GENOME-WIDE ASSOCIATION STUDIES IN A WILD BARLEY NESTED ASSOCIATION MAPPING (NAM) POPULATION TO REVEAL THE GENETIC ARCHITECTURE OF PLANT DEVELOPMENT AND QUALITY TRAITS

DISSERTATION

ZUR ERLANGUNG DES

DOKTORGRADES DER AGRARWISSENSCHAFTEN (DR. AGR.)

DER

NATURWISSENSCHAFTTLICHEN FAKULTÄT III

AGRAR – UND ERNÄHRUNGSWISSENSCHAFTEN
GEOWISSENSCHAFTEN UND INFORMATIK

DER MARTIN – LUTHER – UNIVERSITÄT HALLE – WITTENBERG

VORGELEGT VON

HERRN PAUL HERZIG
GEB. AM 21.08.1988 IN STRALSUND

ERSTGUTACHTER: PROF. DR. KLAUS PILLEN,
MARTIN–LUTHER–UNIVERSITÄT HALLE–WITTENBERG
ZWEITGUTACHTER PROF. DR. JENS LEON,
RHEINISCHE FRIEDRICH-WILHELMS-UNIVERSITÄT BONN

DATUM DER VERTEIDIGUNG: 19.04.2021

Table of contents

1 General introduction

1.1 Global future prospects of agriculture

It is estimated that the world´s population will increase from currently 7.7 billion to up to 9.7 billion in the next 30 years (Figure 1) and could reach its peak at the end of the century at around 11 billion (UN 2019). This has important consequences for the achievement of sustainable development objectives such as the protection of the environment, which is not an end in itself but rather the basis for sustainable and productive agriculture.

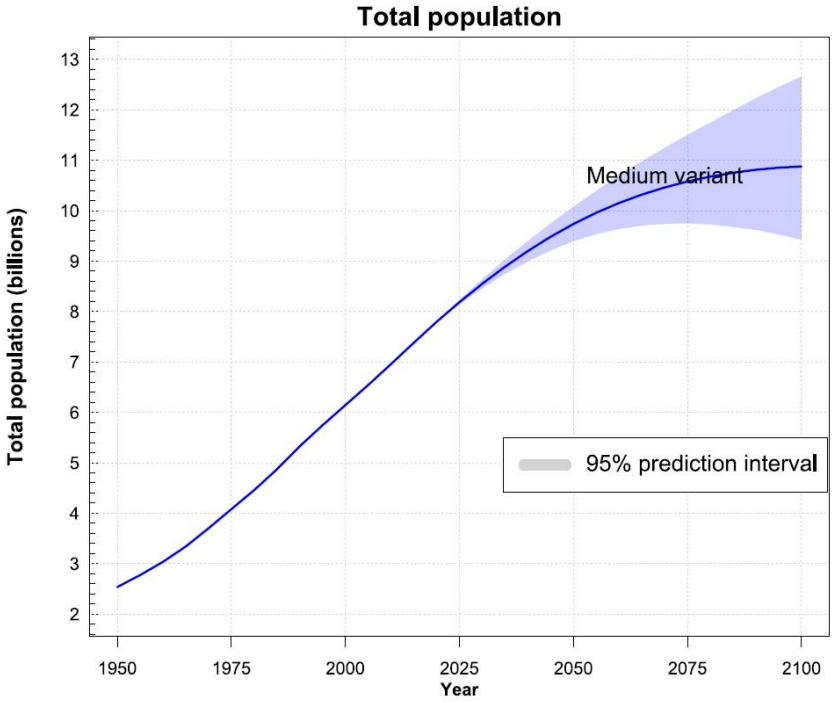

Fig. 1 | Development of the global population in the past and in the future

(Figure from United Nations, World Population Prospects 2019, Volume II: Demographic Profiles)

Moreover, the climate change causes an faster increase in global average temperature, more than in the past 50 million years (Kerr 2007). Even modest climate change scenarios assume an increasing risk of 30% of species extinction in the mid-century (Figure 2). Extreme weather events such as extreme heat and precipitation or coastal flooding, which provoke erosion, soil degradation and desertification (IPCC 2014), are also highly related

to climate change. These major changes challenge agriculture and endanger crop productivity around the globe.

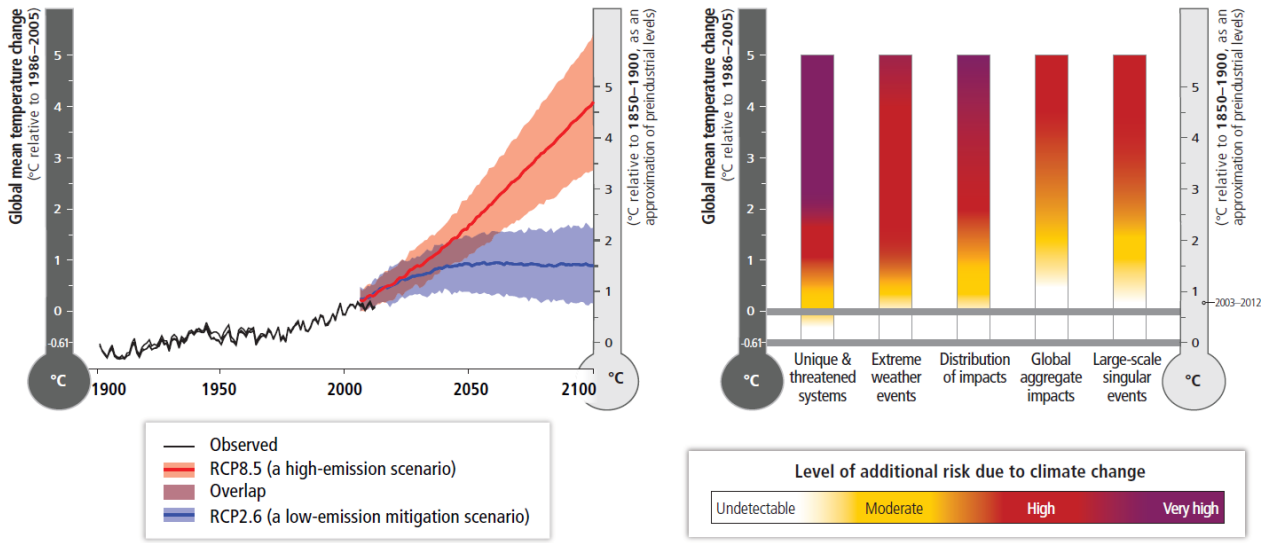

Fig. 2 | Future climate trends and associated risks

Left: Observed change in the average global surface temperature between the average of the period 1850–1900 and 1986–2005 and the predicted changes based on a high (RCP8.5) and a low-emission (RCO2.6) scenario. Right: Climate related risks associated with the global mean temperature change (Figure IPCC (2014)).

Due to drought-related water stress nutrient availability in the soil and the uptake at the root surface will decrease (Brouder & Volenec 2008). Furthermore, rising CO_2 concentration in the atmosphere affects the entire nutrient composition of plants and reduces the amount of essential minerals per calorie (Loladze 2014). Currently, more than 2 billion people worldwide, mainly women and preschool children in developing countries (Figure 3), are affected by micronutrient malnutrition (WHO & FAO 2018). The political will for an effective change in this area was manifested at the annual International Conference on Nutrition and Growth (WHO 2014). However, in the past, the nutritional value has been ignored by conventional breeding.

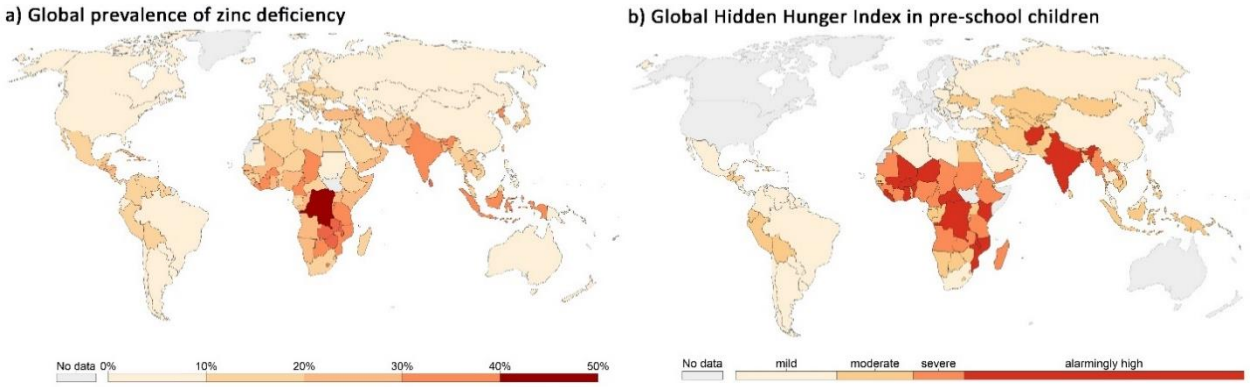

Fig. 3 | Malnutrition on a global scale

a) The global prevalence of zinc deficiency, measured as the share of the total population with intakes below physiological requirements (Wessells & Brown 2012). b) Global Hidden Hunger Index for pre-school children is calculated as the average of three deficiencies (stunting, anemia due to iron deficiency and vitamin-A deficiency) (Muthayya *et al.* 2013). (Figure adapted from Ritchie & Roser (2020))

1.2 Barley

In terms of production, barley is the fourth most important cereal after maize, rice and wheat with a stagnating production over the past two decades. Nevertheless, barley has a large ecological amplitude and is therefore an important crop in marginal environments, which are characterized by abiotic stresses like heat, drought and nutrient deficiency (Baum *et al.* 2007; Tester & Langridge 2010; Rollins *et al.* 2013). Due to its diploid genetics, barley has been established as a model species for temperate cereals.

1.3 Genetics of barley

Cultivated barley (*Hordeum vulgare* ssp. *vulgare*), belongs to the genus *Hordeum*, which is a moderate sized genus with ca. 32 species and altogether ca. 45 taxa. The self-pollinating and diploid cereal has seven chromosomes (2n = 2x = 14) with an estimated genome size of 5.1 Gb (Stein & Muehlbauer 2018). *Hordeum vulgare* ssp. vulgare was domesticated from its wild progenitor *Hordeum vulgare* ssp. *spontaneum*. The initial domestication dating from about 10,000 years ago (Zohary & Hopf 1994) occurred most likely in two different geographical centres. The concept of a second centre of domestication in Central Asia (Figure 4) apart from the Fertile Crescent (Badr *et al.* 2000),

was established by phylogenetic analyses based on nucleotide sequence comparisons (Azhaguvel & Komatsuda 2007; Morrell & Clegg 2007) and is supported by two different rachis mutations which emerged independently in barley (Pourkheirandish *et al.* 2015). Moreover, recently, Civan & Brown (2017) found evidence for a third independent origin of a stiff rachis phenotype of cultivated barley in *brt1b* allele, which draws a more protracted and multiregional domestication scenario.

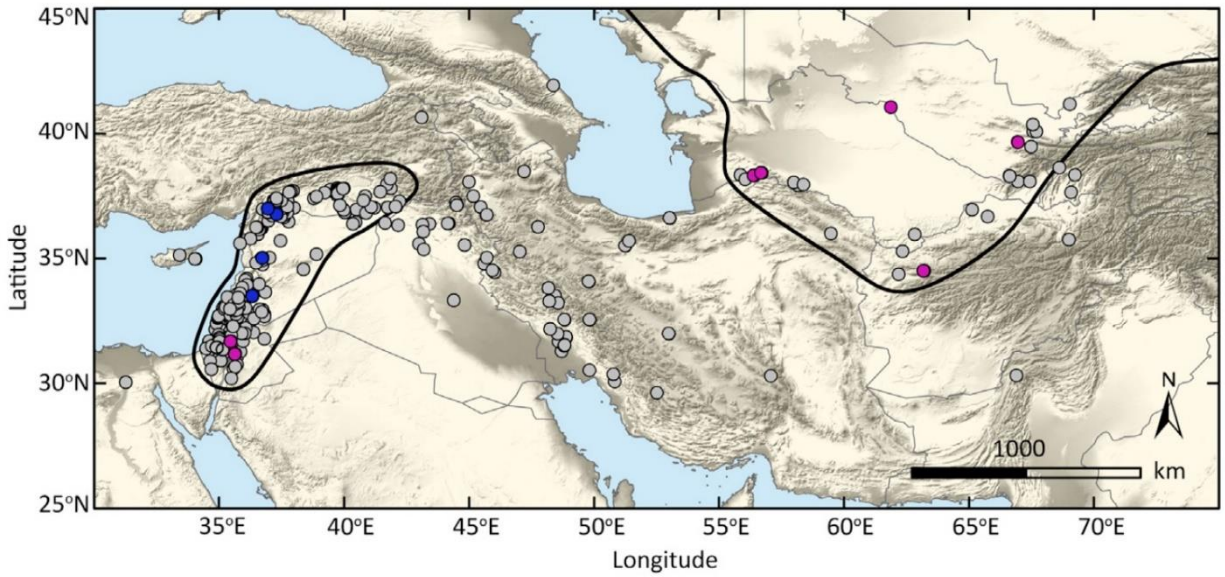

Fig. 4 | Origin of *Hordeum vulgare*

Cultivated barley originated from the Fertile Crescent and Central Asia. Dots are collecting sites of barley accessions harboring *btr1* (purple) and *btr2* (blue). The remaining wild barleys analyzed are indicated with gray dots. (Figure from Pourkheirandish *et al.* (2015)).

The transition from a weed to a crop entailed a ≈50% reduction of genetic diversity, which indicates a strong domestication bottleneck (Pankin *et al.* 2018) (Figure 5). From a breeders view a reduction of genetic diversity of a species compromises the scope of action to cope with future challenges in agriculture. Crop improvement by selection is based on the availability of genetic diversity, which is one of the main resources sustaining human life on earth. Without diversity, crops would be unable to adapt to changing biotic and abiotic conditions. Nevertheless, barley, as one of the first adopted crop species, has a long history of domestication and accompanied human migration paths (Kilian *et al.*

2010). This diversification and adaptation to multiple environmental conditions leads to a wealth of landraces (Pasam *et al.* 2014; Poets *et al.* 2015), which was the basic material for modem plant breeding. Allelic diversity at genes regulating response to day length and cold temperatures favoured the adaptation to different environments.

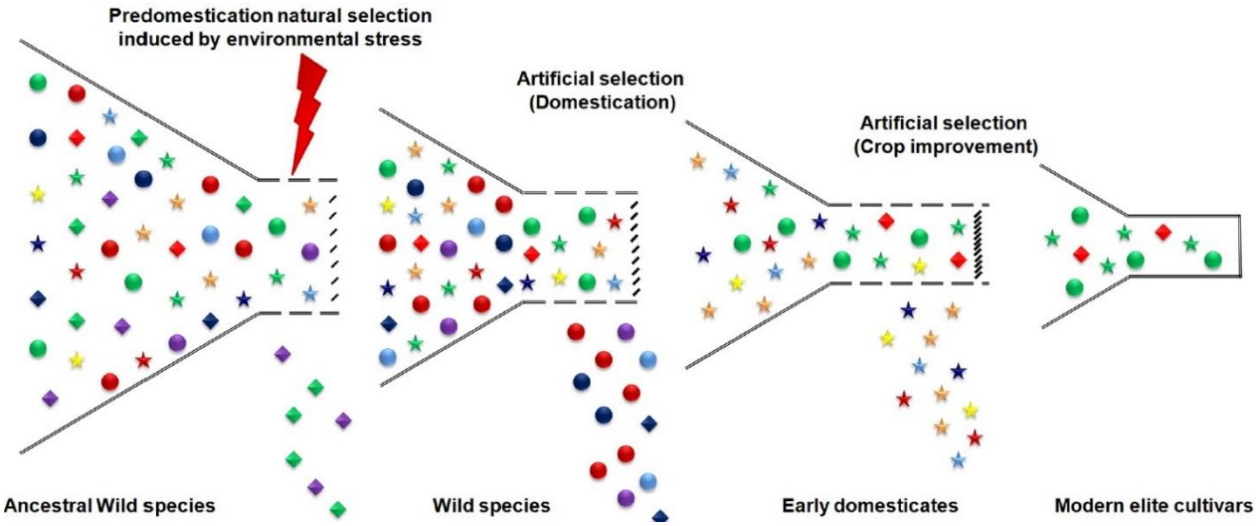

Fig. 5 | Bottleneck effect during domestication and selection

Genetic bottlenecks imposed on crop plants during domestication and modern plant breeding practices. Different classes of genes are represented by different shapes, which were gradually lost through domestication and breeding. Dice shape indicates genes for adaptive traits, circles indicates domestication genes, star shape indicates genes for crop improvement. Such lost alleles can be recovered only by going back to the wild ancestors of our crop species. (Figure from Krishnan *et al.* (2014))

1.4 Plant development and flowering network

Due to these adaptations, barley, as a temperate cereal, can be separated in spring and winter types, bases on the absence and presence of vernalization requirement. It is mainly controlled by the interaction of the two vernalization loci *Vrn-H1* (Yan *et al.* 2003) and *Vrn-H2* (Yan *et al.* 2004)(Figure 6). Winter types capture cold temperature by *cis*-regulatory elements in the promotor region of *Vrn-H1* (Alonso-Peral *et al.* 2011). The resulting up regulation of *Vrn-H1* causes an inactivation of *Vrn-H2*, which leads to an activation of *HvFT1*, the key integrator of transition from the vegetative to the reproductive phase (Yan *et al.* 2006). This ensures flowering only after cold and harsh conditions. Moreover, day

length is another important environmental factor for floral transition. Based on photoperiod, barley varieties can be classified into early flowering under long-day conditions and day-length neutral, late flowering growth types. The photoperiod response is primary controlled by *Photoperiod 1* (*Ppd-H1*), where the sensitive and dominant allele promotes flowering under long days (Laurie *et al.* 1995) and *Photoperiod 2* (*Ppd-H2*), which promotes flowering under short days (Kikuchi *et al.* 2009) (Figure 6). Barley cultivars in short growing seasons are insensitive to vernalization and sensitive to photoperiod to guarantee early flowering (Cockram *et al.* 2007). *Ppd-H1* is also affected by circadian clock genes such as *HvELF3* (Faure *et al.* 2012), *HvLUX* (Campoli *et al.* 2013), and *PHYTOCHROME C* (Pankin *et al.* 2014). The circadian clock is an internal timekeeper that synchronizes physiological and molecular processes with the diurnal cycle (Johansson & Staiger 2015). The proteins of the functional clock genes suppress flowering under non-inductive conditions (short days), whereas non-functional mutants induced flowering through *HvFT1* expression regardless of day length (Stein & Muehlbauer 2018). Also,

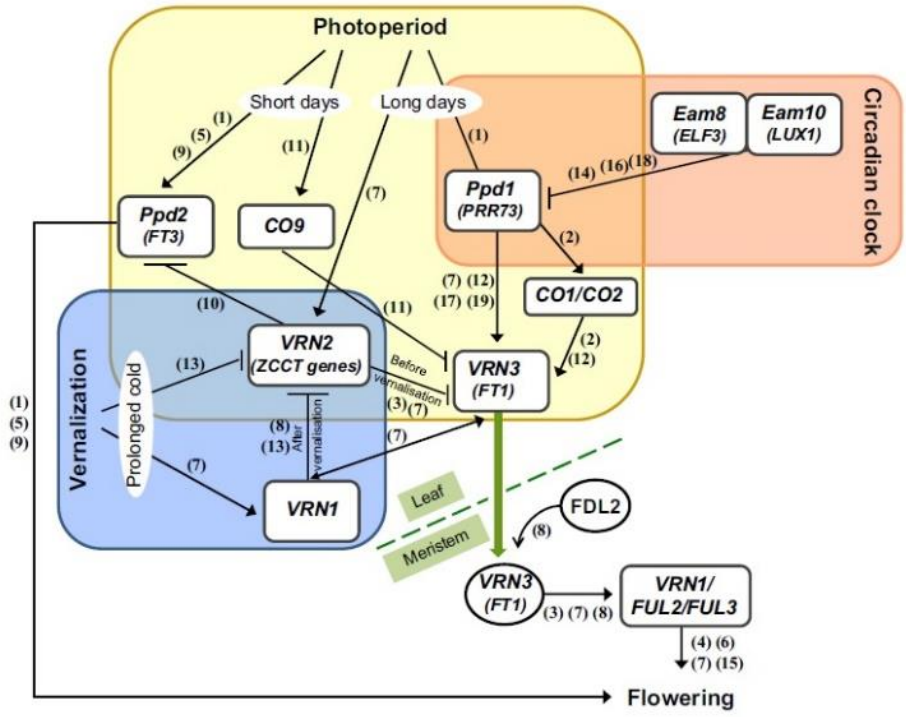

Fig. 6 | Model of flowering time control pathways in wheat and barley.
(Figure from Campoli & von Korff (2014))

phytohormones like gibberellic acid (Boden *et al.* 2014; Youssef *et al.* 2017) are integrated in the complex architecture of floral development.

1.5 Control of nutrient composition in the grains

Early plant development, in particular flowing time, which is a key event and represents the transition from vegetative to generative growth, has a significant influence on yield. Yield, probably the most complex trait, is the product of the so-called yield-related components, which are: total number of ears, grain number per ear and single grain weight (Slafer 2003). Crucial for a maximum final number of fertile florets is a suitable shoot elongation phase and flowing time, which is mediated through the perception of environmental cues (Alqudah & Schnurbusch 2014). The yield of a plant in turn has a decisive influence on the nutrient composition in the grains.

The consequence of the "Green revolution" beginning in the 1960s was a strong increase in grain yields (Evenson & Gollin 2003). A higher yielding plant has a different sink/source relationship, which leads to a reduction in grain quality (Oury *et al.* 2003; Fan *et al.* 2008) and results in a conflict between calories and nutritional value of cereal grains. The so called "dilution effect" is caused by bigger grains due to increased starch accumulation in the endosperm (Kibite & Evans 1984). Most nutrients, in the form of proteins and minerals, are located in the outer layer of cereal grains, especially in the aleurone layer (Lombi *et al.* 2011), whose size is relatively fixed and therefore the relative nutrient concentration is reduced.

Nutrient concentration in barley grains is directly affected by nutrient uptake capacity, transport from root to shoot and by translocation of nutrients from leaves into the grains. The absorption of nutrients into plant tissue is characterized by the biosynthesis and secretion of phytosiderophores and by transporters with high substrate specificity like HvYS1 (Murata *et al.* 2006) and low substrate specificity such as HvHMA2 and HvIRT1 (Pedas *et al.* 2008; Mills *et al.* 2012). Furthermore, the root morphology determines the

soil depth at which the nutrients are available for the plant. The nutrient transfer into the above-ground parts of the plant is considered as one of the rate-limiting processes (Palmgren *et al.* 2008), whereas plant senescence is a decisive step for final grain nutrient concentration. Especially the uppermost leaves are the main source of protein and micronutrients in cereal grains (Barneix 2007; Waters *et al.* 2009). Remobilization efficiency during senescence is the sum of many different factors such as degeneration of macromolecules for instance by proteases, phloem loading of remobilized molecules and passing an apoplastic barrier between mother plant and developing grain. The apoplastic transfer is realized by highly specialized nucellar projection (NP) and endosperm transfer cells (ETC), whose development and differentiation is controlled by gibberellin-to-abscisic acid balances (Weier *et al.* 2014). The strictly ordered progression of leaf senescence is accompanied by changes in expression and activation of a large number of genes in leaves, which is required for the onset of senescence (Buchanan-Wollaston *et al.* 2003). For instance, in *Arabidopsis* 240 transporter genes exhibited increased expression during leaf senescence (van der Graaff *et al.* 2006). In barley, *HvNAM-1*, a key gene in regard to senescence (Uauy *et al.* 2006), affects rapid chlorophyll degradation (Pearce *et al.* 2014) and grain quality traits like protein content and thus significantly influences remobilization efficiency.

In general, there are two strategies to enrich grains with nutrients. Agronomic biofortification is the short-term approach, which means the application of micronutrient fertilisers. The second and most sustainable and cost-effective approach is the genetic biofortification, which is based on the directed breeding of an increased nutrient concentration (Cakmak 2007; Velu *et al.* 2014). This approach could be supported by QTL mapping for grain mineral content to reveal regulatory genes that transmits micronutrient demand to the maternal plant and translocation processes (Waters & Sankaran 2011). However, for this breeding approach, there is a lack of efficient diagnostic methods for the determination of nutrients on a large scale (Pfeiffer & McClafferty 2007).

1.6 Hyperspectral imaging (HSI) - a diagnostic tool

In order to determine nutrient composition in the grain, it is normally necessary to carry out wet chemical analyses. These analyses can be carried out with inductively coupled plasma - optical emission spectrometry (IPC-OES), but there are expensive, time consuming, destructive and require a high labour input. Therefore, it is not possible to perform these analyses on larger sample sizes, which is of crucial importance for plant breeding. In the last decades, the development and improvement of optical and electronic components combined with computers capable of effectively processing information have facilitated the expansion of spectroscopic techniques. A well-known example is the near-infrared spectroscopy (NIRS), which is used for qualitative and quantitative analysis of biological materials (Foley *et al.* 1998) and is also of special interest in early-generation material of cereal breeding programs (Osborne 2006).The basic principle of spectroscopic technologies is based on the interaction between electromagnetic radiation and the object under investigation. There are several ways in which radiation can interact with matter, depending on the wavelength of the radiation and on the nature of the matter. Basically, there are three types of interaction when an electromagnetic wave hit a medium. The wavelengths can be transmitted, reflected or absorbed. Only the reflected part of the radiation is detected by camera sensors, which results in a medium-specific reflectance spectrum.

Hyperspectral imaging sensors by definition have a spectral resolution of at least 20 equally distributed wavelengths and the spectral range can extend beyond the visible range (ultraviolet, infrared). Furthermore, HSI has a spatial resolution of the object in contrast to most NIR technologies. This dataset forms a so-called hyperspectral hypercube (Figure 7). Spectroscopic measurements are always multivariate, especially HSI data, since the intensities of many measured wavelengths form the variables to be evaluated. Therefore, the corresponding mathematical methods must be able to process them. It has been shown that artificial neural networks (ANN) can be used to determine the underlying

function of the relationship, even without explicit knowledge of the relationship between variables (Tetko *et al.* 1995; Miao *et al.* 2006). ANN are also suitable for the modelling of complex non-linear relationships (Vellido *et al.* 1999). The aim is to train a neural network to a calibration dataset (n ≈ 10% of the total dataset) in order to model the relationship between target value and reflection spectrum. The target values of the calibration dataset are determined in the laboratory using standard chemical methods with high precision. The final model created by the ANN is now applied to the entire dataset, with target values estimated based on the spectral information (ElMasry & Sun 2010; Li *et al.* 2014).

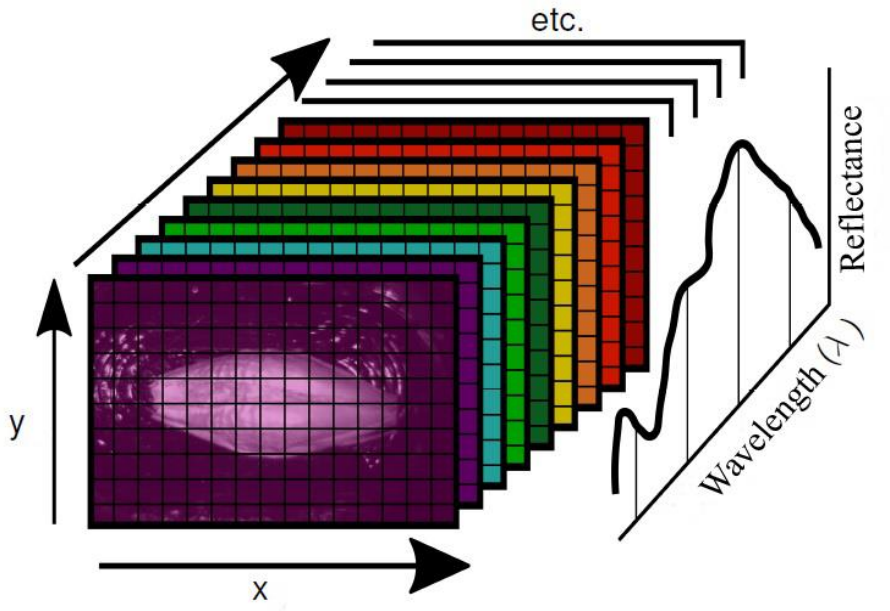

Fig. 7 | Hyperspectral hypercube

A hyperspectral image cube where many images at different wavelengths are stacked to form a cube with two spatial dimensions (x, y) and a spectral dimension (λ) (Figure adapted from Jones & Vaughan (2010).

1.7 Broadening the genetic base by association mapping & allele mining

In times of climate change and world population growth, it is highly important to adapt crops to rapidly changing environmental influences - at best with a gradually increased yield. The importance of conservation and utilization of genetic resources has become a

major issue of global concern (Pistorius 1997) and was matter of the world conference in Rio de Janeiro in 1992. The resulting Convention of Biodiversity has been ratified by 196 countries. The establishment of seed banks in the 1970s around the globe conserved roughly 466,500 *Hordeum* accessions (FAO 2007). Nowadays it is time for utilization and for an effective management of these plant genetic resources (Zamir 2001; Zhang *et al.* 2017).

The development of molecular DNA marker technologies enabled the assay of sequence polymorphisms across the whole genome. With the beginning of the biotechnology era new opportunities for the utilization of plant genetic resources arose by marker-assisted selection (MAS) (Mohan *et al.* 1997). It started in die early 1980s with restriction fragment length polymorphism (RFLP) DNA markers (Botstein *et al.* 1980). Today single nucleotide polymorphism (SNP) are the most common marker system in crops because SNPs are highly abundant across whole genomes (Ganal *et al.* 2009). This marker type was used for the development of SNP arrays, where a large number of markers can be used in parallel for genotyping across the whole genome (Close *et al.* 2009). For instance, the barley 50k Illumina Infinium iSelect genotyping array enables to screen 43,461 SNPs in parallel (Bayer *et al.* 2017).

For the sake of molecular marker mapping, recombination rates from all pairwise marker combinations in a mapping population are determined. This results in linkage groups (chromosomes) and indicates the relative genetic distance between markers. These genetic maps offer the possibility to estimate associations between phenotype and specific genomic regions by so-called marker trait associations. Linkage mapping (LM) is the easiest way to do so, by the use of linkage disequilibrium. In linkage disequilibrium, it is assumed that a significant molecular marker of a trait is genetically linked to the causative gene. LM necessitates the creation of a segregating biparental population, which means the progeny of two intendedly crossed parents differing in the phenotypic trait of interests. The F1 plants are selfed and the resulting F_2 population is genotyped and

phenotyped to detect markers, that are associated with trait variation. If the trait has a quantitative characteristic, significant marker trait associations are defined as quantitative trait loci (QTLs). The detection of QTLs can be achieved by several statistical approaches.

Single-factor analysis, based on the investigation of phenotypic differences between allele combinations of the marker (AA, Aa, aa) is the simplest one. A test of significance can, for instance, be carried out by one-way analysis of variance (ANOVA) or linear regression, where the phenotype is regressed on the number of a certain marker alleles (i.e. AA=2, Aa=1, aa=0) (Lander & Botstein 1989). This analysis is repeated for each marker separately. Interval mapping is an advanced method in regard to a more precise determination of the QTL location. It estimates the QTL location relative to a marker and its flanking markers by including the calculation of a logarithm of odds (LOD) score, where the highest LOD score within a marker interval indicates the most probable location of the QTL (Jansen 1993). In the next step in the evolution of QTL analysis the genetic background noise was reduced by simultaneous analyses of all markers. The so-called composite interval mapping offers a more accurate and efficient mapping method (Zeng 1994). LM is restricted in detection of QTL to the use of bi-parental F_2 or backcross populations.

However, in terms of evaluation and use of genetic resources it would be advantageous to examine a set of diverse wild relatives in a population. To exploit a wider range of genetic variation for traits of interest association mapping (AM) became the method of choice. AM deals with diversity that is present in nature and the time and money consuming development of an artificial mapping population is no longer necessary (Xu *et al.* 2017). Overall, the high number of alleles and the use of natural variation, due to historic recombination events, resulted in smaller linkage blocks, which means a higher mapping resolution in AM compared to bi-parental populations (Myles *et al.* 2009). A disadvantage of AM was a reduced detection power due to low allele frequencies in the population (Korte & Farlow 2013). The major concern in AM is the presence of population structure, which may cause the detection of false-positive marker-trait associations

(MTAs) (Bernardo 2014). This problem is addressed by Yu *et al.* (2006), who implemented a coefficient of coancestry as a measure of relatedness among individuals in a mixed-model approach. The implementation of genetic background noise and genetic relatedness outperformed existing methods in terms of detection power as well as false discovery rate (Segura *et al.* 2012).

Nevertheless, genetic background noise and population structure have still been a present problem. In order to overcome these problems different mapping populations have been developed over time. In regard of the utilization of genetic resources there are more things to consider in population development. Landraces or wild relatives carry a lot of negative genes, which might reduce the agronomic fitness that makes it impossible to phenotype agronomically important traits in the progeny of a genetic resource and a cultivar. To minimize this linkage drag by reducing introgression size several rounds of backcrossing are integrated in population development (Tanksley & Nelson 1996). The advanced backcross populations were used successfully in several organisms like tomato (Tanksley *et al.* 1996), barley (Pillen *et al.* 2003) or rice (Septiningsih *et al.* 2003). Another approach is to build up an introgression library (IL). For this purpose, chromosomal segments of a recipient inbreed cultivar are sequentially replaced by the donor (genetic resource). To achieve complete coverage of the genomes marker-assisted selection is used during backcrossing and selfing. Genetic background noise and epistatic interaction between QTLs can almost be excluded in introgression lines, especially when introgression size is low. In an IL approach favourable detected QTLs could be directly used by breeders. This method was initially established by the introgression of favourable QTL alleles from a wild species into an adapted cultivar in tomato (Eshed & Zamir 1994) and could also be applied in barley (Schmalenbach *et al.* 2008; Schmalenbach & Pillen 2009; Honsdorf *et al.* 2014). Disadvantages of the IL approach are the costly creation of the population and the small number of exotic alleles to be tested.

To combine the statistical power of QTL detection of LM and the use of historic recombination events and high allelic richness offered by AM, a next-generation of mapping resources has emerged over the last years (Ladejobi *et al.* 2016). The so-called multi-parent populations can be divided into two most commonly forms, nested association mapping (NAM) and multi-founder advanced generation inter-cross (MAGIC) populations. NAM populations are built up by crossing a single common parent with a divers set of founder lines resulting in segregating progenies in multiple subpopulations. The first NAM population was created in maize where 25 diverse founders and the common parent resulted in 25 subpopulations with 200 recombinant inbred lines (RILs) each (Yu *et al.* 2008). Thereby the NAM strategy addresses the dissection of complex traits and has been used to study the genetic architecture of a number of morphological and disease resistance traits (Buckler *et al.* 2009; Tian *et al.* 2011; Bajgain *et al.* 2016). The biparental subpopulation structure enables the traceability of detected allele effects to the corresponding donor, which could be helpful when using beneficial wild alleles. The advantage of this kind of mapping population is evident in the high number of NAM populations across all plant genera, such as sorghum (Jordan *et al.* 2011), wheat (Bajgain *et al.* 2016), barley (Maurer *et al.* 2015; Nice *et al.* 2016), rice (Fragoso *et al.* 2017), soybean (Song *et al.* 2017), oilseed rapes (Hu *et al.* 2018) and peanut (Gangurde *et al.* 2019).

The development of MAGIC populations is based on a balanced funnel crossing scheme of an even number of parental lines, resulting in a diverse population whose genomes are fine-scale mosaics of contributions from all founders (Ladejobi *et al.* 2016). MAGIC populations share the advantages of the NAM design and at the same time possess a lower risk of hidden population structure. Due to the absence of subpopulations the MAGIC approach achieves a high detection power also in smaller population sizes compared to NAM populations (Huang *et al.* 2015). Furthermore, in MAGIC populations identified favourable alleles can rarely be referred to the contributing parent if a bi-allelic marker

system (e.g. SNPs) is used. Thus, a direct use of favourable alleles by MAS is more difficult. The first MAGIC population was developed by Churchill *et al.* (2004) in mice followed in plant species by *Arabidopsis thaliana* (Kover *et al.* 2009), wheat (Huang *et al.* 2012; Sannemann *et al.* 2018), barley (Sannemann *et al.* 2015), and rice (Bandillo *et al.* 2013).

In the present thesis an NAM approach was used. For this purpose, 25 highly divergent wild barley accessions act as donors and were used for initial crosses with the barley cultivar Barke. Twenty-four wild barley accessions (*Hordeum vulgare* ssp. *spontaneum*) originated from the Fertile Crescent and one *Hordeum vulgare* ssp *agriocrithon* accession

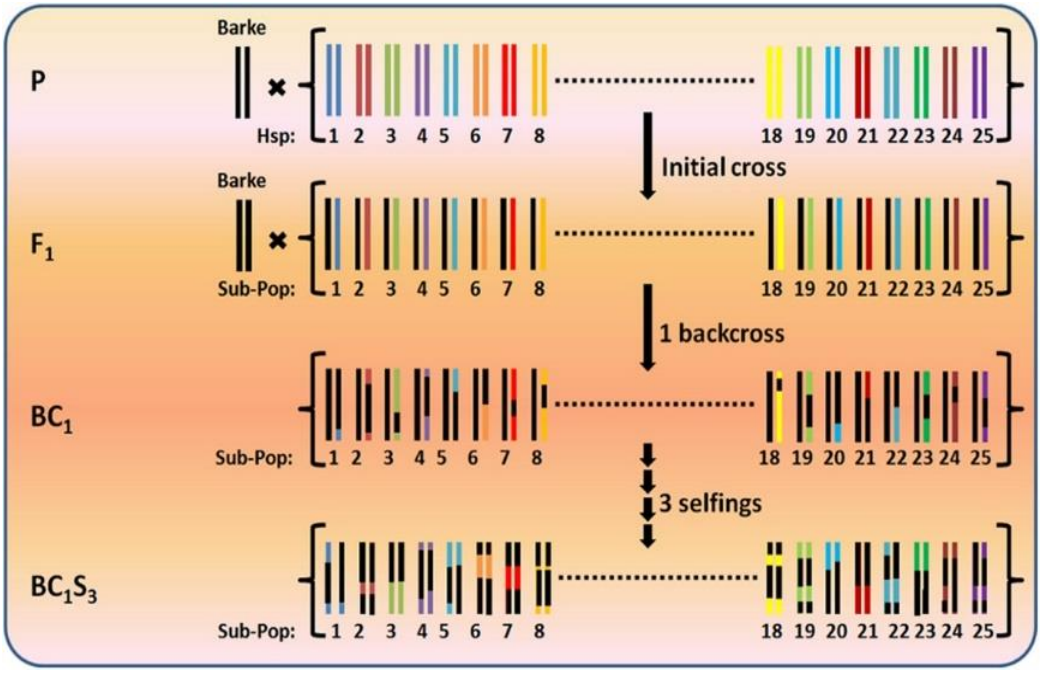

Fig. 8 | Development of the nested association mapping population HEB-25

Colored bars represent chromosome segments originating from a wild donor (Figure from Maurer *et al.* (2015).

from Tibet. The accessions were selected from Badr *et al.* (2000) to represent a maximum part of the genetic diversity. F1 plants were backcrossed to Barke to decrease the proportion of the wild genome in each line to approximately 25%. This modification in the development of the NAM population was necessary to reduce strong negative effects of the donor genomes such as brittleness or bad threshability. After backcrossing, three rounds of selfing resulted in the final population of 1,420 BC_1S_3 lines subdivided in 25

families (Figure 8). The so-called Halle exotic barley 25 (HEB-25) population was introduced by Maurer *et al.* (2015).

1.8 Objectives

This thesis presents the utilization of the NAM population HEB-25 for QTL mapping by genome-wide association studies (GWAS). For the first time, plant development, yield and quality traits were examined simultaneously. The results are discussed with the background of the use of beneficial alleles in the HEB-25. Therefore, four field studies were conducted in Halle (Germany) and Dundee (UK) to investigate the following aspects:

I) Examine the HEB-25 population in two contrasting climatic environments with the focus on location-specific and family-specific impacts to assess the potential ability of the 25 wild barley donors to improve adaptation to future environmental changes.

II) Evaluate the potential of non-invasive HSI to predict mineral element concentration in barley grains.

III) Explore the genetic architecture of element composition to reveal the potential ability of exploiting exotic alleles present in HEB-25 for genetic biofortification in regard to yield and yield related traits.

2 Original Papers

This thesis comprises three original papers in which the HEB-25 population is evaluated for life history traits. Chapters 2.1 (Herzig *et al.* 2018) and 2.2 (Sharma *et al.* 2018) report on the investigation of growth and plant development traits and, respectively, on yield and yield component traits under two contrasting fertilizer treatments studied in two environments. Both chapters focus on explaining the genetic regulation of the studied traits based on genome-wide association studies. Chapter 2.3 (Herzig *et al.* 2019) explore the genetic architecture of element composition of HEB-25 grains by a non-invasive high-throughput hyperspectral imaging (HSI) method to reveal the potential for genetic biofortification using the exotic alleles present in the HEB-25.

2.1 Contrasting genetic regulation of plant development in two European environments revealed by wild barley nested association mapping

2.2 Genome-wide association of yield traits in a nested association map-ping population of barley reveals new gene diversity for future breeding

2.3 Genetic dissection of grain elements predicted by hyperspectral imaging associated with yield-related traits in a wild barley NAM

3 General discussion

3.1 Potential and use of wide phenotypic variation in HEB-25

The phenotypic evaluation of the HEB-25 population resulted in a wide variation among all investigated traits of plant development as well as yield related traits and grain elements. This variation indicates a high genetic diversity, which is an important requirement of breeding and allele mining in genetic resources (Bernardo 2014). Variation of phenotypic traits based on allele diversity deepened the understanding of functions of single genes and revealed a network of genes that determine the phenotype in a particular environment.

Moreover, a comparison between different environments showed different impacts of the same genes, which indicates that phenotypic variation may additionally be caused by epigenetics, triggered by different environmental cues in HEB-25 (Herzig *et al.* 2018). Therefore, to use the full potential of this genetic variation by adapting varieties to future climate scenarios, it is essential to conduct field trials under these environmental conditions, to improve genotypes for future.

Previous studies confirm the wide phenotypic variation of HEB-25 for various traits, like flowering time (Maurer *et al.* 2015), developmental traits (Maurer *et al.* 2016), salinity tolerance (Saade *et al.* 2016) and fungal pathogen resistance (Vatter *et al.* 2017; Vatter *et al.* 2018). The added value of this thesis was the investigation of the whole plant life cycle by including developmental traits (Herzig *et al.* 2018), yield related traits (Sharma *et al.* 2018) and element composition of the grains (Herzig *et al.* 2019) and to detect QTLs for these traits.

3.2 HSI – a technology to relieve the phenotyping bottleneck

During the last two decades, the genomic revolution with next generation sequencing (NGS) technologies has taken place. Reference genome sequences have been published for many crop species (Michael & Jackson 2013) and the resequencing of genomes to assess natural allelic variation has become more and more common, also in barley (Muñoz-Amatriaín & Mascher 2018). Conversely, these achievements highlight new bottlenecks during the plant breeding process. New high-throughput phenotyping techniques are increasingly perceived to ensure the genetic gain in breeding programs (Araus *et al.* 2018), including high-throughput tools and non-invasive quality phenotyping (Fiorani & Schurr 2013).

However, it seems that HSI fulfils all the requirements for screening grains in high-throughput and determine macro and microelement concentrations. The precision accuracy depends on selected elements and was sufficient for subsequent GWAS analysis in the divers HEB-25 panel with wide variation among each element concentration, which could also be a prerequisite for the precision accuracy (Herzig *et al.* 2019). Sample preparation for the non-invasive HSI method is straightforward without any milling procedures, extraction or digestion, which results in a time- and cost-saving way of data acquisition. Due to a quick screening of a large number of genotypes this new method could potentially be used in pre-breeding programs for selecting or rejecting extreme phenotypes in early breeding cycles. Moreover, it is a diagnostic tool, which could catalyse a paradigm shift in the plant breeding communities by allowing direct breeding for macro and micronutrients.

Nevertheless, the currently conducted pilot study showed that the HSI prediction worked well for the HEB-25 population, which is characterized by highly diverse phenotypes. There are indications of limitations of the method for datasets with lower phenotypic variation. In Herzig *et al.* (2019) the carbon concentration in barley grains has the lowest

coefficient of variation of 1.6% and the lowest prediction accuracy of R^2=0.49. Therefore, a screening of genotypes with an expected higher phenotypic similarity is desirable to get more insight into the issue of the required variation for HSI prediction.

For the use of this diagnostic tool in plant breeding communities it has to be noticed that HSI requires deeper understanding of image processing and prediction modelling (which was realized by the Fraunhofer Institute for Factory Operation and Automation (IFF, Magdeburg, Germany) in the framework of this thesis) in addition to the biological point of view. Therefore, a cross-disciplinary cooperation is advisable for a successful establishment of this method. From an economic point of view, the use of the innovative method could also be hampered by high initial costs for the purchase of the HSI camera. Additionally, there is no incentive to breed for crops with enriched nutrients because of absent financial compensation for grains with higher nutritional value.

Nevertheless, the application of HSI seems to be highly versatile in regard to relevant traits for plant breeding. For instance, currently our working group is trying to predict brewing quality parameters like raw protein, malt extract, beta-glucan content, viscosity and β-amylase of barley grains. If prediction models could estimate these traits in a sufficient accuracy, it could be used in early breeding cycles or pre-breeding material to get an insight in brewing quality. So far, an early selection for these traits are impossible because of pricy analysis in the laboratory and insufficient amounts of available plant material. In further experiments, HSI will be used to predict *Fusarium* head blight (*Fusarium graminearum*) disease and the mycotoxin DON (deoxynivalenol) content. A high mycotoxin content is associated with health risks and often not correlated with the visually phenotyping of *Fusarium* head blight by human experts (Birzele *et al.* 2002). Therefore, the research goal is a method with higher degree of objectivity, which is able to distinguish a high and a low DON content. Barbedo *et al.* (2015) demonstrated a HSI method in the 528-1785 nm wavelength range (VR/NIR) for detecting *Fusarium* damaged grains with an error rate of only 9%. In terms of the nutritional value of grains, the

bioavailability of nutrients plays an essential role. Antinutrients like phytic acid, which acts as an effective chelator of cations like P, Mg, Ca, Fe, Zn and Mn (Liu *et al.* 2007), reduce the content of these elements and are thus of great importance in agricultural animal nutrition. The determination of phytic acid could be another interesting field of application for HSI and a crucial step to breed for low phytic acid varieties.

3.3 Association mapping, effect estimation and identification of putative candidate genes

In this thesis, GWAS was conducted with the multiple linear regression 'Model-A' of Liu *et al.* (2011), which showed the highest predictive power in comparison with other joint linkage association mapping models (Würschum *et al.* 2012). It was successfully applied in the NAM population HEB-25 by Maurer *et al.* (2017) and subsequently applied for several traits, for instance leaf sheath hairiness (Saade *et al.* 2017), yield and yield related traits (Sharma *et al.* 2018) and drought stress (Pham *et al.* 2019). 'Model A' makes use of cofactors, which can increase QTL detection power by effectively taking the genetic background into account. The risk of model overfitting was reduced by minimizing the Schwarz Bayesian Criterion (Schwarz 1978) during systematic forward-backward cofactor selection.

To increase the robustness of the QTL detection, a five-fold cross-validation (repeated 20 times) was applied and a QTL was defined if it was significant in at least 25% of cross-validation runs (DR ≥25). This threshold represents a trade-off as on the one hand it assumes rare alleles with a low frequency across all 1420 lines based on a population derived from 25 different wild donors with potentially 25 different alleles. At a DR>25 the risk increases of losing these minor QTLs that could have a weak effect across all families, but a strong effect within families. In regard of favourable allele mining it would be very misleading to lose these interesting minor alleles. For instance, the QTL hotspot 2H-5 in Herzig *et al.* (2019) revealed high variation in family-specific effect sizes. As well as for many other nutrients, the QTL hotspot had an effect range of -15.4% to 3.2% on Mo

concentration in grains and affected TGW by -1.4% to 6.6% with a DR of 'only' 25. On the other hand, a low DR increases the risk of false-positive MTAs and therefore QTLs with low DR have to be treated with more caution. Vilhjalmsson & Nordborg (2013) concluded that population structure in form of genetic background incorporation has to be taken into account for GWAS studies to reduce the false-positive rate, which was done by including cofactors in this thesis.

Because of 25 different wild donor families, allelic variation among families was assumed, which could not be expressed by a single QTL effect. The low number of 22 - 75 genotypes per family limited the possibility of single-family GWAS analysis due to reduced statistical power within families. Therefore, an indirect post-hoc estimation of family-specific QTL effects cumulating significant SNP marker effects as described in Maurer *et al.* (2017) was applied. Each peak marker was placed central in a 26 cM interval and within this interval effect estimation of all significant SNPs were cumulated family-wise. This was done for each cross-validation run and the calculated mean was used as family-specific QTL effect. Thus, QTL definition leads to a low resolution of 26 cM, which increases the risk that two linked QTLs were merged. Otherwise, 26 cM reflects the mean introgression size in HEB-25 (Maurer *et al.* 2015), which makes the choice of a smaller interval implausible. In general, QTL positions should be interpreted with caution. Within the frame of this thesis the HEB-25 population was genotyped with the iSelect 9K chip (Comadran *et al.* 2012) with 5,398 informative SNPs. Meanwhile, HEB-25 was genotyped using the 50k Illumina Infinium iSelect SNP array, developed by Bayer *et al.* (2017), with more than 33,000 informative SNPs in the population. Presumably, this leads to a higher accuracy in QTL localization caused by the higher number of markers segregating in all families. Moreover, this could also enhance family-specific effect estimation.

To identify potential candidate genes and to find physiological mechanisms that are responsible for QTL detection BARLEYMAP (Cantalapiedra *et al.* 2015) was used. Therefore, the peak marker was aligned against the barley physical map (International

Barley Genome Sequencing *et al.* 2012; Mascher *et al.* 2017) and the POPSEQ map (Mascher *et al.* 2013). In rare cases, the peak marker position was located directly in the gene, for instance *Ppd-H1*. Even if the peak marker itself is not a causal SNP, it is most likely that *Ppd-H1* can explain the QTL effect. In many other cases discussed QTLs were located in the peri-centromeric region of the chromosomes where recombination is restricted (Mascher *et al.* 2017). Moreover, candidate genes within the mean introgression size of 7.85 cM around the peak marker revealed several potential candidates and it was unfeasible to discuss all of them. Only the nearest and most probable candidates known from literature were included in the discussion. Subsequently, to achieve the final confirmation of candidates a target-specific gene knock-out or transformation approach for example with CRISPR/Cas9 system (Doudna & Charpentier 2014) should be applied. Another way to fine map detected QTLs in HEB-25 is the use of heterogeneous inbred families (HIFs). For this purpose, near-isogenic lines (NILs) from HIFs, which only segregate for a genomic region of interest, are used for characterization of individual loci as described in Tuinstra *et al.* (1997). As a big advantage, in the HEB-25 population HIFs can directly be created from single HEB lines that segregate at the desired locus due to remaining heterozygosity in BC1S3.

In contrast to molecular genetics, in quantitative genetics it is hardly possible to prove the identity between a gene and a detected QTL. To overcome the problem of speculating and to characterize identified QTLs in more detail exome capture sequencing of all 1420 HEB lines was currently performed. The expected roughly 580,000 SNPs, which captured two-thirds of all high confidence genes in barley by at least one SNPs will dramatically increase the probability to find causal SNPs. Moreover, quantitative genetics is characterized by this direct link between phenotype and genotype in specific environments. This is an advantage of association studies and represents applied breeding research because of using favourable alleles without full knowledge of the gene and his function behind.

3.4 Favourable wild alleles in HEB-25 to improve spring barley pre-breeding pools

On the one hand, GWAS results can be used to deepen the understanding of gene functions encoded by QTLs. On the other hand, GWAS results and QTL effect estimations are useful indicators for the pre-selection of favourable wild alleles in breeding programs. Selection for improved cultivars is only possible if there is genetic variation to start with and wild relatives serve as a source of variation. Therefore, the following examples are discussed in the background of a rapidly changing European environment, which requires more adaptation of cultivars than ever before based on visionary breeding activities.

Since the 'Green Revolution', the use of semi-dwarfing genes in breeding pools is crucial for the development of modern cultivars to increase harvest index and to improve lodging resistance. In the European spring barley gene pool the semi-dwarf locus *sdw1/denso* is predominant (Dahleen *et al.* 2005). The gene of the *sdw1/denso* locus *HvGA20ox*$_2$ encodes a GA-20 oxidase (Jia *et al.* 2009) and is responsible for gibberellic acid (GA) biosynthesis. The *denso* allele has a strong decreasing effect on GA levels in barley (Jia *et al.* 2015) which has far reaching effects. The most favourable effect of *denso* is the semi-dwarf phenotype which causes higher yielding plants (Hellewell *et al.* 2000). The reduced plant height is accompanied by many other negative consequences for instance reduced malting quality and high levels of beta-glucan and reduced seed weight (Li *et al.* 2006; von Korff *et al.* 2006). In the HEB-25 population, several wild alleles showed an increasing agronomic performance at this locus. Maurer *et al.* (2016) detected an accelerating effect up to 5.7 days for shooting on plant development in combination with a delaying effect on shoot elongation phase and ripening phase. As an escape strategy, accelerating plant development could avoid the increasing risk of drought, which is given in future climate scenarios. Interestingly, the accelerating effect is more pronounced in moderate-to-continental than in maritime coastal growing conditions, which indicates a day-length × temperature interaction (Herzig *et al.* 2018). Unfortunately, the increasing wild allele

effect on TGW of this locus, which was found by Maurer *et al.* (2016) could not be verified by Herzig *et al.* (2019) in Halle. In Dundee a significant positive effect on the TGW of up to 6.86% in family 09 could be observed, but this is accompanied by a slightly reduced GNE of 3.5% (Sharma *et al.* 2018). Nevertheless, 10 out of 15 grain quality traits were positively affected by the exotic allele at this locus. The Fe, GCP and Zn concentrations were on average increased by 5.5%, 5.1% and 6.9% compared to Barke without any negative effect on yield or yield components (Herzig *et al.* (2019). This observation is striking, especially due to the fact that grain quality traits are generally negatively correlated with yield and yield components. Allele effects that break this negative correlation seem to be highly interesting in regard to genetic biofortification. It seems very unlikely to find an allele, which combines both characteristics - the reduced plant height and the other positive effects of *sdw1/denso*, like accelerating plant development and increasing grain quality traits - and to exclude the negative effects. It would be more suitable to introgress other semi-dwarf genes which do not affect GA metabolism in barley like the semi-dwarf gene *uzu* from East Asia, where the reduced plant height is based on brassinosteroid-deficiency (Chono *et al.* 2003).

As well as plant height, flowing time is a key agronomic trait and an adaptation on environmental conditions is critical for yield formation. A major locus affecting the photoperiod response is *Ppd-H1* (Laurie *et al.* 1995). Barley landraces predominantly carry the responsive *Ppd-H1* allele in central and northern Europe under long-day conditions (Cockram *et al.* 2007). However, in long growing seasons with moist summers, late flowering as conferred by the *ppd-H1* allele seems to be favourable for grain filling. The well-known interaction between *Ppd-H1* and day-length seems to be not the only gene × environment interaction of *Ppd-H1*. The location-specific effects between Halle and Dundee, which were estimated within the framework of this thesis, revealed differences in the impact of *Ppd-H1* on different growth stages. Especially in the late plant developmental stages like RIP and MAT differences were greater. In single families the

Ppd-H1 effect deviated for RIP up to 2.6 days and for MAT up to 4.7 days with a lower effect amplitude in Dundee. Based on the different day lengths between Dundee and Halle, a stronger Ppd-H1 effect size was expected in Dundee (Herzig *et al.* 2018). The fact, that it occurred the other way around hinted on a *Ppd-H1* interaction with other environmental cues. The higher ambient temperature in Halle is the most likely factor accelerating plant development and potentially interacting with *Ppd-H1*. Higher temperatures resulting in faster reproductive development under long-day conditions based on gene × temperature interaction could be shown by Hemming *et al.* (2012) and Ford *et al.* (2016). Additionally, *Ppd-H1* is integrated in drought stress signals via *HvFT1* (Gol *et al.* 2020) and Ejaz & von Korff (2017) found out that *Ppd-H1* and *Vrn-H1* interact to control inflorescence under high ambient temperature and this is mediated by *Ppd-H1*. This demonstrates that a functional (wild) *Ppd-H1* allele plays an important role to cope with high ambient temperature and to adapt plant development to the respective environment. The non-functional *ppd-H1* allele, prevalent in spring barley cultivars in northern Europe, seems unable to mediate high ambient temperature in this way. In regard to climate change the ecological advantages of *ppd-H1* could thus disappear and in a future scenario the adaptation potential of *Ppd-H1* is getting more important in dry and hot summer locations, which are also predicted in northern Europe. Under current climate conditions in our study the sensitive *Ppd-H1* allele had a slightly negative effect on yield (5%) due to a reduction of GNE by 12.8% on average across families in Halle. The reduction of GNE can be observed in Dundee in the same order of magnitude, but can be masked to some extend by a small increase in the TGW on both locations (Sharma *et al.* 2018). Nevertheless, the accelerating plant growth caused by the functional *Ppd-H1* allele had a positive effect on most of the evaluated grain quality traits. Especially Fe, GPC and Zn concentration were increased by 9.6%, 5.3% and 6.5% compared to Barke. Interestingly, in family 10 yield was only reduced by 1.8% in combination with an increased Fe concentration by 6.8%. The exotic allele of family 10 could be a promising candidate in the

future if the latitudinal decline of the *Ppd-H1* allele moves further north and a functional *Ppd-H1* allele improves agronomic performance by increasing yield and quality traits.

The usefulness of exotic alleles originating from wild accession to increase grain yield, as the key traits in breeding programs, has not been reported frequently. The reason for this is the so-called 'linkage drag' (Feuillet *et al.* 2008). It means that the effects of deleterious wild alleles, which are linked to favourable wild alleles, may mask the much rarer effect of beneficial alleles. This is what makes working with genetic resources in breeding programmes so difficult. In order to obtain the smallest possible coupling groups and a small 'linkage drag', a backcross with the cultivar germplasm is necessary (Li *et al.* 2005). Although HEB-25 was only backcrossed once, Sharma et al. (2018) was able to identify beneficial family specific effects on yield formation. A strong family-specific effect was found in family 01 associated with the QTL hotspot 7-1, which revealed an increased yield of up to 6.8% caused by higher GA and GNE. All other families showed a decreasing effect on yield at this locus. In total, Sharma et al. (2018) defined 14 QTL-hotspots, which exerted effects on yield formation. Ten of these hotspots were located outside of low-recombining peri-centromeric regions and may be easily integrated into breeding programs by crossing and recombination. Also, Wiegmann et al. (2019), was able to identify some HEB lines, which revealed a higher yield than the local elite variety in several locations. These results underline the usefulness of genetic resources and demonstrate its great potential to increase yield and biodiversity of modern elite cultivars.

4 Summary

The UN estimated that the global population would will increase from 7.7 to 9.7 billion people by 2050. This human population growth will exacerbate the problem of global food supply - especially micronutrient malnutrition, the so-called hidden hunger, affects more than one-half of the world's population. Simultaneously, an increasing global average temperature and the occurrence of extreme weather events threaten crop productivity and reduce arable land worldwide. Plant breeding plays an important role in order to cope with these issues. The task for plant breeders is the development of higher-yielding cultivars in a sufficient quality for human diet or animal feed. In this regard, the biggest challenge is to enrich the genetic diversity of the breeding gene pool, which was depleted during domestication, to create phenotypic variation. This variation is a requirement for the breeding selection process and to find new favourable alleles for adaptation in the background of a rapidly changing environment.

In the present thesis agronomic traits of plant development, yield and yield components, as well as grain elements were investigated in the wild barley nested association mapping (NAM) population HEB-25. The mapping population consists of 1420 lines, subdivided into 25 families with up to 75 genotypes. Each line resulted from initial crosses between the German barley cultivar Barke and one of the 25 highly divergent wild barley accessions from the Fertile Crescent and Tibet. This population was used for an exemplary diversification of the barley gene pool and an acquisition of phenotypic data by means of new sensor based methods. The phenotyping of the grain elements was carried out with an innovative hyperspectral imaging screening method. The evaluation of the data, gathered in two years on two different locations in the field, showed a wide phenotypic variation across all traits. The subsequently conducted genome wide associations studies (GWAS) revealed the genetic network, which determine the agronomic traits.

In a comparison of the genetic regulation of plant development within HEB-25 between a moderate-to-continental location and a maritime location highly significant gene ×

environment interactions were detected. Location-specific QTL effects were obtained, for example for *Ppd-H1* indicating not only interactions with day-length but also with temperature. The comparison of the detected location-specific QTL effects could be used for anticipatory selection of favourable wild alleles in regard to future climate scenarios (Braun *et al.* 2010). Moreover, the location-specific QTL effects result in a deeper understanding of the regulation of plant developmental genes.

The results of a novel non-invasive high-throughput phenotyping platform, which was used to determine grain element concentrations by prediction models based on trained artificial neural networks, showed sufficient prediction accuracy for elements of interests, serving as a proof of concept for this method. In comparison with laboratory measurements, this method is fast and cost-effectively and, therefore, could be of interest for gene banks and pre-breeders to phenotypically characterize grains on a large scale. The prediction accuracy (r^2) of the 15 grain elements ranged between r^2= 0.97 for grain protein content and r^2=0.49 for grain carbon concentration. Together with the yield and yield-forming components, the element concentration could be used to uncover nutrient fluxes and relationships among each other. The QTLs, detected by GWAS, confirmed the highly positive correlations between most of the elements, which resulted in 75 QTL hotspots, whereof many of them had significant effects for several grain traits.

Moreover, it was not only possible to estimate a general QTL effect but also 25 family-specific effects for QTLs. The detected wide effect variation reflects the allelic diversity of the multiparental HEB-25 population. The distinction of single alleles enables the selection of favourable exotic alleles for agronomically crucial traits, which could be used for increasing the genetic diversity in modern barley breeding programs. In addition, HEB-25 can also act as a valuable source for the discovery and characterization of genes that underlie traits of agronomic interests.

5 Zusammenfassung

Hochrechnungen der UN zufolge wird die Weltbevölkerung von 7,7 auf 9,7 Milliarden Menschen im Jahre 2050 ansteigen. Dieses Wachstum der Bevölkerung wird das Problem der weltweiten Nahrungsmittelversorgung, insbesondere das der Mangelernährung, wovon heute schon mehr als die Hälfte der Weltbevölkerung betroffen ist, weiter verschärfen. Zeitgleich bedrohen der Anstieg der weltweiten Durchschnittstemperatur und das gehäufte Auftreten von extremen Wetterbedingungen die Pflanzenproduktivität und reduzieren Ackerland weltweit. Die Pflanzenzüchtung spielt eine entscheidende Rolle bei der Bewältigung dieser Probleme. Die hohe Anforderung an den Pflanzenzüchter ist dabei die Entwicklung neuer Hochertragssorten mit hinreichender Qualität für die menschliche Ernährung oder Tierfutter. Die größte Herausforderung ist es dabei, den durch jahrtausendelange Selektion verarmten Genpool wieder mit genetischer Diversität anzureichern, um ausreichend phänotypische Variation zu erzeugen. Diese Variation ist Voraussetzung für den züchterischen Selektionsprozess und für das Finden neuer vorteilhafter Allele für die Anpassung an eine sich rasant ändernde Umwelt.

In der vorliegenden Arbeit wurden agronomisch Merkmale der Pflanzenentwicklung, Ertrag und Ertragskomponenten sowie Korninhaltsstoffe an der Wildgersten- *nested association mapping* (NAM)- Population HEB-25 untersucht. Die Kartierungspopulation setzt sich aus 1420 Linien zusammen, die in 25 Familien mit bis zu 75 Genotypen unterteilt werden kann. Jede Line geht aus der Kreuzung der deutschen Kulturgestensorte Barke und einer von 25 hochdiversen Wildgerstenakzessionen aus dem Fruchtbaren Halbmond und Tibet hervor. Diese Population diente der exemplarischen Diversifizierung des Gersten-Elitegenpools und der Erfassung der phänotypischen Variation mit neuen methodischen Ansätzen. Die Phänotypisierung der Korninhaltsstoffe basierte auf einem innovativen hyperspektralen Bildgebungsverfahren. Die Auswertung der im Feld erhobenen Daten aus zwei Versuchsjahren an zwei unterschiedlichen Standorten ergab eine breite phänotypische Variation über alle Merkmale hinweg. Die anschließend durchgeführten

genomweiten Assoziationsstudien (GWAS) enthüllten das genetische Netzwerk, welches die agronomischen Merkmale determiniert.

Der Vergleich der genetischen Regulation entscheidender Entwicklungsstadien innerhalb der HEB-25 zwischen einem moderat bis kontinental einzuschätzenden Standort und einem maritimen Standort zeigte eine starke Interaktion mit der Umwelt. Standortspezifische QTL-Effekte von beispielsweise *Ppd-H1* lassen auf eine Interaktion nicht nur mit der Tageslänge, sondern auch auf eine Interaktion mit der Temperatur schließen. Der Vergleich der detektierten standortspezifischen QTL-Effekte konnte zum einen dazu genutzt werden, positive Allele herauszuarbeiten um im Hinblick auf zukünftigen klimatischen Bedingungen frühzeitig die richtige Selektionsentscheidung treffen zu können. Zum anderen führten standortspezifische QTL-Effekte zu einem breiteren Verständnis des genetischen Regulationsnetzwerkes der Pflanzenentwicklung.

Die Ergebnisse einer neuartigen nicht-invasiven Hochdurchsatz-Phänotypisierung für die Bestimmung von Korninhaltstoffen mit Hilfe von Vorhersagemodellen aus trainierten künstlichen neuronalen Netzwerken erwies sich als eine hervorragende Methode, um die angestrebten Merkmale mit einer ausreichenden Prädiktionsgüte zu bestimmen. Die im Vergleich zu Labormessungen schnelle und kostengünstige Methode dürfte aufgrund des möglichen hohen Probendurchsatzes bei der Untersuchung von Genbankmaterial oder in Pre-Breeding-Programmen besonders interessant sein. Die Prädiktionsgüte (r^2) der 15 Korninhaltsstoffe lag zwischen r^2= 0,97 für Kornprotein und r^2= 0,49 für die Kohlenstoffkonzentration im Korn. Zusammen mit dem Ertrag und ertragsbildenden Komponenten konnten die Elementkonzentration dazu genutzt werden, Nährstoffflüsse und Beziehungen untereinander aufzudecken. Die mit Hilfe der GWAS ermittelten QTLs bestätigten die starke positive Korrelation vieler Elemente untereinander und ergaben QTL-Hotspots, die einen signifikanten Einfluss auf mehrere Elemente gleichzeitig hatten.

Darüber hinaus konnte nicht nur ein genereller Effekt für jeden QTL berechnet werden, sondern auch 25 familienspezifische Effekte. Die vorhandene breite Effektvariation

spiegelt die eigentliche allelische Diversität der multiparentalen HEB-25-Population wider. Diese Unterscheidung der einzelnen Allele ermöglicht eine Selektion von vorteilhaften exotischen Allelen agronomisch wichtiger Merkmale, welche dazu genutzt werden könnten, die genetische Diversität der modernen Gerstenzüchtung zu erhöhen. Darüber hinaus fungiert die HEB-25 als wertvolle Ressource, die die Entdeckung und Beschreibung von Genen mit agronomischem Interesse ermöglicht.

6 References

Alonso-Peral MM, Oliver SN, Casao MC, Greenup AA, Trevaskis B (2011) The promoter of the cereal VERNALIZATION1 gene is sufficient for transcriptional induction by prolonged cold. PLoS One doi:10.1371/journal.pone.0029456.

Alqudah AM, Schnurbusch T (2014) Awn primordium to tipping is the most decisive developmental phase for spikelet survival in barley. Funct Plant Biol doi:10.1071/Fp13248.

Araus JL, Kefauver SC, Zaman-Allah M, Olsen MS, Cairns JE (2018) Translating High-Throughput Phenotyping into Genetic Gain. Trends Plant Sci doi:10.1016/j.tplants.2018.02.001.

Azhaguvel P, Komatsuda T (2007) A phylogenetic analysis based on nucleotide sequence of a marker linked to the brittle rachis locus indicates a diphyletic origin of barley. Ann Bot doi:10.1093/aob/mcm129.

Badr A, Muller K, Schafer-Pregl R, El Rabey H, Effgen S, Ibrahim HH, Pozzi C, Rohde W, Salamini F (2000) On the origin and domestication history of Barley (Hordeum vulgare). Mol Biol Evol doi:10.1093/oxfordjournals.molbev.a026330.

Bajgain P, Rouse MN, Tsilo TJ, Macharia GK, Bhavani S, Jin Y, Anderson JA (2016) Nested Association Mapping of Stem Rust Resistance in Wheat Using Genotyping by Sequencing. PLoS One doi:10.1371/journal.pone.0155760.

Bandillo N, Raghavan C, Muyco PA*, et al.* (2013) Multi-parent advanced generation inter-cross (MAGIC) populations in rice: progress and potential for genetics research and breeding. Rice (N Y) doi:10.1186/1939-8433-6-11.

Barbedo JGA, Tibola CS, Fernandes JMC (2015) Detecting Fusarium head blight in wheat kernels using hyperspectral imaging. Biosystems Engineering doi:10.1016/j.biosystemseng.2015.01.003.

Barneix AJ (2007) Physiology and biochemistry of source-regulated protein accumulation in the wheat grain. J Plant Physiol doi:10.1016/j.jplph.2006.03.009.

Baum M, Von Korff M, Guo P*, et al.* (2007) Molecular Approaches and Breeding Strategies for Drought Tolerance in Barley. In: Genomics-Assisted Crop Improvement.

Bayer MM, Rapazote-Flores P, Ganal M*, et al.* (2017) Development and Evaluation of a Barley 50k iSelect SNP Array. Front Plant Sci doi:10.3389/fpls.2017.01792.

Bernardo R (2014) Essentials of Plant Breeding. Woodbury, Minnesota: Stemma Press.

Birzele B, Meier A, Hindorf H, Kramer J, Dehne HW (2002) Epidemiology of Fusarium infection and deoxynivalenol content in winter wheat in the Rhineland, Germany. European Journal of Plant Pathology doi:10.1023/A:1020632816441.

Boden SA, Weiss D, Ross JJ, Davies NW, Trevaskis B, Chandler PM, Swain SM (2014) EARLY FLOWERING3 Regulates Flowering in Spring Barley by Mediating Gibberellin Production and FLOWERING LOCUS T Expression. Plant Cell doi:10.1105/tpc.114.123794.

Botstein D, White RL, Skolnick M, Davis RW (1980) Construction of a Genetic-Linkage Map in Man Using Restriction Fragment Length Polymorphisms. American Journal of Human Genetics 32, 314-331.

Braun HJ, Atlin G, Payne T (2010) Multi-location testing as a tool to identify plant response to global climate change. Climate change and crop production. In: Climate change and crop production. Wallingford, UK: CABI.

Brouder SM, Volenec JJ (2008) Impact of climate change on crop nutrient and water use efficiencies. Physiol Plant doi:10.1111/j.1399-3054.2008.01136.x.

Buchanan-Wollaston V, Earl S, Harrison E, Mathas E, Navabpour S, Page T, Pink D (2003) The molecular analysis of leaf senescence -a genomics approach. Plant Biotechnol J doi:10.1046/j.1467-7652.2003.00004.x.

Buckler ES, Holland JB, Bradbury PJ*, et al.* (2009) The genetic architecture of maize flowering time. Science doi:10.1126/science.1174276.

Cakmak I (2007) Enrichment of cereal grains with zinc: Agronomic or genetic biofortification? Plant and Soil doi:10.1007/s11104-007-9466-3.

Campoli C, Pankin A, Drosse B, Casao CM, Davis SJ, von Korff M (2013) HvLUX1 is a candidate gene underlying the early maturity 10 locus in barley: phylogeny, diversity, and interactions with the circadian clock and photoperiodic pathways. New Phytol doi:10.1111/nph.12346.

Campoli C, von Korff M (2014) Genetic Control of Reproductive Development in Temperate Cereals. Molecular Genetics of Floral Transition and Flower Development doi:10.1016/B978-0-12-417162-6.00005-5.

Cantalapiedra CP, Boudiar R, Casas AM, Igartua E, Contreras-Moreira B (2015) BARLEYMAP: physical and genetic mapping of nucleotide sequences and annotation of surrounding loci in barley. Molecular Breeding doi:10.1007/s11032-015-0253-1.

Chono M, Honda I, Zeniya H, Yoneyama K, Saisho D, Takeda K, Takatsuto S, Hoshino T, Watanabe Y (2003) A semidwarf phenotype of barley uzu results from a nucleotide substitution in the gene encoding a putative brassinosteroid receptor. Plant Physiol doi:10.1104/pp.103.026195.

Churchill GA, Airey DC, Allayee H*, et al.* (2004) The Collaborative Cross, a community resource for the genetic analysis of complex traits. Nat Genet doi:10.1038/ng1104-1133.

Civan P, Brown TA (2017) A novel mutation conferring the nonbrittle phenotype of cultivated barley. New Phytol doi:10.1111/nph.14377.

Close TJ, Bhat PR, Lonardi S*, et al.* (2009) Development and implementation of high-throughput SNP genotyping in barley. BMC Genomics doi:10.1186/1471-2164-10-582.

Cockram J, Jones H, Leigh FJ, O'Sullivan D, Powell W, Laurie DA, Greenland AJ (2007) Control of flowering time in temperate cereals: genes, domestication, and sustainable productivity. J Exp Bot doi:10.1093/jxb/erm042.

Comadran J, Kilian B, Russell J*, et al.* (2012) Natural variation in a homolog of Antirrhinum CENTRORADIALIS contributed to spring growth habit and environmental adaptation in cultivated barley. Nat Genet doi:10.1038/ng.2447.

Dahleen LS, Vander Wal LJ, Franckowiak JD (2005) Characterization and molecular mapping of genes determining semidwarfism in barley. J Hered doi:10.1093/jhered/esi125.

Doudna JA, Charpentier E (2014) Genome editing. The new frontier of genome engineering with CRISPR-Cas9. Science doi:10.1126/science.1258096.

Ejaz M, von Korff M (2017) The Genetic Control of Reproductive Development under High Ambient Temperature. Plant Physiol doi:10.1104/pp.16.01275.

ElMasry G, Sun D-W (2010) Principles of Hyperspectral Imaging Technology. In: Hyperspectral Imaging for Food Quality Analysis and Control. Amsterdam, Boston, Heidelberg, London, New York, Qxford, Paris, San Diego, San Francisco, Singapore, Sidney, Tokyo: Elsevier.

Eshed Y, Zamir D (1994) Introgressions fromLycopersicon pennellii can improve the soluble-solids yield of tomato hybrids. Theor Appl Genet doi:10.1007/BF01254002.

Evenson RE, Gollin D (2003) Assessing the impact of the green revolution, 1960 to 2000. Science doi:10.1126/science.1078710.

Fan MS, Zhao FJ, Fairweather-Tait SJ, Poulton PR, Dunham SJ, McGrath SP (2008) Evidence of decreasing mineral density in wheat grain over the last 160 years. J Trace Elem Med Biol doi:10.1016/j.jtemb.2008.07.002.

FAO (2007) The Second Report on the State of the world's Plant Genetic Resources for Food and Agriculture.

Faure S, Turner AS, Gruszka D, Christodoulou V, Davis SJ, von Korff M, Laurie DA (2012) Mutation at the circadian clock gene EARLY MATURITY 8 adapts domesticated barley (Hordeum vulgare) to short growing seasons. Proc Natl Acad Sci U S A doi:10.1073/pnas.1120496109.

Feuillet C, Langridge P, Waugh R (2008) Cereal breeding takes a walk on the wild side. Trends Genet doi:10.1016/j.tig.2007.11.001.

Fiorani F, Schurr U (2013) Future scenarios for plant phenotyping. Annu Rev Plant Biol doi:10.1146/annurev-arplant-050312-120137.

Foley WJ, McIlwee A, Lawler I, Aragones L, Woolnough AP, Berding N (1998) Ecological applications of near infrared reflectance spectroscopy - a tool for rapid, cost-effective prediction of the composition of plant and animal tissues and aspects of animal performance. Oecologia doi:10.1007/s004420050591.

Ford B, Deng W, Clausen J, Oliver S, Boden S, Hemming M, Trevaskis B (2016) Barley (Hordeum vulgare) circadian clock genes can respond rapidly to temperature in an EARLY FLOWERING 3-dependent manner. J Exp Bot doi:10.1093/jxb/erw317.

Fragoso CA, Moreno M, Wang Z*, et al.* (2017) Genetic Architecture of a Rice Nested Association Mapping Population. G3 (Bethesda) doi:10.1534/g3.117.041608.

Ganal MW, Altmann T, Roder MS (2009) SNP identification in crop plants. Curr Opin Plant Biol doi:10.1016/j.pbi.2008.12.009.

Gangurde SS, Wang H, Yaduru S*, et al.* (2019) Nested-association mapping (NAM)-based genetic dissection uncovers candidate genes for seed and pod weights in peanut (Arachis hypogaea). Plant Biotechnol J doi:10.1111/pbi.13311.

Gol L, Haraldsson EB, von Korff M (2020) Ppd-H1 integrates drought stress signals to control spike development and flowering time in barley. J Exp Bot doi:10.1093/jxb/eraa261.

Hellewell KB, Rasmusson DC, Gallo-Meagher M (2000) Enhancing yield of semidwarf barley. Crop Science doi:10.2135/cropsci2000.402352x.

Hemming MN, Walford SA, Fieg S, Dennis ES, Trevaskis B (2012) Identification of high-temperature-responsive genes in cereals. Plant Physiol doi:10.1104/pp.111.192013.

Herzig P, Backhaus A, Seiffert U, von Wiren N, Pillen K, Maurer A (2019) Genetic dissection of grain elements predicted by hyperspectral imaging associated with yield-related traits in a wild barley NAM population. Plant Sci doi:10.1016/j.plantsci.2019.05.008.

Herzig P, Maurer A, Draba V, Sharma R, Draicchio F, Bull H, Milne L, Thomas WTB, Flavell AJ, Pillen K (2018) Contrasting genetic regulation of plant development in wild barley grown in two European environments revealed by nested association mapping. J Exp Bot doi:10.1093/jxb/ery002.

Honsdorf N, March TJ, Hecht A, Eglinton J, Pillen K (2014) Evaluation of juvenile drought stress tolerance and genotyping by sequencing with wild barley introgression lines. Molecular Breeding doi:10.1007/s11032-014-0131-2.

Hu J, Guo C, Wang B*, et al.* (2018) Genetic Properties of a Nested Association Mapping Population Constructed With Semi-Winter and Spring Oilseed Rapes. Front Plant Sci doi:10.3389/fpls.2018.01740.

Huang BE, George AW, Forrest KL, Kilian A, Hayden MJ, Morell MK, Cavanagh CR (2012) A multiparent advanced generation inter-cross population for genetic analysis in wheat. Plant Biotechnol J doi:10.1111/j.1467-7652.2012.00702.x.

Huang BE, Verbyla KL, Verbyla AP, Raghavan C, Singh VK, Gaur P, Leung H, Varshney RK, Cavanagh CR (2015) MAGIC populations in crops: current status and future prospects. Theor Appl Genet doi:10.1007/s00122-015-2506-0.

International Barley Genome Sequencing C, Mayer KF, Waugh R*, et al.* (2012) A physical, genetic and functional sequence assembly of the barley genome. Nature doi:10.1038/nature11543.

IPCC (2014) Climate Change 2014: Impacts, Adaptation, and Vulnerability. Part A: Global and Sectoral Aspects. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge, United Kingdom and New York: Cambridge University

Press.

Jansen RC (1993) Interval mapping of multiple quantitative trait loci. Genetics 135, 205-211.

Jia Q, Li C, Shang Y, Zhu J, Hua W, Wang J, Yang J, Zhang G (2015) Molecular characterization and functional analysis of barley semi-dwarf mutant Riso no. 9265. BMC Genomics doi:10.1186/s12864-015-2116-x.

Jia Q, Zhang J, Westcott S, Zhang XQ, Bellgard M, Lance R, Li C (2009) GA-20 oxidase as a candidate for the semidwarf gene sdw1/denso in barley. Funct Integr Genomics doi:10.1007/s10142-009-0120-4.

Johansson M, Staiger D (2015) Time to flower: interplay between photoperiod and the circadian clock. J Exp Bot doi:10.1093/jxb/eru441.

Jones HG, Vaughan RA (2010) Remote Sensing of Vegetation: Principles, Techniques, and Applications. Oxford: Oxford University Press.

Jordan DR, Mace ES, Cruickshank AW, Hunt CH, Henzell RG (2011) Exploring and Exploiting Genetic Variation from Unadapted Sorghum Germplasm in a Breeding Program. Crop Science doi:10.2135/cropsci2010.06.0326.

Kerr RA (2007) Global warming is changing the world. Science doi:10.1126/science.316.5822.188.

Kibite S, Evans LE (1984) Causes of Negative Correlations between Grain-Yield and Grain Protein-Concentration in Common Wheat. Euphytica doi:10.1007/Bf00021906.

Kikuchi R, Kawahigashi H, Ando T, Tonooka T, Handa H (2009) Molecular and functional characterization of PEBP genes in barley reveal the diversification of their roles in flowering. Plant Physiol doi:10.1104/pp.108.132134.

Kilian B, Martin W, Salamini F (2010) Genetic Diversity, Evolution and Domestication of Wheat and Barley in the Fertile Crescent. Evolution in Action doi:10.1007/978-3-642-12425-9_8.

Korte A, Farlow A (2013) The advantages and limitations of trait analysis with GWAS: a review. Plant Methods doi:10.1186/1746-4811-9-29.

Kover PX, Valdar W, Trakalo J, Scarcelli N, Ehrenreich IM, Purugganan MD, Durrant C, Mott R (2009) A Multiparent Advanced Generation Inter-Cross to fine-map quantitative traits in Arabidopsis thaliana. PLoS Genet doi:10.1371/journal.pgen.1000551.

Krishnan SG, Waters DL, Henry RJ (2014) Australian wild rice reveals pre-domestication origin of polymorphism deserts in rice genome. PLoS One doi:10.1371/journal.pone.0098843.

Ladejobi O, Elderfield J, Gardner KA, Gaynor RC, Hickey J, Hibberd JM, Mackay IJ, Bentley AR (2016) Maximizing the potential of multi-parental crop populations. Appl Transl Genom doi:10.1016/j.atg.2016.10.002.

Lander ES, Botstein D (1989) Mapping Mendelian Factors Underlying Quantitative Traits Using Rflp Linkage Maps. Genetics 121, 185-199.

Laurie DA, Pratchett N, Snape JW, Bezant JH (1995) RFLP mapping of five major genes and eight quantitative trait loci controlling flowering time in a winter x spring barley (Hordeum vulgare L.) cross. Genome doi:10.1139/g95-074.

Li JZ, Huang XQ, Heinrichs F, Ganal MW, Roder MS (2005) Analysis of QTLs for yield, yield components, and malting quality in a BC3-DH population of spring barley. Theor Appl Genet doi:10.1007/s00122-004-1847-x.

Li JZ, Huang XQ, Heinrichs F, Ganal MW, Roder MS (2006) Analysis of QTLs for yield components, agronomic traits, and disease resistance in an advanced backcross population of spring barley. Genome doi:10.1139/g05-128.

Li L, Zhang Q, Huang D (2014) A review of imaging techniques for plant phenotyping. Sensors (Basel) doi:10.3390/s141120078.

Liu K, Peterson KL, Raboy V (2007) Comparison of the phosphorus and mineral concentrations in bran and abraded kernel fractions of a normal barley (Hordeum vulgare) cultivar versus four low phytic acid isolines. J Agric Food Chem doi:10.1021/jf0637776.

Liu W, Gowda M, Steinhoff J, Maurer HP, Wurschum T, Longin CF, Cossic F, Reif JC (2011) Association mapping in an elite maize breeding population. Theor Appl Genet doi:10.1007/s00122-011-1631-7.

Loladze I (2014) Hidden shift of the ionome of plants exposed to elevated CO(2)depletes minerals at the base of human nutrition. Elife doi:10.7554/eLife.02245.

Lombi E, Smith E, Hansen TH, Paterson D, de Jonge MD, Howard DL, Persson DP, Husted S, Ryan C, Schjoerring JK (2011) Megapixel imaging of (micro)nutrients in mature barley grains. J Exp Bot doi:10.1093/jxb/erq270.

Mascher M, Gundlach H, Himmelbach A*, et al.* (2017) A chromosome conformation capture ordered sequence of the barley genome. Nature doi:10.1038/nature22043.

Mascher M, Richmond TA, Gerhardt DJ*, et al.* (2013) Barley whole exome capture: a tool for genomic research in the genus Hordeum and beyond. Plant J doi:10.1111/tpj.12294.

Maurer A, Draba V, Jiang Y, Schnaithmann F, Sharma R, Schumann E, Kilian B, Reif JC, Pillen K (2015) Modelling the genetic architecture of flowering time control in barley through nested association mapping. BMC Genomics doi:10.1186/s12864-015-1459-7.

Maurer A, Draba V, Pillen K (2016) Genomic dissection of plant development and its impact on thousand grain weight in barley through nested association mapping. J Exp Bot doi:10.1093/jxb/erw070.

Maurer A, Sannemann W, Leon J, Pillen K (2017) Estimating parent-specific QTL effects through cumulating linked identity-by-state SNP effects in multiparental populations. Heredity (Edinb) doi:10.1038/hdy.2016.121.

Miao Y, Mulla DJ, Robert PC (2006) Identifying important factors influencing corn yield and grain quality variability using artificial neural networks. Precision Agriculture doi:10.1007/s11119-006-9004-y.

Michael TP, Jackson S (2013) The First 50 Plant Genomes. Plant Genome doi:10.3835/plantgenome2013.03.0001in.

Mills RF, Peaston KA, Runions J, Williams LE (2012) HvHMA2, a P(1B)-ATPase from barley, is highly conserved among cereals and functions in Zn and Cd transport. PLoS One doi:10.1371/journal.pone.0042640.

Mohan M, Nair S, Bhagwat A, Krishna TG, Yano M, Bhatia CR, Sasaki T (1997) Genome mapping,molecularmarkers andmarker-assisted selection in crop plants. Molecular Breeding doi:10.1023/a:1009651919792.

Morrell PL, Clegg MT (2007) Genetic evidence for a second domestication of barley (Hordeum vulgare) east of the Fertile Crescent. Proc Natl Acad Sci U S A doi:10.1073/pnas.0611377104.

Muñoz-Amatriaín M, Mascher M (2018) Sequence Diversity and Structural Variation. In: The Barley Genome. Switzerland: Springer.

Murata Y, Ma JF, Yamaji N, Ueno D, Nomoto K, Iwashita T (2006) A specific transporter for iron(III)-phytosiderophore in barley roots. Plant J doi:10.1111/j.1365-313X.2006.02714.x.

Muthayya S, Rah JH, Sugimoto JD, Roos FF, Kraemer K, Black RE (2013) The global hidden hunger indices and maps: an advocacy tool for action. PLoS One doi:10.1371/journal.pone.0067860.

Myles S, Peiffer J, Brown PJ, Ersoz ES, Zhang Z, Costich DE, Buckler ES (2009) Association mapping: critical considerations shift from genotyping to experimental design. Plant Cell doi:10.1105/tpc.109.068437.

Nice LM, Steffenson BJ, Brown-Guedira GL*, et al.* (2016) Development and Genetic Characterization of an Advanced Backcross-Nested Association Mapping (AB-NAM) Population of Wild x Cultivated Barley. Genetics doi:10.1534/genetics.116.190736.

Osborne BG (2006) Applications of near infrared spectroscopy in quality screening of early-generation material in cereal breeding programmes. Journal of near Infrared Spectroscopy doi:10.1255/jnirs.595.

Oury F-X, P B, M B-H*, et al.* (2003) Yield and grain protein concentration in bread wheat: A review and a study of multi-annual data from a French breeding program. Journal of Genetics and Breeding

Palmgren MG, Clemens S, Williams LE, Kramer U, Borg S, Schjorring JK, Sanders D (2008) Zinc biofortification of cereals: problems and solutions. Trends Plant Sci doi:10.1016/j.tplants.2008.06.005.

Pankin A, Altmüller J, Becker C, von Korff M (2018) Targeted re-sequencing reveals the genomic signatures of multiple barley domestications. New Phytol doi:10.1101/070078.

Pankin A, Campoli C, Dong X*, et al.* (2014) Mapping-by-sequencing identifies HvPHYTOCHROME C as a candidate gene for the early maturity 5 locus modulating the circadian clock and photoperiodic flowering in barley. Genetics doi:10.1534/genetics.114.165613.

Pasam RK, Sharma R, Walther A, Ozkan H, Graner A, Kilian B (2014) Genetic diversity and population structure in a legacy collection of spring barley landraces adapted to a wide range of climates. PLoS One doi:10.1371/journal.pone.0116164.

Pearce S, Tabbita F, Cantu D, Buffalo V, Avni R, Vazquez-Gross H, Zhao R, Conley CJ, Distelfeld A, Dubcovksy J (2014) Regulation of Zn and Fe transporters by the GPC1 gene during early wheat monocarpic senescence. BMC Plant Biol doi:10.1186/s12870-014-0368-2.

Pedas P, Ytting CK, Fuglsang AT, Jahn TP, Schjoerring JK, Husted S (2008) Manganese efficiency in barley: identification and characterization of the metal ion transporter HvIRT1. Plant Physiol doi:10.1104/pp.108.118851.

Pfeiffer WH, McClafferty B (2007) HarvestPlus: Breeding crops for better nutrition. Crop Science doi:10.2135/cropsci2007.09.0020IPBS.

Pham AT, Maurer A, Pillen K, Brien C, Dowling K, Berger B, Eglinton JK, March TJ (2019) Genome-wide association of barley plant growth under drought stress using a nested association mapping population. BMC Plant Biol doi:10.1186/s12870-019-1723-0.

Pillen K, Zacharias A, Leon J (2003) Advanced backcross QTL analysis in barley (Hordeum vulgare L.). Theor Appl Genet doi:10.1007/s00122-003-1253-9.

Pistorius R (1997) Scientists, plants and politics - A history of the plant genetic resources movement. Rome, Italy: International Plant Genetic Resources Institute.

Poets AM, Fang Z, Clegg MT, Morrell PL (2015) Barley landraces are characterized by geographically heterogeneous genomic origins. Genome Biol doi:10.1186/s13059-015-0712-3.

Pourkheirandish M, Hensel G, Kilian B*, et al.* (2015) Evolution of the Grain Dispersal System in Barley. Cell doi:10.1016/j.cell.2015.07.002.

Ritchie H, Roser M (2020) Micronutrient Deficiency. Our World in Data https://ourworldindata.org/micronutrient-deficiency

Rollins JA, Drosse B, Mulki MA, Grando S, Baum M, Singh M, Ceccarelli S, von Korff M (2013) Variation at the vernalisation genes Vrn-H1 and Vrn-H2 determines growth and yield stability in barley (Hordeum vulgare) grown under dryland conditions in Syria. Theor Appl Genet doi:10.1007/s00122-013-2173-y.

Saade S, Kutlu B, Draba V, Forster K, Schumann E, Tester M, Pillen K, Maurer A (2017) A donor-specific QTL, exhibiting allelic variation for leaf sheath hairiness in a nested association mapping population, is located on barley chromosome 4H. PLoS One doi:10.1371/journal.pone.0189446.

Saade S, Maurer A, Shahid M, Oakey H, Schmockel SM, Negrao S, Pillen K, Tester M (2016) Yield-related salinity tolerance traits identified in a nested association mapping (NAM) population of wild barley. Sci Rep doi:10.1038/srep32586.

Sannemann W, Huang BE, Mathew B, Leon J (2015) Multi-parent advanced generation inter-cross in barley: high-resolution quantitative trait locus mapping for flowering time as a proof of concept. Molecular Breeding doi:ARTN 8610.1007/s11032-015-0284-7.

Sannemann W, Lisker A, Maurer A*, et al.* (2018) Adaptive selection of founder segments and epistatic control of plant height in the MAGIC winter wheat population WM-800. BMC Genomics doi:10.1186/s12864-018-4915-3.

Schmalenbach I, Korber N, Pillen K (2008) Selecting a set of wild barley introgression lines and verification of QTL effects for resistance to powdery mildew and leaf rust. Theor Appl Genet doi:10.1007/s00122-008-0847-7.

Schmalenbach I, Pillen K (2009) Detection and verification of malting quality QTLs using wild barley introgression lines. Theor Appl Genet doi:10.1007/s00122-009-0991-8.

Schwarz G (1978) Estimating the Dimension of a Model. Annals of Statistics Vol 6, 461-464.

Segura V, Vilhjalmsson BJ, Platt A, Korte A, Seren U, Long Q, Nordborg M (2012) An efficient multi-locus mixed-model approach for genome-wide association studies in structured populations. Nat Genet doi:10.1038/ng.2314.

Septiningsih EM, Trijatmiko KR, Moeljopawiro S, McCouch SR (2003) Identification of quantitative trait loci for grain quality in an advanced backcross population derived from the Oryza sativa variety IR64 and the wild relative O. rufipogon. Theor Appl Genet doi:10.1007/s00122-003-1376-z.

Sharma R, Draicchio F, Bull H, Herzig P, Maurer A, Pillen K, Thomas WTB, Flavell AJ (2018) Genome-wide association of yield traits in a nested association mapping population of barley reveals new gene diversity for future breeding. J Exp Bot doi:10.1093/jxb/ery178.

Slafer GA (2003) Genetic basis of yield as viewed from a crop physiologist's perspective. Annals of Applied Biology doi:10.1111/j.1744-7348.2003.tb00237.x.

Song Q, Yan L, Quigley C*, et al.* (2017) Genetic Characterization of the Soybean Nested Association Mapping Population. Plant Genome doi:10.3835/plantgenome2016.10.0109.

Stein N, Muehlbauer GJ (2018) The Barley Genome. Switzerland: Springer.

Tanksley SD, Grandillo S, Fulton TM, Zamir D, Eshed Y, Petiard V, Lopez J, BeckBunn T (1996) Advanced backcross QTL analysis in a cross between an elite processing line of tomato and its wild relative L-pimpinellifolium. Theor Appl Genet 92, 213-224.

Tanksley SD, Nelson JC (1996) Advanced backcross QTL analysis: a method for the simultaneous discovery and transfer of valuable QTLs from unadapted germplasm into elite breeding lines. Theor Appl Genet doi:10.1007/BF00223376.

Tester M, Langridge P (2010) Breeding technologies to increase crop production in a changing world. Science doi:10.1126/science.1183700.

Tetko IV, Livingstone DJ, Luik AI (1995) Neural-Network Studies .1. Comparison of Overfitting and Overtraining. Journal of Chemical Information and Computer Sciences doi:10.1021/ci00027a006.

Tian F, Bradbury PJ, Brown PJ, Hung H, Sun Q, Flint-Garcia S, Rocheford TR, McMullen MD, Holland JB, Buckler ES (2011) Genome-wide association study of leaf architecture in the maize nested association mapping population. Nat Genet doi:10.1038/ng.746.

Tuinstra MR, Ejeta G, Goldsbrough PB (1997) Heterogeneous inbred family (HIF) analysis: a method for developing near-isogenic lines that differ at quantitative trait loci. Theor Appl Genet doi:10.1007/s001220050654.

Uauy C, Brevis JC, Dubcovsky J (2006) The high grain protein content gene Gpc-B1 accelerates senescence and has pleiotropic effects on protein content in wheat. J Exp Bot doi:10.1093/jxb/erl047.

UN (2019) Global Sustainable Development Report 2019: The Future is Now – Science for Achieving Sustainable Development. New York: United Nations.

van der Graaff E, Schwacke R, Schneider A, Desimone M, Flugge UI, Kunze R (2006) Transcription analysis of arabidopsis membrane transporters and hormone pathways during developmental and induced leaf senescence. Plant Physiol doi:10.1104/pp.106.079293.

Vatter T, Maurer A, Kopahnke D, Perovic D, Ordon F, Pillen K (2017) A nested association mapping population identifies multiple small effect QTL conferring resistance against net blotch (Pyrenophora teres f. teres) in wild barley. PLoS One doi:10.1371/journal.pone.0186803.

Vatter T, Maurer A, Perovic D, Kopahnke D, Pillen K, Ordon F (2018) Identification of QTL conferring resistance to stripe rust (Puccinia striiformis f. sp. hordei) and leaf rust (Puccinia hordei) in barley using nested association mapping (NAM). PLoS One doi:10.1371/journal.pone.0191666.

Vellido A, Lisboa PJG, Vaughan J (1999) Neural networks in business: a survey of applications (1992-1998). Expert Systems with Applications doi:10.1016/S0957-4174(99)00016-0.

Velu G, Ortiz-Monasterio I, Cakmak I, Hao Y, Singh RP (2014) Biofortification strategies to increase grain zinc and iron concentrations in wheat. Journal of Cereal Science doi:10.1016/j.jcs.2013.09.001.

Vilhjalmsson BJ, Nordborg M (2013) The nature of confounding in genome-wide association studies. Nat Rev Genet doi:10.1038/nrg3382.

von Korff M, Wang H, Leon J, Pillen K (2006) AB-QTL analysis in spring barley: II. Detection of favourable exotic alleles for agronomic traits introgressed from wild barley (H. vulgare ssp. spontaneum). Theor Appl Genet doi:10.1007/s00122-006-0223-4.

Waters BM, Sankaran RP (2011) Moving micronutrients from the soil to the seeds: genes and physiological processes from a biofortification perspective. Plant Sci doi:10.1016/j.plantsci.2010.12.003.

Waters BM, Uauy C, Dubcovsky J, Grusak MA (2009) Wheat (Triticum aestivum) NAM proteins regulate the translocation of iron, zinc, and nitrogen compounds from vegetative tissues to grain. J Exp Bot doi:10.1093/jxb/erp257.

Weier D, Thiel J, Kohl S, Tarkowska D, Strnad M, Schaarschmidt S, Weschke W, Weber H, Hause B (2014) Gibberellin-to-abscisic acid balances govern development and differentiation of the nucellar projection of barley grains. J Exp Bot doi:10.1093/jxb/eru289.

Wessells KR, Brown KH (2012) Estimating the global prevalence of zinc deficiency: results based on zinc availability in national food supplies and the prevalence of stunting. PLoS One doi:10.1371/journal.pone.0050568.

WHO (2014) Second International Conference on Nutrition - Conference Outcome Document: Rome Declaration on Nutrition. Retrieved from http://www.fao.org/about/meetings/icn2/en/

WHO, FAO (2018) The Nutrition Challenge - Food System Solutions.

Würschum T, Liu W, Gowda M, Maurer HP, Fischer S, Schechert A, Reif JC (2012) Comparison of biometrical models for joint linkage association mapping. Heredity doi:10.1038/hdy.2011.78.

Xu Y, Li PC, Yang ZF, Xu CW (2017) Genetic mapping of quantitative trait loci in crops. Crop Journal doi:10.1016/j.cj.2016.06.003.

Yan L, Fu D, Li C, Blechl A, Tranquilli G, Bonafede M, Sanchez A, Valarik M, Yasuda S, Dubcovsky J (2006) The wheat and barley vernalization gene VRN3 is an orthologue of FT. Proc Natl Acad Sci U S A doi:10.1073/pnas.0607142103.

Yan L, Loukoianov A, Blechl A, Tranquilli G, Ramakrishna W, SanMiguel P, Bennetzen JL, Echenique V, Dubcovsky J (2004) The wheat VRN2 gene is a flowering repressor down-regulated by vernalization. Science doi:10.1126/science.1094305.

Yan L, Loukoianov A, Tranquilli G, Helguera M, Fahima T, Dubcovsky J (2003) Positional cloning of the wheat vernalization gene VRN1. Proc Natl Acad Sci U S A doi:10.1073/pnas.0937399100.

Youssef HM, Eggert K, Koppolu R*, et al.* (2017) VRS2 regulates hormone-mediated inflorescence patterning in barley. Nat Genet doi:10.1038/ng.3717.

Yu J, Holland JB, McMullen MD, Buckler ES (2008) Genetic design and statistical power of nested association mapping in maize. Genetics doi:10.1534/genetics.107.074245.

Yu J, Pressoir G, Briggs WH*, et al.* (2006) A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. Nat Genet doi:10.1038/ng1702.

Zamir D (2001) Improving plant breeding with exotic genetic libraries. Nat Rev Genet doi:10.1038/35103590.

Zeng ZB (1994) Precision mapping of quantitative trait loci. Genetics 136, 1457-1468.

Zhang H, Mittal N, Leamy LJ, Barazani O, Song BH (2017) Back into the wild-Apply untapped genetic diversity of wild relatives for crop improvement. Evol Appl doi:10.1111/eva.12434.

Zohary D, Hopf M (1994) Domestication of plants in the Old World: the origin and spread of cultivated plants in West Asia, Europe, and the Nile Valley. Oxford and New York: Clarendon Press.

7 Supplementary files

7.1 Original paper supplements

Contrasting genetic regulation of plant development in two European environments revealed by wild barley nested association mapping

Download link: https://doi.org/10.1093/jxb/ery002

Figure S1. Climate diagrams of both locations during the experimental period.

Figure S2. Frequency distribution of BLUEs for all traits, plotted as density histograms.

Figure S3. Frequency distribution of BLUEs for SHO, HEA and MAT as a function of day length.

Figure S4. Family-specific effects of all discussed QTLs.

Figure S5. Origin of HEB-25 donors in regard to QTL effects of exotic Vrn-H2 alleles on shooting.

Table S1. ANOVA results of phenotypic data within location.

Table S2. GWAS output of all significant SNPs across all cross-validation (CV) runs.

Table S3. Family-specific effect estimation based on the cumulating method.

Table S4. List of QTLs, extreme family effects and candidate genes calculated for eight traits.

Genome-wide association of yield traits in a nested association map-ping population of barley reveals new gene diversity for future breeding

Download link: https://doi:10.1093/jxb/ery178

Figure S1. Correlation matrix of the yield traits across locations and treatments.

Figure S2. Variance components of the yield traits displayed in percentages.

Figure S3. Box plot of thousand grain weight (TGW in g).

Figure S4. Box plot of grain area (GA in mm2).

Figure S5. Box plot of grain length (GL in mm).

Figure S6. Box plot of grain width (GW in mm).

Figure S7. Box plot of grain roundness (width to length ratio in %) (GR).

Figure S8. Box plot of grains per ear (GPE in number).

Figure S9. Box plot of yield (YLD in g, only from Halle).

Figure S10. Box plot of standard error of grain width (SE_GW).

Figure S11. Box plot of standard error of grain length (SE_GL only from Dundee).

Figure S12. Genome-wide association scans of yield traits.

Figure S13. LD of mapped SNP markers in the HEB-25 population.

Figure S14. Distribution of thousand grain weight (TGW) QTLs across barley chromosomes.

Figure S15. Distribution of grain area (GA) QTLs across barley chromosomes.

Figure S16. Distribution of grain length (GL) QTLs across barley chromosomes.

Figure S17. Distribution of grain width (GW) QTLs across barley chromosomes.

Figure S18. Distribution of grains per ear (GPE) QTLs across barley chromosomes.

Figure S19. Distribution of grain roundness (GR) QTLs across barley chromosomes.

Figure S20. Distribution of grain yield (YLD) QTLs from Halle across barley chromosomes.

Figure S21. Distribution of standard error of grain width (SE_GW) QTLs across barley chromosomes.

Figure S22. Distribution of standard error of grain length (SE_GL) QTLs across barley chromosomes.

Table S1. Description of yield traits, locations, and years recorded.

Table S2. Summary statistics of yield-related traits.

Table S3. ANOVA results of phenotypic data within locations.

Table S4. GWAS results for grain traits.

Table S5. Cross-validation results for grain traits.

Table S6. Family-specific effect estimations based on the cumulating method.

Genetic dissection of grain elements predicted by hyperspectral imaging associated with yield-related traits in a wild barley NAM population

Download link: https://doi.org/10.1016/j.plantsci.2019.05.008

Figure S1. Frequency distribution of all investigated traits across years.

Figure S2. Heritability and estimated variance components across treatment.

Figure S3. GWAS results for yield and yield-related traits.

Figure S4. Family-specific effects of all the QTLs discussed.

Figure S5. Effect of thr-1 on grains.

Table S1. BLUEs of all trait for each HEB line across years for N-level and across N-level.

Table S2. ANOVA results of phenotypic data.

Table S3. GWAS output of all significant SNPs across all cross-validation (CV) runs.

Table S4. Number of QTLs, explained phenotypic variance and prediction ability per element for each treatment.

Table S5. Family-specific effect estimation based on the cumulating method.

Table S6. List of QTL-Regions across treatment, which were discussed in the paper

Table S7. List of candidate genes, which are able to affect grain nutrient content.

7.4 Abbreviations

ABA	Abscisic acid
ANOVA	analysis of variance
BLUEs	Best linear unbiased estimates
cM	centiMorgan
CV	cross-validation
DNA	deoxyribonucleic acid
DR	detection rate
FAO	Food and Agriculture Organization of the United Nations
GWAS	genome-wide association study
HEB	Halle exotic barley
HIF	heterogeneous inbred family
HSI	Hyperspectral imaging
LM	Linkage mapping
LSMEANS	least squares means
MAGIC	multi-parent advanced generation inter-cross
MAS	marker-assisted selection
NAM	nested association mapping
Ppd	Photoperiod gene
QTL	quantitative trait locus/loci
RFLP	Restriction Fragment Length Polymorphism
SNP	Single Nucleotide Polymorphism
Vrn	Vernalization gene
WHO	World Health Organization

7.5 List of figures

Acknowledgements

At this point in time I'd like to thank everyone who has supported me during my entire doctoral studies and I would like to thank the following persons in particular:

First of all, I would like to thank Klaus Pillen for providing me the interesting projects and for trusting in my ability to carry out the work properly. His open and confident manner encouraged me in my way of working.

Furthermore, I would like to thank Andreas Maurer, who with his excellent statistical understanding has contributed significantly to the success of my work. The cooperation with Andreas has always been very productive and was also characterized by creative new ideas. I also appreciate him outside of work.

My thanks also go to all the staff of the Chair of Plant Breeding. Among others the colleagues on the Kühnfeld, without them the field trial would not have been possible or my other PhD fellow sufferers from the office, who went through thick and thin with me.

Last but not least I would like to thank my family, my friends and my wonderful girlfriend Hanna who always supported me when I needed it most.

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.

Halle (Saale), 10.06.2020

_____

Unterschrift / *Signature*

Curriculum Vitae

Paul Herzig

01.04.2020

Education

3/2018 – Present	Martin Luther University Halle - PhD student Research associate at the chair of plant breeding BMBF-Initiative "BARLEY BIODIVERSITY - Increasing biodiversity and productivity of barley cultivars using a sensor-based high-throughput phenotyping system in the field."
07/2014 – 06/2017	Martin Luther University Halle - PhD student Research associate at the chair of plant breeding ERA-CAPS project "*BARLEY-NAM – Locating exotic genes that control agronomic traits under stress in a wild barley nested association mapping (NAM) population*"
10/2008 – 04/2014	Ernst Moritz Arndt – University Greifswald Diplom Biology Diploma thesis: "*Divergent growth of the white spruce (Picea glauca) in climate change - the influence of genetic effects"*
1999 – 2008	Hansa-Gymnasium Stralsund

Publication

Paul Herzig, Andreas Backhaus, U. Seiffert, N. von Wirén, K. Pillen and A. Maurer (2019) "*Genetic dissection of grain elements predicted by hyperspectral imaging associated with yield-related traits in a wild barley NAM population*", **Plant Science**, doi:10.1016/j.plantsci.2019.05.008

Rajiv Sharma, Fulvia Draicchio, H. Bull, **P. Herzig**, A. Maurer, K. Pillen, W.T.B. Thomas and A.J. Flavell (2018) "*Genome-wide association of yield traits in a nested association mapping population of barley reveals new gene diversity for future breeding.*", **Journal of Experimental Botany**, doi:10.1093/jxb/ery178

Paul Herzig, Andreas Maurer, V. Draba, R. Sharma, F. Draicchio, H. Bull, L. Milne, W.T.B. Thomas, A.J. Flavell and K. Pillen (2018) "*Contrasting genetic regulation of plant development in wild barley grown in two European environments revealed by nested association mapping*.", **Journal of Experimental Botany**, doi:10.1093/jxb/ery002

Pascal Eusemann, **Paul Herzig**, M. Kieß, S. Ahlgrimm, P. Herrmann, M. Wilmking, M. Schnittler, (2015) "*Three microsatellite multiplex PCR assays allowing high resolution genotyping of white spruce, Picea glauca.*", **Silvae Genetica** 63, doi: 10.1515/sg-2014-0029

Conference contributions

03/2019, 19th Conference of the Genome Research Working Group of the German Plant Breeding Association (GPZ) in Hohenheim, Germany. Poster: *Genetic dissection of ingredients in grains determined by hyperspectral imaging (HSI) associated with yield related traits in a wild barley NAM population.*

06/2016, 12th International Barley Genetics Symposium in Minneapolis, USA. Poster: *Locating exotic genes, which control yield component traits under contrasting nitrogen supply in the wild barley NAM population HEB-25.*

09/2015, 18th Conference of the Genome Research Working Group of the German Plant Breeding Association (GPZ) in Düsseldorf, Germany. Poster: *Characterization of agronomic traits and locating exotic genes, which control agronomic traits under contrasting nitrogen supply in the wild barley NAM population HEB-25.*

09/2014, Main Conference of the Genome Research Working Group of the German Plant Breeding Association (GPZ) in Kiel, Germany. Poster: *Characterization of agronomic traits under contrasting nitrogen supply in the wild barley NAM population HEB-25.*

Halle (Saale), 10.06.2020

_____

Signature