repeat gene cluster on chromosome 1H for pre-anthesis tip degeneration.

Gene IDs (Morex *v1*)	**Gene IDs (Morex *v2*)**	**Gene IDs (Morex *v3*)**	**NCBI database (Johnson et al., 2008)**	**Rice Genome Annotation Database (Kawahara et al., 2013)**	**TAIR database (Lamesch et al., 2012)**	**Known Gene mentioned in Melonek et al. (2019)**
HORVU1Hr1G010970	*HORVU.MOREX.r2.1HG0008450*	*HORVU.MOREX.r3.1HG0011400*	Protein Rf1	LOC_Os4g28300	At1g63130; RPF6	*HvRFL4*
HORVU0Hr1G031920	*HORVU.MOREX.r2.1HG0008510*	*HORVU.MOREX.r3.1HG0011450*	Protein Rf1	LOC_Os4g28300	At1g63130; RPF6	*HvRFL24*
HORVU0Hr1G030540	*HORVU.MOREX.r2.1HG0008580*	*HORVU.MOREX.r3.1HG0011510*	Protein Rf1	LOC_Os4g28300	At1g63130; RPF6	NA
HORVU1Hr1G011150	*HORVU.MOREX.r2.1HG0008630*	*HORVU.MOREX.r3.1HG0011560*	Protein Rf1	LOC_Os4g28300	At1g62670; RPF2	NA
HORVU1Hr1G011160	*HORVU.MOREX.r2.1HG0008650*	*HORVU.MOREX.r3.1HG0011590*	Protein Rf1	LOC_Os4g28300	At1g63130; RPF6	*HvRFL8*
HORVU1Hr1G011250	*HORVU.MOREX.r2.1HG0008710*	*HORVU.MOREX.r3.1HG0011680*	Protein Rf1	LOC_Os4g28300	At1g62670; RPF2	NA
HORVU1Hr1G011240	*HORVU.MOREX.r2.1HG0008730*	*HORVU.MOREX.r3.1HG0011700*	Protein Rf1	LOC_Os4g28234	At1g62670; RPF2	NA
NA	*HORVU.MOREX.r2.1HG0008770*	NA	Protein Rf1	LOC_Os4g28300	At1g12700; RPF1	NA
HORVU1Hr1G011300	*HORVU.MOREX.r2.1HG0008810*	*HORVU.MOREX.r3.1HG0011730*	Protein Rf1	LOC_Os4g28300	At1g63130; RPF6	*HvRFL9*
HORVU1Hr1G011400	*HORVU.MOREX.r2.1HG0008840*	*HORVU.MOREX.r3.1HG0011790*	Protein Rf1	LOC_Os4g28300	At1g12620 RPF8	*HvRFL10*

Rf1 is restorer of fertility 1, RPF is RNA Processing Factor, RFL is Restorer of fertility like and both rice genes, LOC_Os4g28300 and LOC_Os4g28234 are annotated as Rf1, mitochondrial precursor, putative.

(Supplementary Figure S4.24c). The distribution of accessions according to the most significant SNP (*s3_30426221*; T/C) on 3H demonstrated that the accessions harboring the alternate (C; *n*=84) allele were associated with significantly (*P* <0.001) higher spike PTD as compared to those with the reference (T; *n*=326) allele (Figure 4.7a). Furthermore, we checked the transcript expression profile of *HORVU.MOREX.r2.3HG0193400* in the inflorescence meristem (IM, curated from Thiel et al. (2021)) and three immature spike sections, viz., apical, central, and basal over the spike developmental time course (Figure 4.7b-c). Transcript levels of *HORVU.MOREX.r2.3HG0193400* in the IM started to decrease after the stamen primordium stage, while the lowest values were observed for the apical spike part. Our phylogenetic and reciprocal BLAST analyses of *HORVU.MOREX.r2.3HG0193400* revealed that it is a single copy gene with already characterized orthologs in *Arabidopsis* (*AtRAP*) and rice (*Albino Leaf 1*, *AL1;* Figure 4.7d). In rice, *AL1* is involved in chloroplast biogenesis (Zhang et al., 2016) and in *Arabidopsis* in the maturation of chloroplast 16S rRNA (Kleinknecht et al., 2014). This gene encodes for an octotricopeptide repeat protein with an α-helical RNA-binding (RAP) domain. The underlying barley RAP (*HvRAP*) protein has a RAP domain of 55 amino acids (580–635 aa). In total, 21 haplotypes were identified for HvRAP (Supplementary Table S4.20) where first two haplotypes (haplotype1 (*n*=280) and haplotype2 (*n*=51)) accounted for majority of the panel. Sanger-sequencing of *HvRAP* six amino acid substitutions between haplotype1 and haplotype2, one of which resulted in a substitution in a conserved amino acid within the RAP domain (A1751G, from Isoleucine to Methionine in protein position 584; i.e. I584M) (Figure 4.7e). Majority of the haplotype 2 accessions belonged to Europe and were landraces, mainly acquired after 1980 (Supplementary Figure S4.29, Supplementary Table S4.21). All haplotype-2 (G1751 or protein variant 2, M584) carrying accessions also harbored the alternate (minor) allele (C) related to the peak GWAS marker and were in tight LD. Other plant species, such as rice, wheat, *Brachypodium distachyon*, maize, sorghum, and *Arabidopsis*, possess the I584 protein variant, indicating strong structural protein conservation among different plant species for this amino acid substitution while all other five substitutions are quite flexible even between wheat and barley (Supplementary Table S4.22). These two major protein variants also resulted in significant differences for other GY component traits: accessions harboring variant-2 (M584, *n*=51) showed more spike PTD with lower FSN that ultimately translated into bigger grains with more TGW (Figure 4.8a–i; raw haplotype data in Supplementary Table S4.21). Since loss of *AL1* in rice led to chlorophyll defects and albinism, we similarly assessed the chlorophyll concentration in accessions from both variants. Importantly, chlorophyll concentrations were lower for accessions harboring variant-2—both in leaves and immature spike tissues (Figure 4.8j-k), suggesting a lower photosynthetic capacity of these genotypes. These results also point towards a similar function of the orthologs of *HvRAP* as in rice and *Arabidopsis,* i.e., in proper chloroplast development and chlorophyll production.

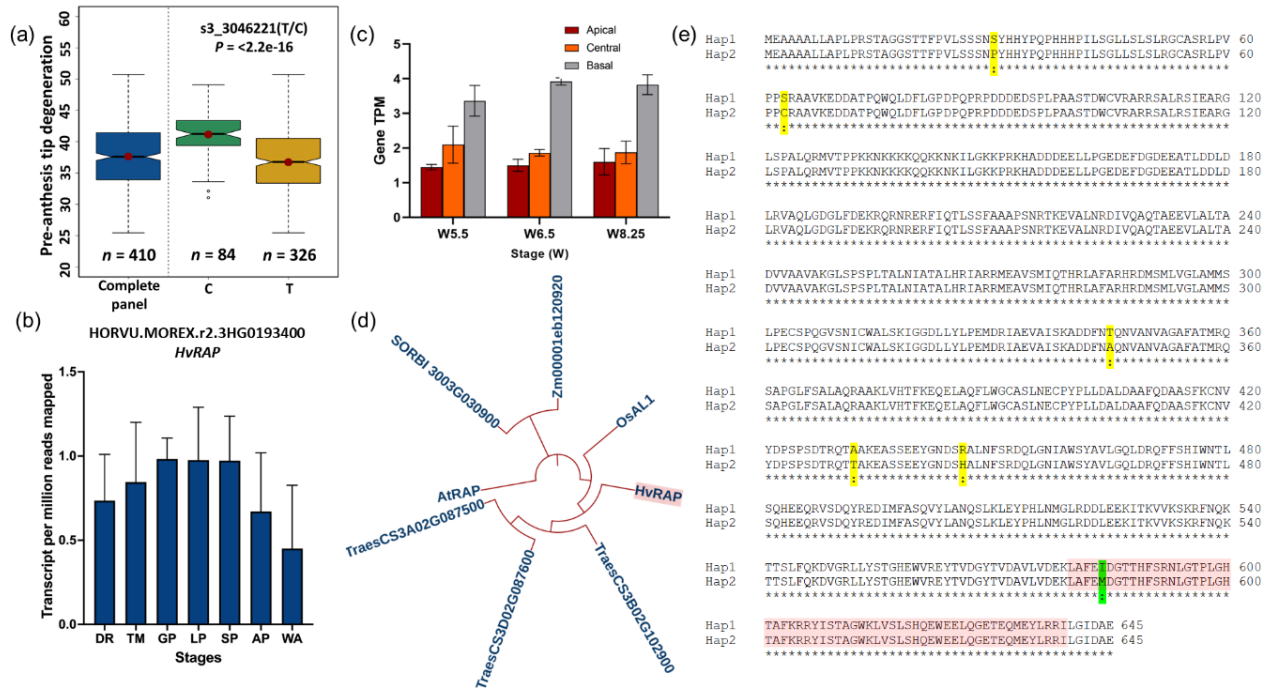


Figure 4.7 Allelic distribution, expression and sequencing analyses reveal association of *HvRAP* with pre-anthesis tip degeneration (PTD) in barley. (a) Allelic distribution for the most significant SNP (s3_30426221) associated with PTD where "T'' and "C" denote reference and alternate alleles, respectively. (b) Expression of the *HvRAP* (*HORVU.MOREX.r2.3HG0193400*) in the inflorescence meristem: DR = double ridge, TM = triple mound, GP = Glume primordium, LP = lemma primordium, SP = stamen primordium, AP = awn primordium, WA = white anther (data retrieved from ePlant browser (http://bar.utoronto.ca/eplant_barley/) (Thiel et al., 2021)) (c) Expression of *HvRAP* in different immature spike sections. (d) Phylogenetic analysis of RAP-domain containing proteins. (e) Two major *HvRAP* haplotypes, including six amino acid substitutions as revealed by Sanger sequencing. Red rectangle box is the RAP domain (580-639aa). Five substitutions outside the RAP domain are highlighted in yellow and amino acid substitution within the RAP domain is highlighted in green.

Since spike PTD is a derived trait from PSN and FSN, both correlation and path analyses accounted for associations among these traits as well as its correlations with AL, GNS, GA and TGW (Kamal et al., 2022a; Kamal et al., 2022b). Based on the previous results, we postulated that it would be possible to detect, at least in parts, shared QTL between PTD and the other traits. However, no overlapping QTL was detected between spike PTD and spikelet number traits, i.e., PSN and FSN, implying that spike PTD appears to have a unique genetic architecture.

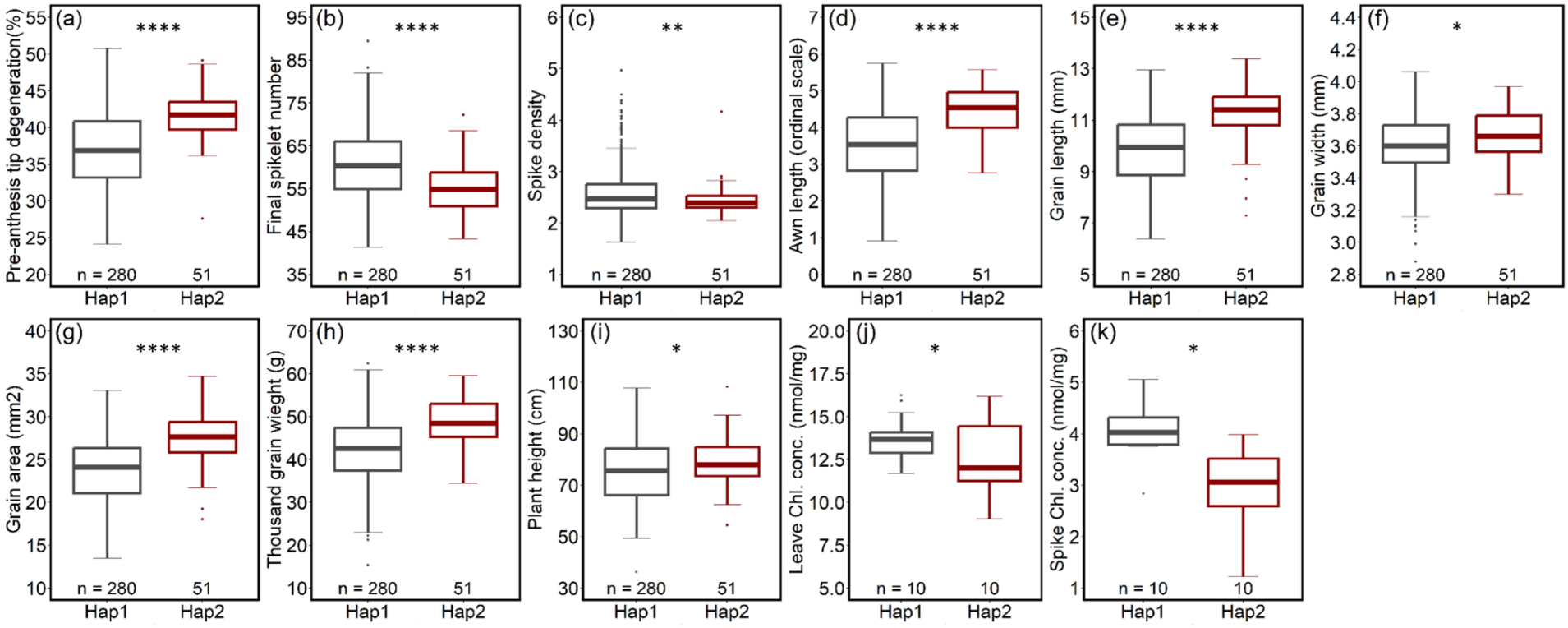


Figure 4.8 Trait comparison and chlorophyll concentration with respect to major haplotypes for pre-anthesis tip degeneration. (a) trait comparison for pre-anthesis tip degeneration, (b) trait comparison final spikelet number, (c) Trait comparison for spike density, (d) trait comparison for awn length, (e) trait comparison for grain length, (f) trait comparison for grain width, (g) trait comparison for grain area, (h) trait comparison for thousand grain weight, (i) trait comparison for plant height, (j) chlorophyll measurement in leave tissues and (k) chlorophyll concentration in immature spike tissues. The comparison was performed using Welch's t-test in Prism Graphpad software. The *p-value* significance level are **** = <0.0001, ***= 0.0001-0.001, ** = 0.001-0.01, * = 0.01-0.05 and ns = >0.05.

4.4.5 *HvRAP affects chloroplast structure, photosystem II efficiency, and spikelet and grain number in barely*

Since we discovered two naturally occurring functional protein variants for *HvRAP*, it was considered a strong candidate gene underlying spike PTD in barley. We thus screened for and identified *HvRAP* TILLING lines in the genetic background of the two-rowed cultivar Sebastian (Szurman-Zubrzycka et al., 2018). In total, six TILLING lines, *viz*., *hvrap.a* (A463V), *hvrap.b* (S394F), *hvrap.c* (D412N), *hvrap.d* (D522N), *hvrap.f* (V561I) and *hvrap.g* (V574M) (details of the lines are mentioned in Supplementary Table S4.23) were retrieved, and M_3 grains were grown; however, two mutant lines, *hvrap.b* (S394F) and *hvrap.d* (D522N) did not germinate and, consequently, could not be evaluated further. Since it was not possible to verify if the lack of viability (germination) in these two lines was due to the homozygosity of the mutations, we predicted their putative effects *in silico*. At least for the first S394F substitution, a deleterious effect was predicted by the PROVEAN software. For all the remaining lines, heterozygous plants were selected and grown in the next generation (i.e., M_4). *hvrap.a* and *hvrap.c* lines were discarded due to discrepancies between genotypic data and phenotype of the plants, i.e., in these lines, majority of the plants grew similar to cv. Sebastian plants but were genotypically labeled as mutant. We detected similar discrepancies in the other lines but to a smaller extent. Nevertheless, the presence of the strongest leaf phenotype (i.e. chlorotic leaves; similar to rice and *Arabidopsis* phenotypes) was observed for the *hvrap.g* allele carrying a missense mutation six amino acids upstream of the RAP domain and, therefore, was selected for further phenotypic, physiological, and histological analyses. For *hvrap.g* plants, we observed a decrease in light absorption, maximum quantum yield, and PSII efficiency compared to the wildtype cv. Sebastian (Figure 4.9). False color images also showed visible differences between cv. Sebastian and *hvrap.g* plants (Figure 4.9). Light microscopic and ultrastructural analysis by transmission electron microscopy revealed that *hvrap.g* (WT) chloroplasts appear broadly similar to cv. Sebastian (Figure 4.10a-f), but the grana stack partly consisted of more thylakoids than cv. Sebastian (Figure 4.10e-f). The mutant plant, however, contained fewer and smaller chloroplasts with varying shapes (Figure 4.10g-h, j-k). In addition, the chloroplasts of mutant plants often harbored fewer granule stacks (Figure 4.10h, k), although some possess a very high number of thylakoid membranes (Figure 4.10h-l). The observed structural changes indicate that this mutation caused severe chloroplast alterations, with most likely impaired photosynthesis. Since spike PTD dictates the number of spikelets retained on the spike after the degeneration process, we were interested to know the effect of *HvRAP* on spikelet and grain number. Three genotypic classes within the *hvrap.g* family, i.e., *HvRAP.g*, *HvRAP.g* (heterozygous) and *hvrap.g* were analyzed for spikelet and grain number differences. There was no significant difference for spikelet number at heading between WT and heterozygous. But, mutant plants had significantly lower spikelet number, which was further translated into lower grain number. Moreover, there was a significant difference for grain number per spike between WT and heterozygous spikes (Figure 4. 11). These results suggest that *HvRAP* not only affects chloroplast biogenesis and PSII efficiency but also spikelet and grain development in barley.

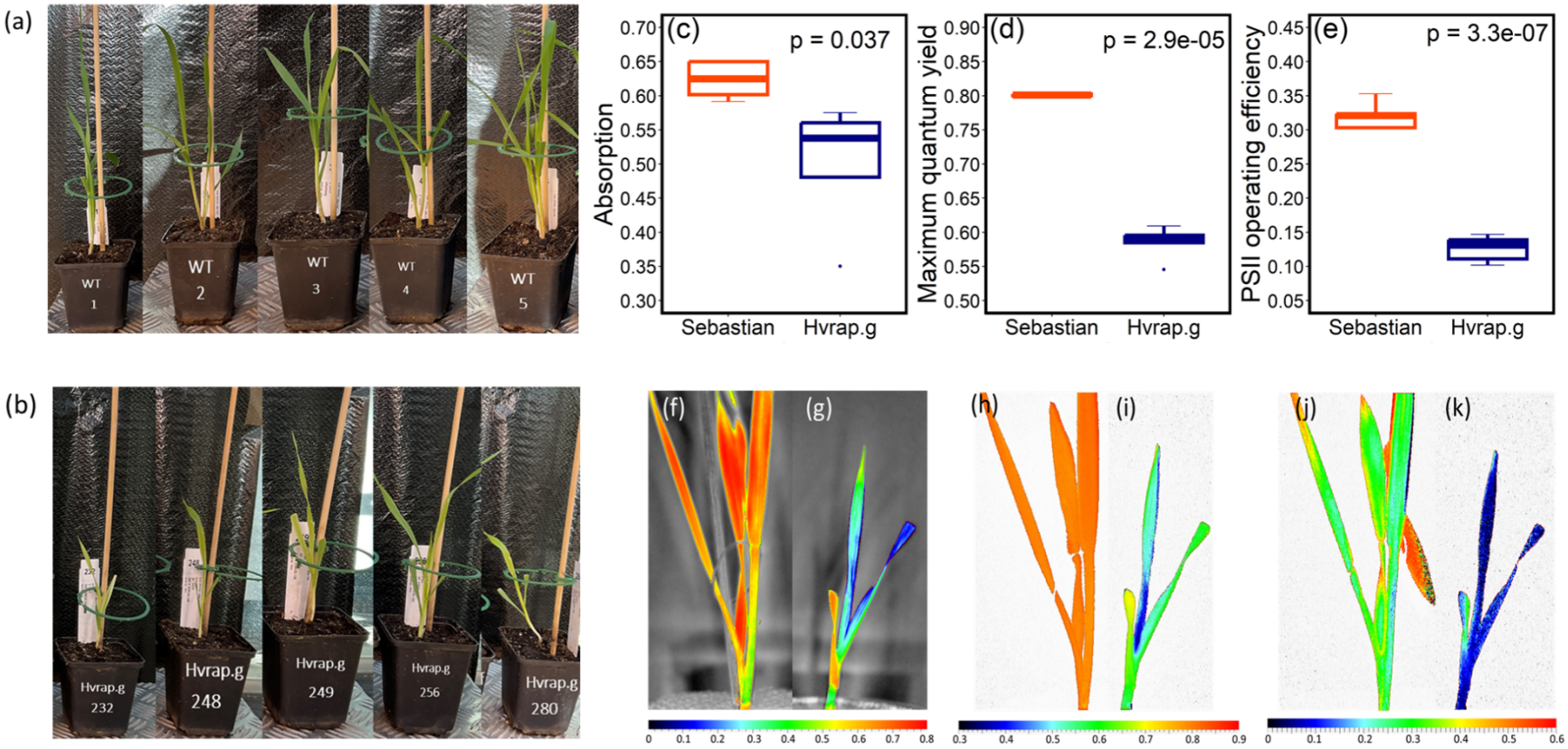


Figure 4.9 FluorCam results for Sebastian (WT) and *hvrap.g* TILLING lines. (a) Wildtype plants (Sebastian) phenotype, (b) *hvrap.g* lines phenotype, (c-e) FluorCam results with respect to absorption, maximum quantum field (chlorophyll fluorescence) and photosystem II operating efficiency; (n=5), (f-k) changes in incident light intensity caused by variation of distance between FluorCam panel and plants. The images f, h and j are false color images for cv. Sebastian and images g, i and k are false color images for *hvrap.g* plants. (f-g) false color images for absorption (Normalized Difference Vegetation Index – NDVI) results, (h-i) false color images for maximum quantum yield results and (j-k) false color images for photosystem II operating efficiency differences.

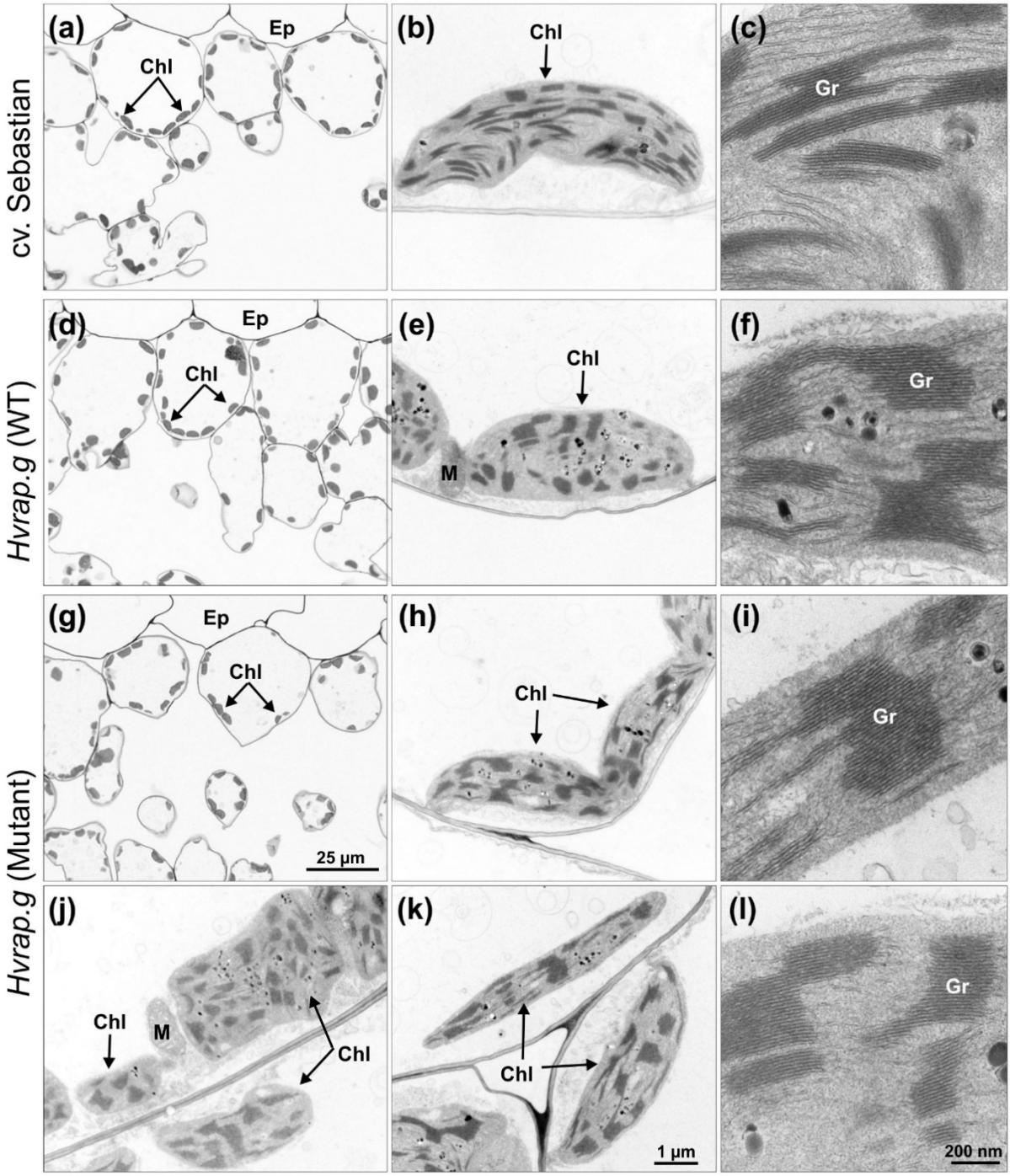


Figure 4.10 Light microscopy (LM) and transmission electronic microscopic (TEM) analysis of chloroplasts of the 3rd leaf of cv. Sebastian and *hvrap.g* TILLING lines. LM (a, d, g) and TEM (b-c, e-f, h-l, k-l) images of cv. Sebastian, *Hvrap.g* (WT) and *Hvrap.g* (Mutant). Chl, chloroplast; Ep, epidermis; Gr, grana stack; M, mitochondria. The sample for LM and TEM were taken from the central part of the third leaf

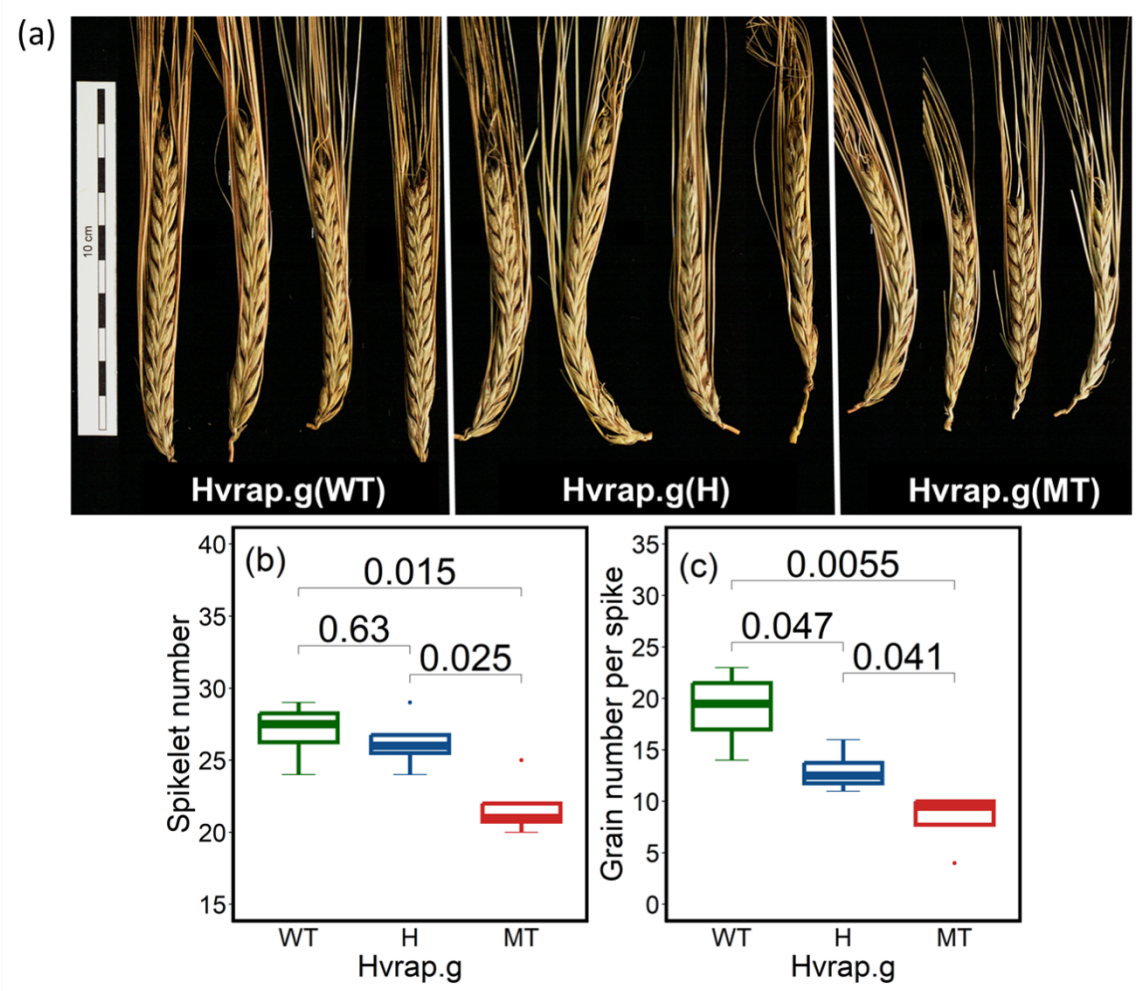


Figure 4.11 Spikelet and grain number analyses of plants from the segregating family of the *hvrap.g* line. (a) spike phenotypes (b) spikelet number differences and (c) grain number phenotypes for three different classes of plants within the family (n =4).

4.5 **Discussion**

Using GWAS, we identified three main-effect QTL for spike PTD, the *HvRFL* gene cluster on chr1H, *HvRAP* on chr3H, and *HvRAN* on chr7H. Most photosynthetic complexes are encoded in the nucleus, and transport systems are needed to convey these complexes from the cytosol into the chloroplast. Importantly, these transport systems include helical repeat proteins (HRP) superfamily members (Macedo-Osorio et al., 2021), such as tetratricopeptide repeat (TPR) proteins, pentatricopeptide repeat proteins (PPR; e.g. found as the detected *HvRFL* gene cluster), and octotricopeptide repeat proteins (OPR, e.g., *HvRAP*). Our study demonstrated that two of these protein classes have been associated with spike PTD, suggesting an important role of HRPs during grain number determination in barley.

We identified two major haplotypes for *HvRAP* where haplotype2 is a naturally occurring, *reduced-function* allele (rather than a *loss-of-function* allele) that remained among accessions but with a lower frequency—perhaps because of its subtle phenotypic effect. We showed that natural sequence variation of accessions carrying *HvRAP*-haplotype-2 had higher PTD, lower chlorophyll in immature spikes, and less FSN (Figure 4.8); similarly, *hvrap.g* TILLING lines showed lower grana number and ruptured chloroplasts that eventually led to lower photosynthetic efficiency and lower grain number per spike in mutant plants (Figures 4.10-4.11). Our results suggest that *HvRAP* is also involved in chloroplast development. In fact, an acropetal chlorophyll gradient exists in an immature spike, where the apical part shows less or no chlorophyll (no rachis greening); however, the gradient increases towards the spike base (more rachis greening) (Huang et al., 2023; Shanmugaraj et al., 2023). Recently, it was postulated that rachis greening is required to sustain floral growth while decreased or no rachis greening in the apical spike part leads to degeneration of apical spikelets (Huang et al., 2023). Since we observed lower chlorophyll content in immature spikes of haplotype2 accessions and the *hvrap.g* TILLING mutant, we hypothesize that the energy supply required to sustain the developing spikelet primordia could be low, which resulted in increased spike PTD. Lowered or insufficient protein functionality of HvRAP may result in a reduced transport of essential chloroplast development genes from nucleus to chloroplast. Thus, it can be assumed that the observed chloroplast impairment and decreased chlorophyll content in immature spikes of haplotype2 accessions and the *hvrap.g* TILLING mutant might result from the hindered transportation of essential genes for chloroplast development; however, this hypothesis requires in depth downstream analyses for further validation. Moreover, it has to be noted that both *HvRAP* alleles are substitutions of conserved amino acids but did not lead to very severe phenotypes, such as albinism or lethality. In fact, both alleles are most likely *HvRAP reduced-function* alleles, primarily compromising greening of the leaf and spike tissues. Why an allele that mainly produces slightly paler green tissues was maintained in the population remains to be shown. Plausible explanations could be related to ameliorated adaptation to high planting density, high irradiation (e.g. by utilizing albedo) or arid climatic conditions (Gommers et al., 2013; Cutolo et al., 2023). Majority of the haplotype2 accessions were landraces that belonged to Southern Europe (n=23) and North Africa (n=14) and were mainly acquired after 1950s (Supplementary Figure S4.29 and Supplementary Table S4.21)—this suggest an advantage of these accessions under Mediterranean conditions and modern agronomic practices as such

better adaptation to higher planting density. As, haplotype2 accession have higher PTD, one could further study the effect of plant density on the extent of PTD.

Secondly, the presence of the *RFL* gene cluster putatively suppresses the effect of mitochondrial genes causing cytoplasmic male sterility (CMS, where plants fail to produce viable pollens) (Melonek et al., 2019). The Rf proteins block the accumulation and expression of CMS-causing mRNA inside mitochondria and restore fertility by ensuring proper pollen development (Melonek & Small, 2022). In the context of spike PTD, it was reported that after the cessation of the IM dome, visible signs of PTD were seen in developing anthers (Shanmugaraj et al., 2023), i.e., anther primordia degenerate before other tissues within florets. The fertility restoration mechanism of the *RFL* genes is well established; however, the mechanism in which RFL proteins might play a role during barley anther development and spike PTD still needs to be elucidated.

We reported another candidate on chr 7H, *HORVU.MOREX.r2.7HG0592180* (annotated as GTP-binding nuclear proteins; Table 4.1) annotated as RAN protein. The small GTPase, RAN, stands for Ras-related nuclear protein and is required for mitotic progression and nucleocytoplasmic transport (Nielsen, 2020). Phylogenetic analyses of *HvRAN* revealed that it is present in the same clade as four *Arabidopsis RAN* genes and two wheat genes, viz., *TraesCS7B02G254000* and *TraesCS7D02G349900*, collectively known as *TaRan2* (Supplementary Figure S4.23b) (Choudhury et al., 2021). The *AtRAN* genes are known to be involved in various developmental processes such as cell division (Chen et al., 2011), mitotic activity (Ciciarello et al., 2007), and gametophytic development, i.e., defects in *RAN* affect both pollen mitosis I as well as megaspores mitosis (Qin et al., 2022). Also, wheat *TaRan2* genes have the highest expression in the reproductive tissues, especially in microspores (Choudhury et al., 2021). Since the onset of spike PTD in barley has been linked with programmed cell death of anther primordia in developing florets (Shanmugaraj et al., 2023), *HvRAN* may represent a reasonable candidate to investigate its role during cell cycle progression of the shoot apical meristem as well as pollen and ovary mitosis within florets.

The duration of heading has an impact on both initiation of spikelet primordia and spikelet development (Alqudah & Schnurbusch, 2014; Digel et al., 2015; Gol et al., 2017) and both the initiated number of spikelet primordia and spikelet number at heading decide the extent of PTD. It was reported that early-maturity mutants have lower PSN than Bowman, and plants with delayed flowering have higher PSN (Huang et al., 2023). Interestingly, the direct effect of HD on PSN and FSN was shown using path analysis, but no direct effect of HD was observed on PTD (Kamal et al., 2022b). In the present study, we detect common QTL region for HD with grain and spikelet number traits, i.e., chr1H QTL for GNS, GS and HD and chr5H QTL for PSN, FSN and HD (Figure 4.5), but no overlapping QTL was detected for PTD and HD. A chr1H QTL shared between GNS, GS and HD (Figure 4.5) harbored three genes: *HORVU.MOREX.r2.1HG0078540*, *HORVU.MOREX.r2.1HG0078570,* and *HORVU.MOREX.r2.1HG0078680*—the first two being annotated as Zinc finger protein *CONSTANS* while the third as Zinc finger CCCH domain protein. These genes showed highest expression during grain development—both at 5 and 15 days after pollination (Supplementary Figure S4.20;Supplementary Table S4.7-4.8). *CONSTANS* plays an important role in the photoperiod pathway and acts between the circadian clock and meristem identity genes

(Griffiths et al., 2003; Campoli et al., 2012; Campoli & von Korff, 2014; Luccioni et al., 2019; Zhou et al., 2021). Within HD significant region, we also detected a major flowering regulator, *Early Flowering 3* (*ELF3*; *HORVU.MOREX.r2.1HG0078390*). However, *ELF3* gene was neither present within the significant genomic interval nor in the LD block with the most significant SNP for GNS and GS. For that reason, the role of Zinc finger protein *CONSTANS* could be studied to associate it with GNS and GS.

Even though we detected an overlapping region between HD and PSN, we focused on the candidate gene on chr7H (as it was the most significant association for PSN). We identified a candidate gene *HORVU.MOREX.r2.7HG0543420* annotated as MADS-box5 on chr7H—(Supplementary Table S4.4). MADS-box genes are involved in spikelet and flower development (Jeon et al., 2000; Cui et al., 2010; Zhang et al., 2013; Hu et al., 2015). Rice *OsMADS5*, for instance, controls spikelet morphogenesis and regulates floral meristem determinacy (Wu et al., 2018). *HORVU.MOREX.r2.7HG0543420* was explicitly expressed in the developing inflorescence and rachis tissues, and a gradual increase in its expression from double-ridge (DR) to white anther stage coincided with the linear increase in spikelet primordia starting from DR stage (Supplementary Figure S4.3b-c). Hence, the available clues make *HORVU.MOREX.r2.7HG0543420* a suitable candidate to study its role in barley spikelet primordia development.

Taken together, in this study, we provide first genetic insights into the possible processes affecting apical spikelet primordia degeneration in barley while narrowing down to promising candidate genes (*HvRAP* and *HvRAN*). We hypothesize functions/pathways in which newly identified QTL could act to alter the respective traits. Since, in Western Europe, spring barley holds a large market share, further fine mapping of these QTL/genes will reveal the mechanism behind the functioning of causative genes and may aid in altering barley GY—a premium goal of any breeding program.

Chapter 5. GENERAL DISCUSSION

Pre-anthesis developmental phase is considered an essential phase in barley growth and development. In combination with the environmental cues, it dictates the number of spikelet primordia initiated on an immature spike and the number of spikelets eventually developing into grains. However, pre-anthesis tip degeneration (PTD)—a phenomenon in which apical spikelet primordia degenerate, thereby decreasing the yield potential of a plant—poses a bottleneck in increasing barley GY to meet the increasing demand. It is, therefore, important to understand the evolutionary reason/s behind floret degeneration, e.g., why plants degenerate floral organs in the first place. Because the natural genotypic variation for PTD has not been exploited before, along with its relationship with agronomically important traits, we studied PTD comprehensively in this dissertation and shed light on the underlying genetic factors.

5.1 **Evolutionary reason behind spikelet/floret degeneration**

Plants have well-developed regulatory mechanisms to cope with short- and long-term environmental perturbations, especially when these perturbations coincide with a costly energy-demanding reproduction stage (Sergio et al., 2011). The growth and reproduction in plants are responsive to both resource as well as non-resource environmental cues such as light quality and photoperiod. The plasticity of a given component could be expected to be negatively related to the cost/benefit associated with the production of that component (Sadras & Slafer, 2012). Hence, plants preferentially produce those components that have lower production costs. It was reported that the cost of floret primordia production is small (Stephenson, 1981), and the plant could afford to overproduce energetically inexpensive floret primordia. However, when highly resource-intensive floret development begins, the number of floret primordia that become fertile is adjusted to the actual assimilate availability (Sadras & Slafer, 2012). Therefore, the decisive reason behind the floret reduction mechanism (distal developing floret primordia starve to death) is natural selection that favors uniform allocation of resources under constraints imposed by plant morphology and genomic conflict (Sadras & Denison, 2009).

Sakai (2007) explained the overproduction of the floret by studying the overproduction of ovules within flowers. Ovule overproduction could actually be beneficial as it allows for selective ovule abortion. Non-uniform seeds are produced if there exists a variance in the resource absorption ability among the fertilized ovules. Therefore, by overproducing the ovules, females could select fertilized ovules with similar resource absorption rates, resulting in uniform seed size. By selectively aborting the fertilized ovules, plants can reduce resource competition among the other developing ovules.

In barley, the same principle could be applied to explain the overproduction of spikelet primordia. Like wheat floret primordia, the overproduction of spikelet primordia in barley could be due to their energetically inexpensive production. However, due to the difference in the resource absorption ability of spikelet primordia during the juvenile spike growth, higher resource absorption takes place for basal and middle spikelets and lower for apical spikelets, leading to the initiation of degeneration. A wide range of degenerated spikelets was also observed in the barley panel used in this dissertation (Kamal et al., 2022b). It could be hypothesized that genotypes with higher resource absorption should have higher spikelet

survival and vice-versa; however, the resource absorption capacity in different genotypes must be checked to affirm the hypothesis.

5.2 **Proxy traits for PTD and importance of panel selection to investigate PTD**

In this study, correlation analyses among PTD and 16 other investigated traits revealed a moderate positive (r = 0.3−0.38) correlation of AL, GL, GA, and TGW with PTD; whereas, traits like SL (r = -0.19) showed low negative, SD (r = -0.32) and GNS (r = -0.3) showed moderately negative, and FSN (r = -0.6) showed a high negative correlation with PTD. A non-significant correlation was observed for PSN and other remaining traits. Path analysis, interestingly, showed a positive impact of PSN on PTD. (Kamal et al., 2022a; Kamal et al., 2022b). Using PTD in the breeding program can help realize the actual yield potential of barley. As mentioned above, PSN is difficult to phenotype manually; nevertheless, the closest trait that could be used as the proxy for PTD is FSN. Because of the high negative correlation between FSN and PTD, it could be extrapolated that spikes with higher spikelet number at heading have fewer spikelet primordia degenerated during the degeneration process and *vice-versa*. Also, FSN showed a high correlation with PSN, this indicates that spikes with high FSN also show generally higher PSN. FSN alone explained 63% and 35% of variation in PSN and PTD (Kamal et al., 2022b)—therefore, FSN can therefore be used to predict the PTD variation. Another trait whose role must be assessed in PTD is AL. We observed a moderate positive correlation of PTD with AL. Asian accessions with lower AL showed lower PTD, whereas African and European accessions with higher AL show high PTD. This further points to a potential role of awn development in affecting the extent of degeneration. Nonetheless, precautions must be taken before using AL as a proxy trait. Further experiments with isogenic lines only differing for AL could be useful to shed light on this relationship and the possibility of using AL as a proxy trait for PTD.

The population structure in the panel was controlled by selecting the six-rowed accessions, or, in other words, by eliminating the effect of row-type genes. However, population structure still existed because the accessions belonged to different continents such as Asia, Africa, Europe, North America and South America. The presence of a diverse population in this study lead to significant difference among the five subgroups (based on the continents) for most of the investigated traits. Most of the difference were observed for African, Asian, and European accessions (Kamal et al., 2022a; Kamal et al., 2022b). For GWAS, the genotypic diversity is a pre-requisite, and, thus, the studied panel presented a rich source of allelic variation to capture suitable loci associated with traits and study the corresponding genes. Nevertheless, from a practical standpoint, it can be advised to capture “within-continent” or even “target-environment” variation—this will help find the genotypes that are relatively closer to the existing germplasm in running breeding programs.

5.3 **Image analyses pipelines and genomic selection may help exploit PTD in breeding and genetics studies**

The genetic nature of the natural variation present for barley PTD in a panel of 417 six-rowed barley accessions were studied in this dissertation. PTD varied from 13%–51%, i.e., in some barley accession, ~50% of the potential yield is lost due to the degeneration process. In wheat, floret abortion and fertility or overall spike fertility are extensively studied as a putative trait to increase the wheat yield to meet the increasing demand. However, floret abortion as a trait has

not been introduced into the breeding programs probably due to its laborious phenotyping. Similarly, in barley breeding programs, much of the focus is on the final grain set instead of the pre-anthesis development phases that decide the number of survived spikelets. Phenotyping PTD in barley is a laborious task as its calculation depends on two different developmental time points. The first one at MYP stage to calculate the maximum number of spikelet primordia initiated on the spike (potential spikelet number; PSN) and another at heading to know the number of survived spikelets (final spikelet number; FSN). Because of the lengthy process of data collection, previous studies in cereals were confined to a small number of genotypes that did not represent the breadth of genotypic variation for PTD.

To calculate the variation for PTD in the present study, thousands of plants were manually dissected which was both labor- and time-intensive. Therefore, to investigate PTD in a larger panel, efficient methods or strategies must be followed to reduce the effort needed to phenotype PTD at the microscopic level. One of such strategies include development of the state-of-the-art deep learning models and image analyses algorithms for automation of PSN and FSN calculations. A variety of image analyses software analyzing different plant parts are available (Hund et al., 2019; Sun, D et al., 2022; Sun, G et al., 2022); however, no image analyses software is capable of counting the young spikelet primordia or the mature spikelets on the spike. Therefore, setting up a computerized pipeline for the automated calculation of spikelet primordia at MYP stage as well as the FSN at heading or harvest is essential. To set up an image analysis pipeline, a representative training data set of the spike images at MYP is needed to train the convolutional neural network (CNN). After successfully training the network, an adapted spikelet number CNN is generated. The performance of the generated CNN could be evaluated by applying the network to a set of images and manually analyzing the same image data set using image analysis software (e.g., Axio vision, Fiji, etc) to correlate the CNN and manual curated data points.

Another strategy is the deployment of genomic selection for PTD. Genomic selection—a powerful tool in plant breeding—provides advantages in crop breeding by shortening of the generation interval by accelerating the genetic gain per unit time. In genomic selection, a prediction model is built using a training set with genotypic and phenotypic data and genomic estimated breeding/genetic values (GEBVs or GEGVs) are predicted in a target set of lines having only the genotypic data (Meuwissen et al., 2001). GEGVs are particularly useful in case of highly complex and laborious-to-phenotype traits (Bernardo & Yu, 2007; Lorenzana & Bernardo, 2009; Heslot et al., 2012; Schmid & Thorwarth, 2014; He et al., 2016; Sallam & Smith, 2016). There are several approaches to perform genomic selection such as genomic best linear unbiased predictions. Bayesian methods (De Los Campos et al., 2009), reproducing kernels Hilbert spaces regression (Gianola & Van Kaam, 2008), etc. Previous success of genomic selection in predicting the genetic values in barley warrant its use in predicting both PSN and PTD. Given the robust phenotypic and genotypic data available for 417 accessions in this dissertation, prospects of genome-wide prediction can be assessed. High heritability values and high-quality phenotypic data promise high prediction accuracies for the investigated traits, most importantly PSN and PTD.

5.4 *RFL* genes in relation to pre-anthesis tip degeneration

For PTD, three QTL on chr1H, chr3H and chr7H were identified. The chr3H QTL spanned a few kb and contained three genes, of which, one gene annotated as RAP (octotricopeptide repeat) was selected as the candidate gene for barley spike PTD. The detailed description and analyses of *HvRAP* is mentioned in Chapter 4. The chr1H QTL spanned around 1.51Mb and is associated with 10 high-confidence genes annotated as Pentatricopeptide repeats (PPR). The PPRs on chr 1H belong to a family of helical repeat proteins and are known to be involved in various development processes such as chloroplast development, flowering time, RNA metabolism, tassel branch number and kernel development in maize, pollen development, and other plant growth and development process (Barkan & Small, 2014; Manna, 2015; Wei & Han, 2016; Liu et al., 2017; Chen et al., 2018; Emami & Kempken, 2019; Yang et al., 2020; Zhang et al., 2020). Furthermore, mutation in PPR protein-coding genes leads to defective seed development, pollen abortion, and retarded growth (Li et al., 2021).

The annotation analyses revealed that chr1H PPR gene cluster is the mitochondria-located restorer of fertility-like (*RFL*) genes and these *RFL* genes function as a restorers of fertility in cytoplasmic male sterility/fertility restorer systems (Hölzle et al., 2011). The majority of the *RFL* genes identified so far belong to P-class (proteins containing series of only canonical PPR motifs) of PPR proteins. The P-class PPRs are involved in RNA stabilization and processing, including 5' and 3' RNA cleavage and intron splicing and also in the initiation of mRNA translation (Pfalz et al., 2009; Prikryl et al., 2011). It was reported that *RFL* genes are mostly ubiquitous in plant genomes even when the plant has no relationship to cytoplasmic male sterility (Kubo et al., 2020). The *RFL* genes in *Arabidopsis* are not involved in cytoplasmic male sterility, indicating that there exist functions of *RFL* genes other than restoring fertility. For instance, the *RFL8* gene is essential for plant embryo development (Nguyen et al., 2021).

The PPR gene cluster reported in the present study coincides with an already reported high-density *RFL* gene cluster on chr1H (Melonek et al., 2019). Five of the 10 detected PPR genes were previously annotated as *HvRFL4*, *HvRFL24*, *HvRFL8*, *HvRFL9*, and *HvRFL10* (Table 5.1). The most significant SNP associated with chr1H QTL was an intergenic SNP between *HvRFL9* and *HvRFL10*; however, the entire region was in LD. This warrant, first and foremost, additional analyses to identify the exact gene associated with barley PTD. Based on the available expression databases in barley, the detected PPR gene cluster show the highest expression in inflorescence tissues (both young and developing) and developing caryopsis and lodicules (Figures 5.1 and 5.2).

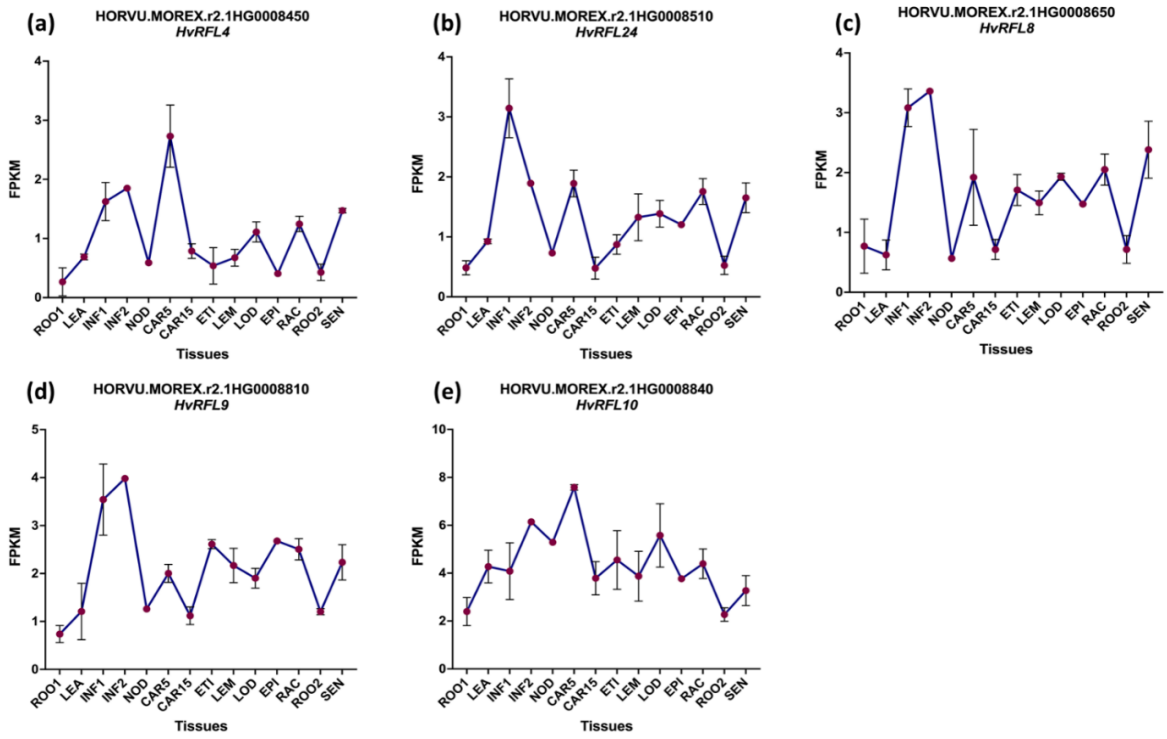


Figure 5.1 Expression pattern of known barley *HvRFL* genes. Expression pattern of (a) *HvRFL4* (b) *HvRFL24* (c) *HvRFL8* (d) *HvRFL9*, and (e) *HvRFL10*. Raw data were retrieved from the IPK BRIDGE web portal https://bridge.ipk-gatersleben.de/#snpbrowser.
ROO1 = roots from seedlings (10cm shoot stage),
LEA = shoots from seedlings (10cm shoot stage), INF1 = young inflorescence (5mm),
INF2 = developing inflorescence (1–1.5cm), NOD = developing tillers,
3rd internode (42 DAP), CAR5 = developing grain (5 DAP),
CAR15 = developing grain (15 DAP), ETI = etiolated seedling, dark condition (10 DAP),
LEM = inflorescences, lemma (42 DAP), LOD = inflorescences, lodicule (42DAP),
EPI = epidermal strips (28 DAP), RAC = inflorescence, rachis (35 DAP),
ROO2 = roots (28 DAP), SEN = senescing leaves (58 DAP), and
FPKM = fragments per kilobase of exon per million mapped fragments.

Apart from these tissues, these genes are also expressed in anthers and apical meristem (Table 5.1, the expression data was retrieved from https://ics.hutton.ac.uk/eorna/index.html), indicating their involvement in inflorescence and floral organ development. As *RFL* genes are located in mitochondria, and anthers are the first organ to show the visible effects of barley PTD (Huang et al., 2023; Shanmugaraj et al., 2023), exploration of mitochondria-associated *RFL* gene functioning within the anthers will shed more light on relationship of these genes with barley PTD. Also, this will broaden the knowledge related to the functions of *RFL* genes in barley apart from fertility restoration.

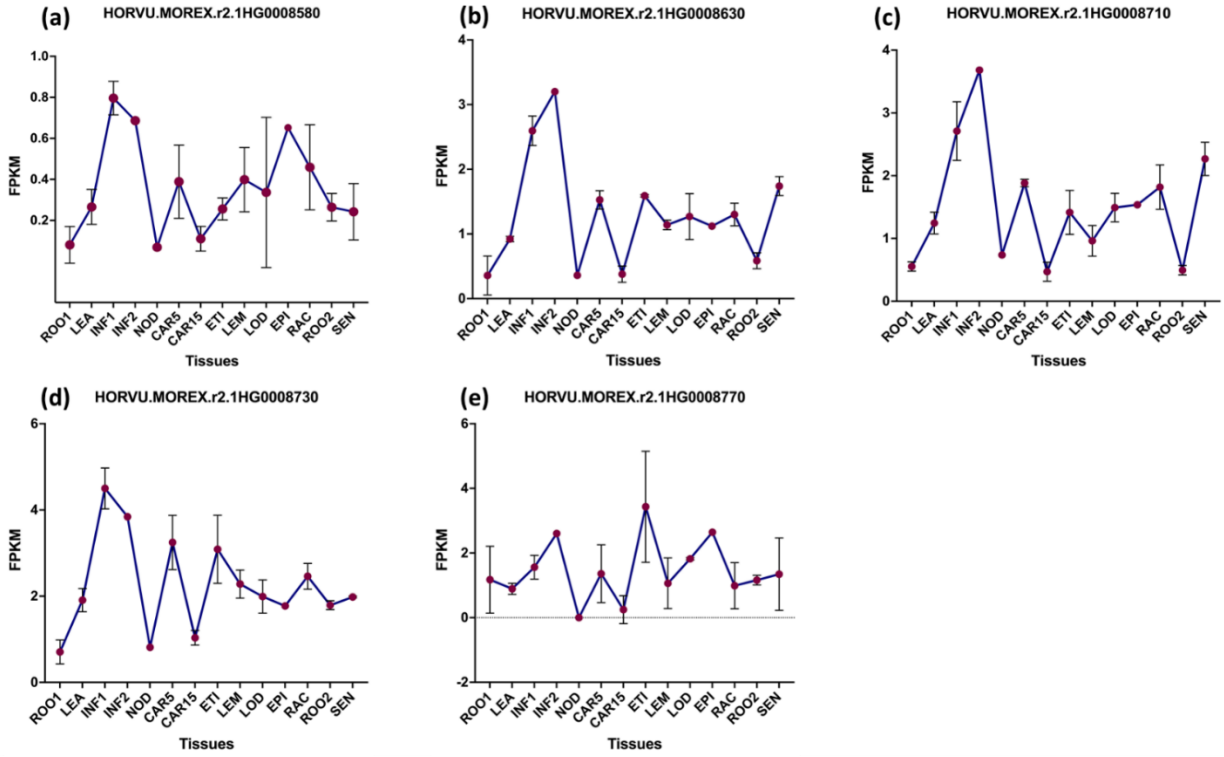


Figure 5.2 Expression pattern of unknown barley *RFL* gene on chromosome 1H. Expression pattern of (a) *HORVU.MOREX.r2.1HG0008580*,
(b) *HORVU.MOREX.r2.1HG0008630*, (c) *HORVU.MOREX.r2.1HG0008710*,
(d) *HORVU.MOREX.r2.1HG0008730*, and (e) *HORVU.MOREX.r2.1HG0008770*. Raw data were retrieved from the IPK BRIDGE web portal; https://bridge.ipk-gatersleben.de/#snpbrowser.
ROO1 = roots from seedlings (10cm shoot stage),
LEA = shoots from seedlings (10cm shoot stage), INF1 = young inflorescence (5mm),
INF2 = developing inflorescence (1–1.5cm), NOD = developing tillers,
3rd internode (42 DAP), CAR5 = developing grain (5 DAP),
CAR15 = developing grain (15 DAP), ETI = etiolated seedling, dark condition (10 DAP),
LEM = inflorescences, lemma (42 DAP), LOD = inflorescences, lodicule (42DAP),
EPI = epidermal strips (28 DAP), RAC = inflorescence, rachis (35 DAP),
ROO2 = roots (28 DAP), SEN = senescing leaves (58 DAP), and
FPKM = fragments per kilobase of exon per million mapped fragments.

Table 5.1 BART IDs along with tissue expression for the PPR gene cluster on chromosome 1H. BART IDs and expression data were retrieved from EoRNA and presented in descending order.

Barley PPR gene-IDs and the known gene names	BART IDs	Highest expression in tissues/stages
HORVU.MOREX.r2.1HG0008450 ***HvRFL4***	BART1_0-p00904	• Anthers (0.5-0.9mm) • Apical meristem, W 5.5 • Developing grains (5DPA)
HORVU.MOREX.r2.1HG0008510 ***HvRFL24***	BART1_0-p00898	• Developing inflorescence (1–1.5cm) • Young inflorescence (5mm) • Developing tillers at six leaf stage
HORVU.MOREX.r2.1HG0008580	BART1_0-p00904	• Anthers (0.5–0.9mm) • Apical meristem, W 5.5 • Developing grains (5DPA)
HORVU.MOREX.r2.1HG0008630	BART1_0-p00913	• Apical meristem, W3.5 • AP stage • Shoot apex, W3.5
HORVU.MOREX.r2.1HG0008650 ***HvRFL8***	BART1_0-p00907	• Apical meristem, W5.5 • WA stage • Developing inflorescence
HORVU.MOREX.r2.1HG0008710	BART1_0-p00898	• Developing inflorescence (1–1.5cm) • Young inflorescence (5mm) • Developing tillers at six leaf stage
HORVU.MOREX.r2.1HG0008730	BART1_0-p00923	• Developing inflorescence (1–1.5cm) • WA stage • Apical meristem, W5.5
HORVU.MOREX.r2.1HG0008770	BART1_0-p00923	• Developing inflorescence (1–1.5cm) • WA stage • Apical meristem, W5.5
HORVU.MOREX.r2.1HG0008810 ***HvRFL9***	BART1_0-p00923	• Developing inflorescence (1–1.5cm) • WA stage • Apical meristem, W5.5
HORVU.MOREX.r2.1HG0008840 ***HvRFL10***	BART1_0-p00950	• Young inflorescence • Developing grains (5DPA) • Root zone

5.5 **Potential association of *HvRAN2* with barley pre-anthesis tip degeneration**

The PTD QTL on chr7H harbor a candidate gene, *HORVU.MOREX.r2.7HG0592180* (*HvRAN2*) that is annotated as GTP binding nuclear protein. Phylogenic analysis revealed that *HvRAN2* is present in the same clade as wheat *TaRAN2* (*TraesCS7B02G254000* and *TraesCS7D02G349900*), rice *OsRAN1* (LOC_Os06g39875) and *Arabidopsis RAN* (*AtRAN1-AtRAN4*) genes (Choudhury et al., 2021). *RAN* genes are a family of Ras-related GTPases that regulate the transport of RNA and proteins between the nucleus and cytoplasm during interphase and modulate several aspects of mitosis. Their function is mediated by regulatory factors, RanGEFs (Ran Guanine exchange factors) and RanGAPs (Ran GTPases activating

proteins), collectively known as Ran-binding proteins (RanBPs) (Nielsen, 2020). RanGEFs switch on whereas RanGAPs switch off the activity of *RAN*.

The *RAN* genes are involved in various developmental processes such as cell division (Chen et al., 2011), mitotic activity (Ciciarello et al., 2007), spindle assembly (Petr Kalab, 1999), nucleo-cytoplasmic transport (Görlich & Kutay, 1999; Melchior, 2001; Cavazza & Vernos, 2015), regulation of nuclear structure (Clarke & Zhang, 2001), cold resistance (Xu & Cai, 2014), endosperm cellularization (Liu et al., 2014), stress tolerance (Xu et al., 2016), leaf senescence (Pham et al., 2022), and lastly, in male and female gametogenesis (Qin et al., 2022). In *Arabidopsis*, all *RAN* genes expressed highly in reproductive tissues such as ovules and pollens, and the loss-of-function of one *RAN* gene (*RAN1* in *Arabidopsis*) showed no defect in the gametophytic development. (Qin et al., 2022). The lack of phenotypic defect was likely due to the redundant function of all the *RAN* genes. Changes in the dynamic activity of *RAN* GTPases caused the arrest of pollen mitosis I as well as the arrest of megaspores mitosis, reducing both male and female fertility. The mechanism in which *RAN* along with RanBPs regulate the progression of mitosis was described by Kim et al. (2001). *RAN* functions in two forms, Ran-GDP and Ran-GTP: Ran-GTP is present in the nucleus while Ran-GDP in the cytoplasm. RanBPs are essential for hydrolyses of GTP to GDP. For example, a decrease in RanBPs concentration results in an increase in the ratio of Ran-GTP to Ran-GDP in the cytoplasm leading to a decrease in the delivery rate of key proteins involved in the mitotic cell cycle to the nucleus. This results in failure of sister chromatid separation, freezing many cells at metaphase, and ultimately, the failure in the progression of the cell cycle.

To date, two *RAN* genes have been identified in barley, one on chr1H (*HORVU.MOREX.r2.1HG0074850*; *HvRAN1*) and another on chr7H. Unlike *Arabidopsis RAN*, both genes in barley have different expression patterns. *HvRAN1* is expressed in the young and developing inflorescence and only developing grain (five days pollination), whereas, *HvRAN2* is expressed exclusively in the developing grains, both five and 15 days after pollination and microspore tissues (Figure 5.3). Hence, it could be speculated that both barley *RAN* genes do not function redundantly and the activation and deactivation mechanism of these genes at a specific developmental stage must be studied in detail. *HvRAN2* expression decreases sharply in the inflorescence meristem (shown in Chapter 4), and the onset of spike PTD in barley has been linked with programmed cell death of anther primordia in developing florets (Shanmugaraj et al., 2023). Therefore, based on the known functions of *RAN*, it becomes crucial to study the role of *RAN* in mitotic activity of the inflorescence meristem and in the later stages, for instance, in anther and ovary mitosis. Based on the GWAS results, two haplotypes were constructed for *HvRAN2*; here, haplotype2 showed a smaller number of degenerated spikelets and, consequently, higher survived spikelets (Figure 5.4). Deeper genetic and molecular analyses should reveal the DNA or protein level differences between both haplotypes responsible for the phenotypic differences. RanBPs are essential for the activity of *RAN* GTPases; in barley, no RanBPs have been identified so far. Therefore, to understand the role of *RAN* genes in barley, it is essential to determine its regulatory factors as well. In conclusion, although *HvRAN2* is identified as a candidate gene for PTD, more work is needed to understand its role in barley development thoroughly.

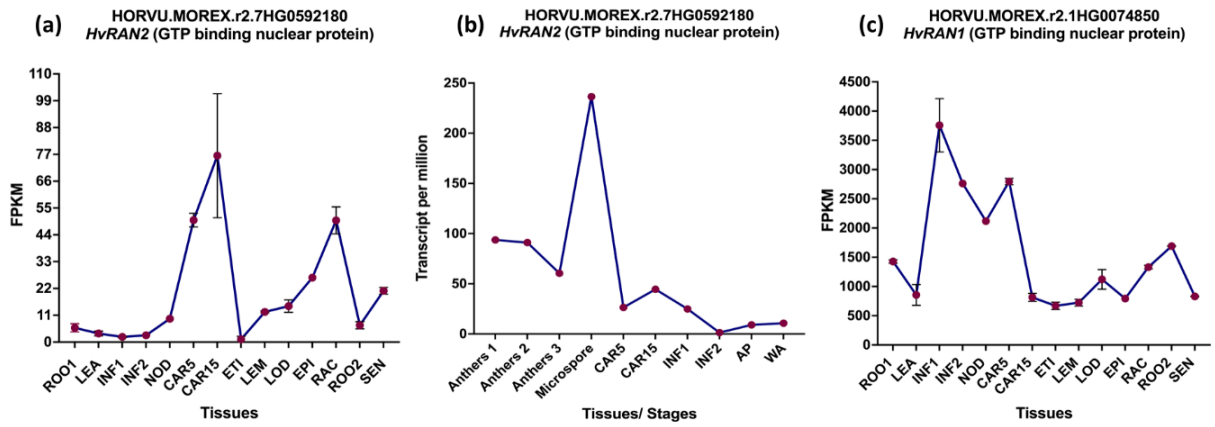


Figure 5.3 Expression pattern for barley *RAN* genes. Expression pattern for (a) *HvRAN2* (raw data from IPK BRIDGE web portal; https://bridge.ipk-gatersleben.de/#snpbrowser), (b) *HvRAN2* (raw data from EoRNA; https://ics.hutton.ac.uk/eorna/index.html), and (c) *HvRAN1* (raw data from IPK BRIDGE web portal; https://bridge.ipk-gatersleben.de/#snpbrowser). ROO1 = roots from seedlings (10cm shoot stage), LEA = shoots from seedlings (10cm shoot stage), INF1 = young inflorescence (5mm), INF2 = developing inflorescence (1–1.5cm), NOD = developing tillers, 3rd internode (42 DAP), CAR5 = developing grain (5 DAP), CAR15 = developing grain (15 DAP), ETI = etiolated seedling, dark condition (10 DAP), LEM = inflorescences, lemma (42 DAP), LOD = inflorescences, lodicule (42DAP),EPI = epidermal strips (28 DAP), RAC = inflorescence, rachis (35 DAP), ROO2 = roots (28 DAP), SEN = senescing leaves (58 DAP), anthers 1 = anther size 0.5–0.9mm, anthers 2 = anther size 1–1.2mm, anthers 3 = anther size 1.3–1.4 mm, AP = awn primordium stage, WA = white anther stage, and FPKM = fragments per kilobase of exon per million mapped fragments.

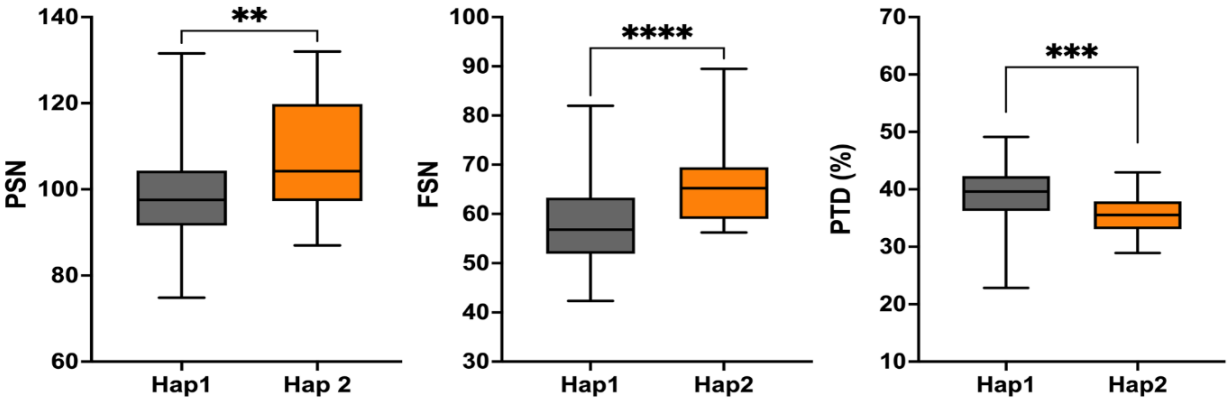


Figure 5.4 Comparison for potential spikelet number (PSN), final spikelet number (FSN) and pre-anthesis tip degeneration (PTD) with respect to major haplotypes (Hap1: *n* = 201; Hap2: *n* = 27) for the gene *HvRAN2*.

5.6 **Putative candidate genes for final spikelet number, plant height, and culm dry weight**

In total, 132 QTL identified for all the investigated traits. Based on the available literature and expression databases, putative candidate genes could be shortlisted for a few of the traits. The putative candidate gene for PSN, AL, HD, GNS and GS are mentioned in Chapter 4.

A strong association on chr3H for final spikelet number was identified. In the associated region, gene *HORVU.MOREX.r2.3HG0256830*, annotated as GRAS family transcription factor, which is part of *SCARECROW-LIKE (SCL)* GRAS transcription factors seems to be of interest in understanding the mechanism behind FSN. The GRAS proteins are involved in GA responses that control flowering and regulate apical meristem development (Bolle, 2004). In *Arabidopsis*, microRNA171 (*miR171*) represses the differentiation of axillary meristems by suppressing the expression of *SCL* transcription factors. *miR171* is a well-conserved miRNA family known to regulate members of the *SCL* transcription factor family. In barley, over-expression of *miR171* affects phase transitions and floral meristem determinacy where the transgenic plants showed a delay in juvenile-to-adult transition, dwarfism due to reduced internode length, and partially sterile spikelets (Curaba et al., 2013). Since the SCL genes are involved in the spikelet meristem determinacy and overall plant growth, the association of FSN on chr3H warrants deeper study related to SCL genes in barley.

Among the HC genes on chr5H for PH, two genes (*HORVU.MOREX.r2.5HG0415910* and *HORVU.MOREX.r2.5HG0416190*) were annotated as Cytochrome P450. Plant Cytochrome P450 is known to be involved in various biosynthetic reactions and plays important roles in plant development. In rice, two studies described the role of Cytochrome P450 in internode elongation which was further translated into overall decreased plant height in the mutants (Luo et al., 2006; Ramamoorthy et al., 2011). The *EUI1* locus in rice encoding Cytochrome P450 was mapped on chromosome 5 and played a negative role in gibberellin-mediated regulation of cell elongation in the uppermost internode. Bai et al. (2021) also identified the PH QTL on chr2H in barley corresponding to the gene annotated as Cytochrome P450. For CDW, the most significant peak on chr7H contained 10 HC-genes, of which, seven were annotated as ABC transporter. ABC (ATP binding cassette) transporters are important for plant development as these influence gametogenesis, seed development, seed germination, organ formation, and secondary growth. Also, ABC transporters are essential for transporting phytohormones such as auxin, cytokinin, ABA, and jasmonic acid. Recently, it was shown that the mutants for auxin transporting ABC transporter showed reduced growth and decreased apical dominance (Kang et al., 2011; Do et al., 2018). Since phytohormones are indispensable for the growth and development of plants, the role of ABC transporters in plant development in barley is worth studying. Future studies regarding the putative candidate genes for FSN, PH, and CDW as well as for the traits mentioned in Chapter 4 are expected to aid in understanding the genetic architecture of these traits.

Chapter 6. SUMMARY

Barley (*Hordeum vulgare* L.) is a cereal plant of the grass family *Poaceae* with unbranched inflorescence known as spike. The barley spike is indeterminate in nature, i.e., it produces multiple spikelet primordia without terminating into a terminal spikelet. The barley spikelets are single-flowered grain bearing structures produced in a distichous manner on the spike. Barley possesses three growth phases, namely, vegetative, reproductive (further divided into early and late reproductive), and grain filling phase. During the early reproductive phase, the immature spike reaches maximum yield potential (MYP) stage. After MYP stage, no new spikelet primordia are initiated on the spike, and a degeneration process is triggered in which the apical spikelet primordia start to degenerate from the tip basipetally (basipetal degeneration). This degeneration process is known as pre-anthesis tip degeneration (PTD) which decreases the yield potential of barley by reducing the final spikelet number per spike. Reducing the extent of PTD, hence, represents an opportunity to realize true yield potential in barley.

In this dissertation, naturally present variations of PTD and other spike, grain, and shoot traits in a barley association panel of 417 six-rowed barley accessions collected from the German Federal *ex-situ* Genebank maintained at IPK, Gatersleben, Germany, were studied. The samples were collected at three development stages, namely, MYP stage, heading and harvest. The phenotypic data were collected only from the main culms as they are known to be less affected by the environmental perturbations. The phenotypic data analyses showed a significantly large within- and across-years genotypic variation with high broad-sense heritability estimates for all the investigated traits. High heritability estimates showed the high quality of phenotypic data for further genetic and genomic studies. PTD varied from 13 to 51% depending on the accessions and their geographical origin. Among the spike traits, PTD was negatively correlated with final spikelet number, spike length, and density, while, positively with awn length. PTD showed non-significant correlation with potential spikelet number. However, path analysis revealed that PSN has a significant positive effect on PTD, and it acts as a suppressor variable. Here, PSN together with FSN explained about 93% of the observed phenotypic variability for PTD. The positive correlation between PTD and awn length suggested a plausible role of the rapidly growing awns during the spikelet abortion process—especially after Waddington stage 5. In addition, moderate positive correlations with PTD and grain traits such as grain length, grain area, and thousand-grain weight were observed. Interestingly, opposite trends for different traits in the accessions belonging to different geographical origin were observed. For example, European accessions displayed higher PTD, earlier heading, and higher values for grain morphometric traits. On the other hand, Asian accessions displayed the lowest PTD indicating the presence of favorable alleles that may be exploited in breeding programs.

Genome-wide association studies using ~ 22 million SNPs for each investigated trait to identify trait-underlying-QTL or genes were performed within the purview of this dissertation. A total of 132 QTL were detected for the investigated traits along with a few overlapping QTL for different traits which may mark the pleiotropy among traits. As identifying PTD associated QTL was one of the major objectives of this dissertation, three PTD-QTL located on chr1H, 3H and 7H were identified. Delineating chr3H PTD-QTL revealed a promising candidate gene

annotated as barley octotricopeptide repeat protein with α-helical RNA-binding (RAP) domain. It is a single copy gene essential for the maturation of 16s rRNA, chloroplast biogenesis and chlorophyll synthesis. The barley RAP (*HvRAP*) has three exons, and the RAP domain ranges from 580 to 635 amino acids. Natural sequence variations here, i.e., an amino acid substitution (I584M) within the RAP domain, lead to enhanced PTD. Other plant species such as wheat, rice, maize, and *Arabidopsis* contain Isoleucine at that amino acid site. *HvRAP* showed decreased expression starting from the stamen primordia stage and has lower expression in the apical spike section (that is destined to abort) of the spike. The *HvRAP* TILLING lines showed reduced photosynthetic efficiency, chloroplast structural defects, and decreased spikelet as well as grain number. Along with *HvRAP*, two another putative candidate genes/regions for PTD, one on chromosome 7H (*HvRAN2*) and another on chromosome 1H (*HvRFL*) were reported. Further in-depth molecular studies for the identified candidate genes for barley spike PTD are needed to reveal their exact role in PTD. In addition to PTD, putative candidate genes for other important traits were also identified. Taken together, based on a large set of diverse barley accessions, this dissertation provides a deeper understanding of the genetic nature of PTD, its association with traits of high agronomic importance, and putative candidate genes for further gene cloning studies.

Chapter 7. **Zusammenfassung**

Gerste (*Hordeum vulgare* L.) ist eine Getreidepflanze aus der Familie der Gräser (*Poaceae*) mit unverzweigtem Blütenstand, der als Ähre bezeichnet wird. Die Gerstenähre ist von Natur aus unbestimmt, d. h. sie bildet mehrere Ährchen-Primordien aus, ohne in ein endständiges Ährchen zu münden. Die Ährchen der Gerste sind einblütige, korntragende Gebilde, die in einer distichartigen Weise an der Ähre entstehen. Die Gerste besitzt drei Wachstumsphasen, nämlich die vegetative, die reproduktive (weiter unterteilt in frühe und späte reproduktive Phase) und die Kornfüllungsphase. Während der frühen Reproduktionsphase erreicht die unreife Ähre das Stadium des maximalen Ertragspotenzials (MYP). Nach dem MYP-Stadium bilden sich an der Ähre keine neuen Ährchenprimordien mehr, und es wird ein Degenerationsprozess ausgelöst, bei dem die apikalen Ährchenprimordien von der Spitze aus basipetal zu degenerieren beginnen (basipetale Degeneration). Dieser Degenerationsprozess wird als "pre-anthesis tip degeneration" (PTD) bezeichnet, die das Ertragspotenzial der Gerste durch eine Verringerung der endgültigen Ährchenanzahl pro Ähre verringert. Die Verringerung des Ausmaßes der PTD stellt daher eine Möglichkeit dar, das tatsächliche Ertragspotenzial der Gerste auszuschöpfen.

In dieser Dissertation wurden natürlich vorhandene Variationen der PTD und anderer Ähren-, Korn- und Sprossmerkmale in einem Gersten-Assoziationspanel aus 417 sechsreihigen Gerstenakzessionen untersucht, die aus der Ex-situ-Genbank des Bundes am IPK in Gatersleben, Deutschland, stammen. Die Proben wurden in drei Entwicklungsstadien entnommen, nämlich im MYP-Stadium, zur Ernte und zur Ernte. Die phänotypischen Daten wurden nur an den Hauptstängeln erhoben, da diese bekanntermaßen weniger von Umweltstörungen betroffen sind. Die Analysen der phänotypischen Daten zeigten eine signifikant große genotypische Variation innerhalb und zwischen den Jahren mit hohen Heritabilitätsschätzungen für alle untersuchten Merkmale. Die hohen Heritabilitätsschätzungen zeigen die hohe Qualität der phänotypischen Daten für weitere genetische und genomische Studien. Die PTD variierte je nach Akzessionen und deren geografischer Herkunft zwischen 13 und 51 %. Von den Ährenmerkmalen korrelierte die PTD negativ mit der Anzahl der letzten Ährchen, der Ährenlänge und der Dichte, während sie positiv mit der Grannenlänge korrelierte. PTD zeigte eine nicht signifikante Korrelation mit der potenziellen Ährchenzahl. Die Pfadanalyse ergab jedoch, dass PSN einen signifikant positiven Effekt auf PTD hat und als Suppressor-Variable wirkt. Hier erklärte PSN zusammen mit FSN etwa 93 % der beobachteten phänotypischen Variabilität für PTD. Die positive Korrelation zwischen PTD und Grannenlänge deutet auf eine plausible Rolle der schnell wachsenden Grannen während des Ährchenabtreibungsprozesses hin - insbesondere nach dem Waddington-Stadium 5. Darüber hinaus wurden mäßig positive Korrelationen zwischen PTD und Kornmerkmalen wie Kornlänge, Kornfläche und Tausendkorngewicht beobachtet. Interessanterweise wurden bei den Akzessionen unterschiedlicher geografischer Herkunft entgegengesetzte Trends für verschiedene Merkmale beobachtet. So wiesen die europäischen Akzessionen beispielsweise eine höhere PTD, einen früheren Kornansatz und höhere Werte für die morphometrischen Eigenschaften der Körner auf. Andererseits wiesen die asiatischen Akzessionen die niedrigste PTD auf, was auf das Vorhandensein günstiger Allele hinweist, die in Züchtungsprogrammen genutzt werden können.

Im Rahmen dieser Dissertation wurden genomweite Assoziationsstudien unter Verwendung von ca. 22 Millionen SNPs für jedes untersuchte Merkmal durchgeführt, um dem Merkmal zugrundeliegende QTL oder Gene zu identifizieren. Insgesamt wurden 132 QTL für

die untersuchten Merkmale sowie einige sich überschneidende QTL für verschiedene Merkmale entdeckt, die möglicherweise die Pleiotropie zwischen den Merkmalen markieren. Da die Identifizierung von PTD-assoziierten QTL eines der Hauptziele dieser Dissertation war, wurden drei PTD-QTL auf chr1H, 3H und 7H identifiziert. Die Abgrenzung des PTD-QTL auf chr3H ergab ein vielversprechendes Kandidatengen, das als Barley-Octotricopeptid-Repeat-Protein mit α-helicaler RNA-Bindungsdomäne (RAP) annotiert ist. Es handelt sich um ein Gen mit einer einzigen Kopie, das für die Reifung der 16s rRNA, die Chloroplastenbiogenese und die Chlorophyllsynthese wesentlich ist. Das Gersten-RAP (*HvRAP*) hat drei Exons, und die RAP-Domäne reicht von 580 bis 635 Aminosäuren. Natürliche Sequenzvariationen, d. h. eine Aminosäuresubstitution (I584M) innerhalb der RAP-Domäne, führen hier zu einer erhöhten PTD. Andere Pflanzenarten wie Weizen, Reis, Mais und *Arabidopsis* enthalten Isoleucin an dieser Aminosäureposition. *HvRAP* zeigte eine verminderte Expression ab dem Stadium der Staubblattprimordien und eine geringere Expression im apikalen Ährenabschnitt (der zum Abbruch bestimmt ist) der Ähre. Die *HvRAP* TILLING-Linien zeigten eine verringerte photosynthetische Effizienz, strukturelle Defekte der Chloroplasten und eine verringerte Ährchen- und Kornzahl. Neben *HvRAP* wurden zwei weitere mutmaßliche Kandidatengene/-regionen für PTD, eines auf Chromosom 7H (*HvRAN2*) und ein weiteres auf Chromosom 1H (*HvRFL*), entdeckt. Weitere eingehende molekulare Studien zu den identifizierten Kandidatengenen für PTD in Gerstenähren sind erforderlich, um ihre genaue Rolle bei PTD aufzudecken. Neben der PTD wurden auch mutmaßliche Kandidatengene für andere wichtige Merkmale identifiziert. Insgesamt bietet diese Dissertation auf der Grundlage eines großen Satzes verschiedener Gerstenakzessionen ein tieferes Verständnis der genetischen Natur der PTD, ihrer Verbindung zu agronomisch wichtigen Merkmalen und möglicher Kandidatengene für weitere Genklonierungsstudien.

REFERENCES

Abd El-Mohsen AA, Abo-Hegazy SR. 2013. Comparing the relative efficiency of two experimental designs in wheat field trials. *Scientific Research and Review Journal* **1**(3): 101-109.

AHDB, Kendall S, Berry P, Sylvester-Bradley R, White C, Clarke S, Bingham I, Hoad S, Knight S, Ober E 2023. Barley growth guide https://ahdb.org.uk/knowledge-library/barley-growth-guide.

Ali A, Wu T, Xu Z, Riaz A, Alqudah AM, Iqbal MZ, Zhang H, Liao Y, Chen X, Liu Y, et al. 2022. Phytohormones and transcriptome analyses revealed the dynamics involved in spikelet abortion and inflorescence development in rice. *International Journal of Molecular Sciences* **23**(14).

Ali A, Xu P, Riaz A, Wu X. 2019. Current advances in molecular mechanisms and physiological basis of panicle degeneration in rice. *International Journal of Molecular Sciences* **20**(7): 1613.

Ali M, Hussain M, Khan M, Ali Z, Zulkiffal M, Anwar J, Sabir W, Zeeshan M. 2010. Source-sink relationship between photosynthetic organs and grain yield attributes during grain filling stage in spring wheat (Triticum aestivum). *Int. J. Agric. Biol* **12**(4): 509-515.

Ali M, Li J, Hu W, Yu J, Khan SU, Khan MHU, Xie G, Wang J, Wang L. 2021. Uncovering genomic regions controlling plant architectural traits in hexaploid wheat using different GWAS models. *Scientific Reports* **11**(1): 6767.

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

Anderson P, Oelke E, Simmons S. 1995. *Growth and development guide for spring barley*: St. Paul, MN: University of Minnesota Extension Service.

Appleyard M, Kirby E, Fellowes G. 1982. Relationships between the duration of phases in the pre-anthesis life cycle of spring barley. *Australian Journal of Agricultural Research* **33**(6): 917-925.

Arisnabarreta S, Miralles DJ. 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 DJ. 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.

Arisnabarreta S, Miralles DJ. 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.

Aspinall D. 1966. Effects of day length and light intensity on growth of barley IV. Genetically controlled variation in response to photoperiod. *Australian Journal of Biological Sciences* **19**(4): 517-534.

Badr A, M K, Sch R, Rabey HE, Effgen S, Ibrahim HH, Pozzi C, Rohde W, Salamini F. 2000. On the origin and domestication history of barley (*Hordeum vulgare*). *Molecular Biology and Evolution* **17**(4): 499-510.

Baethgen WE, Christianson CB, Lamothe AG. 1995. Nitrogen fertilizer effects on growth, grain yield, and yield components of malting barley. *Field Crops Research* **43**(2-3): 87-99.

Bai J, Zhu X, Wang Q, Zhang J, Chen H, Dong G, Zhu L, Zheng H, Xie Q, Nian J, et al. 2015. Rice *TUTOU1* encodes a suppressor of cAMP Receptor-Like Protein that is important for actin organization and panicle development. *Plant Physiology* **169**(2): 1179-1191.

Bai Y, Zhao X, Yao X, Yao Y, An L, Li X, Wang Y, Gao X, Jia Y, Guan L, et al. 2021. Genome wide association study of plant height and tiller number in hulless barley. *PLoS One* **16**(12): e0260723.

Baik B-K, E. Ullrich S. 2008. Barley for food: characteristics, improvement, and renewed interest. *Journal of Cereal Science* **48**: 233-242.

Bancal P. 2008. Positive contribution of stem growth to grain number per spike in wheat. *Field Crops Research* **105**(1): 27-39.

Bancal P. 2009. Early development and enlargement of wheat floret primordia suggest a role of partitioning within spike to grain set. *Field Crops Research* **110**(1): 44-53.

Barkan A, Small I. 2014. Pentatricopeptide repeat proteins in plants. *Annual Review of Plant Biology* **65**: 415-442.

Bates D, Mächler M, Bolker B, Walker S. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* **67**(1): 1-48.

Bera D, Chatterjee ND, Bera S. 2021. Comparative performance of linear regression, polynomial regression and generalized additive model for canopy cover estimation in the dry deciduous forest of West Bengal. *Remote Sensing Applications: Society and Environment* **22**: 100502.

Bernardo R, Yu J. 2007. Prospects for genomewide selection for quantitative traits in maize. *Crop Science* **47**(3): 1082-1090.

Beyer S, Daba S, Tyagi P, Bockelman H, Brown-Guedira G, com Iee, Mohammadi M. 2019. Loci and candidate genes controlling root traits in wheat seedlings—a wheat root GWAS. *Functional & Integrative Genomics* **19**: 91-107.

Bobbitt Z 2020. How to Perform Quadratic Regression in R

Bolle C. 2004. The role of GRAS proteins in plant signal transduction and development. *Planta* **218**: 683-692.

Bonnett OT. 1966. Inflorescences of maize, wheat, rye, barley, and oats: their initiation and development/721. *Bulletin (University of Illinois (Urbana-Champaign campus). Agricultural Experiment Station); no. 721*.

Boonchoo S, Fukai S, Hetherington SE. 1998. Barley yield and grain protein concentration as affected by assimilate and nitrogen availability. *Australian Journal of Agricultural Research* **49**(4): 695-706.

Bothmer Rv. 1992. The wild species of Hordeum: relationships and potential use for improvement of cultivated barley. *Barley: Genetics, Biochemistry, Molecular Biology, and Biotechnology*: 3-18.

Boussora F, Allam M, Guasmi F, Ferchichi A, Rutten T, Hansson M, Youssef HM, Börner 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): 1-11.

Brooking I, Kirby E. 1981. Interrelationships between stem and ear development in winter wheat: the effects of a Norin 10 dwarfing gene, Gai/Rht 2. *Journal of Agricultural Science* **97**(2): 373-381.

Burger W, LaBerge D. 1985. Malting and brewing quality. *Barley*(barley): 367-401.

Cavazza T, Vernos I. 2015. The RanGTP Pathway: From nucleo-cytoplasmic transport to spindle assembly and beyond. *Frontiers in Cell and Developmental Biology* **3**: 82.

Chen L, Li YX, Li C, Shi Y, Song Y, Zhang D, Li Y, Wang T. 2018. Genome-wide analysis of the pentatricopeptide repeat gene family in different maize genomes and its important role in kernel development. *BMC Plant Biology* **18**(1): 366.

Chen N, Xu Y, Wang X, Du C, Du J, Yuan M, Xu Z, Chong K. 2011. *OsRAN2*, essential for mitosis, enhances cold tolerance in rice by promoting export of intranuclear tubulin and maintaining cell division under cold stress. *Plant, Cell & Environment* **34**(1): 52-64.

Choudhury S, Mansi, Muthusamy SK, Padaria JC, Dalal M. 2021. Genome-wide identification of Ran GTPase family genes from wheat (*T. aestivum*) and their expression profile during developmental stages and abiotic stress conditions. *Functional & Integrative Genomics* **21**(2): 239-250.

Ciciarello M, Mangiacasale R, Lavia P. 2007. Spatial control of mitosis by the GTPase Ran. *Cell and Molecular Life Sciences* **64**(15): 1891-1914.

Clarke PR, Zhang C. 2001. Ran GTPase: a master regulator of nuclear structure and function during the eukaryotic cell division cycle? *Trends in Cell Biology* **11**(9).

Corbesier L, Vincent C, Jang S, Fornara F, Fan Q, Searle I, Giakountis A, Farrona S, Gissot L, Turnbull C. 2007. FT protein movement contributes to long-distance signaling in floral induction of Arabidopsis. *Science* **316**(5827): 1030-1033.

Corey DM, Dunlap WP, Burke MJ. 1998. Averaging correlations: Expected values and bias in combined Pearson rs and Fisher's *z* transformations. *The Journal of General Psychology* **125**(3): 245-261.

Cotterell J, Dale J, Jeffcoat B. 1981a. Endogenous control of spikelet initiation and development in barley. *Br. Plant Growth Regul. Group Monogr* **7**: 130-139.

Cotterell J, Dale J, Jeffcoat B. 1981b. Endogenous control of spikelet initiation and development in barley. *British Plant Growth Regulator Group Monograph* **7**: 130-139.

Cottrell JE, Easton RH, Dale JE, Wadworth AC, Adam JS, Child RD, Hoad GV. 1985. A comparison of spike and spikelet survival in mainstem and tillers of barley. *Annals of Applied Biology* **106**(2): 365-377.

Curaba J, Talbot M, Li Z, Helliwell C. 2013. Over-expression of microRNA171 affects phase transitions and floral meristem determinancy in barley. *BMC Plant Biology* **13**(1): 1-10.

Dai D, Zhang H, He L, Chen J, Du C, Liang M, Zhang M, Wang H, Ma L. 2022. *Panicle Apical Abortion 7* regulates panicle development in rice (*Oryza sativa* L.). *International Journal of Molecular Sciences* **23**(16).

Dai F, Nevo E, Wu D, Comadran J, Zhou M, Qiu L, Chen Z, Beiles A, Chen G, Zhang G. 2012. Tibet is one of the centers of domestication of cultivated barley. *Proceedings of the National Academy of Sciences* **109**(42): 16969-16973.

De Los Campos G, Naya H, Gianola D, Crossa J, Legarra A, Manfredi E, Weigel K, Cotes JM. 2009. Predicting quantitative traits with regression models for dense molecular markers and pedigree. *Genetics* **182**(1): 375-385.

del Moral LFGa, del Moral MBGa, Molina-Cano JL, Slafer GA. 2003. Yield stability and development in two-and six-rowed winter barleys under Mediterranean conditions. *Field Crops Research* **81**(2-3): 109-119.

Del Moral LG, De la Morena I, Ramos J. 1999. Effects of nitrogen and foliar sulphur interaction on grain yield and yield components in barley. *Journal of Agronomy and Crop Science* **183**(4): 287-295.

Digel B, Pankin A, von Korff M. 2015a. Global transcriptome profiling of developing leaf and shoot apices reveals distinct genetic and environmental control of floral transition and inflorescence development in barley. *Plant Cell* **27**(9): 2318-2334.

Digel B, Pankin A, von Korff M. 2015b. Global transcriptome profiling of developing leaf and shoot apices reveals distinct genetic and environmental control of floral transition and inflorescence development in barley. *The Plant Cell* **27**(9): 2318-2334.

Do THT, Martinoia E, Lee Y. 2018. Functions of ABC transporters in plant growth and development. *Current Opinion in Plant Biology* **41**: 32-38.

Dubcovsky J, Chen C, Yan L. 2005. Molecular characterization of the allelic variation at the *VRN-H2* vernalization locus in barley. *Molecular Breeding* **15**: 395-407.

Ellis R, Kirby E. 1980. A comparison of spring barley grown in England and in Scotland. 2. Yield and its components. *The Journal of Agricultural Science* **95**(1): 111-115.

Emami H, Kempken F. 2019. PRECOCIOUS1 (POCO1), a mitochondrial pentatricopeptide repeat protein affects flowering time in *Arabidopsis thaliana*. *The Plant Journal* **100**(2): 265-278.

FAO F. 2021. Food and agriculture organization of the United Nations. *Rome, URL: http://faostat.fao.org*.

Feekes W. 1941. De tarwe en haar milieu. *Vers XVII Tech. Tarwe. Comm., Groningen*.

Ferrante A, Savin R, Slafer GA. 2010. Floret development of durum wheat in response to nitrogen availability. *Journal of Experimental Botany* **61**(15): 4351-4359.

Ferrante A, Savin R, Slafer GA. 2013a. Floret development and grain setting differences between modern durum wheats under contrasting nitrogen availability. *Journal of Experimental Botany* **64**(1): 169-184.

Ferrante A, Savin R, Slafer GA. 2013b. Is floret primordia death triggered by floret development in durum wheat? *Journal of Experimental Botany* **64**(10): 2859-2869.

Ferrante A, Savin R, Slafer GA. 2015. Relationship between fruiting efficiency and grain weight in durum wheat. *Field Crops Research* **177**: 109-116.

Fu D, Szűcs P, Yan L, Helguera M, Skinner JS, Von Zitzewitz J, Hayes PM, Dubcovsky J. 2005. Large deletions within the first intron in *VRN-1* are associated with spring growth habit in barley and wheat. *Molecular Genetics and Genomics* **273**: 54-65.

Gallagher J, Biscoe P. 1978. Radiation absorption, growth and yield of cereals. *The Journal of Agricultural Science* **91**(1): 47-60.

Ganeshaiah KN, Uma Shaanker R. 1994. Seed and fruit abortion as a process of self organization among developing sinks. *Physiologia Plantarum* **91**(1): 81-89.

Garcıa del Moral L, Miralles D, Slafer G. 2002. Initiation and appearance of vegetative and reproductive structures throughout barley development. *Barley science: recent advances from molecular biology to agronomy of yield and quality*: 243-268.

Ghiglione HO, Gonzalez FG, Serrago R, Maldonado SB, Chilcott C, Cura JA, Miralles DJ, Zhu T, Casal JJ. 2008. Autophagy regulated by day length determines the number of fertile florets in wheat. *The Plant Journal* **55**(6): 1010-1024.

Gianola D, Van Kaam JB. 2008. Reproducing kernel Hilbert spaces regression methods for genomic assisted prediction of quantitative traits. *Genetics* **178**(4): 2289-2303.

Gol L, Tome F, von Korff M. 2017. Floral transitions in wheat and barley: interactions between photoperiod, abiotic stresses, and nutrient status. *J Exp Bot* **68**(7): 1399-1410.

Gol L, Tomé F, von Korff M. 2017. Floral transitions in wheat and barley: interactions between photoperiod, abiotic stresses, and nutrient status. *Journal of Experimental Botany* **68**(7): 1399-1410.

González-Navarro OE, Griffiths S, Molero G, Reynolds MP, Slafer GA. 2015. Dynamics of floret development determining differences in spike fertility in an elite population of wheat. *Field Crops Research* **172**: 21-31.

Gonzalez FG, Miralles DJ, Slafer GA. 2011. Wheat floret survival as related to pre-anthesis spike growth. *Journal of Experimental Botany* **62**(14): 4889-4901.

González FG, Slafer GA, Miralles DJ. 2003a. Floret development and spike growth as affected by photoperiod during stem elongation in wheat. *Field Crops Research* **81**: 29-38.

González FG, Slafer GA, Miralles DJ. 2003b. Floret development and spike growth as affected by photoperiod during stem elongation in wheat. *Field Crops Research* **81**(1): 29-38.

González FG, Slafer GA, Miralles DJ. 2003c. Grain and floret number in response to photoperiod during stem elongation in fully and slightly vernalized wheats *Field Crops Research* **81** 17-27.

González FG, Slafer GA, Miralles DJ. 2005a. Floret development and survival in wheat plants exposed to contrasting photoperiod and radiation environments during stem elongation. *Functional Plant Biology* **32**(3).

González FG, Slafer GA, Miralles DJ. 2005b. Photoperiod during stem elongation in wheat: is its impact on fertile floret and grain number determination similar to that of radiation? *Functional Plant Biology* **32**(3): 181-188.

González FG, Slafer GA, Miralles DJ. 2006. Pre-anthesis development and number of fertile florets in wheat as affected by photoperiod sensitivity genes *Ppd-D1* and *Ppd-B1*. *Euphytica* **146**(3): 253-269.

Görlich D, Kutay U. 1999. Transport between the cell nucleus and the cytoplasm. *Annual Review of Cell and Developmental Biology* **15(1)**: 607–660.

Griffey C, Brooks W, Kurantz M, Thomason W, Taylor F, Obert D, Moreau R, Flores R, Sohn M, Hicks K. 2010. Grain composition of Virginia winter barley and implications for use in feed, food, and biofuels production. *Journal of Cereal Science* **51**(1): 41-49.

Guo Z, Chen D, Alqudah AM, Roder MS, Ganal MW, Schnurbusch T. 2017. Genome-wide association analyses of 54 traits identified multiple loci for the determination of floret fertility in wheat. *New Phytologist* **214**(1): 257-270.

Guo Z, Schnurbusch T. 2015. Variation of floret fertility in hexaploid wheat revealed by tiller removal. *Journal of Experimental Botany* **66**(19): 5945-5958.

Guo Z, Schnurbusch T. 2015. Variation of floret fertility in hexaploid wheat revealed by tiller removal. *Journal of Experimental Botany* **66**(19): 5945-5958.

Guo Z, Schnurbusch T. 2016. Costs and benefits of awns. *Journal of Experimental Botany* **67**(9): 2533.

Guo Z, Slafer GA, Schnurbusch T. 2016. Genotypic variation in spike fertility traits and ovary size as determinants of floret and grain survival rate in wheat. *Journal of Experimental Botany* **67**(14): 4221-4230.

Gupta PK, Kulwal PL, Jaiswal V 2014. Association mapping in crop plants: opportunities and challenges. In: Friedmann T, Dunlap JC, Goodwin SF eds. *Advances in Genetics*: Academic Press, 109-147.

Hall D, Tegström C, Ingvarsson PK. 2010. Using association mapping to dissect the genetic basis of complex traits in plants. *Briefings in Functional Genomics* **9**(2): 157-165.

Harlan JR, de Wet JMJ. 1971. Toward a rational classification of cultivated plants. *Taxon* **20(4)**: 509-517.

Harlan JR, Zohary D. 1966. Distribution of wild wheats and barley. *Science* **153**(3740): 1074-1080.

Hastie TJ. 2017. *Generalized additive models*: Routledge.

Haun JR. 1973. Visual quantification of wheat development. *Agronomy Journal* **65**(1): 116-119.

He S, Schulthess AW, Mirdita V, Zhao Y, Korzun V, Bothe R, Ebmeyer E, Reif JC, Jiang Y. 2016. Genomic selection in a commercial winter wheat population. *Theoretical and Applied Genetics* **129**: 641-651.

Heng Y, Wu C, Long Y, Luo S, Ma J, Chen J, Liu J, Zhang H, Ren Y, Wang M, et al. 2018. *OsALMT7* maintains panicle size and grain yield in rice by mediating malate transport. *The Plant Cell* **30**(4): 889-906.

Herzig P, Maurer A, Draba V, Sharma R, Draicchio F, Bull H, Milne L, Thomas WT, Flavell AJ, Pillen K. 2018. Contrasting genetic regulation of plant development in wild barley grown in two European environments revealed by nested association mapping. *Journal of Experimental Botany* **69**(7): 1517-1531.

Heslot N, Yang HP, Sorrells ME, Jannink JL. 2012. Genomic selection in plant breeding: a comparison of models. *Crop Science* **52**(1): 146-160.

Hölzle A, Jonietz C, Törjek O, Altmann T, Binder S, Forner J. 2011. A RESTORER OF FERTILITY-like PPR gene is required for 5′-end processing of the *nad4* mRNA in mitochondria of *Arabidopsis thaliana*. *The Plant Journal* **65**(5): 737-744.

Huang X, Sang T, Zhao Q, Feng Q, Zhao Y, Li C, Zhu C, Lu T, Zhang Z, Li M. 2010. Genome-wide association studies of 14 agronomic traits in rice landraces. *Nature Genetics* **42**(11): 961.

Huang Y, Kamal R, Shanmugaraj N, Rutten T, Thirulogachandar V, Zhao S, Hoffie I, Hensel G, Rajaraman J, Moya YAT, et al. 2023. A molecular framework for grain number determination in barley. *Science Advances* **9**(9): eadd0324.

Hund A, Kronenberg L, Anderegg J, Yu K, Walter A 2019. Non-invasive field phenotyping of cereal development. *Advances in breeding techniques for cereal crops*: Burleigh Dodds Science Publishing, 249-292.

Jones H, Leigh FJ, Mackay I, Bower MA, Smith LM, Charles MP, Jones G, Jones MK, Brown TA, Powell W. 2008. Population-based resequencing reveals that the flowering time adaptation of cultivated barley originated east of the Fertile Crescent. *Molecular Biology and Evolution* **25**(10): 2211-2219.

K. IP, R. SN. 2011. Association genetics of complex traits in plants. *New Phytologist* **189**(4): 909-922.

Kamal R, Muqaddasi QH, Schnurbusch T. 2022a. Genetic association of spikelet abortion with spike, grain, and shoot traits in highly-diverse six-rowed barley. *Frontiers in Plant Science* **13**: 1015609.

Kamal R, Muqaddasi QH, Trautewig C, Huang Y, Shanmugaraj N, Melzer M, Tschiersch H, Chmielewska B, Szurman-Zubrzycka M, Szarejko I, et al. 2023. Insights into the genetic basis of pre-anthesis tip degeneration and related traits in barley. **submitted**.

Kamal R, Muqaddasi QH, Zhao Y, Schnurbusch T. 2021. Spikelet abortion in six-rowed barley is mainly influenced by final spikelet number, with potential spikelet number acting as a suppressor trait. *Journal of Experimental Botany*: erab529.

Kamal R, Muqaddasi QH, Zhao Y, Schnurbusch T. 2022b. Spikelet abortion in six-rowed barley is mainly influenced by final spikelet number, with potential spikelet number acting as a suppressor trait. *Journal of Experimental Botany* **73**(7): 2005-2020.

Kang J-W, Kabange NR, Phyo Z, Park S-Y, Lee S-M, Lee J-Y, Shin D, Cho JH, Park D-S, Ko J-M. 2020. Combined linkage mapping and Genome-Wide Association study identified QTLs associated with grain shape and weight in rice (*Oryza sativa* L.). *Agronomy* **10**(10): 1532.

Kang J, Park J, Choi H, Burla B, Kretzschmar T, Lee Y, Martinoia E. 2011. Plant ABC Transporters. *Arabidopsis Book/American Society of Plant Biologists* **9**: e0153.

Karunarathne SD, Han Y, Zhang X-Q, Zhou G, Hill CB, Chen K, Angessa T, Li C. 2020. Genome-wide association study and identification of candidate genes for nitrogen use efficiency in barley (*Hordeum vulgare* L.). *Frontiers in Plant Science* **11**: 1361.

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

Kilian B, Özkan H, Pozzi C, Salamini F 2009. Domestication of the *Triticeae* in the Fertile Crescent. *Genetics and Genomics of the Triticeae*, 81-119.

Kim SH, Arnold D, Lloyd A, Roux SJ. 2001. Antisense expression of an Arabidopsis ran binding protein renders transgenic roots hypersensitive to auxin and alters auxin-induced root growth and development by arresting mitotic progress. *The Plant Cell* **13**(12): 2619-2630.

Kirby E. 1988. Analysis of leaf, stem and ear growth in wheat from terminal spikelet stage to anthesis. *Field Crops Research* **18**(2-3): 127-140.

Kirby E, Appleyard M. 1984. Cereal development guide 2nd edition. *Stoneleigh: National Agricultural Centre Arable Unit*.

Kirby E, Appleyard M. 1987. Cereal development guide.

Kirby E, Faris D. 1972. The effect of plant density on tiller growth and morphology in barley. *The Journal of Agricultural Science* **78**(2): 281-288.

Kirby E, Riggs T. 1978. Developmental consequences of two-row and six-row ear type in spring barley: 2. Shoot apex, leaf and tiller development. *The Journal of Agricultural Science* **91**(1): 207-216.

Kirby EJM, Appleyard M 1987. Development and structure of the wheat plant. *Wheat Breeding*: Springer, Dordrecht, 287-311.

Kitchen BM, Rasmusson DC. 1983. Duration and Inheritance of Leaf Initiation, Spike Initiation, and Spike Growth in Barley 1. *Crop Science* **23**(5): 939-943.

Knott DR. 1986. Effect of genes for photoperiodism, semidwarfism, and awns on agronomic characters in a wheat cross. *Crop Science* **26**(6): 1158-1162.

Komatsuda T, Pourkheirandish M, He C, Azhaguvel P, Kanamori H, Perovic D, Stein N, Graner A, Wicker T, Tagiri A. 2007. Six-rowed barley originated from a mutation in a homeodomain-leucine zipper I-class homeobox gene. *Proceedings of the National Academy of Sciences* **104**(4): 1424-1429.

Komatsuda T, Tanno K-i, Salomon B, Bryngelsson T, Bothmer Rv. 1999. Phylogeny in the genus Hordeum based on nucleotide sequences closely linked to the vrs1 locus (row number of spikelets). *Genome* **42**(5): 973-981.

Koppolu R, Chen S, Schnurbusch T. 2022. Evolution of inflorescence branch modifications in cereal crops. *Current Opinion in Plant Biology* **65**: 102168.

Koppolu R, Schnurbusch T. 2019. Developmental pathways for shaping spike inflorescence architecture in barley and wheat. *Journal of Integrative Plant Biology* **61**(3): 278-295.

Kubo T, Arakawa T, Honma Y, Kitazaki K. 2020. What does the molecular genetics of different types of restorer-of-fertility genes imply? *Plants* **9**(3): 361.

Lê S, Josse J, Husson F. 2008. FactoMineR: an R package for multivariate analysis. *Journal of Statistical Software* **25**(1): 1-18.

Li W-c, Meng Z-d, Zhang F-j, Sun Q, Mu C-h, Ding Z-h, Guo J-m, Xu Z. 2008. Gene effect and genetic variation of maize ear tip-barren trait. *Acta Agriculturae Boreali-Sinica* **23**: 146-148.

Li X, Sun M, Liu S, Teng Q, Li S, Jiang Y. 2021. Functions of PPR proteins in plant growth and development. *International Journal of Molecular Sciences* **22**(20): 11274.

Li X, Wang H, Li H, Zhang L, Teng N, Lin Q, Wang J, Kuang T, Li Z, Li B. 2006. Awns play a dominant role in carbohydrate production during the grain-filling stages in wheat (Triticum aestivum). *Physiologia Plantarum* **127**(4): 701-709.

Lipka AE, Gore MA, Magallanes-Lundback M, Mesberg A, Lin H, Tiede T, Chen C, Buell CR, Buckler ES, Rocheford T. 2013. Genome-wide association study and pathway-level analysis of tocochromanol levels in maize grain. *G3: Genes, Genomes, Genetics* **3**(8): 1287-1299.

Liu P, Qi M, Wang Y, Chang M, Liu C, Sun M, Yang W, Ren H. 2014. Arabidopsis RAN1 mediates seed development through its parental ratio by affecting the onset of endosperm cellularization. *Molecular Plant* **7**(8): 1316-1328.

Liu YJ, Liu X, Chen H, Zheng P, Wang W, Wang L, Zhang J, Tu J. 2017. A plastid-localized pentatricopeptide repeat protein is required for both pollen development and plant growth in rice. *Scientific Reports* **7**(1): 11484.

Lorenzana RE, Bernardo R. 2009. Accuracy of genotypic value predictions for marker-based selection in biparental plant populations. *Theoretical and Applied Genetics* **120**: 151-161.

Lund B, Ortiz R, Skovgaard I, Waugh R, Andersen S. 2003. Analysis of potential duplicates in barley gene bank collections using re-sampling of microsatellite data. *Theoretical and Applied Genetics* **106**(6): 1129-1138.

Luo A, Qian Q, Yin H, Liu X, Yin C, Lan Y, Tang J, Tang Z, Cao S, Wang X, et al. 2006. EUI1, encoding a putative cytochrome P450 monooxygenase, regulates internode elongation by modulating gibberellin responses in rice. *Plant Cell Physiology* **47**(2): 181-191.

Manna S. 2015. An overview of pentatricopeptide repeat proteins and their applications. *Biochimie* **113**: 93-99.

Mansfeld R. 1950. Das morphologische System der Saatgerste, *Hordeum vulgare* L. sl. *Der Züchter* **20**: 8-24.

Marcillo GS, Martin NF, Diers BW, Da Fonseca Santos M, Leles EP, Chigeza G, Francischini JH. 2021. Implementation of a Generalized Additive Model (GAM) for Soybean Maturity Prediction in African Environments. *Agronomy* **11**(6): 1043.

Martin J, Carver BF, Hunger RM, Cox T. 2003. Contributions of leaf rust resistance and awns to agronomic and grain quality performance in winter wheat.

Masood MA, Qamar M, Raza I. 2018. Comparative efficiency of alpha lattice design versus randomized complete block design in wheat field trials. *International Journal of Scientific and Engineering Research* **9**(11): 646-650.

McKenzie H. 1972. Adverse influence of awns on yield of wheat. *Canadian Journal of Plant Science* **52**(1): 81-87.

Melchior F. 2001. Ran GTPase cycle: One mechanism — two functions. *Current Biology* **11(7)**: 257-260.

Melonek J, Zhou R, Bayer PE, Edwards D, Stein N, Small I. 2019. High intraspecific diversity of Restorer-of-fertility-like genes in barley. *The Plant Journal* **97**(2): 281-295.

Meng Z-D, Zhang F-J, Ding Z-H, Qi S, Liming W, Guo Q-F, Wang H-G. 2007. Inheritance of ear tip-barrenness trait in maize. *Agricultural Sciences in China* **6**(5): 628-633.

Meuwissen TH, Hayes BJ, Goddard M. 2001. Prediction of total genetic value using genome-wide dense marker maps. *Genetics* **157(4)**: 1819-1829.

Milner SG, Jost M, Taketa S, Mazón ER, Himmelbach A, Oppermann M, Weise S, Knüpffer H, Basterrechea M, König P, et al. 2019. Genebank genomics highlights the diversity of a global barley collection. *Nature Genetics* **51**(2): 319-326.

Miralles DJ, Katz SD, Colloca A, Slafer GA. 1998. Floret development in near isogenic wheat lines differing in plant height. *Field Crops Research* **59**(1): 21-30.

Miralles DJ, Richards RA. 2000. Responses of Leaf and Tiller Emergence and Primordium Initiation in Wheat and Barley to Interchanged Photoperiod. *Annals of Botany* **85**(5): 655-663.

Morrell PL, Clegg MT. 2007. Genetic evidence for a second domestication of barley (*Hordeum vulgare*) east of the Fertile Crescent. *Proceedings of the National Academy of Sciences* **104(9)**: 3289-3294.

Motzo R, Giunta F. 2002. Awnedness affects grain yield and kernel weight in near-isogenic lines of durum wheat. *Australian Journal of Agricultural Research* **53**(12): 1285-1293.

Mulki MA, Bi X, von Korff M. 2018. FLOWERING LOCUS T3 controls spikelet initiation but not floral development. *Plant Physiology* **178**(3): 1170-1186.

Muqaddasi QH, Brassac J, Ebmeyer E, Kollers S, Korzun V, Argillier O, Stiewe G, Plieske J, Ganal MW, Röder MS. 2020. Prospects of GWAS and predictive breeding for european winter wheat's grain protein content, grain starch content, and grain hardness. *Scientific Reports* **10**(1): 1-17.

Muqaddasi QH, Brassac J, Koppolu R, Plieske J, Ganal MW, Röder MS. 2019. *TaAPO-A1*, an ortholog of rice *ABERRANT PANICLE ORGANIZATION 1*, is associated with total spikelet number per spike in elite European hexaploid winter wheat (*Triticum aestivum* L.) varieties. *Scientific Reports* **9**(1): 1-12.

Neale DB, Savolainen O. 2004. Association genetics of complex traits in conifers. *Trends in Plant Science* **9**(7): 325-330.

Nevo E, Beiles A, Zohary D. 1986. Genetic resources of wild barley in the Near East: structure, evolution and application in breeding. *Biological Journal of the Linnean Society* **27**(4): 355-380.

Nguyen T-T, Planchard N, Dahan J, Arnal N, Balzergue S, Benamar A, Bertin P, Brunaud V, Dargel-Graffin C, Macherel D. 2021. A case of gene fragmentation in plant mitochondria fixed by the selection of a compensatory restorer of fertility-like PPR gene. *Molecular Biology and Evolution* **38**(8): 3445-3458.

Nielsen E. 2020. The small GTPases superfamily in plants: A conserved regulatory module with novel functions. *Annual Reveiw of Plant Biology* **71**: 247-272.

Nordborg M, Tavaré S. 2002. Linkage disequilibrium: what history has to tell us. *Trends in Genetics* **18**(2): 83-90.

Ochagavia H, Prieto P, Savin R, Griffiths S, Slafer G. 2018. Dynamics of leaf and spikelet primordia initiation in wheat as affected by Ppd-1a alleles under field conditions. *Journal of Experimental Botany* **69**(10): 2621-2631.

Orabi J, Backes G, Wolday A, Yahyaoui A, Jahoor A. 2007. The Horn of Africa as a centre of barley diversification and a potential domestication site. *Theoretical and Applied Genetics* **114**: 1117-1127.

Pasam RK, Sharma R, Malosetti M, van Eeuwijk FA, Haseneyer G, Kilian B, Graner A. 2012. Genome-wide association studies for agronomical traits in a world wide spring barley collection. *BMC Plant Biology* **12**(1): 1-22.

Pei Y, Deng Y, Zhang H, Zhang Z, Liu J, Chen Z, Cai D, Li K, Du Y, Zang J, et al. 2022. *EAR APICAL DEGENERATION1* regulates maize ear development by maintaining malate supply for apical inflorescence. *The Plant Cell* **34(6)**: 2222-2241.

Petr Kalab RTPaMD. 1999. The Ran GTPase regulates mitotic spindle assembly. *Current Biology* **9**: 481-484.

Pfalz J, Bayraktar OA, Prikryl J, Barkan A. 2009. Site-specific binding of a PPR protein defines and stabilizes 5′ and 3′ mRNA termini in chloroplasts. *The EMBO journal* **28**(14): 2042-2052.

Pham G, Shin DM, Kim Y, Kim SH. 2022. Ran-GTP/-GDP-dependent nuclear accumulation of NONEXPRESSOR OF PATHOGENESIS-RELATED GENES1 and TGACG-BINDING FACTOR2 controls salicylic acid-induced leaf senescence. *Plant Physiology* **189**(3): 1774-1793.

Pins J, Kaur H. 2006. A review of the effects of barley beta-glucan on cardiovascular and diabetic risk. *Cereal Foods World* **51**(1): 8.

Pourkheirandish M, Hensel G, Kilian B, Senthil N, Chen G, Sameri M, Azhaguvel P, Sakuma S, Dhanagond S, Sharma R, et al. 2015. Evolution of the Grain Dispersal System in Barley. *Cell* **162**(3): 527-539.

Prieto P, Ochagavía H, Griffiths S, Slafer GA. 2020. Earliness per se × temperature interaction: consequences on leaf, spikelet, and floret development in wheat. *Journal of Experimental Botany* **71**(6): 1956-1968.

Prieto P, Ochagavía H, Savin R, Griffiths S, Slafer GA. 2018. Dynamics of floret initiation/death determining spike fertility in wheat as affected by *Ppd* genes under field conditions. *Journal of Experimental Botany* **69**(10): 2633-2645.

Prikryl J, Rojas M, Schuster G, Barkan A. 2011. Mechanism of RNA stabilization and translational activation by a pentatricopeptide repeat protein. *Proceedings of the National Academy of Sciences* **108**(1): 415-420.

Qilin T, Tingzhao R, Hunag Y. 1999. The study of barren ear tip in maize. II. The growth rhythm of flower and kernel in different region of maize ear. *Journal of Sichuan Agricultural University* **17(2)**: 162-166.

Qin Z, Wu YN, Sun TT, Ma T, Xu M, Pang C, Li SW, Li S. 2022. *Arabidopsis* RAN GTPases are critical for mitosis during male and female gametogenesis. *FEBS letters*.

Quinde Z, Ullrich SE, Baik BK. 2004. Genotypic variation in color and discoloration potential of barley-based food products. *Cereal Chemistry* **81**(6): 752-758.

Ramamoorthy R, Jiang SY, Ramachandran S. 2011. *Oryza sativa* cytochrome P450 family member OsCYP96B4 reduces plant height in a transcript dosage dependent manner. *PLoS One* **6**(11): e28069.

Ramsay L, Comadran J, Druka A, Marshall DF, Thomas WT, Macaulay M, MacKenzie K, Simpson C, Fuller J, Bonar N. 2011. *INTERMEDIUM-C*, a modifier of lateral spikelet fertility in barley, is an ortholog of the maize domestication gene *TEOSINTE BRANCHED 1*. *Nature Genetics* **43**(2): 169-172.

Rawson HM, Evans LT. 1970. The pattern of grain growth within the ear of wheat. *Australian Journal of Biological Sciences* **23 (4)**: 753-764.

Rebetzke G, Bonnett D, Reynolds M. 2016. Awns reduce grain number to increase grain size and harvestable yield in irrigated and rainfed spring wheat. *Journal of Experimental Botany* **67**(9): 2573-2586.

Remington DL, Thornsberry JM, Matsuoka Y, Wilson LM, Whitt SR, Doebley J, Kresovich S, Goodman MM, Buckler ES. 2001. Structure of linkage disequilibrium and phenotypic associations in the maize genome. *Proceedings of the National Academy of Sciences* **98**(20): 11479-11484.

Ren X, Nevo E, Sun D, Sun G. 2013. Tibet as a potential domestication center of cultivated barley of China. *PLoS One* **8**(5): e62700.

Reynolds M, Foulkes J, Furbank R, Griffiths S, King J, Murchie E, Parry M, Slafer G. 2012. Achieving yield gains in wheat. *Plant, Cell & Environment* **35**(10): 1799-1823.

Risch N, Merikangas K. 1996. The future of genetic studies of complex human diseases. *Science* **273**(5281): 1516.

Rosseel Y. 2012. Lavaan: An R package for structural equation modeling and more. Version 0.5–12 (BETA). *Journal of Statistical Software* **48**(2): 1-36.

Sadras VO, Denison RF. 2009. Do plant parts compete for resources? An evolutionary viewpoint. *New Phytologist* **183**(3): 565-574.

Sadras VO, Slafer GA. 2012. Environmental modulation of yield components in cereals: Heritabilities reveal a hierarchy of phenotypic plasticities. *Field Crops Research* **127**: 215-224.

Sakai S. 2007. A new hypothesis for the evolution of overproduction of ovules: an advantage of selective abortion for females not associated with variation in genetic quality of the resulting seeds. *Evolution* **61**(4): 984-993.

Sakuma S, Golan G, Guo Z, Ogawa T, Tagiri A, Sugimoto K, Bernhardt N, Brassac J, Mascher M, Hensel G. 2019. Unleashing floret fertility in wheat through the mutation of a homeobox gene. *Proceedings of the National Academy of Sciences* **116**(11): 5182-5187.

Sakuma S, Koppolu R. 2023. Form follows function in Triticeae inflorescences. *Breeding Science* **73**(1): 46-56.

Sakuma S, Pourkheirandish M, Hensel G, Kumlehn J, Stein N, Tagiri A, Yamaji N, Ma JF, Sassa H, Koba T. 2013. Divergence of expression pattern contributed to neofunctionalization of duplicated HD-Zip I transcription factor in barley. *New Phytologist* **197**(3): 939-948.

Sakuma S, Schnurbusch T. 2019. Of floral fortune: tinkering with the grain yield potential of cereal crops. *New Phytologist*.

Sallam AH, Smith KP. 2016. Genomic selection performs similarly to phenotypic selection in barley. *Crop Science* **56**(6): 2871-2881.

Sanchez-Bragado R, Kim J, Rivera-Amado C, Molero G, Araus J, Savin R, Slafer GA. 2020. Are awns truly relevant for wheat yields? A study of performance of awned/awnless isogenic lines and their response to source–sink manipulations. *Field Crops Research* **254**: 107827.

Sasani S, Hemming MN, Oliver SN, Greenup A, Tavakkol-Afshari R, Mahfoozi S, Poustini K, Sharifi H-R, Dennis ES, Peacock WJ. 2009. The influence of vernalization and daylength on expression of flowering-time genes in the shoot apex and leaves of barley (*Hordeum vulgare* L.). *Journal of Experimental Botany* **60**(7): 2169-2178.

Schaller C, Qualset C. 1975. Isogenic Analysis of Productivity in Barley: Interaction of Genes Affecting Awn Length and Leaf-spotting. *Crop Science* **15**(3): 378-382.

Schaller C, Qualset CO, Rutger JN. 1972. Isogenic Analysis of the Effects of the Awn on Productivity of Barley. *Crop Science* **12**(4): 531-535.

Schmid KJ, Thorwarth P. 2014. Genomic selection in barley breeding. *Biotechnological approaches to barley improvement*: 367-378.

Sergio F, Blas J, López L, Tanferna A, Díaz-Delgado R, Donázar JA, Hiraldo F. 2011. Coping with uncertainty: breeding adjustments to an unpredictable environment in an opportunistic raptor. *Oecologia* **166**: 79-90.

Shanmugaraj N, Rajaraman J, Kale S, Kamal R, Huang Y, Thirulogachandar V, Garibay-Hernández A, Budhagatapalli N, Moya YT, Hajirezaei M-R, et al. 2023. Multilayered regulation of developmentally programmed pre-anthesis tip degeneration of the barley inflorescence. *The Plant Cell*: koad164.

Shaw LM, Lyu B, Turner R, Li C, Chen F, Han X, Fu D, Dubcovsky J. 2018. FLOWERING LOCUS T2 regulates spike development and fertility in temperate cereals. *Journal of Experimental Botany* **70**(1): 193-204.

Slafer GA, Molina-Cano JL, Savin R, Araus JL, Romagosa I. 2002. *Barley science: Recent advances from molecular biology to agronomy of yield and quality*: CRC Press.

Sreenivasulu N, Schnurbusch T. 2012. A genetic playground for enhancing grain number in cereals. *Trends in Plant Science* **17**(2): 91-101.

Stephenson AG. 1981. Flower and fruit abortion: proximate causes and ultimate functions. . *Annual Review of Ecology and Systematics,* **12(1)**: 253-279.

Streiner DL. 2005. Finding our way: an introduction to path analysis. *The Canadian Journal of Psychiatry* **50**(2): 115-122.

Su X, Yan X, Tsai CL. 2012. Linear regression. *Wiley Interdisciplinary Reviews: Computational Statistics* **4**(3): 275-294.

Suhr D 2008. Step your way through path analysis. *Western users of SAS software conference proceedings*. 2017.

Sun D, Robbins K, Morales N, Shu Q, Cen H. 2022. Advances in optical phenotyping of cereal crops. *Trends in Plant Science* **27**(2): 191-208.

Sun G, Lu H, Zhao Y, Zhou J, Jackson R, Wang Y, Xu LX, Wang A, Colmer J, Ober E, et al. 2022. AirMeasurer: open-source software to quantify static and dynamic traits derived from multiseason aerial phenotyping to empower genetic mapping studies in rice. *New Phytologist* **236**(4): 1584-1604.

Szűcs P, Skinner JS, Karsai I, Cuesta-Marcos A, Haggard KG, Corey AE, Chen TH, Hayes PM. 2007. Validation of the V*RN-H2/VRN-H1* epistatic model in barley reveals that intron length variation in *VRN-H1* may account for a continuum of vernalization sensitivity. *Molecular Genetics and Genomics* **277**: 249-261.

Team RC. 2013. R: A language and environment for statistical computing.

Terzi V, Tumino G, Pagani D, Rizza F, Ghizzoni R, Morcia C, Stanca A. 2017. Barley Developmental Mutants: The High Road to Understand the Cereal Spike Morphology. *Diversity* **9**(2).

Thabet SG, Moursi YS, Karam MA, Börner A, Alqudah AM. 2020. Natural variation uncovers candidate genes for barley spikelet number and grain yield under drought stress. *Genes* **11**(5): 533.

Thirulogachandar V, Koppolu R, Schnurbusch T. 2020. Modes of grain number determination differentiate barley row-types. *bioRxiv*.

Thirulogachandar V, Koppolu R, Schnurbusch T. 2021. Strategies of grain number determination differentiate barley row types. *J Exp Bot* **72**(22): 7754-7768.

Thirulogachandar V, Schnurbusch T. 2021. 'Spikelet stop' determines the maximum yield potential stage in barley *Journal of Experimental Botany* **erab342**: dx.doi.org/10.1093/jxb/erab1342.

Thompson FT, Levine DU. 1997. Examples of easily explainable suppressor variables in multiple regression research. *Multiple Linear Regression Viewpoints* **24**(1): 11-13.

Traas J, Vernoux T. 2002. The shoot apical meristem: the dynamics of a stable structure. *Philosophical Transactions of the Royal Society B: Biological Sciences* **357**(1422): 737-747.

Tricase C, Amicarelli V, Lamonaca E, Leonardo Rana R. 2018. Economic analysis of the barley market and related uses. *Grasses as food and feed* **10**: 25-46.

Turner A, Beales J, Faure S, Dunford RP, Laurie DA. 2005. The pseudo-response regulator *Ppd-H1* provides adaptation to photoperiod in barley. *Science* **310**(5750): 1031-1034.

Vervelde G. 1953. The agricultural value of awns in cereals. *Netherlands Journal of Agricultural Science* **1**(1): 2-10.

von Zitzewitz J, Szűcs P, Dubcovsky J, Yan L, Francia E, Pecchioni N, Casas A, Chen TH, Hayes PM, Skinner JS. 2005. Molecular and structural characterization of barley vernalization genes. *Plant Molecular Biology* **59**: 449-467.

Waddington SR, Cartwright PM, Wall PC. 1983. A quantitative scale of spike initial and pistil development in barley and wheat. *Annals of Botany* **51**(1): 119-130.

Wang Y, Ye H, Bai J, Ren F. 2021. The regulatory framework of developmentally programmed cell death in floral organs: A review. *Plant Physiology and Biochemistry* **158**: 103-112.

Wang YL, Ye H, Liu L, Wu JH, Ru WM, Sun GL. 2019. Molecular insights on the domestication of barley (*Hordeum vulgare* L.). *Critical Reviews in Plant Sciences* **38**(4): 280-294.

Wang Z-q, Zhang W-y, Yang J-c. 2018. Physiological mechanism underlying spikelet degeneration in rice. *Journal of Integrative Agriculture* **17**(7): 1475-1481.

Wei K, Han P. 2016. Pentatricopeptide repeat proteins in maize. *Molecular Breeding* **36**(12): 1-18.

Weyhrich RA, Carver BF, Smith EL. 1994. Effect of awn suppression on grain yield and agronomic traits in hard red winter wheat. *Crop Science* **34**(4): 965-969.

Whingwiri E, Stern W. 1982. Floret survival in wheat: significance of the time of floret initiation relative to terminal spikelet formation. *The Journal of Agricultural Science* **98**(2): 257-268.

Whingwiri EE, Stern WR. 1982. Floret survival in wheat: significance of the time of floret initiation relative to terminal spikelet formation. *Journal of Agricultural Sciences (Cambridge)* **98**: 257-268.

Wolde GM, Schnurbusch T. 2019. Inferring vascular architecture of the wheat spikelet based on resource allocation in the branched headt (*bht-A1*) near isogenic lines. *Functional Plant Biology* **46**(11): 1023-1035.

Wood S, Wood MS. 2015. Package 'mgcv'. *R package version* **1**: 29.

Wood SN. 2006. *Generalized additive models: an introduction with R*: chapman and hall/CRC.

Wood SN, Pya N, Säfken B. 2016. Smoothing parameter and model selection for general smooth models. *Journal of the American Statistical Association* **111**(516): 1548-1563.

Xu P, Cai W. 2014. *RAN1* is involved in plant cold resistance and development in rice (*Oryza sativa*). *Journal of Experimental Botany* **65**(12): 3277-3287.

Xu P, Zang A, Chen H, Cai W. 2016. The Small G protein *AtRAN1* regulates vegetative growth and stress tolerance in *Arabidopsis thaliana*. *PLoS One* **11**(6): e0154787.

Yang W, Zheng L, He Y, Zhu L, Chen X, Tao Y. 2020. Fine mapping and candidate gene prediction of a major quantitative trait locus for tassel branch number in maize. *Gene* **757**: 144928.

Yano K, Morinaka Y, Wang F, Huang P, Takehara S, Hirai T, Ito A, Koketsu E, Kawamura M, Kotake K, et al. 2019. GWAS with principal component analysis identifies a gene comprehensively controlling rice architecture. *Proceedings of the National Academy of Sciences* **116**(42): 21262-21267.

Youssef HM. 2016. *Genotypic and phenotypic analysis of the spike row-type in barley (Hordeum vulgare L.).*Cumulative dissertation, Martin-Luther-Universität Halle-Wittenberg.

Youssef HM, Eggert K, Koppolu R, Alqudah AM, Poursarebani N, Fazeli A, Sakuma S, Tagiri A, Rutten T, Govind G, et al. 2017. *VRS2* regulates hormone-mediated inflorescence patterning in barley. *Nature Genetics* **49**(1): 157-161.

Yu J, Buckler ES. 2006. Genetic association mapping and genome organization of maize. *Current Opinion in Biotechnology* **17**(2): 155-160.

Yuan J, Wang X, Zhao Y, Khan NU, Zhao Z, Zhang Y, Wen X, Tang F, Wang F, Li Z. 2020. Genetic basis and identification of candidate genes for salt tolerance in rice by GWAS. *Scientific Reports* **10**(1): 1-9.

Zadoks JC, Chang TT, Konzak CF. 1974. A decimal code for the growth stages of cereals. *Weed Research* **14**(6): 415-421.

Zafar SA, Patil SB, Uzair M, Fang J, Zhao J, Guo T, Yuan S, Uzair M, Luo Q, Shi J, et al. 2020. *DEGENERATED PANICLE AND PARTIAL STERILITY 1* (*DPS1*) encodes a cystathionine beta-synthase domain containing protein required for anther cuticle and panicle development in rice. *New Phytologist* **225**(1): 356-375.

Zhang Q, Xu Y, Huang J, Zhang K, Xiao H, Qin X, Zhu L, Zhu Y, Hu J. 2020. The rice pentatricopeptide repeat protein PPR756 is involved in pollen development by affecting multiple RNA editing in mitochondria. *Frontiers in Plant Science* **11**: 749.

Zhang Z, Li J, Hu N, Li W, Qin W, Li J, Gao Y, Liu Y, Sun Z, Yu K. 2021. Spike growth affects spike fertility through the number of florets with green anthers before floret abortion in wheat. *Field Crops Research* **260**: 108007.

Zhou D, Shen W, Cui Y, Liu Y, Zheng X, Li Y, Wu M, Fang S, Liu C, Tang M, et al. 2021. *APICAL SPIKELET ABORTION* (*ASA*) controls apical panicle development in rice by regulating salicylic acid biosynthesis. *Frontiers in Plant Science* **12**: 636877.

Zhu Y, Chu J, Dai X, He M. 2019. Delayed sowing increases grain number by enhancing spike competition capacity for assimilates in winter wheat. *European Journal of Agronomy* **104**: 49-62.

Zhu Z-c, Luo S, Lei B, Li X-y, Cheng Z-j. 2022. Locus *TUTOU2* determines the panicle apical abortion phenotype of rice (*Oryza sativa* L.) in *tutou2* mutant. *Journal of Integrative Agriculture* **21**(3): 621-630.

Zohary D, Hopf M. 2000. *Domestication of plants in the Old World: The origin and spread of cultivated plants in West Asia, Europe and the Nile Valley*: Oxford university press.

Zscheischler J, Fischer EM. 2020. The record-breaking compound hot and dry 2018 growing season in Germany. *Weather and Climate Extremes* **29**: 100270.

Zwirek M, Waugh R, McKim SM. 2019. Interaction between row-type genes in barley controls meristem determinacy and reveals novel routes to improved grain. *New Phytologist* **221**(4): 1950-1965.

SUPPLEMENTARY TABLES

Table S2.1 ANOVA for the investigated traits in 2018

	(A) Potential spikelet number							
	Df	**Sum Sq**	**Mean Sq**	***F*-value**	**Pr (>*F*)**	**Sig.**	**σ^2**	**SD.**
genotype	287	268616	936	11.04	<2e-16	***	121.96	11.04
replication	1	11454	11454	135.1	<2e-16	***	23.21	4.82
Residuals	1733	146919	85				79.77	8.93
	(B) Final spikelet number							
	Df	**Sum Sq**	**Mean Sq**	***F*-value**	**Pr (>*F*)**	**Sig.**	**σ^2**	**SD.**
genotype	287	161978	564	17.53	<2e-16	***	69.79	8.35
replication	1	4597	4597	142.83	4.28E-08	***	9.29	3.05
Residuals	1867	60094	32				29.96	5.47
	(C) Heading date							
	Df	**Sum Sq**	**Mean Sq**	***F*-value**	**Pr (>*F*)**	**Sig.**	**σ^2**	**SD.**
genotype	281	16900	60.14	3.26	5.44E-15	***	25.38	5.04
replication	1	202	202.24	10.96	0.0012	**	2.44	1.56
Residual	156	2879	18.45				19.48	4.41

Df = Degree of freedom; Sq = squares; Sig = significance codes; σ^2 = variance; SD = standard deviation; ***, **, and * = significant at the 0.001, 0.01, and 0.05 probability level, respectively.

Table S2.2 ANOVA for the investigated traits in 2019

	(A) Potential spikelet number							
	Df	**Sum Sq**	**Mean Sq**	***F*-value**	**Pr (>*F*)**	**Sig.**	**σ^2**	**SD.**
genotype	416	479109	1151.7	36.62	<2e-16	***	128.99	11.36
replication	1	82	82.4	2.62	0.105	–	–	–
replication(block)	1	1443	1442.8	45.88	1.48E-11	***	3.17	1.78
Residuals	3245	102042	31.4				29.57	5.44
	(B) Final spikelet number							
	Df	**Sum Sq**	**Mean Sq**	***F*-value**	**Pr (>*F*)**	**Sig.**	**σ^2**	**SD.**
genotype	416	263064	632.4	33.81	<2e-16	***	72.18	8.49
replication	1	1126	1125.7	60.19	1.14E-14	***	0.4	0.64
replication(block)	1	106	106.2	5.68	0.0172	*	3.33	1.82
Residuals	3280	61346	18.7				16.33	4.04
	(C) Heading date							
	Df	**Sum Sq**	**Mean Sq**	**F-value**	**Pr (>F)**	**Sig.**	**σ^2**	**SD.**
genotype	411	37592	91.47	23.28	<2e-16	***	17.28	4.16
replication	1	6	5.59	1.43	0.23	–	0.04	0.2
replication(block)	1	0	0	0.001	0.98	–	1.67	0.41
Residuals	802	3152	3.93				12.93	3.59
	(D) Plant height							
	Df	**Sum Sq**	**Mean Sq**	***F*-value**	**Pr (>*F*)**	**Sig.**	**σ^2**	**SD.**
genotype	416	561977	1351	62.83	<2e-16	***	149.78	12.24
replication	1	4176	4176	194.22	<2e-16	***	1.84	1.36
replication(block)	1	1000	1000	46.53	1.07E-11	***	4.99	2.23
Residual	3280	70518	21				18.01	4.24

Df = Degree of freedom; Sq = squares; Sig = significance codes; σ^2 = variance; SD = standard deviation;
***, **, and * = significant at the 0.001, 0.01, and 0.05 probability level, respectively.

Table S2.3 ANOVA for the investigated traits in 2020

	(A) Potential spikelet number							
	Df	**Sum Sq**	**Mean Sq**	***F*-value**	**Pr (>*F*)**	**Sig.**	**σ^2**	**SD.**
genotype	416	26718	64.23	4.89	<2e-16	***	131.35	11.46
replication	1	8	7.74	0.59	0.44	–	–	–
replication(block)	1	3	3.1	0.24	0.63	–	1.25	1.12
Residuals	824	10822	13.13				17.4	4.17
	(B) Final spikelet number							
	Df	**Sum Sq**	**Mean Sq**	***F*-value**	**Pr (>*F*)**	**Sig.**	**σ^2**	**SD.**
genotype	416	309617	744.3	88.86	<2e-16	***	83.42	9.13
replication	1	1115	1115.4	133.16	<2e-16	***	0.74	0.86
replication(block)	1	86	185.6	510.22	0.001	**	0.8	0.89
Residuals	3310	27724	8.4				7.57	2.75
	(C) Heading date							
	Df	**Sum Sq**	**Mean Sq**	**F-value**	**Pr (>F)**	**Sig.**	**σ^2**	**SD.**
genotype	416	26718	64.23	4.89	<2e-16	***	17.28	4.16
replication	1	8	7.74	0.59	0.44	–	0.04	0.2
replication(block)	1	3	3.1	0.24	0.63	–	1.67	0.41
Residuals	824	10822	13.13				12.93	3.59
	(D) Plant height							
	Df	**Sum Sq**	**Mean Sq**	***F*-value**	**Pr (>*F*)**	**Sig.**	**σ^2**	**SD.**
genotype	416	502014	1207	93	<2e-16	***	134.74	11.61
replication	1	5077	5077	391.25	<2e-16	***	2.14	1.46
replication(block)	1	343	343	26.44	2.88E-07	***	2.09	1.45
Residual	3310	42953	13				11.34	3.37

Df = Degree of freedom; Sq = squares; Sig = significance codes; σ^2 = variance; SD = standard deviation; ***, **, and * = significant at the 0.001, 0.01, and 0.05 probability level, respectively.

Table S2.4 Across years ANOVA for the investigated traits

	(A) Potential spikelet number							
	Df	**Sum Sq**	**Mean Sq**	***F*-value**	**Pr (>*F*)**	**Sig.**	**σ^2**	**SD.**
Genotype	416	903668	2172	86.813	<2e-16	***	115.45	10.75
Year	1	11411	11411	456.023	<2e-16	***	3.05	1.75
Genotype×Year	416	63298	152	6.081	<2e-16	***	14.68	3.83
Year×Replication×Block	1	614	614	24.531	7.57E-07	***	2.17	1.47
Residuals	663	162822	25				23.49	4.85
	(B) Final spikelet number							
	Df	**Sum Sq**	**Mean Sq**	***F*-value**	**Pr (>*F*)**	**Sig.**	**σ^2**	**SD.**
Genotype	416	551552	1325.8	95.553	<2e-16	***	72.97	8.54
Year	1	417	417.5	30.088	4.28E-08	***	0.06	0.25
Genotype×Year	416	21183	50.9	3.67	<2e-16	***	4.69	2.17
Year×Replication×Block	1	22	21.8	1.573	0.21	ns	2.42	1.56
Residuals	6593	91482	13.9				11.97	3.46
	(C) Pre-anthesis tip degeneration							
	Df	**Sum Sq**	**Mean Sq**	***F*-value**	**Pr (>*F*)**	**Sig.**	**σ^2**	**SD.**
Genotype	416	30293	73	5.09	<2e-16	***	22.69	4.76
Year	1	3627	3627	253.55	<2e-16	***	7.72	2.78
Residual	673	9628	14				13.63	3.69
	(D) Heading date							
	Df	**Sum Sq**	**Mean Sq**	***F*-value**	**Pr (>*F*)**	**Sig.**	**σ^2**	**SD.**
Genotype	416	51960	125	14.544	<2e-16	***	17.65	4.2
Year	1	28171	28171	3280.42	<2e-16	***	22.93	4.79
Genotype×Year	411	11335	28	3.211	<2e-16	***	6.43	2.54
Year×Replication×Block	1	1	1	1.104	0.747	ns	0.15	0.39
Residuals	1629	13989	9				14.65	3.83
	(E) Plant height							
	Df	**Sum Sq**	**Mean Sq**	**F-value**	**Pr (>F)**	**Sig.**	**σ^2**	**SD.**
Genotype	416	1018391	2448	130.737	<2e-16	***	132.94	11.53
Year	1	33421	33421	1784.8	<2e-16	***	8.92	2.99
Genotype×Year	416	44259	106	5.682	<2e-16	***	9.62	3.1
Year×Replication×Block	1	613	613	32.742	1.10E-08	***	5.07	2.52
Residuals	6593	123454	19				14.65	3.83

Df = Degree of freedom; Sq = squares; Sig = significance codes; σ^2 = variance; SD = standard deviation;
***, **, and * = significant at the 0.001, 0.01, and 0.05 probability level, respectively.

Table S2.5 Weather data during the 2018 growing season at IPK, Gatersleben. (Submitted in a separate Excel file)

Table S2.6 Weather data during the 2019 growing season at IPK, Gatersleben. (Submitted in a separate Excel file)

Table S2.7 Weather data during the 2020 growing season at IPK, Gatersleben. (Submitted in a separate Excel file)

Table S2.8 Path analysis results for potential spikelet number (PSN) as an independent variable for pre-anthesis tip degeneration (PTD)

Model used <-
PSN ~ HD + PH
FSN ~ PSN + HD
PTD ~ PSN '

Number of observations	**417**			
Model test user model				
Test statistics	**8.94**			
Degree of freedom	**3**			
P- value (chi-square)	**0.3**			
Model test baseline model				
Test statistics	**1823.99**			
Degree of freedom	**9**			
P- value (chi-square)	**0**			
User model versus baseline model				
Comparative Fit Index (CFI)	**0.99**			
Tucker-Lewis Index (TLI)	**0.99**			
Root Mean Square Error of Approximation (RMSEA)	**0.07**			
Standardised Root Mean Square Residual (SRMR)	**0.03**			
Regression				
	Estimate	**Std. error**	***z*-value**	**P (> |z|)**
PSN ~				
HD	**1.26**	**0.09**	**14.07**	**0.00**
PH	**0.26**	**0.04**	**7.28**	**0.00**
FSN ~				
PSN	**0.59**	**0.02**	**25.42**	**0.00**
HD	**0.09**	**0.02**	**4.64**	**0.00**
PTD ~				
PSN	**-0.01**	**0.02**	**-0.31**	**0.76**
R-square				
PSN	**0.49**			
FSN	**0.64**			
PTD	**0**			

Table S2.9 Path analysis results for final spikelet number (PSN) as an independent variable for pre-anthesis tip degeneration (PTD)

Model used <-
PSN ~ HD + PH
FSN ~ PSN + HD
PTD ~ FSN '

Number of observations	**417**			
Model test user model				
Test statistics	**928.1**			
Degree of freedom	**4**			
P- value (chi-square)	**0.3**			
Model test baseline model				
Test statistics	**1823.99**			
Degree of freedom	**9**			
P- value (chi-square)	**0.00**			
User model versus baseline model				
Comparative Fit Index (CFI)	**0.49**			
Tucker-Lewis Index (TLI)	**-0.15**			
Root Mean Square Error of Approximation (RMSEA)	**0.74**			
Standardised Root Mean Square Residual (SRMR)	**0.14**			
Regression				
	Estimate	**Std. error**	***z*-value**	**P (> |z|)**
PSN ~				
HD	**1.26**	**0.09**	**14.07**	**0.00**
PH	**0.26**	**0.04**	**7.28**	**0.00**
FSN ~				
PSN	**0.53**	**0.03**	**18.15**	**0.00**
HD	**0.28**	**0.07**	**4.08**	**0.00**
PTD ~				
FSN	**-0.37**	**0.02**	**-15.13**	**0.00**
R-square				
PSN	**0.49**			
FSN	**0.65**			
PTD	**0.35**			

Table S3.1 ANOVA for the spike, grain and shoot traits in 2018

Spike traits

(a) Potential spikelet number (PSN)

	Df	Sum Sq	Mean Sq	F-value	Pr (>F)	Sig.	σ2	SD.
Genotype	293	275976	942	11.14	<2e-16	***	126.4	11.243
Replication	1	11454	11454	135.5	<2e-16	***	23.1	4.806
Residuals	1745	147501	85				79.59	8.921

(b) Final spikelet number (FSN)

	Df	Sum Sq	Mean Sq	F-value	Pr (>F)	Sig.	σ2	SD.
Genotype	290	163100	562	17.48	<2e-16	***	69.608	8.3434
Replication	1	4608	4608	143.22	<2e-16	***	9.303	30.5
Residuals	1876	60353	32				29.963	5.474

(c) Spike length (SL)

	Df	Sum Sq	Mean Sq	F-value	Pr (>F)	Sig.	σ2	SD.
Genotype	287	3940	13.73	12.37	<2e-16	***	1.63	12.8
Replication	1	74	73.57	66.29	7.05E-16	***	0.23	0.48
Residuals	1863	2067	1.11	4.96			1.02	1.01

(d) Spike weight (SW)

	Df	Sum Sq	Mean Sq	F-value	Pr (>F)	Sig.	σ2	SD.
Genotype	287	999.6	3.48	7.86	<2e-16	***	0.42	0.64
Replication	1	82.2	82.17	185.37	<2e-16	***	0.13	0.36
Residuals	1867	827.6	0.44				0.41	0.64

(e) Awn length (AL)

	Df	Sum Sq	Mean Sq	F-value	Pr (>F)	Sig.	σ2	SD.
Genotype	283	759	2.68	4.22	<2e-16	***	1.05	1.07
Replication	1	54.9	54.94	86.50	<2e-16	***	0.21	0.46
Residuals	263	167.1	0.64				0.64	0.79

Grain traits

(a) Grain number per spike (GNS)

	Df	Sum Sq	Mean Sq	F-value	Pr (>F)	Sig.	σ2	SD.
Genotype	284	184143	648	6.19	<2e-16	***	79.32	8.91
Replication	1	31120	31120	296.97	<2e-16	***	43.22	6.58
Residuals	1604	168089	105				99.38	9.97

(b) Grain length (GL)

	Df	Sum Sq	Mean Sq	F-value	Pr (>F)	Sig.	σ2	SD.
Genotype	284	7302	25.712	40.55	<2e-16	***	3.65	1.91
Replication	1	29	29.447	46.44	1.33E-11	***	0.14	0.37
Residuals	1604	1017	0.634				0.57	0.75

(c) Grain width (GWi)

	Df	Sum Sq	Mean Sq	F-value	Pr (>F)	Sig.	σ2	SD.
Genotype	284	106.21	0.374	11.26	<2e-16	***	0.05	0.23
Replication	1	0.05	0.0547	1.65	0.20	ns	0.00	0.02
Residuals	1604	53.38	0.0332				0.03	0.18

(d) Grain area (GA)

	Df	Sum Sq	Mean Sq	F-value	Pr (>F)	Sig.	σ2	SD.
Genotype	284	45946	161.78	35.73	<2e-16	***	23.05	4.80
Replication	1	205	204.57	45.19	2.48E-11	***	0.58	0.76
Residuals	1604	7262	4.53				4.31	2.08

(e) Grain weight (GWe)

	Df	Sum Sq	Mean Sq	F-value	Pr (>F)	Sig.	σ2	SD.
Genotype	284	831.4	2.93	9.54	<2e-16	***	0.40	0.63
Replication	1	54.1	54.14	176.44	<2e-16	***	0.08	0.27
Residuals	1601	491.2	0.31				0.30	0.54

(f) Thousand-grain weight (TGW)

	Df	Sum Sq	Mean Sq	F-value	Pr (>F)	Sig.	σ2	SD.
Genotype	284	206402	726.8	11.91	<2e-16	***	104.1748	10.2066
Replication	1	203	203	3.33	6.83E-02	.	0.4602	0.6784
Residuals	1601	97682	61				60.8168	7.7985

Shoot trait

(a) Heading date (HD)

	Df	Sum Sq	Mean Sq	F-value	Pr (>F)	Sig.	σ2	SD.
Genotype	281	16900	60.14	3.259	5.44E-15	***	25.378	5.038
Replication	1	202	202.24	10.96	0.00116	**	2.439	1.562
Residuals	156	2879	18.45				19.478	4.413

Df = Degree of freedom; Sq = squares; Sig = significance codes; σ^2 = variance; SD = standard deviation; ***, **,*, and . = significant at the 0.001, 0.01, 0.5 and 1 probability level , respectively; ns = nonsignificant association.

Table S3.2 ANOVA for the spike, grain and shoot traits in 2019

Spike traits

(a) Potential spikelet number (PSN)

	Df	Sum Sq	Mean Sq	F-value	Pr (>F)	Sig.	σ2	SD.
Genotype	422	494065	1170.8	37.24	< 2e-16	***	131.75	11.48
Replication	1	111	111.1	3.53	0.0602	.	2.92	1.71
Replication:block	1	1391	1391.4	44.25	3.37E-11	***	0.19	0.43
Residuals	3285	103285	31.4				29.60	5.44

(b) Final spikelet number (FSN)

	Df	Sum Sq	Mean Sq	F-value	Pr (>F)	Sig.	σ2	SD.
Genotype	416	263064	632.4	33.81	< 2e-16	***	72.18	8.50
Replication	1	1126	1125.7	60.19	1.14E-14	***	3.33	1.83
Replication:block	1	106	106.2	5.68	0.0172	*	0.40	0.64
Residuals	3280	61346	18.7				16.33	4.04

(c) Spike length (SL)

	Df	Sum Sq	Mean Sq	F-value	Pr (>F)	Sig.	σ2	SD.
Genotype	416	5216	12.54	30.51	<2e-16	***	1.39	1.18
Replication	1	17	17.21	41.87	1.12E-10	***	0.03	0.18
Replication:block	1	0	0.02	0.038	0.845	ns	0.01	0.09
Residuals	3274	1346	0.41				0.38	0.62

(d) Spike weight (SW)

	Df	Sum Sq	Mean Sq	F-value	Pr (>F)	Sig.	σ2	SD.
Genotype	416	1157.8	2.78	18.3	<2e-16	***	0.30	0.55
Replication	1	14.3	14.28	93.89	<2e-16	***	0.01	0.10
Replication:block	1	3	30.01	19.77	9.04E-06	***	0.01	0.08
Residuals	3276	498.3	0.15				0.15	0.38

(e) Awn length (AL)

	Df	Sum Sq	Mean Sq	F-value	Pr (>F)	Sig.	σ2	SD.
Genotype	416	3772	9.07	21.101	<2e-16	***	0.98	0.99
Replication	1	2	1.84	4.287	0.0385	*	0.03	0.19
Replication:block	1	11	10.79	25.111	5.70E-07	***	0.00	0.04
Residuals	3272	1406	0.43				0.41	0.64

Grain traits

(a) Grain number per spike (GNS)

	Df	Sum Sq	Mean Sq	F-value	Pr (>F)	Sig.	σ2	SD.
Genotype	416	322682	776	14.01	<2e-16	***	84.78	9.21
Replication	1	8489	84889	153.35	<2e-16	***	4.10	2.02
Replication:block	1	2239	2239	40.44	2.31E-10	***	3.50	1.88
Residuals	3253	180081	55				52.93	7.28

(b) Grain length (GL)

	Df	Sum Sq	Mean Sq	F-value	Pr (>F)	Sig.	σ2	SD.
Genotype	416	8920	21.49	57.73	<2e-16	***	2.43	1.56

Replication	1	207	207.49	556.56	<2e-16	***	0.07	0.27
Replication:block	1	8	8.07	21.68	3.35E-06	***	0.08	0.29
Residuals	3253	1211	0.37				0.32	0.57

(c) Grain width (GWi)

	Df	Sum Sq	Mean Sq	F-value	Pr (>F)	Sig.	σ2	SD.
Genotype	416	164.61	0.40	24.408	<2e-16	***	0.04	0.21
Replication	1	0.58	0.58	35.836	2.38E-09	***	0.00	0.03
Replication:block	1	0.02	0.02	1.478	2.24E-01	ns	0.00	0.01
Residuals	3253	52.67	0.02				0.02	0.12

(d) Grain area (GA)

	Df	Sum Sq	Mean Sq	F-value	Pr (>F)	Sig.	σ2	SD.
Genotype	416	65764	158.09	61.9	<2e-16	***	17.83	4.22
Replication	1	234	233.9	91.58	<2e-16	***	0.35	0.59
Replication:block	1	44	44.02	17.24	3.38E-05	***	0.09	0.30
Residuals	3253	8308	2.55				2.31	1.52

(e) Grain weight (GWe)

	Df	Sum Sq	Mean Sq	F-value	Pr (>F)	Sig.	σ2	SD.
Genotype	416	1032.5	2.48	17.84	<2e-16	***	0.27	0.52
Replication	1	12.7	12.68	91.15	<2e-16	***	0.01	0.09
Replication:block	1	3.6	3.58	25.74	4.31E-07	***	0.01	0.07
Residuals	3252	452.5	0.14				0.13	0.37

(f) Thousand-grain weight (TGW)

	Df	Sum Sq	Mean Sq	F-value	Pr (>F)	Sig.	σ2	SD.
Genotype	416	250288	601.7	33.209	<2e-16	***	68.14	8.25
Replication	1	101	101	5.577	0.0183	*	1.12	1.06
Replication:block	1	39	38.9	2.145	1.43E-01	ns	0.05	0.22
Residuals	3252	58917	18.1				17.31	4.16

Shoot traits

(a) Heading date (HD)

	Df	Sum Sq	Mean Sq	F-value	Pr (>F)	Sig.	σ2	SD.
Genotype	412	37607	91.28	23.228	<2e-16	***	31.18	5.58
Replication	1	6	5.59	1.423	0.233	ns	0.02	0.14
Replication:block	1	0	0	0.001	0.981	ns	0.13	0.36
Residuals	802	3152	3.93				3.79	1.95

(b) Plant height (PH)

	Df	Sum Sq	Mean Sq	F-value	Pr (>F)	Sig.	σ2	SD.
Genotype	417	562697	1349	62.76	< 2e-16	***	149.652	12.233
Replication	1	4261	4261	198.15	< 2e-16	***	1.875	1.369
Replication:block	1	978	978	45.49	1.80E-11	***	4.957	2.226
Residuals	3288	70699	22				18.031	4.246

(c) Culm dry weight (CDW)

	Df	Sum Sq	Mean Sq	F-value	Pr (>F)	Sig.	σ2	SD.
Genotype	416	1042.8	2.5068	20.27	<2e-16	***	0.28	0.53

Replication	1	2.7	2.6868	21.73	3.27E-06	***	0.01	0.10
Replication:block	1	0	0.0058	0.05	8.29E-01	ns	0.00	0.05
Residuals	3275	405	0.1237				0.12	0.34

Df = Degree of freedom; Sq = squares; Sig = significance codes; σ^2 = variance; SD = standard deviation; ***, **,*, and . = significant at the 0.001, 0.01, 0.5 and 1 probability level , respectively; ns = nonsignificant association.

Table S3.3 ANOVA for the spike, grain and shoot traits in 2020

Spike traits

(a) Potential spikelet number (PSN)

	Df	Sum Sq	Mean Sq	F-value	Pr (>F)	Sig.	σ2	SD.
Genotype	416	48803	1173.1	63.87	<2e-16	***	131.35	11.46
Replication	1	12	12	0.65	0.42	ns	–	–
Replication:block	1	3	3.2	0.17	0.68	ns	1.25	1.12
Residuals	3259	59854	18.4				17.4	4.17

(b) Final spikelet number (FSN)

	Df	Sum Sq	Mean Sq	F-value	Pr (>F)	Sig.	σ2	SD.
Genotype	416	309617	744.3	88.86	<2e-16	***	83.42	9.13
Replication	1	1115	1115.4	133.16	<2e-16	***	0.74	0.86
Replication:block	1	86	185.6	510.22	0.001	**	0.8	0.89
Residuals	3310	27724	8.4				7.57	2.75

(c) Spike length (SL)

	Df	Sum Sq	Mean Sq	F-value	Pr (>F)	Sig.	σ2	SD.
Genotype	416	5396	12.97	49.52	<2e-16	***	1.42	1.19
Replication	1	25	24.59	93.89	<2e-16	***	0.01	0.11
Replication:block	1	25	21.15	96.02	<2e-16	***	0.03	0.18
Residuals	3293	863	0.26				0.24	0.49

(d) Spike weight (SW)

	Df	Sum Sq	Mean Sq	F-value	Pr (>F)	Sig.	σ2	SD.
Genotype	416	1148.3	2.76	15.14	<2e-16	***	0.29	0.54
Replication	1	1.9	1.87	10.275	0.00136	**	4.99E-03	0.07
Replication:block	1	1.6	1.56	8.566	3.45E-03	**	3.98E-03	0.06
Residuals	3299	601.5	0.18				0.18	0.42

(e) Awn length (AL)

	Df	Sum Sq	Mean Sq	F-value	Pr (>F)	Sig.	σ2	SD.
Genotype	416	4659	11.20	22.791	<2e-16	***	1.20	1.10
Replication	1	2	1.61	3.269	0.0707	.	0.03	0.16
Replication:block	1	9	8.93	18.173	2.07E-05	***	0.00	0.05
Residuals	3281	1612	0.49				0.47	0.69

Grain traits

(a) Grain number per spike (GNS)

	Df	Sum Sq	Mean Sq	F-value	Pr (>F)	Sig.	σ2	SD.
Genotype	416	405759	975	14.91	<2e-16	***	106.85	10.34
Replication	1	446	446	6.82	0.00906	**	3.81	1.95
Replication:block	1	4759	4759	72.77	<2e-16	***	0.57	0.75
Residuals	3208	209808	65				63.71	7.98

(b) Grain length (GL)

	Df	Sum Sq	Mean Sq	F-value	Pr (>F)	Sig.	σ2	SD.
Genotype	416	9120	21.92	38.561	<2e-16	***	2.46221	1.5691

Replication	1	89	88.63	155.898	<2e-16	***	0.07393	0.2719
Replication:block	1	0	0.03	0.047	0.828	ns	0.0342	0.1849
Residuals	3208	1824	0.57				0.51413	0.717
(c) Grain width (GWi)								
	Df	Sum Sq	Mean Sq	F-value	Pr (>F)	Sig.	σ2	SD.
Genotype	416	141.07	0.3391	17.058	<2e-16	***	0.04	0.19
Replication	1	0.16	0.16	8.05	0.00458	**	0.00	0.03
Replication:block	1	0.11	0.114	5.737	1.67E-02	*	0.00	0.02
Residuals	3208	63.78	0.0199				0.02	0.14
(d) Grain area (GA)								
	Df	Sum Sq	Mean Sq	F-value	Pr (>F)	Sig.	σ2	SD.
Genotype	416	58938	141.7	41.476	<2e-16	***	15.95	3.99
Replication	1	516	516.4	151.179	<2e-16	***	0.37	0.61
Replication:block	1	11	11.1	3.245	7.17E-02	.	0.23	0.48
Residuals	3208	10958	3.4				3.13	1.77
(e) Grain weight (GWe)								
	Df	Sum Sq	Mean Sq	F-value	Pr (>F)	Sig.	σ2	SD.
Genotype	416	1212.4	2.92	16.72	<2e-16	***	0.32	0.57
Replication	1	3.2	3.19	18.31	0.0000193	***	0.01	0.08
Replication:block	1	7.5	7.49	43	6.35E-11	***	0.00	0.06
Residuals	3206	558.7	0.17				0.17	0.41
(f) Thousand-grain weight (TGW)								
	Df	Sum Sq	Mean Sq	F-value	Pr (>F)	Sig.	σ2	SD.
Genotype	416	252222	606.3	25.66	<2e-16	***	68.48	8.28
Replication	1	541	540.5	22.88	1.86-06	***	1.20	1.10
Replication:block	1	0	0	0	9.94E-01		0.59	0.77
Residuals	3206	75746	23.6				22.46	4.74

Shoot traits

(a) Heading date (HD)								
	Df	Sum Sq	Mean Sq	F-value	Pr (>F)	Sig.	σ2	SD.
Genotype	416	26718	64.23	4.89	<2e-16	***	17.28	4.16
Replication	1	8	7.74	0.589	0.443	ns	0.04	0.2
Replication:block	1	3	3.1	0.236	0.627	ns	1.67	0.41
Residuals	824	10822	13.13				12.93	3.59
(b) Plant height (PH)								
	Df	Sum Sq	Mean Sq	F-value	Pr (>F)	Sig.	σ2	SD.
Genotype	416	502014	1207	93	<2e-16	***	134.74	11.61
Replication	1	5077	5077	391.25	<2e-16	***	2.14	1.46
Replication:block	1	343	343	26.44	2.88E-07	***	2.09	1.45
Residuals	3310	42953	13				11.34	3.37
(c) Culm dry weight (CDW)								
	Df	Sum Sq	Mean Sq	F-value	Pr (>F)	Sig.	σ2	SD.
Genotype	416	1117.8	2.69	29.11	<2e-16	***	0.29	0.54

Replication	1	77.9	77.88	843.69	<2e-16	***	0.01	0.08
Replication:block	1	0	0	0.02	8.87E-01	ns	0.03	0.18
Residuals	3293	304	0.09				0.09	0.29

Df = Degree of freedom; Sq = squares; Sig = significance codes; σ^2 = variance; SD = standard deviation; ***, **,*, and . = significant at the 0.001, 0.01, 0.5 and 1 probability level , respectively; ns = nonsignificant association.

Table S3.4 Across year ANOVA for the spike, grain and shoot traits

Spike traits

(a) Potential spikelet number (PSN)

	Df	Sum Sq	Mean Sq	F-value	Pr (>F)	Sig.	σ2	SD.
Genotype	416	903668	2172	86.813	<2e-16	***	115.45	10.75
Replication	1	11411	11411	456.023	<2e-16	***	3.05	1.75
Genotype x year	416	63298	152	6.081	<2e-16	***	14.68	3.83
Year x replication x block	1	614	614	24.531	7.57E-07	***	2.17	1.47
Residuals	6507	162822	25				23.49	4.85

(b) Final spikelet number (FSN)

	Df	Sum Sq	Mean Sq	F-value	Pr (>F)	Sig.	σ2	SD.
Genotype	416	551552	1325.8	95.553	<2e-16	***	72.97	8.54
Replication	1	417	417.5	30.088	4.28E-08	***	0.06	0.25
Genotype x year	416	21183	50.9	3.67	<2e-16	***	4.69	2.17
Year x replication x block	1	22	21.8	1.573	0.21	ns	2.42	1.56
Residuals	6593	91482	13.9				11.97	3.46

(c) Pre-anthesis tip degeneration (PTD)

	Df	Sum Sq	Mean Sq	F-value	Pr (>F)	Sig.	σ2	SD.
Genotype	416	30293	73	5.09	<2e-16	***	22.69	4.76
Replication	1	3627	3627	235.55	<2e-16	***	7.72	2.78
Residuals	673	9628	14				13.63	3.69

(d) Spike length (SL)

	Df	Sum Sq	Mean Sq	F-value	Pr (>F)	Sig.	σ2	SD.
Genotype	416	9907	23.8	68.94	<2e-16	***	1.25	1.11
Replication	1	351	350.6	1014.88	<2e-16	***	0.09	0.31
Genotype x year	416	713	1.7	4.96	<2e-16	***	0.15	0.39
Year x replication x block	1	6	5.6	16.11	6.06E-05	***	0.04	0.21
Residuals	6570	2270	0.3				0.31	0.56

(e) Spike weight (SW)

	Df	Sum Sq	Mean Sq	F-value	Pr (>F)	Sig.	σ2	SD.
Genotype	416	1939.2	4.66	27.58	<2e-16	***	0.21	0.46
Replication	1	127.3	127.25	752.94	<2e-16	***	0.03	0.18
Genotype x year	416	361.9	0.87	5.148	<2e-16	***	0.08	0.28
Year x replication x block	1	8.7	8.75	51.76	6.96E-13	***	0.01	0.11
Residuals	6578	1111.7	0.17					

(f) Spike density (SD)

	Df	Sum Sq	Mean Sq	F-value	Pr (>F)	Sig.	σ2	SD.
Genotype	416	244.75	0.59	19.67	<2e-16	***	0.22	0.47
Replication	1	16.7	16.69	558.29	<2e-16	***	0.03	0.16
Residuals	674	20.16	0.03				0.03	0.17

(g) Awn length (AL)

	Df	Sum Sq	Mean Sq	F-value	Pr (>F)	Sig.	σ2	SD.
Genotype	416	7811	18.78	40.47	<2e-16	***	0.98	0.99
Replication	1	5	5.33	11.46	7.06E-04	***	6.90E-04	0.03
Genotype x year	416	621	1.49	3.22	<2e-16	***	0.11	0.34
Year x replication x block	1	0	0.1	0.22	0.64	ns	0.03	0.17
Residuals	6556	3041	0.46				0.44	0.66

Grain traits

(a) Grain number per spike (GNS)

	Df	Sum Sq	Mean Sq	F-value	Pr (>F)	Sig.	σ2	SD.
Genotype	416	577318	1388	22.65	<2e-16	***	60.67	7.79
Replication	1	47740	47740	779.35	<2e-16	***	12.61	3.55
Genotype x year	416	149917	360	5.88	<2e-16	***	35.32	5.94
Year x replication x block	1	9857	9857	160.91	<2e-16	***	5.65	2.38
Residuals	6464	395965	61				58.29	7.63

(b) Grain length (GL)

	Df	Sum Sq	Mean Sq	F-value	Pr (>F)	Sig.	σ2	SD.
Genotype	416	17329	41.7	84.59	<2e-16	***	2.27	1.51
Replication	1	1294	1293.5	2626.71	<2e-16	***	0.35	0.59
Genotype x year	416	767	1.8	3.74	<2e-16	***	0.17	0.41
Year x replication x block	1	156	155.7	316.11	<2e-16	***	0.11	0.33
Residuals	6464	3183	0.5				0.42	0.64

(c) Grain width (GWi)

	Df	Sum Sq	Mean Sq	F-value	Pr (>F)	Sig.	σ2	SD.
Genotype	416	273.38	0.66	36.21	<2e-16	***	0.03	0.18
Replication	1	2.69	2.69	148.13	<2e-16	***	6.90E-04	0.03
Genotype x year	416	32	0.08	4.24	<2e-16	***	6.90E-03	0.08
Year x replication x block	1	0	0.007	0.09	0.76	ns	1.30E-03	0.04
Residuals	6464	117.32	0.02				0.02	0.13

(d) Grain area (GA)

	Df	Sum Sq	Mean Sq	F-value	Pr (>F)	Sig.	σ2	SD.
Genotype	416	119189	287	94.59	<2e-16	***	15.57	3.97
Replication	1	4139	4139	1366.39	<2e-16	***	1.13	1.06
Genotype x year	416	5660	14	4.49	<2e-16	***	1.32	1.15
Year x replication x block	1	493	493	162.71	<2e-16	***	0.45	0.67
Residuals	6464	19579	3				2.73	1.65

(e) Grain weight (GWe)

	Df	Sum Sq	Mean Sq	F-value	Pr (>F)	Sig.	σ2	SD.
Genotype	416	1853.3	4.46	28.06	<2e-16	***	0.21	0.45
Replication	1	118.1	118.14	744.13	<2e-16	***	0.03	0.18
Genotype x year	416	387.1	0.93	5.86	<2e-16	***	0.09	0.30

Year x replication x block	1	12.4	12.41	78.19	<2e-16	***	0.01	0.11
Residuals	6461	1025.8	0.16				0.15	0.39
(f) Grain set (GS)								
	Df	Sum Sq	Mean Sq	F-value	Pr (>F)	Sig.	σ2	SD.
Genotype	416	166718	401	3.13	<2e-16	***	111.54	10.56
Replication	1	3481	3481	27.19	2.46E-07	***	20.35	4.51
Residuals	672	86041	128				108.74	10.43
(f) Thousand-grain weight (TGW)								
	Df	Sum Sq	Mean Sq	F-value	Pr (>F)	Sig.	σ2	SD.
Genotype	416	456235	1096.7	52.48	<2e-16	***	57.31	3.31
Replication	1	290	290.2	13.88	1.96E-04	***	0.05	7.57
Genotype x year	416	46157	111	5.31	<2e-16	***	10.93	1.14
Year x replication x block	1	323	322.7	15.44	8.60E-05	***	1.31	0.23
Residuals	6461	135021	20.9				19.91	4.46

Shoot traits

(a) Heading date (HD)								
	Df	Sum Sq	Mean Sq	F-value	Pr (>F)	Sig.	σ2	SD.
Genotype	416	51960	125	14.54	<2e-16	***	17.65	4.2
Replication	1	28171	28171	3280.42	<2e-16	***	22.93	4.79
Genotype x year	411	11335	28	3.21	<2e-16	***	6.43	2.56
Year x replication x block	1	1	1	0.1	0.75	ns	0.15	0.39
Residuals	1629	13989	9				8.45	2.91
(b) Plant height (PH)								
	Df	Sum Sq	Mean Sq	F-value	Pr (>F)	Sig.	σ2	SD.
Genotype	416	101839 1	2448	130.74	<2e-16	***	132.94	11.53
Replication	1	33421	33421	1784.8	<2e-16	***	8.92	2.99
Genotype x year	416	44259	106	5.68	<2e-16	***	9.63	3.1
Year x replication x block	1	613	613	32.74	1.10E-08	***	5.07	2.25
Residuals	6593	123454	19				14.69	3.83
(c) Culm dry weight (CDW)								
	Df	Sum Sq	Mean Sq	F-value	Pr (>F)	Sig.	σ2	SD.
Genotype	416	1946.8	4.68	39.32	<2e-16	***	0.24	0.21
Replication	1	94.9	94.9	797.4	<2e-16	***	0.03	0.16
Genotype x year	416	210.4	0.51	4.43	<2e-16	***	0.04	0.21
Year x replication x block	1	7.5	7.51	63.04	2.34E-15	***	0.02	0.14
Residuals	6571	782	0.12				0.1	0.32

Df = Degree of freedom; Sq = squares; Sig = significance codes; σ^2 = variance; SD = standard deviation; ***, **,*, and . = significant at the 0.001, 0.01, 0.5 and 1 probability level , respectively; ns = nonsignificant association.

Table S3.5 Multiple regression results for all the traits underlining the trait displaying the highest regression coefficient for a selected trait

Response trait	Predictor trait	Multiple regression model	Adj. R^2
PSN	FSN, PTD, SW	-51.75 + 1.56FSN + 1.47PTD + 1.02SW	0.96
FSN	PSN, PTD, SL, SD, HD	7.62 + 0.47PSN – 0.74PTD + 1.62 SL + 4.23SD + 0.06HD	0.98
PTD	PSN, FSN	27.99 + 0.59PSN – 1.00FSN	0.93
SL	FSN, SD, PH, CDW	7.45 + 0.119FSN – 2.55SD – 0.005PH + 0.09CDW	0.94
SW	SL, AL, GNS, GL, GWe, GS, TGW, PH, CDW	0.31 + 0.06SL + 0.09AL + 0.02GNS – 0.06GL + 0.71GWe – 0.01GS + 0.02TGW – 0.01PH + 0.27CDW	0.96
SD	FSN, SL, GL, PH	2.55 + 0.05FSN – 0.36SL + 0.02GL – 0.002PH	0.95
AL	PSN, FSN, SL, SW, GNS, GL, GA, GWe, GS, PH, CDW	0.17 + 0.02PSN – 0.03FSN – 0.34SL + 3.71SW – 0.08GNS + 0.65GL - 0.22GA – 2.37GWe + 0.02GS + 0.03PH – 0.45CDW	0.58
GNS	FSN, GWe, GS, TGW, HD, PH, CDW	18.04 + 0.27FSN + 14.24GWe + 0.23GS – 0.46TGW – 0.1HD + 0.06PH – 1.94CDW	0.97
GL	SW, AL, GNS, GWi, GA	7.47 – 0.45SW + 0.05AL + 0.02GNS – 2.52GWi + 0.48GA	0.98
GWi	SL, GNS, GL, GA	2.97 – 0.01SL + 0.002GNS – 0.26GL + 0.13GA	0.91
GA	SW, AL, GNS, GL, GWi, TGW	-14.09 + 0.82SW – 0.11AL – 0.05GNS + 1.96GL + 4.88GWi + 0.04TGW	0.99
GWe	SL, SW, AL, GNS, GS, TGW, HD	-2.02 – 0.01SL + 0.24SW – 0.02AL + 0.03GNS + 0.003GS + 0.03TGW+ 0.01HD	0.98

Response trait	Predictor trait	Multiple regression model	Adj. R^2
GS	FSN, SL, SD, GNS, GWi, GWe	81.19 – 0.62FNS – 1.42SL – 4.06SD + 1.19GNS – 3.53GWi + 4.85GWe	0.93
TGW	SL, SW, GNS, GL, GA, GWe, HD, PH, CDW	51.09 – 0.32SL + 2.47SW – 0.57GNS – 1.51GL + 1.35GA + 13.07GWe – 0.2HD + 0.09PH – 2.03CDW	0.96
HD	FSN, PTD, SW, GNS, GL, GWe, TGW, CDW	138.47 + 0.37FSN + 0.22 PTD – 3.38SW – 0.33GNS + 0.65GL + 8.59GWe – 0.35TGW + 3.49CDW	0.67
PH	PSN, SW, SD, AL, GNS, GA, GWe, GS, TGW, CDW	36.35 + 0.17PSN – 30.92SW – 7.64SD + 2.41AL + 1.27GNS – 0.84GA + 15.58GWe – 0.27GS + 0.88TGW + 18.49CDW	0.75
CDW	SW, SD, AL, GNS, GA, TGW, HD, PH	-2.61 + 1.16SW + 0.15SD – 0.06AL – 0.06GNS + 0.03GA – 0.04TGW + 0.01HD + 0.03PH	0.82

Spike traits: potential spikelet number (PSN), final spikelet number (FSN), pre-anthesis tip degeneration (PTD in %), spike length (SL in cm), spike weight (SW in g), spike density (SD) and awn length (AL)

Grain traits: grain number per spike (GNS), grain length (GL in cm), grain width (GWi in cm), grain area (GA in cm^2), grain weight (GWe in g), grain set (GS in %), thousand-grain weight (TGW in g)

Shoot traits: heading date (HD in days from January 1st), plant height (PH in cm) and culm dry weight (CDW in g)

Table S3.6 Comparison between linear, quadratic and generalised additive model (GAM) to study the relationship between each investigated trait and pre-anthesis tip degeneration

Dependent trait	Independent trait	Linear model (m1)	Quadratic model (m2)	GAM model (m3)	*P*-value for m1 *vs* m2 ANOVA	Description	*P*-value for m1 *vs* m3 ANOVA	Description	*P*-value for m2 *vs* m3 ANOVA	Description
		R^2	R^2	R^2						
Spike traits										
PTD	PSN	0.0002	0.0048	0.035	0.17	Accept H_0, no significant difference between m1 and m2	1.50E-03	Reject H_0, m3 performs better than m1	1.50E-03	Reject H_0, m3 performs better than m2
PTD	FSN	0.35	0.36	0.35	0.28	Accept H_0, no significant difference between m1 and m2	0.12	Accept H_0, no significant difference between m1 and m3	0.48	Accept H_0, no significant difference between m2 and m3
PTD	SL	0.036	0.049	0.056	1.40E-02	Reject H_0, m2 performs better than m1	3.00E-03	Reject H_0, m3 performs better than m1	1.90E-02	Reject H_0, m3 performs better than m2
PTD	SW	0.016	0.016	0.013	0.73	Accept H_0, no significant difference between m1 and m2	4.09E-07	Reject H_0, m3 performs better than m1	0.73	Accept H_0, no significant difference between m2 and m3
PTD	SD	0.103	0.12	0.14	1.80E-02	Reject H_0, m2 performs better than m1	5.60E-04	Reject H_0, m3 performs better than m1	2.60E-03	Reject H_0, m3 performs better than m2
PTD	AL	0.12	0.12	0.13	0.89	Accept H_0, no significant difference between m1 and m2	9.00E-03	Reject H_0, m3 performs better than m1	3.00E-03	Reject H_0, m3 performs better than m2
Grain traits										
PTD	GNS	0.09	0.11	0.13	3.00E-03	Reject H_0, m2 performs better than m1	9.42E-05	Reject H_0, m3 performs better than m1	1.80E-03	Reject H_0, m3 performs better than m2
PTD	GL	0.14	0.16	0.17	1.00E-02	Reject H_0, m2 performs better than m1	3.00E-03	Reject H_0, m3 performs better than m1	2.00E-02	Reject H_0, m3 performs better than m2

Dependent trait	Independent trait	Linear model (m1)	Quadratic model (m2)	GAM model (m3)	*P*-value for m1 *vs* m2 ANOVA	Description	*P*-value for m1 *vs* m3 ANOVA	Description	*P*-value for m2 *vs* m3 ANOVA	Description
PTD	GWi	0.02	0.02	0.02	0.3	Accept H_0, no significant difference between m1 and m2	1.89E-08	Reject H_0, m3 performs better than m1	0.3	Accept H_0, no significant difference between m2 and m3
PTD	GA	0.12	0.12	0.12	0.08	Accept H_0, no significant difference between m1 and m2	0.05	Accept H_0, no significant difference between m1 and m3	0.04	Reject H_0, m3 performs better than m2
PTD	GWe	0.004	0.01	0.007	0.1	Accept H_0, no significant difference between m1 and m2	0.08	Accept H_0, no significant difference between m1 and m3	0.12	Accept H_0, no significant difference between m2 and m3
PTD	GF	0.004	0.011	0.015	0.07	Accept H_0, no significant difference between m1 and m2	0.02	Reject H_0, m3 performs better than m1	0.03	Reject H_0, m3 performs better than m2
PTD	TGW	0.093	0.095	0.09	0.32	Accept H_0, no significant difference between m1 and m2	3.24E-09	Reject H_0, m3 performs better than m1	0.32	Accept H_0, no significant difference between m2 and m3
Shoot traits										
PTD	HD	0.014	0.015	0.012	0.47	Accept H_0, no significant difference between m1 and m2	9.76E-09	Reject H_0, m3 performs better than m1	0.47	Accept H_0, no significant difference between m2 and m3
PTD	PH	0.002	0.006	0.006	0.19	Accept H_0, no significant difference between m1 and m2	0.11	Accept H_0, no significant difference between m1 and m3	0.11	Accept H_0, no significant difference between m2 and m3
PTD	CDW	0.013	0.013	0.011	0.91	Accept H_0, no significant difference between m1 and m2	5.64E-09	Reject H_0, m3 performs better than m1	0.91	Accept H_0, no significant difference between m2 and m3

SUPPLEMENTARY FIGURES

Figure S2.1 Representative immature barley spike at maximum yield potential (MYP) stage. "*" represents the individual nodes on a spike that were counted to determine its yield potential. The number of nodes was later multiplied by three to calculate the potential spikelet number at MYP.

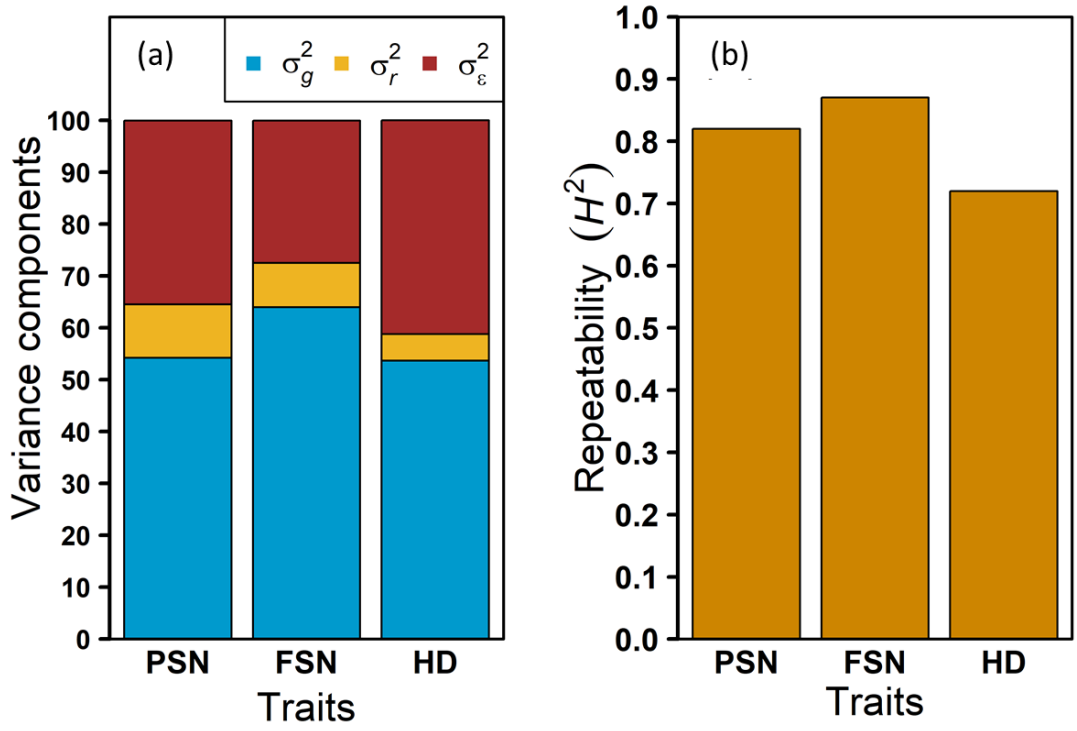


Figure S2.2 Variance component and repeatability analyses in 2018. (a) Proportion of the different variance components for each investigated trait where σ_g^2 is the genotypic variance, σ_r^2 is the replication variance and σ_ε^2 is the error or residual variance. (b) Repeatability (H^2) for each trait. PSN, FSN, and HD represent potential spikelet number, final spikelet number, and heading date (days from January 1st), respectively.

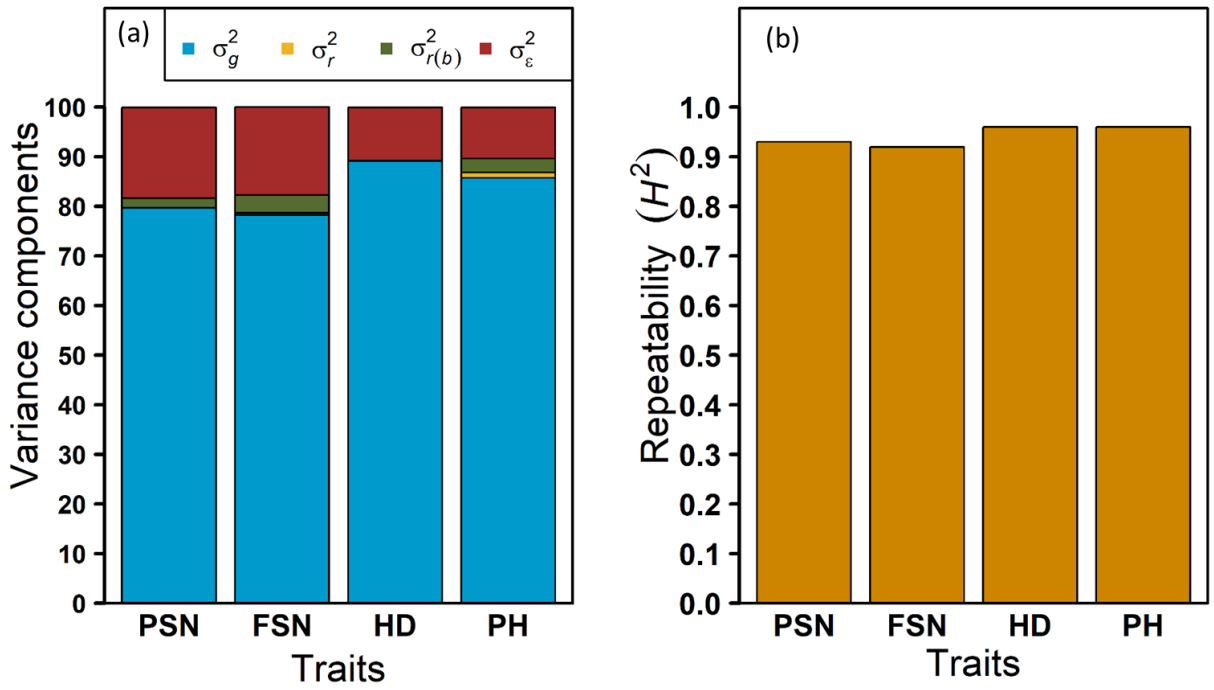


Figure S2.3 Variance component and repeatability analyses in 2019. (a) Proportion of the different variance components for each investigated trait where σ_g^2 is the genotypic variance, σ_r^2 is the replication variance, $\sigma_{r(b)}^2$ is the replication and block interaction variance and σ_ε^2 is the error or residual variance. (b) Repeatability for each trait. PSN, FSN, HD and PH represent potential spikelet number, final spikelet number, heading date (days from January 1st) and plant height (cm), respectively.

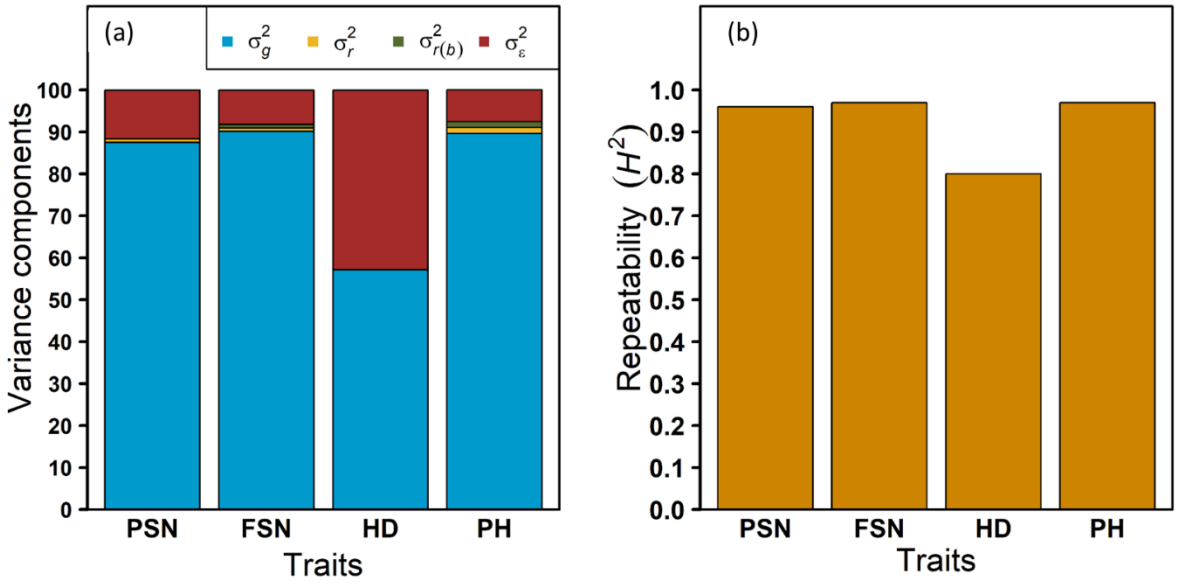


Figure S2.4 Variance component and repeatability analyses in 2020. (a) Proportion of the different variance component for each investigated trait where σ_g^2 is the genotypic variance, σ_r^2 is the replication variance, $\sigma_{r(b)}^2$ is the replication and block interaction variance and σ_ε^2 is the error or residual variance. (b) Repeatability for each trait. PSN, FSN, HD and PH represents potential spikelet number, final spikelet number, heading date (days from January 1st) and plant height (cm), respectively.

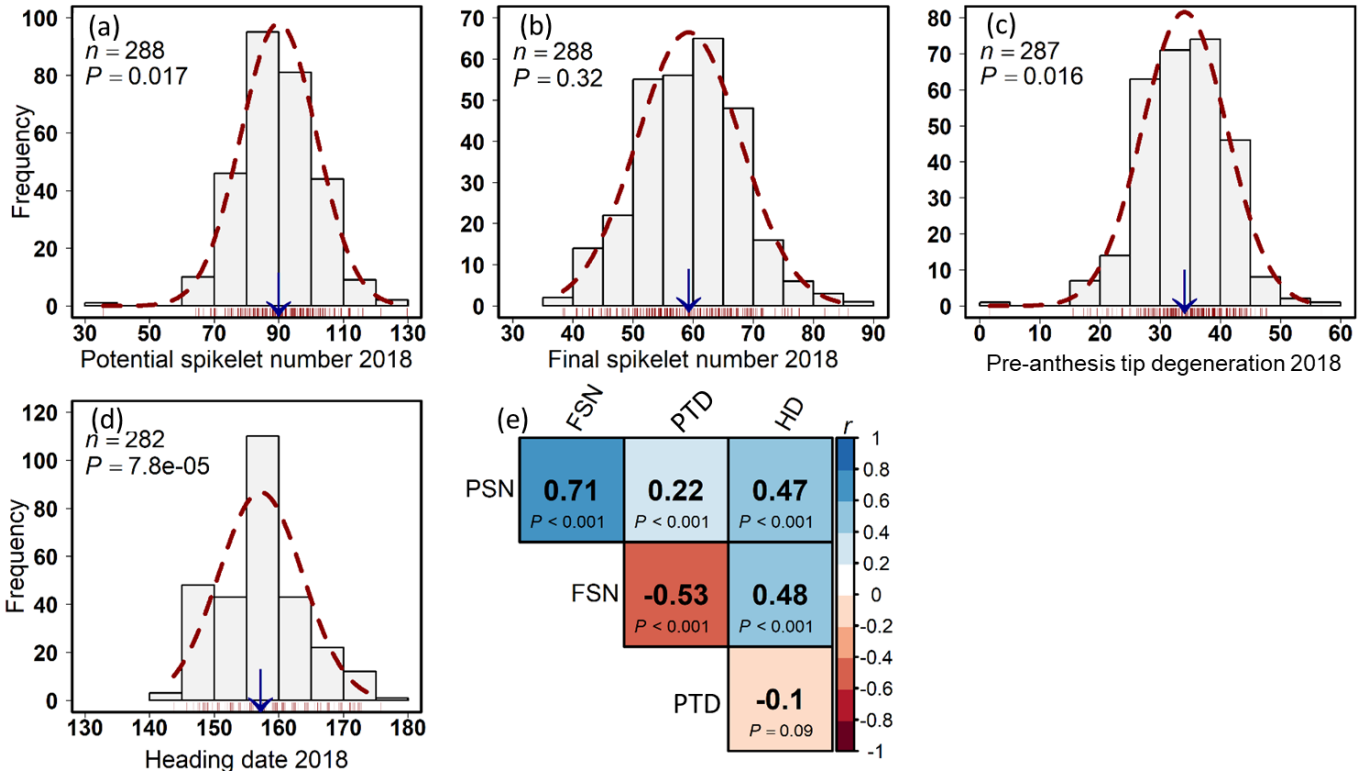


Figure S2.5 Phenotypic distribution and correlation analysis of the investigated traits in a panel of ~288 six-rowed spring barley accessions in 2018. Frequency distribution of (a) potential spikelet number (PSN), (b) final spikelet number (FSN), (c) pre-anthesis tip degeneration (PTD in %), and (d) heading date (HD as days from January 1st). The *x*- and *y*-axis of each histogram denotes the individual trait and number of accessions (frequency), respectively. '*n*' is the number of accessions and '*P*' represents Shapiro-Wilk's test results. The downward blue arrow indicates the mean value for the respective trait and the dashed curve shows the fit of the normal distribution curve. (e) Pearson's product-moment correlation (*r*) among the traits. *P*-value in the plot (e) denotes the significance level of the respective correlation.

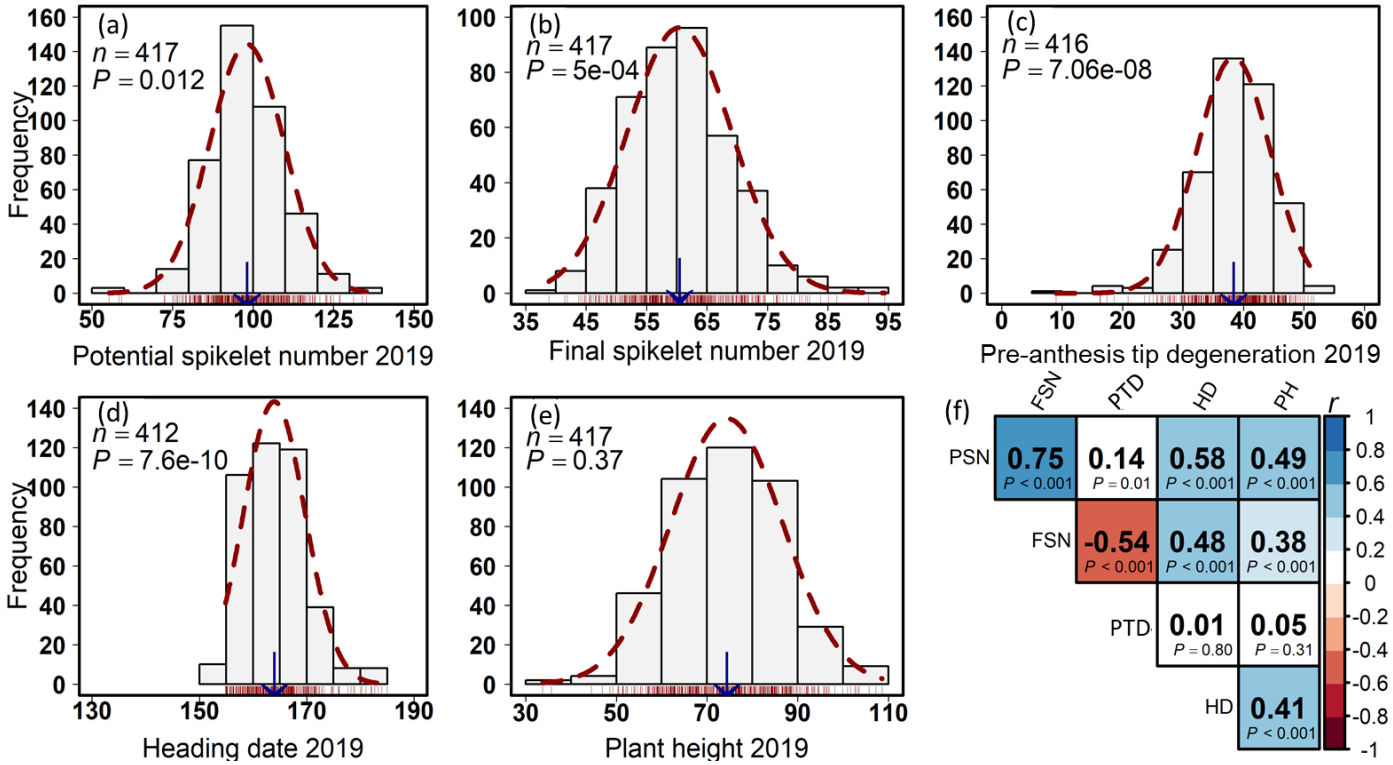


Figure S2.6 Phenotypic distribution and correlation analysis of the investigated traits in a panel of 417 six-rowed spring barley accessions in 2019. Frequency distribution of (a) potential spikelet number (PSN), (b) final spikelet number (FSN), (c) pre-anthesis tip degeneration (PTD in %), (d) heading date (HD as days from January 1st) and (e) plant height (PH in cm). The *x*- and *y*-axis of each histogram denotes the individual trait and number of accessions (frequency), respectively. '*n*' is the number of accessions and '*P*' represents Shapiro-Wilk's test results. The downward blue arrow indicates the mean value for the respective trait and the dashed curve shows the fit of the normal distribution curve. (f) Pearson's product-moment correlation (*r*) among the traits. *P*-value in the plot (f) denotes the significance level of the respective correlation.

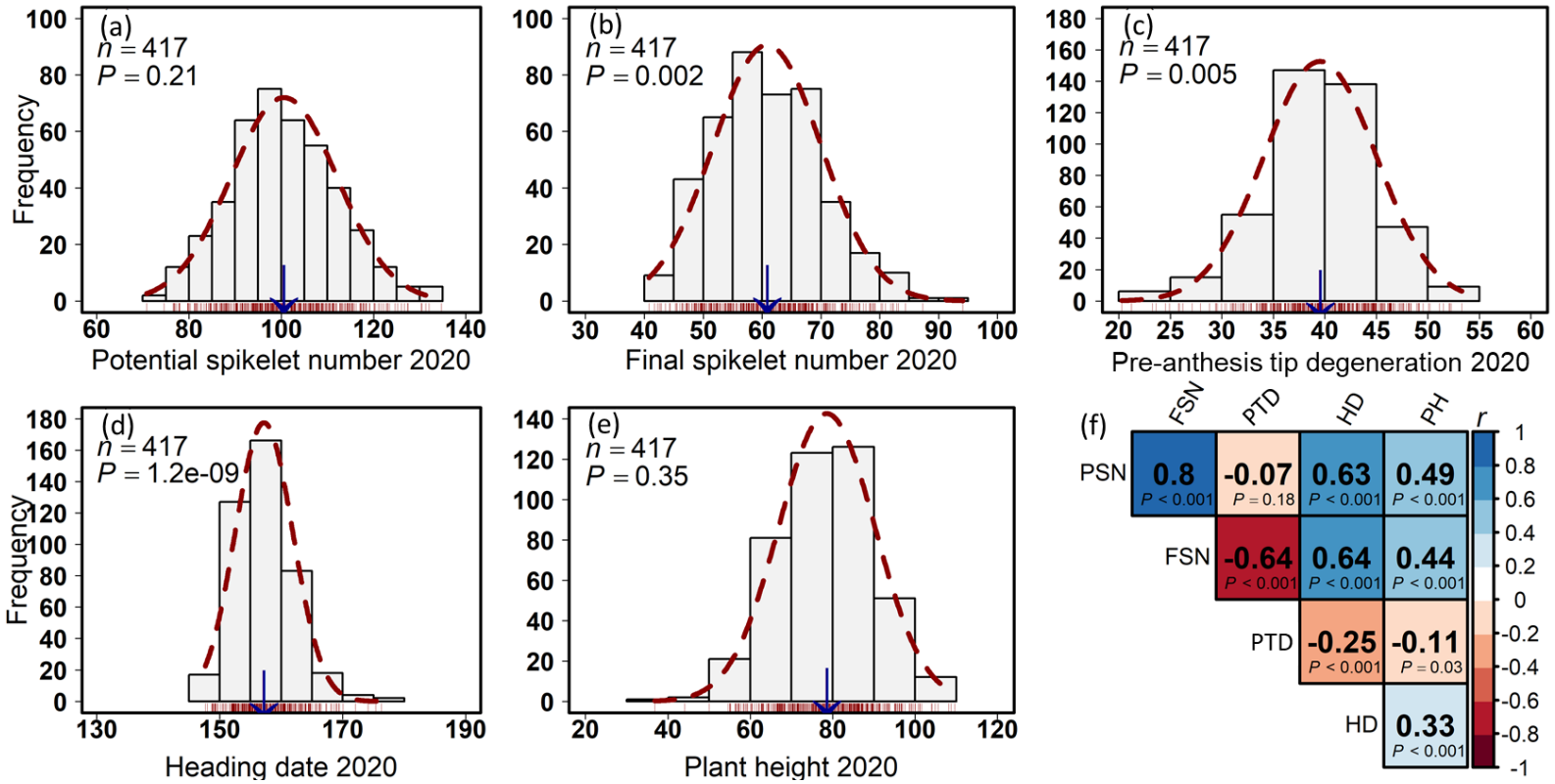


Figure S2. 7 Phenotypic distribution and correlation analysis of the investigated traits in a panel of 417 six-rowed spring barley accessions in 2020. Frequency distribution of (a) potential spikelet number (PSN), (b) final spikelet number (FSN), (c) pre-anthesis tip degeneration (PTD in %), (d) heading date (HD as days from January 1st) and (e) plant height (PH in cm). The *x*- and *y*-axis of each histogram denotes the individual trait and number of accessions (frequency), respectively. '*n*' is the number of accessions and *P* represents the result of Shapiro-Wilk's test (normality test). The downward blue arrow indicates the mean value for the respective trait and the dashed curve shows the fit of the normal distribution curve. (f) Pearson's product-moment correlation (*r*) among the traits. *P*-value in the plot (f) denotes the significance level of the respective correlation.

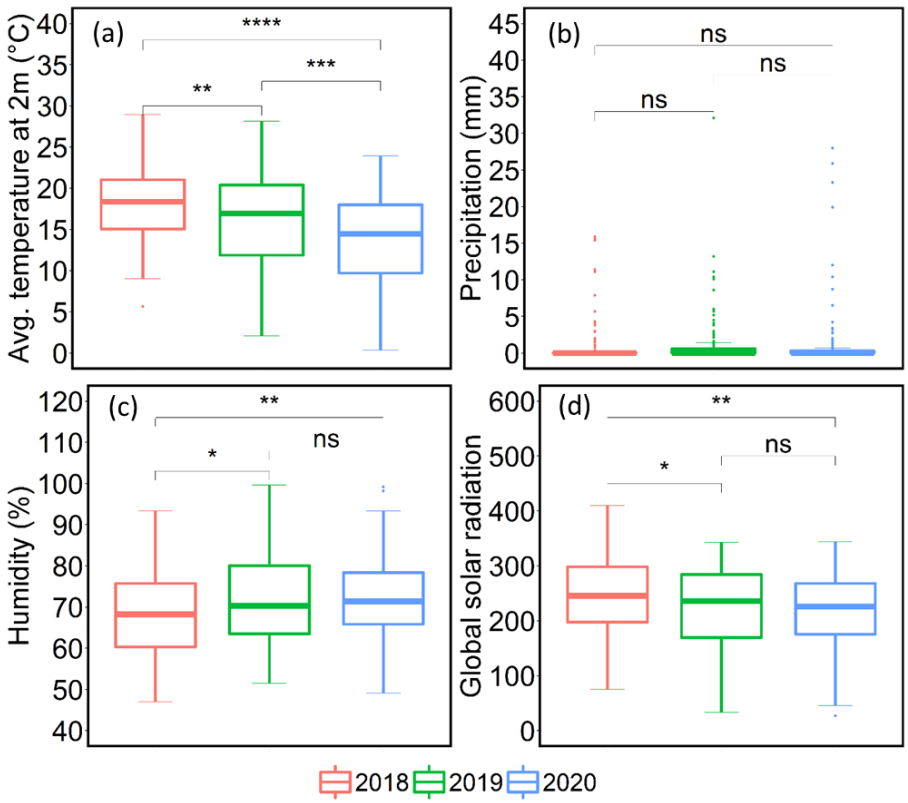


Figure S2.8 Comparison of weather data in three consecutive growing seasons from 2018-2020. (a) difference between the average temperature (°C) at 2m above the soil surface, (b) difference between the precipitation level, (c) difference between the humidity levels and (d) difference between the global solar radiation (Wm^{-2}) in three growing seasons. ****, ***, **, and * = significance level at the 0.0001, 0.001, 0.01, and 0.05 probability level, respectively and 'ns' represents insignificant differences between the growing seasons.

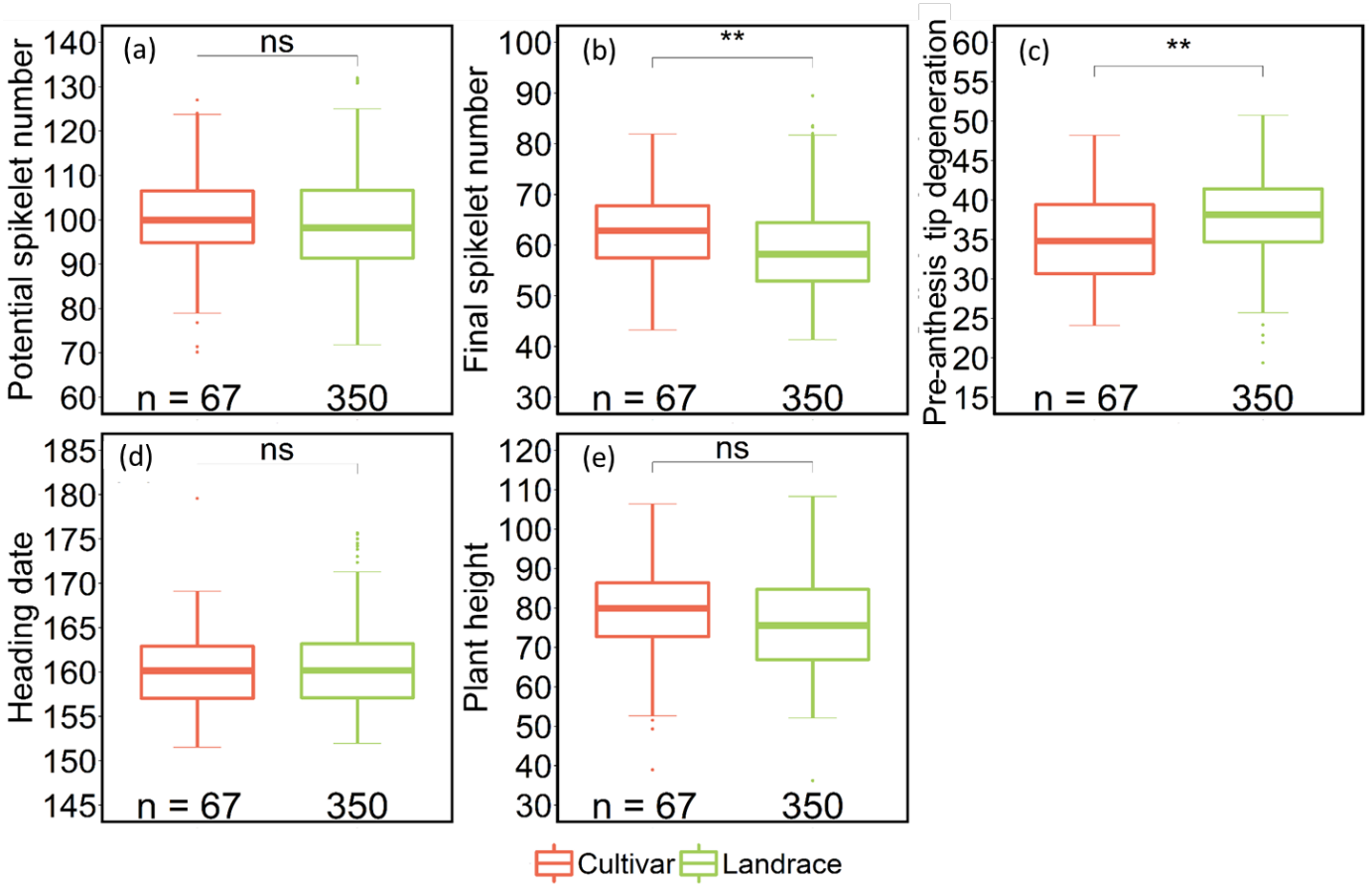


Figure S2. 9 Comparison of accessions with respect to accession nature, viz., cultivar and landraces (a) potential spikelet number (PSN), (b) final spikelet number (FSN), (c) pre-anthesis tip degeneration (PTD in %), (d) heading date (HD as days from January 1st) and (e) plant height (PH in cm). '*n*' denotes the number of accessions; **= 0.01 probability level and 'ns' represents insignificant differences.

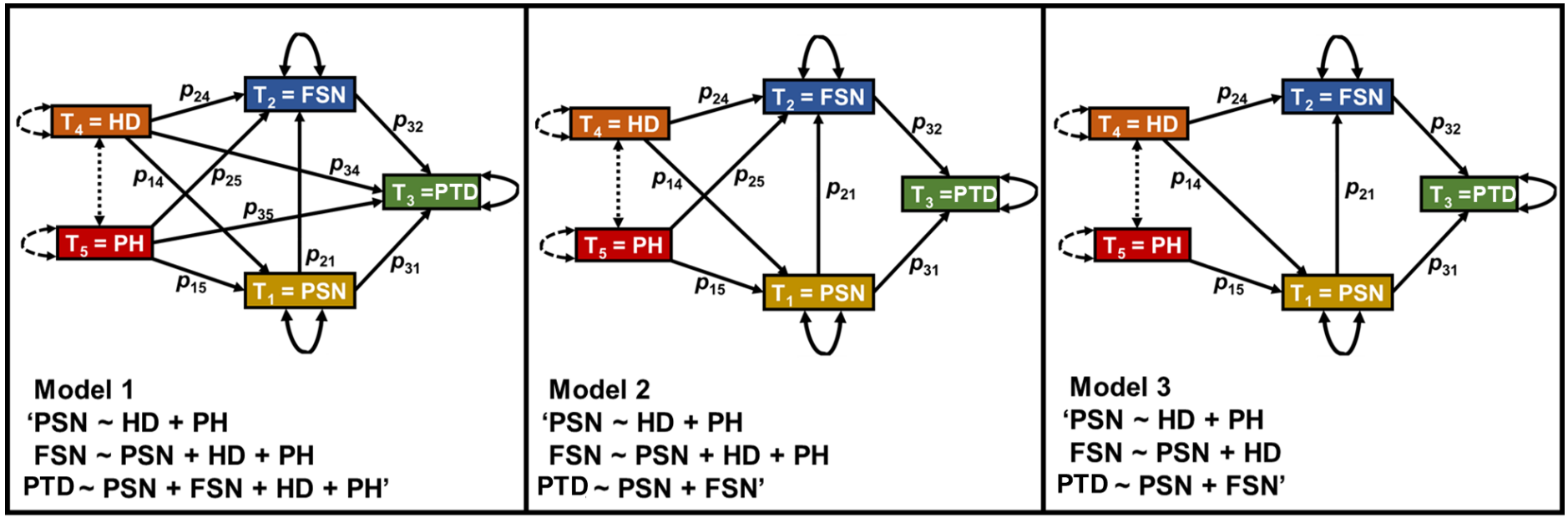


Figure S2.10 Path analysis models used to study the relationship between the investigated traits. The putative relationship is represented in the form of a path diagram. The directional arrows imply the direct effect of one trait on the another, the dashed double-headed arrow implies a covariance between two traits, the circular curved arrows represent the variance of a trait and the dashed circular curved arrows represents the variance of a trait that was not specified in the model. T_1 – T_5 denotes the traits number from 1-5. PSN, FSN, PTD, HD and PH are the traits viz., potential spikelet number, final spikelet number, pre-anthesis tip degeneration (%), heading date (days from 1 January) and plant height (cm). All the directional arrows are labeled as paths, *p*, and denote where the path begins and where it ends. The path coefficients describe as p_{15} = path began at trait T_5 and ended at trait T_1, p_{14} = path began at trait T_4 and ended at trait T_1, p_{21} = path began at trait T_1 and ended at trait T_2, p_{24} = path began at trait T_4 and ended at trait T_2, p_{25} = path began at trait T_5 and ended at trait T_2, p_{31} = path began at trait T_1 and ended at trait T_3, p_{32} = path began at trait T_2 and ended at trait T_3, p_{34} = path began at trait T_4 and ended at trait T_3, and p_{35} = path began at trait T_5 and ended at trait T_3.

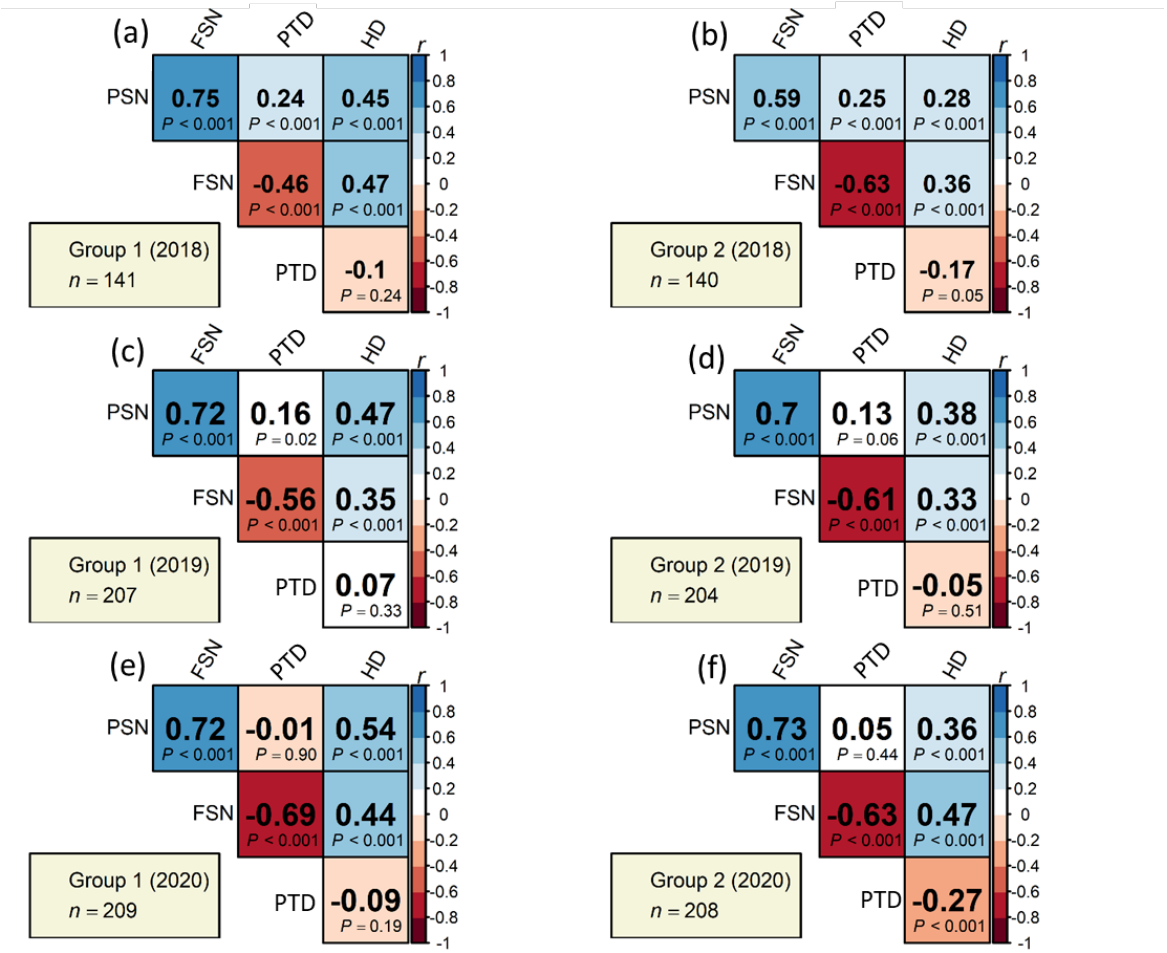


Figure S2.11 Within year Pearson's product-moment correlation (*r*) among the traits by dividing the panel into two sub-groups viz., group 1 and group 2. (a) correlation plot for group 1 accessions in 2018, (b) correlation plot for group 2 accessions in 2018, (c) correlation plot for group 1 accessions in 2019, (d) correlation plot for group 2 accessions in 2019, (e) correlation plot for group 1 accessions in 2020 and (f) correlation plot for group 2 accessions in 2020. '*n*' is the number of accessions in the respective sub-panel and *P*-value denotes the significance level of the respective correlation.

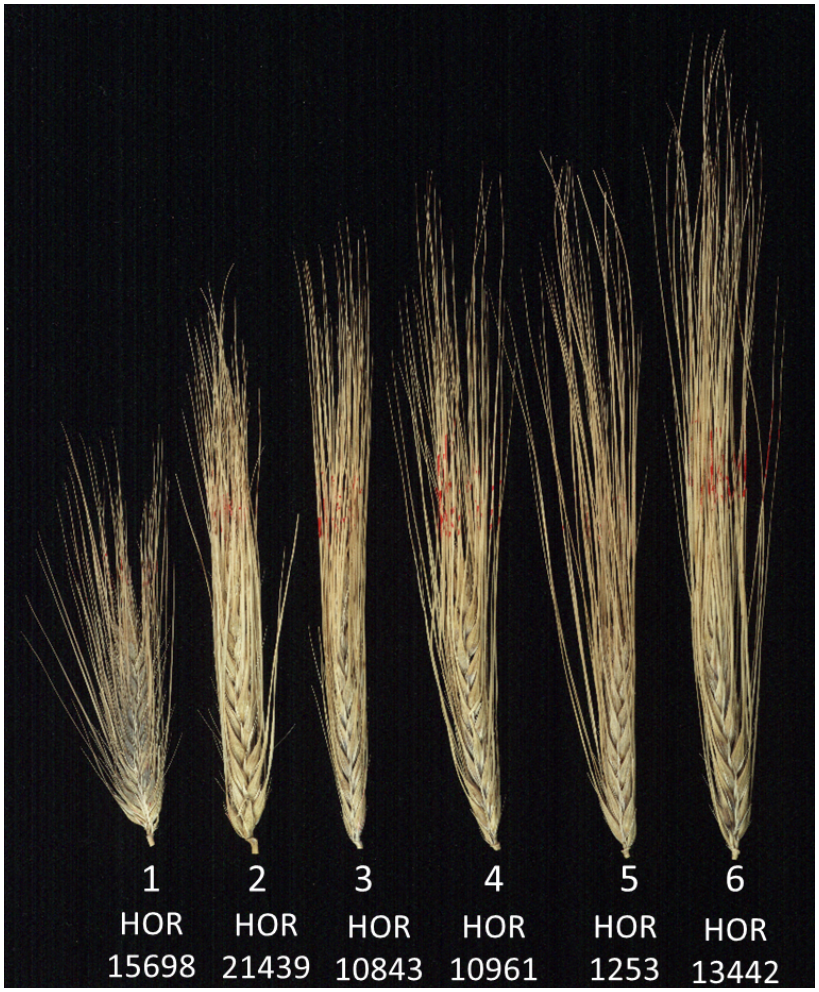


Figure S3.1 Representative samples for awn length in six-rowed barley panel. The awn length range from one to six. The accession names for the respective awn length number are mentioned in the image.

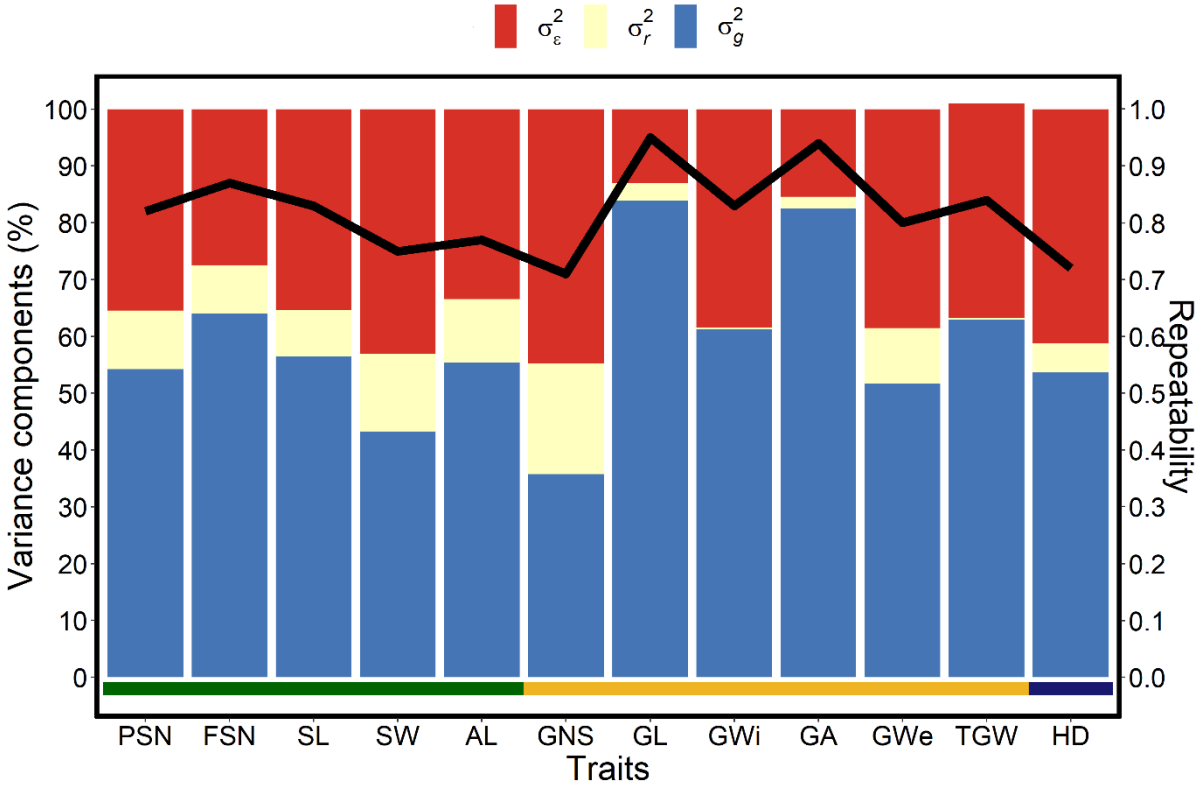


Figure S3.2 Proportion of the different variance components and heritability for each investigated trait in a panel of 288 six-rowed spring barley accessions in the year 2018. The *x*-axis represents all investigated traits, the left *y*-axis denotes the proportion of the variance components in percent, and the right *y*-axis represents the repeatability scores. The black line represents the repeatability value for the respective trait, σ_g^2 is the genotypic variance, σ_r^2 is the replication variance, and σ_ε^2 is the error or residual variance. The spike traits indicated by the green horizontal line include potential spikelet number (PSN), final spikelet number (FSN), spike length (SL in cm), spike weight (SW in g), and awn length (AL). Grain traits represented by the yellow horizontal line include grain number per spike (GNS), grain length (GL in cm), grain width (GWi in cm), grain area (GA in cm^2), grain weight (GWe in g), thousand-grain weight (TGW in g) and the shoot trait represented by blue horizontal line include heading date (HD in days from January 1st).

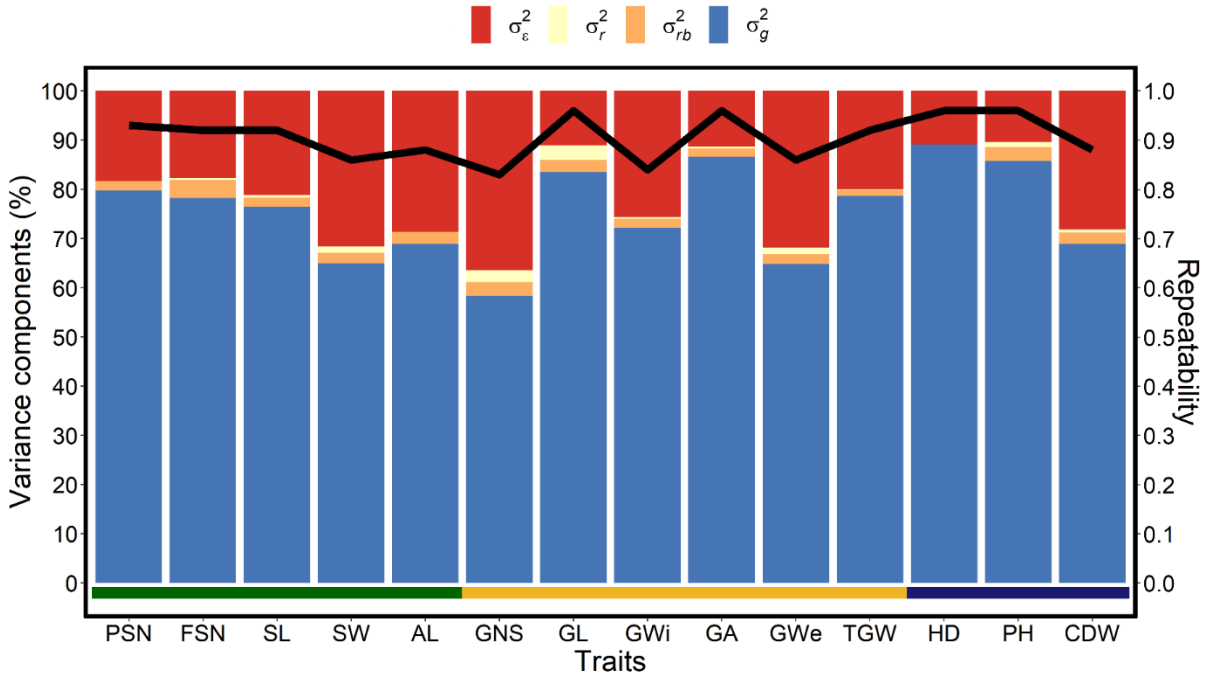


Figure S3.3 Proportion of the different variance components and heritability for each investigated trait in a panel of 417 six-rowed spring barley accessions in the year 2019. The *x*-axis represents all investigated traits, the left *y*-axis denotes the proportion of the variance components in percent, and the right *y*-axis represents the repeatability scores. The black line represents the repeatability value for the respective trait, σ_g^2 is the genotypic variance, σ_r^2 is the replication variance, σ_{rb}^2 is the replication nested in blocks variance and σ_ε^2 is the error or residual variance. The spike traits indicated by the green horizontal line include potential spikelet number (PSN), final spikelet number (FSN), spike length (SL in cm), spike weight (SW in g), and awn length (AL). Grain traits represented by the yellow horizontal line include grain number per spike (GNS), grain length (GL in cm), grain width (GWi in cm), grain area (GA in cm^2), grain weight (GWe in g), thousand-grain weight (TGW in g) and the shoot traits represented by blue horizontal line include heading date (HD in days from January 1st), plant height (PH in cm) and culm dry weight (CDW in g).

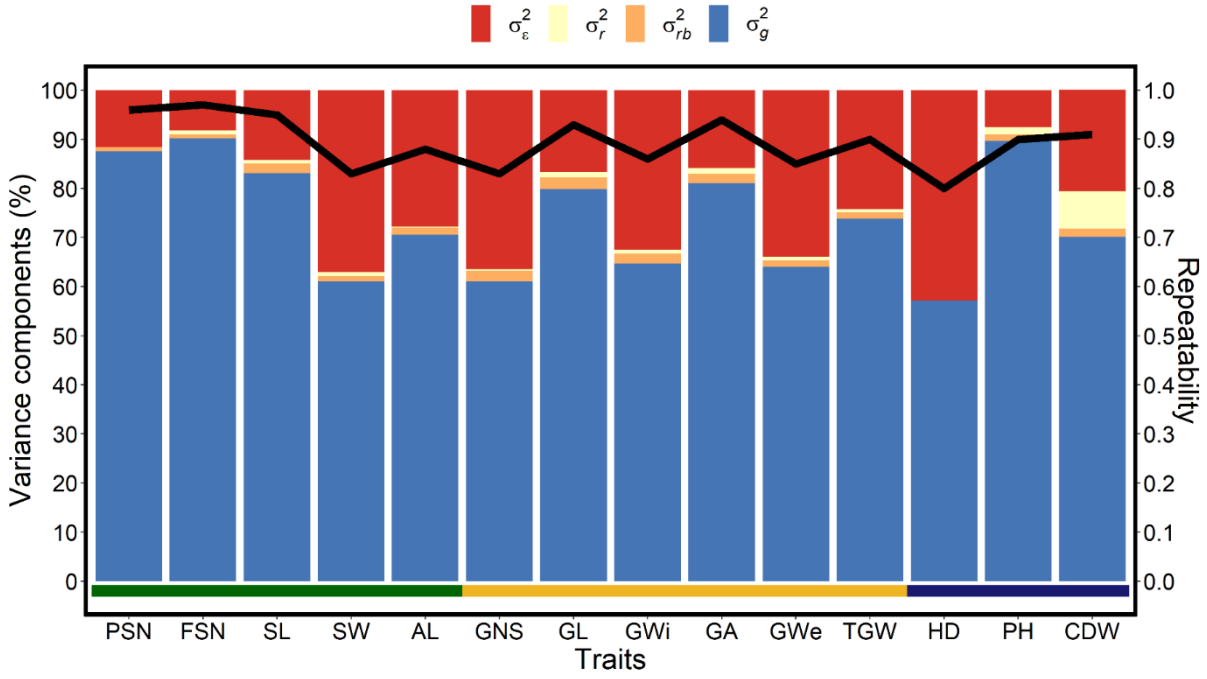


Figure S3.4 Proportion of the different variance components and heritability for each investigated trait in a panel of 417 six-rowed spring barley accessions in the year 2020. The *x*-axis represents all investigated traits, the left *y*-axis denotes the proportion of the variance components in percent, and the right *y*-axis represents the repeatability scores. The black line represents the repeatability value for the respective trait, σ_g^2 is the genotypic variance, σ_r^2 is the replication variance, σ_{rb}^2 is the replication nested in blocks variance and σ_ε^2 is the error or residual variance. The spike traits indicated by the green horizontal line includes potential spikelet number (PSN), final spikelet number (FSN), spike length (SL in cm), spike weight (SW in g), and awn length (AL). Grain traits represented by the yellow horizontal line includes grain number per spike (GNS), grain length (GL in cm), grain width (GWi in cm), grain area (GA in cm^2), grain weight (GWe in g), thousand-grain weight (TGW in g) and the shoot traits represented by blue horizontal line includes heading date (HD in days from January 1st), plant height (PH in cm) and culm dry weight (CDW in g).

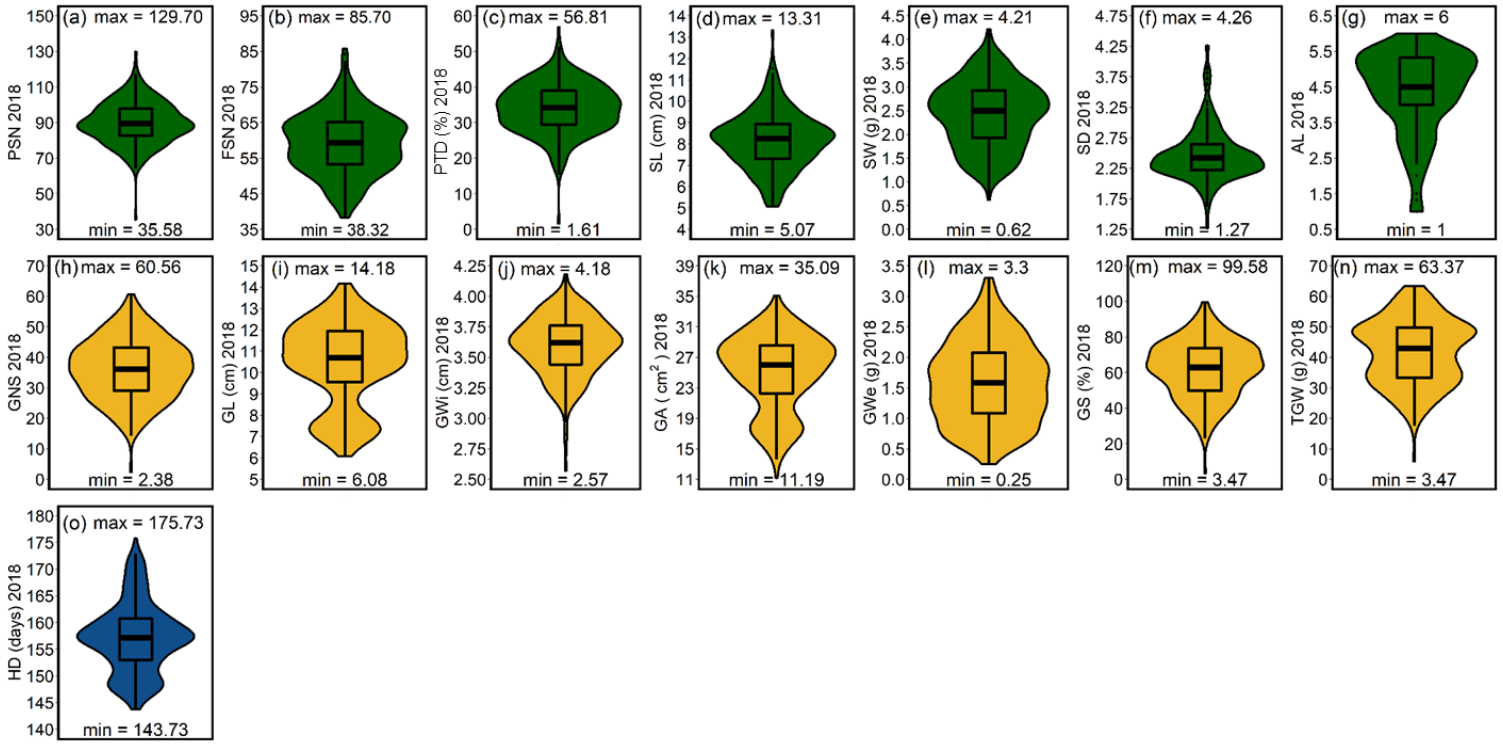


Figure S3.5 Phenotypic distribution of the investigated traits in a panel of 288 six-rowed spring barley accessions in 2018. (a-g) frequency distribution for the spike traits, (h-n) frequency distribution for grain traits and (o) frequency distribution for the shoot trait. "max" and "min" represents the maximum and minimum value for each investigated trait and the box plot within the violin plots represents the lower quartile, median and upper quartile for each trait. (a) potential spikelet number (PSN), (b) final spikelet number (FSN), (c) pre-anthesis tip degeneration (PTD in %), (d) spike length (SL in cm), € spike weight (SW in g), (f) spike density (SD); (g) awn length (AL), (h) grain number per spike (GNS), (i) grain length (GL in cm), (j) grain width (GWi in cm), (k) grain area (GA in cm^2), (l) grain weight (GWe in g), (m) grain set (GS in %), (n) thousand-grain weight (TGW in g), (o) heading date (HD in days from January 1^{st}), (p) plant height (PH in cm) and (q) culm dry weight (CDW in g)

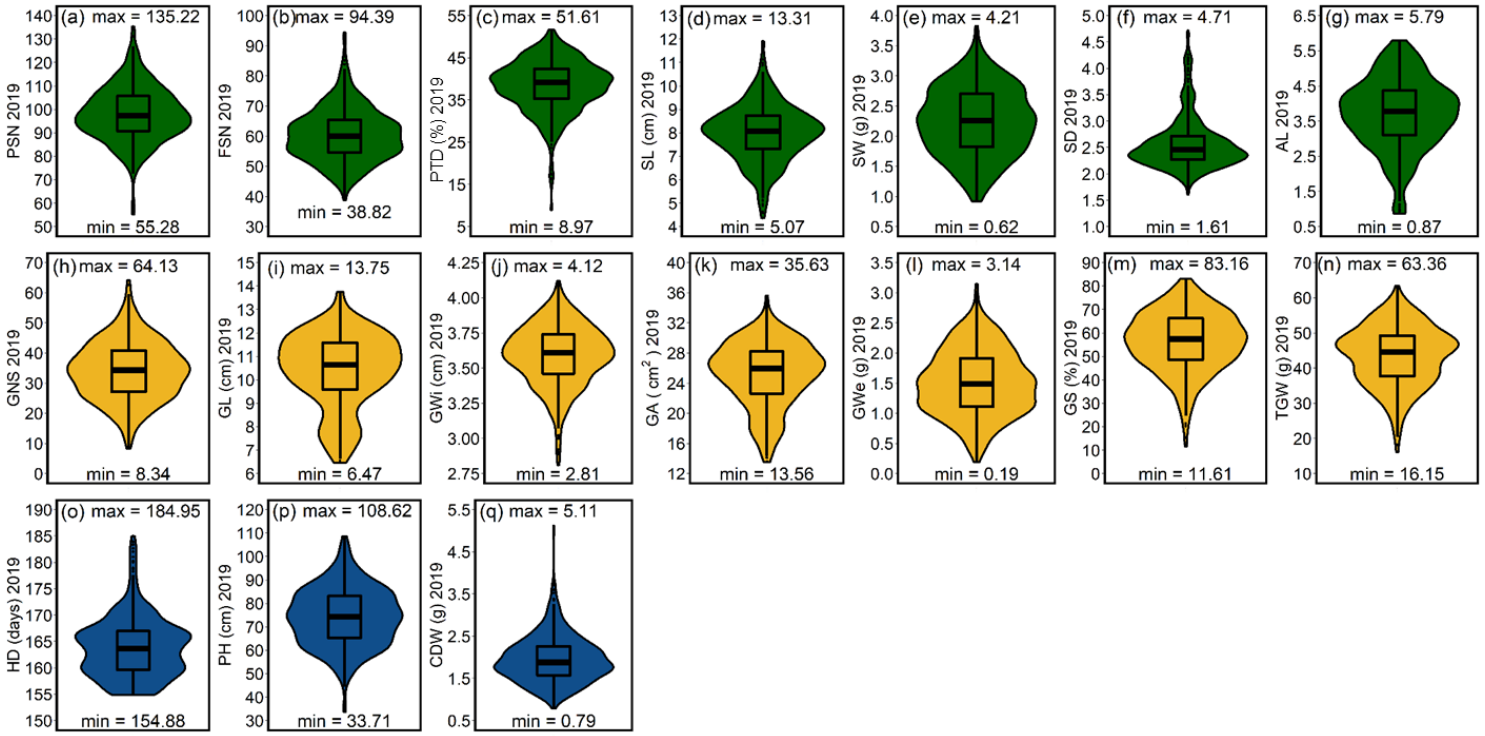


Figure S3.6 Phenotypic distribution of the investigated traits in a panel of 417 six-rowed spring barley accessions in 2019. (a-g) frequency distribution for the spike traits, (h-n) frequency distribution for grain traits and (o-q) frequency distribution for shoot traits. "max" and "min" represents the maximum and minimum value for each investigated trait and the box plot within the violin plots represents the lower quartile, median and upper quartile for each trait. (a) potential spikelet number (PSN), (b) final spikelet number (FSN), (c) pre-anthesis tip degeneration (PTD in %), (d) spike length (SL in cm), € spike weight (SW in g), (f) spike density (SD); (g) awn length (AL), (h) grain number per spike (GNS), (i) grain length (GL in cm), (j) grain width (GWi in cm), (k) grain area (GA in cm^2), (l) grain weight (GWe in g), (m) grain set (GS in %), (n) thousand-grain weight (TGW in g), (o) heading date (HD in days from January 1^{st}), (p) plant height (PH in cm) and (q) culm dry weight (CDW in g).

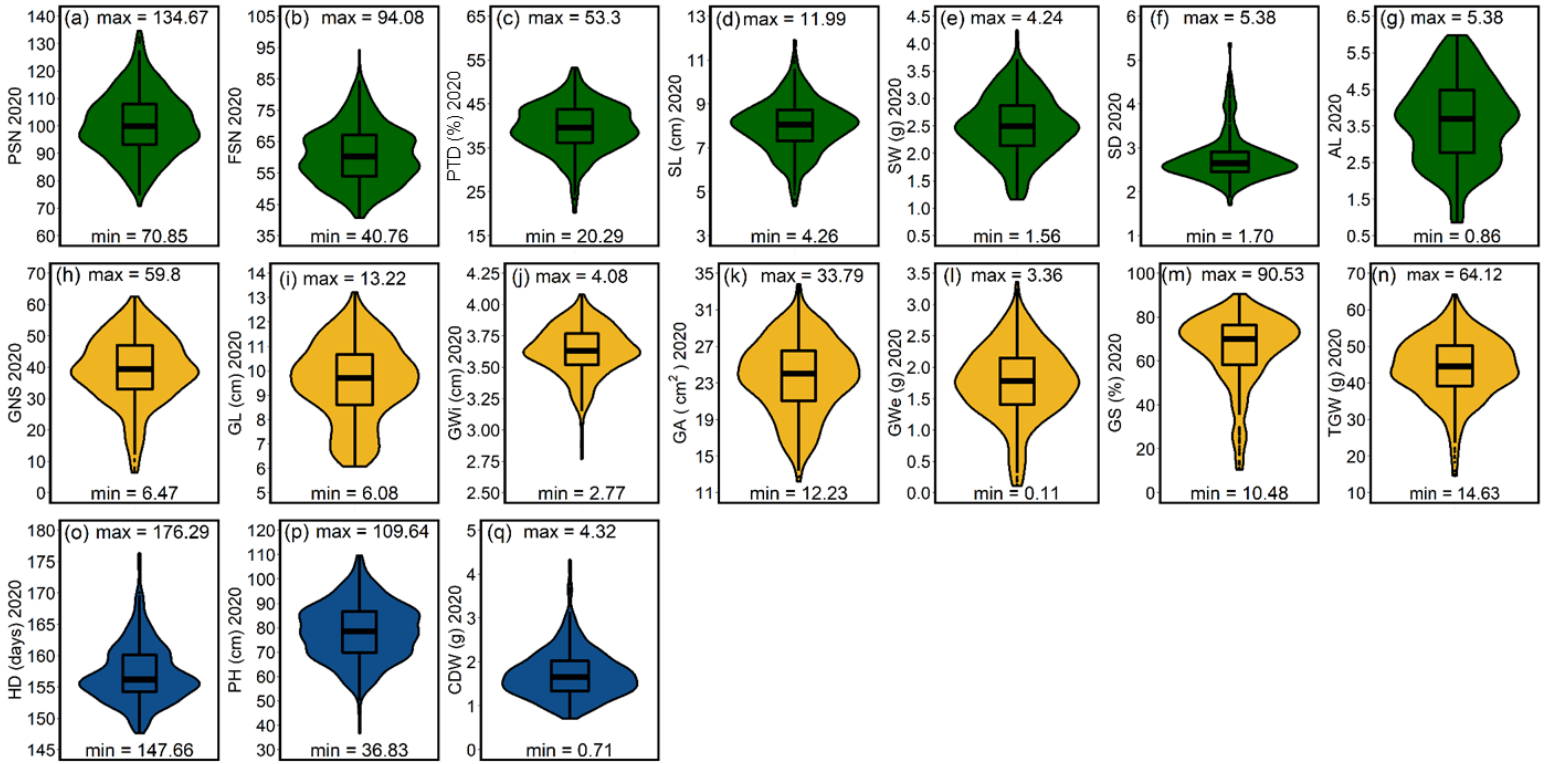


Figure S3.7 Phenotypic distribution of the investigated traits in a panel of 417 six-rowed spring barley accessions in 2020. (a-g) frequency distribution for the spike traits, (h-n) frequency distribution for grain traits and (o-q) frequency distribution for shoot traits. "max" and "min" represents the maximum and minimum value for each investigated trait and the box plot within the violin plots represents the lower quartile, median and upper quartile for each trait. (a) potential spikelet number (PSN), (b) final spikelet number (FSN), (c) pre-anthesis tip degeneration (PTD in %), (d) spike length (SL in cm), € spike weight (SW in g), (f) spike density (SD); (g) awn length (AL), (h) grain number per spike (GNS), (i) grain length (GL in cm), (j) grain width (GWi in cm), (k) grain area (GA in cm^2), (l) grain weight (GWe in g), (m) grain set (GS in %), (n) thousand-grain weight (TGW in g), (o) heading date (HD in days from January 1^{st}), (p) plant height (PH in cm) and (q) culm dry weight (CDW in g).

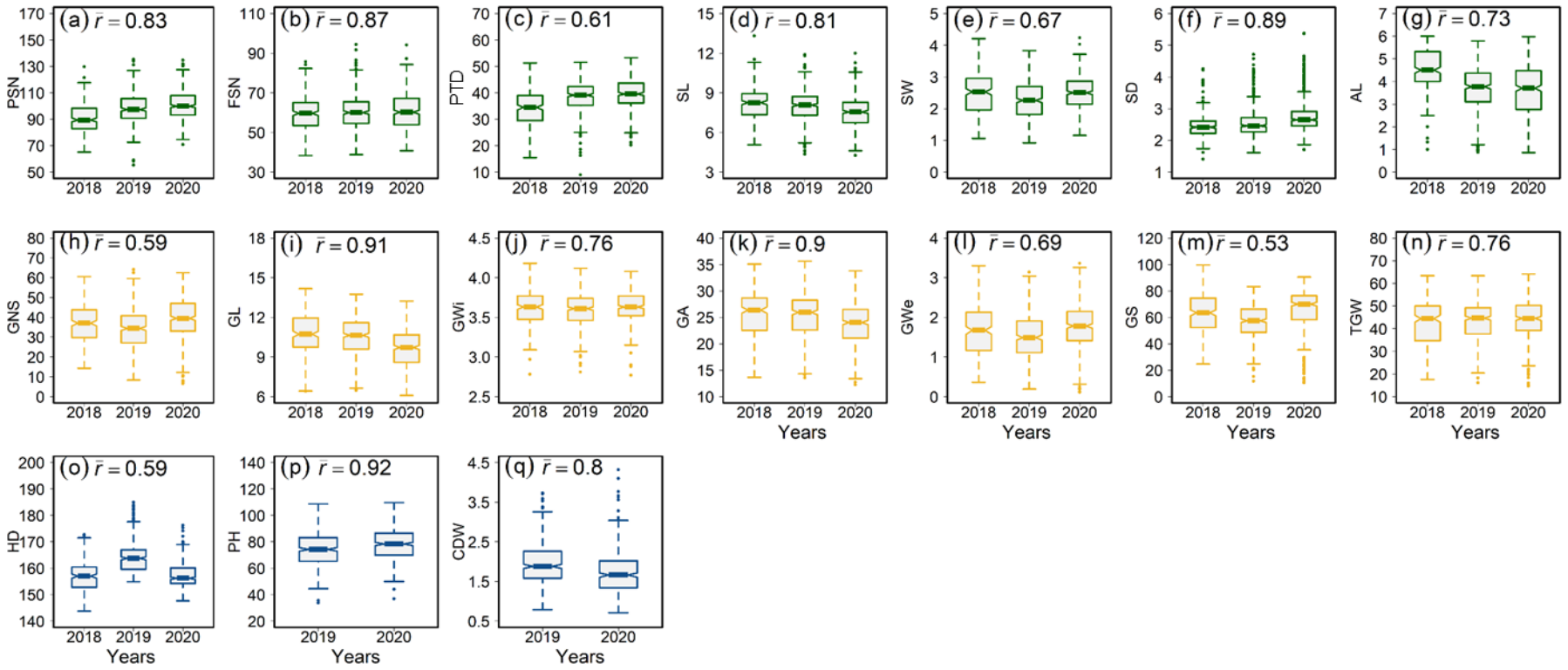


Figure S3.8 Environment (year) specific phenotypic distribution of the investigated traits in a panel of 417 six-rowed spring barley accessions with between years average trait correlation ($\bar{r}$) calculated by performing the Fisher's *z* transformation. (a-g) average correlation for the spike traits, (h-n) average correlation for grain traits and (o-q) average correlation for shoot traits. The *x*- and *y*-axis of each plot indicate the years and the particular studied trait, respectively. (a) potential spikelet number (PSN), (b) final spikelet number (FSN), (c) pre-anthesis tip degeneration (PTD in %), (d) spike length (SL in cm), € spike weight (SW in g), (f) spike density (SD); (g) awn length (AL), (h) grain number per spike (GNS), (i) grain length (GL in cm), (j) grain width (GWi in cm), (k) grain area (GA in cm^2), (l) grain weight (GWe in g), (m) grain set (GS in %), (n) thousand-grain weight (TGW in g), (o) heading date (HD in days from January 1st), (p) plant height (PH in cm) and (q) culm dry weight (CDW in g)

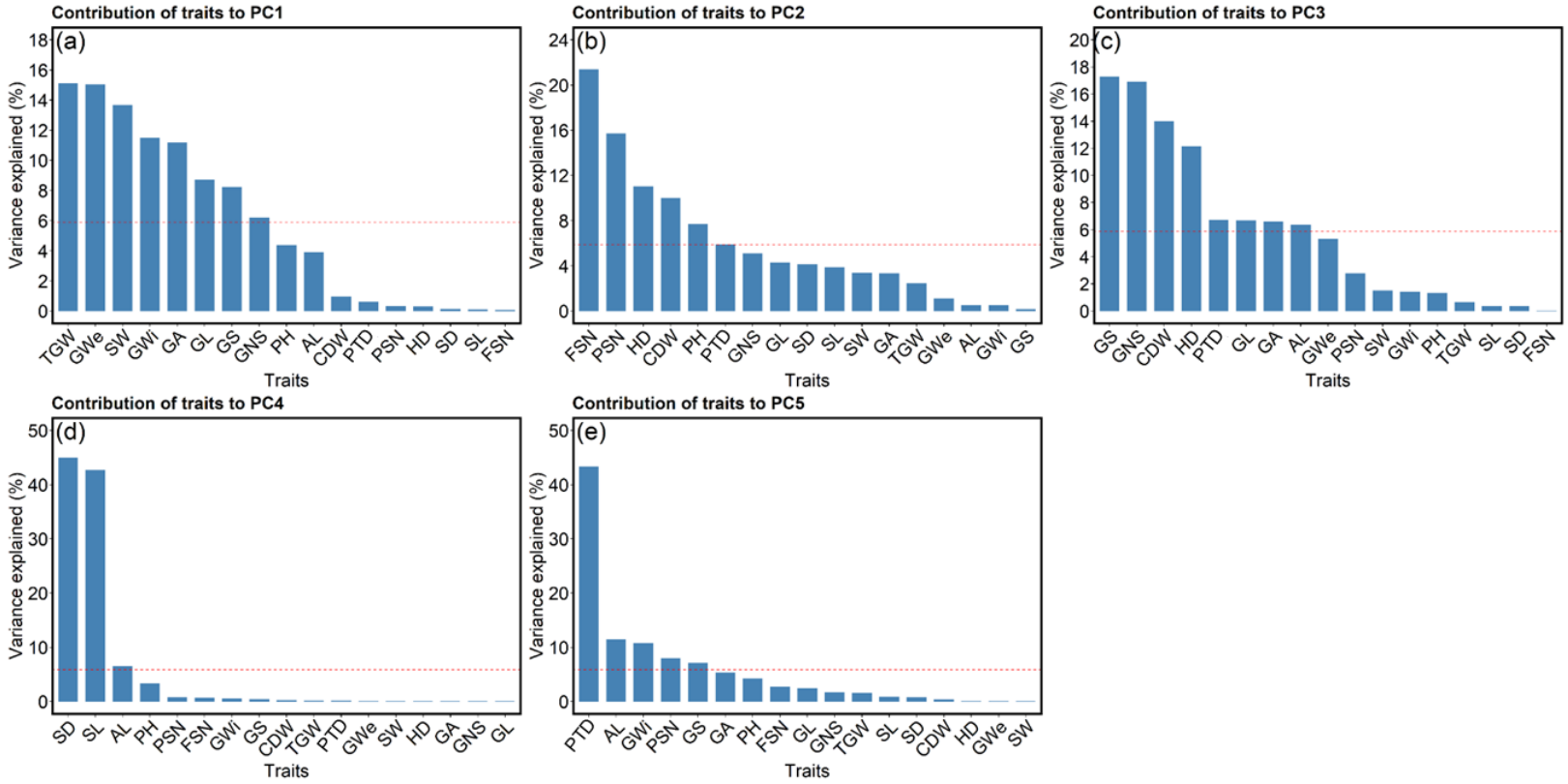


Figure S3.9 Contribution of the traits to the individual principal components (PCs). The *x*- and *y*-axis in each graph explains the traits and variance explained in percent, respectively. (a) Contribution of the traits to PC1, (b) contribution of the traits to PC2, (c) contribution of the traits to PC3, (d) contribution of the traits to PC4 and (e) contribution of the traits to PC5. PSN = potential spikelet number; FSN = final spikelet number; PTD = pre-anthesis tip degeneration (PTD in %); SL = spike length (in cm); SW = spike weight (in g); SD = spike density (in %); AL = awn length; GNS = grain number per spike; GL = grain length (in cm); GWi = grain width (in cm); GA = grain area (in cm^2); GWe = grain weight (in g); GS = grain set (in %); TGW = thousand-grain weight (in g); HD = heading date (in days from January 1st); PH = plant height (in cm) and CDW = culm dry weight (in g).

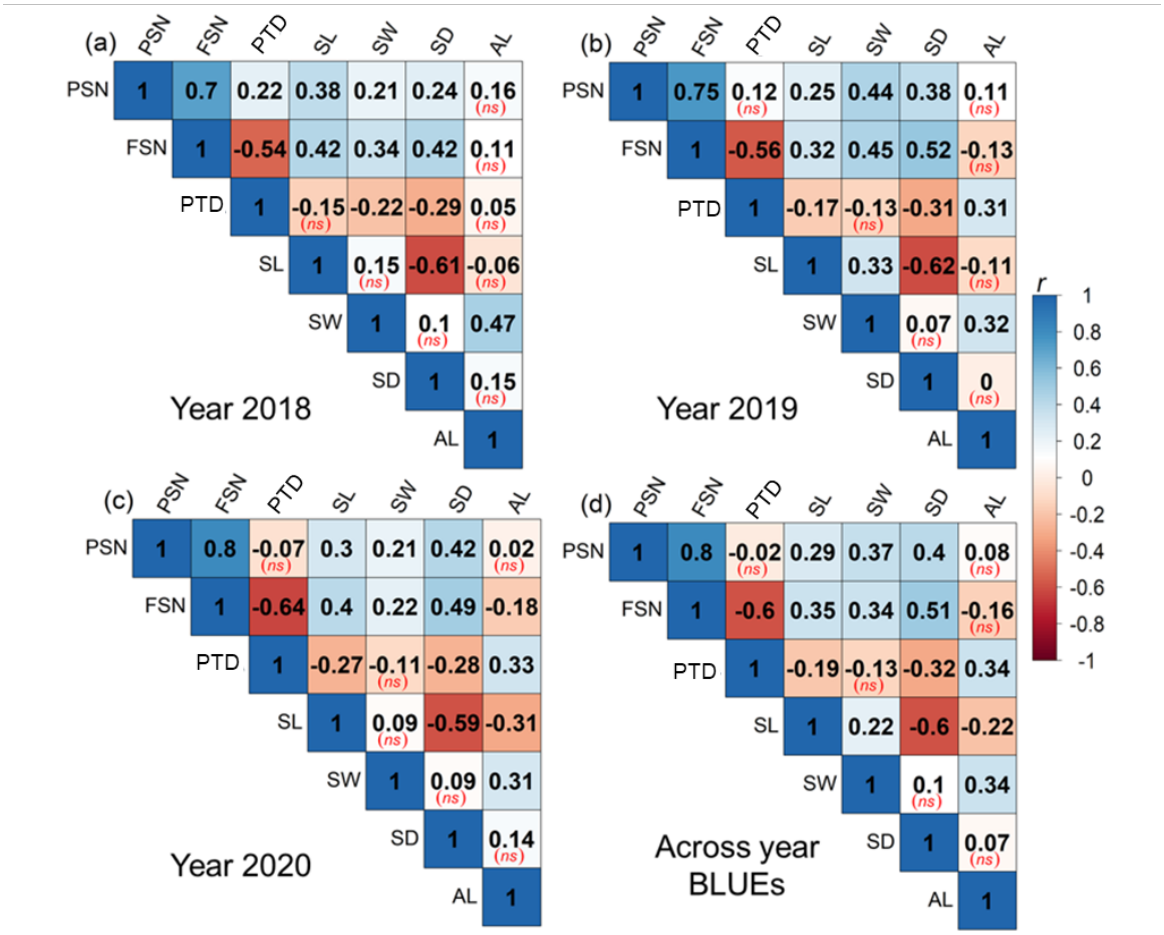


Figure S3.10 Within- and across-year Pearson-product moment correlation (*r*) analyses for spike traits. (a) correlation among the traits in 2018, (b) correlation among the traits in 2019, (c) correlation among the traits in 2020 and (d) correlation among the best linear unbiased estimations (BLUEs). *'(ns)'* shows the insignificant correlation among the traits at $P < 0.001$. The spike traits include potential spikelet number (PSN), final spikelet number (FSN), pre-anthesis tip degeneration (PTD in %), spike length (SL in cm), spike weight (SW in g), spike density (SD) and awn length (AL).

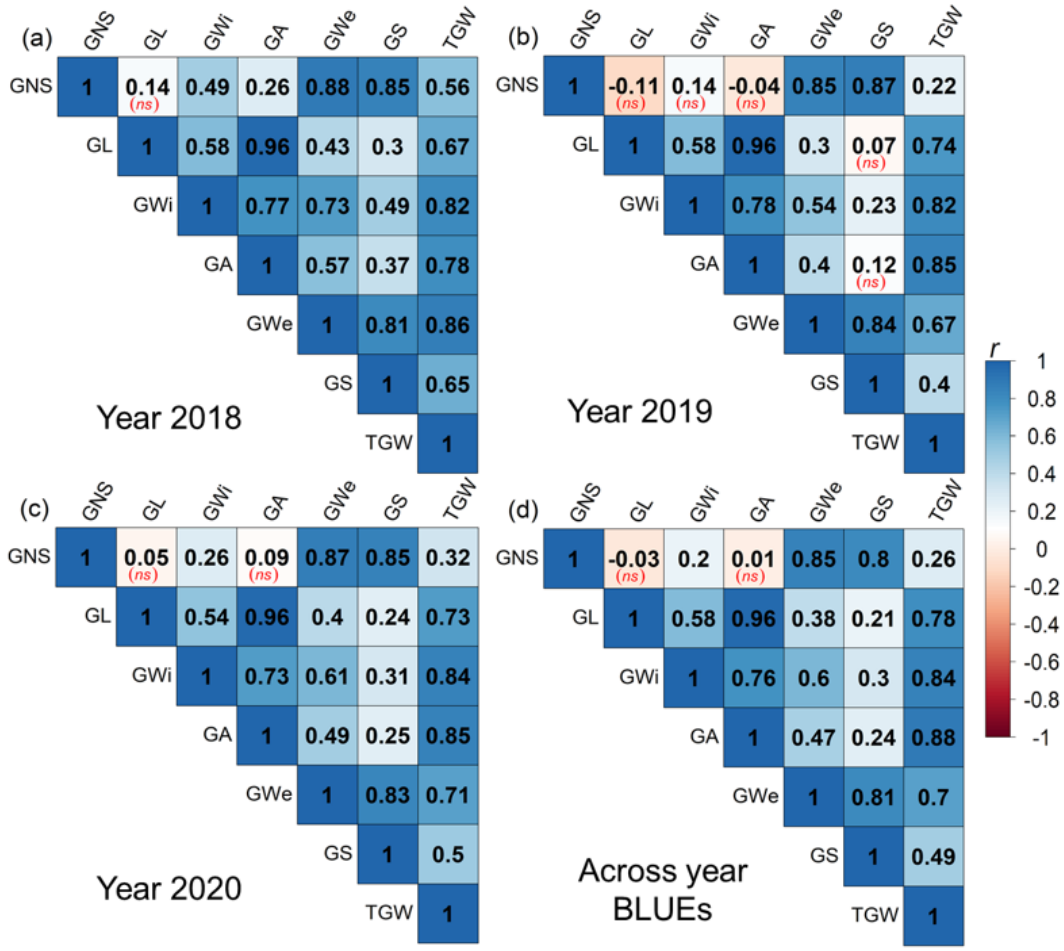


Figure S3.11 Within- and across-year Pearson-product moment correlation (*r*) analyses for grain traits. (a) correlation among the traits in 2018, (b) correlation among the traits in 2019, (c) correlation among the traits in 2020 and (d) correlation among the best linear unbiased estimations (BLUEs). *'(ns)'* shows the insignificant correlation among the traits at $P < 0.001$. Grain traits include grain number per spike (GNS), grain length (GL in cm), grain width (GWi in cm), grain area (GA in cm^2), grain weight (GWe in g), grain set (GS in %), and thousand-grain weight (TGW in g).

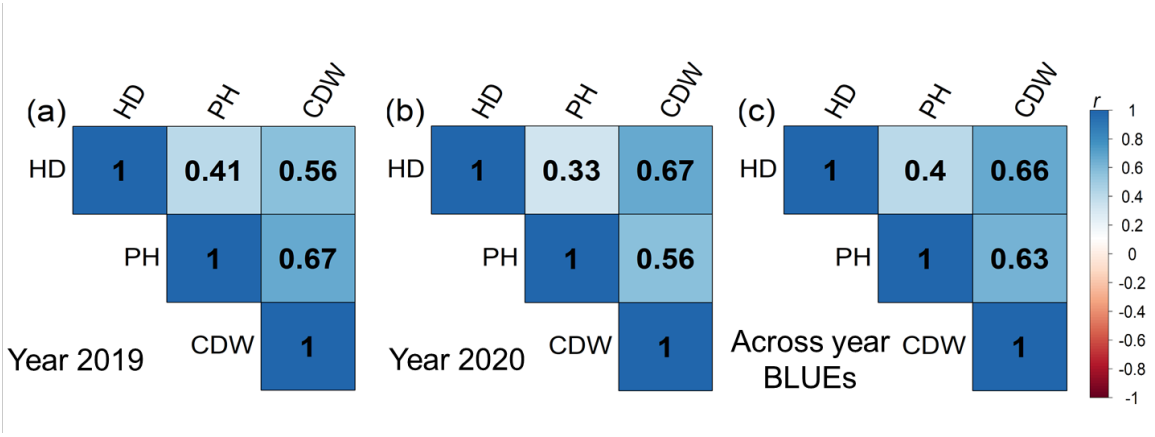


Figure S3.12 Within- and across-year Pearson-product moment correlation (*r*) analyses for shoot traits. (a) correlation among the traits in 2019, (b) correlation among the traits in 2020, and (c) correlation among the best linear unbiased estimations (BLUEs). The agronomic traits include in heading date (HD in days from January 1st), plant height (PH in cm) and culm dry weight (CDW in g).

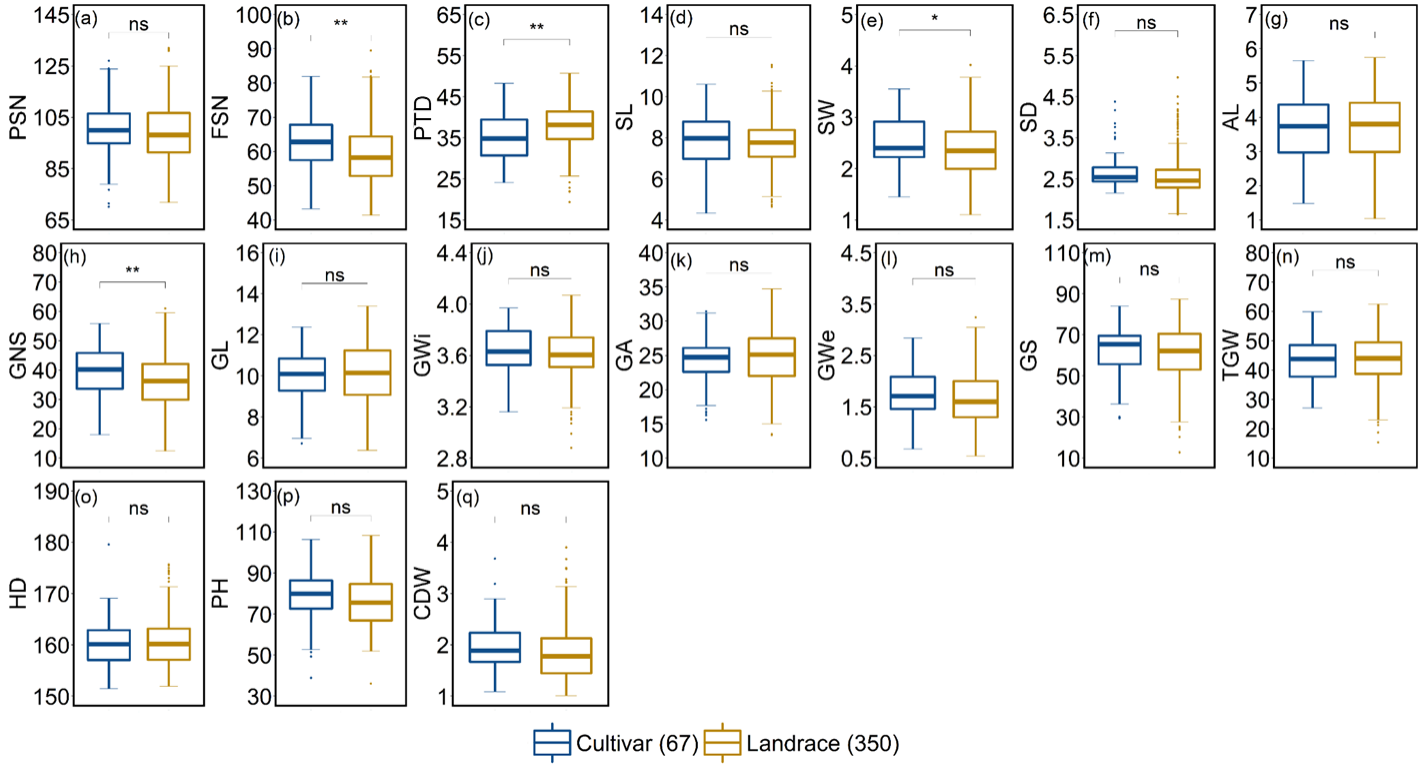


Figure S3.13 Comparison of accessions with respect to accession nature, viz., cultivar and landraces (a) potential spikelet number (PSN), (b) final spikelet number (FSN), (c) pre-anthesis tip degeneration (PTD in %), (d) spike length (SL in cm), € spike weight (SW in g), (f) spike density (SD); (g) awn length (AL), (h) grain number per spike (GNS), (i) grain length (GL in cm), (j) grain width (GWi in cm), (k) grain area (GA in cm^2), (l) grain weight (GWe in g), (m) grain set (GS in %), (n) thousand-grain weight (TGW in g), (o) heading date (HD in days from January 1st), (p) plant height (PH in cm) and (q) culm dry weight (CDW in g). '*n*' denotes the number of accessions; **= 0.01 probability level; *= 0.05 and 'ns' represents insignificant differences.

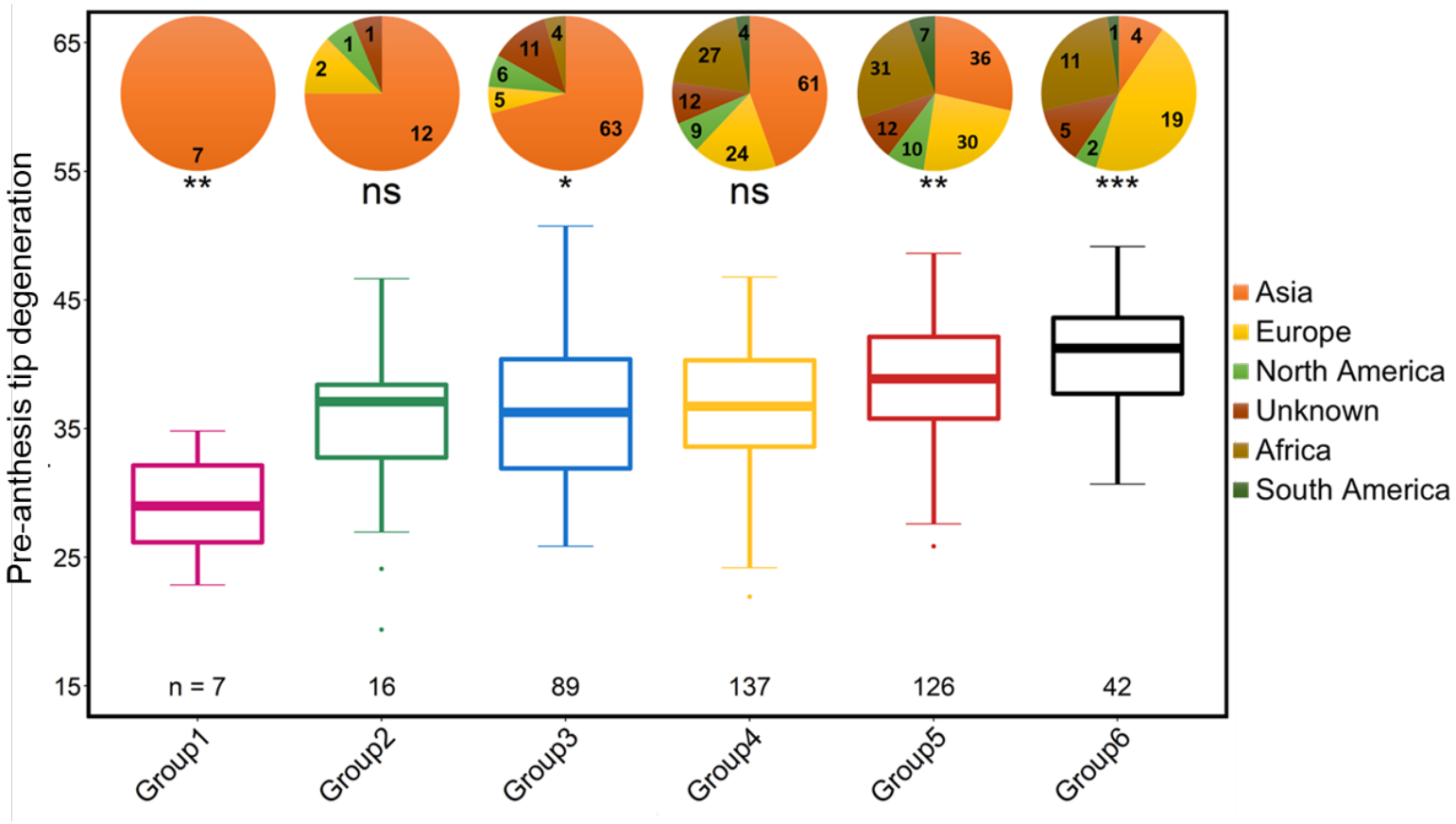


Figure S3.14 Distribution (grouping) of the 417 barley accessions based on their awn length (AL) on 1 to 6 ordinal scale. The Pie charts within the figure represent the continent-wise distribution of AL in each group. The accessions are grouped with AL ranging from 0–1, 1–2, 2–3, 3–4, 4–5, and 5–6. *, **, ***, and ns represent the significance (*P*) values based on student's t-test as $P < 0.05$, $P < 0.01$, $P < 0.001$, and $P > 0.5$, respectively.

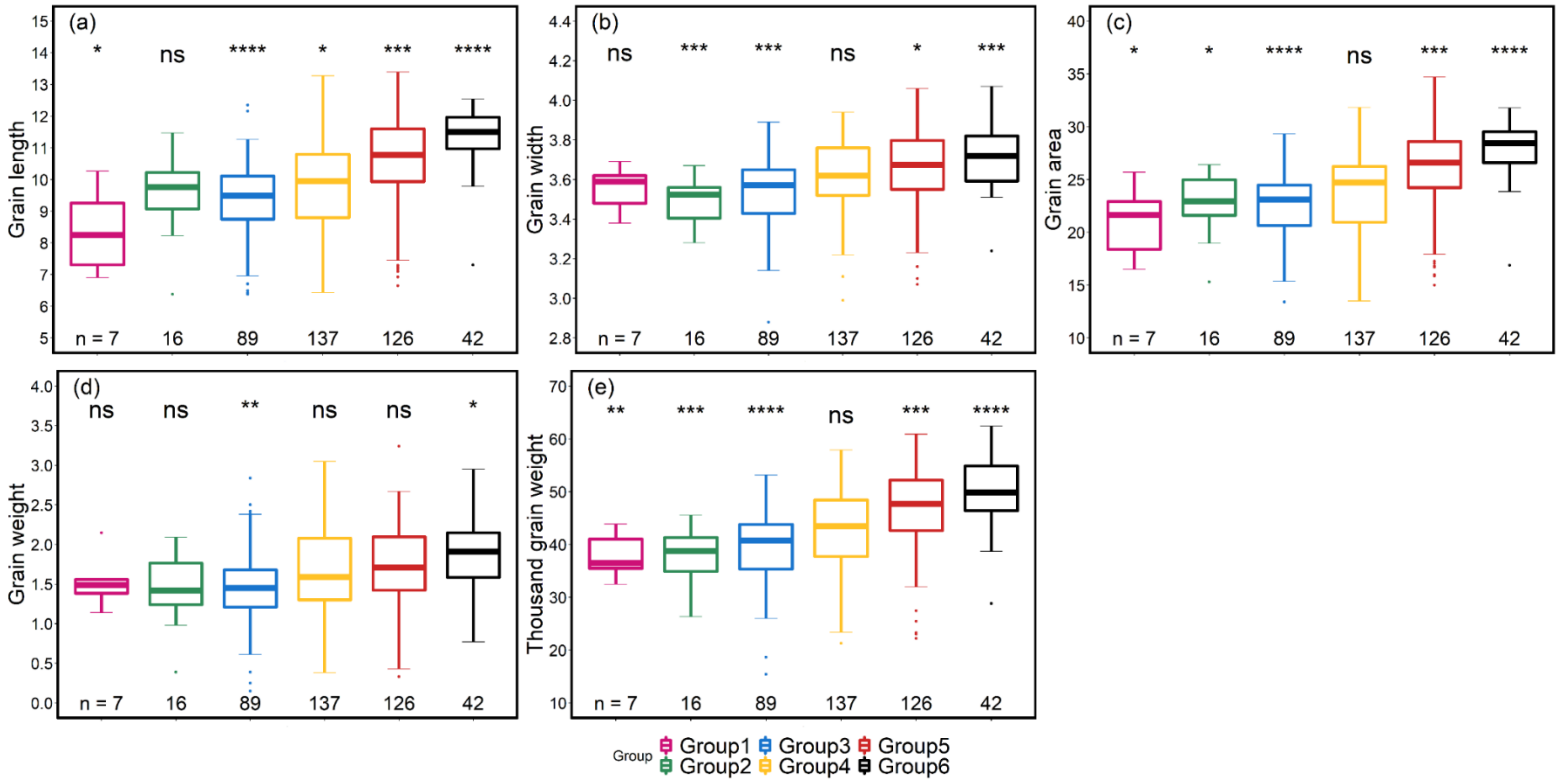


Figure S3.15 Effect of awn length (AL) of the grain traits. The accessions are grouped with AL ranging from 0–1 (group1) , 1–2 (group2) , 2–3 (group3) , 3–4 (group4) , 4–5 (group5) , and 5–6 (group6) . *, **, ***, and ns represent the significance (*P*) values based on student's t-test as $P < 0.05$, $P < 0.01$, $P < 0.001$, and $P > 0.5$, respectively.

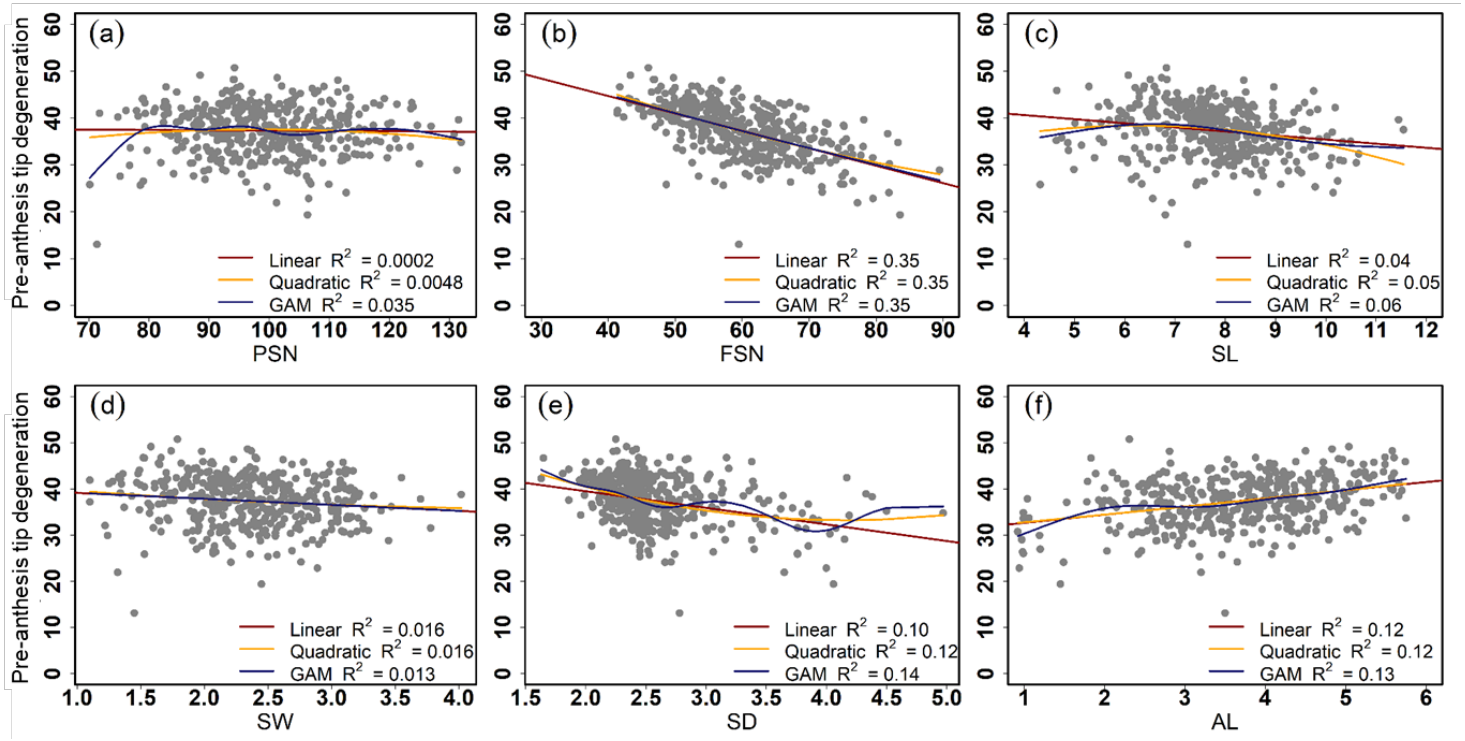


Figure S3.16 Comparison of different regression models explaining the relationship between pre-anthesis tip degeneration (PTD) and all the investigated spike traits. The *x*- and *y*-axis represent the different spike traits and PTD respectively. (a) regression analysis between potential spikelet number (PSN) and PTD, (b) regression analysis between final spikelet number (FSN) and PTD, (c) regression analysis between spike length (SL) and PTD, (d) regression analysis between spike weight (SWe) and PTD, (e) regression analysis between spike density (SD) and PTD and (f) regression analysis between awn length (AL) and PTD. The red line shows the linear regression, the yellow line indicates the quadratic regression, the blue line indicates the generalized additive model (GAM) and R^2 is the coefficient of determination.

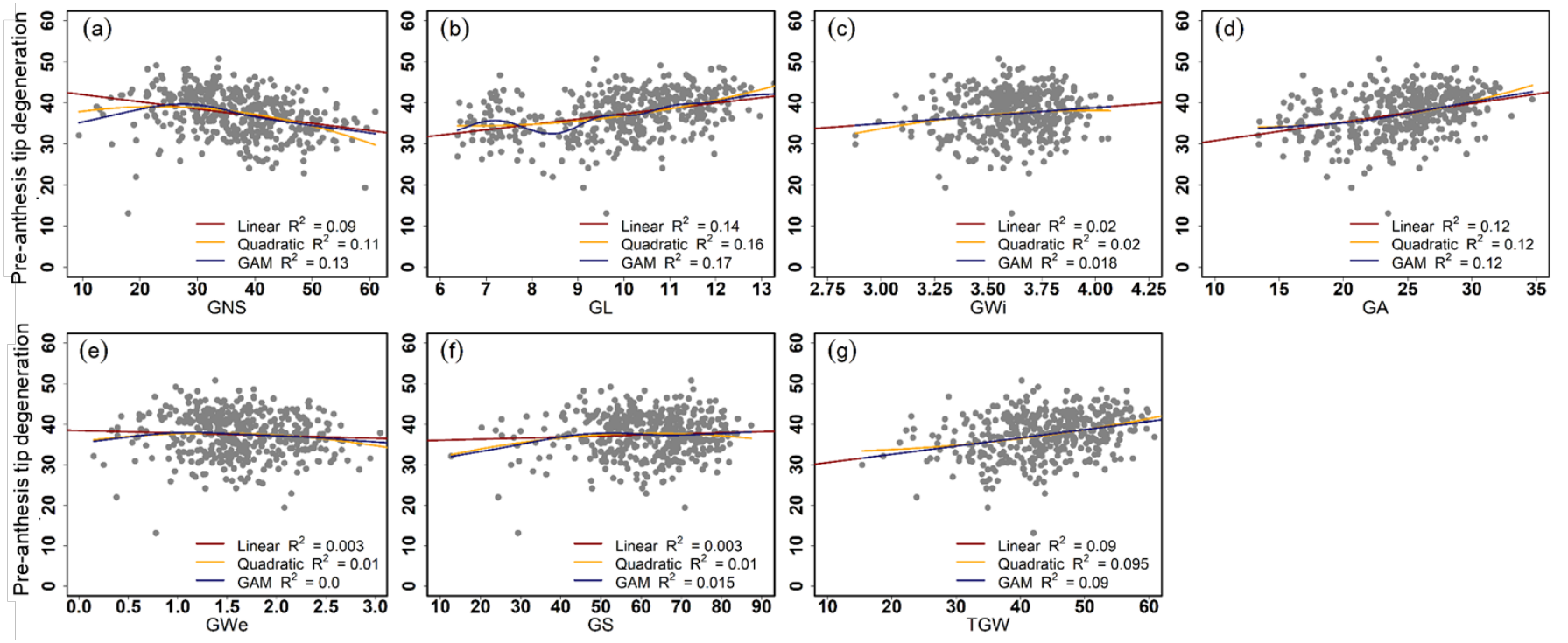


Figure S3.17 Comparison of different regression models explaining the relationship between pre-anthesis tip degeneration (PTD) and all the investigated grain traits. The *x*- and *y*-axis represent the different grain traits and PTD respectively. (a) regression analysis between grain number per spike (GNS) and PTD, (b) regression analysis between grain length (GL) and PTD, (c) regression analysis between grain width (GWi) and PTD, (d) regression analysis between grain area (GA) and PTD, (e) regression analysis between grain weight (GWe) and PTD, (f) regression analysis between grain set (GS) and PTD and (g) regression analysis between thousand-grain weight (TGW) and PTD. The red line shows the linear regression, the yellow line indicates the quadratic regression, the blue line indicates the generalized additive model (GAM) and R2 is the coefficient of determination.

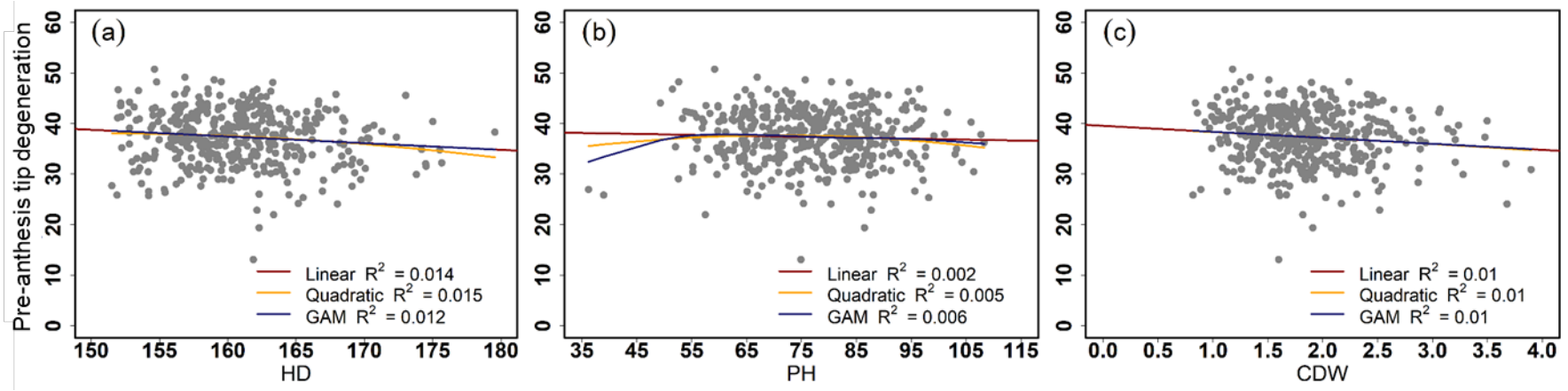


Figure S3.18 Comparison of different regression models explaining the relationship between pre-anthesis tip degeneration (PTD) and all the investigated shoot traits. The *x*- and *y*- axis represents the different agronomic traits and spikelet abortion respectively. (a) regression analysis between heading date (HD) and PTD, (b) regression analysis between plant height (PH) and PTD, and (c) regression analysis between culm dry weight (CDW) and PTD. The red line shows the linear regression, the yellow line indicates the quadratic regression, the blue line indicates the generalized additive model (GAM) and R2 is the coefficient of determination.

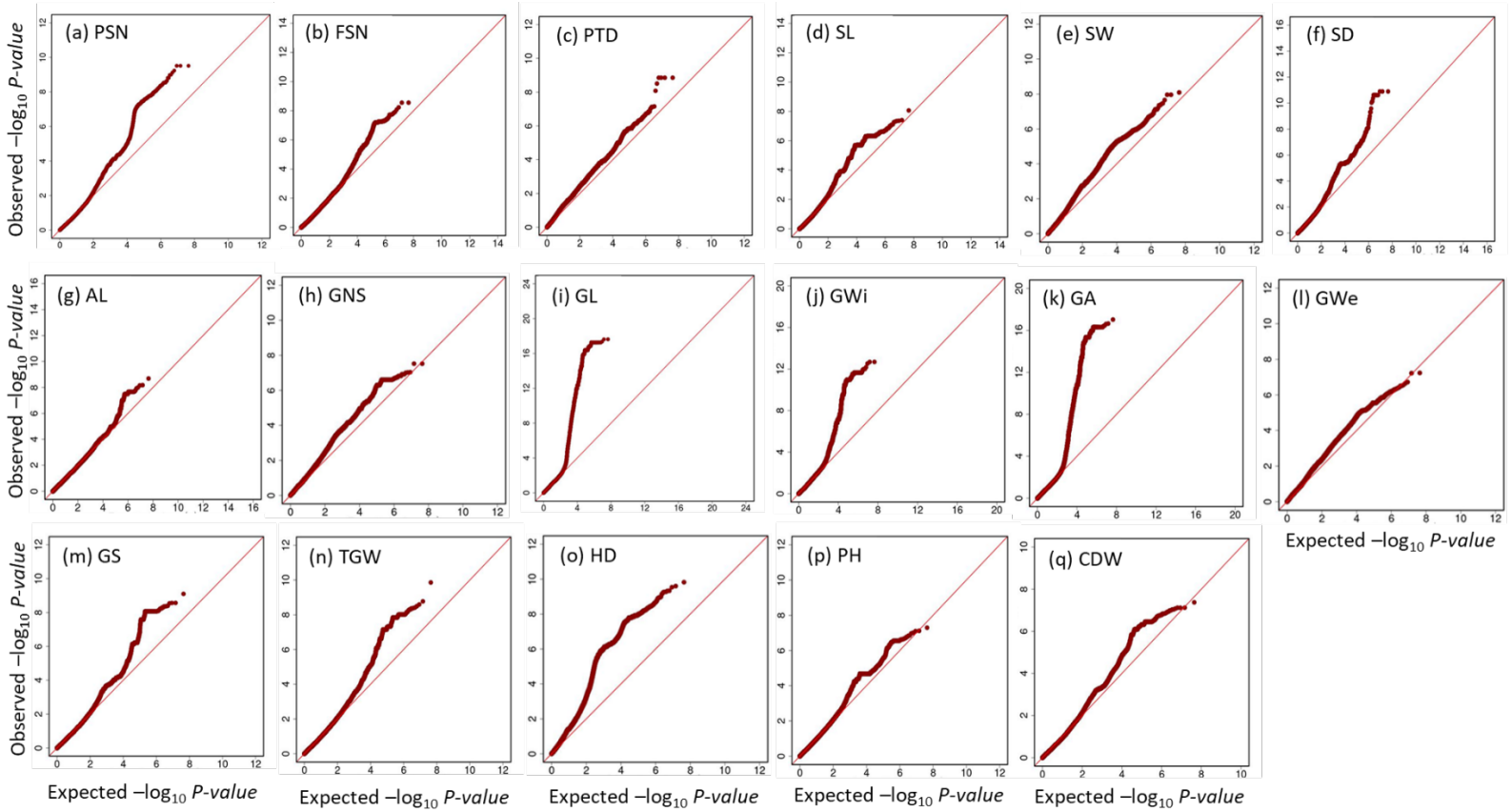


Figure S4.1 Quantile-Quantile (QQ) plots for all the investigated traits. QQ plots for (a) potential spikelet number, (b) final spikelet number, (c) pre-anthesis tip degeneration, (d) spike length, (e) spike weight, (f) spike density, (g) awn length, (h) grain number per spike, (i) grain length, (j) grain width, (k) grain area, (l) grain weight, (m) grain set, (n) thousand grain weight, (o) heading date, (p) plant height, and (q) culm dry weight.

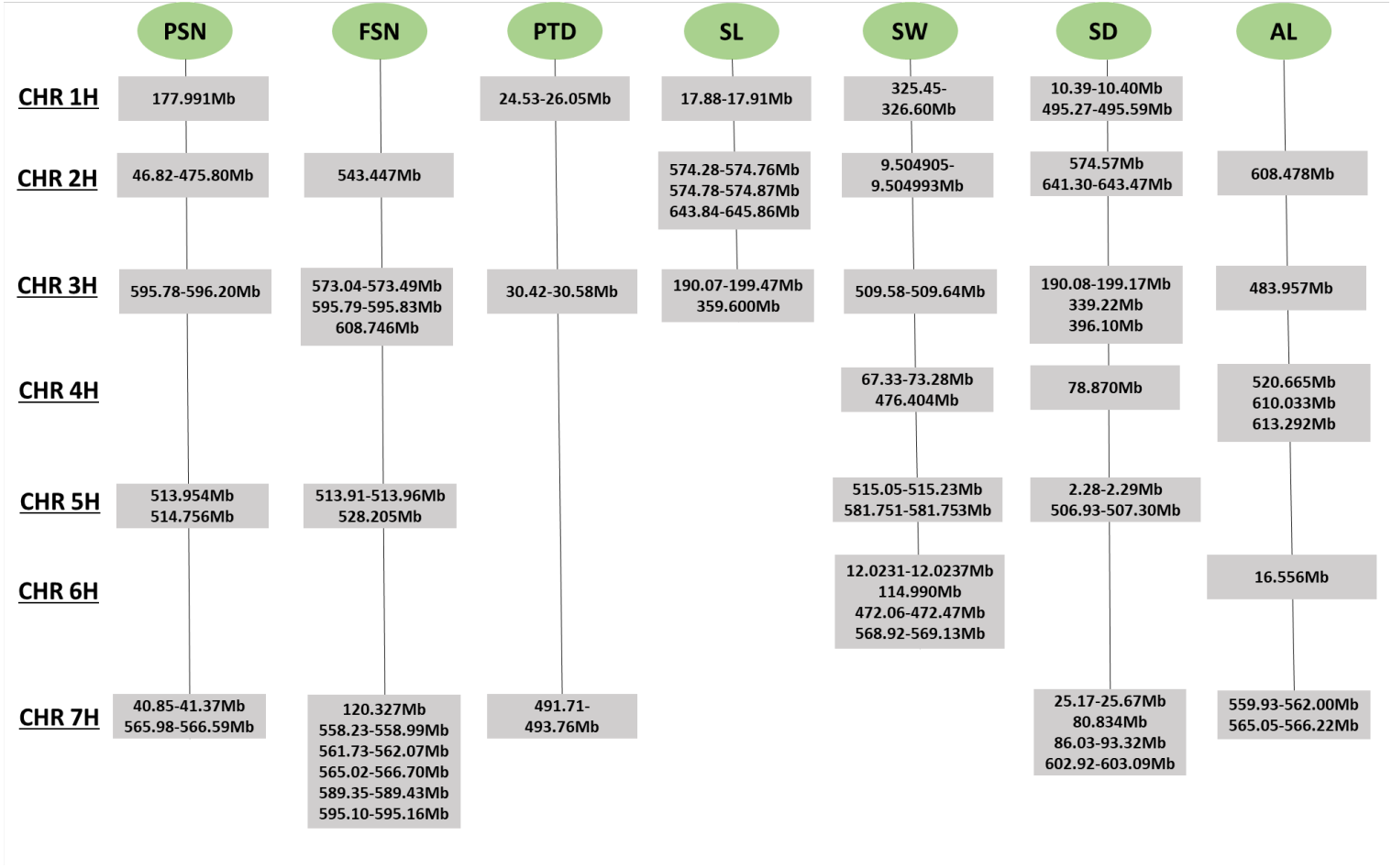


Figure S4. 2 Unique QTL for spike traits distributed on all seven chromosomes of barley where PSN = potential spikelet number, FSN = final spikelet number, PTD = pre-anthesis tip degeneration, SL = spike length, SW = spike weight, SD = spike density, and AL = awn length.

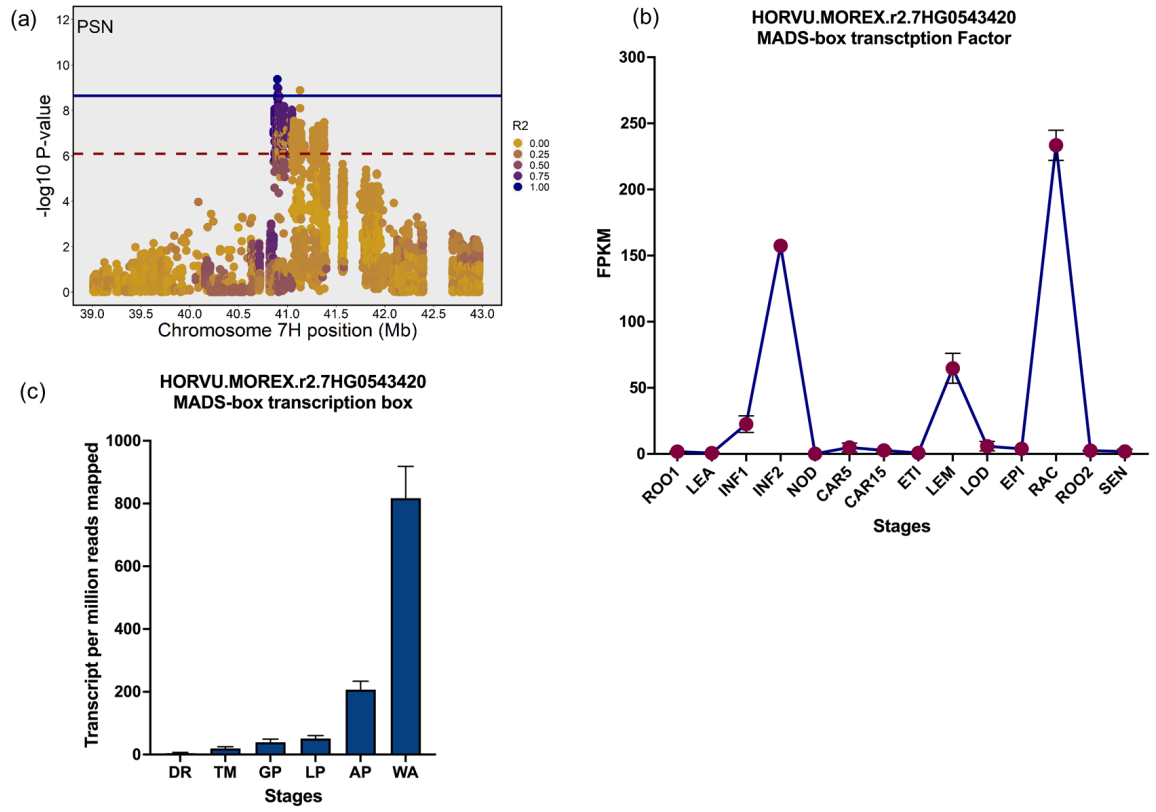


Figure S4. 3 Linkage disequilibrium and expression pattern for *HORVU.MOREX.r2.7HG0543420*. (a) LD plots for potential spikelet number on chromosome 7HS. Blue line is the Bonferroni correction at alpha level of 0.05 (BC = 8.64) and red dashed line is false discovery rate at alpha level of 0.05 (FDR = 6.09). (b) Gene expression for *HORVU.MOREX.r2.7HG0543420* in different tissues. (c) Gene expression for *HORVU.MOREX.r2.7HG0543420* in inflorescence meristem (IM). ROO1 = roots from seedlings (10cm shoot stage), LEA = shoots from seedlings (10cm shoot stage), INF1 = young inflorescence (5mm), INF2 = developing inflorescence (1-1.5cm), NOD = developing tillers, 3rd internode (42 DAP), CAR5 = developing grain (5 DAP), CAR15 = developing grain (15 DAP), ETI = etiolated seedling, dark condition (10 DAP), LEM = inflorescences, lemma (42 DAP), LOD = inflorescences, lodicule (42DAP), EPI = epidermal strips (28 DAP), RAC = inflorescence, rachis (35 DAP), ROO2 = roots (28 DAP), SEN = senescing leaves (58 DAP) and FPKM = fragments per kilobase of exon per million mapped fragments, DR = double ridge, TM = triple mound, GP = Glume primordia, LP = lemma primordia, SP = stamen primordia, AP = awn primordia, WA = white anther.

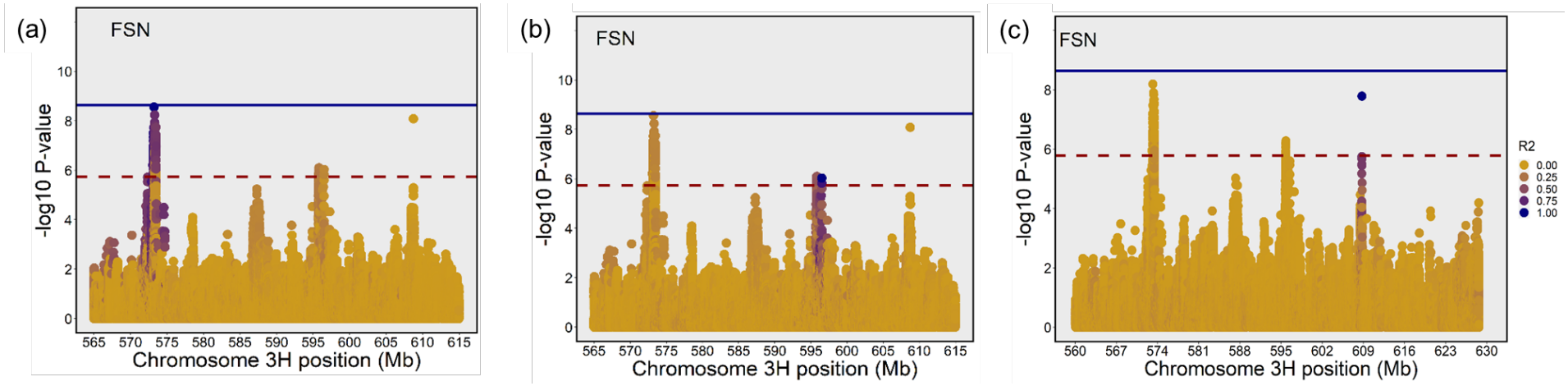


Figure S4.4 Linkage disequilibrium plots for final spikelet number on chromosome 3HL. Blue line is the Bonferroni correction at alpha level of 0.05 (BC = 8.64) and red dashed line is false discovery rate at alpha level of 0.05 (FDR = 5.78).

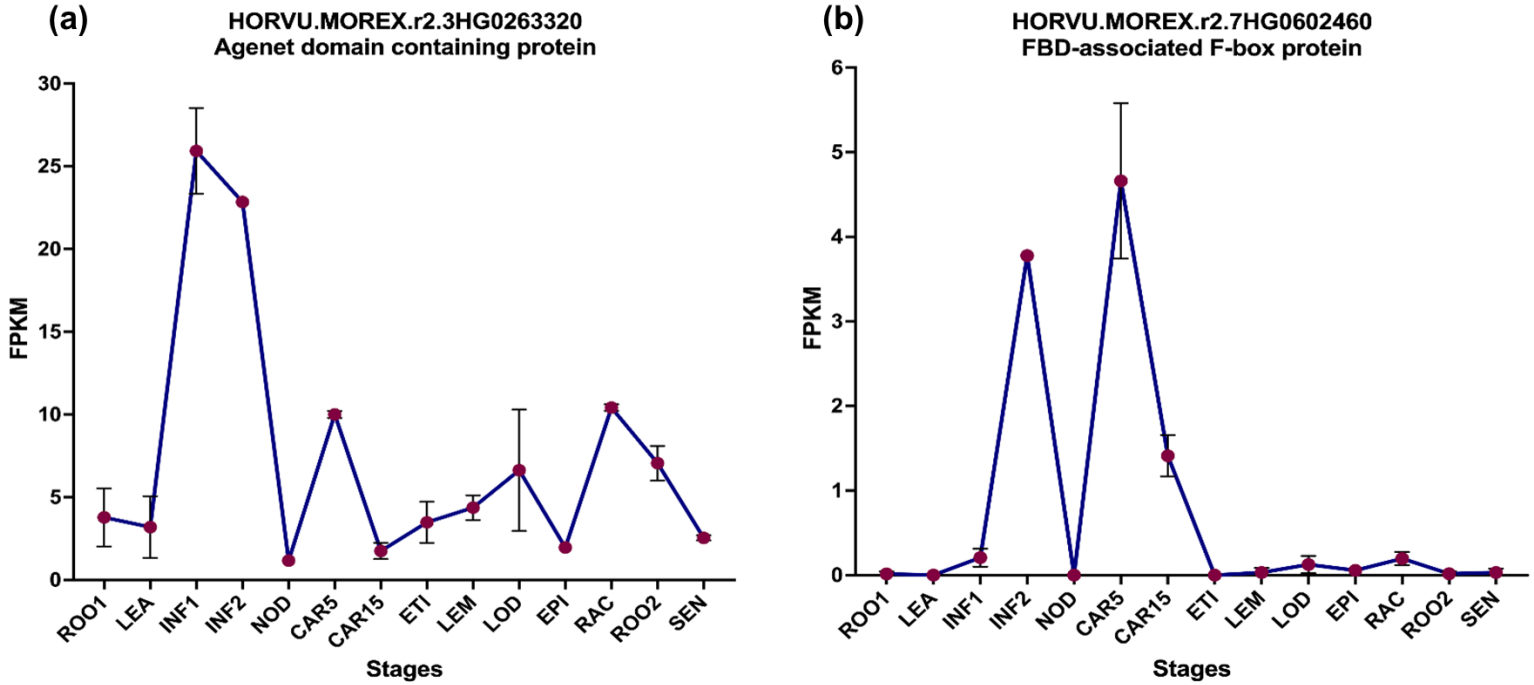


Figure S4.5 Expression pattern for (a) *HORVU.MOREX.r2.3HG0263320*, and (b) *HORVU.MOREX.r2.7HG0602460* in different tissues. ROO1 = roots from seedlings (10cm shoot stage), LEA = shoots from seedlings (10cm shoot stage), INF1 = young inflorescence (5mm), INF2 = developing inflorescence (1-1.5cm), NOD = developing tillers, 3rd internode (42 DAP), CAR5 = developing grain (5 DAP), CAR15 = developing grain (15 DAP), ETI = etiolated seedling, dark condition (10 DAP), LEM = inflorescences, lemma (42 DAP), LOD = inflorescences, lodicule (42DAP), EPI = epidermal strips (28 DAP), RAC = inflorescence, rachis (35 DAP), ROO2 = roots (28 DAP), SEN = senescing leaves (58 DAP) and FPKM = fragments per kilobase of exon per million mapped fragments. The expression data was obtained from Colmsee et al 2015.

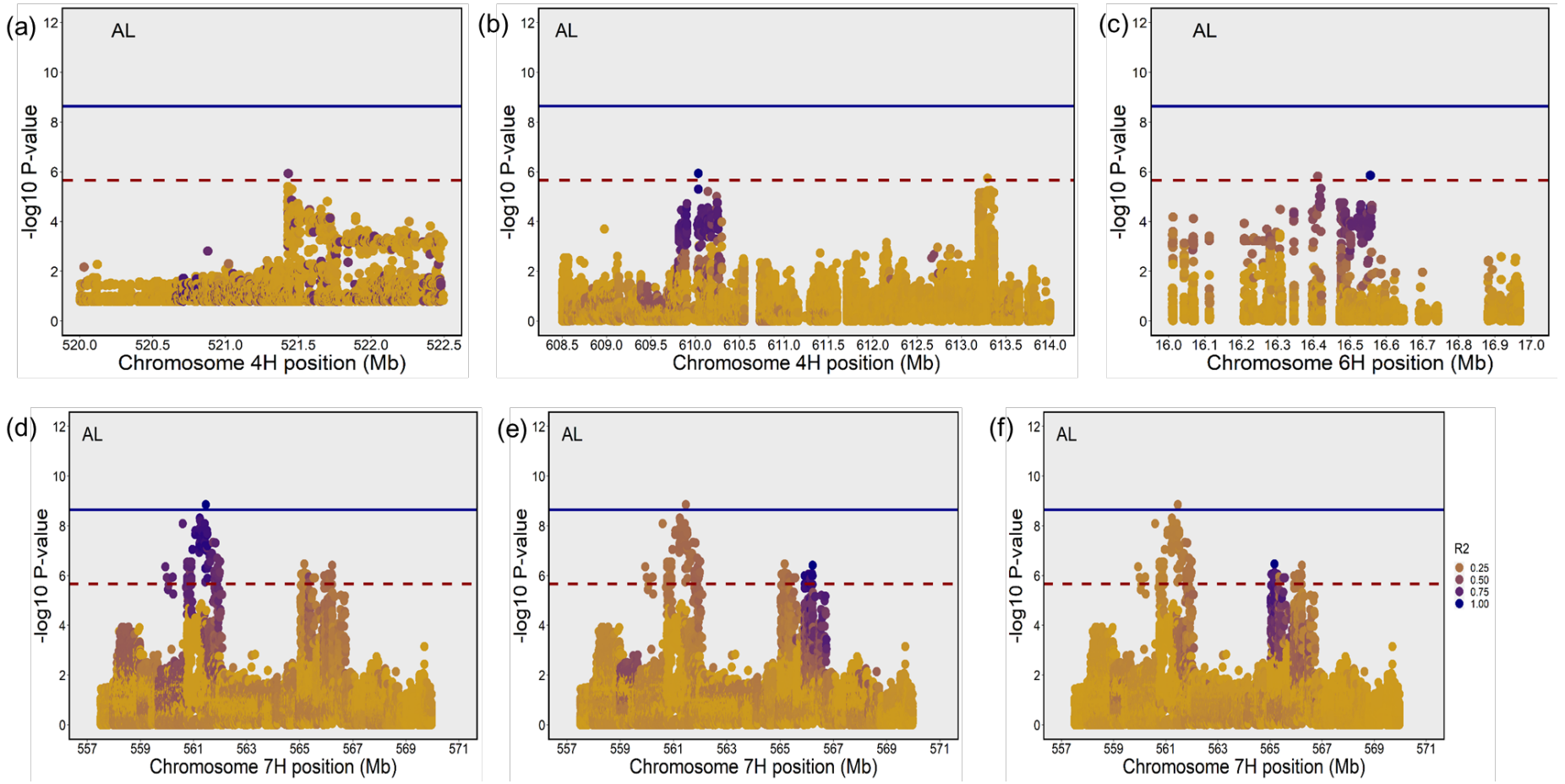


Figure S4.6 Linkage disequilibrium plots for awn length on (a,b) chr4H, (c) chr6H, and (d–f) chr7H. Blue line is the Bonferroni correction at alpha level of 0.05 (BC = 8.64) and red dashed line is false discovery rate at alpha level of 0.05 (FDR = 5.66).

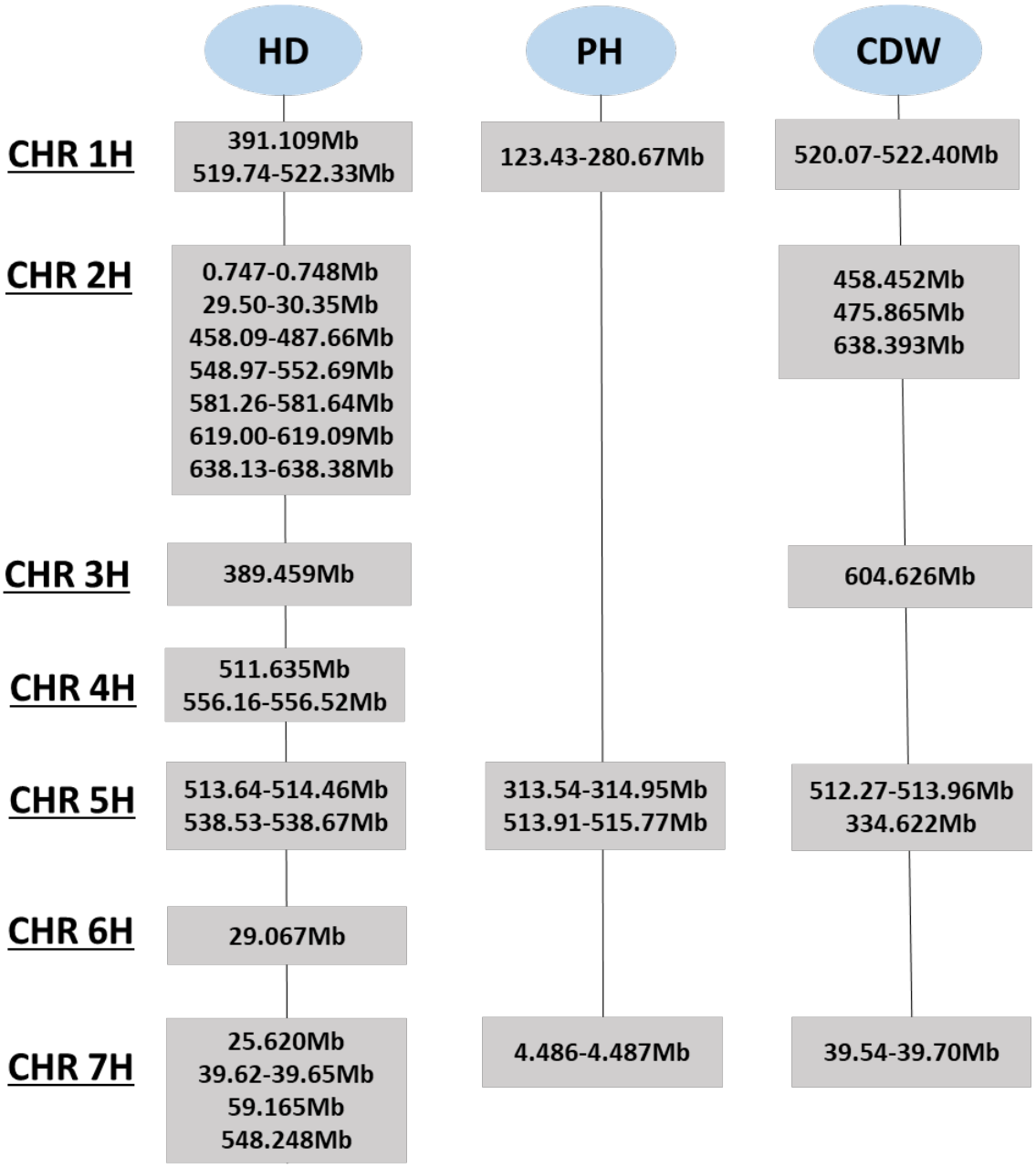


Figure S4.7 Shoot traits' unique QTL distribution on seven barley chromosomes. HD = heading date, PH = plant height, and CDW = culm dry weight

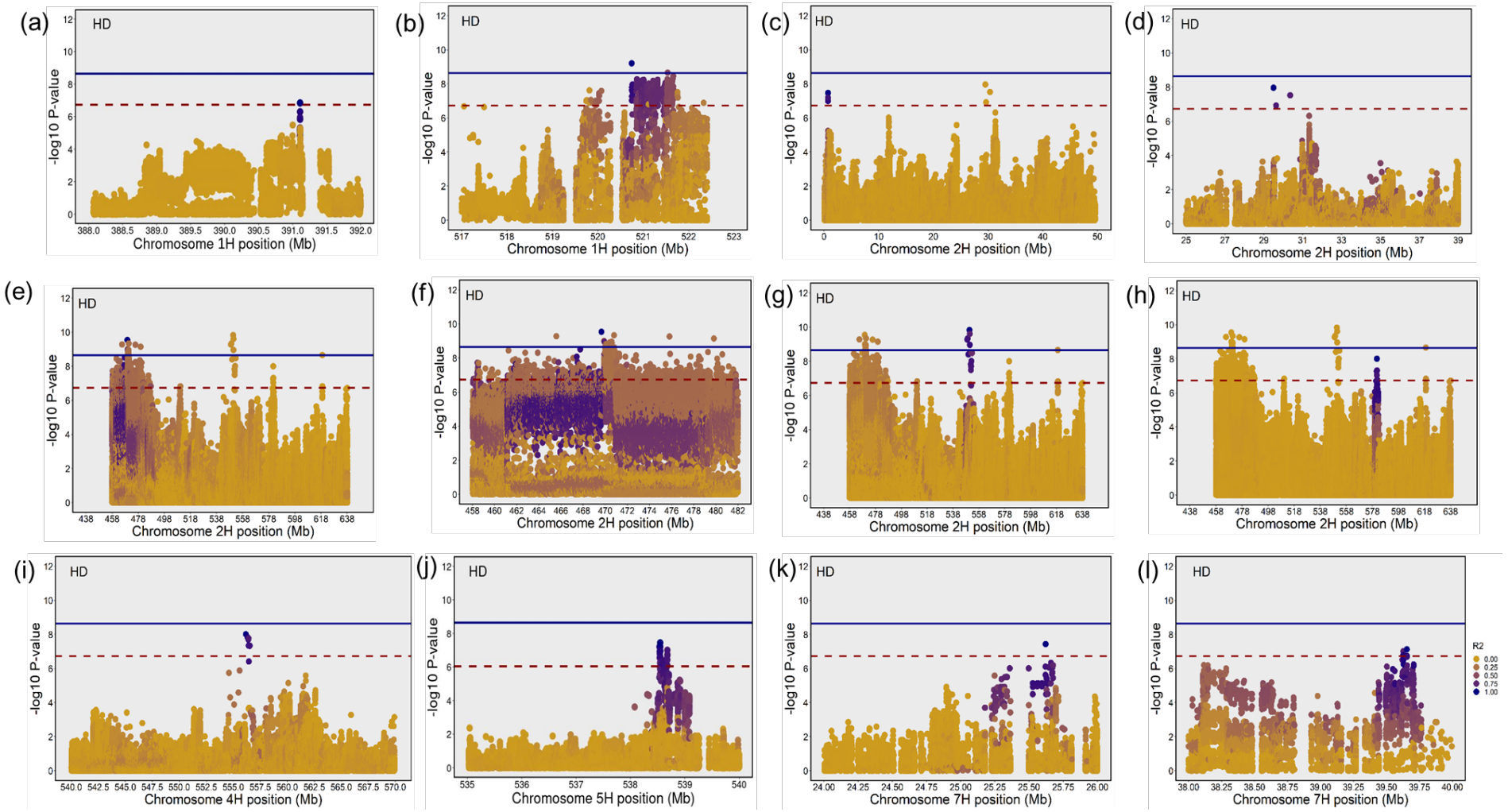


Figure S4.8 Linkage disequilibrium plots for heading date on (a,b) chr1H, (c–h) chr2H, (i) chr4H, (j) chr5H, and (k,l) on chr7H. Blue line is the Bonferroni correction at alpha level of 0.05 (BC = 8.64) and red dashed line is false discovery rate at alpha level of 0.05 (FDR = 6.73).

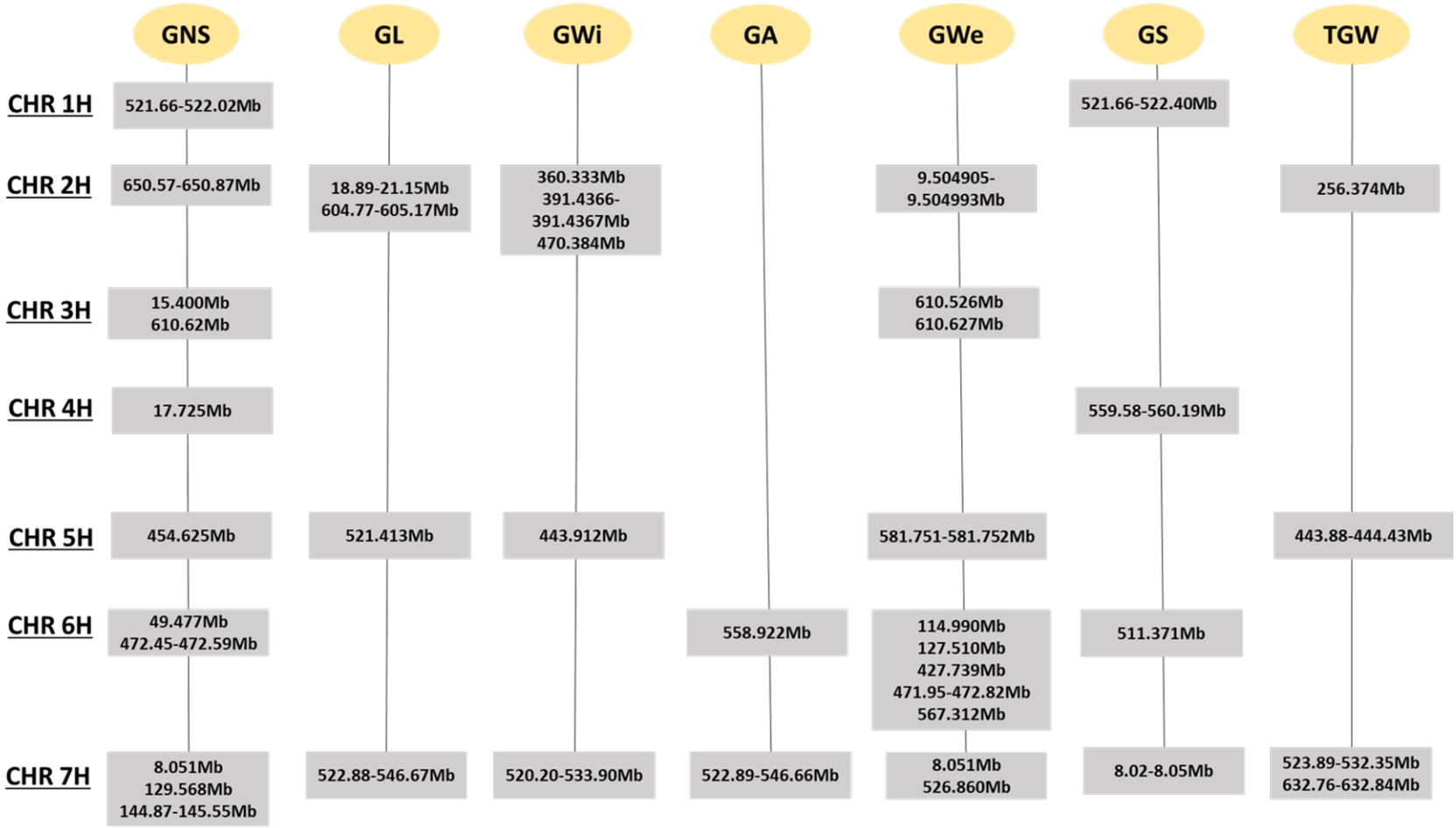


Figure S4.9 Grain traits' unique QTL distribution on seven barley chromosomes. GNS = grain number per spike, GL = grain length, GWi = grain width, GA = grain area, GWe = grain weight, GS = grain set, and TGW = thousand grain weight.

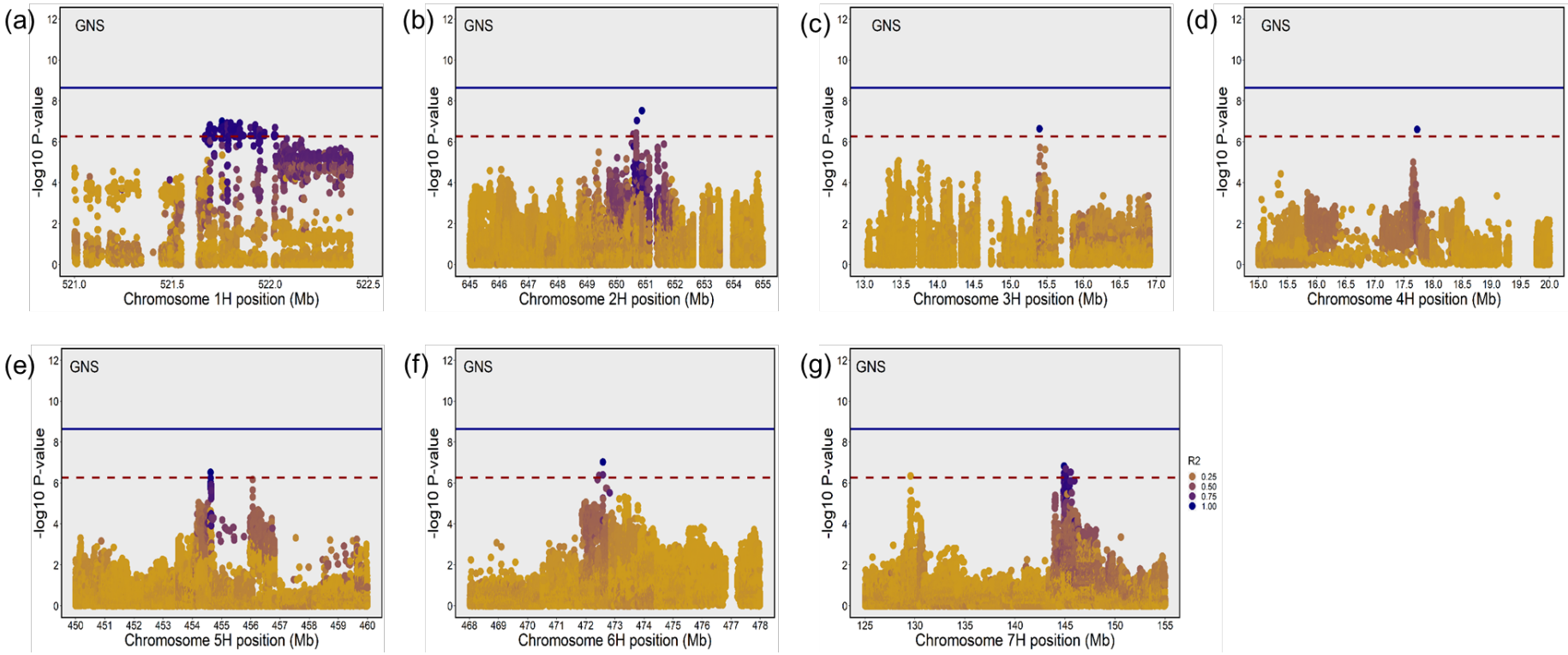


Figure S4.10 Linkage disequilibrium plots for grain number spike on (a) chr1H, (b) chr2H, (c) chr3H, (d) chr4H, (e) chr5H, (f) chr6H, and (g) chr7H. Blue line is the Bonferroni correction at alpha level of 0.05 (BC = 8.64) and red dashed line is false discovery rate at alpha level of 0.05 (FDR = 6.27).

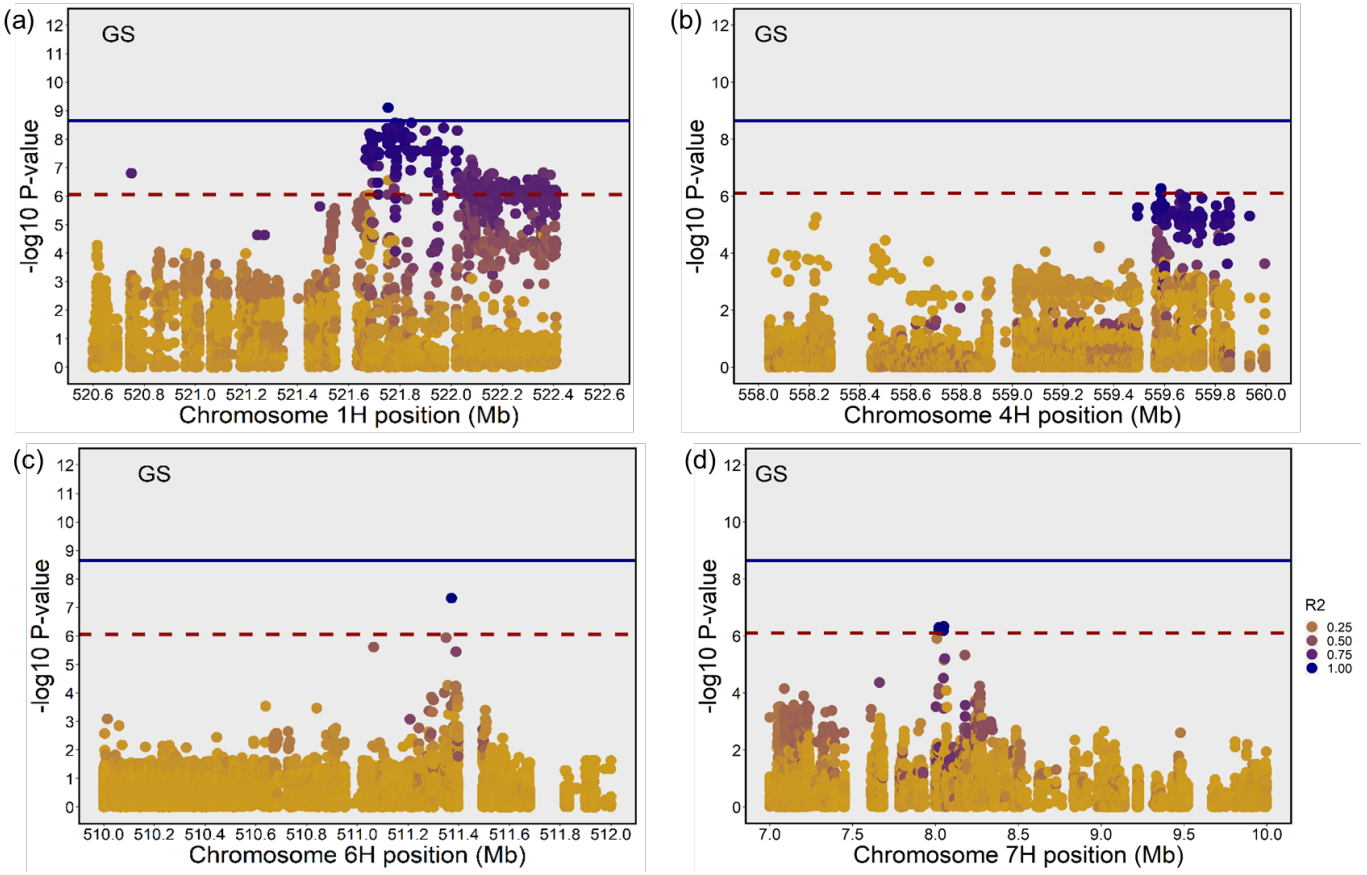


Figure S4.11 Linkage disequilibrium plots for grain set on (a) chr1H, (b) chr4H, (c) chr6H, and (d) chr7H. Blue line is the Bonferroni correction at alpha level of 0.05 (BC = 8.64) and red dashed line is false discovery rate at alpha level of 0.05 (FDR = 6.11).

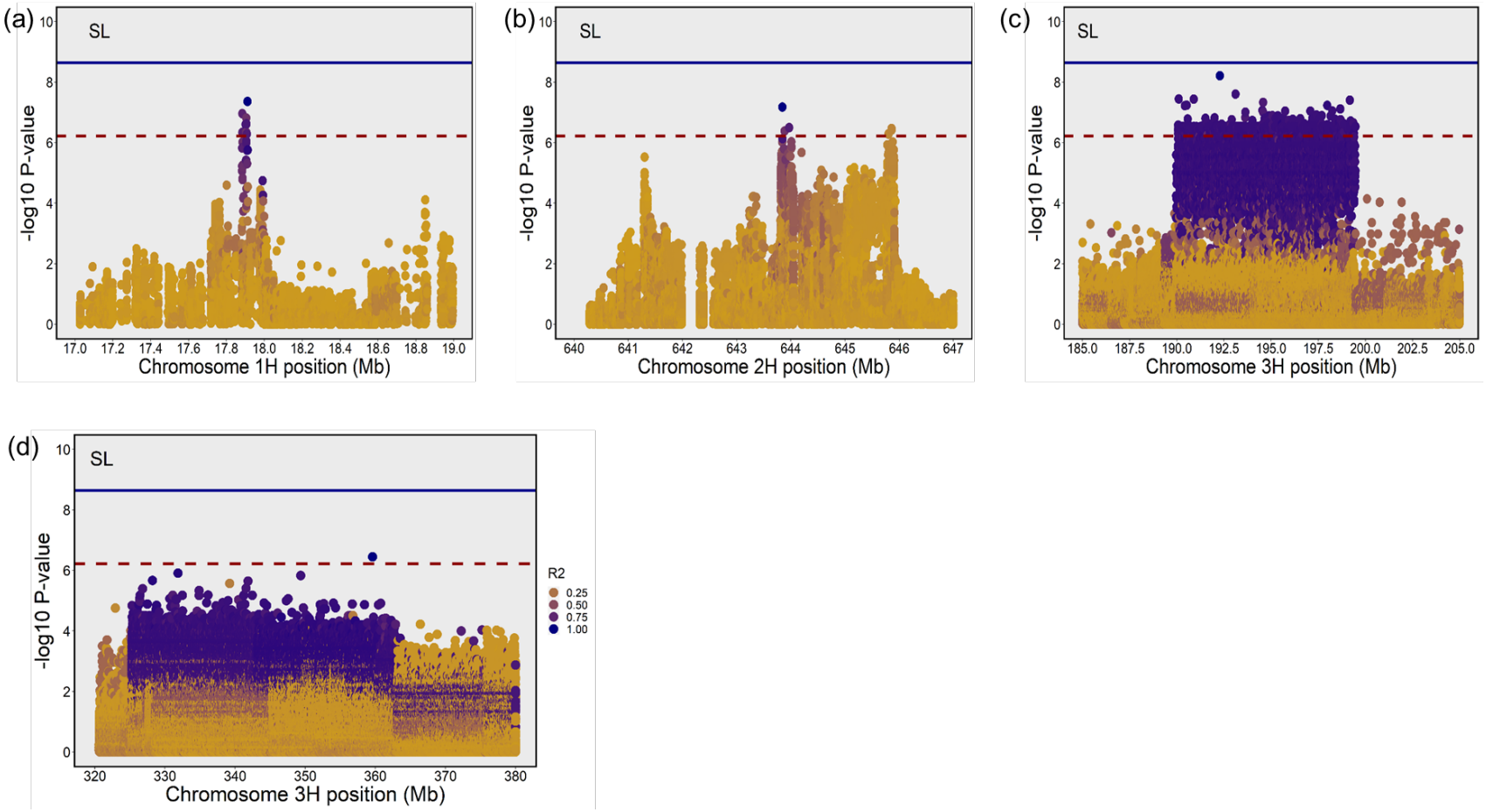


Figure S4.12 Linkage disequilibrium plots for spike length on (a) chr1HS, (b) chr2HL, and (c,d) on chr3H. Blue line is the Bonferroni correction at alpha level of 0.05 (BC = 8.64) and red dashed line is false discovery rate at alpha level of 0.05 (FDR = 6.22).

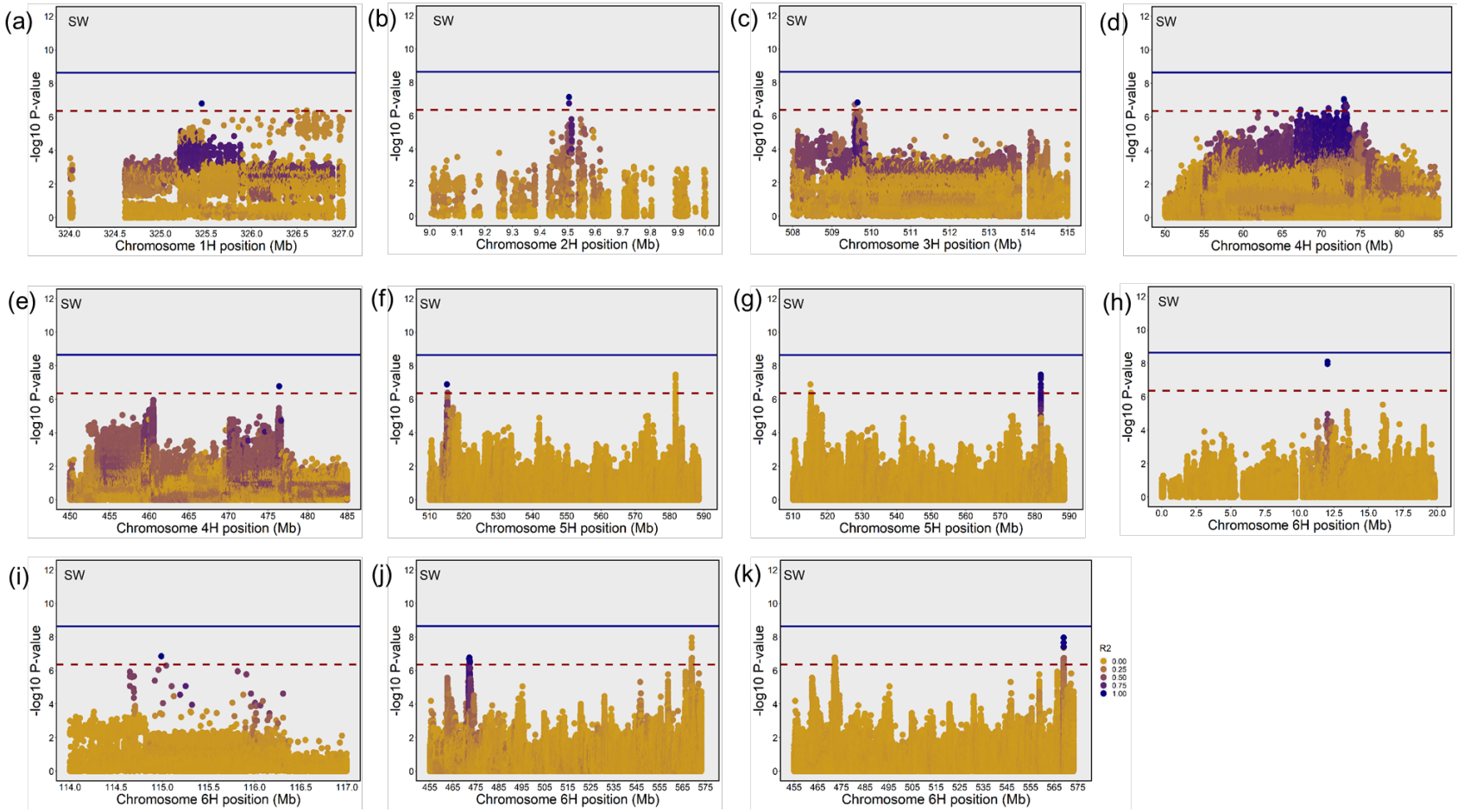


Figure S4.13 Linkage disequilibrium plots for spike weight on(a) ch1H, (b) chr2H, (c) chr3H, (d,e) chr4H, (f,g) chr5H and (h–k) on chr6H. Blue line is the Bonferroni correction at alpha level of 0.05 (BC = 8.64) and red dashed line is false discovery rate at alpha level of 0.05 (FDR = 6.36).

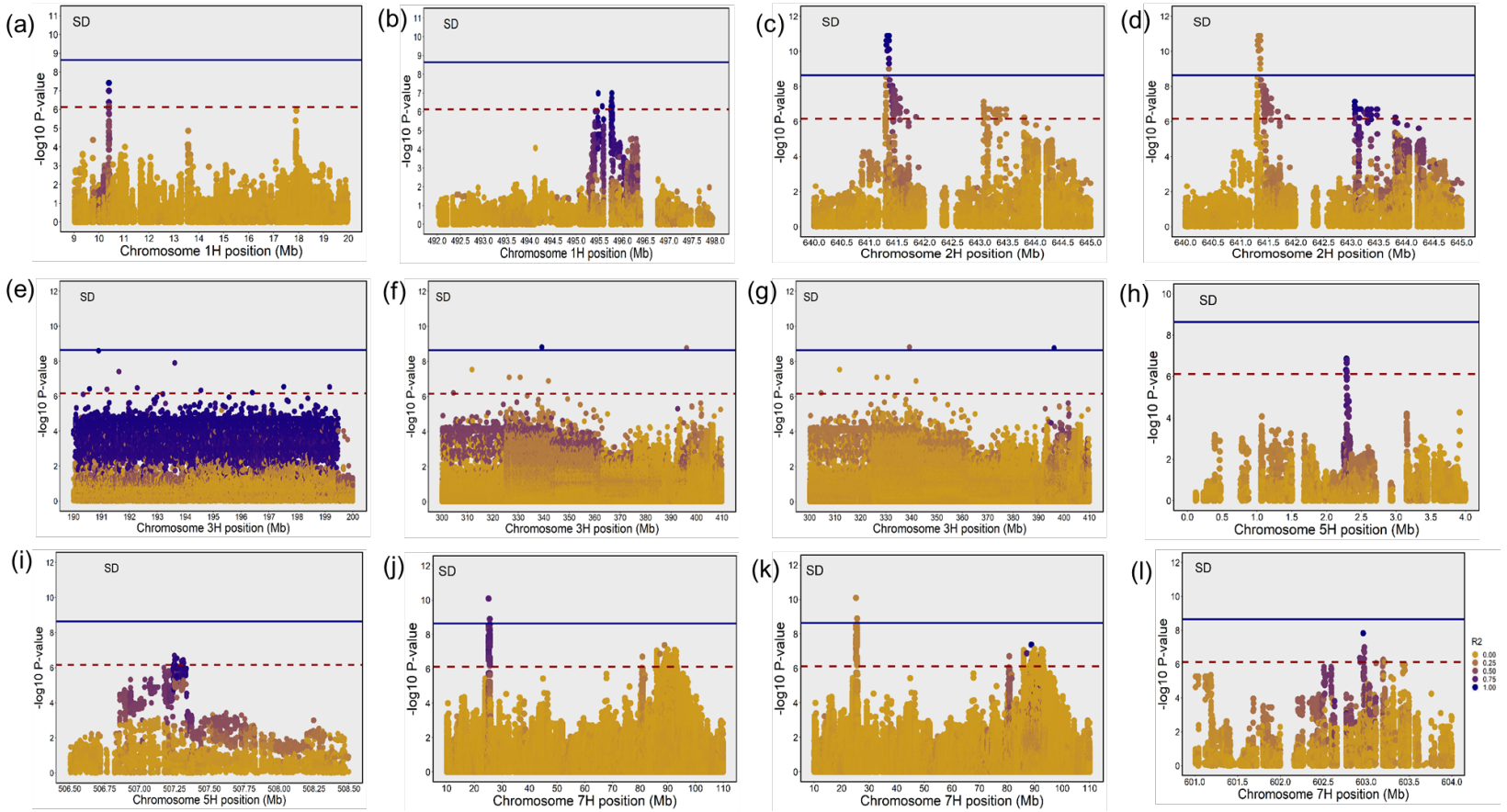


Figure S4.14 Linkage disequilibrium plots for spike density on (a,b) chr1H, (c,d) chr2H, (e-g) chr3H, (h,i) chr5H, and (j–l) on chr7H. Blue line is the Bonferroni correction at alpha level of 0.05 (BC = 8.64) and red dashed line is false discovery rate at alpha level of 0.05 (FDR = 6.12).

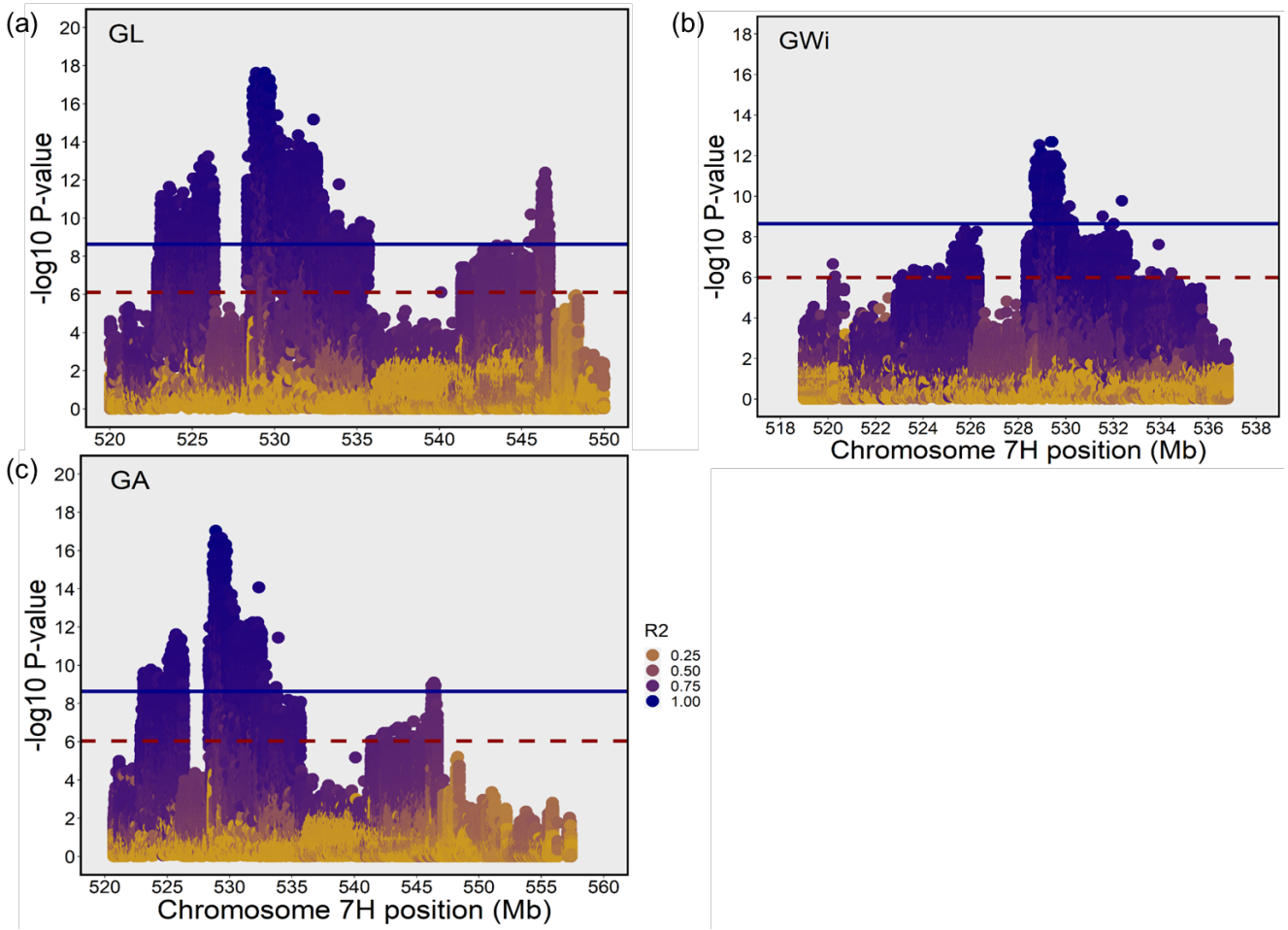


Figure S4.15 Linkage disequilibrium (LD) plots for grain morphometric traits. LD plot for (a) grain length on chr7H (FDR = 6.10), (b) grain width on chr7H (FDR = 6.02), and (c) grain area on chr7H (FDR = 6.05). Blue line is the Bonferroni correction at alpha level of 0.05 (BC = 8.64) and red dashed line is false discovery rate at alpha level of 0.05.

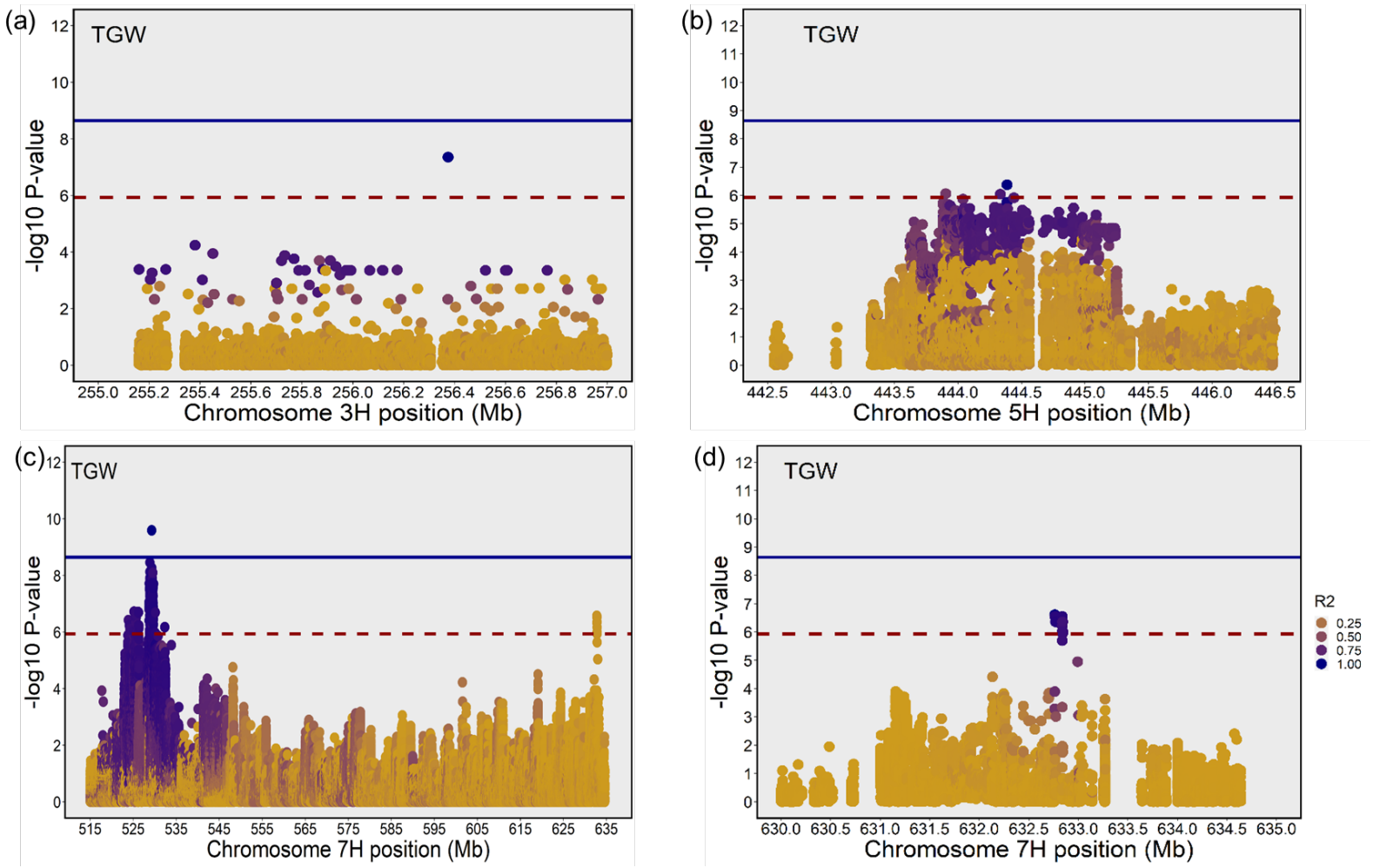


Figure S4.16 Linkage disequilibrium plots for thousand grain weight on (a) chr3H, (b) on chr5H, and (c,d) on chr7H. Blue line is the Bonferroni correction at alpha level of 0.05 (BC = 8.64) and red dashed line is false discovery rate at alpha level of 0.05 (FDR = 5.94).

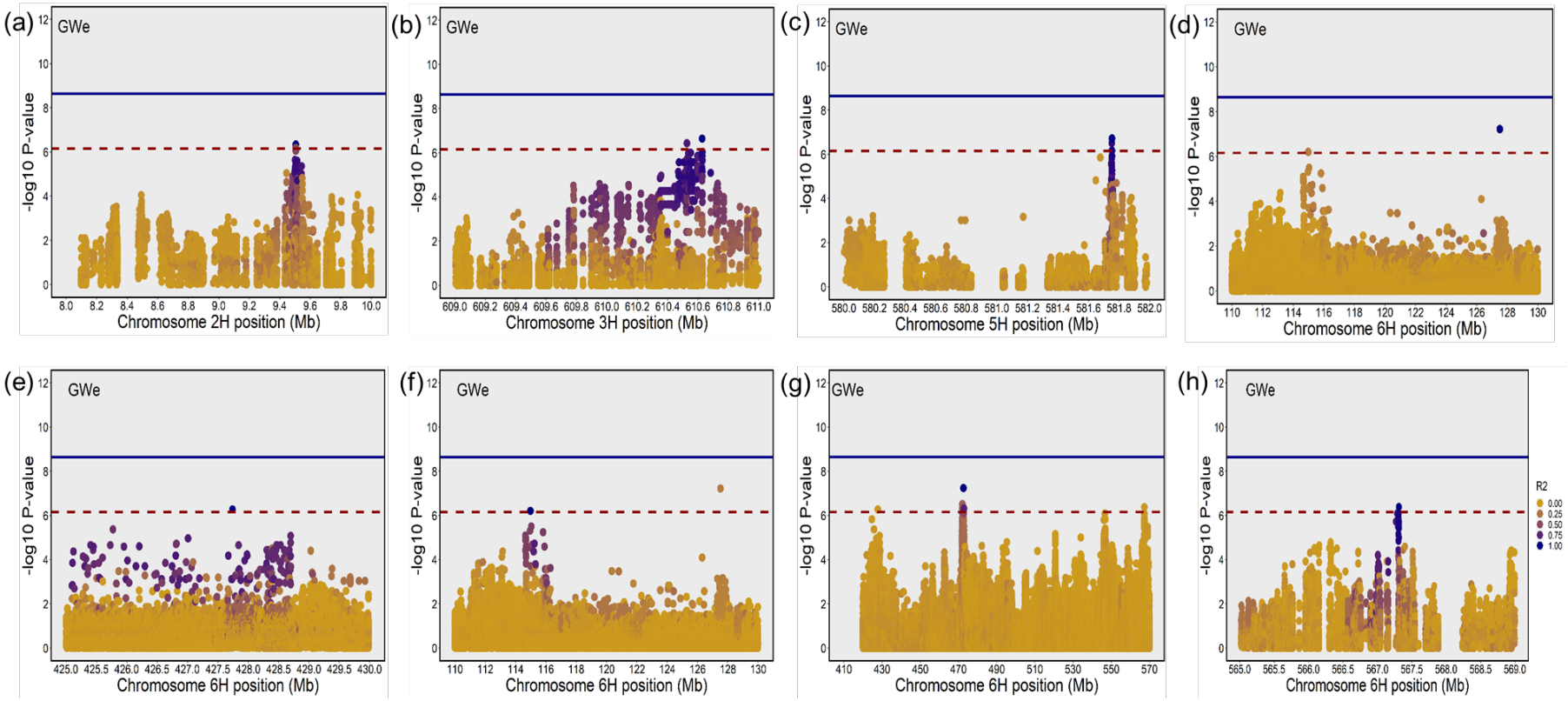


Figure S4.17 Linkage disequilibrium plots for grain weight on (a) chr2H, (b) chr3H, (c) chr5H, and (d-h) chr6H. Blue line is the Bonferroni correction at alpha level of 0.05 (BC = 8.64) and red dashed line is false discovery rate at alpha level of 0.05 (FDR = 6.16).

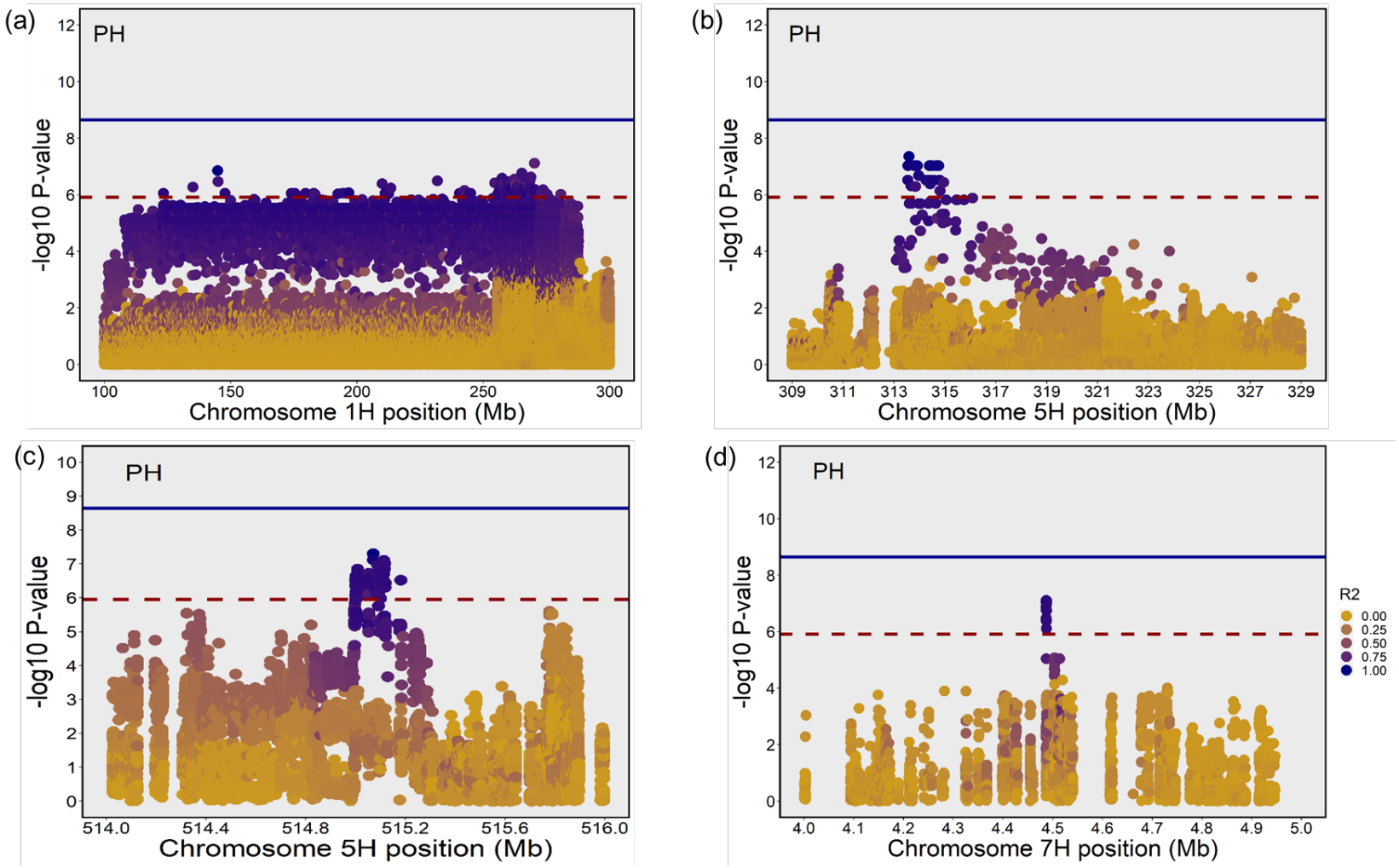


Figure S4.18 Linkage disequilibrium plots for plant height on (a) chr1H, (b,c) chr5H, and (d) chr7H. Blue line is the Bonferroni correction at alpha level of 0.05 (BC = 8.64) and red dashed line is false discovery rate at alpha level of 0.05 (FDR = 5.91).

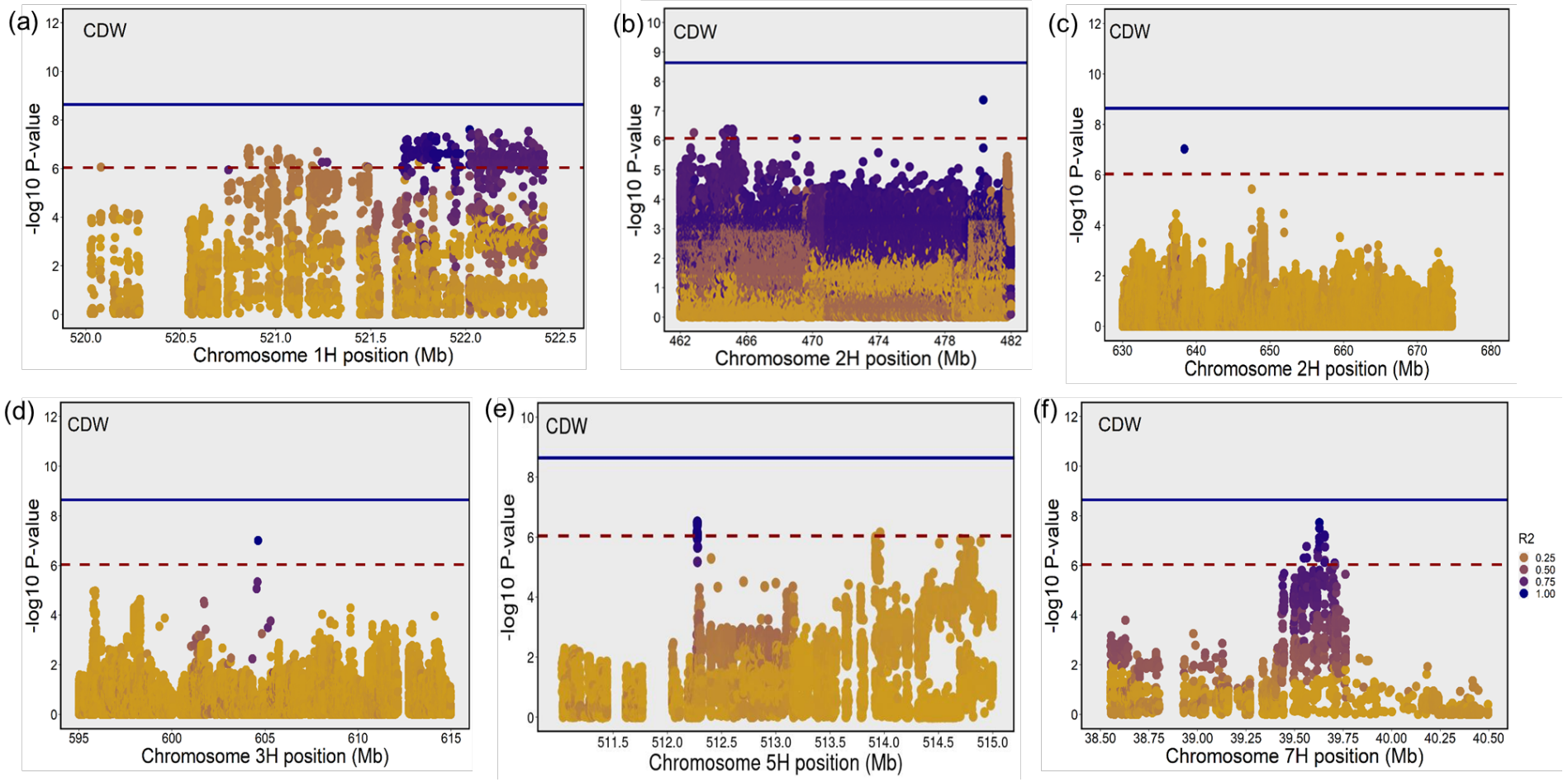


Figure S4.19 Linkage disequilibrium plots for culm dry weight on (a) chr1H, (b,c) on chr2H, (d) on chr3H, (e) on chr5H, and (f) on chr7H. Blue line is the Bonferroni correction at alpha level of 0.05 (BC =8.64) and red dashed line is false discovery rate at alpha level of 0.05 (FDR = 6.04).

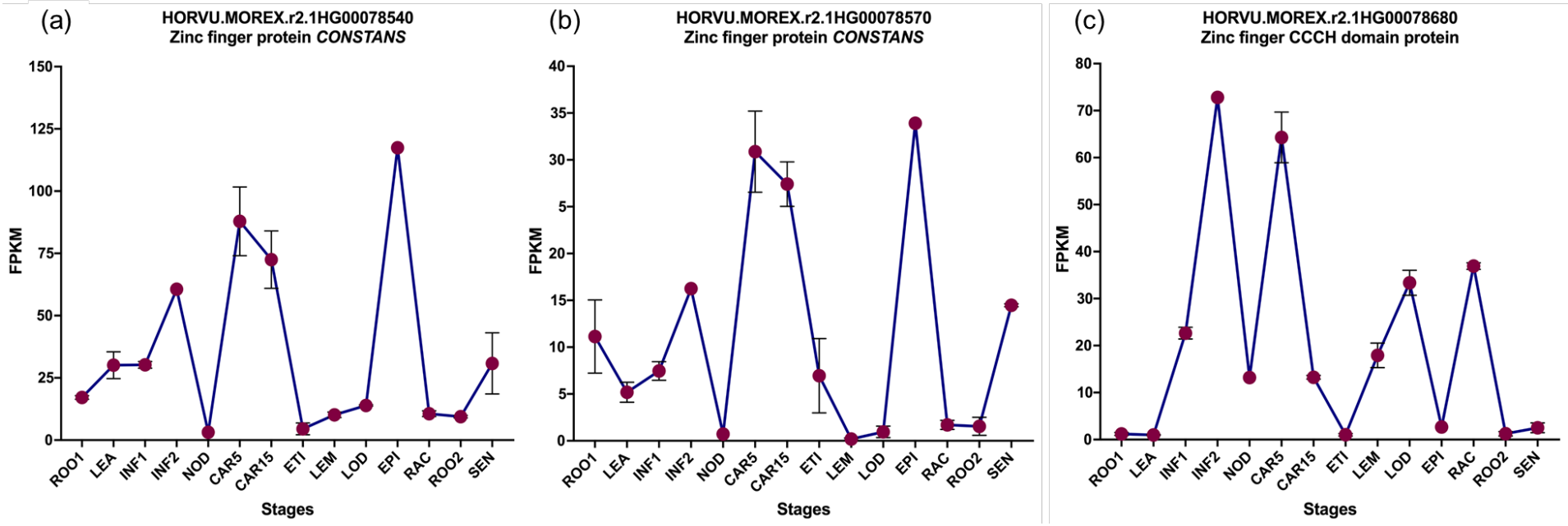


Figure S4.20 Expression pattern of (a) *HORVU.MOREX.r2.1HG0078540* annotated as zinc finger protein CONSTANS, (b) *HORVU.MOREX.r2.1HG0078570* annotated as zinc finger protein CONSTANS, and (c) *HORVU.MOREX.r2.1HG0078680* annotated as zinc finger CCCH domain protein in different tissues. ROO1 = roots from seedlings (10cm shoot stage), LEA = shoots from seedlings (10cm shoot stage), INF1 = young inflorescence (5mm), INF2 = developing inflorescence (1-1.5cm), NOD = developing tillers, 3rd internode (42 DAP), CAR5 = developing grain (5 DAP), CAR15 = developing grain (15 DAP), ETI = etiolated seedling, dark condition (10 DAP), LEM = inflorescences, lemma (42 DAP), LOD = inflorescences, lodicule (42DAP), EPI = epidermal strips (28 DAP), RAC = inflorescence, rachis (35 DAP), ROO2 = roots (28 DAP), SEN = senescing leaves (58 DAP) and FPKM = fragments per kilobase of exon per million mapped fragments.

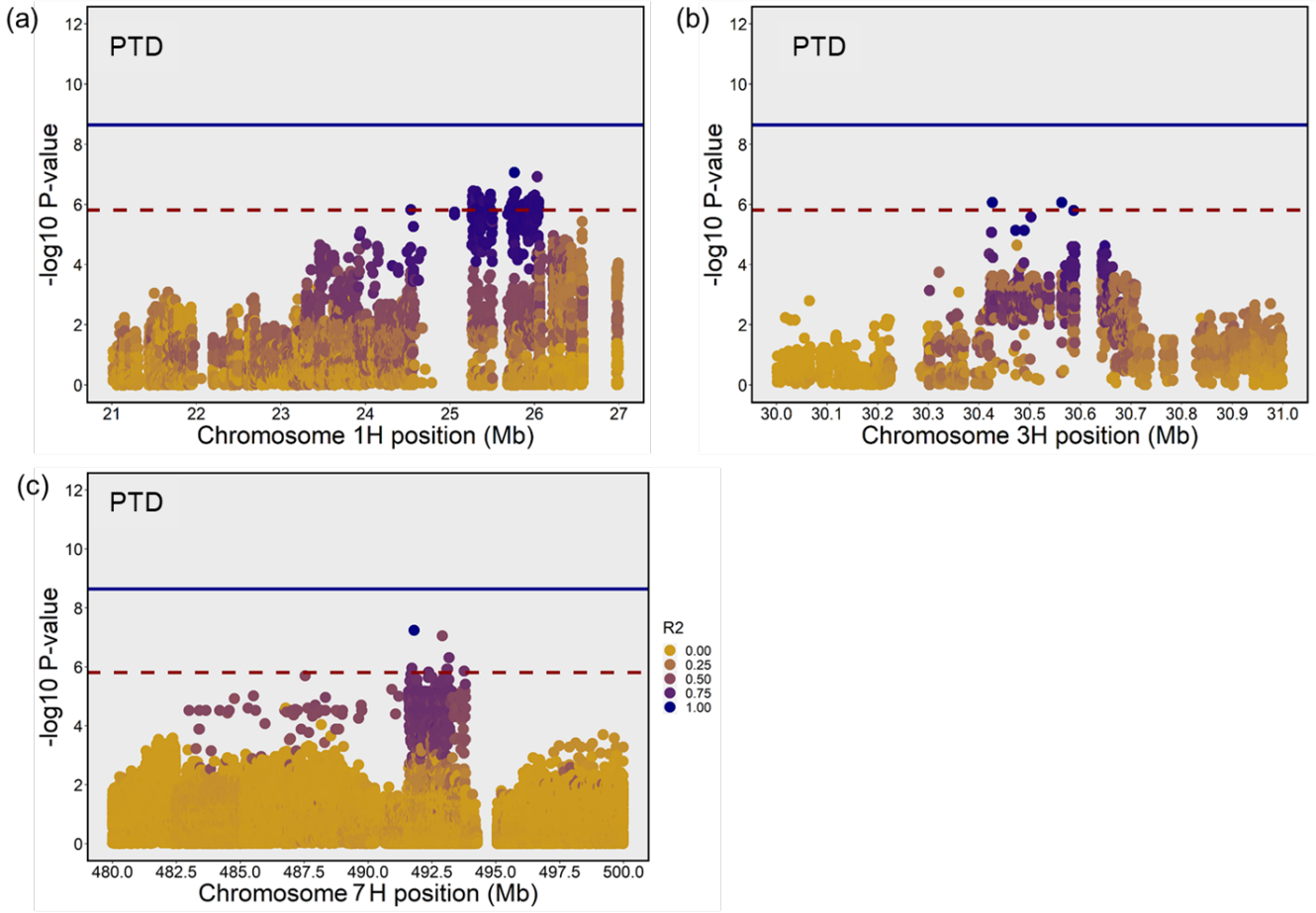


Figure S4.21 Linkage disequilibrium plots for spike pre-anthesis tip degeneration (PTD) on (a) chr 1H, (b) chr3HL, and (c) chr7H. Blue line is the Bonferroni correction at alpha level of 0.05 (BC = 8.64) and red dashed line is false discovery rate at alpha level of 0.05 (FDR = 5.81).

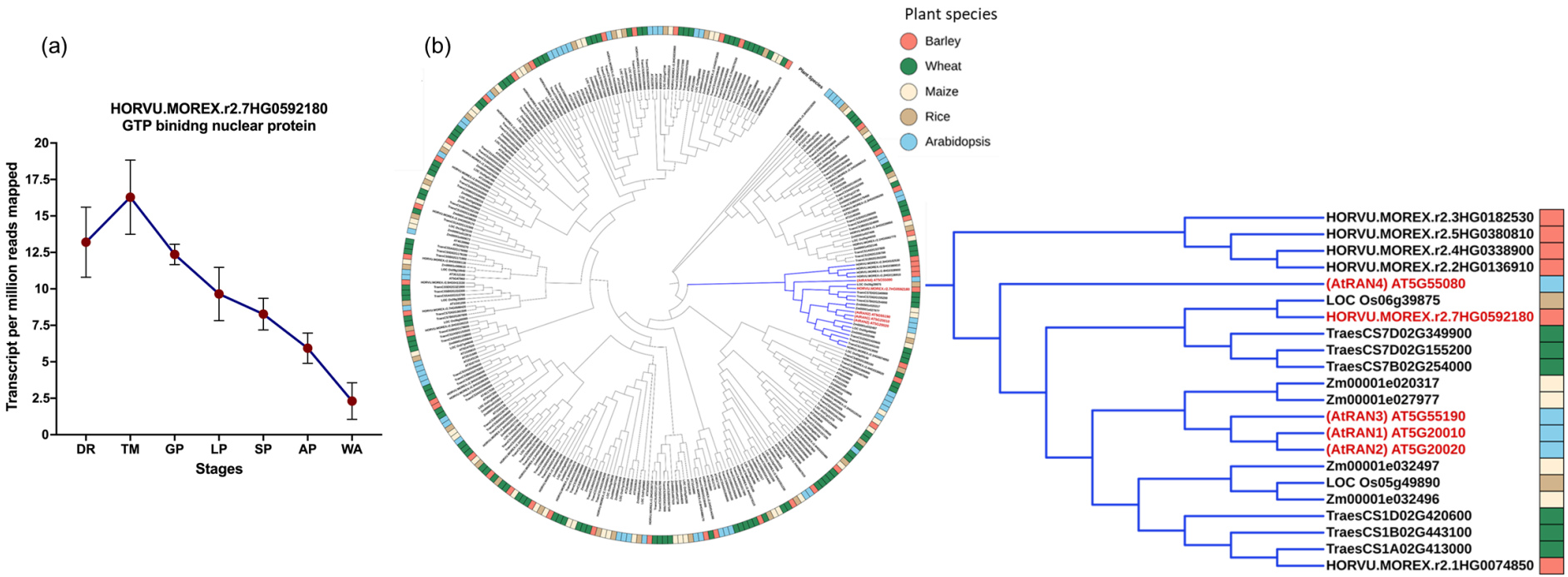


Figure S4.22 Gene expression and phylogenetic analysis of *HvRAN2* (*HORVU.MOREX.r2.7HG0592180*). (a) expression pattern of *HvRAN2* in inflorescence meristem, and (b) phylogenetic analysis of *HvRAN2*. DR = double ridge, TM = triple mound, GP = Glume primordium, LP = lemma primordium, SP = stamen primordium, AP = awn primordium, WA = white anther.

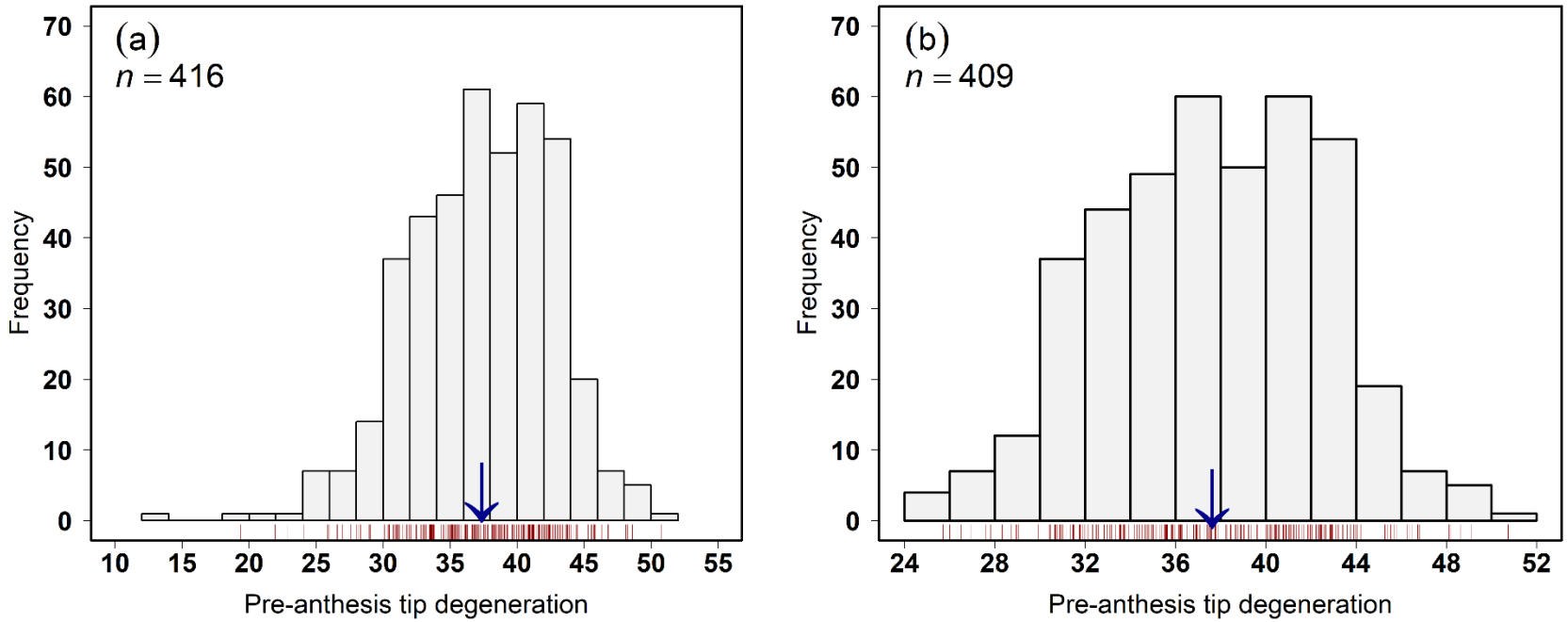


Figure S4.23 Frequency distribution for pre-anthesis tip degeneration (PTD) (a) in a panel of 416 six-rowed barley accessions and (b) after outlier removal keeping 409 accessions. x-axis depicts the percentage of PTD and y-axis represents the number of accessions. The downwards blue arrow is the mean value.

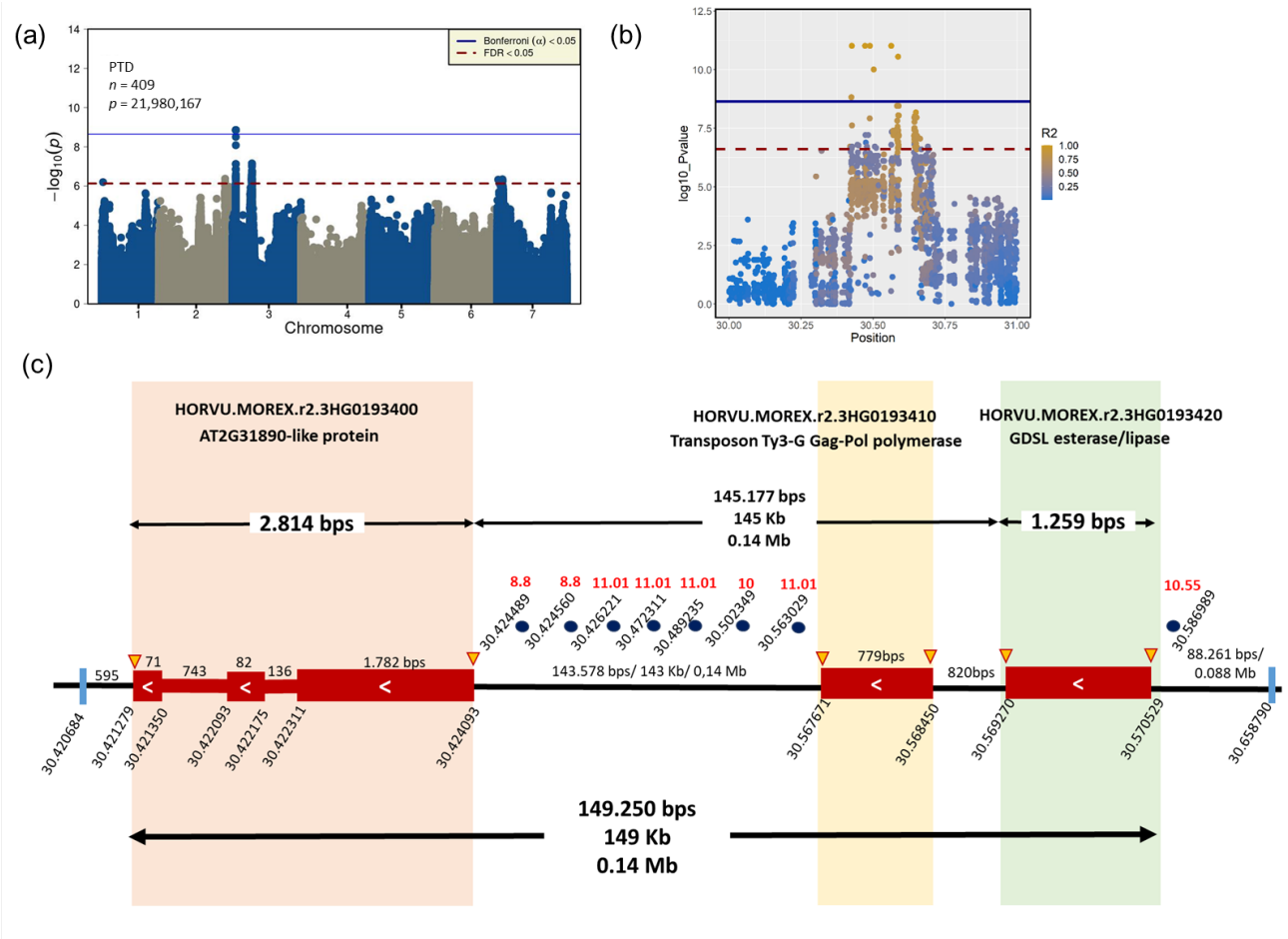


Figure S4.24 GWAS results for spike pre-anthesis tip degeneration (PTD) in 409 six-rowed spring barley accessions. (a) manhattan plot for 409 accessions, (b) linkage disequilibrium plot for chr3H, and (c) detailed description of the three genes present within the interval of significant SNPs on chr3HS. The blue vertical lines mark the start and end of the QTL region. The structure of each gene is described with red boxes; blue circle are the position of the significant SNPs above Bonferroni correction while red numbers are the corresponding $-\log_{10}$ *P*-values of these significant SNPs

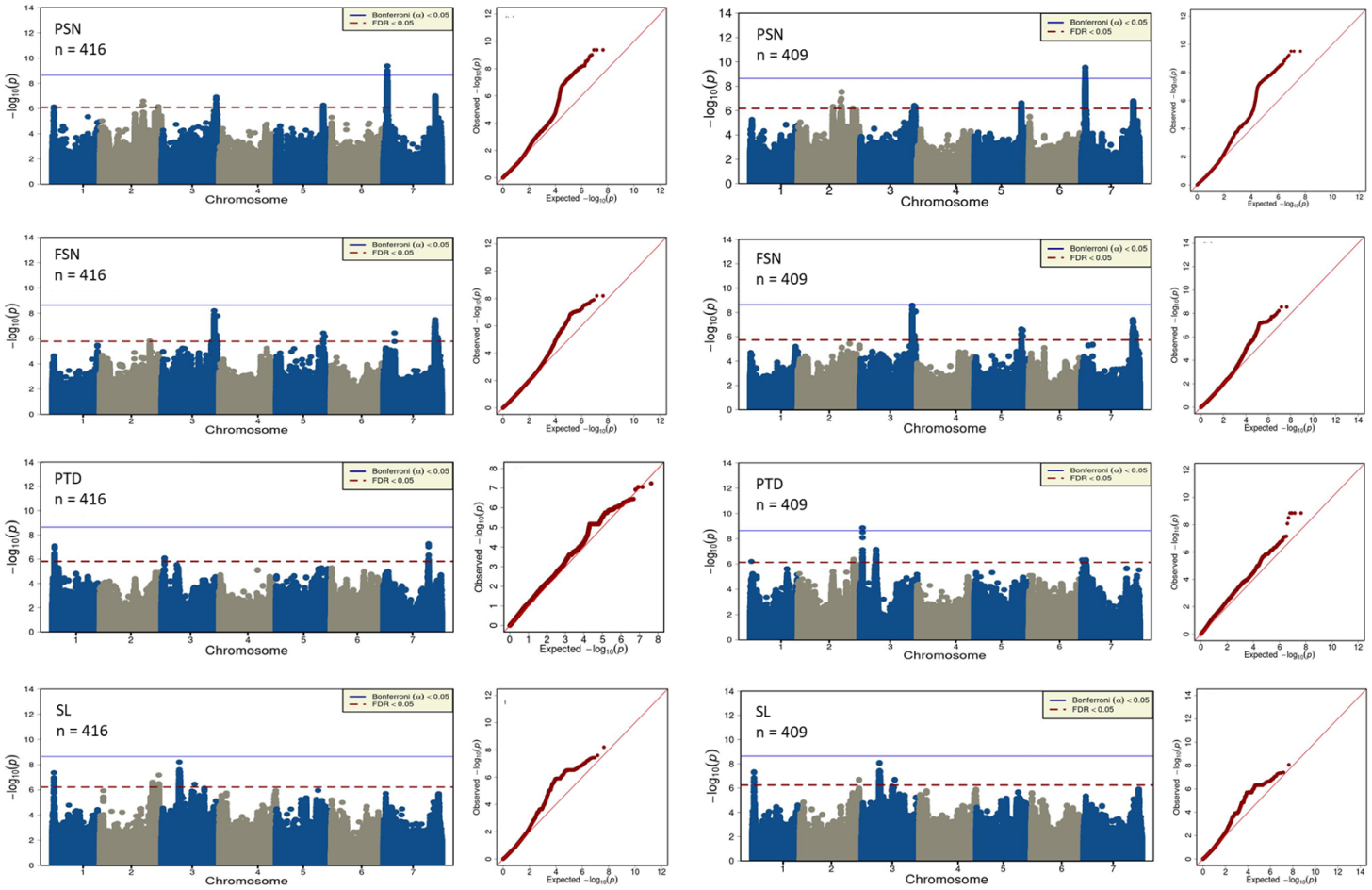


Figure S4.25 Comparison of GWAS results for PSN, FSN, PTD and SL between panel with 416 accessions and panel with 409 accessions. The x-axis shows seven barley chromosomes and y-axis shows the marker significances as $-\log_{10}$ (P). A Bonferroni α level of 0.05 and false discovery rate (FDR) level of 0.05 was used to correct for multiple testing and identify significant markers. PSN = potential spikelet number, FSN = final spikelet number, PTD = pre-anthesis tip degeneration, and SL = spike length

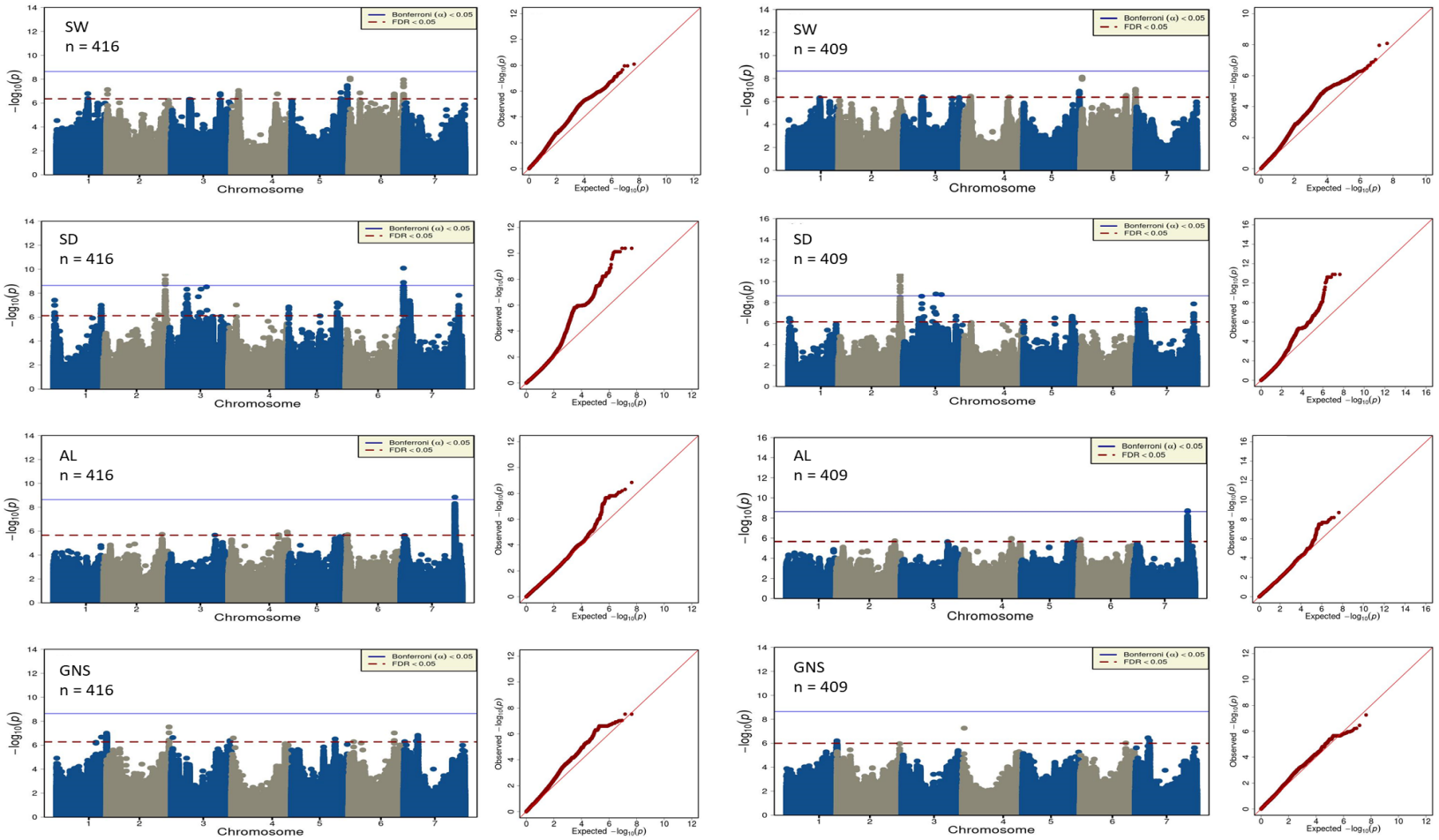


Figure S4.26 Comparison of GWAS results for SW, SD, AL and GNS between panel with 416 accessions and panel with 409 accessions. The x-axis shows seven barley chromosomes and y-axis shows the marker significances as $-\log_{10}$ (P). A Bonferroni α level of 0.05 and false discovery rate (FDR) level of 0.05 was used to correct for multiple testing and identify significant markers. SW = spike weight, SD = spike density, AL = awn length, and GNS = grain number per spike

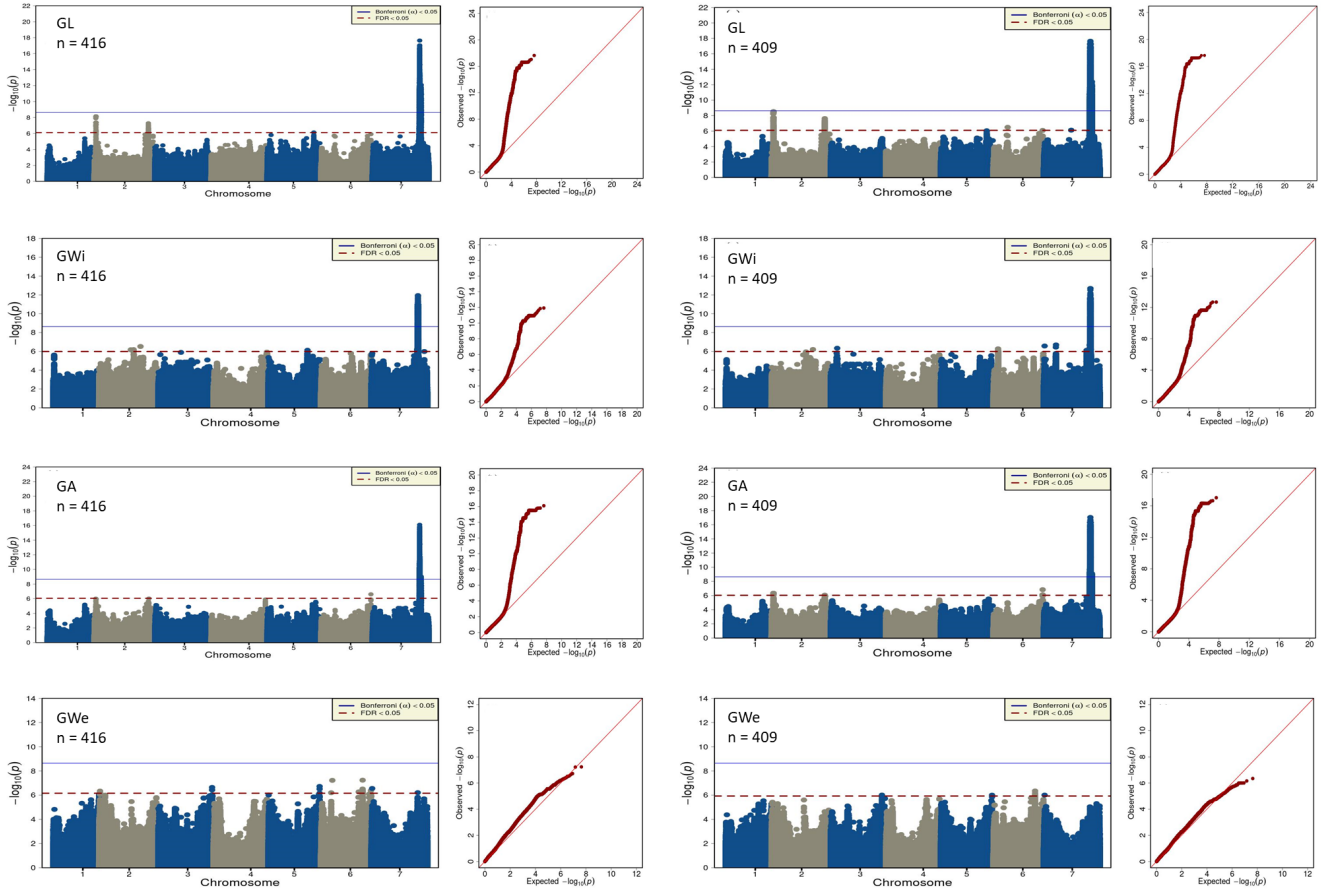


Figure S4.27 Comparison of GWAS results for GL, GWi, GA and GWe between panel with 416 accessions and panel with 409 accessions. The x-axis shows seven barley chromosomes and y-axis shows the marker significances as $-\log_{10}$ (P). A Bonferroni α level of 0.05 and false discovery rate (FDR) level of 0.05 was used to correct for multiple testing and identify significant markers. GL = grain length, GWi = grain width, GA = grain area, and GWe = grain weight

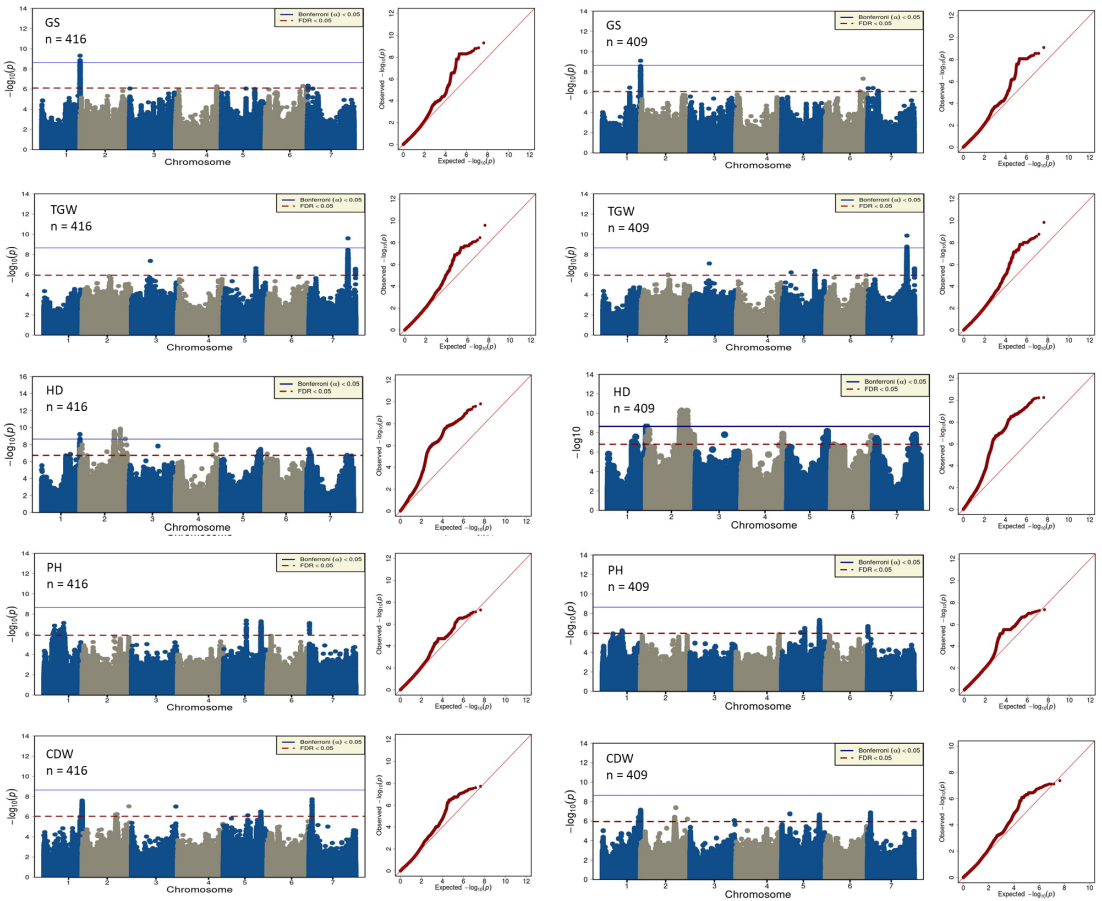


Figure S4.28 Comparison of GWAS results for GS, TGW, HD, PH and CDW between panel with 416 accessions and panel with 409 accessions. The x-axis shows seven barley chromosomes and y-axis shows the marker significances as $-\log_{10}$ (P). A Bonferroni α level of 0.05 and false discovery rate (FDR) level of 0.05 was used to correct for multiple testing and identify significant markers. GS = grain set, TGW = thousand grain weight, HD = heading date, PH = plant height and CDW = culm dry weight

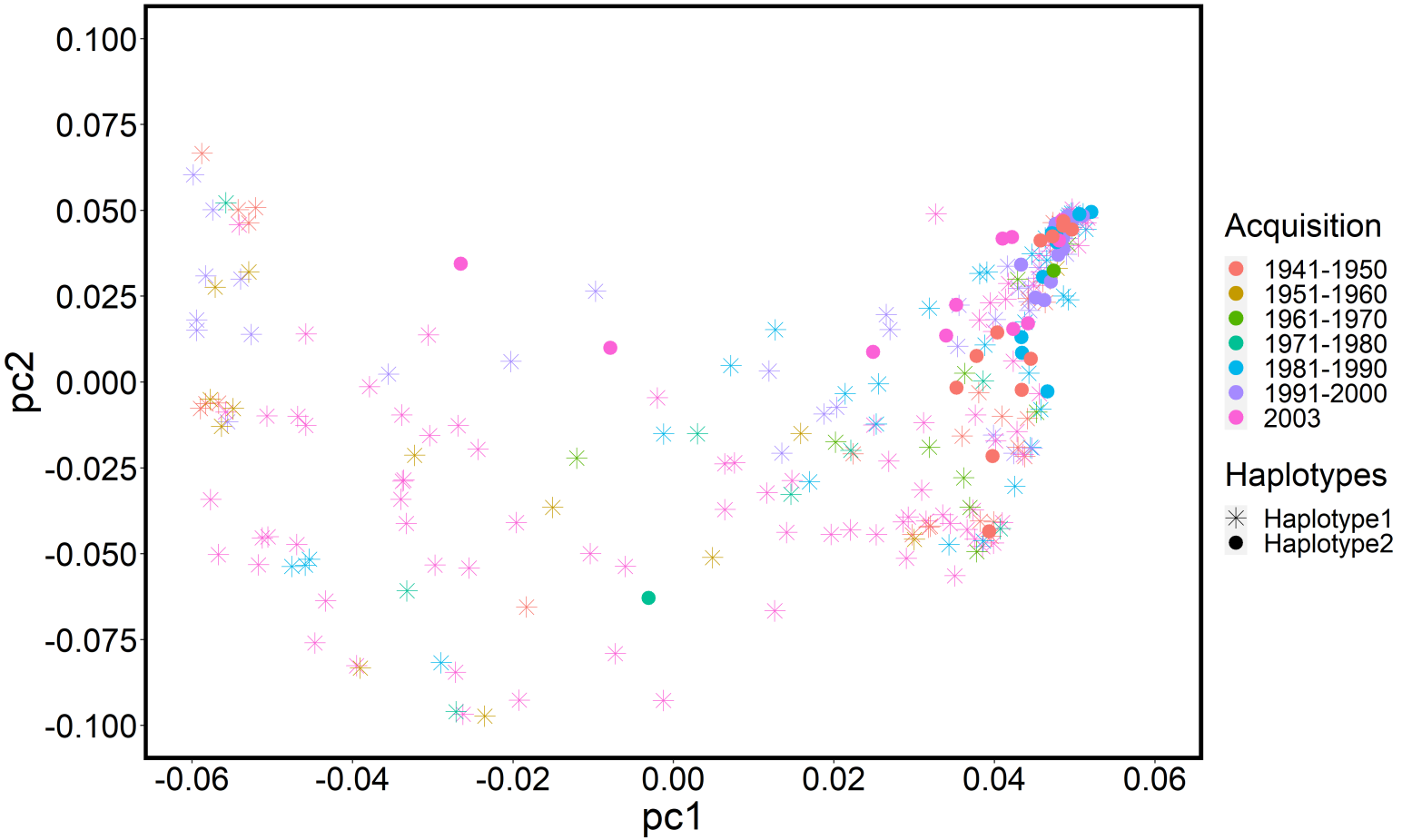


Figure S4.29 PCA plot distinguishing the major haplotypes for PTD and accessions based on the acquisition year

EIDESSTATTLICHE ERKLÄRUNG

DECLARATION UNDER OATH

Ich erkläre an Eides statt, dass ich die Arbeit selbstständig und ohne fremde Hilfe verfasst, keine anderen als die von mir angegebenen Quellen und Hilfsmittel benutzt und die den benutzten Werken wörtlich oder inhaltlich entnommenen Stellen als solche kenntlich gemacht habe.

I declare under penalty of perjury that this thesis is my own work entirely and has been written without any help from other people. I used only the sources mentioned and included all the citations correctly both in word or content.

Datum / Date	Unterschrift des Antragstellers / Signature of the applicant

ERKLÄRUNG ÜBER BESTEHENDE VORSTRAFEN UND ANHÄNGIGE ERMITTLUNGSVERFAHREN

DECLARATION CONCERNING CRIMINAL RECORD AND PENDINGINVESTIGATIONS

Hiermit erkläre ich, dass ich weder vorbestraft bin noch dass gegen mich Ermittlungsverfahren anhängig sind.

I hereby declare that I have no criminal record and that no preliminary investigationsare pending against me.

Datum / Date	Unterschrift des Antragstellers / Signature of the applicant

Roop Kamal Muqaddasi

PREVIOUS EMPLOYMENT

LEIBNIZ INSTITUTE OF PLANT GENETICS AND CROP PLANT RESEARCH (IPK), GATERSLEBEN, GERMANY
BREEDING RESEARCH DEPARTMENT

Post-doctoral Scientist (January 2023–September 2024)
Ph.D. Researcher (November 2017–September 2022)

EDUCATION

MARTIN-LUTHER-UNIVERSITY HALLE-WITTENBERG, HALLE (SAALE), GERMANY
INSTITUTE OF AGRICULTURAL AND NUTRITIONAL SCIENCES

Ph.D. Agricultural Sciences (November 2017–April 2025)

PUNJAB AGRICULTURAL UNIVERSITY, LUDHIANA, PUNJAB, INDIA
DEPARTMENT OF PLANT BREEDING AND GENETICS, COLLEGE OF AGRICULTURE

M.Sc. Plant Breeding and Genetics (July 2015–November 2017)
B.Sc. (Hons.) Agriculture (July 2011–May 2015)

PUBLICATIONS

Muqaddasi RK, Muqaddasi QH, Trautewig C, Huang Y, Shanmugaraj N, Melzer M, Tschiersch H, Chmielewska B, Szurman-Zubrzycka M, Szarejko I, Himmelbach A, Jayakodi M, Stein N, Mascher M, Schnurbusch T (2023) Insights into the genetic basis of apical spikelet primordia degeneration and related traits in barley. Submitted.

Muqaddasi QH, **Muqaddasi RK,** Ebmeyer E, Korzun V, Argillier O, Mirdita V, Reif JC, Ganal MW, Röder MS (2023) Genetic control and prospects of predictive breeding for European winter wheat's Zeleny sedimentation values and Hagberg-Perten falling number. *Theoretical and Applied Genetics*. Accepted.

Shanmugaraj N, Rajaraman J, Kale S, **Kamal (Muqaddasi) R,** Huang Y, Thirulogachandar V, Garibay-Hernández A, Budhagatapalli N, Moya YAT, Hajirezaei MR, Rutten T, Hensel G, Melzer M, Kumlehn J, von Wirén N, Mock HP, Schnurbusch T (2023) Multilayered regulation of developmentally programmed pre-anthesis tip degeneration of the barley inflorescence. *The Plant Cell*. koad164

Huang Y, **Kamal (Muqaddasi) R,** Shanmugaraj N, Rutten T, Thirulogachandar V, Zhao S, Hoffie I, Hensel G, Rajaraman J, Moya YAT, Hajirezaei MR, Himmelbach A, Poursarebani N, Lundqvist U, Kumlehn J, Stein N, von Wirén N, Mascher M, Melzer M, Schnurbusch T (2023) A molecular framework for grain number determination in barley. *Science Advances*. 9(9): eadd0324

Kamal (Muqaddasi) R, Muqaddasi QH, Schnurbusch T (2022) Genetic association of spikelet abortion with spike, grain, and shoot traits in highly-diverse six-rowed barley. *Frontiers in Plant Sciences*. 13:1015609

Kamal (Muqaddasi) R, Muqaddasi QH, Zhao Y, Schnurbusch T (2022) Spikelet abortion in six-rowed barley is mainly predicted by the final spikelet number with potential spikelet number acting as a suppressor trait. *Journal of Experimental Botany*. 73:2005–2020

Kaur B, Sandhu KS, **Kamal (Muqaddasi) R,** Kaur K, Singh J, Röder MS, Muqaddasi QH (2021) Omics for the improvement of abiotic, biotic, and agronomic traits in major cereal crops: applications, challenges, and prospects. *Plants*. 10:1989

Hussain M, Gul M, **Kamal (Muqaddasi) R,** Iqbal MA, Zulfiqar S, Abbas A, Röder MS, Muqaddasi QH, Rahman M-u (2021) Prospects of developing novel genetic resources by chemical and physical mutagenesis to enlarge the genetic window in bread wheat varieties. *Agriculture*. 11:621

Muqaddasi QH, **Kamal (Muqaddasi) R,** Mirdita V, Rodemann B, Ganal MW, Reif JC, Röder MS (2021) Genome-wide association studies and prediction of tan spot (*Pyrenophora tritici-repentis*) infection in European winter wheat via different marker platforms. *Genes*. 12:490

Kamal (Muqaddasi) R, Sharma P, Kaur G, Srivastava P (2020) Generation and characterization of doubled haploid plants in maize. *Agricultural Research Journal*. 57 (2): 148-153

PARTICIPATIONS

Monogram 2023 (April 3-5, 2023), University of Reading, Reading, UK

The Next Generation Sequencing workshop (August 14–25, 2022), Tel-Hai College, Israel (online)

13th International Barley Genetics Symposium (July3–7, 2022), UL Academic Centre, University of Lativa, Riga, Latvia

Know your Competencies—find your job (June 21–22, 2021), International Graduate Academy, Martin-Luther-University Halle-Wittenberg, Halle (Salle), Germany

Plant Science Student Conference, PSSC (June 15–18, 2021), Leibniz Institute of Plant Genetics and Crop Plant Research (IPK), Gatersleben, Germany

Introduction to Genome-Wide Association Studies (GWAS) (March 1–4, 2020), Physalia Courses, Freie Universität Berlin, Germany

Plant and Animal Genome XXVIII Conference (PAG) (January 11–15, 2020), San Diego, California, United States.

The Great Presentation (Novemeber 4–5, 2019), International Graduate Academy, Martin-Luther-University Halle-Wittenberg, Halle (Salle), Germany

Plant Science Student Conference, PSSC (June 18–21, 2019), Leibniz Institute of Plant Biochemistry (IPK), Halle-Saale, Germany

Breeding Symposium: Breeding for the next 200 years-prospects and challenges (December 5, 2018), University of Hohenheim, Germany

Good scientific practice: Ethical and legal orientation in everyday research (July 17, 2018), Leibniz Institute of Plant Genetics and Crop Plant Research (IPK), Gatersleben, Germany

PROGRAMMING LANGUAGES

R and Linux

- - -