Genome-wide analysis identified candidate genes underlying mineral concentrations in wheat grains

Dissertation
zur Erlangung des
Doktorgrades der Agrarwissenschaften (Dr. agr.)

Der

Naturwissenschaftlichen Fakultät III

Agrar- und Ernährungswissenschaften,

Geowissenschaften und Informatik

der Martin-Luther-Universität Halle-Wittenberg

vorgelegt von

Frau Dalia Zakaria Saleh Alomari

Geb. am 19.05.1984 in Dammam, Saudi Arabia

Gutachter:
Prof. Dr. Klaus Pillen (Halle)
Prof. Dr. Volker Mohler (Munich)

Verteidigt am:
01st July 2019

Table of Contents

Acknowledgements

First of all, I would like to express my deepest appreciation, sincere gratitude and special thanks to my advisor, Dr. Marion Röder for giving me the opportunity to conduct my research in her lab with scientific freedom and for her guidance, academic support, kindness, encouragements, critical reading and valuable suggestions on the manuscripts and the thesis. I have been truly blessed to work with Dr. Röder over the past 3.3 years.

I am grateful to Prof. Dr. Klaus Pillen for accepting me as a Ph.D. student.

I am thankful to Prof. Dr. Nicolaus von Wirén for the collaboration in this project and for his help and advices in all the co-authors of the publications in this thesis and for Dr. Kai Eggert and Yudelsy Antonia Tandron Moya for the technical assistance and collaboration.

I am thankful to my group members Ellen Weiß, Sonja Allner, Annette Heber, Rosemarie Czihal for their help in data collection, support and technical assistances. I would like to extend my thanks to my colleague in Gene and Genome mapping group Quddoos Muqaddasi for his help especially at the beginning of my project and my friend Rasha Tarawneh for her support.

I would like to thank Dr. Britt Leps for her encouragement, endless help and social supports.

My family has been unlimited and extremely supportive of everything that I have set out to do in my life. Words cannot describe nor articulate

my sincere indebtedness and love to my parents Zakaria Alomari and Khitam Ghannam who raised and educated me and give me the continuous encouragements and support during each educational steps and making me the person who I am today. My brothers Ashraf and Ahmad, my sisters Tasneem and Raneem and my brother in law Mohannad are thanked for all their help in driving me forward to achieve my goals. My thanks are extended to parent-in-laws, brothers and sisters in law for their support.

I could not have completed this thesis if not for my dearest husband, Ahmad Alqudah. From the beginning of my project until now, you gave me an endless help, advices, support and encouragement. I would like to thank you truly for taking care of me and our kids. I would like to be in awe of my daughter Yomna and son Abdallah, for refreshing me from my tiresome research work with their sweet smile and affectionate hug.

At the last but not least, I praise Allah for giving me the strength to start and finish this work. There were many times that I called on his name for help, comfort, and guidance; he was always right at my side.

This dissertation is dedicated to my parents, my Husband and my sweet kids Yomna, Abdallah and Mila.

List of Abbreviations

AFLP	Amplified fragment length polymorphisms
BLUEs	Best linear unbiased estimates
bp	Base pair
BZIP	Basic leucine zipper
Ca	Calcium
CAPS	Cleavable amplified polymorphic sequences
CGIAR	Consultative Group on International Agricultural Research
CIAT	Center for Tropical Agriculture
Cu	Copper
DW	Dry weight
FAO	Food and Agriculture Organization
Fe	Iron
GBLUP	Genomic best linear unbiased predictions
GEBVs	Genomic estimated breeding values
GMO	Genetically modified organisms
GP	Genomic prediction
GS	Genomic selection
GWAS	Genome-wide association studies
H^2	Broad-sense heritability
HOX4	Homeobox-leucine zipper protein
I	Iodine
ICP-OES	Inductively coupled plasma optical emission spectrometry
IFPRI	International Food Policy Research Institute
IWGSC	International Wheat Genome Sequencing Consortium
LD	Linkage disequilibrium
MAF	Minor allele frequency
MAPK	Mitogen-activated protein kinase
MAS	Marker-assisted selection
Mg	Magnesium
MLM	Mixed linear model
MLM	Mixed linear model
MTAs	Marker-trait associations
NA	Nicotianamine
NAM	No apical meristem
NIL	Near isogenic lines
PCA	Principal component analysis
PoU	Prevalence of undernourishment
QTL	Quantitative trait loci
RAPD	Random amplified polymorphic sequences
RFLP	Restriction fragment length polymorphisms
RIL	Recombinant inbred lines
RRBLUP	Ridge regression best linear unbiased predictions
$r2$	Squared allele frequency correlation

Se	Selenium
SNPs	Single nucleotide polymorphisms
SSCP	Single-strand conformation polymorphisms
SSR	Microsatellite or simple sequence repeat length polymorphisms
SWAP	Suppressor-of-white-apricot/surp
TF	Transcription factor
TKW	Thousand kernel weight
Zn	Zinc

1 General Introduction

1.1 Wheat

Wheat is one of the oldest and most significant staple crops in the world. Currently, wheat has a great social, economic and biological value since it is a major source of food for humans and the most widely grown cereal crop with diverse uses such as food, animal feed and biofuel that makes it as a key factor in the world crop trade (IDRC, 2015).

Genetically, wheat has different ploidy levels; its progenitors belong to the *Poaceae* family, genus *Triticum (T. urartu* and *T. monococcum,* AA genome; 2n = 2x = 14*)* and *Aegilops (Ae. speltoides*, BB genome; 2n = 2x = 14*)* and due to hybridization of diploid wheat, the tetraploid durum wheat (*T. turgidum ssp. durum,* AABB genome; 2n = 4x = 28) was formed in the Fertile Crescent around 500,000 years ago (Salamini et al., 2002; Matsuoka, 2011; Marcussen et al., 2014). Durum wheat is mainly used in making pasta (spaghetti and macaroni), burghul and couscous. Hexaploid bread wheat developed by hybridization of tetraploid wheat (AABB) with diploid Aegilops species, (*Ae. tauschii*, DD genome; 2n = 2x = 14) to form *Triticum aestivum* (AABBDD genome; 2n = 6x = 42) which is one of the most important cereal grain crops used to make bread (Kihara, 1924; International Wheat Genome Sequencing, 2014; Rasheed et al., 2018).

The first appearance of hexaploid wheat was around 9,000 years ago in the Fertile Crescent (Shewry, 2009) and nowadays it is the third produced crop after rice and maize. Wheat is mostly grown by Asian and Europeans countries (Figure 1a). According to The United Nations Food and Agriculture Organization (FAO), the global annual wheat production was dramatically increased over the last 25 year and reached around 775 million tons at 2017 compared to around 500 million tons in 1994 (Figure 1b) (FAOSTAT, 2018). The increment in wheat yield production was mostly due to the efficient progress in breeding programs for producing high yielding cultivars since the harvested area did not significantly change from 1994 (215 million hectare) to 2017 (218 million hectare, Figure 1b) (FAOSTAT, 2018). Nowadays, hexaploid wheat grain production constitutes 95% of the produced wheat and the remaining 5% are tetraploid wheat (Shewry, 2009).

The global distribution of bread wheat is attributed to its unique characteristic to form a gluten network that is composed of two protein classes: glutenin and gliadin proteins and these

characteristics enable wheat flour to have superior rheological properties of the resulting dough and then high quality baked products like bread, cookies, and cakes (Ando et al., 2002).

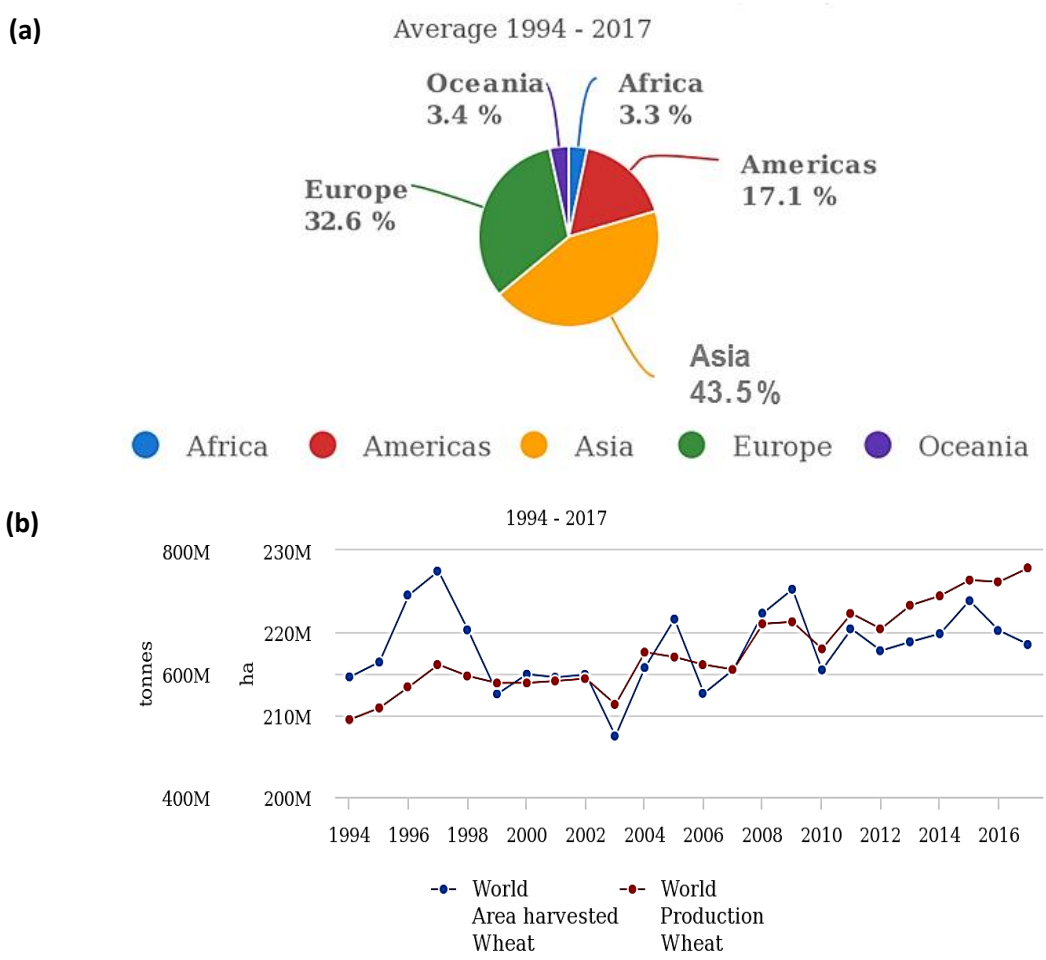


Figure1. (a) Wheat production in different continents, (b) wheat global harvested area and production in the world. FAOSTAT | FAO Statistics Division 2018 (FAOSTAT, 2018).

1.2 Grain anatomy

Wheat as a member of the grass family *Gramineae* (synonym to *Poaceae*) produces dry, one-seeded fruits that are called caryopsis, grain or kernel. Wheat grain is the nutrients' storage place which is coated with a hull (husk) which is separated from the grain during threshing (Figure 2). Grain structural components are divided into three major parts: 80-85% endosperm (the largest tissue in the wheat grain), 13-17% bran and 2-3% germ (Yadav, 2011). The endosperm consists of starch granules that are embedded in a protein matrix in addition to a very small amount of minerals. The starchy endosperm is considered as the main source of white flour which results after the milling process. The outer layer of the grain is called bran which is consisting of several

layers including nucellar tissue, testa and pericarp and its function is grain protection. More than half of the bran consists of fibers as well as a number of minerals and vitamin B. Bran is considered as a beneficial source for human nutrition and therefore its integration in food helps to improve health and to prevent some diseases, such as colon cancer and cardiovascular diseases (Björck et al., 2012; Balandrán-Quintana et al., 2015). The germ or embryo is the main source of fat but also contains a small amount of protein, sugar, and minerals (Delcour and Hoseney, 2010). Bran and its layers including aleurone layer are removed during the first step of the milling process (Sramkova et al., 2009) that leads to loss of nutrient value of wheat flour (Sramkova et al., 2009). In contrast, whole wheat flour that contains the starchy endosperm, germ and bran provides high nutritious flour.

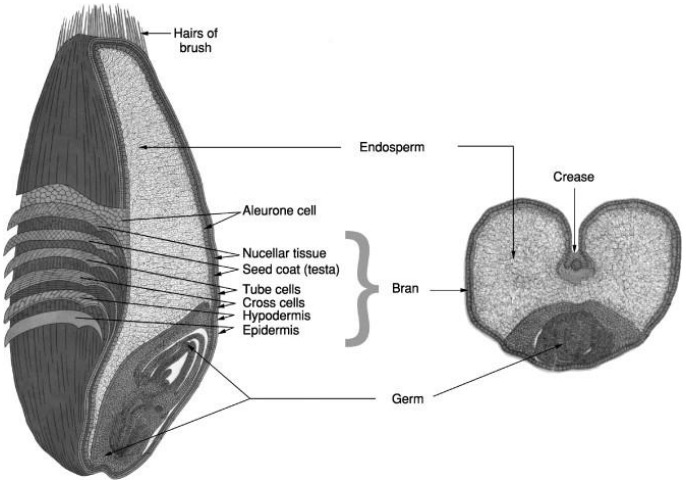


Figure 2. Wheat grain structure and layers (source: Encyclopedia of Food Sciences and Nutrition).

1.3 Wheat nutrients and mineral distribution in the grain

Wheat grain contains 78% carbohydrates, 15% proteins, 2.1% fats and 2.1-1.6% minerals (Yadav, 2011). As well, bread wheat serves as a staple food for 35% of the world's population and even more for the population who relays basically on bread for surviving (Paux et al., 2008; Shewry, 2009). Wheat is considered an insufficient source of minerals like iron (Fe), zinc (Zn) and calcium (Ca) (Tang et al., 2008). Fe is found in the outer layer of the bran and in the aleurone (Singh et al., 2013; De Brier et al., 2015) and Zn is located mainly in the aleurone and embryo (Ozturk et al., 2006) with a very small concentration of Zn in the endosperm (Ozturk et al., 2006). The highest Ca concentration is distributed in the outer layer of the bran (De Brier et

al., 2015). Wheat grain is often consumed after milling process, which leads to simultaneous removal of the embryo and bran layer (pericarp, seed coat, nucellar epidermis and aleurone) (Barron et al., 2007; Delcour and Hoseney, 2010). That means loss of mineral-rich parts and leaving the starchy endosperm (the poor source of minerals) to produce white flour, therefore the overall minerals concentration is decreased from 1.6 to 0.4% after milling (Fujino et al., 1996; De Brier et al., 2016). Accordingly, whole wheat flour is found to have a higher concentration of the minerals related to the wheat bran and aleurone fractions; therefore it is better for the human health and can remarkably improve the nutritive value of the daily dietary intake (Lopez et al., 2002; Björck et al., 2012; De Brier et al., 2015).

1.4 Minerals deficiency and consequences on plant nutrition and human health

In the same context, the minerals that are most frequently lacking in human diet are Fe, Zn and iodine (I), in addition to Ca, magnesium (Mg), copper (Cu) and selenium (Se) that can be deficient for some population who is relying only on the staple food consumption without diet diversification. The State of Food Security and Nutrition in the World 2018 has mentioned that about 821 million people (~11% of the total world's population) worldwide were malnourished at 2017 (Figure 3, (FAO, 2018)). According to FAO (2018), Africa is still the continent with the highest prevalence of undernourishment (PoU) in the world with almost 21% of undernourished people (more than 256 million people). The situation is also deteriorating in South America, where around 5.0% of the population (21.5 million) were undernourished at 2017 compared to 4.7% (19.3 million) in 2014 (FAO, 2018). In general, the increment of PoU in the world may be related to climate changes and economic slowdowns in addition to persistent political instabilities leading to worsened food security. Therefore, more breeding strategies focusing on improving the yield and yield stability to overcome the undernourishment problem and eradicating hunger in the world are imperative.

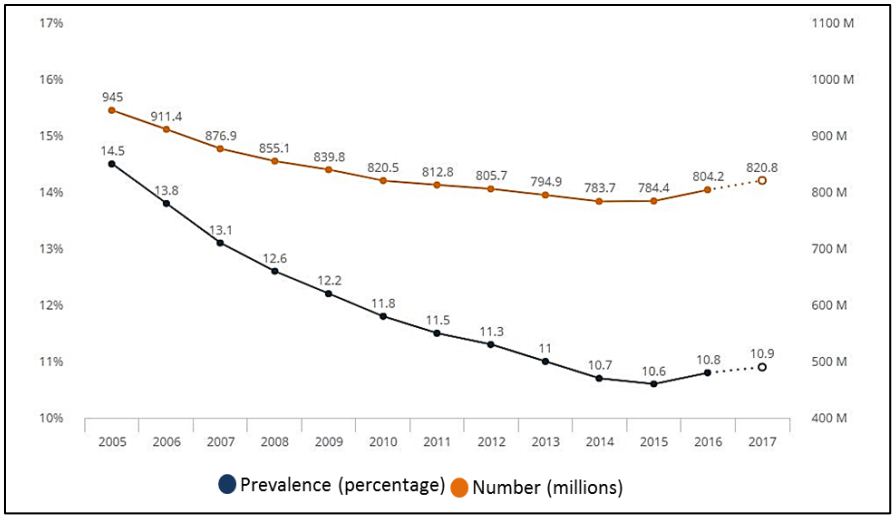


Figure 3. The number of undernourished people in the world has been on the rise since 2014, reaching an estimated 821 million in 2017. FAO | The State of Food Security and Nutrition in the World 2018 (FAO, 2018).

The most prevalent mineral deficiency occurrence is related to Fe, Zn or Ca deficiency where around 2 billion people are suffering (Figure 4 and Figure 5). Also it is quite obvious that the population who changed from bean-based diet to cereal-based diet has a high prevalence of Ca deficiency because of the chemical structure of the cell wall of cereal (monocot) which has low Ca concentration in contrast to bean species (eudicot) under the same conditions (Figure 6) (White and Broadley, 2005). At the same time, these minerals (Fe, Zn and Ca) are essential not only for human health but also for plant nutrition. Zn is vital for plant growth due to its inclusion in several enzymatic reactions and protein synthesis as well as in metabolic processes, regulation of auxin synthesis, pollen formation and oxidation-reduction reactions (Cakmak, 2000; 2002). Its deficiency resulted in the development of abnormalities in plants such as stunted growth, chlorosis, smaller leaves and spikelet sterility; as well, Zn deficiency adversely affects the plant yield, increases the sensitivity of plants to injury by high light or temperature and to infection by fungal diseases (Cakmak, 2000; 2002). On the other hand, insufficient dietary intake of Zn and low absorption of Zn in the human intestine may lead to Zn deficiency. Zn deficiency symptoms include hair and memory loss, skin problems, and weakness in body muscles. During pregnancy, it causes stunted brain development of the fetus (Lukaski, 2004; Shenkin, 2006).

Fe plays a crucial role in the plant metabolic processes such as DNA synthesis, respiration, and photosynthesis; also it is essential for the activation of several metabolic pathways and acts as a co-factor for many enzymes. When plants absorb the low amount of Fe from the soil, this leads to Fe chlorosis which is characterized by yellow pigments on the plants leaves (Rout and Sahoo, 2015). Fe deficiency also affects human health by causing retardation in the physical growth and affecting the motoric development, leading to fatigue with low productivity, as well Fe deficiency is the main cause of anemia (Bouis, 2002; 2007).

Ca acts as strengthening agent for the structure of plant cell wall, plant architecture, quality, yield formation and as a secondary messenger for different signals, while its deficiency makes the plant more sensitive to biotic and abiotic stresses (Dayod et al., 2010). Adequate Ca intake is essential for human health especially during adolescence, because it is critical in reducing the rate of bone loss, rickets and osteoporosis. However, lower Ca intake provokes health risks such as hypocalcemia, hypertension, colorectal cancer as well as bone weakness and fractures accompanied with aging (Centeno et al., 2009; Dayod et al., 2010; Piste et al., 2012).

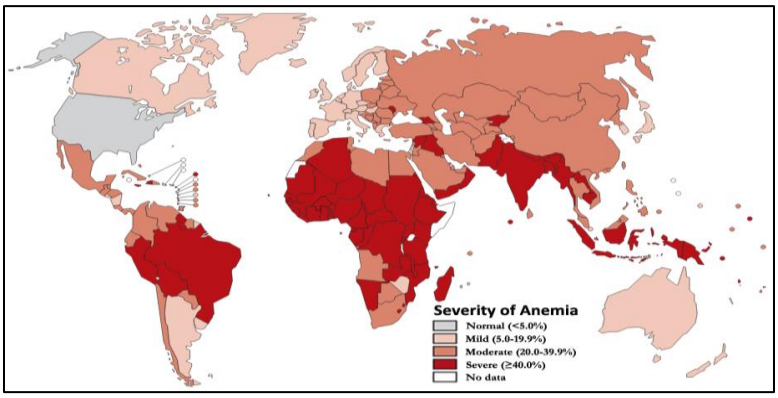


Figure 4. Iron (Fe) deficiency and distribution in the world (WHO, 2011).

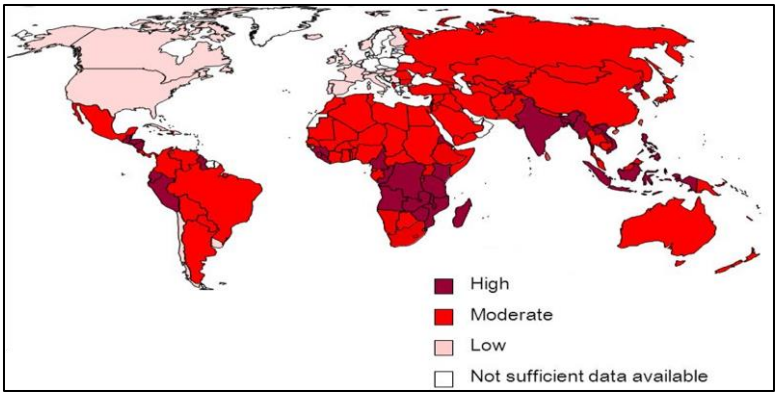


Figure 5. Worldwide Zinc (Zn) deficiency prevalence (Duffner et al. (2014).

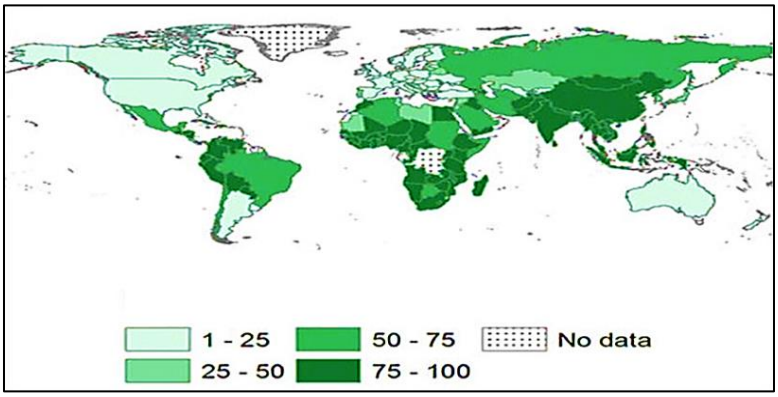


Figure 6. Calcium (Ca) deficiency risk worldwide (Kumssa et al. (2015))

1.5 Biofortification

Biofortification or biological fortification is a strategy for improving the nutritional quality including vitamin and mineral density in the crop's edible parts by breeding, agronomy and genetic modification (Bouis et al., 2011). Use of traditional short term strategies like supplementation, food fortification or dietary diversification programs, that may help in overcoming malnutrition problems, is not always a practical solution particularly in the developing countries where poverty is spreading and people earn under US$2 per day (White and Broadley, 2005; Zhu et al., 2007). Therefore, wheat is among the targeted crops in biofortification programs which can be biofortified by using the previously mentioned methods (agronomical practices, transgenic modification and breeding programs) (Zhu et al., 2007; Garg et al., 2018). Wheat agronomic biofortification for Fe was applied by the inclusion of Fe in the foliar urea fertilizers instead of using Fe fertilizers due to the low mobility of Fe in the soil.

Therefore, Fe fertilizers have little impact on Fe levels in the plants. For that reason, using a foliar spray of $FeSO_4$ or iron chelates enhances the uptake and absorption of Fe by the plant and returns with a positive effect on iron accumulation in the plant (Aciksoz et al., 2011). In contrast, Zn is mobile in the soil and applying the Zn fertilizers such as $ZnSO_4$ led to increasing grain Zn concentration and returned with positive consequences on human health (Cakmak, 2008). Selenium and iodine fertilizers are also mobile in soil and plants; therefore, the application of these fertilizers was successful in increasing the values of these minerals in the plant (Garg et al., 2018). Agronomical biofortification by the supplementation of fertilizers is the simplest way to improve the quality and nutrients of the resulting grains or fruits. Agronomical biofortification strategy can be directly applied but still, this method has a lot of limitations due to the differences in soil compositions and properties in addition to the mineral mobility and accumulation sites within the plant (Zhu et al., 2007; Garg et al., 2018). Furthermore, it is not always possible to target the micronutrient into the plant edible parts, because the nutrients may accumulate in the non-edible portions of plants like leaves. Therefore, agronomic biofortification is only successful for certain minerals and specific plant species (Zhu et al., 2007).

Another biofortification approach is known as conventional breeding which can be applied if there is genetic diversity among the genotypes related to the desired characteristics of nutrient and agronomic traits. Then breeders can utilize this variation to select parental lines with the desired traits for the crossing which requires several generations to produce varieties with the desired traits (Garg et al., 2018). This method is sustainable, convenient and not costly for the consumer. Therefore, several international organizations have initiated breeding programs to improve the nutritional quality of crops. For example, the Health Grain Project (Health grain project) was established in the European Union (2005–2010) gathering several partners from different countries to promote human health by introducing safe and high quality of cereal foods. Another international organization is called HarvestPlus (Harvestplus) which is initiated by cooperation between the Consultative Group on International Agricultural Research (CGIAR) along with the International Center for Tropical Agriculture (CIAT) and the main branch is located in USA, while other branches are situated in Africa, Asia, and Latin America. The International Food Policy Research Institute (IFPRI) was founded in USA as the main branch with other branches distributed in Africa Region, South Asia Region and East and Central Asia. This organization is focusing on biofortification of three key nutrients: vitamin A, iron and zinc

in wheat, rice, maize, cassava, pearl millet, beans and sweet potato in Asia and Africa. Several high Zn wheat varieties were released by HarvestPlus with higher Zn content, six wheat varieties were released in India (2014) and then four wheat varieties in Pakistan (2015) (Garg et al., 2018). Another wheat variety with high Zn has been released recently by Punjab Agricultural University, India and another wheat variety with high Zn and Fe content has been released by Indian Institute of Wheat and Barley Research in India (Garg et al., 2018).

The transgenic approach is one of the alternative options for developing biofortified crops and researchers turned into this approach, if there is limited or no genetic variation of nutrients content among plant varieties or when a particular micronutrient does not naturally exist in crops (Zhu et al., 2007). This approach is based on transferring the desirable genes from one plant species to another plant species. One of the successful examples of the transgenic approach is enhancing the Fe content in wheat by the expression of a ferritin gene from soybean (Garg et al., 2018). Meanwhile, the biofortification approach has a number of limitations because of the low acceptability of genetically modified (GMO) food or crops by the farmers and people. Another limitation of GMOs is that countries have complicated and expensive systems with long regulation processes for the acceptance and commercialization of transgenic crops. For example, golden rice which is the most popular transgenic crop is still not available for farmers due to the non-existence of governmental approval for releasing it (Garg et al., 2018).

1.6 Genome-wide analyses

Marker-assisted selection (MAS) is a component of the new discipline of molecular breeding that based on the wide availability of DNA markers to detect the allelic variation linked with genes that underlie a trait (phenotype) of interest (Collard Bertrand and Mackill David, 2008). MAS is used in plant improvement since the 1990s and aimed to accelerate the breeding process by studying the genetic basis of the phenotype and the relationship between genotype and phenotype (Pérez-de-Castro et al., 2012). The development of molecular markers with different molecular techniques such as restriction fragment length polymorphisms (RFLP), amplified fragment length polymorphisms (AFLP), microsatellite or simple sequence repeat length polymorphisms (SSR), random amplified polymorphic sequences (RAPD), cleavable amplified polymorphic sequences (CAPS), single-strand conformation polymorphisms (SSCP) and single nucleotide polymorphisms (SNPs) in parallel with the dramatic decrease in the cost of DNA sequencing has led to several quantitative genetic studies. Recently, this trend increased by the

development of various SNP-arrays for whole genome genotyping of many plant lines (Rasheed et al., 2017), allowing selection at genomic level to be realistic (Mohan et al., 1997; Rafalski, 2002; Pérez-de-Castro et al., 2012). One of the relatively new approaches that are used for assessing the genotype-phenotype relationship by association mapping or linkage disequilibrium (LD) mapping is the so-called genome-wide association studies (GWAS) which are used for detecting associated markers of quantitative trait loci (QTL) (Hu et al., 2018). This approach is more applicable over the traditional approach which is known as a bi-parental analysis or linkage mapping that is based on using bi-parental populations resulting from a cross between two inbred parents. The development of suitable mapping populations is costly and time-consuming (Pérez-de-Castro et al., 2012; Hu et al., 2018). In linkage mapping studies, the used populations are generally $F2$ or its derivatives ($F3$, $F4$ etc.,), backcrossed lines, doubled-haploids, recombinant inbred lines (RIL), immortalized $F2$ and near isogenic lines (NIL). In contrast, the GWAS approach is based on natural variation using a diverse collection of varieties or accessions with a high dense map in order to detect marker-trait associations (MTAs) are underlying such variation. Association mapping is mostly influenced by population structure among individuals that requires special statistical models to reduce the rate of false positive associations, improve mapping precision and power (Weir, 2010). Meanwhile, using the mixed linear model (MLM) is a powerful method that implements the correction of the population structure by principal component analysis (PCA) or kinship matrix model in order to reduce the rate of false positives results (Yu et al., 2006; Li et al., 2017). Therefore, association mapping is capable to identify the markers underlying the natural variation with their allelic effects for the given trait.

Recently, another genome-wide analysis appeared to be suitable for the selection of polygenic traits for accelerating the breeding process which is known as genomic prediction (GP). GP approach aimed to use both large and small effects of genome-wide markers to predict the genomic estimated breeding values (GEBVs) of a complex trait (Meuwissen et al., 2001). In this type of analysis, the genotypic and phenotypic data are provided for the training population to predict GEBVs in a test population upon their genotypic data (Meuwissen et al., 2001; Habier et al., 2011). Several statistical models have been proposed to estimate the marker effects in the training population where the most robust ones are: Bayesian models (Meuwissen et al., 2001; Habier et al., 2011), ridge regression best linear unbiased predictions (rrBLUP) (Piepho et al., 2012; Calus et al., 2014) and Genomic best linear unbiased prediction (GBLUP) (Ober et al.,

2011; Calus et al., 2014). Therefore, the GP approach became widely used in recent quantitative genetics research as a promising direction in breeding programs (Crossa et al., 2017).

1.7 Organization and objectives of the dissertation

The dissertation is written in the form of a cumulative thesis, which includes a general introduction (chapter 1), three scientific peer-reviewed articles (chapters 2.1, 2.2 and 2.3) and a general discussion based on all results (chapter 3). In chapter 1 (general introduction), the general information about wheat, grain anatomy, wheat nutrients and mineral distribution in the grain, mineral deficiency and consequences on human health and plant nutrition, biofortification and genome-wide analyses are provided. Chapters 2.1, 2.2 and 2.3 are published peer-reviewed papers and are in general self-contained i.e. they contain own introduction, materials and methods, results and discussion parts. In chapter 3, the major results from chapters 2.1, 2.2 and 2.3 are discussed in order to explain how the presented findings are related to improving nutrient quality traits.

The overall aim of this study was to perform the genome-wide analysis of Fe, Zn and Ca concentrations in hexaploid (*Triticum aestivum* L.) wheat grains using the latest high-dense SNP arrays in diverse germplasm including elite wheat varieties (355 winter wheat lines and 14 spring wheat lines).

The specific objectives of this thesis are:

1. Investigation of the natural phenotypic variation of Ca, Zn and Fe concentration in wheat grains
2. Genome-wide association analysis of Ca, Zn and Fe concentrations in wheat grains in the wheat panel by using different marker arrays
3. Identification of putative genomic target regions and identification of putative candidate genes of Ca, Zn and Fe concentrations
4. Estimating the genome-wide prediction of Fe concentration.

2 Research Publications

2.1 Genome-Wide Association Study of Calcium Accumulation in Grains of European Wheat Cultivars

Published in: Frontiers in Plant Science (2017) 8: 1797.

DOI: 10.3389/fpls.2017.01797

Authors: Dalia Z. Alomari*, Kai Eggert, Nicolaus von Wirén, Klaus Pillen and Marion S. Röder

The original publication is available online at:

https://www.frontiersin.org/articles/10.3389/fpls.2017.01797/full

frontiers in Plant Science

ORIGINAL RESEARCH
published: 27 October 2017
doi: 10.3389/fpls.2017.01797

Genome-Wide Association Study of Calcium Accumulation in Grains of European Wheat Cultivars

Dalia Z. Alomari[1], Kai Eggert[1], Nicolaus von Wirén[1], Klaus Pillen[2] and Marion S. Röder[1]*

[1] Leibniz Institute of Plant Genetics and Crop Plant Research, Gatersleben, Germany, [2] Institute of Agricultural and Nutritional Sciences, Martin Luther University Halle-Wittenberg, Halle, Germany

Mineral concentrations in cereals are important for human health, especially for people who depend mainly on consuming cereal diet. In this study, we carried out a genome-wide association study (GWAS) of calcium concentrations in wheat (*Triticum aestivum* L.) grains using a European wheat diversity panel of 353 varieties [339 winter wheat (WW) plus 14 of spring wheat (SW)] and phenotypic data based on two field seasons. High genotyping densities of single-nucleotide polymorphism (SNP) markers were obtained from the application of the 90k iSELECT ILLUMINA chip and a 35k Affymetrix chip. Inductively coupled plasma optical emission spectrometry (ICP-OES) was used to measure the calcium concentrations of the wheat grains. Best linear unbiased estimates (BLUEs) for calcium were calculated across the seasons and ranged from 288.20 to 647.50 among the varieties ($\mu g\ g^{-1}$ DW) with a mean equaling 438.102 ($\mu g\ g^{-1}$ DW), and the heritability was 0.73. A total of 485 SNP marker–trait associations (MTAs) were detected in data obtained from grains cultivated in both of the two seasons and BLUE values by considering associations with a $-\log_{10}$ (*P*-value) $\geq$3.0. Among these SNP markers, we detected 276 markers with a positive allele effect and 209 markers with a negative allele effect. These MTAs were found on all chromosomes except chromosomes 3D, 4B, and 4D. The most significant association was located on chromosome 5A (114.5 cM) and was linked to a gene encoding cation/sugar symporter activity as a potential candidate gene. Additionally, a number of candidate genes for the uptake or transport of calcium were located near significantly associated SNPs. This analysis highlights a number of genomic regions and candidate genes for further analysis as well as the challenges faced when mapping environmentally variable traits in genetically highly diverse variety panels. The research demonstrates the feasibility of the GWAS approach for illuminating the genetic architecture of calcium-concentration in wheat grains and for identifying putative candidate genes underlying this trait.

Keywords: wheat, calcium, GWAS, MTAs, mineral concentration

OPEN ACCESS

Edited by:
Soren K. Rasmussen, University of Copenhagen, Denmark

Reviewed by:
Toi J. Tsilo, Agricultural Research Council of South Africa (ARC-SA), South Africa
Harsh Raman, NSW Department of Primary Industries, Australia

****Correspondence:***
Dalia Z. Alomari
alomari@ipk-gatersleben.de

Specialty section:
This article was submitted to Plant Breeding, a section of the journal Frontiers in Plant Science

Received: *22 June 2017*
Accepted: *03 October 2017*
Published: *27 October 2017*

Citation:
Alomari DZ, Eggert K, von Wirén N, Pillen K and Röder MS (2017) Genome-Wide Association Study of Calcium Accumulation in Grains of European Wheat Cultivars. Front. Plant Sci. 8:1797. doi: 10.3389/fpls.2017.01797

INTRODUCTION

Hexaploid wheat (*Triticum aestivum* L.) is one of the most essential and widely planted crops worldwide with its products feeding most of the global population (FAO, 2016)[1]. Many people relying strongly on wheat-based food stuff suffer from nutrient deficiencies, especially of Fe, Zn, Ca, and Mg (Welch and Graham, 2004; White and Broadley, 2005, 2009; Yano et al., 2016), because

[1]http://www.fao.org/worldfoodsituation/csdb/en/

wheat grains contain low amounts of these nutrients. Genetic biofortification is one strategy involving plant breeding, which offers a sustainable and long-term approach for developing mineral-rich crop varieties (Bouis, 2007; Velu et al., 2014). This requires a better understanding of the genetic basis of mineral element accumulation in wheat grains that improves wheat quality and its value for human dietary consumption. Calcium plays an important role in cell wall structure, plant architecture, quality, and yield formation, while its deficiency makes the plant more sensitive to biotic and abiotic stresses (Dayod et al., 2010). Most of calcium dietary consumption in humans is lower than the recommended daily intake (RDI) of 800–1,300 mg per capita (Kranz et al., 2007). Adequate calcium intake especially during adolescence is critical to reduce the rate of bone loss, rickets, and osteoporosis, while lower intake provokes health risks, such as hypocalcemia, hypertension, colorectal cancer as well as bone weakness and fractures accompanied with aging (Centeno et al., 2009; Dayod et al., 2010; Pravina et al., 2013). Increasing Ca accumulation in wheat grains is thus an important goal in wheat breeding.

Several studies on different plant species and crops identified putative quantitative trait loci (QTL) for calcium in grains of wheat, rice, sorghum, barley, maize, pearl millet, or beans (Peleg et al., 2009; Zhang et al., 2009; Goel et al., 2011; Orazaly et al., 2015; Fedorowicz-Strońska et al., 2017; Sharma et al., 2017). In a tetraploid wheat population of recombinant inbred lines (RILs), derived from a cross between durum wheat and wild emmer, nine significant QTLs were associated with calcium concentration in grains (Peleg et al., 2009). Goel et al. (2011) reported that 31 genes are responsible for calcium accumulation in rice and 28 genes in sorghum. Five QTLs were identified in *Arabidopsis thaliana*, in which they explained 36.4% of the variation in calcium content (Vreugdenhil et al., 2004).

European countries are among the top wheat producers and exporters in the world (FAO, 2017)[1]; thus we chose a panel of recent European wheat varieties to explore the genetic variation of calcium in 353 varieties [339 winter wheat (WW) and 14 spring wheat (SW) varieties], and to identify QTLs associated with this trait by using a genome-wide association study (GWAS) in order to detect potential candidate genes.

MATERIALS AND METHODS

Plant Materials and Growth Conditions

A European wheat panel consisting of 353 varieties mainly coming from Germany and France was used in this study. This panel included 339 WW and 14 varieties of SW. Field experiments were carried out at the Leibniz Institute of Plant Genetics and Crop Plant Research in Gatersleben, Germany (51°490N, 11°170E, 112 m), during two consecutive seasons (2014/2015 and 2015/2016). The individual plot size was 1 m × 1.5 m with four rows spaced 0.20 m apart. All varieties were sown in autumn and subjected to standard agronomic wheat management practices.

Determination of Calcium Concentrations

Phenotypic analysis was conducted for the whole set of wheat varieties in each season. For each sample, 50 kernels were counted using a digital seed analyzer/counter Marvin (GTA Sensorik GmbH, Neubrandenburg, Germany) and the thousand-grain weight (TGW) was estimated. The samples were milled using a Retsch mill (MM300, Germany) and the milled samples were dried overnight at 40°C. Calcium concentrations were measured by inductively coupled plasma optical emission spectrometry (ICP-OES, iCAP 6000, Thermo Fisher Scientific, Germany) combined with a CETAC ASXPRESS^{TM} PLUS rapid sample introduction system and a CETAC autosampler (CETAC Technologies, Omaha, United States). Fifty micrograms of dried and ground samples from each variety were wet digested in 2 ml nitric acid (HNO_3, 69%, Bernd Kraft GmbH, Germany) using a high-performance microwave reactor (UltraClave IV, MLS, Germany). Digested samples were filled up to 15 ml final volume with de-ionized distilled water (Milli-Q® Reference System, Merck, Germany). Element standards were prepared from Bernd Kraft multi-element standard solution (Germany). Calcium as an external standard and Y (ICP Standard Certipur®, Merck, Germany) were used as internal standards for matrix correction.

Phenotyping and Statistical Analysis

The resulting calcium values for wheat grains of each variety and environment were used to calculate the best linear unbiased estimates (BLUEs), by applying the residual maximum likelihood (REML) algorithm with mixed linear models (MLMs) function (Yu et al., 2006) and considering genotype as fixed effect and environment as random effect. These calculations were accomplished using GenStat v16 software (VSN International, Hemel Hempstead, Hertfordshire, United Kingdom).

The broad sense heritability of Ca was calculated using the equation:

$$H^2 = \sigma_G^2/(\sigma_G^2 + (\sigma_e^2/nE)$$

where σ_G^2 is the variance of the genotype, σ_e^2 represents the variance of the residual and nE is the number of the environments.

Analyses of variance (ANOVA) and Pearson's correlation coefficient were calculated for the calcium trait across the two environments with SigmaPlot package 13.

Genotyping

All wheat varieties were genotyped by the company Trait Genetics GmbH, Gatersleben, Germany[2] using a new 90k iSELECT Infinium array (Wang et al., 2014) which contained 7761 mapped polymorphic single-nucleotide polymorphism (SNP) markers and a 35k Affymetrix-SNP array (Axiom® Wheat Breeder's Genotyping Array[3]) which contained 7762 mapped polymorphic SNPs. For the reference map, the ITMI-DH

[2]www.traitgenetics.com

[3]http://www.cerealsdb.uk.net/

population (Sorrells et al., 2011; Poland et al., 2012) was used to anchor all SNP-markers. Only mapped markers with a minor allele frequency (MAF) ≥3% (equaling 11 varieties out of 353) were used for association analysis.

GWAS Mapping and Linkage Disequilibrium Characteristics

GWAS analysis for the phenotypic and genotypic dataset was performed by using GenStat v16 software (VSN International, Hemel Hempstead, Hertfordshire, United Kingdom). Association analysis was performed using the "Single trait association analysis" function with Kinship matrix (K) as a relationship model to control for population structure by GenStat v16, though no obvious population structure was observed in the described population (Kollers et al., 2013).

To declare the significant marker–traits associations (MTAs), we considered a threshold P-value of $-\log_{10}(P) \geq 3$. When Bonferroni correction with $P < 0.05$ was applied the resulting $-\log_{10}(P)$ threshold rose to 5.49. The proportion of the phenotypic variation (R^2) was calculated using the software package TASSEL 5.0. Marker effects (positive/negative) were estimated by GenStat v16 based on the effect of specific allele in the varieties.

BLUEs of the trait, each variety and across the seasons (2015 and 2016) were calculated by applying the "mixed models REML" module with the "linear mixed models" of GenStat v16.

Linkage disequilibrium (LD) which is the non-random association between pairs of loci was studied in the whole panel, observed by using squared allele frequency correlation and calculated within each chromosome. Loci in the LD region were determined according to the squared allele frequency correlations (r^2) and were considered to be in significant LD when $r^2 \geq 0.2$. LD plots were performed by GenStat (v16) to examine the average LD decay within each chromosome.

Physical Mapping Resources of Wheat and Identification of Putative Candidate Genes

While in GWAS analysis the marker data were connected with the phenotypic data in order to identify significant MTAs, in this step we identified the flanking sequence of SNP markers defining significant associations with the calcium trait. Markers which were located in significant LD regions were obtained from the wheat 90k database (Wang et al., 2014) and 35k database (see text footnote 3). The wheat marker sequences were blasted on the wheat genome assembly IWGSC1 and POPSEQ (The International Barley Genome Sequencing Consortium, 2012) and the website of Ensemble Plants[4] to obtain their corresponding genes, transcripts, and gene identifiers (IDs). Related regions for these significant associations were anchored using the wheat sequence assembly. For the resulting gene IDs the Human-Readable Descriptions were selected to define annotated gene functions by ftp://ftpmips.helmholtz-muenchen.de/plants/wheat/IWGSC/genePrediction_v2.2. The whole set of marker sequences was blasted using the software package Geneious 10[5] and the most significant hit was selected (Kearse et al., 2012). A similar strategy was applied to find a candidate gene for eyespot resistance in wheat (Zanke et al., 2017).

RESULTS

Description of Phenotypic Data

Calcium measurements were performed for the whole set of European wheat varieties (WW = 339, SP = 14) grown in two seasons (2015 and 2016) (Supplementary Table S1). In each season calcium concentrations covered a wide range of variation (**Figure 1A**). In the season 2015, the highest measured grain Ca concentration was 797 $\mu g\ g^{-1}$ DW, while in 2016 the highest Ca value was around 647 $\mu g\ g^{-1}$ DW (**Figure 1B**). Estimated BLUEs ranged from 288.2 to 647.5 $\mu g\ g^{-1}$ DW with a mean of 438.1 $\mu g\ g^{-1}$ DW (**Figure 1B**). Based on BLUEs the highest scored value for calcium in the whole set of wheat varieties was 647.5 $\mu g\ g^{-1}$ DW for the variety Nirvana from France (**Figure 2**). The ANOVA showed significant effects of the genotype and the environment on calcium concentrations in the grain (Supplementary Table S2). The Pearson's correlation measured for calcium trait among the growing environments and BLUEs, ranged from 0.59 to 0.91 ($P < 0.001$, **Figure 1C**). The highest correlation was between season 2015 and BLUEs ($r = 0.91$, $P < 0.01$), while the lowest but still significant correlation was between seasons 2015 and 2016 ($r = 0.59$, $P < 0.01$). The broad sense heritability equaled 0.73 across the two environments for 353 varieties indicating that the phenotypic values in the two years are relatively stable for the different varieties.

Detection of MTAs

GWAS analysis was performed using a MLM with 90k and 35k SNP markers for the calcium data from the two growing seasons 2015 and 2016. Additionally the MTAs for the BLUEs from both years were calculated. Our analysis detected 485 significant [$-\log_{10}$ (P-value) ≥3] association signals for both environments and BLUEs (Supplementary Table S2). A number of 276 significant markers showed a positive allele effect, while the remaining markers (209 markers) had negative allele effects. These MTAs were located on all chromosomes except chromosomes 3D, 4B, and 4D (**Figure 3**). The most significant association was detected on chromosome 5A. On the other hand, most of the significant MTAs for grain calcium were identified on chromosome 2A (111 MTAs) and chromosome 5B (127 MTAs). On chromosome 2A, most of the MTAs were located in the genomic region of 64.3–67.4 cM. Based on the analysis, we found 31 consistent associations, which were present in both environments plus BLUEs and 20 consistent associations are above the

[4]http://plants.ensembl.org/Triticum_aestivum/Tools/Blast

[5]http://www.geneious.com

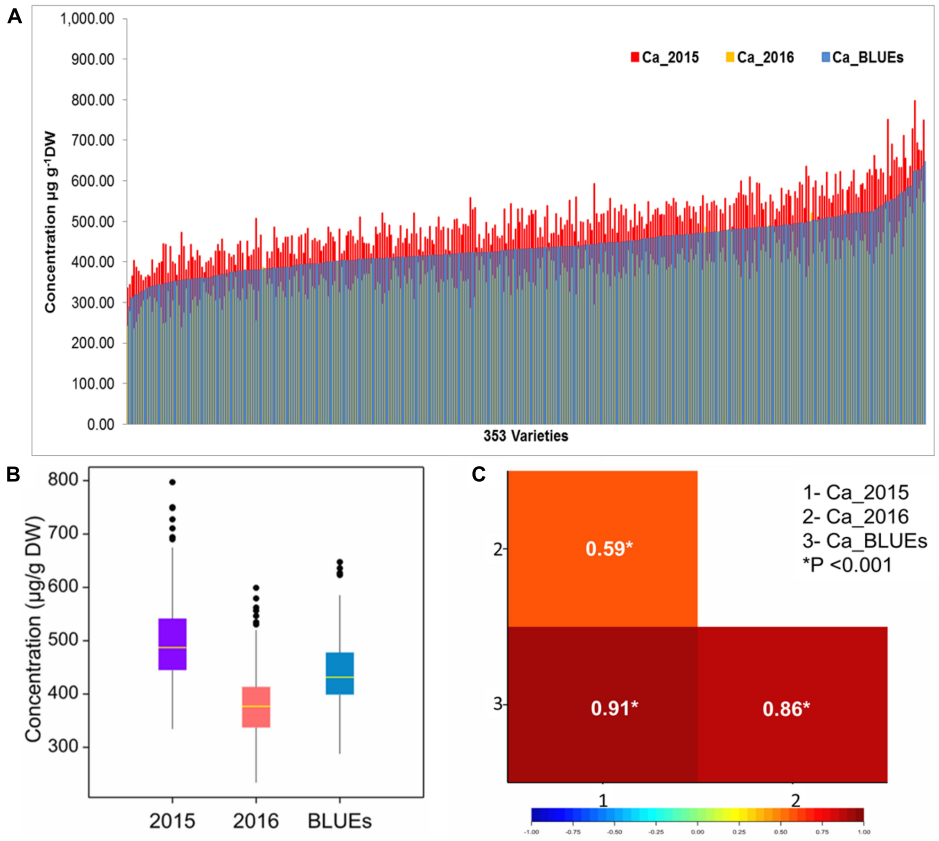


FIGURE 1 | (A) Calcium concentration (µg g^{-1} DW) distribution among 353 European wheat varieties in two different seasons (2015 and 2016) and BLUEs. **(B)** Box plots of calcium values for two single season and BLUEs in 353 varieties. **(C)** Pearson correlation coefficients among (2015/2016) seasons and BLUE value.

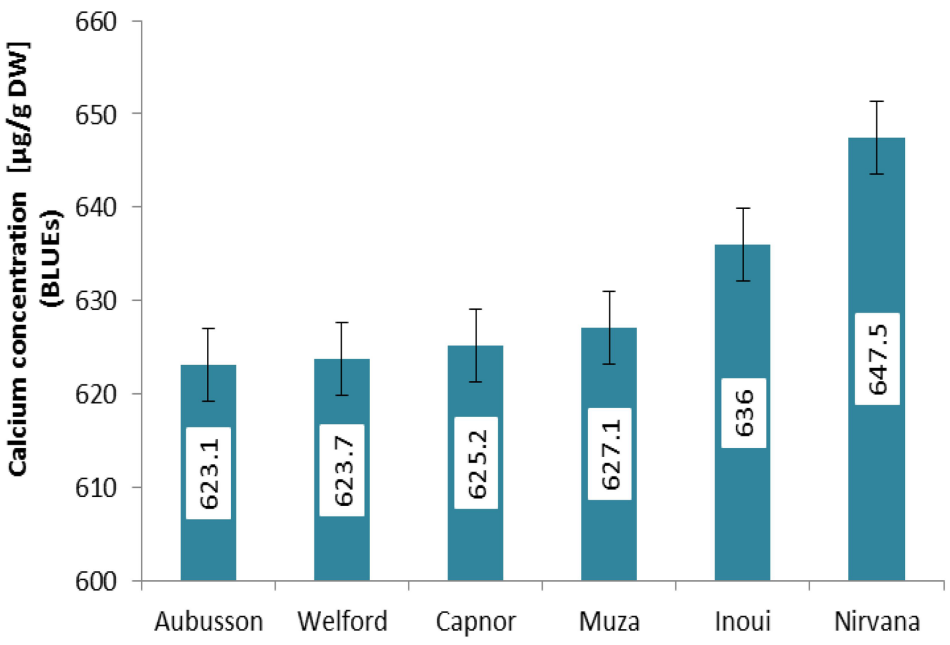


FIGURE 2 | Calcium concentrations of the top 6 accessions [Ca > 600 (µg g^{-1} DW)] based on BLUE values ± SE. Numbers on the *y*-axis represent the calcium concentration.

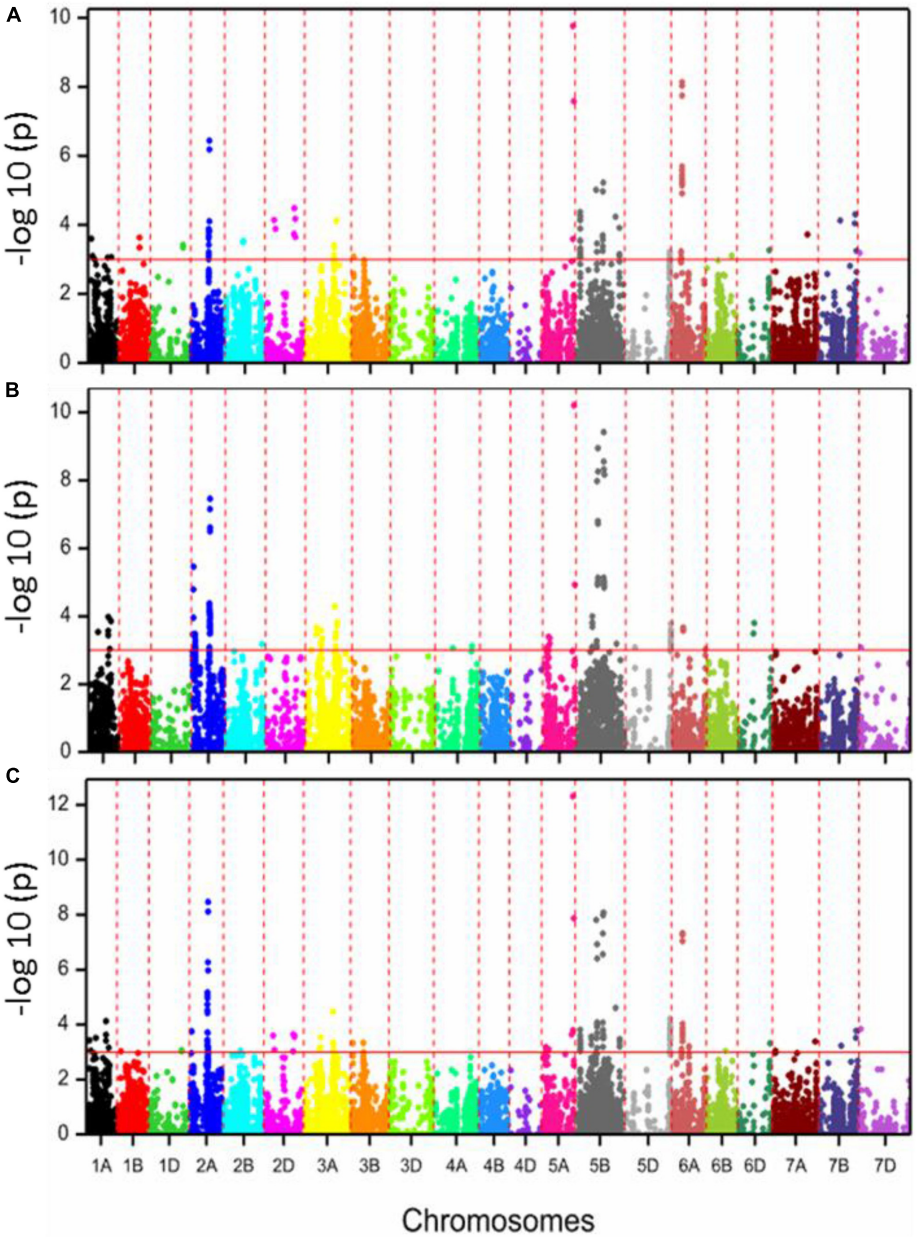


FIGURE 3 | Manhattan plots with –log *P*-values of SNPs associated with calcium concentration values for 353 European wheat varieties based on 2015 **(A)**, 2016 **(B)**, and BLUE values **(C)**. **(A–C)** The red line color in the figure shows the threshold of $-\log_{10}$ (*P*-value) of three and all the significantly associated SNP markers are above the red line.

Bonferroni correction threshold which equals 5.49 (**Table 1**). The explained phenotypic variances (R^2) ranged from 0.81 to 11.27%.

The additive effects of five representative significant markers based on BLUEs are depicted in **Figure 4**. Two significant markers [RAC875_c8642_231 marker (M2) (MAF = 0.09) and wsnp_Ex_c17575_26299925 marker (M5) (MAF = 0.16)] had positive allele effects. Another three significant markers [BS00049644_51 (M1), GENE-0168_7 (M3), and AX-94644169 (M4)] showed negative effects. Marker RAC875_c8642_231 (M2) had a highly significant positive effect ($P < 0.001$) with 43.3 $\mu g\ g^{-1}$ DW (**Figure 4**).

Connection of the Significant Markers to the Wheat Genome Sequence and Identification of Candidate Genes

In order to identify potential candidate genes for calcium concentrations in wheat grains, the significant SNP markers [$-\log_{10}$ (*P*-value) ≥ 3], together with other markers in LD

TABLE 1 | Summary of consistently significant markers detected in the two environments and BLUEs.

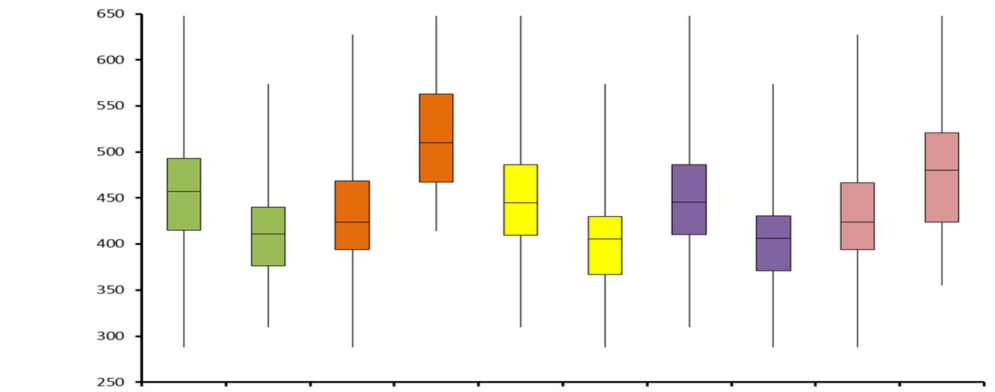
Marker*	Chromosome	Position (cM)	$-\log_{10}$ (*P*) BLUEs	Effect BLUEs	% R^2 BLUEs
BS00049644_51	2A	66.6	8.46	−20.89	4.60
RAC875_c24517_558	2A	64.3	4.73	−15.17	1.68
Kukri_c40035_258	2A	64.3	4.98	−15.60	1.85
AX-95169653	2A	66.6	6.27	18.27	3.12
AX-94940052	2A	66.6	5.97	−17.96	1.61
AX-94536561	2A	66.6	8.11	−20.42	3.44
AX-94881950	2A	64.3	5.17	−16.02	2.25
AX-94850365	2A	64.3	5.06	−15.85	2.19
AX-94560505	2A	64.3	4.42	−14.61	1.70
AX-94544896	2A	64.3	4.53	−14.78	1.78
AX-94404038	2A	64.3	4.96	−15.55	2.13
RFL_Contig1175_354	3A	109	4.47	19.60	0.81
wsnp_Ex_c20899_30011827	5A	117.7	7.87	45.80	7.83
RAC875_c8642_231	5A	114.5	12.31	43.28	11.27
AX-95077733	5A	117.7	7.87	45.80	8.26
snp_CAP8_c1210_739429	5B	149.8	4.61	−33.57	1.86
CAP7_c5481_96	5B	149.8	4.61	−33.57	1.86
RAC875_c30011_426	5B	78.7	6.93	−18.44	3.44
BS00062731_51	5B	78.7	6.41	−17.70	1.03
GENE-0168_7	5B	75.5	7.81	−21.80	4.67
AX-94644169	5B	103.2	8.08	−21.88	4.43
AX-94541836	5B	101.7	7.98	−20.14	2.80
AX-94547820	5B	100.9	7.32	−18.95	3.61
AX-94452355	5B	100.9	6.56	−17.89	2.40
Jagger_c8037_96	5D	167	3.87	13.73	1.64
wsnp_Ex_c17575_26300030	6A	37.3	7.27	24.98	3.55
wsnp_Ex_c17575_26299925	6A	37.3	7.33	25.12	3.72
Tdurum_contig62141_496	6A	37.3	7.27	24.98	3.55
Kukri_rep_c104648_439	6A	37.3	7.27	24.98	3.55
Kukri_c35661_63	6A	37.3	7.27	24.98	3.55
AX-94415776	6A	37.3	7.04	24.69	3.88

**Consistently significant markers appeared in the two environments (2015 and 2016) and BLUE values with threshold of $-\log_{10}$ (P-value ≥3). The highlighted markers with gray color are above the Bonferroni threshold which equals 5.49.*

(r^2 ≥0.2), were used to query the available wheat genome sequence data in the POPSEQ assembly (Supplementary Table S4). The rationale behind this strategy was that the available DNA chips did not cover all genes in the wheat genome and that a highly significant marker may either be the causative gene itself or in close linkage to the causative gene.

Blast analysis of these markers using POPSEQ showed that chromosomes 2A, 5A, 5B, 5D, and 6A harbored many calcium-transporting genes. The most significant SNP (RAC875_c8642_231) in our analysis was located on chromosome 5A (114.5 cM). The gene underlying this marker encodes a cation/sugar symporter, while the second significant locus (wsnp_Ex_c20899_30011827) on the same chromosome (117.7 cM) carries a gene that encodes an AP2-type transcription factor. We further detected two genes, which may be related to calcium transport near this significant region (114.5–117.7 cM): one gene (Traes_5AL_898DAA873) is related to plasma membrane ATPases while another gene (Traes_5AL_637EB761F) encodes an H^+-ATPase. Furthermore, in the same region, we found a gene (Traes_5AL_AE6B41A0A) related to divalent metal cation transport together with two further genes (Traes_5AL_6C9A5537F and Traes_5AL_6C8BD96CB) related to heavy metal transport/detoxification. Along this chromosome we were able to find loci associated with Ca-permeable ion channels, such as Traes_5AL_E1F7DD9EA and Traes_5AL_C89AC9640 coding for cyclic nucleotide-gated channels (CNGCs) or Traes_5AL_F1522B81F and Traes_5AL_98814295D encoding mechanosensitive ion channels. However, these genes were not closely located to the significant markers (Supplementary Table S4).

In the LD region of chromosome 2A, we found a number of genes, which are related to calcium transport functions, such as Traes_2AL_72F83E7B0 and Traes_2AL_6069A8864 (mechanosensitive ion channel family protein), Traes_2AL_D33454518 (cation/H^+ antiporter), and Traes_2AL_CF9F964E6 underlying a calcium-transporting ATPase. The most significant association on chromosome 5B was the locus Traes_5BL_DF8D1B819, which is related to an ammonium transporter.



Marker	M1 BS00049644_51		M2 RAC875_c8642_231		M3 GENE-0168_7		M4 AX-94644169		M5 wsnp_Ex_c17575_26 299925	
SNP	A	G	C	T	A	C	A	C	C	T
Ch.	2A		5A		5B		5B		6A	
N	192	158	322	31	93	250	96	252	57	293
Allele frequancy	54.3	44.7	91.2	8.7	70.8	26.3	71.3	27.1	83.0	16.1
Ca-Mean	458	413	430	516	450	405	451	405	429	479
Additive effect	-20.8	-	43.2	-	-21.7	-	-21.8	-	25.1	-
T.test	***	-	-	***	***	-	***	-	-	***

FIGURE 4 | Box plot analysis for five significant markers based on BLUEs for calcium trait (MAF = AF). ***denotes 0.001.

Other closely linked genes are Traes_5BS_4AEE5C2AE encoding a mechanosensitive ion channel family protein and Traes_5BS_98C73F5CA, Traes_5BS_06F7D0060, Traes_5BS_272FDBF9D, Traes_5BL_4AACBDDAA, and Traes_5BL_E4BE45756, all related to cation/H^+ antiporters. Near to the highly significant markers of chromosome 5B (depicted in yellow in Supplementary Table S4) a gene for a CNGC (Traes_5BL_411EF97B9) is located, while a cation/calcium exchanger (Traes_5BL_6A7BE3F0C) is located quite distantly from the significant markers and are therefore not likely as candidate genes.

DISCUSSION

European Wheat Germplasm Harbors a Large Genotypic Variability in Calcium Accumulation

Genetic fortification strategies are highly suitable for developing wheat varieties with high mineral element contents. Therefore, this study focused on investigating the natural genetic variation in European wheat varieties and on identifying candidate genes contributing to calcium accumulation in wheat grains. Phenotypic analysis for calcium concentrations showed a wide variation between the varieties based on BLUEs which ranged from 288.2 to 647.5 $\mu g\ g^{-1}$ DW. The heritability was high (0.73) indicating that the major part of the variability was due to genotypic effects, which is in agreement with previous studies (Vreugdenhil et al., 2004; Garcia-Oliveira et al., 2009; Peleg et al., 2009). Very strong, significant correlation coefficients were detected between the two seasons indicating that the phenotypic measurements were quite stable in the different years. This conclusion was also supported by a high heritability. Considering the analysis across the two growth seasons, the results showed that genotypic variances due to genotypes were significant (at $P < 0.01$). Marker effects (R^2) which explained the proportion of phenotypic variance for consistently significant markers (appearing significant in both seasons and BLUEs) contributed a modest proportion ranging from 0.81 to 11.27%. The ANOVA results indicated that genotypes and environmental factors have a significant effect on calcium concentration in wheat grains. A similar conclusion was reached by (Gu et al., 2015) for grain Ca in maize.

Calcium Accumulation in Grains Is Controlled by Multiple Loci

In the present study, genome mapping revealed that most of the significant MTAs for the consistently significant markers in 2015, 2016, and BLUEs (**Table 1**) are conferred mostly by genome A (chromosomes 2A, 3A, 5A, and 6A), while one locus was related to the B genome (chromosome 5B) and another one related to the D genome (chromosome 5D). A mapping study of a RIL population in tetraploid wheat detected significant QTL for calcium concentration on chromosomes 1A, 4A, 6A, 2B, 4B, 5B, 6B, and 7B (Peleg et al., 2009). Another study on bread wheat for calcium-dependent protein kinases (CDPKs) which are crucial

TABLE 2 | Putative candidate genes for Ca in the vicinity of highly significant associated SNP-marker.

Significant marker	Chromosome, position (cM)	(r^2) bw_SNP	$-\log_{10}(p)$	Effect	% R^2 BLUES	Gene name (IWGSC_v1)	Human_Readable_Description
RAC875_c8642_231	5A, −114.5	0.46	12.31	43.72	11.27	Traes_5AL_F49663738	Sugar transporter/solute-cation symport
wsnp_Ex_c20899_30011827	5A, −117.7	1.00	7.87	46.14	7.83	Traes_5AL_19637DE03	AP-2 complex subunit alpha-1
–	5A	–	–	–	–	Traes_5AL_6C8BD96CB	Heavy metal transport/detoxification superfamily protein
–	5A	–	–	–	–	Traes_5AL_898DAA873	Plasma membrane ATPase 1
–	5A	–	–	–	–	Traes_5AL_637EB761F	H(+)-ATPase 11
–	5A	–	–	–	–	Traes_5AL_AE6B41A0A	Divalent metal cation transporter MntH
–	5A	–	–	–	–	Traes_5AL_6C9A5537F	Heavy metal transport/detoxification
–	5A	–	–	–	–	Traes_5AL_6C8BD96CB	Heavy metal transport/detoxification
AX-94940052	2A, −66.6	0.48	5.97	−17.96	1.61	Traes_2AL_156C770EE	Receptor-like protein kinase 2
–	2A	–	–	–	–	Traes_2AL_542E74269	Potassium channel AKT1
Kukri_c40035_258	2A, −64.0	0.10	4.98	−15.60	1.85	Traes_2AS_0C87833C6	Phosphatidylinositol-4-phosphate 5-kinase family protein
–	2A	–	–	–	–	Traes_2AL_72F83E7B0	Mechanosensitive ion channel family protein
–	2A	–	–	–	–	Traes_2AL_6069A8864	Mechanosensitive ion channel family protein
–	2A	–	–	–	–	Traes_2AL_D33454518	Cation/H^+ antiporter
–	2A	–	–	–	–	Traes_2AS_B264257CD	Flavin-binding monooxygenase family protein
–	2A	–	–	–	–	Traes_2AL_CF9F964E6	Calcium-transporting ATPase
–	2A	–	–	–	–	Traes_2AS_95611CAD2	Heavy metal transport/detoxification superfamily
–	2A	–	–	–	–	Traes_2AL_6DD37E6BE	Heavy metal transport/detoxification superfamily
–	2A	–	–	–	–	Traes_2AL_9B175F3Da	Heavy metal transport/detoxification superfamily
–	2A	–	–	–	–	Traes_2AL_F360E3FE3	Heavy metal transport/detoxification superfamily
–	2A	–	–	–	–	Traes_2AL_13CBA4FEA	Heavy metal transport/detoxification superfamily
–	2A	–	–	–	–	Traes_2AS_AA84E72D4	Heavy metal transport/detoxification superfamily
–	2A	–	–	–	–	Traes_161086245	Heavy metal transport/detoxification superfamily
AX-94547820	5B, −100.9	0.03	7.20	−18.80	3.61	Traes_5BL_DF8D1B819	Ammonium transporter 2
–	5B	–	–	–	–	Traes_5BL_411EF97B9	Cyclic nucleotide-gated channel

sensors of calcium concentration changes in plant cells, identified 20 CDPK genes (Li et al., 2008). To our knowledge, this is the first report on GWAS for grain calcium concentration in hexaploid wheat. Thus, further genetic and functional analysis of associated genomic regions may shed further light on the genetic basis of improved calcium concentration in wheat grains.

Putative Candidate Genes for Ca-related QTLs

In general, calcium transporters are involved in the cellular compartmentalization of calcium in different plant organs. Three major gene families of calcium transporter proteins have been described: (i) Ca^{2+}-transporting P-type-ATPases [endoplasmic reticulum-type Ca^{2+}-ATPase (ECA/IIA Type) and autoinhibited Ca^{2+}-ATPase (ACA/IIB-type)], (ii) divalent cation–H^+ antiporters/exchangers [cation/H^+ antiporters (CAX), CCX and CHX], and (iii) Ca-permeable ion channels that include mechanosensitive calcium-permeable channels (MSCCs), glutamate receptors (GLRs), CNGCs and two-pore channels (TPC) (Vinoth and Ravindhran, 2017). The highly significant SNP-markers (**Table 1** and Supplementary Table S3) could either be derived from the causative genes themselves or be in linkage to the causative genes for the identified QTLs for Ca. In **Table 2**, we compiled the list of Ca-related genes or transporters which are in close vicinity to highly significant SNP-markers and which are therefore potential candidate genes for the Ca-related QTLs.

Based on our investigations, we found 41 potential calcium-transporting genes distributed over six chromosomes (2A, 3A, 5A, 5B, 5D, and 6A). These include 8 Ca/proton exchangers, 4 Ca-ATPases, and 31 channels in addition to other genes that are putatively related to calcium transport, such as Traes_5AL_6C9A5537F which is annotated as heavy metal transporter (Supplementary Table S4). On chromosome 5A, the most significantly associated MTA with a R^2 value equaling to 11.27% and a favorable additive effect, is related to Traes_5AL_F49663738 gene encoding a putative cation/sugar symporter. The second significant gene (Traes_5AL_19637DE03) with R^2 equaling 7.83 encodes an AP-2 complex subunit alpha-2-like protein that is possibly related to calcium transport function (Matros et al., 2017). Another significant SNP-marked gene (Traes_5AL_320913F7A) also located on the same chromosome and related to a gene of the 2S albumin superfamily, which encodes as a storage protein (Yamazaki et al., 2008). All of these three genes are located on a region between 114.5 and 117.7 cM on chromosome 5AL. In addition to six Ca^{+2} channels associated with these markers: Traes_5AL_BBFBC2F48, Traes_5AL_B598F5A0D, Traes_5AL_E1F7DD9EA, Traes_5AL_C89AC9640, Traes_5AL_F1522B81F, and Traes_5AL_98814295D distributed along chromosome 5A (**Table 2** and Supplementary Table S4), we also detected on this chromosome, two genes (Traes_5AL_6C9A5537F and Traes_5AL_6C8BD96CB) that encode for heavy metal transport/detoxification superfamily proteins involved in metal ion binding (Hall, 2002). Near the significant region, the Traes_5AL_AE6B41A0A marker relates to divalent metal cation transporters that may also act as calcium transporter. Significant associations were also noted on chromosome 2A with 11 SNP markers located within this region (64–66.6 cM) and some of them encoding a disease resistance protein, CBS domain-containing protein, receptor-like protein kinase 2, phosphatidylinositol-4-phosphate 5-kinase family protein, NHL domain-containing protein or Rho GTPase-activating protein besides other genes with unknown function. The LD region on chromosome 2A is widely spread on the physical map of the genome assembly of IWGSC1 extending to the long and the short arm of chromosome 2A. Discrepancies in the order of the contigs in this genome assembly were already described in Zanke et al. (2017). This region contains a number of genes potentially related to calcium-accumulation such as mechanosensitive ion channel family proteins (Traes_2AL_6069A884, Traes_2AL_72F83E7B0) and a number of heavy metal transport/detoxification superfamily proteins (Traes_2AS_95611CAD2, Traes_2AL_6DD37E6BE, Traes_2AL_9B175F3Da, Traes_2AL_F360E3FE3, Traes_2AL_13CBA4FEA, Traes_2AS_AA84E72D4, and Traes_161086245). Nine significant SNPs occurred on chromosome 5B encoding for different functions and some of them may be involved in calcium transport, like Traes_5BL_DF8D1B819 gene which is located on 100.9 cM and is encodes an ammonium transporter. On chromosome 5D, there were two significant markers: Jagger_c8037_96 and BS00032035_51 with unknown functions. On chromosome 6A are located six significant SNP markers, which are related to two genes encoding histone superfamily proteins with a role in the activation of calcium/calmodulin-dependent protein kinases (Davis et al., 2003). Based on our results, the annotated functions of significant genes and genes in the LD region suggested the presence of several genes controlling the calcium uptake. These genes can be considered as putative candidate genes for calcium accumulation in wheat grains and provide a solid resource for future work. However, further functional validation of these genes and their role in calcium uptake in wheat grains is still needed.

CONCLUSION

Apart from focusing on the concentrations of iron and zinc in wheat, which has taken much attention in previous studies, only few genetic studies are available on calcium concentrations in wheat grains are available. Improving levels of grain calcium concentration in hexaploid wheat remains one of the most important breeding objectives for the nutritional security of the whole population and especially for the poor from the nations where wheat is the main source of calories. Overall, through measurable phenotypic and genotypic variation for grain calcium concentrations as well as by considering the detected favorable QTLs distributed across various chromosomes and potentially responsible genes in the current research, we aimed to deepen the understanding of the genetic basis of calcium accumulation in wheat grains and to open the door to more efficient ways to increase calcium concentration in the grain and thereby overall wheat quality.

AUTHOR CONTRIBUTIONS

DA performed the data analysis including genome-wide association scan and related analyses. KE and NvW participated in calcium concentration measurements. KP and MR designed the experiment. MR conceived the idea and participated in the interpretation of results. DA and MR wrote the manuscript. All authors read and approved the final manuscript.

FUNDING

The project was funded by internal financial support of IPK Gatersleben.

ACKNOWLEDGMENTS

The authors would like to acknowledge Ahmad M. Alqudah for the helpful assistance in software learning and giving good advices. Authors also thank Quddoos Muqaddasi for his help in the software tools. We also thank Ellen Weiß and Yudelsy Antonia Tandron Moya for excellent technical assistance.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2017.01797/full#supplementary-material

REFERENCES

Bouis, H. E. (2007). Micronutrient fortification of plants through plant breeding: Can it improve nutrition in man at low cost? *Proc. Nutr. Soc.* 62, 403–411. doi: 10.1079/PNS2003262

Centeno, V., de Barboza, G. D., Marchionatti, A., Rodriguez, V., and Tolosa de Talamoni, N. (2009). Molecular mechanisms triggered by low-calcium diets. *Nutr. Res. Rev.* 22, 163–174. doi: 10.1017/S0954422409990126

Davis, F. J., Gupta, M., Camoretti-Mercado, B., Schwartz, R. J., and Gupta, M. P. (2003). Calcium/calmodulin-dependent protein kinase activates serum response factor transcription activity by its dissociation from histone deacetylase, HDAC4. Implications in cardiac muscle gene regulation during hypertrophy. *J. Biol. Chem.* 278, 20047–20058. doi: 10.1074/jbc.M209998200

Dayod, M., Tyerman, S. D., Leigh, R. A., and Gilliham, M. (2010). Calcium storage in plants and the implications for calcium biofortification. *Protoplasma* 247, 215–231. doi: 10.1007/s00709-010-0182-0

Fedorowicz-Strońska, O., Koczyk, G., Kaczmarek, M., Krajewski, P., and Sadowski, J. (2017). Genome-wide identification, characterisation and expression profiles of calcium-dependent protein kinase genes in barley (*Hordeum vulgare* L.). *J. Appl. Genet.* 58, 11–22. doi: 10.1007/s13353-016-0357-2

Garcia-Oliveira, A. L., Tan, L., Fu, Y., and Sun, C. (2009). Genetic identification of quantitative trait loci for contents of mineral nutrients in rice grain. *J. Integr. Plant Biol.* 51, 84–92. doi: 10.1111/j.1744-7909.2008.00730.x

Goel, A., Taj, G., Pandey, D., Gupta, S., and Kumar, A. (2011). Genome-wide comparative *in silico* analysis of calcium transporters of rice and sorghum. *Genomics Proteomics Bioinformatics* 9, 138–150. doi: 10.1016/S1672-0229(11)60017-X

Gu, R., Chen, F., Liu, B., Wang, X., Liu, J., Li, P., et al. (2015). Comprehensive phenotypic analysis and quantitative trait locus identification for grain mineral concentration, content, and yield in maize (*Zea mays* L.). *Theor. Appl. Genet.* 128, 1777–1789. doi: 10.1007/s00122-015-2546-5

Hall, J. L. (2002). Cellular mechanisms for heavy metal detoxification and tolerance. *J. Exp. Bot.* 53, 1–11. doi: 10.1093/jexbot/53.366.1

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., et al. (2012). Geneious basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* 28, 1647–1649. doi: 10.1093/bioinformatics/bts199

Kollers, S., Rodemann, B., Ling, J., Korzun, V., Ebmeyer, E., Argillier, O., et al. (2013). Whole genome association mapping of fusarium head blight resistance in European winter wheat (*Triticum aestivum* L.). *PLOS ONE* 8:e57500. doi: 10.1371/journal.pone.0057500

Kranz, S., Lin, P. J., and Wagstaff, D. A. (2007). Children's dairy intake in the United States: Too little, too fat? *J. Pediatr.* 151, 642–646, 646.e1–646.e2. doi: 10.1016/j.jpeds.2007.04.067

Li, A.-L., Zhu, Y.-F., Tan, X.-M., Wang, X., Wei, B., Guo, H.-Z., et al. (2008). Evolutionary and functional study of the CDPK gene family in wheat (*Triticum aestivum* L.). *Plant Mol. Biol.* 66, 429–443. doi: 10.1007/s11103-007-9281-5

Matros, A., Liu, G., Hartmann, A., Jiang, Y., Zhao, Y., Wang, H., et al. (2017). Genome–metabolite associations revealed low heritability, high genetic complexity, and causal relations for leaf metabolites in winter wheat (*Triticum aestivum*). *J. Exp. Bot.* 68, 415–428. doi: 10.1093/jxb/erw441

Orazaly, M., Chen, P., Zeng, A., and Zhang, B. (2015). Identification and confirmation of quantitative trait loci associated with soybean seed hardness. *Crop Sci.* 55, 688–694. doi: 10.2135/cropsci2014.03.0219

Peleg, Z., Cakmak, I., Ozturk, L., Yazici, A., Jun, Y., Budak, H., et al. (2009). Quantitative trait loci conferring grain mineral nutrient concentrations in durum wheat × wild emmer wheat RIL population. *Theor. Appl. Genet.* 119, 353–369. doi: 10.1007/s00122-009-1044-z

Poland, J. A., Brown, P. J., Sorrells, M. E., and Jannink, J.-L. (2012). Development of high-density genetic maps for barley and wheat using a novel two-enzyme genotyping-by-sequencing approach. *PLOS ONE* 7:e32253. doi: 10.1371/journal.pone.0032253

Pravina, P., Sayaji, D., and Avinash, M. (2013). Calcium and its role in human body. *Int. J. Res. Pharm. Biomed. Sci.* 4, 659–668.

Sharma, D., Jamra, G., Singh, U. M., Sood, S., and Kumar, A. (2017). Calcium biofortification: three pronged molecular approaches for dissecting complex trait of calcium nutrition in finger millet (*Eleusine coracana*) for devising strategies of enrichment of food crops. *Front. Plant Sci.* 7:2028. doi: 10.3389/fpls.2016.02028

Sorrells, M. E., Gustafson, J. P., Somers, D., Chao, S., Benscher, D., Guedira-Brown, G., et al. (2011). Reconstruction of the Synthetic W7984 × Opata M85 wheat reference population. *Genome* 54, 875–882. doi: 10.1139/g11-054

The International Barley Genome Sequencing Consortium (2012). A physical, genetic and functional sequence assembly of the barley genome. *Nature* 491, 711–716. doi: 10.1038/nature11543

Velu, G., Ortiz-Monasterio, I., Cakmak, I., Hao, Y., and Singh, R. P. (2014). Biofortification strategies to increase grain zinc and iron concentrations in wheat. *J. Cereal Sci.* 59, 365–372. doi: 10.1007/s00122-016-2726-y

Vinoth, A., and Ravindhran, R. (2017). Biofortification in millets: a sustainable approach for nutritional security. *Front. Plant Sci.* 8:29. doi: 10.3389/fpls.2017.00029

Vreugdenhil, D., Aarts, M. G. M., Koornneef, M., Nelissen, H., and Ernst, W. H. O. (2004). Natural variation and QTL analysis for cationic mineral content in seeds of *Arabidopsis thaliana*. *Plant Cell Environ.* 27, 828–839. doi: 10.1111/j.1365-3040.2004.01189.x

Wang, S., Wong, D., Forrest, K., Allen, A., Chao, S., Huang, B. E., et al. (2014). Characterization of polyploid wheat genomic diversity using a high-density 90,000 single nucleotide polymorphism array. *Plant Biotechnol. J.* 12, 787–796. doi: 10.1111/pbi.12183

Welch, R. M., and Graham, R. D. (2004). Breeding for micronutrients in staple food crops from a human nutrition perspective. *J. Exp. Bot.* 55, 353–364. doi: 10.1093/jxb/erh064

White, P. J., and Broadley, M. R. (2005). Biofortifying crops with essential mineral elements. *Trends Plant Sci.* 10, 586–593. doi: 10.1016/j.tplants.2005.10.001

White, P. J., and Broadley, M. R. (2009). Biofortification of crops with seven mineral elements often lacking in human diets–iron, zinc, copper, calcium, magnesium, selenium and iodine. *New Phytol.* 182, 49–84. doi: 10.1111/j.1469-8137.2008.02738.x

Yamazaki, M., Shimada, T., Takahashi, H., Tamura, K., Kondo, M., Nishimura, M., et al. (2008). *Arabidopsis* VPS35, a retromer component, is required for vacuolar protein sorting and involved in plant growth and leaf senescence. *Plant Cell Physiol.* 49, 142–156. doi: 10.1093/pcp/pcn006

Yano, K., Yamamoto, E., Aya, K., Takeuchi, H., Lo, P. C., Hu, L., et al. (2016). Genome-wide association study using whole-genome sequencing rapidly identifies new genes influencing agronomic traits in rice. *Nat. Genet.* 48, 27–34. doi: 10.1038/ng.3596

Yu, J., Pressoir, G., Briggs, W. H., Vroh Bi, I., Yamasaki, M., Doebley, J. F., et al. (2006). A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. *Nat. Genet.* 38, 203–208. doi: 10.1038/ng1702

Zanke, C. D., Rodemann, B., Ling, J., Muqaddasi, Q. H., Plieske, J., Polley, A., et al. (2017). Genome-wide association mapping of resistance to eyespot disease (*Pseudocercosporella herpotrichoides*) in European winter wheat (*Triticum aestivum* L.) and fine-mapping of Pch1. *Theor. Appl. Genet.* 130, 505–514. doi: 10.1007/s00122-016-2830-z

Zhang, B., Chen, P., Shi, A., Hou, A., Ishibashi, T., and Wang, D. (2009). Putative quantitative trait loci associated with calcium content in soybean seed. *J. Hered.* 100, 263–269. doi: 10.1093/jhered/esn096

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

2.2 Identifying Candidate Genes for Enhancing Grain Zn Concentration in Wheat

Published in: Frontiers in Plant Science (2018) 9: 1313.

DOI: 10.3389/fpls.2018.01313

Authors: Dalia Z. Alomari*, Kai Eggert, Nicolaus von Wirén, Ahmad M. Alqudah, Andreas Polley, Jörg Plieske, Martin W. Ganal, Klaus Pillen and Marion S. Röder

The original publication is available online at:

https://www.frontiersin.org/articles/10.3389/fpls.2018.01313/full

frontiers in Plant Science

ORIGINAL RESEARCH
published: 10 September 2018
doi: 10.3389/fpls.2018.01313

Identifying Candidate Genes for Enhancing Grain Zn Concentration in Wheat

Dalia Z. Alomari[1], Kai Eggert[1†], Nicolaus von Wirén[1], Ahmad M. Alqudah[1], Andreas Polley[2], Jörg Plieske[2], Martin W. Ganal[2], Klaus Pillen[3] and Marion S. Röder[1]*

[1] Leibniz Institute of Plant Genetics and Crop Plant Research (IPK), Gatersleben, Germany, [2] SGS TraitGenetics GmbH, Gatersleben, Germany, [3] Institute of Agricultural and Nutritional Sciences, Martin-Luther-University Halle-Wittenberg, Halle, Germany

Wheat (*Triticum aestivum* L.) is one of the major staple food crops worldwide. Despite efforts in improving wheat quality, micronutrient levels are still below the optimal range for human nutrition. In particular, zinc (Zn) deficiency is a widespread problem in human nutrition in countries relying mainly on a cereal diet; hence improving Zn accumulation in grains is an imperative need. This study was designed to understand the genetic architecture of Zn grain concentrations in wheat grains. We performed a genome-wide association study (GWAS) for grain Zn concentrations in 369 European wheat genotypes, using field data from 3 years. The complete wheat panel was genotyped by high-density arrays of single nucleotide polymorphic (SNP) markers (90k iSELECT Infinium and 35k Affymetrix arrays) resulting in 15,523 polymorphic markers. Additionally, a subpanel of 183 genotypes was analyzed with a novel 135k Affymetrix marker array resulting in 28,710 polymorphic SNPs for high-resolution mapping of the potential genomic regions. The mean grain Zn concentration of the genotypes ranged from 25.05–52.67 $\mu g\ g^{-1}$ dry weight across years with a moderate heritability value. Notably, 40 marker-trait associations (MTAs) were detected in the complete panel of varieties on chromosomes 2A, 3A, 3B, 4A, 4D, 5A, 5B, 5D, 6D, 7A, 7B, and 7D. The number of MTAs in the subpanel was increased to 161 MTAs whereas the most significant and consistent associations were located on chromosomes 3B (723,504,241–723,611,488 bp) and 5A (462,763,758–466,582,184 bp) having major effects. These genomic regions include newly identified putative candidate genes, which are related to Zn uptake and transport or represent bZIP and mitogen-activated protein kinase genes. These findings provide the basis for understanding the genetic background of Zn concentration in wheat grains that in turn may help breeders to select high Zn-containing genotypes to improve human health and grain quality.

Keywords: Zinc, *Triticum aestivum*, wheat quality, micronutrient, GWAS

OPEN ACCESS

Edited by:
Pierre Sourdille, INRA Centre Auvergne Rhône-Alpes, France

Reviewed by:
Zhaohui Wang, Northwest A&F University, China
Benoit Darrier, The University of Adelaide, Australia
Sabina Vitalievna Chebotar, Odessa University, Ukraine

****Correspondence:***
Dalia Z. Alomari
alomari@ipk-gatersleben.de; alamridalia@gmail.com

† Deceased

Specialty section:
This article was submitted to Plant Breeding, a section of the journal Frontiers in Plant Science

Received: *14 June 2018*
Accepted: *20 August 2018*
Published: *10 September 2018*

Citation:
Alomari DZ, Eggert K, von Wirén N, Alqudah AM, Polley A, Plieske J, Ganal MW, Pillen K and Röder MS (2018) Identifying Candidate Genes for Enhancing Grain Zn Concentration in Wheat. Front. Plant Sci. 9:1313. doi: 10.3389/fpls.2018.01313

INTRODUCTION

Wheat is among the primary staple crops in the world and its production reached almost 750 million tons per year (FAOSTAT, 2016[1]), while 68% of the yield is used for human nutrition (FAOSTAT, 2012). Wheat provides substantial amounts of mineral elements, which are beneficial for human health. Several reports emphasize that over 2 billion of people are suffering from hidden

[1] http://faostat.fao.org

hunger (Welch and Graham, 2004), i.e., Zinc (Zn) and Iron (Fe) deficiency, mainly in middle- or low-income countries where staple crops are the major food source (Sands et al., 2009); recently, the problem was also reported in developed countries (Pandey et al., 2016).

Zn plays significant roles in different metabolic processes and is an essential cofactor for many enzymes and regulatory proteins. The symptoms of insufficient dietary Zn intake for humans can be observed as growth and development retardation, excessive weight loss, diarrhea, and depression (Ozturk et al., 2006; Kambe et al., 2014; Krishnappa et al., 2017). Consequently, improving the nutritional quality of wheat grains by enhancing Zn concentrations is a long-term goal for breeding novel wheat cultivars with a positive effect on grain yield, nutritional quality of the plant, as well as human health (Cakmak, 2008; Genc et al., 2008; Crespo-Herrera et al., 2016).

Since Zn accumulation in grains is a genetically complex trait, genome-wide association study (GWAS) is a powerful tool to detect the genetic factors underlying the natural variation in such complex traits (Hamblin et al., 2011). Several studies identified quantitative trait loci (QTL) for micronutrients, such as Fe and Zn, or macronutrients like Ca in wheat (Morgounov et al., 2007; Tiwari et al., 2009; Crespo-Herrera et al., 2016; Alomari et al., 2017). Peleg et al. (2009) found six QTLs on chromosomes 2A, 2B, 3A, 4B, 5A, 6A, 6B, 7A, and 7B for Zn in a durum wheat × emmer wheat recombinant inbred lines (RILs) population. Four QTLs for grain Zn concentration were identified by Genc et al. (2008) on chromosomes 3D, 4B, 6B, and 7A in a doubled haploid wheat population. Another study mentioned seven QTLs located on chromosomes 1A, 2D, 3A, 4A, 4D, 5A, and 7A for Zn content in wheat grains of which four QTLs are shared with Zn concentration (Shi et al., 2008). Shi et al. (2013) found that chromosome 4D and 5A probably very vital in controlling mineral status in wheat grains.

Previous studies on Zn concentration mainly used bi-parental population, for instance, RIL (Xu et al., 2012; Pu et al., 2014; Srinivasa et al., 2014) but a few studies have used GWAS with high dense single nucleotide polymorphic (SNP) arrays to investigate the genomic regions underlying the accumulation of micronutrients including Zn in the grains of major cereals like wheat (Guttieri et al., 2015). Therefore, understanding the genetic background of Zn accumulation in wheat grains by GWAS provides the basis for devising the plant breeding strategies and for improving the grain Zn status by introducing the putative candidate genes based on the newly available wheat reference (IWGSC RefSeq v1.0) and using advanced bioinformatics tools.

The main goals of this study were (i) to investigate the natural phenotypic variation on grain Zn concentrations for 369 wheat varieties of 3 years field experiments, (ii) to study the genetic architecture of Zn grain concentration by GWAS analysis with three different high dense SNP arrays including 44,233 SNPs providing a high-resolution genetic map, and (iii) to identify the genomic regions and potential candidate genes for consistently significant QTLs.

MATERIALS AND METHODS

Plant Material and Field Trials

In this study, we used 369 European elite wheat varieties including 355 genotypes of winter wheat and 14 spring wheat genotypes, originating from Germany, France, Poland, Denmark, Austria, Czech Republic, United Kingdom, Sweden, Switzerland, Hungary, Italy, Belgium, and Netherlands described in (Kollers et al., 2013). Field trials were conducted at IPK, Gatersleben, Germany within 3 years (2014/2015 for 358 genotypes, 2015/2016 for 365 genotypes, and 2016/2017 for 360 genotypes). Few genotypes were missing in each individual year due to poor performance and loss in the field. Each plot size was 2 m × 2 m with six rows spaced 0.20 m apart. Plants were grown in clayey loam soil with phosphorus ranges between 7.1–9.0 $\mu g\ g^{-1}$ and pH ≈ 7 across years. Standard agronomic wheat management practices were applied without using fertilizers to avoid the effect of additional fertilizers on the actual Zn concentrations.

Wheat Grain Samples Preparation and Milling

The complete panel of genotypes was analyzed for each individual year. For each genotype, thousand kernel weight (TKW) was measured using a digital seed analyzer/counter Marvin (GTA Sensorik GmbH, Neubrandenburg, Germany). Grains were milled using a Retsch mill (MM300, Germany) and the complete panel of the milled samples was dried by incubating overnight at 40°C.

Measuring Grain Zinc Concentration

Fifty milligrams of dried and milled wheat grain flour was taken to be digested by (2 ml) nitric acid (HNO_3 69%, Bernd Kraft GmbH, Germany). The digestion process was performed using a high-performance microwave reactor (UltraClave IV, MLS, Germany). All digested samples were filled up to 15 ml final volume with de-ionized distilled (Milli-Q®) water (Milli-Q Reference System, Merck, Germany). Element standards were prepared from Bernd Kraft multi-element standard solution (Germany). Zinc as an external standard and yttrium (Y) (ICP Standard Certipur® Merck Germany) were used as internal standards for matrix correction. Zinc concentrations were measured by inductively coupled plasma optical emission spectrometry (ICP-OES, iCAP 6000, Thermo Fisher Scientific, Germany) combined with a CETAC ASXPRESS™ PLUS rapid sample introduction system and a CETAC autosampler (CETAC Technologies, Omaha, NE, United States).

Statistical Analysis

The broad-sense heritability (H^2) was calculated using the equation:

$$H^2 = \sigma_G^2/(\sigma_G^2 + (\sigma_e^2/nE) \quad (1)$$

where σ_G^2 is the variance of the genotype, σ_e^2 represents the variance of the residual, and nE is the number of the environments.

Analyses of variance (ANOVA) and Pearson's correlation coefficient were calculated for the grain Zn trait across 3 years with Sigma Plot package 13.

Best linear unbiased estimates (BLUEs) based on mixed linear models (MLMs) function with applying the residual maximum likelihood (REML) algorithm were calculated to analyze the phenotypic data and estimate the mean of each individual over the years (Yu et al., 2006). To this end, the genotype term was considered as a fixed effect and we denote year as environment term, which was considered as a random effect. These calculations were accomplished using GenStat v16 software (VSN International, Hemel Hempstead, Hertfordshire, United Kingdom).

SNP Genotyping and GWAS Analysis

The complete wheat panel consisting of 369 varieties was genotyped by TraitGenetics GmbH, Gatersleben, Germany[2] using two marker arrays: a 90k iSELECT Infinium array (Wang et al., 2014) and a 35k Affymetrix-SNP array (Axiom® Wheat Breeder's Genotyping Array[3]; Allen et al., 2017). Additionally, a novel 135k Affymetrix array designed by TraitGenetics was used to genotype a subpanel of 183 genotypes from the complete panel of genotypes (Zanke et al., 2017). For the reference map, the ITMI-DH population (Sorrells et al., 2011; Poland et al., 2012) was used to anchor the SNP-markers of the 90k and 35k arrays. The 135k array markers were genetically mapped on four different F_2-populations and then physically anchored on the reference sequence RefSeq v1.0 of hexaploid wheat[4] from International Wheat Genome Sequencing Consortium (IWGSC). For SNP markers quality control, we applied a minor allele frequency (MAF) $\leq$ 3% (equaling 11 varieties out of 369) with rejecting SNPs having missing values or heterozygosity $\geq$ 3%, resulting in 7,761 mapped polymorphic SNP markers from the 90k iSELECT, 7,762 SNPs from the 35k Affymetrix-SNP, and 28,710 from the 135k Affymetrix, which were used for association analysis. The investigated genotype panel and its population structure were described in a previous study by Kollers et al. (2013).

Association mapping based on a MLM was conducted primarily using the Genome Association and Prediction Integrated Tool (GAPIT; Lipka et al., 2012) in R: a language and environment for statistical computing. It includes the phenotypic data with SNP markers coming from the high-density arrays. We incorporated PCA for population correction and stratification. For significant marker-trait associations (MTAs) detection, we set a threshold P-value of $-\log_{10}(P) \geq 3$. Quantile-quantile plots were drawn based on the observed and expected $-\log_{10}(P)$ values. Explained phenotypic variance (R^2) and marker effects (positive/negative) were extracted from GWAS results.

Connecting Significant SNPs With the Physical Sequence of Wheat

The flanking sequence of significant SNP markers defining significant associations with the grain Zn concentration trait was obtained from the wheat 90k database (Wang et al., 2014), 35k database[5] and 135k Affymetrix array (unpublished data, TraitGenetics). These flanking sequences were blasted by Galaxy software, which is an IPK-internal web-based platform[6] by using megablast to fetch the whole sequence of the genomic region of interest based on IWGSC RefSeq v1.0. The extracted sequences were submitted to the annotation pipeline MEGANTE[7] in order to identify potential candidate genes and their gene ontologies.

RESULTS

Natural Phenotypic Variation of Grain Zn Concentrations in Two Wheat Panels

Zn measurements were obtained from grain samples of 369 European wheat varieties, which were grown under field conditions in three consecutive years (2015, 2016, and 2017). Zn concentrations of each individual wheat genotype for the complete panel of 369 genotypes and for the subpanel with 183 genotypes are presented in **Supplementary Table S1**. The phenotypic distribution of the Zn concentrations in the individual years appeared to be normally distributed (**Supplementary Figure S1**). A wide range of variation in the Zn concentration was observed for the complete panel (**Figure 1A**) and the subpanel (**Figure 1B**) in all 3 years and most of the variation within the complete panel was also captured in the subpanel (**Figure 2A** and **Table 1**). The results of BLUEs across 3 years' data ranged from 25.05 to 52.67 $\mu g\ g^{-1}$DW with a mean of 34.92 $\mu g\ g^{-1}$ DW. The genotype "Haven" had the highest Zn concentration equaling 52.67 $\mu g\ g^{-1}$ DW in the complete panel of wheat grain genotypes based on the BLUEs (**Figure 2B**). A significant positive Pearson's correlation ranging from $r = 0.18$ to 0.39 ($P < 0.001$) among the years (**Figure 2C**) indicated a relatively stable measurement of the phenotypes. A significant positive Pearson's correlation was found between Zn and TKW in all 3 years (**Supplementary Figure S2**). The broad-sense heritability for Zn concentration across the years was $H^2 = 0.54$. The results of ANOVA for Zn concentration indicated significant effects of genotype and environment, i.e., years (**Supplementary Table S2**).

Association Mapping of Grain Zn Concentrations in Two Diverse Wheat Panels

Genome-wide association mapping was performed for the complete panel and subpanel of wheat genotypes with Zn concentration data for each individual year in addition to BLUEs, using the implemented MLM with applying principal component

[2]http://www.traitgenetics.com

[3]http://www.cerealsdb.uk.net/

[4]https://urgi.versailles.inra.fr/WheatMine/begin.do

[5]http://www.cerealsdb.uk.net

[6]http://www.galaxyproject.org/

[7]https://megante.dna.affrc.go.jp/

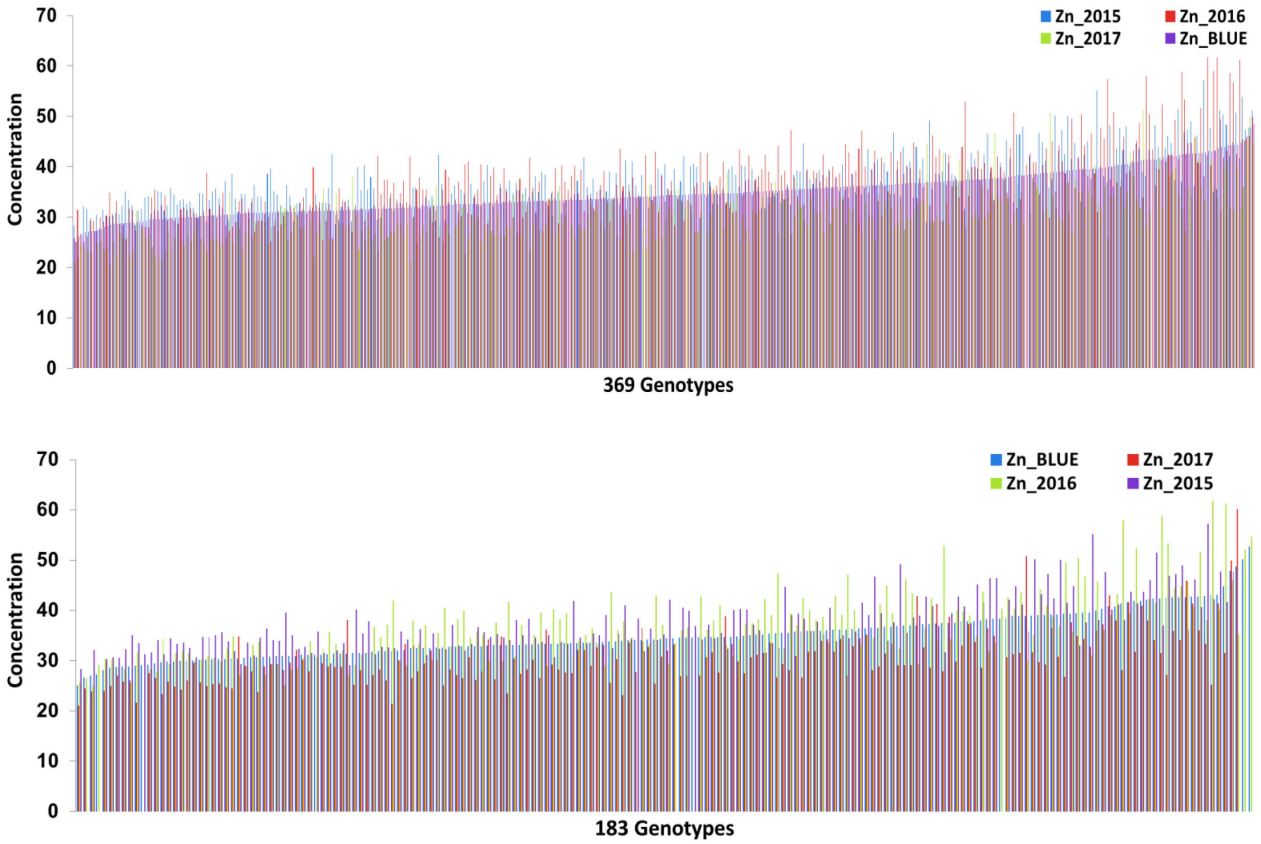


FIGURE 1 | (A) Zn concentration distribution for the whole wheat genotypes panel (369) in the 3 years (2015/2016/2017) and BLUE. **(B)** Zinc concentration distribution for the subpanel of wheat genotypes (183) in the 3 years (2015/2016/2017) and BLUE.

analysis (PCA) as a correction factor for population structure. The complete panel of wheat genotypes was analyzed by a combination of markers from the 90K iSELECT INFINIUM array and the 35K Affymetrix array resulting in 15,523 polymorphic SNP markers which were anchored in a genetic reference map. The subpanel was analyzed by merging 90K iSELECT array, 35K and 135k Affymetrix arrays resulting in a total of 44,233 polymorphic SNP markers based on their physical locations in order to increase the density of markers, achieve good mapping resolution, and to further enhance the power of GWAS output within the germplasm panel. Significant MTAs were detected above the threshold of –log10 (P-value) ≥ 3 as shown in Manhattan plots for both panels (**Figures 3A**, **4A**). The GWAS results were presented along with the QQ plots for SNPs, revealing that the distributions of observed association P-values were close to the distribution of expected associations (**Figures 3B**, **4B**). A total of 40 MTAs were detected in the complete panel on chromosomes 2A, 3A, 3B, 4A, 4D, 5A, 5B, 5D, 6D, 7A, 7B, and 7D with R^2-values ranging from 2.5 to 5.2%. A total of 21 MTAs had positive effects related to the minor allele and 19 MTAs had negative effects (**Supplementary Table S3**). While most MTAs were only detected in 1 year, an MTA on chromosome 3B was detected in all 3 years in similar mapping locations of 64.5 to 66.8 cM. The most significant MTA was detected on chromosome 5A with −log (p) value equaling 4.87 in the genomic region of 114.5 cM and explaining an R^2 value of 5.2%. The number of MTAs in the subpanel was increased to 161 including 31 unmapped markers on chromosomes 1A, 1B, 2A, 2B, 3A, 3B, 3D, 4A, 4D, 5A, 5B, 6A, 6B, 7A, and 7B with R^2-values ranging from 5.5 to 13.7% (**Supplementary Table S4**). A genomic region on chromosome 3B between the physical location of 716,993,339 and 736,712,355 (IWGSC RefSeq v1.0) is defined by 26 MTAs in the years 2016, 2017 and BLUEs with the highest R^2 of 11.3% at AX-95129199. A continuous range of 27 significant MTAs was detected on chromosome 5A ranging from physical location 353,989,023–698,510,016 including all 3 years and BLUEs. The most significant marker AX-158550766 located at position 464,479,275 explained 12.3% of phenotypic variation. A total of six markers for chromosome 3B (64.5–66.8 cM) and two markers for chromosome 5A (98.1–114.5 cM) were shared between the complete panel of varieties and the subpanel.

Defining Physical Regions of Candidate Genes Underlying Zn Accumulation in Wheat Grains

The highly significant SNP markers that located on chromosome 3B and 5A (**Figure 5**) were selected for BLAST analysis,

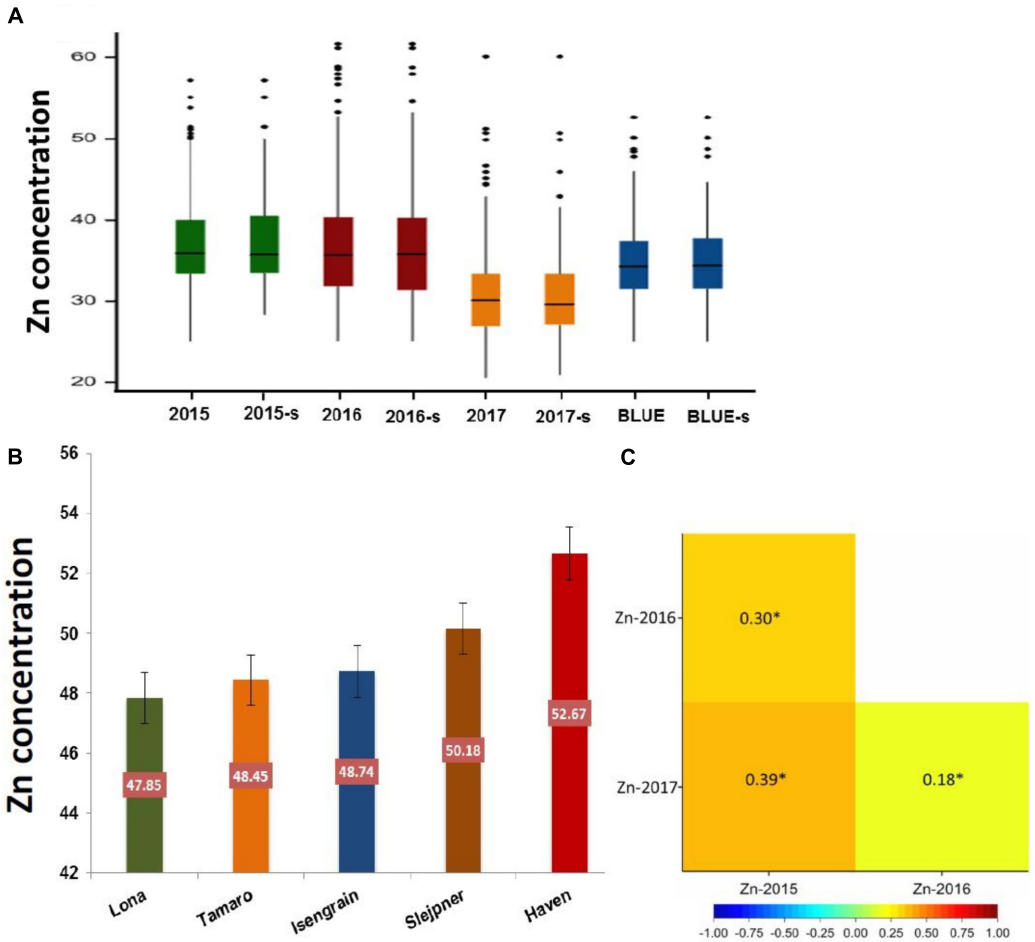


FIGURE 2 | (A) The boxplots of Zn concentration for the whole panel (369 genotypes) and a subpanel (183 genotypes) among 3 years and BLUE. **(B)** The scale of the top five genotypes with the highest Zn concentration value based on BLUE values. **(C)** Person correlation between years.

TABLE 1 | Grain Zn concentration mean, median, minimum, and maximum values within the complete and subpanel of wheat genotypes for each individual year.

Year-panel	Observation number	Mean μg g^{-1}	Median μg g^{-1}	Minimum μg g^{-1}	Maximum μg g^{-1}
2015-Complete panel	358	37.10	35.90	25.00	57.17
2015-Subpanel	176	37.34	35.75	28.34	57.17
2016-Complete panel	365	36.92	35.70	25.10	61.70
2016-Subpanel	180	36.83	35.80	25.10	61.70
2017-Complete panel	359	30.61	30.12	20.56	60.09
2017-Subpanel	176	30.80	29.64	21.00	60.09

using the web-based platform Galaxy[6]. The physical region of these SNPs at chromosome 3B located between 723,504,241 to 723,611,488 bp and for 5A on 462,763,758 to 466,582,184 bp (**Figure 5**) that were queried against IWGSC RefSeq v1.0. The fetched sequence output from Galaxy was submitted to MEGANTE[7], which is a web-based system for integrated plant genome annotation to perform genome annotations. On chromosomes 3B and 5A, we found a number of genes encoding proteins with known functions and others reported as hypothetical proteins (**Supplementary Table S5**). Putative candidate genes based on their function included a transcription factor (TF) belonging to the basic leucine zipper (bZIP) family and the TF bHLH76, a homeobox-leucine zipper protein HOX4, a SWAP (suppressor-of-white-apricot)/surp domain-containing protein and several genes related to the mitogen-activated protein kinase (MAPK) gene family (**Table 2**). Thus, we conclude that these two genomic regions on chromosomes 3B and 5A harbor a number of putative candidate genes, which may have a significant role in the process of grain Zn accumulation.

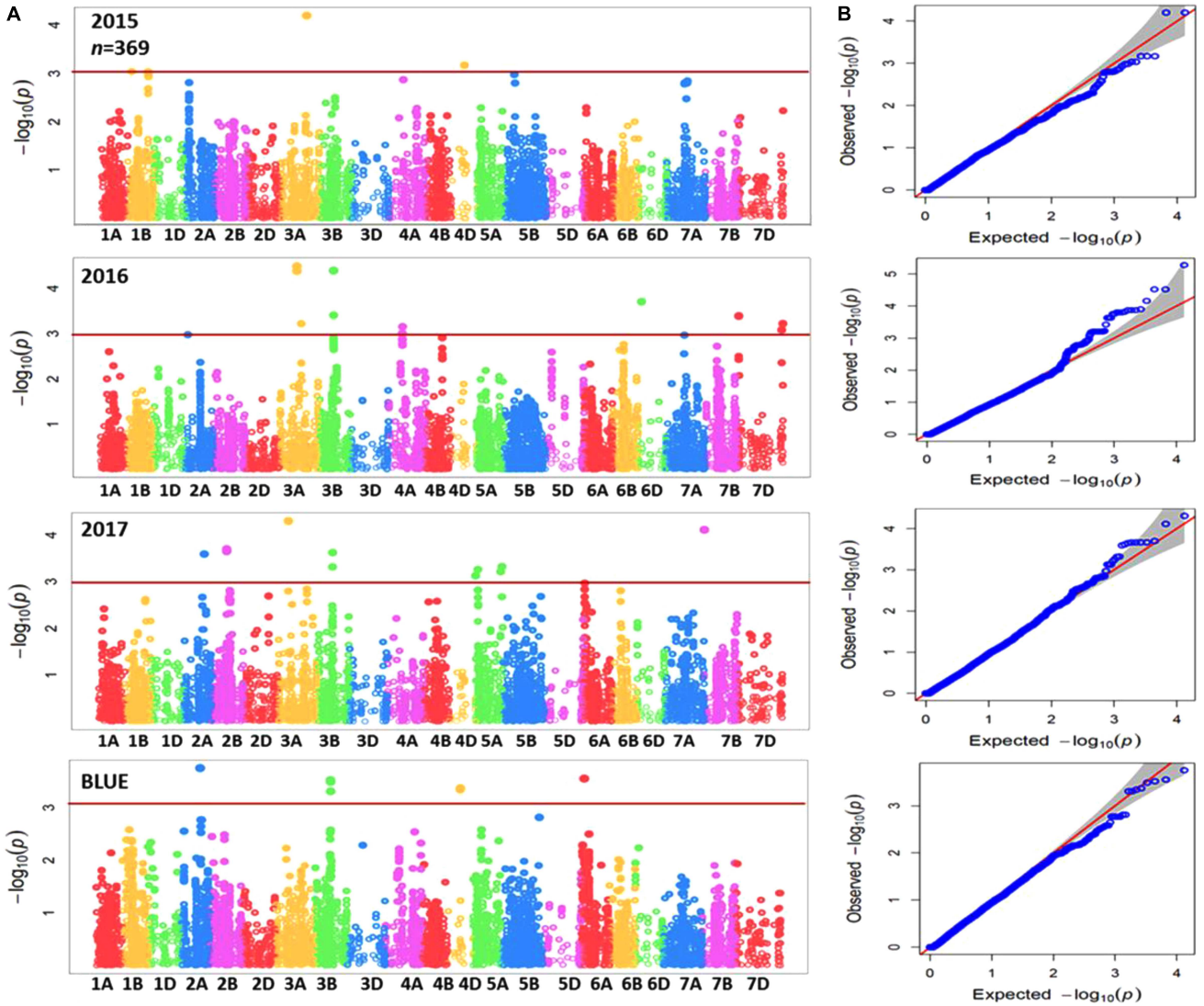


FIGURE 3 | (A) Summary of genome-wide association plots output of Zn concentration for the whole panel of wheat genotypes (369) which analyzed by using 90k and35k for each year (2015/2016/2017) and BLUEs with mixed linear model and PCA model. **(B)** Quantile-quantile scale representing expected versus observed *P*-values at −log10 (*P*).

DISCUSSION

Wide Variation for Zn Accumulation in Wheat Grains

The poor bioavailability of essential nutrients in cereal grains leads the breeders to use plant breeding which is a seed-based approach to develop cultivars with improved and adequate levels of nutrients (Tiwari et al., 2016). Plant breeding or genetic biofortification was found to be comparative with other costly and non-sustainable approaches such as agronomic biofortification which is based on using fertilizers or other approaches that are based on food fortification and daily consumed supplementations (Garcia-Oliveira et al., 2018). Therefore, genetic biofortification is considered as one of the vital approaches that can help to overcome malnutrition problems either by classical plant breeding or approaches involving GMOs (genetically modified organisms) (Borrill et al., 2014; Singh et al., 2017). Many reports mentioned that the targeted range for biofortified grains and to develop cultivars with high Zn concentration is between 40–50 $\mu g\ g^{-1}$ (Howarth et al., 2011; Cakmak and Kutman, 2017). The phenotypic variation that found in our germplasm ranged from 25.05–52.65 $\mu g\ g^{-1}$ which is compatible with the target range and provides the chance to use the highest grain Zn-containing genotypes in breeding programs. Similar Zn concentration ranges were also reported by Graham et al. (1999) and Guttieri et al. (2015) who found that Zn concentrations in 132 bread wheat genotypes ranged between 25–53 and 13.1–45.2 $\mu g\ g^{-1}$ in hexaploid wheat. Additionally, in

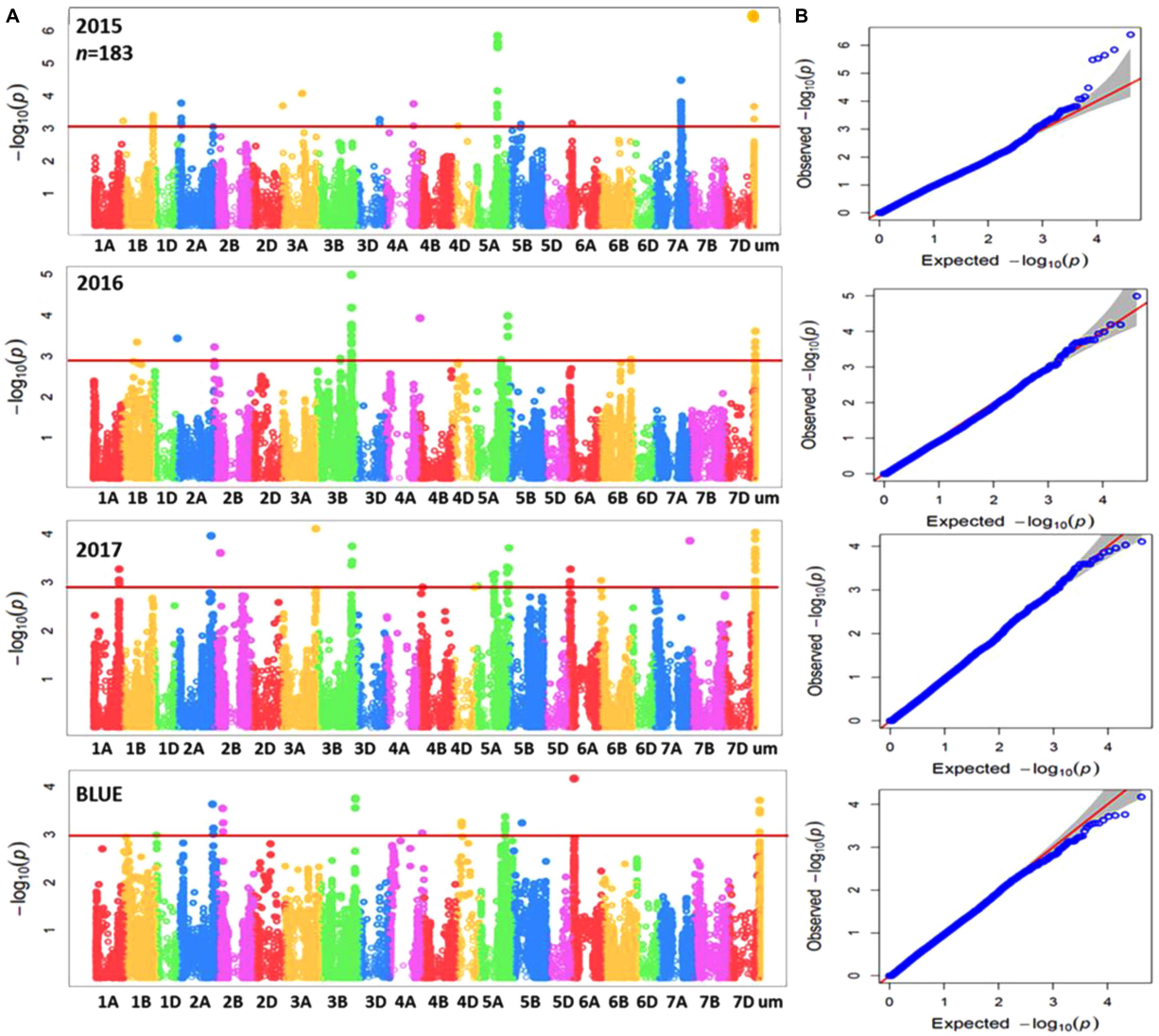


FIGURE 4 | (A) Summary of genome-wide association plots output of Zn concentration for the subpanel of wheat genotypes (183) which analyzed by 90k, 35k, and 135k for each year (2015/2016/2017) and BLUEs using mixed linear model and PCA model. **(B)** Quantile-quantile scale representing expected versus observed *P*-values at −log10 (*P*).

durum wheat, the variation ranged from 24.8–48.8 $\mu g\ g^{-1}$ for Zn which is comparable with our observations (Magallanes-Lopez et al., 2017).

Grain Zn concentrations across years were weakly to moderately correlated (r = 0.18–0.39; $P < 0.001$) which may be attributed to environmental effects across years and its interaction with genotypes as was also reported for grains of other crops (Singh et al., 2017). The calculated heritability (H^2 = 0.54) for our trait of interest represents a moderate contribution of the genotype to the overall variation in grain Zn concentration, which was also affected by the environment in experiments being conducted across 3 years. Similarly, Tiwari et al. (2016) and Khokhar et al. (2018) found the moderate effect of genotypes on wheat grain Zn concentrations. We observed a significant positive correlation between Zn and TKW, which implies that both traits improve simultaneously each other and this observation has also been made in other studies with wheat (Morgounov et al., 2007; Peleg et al., 2009; Krishnappa et al., 2017). Our findings provide a list of improved cultivars with high Zn concentration that can be utilized in future breeding programs for boosting grain quality.

Zn Grain Concentration as a Complex Trait

Genetic dissection for grain Zn concentration in our diversity panel showed that this trait is under control of many genetic loci. The constant significant MTAs across the years 2015, 2016, 2017 and BLUE values are conferred by loci

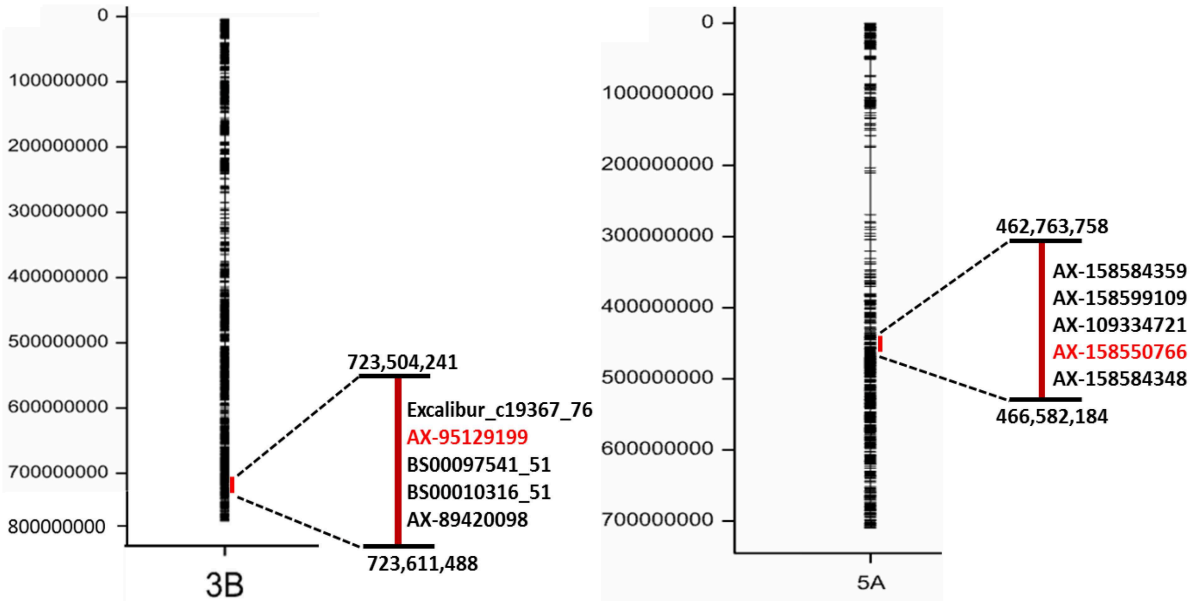


FIGURE 5 | Alignment of the significant SNP markers (in black) to chromosome 3BS and 5AL.The most significant SNP (in red) with –log *(P)* value equaling 5.84.

TABLE 2 | Summary of putative candidate genes associated with Zn grain concentration in wheat.

Gene number	Description	GO ID	GO term	GO category	Chr.
mRNA_2.1	Mitogen- activated	GO:0004672	Protein kinase activity,	Molecular	3B
mRNA_3.1	protein kinase kinase	GO:0006468	protein phosphorylation,	function,	
mRNA_10.1	kinase	GO:0015031	protein transport	biological	
mRNA_23.1				process	
mRNA_24.1					
mRNA_32.1	SWAP (suppressor-of-white-APRICOT)/surp domain-containing protein	GO:0003723	RNA binding	Molecular function	3B
mRNA_11.1	Homeobox-leucine zipper protein HOX4	GO:0003677/ GO:0003700	DNA binding/DNA binding transcription factor activity	Molecular function	5A
mRNA_34.1	Protein FAR1-RELATED SEQUENCE 11 (FRS11)	GO:0008270	Zinc ion binding	Molecular function	5A
mRNA_42.1	BZIP protein	GO:0003700	DNA binding transcription factor activity	Molecular function	5A
mRNA_44.1	Transcription factor bHLH76	GO:0046983	Protein dimerization activity	Molecular function	5A

on chromosomes 3B (723,504,241–723,611,488 bp) and 5A (462,763,758–466,582,184 bp) in the complete panel as well as in the subpanel of wheat genotypes (**Figure 5**). Previously, a QTL for Zn concentration was reported in a similar location on chromosome 3B by Crespo-Herrera et al. (2017) in a population of hexaploid wheat RILs. Another study detected on chromosome 5A a QTL for grain Zn concentration in a RIL population derived from a cross between durum wheat and wild emmer wheat (Peleg et al., 2009). A previous study in rice mentioned that many QTLs for grain Zn have been mapped based on eight different mapping populations, where the most constant QTL for grain Zn content across environments was located on chromosome 12 (Swamy et al., 2016), which has synteny to chromosome 5A in wheat (Salse et al., 2009). Therefore, our results indicate potential genomic regions controlling Zn in wheat that can be used in further genetic investigations.

Identification of Candidate Genes

The gene content of the two genomic regions on chromosomes 3B and 5A harbors many hypothetical and functionally annotated genes or proteins including TFs and transporter proteins (**Supplementary Table S5**). For instance, we found five genes on chromosome 3B related to the MAPK family (**Table 2**) and this gene is well documented in biotic and abiotic stress signaling (Xu and Zhang, 2015). Recently, several publications reported that different MAPK genes play major roles in sugar, nitrogen, phosphate, iron, potassium, or Zn signaling pathways (Lastdrager et al., 2014; Briat et al., 2015; Chardin et al., 2017), which makes them promising candidates for being involved in grain Zn accumulation. The gene annotations of the MEGANTE pipeline showed that one of the MAPK-related genes encoded a vacuolar protein sorting-associated protein. Interestingly, a recent report showed that vacuolar protein sorting-associated protein was

identified as one of the candidate genes mediating elevated Zn concentrations in chickpea seeds (Upadhyaya et al., 2016). In the same study, a SWAP/surp domain-containing protein was reported to be linked with seed Zn concentration in chickpea and a SWAP was found in the present study as putative candidate gene on chromosome 3B (**Table 2**).

On chromosome 5A, a homeobox-leucine zipper protein HOX4 that annotated as *TaHDZIP1* was found to be associated with grain Zn concentrations in the used panel. Another regulatory element detected in this genomic region is the putative TF *bHLH76* and it has been reported that bHLH is one of the binding factors of the cis-element G-box which was found in promoter regions of all *TaMTPs* (metal tolerance proteins), which are involved in trace metal homeostasis and have a potential role in cereal grain biofortification with essential micronutrients including Zn (Menguer et al., 2017; Vatansever et al., 2017). Additionally, we found that chromosome 5A harbored a TF belonging to the bZIP (basic-region leucine-zipper) family which have a crucial role in nutrient and Zn homeostasis (Ishimaru et al., 2011; Evens et al., 2017; Cifuentes-Esquivel et al., 2018). In *Arabidopsis thaliana,* the TFs *bZIP19* and *bZIP23* were shown to regulate the adaption to Zn deficiency in roots (Assunção et al., 2010; Inaba et al., 2015). A total of 187 *TabZIP* genes have been identified in wheat (Li et al., 2015) and a specific group of *TabZIP* genes conferred functional complementation of Zn deficiency-hypersensitive such as *bzip19 bzip23* (Evens et al., 2017; Henríquez-Valencia et al., 2018). So far, most functional studies of bZIPs were related to roots or leaves while little information about bZIP-dependent regulatory mechanisms is available for grains. Therefore, novel bZIP genes could play a critical role in improving Zn accumulation in grains. However, this requires further genetic and functional validation. Finally, the FAR1 protein detected on chromosome 5A (**Supplementary Table S5**) and its molecular function based on gene ontology analysis is related to zinc ion binding (**Table 2**), which also makes it a potential candidate gene.

CONCLUSION

The present analysis showed the power of the GWAS approach for identifying putative candidate genes for grain Zn accumulation in wheat. This study discovered genetic factors controlling grain Zn accumulation that may establish the basis for further breeding and genetic work in cereals. Two physically anchored chromosomal segments 3B and 5A harbor many putative candidate genes like MAPK and bZIP genes which are proposed as candidates conferring enhanced grain Zn concentrations. Further validation and functional characterization are required to elucidate the role of these genes for Zn homeostasis in wheat.

AUTHOR CONTRIBUTIONS

DA performed the data analysis including genome-wide association scan, candidate genes identification, and statistical analysis. KE and NvW participated in Zn concentration measurements. AA helped in manuscript modification and statistical analysis. KP and MR designed the experiment. MR conceived the idea and participated in the interpretation of results. DA and MR wrote the manuscript. All authors read and approved the final manuscript.

FUNDING

The project was funded by the internal financial support of IPK Gatersleben.

ACKNOWLEDGMENTS

We thank Ellen Weiß and Yudelsy Antonia Tandron Moya for excellent technical assistance. The genotyping data were created in the frame of the projects VALID and SELECT (project numbers 0314947 and 0315949) funded by the Plant Biotechnology Program of the German Federal Ministry of Education and Research (BMBF). We are grateful for the IWGSC for pre-publication access to IWGSC RefSeq v1.0 for data analysis during the development of this manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2018.01313/full#supplementary-material

REFERENCES

Allen, A. M., Winfield, M. O., Burridge, A. J., Downie, R. C., Benbow, H. R., Barker, G. L., et al. (2017). Characterization of a wheat breeders' array suitable for high-throughput snp genotyping of global accessions of hexaploid bread wheat (*Triticum aestivum*). *Plant Biotechnol. J.* 15, 390–401. doi: 10.1111/pbi.12635

Alomari, D. Z., Eggert, K., von Wiren, N., Pillen, K., and Roder, M. S. (2017). Genome-wide association study of calcium accumulation in grains of european wheat cultivars. *Front. Plant Sci.* 8:1797. doi: 10.3389/fpls.2017.01797

Assunção, A. G. L., Herrero, E., Lin, Y.-F., Huettel, B., Talukdar, S., Smaczniak, C., et al. (2010). Arabidopsis thaliana transcription factors bZIP19 and bZIP23 regulate the adaptation to zinc deficiency. *Proc. Natl. Acad. Sci. U.S.A.* 107, 10296–10301. doi: 10.1073/pnas.1004788107

Borrill, P., Connorton, J., Balk, J., Miller, T., Sanders, D., and Uauy, C. (2014). Biofortification of wheat grain with iron and zinc: integrating novel genomic resources and knowledge from model crops. *Front. Plant Sci.* 5:53. doi: 10.3389/fpls.2014.00053

Briat, J. F., Rouached, H., Tissot, N., Gaymard, F., and Dubos, C. (2015). Integration of P, S, Fe, and Zn nutrition signals in *Arabidopsis thaliana*: potential involvement of phosphate starvation response 1 (PHR1). *Front. Plant Sci.* 6:290. doi: 10.3389/fpls.2015.00290

Cakmak, I. (2008). Enrichment of cereal grains with zinc: agronomic or genetic biofortification? *Plant Soil* 302, 1–17. doi: 10.1007/s11104-007-9466-3

Cakmak, I., and Kutman, U. B. (2017). Agronomic biofortification of cereals with zinc: a review. *Eur. J. Soil Sci.* 69, 172–180. doi: 10.1111/ejss.12437

Chardin, C., Schenk, S. T., Hirt, H., Colcombet, J., and Krapp, A. (2017). Review: mitogen-activated protein kinases in nutritional signaling in Arabidopsis. *Plant Sci.* 260, 101–108. doi: 10.1016/j.plantsci.2017.04.006

Cifuentes-Esquivel, N., Celiz-Balboa, J., Henriquez-Valencia, C., Mitina, I., Arra no-Salinas, P., Moreno Adrián, A., et al. (2018). bZIP17 regulates the expression of genes related to seed storage and germination, reducing seed susceptibility to osmotic stress. *J. Cell. Biochem.* 119, 6857–6868. doi: 10.1002/jcb.26882

Crespo-Herrera, L. A., Govindan, V., Stangoulis, J., Hao, Y., and Singh, R. P. (2017). QTL mapping of grain zn and fe concentrations in two hexaploid wheat RIL populations with ample transgressive segregation. *Front. Plant Sci.* 8:1800. doi: 10.3389/fpls.2017.01800

Crespo-Herrera, L. A., Velu, G., and Singh, R. P. (2016). Quantitative trait loci mapping reveals pleiotropic effect for grain iron and zinc concentrations in wheat. *Ann. Appl. Biol.* 169, 27–35. doi: 10.1111/aab.12276

Evens, N. P., Buchner, P., Williams, L. E., and Hawkesford, M. J. (2017). The role of ZIP transporters and group F bZIP transcription factors in the Zn-deficiency response of wheat (*Triticum aestivum*). *Plant J.* 92, 291–304. doi: 10.1111/tpj.13655

FAOSTAT (2012). *Crop Statistics*. Available at: http://www.fao.org/faostat

FAOSTAT (2016). *Crop Statistics*. Available at: http://www.fao.org/faostat

Garcia-Oliveira, A. L., Chander, S., Ortiz, R., Menkir, A., and Gedil, M. (2018). Genetic basis and breeding perspectives of grain iron and zinc enrichment in cereals. *Front. Plant Sci.* 9:937. doi: 10.3389/fpls.2018.00937

Genc, Y., Verbyla, A. P., Torun, A. A., Cakmak, I., Willsmore, K., Wallwork, H., et al. (2008). Quantitative trait loci analysis of zinc efficiency and grain zinc concentration in wheat using whole genome average interval mapping. *Plant Soil* 314:49. doi: 10.1007/s11104-008-9704-3

Graham, R., Senadhira, D., Beebe, S., Iglesias, C., and Monasterio, I. (1999). Breeding for micronutrient density in edible portions of staple food crops: conventional approaches. *Field Crops Res.* 60, 57–80. doi: 10.1016/S0378-4290(98)00133-6

Guttieri, M. J., Baenziger, P. S., Frels, K., Carver, B., Arnall, B., Wang, S., et al. (2015). Prospects for selecting wheat with increased zinc and decreased cadmium concentration in grain. *Crop Sci.* 55, 1712–1728. doi: 10.2135/cropsci2014.08.0559

Hamblin, M. T., Buckler, E. S., and Jannink, J.-L. (2011). Population genetics of genomics-based crop improvement methods. *Trends Genet.* 27, 98–106. doi: 10.1016/j.tig.2010.12.003

Henríquez-Valencia, C., Arenas-M, A., Medina, J., and Canales, J. (2018). Integrative transcriptomic analysis uncovers novel gene modules that underlie the sulfate response in *Arabidopsis thaliana*. *Front. Plant Sci.* 9:470. doi: 10.3389/fpls.2018.00470

Howarth, E. B., Christine, H., Bonnie, M., Meenakshi, J. V., and Wolfgang, H. P. (2011). Biofortification: a new tool to reduce micronutrient malnutrition. *Food Nutr. Bull.* 32, S31–S40. doi: 10.1177/15648265110321S105

Inaba, S., Kurata, R., Kobayashi, M., Yamagishi, Y., Mori, I., Ogata, Y., et al. (2015). Identification of putative target genes of bZIP19, a transcription factor essential for Arabidopsis adaptation to Zn deficiency in roots. *Plant J.* 84, 323–334. doi: 10.1111/tpj.12996

Ishimaru, Y., Bashir, K., and Nishizawa, N. K. (2011). Zn uptake and translocation in rice plants. *Rice* 4, 21–27. doi: 10.1007/s12284-011-9061-3

Kambe, T., Hashimoto, A., and Fujimoto, S. (2014). Current understanding of ZIP and ZnT zinc transporters in human health and diseases. *Cell. Mol. Life Sci.* 71, 3281–3295. doi: 10.1007/s00018-014-1617-0

Khokhar, J. S., Sareen, S., Tyagi, B. S., Singh, G., Wilson, L., King, I. P., et al. (2018). Variation in grain Zn concentration, and the grain ionome, in field-grown Indian wheat. *PLoS One* 13:e0192026. doi: 10.1371/journal.pone.0192026

Kollers, S., Rodemann, B., Ling, J., Korzun, V., Ebmeyer, E., Argillier, O., et al. (2013). Whole genome association mapping of fusarium head blight resistance in European winter wheat (*Triticum aestivum* L.). *PLoS One* 8:e57500. doi: 10.1371/journal.pone.0057500

Krishnappa, G., Singh, A. M., Chaudhary, S., Ahlawat, A. K., Singh, S. K., Shukla, R. B., et al. (2017). Molecular mapping of the grain iron and zinc concentration, protein content and thousand kernel weight in wheat (*Triticum aestivum* L.). *PLoS One* 12:e0174972. doi: 10.1371/journal.pone.0174972

Lastdrager, J., Hanson, J., and Smeekens, S. (2014). Sugar signals and the control of plant growth and development. *J. Exp. Bot.* 65, 799–807. doi: 10.1093/jxb/ert474

Li, X., Gao, S., Tang, Y., Li, L., Zhang, F., Feng, B., et al. (2015). Genome-wide identification and evolutionary analyses of bZIP transcription factors in wheat and its relatives and expression profiles of anther development related *TabZIP* genes. *BMC Genomics* 16:976. doi: 10.1186/s12864-015-2196-7

Lipka, A. E., Tian, F., Wang, Q., Peiffer, J., Li, M., Bradbury, P. J., et al. (2012). GAPIT: genome association and prediction integrated tool. *Bioinformatics* 28, 2397–2399. doi: 10.1093/bioinformatics/bts444

Magallanes-Lopez, A. M., Hernandez-Espinosa, N., Velu, G., Posadas-Romano, G., Ordonez-Villegas, V. M. G., Crossa, J., et al. (2017). Variability in iron, zinc and phytic acid content in a worldwide collection of commercial durum wheat cultivars and the effect of reduced irrigation on these traits. *Food Chem.* 237, 499–505. doi: 10.1016/j.foodchem.2017.05.110

Menguer, K., Vincent, T., Miller Anthony, J., Brown James, K. M., Vincze, E., Borg, S., et al. (2017). Improving zinc accumulation in cereal endosperm using HvMTP1, a transition metal transporter. *Plant Biotechnol. J.* 16, 63–71. doi: 10.1111/pbi.12749

Morgounov, A., Gómez-Becerra, H. F., Abugalieva, A., Dzhunusova, M., Yessimbekova, M., Muminjanov, H., et al. (2007). Iron and zinc grain density in common wheat grown in Central Asia. *Euphytica* 155, 193–203. doi: 10.1007/s10681-006-9321-2

Ozturk, L., Yazici, M. A., Yucel, C., Torun, A., Cekic, C., Bagci, A., et al. (2006). Concentration and localization of zinc during seed development and germination in wheat. *Physiol. Plant.* 128, 144–152. doi: 10.1111/j.1399-3054.2006.00737.x

Pandey, A., Khan, M. K., Hakki, E. E., Thomas, G., Hamurcu, M., Gezgin, S., et al. (2016). Assessment of genetic variability for grain nutrients from diverse regions: potential for wheat improvement. *Springerplus* 5:1912. doi: 10.1186/s40064-016-3586-2

Peleg, Z., Cakmak, I., Ozturk, L., Yazici, A., Jun, Y., Budak, H., et al. (2009). Quantitative trait loci conferring grain mineral nutrient concentrations in durum wheat × wild emmer wheat RIL population. *Theor. Appl. Genet.* 119, 353–369. doi: 10.1007/s00122-009-1044-z

Poland, J. A., Brown, P. J., Sorrells, M. E., and Jannink, J.-L. (2012). Development of high-density genetic maps for barley and wheat using a novel two-enzyme genotyping-by-sequencing approach. *PLoS One* 7:e32253. doi: 10.1371/journal.pone.0032253

Pu, Z.-E., Yu, M., He, Q.-Y., Chen, G.-Y., Wang, J.-R., Liu, Y.-X., et al. (2014). Quantitative trait loci associated with micronutrient concentrations in two recombinant inbred wheat lines. *J. Integr. Agric.* 13, 2322–2329.

Salse, J., Abrouk, M., Bolot, S., Guilhot, N., Courcelle, E., Faraut, T., et al. (2009). Reconstruction of monocotelydoneous proto-chromosomes reveals faster evolution in plants than in animals. *Proc. Natl. Acad. Sci. U.S.A.* 106, 14908–14913. doi: 10.1073/pnas.0902350106

Sands, D. C., Morris, C. E., Dratz, E. A., and Pilgeram, A. (2009). Elevating optimal human nutrition to a central goal of plant breeding and production of plant-based foods. *Plant Sci.* 117, 377–389. doi: 10.1016/j.plantsci.2009.07.011

Shi, R., Li, H., Tong, Y., Jing, R., Zhang, F., and Zou, C. (2008). Identification of quantitative trait locus of zinc and phosphorus density in wheat (*Triticum aestivum* L.) grain. *Plant Soil* 306, 95–104.

Shi, R.-L., Tong, Y.-P., Jing, R.-L., Zhang, F.-S., and Zou, C.-Q. (2013). Characterization of quantitative trait loci for grain minerals in hexaploid wheat (*Triticum aestivum* L.). *J. Integr. Agric.* 12, 1512–1521.

Singh, A., Sharma, V., Dikshit, H. K., Aski, M., Kumar, H., Thirunavukkarasu, N., et al. (2017). Association mapping unveils favorable alleles for grain iron and zinc concentrations in lentil (*Lens culinaris* subsp. *culinaris*). *PLoS One* 12:e0188296. doi: 10.1371/journal.pone.0188296

Sorrells, M. E., Gustafson, J. P., Somers, D., Chao, S., Benscher, D., Guedira-Brown, G., et al. (2011). Reconstruction of the synthetic W7984 × Opata M85 wheat reference population. *Genome* 54, 875–882. doi: 10.1139/g11-054

Srinivasa, J., Arun, B., Mishra, V. K., Singh, G. P., Velu, G., Babu, R., et al. (2014). Zinc and iron concentration QTL mapped in a Triticum spelta × *T. aestivum* cross. *Theor. Appl. Genet.* 127, 1643–1651. doi: 10.1007/s00122-014-2327-6

Swamy, B. P. M., Rahman, M. A., Inabangan-Asilo, M. A., Amparado, A., Manito, C., Chadha-Mohanty, P., et al. (2016). Advances in breeding for high grain Zinc in Rice. *Rice* 9:49. doi: 10.1186/s12284-016-0122-5

Tiwari, C., Wallwork, H., Arun, B., Mishra, V. K., Velu, G., Stangoulis, J., et al. (2016). Molecular mapping of quantitative trait loci for zinc, iron and protein content in the grains of hexaploid wheat. *Euphytica* 207, 563–570. doi: 10.1007/s10681-015-1544-7

Tiwari, V. K., Rawat, N., Chhuneja, P., Neelam, K., Aggarwal, R., Randhawa, G. S., et al. (2009). Mapping of quantitative trait loci for grain iron and zinc concentration in diploid a genome wheat. *J. Hered.* 100, 771–776. doi: 10.1093/jhered/esp030

Upadhyaya, H. D., Bajaj, D., Das, S., Kumar, V., Gowda, C. L., Sharma, S., et al. (2016). Genetic dissection of seed-iron and zinc concentrations in chickpea. *Sci. Rep.* 6:24050. doi: 10.1038/srep24050

Vatansever, R., Filiz, E., and Eroglu, S. (2017). Genome-wide exploration of metal tolerance protein (MTP) genes in common wheat (*Triticum aestivum*): insights into metal homeostasis and biofortification. *Biometals* 30, 217–235. doi: 10.1007/s10534-017-9997-x

Wang, S., Wong, D., Forrest, K., Allen, A., Chao, S., Huang, B. E., et al. (2014). Characterization of polyploid wheat genomic diversity using a high-density 90,000 single nucleotide polymorphism array. *Plant Biotechnol. J.* 12, 787–796. doi: 10.1111/pbi.12183

Welch, R. M., and Graham, R. D. (2004). Breeding for micronutrients in staple food crops from a human nutrition perspective. *J. Exp. Bot.* 55, 353–364. doi: 10.1093/jxb/erh064

Xu, J., and Zhang, S. (2015). Mitogen-activated protein kinase cascades in signaling plant growth and development. *Trends Plant Sci.* 20, 56–64. doi: 10.1016/j.tplants.2014.10.001

Xu, Y., An, D., Liu, D., Zhang, A., Xu, H., and Li, B. (2012). Molecular mapping of QTLs for grain zinc, iron and protein concentration of wheat across two environments. *Field Crops Res.* 138, 57–62.

Yu, J., Pressoir, G., Briggs, W. H., Vroh Bi, I., Yamasaki, M., Doebley, J. F., et al. (2006). A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. *Nat. Genet.* 38, 203–208.

Zanke, C. D., Rodemann, B., Ling, J., Muqaddasi, Q. H., Plieske, J., Polley, A., et al. (2017). Genome-wide association mapping of resistance to eyespot disease (*Pseudocercosporella herpotrichoides*) in European winter wheat (*Triticum aestivum* L.) and fine-mapping of Pch1. *Theor. Appl. Genet.* 130, 505–514. doi: 10.1007/s00122-016-2830-z

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

2.3 Whole-Genome Association Mapping and Genomic Prediction for Iron Concentration in Wheat Grains

Published in: International Journal of Molecular Sciences (2019) 20(1), 76.

DOI: 10.3390/ijms20010076

Authors: Dalia Z. Alomari*, Kai Eggert, Nicolaus von Wirén, Andreas Polley, Jörg Plieske, Martin W. Ganal, Fang Liu, Klaus Pillen and Marion S. Röder

The original publication is available online at:

https://www.mdpi.com/1422-0067/20/1/76

International Journal of
Molecular Sciences

MDPI

Article

Whole-Genome Association Mapping and Genomic Prediction for Iron Concentration in Wheat Grains

Dalia Z. Alomari [1,*], Kai Eggert [1,†], Nicolaus von Wirén [1], Andreas Polley [2], Jörg Plieske [2], Martin W. Ganal [2], Fang Liu [1], Klaus Pillen [3] and Marion S. Röder [1]

1 Leibniz Institute of Plant Genetics and Crop Plant Research (IPK), Corrensstrasse 3, D-06466 Stadt Seeland, OT Gatersleben, Germany; eggert@ipk-gatersleben.de (K.E.); vonwiren@ipk-gatersleben.de (N.v.W.); liuf@ipk-gatersleben.de (F.L.); roder@ipk-gatersleben.de (M.S.R.)

2 SGS TraitGenetics GmbH, D-06466 Stadt Seeland, OT Gatersleben, Germany; andreas@traitgenetics.de (A.P.); joerg@traitgenetics.de (J.P.); ganal@traitgenetics.de (M.W.G.)

3 Institute of Agricultural and Nutritional Sciences, Martin-Luther-University Halle-Wittenberg, Betty-Heimann-Str. 3, 06120 Halle/Saale, Germany; klaus.pillen@landw.uni-halle.de

* Correspondence: alomari@ipk-gatersleben.de

† Deceased.

Received: 28 November 2018; Accepted: 20 December 2018; Published: 25 December 2018

Abstract: Malnutrition of iron (Fe) affects two billion people worldwide. Therefore, enhancing grain Fe concentration (GFeC) in wheat (*Triticum aestivum* L.) is an important goal for breeding. Here we study the genetic factors underlying GFeC trait by genome-wide association studies (GWAS) and the prediction abilities using genomic prediction (GP) in a panel of 369 European elite wheat varieties which was genotyped with 15,523 mapped single-nucleotide polymorphism markers (SNP) and a subpanel of 183 genotypes with 44,233 SNP markers. The resulting means of GFeC from three field experiments ranged from 24.42 to 52.42 $\mu g \cdot g^{-1}$ with a broad-sense heritability (H^2) equaling 0.59 over the years. GWAS revealed 41 and 137 significant SNPs in the whole and subpanel, respectively, including significant marker-trait associations (MTAs) for best linear unbiased estimates (BLUEs) of GFeC on chromosomes 2A, 3B and 5A. Putative candidate genes such as NAC transcription factors and transmembrane proteins were present on chromosome 2A (763,689,738–765,710,113 bp). The GP for a GFeC trait ranged from low to moderate values. The current study reported GWAS of GFeC for the first time in hexaploid wheat varieties. These findings confirm the utility of GWAS and GP to explore the genetic architecture of GFeC for breeding programs aiming at the improvement of wheat grain quality.

Keywords: Wheat; mineral; iron; GWAS; SNP; candidate genes

1. Introduction

Wheat is the second most produced and consumed food crop worldwide and its products form a fundamental diet in the daily life for people in the whole world (FAOSTAT 2016; http://faostat.fao.org). Wheat grains contain mainly carbohydrates with a small proportion of proteins and essential micronutrients such as iron (Fe) and zinc (Zn) [1–3]. Micronutrient deficiency including Fe and Zn are among the most prevalent deficiencies in the developing countries and high-risk groups are women and children [4]. More than 2 billion people are affected with Fe deficiency which has an adverse effect on health, such as retarding the physical growth and affecting the motoric development, leading to fatigue and low productivity [5,6]. Therefore, in regions where the people depend mostly on cereal-based foods, deficiencies in micronutrients become a challenge. On the other side, improving Fe concentrations in the edible part of crops are linked with positive consequences on both grain yield and nutritional status as well as a positive effect on human health [7].

Understanding the genetic basis of Fe concentration in wheat grains is imperative for enhancing Fe values in newly developed varieties. Therefore, we performed a genome-wide association study (GWAS) approach which is one of the main approaches for dissecting complex traits including nutritional quality traits that are controlled by many genes and influenced by the environment [8,9]. Several genetic regions controlling mineral concentration traits in wheat have been identified by applying traditional quantitative trait loci (QTL) analysis using bi-parental mapping populations. For instance, Peleg et al. [10] detected five QTLs on chromosomes 2A, 3B, 5A, 6B and 7A for Fe concentration in a tetraploid wild emmer × durum wheat recombinant inbred lines (RILs) population. Another study identified five QTLs underlying grain Fe concentration (GFeC) in a *Triticum spelta* × *T. aestivum* RIL population, of which three mapped to chromosome 1A while two QTLs mapped to chromosomes 2A and 3B [11]. To our knowledge two GWAS studies have been reported on synthetic wheat lines. Gorafi et al. [12] studied grain iron content in 47 synthetic hexaploid wheat germplasm lines and Bhatta et al. [13] performed GWAS for various grain minerals including Fe on 123 synthetic hexaploid wheat lines. Association mapping for seed Fe content was also performed for other crops such as pearl millet [14] and seed Fe concentration in chickpea [15]; however, no GWAS study on released wheat varieties is available to our knowledge.

Recently, genomic prediction (GP) or genomic selection (GS) approaches were developed based on genome-wide marker information to predict the breeding value of complex traits for which only genotyping data are provided (test population) [16]. These predicted values are called genome estimate breeding values (GEBVs) and are based on actual phenotypic data related to genotypes in a training population [17]. Several methods were adopted for GP or GS calculation such as Bayesian methods, rrBLUP and Genomic best linear unbiased prediction (GBLUP), while the main affecting factor within these methods is the density of the markers [18]. Application of GP will be helpful particularly for complex traits and for traits that are costly to phenotype; therefore, applying GP could speed up the genetic gains in the development of nutrient-dense wheat varieties. To date, numerous plant breeding studies have been published to investigate complex traits such as nutritional quality traits in wheat [19,20].

The goals of this study were (i) to study the natural phenotypic variation of wheat GFeC in a panel of 369 elite wheat varieties grown for three years in the field, (ii) to investigate the genetic architecture of this trait and to identify QTLs by applying a GWAS approach, (iii) to define the gene content in the respective genomic region of the wheat reference sequence as well as to identify potential candidate genes, and (iv) to examine the prediction ability in the present wheat panel by using different statistical models.

2. Results

2.1. Phenotypic Analysis and Correlations

The analyses of variance (ANOVA) for Fe concentration in grains showed a significant effect of both genotype and years ($p < 0.001$) (Table S1). Wide genetic variation of GFeCs was found between the genotypes in both the whole panel and subpanel in each year (Figure S1, Table S2). The genotypic variation of Fe concentrations in each year appeared to be normally distributed (Figure S2). The average of grain Fe based on BLUE values was about 34 $\mu g \cdot g^{-1}$ dry weight (DW) in the whole and subpanel of genotypes (Figure 1) with a range of 24.42–52.42 $\mu g \cdot g^{-1}$ DW in the whole panel and 26.99–48.52 $\mu g \cdot g^{-1}$ DW in the subpanel of genotypes (Figure 1). This trend of GFeC decrease among years may be attributed to environmental effects including rain fall and temperature (Figure S3). In fact, this conclusion was supported by the resulting heritability for Fe concentration across the years which is equal to $H^2 = 0.59$.

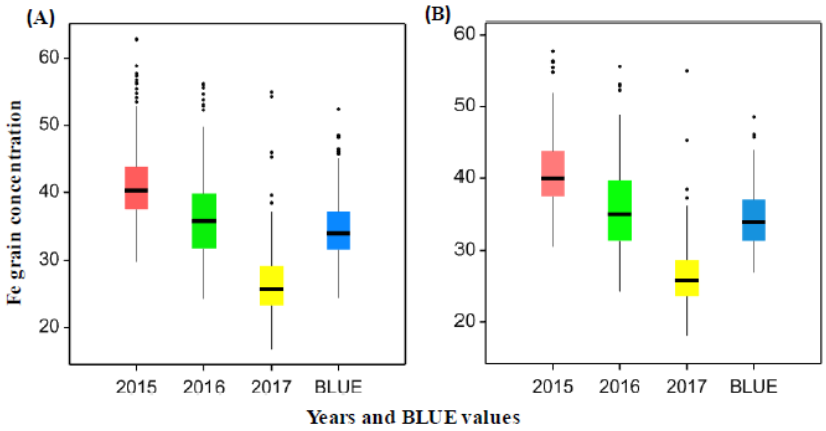


Figure 1. (**A**) Grain iron concentration (μg·g^{-1}) in all wheat genotypes of the whole panel for the years 2015, 2016 and 2017 and BLUE values. (**B**) Grain iron concentration (μg·g^{-1}) for wheat genotypes in the subpanel for the years 2015, 2016 and 2017 and BLUE values.

A significant positive correlation ranging from r = 0.26 to 0.39 ($p < 0.001$) was found for grain Fe in all three years (Figure 2). As well, a significant positive correlation (0.11–0.26, $p < 0.001$) was present between Fe and thousand kernel weight (TKW) in all three years (Figure S4) and a strong correlation was found between Fe and Zn with values ranging between 0.51–0.68 ($p < 0.001$) over years (Figure S4). In the whole wheat panel, genotype "SW Tataros" showed the highest Fe concentration equaling 52.67 μg·g^{-1} DW based on the BLUEs (Figure 3).

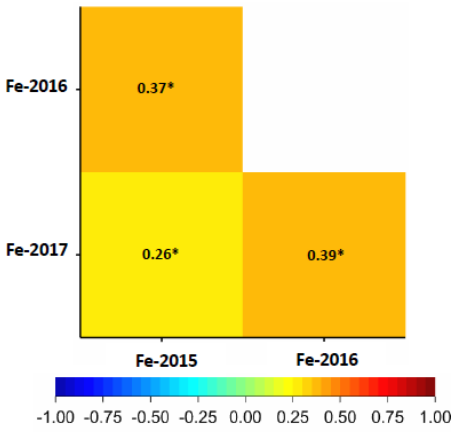


Figure 2. Pearson correlation between Fe grain concentrations (μg·g^{-1}) in the years 2015, 2016 and 2017. The degree of significance indicated as * $p \leq 0.05$.

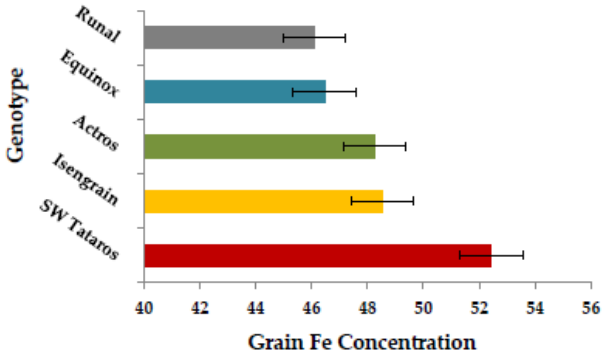


Figure 3. The scale of the top five genotypes with the highest Fe concentration (μg·g^{-1}) value crossing years (BLUE).

2.2. Genetic Analysis and Genes Underlying GFeC Trait

GWAS analysis for the whole panel identified 41 significant MTAs ($-\log_{10}$ (p-value) $\geq$ 3) (Figure 4A) which were distributed over the genome with R^2 values ranging between 2.7% to 5.22%. In total 41 MTAs, of which 17 were located on chromosome 3B between 46.6 to 59.8 cM (Table S3). Due to no common associations among years, our analyses were based on BLUEs by including the most significant 3 SNPs for further analysis (Table S3). In the subpanel, the number of significant associations including unmapped markers was higher and mounted to 137 MTAs (Figure 5A) with R^2 values ranging from 5.60% to 13.09% (Table S4). The highest phenotypic variation was related to unmapped markers (AX-158577508 and AX-158577509) and equaled 10.38% and 13.09%, respectively. Fifteen, four and two significant SNPs which were present on chromosomes 2A (763,689,738–765,710,113 bp), 3B (731,263,238–731,264,585 bp) and 5A (538,758,878–539,958,539 bp) were targeted for further analysis.

The QQ plots for SNP results revealed that the distribution of observed association p-values were close to the distribution of expected associations (Figures 4B and 5B); that means the model which we implemented for GWAS was sufficiently stringent to control for false positive associations. In a previous study a total of 8 markers in the whole panel and 31 markers in the subpanel (Tables S3 and S4) had been found significant for grain Zn concentration in the same germplasm [9]. Based on BLUEs, significant markers from the whole and subpanel were selected for a query against IWGSC RefSeq annotation v1.0 to get their annotations.

In the subpanel, we detected several potential candidate genes that located on chromosome 2A (763,689,738–765,710,113 bp) (Table S5). Based on the functional annotation, we found genes which encode either a transcription factor (TF) related to the NAC (NAM (no apical meristem)) domain family or a transmembrane protein (Table 1). These genes are well known to play a role in nutrient remobilization in plants [21,22]. Therefore, we conclude that this genomic region harbors several putative candidate genes, which may have a significant role in grain Fe accumulation.

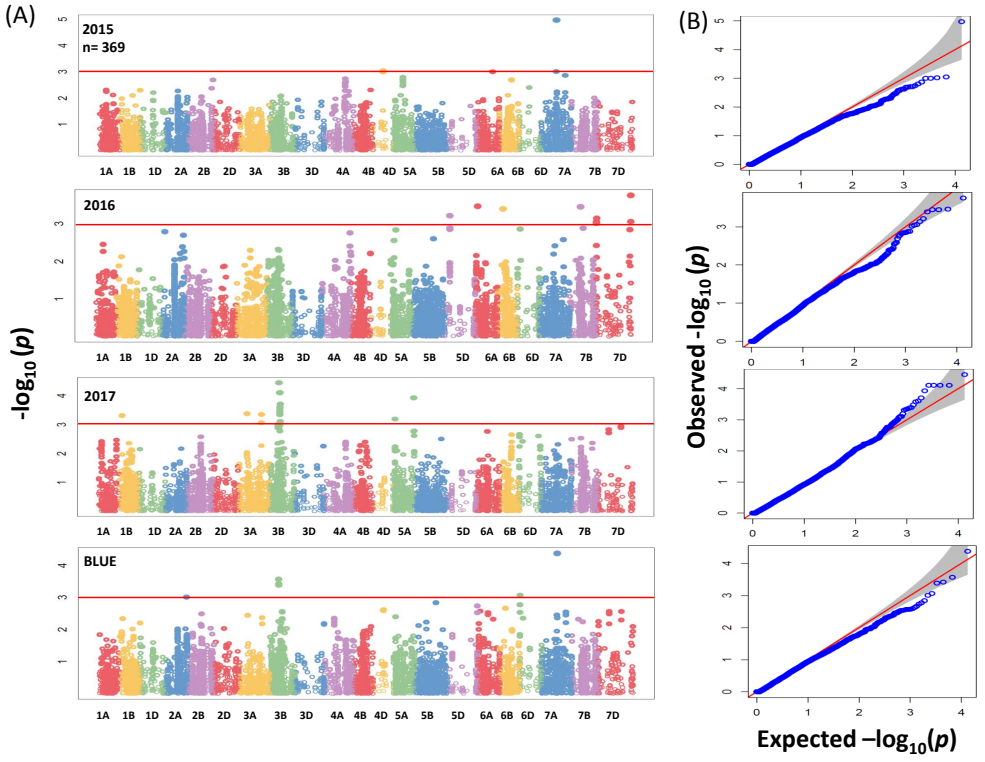


Figure 4. (**A**) Summary of genome-wide association scans for the whole panel of wheat genotypes (369) which were analyzed by using the 90K iSELECT Infinium array and the 35K Affymetrix SNP array for each year (2015/2016/2017) and BLUEs. The horizontal red color line indicated the threshold of $-\log_{10}$ (p-value) of 3. (**B**) Quantile-quantile scale representing expected versus observed $-\log_{10}$ (p-value).

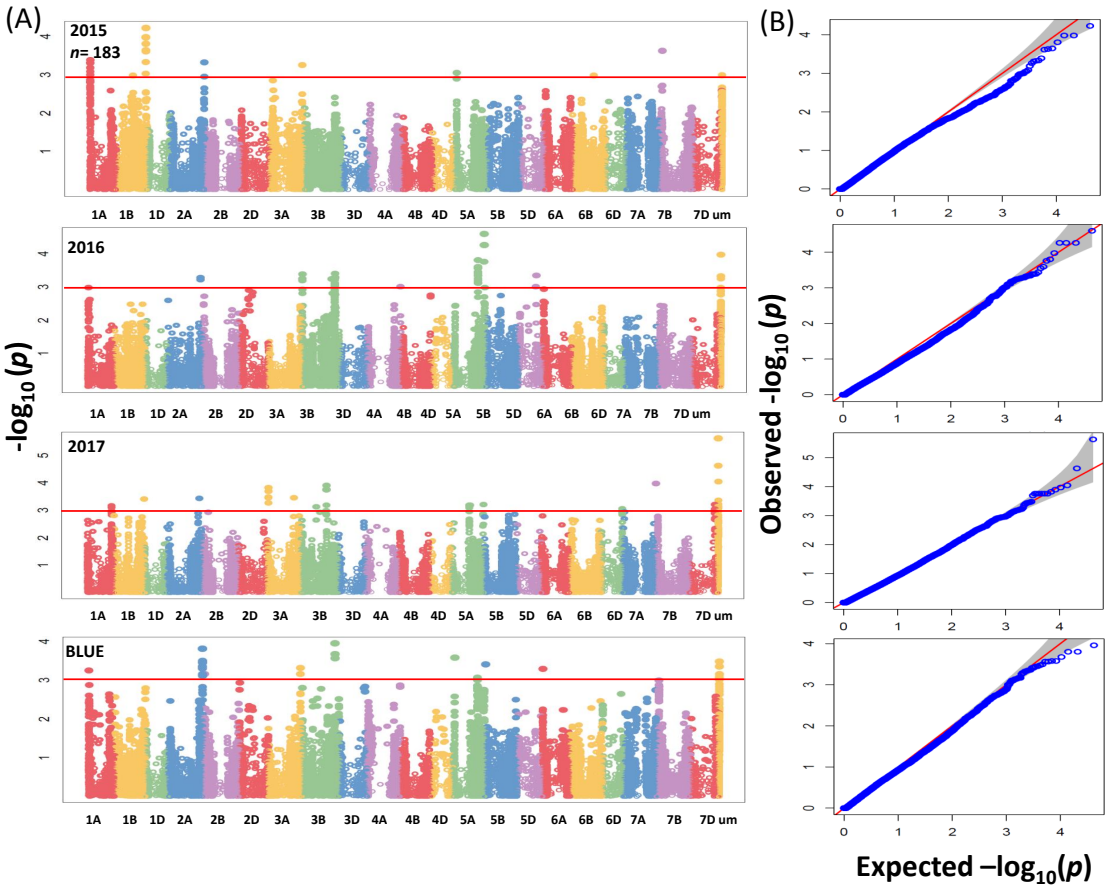


Figure 5. (**A**) Summary of genome-wide association scans for the subpanel of wheat genotypes (183) which analyzed by using the 90K iSELECT Infinium array, the 35K Affymetrix SNP array and the 135K Affymetrix SNP array for each year (2015/2016/2017) and BLUEs. The horizontal red color line indicated the threshold of $-\log_{10}$ (*p*-value) of 3. (**B**) Quantile-quantile scale representing expected versus observed $-\log_{10}$ (*p*-value).

Table 1. Potential candidate genes underlying GFeC trait in wheat.

Gene ID	Gene Annotation	Chr.	Start (bp)	End (bp)
TraesCS2A01G562600, TraesCS2A01G562700	transmembrane protein, (DUF247)	2A	763,796,420 763,802,755	763,799,183 763,804,683
TraesCS2A01G563600, TraesCS2A01G565000	transmembrane protein, (DUF594)	2A	764,149,111 764,898,033	764,150,898 764,900,078
TraesCS2A01G565900, TraesCS2A01G566000, TraesCS2A01G566100, TraesCS2A01G566200, TraesCS2A01G566300, TraesCS2A01G566400	NAC domain-containing protein	2A	765,277,860 765,373,519 765,392,440 765,441,104 765,514,989 765,546,770	765,278,647 765,375,363 765,393,650 765,442,258 765,518,243 765,547,909

2.3. Genomic Prediction of GFeC Trait

GP was evaluated for GFeC trait with three statistical models including GBLUP, ridge regression best linear unbiased prediction (RR-BLUP) and Bayes-Cπ in the whole panel. Prediction ability values were 0.29 to 0.38, 0.27 to 0.35, and 0.20 to 0.35 based on using these methods: GBLUP, RR-BLUP and Bayes-Cπ respectively (Figure 6). The highest value is equal 0.38 (GBLUP) and 0.35 (RR-BLUP and Bayes-Cπ) based on Fe BLUE values. The prediction values within years were almost the same and

equaled around 0.2 (Figure 6). Based on the GP results, more accurate estimates of breeding values through marker-based relationship matrices could be obtained by increasing the number of genotypes in the training data.

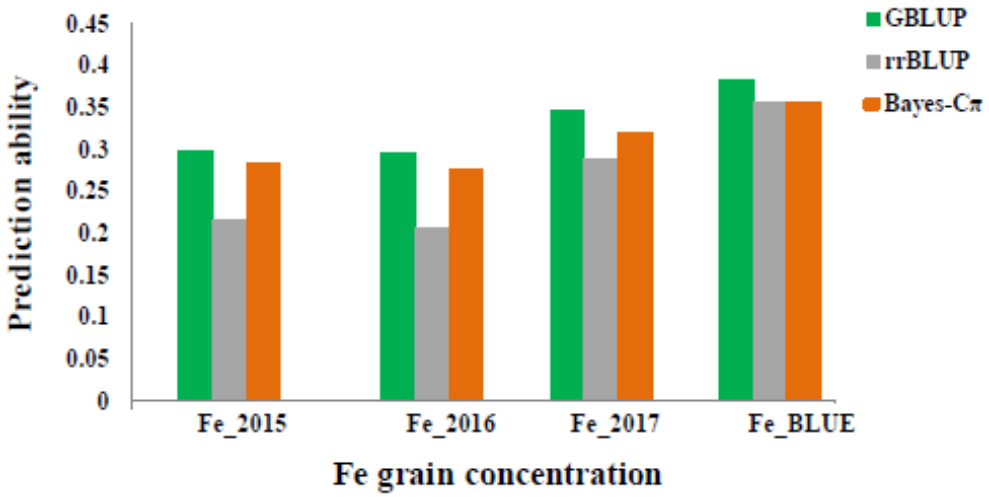


Figure 6. Prediction ability values for grain Fe concentration (μg·g^{-1}) according to different years (2015, 2016 and 2017) and BLUEs by using three different statistical models: GBLUP, rrBLUP and BayesC.

3. Discussion

3.1. The Usefulness of the Natural Phenotypic Variation

In human nutrition, the estimated average requirement (EAR) of Fe is 1460 μg/day/person, while the target level for sufficient Fe concentration in wheat grains was established as 52 μg·g^{-1} [23]. In our mapping panel, we observed high phenotypic variation in grain Fe concentrations ranging between 16.77–62.87 μg·g^{-1} among years and identifying 23 lines equal or above the required target (≥52 μg·g^{-1}). A similar range of GFeCs was reported for a wheat RIL population (17.8–69.0 μg·g^{-1}) which resulted from crossing *Triticum boeoticum* with *Triticum monococcum* [24]. Morgounov et al. [25] found GFeCs in the range of 34–43 μg·g^{-1} for 41 winter w·μg heat cultivars except for one spring wheat cultivar, that had GFeC of 56 μg·g^{-1}. Also, the range of Fe concentrations in Indian and Pakistan hexaploid wheat grains was found to be in the range of (9.2 to 49.7 μg·g^{-1}) [26]. Therefore, using lines with elevated GFeC are important to develop new varieties for crop improvement.

The heritability of Fe concentration among the years was moderate equaling $H^2 = 0.59$, suggesting a quantitative nature of inheritance and a considerable environmental influence on the expression of the trait. Gorafi et al. [12] reported a broad-sense heritability value of Fe grain concentration in synthetic hexaploid wheat germplasm of 0.80. Khokhar et al. [27] reported broad-sense heritability equal to 0.75 for Fe grain concentrations in field-grown Indian wheat.

GFeCs showed a significant positive correlation among years ($r = 0.25$–0.38, $p < 0.001$), indicating a relatively stable measurement of the phenotypic data. The resulting correlation values were moderate; that may be attributed to the influence of genetics and environment on the GFe accumulation which can also explain the moderate heritability value (0.59) of GFe. Tiwari et al. [24] found a constant correlation between different locations for grain Fe concentrations which is compatible with our results.

The positive and highly significant correlation between Fe and Zn in addition to a significant positive correlation between Fe, Zn and TKW found in the current study, was also reported in earlier studies in wheat [10,25]. For instance, our results agree with Pandey et al. [26], who reported a positive correlation between GFeCs and Zn concentrations in 150 bread wheat lines. Additionally, Fe and Zn have the same families of transporter proteins in several steps during the transportation from the soil to the grain, for example nicotianamine (NA) related enzymes are important for both of Fe and Zn radial movement through the root [28,29]. As well, such a high correlation between Fe, Zn and

TKW suggest that these traits (Fe and Zn) may have the same genetic basis and could be improved simultaneously with TKW or TKW determinants such as starch and protein. Krishnappa et al. [30] found common genetic regions between Fe, Zn, TKW and protein content. Peleg et al. [10] showed a positive correlation between Fe, Zn, TKW and protein content in wheat. Therefore, it is important to shed light on the genetic makeup of these traits together.

3.2. Putative Candidate Genes

Based on GWAS analysis, we found that different loci are controlling Fe accumulation in wheat grains indicating that it is a complex trait with polygenic control. In the whole panel, 137 significant associations were underlying grain Fe and were distributed on chromosomes 1A, 1B, 2A, 2B, 3A, 3B, 4A, 5A, 5B, 5D, 6A, 6D, 7B and 7D of which 3 significant SNPs were located on 3B (46.60–47.42 cM). There were no obvious candidate genes detected within the aforementioned region. In the subpanel, three physical regions contained significant SNPs, on chromosomes 2A (763,689,738–765,710,113 bp), 3B (731,263,238–731,264,585 bp) and 5A (538,758,878–539,958,539 bp), but only the 2A region conferred candidate genes involved in iron uptake or homeostasis.

Chromosome 2A conferred six putative genes related to the NAC (NAM (no apical meristem)) domain family proteins (Table 1), which are well known to be involved in accelerated senescence and an increase of nutrient remobilization from leaves to grains. Several studies reported about NAC gene and increasing Fe and Zn content in the grains of wheat [13,21,31,32]. Uauy et al. [21] described that a NAC TF (*NAM-B1*) accelerated senescence and nutrient remobilization from leaves to grains. The reduction in RNA levels of the multiple *NAM* homologs by RNA interference delayed the senescence process and reduced wheat grain protein, Zn, and Fe content by more than 30%. In the same context, Ricachenevsky et al. [31] showed that *NAM-B1* which is one of the NAC TFs has a major role in regulating key genes responsible for the senescence process which leads to higher Fe and Zn concentrations in wheat grains.

Another four genes encoded transmembrane proteins on chromosome 2A. It has been reported that transmembrane proteins are responsible for nutrient uptake in plants and play an important role in enhancing the micronutrient content of grains [22,33]. Therefore, these genes could be important for GFeC in wheat; however, functional characterization studies are required to validate the function of these genes.

3.3. Genome-Wide Prediction Accuracy

GP or GS has been proposed as a method to improve the breeding efficiency of quantitative and complex traits. Therefore, we extended our analyses and included GP as a suggested tool for improving a polygenic trait such as GFeC in wheat. Our predictability results showed low to moderate values according to three different years and BLUE values, which agrees with another report that obtained low to moderate predictability values for the macro- and micro-nutrients including Fe in wheat landraces [19]. In spring wheat, GP showed moderate to high prediction accuracy for grain Fe by imputing different statistical models [20]. Based on our findings, GP may be considered as a promising approach for enhancing GFeC in wheat especially when larger size germplasm panels with additional genotypes are used to have more accurate estimates of breeding values.

4. Materials and Methods

4.1. Plant Germplasm

A population comprised of 369 elite European wheat varieties including 355 genotypes of winter wheat and 14 genotypes of spring wheat, mainly from Germany and France was used in this study. Field experiments were carried out at IPK, Gatersleben, Germany over three consecutive years (2014/2015, 2015/2016 and 2016/2017) using plot with a size of 2 × 2 m for each genotype with six rows spaced 0.20 m apart and more details were described in a previous study by Alomari et al. [9].

The grains were collected randomly from more than 250 plants of each plot to be used in the study. Standard agronomic wheat management practices were subjected without applying fertilizers to the soil.

4.2. Milling Process

Three hundred sixty-nine wheat genotypes harvested from three different field experimental trials were prepared for milling process by collecting 50 kernels for each genotype to measure thousand-grain weights (TGW) using a digital seed analyzer/counter Marvin (GTA Sensorik GmbH, Neubrandenburg, Germany). Wheat grains were milled by using a Retsch mill (MM300, Mettmann, Germany), afterward, the whole panel of the milled wheat grains was dried by incubating overnight at 40 °C in the oven.

4.3. Iron Concentration Measurements

Fifty mg of dried and milled wheat grain flour was taken to be digested by (2 mL) nitric acid (HNO_3 69%, Bernd Kraft GmbH, Germany). The digestion process was completed using a high-performance microwave reactor (UltraClave IV, MLS, Leutkirch im Allgäu, Baden-Württemberg, Germany). All digested samples were filled up to 15 mL final volume with de-ionized distilled (Milli-Q) water (Milli-Q® Reference System, Merck, Germany). Element standards were prepared from Bernd Kraft multi-element standard solution (Germany). Fe as an external standard and Yttrium (Y) (ICP Standard Certipur® Merck, Germany) were used as internal standards for matrix correction. Fe concentrations were measured by Inductively Coupled Plasma Optical Emission Spectrometry (ICP-OES, iCAP 6000, Thermo Fisher Scientific, Dreieich, Germany) combined with a CETAC ASXPRESS™ PLUS rapid sample introduction system and a CETAC autosampler (CETAC Technologies, Omaha, NE, USA).

4.4. Statistical Analysis

We used Sigma Plot package 13 to perform the ANOVA and Pearson's correlation coefficient (r) which were calculated for the grain Fe data resulted from the three years. The broad-sense heritability was calculated using the equation:

$$H^2 = {\sigma^2}_G / ({\sigma^2}_G + ({\sigma^2}_e / nE),$$

where ${\sigma^2}_G$ is the genotype variance, ${\sigma^2}_e$ represents the variance of the residual and nE is the environments number.

Mixed linear model function and the residual maximum likelihood (REML) algorithm were applied to calculate the Best linear unbiased estimates (BLUEs) of Fe concentration in wheat grains for each genotype across the years [34] by considering the genotype as a fixed effect and the environment as a random effect. All these calculations were accomplished using GenStat v18 software (VSN International, Hemel Hempstead, UK).

4.5. Genotyping

The whole wheat germplasm (369 varieties) was genotyped using two marker arrays: a 90K iSELECT Infinium array [35] and a 35K Affymetrix SNP array (Axiom® Wheat Breeder's Genotyping Array, http://www.cerealsdb.uk.net/) [36] and these two arrays were genotyped by TraitGenetics GmbH, Gatersleben, Germany (www.traitgenetics.com). Moreover, a novel 135K Affymetrix array was used to genotype a subpanel of 183 genotypes from the whole genotypes panel [9,37] and this chip was designed by TraitGenetics GmbH. As a reference map, the ITMI-DH population [38,39] was used to anchor the SNP markers of the 90K and 35K chips. The 135K chip markers were genetically mapped on four different F_2-populations and then physically anchored on the chromosome-based sequence of hexaploid wheat [40].

4.6. GWAS Analysis

To identify the MTA and QTL (i.e., genomic regions) for Fe concentration in wheat grains, association analyses were conducted between SNP markers and Fe data for each genotype in each year and BLUEs value. For SNP markers, quality control was applied by considering a minor allele frequency (MAF) $\leq$3% (equaling 11 varieties out of 369) with rejecting SNPs having missing values or heterozygosity $\geq$3%, resulting in 15,523 polymorphic SNP markers from both of the 90K iSELECT array and the 35K Affymetrix array and 28,710 polymorphic SNP markers from the 135K Affymetrix array, which were used for association analysis.

GWAS was carried out for Fe concentration data from both panels (whole and subpanel) over individual year plus BLUE values by applying the implemented mixed linear model (MLM) and principal component analysis (PCA) as a correction factor for population structure. Whole wheat genotypes panel was analyzed by using the combination of two SNP chips (90K and the 35K chips) based on their genetic reference map whereas, the subpanel was analyzed by the combination of 90K, 35K and 135K chips which were anchored based on physical locations. The purpose of combining the SNP chips was to increase the density of the used markers, achieving good mapping resolution and to further enhance the power of GWAS output within the germplasm panel. All the detected marker-trait associations (MTAs) above the threshold of $-\log_{10}$ (p-value) ≥ 3 where considered as a significant MTA.

GWAS analysis was computed based on a MLM and PCA which was used for population correction and stratification by using Genome Association and Prediction Integrated Tool (GAPIT) in R [41]. The appropriateness of the used model was evaluated through Q–Q plots that were obtained by plotting "expected$-\log_{10}$ (p-values)" on the x-axis and "observed$-\log_{10}$ (p-values)" on the y-axis. The population structure of the investigated genotypes panel was described in a previous study by Kollers et al. [42].

4.7. Blasting and Annotation

The significant SNP markers which defined the significant associations underlying GFeC trait were listed to obtain their annotation based on the newly released reference genome sequence of Chinese Spring by blasting their sequance against IWGSC RefSeq annotation v1.0 to detect potential candidate genes [43,44].

4.8. Genomic Prediction

4.8.1. GBLUP

We used GBLUP to impute GP for GFeC trait data by using Tassel version 5.2.10 [45]. In this model, we evaluated the prediction accuracy by using fivefold cross-validation with 20 iterations as implemented in Tassel software.

4.8.2. RR-BLUP and Bayes-Cπ

We evaluate the prediction ability with the two GS models that are ridge regression best linear unbiased prediction (RR-BLUP) and Bayes-Cπ [16,46]. For both models, GSs were implemented in R using a fivefold cross-validation as described in previous literature Jiang et al. [47]. Simply, all the individuals were randomly divided into five subsets, in which four of the five were used as estimation set and the remaining one were used as test set. After all the genotypic values of individuals were obtained, we calculate the prediction ability that is the correlation between observed and predicted values. The whole process was repeated 100 times and then the mean value was used as the final prediction ability.

5. Conclusions

This study characterized many lines of a diverse wheat panel for GFeCs to understand the natural diversity that exists for Fe grain trait and to identify potential genes that contribute to this phenotypic variation and to examine the prediction accuracy. Broad-sense heritability calculation revealed moderate variation that could be attributed to both genetic and environmental effects. Overall, the resources generated in this study can be used to identify suitable candidate genes for further validation analysis. Results of applying GP models to GFeC showed that the correlation between observed and predicted values was relatively moderate; therefore, it would be useful to study the effects of GxE interactions that may improve the predictability value.

Supplementary Materials: Supplementary Materials can be found at http://www.mdpi.com/1422-0067/20/1/76/s1.

Author Contributions: D.Z.A. performed the data analysis including genome-wide association scan, candidate genes identification, genomic prediction (GBLUP) and other statistical analysis. K.E. and N.v.W. participated in Fe concentration measurements. F.L. performed the genomic prediction analysis (rrBLUB and BGLR). K.P. and M.S.R. designed the experiment. M.S.R. conceived the experiment and participated in the result interpretation. D.Z.A. and M.S.R. wrote the manuscript. All authors read and approved the final manuscript.

Funding: The project was funded by internal financial support of IPK Gatersleben.

Acknowledgments: We would like to acknowledge Ahmad M. Alqudah for critically reading the manuscript, as well as Murukarthick Jayakodi for sharing the physical position of markers. The authors are thankful for Ellen Weiß and Yudelsy Antonia Tandron Moya for excellent technical assistance. The genotyping data were created in the frame of the projects VALID and SELECT (project numbers 0,314,947 and 0315949) funded by the Plant Biotechnology Program of the German Federal Ministry of Education and Research (BMBF).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Lafiandra, D.; Riccardi, G.; Shewry, P.R. Improving cereal grain carbohydrates for diet and health. *J. Cereal Sci.* **2014**, *59*, 312–326. [CrossRef] [PubMed]
2. Al Omari, D.Z.; Abdul-Hussain, S.S.; Ajo, R.Y. Germinated lupin (*Lupinus albus*) flour improves Arabic flat bread properties. *Qual. Assur. Saf. Crops Foods* **2016**, *8*, 57–63. [CrossRef]
3. Alomari, D.Z.; Abdul-Hussain, S.S. Effect of Lupin Flour Supplementation on Chemical, Physical and Sensory Properties of Mediterranean Flat Bread. *Int. J. Food Sci. Nutr. Eng.* **2013**, *3*, 49–54. [CrossRef]
4. Graham, R.D. Micronutrient Deficiencies in Crops and Their Global Significance. In *Micronutrient Deficiencies in Global Crop Production*; Alloway, B.J., Ed.; Springer: Dordrecht, The Netherlands, 2008; pp. 41–61.
5. Bouis, H.E. Micronutrient fortification of plants through plant breeding: Can it improve nutrition in man at low cost? *Proc. Nutr. Soc.* **2007**, *62*, 403–411. [CrossRef]
6. Bouis, H.E. Plant breeding: A new tool for fighting micronutrient malnutrition. *J. Nutr.* **2002**, *132*, 491S–494S. [CrossRef] [PubMed]
7. Graham, R.; Senadhira, D.; Beebe, S.; Iglesias, C.; Monasterio, I. Breeding for micronutrient density in edible portions of staple food crops: Conventional approaches. *Field Crops Res.* **1999**, *60*, 57–80. [CrossRef]
8. Alomari, D.Z.; Eggert, K.; von Wiren, N.; Pillen, K.; Roder, M.S. Genome-Wide Association Study of Calcium Accumulation in Grains of European Wheat Cultivars. *Front. Plant Sci.* **2017**, *8*, 1797. [CrossRef]
9. Alomari, D.Z.; Eggert, K.; von Wirén, N.; Alqudah, A.M.; Polley, A.; Plieske, J.; Ganal, M.W.; Pillen, K.; Röder, M.S. Identifying Candidate Genes for Enhancing Grain Zn Concentration in Wheat. *Front. Plant Sci.* **2018**, *9*. [CrossRef]
10. Peleg, Z.; Cakmak, I.; Ozturk, L.; Yazici, A.; Jun, Y.; Budak, H.; Korol, A.B.; Fahima, T.; Saranga, Y. Quantitative trait loci conferring grain mineral nutrient concentrations in durum wheat x wild emmer wheat RIL population. *Theor. Appl. Genet.* **2009**, *119*, 353–369. [CrossRef] [PubMed]
11. Srinivasa, J.; Arun, B.; Mishra, V.K.; Singh, G.P.; Velu, G.; Babu, R.; Vasistha, N.K.; Joshi, A.K. Zinc and iron concentration QTL mapped in a Triticum spelta × T. aestivum cross. *Theor. Appl. Genet.* **2014**, *127*, 1643–1651. [CrossRef]

12. Gorafi, Y.S.A.; Ishii, T.; Kim, J.-S.; Elbashir, A.A.E.; Tsujimoto, H. Genetic variation and association mapping of grain iron and zinc contents in synthetic hexaploid wheat germplasm. *Plant Genet. Resour. Charact. Utili.* **2016**, *16*, 9–17. [CrossRef]
13. Bhatta, M.; Baenziger, P.S.; Waters, B.M.; Poudel, R.; Belamkar, V.; Poland, J.; Morgounov, A. Genome-Wide Association Study Reveals Novel Genomic Regions Associated with 10 Grain Minerals in Synthetic Hexaploid Wheat. *Int. J. Mol. Sci.* **2018**, *19*, 3237. [CrossRef] [PubMed]
14. Anuradha, N.; Satyavathi, C.T.; Bharadwaj, C.; Nepolean, T.; Sankar, S.M.; Singh, S.P.; Meena, M.C.; Singhal, T.; Srivastava, R.K. Deciphering Genomic Regions for High Grain Iron and Zinc Content Using Association Mapping in Pearl Millet. *Front. Plant Sci.* **2017**, *8*, 412. [CrossRef] [PubMed]
15. Upadhyaya, H.D.; Bajaj, D.; Das, S.; Kumar, V.; Gowda, C.L.; Sharma, S.; Tyagi, A.K.; Parida, S.K. Genetic dissection of seed-iron and zinc concentrations in chickpea. *Sci. Rep.* **2016**, *6*, 24050. [CrossRef] [PubMed]
16. Meuwissen, T.H.E.; Hayes, B.J.; Goddard, M.E. Prediction of Total Genetic Value Using Genome-Wide Dense Marker Maps. *Genetics* **2001**, *157*, 1819–1829. [PubMed]
17. Roos, A.; Boron, W.F. Intracellular pH transients in rat diaphragm muscle measured with DMO. *Am. J. Physiol.* **1978**, *235*, C49–C54. [CrossRef] [PubMed]
18. Jia, C.; Zhao, F.; Wang, X.; Han, J.; Zhao, H.; Liu, G.; Wang, Z. Genomic Prediction for 25 Agronomic and Quality Traits in Alfalfa (*Medicago sativa*). *Front. Plant Sci.* **2018**, *9*, 1220. [CrossRef]
19. Manickavelu, A.; Hattori, T.; Yamaoka, S.; Yoshimura, K.; Kondou, Y.; Onogi, A.; Matsui, M.; Iwata, H.; Ban, T. Genetic Nature of Elemental Contents in Wheat Grains and Its Genomic Prediction: Toward the Effective Use of Wheat Landraces from Afghanistan. *PLoS ONE* **2017**, *12*, e0169416. [CrossRef]
20. Velu, G.; Crossa, J.; Singh, R.P.; Hao, Y.; Dreisigacker, S.; Perez-Rodriguez, P.; Joshi, A.K.; Chatrath, R.; Gupta, V.; Balasubramaniam, A.; et al. Genomic prediction for grain zinc and iron concentrations in spring wheat. *Theor. Appl. Genet.* **2016**, *129*, 1595–1605. [CrossRef]
21. Uauy, C.; Distelfeld, A.; Fahima, T.; Blechl, A.; Dubcovsky, J. A NAC Gene regulating senescence improves grain protein, zinc, and iron content in wheat. *Science* **2006**, *314*, 1298–1301. [CrossRef]
22. Chrispeels, M.J.; Crawford, N.M.; Schroeder, J.I. Proteins for transport of water and mineral nutrients across the membranes of plant cells. *Plant Cell* **1999**, *11*, 661–676. [CrossRef]
23. Bouis, H.E.; Welch, R.M. Biofortification—A Sustainable Agricultural Strategy for Reducing Micronutrient Malnutrition in the Global South. *Crop Sci.* **2010**, *50*, S-20. [CrossRef]
24. Tiwari, V.K.; Rawat, N.; Chhuneja, P.; Neelam, K.; Aggarwal, R.; Randhawa, G.S.; Dhaliwal, H.S.; Keller, B.; Singh, K. Mapping of Quantitative Trait Loci for Grain Iron and Zinc Concentration in Diploid A Genome Wheat. *J. Hered.* **2009**, *100*, 771–776. [CrossRef] [PubMed]
25. Morgounov, A.; Gómez-Becerra, H.F.; Abugalieva, A.; Dzhunusova, M.; Yessimbekova, M.; Muminjanov, H.; Zelenskiy, Y.; Ozturk, L.; Cakmak, I. Iron and zinc grain density in common wheat grown in Central Asia. *Euphytica* **2007**, *155*, 193–203. [CrossRef]
26. Pandey, A.; Khan, M.K.; Hakki, E.E.; Thomas, G.; Hamurcu, M.; Gezgin, S.; Gizlenci, O.; Akkaya, M.S. Assessment of genetic variability for grain nutrients from diverse regions: Potential for wheat improvement. *Springerplus* **2016**, *5*, 1912. [CrossRef]
27. Khokhar, J.S.; Sareen, S.; Tyagi, B.S.; Singh, G.; Wilson, L.; King, I.P.; Young, S.D.; Broadley, M.R. Variation in grain Zn concentration, and the grain ionome, in field-grown Indian wheat. *PLoS ONE* **2018**, *13*, e0192026. [CrossRef]
28. Rellán-Álvarez, R.; Giner-Martínez-Sierra, J.; Orduna, J.; Orera, I.; Rodríguez-Castrillón, J.Á.; García-Alonso, J.I.; Abadía, J.; Álvarez-Fernández, A. Identification of a Tri-Iron(III), Tri-Citrate Complex in the Xylem Sap of Iron-Deficient Tomato Resupplied with Iron: New Insights into Plant Iron Long-Distance Transport. *Plant Cell Physiol.* **2010**, *51*, 91–102. [CrossRef]
29. Deinlein, U.; Weber, M.; Schmidt, H.; Rensch, S.; Trampczynska, A.; Hansen, T.H.; Husted, S.; Schjoerring, J.K.; Talke, I.N.; Krämer, U.; et al. Elevated Nicotianamine Levels in Arabidopsis Halleri Roots Play a Key Role in Zinc Hyperaccumulation. *Plant Cell* **2012**, *24*, 708. [CrossRef]
30. Krishnappa, G.; Singh, A.M.; Chaudhary, S.; Ahlawat, A.K.; Singh, S.K.; Shukla, R.B.; Jaiswal, J.P.; Singh, G.P.; Solanki, I.S. Molecular mapping of the grain iron and zinc concentration, protein content and thousand kernel weight in wheat (*Triticum aestivum* L.). *PLoS ONE* **2017**, *12*, e0174972. [CrossRef]

31. Ricachenevsky, F.K.; Menguer, P.K.; Sperotto, R.A. kNACking on heaven's door: How important are NAC transcription factors for leaf senescence and Fe/Zn remobilization to seeds? *Front. Plant Sci.* **2013**, *4*, 226. [CrossRef]
32. Nadolska-Orczyk, A.; Rajchel, I.K.; Orczyk, W.; Gasparis, S. Major genes determining yield-related traits in wheat and barley. *Theor. Appl. Genet.* **2017**, *130*, 1081–1098. [CrossRef] [PubMed]
33. Schroeder, J.I.; Delhaize, E.; Frommer, W.B.; Guerinot, M.L.; Harrison, M.J.; Herrera-Estrella, L.; Horie, T.; Kochian, L.V.; Munns, R.; Nishizawa, N.K.; et al. Using membrane transporters to improve crops for sustainable food production. *Nature* **2013**, *497*, 60–66. [CrossRef] [PubMed]
34. Yu, J.; Pressoir, G.; Briggs, W.H.; Vroh Bi, I.; Yamasaki, M.; Doebley, J.F.; McMullen, M.D.; Gaut, B.S.; Nielsen, D.M.; Holland, J.B.; et al. A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. *Nat. Genet.* **2006**, *38*, 203–208. [CrossRef] [PubMed]
35. Wang, S.; Wong, D.; Forrest, K.; Allen, A.; Chao, S.; Huang, B.E.; Maccaferri, M.; Salvi, S.; Milner, S.G.; Cattivelli, L.; et al. Characterization of polyploid wheat genomic diversity using a high-density 90,000 single nucleotide polymorphism array. *Plant Biotechnol. J.* **2014**, *12*, 787–796. [CrossRef] [PubMed]
36. Allen, A.M.; Winfield, M.O.; Burridge, A.J.; Downie, R.C.; Benbow, H.R.; Barker, G.L.; Wilkinson, P.A.; Coghill, J.; Waterfall, C.; Davassi, A.; et al. Characterization of a Wheat Breeders' Array suitable for high-throughput SNP genotyping of global accessions of hexaploid bread wheat (*Triticum aestivum*). *Plant Biotechnol. J.* **2017**, *15*, 390–401. [CrossRef]
37. Zanke, C.D.; Rodemann, B.; Ling, J.; Muqaddasi, Q.H.; Plieske, J.; Polley, A.; Kollers, S.; Ebmeyer, E.; Korzun, V.; Argillier, O.; et al. Genome-wide association mapping of resistance to eyespot disease (Pseudocercosporella herpotrichoides) in European winter wheat (*Triticum aestivum* L.) and fine-mapping of Pch1. *TAG. Theor. Appl. Genet.* **2017**, *130*, 505–514. [CrossRef] [PubMed]
38. Poland, J.A.; Brown, P.J.; Sorrells, M.E.; Jannink, J.L. Development of high-density genetic maps for barley and wheat using a novel two-enzyme genotyping-by-sequencing approach. *PLoS ONE* **2012**, *7*, e32253. [CrossRef]
39. Sorrells, M.E.; Gustafson, J.P.; Somers, D.; Chao, S.; Benscher, D.; Guedira-Brown, G.; Huttner, E.; Kilian, A.; McGuire, P.E.; Ross, K.; et al. Reconstruction of the synthetic W7984 x Opata M85 wheat reference population. *Genome* **2011**, *54*, 875–882. [CrossRef]
40. Appels, R.; Eversole, K.; Feuillet, C.; Keller, B.; Rogers, J.; Stein, N.; Pozniak, C.J.; Stein, N.; Choulet, F.; Distelfeld, A.; et al. Shifting the limits in wheat research and breeding using a fully annotated reference genome. *Science* **2018**, *361*. [CrossRef]
41. Lipka, A.E.; Tian, F.; Wang, Q.; Peiffer, J.; Li, M.; Bradbury, P.J.; Gore, M.A.; Buckler, E.S.; Zhang, Z. GAPIT: Genome association and prediction integrated tool. *Bioinformatics* **2012**, *28*, 2397–2399. [CrossRef]
42. Kollers, S.; Rodemann, B.; Ling, J.; Korzun, V.; Ebmeyer, E.; Argillier, O.; Hinze, M.; Plieske, J.; Kulosa, D.; Ganal, M.W.; et al. Whole genome association mapping of Fusarium head blight resistance in European winter wheat (*Triticum aestivum* L.). *PLoS ONE* **2013**, *8*, e57500. [CrossRef] [PubMed]
43. BLAST. Available online: https://wheat-urgi.versailles.inra.fr/Seq-Repository/BLAST (accessed on 22 November 2018).
44. IWGSC Reference Sequence v1.0 Browser Now Available at URGI. Available online: https://www.wheatgenome.org/News/Latest-news/IWGSC-Reference-Sequence-v1.0-browser-now-available-at-URGI (accessed on 22 November 2018).
45. Bradbury, P.J.; Zhang, Z.; Kroon, D.E.; Casstevens, T.M.; Ramdoss, Y.; Buckler, E.S. TASSEL: Software for association mapping of complex traits in diverse samples. *Bioinformatics* **2007**, *23*, 2633–2635. [CrossRef] [PubMed]

46. Habier, D.; Fernando, R.L.; Kizilkaya, K.; Garrick, D.J. Extension of the bayesian alphabet for genomic selection. *BMC Bioinform.* **2011**, *12*, 186. [CrossRef] [PubMed]
47. Jiang, Y.; Zhao, Y.; Rodemann, B.; Plieske, J.; Kollers, S.; Korzun, V.; Ebmeyer, E.; Argillier, O.; Hinze, M.; Ling, J.; et al. Potential and limits to unravel the genetic architecture and predict the variation of Fusarium head blight resistance in European winter wheat (*Triticum aestivum* L.). *Heredity* **2014**, *114*, 318. [CrossRef] [PubMed]

3. General Discussion

3.1 Strategies for dealing with complex traits like nutritional quality traits

3.1.1. General prerequisites for GWAS

Robustness and effectiveness of GWAS for dissecting complex traits in crops relies on several factors (Myles et al., 2009; Korte and Farlow, 2013). For instance, sufficient natural phenotypic variation with medium to high broad-sense heritability of a given trait within the population is imperative for performing GWAS and finding significant associations to identify the causative loci/gene(s) (Soto-Cerda and Cloutier, 2012). A GWAS study is more effective and productive when several environments such as years or locations are used to test a selected diverse population in order to identify stable QTL among the different environments. The investigated population is preferred to consist of a high number of individuals /accessions, i.e. hundreds of lines (Korte and Farlow, 2013). To reach the GWAS aims of detecting the responsible loci/gene(s) underlying the trait of interest, a high dense map (several thousands of SNPs) with high-resolution of physical or genetic positions and adequate allele frequencies are also required. Population structure must be considered during the GWAS to clarify the confounding effects among the used individuals based on their genetic relatedness (Myles et al., 2009). Therefore, applying the proper statistical methods that take into account all of the aforementioned factors especially population structure are crucial for having true associations (Myles et al., 2009; Korte and Farlow, 2013). Another analysis which is important to be considered during the GWAS is LD analysis (Myles et al., 2009). A combination of all of these conditions empowers researchers to reveal QTLs for natural variation that can be analyzed using state of the art bioinformatics tools to detect the underlying gene(s).

3.1.2 Plant material, population size and phenotypic analysis

In the three studies presented at chapter two, we exploited European elite wheat germplasm of 369 genotypes including 355 winter wheat and 14 spring wheat varieties. The wheat germplasm was grown in the field located within IPK campus for three consecutive years (2014/2015, 2015/2016 and 2016/2017). Planting was done within plots and each plot was containing around 250 plants to ensure an ideal plant density for each genotype and then wheat grains were

collected from each plot. Fifty wheat grains were taken randomly for each genotype and year and prepared for phenotypic analysis and further measurements. Using representative samples of three years of field experiments helped to ensure an accurate and precise estimation of the phenotypes, i.e. the mineral grain concentrations. The soil texture was clay loam and the soil analysis of the three field experiments demonstrated that the soil samples were homogeneous with the same characteristics, such as pH ≈ 7 across years. The field was planted during the fall under rainfed conditions without adding fertilizers in order to ensure standard agronomic wheat management practices and to avoid any effects on the realistic natural variation of minerals accumulation in wheat grains (Table 1). The same criteria of germplasm size and growing plants across different years were also used in a previous study, e.g. Velu et al. (2018) who used 330 bread wheat lines to characterize grain Zn concentrations across multiple environments. In contrast, most of the GWAS studies in wheat used a population panel up to 250 individuals which may lead to reduced phenotypic and genotypic variation (Wang et al., 2017; Liu et al., 2018; Maulana et al., 2018). Therefore, the power of GWAS depends on the population size (sufficient number of individuals) in order to identify the highly significant MTAs with a larger effect, acceptable frequency within the population and to detect rare variants. Recently, Bhatta et al. (2018) used almost the same conditions of growing the plants and collecting grains for GWAS study of grain minerals in a diverse panel containing 123 synthetic hexaploid wheat lines which were grown in clay loam soil with a pH around 7-8 and without adding fertilizers for two years of field experiments. In our study, we applied a realistic criteria to ensure high quality of analysis and output in the investigated European winter wheat elite varieties where we found a wide natural genetic diversity for the studied traits (Fe, Zn and Ca) during three years (Table 1) (Alomari et al., 2017; Alomari et al., 2018; Alomari et al., 2019). Previous GWAS studies that characterized minerals in wheat used several wheat populations without including released varieties. For instance, different recent GWAS studies characterized minerals in synthetic hexaploid wheat, spring wheat and in *Aegilops tauschii* (Bhatta et al., 2018; Velu et al., 2018; Arora et al., 2019). Another QTL analysis study investigated minerals in winter wheat without including elite wheat varieties (Shi et al., 2013). So far, our presented results are the first GWAS to characterize minerals in grains of European elite winter wheat varieties.

Table1. Phenotypic variation for the measured traits during three years and the calculated heritability

Trait		2014/2015	2015/2016	2016/2017	H^2
Fe	Range	29.83-62.87	24.30-56.20	16.77-54.97	0.59
	Mean	41.00	36.21	26.44	
Zn	Range	25.07-57.17	25.10-61.70	20.56-60.09	0.54
	Mean	37.10	36.92	30.61	
Ca	Range	334.6-797.2	234.3-797.2	208.5-597.8	0.73
	Mean	497.0	438.1	314.8	
TKW	Range	31.20-63.33	37.55-65.60	40.60-65.58	0.74
	Mean	48.51	52.96	52.59	

3.1.3 Analysis of phenotypic data and environmental effects

Broad-sense heritability (H^2) is usually used by plant breeders to measure the strength of the relationship between the phenotypic performance and the genetic effect of an individual genotype or in other words it is the genetic contribution to phenotypic variability. The measured heritability values may be ranged between 0 (no genetic contribution) and 1 (variation under genetic control) while values within the range of 0.30 to 0.60 are considered as moderate. Broad-sense heritability (H^2) estimates for Fe and Zn were moderate (0.59 and 0.54 respectively) and high (0.73) for Ca (Table 1) (Alomari et al., 2017; Alomari et al., 2018; Alomari et al., 2019), indicating high to moderate contribution of genotypic effect to the overall variation in the minerals' accumulation in wheat grains. Our findings corresponded with Crespo-Herrera et al. (2016) who found a moderate heritability for Fe and Zn concentration in wheat and with Khokhar et al. (2018) who found high heritability for grain Ca in wheat. In conclusion, wide natural genetic diversity with moderate to high heritability is a prerequisite to QTL detection and makes the breeding process feasible by selecting the superior genotypes.

A significant positive correlations were found between thousand kernel weight (TKW) with grain Fe concentration (0.11-0.25; $P<0.001$) and between TKW with grain Zn concentration (0.01-0.24; $P<0.001$), while TKW was not correlated or negatively correlated with grain Ca concentration (-0.28-0.04; $P<0.001$) (Figure 7). These results are consistent with some other reports; Shi et al. (2013) revealed a significant positive correlation between TKW and grain Fe whereas a negative correlation between TKW with grain Ca in wheat. Another recent report

observed a weak to almost no correlation between TKW with Fe and Zn in wheat grains (Arora et al., 2019).

The correlation analysis of the mineral concentrations among each other revealed that grain Ca, Fe and Zn were ranging between almost no correlations (0.06; $P<0.001$) to moderate significant and positive correlation (0.45; $P<0.001$). A significant moderate correlation between Ca and Zn was observed in two years (0.30, 0.45; $P<0.001$) in contrast to a correlation between Ca and Fe which was significantly low in two years (0.06, 0.08; $P<0.001$). Khokhar et al. (2018) reported a significant positive correlation between Ca with Zn in wheat grains equaling 0.36 with $P<0.001$ and Shi et al. (2013) reported about a slightly negative correlation between grain Ca and grain Fe concentrations equaling -0.08 in wheat. A strong significant and positive correlation (0.51 to 0.69; $P<0.001$) was detected between Fe and Zn suggesting the presence of common genetic factors affecting the accumulation of these minerals in grains and eventually the improvement of one of these minerals could positively affect the other one (Figure 7). Our findings support what is found repeatedly in previous studies (Zhao et al., 2009; Xu et al., 2012; Srinivasa et al., 2014; Khokhar et al., 2018) about the existence of a highly positive correlation between Fe and Zn. Therefore it might be possible to breed for both minerals (high Fe and Zn) in parallel.

A comparison between the top 15 accessions with highest concentrations of Ca, Fe and Zn (Table 2) observed that six accessions were common between Fe and Zn while only one accession was in common between Ca and Zn. These results confirmed what we found in the aforementioned correlation analysis between Ca and Zn and between Fe and Zn.

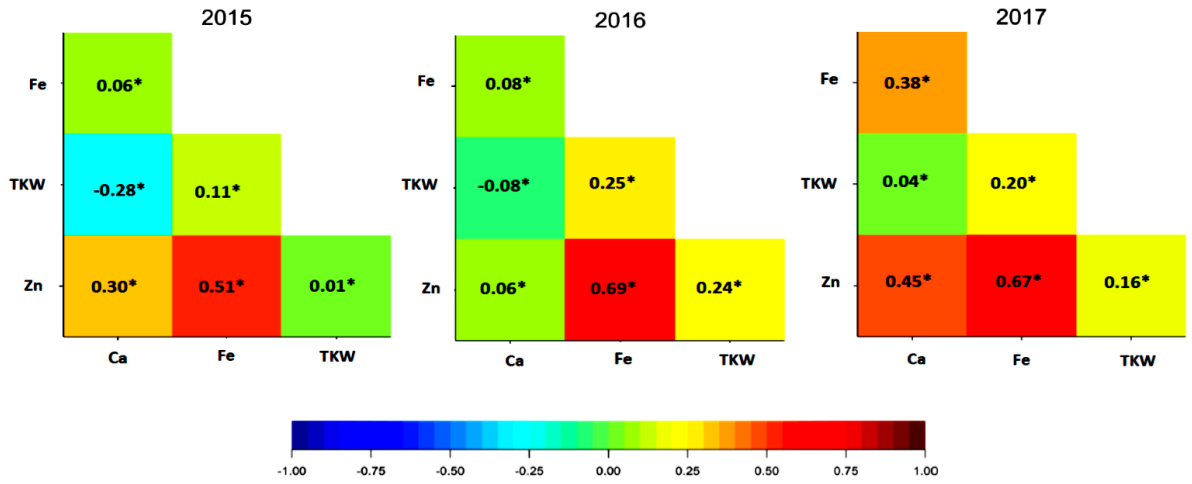


Figure 7. Person correlation for calcium (Ca), iron (Fe) and zinc (Zn) with thousand kernel weight (TKW) across three years, *P<0.001.

Table 2. List of wheat accessions with highest grain zinc (Zn), iron (Fe) and calcium (Ca) concentrations (µg g^{-1} dry weight (DW)) based on BLUEs.

Variety Name	Genotype	Zn	Variety name	Genotype	Fe	Variety name	Genotype	Ca
Fridolin	GW0251	56.41	SW Tataros	GW0334	56.98	Nirvana	GW0169	647.5
Ares	GW0337	51.98	Actros	GW0001	56.8	Inoui	GW0156	636
Rainer	GW0253	50.81	Tamaro	GW0197	53.37	Muza	GW0276	627.1
Tamaro	GW0197	50.52	Baguette 11	GW0376	49.74	Capnor	GW0018	625.2
Lars	GW0293	50.36	Runal	GW0189	49.73	Welford	GW0333	623.7
Vitus	GW0254	50.32	Heroldo	GW0040	49.33	Autan	GW0119	623.1
Xenos	GW0255	50.1	Ludwig	GW0050	48.93	Lona	GW0163	585.4
Zyta	GW0290	49.68	Alidos	GW0004	48.86	Pytagor	GW0183	584
SW Tataros	GW0334	48.6	Muza	GW0276	48.75	Savannah	GW0355	582.1
Jafet	GW0380	48.07	Zyta	GW0290	48.55	Ostka	GW0281	573.5
Heroldo	GW0040	47.88	Lucius	GW0049	48.39	H05581A	GW0209	572.4
Fregata	GW0268	47.64	Vitus	GW0254	47.86	Samurai	GW0190	565.6
Aperitiv	GW0291	47.13	Fregata	GW0268	47.82	Calisto	GW0129	563.1

3.1.4. The effectiveness of high-density SNP arrays

High-density single nucleotide polymorphism (SNP) genotyping arrays are a powerful tool for studying genomic diversity, understanding ancestral relationships among individuals in populations and detecting MTAs in mapping populations (Tian et al., 2011; Zhao et al., 2011).

The robustness and effectiveness of SNPs are based on several parameters such as the coverage percentage of the markers across the genome, minor allele frequency (MAF) and LD (Rasheed et al., 2017). In our study, we used one of the most common arrays in wheat that provides a dense coverage of the wheat genome which is known as 90k iSELECT Infinium array and provided an initial number of 46,977 SNP markers (Wang et al., 2014) and 7,761 SNPs after applying the quality control test with MAF ≥3%. Resulting in a density coverage equaling 2.7 SNP markers per cM in our dataset (Alomari et al., 2017; Alomari et al., 2018; Alomari et al., 2019). This array has been widely used for GWAS and QTL-mapping in wheat for studying various agronomic traits (Wen et al., 2017; Sukumaran et al., 2018; Velu et al., 2018). In the GWAS studies presented in chapter 2.1, 2.2 and 2.3, the map was improved by using a newly released array in wheat, the so-called 35k Axiom® Wheat Breeders' Genotyping Array which includes 35,143 SNP markers (Allen et al., 2017). After applying the quality control test including MAF ≥3%, the number reached 7,762 SNPs in our dataset and the marker density equaled 2.6 markers per cM (Alomari et al., 2017). Therefore, the combination of both arrays increased the marker density to reach 4.5 markers per cM that resulted in 15,523 polymorphic markers and led to increase the resolution of association mapping and to facilitate the detection of the candidate genes underlying Ca and Zn (Alomari et al., 2017; Alomari et al., 2018).

In chapters 2.2 and 2.3, a third novel array was added to the analysis, called 135k Affymetrix marker array and designed with an initial number of markers equaling 136,780 SNPs; the number of functional markers then reduced to 77,765 high-quality markers and after applying the MAF ≥3% we had in total 28,710 mapped SNPs which were involved in the genetic analyses. This array was applied to a subpanel of 183 genotypes from the whole genotype panel. A combination of all of these arrays increased the density up to 14 markers per cM which is to our knowledge the first time to present such a high-resolution map in a wheat association panel. Another feature of using these arrays is that the SNPs could be anchored to the wheat physical map providing the possibility to identify physical positions of candidate genes precisely.

The rationale behind incorporating several high-dense SNP arrays was to increase the marker numbers and density and to ensure a high resolution of the association mapping output which was confirmed later by the resulted findings. In conclusion, using high quality and density of SNPs along with a diverse panel of wheat genotypes enabled us to identify highly significant marker-trait associations and candidate genes (Alomari et al., 2018; Alomari et al., 2019).

Other important parameters to improve the mapping resolution in our studies (chapter 2.1, 2.2 and 2.3) are relying on the size and structure of the mapping population, a sharp decay in LD and the extent to which these markers are in LD to the causative genes (Mackay, 2001; Zondervan and Cardon, 2004). Population structure and genetic relatedness among the individuals and the lack of appropriate correction for population structure can lead to spurious associations. It is, therefore, important to know the structure of the population with respect to genetic relatedness among pairs of all the individuals used in the study in order to select the optimal GWAS model (Pritchard et al., 2000). In our study, there was no obvious population structure (Kollers et al., 2013) which can be regarded as a beneficial point to the analysis even though different correction methods were incorporated with mixed linear model (MLM) analysis. The kinship matrix model was used to deal with the Ca grain trait while correction with PCA was used to deal with Fe and Zn traits revealing associations with more stringent QQ-plots (Alomari et al., 2017; Alomari et al., 2018; Alomari et al., 2019).

LD is the nonrandom association of alleles at two or more loci in the population. It is measured as the squared allele frequency correlation ($r2$). The genetic or physical distance over which LD decays gives a suitable measure for the number and density of markers required determines the resolution in association mapping studies. If the LD exists between a marker or locus associated with a trait, then specific marker alleles of haplotypes can be associated with phenotypes at a high level of statistical significance (Cardon and Bell, 2001). Values of LD between the linked markers decreased with the increase in distance between them. In our investigated germplasm panel, we found the genome-wide LD extend of ~5cM at $r^2 \geq 0.2$ and this result corresponded to other literature findings where the LD decay ranged between 5-10 cM in wheat (Chao et al., 2010). The aforementioned features in our studies resulted in successfully identified numerous genetic loci that are associated with phenotypic traits of mineral grain concentrations.

3.2 Quantitative nature of inheritance for Ca, Fe and Zn concentrations in wheat grains

Complex traits are generally controlled by multiple loci and influenced by environmental factors; most of the agronomic and developmental traits including mineral concentrations in grains are quantitative and complex. One of the most widely used tools in genetic mapping and characterization in order to dissect the complex traits are linkage mapping and association mapping (Hu et al., 2018). Most of the mapping studies published on Ca, Fe and Zn in wheat are linkage mapping studies. For instance, earlier studies have reported QTLs based on linkage

mapping populations for Ca concentration on chromosomes 1A (Peleg et al., 2009; Shi et al., 2013), 2A (Shi et al., 2013), 2D (Shi et al., 2013), 2B, 4A, 4B, 5B, 6A, and 7B (Peleg et al., 2009). A very recent GWAS study about grain minerals accumulation in synthetic hexaploid wheat found that chromosomes 1B, 2B, 2D, 3A, 3B, 3D, 6A, 6B and 7A were controlling Ca concentration in wheat grains (Bhatta et al., 2018), while three MTAs for Fe concentration were identified on chromosomes 1A and 3A and 13 MTAs on chromosomes 1A, 2A, 3A, 3B, 4A, 4B, 5A, and 6B were found to be associated with Zn concentration.

It is difficult to align our findings with earlier studies because of the employment of different marker systems (SSR versus SNPs) (Krishnappa et al., 2017) and the lack of precise location information in previous literature due to different types of populations (Manickavelu et al., 2017). The identified genomic regions for Ca, Fe and Zn indicated that some of the significant markers are shared between these traits on the same chromosomes and positions. For instance, there were several markers that were found to be associated with Fe and Zn, e.g. 11 SNPs on chromosome 3B, 6 SNPs on chromosome 1B, 6 SNPs on chromosome 5A, and 1 SNP on chromosome 2A (Table 3). As well we observed one significant marker shared between Ca, Fe and Zn on chromosome 5A at 114.5 cM which will be discussed in the following section related to the candidate genes.

Table 3. List of significantly shared SNP makers between studied traits.

Traits	*Chrom.*	*Position (bp)*	*SNP ID*	*Shared SNP Number*
Fe-Zn	3B	723,276,117-731,264,585	BS00022039_51 BS00022025_51 Ra_c23717_305 Excalibur_c19367_76 AX-95129199 CAP8_c1113_199 IACX3169 Excalibur_c2850_126 RAC875_rep_c118396_333 RAC875_rep_c117294_342 AX-94457592	**11**
Fe-Zn	1B	655,347,159-655,797,883	AX-158556877 AX-158537662 AX-158595781 AX-109326596 AX-109510374	**5**
Fe-Zn	5A	673,720,244-676,890,191	AX-94416605 AX-94995722 AX-94857628 Kukri_c49033_52 AX-94723827	**5**
Fe-Zn	2A	719,566,769	AX-94461119	**1**
Ca-Fe-Zn	5A	698,509,966	RAC875_c8642_231	**1**

3.3 Identification of putative candidate genes for wheat grain mineral concentrations

The ultimate goal of association mapping analysis is to dissect complex traits in order to identify the functional genes or alleles that are responsible for natural phenotypic variation. Identification of significant genomic regions by GWAS analysis includes the genes based on highly associated markers in addition to genes within the LD region ($r^2 \geq 0.2$) in order to study all potential candidate genes corresponding to the desired trait. This requires the presence of a reference sequence or at least a physical map. For wheat the Genome Zipper which provided an ordered scaffold of wheat genes based on synteny of well-established genomes, such as rice (*Oryza sativa*), *Sorghum bicolor* and *Brachypodium distachyon* became available in 2013 (Spannagl et al., 2013); later a wheat draft sequence was published by the International Wheat Genome Sequencing Consortium (IWGSC) (2014), while the wheat reference sequence only became available in August 2018 (Appels et al., 2018).

Accordingly, in the first study we exploited the Genome Zipper together with the first wheat genome assemblies IWGSC1 and POPSEQ (2014) in order to identify the candidate genes for Ca

concentration in wheat grains (Alomari et al., 2017), while in chapters 2.2 and 2.3, we were able to exploit the pre-published IWGSC RefSeq v1.0 and the published wheat genome reference sequence v 1.0 respectively for data analysis during the development of these papers (Alomari et al., 2018; Alomari et al., 2019).

We were capable to detect numerous putative candidate genes underlying the natural variation of Ca, Fe and Zn concentrations in wheat grains (Table 4); the detected putative candidate genes were involved in metal uptake, transportation and hemostasis. Generally, a limited number of studies focused on studying genes for nutrient accumulation in wheat grains. A significant SNP (RAC875_c8642_231) marked a candidate gene (TraesCS5A02G542600) that annotated as transmembrane transporter activity which possibly plays a role for Ca transportation within the plant. The significant marker (RAC875_c8642_231) was common for all three investigated elements Ca, Fe and Zn on chromosome 5A at 114.5 cM indicating that it may play an important role not only in Ca accumulation but also for Fe and Zn accumulation in wheat grains.

Other candidate genes identified for Ca transportation included plasma membrane ATPase, H(+)-ATPase, heavy metal transport/detoxification, divalent metal cation transporter, calcium-transporting ATPase, mechanosensitive ion channel family protein and cyclic nucleotide-gated channel (Goel et al., 2011; Vinoth and Ravindhran, 2017).

Several putative candidate genes found to be associated with grain Zn accumulation based on their function in Zn signaling pathways included a transcription factor (TF) belonging to the basic leucine zipper (bZIP) family and the TF bHLH76, a homeobox-leucine zipper protein HOX4, a SWAP (suppressor-of-white-apricot)/surp domain-containing protein and several genes related to the mitogen-activated protein kinase (MAPK) gene family (Upadhyaya et al., 2016; Chardin et al., 2017).

Based on the functional annotation, we found genes which encoded a transcription factor (TF) related to the NAC (NAM (no apical meristem)) domain family which is involved in accelerated senescence and nutrient remobilization from leaves to grains as well nicotianamine (NA) related enzymes which are important for both of Fe and Zn radial movement through the root; several studies reported about NAC genes and increasing Fe and Zn content in the grains of wheat (Uauy et al., 2006; Ricachenevsky et al., 2013; Nadolska-Orczyk et al., 2017; Bhatta et al., 2018). Although the further investigation and functional validation of the roles of these genes in wheat

grains are still needed, the detected genes can be considered as potential candidate genes which provide a resource for enhancing nutrients in wheat grains in future breeding programs.

Table 4. Putative candidate gene predicted in genomic regions harboring grain minerals marker-trait associations.

Candidate gene description	Chromosome	Trait
Mechanosensitive ion channel family protein	2A	Ca
Cation/H^+ antiporter	2A	Ca
Divalent metal cation transporter MntH	5A	Ca
Plasma membrane ATPase	5A	Ca
H(+)-ATPase	5A	Ca
Heavy metal transport /detoxification	5A	Ca
Divalent metal cation transporter	5A	Ca
Cyclic nucleotide-gated channel	5B	Ca
SWAP	3B	Zn
MAPK	3B	Zn
bZIP	5A	Zn
bHLH76	5A	Zn
HOX4	5A	Zn
NAC	2A	Fe

3.4 Examination of the potential for applying Genomic Prediction to grain Fe concentration

Since GP approach has a great potential to accelerate the genetic gains for traits that are costly to phenotype, such as grain quality and nutritional traits; GP can increase the genetic gain per generation through early prediction of the breeding values for individuals. Therefore, we extended our analysis to include GP to predict the breeding values for grain Fe concentration. Different wheat studies revealed that the majority of GP accuracies are within the range of 0.05 and 0.8, depending on the traits, statistical methods, and experimental designs. Cross-validated prediction accuracies of grain Fe based on BLUEs in the used wheat panel was amounting to 0.38 by applying GBLUP statistical model and 0.35 by applying RR-BLUP or Bayes-Cπ model. Our result is in the range of findings of the breeding value prediction accuracies of grain Fe concentration in wheat landraces which showed a moderate degree of accuracy (Manickavelu et al., 2017). Velu et al. (2016) examined the prediction ability for grain Fe in spring wheat for two years and observed predicted breeding values which were ranged between 0.324 to 0.734 for five different environments with an average of 0.44 which is similar to our results. Low to moderate prediction accuracies could be improved by using a bigger reference germplasm panel, a higher density of SNP markers, and more powerful statistical tools.

3.5. Conclusions for application in breeding and genetic biofortification

Major food crops like wheat are insufficient sources of micronutrients and minerals required for normal human nutrition. Biofortification using different strategies such as breeding, agronomy, and genetic modification has been suggested as a promising approach to improve micronutrient and mineral contents in crop plants. Several biofortified food crops such as cereals, legumes, vegetables, and fruits were found to be successful in providing sufficient levels of micronutrients and minerals to the targeted populations (Garg et al., 2018). Considerations must also include the micronutrient and mineral accumulations in the edible portions of crops in addition to the amount of nutrients that can be absorbed by the targeted people after processing and cooking. To achieve this goal, collaboration between plant breeders, nutrition scientists, genetic engineers, and molecular biologists is essential. In spite of challenges which biofortified crops are still facing, there is a promising future to diminish malnutrition by using this approach.

4 Summary

Minerals are important components required by humans in their daily food. The genetic basis underlying the natural variations in the minerals' accumulation in wheat grains remains largely unknown. The studies were designed to discover the phenotypic and genotypic variation of Ca, Zn and Fe. To this end, a large hexaploid wheat (*Triticum aestivum* L.) germplasm panel consisting of 369 elite varieties was grown under field conditions for three years. Mineral concentrations in wheat grains were measured using inductively coupled plasma optical emission spectrometry (ICP-OES). High-throughput SNP arrays were implemented in GWAS analysis in order to detect the causative loci/genes underlying the mineral accumulation for a deeper understanding of genetic factors.

Wide natural phenotypic variation was found for the studied traits among the genotypes over the three years with moderate heritability for Fe and Zn to high heritability for Ca, indicating that these minerals are mostly controlled by genetic factors. Correlation analysis revealed a strong positive relationship between grain Fe and Zn concentrations which were both positively correlated with TKW. Even though the correlation between Ca with Zn and Fe concentrations was positive, Ca concentration was slightly negatively correlated with TKW. These findings suggest common genetic factors controlling Zn, Fe and TKW that may enable to improve them simultaneously in future breeding programs.

GWAS results indicated several QTLs distributed on several chromosomes contributing to the concentrations of the investigated minerals. While one QTL region located on chromosome 5A (698,509,966 - 698,510,066 bp) appeared as a vital common genetic region for all three minerals, a genomic region on chromosome 3B (723,276,117- 731,264,585 bp) was shared between Fe and Zn. The first wheat genome assemblies IWGSC1 along with Genome Zipper and POPSEQ were utilized to detect candidate genes associated with Ca trait. For Zn and Fe concentration with their related genes, we used the recently published wheat genome reference sequence v 1.0. The identified putative candidate genes are annotated for metal hemostasis, transportation and uptake, while some of them are known to be involved in minerals' accumulation process and other new candidates still need further characterization. For instance, calcium-transporting ATPase was linked with Ca, bZIP transcription factors and mitogen-activated protein kinase genes were associated with Zn whereas NAC gene was a candidate for Fe.

The results of genomic prediction (GP) revealed moderate prediction accuracy for grain Fe by applying different statistical models like Bayes-Cπ, GBLUP and RR-BLUP. These findings could be a promising result for using genomic selection (GS) to accelerate the breeding programs in the area of enhancing minerals in wheat grains. Our study provides crucial insights into the genetic basis of minerals' variation and accumulation in wheat grains and serves as an important foundation for further genetic and molecular mechanisms analysis.

Zusammenfassung

Mineralien sind wichtige Komponenten der menschlichen Ernährung. Jedoch ist wenig über die genetische Basis, welche der Akkumulation von Mineralien in Weizenkörnern zu Grunde liegt, bekannt. In dieser Arbeit sollte die phänotypische und genotypische Variation von Ca, Zn und Fe studiert werden. Dazu wurde eine große Kollektion von Weizenlinien (*Triticum aestivum* L.), welche 369 Elitesorten beinhaltete, für drei Jahre auf dem Feld angebaut. Die Mineralienkonzentrationen in den Weizenkörnern wurden mittels optischer Emissionsspektrometrie mit induktiv gekoppeltem Plasma (ICP-OES) bestimmt. Hochdurchsatz SNP-chips wurden für die GWAS Analyse genutzt, um die ursächlichen Loci/Gene für die Mineralienakkumulation zu finden und die genetischen Faktoren besser zu verstehen.

Für die untersuchten Merkmale wurde eine weite natürliche phänotypische Variation zwischen den Genotypen über die drei Jahre hinweg beobachtet; dabei waren die Heritabilitäten für Fe und Zn moderat und hoch für Ca, was eine starke Kontrolle von genetischen Faktoren impliziert. Hohe positive Korrelationen wurden für Fe und Zn Konzentrationen in den Körnern gefunden, welche beide positiv mit TKG korreliert waren. Obwohl die Korrelationen von Ca Konzentration mit Fe und Zn positiv waren, war die Ca Konzentration leicht negativ mit TKG korreliert. Daraus kann man schließen, dass gemeinsame genetische Faktoren für die Kontrolle von Zn, Fe und TKG existieren und diese Merkmale daher gemeinsam züchterisch bearbeitet werden können.

GWAS resultierte in mehreren QTL auf verschiedenen Chromosomen für die Konzentrationen der untersuchten Mineralien. Während eine QTL-Region auf Chromosom 5A (698.509.966-698.510.066 bp) relevant für alle drei Mineralien war, war eine genomische Region auf Chromosom 3B (723.276.117- 731.264.585 bp) signifikant für Fe und Zn. Die erste Weizengenomassemblierung IWGSC1, sowie der Genome Zipper und POPSEQ wurden für die Detektion von Kandidatengenen für Ca genutzt. Für die Auffindung der Gene für Zn und Fe Konzentrationen nutzten wir die erst kürzlich publizierte Referenzsequenz v1.0. Die identifizierten möglichen Kandidatengene hatten Annotationen für Metallhämostase, Transport, Aufnahme und Akkumulierung, während andere Kandidaten noch weiterer Charakterisierung bedürfen. So waren eine ‚calcium-transporting ATPase' mit Ca, bZIP Transkriptionsfaktoren und

‚mitogen-activated protein kinase genes' mit Zn und ein ‚NAC gene' mit Fe Konzentration assoziiert.

Ergebnisse für ‚genomic prediction' (GP) für Fe Konzentration im Korn zeigten moderate Vorhersagegenauigkeiten für verschiedene statistische Modelle einschließlich Bayes-Cπ, GBLUP and RR-BLUP. Diese Ergebnisse beleuchten den Einsatz von genomischer Selektion (GS) um Züchtungsprogramme für einen erhöhten Mineralstoffgehalt in Weizenkörnern zu beschleunigen. Diese Studie gibt entscheidende Einblicke in die genetische Basis der Variation und Akkumulierung von Mineralien in Weizenkörnern und stellt eine wichtige Grundlage für weitere genetische und molekulare Studien zur Aufklärung der Mechanismen dar.

References

(2014). A chromosome-based draft sequence of the hexaploid bread wheat (Triticum aestivum) genome. *Science* 345**,** 1251788.

Aciksoz, S.B., Yazici, A., Ozturk, L., and Cakmak, I. (2011). Biofortification of wheat with iron through soil and foliar application of nitrogen and iron fertilizers. *Plant and Soil* 349**,** 215-225.

Allen, A.M., Winfield, M.O., Burridge, A.J., Downie, R.C., Benbow, H.R., Barker, G.L.A., Wilkinson, P.A., Coghill, J., Waterfall, C., Davassi, A., Scopes, G., Pirani, A., Webster, T., Brew, F., Bloor, C., Griffiths, S., Bentley, A.R., Alda, M., Jack, P., Phillips, A.L., and Edwards, K.J. (2017). Characterization of a Wheat Breeders' Array suitable for high-throughput SNP genotyping of global accessions of hexaploid bread wheat (Triticum aestivum). *Plant Biotechnology Journal* 15**,** 390-401.

Alomari, D.Z., Eggert, K., Von Wirén, N., Alqudah, A.M., Polley, A., Plieske, J., Ganal, M.W., Pillen, K., and Röder, M.S. (2018). Identifying Candidate Genes for Enhancing Grain Zn Concentration in Wheat. *Frontiers in Plant Science* 9.

Alomari, D.Z., Eggert, K., Von Wiren, N., Pillen, K., and Roder, M.S. (2017). Genome-Wide Association Study of Calcium Accumulation in Grains of European Wheat Cultivars. *Front Plant Sci* 8**,** 1797.

Alomari, Z.D., Eggert, K., Von Wirén, N., Polley, A., Plieske, J., Ganal, W.M., Liu, F., Pillen, K., and Röder, S.M. (2019). Whole-Genome Association Mapping and Genomic Prediction for Iron Concentration in Wheat Grains. *International Journal of Molecular Sciences* 20.

Ando, H., Sugi, K., Watanabe, K., Morita, N., and Mitsunaga, T. (2002). Distribution of Food Components in Each Fraction of Wheat Grain. *Food Science and Technology Research* 8**,** 10-13.

Appels, R., Eversole, K., Feuillet, C., Keller, B., Rogers, J., Stein, N., Pozniak, C.J., Stein, N., Choulet, F., Distelfeld, A., Eversole, K., Poland, J., Rogers, J., Ronen, G., Sharpe, A.G., Pozniak, C., Ronen, G., Stein, N., Barad, O., Baruch, K., Choulet, F., Keeble-Gagnère, G., Mascher, M., Sharpe, A.G., Ben-Zvi, G., Josselin, A.-A., Stein, N., Mascher, M., Himmelbach, A., Choulet, F., Keeble-Gagnère, G., Mascher, M., Rogers, J., Balfourier, F., Gutierrez-Gonzalez, J., Hayden, M., Josselin, A.-A., Koh, C., Muehlbauer, G., Pasam, R.K., Paux, E., Pozniak, C.J., Rigault, P., Sharpe, A.G., Tibbits, J., Tiwari, V., Choulet, F., Keeble-Gagnère, G., Mascher, M., Josselin, A.-A., Rogers, J., Spannagl, M., Choulet, F., Lang, D., Gundlach, H., Haberer, G., Keeble-Gagnère, G., Mayer, K.F.X., Ormanbekova, D., Paux, E., Prade, V., Šimková, H., Wicker, T., Choulet, F., Spannagl, M., Swarbreck, D., Rimbert, H., Felder, M., Guilhot, N., Gundlach, H., Haberer, G., Kaithakottil, G., Keilwagen, J., Lang, D., Leroy, P., Lux, T., Mayer, K.F.X., Twardziok, S., Venturini, L., Appels, R., Rimbert, H., Choulet, F., Juhász, A., Keeble-Gagnère, G., Choulet, F., Spannagl, M., Lang, D., Abrouk, M., Haberer, G., Keeble-Gagnère, G., Mayer, K.F.X., Wicker, T., Choulet, F., Wicker, T., Gundlach, H., Lang, D., Spannagl, M., Lang, D., Spannagl, M., Appels, R., et al. (2018). Shifting the limits in wheat research and breeding using a fully annotated reference genome. *Science* 361**,** eaar7191.

Arora, S., Cheema, J., Poland, J., Uauy, C., and Chhuneja, P. (2019). Genome-Wide Association Mapping of Grain Micronutrients Concentration in Aegilops tauschii. *Frontiers in Plant Science* 10.

Balandrán-Quintana, R.R., Mercado-Ruiz, J.N., and Mendoza-Wilson, A.M. (2015). Wheat Bran Proteins: A Review of Their Uses and Potential. *Food Reviews International* 31**,** 279-293.

Barron, C., Surget, A., and Rouau, X. (2007). Relative amounts of tissues in mature wheat (Triticum aestivum L.) grain and their carbohydrate and phenolic acid composition. *Journal of Cereal Science* 45**,** 88-96.

Bhatta, M., Baenziger, P.S., Waters, B.M., Poudel, R., Belamkar, V., Poland, J., and Morgounov, A. (2018). Genome-Wide Association Study Reveals Novel Genomic Regions Associated with 10 Grain Minerals in Synthetic Hexaploid Wheat. *Int J Mol Sci* 19.

Björck, I., Östman, E., Kristensen, M., Mateo Anson, N., Price, R.K., Haenen, G.R.M.M., Havenaar, R., Bach Knudsen, K.E., Frid, A., Mykkänen, H., Welch, R.W., and Riccardi, G. (2012). Cereal grains for nutrition and health benefits: Overview of results from in vitro, animal and human studies in the HEALTHGRAIN project. *Trends in Food Science & Technology* 25**,** 87-100.

Bouis, H.E. (2002). Plant Breeding: A New Tool for Fighting Micronutrient Malnutrition. *The Journal of Nutrition* 132**,** 491S-494S.

Bouis, H.E. (2007). Micronutrient fortification of plants through plant breeding: can it improve nutrition in man at low cost? *Proceedings of the Nutrition Society* 62**,** 403-411.

Bouis, H.E., Hotz, C., Mcclafferty, B., Meenakshi, J.V., and Pfeiffer, W.H. (2011). Biofortification: A New Tool to Reduce Micronutrient Malnutrition. *Food and Nutrition Bulletin* 32**,** S31-S40.

Cakmak, I. (2000). Tansley Review No. 111 Possible roles of zinc in protecting plant cells from damage by reactive oxygen species. *New Phytologist* 146**,** 185-205.

Cakmak, I. (2002). Plant nutrition research: Priorities to meet human needs for food in sustainable ways. *Plant and Soil* 247**,** 3-24.

Cakmak, I. (2008). Enrichment of cereal grains with zinc: Agronomic or genetic biofortification? *Plant and Soil* 302**,** 1-17.

Calus, M.P.L., Huang, H., Vereijken, A., Visscher, J., Ten Napel, J., and Windig, J.J. (2014). Genomic prediction based on data from three layer lines: a comparison between linear methods. *Genetics Selection Evolution* 46**,** 57.

Cardon, L.R., and Bell, J.I. (2001). Association study designs for complex diseases. *Nature Reviews Genetics* 2**,** 91.

Centeno, V., Díaz De Barboza, G., Marchionatti, A., Rodríguez, V., and Tolosa De Talamoni, N. (2009). Molecular mechanisms triggered by low-calcium diets. *Nutrition Research Reviews* 22**,** 163-174.

Chao, S., Dubcovsky, J., Dvorak, J., Luo, M.-C., Baenziger, S.P., Matnyazov, R., Clark, D.R., Talbert, L.E., Anderson, J.A., Dreisigacker, S., Glover, K., Chen, J., Campbell, K., Bruckner, P.L., Rudd, J.C., Haley, S., Carver, B.F., Perry, S., Sorrells, M.E., and Akhunov, E.D. (2010). Population- and genome-specific patterns of linkage disequilibrium and SNP variation in spring and winter wheat (Triticum aestivum L.). *BMC Genomics* 11**,** 727.

Chardin, C., Schenk, S.T., Hirt, H., Colcombet, J., and Krapp, A. (2017). Review: Mitogen-Activated Protein Kinases in nutritional signaling in Arabidopsis. *Plant Sci* 260**,** 101-108.

Collard Bertrand, C.Y., and Mackill David, J. (2008). Marker-assisted selection: an approach for precision plant breeding in the twenty-first century. *Philosophical Transactions of the Royal Society B: Biological Sciences* 363**,** 557-572.

Crespo-Herrera, L.A., Velu, G., and Singh, R.P. (2016). Quantitative trait loci mapping reveals pleiotropic effect for grain iron and zinc concentrations in wheat. *Annals of Applied Biology* 169**,** 27-35.

Crossa, J., Pérez-Rodríguez, P., Cuevas, J., Montesinos-López, O., Jarquin, D., De Los Campos, G., Burgueño, J., M. Camacho-González, J., Perez-Elizalde, S., Beyene, Y., Dreisigacker, S., Singh, R., Zhang, X., Gowda, M., Roorkiwal, M., Rutkoski, J., and K. Varshney, R. (2017). *Genomic Selection in Plant Breeding: Methods, Models, and Perspectives.*

Dayod, M., Tyerman, S.D., Leigh, R.A., and Gilliham, M. (2010). Calcium storage in plants and the implications for calcium biofortification. *Protoplasma* 247**,** 215-231.

De Brier, N., Gomand, S.V., Donner, E., Paterson, D., Delcour, J.A., Lombi, E., and Smolders, E. (2015). "Distribution of minerals in wheat grains (Triticum aestivum L.) and in roller milling fractions affected by pearling". US American Chemical Society).

De Brier, N., Gomand, S.V., Donner, E., Paterson, D., Smolders, E., Delcour, J.A., and Lombi, E. (2016). Element distribution and iron speciation in mature wheat grains (Triticum aestivum L.) using synchrotron X-ray fluorescence microscopy mapping and X-ray absorption near-edge structure (XANES) imaging. *Plant, Cell & Environment* 39**,** 1835-1847.

Delcour, J.A., and Hoseney, R.C. (2010). *Principles of Cereal Science and Technology.* AACC International.

Duffner, A., Hoffland, E., Stomph, T.J., Melse, A., and Bindraban, P.S. (2014). "Eliminating Zinc Deficiency in Rice-Based Systems".).

Fao (2018). "THE STATE OF THE WORLD", in: *The State of Food Security and Nutrition in the World 2018.*).

Faostat (2018). "Food and agriculture organization of the united nations, satatistical database". (Italy).

Fujino, Y., Kuwata, J., Mano, Y., and Ohnishi, M. (1996). "Other grain components."), 289-317.

Garg, M., Sharma, N., Sharma, S., Kapoor, P., Kumar, A., Chunduri, V., and Arora, P. (2018). Biofortified Crops Generated by Breeding, Agronomy, and Transgenic Approaches Are Improving Lives of Millions of People around the World. *Frontiers in Nutrition* 5.

Goel, A., Taj, G., Pandey, D., Gupta, S., and Kumar, A. (2011). Genome-Wide Comparative in silico Analysis of Calcium Transporters of Rice and Sorghum. *Genomics, Proteomics & Bioinformatics* 9**,** 138-150.

Habier, D., Fernando, R.L., Kizilkaya, K., and Garrick, D.J. (2011). Extension of the bayesian alphabet for genomic selection. *BMC Bioinformatics* 12**,** 186.

Hu, X., Zuo, J., Wang, J., Liu, L., Sun, G., Li, C., Ren, X., and Sun, D. (2018). Multi-Locus Genome-Wide Association Studies for 14 Main Agronomic Traits in Barley. *Frontiers in Plant Science* 9.

Idrc (2015). *International Development Research Centre, Facts & Figures on Food and Biodiversity* [Online]. [Accessed].

International Wheat Genome Sequencing, C. (2014). A chromosome-based draft sequence of the hexaploid bread wheat (Triticum aestivum) genome. *Science* 345**,** 1251788.

Khokhar, J.S., Sareen, S., Tyagi, B.S., Singh, G., Wilson, L., King, I.P., Young, S.D., and Broadley, M.R. (2018). Variation in grain Zn concentration, and the grain ionome, in field-grown Indian wheat. *PLoS One* 13**,** e0192026.

Kihara, H. (1924). Cytologische und genetische Studien bei wichtigen Getreidearten mit besonderer Rucksicht auf das Verhalten der Chromosomen und die Sterilitat in den Bastarden. *Mem. Coll. Sci., Kyoto Imp. Univ.* 1**,** 1-200.

Kollers, S., Rodemann, B., Ling, J., Korzun, V., Ebmeyer, E., Argillier, O., Hinze, M., Plieske, J., Kulosa, D., Ganal, M.W., and Roder, M.S. (2013). Whole genome association mapping of Fusarium head blight resistance in European winter wheat (Triticum aestivum L.). *PLoS One* 8**,** e57500.

Korte, A., and Farlow, A. (2013). The advantages and limitations of trait analysis with GWAS: a review. *Plant methods* 9**,** 29-29.

Krishnappa, G., Singh, A.M., Chaudhary, S., Ahlawat, A.K., Singh, S.K., Shukla, R.B., Jaiswal, J.P., Singh, G.P., and Solanki, I.S. (2017). Molecular mapping of the grain iron and zinc concentration, protein content and thousand kernel weight in wheat (Triticum aestivum L.). *PLoS One* 12**,** e0174972.

Kumssa, D.B., Joy, E.J.M., Ander, E.L., Watts, M.J., Young, S.D., Walker, S., and Broadley, M.R. (2015). Dietary calcium and zinc deficiency risks are decreasing but remain prevalent. *Scientific Reports* 5**,** 10974.

Li, H., Su, G., Jiang, L., and Bao, Z. (2017). An efficient unified model for genome-wide association studies and genomic selection. *Genetics Selection Evolution* 49**,** 64.

Liu, J., Xu, Z., Fan, X., Zhou, Q., Cao, J., Wang, F., Ji, G., Yang, L., Feng, B., and Wang, T. (2018). A Genome-Wide Association Study of Wheat Spike Related Traits in China. *Frontiers in Plant Science* 9.

Lopez, H.W., Leenhardt, F., Coudray, C., and Remesy, C. (2002). Minerals and phytic acid interactions: is it a real problem for human nutrition? *International Journal of Food Science & Technology* 37**,** 727-739.

Lukaski, H.C. (2004). Vitamin and mineral status: effects on physical performance. *Nutrition* 20**,** 632-644.

Mackay, T.F.C. (2001). The Genetic Architecture of Quantitative Traits. *Annual Review of Genetics* 35**,** 303-339.

Manickavelu, A., Hattori, T., Yamaoka, S., Yoshimura, K., Kondou, Y., Onogi, A., Matsui, M., Iwata, H., and Ban, T. (2017). Genetic Nature of Elemental Contents in Wheat Grains and Its Genomic Prediction: Toward the Effective Use of Wheat Landraces from Afghanistan. *PLOS ONE* 12**,** e0169416.

Marcussen, T., Sandve, S.R., Heier, L., Spannagl, M., Pfeifer, M., International Wheat Genome Sequencing, C., Jakobsen, K.S., Wulff, B.B., Steuernagel, B., Mayer, K.F., and Olsen, O.A. (2014). Ancient hybridizations among the ancestral genomes of bread wheat. *Science* 345**,** 1250092.

Matsuoka, Y. (2011). Evolution of Polyploid Triticum Wheats under Cultivation: The Role of Domestication, Natural Hybridization and Allopolyploid Speciation in their Diversification. *Plant and Cell Physiology* 52**,** 750-764.

Maulana, F., Ayalew, H., Anderson, J.D., Kumssa, T.T., Huang, W., and Ma, X.-F. (2018). Genome-Wide Association Mapping of Seedling Heat Tolerance in Winter Wheat. *Frontiers in Plant Science* 9.

Meuwissen, T.H.E., Hayes, B.J., and Goddard, M.E. (2001). Prediction of Total Genetic Value Using Genome-Wide Dense Marker Maps. *Genetics* 157**,** 1819-1829.

Mohan, M., Nair, S., Bhagwat, A., Krishna, T.G., Yano, M., Bhatia, C.R., and Sasaki, T. (1997). Genome mapping, molecular markers and marker-assisted selection in crop plants. *Molecular Breeding* 3**,** 87-103.

Myles, S., Peiffer, J., Brown, P.J., Ersoz, E.S., Zhang, Z., Costich, D.E., and Buckler, E.S. (2009). Association mapping: critical considerations shift from genotyping to experimental design. *Plant Cell* 21**,** 2194-2202.

Nadolska-Orczyk, A., Rajchel, I.K., Orczyk, W., and Gasparis, S. (2017). Major genes determining yield-related traits in wheat and barley. *Theor Appl Genet* 130**,** 1081-1098.

Ober, U., Erbe, M., Long, N., Porcu, E., Schlather, M., and Simianer, H. (2011). Predicting Genetic Values: A Kernel-Based Best Linear Unbiased Prediction With Genomic Data. *Genetics* 188**,** 695.

Ozturk, L., Yazici, M.A., Yucel, C., Torun, A., Cekic, C., Bagci, A., Ozkan, H., Braun, H.-J., Sayers, Z., and Cakmak, I. (2006). Concentration and localization of zinc during seed development and germination in wheat. *Physiologia Plantarum* 128**,** 144-152.

Paux, E., Sourdille, P., Salse, J., Saintenac, C., Choulet, F., Leroy, P., Korol, A., Michalak, M., Kianian, S., Spielmeyer, W., Lagudah, E., Somers, D., Kilian, A., Alaux, M., Vautrin, S., Bergès, H., Eversole, K., Appels, R., Safar, J., Simkova, H., Dolezel, J., Bernard, M., and Feuillet, C. (2008). A Physical Map of the 1-Gigabase Bread Wheat Chromosome 3B. *Science* 322**,** 101.

Peleg, Z., Cakmak, I., Ozturk, L., Yazici, A., Jun, Y., Budak, H., Korol, A.B., Fahima, T., and Saranga, Y. (2009). Quantitative trait loci conferring grain mineral nutrient concentrations in durum wheat × wild emmer wheat RIL population. *Theoretical and Applied Genetics* 119**,** 353-369.

Pérez-De-Castro, A.M., Vilanova, S., Cañizares, J., Pascual, L., Blanca, J.M., Díez, M.J., Prohens, J., and Picó, B. (2012). Application of genomic tools in plant breeding. *Current genomics* 13**,** 179-195.

Piepho, H.-P., Ogutu, J., Schulz-Streeck, T., Estaghvirou, B., Gordillo, A., and Technow, F. (2012). *Efficient Computation of Ridge-Regression Best Linear Unbiased Prediction in Genomic Selection in Plant Breeding.*

Piste, P., Sayaji, D., and Avinash, M. (2012). *Calcium and its Role in Human Body.*

Pritchard, J.K., Stephens, M., Rosenberg, N.A., and Donnelly, P. (2000). Association Mapping in Structured Populations. *The American Journal of Human Genetics* 67**,** 170-181.

Rafalski, A. (2002). Applications of single nucleotide polymorphisms in crop genetics. *Current Opinion in Plant Biology* 5**,** 94-100.

Rasheed, A., Hao, Y., Xia, X., Khan, A., Xu, Y., Varshney, R.K., and He, Z. (2017). Crop Breeding Chips and Genotyping Platforms: Progress, Challenges, and Perspectives. *Molecular Plant* 10**,** 1047-1064.

Rasheed, A., Ogbonnaya, F.C., Lagudah, E., Appels, R., and He, Z. (2018). The goat grass genome's role in wheat improvement. *Nature Plants* 4**,** 56-58.

Ricachenevsky, F., Menguer, P., and Sperotto, R. (2013). kNACking on heaven's door: how important are NAC transcription factors for leaf senescence and Fe/Zn remobilization to seeds? *Frontiers in Plant Science* 4.

Rout, G., and Sahoo, S. (2015). *Role of iron in plant growth and metabolism.*

Salamini, F., Özkan, H., Brandolini, A., Schäfer-Pregl, R., and Martin, W. (2002). Genetics and geography of wild cereal domestication in the near east. *Nature Reviews Genetics* 3**,** 429.

Shenkin, A. (2006). Micronutrients in health and disease. *Postgraduate medical journal* 82**,** 559-567.

Shewry, P.R. (2009). Wheat. *Journal of Experimental Botany* 60**,** 1537-1553.

Shi, R.-L., Tong, Y.-P., Jing, R.-L., Zhang, F.-S., and Zou, C.-Q. (2013). Characterization of Quantitative Trait Loci for Grain Minerals in Hexaploid Wheat (Triticum aestivum L.). *Journal of Integrative Agriculture* 12**,** 1512-1521.

Singh, S.P., Vogel-Mikuš, K., Arčon, I., Vavpetič, P., Jeromel, L., Pelicon, P., Kumar, J., and Tuli, R. (2013). Pattern of iron distribution in maternal and filial tissues in wheat grains with contrasting levels of iron. *Journal of Experimental Botany* 64**,** 3249-3260.

Soto-Cerda, B.J., and Cloutier, S. (2012). "Association mapping in plant genomes," in *Genetic diversity in plants*. InTech).

Spannagl, M., Martis, M.M., Pfeifer, M., Nussbaumer, T., and Mayer, K.F.X. (2013). Analysing complex Triticeae genomes – concepts and strategies. *Plant Methods* 9**,** 35.

Sramkova, Z., Gregová, E., and Šturdík, E. (2009). *Chemical composition and nutritional quality of wheat grain-Review.*

Srinivasa, J., Arun, B., Mishra, V.K., Singh, G.P., Velu, G., Babu, R., Vasistha, N.K., and Joshi, A.K. (2014). Zinc and iron concentration QTL mapped in a Triticum spelta × T. aestivum cross. *Theoretical and Applied Genetics* 127**,** 1643-1651.

Sukumaran, S., Lopes, M., Dreisigacker, S., and Reynolds, M. (2018). Genetic analysis of multi-environmental spring wheat trials identifies genomic regions for locus-specific trade-offs for grain weight and grain number. *Theoretical and Applied Genetics* 131**,** 985-998.

Tang, J., Zou, C., He, Z., Shi, R., Ortiz-Monasterio, I., Qu, Y., and Zhang, Y. (2008). Mineral element distributions in milling fractions of Chinese wheats. *Journal of Cereal Science* 48**,** 821-828.

Tian, F., Bradbury, P.J., Brown, P.J., Hung, H., Sun, Q., Flint-Garcia, S., Rocheford, T.R., Mcmullen, M.D., Holland, J.B., and Buckler, E.S. (2011). Genome-wide association study of leaf architecture in the maize nested association mapping population. *Nature Genetics* 43**,** 159.

Uauy, C., Distelfeld, A., Fahima, T., Blechl, A., and Dubcovsky, J. (2006). A NAC Gene Regulating Senescence Improves Grain Protein, Zinc, and Iron Content in Wheat. *Science* 314**,** 1298.

Upadhyaya, H.D., Bajaj, D., Das, S., Kumar, V., Gowda, C.L., Sharma, S., Tyagi, A.K., and Parida, S.K. (2016). Genetic dissection of seed-iron and zinc concentrations in chickpea. *Sci Rep* 6**,** 24050.

Velu, G., Crossa, J., Singh, R.P., Hao, Y., Dreisigacker, S., Perez-Rodriguez, P., Joshi, A.K., Chatrath, R., Gupta, V., Balasubramaniam, A., Tiwari, C., Mishra, V.K., Sohu, V.S., and Mavi, G.S. (2016). Genomic prediction for grain zinc and iron concentrations in spring wheat. *Theoretical and Applied Genetics* 129**,** 1595-1605.

Velu, G., Singh, R.P., Crespo-Herrera, L., Juliana, P., Dreisigacker, S., Valluru, R., Stangoulis, J., Sohu, V.S., Mavi, G.S., Mishra, V.K., Balasubramaniam, A., Chatrath, R., Gupta, V., Singh, G.P., and Joshi, A.K. (2018). Genetic dissection of grain zinc concentration in spring wheat for mainstreaming biofortification in CIMMYT wheat breeding. *Scientific Reports* 8**,** 13526.

Vinoth, A., and Ravindhran, R. (2017). Biofortification in Millets: A Sustainable Approach for Nutritional Security. *Frontiers in Plant Science* 8.

Wang, S.-X., Zhu, Y.-L., Zhang, D.-X., Shao, H., Liu, P., Hu, J.-B., Zhang, H., Zhang, H.-P., Chang, C., Lu, J., Xia, X.-C., Sun, G.-L., and Ma, C.-X. (2017). Genome-wide association study for grain yield and related traits in elite wheat varieties and advanced lines using SNP markers. *PLOS ONE* 12**,** e0188662.

Wang, S., Wong, D., Forrest, K., Allen, A., Chao, S., Huang, B.E., Maccaferri, M., Salvi, S., Milner, S.G., Cattivelli, L., Mastrangelo, A.M., Whan, A., Stephen, S., Barker, G., Wieseke, R., Plieske, J., International Wheat Genome Sequencing, C., Lillemo, M., Mather, D., Appels, R., Dolferus, R., Brown-Guedira, G., Korol, A., Akhunova, A.R., Feuillet, C., Salse, J., Morgante, M., Pozniak, C., Luo, M.C., Dvorak, J., Morell, M., Dubcovsky, J., Ganal, M., Tuberosa, R., Lawley, C., Mikoulitch, I., Cavanagh, C., Edwards, K.J., Hayden, M., and Akhunov, E. (2014). Characterization of polyploid wheat genomic diversity using a high-density 90,000 single nucleotide polymorphism array. *Plant Biotechnol J* 12**,** 787-796.

Weir, B.S. (2010). Statistical genetic issues for genome-wide association studiesThis article is one of a selection of papers from the conference "Exploiting Genome-wide Association in Oilseed Brassicas: a model for genetic improvement of major OECD crops for sustainable farming". *Genome* 53**,** 869-875.

Wen, W., He, Z., Gao, F., Liu, J., Jin, H., Zhai, S., Qu, Y., and Xia, X. (2017). A High-Density Consensus Map of Common Wheat Integrating Four Mapping Populations Scanned by the 90K SNP Array. *Frontiers in Plant Science* 8.

White, P.J., and Broadley, M.R. (2005). Biofortifying crops with essential mineral elements. *Trends in Plant Science* 10**,** 586-593.

Xu, Y., An, D., Liu, D., Zhang, A., Xu, H., and Li, B. (2012). Molecular mapping of QTLs for grain zinc, iron and protein concentration of wheat across two environments. *Field Crops Research* 138**,** 57-62.

Yu, J., Pressoir, G., Briggs, W.H., Vroh Bi, I., Yamasaki, M., Doebley, J.F., Mcmullen, M.D., Gaut, B.S., Nielsen, D.M., Holland, J.B., Kresovich, S., and Buckler, E.S. (2006). A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. *Nat Genet* 38**,** 203-208.

Zhao, F.J., Su, Y.H., Dunham, S.J., Rakszegi, M., Bedo, Z., Mcgrath, S.P., and Shewry, P.R. (2009). Variation in mineral micronutrient concentrations in grain of wheat lines of diverse origin. *Journal of Cereal Science* 49**,** 290-295.

Zhao, K., Tung, C.-W., Eizenga, G.C., Wright, M.H., Ali, M.L., Price, A.H., Norton, G.J., Islam, M.R., Reynolds, A., Mezey, J., Mcclung, A.M., Bustamante, C.D., and Mccouch, S.R. (2011). Genome-wide association mapping reveals a rich genetic architecture of complex traits in Oryza sativa. *Nature communications* 2**,** 467-467.

Zhu, C., Naqvi, S., Gomez-Galera, S., Pelacho, A.M., Capell, T., and Christou, P. (2007). Transgenic strategies for the nutritional enhancement of plants. *Trends in Plant Science* 12**,** 548-555.

Zondervan, K.T., and Cardon, L.R. (2004). The complex interplay among factors that influence allelic association. *Nature Reviews Genetics* 5**,** 89.

http://www.fao.org/docrep/018/i3107e/i3107e03.pdf

https://ourworldindata.org/micronutrient-deficiency

Eidesstattliche Erklärung
Declaration under Oath

Ich erkläre an Eides statt, dass ich die Arbeit selbstständig und ohne fremde Hilfe verfasst, keine anderen als die von mir angegebenen Quellen und Hilfsmittel benutzt und die den benutzten Werken wörtlich oder inhaltlich entnommenen Stellen als solche kenntlich gemacht habe.

I declare under penalty of perjury that this thesis is my own work entirely and has been written without any help from other people. I used only the sources mentioned and included all the citations correctly both in word or content.

_____ _____

Datum / Date Unterschrift des Antragstellers / Signature of the applicant

Erklärung über bestehende Vorstrafen und anhängige Ermittlungsverfahren
Declaration concerning Criminal Record and Pending Investigations

Hiermit erkläre ich, dass ich weder vorbestraft bin noch dass gegen mich Ermittlungsverfahren anhängig sind.

I hereby declare that I have no criminal record and that no preliminary investigations are pending against me.

_____ _____

Datum / Date Unterschrift des Antragstellers / Signature of the applicant

Curriculum vitae

Dalia Zakaria Saleh Alomari

Gene and genome mapping, Breeding research department, Leibniz Institute of Plant Genetics and Crop Plant Research (IPK)
Corrensstraße 3,
OT Gatersleben
06466 Stadt Seeland, Germany.

Gender: Female
Marital status: Married
Nationality: Jordanian
Place of Birth: Saudi Arabia
Date of Birth: May 19th 1984

Qualification

Ph.D. Oct., 2015- July, 2019

Ph.D. work carried out at IPK Gatersleben under supervision of Dr. Marion Roeder, Gene and genome mapping Group, Breeding research department.

M.Sc. 2011
Nutrition and Food Technology Science with a specialization in Nutrition Science, Jordan University of Science and Technology (JUST), Irbid- Jordan.

B.Sc. 2006
Nutrition and Food Technology Science, Jordan University of Science and Technology (JUST), Irbid- Jordan.

Experience

- **Researcher**

- **Applied Biochemistry Group**, Leibniz Institute of Plant Genetics and Crop Plant Research (IPK) (Dec. 2012 – Dec. 2013).
- **Biochemistry Lab, Chemical Analysis and Food Microbiology Labs**, Department of Nutrition and Food Technology (JUST) (April 2009 – Feb. 2011).

- **Food Microbiology and Chemical Analysis Labs**, Department of Nutrition and Food Technology (JUST) (May 2007- March 2009).

- **Teaching Assistant**

- **Diet Therapy; Microbiology; Food Technology; Fruit and Vegetable; Meal Planning; Dairy Technology labs.** Department of Nutrition and Food Technology (Faculty of Agriculture, JUST) (Sep. 2006-Jan, 2009).

Publications

1. **Alomari, D.Z.,** Eggert, K., Von Wirén, N., Polley, A., Plieske, J., Ganal, W.M., Liu, F., Pillen, K., and Röder, S.M. (**2019**). Whole-Genome Association Mapping and Genomic Prediction for Iron Concentration in Wheat Grains. *International Journal of Molecular Sciences* 20.

2. **Alomari, D.Z.,** Eggert, K., Von Wirén, N., Alqudah, A.M., Polley, A., Plieske, J., Ganal, M.W., Pillen, K., and Röder, M.S. (**2018**). Identifying Candidate Genes for Enhancing Grain Zn Concentration in Wheat. *Frontiers in Plant Science* 9.

3. **Alomari, D.Z.,** Eggert, K., Von Wiren, N., Pillen, K., and Roder, M.S. (**2017**). Genome-Wide Association Study of Calcium Accumulation in Grains of European Wheat Cultivars. *Front Plant Sci* 8**,** 1797.

4. **Alomari, D.Z.,** Abdul-Hussain, S.S., Ajo, R.Y. (**2016**) Germinated lupin (Lupinus albus) flour improves Arabic flat bread properties. Qual. Assur. Saf. Crops Foods 8, 57–63. **DOI:** http://dx.doi.org/10.3920/QAS2014.0441

5. **Alomari, D.Z.,** Abdul-Hussain, S.S. (**2013**) Effect of Lupin Flour Supplementation on Chemical, Physical and Sensory Properties of Mediterranean Flat Bread. Int. J. Food Sci. Nutr. Eng. 3, 49–54. **DOI:** 10.5923/j.food.20130304.01

6. Obeidat B. A., Abdul-Hussain, S.S., and **Alomari, D.Z. (2012)** "Effect of addition of germinated lupin flour on the physiochemical and organoleptic properties of cookies". Journal of Food Processing and Preservation, vol. 37, pp. 637– 643. **DOI:**10.1111/j.1745-4549.2012.00688.x

Oral &Poster presentations

Dalia Z. Alomari; Kai Eggert; Nicolaus von Wirén; Klaus Pillen; Marion S. Röder. Genome-Wide Association Study of Calcium Accumulation in Grains of European Wheat Cultivars "14th Plant Science Student Conference"/**oral presentation**. IPK, Gatersleben/Germany, 19.06.2018 – 22.06.2018.

Dalia Z. Alomari; Kai Eggert; Nicolaus von Wirén; Klaus Pillen; Marion S. Röder. Genome-wide association mapping of mineral concentrations in grains of bread wheat (*Triticum aestivum L.*). "German plant breeding conference". Wernigrode/Germany, 28.02.2018 – 02.03.2018.

Dalia Z. Alomari; Kai Eggert; Nicolaus von Wirén; Klaus Pillen; Marion S. Röder. Genome-wide association mapping of mineral concentrations in grains of bread wheat (*Triticum aestivum L.*). "13th international Wheat Genetics Symposium". Tulln/Austria, 24 – 27 April 2017.

Dalia Z. Alomari; Kai Eggert; Nicolaus von Wirén; Klaus Pillen; Marion S. Röder. Genome-wide association mapping of mineral contents in bread wheat (Triticum aestivum L.) grain. "5. Quedlinburger Pflanzenzüchtungstage". The Leibniz Institute of Plant Genetics and Crop Plant Research (IPK) in Gatersleben, 01-03 March 2017.

Selma S. Abdul-Hussain, Bayan A. Obeidat, **Dalia Z. Alomari.** Effect of addition of germinated Lupin flour on the physiochemical and organoleptic properties of cookies. The Jordanian Nutritional Conference. Jordan. 28-29 April, 2010.