Genotypic and phenotypic analysis of the spike row-type in barley (*Hordeum vulgare* L.)

Kumulative 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

Helmy Mohamed Youssef Ibrahim

geb. am: 14.06.1973 in Giza, Egypt

Gutachter:
PD Dr. Thorsten Schnurbusch
Prof. Dr. Andreas Graner
Prof. Dr. Takao Komatsuda
Verteidigt am:
01st February 2016

Acknowledgements

In the name of Allah, The Most Gracious, The Most Merciful

I am grateful to Allah, our Lord and Cherisher, for guiding me to conceptualize, develop and complete the thesis. Indeed, without his help and will, nothing is accomplished.

I would like to convey my deepest thanks to my supervisor and mentor, Dr. Thorsten Schnurbusch (DFG Heisenberg Plant Architecture Group leader, IPK) for giving me the opportunity to join the group and work on this interesting topic. Without his advices, support, fruitful scientific discussions, and encouragement neither I nor my work could have develop on personal and professional level.

I would like to thank all group members from DFG Heisenberg Plant Architecture Group for creating a friendly and scientific environment and constant support and discussions specially Dr. Ahmad Alqudah, who I consider as my brother. My special thanks to office mates Dr. Ravi Koppolu, and Dr. Johannes Thiel for the useful scientific discussion and advices. I would like also to thank very much Dr. Britt Leps for all of her help related to administration stuff and visa as well as personal help when ever we need.

I sincerely and greatly acknowledge the technical assistance rendered by, Corinna Trautewig, Claudia Weissleder, and Mechthild Pürschel. My sincere thanks are also to the technical staff from the research group Genome Diversity. The technical help offered by the greenhouse staff is invaluable.

I especially thank my big family; my parents and brothers and sisters and I convey my heartfelt thanks to my small family; My dear wife Hend, my sons Abdallah, Abdelrahman and Mohamed and my daughter Rofaida for their support and encouragement during my study.

Finally I would like to share what I believe with all of you:

“If you miss an opportunity don’t fill the eyes with tears, it will hide another better opportunity in front of you”

Table of Contents

List of Tables:

List of Figures:

List of Supplementary Tables

List of Supplementary Figures

List of Abbreviations

ADLFM	Arrested Development of Lateral Floret Meristem
AFLP	Amplified Fragment Length Polymorphism
BAC	Bacterial Artificial Chromosome
BC	Before Crescent
BLAST	Basic Local Alignment Search Tool
BOPA	Barley Oligo Pool Assay
bp	base pair
BW–NIL	Bowman Near Isogenic Line
CAPS	Cleaved Amplified Polymorphic Sequence
cM	centi Morgan
CS	Central Spikelet
˚C	Celsius
DArT	Diversity Arrays Technology
DH	Double Haploids
DLS	Developed Lateral Spikelets
DLSFM	Developed Lateral Spikelet and Floral Meristems
DNA	De-oxyribo Nucleic Acid
DS	Developed Spikelets,
EST	Expressed Sequence Tag
F2	Second filial generation
FLcDNA	Full Length cDNA
FLS	Filled Lateral Spikelets,
Gbp	Gigabase
GWAS	Genome Wide Association Study
Hap	haplotype
HD–Zip	homeodomain–leucine zipper
HTP	high-throughput
IBSC	International Barley Sequencing Consortium
int–c	*intermedium–c* gene/locus
IPK	Leibniz Institute of Plant Genetics and Crop Plant Research
lab	*labile* gene/locus
LOB	Lateral Organ Boundary
LS	Lateral Spikelet
Mbp	megabase
MCFM	Missing Central Foret Meristem
MLFM	Missing Lateral Floret Meristem
mRNA	messenger RNA
MTP	Minimum Tiling Path
μg	Microgram
μl	Microlitre
NCBI	National Centre for Biotechnology Information
NGS	Next-Generation Sequencing
PacBio	Pacific Biosciences
PCR	Polymerase Chain Reaction
PLS	Potential Lateral Spikelet
POPSEQ	population sequencing
qRT–PCR	quantitative Real Time-Polymerase Chain Reaction
QTL	Quantitative Trait Locus
RAPD	Random Amplification of Polymorphic DNA

RFLPs	Restriction Fragment Length Polymorphisms
RILs	Recombinant Inbred Lines
RNAi	RNA interference
SbL	Sequencing by Ligation
SBS	Sequencing by Synthesis
SEM	Scanning Electron Microscopy
SMRT	Single-Molecule sequencing chemistry with Real Time detection
SNP	Single Nucleotide Polymorphism
ssp.	subspecies
SSRs	Microsatellites or Simple Sequence Repeats
Sup.	Supplementary
TAS	Targeted Amplicon re-Sequencing
TCP	transcription factors named after: **T**eosinte branched 1 (tb1), **C**ycloidea (cyc) and **P**CF
TF	Transcription Factor
TILLING	Targeting Induced Local Lesions IN Genomes
UCS	Unfilled Central Spikelets,
ULS	Unfilled Lateral Spikelets,
USDA	United States Department of Agriculture
vrs1	*six–rowed spike1* gene/locus
vrs2	*six–rowed spike2* gene/locus
vrs3	*six–rowed spike3* gene/locus
vrs4	*six–rowed spike4* gene/locus
WGS	Whole-Genome Shotgun Sequence
WT	Wild Type

1.0 CHAPTER ONE: General Introduction

1.1 Importance of barley

Cereal crops are members of the grass family (Poaceae) with a high socio-economic importance. Crops such as wheat (*Triticum spp*.), maize (*Zea mays L*.) and rice (*Oryza sativa* L.) form the main constituents of most dietary system in the world. Also barley (*Hordeum vulgare* L.) is still in the core of the dietary regimes of many developing countries in Central and West Asia, as well as Northern Africa (Grando and Macpherson 2005 and Newman and Newman 2006). Globally, however, only 5% of barley production is used as food, 75% is used as fodder while the remaining 20% finds its way in the industries of beverage brewing and digestion-enhancing health foods (Blake et al., 2011 and Baik and Ullrich 2008). The growing demand for barley by the food industry is mainly because of its health-promoting beta-glucan, acetylcholine, lysine, thiamine and riboflavin contents, and the easy digestibility (Marwat et al. 2012).

Adding to its economic importance, barley not only has a high yield potential under optimal circumstances, but also can be successfully cultivated under a wide range of climatic conditions and its easy to use in cross-breeding (Saisho and Takeda 2011). The self-pollinating barley has a short life cycle. It has a low chromosome number (2n = 14) and relatively large genome of more than 5 Gbp (Bennett and Smith 1976). The complete barley genome sequence was recently made available by the International Barley Sequencing Consortium (IBSC 2012). These criteria make barley an important crop model species for investigating the potentials of cereal breeding.

1.2 Barley classification, inflorescence and row-types

The Tribe Triticeae is one of the most successful branches of the grass family (Poaceae). The Triticeae comprise a large number of species and genera (von Bothmer et al 2003). Despite a worldwide distribution Triticeae have a center in temperate climates. The monophyletic Tribe Triticeae displays a significant biological diversity, with individual members showing various degrees of polyploidy, versatile forms of life, as well as multiple patterns of reproduction and dispersal (Dewey, 1984). All species of the tribe Triticeae are characterized by a special form of inflorescence called the spike in which the sessile spikelets are produced directly on the inflorescence axis (von Bothmer et al. 1985). The meristem differentiation and organization is comparatively less complex compared to the inflorescences of other grasses such as rice and maize.

Barley belongs to the genus *Hordeum* which comprises about 32 species (von Bothmer et al 1995) that share a similar set of diagnostic and morphological properties. Particularly, the genus is characterized by the presence of one central and two lateral single-flowered spikelets at each rachis node (von Bothmer et al. 1985). However, the barley spike shows an unusual meristem differentiation and organization. The inflorescence meristem differentiates into a double ridge meristem but then develops into a *Hordeum*-specific, determinate triple spikelet meristem. Further, each spikelet meristem produces just one floret meristem, resulting in three single-flowered spikelets per rachis node.

Differences in the fertility of the *Hordeum*-specific spikelet triplet confer a unique row–type identity to barley spikes. One way to classify barley therefore is based on whether the spike bears two, four or six rows of grains (Young 2001). Based on this, wild barley is considered as two-rowed, and cultivated forms either as two- or six-rowed. Another way to classify barley is to describe the awns covering the kernels (U.S. Grains Council, 2006). In the

barley germplasm database awns are described along the following morphology: long awned, short awned, awnless, hooded, elevated hooded, subjacent hooded, long awned in central row, and awnletted or awnless in lateral rows, short awned in central row, and awnletted or awnless in lateral rows, awnless or awnletted in central and lateral rows, elevated hoods in central row and awnless in lateral rows. Classification of barley cultivars started in the late 19th century with the German botanist Körnicke. In his systematic work on barley Körnicke described 44 botanical forms of barley using spikelet fertility, color, nature of the awn and glume, and the adherence or nonadherence of the palea (Körnicke,1885). More than half a century later, Mansfeld (1950) classified known cultivated barley into five groups: (i) two-rowed barley with fully fertile central spikelets but the two lateral spikelets sterile, (ii) the deficiens-barley, a two-rowed barley with no or extremely reduced lateral spikelets, (iii) six-rowed barley in which all three spikelets are fully fertile producing three seeds per rachis node, (iv) Intermedium barley which has a fully fertile central spikelet whereas the two lateral spikelets are enlarged or developed, and vary from rarely fertile to fertile or setting small seeds, and (v) labile-barleys which can have the lateral spikelet developed or absent, fertile or sterile even within one spike of the same genotype (**Fig 1-1** and **Table 1-1**).

Fig. 1-1: Phenotypes of barley spike row-types.

Table 1-1: Barley spike row-types and number of central and lateral fertile spikelets at each rachis internode.

Row-type	Central spikelet fertility	Lateral spikelet fertility	No of seeds/rachis node	Phenotype
Two-rowed	Fertile, developing into kernels	Sterile	1	
Six-rowed	Fertile, developing into kernels	Fertile and develop into kernels	3	
Deficiens	Fertile developing into kernels	Extremely reduced/absent	1	
Intermedium	Fertile developing into kernels	Fertile and develop into small kernels	1-3	
Labile **(Irregular)**	Fertile/Sterile may or may not develop into kernels or missing	fertile and/or sterile within one spike	0- 3	

1.3 Barley inflorescence development and its regulation

The inflorescence architecture of barley is unique among the species of Triticeae due to the presence of triple spikelet meristem that gives rise to three spikelets at each rachis node. Little is known about the genetic regulation of the spike meristem initiation, differentiation and its development during the spike growth. A major partition in barley germplasm is based on two– and six–rowed cultivars which partly reflects breeding history in the past (Cockram et al. 2010). It has been hypothesized that barley was domesticated from two–rowed types found in wild barley (*Hordeum vulgare ssp. spontaneum*) (Sakuma et al. 2011), as explained further on. The barley row–type is regulated by at least six known genetic loci that include *six–rowed spike1* (*vrs1*), *vrs2*, *vrs3*, *vrs4*, *intermedium–c* (*int–c*) and *labile* (**Table 1-2**). All

these six loci display clearly distinguishable spike phenotypes (**Fig. 1-2**). In cultivated barleys, the six–rowed spike types arose by loss–of–function mutations in *Vrs1*, which encodes a Class I homeodomain–leucine zipper (HD–Zip) transcription factor (HvHOX1) that functions as a negative regulator of lateral spikelet development (Komatsuda et al. 2007). The loss–of–function mutant, the *vrs1.a* allele, promotes lateral spikelet fertility resulting in a complete six–rowed phenotype. The wild type allele, *Vrs1.b*, functions as a negative regulator of lateral spikelet fertility and produces the two–rowed phenotype. Consistent with this, *Vrs1* is highly expressed in the two lateral spikelets. It has been proposed that *Vrs1* arose after duplication of a closely related HD–Zip I paralog, *HvHOX2*, which has a global expression pattern. Following this gene duplication, the *Vrs1* acquired a unique expression pattern in lateral spikelets thus suppressing their fertility (Sakuma et al. 2010). Mutant analysis revealed that in two–rowed barleys (*Vrs1.b* allele) the fertility of lateral spikelets can be modified by up to ten independent *Intermedium* loci distributed across the barley genome. When homozygous, these loci generate a partial or complete six–rowed phenotype (Lundqvist and Lundqvist 1988). The *int–c* locus has been associated with the natural quantitative variation for row–type (Lundqvist et al. 1997). The gene underlying *Int–c* encodes the barley ortholog of maize *TEOSINTE BRANCHED1*, *HvTB1* (Ramsay et al. 2011). Alleles at the *int–c* locus modify the fertility of the lateral spikelets with respect to the allelic constitution at *Vrs1*. The *int–c.b* allele is generally found in two–rowed barleys (*Vrs1.b*), where it inhibits anther development in lateral spikelets, while its presence in six–rowed barleys (*vrs1.a*) results in smaller lateral spikelets (Lundqvist et al. 1997). The *Int–c.a* allele is present in six–rowed barleys (*vrs1.a*), but in two-rowed barley the same allele can produce partially fertile lateral spikelets resulting in an intermediate phenotype between two– and six–rowed barleys termed *Intermedium* spike phenotype. The other barley row–type loci, are currently being studied by different research groups; *vrs2* and *lab* at (Plant Architecture Group, IPK-Gatersleben), and

vrs3 at (James Hutton Inst., Dundee, Scotland) for the underlying genes and their interaction with known row–type loci.

Table 1-2: Known barley row-type loci and chromosomal location.

Locus	Occurrence	Chr . Location	Study
vrs1	Natural alleles; Induced mutants	2HL	Komatsuda *et al.* 2007
vrs2	Induced mutant	5HL	Youssef *et al.* unpublished
vrs3	Induced mutant	1HS	Bull *et al.* unpublished
vrs4	Induced mutant	3HS	Koppolu *et al.* 2013
***int–c* (*vrs5*)**	Natural alleles; Induced mutants	4HS	Ramsay *et al.* 2011

Koppolu et al. (2013) identified the *Vrs4* gene as an ortholog of maize *RAMOSA2* which encodes a lateral organ boundaries (LOB) domain-containing transcriptional regulator. Interestingly, *Vrs4* also controls barley row-type because all of the induced *vrs4* mutants analyzed showed either a complete or partial six–rowed phenotype (Koppolu et al. 2013). Most importantly, the results also identified *Vrs4* as an important regulator of *Vrs1*. These results clearly suggest that not only *Vrs1* and *Int-c* but also *Vrs4* controls spikelet fertility in barley.

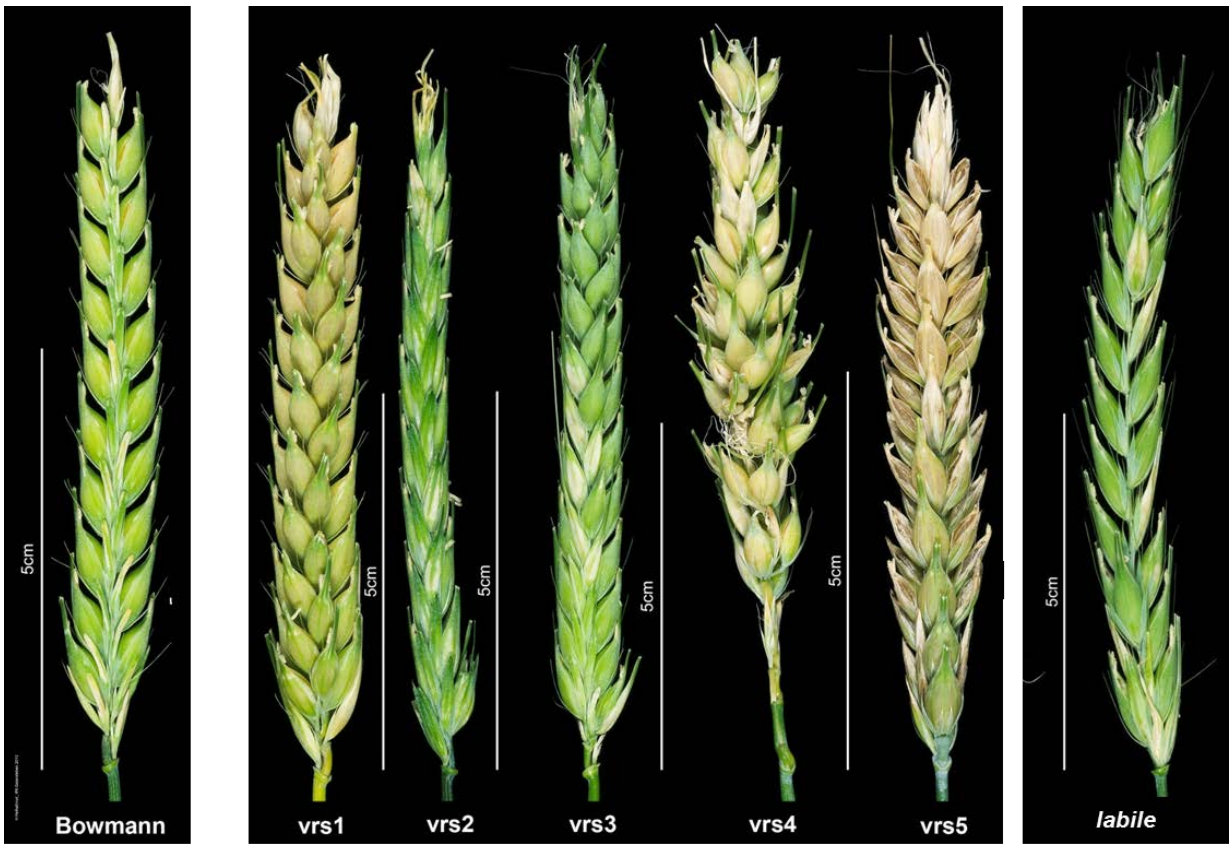


Fig. 1-2: Spike form of Bowman (*Vrs1.b*), Morex (*vrs1.a*), *vrs2*, *vrs3*, *vrs4*, *vrs5*, *lab* mutants

1.4 Ancestor and centers of barley domestication

The modification and development of crops in the last 12,000 years before present to adapt to the environment and meet human's needs is known as crop domestication. Studying centers of crop domestication as well as the domestication processes and period will help in understanding human civilization (Badr and El-Shazly 2012; Pourkheirandish and Komatsuda 2007). Modifications of wild plants happened mainly through selections made by ancient farmers. Barley domestication comprised many trait modifications related to spike and plant architecture as well as flowering time (Pourkheirandish and Komatsuda 2007). Three of these traits are considered the main traits of barley domestication; i.e. loss of seed shattering (non-brittle rachis), increased grain number per spike (six-rowed spike) and seed form and shape (naked caryopsis) (Salamini et al 2002). These three genetic changes are associated with the evolution of cultivated barley from wild barley.

The capability of barley to adapt to new environment enabled its distribution and spread into different geographic domestication areas (Jones et al 2008). For these reasons the origin of barley is still a matter of debate suggesting Egypt and Ethiopia in Africa, the Near East or even Tibet in the Far East as possible centers of origins (Duke 1983). Nonetheless, along with the domestication of wheat, barley certainly was one of the earliest grains to be cultivated. It is proposed that barley was already cultivated in the Middle East (ca. 10000 BC) prior to its arrival in China and India (Kling 2004). The German botanist Carl Koch (1809-1879) described the first ancestor of cultivated barley found in Turkey as a separate species, *H. spontaneum*. However, in respect to the biological species concept described in the late 20th century (von Bothmer *et al.*, 1995), this ancestor is regarded as subspecies [ssp. *spontaneum* (C. Koch) Thell.] within cultivated barley (ssp. *vulgare*) (Bothmer *et al.*, 2003).

It was long held that distribution and domestication of wild barley took place in the near East Fertile Crescent (Harlan and Zohary 1966; Nevo et al 1984; Nevo et al 1986; Badr et al 2000; Nevo 2006 and Zohary et al 2012). However, the appearance of *H. vulgare* ssp. *spontaneum* in Africa (Morocco, Libya, Egypt, Ethiopia), Central Asia and Tibet has changed the previous believe of a single origin theory of barley (Molina-Cano et al 2005; Azhaguvel and Komatsuda 2007; Morrell and Clegg 2007; von Bothmer and Komatsuda 2011 and Dai et al 2012). *H. vulgare* ssp. *spontaneum* was already spread in some west Asia and east-Mediterranean countries before it entered Central and East Asia (von Bothmer et al 1995 and Nevo 2012). According to Morrell and Clegg (2007) and Dai et al (2012) recent molecular evidence suggests two additional centers of wild barley domestication viz. Central Asia, 1,500–3,000 km farther east from the Fertile Crescent and the Tibet region in China.

The appearance of *H. agriocrithon* Åberg - which is considered the closest wild relative of barley - and *H. vulgare* ssp. *spontaneum* on the Tibet plateau in China was the reason for researchers recently to focus on the origin of Chinese cultivated barley and gave it

more attention (Xu 1985 and Yang and Yen 1985). The contrasting environmental conditions of the Near East (warm and dry) with those of the Tibet plateau (cold and dry) suggests the presence of different forms of wild barley in the fertile Crescent and Tibet. It thus appears that barley might have two different centers of domestication which nevertheless were somehow connected to the first center of domestication within the Fertile Crescent. In contrast to wheat and other Fertile Crescent founder crops, once domesticated the natural range of wild barley - the progenitor of cultivated barley - extended east into Central Asia to present day Kyrgyzstan, Afghanistan, and western Pakistan (Harlan and Zohary 1966) causing the second center of domestication in East Asia and Tibet. From these two central areas wild and cultivated barleys were distributed to other parts of the world (**Fig. 1-3**).

Fig. 1-3: Centers of barley domestication: the Fertile Crescent (left) and the Tibet plateau (right). The black arrows show the possible ways seeds were transported either between the two centers or to other places (Giles and Bothmer, 1985; Zohary and Hopf, 1993; Duke 1983).

1.5 Barley genomic resources

Mainly the continuous selection for plants optimally adapted to local conditions over the ages created a great variety of landraces which now form the basis for the improvement of modern cultivars suitable for different environmental conditions. In the last century, genetic approaches strongly enhanced barley grain yield (Schuster 1997). However, climate change

and a fast growing world population maintain an unremittingly demand for further yield improvements. To address these demands, new genomic and genetic technologies are expected to play a main role in identifying and isolating agronomically important genes controlling plant performance (Sreenivasulu et al. 2008).

Recent improvements on the level of chromosomal genetics such as chromosome physical map sequences, highly contributed to our understanding of the gene functions and their relation to agronomical important phenotypes in barley. In addition, genomics resources such as genetic and physical maps, Expressed Sequence Tags (EST), Full Length cDNA (FLcDNA), and DNA libraries have become widespread available. Both EST and FLcDNA sequences are important sources for deriving gene based markers as well as for the identification of gene expression patterns and are freely available at GenBank (http://www.ncbi.nlm.nih.gov/) and at (http://www.schigen.nig.sc.jp/barley/), respectively (Koppolu 2014).

Large insert genomic DNA libraries are a main source for gene discovery in map–based gene isolation projects (Schulte et al. 2011). Barley researchers have access to the data of nine Bacterial Artificial Chromosome (BAC) libraries. Six of these use cv. Morex as a source (Schulte et al. 2011; Yu et al. 2000), the remaining three are based on the cv. Haruna Nijo (Saisho et al. 2007), from the cv. Cebada Capa (Isidore et al. 2005), and the double haploid line CS134 (Shi et al. 2010). All genetic maps which form the basis for cloning studies to identify genes of interest are highly dependent depend on these resources (Poursarebani 2012). A whole-genome shotgun sequence (WGS) assembly and barley physical map of 4.98 Gb with more than 3.9 Gb anchored to the high resolution genetic map was available at the end of 2012 (IBSC 2012). The sequence data resulting from BAC ends and full sequencing, WGS, and shotgun sequencing of sorted chromosome arms helped in locating almost half of the physical contigs through integration of a genome-wide physical map providing more information about 26,159 high-confidence gene models (Ariyadasa

et al 2014). Recently, a population sequencing (POPSEQ) method developed for ordering WGS assemblies in the absence of highly developed genomic resources (Ariyadasa et al 2014). As an extension and improvement of the earlier published physical and genetic basis of barley by IBSC (2012), the sequence-ready minimum tiling path (MTP) of 66,772 minimally overlapping clones reported by Ariyadasa et al (2014) serve as a template for hierarchical clone-by-clone map-based shotgun sequencing. A web-based application called BARLEX has developed based upon requests from barley research community members to access the improved genomic database of barley (Colmsee et al 2015). BARLEX is focused on barley genome-wide physical map combined with sequence assemblies and genetic maps (IBSC, 2012 and Ariyadasa et al., 2014) and it is freely accessible at http://barlex.barleysequence.org.

1.6 Molecular markers systems

The molecular markers which are used in genetics and plant breeding can be classified into two different types. The first type is classical markers which comprise morphological, cytological and biochemical markers. The second class is DNA-based markers (Xu, 2010). Based on the way of detecting the polymorphism in the genome, different methods and techniques are used in this marker system such as; southern or northern blotting, PCR ‑ polymerase chain reaction, and DNA sequencing (Collard et al., 2005). Depending on which method is used in detecting the polymorphism, Mammadov et al (2012) noted that, molecular markers can be divided into three groups: I) low-throughput, hybridization-based markers i.e. restriction fragment length polymorphisms (RFLPs) Bernatsky and Tanksley (1986); II) medium-throughput, PCR-based markers such as random amplification of polymorphic DNA (RAPD) (Welsh and McClelland 1990), amplified fragment length polymorphism (AFLP) (Vos et al 1995), simple sequence repeats (SSRs) (Jacob et al (1991); III) high-throughput (HTP) sequence-based markers: single nucleotide polymorphisms SNPs (Wang et al 1998).

In the last 30 years, based on the molecular marker system development, the genetic mapping has moved from technology platform to another. The work of Bernatsky and Tanksley (1986) on tomato plants marked the start of using molecular markers in plant breeding. Restriction fragment length polymorphisms (RFLPs) marker system was the first to be use widely for genetic map construction. The RFLP marker system has some strength such as; codominant inheritance, and good transferability between labs in addition to locus-specificity but it has also weaknesses such as; high quantity and quality DNA required, low polymorphism and time consuming as well as it is considered hazardous (requires radioactive detection) (Semagn et al 2006). Microsatellites (SSR) and inter simple sequence repeat were the next marker systems to be used in genetic mapping in beginning of 1990s and it was the main marker system for genetic mapping at that time (Powell et al., 1996). Whole Genomic coverage with high level of polymorphism, avoiding using radioactive detection and small amount of DNA required as well as co-dominant inheritance, all considered as advantages of these systems. But SSR was limited only to use in some of the important crops because of the high cost and specificity of the primers (Squirrell et al 2003; Jian 2013). While co-migration and non-reproducibility are considered the limitations for ISSR system (Moreno et al 1998). Random amplified polymorphic DNA (RAPD) and amplified fragment length polymorphism (AFLP) are marker systems used also in linkage mapping but they have not found to be widely accepted. ISSR, RAPD and AFLP share disadvantages which causing their limitation of use for mapping F2 populations. In all of them high quality of DNA is required in addition to incapability to detect genomic differences in heterozygotes individuals. Recently, SNP became the preferable marker system in plant genetics (Batley and Edwards, 2007) because of the high frequency of SNPs in eukaryotic genomes and compared to simple sequence repeat (SSR) it is display a minor rate of mutation and thus are not difficult in population genetic analyzes (Xu et al., 2005). The SNP densities are found to be different based on the species and tested cultivars numbers as well as the tested DNA region (coding or noncoding)

(Ossowski et al., 2008; Russell et al., 2004). In maize, as an example, SNPs occur in a density of 1 per 124 bp and per 31 bp in coding and noncoding regions, respectively (Ching et al., 2002). In barley, based on re-sequencing 877 unigenes fragments from eight diverse accessions, Rostoks et al (2005) found 1 SNP per 200 bp as an average, however, the SNP frequency was extremely variable within European cultivated germ-plasm showing the selection effect during the breeding programs. Technically, numerous numbers of methods such as expressed sequence tag (EST) (Barbazuk et al., 2007) and targeted amplicon re-sequencing (TAS) (Bundock et al., 2009) described for the early detection of SNPs in plant genomes. Especially for the barley researchers community, a very rich resource of SNP markers is available in the form of BOPA SNP markers (Close et al. 2009) and barley iSELECT chip (Comadran et al. 2012), for most of these SNPs chromosomal positions were already assigned. Illumina's Goldengate assay provided SNP genotyping for genome-wide marker profiling. This technology contains GoldenGate BeadArray, GoldenGate VeraCode and GoldenGate Indexing. GoldenGate VeraCode was considered one of the most efficient and plastic platforms for SNP genotyping in plant breeding. In this technology the users can combine 48, 96, 144, 192 and 384 SNP within a single well of a standard microplate. This technology used for different applications in plant breeding such association mapping, linkage mapping and crop diversity analysis (Thompson et al 2011).

1.7 Sequencing technologies and their uses

DNA sequencing methods without which modern molecular biology cannot be imaged, have developed over a span of just 40 years. The start was inaugurated in 1973 by Gilbert and Maxam. Using a method known as wandering- spot analysis they published the very first DNA sequence measuring 24 base pairs. Sanger DNA sequencing with chain-terminating inhibitors was first obtained in 1977 and since that time Sanger's sequencing method has remained one of the most powerful innovations in genetics research. About 20 years ago the arsenal of sequencing methods was augmented with the bioluminescence

sequencing-by-synthesis approach (Nyren 2007). The never-ending demand for ever faster and cheaper sequencing systems lead to the development of high- throughput sequencing technologies (Cutler et al 2001). The parallelized version of Pyrosequencing by 454 Life Sciences produces about five hundred million bases of raw sequence an hour (Margulies et al 2005). Sanger sequencing in its current status cannot match this throughput sequencing technique. However, in 2008, Illumina and Applied Biosystems introduced the Illumina (Solexa) sequencing technology based on sequencing by synthesis (SBS) producing a higher throughput than the 454 system and generating billions of bases in one run (Pettersson et al 2009; Llaca 2012). This Illumina HiSeq 2000 sequencer platform is at the moment the device of choice for Next-Generation sequencing (NGS) system in plant breeding and biotechnology. While compared to the Sanger method, the parallelized version of 454 could sequence thousands of bases per second (Pettersson et al 2009), the Illumina HiSeq 2000 sequencer can finish sequencing 540-600 Gbp in a single 2-flow cell within 8.5-day at a cost of about 2 cents per Mbp (http://www.illumina.com/systems/hiseq_2000.ilmn). A similar high throughput at similar cost is The Support Oligonucleotide Ligation Detection (SOLiD) system which is based on Sequencing by Ligation (SbL) chemistry and is produced by Life Technologies (Llaca 2012). One on the NGS platforms also produced by Life Technologies, the Ion Torrent (Rothberg et al, 2011; http://www.iontorrent.com). This technology has not been as much as 454 and Illumina widely used in plant breeding and biotechnology. Ion Torrent differs from the other NGS in that it ios based on pH and does not require fluorescence. One major disadvantage of the IonTorrent is the low throughput, though it is considered the fastest turnaround times among the available NGS systems (Rothberg et al, 2011). An assessment of the Sanger, SOLiD, 454, HiSeq2000 and Ion Torrent sequence systems can be found in table 3.

Table 1-3: Comparison between Sanger, SOLiD, 454, HiSeq2000 and Ion Torrent sequence technology (developed from Llaca 2012).

	Sanger	**Roche 454**	**SOLiD**	**HiSeq2000**	**Ion Torrent**
Sequence-based	Dideoxy	Pyrosequencing	Sequencing by ligation	Sequencing by synthesis	pH
Detection	Fluorescence	Fluorescence	Fluorescence	Fluorescence	pH
Run time (hours)	~ 2	12	192	192	~ 2
Max. TP/Run (Mbp)	0.08	800	310000	600000	100
Max. Reads/Run (Million)	0.000096	1	5167	3000	1
Cost/Mbp (Euros)	3600	7.2	0.04	0.02	9

Whole-transcriptome sequencing helps in the analysis of gene expression through the sequencing of DNA extracted by antibodies targeting DNA-binding proteins (ChIP-Seq) (Barski et al (2007); Mikkelsen et al (2007)). Recently, Mayer et al. (2011) developed the barley genome zipper using the gene-based barley 454 sequence reads. These 454 reads, which were aligned in synteny with the genes of *Brachypodium*, rice, and sorghum, were positioned in between the gene-based SNP markers reported in the consensus map developed by (Close et al. 2009b). As a result, a total of 27,581 syntenic genes of barley with rice, *Brachypodium* and sorghum were located onto this genetic map, indicating a linear gene order on the seven barley chromosomes. These genomic data are an excellent source for gene mapping, marker development, gene identification and cloning.

By mid of 2011 the first commercially 3rd-Generation sequencing system, called the Pacific Biosciences (PacBio) RS platform, became available. This function of this system is mainly based on a Single-Molecule sequencing chemistry with Real Time detection (SMRT).

In the sequencing cell, DNA polymerases arte attached to nanowells and exposed to single molecule templates and labeled NTPs. No terminators are used in this system, although the polymerization takes time to be detected by a charge coupled device (CCD) camera (Llaca 2012).

1.8 Genetic (Linkage) mapping

Bateson W. and Punnett R. in 1904 were the first to discover the genetic linkage between two flower color genes in peas (Edwards 2012). However, about 9 years later the understanding of genetic linkage developed by Thomas Morgan and his student Alfred Sturtevant when they studied the inheritance of characters in the common fruit fly, *Drosophila melanogaster* (Smith 2013). The importance of genetic markers is directly related to the known location on the genome. In this regard, genetic markers can be used in either physical maps or genetic maps (White et al 2007). Genetic mapping, which is extensively used in domesticated plants (Hall et al. 2010), is applied to find the exact location of a gene that controls a specific phenotype. Selecting mapping populations is the main step for genetic mapping. Nowadays, several numbers and types of mapping populations (second filial generation (F2), double haploids (DHs), recombinant inbred lines (RILs)), marker systems as well as statistical methods used for stablishing linkage maps in plant breeding (Ferreira et al., 2006). The latter found that the minimum number of individuals to get precise maps from all different types of populations is 200 individuals and preferentially more. Once a mapping population is selected, one needs a marker system to locate the polymorphic data from the population. Following, it is necessary to select statistical methods for linkage analysis to locate the targeted marker/phenotypes in the right position on the chromosome (White et al 2007). This last step in genetic linkage mapping involves software packages such as Join-Map (Stam, 1993a) which is a commercial program, or free programs such as LINKAGE (Suiter et al., 1983), MAPMAKER/EXP (Lander et al., 1987), GMENDEL (Echt et al., 1992), and Map

Manager QTX (Manly et al., 2001). The generation of high-density genetic maps using markers such as SNP has greatly improved the ability to identify genes or QTLs (Stein et al. *2007*).

1.9 Organization and aims of the dissertation

The inflorescence architecture of barley (*Hordeum vulgare* L.) is unique among the Triticeae, which also include wheat, rye and triticale. Barley row-types include two-rowed, six-rowed, *deficiens*, *labile* and Intermedium. To better understand the importance and role of the genes controlling the spike architecture and come up with a clear relation between the cultivated barley and its ancestor the main goal of the present study was to unravel the genetic basis controlling variation in *labile* and Intermedium barley spike architecture using natural collections.

The present dissertation comprises five chapters. After a general introduction to the importance, the domestication, the genetics and genomic resources of barley, the second and third chapters summarize the results of two published articles. The second chapter deals with the re-sequencing of *vrs1* and *int-c* loci in labile barley collection. The aim was to investigate whether *labile*-barleys have a two-rowed genetic background, resulting in increased lateral spikelet fertility, and whether they show reduced lateral spikelet fertility in the presence of a six-rowed genetic background. In the third chapter two mapping populations are described from crosses between labile and six-rowed barley. Using available SNP resources a whole genome genetic linkage map is created to genetically locate the *lab* locus onto a barley chromosome arm. The fourth chapter focusses on the Intermedium row-type and its putative role in barley evolution and domestication. By re-sequencing the *vrs1*, *int-c* and *vrs4* loci in the Intermedium collection new light is shed in the relationship between two- and six-rowed cultivated barley and its ancestors during the domestication process. The fifth chapter

provides a general discussion on how the results presented here have increased our understanding of the genetic background of labile and Intermedium barley row-types, concluding with outlook of future work.

CHAPTER TWO

Genetic Resources and Crop Evolution

October 2012, Volume 59, Issue 7, pp 1319-1328 **(DOI:** 10.1007/s10722-011-9759-5)

Re-sequencing of *vrs1* and *int-c* loci shows that *labile* barleys (*Hordeum vulgare* convar. *labile*) have a six-rowed genetic background

Helmy M Youssef [1,2]; Ravi Koppolu[1] and Thorsten Schnurbusch[1]

[1]Leibniz Institute of Plant Genetics and Crop Plant Research (IPK), Genebank Department, Research Group Plant Architecture
06466 Gatersleben, Germany.

[2] Plant Physiology, Faculty of Agriculture, Cairo University, 12613 Giza, Egypt

2.1 Abstract

Labile-barleys (*Hordeum vulgare* convar. *labile* (Schiem.) Mansf.) are found in the highlands of Ethiopia, Eretria and North India-Pakistan districts. They represent a distinct spike form showing row-type alterations even within individual spikes of the same genotypes. Variation at the *six-rowed spike 1* (*vrs1*) locus is sufficient to control barley lateral spikelet fertility, which is also modified by alleles at the *Intermedium -c* (*int-c*) locus. This study aimed at re-sequencing these two loci to investigate whether *labile*-barleys have a two-rowed genetic background, resulting in increased lateral spikelet fertility, or show reduced lateral spikelet fertility if they possess a six-rowed genetic background. The *Vrs1* re-sequencing results of 221 supposedly *labile*-barley accessions from Ethiopia revealed 13 accessions with two novel *vrs1.a1* haplotypes. Following the current nomenclature of *vrs1* haplotypes, the new haplotypes were named as haplotypes 66 and 67. Re-sequencing at the *int-c* locus showed that 118 of the *labile*-barleys possessed the previously described *Int-c.a* allele but only one accession was found having a novel *Int-c.a* haplotype in the homozygous state (termed *Int-c.a haplotype1*; *Hap_1*). Interestingly, 101 *labile*-barleys carried the *Int-c.a* allele and *Int-c.a haplotype1* simultaneously, suggesting maintained heterozygosity or recent gene duplication at this locus. Only one accession had a two-rowed haplotype (*Vrs1.b3*, *int-c.b1*) and one accession possessed the *Vrs1.t* (*deficiens*) and *Int-c.a* alleles (six-rowed). These two accessions were considered as misclassified *labile* genotypes and not included in further analysis. Thus, these results confirmed that all of the 219 *labile* accessions studied in this work showed six-rowed alleles at *vrs1* but reduced lateral spikelet fertility. This reduction is most likely caused by the recessive *labile* (*lab*) locus which we are in the process to characterize further.

Keywords *Hordeum vulgare* convar. *labile int-c Labile*-barleys Lateral spikelet fertility Novel haplotype *vrs1*

2.2 Introduction

The inflorescence architecture of barley (*Hordeum vulgare* L.) is unique among Triticeae family members, which also include wheat, rye and triticale. The barley spike is characterized by a triple spikelet meristem (one central and two lateral spikelets) at each rachis node which bears three one flowered spikelets (von Bothmer et al. *1985*). Based on the central and lateral spikelet fertility, barley could be classified into four different groups: (i) two-rowed barley which has fully fertile central spikelets but the two lateral spikelets are sterile and produce only one seed per rachis node, (ii) the *deficiens*-barley includes two-rowed barleys from Ethiopia which have no or extremely reduced lateral spikelets and it produces one seed per rachis node, (iii) six-rowed barley which has three fully fertile spikelets and produces three seeds per rachis node, and (iv)*labile*-barleys which can have the lateral spikelets developed or absent, fertile or sterile even within one spike of the same genotype (**Fig. 2-1**).

Labile-barley (*Hordeum vulgare* L. convar. *labile* (Schiem.) Mansf.), previously known as '*irregular*' barley, was found among barleys originating from the highlands of Ethiopia, Eritrea (Åberg and Wiebe *1945*) and also the North India-Pakistan districts (Takeda and Saito *1988*). The *labile* row-type has been considered as a distinct spike character especially among Ethiopian barleys (Bjørnstad and Abay *2010*) and 'irregular spike'-forms have been growing in most of the barley cropping areas throughout Northern Ethiopia (Abay and Bjørnstad *2009*; Hadado et al. *2009*). Recent phenotypic and molecular evidence linked the occurrences of *labile*-barleys mainly to higher altitudes above 2,800 m a.s.l. (subpopulation T6; Hadado et al. *2010*). However, among all the naturally occurring row-type variants in barley, the *labile*-barleys are genetically least described probably owing to their high phenotypic plasticity, which caused difficulties while classifying them (Mansfeld *1950*).

Fig. 2-1: Barley spikes showing different row-types; (**a**) two-rowed barley, (**b**) six-rowed barley, (**c**) *deficiens*, (**d**) spikelets from *top* to *bottom*: two-rowed, six-rowed and *deficiens* and (**e**) different spike forms of *labile*-barleys showing increased lateral fertility from *left* to *right* (Awns were clipped off for clarity)

The *labile*-phenotype can vary from spike to spike even within a single plant; for example one spike can be completely reduced to a *deficiens* phenotype with other spikes expressing various degrees of lateral spikelet fertility within an individual plant. There is a continuous variation

in the number of fertile lateral spikelets from genotype to genotype (Djalali et al. *1970*). It has been suggested that two genetic factors are necessary for the manifestation of the *labile* character: (i) the recessive allele for the six-rowed phenotype (*vrs1.a*), and (ii) the recessive allele at the *lab* gene for the *labile* character (Djalali *1970*). The two-rowed allele at *Vrs1* appeared to be epistatic to the *lab* gene and the *lab* gene contributes to variable expression in lateral spikelets (Djalali et al. *1970*). The F_1 and F_2 plants arising from crosses between *labile* genotypes were also found to show the *labile* character, indicating the complete penetrance of the *labile* phenotype. However, lateral spikelets of F_1 plants from *labile* and two-rowed barley crosses were completely developed but sterile. F_2 individuals segregated for two-rowed, six-rowed and *labile* phenotypes (Djalali*1970*).

Understanding the developmental genetics of the barley inflorescence such as spikelet initiation, abortion and fertility has started to emerge relatively recently. Komatsuda et al. (*2007*) provided the first step in the elucidation of lateral spikelet fertility in barley and showed that loss-of-function of the wild-type *Vrs1* gene (responsible for two-rowed phenotype) resulted in complete fertility of lateral spikelets displaying the six-rowed phenotype. The *Vrs1* gene belongs to the HD-ZIP I class of homeobox transcription factors. Lack of the VRS1 protein in lateral spikelet primordia enabled complete fertility, suggesting that VRS1 protein suppresses the development of lateral spikelets in barley. Previous and recent studies revealed that different alleles at the *vrs1* locus are responsible for the size and fertility of lateral spikelets, for example barleys classified as convar. *hexastichon* (L.) Alef. possess the *vrs1.a* allele, that of convar. *deficiens* (Steud.) Mansf. carry the *Vrs1.t* allele, convar. *distichon* (L.) Alef. have the *Vrs1.b* allele and some barleys belonging to the convar. *Intermedium* display the *int-d* allele (Komatsuda et al. *2007*; Lundqvist and Lundqvist *1989*).

Lundqvist and Lundqvist (*1988*) showed that the phenotypic effect of *Vrs1.b* can be influenced by ten independent *Intermedium* (*int*) genes distributed all over the barley genome. In addition to this, natural quantitative variation in the size and fertility of the lateral spikelets has also been observed, particularly in progenies of two-rowed by six-rowed crosses (Lundqvist and Lundqvist *1989*). Genetic studies indicated that this quantitative variation is largely due to the effect of alleles at the *int-c* locus. Alleles at the *int-c* locus either complement or repress the fertility of lateral spikelets based on the allelic status at the *Vrs1* gene. For example in six-rowed barley, the loss-of-function *vrs1.a* allele is generally complemented by the *Int-c.a* allele, and in two-rowed barley the *Vrs1.b* allele is complemented by the *int-c.b* allele (Lundqvist et al. *1997*). Recently Ramsay et al. (*2011*) identified *int-c* as an orthologue of the maize (*Zea mays* L.) domestication gene, *Teosinte branched 1* (*ZmTB1*). *ZmTB1* is mainly involved in the control of axillary organ growth and also in female inflorescence development in maize (Doebley et al. *1997*) rather than inflorescence architecture. It was observed that the induced mutant allele *int-c.5* significantly increases the tiller number during the juvenile stages (Ramsay et al. *2011*). However, it is presumed that tillering mediated by *Int-c.a* in six-rowed cultivars is under the masking effect of the reduction in tiller number associated with six-rowed alleles at the *Vrs1* gene (Kirby and Riggs *1978*). Apart from the naturally occurring six-rowed mutants *vrs1*and *int-c*, there are three induced mutants *vrs2*, *vrs3*, *vrs4* which can individually convert two-rowed to six-rowed barley (Druka et al. *2011*).

Until today, almost all cultivated six-rowed barleys are known to carry the recessive (loss-of-function) *vrs1.a* allele (Komatsuda et al. *2007*) and the alternative *Int-c.a* allele (Ramsay et al. *2011*), enabling complete lateral spikelet fertility. In contrast, two-rowed barleys, carrying the functional *Vrs1.b* allele (Komatsuda et al.*2007*), and *int-c.b* allele (Ramsay et al. *2011*), develop always infertile lateral spikelets, and therefore, produce only

one fertile central spikelet. Previous genetic studies including *labile*-barleys found contradicting results, suggesting that *labile*-barleys were either derived from two-rowed (Breitenfeld *1957*) or six-rowed barleys (Nötzel *1952*). In an attempt to reveal the haplotype structure at the *vrs1* locus in *labile*-barleys, a set of 14 *labile* accessions had been analyzed and all carried the *vrs1.a* allele, (Saisho et al. *2009*). However, the sample size used in the Japanese study was very limited and may not provide a representative and sufficient coverage of the available haplotypes. In order to get a better understanding in which genotypic background the *labile* character is most reliably detectable we initiated the present study to determine the genotypic status of 221 Ethiopian barleys at the two known row-type genes, *Vrs1* and *Int-c*.

2.3 Materials and methods

2.3.1 Plant materials

221 Ethiopian barley accessions categorized as *labile* and maintained at the IPK Gene bank, Gatersleben, Germany, were selected for the present study (**Supplementary table 2-1**). Two two-rowed (Barke and Ametyst) and two six-rowed (Morex and Streptoe) spring barley cultivars were grown alongside with all *labile* accessions as controls.

2.3.2 Growing conditions and spike phenotyping

The expression of the *labile* character was found to be influenced by day length. Djalali (*1970*) found that, a period of short day (12 h light and 12 h dark) treatment for 20–30 days after seed germination resulted in reduced lateral fertility. It was also noted that the expression of the *labile* character is more pronounced under continuous short days. To account for this observation, *labile* accessions and controls were grown under 12 h/12 h (day/night) light conditions and a temperature of ~14°C during the day and ~12°C during the

night. After anthesis plants were scored for visible phenotypes such as Filled Lateral Spikelets (FLS) with developed kernels, Unfilled Lateral Spikelets (ULS) with enlarged lemma and palea with awns, but without developed kernels, Developed Lateral Spikelets (DLS) comprising both FLS and ULS, Potential Spikelets (PS = Number of rachis nodes/spike (in case of expected two-rowed or *deficiens* accessions) and number of rachis nodes/spike* 3 (in case of expected six-rowed accessions)), Potential Lateral Spikelets (PLS = number of rachis nodes/spike* 2 (in case of expected six-rowed accessions) and zero (in case of expected two-rowed or *deficiens* accessions)), Unfilled Central Spikelets (UCS) and Developed Spikelets (DS) which comprise of filled, unfilled central spikelets and developed lateral spikelets; all calculations were performed on the total number of heads averaging 5 heads on a single plant.

2.3.3 Genomic DNA isolation

Leaf samples were collected from single plants (at 3–5 leaf stage) of each accession for DNA extraction. The total genomic DNA was extracted using the Doyle and Doyle (1990) method. DNA quality and quantity were checked on 0.8% agarose gels and the concentration was adjusted to ~20 ng/µl for PCR.

2.3.4 PCR amplification, sequencing and sequence analysis

Three primer pairs were designed to cover the 2,062 bp fragment at the *vrs1* locus in order to obtain sequence data for the whole gene by Sanger sequencing. The three primer pairs include *Vrs1-1*F (5′-TATCTAGAGGAACTCGATGAACTTGAG-3′), *Vrs1-1*R (5′-GTACCATTGGCCGCGAA-3′) covering promoter and 5′ untranslated region (5′ UTR), *Vrs1-2*F (5′-ACACCAACAGGCAACAGAACAACCTA-3′), *Vrs1-2*R (5′-GGACGCACATCATCAGGTCATCGT-3′), covering exon1, exon2, exon3 and *Vrs1-3*F (5′-CAAACATATGGCCAGCTGCT-3′), *Vrs1-3*R (5′-TGATCTTCAAGAGAGCTGCCA-3′)

covering the 3′ UTR. For the *HvTB1* gene a single primer pair was designed to amplify a 1,074 bp fragment. The primer pair for this locus include *HvTB1*F (5′-TCCTTTCTATGATTCCCCAAGCCCC-3′) and *HvTB1*R (5′-CCACTCCACCGAGCTCCC-3′). PCR amplifications with individual primer pairs for *Vrs1* and *Int-c* were performed in all 221 barley accessions.

PCR amplifications were carried out in a 25 µl reaction volume containing 20 ng of DNA, 2.5 µl of PCR buffer (10X) (Qiagen, Hilden, Germany), 5 mM dNTPs, 5 µl of Q-solution (Qiagen, Hilden, Germany), 5 pM primers, and 1U of *Taq* polymerase (Qiagen, Hilden, Germany). The PCRs were conducted using a thermal cycler (SensoQuest Thermal Cycler, USA) and the touchdown PCR amplification profile has an initial denaturation step for 3 min at 94°C followed first by 8 cycles of 94°C for 40 s, 61°C for 40 s (for *vrs1-1* and *vrs1-3*) or 65°C (for *Vrs1-2* and *HvTB1*) and 72°C for 2 min, with 1°C decrement in temperature per each cycle, then followed by 45 cycles of 94°C for 40 s with constant annealing temperatures (55°C- for *vrs1-1* and *vrs1-3* or 60°C for *Vrs1-2* and *HvTB1*) for 40 s and 72°C for 2 min, followed by a final extension for 10 min at 72°C. The PCR products were tested on 1.2% agarose gels to check the amplification.

For direct-sequencing of PCR products, the PCR products were first purified using MinElute 96UF PCR purification kit (Qiagen, Hilden, Germany), then sequenced using BigDye Terminator v3.1 cycle sequencing Kits (Applied Biosystems, USA). DNA sequence analysis, quality score assignments and the construction of contigs were achieved using Sequencher 4.7 DNA sequence assembly software. Multiple sequence alignments were carried out using ClustalW2 (http://www.ebi.ac.uk/Tools/msa/clustalw2/). The *Vrs1.b2*, *Vrs1.b3*, *Vrs1.t*, *vrs1.a1*, *vrs1.a2* and *vrs1.a3* alleles (Saisho et al. *2009*) for the *vrs1* locus and *Int-c.a*, *int-c.b1* and *int-c.b2* alleles (Ramsay et al. *2011*) for the *int-c* locus were considered as the reference alleles.

2.4 Results

2.4.1 Genotypic status of the *vrs1* locus in *labile*-barleys

A total of 221 Ethiopian barley accessions classified as *labile* (**Sup. table 2-1**) were sequenced at the *vrs1* locus. The sequence analysis revealed that 217 accessions had the *vrs1.a1* allele which is responsible for the six-rowed phenotype (**Table 2-1**). Meanwhile, two accessions (HOR3529 and HOR3587) carried the *vrs1.a3* allele, one accession possessed the two-rowed allele *Vrs1.b3* (HOR5281) and one accession (HOR5471) had the *deficiens* allele *Vrs1.t*. Since *Vrs1* seems to be epistatic to the *labile* locus, these two accessions possessing two-rowed alleles (HOR5281 & HOR5471) can be considered as misclassified *labile* genotypes, and hence, were omitted from further analyses, thus reducing the effective number of *labile* accessions to 219.

Table 2-1: *Vrs1* and *Int-c* alleles and haplotypes in all studied 221 Ethiopian barley accessions.

		Int-c alleles/haplotypes				
		Int-c.a	*Int-c.a/Hap_1*	*Int-c.a/Hap_2*	*int-c.b1*	Total
Vrs1 alleles	*vrs1.a1*	115	1	101		217
	vrs1.a3	2	-	-	-	2
	Vrs1.b3	-	-	-	1	1
	Vrs1.t	1	-	-	-	1
	Total	118	1	101	1	221

*Two-rowed and *deficience* barleys were misclassified as *labile*-barleys (confirmed by re-sequencing of *vrs1*)

Interestingly, 13 of the 217 accessions carrying the *vrs1.a1* allele showed two novel haplotypes either in the promoter region or in the highly conserved HD domain. Among the 13 accessions nine of them (HOR6178, HOR6179, HOR6180, HOR6279, HOR7729, HOR7734, HOR9405, HOR10421 and HOR10490) had a unique single nucleotide polymorphism (SNP) within the promoter region at 230 base pair (bp) (**Fig. 2-2**). The

sequence carrying an SNP at 230 bp in different accessions has been named as *haplotype 66* of *vrs1.a1* (*Hap_66*) in continuation to the previously named haplotypes of *Vrs1* (Saisho et al. *2009*). The remaining four of the 13 accessions, HOR5172, HOR6400, HOR6420 and HOR6440, showed an SNP at bp position 987 (**Fig. 2-2**) within the conserved HD domain that resulted in an amino acid substitution from alanine to aspartic acid (**supplementary Figure 2-1**). Thus, the novel haplotype identified by the amino acid substitution in the HD domain has been named as *haplotype 67* of *vrs1.a1* (*Hap_67*) (**Fig. 2-2**). Sequence data of these two novel haplotypes (*Hap_66 and Hap_67*) were submitted to the National Center for Biotechnology Information (NCBI) and is available under the following accession numbers JF904736 and JF904737, respectively. In summary, and in accordance with Saisho et al. (*2009*), all the *labile* accessions (100%) showed six-rowed alleles at the *vrs1* locus.

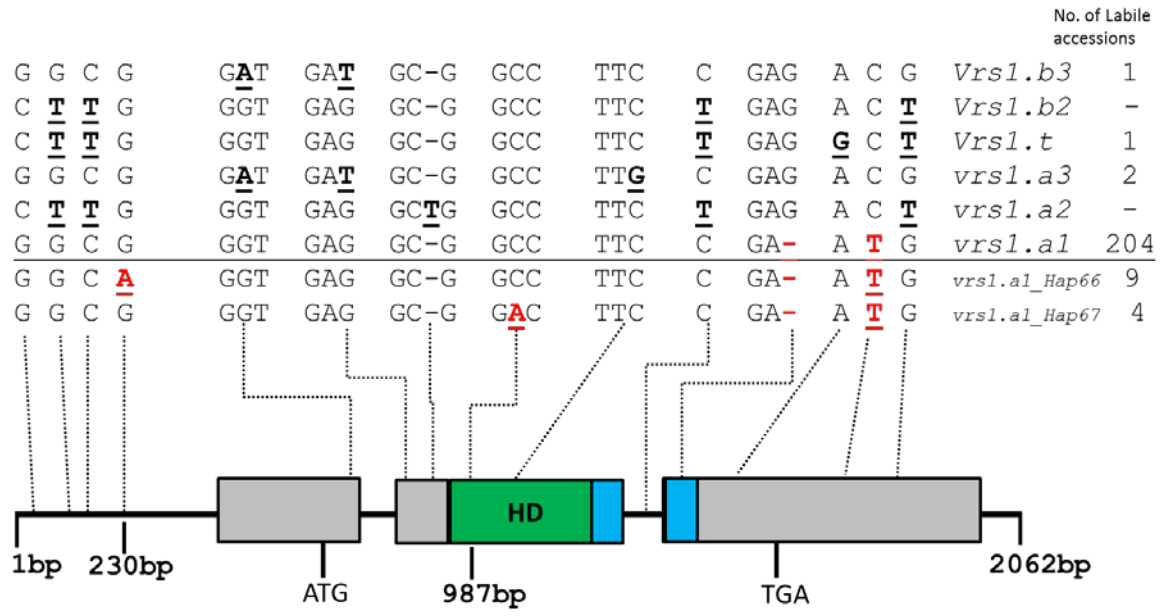


Fig. 2-2: Alleles and haplotypes at the *vrs1* locus in two- and six-rowed cultivars as well as *labile*-barley accessions.*Vrs1.b3* (two-rowed allele) was identified in one accession, *Vrs1.t.* (*deficiens* allele) in one accession,*vrs1.a3* (six-rowed allele) in two accessions, *vrs1.a1* (six-rowed allele) in 217 accessions were identified. Three overlapping fragments were aligned to cover the *vrs1* locus, spanning in total 2,062 bp

2.4.2 Genotypic status of the *int-c* locus in *labile* barleys

Sequencing of the *int-c* locus in 221 Ethiopian barley accessions (**supplementary Table 2-1**) revealed that 220 of them carried dominant, functional *Int-c* alleles and only one accession possessed the recessive two-rowed *int-c.b1* allele (HOR5281). Of the 220 accessions having dominant *Int-c* alleles, 118 showed the previously described and lateral fertility promoting *Int-c.a* allele (**Table 2-1**). Apart from this, two interesting sequence variations at this locus were observed in the remaining accessions; one of them included a non-synonymous nucleotide change at position 437 bp from the start codon within the TCP domain of *HvTB1*that lead to an amino acid substitution from aspartic acid to valine (**supplementary Figure 2-2**). This kind of sequence variation within the previously described dominant allele of *Int-c.a* was observed in one *labile* accession, HOR1643, and hence has been denoted as a novel *Int-c.a haplotype1* (*Hap_1*) (**Fig. 2-3**). The GenBank accession number of the novel *Int-c.a haplotype1* is JF904738. The second sequence variant showed a highly consistent, putatively heterozygous position at the *Int-c.a* allele, where the characteristic SNP of the *Int-c.a haplotype1* (*Hap_1*) (SNP "T" at 437 bp instead of "A" in *Int-c.a1*) occurred simultaneously. This feature was observed in 101 *labile* accessions and has been denoted as *Int-c.ahaplotype2* (*Int-c.a/Int-c.a haplotype1*) (*Hap_2*). Sequence analysis of the *labile* accessions also revealed that all of them (100%) carried the lateral fertility promoting alleles at the *int-c* locus.

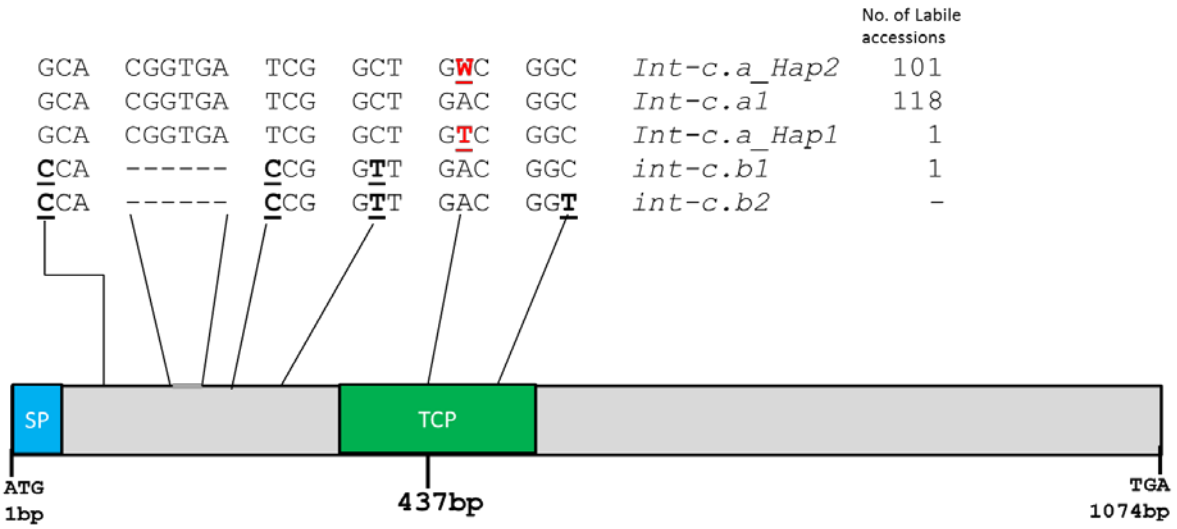


Fig. 2-3: Allele and haplotype analysis of the *int-c* locus in two- and six-rowed cultivars as well as *labile*-barley accessions. *Int-c.a1* (promoting lateral spikelets) in 118 accessions, *Int-c.a2* (promoting lateral spikelets) in one accession, *Int-c.a1/.a2* (both promoting lateral spikelets) in 101 accessions, and i*nt-c.b1* (impairing lateral spikelets) in one accession were identified. Following the International Union of Pure and Applied Chemistry (IUPAC) notation for degenerate base symbols the "W" at bp position 437 represents the two nucleotides A and T. Re-sequencing comprised one PCR fragment of 1,074 bp in length.

2.4.3 *vrs1* and *int-c* allele/haplotype combinations in *labile* accessions

In order to have a better understanding of the row-type status in *labile*-barleys, the haplotypes at *vrs1* and *int-c* loci were compared with that of normal two-rowed (*Vrs1.b*, *int-c.b*) and six-rowed (*vrs1.a*, *Int-c.a*) haplotypes. As shown in **Table 2-1**, of the 221 barley accessions studied, 115 accessions had the *vrs1.a1*, *Int-c.a* haplotypes (six-rowed alleles), 101 accessions carried the *vrs1.a1*, *Hap_2* haplotypes (six-rowed alleles) and one accession (HOR1643) showed the *vrs1.a1*, *Hap_1* haplotypes (six-rowed alleles). Apart from this, two accessions (HOR3529 and HOR3587) displayed the *vrs1.a3*, *Int-c.a* combination (six-rowed alleles). Among the remaining two accessions, interestingly, one accession (HOR5281) showed both two-rowed alleles, i.e. *Vrs1.b3*, *int-c.b*, and the other accession (HOR5471) showed a *deficiens* allele background*Vrs1.t*, *Int-c.a* at *vrs1* and *int-c* loci. These two accessions have been considered as misclassified *labile* genotypes and were omitted from

further analyses. In summary, all 219 *labile* accessions (100%) revealed a six-rowed genotypic background (*vrs1.a*, *Int-c.a*) at the two loci. Thus, data generated in this study can provide concrete evidence that the *labile* accessions possess a six-rowed rather than two-rowed genotypic background.

2.4.4 Haplotype combination/phenotype relationships

The phenotypic data (**Table 2-2**) clearly showed that, the *vrs1.a3*, *Int-c.a1* haplotype combination in *labile* accessions resulted in 99.5% of DS, however, it should be noted that about 10% of the DS were UCS. At the same time, the other six-rowed haplotypes resulted in 41.8% (*vrs1.a1*, *Hap_2*) to 51.6% (*vrs1.a1*, *Hap_1*) of DS with a decrease in the number of UCS (1.0–8.5). Regarding the lateral spikelet development in different haplotypes, almost all lateral spikelets were developed (99.3% of DLS/PLS) in accessions having*vrs1.a3*, *Int-c.a* haplotype (six-rowed haplotype), 54.5% of them include ULS. Meanwhile, the other expected six-rowed haplotypes ranged from 13.7 (*vrs1.a1*, *Int-c.a_Hap2*) to 27.4% (*vrs1.a1*, *Int-c.a_Hap_1*) DLS. On the other hand, 101 *labile* accessions had the *vrs1.a1*, *Int-c.a1/.a2* haplotype and showed 13.7% DLS compared to 14.3% DLS in the typical six-rowed haplotype *vrs1.a1*, *Int-c.a1* which was found in 115 accessions. There was no significant phenotypic effect of the putative heterozygous *Int-c.a* allele (*Int-c.a/Hap_2*) on the number of DLS compared to the homozygous *Int-c.a1* allele (*P*-value = 0.69), suggesting that the altered protein sequence within the TCP domain of *Int-c.a/Hap_2* might not have any deleterious or beneficial phenotypic effect on lateral spikelet fertility.

2.5 Discussion

Lateral spikelet fertility forms an important basis for grouping barleys into two-rowed (unfertile lateral spikelets) and six-rowed (fertile lateral spikelets) forms. The *labile*-barleys show an intermediate phenotype between two-rowed and six-rowed condition, where the

lateral spikelets are occasionally developed but in other cases development is aborted. Variation in lateral spikelet fertility within the spike, and from spike to spike on a single plant, and even from plant to plant within one progeny in *labile*-barley poses an interesting question on its row-type genotypic background (Djalali et al. *1970*). So, studying the genetic mechanism(s) underlying lateral spikelet fertility may similarly contribute to a better general understanding of spikelet development in barley or related grass species. As a first step towards this goal, the current study has been initiated to reveal the allele/haplotype structure of *labile*-barleys at two of the known row-type genes *Vrs1* and *Int-c*.

Sequence analysis of 221 Ethiopian barleys at the *vrs1* locus revealed four previously reported alleles and two new haplotypes at *vrs1.a1*. The Four previously reported alleles include *vrs1.a1*, *vrs1.a3* (six-rowed alleles), *Vrs1.b3* (two-rowed allele) and *Vrs1.t* (*deficiens* allele), were identified in an earlier study by Komatsuda et al. (*2007*). All of the 219 *labile* accessions in the present study showed six-rowed alleles at*vrs1* (*vrs1.a1*- 99.08%, *vrs1.a3* − 0.92%) and only two misclassified accessions had either of the two-rowed alleles *Vrs1.b3* (HOR5281) or *Vrs1.t* (HOR5471). The *vrs1.a1* and *vrs1.a2* alleles identified by Komatsuda et al. (*2007*) are due to a premature stop codon resulting in a truncated and non-functional protein; whereas*vrs1.a3* has an amino acid substitution from phenylalanine to lysine in the conserved HD domain.

The *vrs1.a1* allele is found in six-rowed barleys from all over the world, whereas *vrs1.a2* and *vrs1.a3* are specific to the regions of Western Mediterranean and East Asia, respectively (Komatsuda et al. *2007*). The novel *vrs1.a1 haplotype 67* (*Hap_67*) identified in the present study is caused by an amino acid substitution (aspartic acid to valine) within the conserved HD domain and has been observed in four of the *labile* accessions, whereas the other novel *vrs1.a1* haplotype (*Hap_66*) was due to an SNP in the promoter region. The newly found haplotypes of *vrs1.a1* could be specific to the highlands of Ethiopia

and Eretria, which *labile*-barleys originated from; however, such hypothesis can be tentative since the number of accessions carrying these alleles are very limited (*Hap_66*- nine accessions and *Hap_67*- four accessions). The low number of accessions carrying these two haplotypes may suggest that these evolutionary events were rather recent. Further studies with a larger set of *labile* and non-*labile* accessions may provide a deeper insight into the rise of the two novel haplotypes.

The *Vrs1.t* allele, specific to *deficiens* barleys from Ethiopia, was found in one of the accessions and is most likely due to a misclassification as *labile*-barley. To confirm this, the corresponding accession (HOR5471) was re-grown under long day and short day conditions and the similar *deficiens* phenotype was observed again. The recessive *vrs1.a2* (Komatsuda et al. *2007*) and *vrs1.a4* allele, identified by Cuesta-Marcos et al. (*2010*), were not observed in the present set of *labile* accessions. The *vrs1.a2* allele was identified as being specific to accessions originating from the Western Mediterranean region (Komatsuda et al. *2007*); whereas in case of *vrs1.a4*, of the 102 accessions used in Cuesta-Marcos study four accessions showed the *vrs1.a4* allele and all four accessions originated from Northern America (three from Virginia and one from Washington), suggesting that the *vrs1.a4* allele could be unique to the Northern American region.

By re-sequencing of the *int-c* locus in an association panel of two-rowed and six-rowed barleys (190 lines) Ramsay et al. (*2011*) identified that all two-rowed barleys used in their study carried the recessive *int-c* allele and, conversely, almost all six-rowed barleys had the alternative *Int-c.a* allele, irrespective of their particular*vrs1.a* (six-rowed) allele. Another study reported that the specific dominant *Int-c.a* allele is necessary for the commercial six-rowed phenotype, but dominant or recessive *int-c.b* alleles are found in cultivated two-rowed germplasm (Cuesta-Marcos et al. *2010*). In our study, re-sequencing of *int-c* in *labile* revealed

that all (100%) the *labile* accessions carried *Int-c.a* allele (*Int-c.a*- 118 accessions, *Int-c.a_Hap1*- one accession and *Int-c.a/Int-c.ahaplotype1*- 101 accessions) which further confirms that *labile*-barleys have a six-rowed genetic background; only a single accession showed the two-rowed allele *int-c.b1* at the *int-c* locus along with dominant *Vrs1* allele (misclassified *labile*). The newly identified *Int-c.a* haplotype (*Hap_1*) is due to an amino acid substitution in the conserved TCP domain, it occurred in only one of the accessions in the homozygous state and it is too early to infer that it is specific to *labile*-barleys at this stage. In fact, across 101 accessions the presences of the *Int-c.a* allele and *Hap_1* were simultaneously detected, suggesting that heterozygosity had been maintained at the *int-c* locus. Why heterozygosity appeared to be maintained at this particular locus remains puzzling, given that barley is a highly self-fertilizing species. Hence, another possibility might be that the *int-c* locus has been duplicated in the 101 accessions and that both *int-c* genes differ by only one SNP; however, future work is required to clarify these hypotheses. From our phenotypic analyses it became evident that the apparent heterozygosity in 101 accessions did not significantly alter lateral spikelet fertility compared with the 118 homozygous *Int-c.a* accessions.

The haplotype combinations observed at the *vrs1* and *int-c* loci indicated that all the *labile*-barleys showed preference of six-rowed alleles over two-rowed alleles, i.e. 100% of the accessions have one or the other of the aforementioned six-rowed allele combination. The phenotypic data obtained from the *labile* accessions and its comparison to the observed allele/haplotypes combinations (**Table 2-2**) showed that, in spite of the*vrs1.a* and *Int-c.a* (genotypically six-rowed alleles) being present in the majority of the analyzed accessions, the observed phenotypic data did not support the expected six-rowed phenotype in *labile*. The *labile*-barley spike phenotype displays a variable number of fertile lateral spikelets at each rachis node (0–2 seeds/rachis node; the present study, but also Djalali et al. *1970*;

Takeda and Saito *1988*). Except for the *vrs1.a3, Int-c.a* haplotype combination, which was found in two accessions that gave 99.5% DS, all six-rowed haplotypes, irrespective of their particular allele combinations, showed a maximum of 51.6% of DS. In the *vrs1.a3, Int-c.a* haplotype (two accessions) 99.3% of the lateral spikelets were fertile, but in the other six-rowed haplotypes (217 accessions) only from 13.7 to 27.4% of the lateral spikelets were developed. The *Vrs1.b3,int-c.b1* haplotype combination (two-rowed phenotype) showed an unexpected increase in the number of DS, and subsequently increased developed and fertile lateral spikelet number. This phenotype could potentially be associated with a naturally derived mutation at any of the three other *six-rowed spike* loci such as *vrs2, vrs3* or *vrs4* in this accession.

From previous work it was assumed that the *labile* locus interacts with the recessive six-rowed *vrs1* locus to produce irregular spikes with variable abortion of lateral spikelets (Lundqvist and Franckowiak *2003*). Djalali et al. (*1970*) reported a recessive mode of gene action for *labile* in F_1 and F_2 plants when they crossed *labile*-barley with two-rowed barleys. Currently we are in the process to better understand the *labile* genetics and its corresponding phenotype by developing crosses between six-rowed and *labile*-barleys. Future work is required to elucidate the molecular basis for the apparently random spikelet fertility in *labile*-barleys.

2.6 Acknowledgments:

The authors gratefully acknowledge careful maintenance and supply of the 221 Ethiopian barley accessions from IPK Genebank, Gatersleben, Germany. We are grateful to Dr. Jochen Kumlehn for providing the phyto chamber facility with short day growing conditions for our plants. We also would like to acknowledge funding from the DFG (Deutsche Forschungsgemeinschaft, FKZ SCHN 768-2-1-569091) and the BMBF (German Federal Ministry of Education and Research, FKZ 0315071_GABI-FUTURE Start) during the course of this study. The authors thank Mrs. S. König for excellent technical support.

Table 2-2: *Vrs1* and *Int-c* allele/haplotype combinations and their corresponding spike phenotypes of 219 *labile* accessions after anthesis. Spikelet numbers were calculated from all heads per plant.

			All spikelets				Lateral spikelets				
Allele/haplotype combinations	Number of accessions	Expected row-type	PS	DS	UCS	%DS/ PS	PLS	DLS	ULS	FLS	% DLS/PLS
vrs1.a1, Int-c.a	115	6-r	201.7±69	84.6±31	6.5±7	42	134.5±46	19.3±16	7.2±7	12.1±13	14.3
vrs1.a1, Int-c.a/Hap_2	101	6-r	185.0±72	77.4±31	8.5±8	41.8	123.4±48	16.9±17	5.2±4	12.2±15	13.7
vrs1.a1, Int-c.a_Hap_1	1	6-r	126	65	1	51.6	84	23	1	22	27.4
vrs1.a3, Int-c.a	2	6-r	201.0±81	200.0±76	20.0±22	99.5	134.0±53	133.0±49	54.5±62	78.5±12	99.3
vrs1.a, Int-c.a **(six-rowed controls)**	-	6-r	238.8±6.6	239.6±8.3	0.4±0.5	100.3	160.0±5.7	159.6±5.5	11.0±7.8	148.6±12.5	99.8
Vrs1.b, int-c.b **(two-rowed controls)**	-	2-r	111.7±10.4	111.7±10.4	0.5±0.5	100.0	0	0	0	0	0

PS= Potential spikelets (= rachis nodes number * 3 (in case of six-rowed haplotype) and * 1 (in case of two-rowed or *deficiens*)); as total of all spikes per plant

DS= Developed Spikelets, FLS = Filled Lateral Spikelets, ULS= Unfilled Lateral Spikelets,

UCS= Unfilled Central Spikelets, DLS = Developed Lateral Spikelets PLS= Potential Lateral Spikelet

CHAPTER THREE

Theoretical and Applied Genetics

May 2014, Volume 127, Issue 5, pp 1123-1131 (doi:10.1007/s00122-014-2284-0)

Genetic mapping of the *labile* (*lab*) gene: a recessive locus causing irregular spikelet fertility in *labile*-barley (*Hordeum vulgare* convar. *labile*)

Helmy M Youssef [1,2],Ravi Koppolu[1],Twan Rutten[1],Viktor Korzun[3],Patrick Schweizer[1] and Thorsten Schnurbusch[1]

[1]Leibniz Institute of Plant Genetics and Crop Plant Research (IPK), 06466 Gatersleben, Germany.

[2] Plant Physiology, Faculty of Agriculture, Cairo University, 12613 Giza, Egypt

[3]Cereals Biotechnology, KWS LOCHOW GMBH, D-37574 Einbeck, Germany

Communicated by T. Komatsuda.

3.1 Abstract

Key message

The recessive *labile* locus mapped on chromosome 5HL causes irregular spikelet fertility and controls floret development as well as row-type in barley.

The *labile*-barley displays a variable number of fertile spikelets at each rachis internode (0–3 fertile spikelets/rachis internode) which is intermediate between that observed in two- or six-rowed types. Previous re-sequencing of *Vrs1* in 219 *labile*-barley (*Hordeum vulgare* L. convar. *labile*) accessions showed that all carried a six-rowed specific allele. We therefore hypothesized that this seemingly random reduction in spikelet fertility is most likely caused by the *labile* (*lab*) locus, which we aimed to phenotypically and genetically define. Here, we report a detailed phenotypic analysis of spikelet fertility in *labile*-barleys in comparison to two- and six-rowed genotypes using scanning electron microscopy analysis. We found that the first visible morphological deviation occurred during the stamen primordium stage, when we regularly observed the appearance of arrested central floral primordia in *labile* but not in two- or six-rowed barleys. At late stamen and early awn primordium stages, lateral florets in two-rowed and only some in *labile*-barley showed retarded development and reduction in size compared with fully fertile lateral florets in six-rowed barley. We used two F_2mapping populations to generate whole genome genetic linkage maps and ultimately locate the *lab* locus as a recessive Mendelian trait to a 4.5–5.8 cM interval at approximately 80 cM on chromosome 5HL. Our results will help identifying the role of the *lab* gene in relation to other spikelet fertility factors in barley.

3.2 Introduction

The inflorescence of cultivated barley (*Hordeum vulgare* L.) is an indeterminate spike that produces three single-flowered spikelets at each rachis internode with one central and two lateral spikelets (Harlan *1914*; Bonnett *1935*; von Bothmer et al. *1985*; Forster et al. *2007*; Sreenivasulu and Schnurbusch *2012*). Based upon lateral spikelet fertility, barley is classified into two- and six-rowed varieties (Mansfeld *1950*). In two-rowed barley, to which wild barley (*H. spontaneum*) and some strains of cultivated barley belong, only the central spikelet is fertile and sets seed, while the florets of the two lateral spikelets remain empty. In six-rowed barley, all three spikelets are fertile and produce grains (von Bothmer et al. *1985*).

Apart from two- and six-rowed barleys there is another row-type class, which is better known as *labile*-barley (*Hordeum vulgare* L. convar. *labile* (Schiem.) Mansf.) originally described as an irregular row-type of Abyssinian barley (Åberg and Wiebe *1945*). It was first identified in 1848 and regarded as a transition form between two- and six-rowed barley. Harlan (*1914*) as well as Åberg and Wiebe (*1945*) used the term '*irregular*' because these barleys showed irregular fertility of lateral spikelets along the spike, whereas infertile lateral spikelets were completely reduced to glumes without visible floral residues, rather resembling *deficiens* barleys (*H. vulgare* L. convar. *deficiens* (Schiem.) Mansf.). The *labile* row-type has been considered as a distinct spike character especially among Ethiopian barleys (Bjørnstad and Abay *2010*) and has been grown in most of the barley cropping areas throughout Northern Ethiopia (Abay and Bjørnstad *2009*; Hadado et al.*2009*). Among all the naturally occurring row-type variants, the *labile*-barleys are genetically least described probably owing to their high phenotypic plasticity, which complicated their classification (Mansfeld *1950*).

Genetic mapping and identification of genes controlling spikelet fertility and row-type are crucial for a better understanding of barley inflorescence development, including spikelet

initiation, fertility and abortion. Until today, we know that the row-type phenotype is controlled by at least five independent loci that include *six-rowed spike1* (*vrs1*), *vrs2*, *vrs3*, *vrs4* and *Intermedium -c* (*Int-c*) mapping on barley chromosomes 2HL, 5HL, 1HS, 3HS and 4HS, respectively (Pourkheirandish and Komatsuda *2007*). Variation at the *vrs1* locus, which belongs to the HD-ZIP I class of homeobox transcription factors, is sufficient to control complete lateral spikelet fertility (Komatsuda et al. *2007*). The functional *Vrs1.b* acts as a negative regulator of lateral spikelet fertility resulting in a two-rowed phenotype, whereas the non-functional *vrs1.a* promotes lateral spikelet fertility resulting in a six-rowed phenotype. However, spikelet fertility in two- and six-rowed barley can be modified through the presence of different *Int-c* alleles. Ramsay et al. (*2011*) identified *Int-c* as an ortholog of the maize (*Zea mays* L.) domestication gene, *Teosinte branched1* (*ZmTB1*). They found that two-rowed barleys (*Vrs1*) usually possess the *int-c.b* allele, whereas six-rowed barleys (*vrs1*) often possess *Int-c.a*. Moreover, *Vrs1* also appears to be under the transcriptional control of *Vrs4*, which functions as a central regulator of spikelet meristem determinacy and row-type (Koppolu et al. *2013*).

Djalali (*1970*) noted that, the *labile*-barleys display a continuous variation in the number of fertile lateral spikelets from genotype to genotype. Previous genetic studies suggested that the *labile* phenotype is a constant and heritable attribute (Engledow *1924*) either derived from two- (Breitenfeld *1957*) or six-rowed barleys (Nötzel *1952*). For better understanding the inheritance of the *labile* phenotype, Djalali et al. (*1970*) tested the crosses between *labile* and two-rowed (*H. vulgare* L. convar. *distichon*) barley and reported that two genetic factors are necessary for the manifestation of the *labile* character, they include: the recessive allele for the six-rowed phenotype and the recessive allele at the *lab* locus for the *labile* character. In an attempt to reveal the haplotype structure at the *vrs1* locus in *labile*-barleys, Saisho et al. (*2009*) analyzed a set of 14 *labile* accessions, which showed reduction in

lateral spikelet fertility and found that all carried the*vrs1.a* allele. They suggested that the complete six-rowed spike phenotype of *labile* resulted from a sequence variant in the HD-ZIP motif, whereas the irregular spikelet phenotype is controlled by another genetic factor. Also, *Vrs1* re-sequencing results in 219 *labile*-barley accessions from Ethiopia revealed two six-rowed alleles at *vrs1* (*vrs1.a1* and *vrs1.a3*), but reduced lateral spikelet fertility as well as the occasional missing of central florets (Youssef et al. *2012*). We hypothesized that this reduction in lateral and central spikelet fertility in *labile*-barleys is most likely caused by the recessive *lab* locus (Youssef et al. *2012*).

The generation of high-density genetic maps using markers such as single nucleotide polymorphisms (SNP) has greatly improved the ability to identify genes or QTLs (Stein et al. *2007*). In the recent past, SNPs have become the markers of choice for genetic mapping because they are robust, simple to generate, co-dominant and highly reproducible (Manikanda *2012*). For the barley research community, a very rich resource of SNP markers is available in the form of BOPA SNP markers (Close et al. *2009*) and barley iSELECT chip (Comadran et al. *2012*) for most of these SNPs chromosomal positions were already assigned.

In the present study, we used two mapping populations from crosses between *labile* and six-rowed barley to utilize the available SNP resources and to generate a whole genome genetic linkage map to genetically locate the *lab* locus onto a barley chromosome arm. The recessive *lab* locus mapped at approximately 80 cM on the long arm of chromosome 5H within a genetic interval of four to six cM. We also found that at late stamen and early awn primordium stages, lateral and occasionally central florets in *labile*-barleys showed retarded differentiation and reduction in size compared with fully fertile lateral florets in six-rowed barley. The long-term goal of our research in *labile*-barleys is to identify the underlying gene for the *lab* locus and elucidate its molecular function and relationship with other known *Vrs* genes that control barley spikelet fertility and row-type.

3.3 Materials and methods

3.3.1 Plant materials, growing conditions and spike phenotyping

Based upon our previous work (Youssef et al. *2012*), crosses between different six-rowed barley cultivars and *labile* accessions carrying *vrs1.a* and *Int-c.a* alleles at *vrs1* and *int-c* loci, respectively, were performed. F_1 plants from these crosses were tested for heterozygous plants using SSR markers according to Li et al. (*2003*). True F_1 plants were grown under greenhouse conditions at IPK, Gatersleben; 12/12 h (day/night) light and 14/12 °C (day/night) and F_2 seeds were harvested. Two segregating F_2 populations consisting of 130 individuals derived from the cross between Morex × HOR2573 (M/H2) and 96 individuals derived from the cross between Shimabara × HOR5465 (SH/H5) were used for linkage analysis. The F_1 and F_2 plants were scored for spike phenotypes (either six-rowed or *labile* phenotype) after anthesis.

3.3.2 Scanning electron microscopy (SEM)

For SEM analysis, immature barley spikes at triple mound, lemma, stamen and awn primordium stages (Kirby and Appleyard *1987*) were collected from *labile* and wild-type plants (WT, i.e. Morex, six-rowed; Bowman, two-rowed). Plants were grown under greenhouse conditions as described previously. Immature spikes were fixed with 4 % formaldehyde in 50 mM phosphate buffer, pH 7.0 overnight. After dehydration in a graded ethanol series and critical point drying in a Bal-Tec critical point dryer (Bal-Tec AG, Balzers, Switzerland), spikes were gold sputtered in an Edwards S150B sputter coater (Edwards High Vacuum Inc., Crowley, West Sussex, UK) and examined in a Hitachi S-4100 SEM (Hisco Europe, Ratingen, Germany) at 5 kV acceleration voltage. Digital recordings were made and stored as Tiff-image files.

3.3.3 Genomic DNA isolation

For DNA extraction, leaf samples from the two mapping populations and respective parental genotypes were collected from single plants at the three to five leaf stage. Total genomic DNA was extracted according to Doyle and Doyle (*1990*). DNA quality and quantity were checked on 0.8 % agarose gels. For PCR amplification, DNA concentration was adjusted to 50 ng/µl.

3.3.4 SNP selection and marker development

Two different VeraCode SNP oligo pools comprising of 384 BOPA SNP markers mapped at regular intervals on seven barley chromosomes (see supplementary data) were custom designed for Illumina GoldenGate genotyping on the Bead express reader. The SNP markers with minor allele frequencies (MAF) ≥2.0 were selected to maximize the polymorphism rate (Close et al. *2009*). The raw data from the SNP GoldenGate assay were analyzed using GenomeStudio v2010.3. For further marker development in the defined *labile* map interval (based on VeraCode SNP genotyping data), we relied on the barley genome zipper (Mayer et al.*2011*). Gene sequences from syntenic interval were extracted from *Brachypodium* genome browser server (http://www.phytozome.net/cgi-bin/gbrowse/brachy/). Syntenic *Brachypodium* gene sequences were BLASTed against IPK Barley BLAST server (http://webblast.ipk-gatersleben.de/barley/viroblast.php) to obtain barley sequences for respective *Brachypodium* genes. Primers designed from the barley sequences all had annealing temperatures of 60 ± 1 °C according to Gawroński and Schnurbusch (*2012*).

PCR amplifications were carried out in 25 µl reaction volume containing 20 ng of DNA, 2.5 µl of PCR buffer (10X) (Qiagen, Hilden, Germany), 5 mM dNTPs, 5 µl of Q-solution (Qiagen, Hilden, Germany), 5 pM primers, and 1U of *Taq* polymerase (Qiagen, Hilden, Germany) using a thermal cycler (SensoQuest Thermal Cycler, USA). The PCR profile included an initial denaturation step for 3 min at 94 °C followed by 45 cycles of 94 °C

for 40 s denaturation and an annealing step with constant annealing temperature of 60 °C for 40 s and an extension step at 72 °C for 2 min, followed by a final extension for 10 min at 72 °C. PCR products were tested on 1.2 % agarose gels. For Sanger-sequencing, PCR products were purified using MinElute 96UF PCR purification kit (Qiagen, Hilden, Germany) and sequenced with BigDye Terminator v3.1 cycle sequencing Kits (Applied Biosystems, USA). DNA sequence analysis, quality score assignments and construction of contigs were done with Sequencher 4.7 DNA sequence assembly software. SNP polymorphisms identified from sequencing data were converted to restriction enzyme-based CAPS markers (Vincze et al. *2003*) (http://tools.neb.com/NEBcutter2/). For mapping the *lab* locus, the *labile* phenotype was scored as a monogenic Mendelian trait, and linkage analysis of the phenotype with polymorphic markers was carried out using JoinMap3.0 (Van Ooijen *2006*).

3.4 Results

3.4.1 Spikelet and floret development for the *labile* row-type

SEM image analysis at triple mound, lemma, stamen and awn primordium stages revealed that up until the lemma primordium stage no morphological differences among the three tested row-types became apparent, suggesting that two-rowed, six-rowed and *labile*-barleys initially go through a very similar succession of spike developmental processes (**Fig. 3-1a, b**). The first visible morphological deviation between two- and six-rowed cultivars was found in late stamen primordium and early awn primordium stages (i.e. staging according to the central spikelet and floret development!), when lateral florets in two-rowed showed retarded development and reduction in size compared with lateral florets in six-rowed barley (**Fig. 3-1c1, c2**). The clearly observed developmental difference in lateral florets for these two row-type classes is diagnostic for the allelic differences at the *vrs1* locus. In *labile*-barley at stamen primordium stage, however, we regularly observed the occurrence of arrested central floral primordia (**Fig. 3-1f3**), a feature generally not found in two- and six-rowed barleys

(**Fig. 3-1f1, f2**). Next to that, at late stamen primordium and early awn primordium stages, we found that only a few lateral florets showed retarded differentiation and reduction in size, producing an irregular pattern of spikelet fertility along the spike (**Fig. 3-1c3**). Importantly, glume primordia developed normally in all three row-types at these stages (**Fig. 3-1d1–d3**), indicating that all row-types are not affected in their spikelet development. At the awn primordium stage, the *labile* spike displayed a mosaic of lateral spikelet fertility, eventually setting seeds in lateral florets (**Fig. 3-1g3**); however, fertile lateral florets in *labile*-barley did not differ compared to those in six-rowed barley (**Fig. 3-1e2, g2**). Moreover, glumes appeared to be always present regardless of floral status (i.e. fertile or sterile floret), clearly indicating that the *lab* locus is primarily affecting floret meristem development in a random fashion (**Fig. 3-1e3**). In two-rowed barley, all lateral florets had stopped differentiating at the awn primordium stage (**Fig. 3-1e1–g1**). Most interestingly, in *labile*-barleys, arrested floral development was not restricted to the lateral florets. These morphological features make the *labile* spike (**Fig. 3-1h3**) a mosaic form between the two- (**Fig. 3-1h1**) and six-rowed spike (**Fig. 3-1h2**), but also showed that spikelet fertility in *labile*-barley seems to be generally affected and is not restricted to lateral floral meristems as reported previously (Djalali *1970*; Takeda and Saito *1988*; Saisho et al. *2009*).

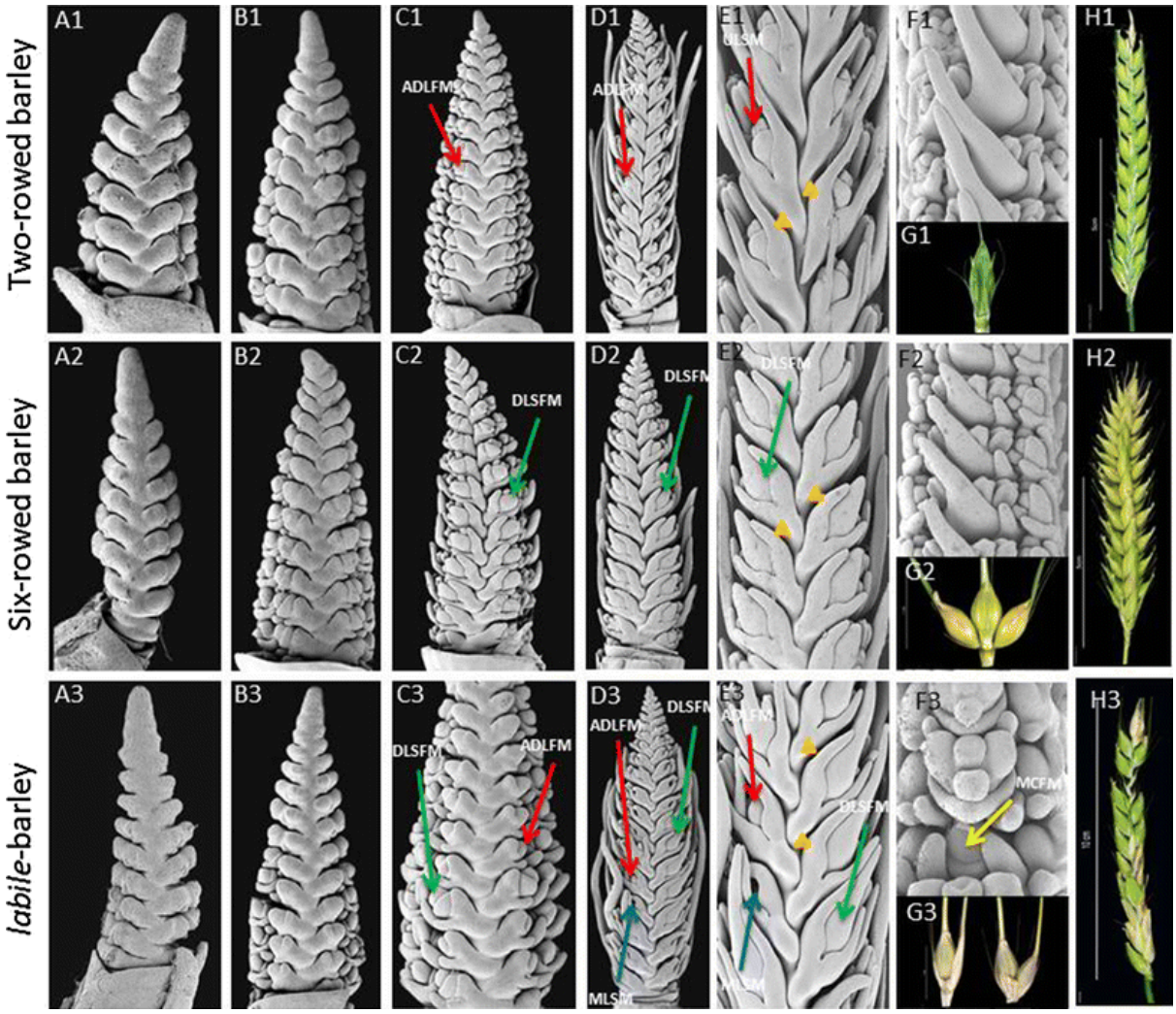


Fig. 3-1: SEM analysis of two-rowed, six-rowed and *labile*-barley spikes. (*A1-3 and B1-3*) Immature spikes at lemma primordium and early stamen primordium stages without visible differences. *C* Lateral view of inflorescences at early awn primordium stage; **C1** two-rowed spike shows arrested development of lateral floret meristem (ADLFM) *red arrows*, **C2** six-rowed spike shows developed lateral spikelet and floral meristems (DLSFM)*green arrows*, **C3** *labile* spike shows DLSFM and ADLFM, *D* lateral view of spikes at late awn primordium stage; **D1** two-rowed spike shows clearly ADLFM, **D2** six-rowed spike shows DLSFM, **D3** *labile* spike shows DLSM and ADLFM in addition to missing lateral floret meristem (MLFM) *blue arrows*, *E* high magnification of (**D**) photos shows clearly ADLFM, DLSFM and MLFM in *labile* spikes comparing with two-rowed and six-rowed barley spikes at late awn primordium stage. At this stage, glume primordium (*GPorange arrows heads*) were developed in two-rowed, six-rowed and *labile* spikes. (*F*) dorsal view of spikes at late awn primordium stage shows no differences between two- and six-rowed barley (**F1** and **F2**). (**F3**) *labile*spike with missing central floret meristem (MCFM) *yellow arrow*. **G** Number of set seeds per rachis internode; **G1** in two-rowed barley only the central spikelet is setting seed. **G2** in six-rowed three spikelets set seeds (one central and two lateral), **G3** *labile*-barley spikelets setting from 0 to 3 seeds per rachis internode. **H** Spikes of two-rowed, six-rowed and *labile*-barley; **H1** two-rowed spike, **H2** six-rowed spike, **H3***labile* spike

3.4.2 Phenotypes of F_1 plants and F_2 populations

Phenotyping of 38 true F_1 plants obtained from the crosses, Morex × HOR2573 (M/H2) and Shimabara × HOR5465 (SH/H5) showed a complete six-rowed phenotype in all F_1 plants, indicating that *labile* is a recessive trait. The F_2 populations of the respective crosses (130 M/H2 plants and 96 SH/H5 plants) segregated for two different spike forms, i.e. either *labile* or six-rowed spike. The *labile* spike phenotype was found in 29 (22.3 %) plants ($\chi^2 = 0.50$) of M/H2 and 22 (22.9 %) plants ($\chi^2 = 0.50$) of the SH/H5 populations (see **Table 3-1**). This near to 1:3 segregation ratio of *labile* versus six-rowed spike phenotype in F_2 confirmed that the *lab* locus segregated as a monogenic recessive gene.

Table 3-1: Spike phenotypes in F_1 and F_2 crosses of Morex × HOR2573 and Shimabara × HOR5465.

	***labile* phenotype**	**six-rowed phenotype**	**Total**	**% of the *labile* phenotype***	**χ2/P value for 3:1**
F_1 (Morex × HOR2573)	-	30	30	0 %	-
F_1 (Shimabara × HOR5465)	-	8	8	0 %	-
F_2 (Morex × HOR2573)	29	101	130	22.3%	0.50/0.48
F_2 (Shimabara × HOR5465)	22	74	96	22.9%	0.50/0.47

* *labile* barleys show an irregular spike phenotype with either missing, undeveloped, and sterile central and lateral florets along a spike.

3.4.3 Genetic mapping of the *labile* (*lab*) locus

To identify the *lab* locus, a set of 381 selected SNP markers were used to perform whole genome mapping in barley. To this end, eight linkage groups were built based on 173 polymorphic markers, which were mapped on the seven barley chromosomes (one linkage group for chromosomes 1H to 6H and two for chromosome 7H) (**Table 3-2**). Mapping of the *labile* phenotype from the F_2 populations located it on chromosome 5H. The *lab* locus

mapped in the genomic region between two linked markers ge00218s01 and ge00066s01 (23.2 cM) (**Fig. 3-2**).

Table 3-2: The whole barley genome genetic linkage analysis.

	No. of selected markers	**No. of polymorphic markers**	**Length of linkage group (cM)**	**Average marker interval (cM)**
Chromosome 1H	40	24	113.8	4.7
Chromosome 2H	61	24	129.1	5.4
Chromosome 3H	53	31	141.1	4.6
Chromosome 4H	45	21	123.7	5.9
Chromosome 5H	67	30	171.6	5.7
Chromosome 6H	51	24	115.5	4.8
Chromosome 7HL	21	12	77.5	6.5
Chromosome 7HS	22	7	40.6	5.8
***Total / average**	**360**	**173**	**912.9**	**5.4**

***** The numbers for No. of selected markers, No. of polymorphic markers, Length of linkage group (cM), represent total values for the seven chromosomes, whereas the number for Average marker interval (cM) represents average value for all chromosomes.

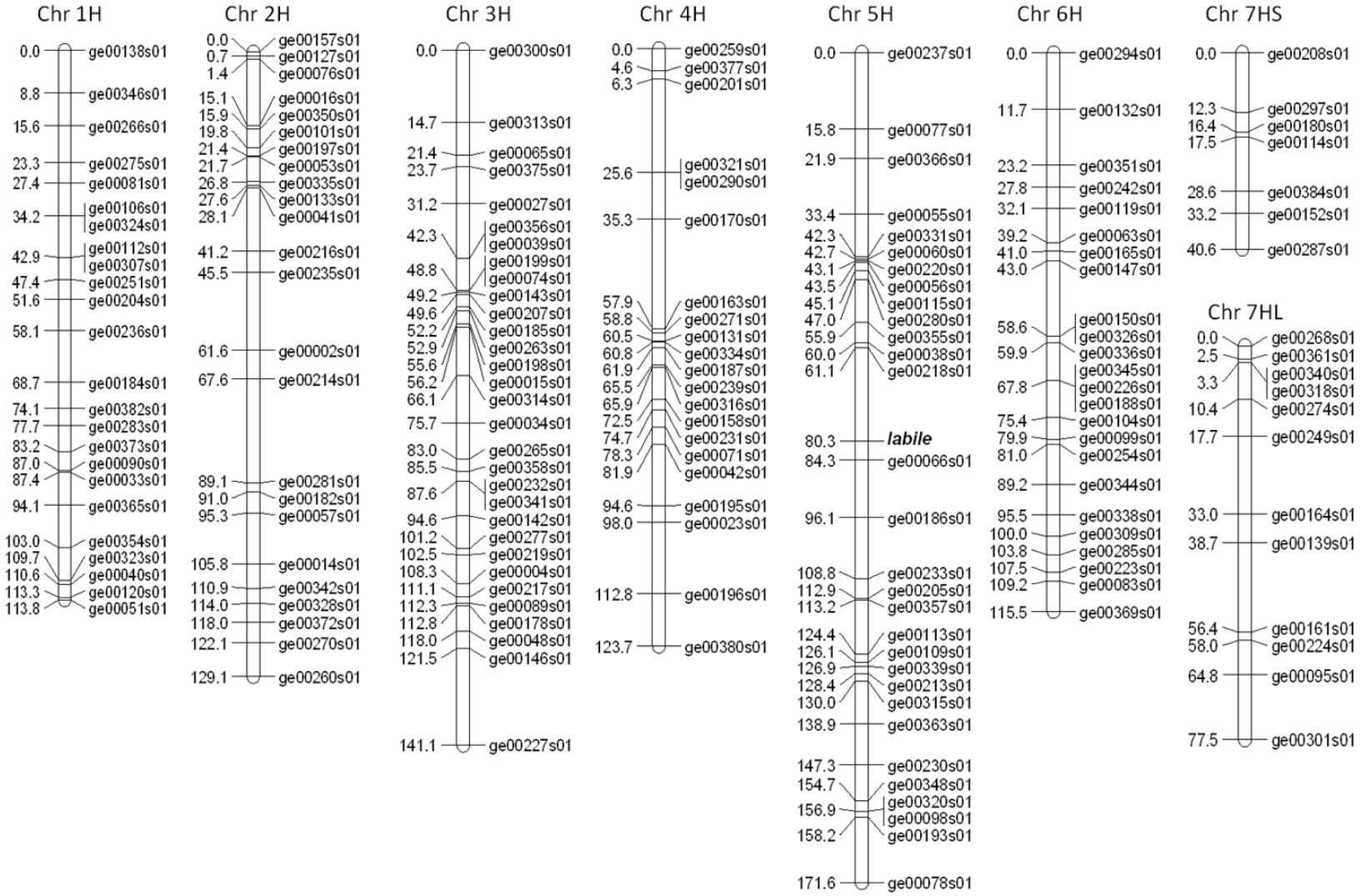


Fig. 3-2: Whole genome mapping of the *labile* phenotype in the F_2 population derived from Morex × HOR2573

We further narrowed down the mapping interval containing *lab*, using 50 primer pairs designed from 18 barley genes, which are in synteny with *Brachypodium* chromosome 4. The syntenic genes were extracted based on the virtual gene order reported in the barley genome zipper (Mayer et al. *2011*). The 18 genes selected for marker design were spaced at a regular interval of 1.3 cM according to genome zipper. Genetic mapping results localized the *lab* locus at approximately 80 cM on the long arm of chromosome 5H to an interval of 5.7 cM in the M/H2 population and 4.6 cM in the SH/H5 population between the closely linked markers BAR and ge00066s01 (**Fig. 3-3**). Five common markers ge00355s01, ABC1, BAR, ge00066s01 and ge00186s01 were mapped in both populations in the *labile* genomic region (**Fig. 3-3**).

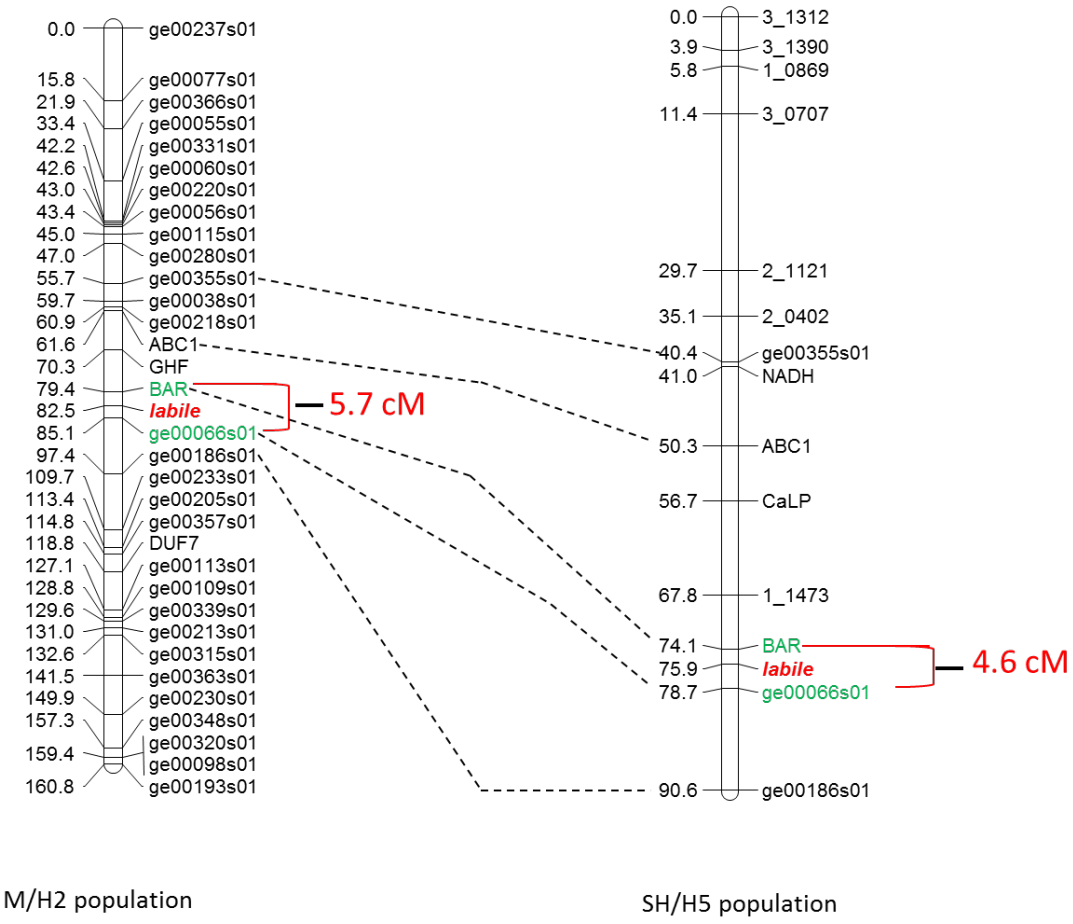


Fig. 3-3: Genetic linkage maps of the *labile* (*lab*) locus on barley chromosome 5H. Linkage analysis was performed on 130 and 96 F_2 plants from the crosses, Morex × HOR 2573 and Shimabara × HOR5465, respectively

3.5 Discussion

Possibly due to their high phenotypic row-type plasticity and restricted regional occurrence, *labile*-barleys (Mansfeld *1950*) are genetically least described among all the naturally occurring row-type variants. This study is the first showing that the *lab* locus can be mapped as a distinct Mendelian trait located on the long arm of chromosome 5H at approximately 80 cM to an interval of 4.5–5.8 cM. The segregating F_2 populations showed either the *labile* or six-rowed spike phenotype. This segregation pattern (1:3) in conjunction with reports by Djalali (*1970*) confirmed that the reliable detection of the *labile* character is primarily dependent upon two conditions: (1) the six-rowed spike phenotype (i.e. genetically constituted as *vrs1.a* + *Int-c.a*), and (2) the recessive allele at the *lab* locus for the *labile* character. Furthermore, we show that the newly identified *lab* locus is genetically distinct from previously known loci conferring the six-rowed spike phenotype in barley (i.e. *vrs1*, *vrs2*, *vrs3*, *vrs4* and *int-c* on chromosomes 2HL, 5HL, 1HS, 3HS and 4HS, respectively; Pourkheirandish and Komatsuda *2007*); but resides close to the *vrs2* locus, which is located in between the flanking markers ge00066s01 and ge00186s01 on the same chromosome arm (data not published). The *vrs2* mutant shows occasional lateral spikelet fertility without any missing central or lateral florets, the characteristic features of *labile* spikes. Moreover, the *vrs2* mutant phenotype does not require the presence of the six-rowed genetic background to become apparent. So, these clear differences between *vrs2*and *labile* mutants strongly suggest that they are two independent genetic loci located on the same chromosome arm (Youssef et al. *2013*).

Among all row-type mutants, the *labile*-barleys are unique in their spike architecture showing a mosaic of six- and two-rowed spike phenotypes (irregular spikelet fertility). The analysis of our scanning electron micrographs revealed that up until the lemma primordium stage there were no visible differences in spike development between two-rowed, six-rowed

and *labile* spikes. During these early stages, the central floral primordia were more developed than those of the lateral floral primordia (see also Komatsuda et al. *2007*). At late stamen and early awn primordium, first morphological differences in spike development became apparent. In the case of six-rowed barleys, central and lateral spikelets/florets displayed complete development, whereas in two-rowed barleys, a clearly retarded development was consistently observed in all lateral florets. Komatsuda et al. (*2007*) provided the first step in the elucidation of lateral floret fertility and showed that loss-of-function of the wild-type *Vrs1* gene (responsible for the two-rowed phenotype) resulted in complete fertility of lateral florets displaying the six-rowed spike phenotype. The *Vrs1* gene belongs to the HD-ZIP I class of homeobox transcription factors. Loss-of-function of the VRS1 protein in lateral floral primordia enabled complete fertility, suggesting that VRS1 suppresses the development of lateral florets. In contrast to two-rowed barleys, infertile lateral florets of *labile*-barleys are completely reduced without any floral development except glumes (**Fig. 3-1d3, e3**), rather resembling the *deficiens* phenotype (Mansfeld *1950*). Moreover, *labile*-barleys showed another interesting feature whereby some central spikelets also remained reduced only to glumes without any floral meristem development (**Fig. 3-1f3**). These *labile*-specific features resulted in a variable number of fertile lateral (Djalali et al. *1970*; Takeda and Saito *1988*) and central (Youssef et al. *2012*) spikelets at each rachis internode (0–3 per rachis internode). This unique spikelet fertility phenotype only observed in *labile*-barleys clearly suggests that the *labile* gene seems to be important for floral meristem identity and development.

The development of six-rowed spikes is often controlled by a non-functional *Vrs1*, (Komatsuda et al. *2007*) complemented by *Int-c.a* (Ramsay et al. *2011*). The spatial and temporal specificity of *Vrs1* gene expression suggests that VRS1 is involved in the development (i.e. suppression) of lateral florets complemented by the presence of the *int-c.b* allele in two-rowed barley (Komatsuda et al. *2007*; Ramsay et al. *2011*). We established

that in both two-rowed and *labile*-barley this suppression of lateral spikelet development starts in late stamen primordium (see **Fig. 3-1c1, c3**). According to Komatsuda et al. (*2007*), loss-of-function in *Vrs1*during early spike development leads to the formation of the six-rowed spike. Our previous study revealed that *labile*-barleys carry six-rowed alleles (*vrs1.a, Int-c.a)* at *vrs1* and *int-c* loci (Youssef et al. *2012*), displaying a mosaic spike phenotype between the two- and six-rowed condition. Apparently, the lack of *Vrs1*gene function in combination with *Int-c* may be sufficient to explain the lateral spikelet fertility in *labile*-barleys, but seems rather insufficient to clarify the random floral sterility seen in lateral as well as central spikelets. Thus, identification of the underlying gene for the *lab* locus, using a map-based and mutant analysis approaches seems promising. Molecular genetic results in combination with the examination of detailed lateral and central spikelet development in *labile*- and other row-types may help elucidate the role of the *lab* gene in relation to other floret development and fertility factors in barley.

3.6 Acknowledgments:

The authors gratefully acknowledge Dr. Marion Röder for providing the SSR markers and lab facilities to test the F1 plants. We also thank Mrs. S. König and I. Walde for excellent technical support and H. Ernst for taking spike photos. We also would like to acknowledge funding from the DFG (Deutsche Forschungsgemeinschaft, FKZ SCHN 768/2-1) and the BMBF (German Federal Ministry of Education and Research, GABI-FUTURE Start Program grant number FKZ 0315071) during the course of this study.

CHAPTER FOUR

Submitted to JXB

Genotypic and phenotypic analyses of wild and Intermedium barleys highlights two distinct row-types in *Hordeum spontaneum*

Helmy M Youssef [1,2]; Martin Mascher [3]; Mohammad Ayoub[1, 2]; Benjamin Kilian[4,5]; Nils Stein[4] and Thorsten Schnurbusch[1,]

[1] Leibniz Institute of Plant Genetics and Crop Plant Research (IPK), HEISENBERG-Research Group Plant Architecture, 06466 Gatersleben, Germany.

[2] Faculty of Agriculture, Cairo University, 12613 Giza, Egypt

[3] Leibniz Institute of Plant Genetics and Crop Plant Research (IPK), Research Group Domestication Genomics, 06466 Gatersleben, Germany.

[4] Leibniz Institute of Plant Genetics and Crop Plant Research (IPK), Research Group Genome Diversity, 06466 Gatersleben, Germany.

[5]Current address: Bayer CropScience NV Ghent, Belgium.

4.1 Abstract

Cultivated barley (*Hordeum vulgare* L.) was domesticated from its wild two-rowed ancestor *H. vulgare* ssp. *spontaneum*. The row-type of the barley inflorescence is controlled by at least six independent loci, including *six-rowed-spike1* (*vrs1*), *vrs2*, *vrs3*, *vrs4*, *Intermedium -c* (*int-c*) and *labile* (*lab*). To clarify the relationship between two- and six-rowed cultivated barley, and to better understand the importance and role of the 'Intermedium' row-type during its evolution and domestication, phenotypic analyses and re-sequencing of three known row-type genes (*Vrs1*, *Vrs4* and *Int-c*) was performed in a worldwide 'Intermedium' collection (302 accessions), wild barley ssp. s*pontaneum* (100) and ssp. *agriocrithon* (10) accessions. Measuring lateral spikelet length and width revealed the presence of two distinct row-type groups within ssp. *spontaneum*; one 'Intermedium' row-type form with enlarged but infertile lateral spikelets, designated as ssp. *spontaneum* (Intermedium), and one group with significantly smaller lateral spikelets similar to cultivated two-rowed barleys, designated as ssp. *spontaneum* (Distichon). 93.7% of the 'Intermedium' collection and all of ssp. *spontaneum* (Intermedium) carried two-rowed allele *Vrs1.b* in combination with the six-rowed allele *Int-c.a*. All ssp. *spontaneum* (Distichon) accessions possessed two-rowed alleles at both *Vrs1* and *Int-c*. On the contrary, ssp. *agriocrithon* carried six-rowed specific alleles at all tested row-type loci. *Vrs1* transcript levels in wild and Intermedium barleys correlated negatively with lateral spikelet fertility as expected; but with a few exceptions. This study provides new insights into the relationship between wild and cultivated barleys during domestication, and identify 'Intermedium' barleys as a rich genetic resource for increasing lateral spikelet fertility.

Key words: domestication, Intermedium barley, *Int-c*, ssp. *spontaneum*, spikelet fertility, *Vrs1*

4.2 Introduction

Barley (*Hordeum vulgare* L.) was among the earliest crops to be exploited during the Mesolithic Neolithic transition (Weiss et al. 2004) and the first to become domesticated about 9500 years before present (BP) (Zohary and Hopf 2000). The immediate ancestor of two- and six-rowed cultivated barley is still abundant in nature. It was first discovered and described as '*Hordeum spontaneum'* from Turkey by the German botanist Carl Koch (1809-1879). This progenitor form is nowadays regarded as a subspecies (ssp. *spontaneum* (C. Koch) Thell.) of *H. vulgare* L., similar to cultivated barley (ssp. *vulgare*) (Bothmer et al. 2003). Although *H. vulgare* ssp. *spontaneum* was regarded a direct ancestor of cultivated two-rowed barley based on their close similarity (De Candolle, 1959), the origin of six-rowed cultivated barley remained unclear. With the discovery of *H. agriocrithon*, a six-rowed form with brittle rachis from western China (Åberg 1938) a direct ancestor of the six-rowed cultivated barley appeared to have been found (Åberg, 1940; Bell, 1965). However, most barley scientists nowadays favor the theory of a single evolutionary line from ssp. *spontaneum* to ssp. *vulgare* (Pourkheirandish and Komatsuda 2007). Remains of six-rowed barley appear in aceramic Neolithic beds from 8800 BP onwards (Helbaek, 1959; Zohary and Hopf, 2000) and in 9000 year old sediments at the early agricultural site Ali Kosh; two-rowed barley materials were interspersed with traces of six-rowed types (Helbaek, 1969). This seems to identify two-rowed barley as the main source of six-rowed forms during domestication (Pourkheirandish and Komatsuda 2007). Although Zhang et al. (1994) suggested that Tibetan six-rowed wild barley may be a direct ancestor of Tibetan six-rowed and two-rowed cultivated barley, Feng et al. (2006) argued that Chinese cultivated barley originated from two-rowed wild barley from Tibet. In their view the Tibetan six-rowed wild barley is an intermediate form in the process of transformation from two-rowed wild barley to cultivated barley.

The inflorescence of cultivated barley is an indeterminate spike that produces three single-flowered spikelets at each rachis internode; a unique feature of the genus *Hordeum*. Classification of cultivated barley started in late 19th century with the systematic work of Körnicke (1885). Present day classification is based on Mansfeld (1950) who split cultivated barley into five row-type groups: (i) two-rowed barley (convar. *distichon*) with fully fertile central spikelets and two sterile lateral spikelets, (ii) deficiens-barley (convar. *deficiens*), a two-rowed barley with no or extremely reduced lateral spikelets, (iii) six-rowed barley (convar. *hexastichon*) with three fully fertile spikelets producing three seeds per rachis node, (iv) 'Intermedium' barley (convar. *intermedium*) with fully fertile central spikelets but enlarged and developed lateral spikelets varying from rarely fertile to complete fertile producing small seeds, and (v) labile-barleys (convar. *labile*) with lateral spikelets developed or absent, fertile or sterile even within one spike of the same genotype showing an irregular or random pattern of fertility at all spikelet positions (Youssef et al. 2014).

Barley's row-type is at least controlled by six independent loci that include *six-rowed-spike1* (*vrs1*), *vrs2*, *vrs3*, *vrs4*, *Intermedium -c* (*int-c*) and *lab* mapping on chromosomes 2HL, 5HL, 1HS, 3HS, 4HS and 5HL, respectively (Pourkheirandish and Komatsuda 2007; Youssef et al. 2014). The genes *vrs1*, a HD-ZIP I class of homeobox transcription factor, and *int-c*, an ortholog of the maize (*Zea mays* L.) domestication gene, *Teosinte branched 1* (*ZmTB1*), are sufficient to control lateral spikelet fertility. However, modification of lateral spikelet fertility by alleles at the *int–c* locus depends on the allelic constitution at *Vrs1* (Komatsuda et al. 2007; Ramsay et al. 2011). The *int–c.b* allele for example represses the development of lateral spikelets in cultivated two–rowed barley (*Vrs1.b*), whereas in cultivated six–rowed types (i.e. *vrs1.a* genotype) it results in reduced lateral spikelet fertility and small seeds. The *Int–c.a* allele is predominantly present in cultivated six–rowed barley, but in combination with *Vrs1.b* (i.e. two-rowed) it causes partially fertile lateral spikelets, resulting in an 'Intermedium' spike phenotype (Ramsay et al., 2011).

Koppolu et al. (2013) identified the *vrs4* locus as an ortholog of maize *RAMOSA2* (*RA2*) which encodes a Lateral Organ Boundaries (LOB) domain-containing transcriptional regulator. Interestingly, *Vrs4* (*HvRA2*) controls barley row-type because all of the induced *vrs4* mutants analyzed showed either a complete or partial six–rowed phenotype (Koppolu et al. 2013), clearly suggesting that *Vrs4* is functionally upstream of *Vrs1*.

Understanding the genetic basis controlling variation in inflorescence architecture will also help to understand the foundation of morphological diversity during barley domestication (Malcomber et al., 2006). This is important because the domestication process leads to reduced genetic diversity relative to the wild ancestor across the genome (Nevo 2006). Gene sequence variations reflect the genetic and evolutionary history of organisms (Maeso et al. 2012). Single nucleotide polymorphism (SNP) analysis is currently one of the most powerful tools for defining genetic variation between cultivated barley and its wild ancestor. In the current study, we therefore re-sequenced the loci controlling barley row-type (*vrs1*, *vrs4* and *int-c*) in a worldwide barley collection classified as 'Intermedium' with the aim of (*i*) clarifying the relationship of cultivated two- and six-rowed types and its ancestors during the domestication process, and (*ii*) understanding the 'Intermedium' row-type and its role during barley evolution.

4.3 Materials and Methods

4.3.1 Plant materials, growing conditions and spike phenotyping

Three hundred and two worldwide 'Intermedium' barley (*Hordeum vulgare* L. convar. *Intermedium* -Körn- Mansf) accessions, ten accessions of *H. vulgare* ssp. *agriocrithon* and 100 accessions of *Hordeum vulgare* ssp. *spontaneum* used in this study were obtained from the gene bank and genome diversity group at IPK-Gatersleben, Germany (**supplementary (sup.) Fig. 4-1**). Three replicates from each accession of 'Intermedium' barley were

germinated in trays and kept for 10-15 days. At the 3 leaves stage plantlets were vernalized for four weeks at 4ºC followed by hardness for one week under light conditions 12/12 h day/night and temperature 12±2 ºC day and night. Afterwards, plants were grown in greenhouse under long day conditions, 16 h/8 h (day/night) light and temperature 20±2 ºC during the day and 16±2 ºC during the night. At heading stage plants were scored for phenotypes such as spike size (normal or dense), central and lateral spikelet size and fertility (long or small, filled or unfilled spikelet), length and form of awns (long, short or absent, normal or hooded). Additionally, from each ssp. *spontaneum* accession length and width of 10 lateral spikelets/five rachis internodes from the middle of the spike was measured after harvest from field-grown spikes in 2014.

4.3.2 Genomic DNA isolation

Leaf samples were collected at 3-5 leaf stage. Total genomic DNA was extracted according to Doyle and Doyle (1990). DNA quality and quantity were checked on 1% agarose gels and concentration adjusted to ~40ng/µl for PCR work.

4.3.3 PCR amplification, sequencing and sequence analysis

Three primer pairs were designed (with 120 to 150 bp overlapping between each two primer pairs) to cover the 2.062 kb fragment at the *vrs1* locus in order to obtain sequence data for the whole gene by Sanger sequencing. Regarding *vrs4* locus, a single primer pair was designed to amplify a 1.2 kb fragment covering the single exon (774 bp) as well as 264 bp upstream of ATG and 162 bp downstream of TGA. A single primer pair was designed for the *int-c* locus to amplify the single exon 1.07 kb fragment. Primer sequence, annealing temperature and fragment length for all loci are described in **sup. Table 4-1**. PCR amplifications with individual primer pairs for *vrs1*, *vrs2*, *vrs4* and *int*-c were performed in all 302 ‘Intermedium’ barley accessions.

PCR amplifications were performed in 25 µl total reaction volume containing 3 µl DNA (40 ng/µl), 2.5 µl PCR buffer 10X (Qiagen, Hilden, Germany), 5 mM dNTPs, 5 µl Q-solution (Qiagen, Hilden, Germany), 5 pM primers, and 1U of Taq polymerase (Qiagen, Hilden, Germany) using thermal cycler (SensoQuest Thermal Cycler, USA). Touchdown PCR amplification profile with initial denaturation step 3 min at 95°C followed by 7 cycles of 95°C for 40 s. Annealing step were 63°C for 40 s (for *Vrs1-2*, *Vrs4-1* and *Int-c-1*), 59°C for 40 s (for *Vrs1-1* and *Vrs1-3*). Extension was at 72°C for 2 min, with 1°C decrement in annealing temperature per cycle, followed by main program for 40 cycles at 95°C for 40 s with constant annealing temperatures, 60°C (for *Vrs1-2*, *Vrs4*, *Int-c-1*), 56°C (for *Vrs1-1* and *Vrs1-3*) for 40 s and extension at 72°C for 2 min, followed by a final extension for 10 min at 72°C. Amplification of PCR products were tested on 1% agarose gels. For Sanger-sequencing, PCR products were initially purified using MinElute 96UF PCR purification kit (Qiagen, Hilden, Germany), then sequenced using BigDye Terminator v3.1 cycle sequencing Kits (Applied Biosystems, USA). DNA sequence analysis, quality score assignments and construction of contigs were achieved using Sequencher 5.2.3 DNA sequence assembly software. Multiple sequence alignments were carried out using ClustalW2 software (http:// www.ebi.ac.uk/Tools/msa/clustalw2/.

4.3.4 Exome capture data

Exome capture data (Mascher et al., 2013) was generated for (among others) 100 ssp. *spontaneum* and 10 ssp. *agriocrithon* accessions and included in this study. Reads were mapped against the whole genome shotgun assembly of barley cv. Morex (IBSC, 2012) with BWA (Li and Durbin 2009). Reads mapping to *vrs1*, *vrs4* and *int-c* loci were extracted with SAM tools and assembled with CLC assembly cell (http:// www.clcbio.com) using default parameters.

4.3.5 RNA extraction and qPCR analysis

Total RNA was extracted from spike meristems collected at awn primordium stage from four *spontaneum* accession (FT11, FT144, FT237 and 332) and six Intermedium accessions (HOR804, HOR4838, HOR7076, HOR7078, HOR7211 and HOR13728) using the Absolutely RNA Microprep Kit (Agilent). RNA concentrations were measured using NanoDrop ND-1000 spectrophotometer (Peqlab, Biotechnologie GmbH, Germany). Reverse transcription and cDNA synthesis were carried out using SuperScript® III Reverse Transcriptase (RT) kit (Invitrogen). Real-time PCR was performed using the Power SYBR Green PCR Master Mix Kit (Applied Biosystems) and the ABI Prism 7900HT sequence detection system (Applied Biosystems). Quantitative RT-PCR primer sequences were as follow; F-Vrs1_ 3'UTR 5'-TGGGCTGTGATTTGAAGTTG-3', R-Vrs1_ 3'UTR 5'-AGTAATCGATCAGCGTGCAG-3', F-int-c 5'-GCTTGTGAATTAGCTCCGTGC-3' and R-int-c 5'-CAATACTGCATCATGGAATGACC-3'. RT-PCR results were analyzed using SDS2.3 software and PCR efficiency was calculated using LinRegPCR 2015.3 software.

4.4 Results

4.4.1 Phenotypic analyses of spike and spikelets in 'Intermedium' and wild barleys

Measuring lateral spikelet length and width in 100 wild barley ssp. *spontaneum* accessions revealed the presence of two distinct row-type groups (**Fig. 4-1a and b**). The first group containing 52 accessions had enlarged lateral spikelets (average length 17.6 mm and width 1.95 mm) and was defined as ssp. *spontaneum* 'Intermedium' (**Fig. 4-2a and b**). The second group comprising 48 accessions had significantly smaller lateral spikelets (length 11.5 mm and width 1.23 mm) and was classified as ssp. *spontaneum* 'Distichon' (**Fig. 4-2a and b**). Wild barley ssp. *agriocrithon* always showed a complete six-rowed spike phenotype with

three fully developed seeds at each rachis internode (**Fig. 4-1c**) not distinguishable from six-rowed forms of cultivated barleys.

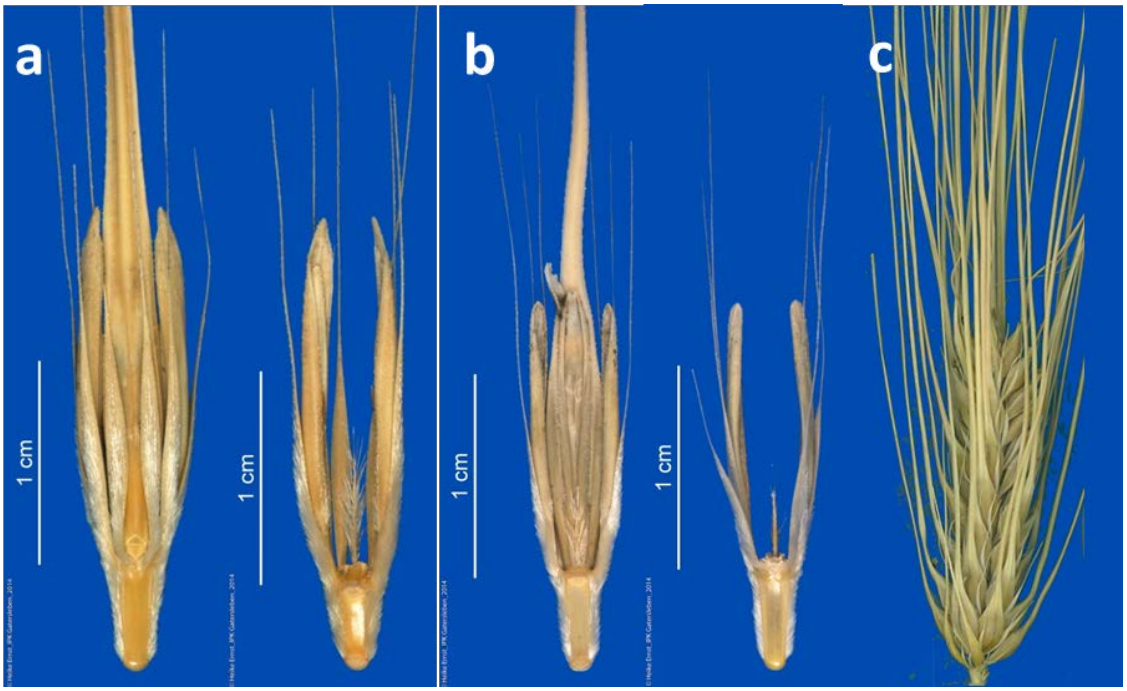


Fig. 4-1: Images of wild barley spikelets. (a) the three spikelets at the rachis internode (left) and only the lateral (right) in the *spontaneum* (Intermedium) in comparison with (b) two-rowed *spontaneum* (Distichon), (c) six-rowed spike of ssp. *agriocrithon* (image modified from USDA website: http://www.ars-grin.gov/cgi-bin/npgs/html/dispimage.pl?313529.

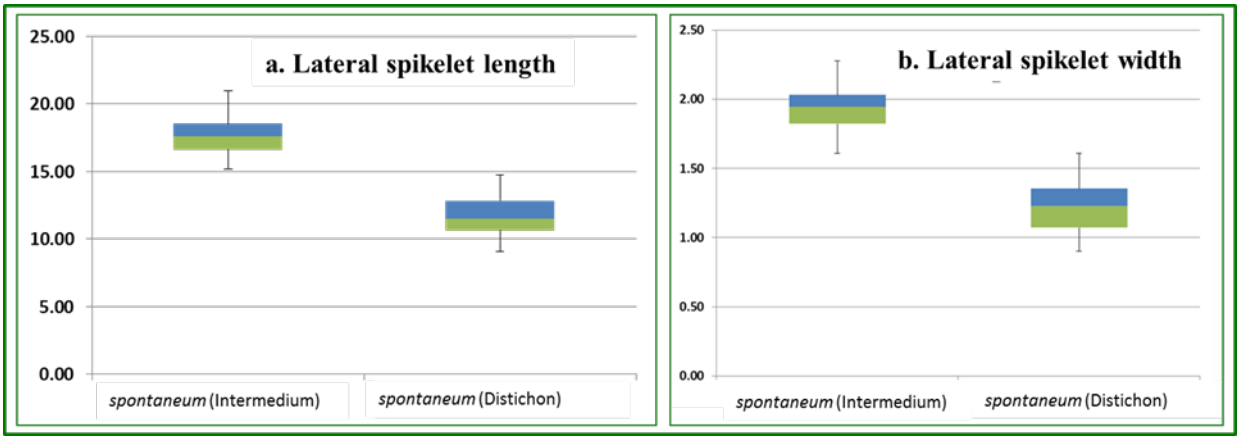


Fig. 4-2: Box-plot diagram showing the lateral spikelet length and width (mm) of 100 accessions of wild barley ssp *spontaneum*. (a) Box-plot of spikelet length showing the group of ssp. *spontaneum* 'Intermedium' (n=52; median 17.6, range 15.2-20.9), ssp *spontaneum* 'Distichon' (n=48; median 11.5, range 9.1-14.8) and (b) Box-plot of spikelet width, ssp. *spontaneum* 'Intermedium' (n=52; median 1.95, range 1.61-2.28) ssp *spontaneum* 'Distichon' (n=48; median 1.23, range 0.90-1.61).

The 302 barley accessions were classified as 'Intermedium' based on Mansfeld (1950). Spike phenotype in this collection is not uniform differing in many phenotypic characters such as spike phenotype and lateral size (vary from complete six-rowed to complete two rowed; **Fig. 4-3**) spike length, grain form (naked or covered), grain and ear color (yellow or colored), central and lateral spikelet shape (long or short awns, awns pointed, awnless or hooded). Phenotypic data showed a wide variation in lateral and central spikelet fertility ranging from developed but un-filled spikelets/florets to complete six-rowed spike. Within each spike, lateral spikelets always were distinctly smaller than central spikelets. Among the 302 accessions, only four accessions (HOR3825, HOR10281, HOR14233 and HOR17788) showed a normal two-rowed phenotype, while further 11 accessions (HOR7040, HOR7041, HOR7129, HOR7512, HOR8746, HOR8755, HOR8771, HOR8773, HOR8779, HOR11414 and HOR14237) showed similarity to the *vrs3* mutant phenotype (the basal third of the spike showing two-rowed spike phenotype with enlarged lateral spikelets whereas the apical two-third of the spike were completely six-rowed).

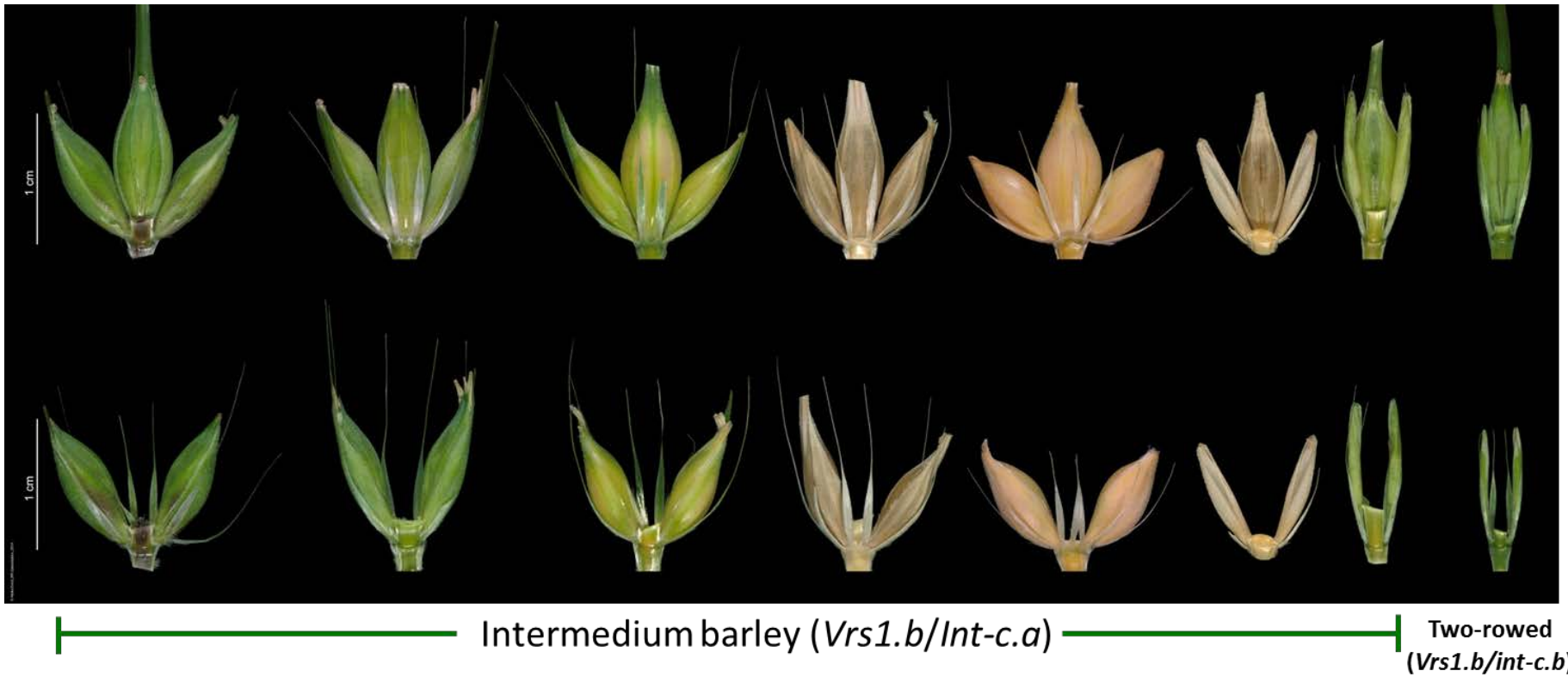


Fig. 4-3: Spikelet triplet at each rachis node in the Intermedium collection (*Vrs1.b/Int-c.a*) showing the variation in fertility and size of lateral spikelets from complete six-rowed (left) to complete two-rowed (penultimate right); two-rowed barley (*Vrs1.b/Int-c.b*).

4.4.2 Genotypic status and network analysis at the *vrs1* locus in 'Intermedium' and wild barleys

Sequence analysis of the *vrs1* locus revealed 20 new haplotypes compared with known alleles. These known alleles and the new haplotypes could be divided into two groups based on the sequence at the positions (-597 bp), (-531 bp) and (-513 bp) upstream of the start codon (ATG). The first group carries C, T and T at these positions. Within this group three known alleles, including the six-rowed allele *vrs1.a2* and the two-rowed alleles *Vrs1.b2* and *Vrs1.t* are present next to 12 new haplotypes including 11 haplotypes of the functional two-rowed allele *Vrs1.b2* named *Haplotypes 1-11* of *Vrs1.b2* (*Hap_1-Hap_11*) and one haplotype showed the same deletion as in *vrs1.a1* (*Hap_12*). The second group displays G, G and C at the positions mentioned above, and contains the other four known alleles including the six-rowed alleles (*vrs1.a1* and *vrs1.a3*) and the two-rowed alleles (*Vrs1.b1* and *Vrs1.b3*) next to eight new haplotypes. The latter have been denoted as *Haplotypes 13* and *15-20* of *Vrs1.b3* (*Hap_13*, and *Hap_15 - Hap_20*), with *Hap_20* being unusual in having a 'T' instead of a 'C' at (-513 bp), while *Hap_14* is considered a haplotype of *vrs1.a1*. (**sup. Table 4-2**). Within the first group, we found that 13 accessions from wild barley ssp. *spontaneum* possessed the two-rowed allele *Vrs1.b2*. Among the 302 'Intermedium' accessions, 260 belong to the first group of haplotypes comprising 182, 41, 23 and 4 accessions possessing the sequence of *Hap_1*, *2*, *3* and *4*, respectively, two accessions carried the sequence of each haplotype from *Hap_5* to *Hap_7*, one accession belongs to each haplotype from *Hap_8* to *Hap_11*. However, *Hap_12* sequence is defined by five accessions from ssp. *agriocrithon*. The sequence analysis of the second group revealed that two Intermedium and five ssp. *agriocrithon* accessions have the six-rowed allele *vrs1.a1*, a further 30 'Intermedium' accessions carried the two-rowed allele (*Vrs1.b3*) and a single 'Intermedium' accession the six-rowed allele (*vrs1.a3*). The remaining nine 'Intermedium' accessions carried the sequence of *Hap_13* to *Hap_16* (one accession for

each), *Hap_17* (two accessions) and *Hap_18* (three accessions). Among the 100 accessions of ssp. *spontaneum*, 66 and 21 of them possessed *Hap_19* and *Hap_20*, respectively. In the coding region, we found that the SNP differences at 23 bp and 202 bp downstream of the start codon cause amino acid substitutions from glycine (G) and glutamic acid (E) to aspartic acid (D), respectively, and are present either in two- (*Vrs1.b*) or six-rowed (*vrs1.a*) genotypes (Table 1). None of the 20 new haplotypes contained specific SNPs at location 243 bp and 349 bp downstream of the start codon, responsible for the difference between the six-rowed alleles (*vrs1.a2* and *vrs1.a3*) and the known two-rowed alleles (**Table 4-1**).

The *vrs1.a1* deletion at 680 bp producing a truncated protein2 (TP2) was found in *Hap_12* (in 5 ssp. *agriocrithon* accessions) and *Hap_14* (in only one 'Intermedium' accession; HOR6408). Of the new 20 haplotypes only one (*Hap_8*; in HOR12251) had a unique single nucleotide polymorphism (SNP) within the coding region at 396 bp causing amino acid change from alanine to glycine. Interestingly, another haplotype (*Hap_16*; in HOR10255) showed highly consistent, putatively heterozygous positions at the *Vrs1.b3* and *vrs1.a1* alleles, where the characteristic SNPs of the v*rs1.a1* (SNPs A to G at 23 pb, T to G at 202 bp and deletion of G at 680 bp in *Vrs1.b3*) occurred simultaneously. The network analysis (**Fig. 4-4**) of these new 20 haplotypes and seven known alleles at *Vrs1* showed that the majority of the new haplotypes present in the Intermedium collection (*Hap_1* to *Hap_11*) in addition to known alleles (*vrs1.a2* and *vrs1.t*) are clearly derived from the *Vrs1.b2* which is present in 11 ssp. *spontaneum* 'Distichon' collected from the western part of Fertile Crescent (Morocco, Egypt Syria, Iran and Turkey) and 2 ssp. *spontaneum* 'Intermedium' accessions collected from Turkey (**Fig. 4-4**). While, the rest of the new Haplotypes (*Hap_12* to *Hap_18*) in addition to three known alleles (*vrs1.a1*, *vrs1.a3* and *Vrs1.b3*) are derived from the wild barley *spontaneum* haplotypes; *Hap_19* and *Hap_20*. In summary, 298 (98,7%) of the 302 'Intermedium' accessions, and all 100 ssp. *spontaneum* accessions (100%) showed known

two-rowed alleles or two-rowed related haplotypes at *Vrs1*. In contrast, all 10 tested ssp. *agriocrithon* accessions (100%) carried the six-rowed allele *vrs1.a1* or its derived haplotype (*Hap_12*).

Table 4-1: Amino acid changes and positions in the known alleles and new haplotypes and the accessions number carrying these haplotypes at *vrs1* locus

SNP position	**23**	**202**	**243**	**349**	**396**	**680**	**757**	**775**	**Number of accessions**		
Amino acid position	**8**	**68**	**TP1**	**79**	**91**	**TP2**	**178**	**184**	**Inter.**	**Agrio.**	**Spont.**
vrs1.a1	G	E	-	R	A	TP2	L	A	2	5	
vrs1.a2	G	E	TP1	-	-	-	-	-	-		
vrs1.a3	D	D	-	G	A	-	V	S	1		
Vrs1.b1	G	E	-	R	A	-	I	S	-		
Vrs1.b2	G	E	-	R	A	-	V	S	-		13
Vrs1.b3	D	D	-	R	A	-	V	S	30		
Vrs.t	G	E	-	R	A	-	V	G	-		
Hap_1	G	E	-	R	A	-	V	S	182		
Hap_2	G	E	-	R	A	-	V	S	41		
Hap_3	G	E	-	R	A	-	V	S	23		
Hap_4	G	E	-	R	A	-	V	S	4		
Hap_5	G	E	-	R	A	-	V	S	2		
Hap_6	D	D	-	R	A	-	V	S	2		
Hap_7	D	D	-	R	A	-	V	S	2		
Hap_8	G	E	-	R	G	-	V	S	1		
Hap_9	G	E/D	-	R	A	-	V	S	1		
Hap_10	G	E	-	R	A	-	V	S	1		
Hap_11	G	E	-	R	A	-	V	S	1		
Hap_12	G	E	-	R	A	TP2	L	A	-	5	
Hap_13	D	E	-	R	A	-	V	S	1		
Hap_14	D	D	-	R	A	TP2	L	A	1		
Hap_15	G	E	-	R	A	-	V	S	1		
Hap_16	G/D	E/D	-	R	A	TP2/FP	V/L	S/A	1		
Hap_17	D	FT/RG	-	R	A	-	V	S	2		
Hap_18	G	E	-	R	A	-	V	S	3		
Hap_19	G	E	-	R	A	-	V	S	-		66
Hap_20	G	E	-	R	A	-	V	S	-		21
Total									302	10	100

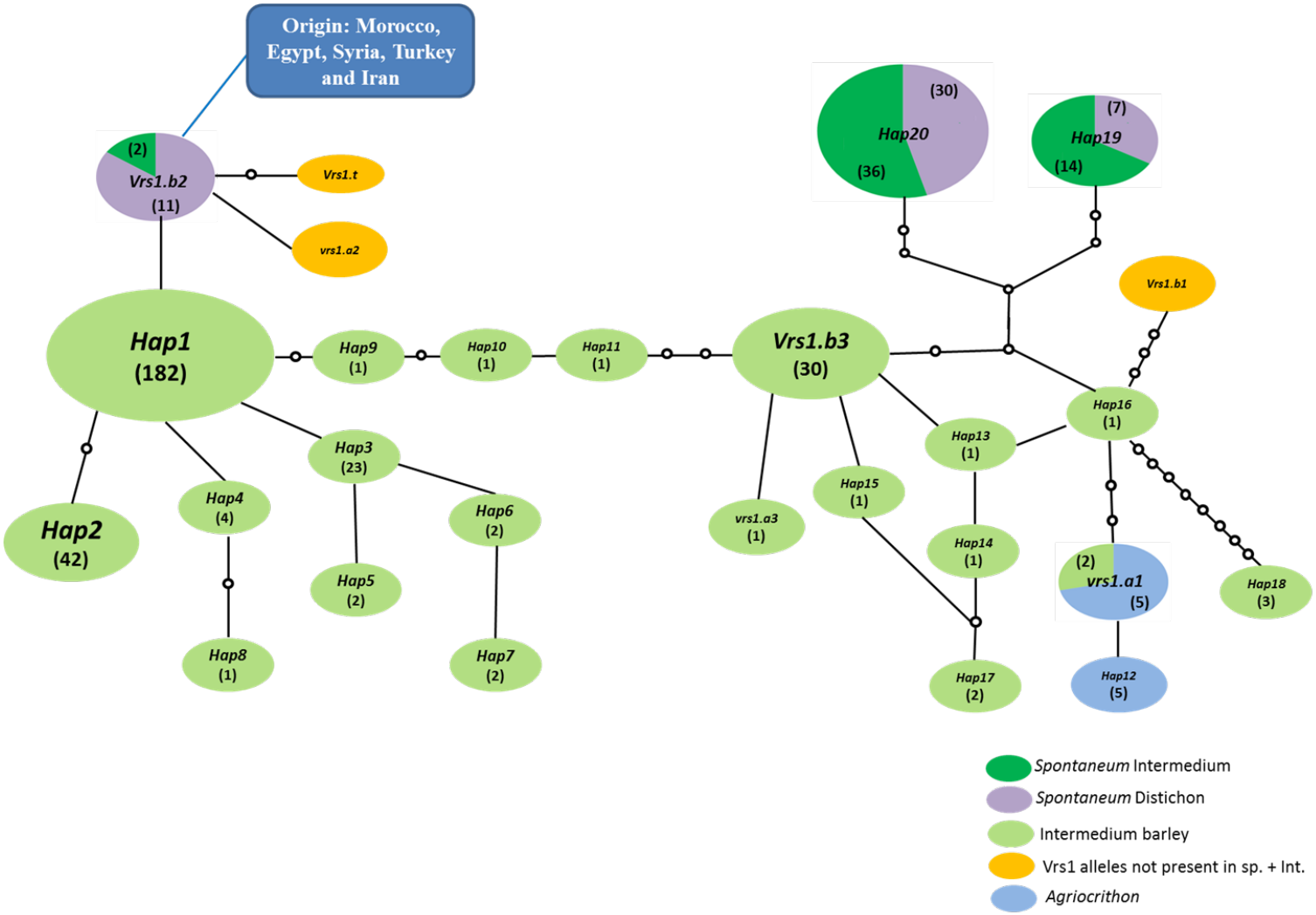


Fig. 4-4: Biological network analysis of new 20 haplotypes and seven known alleles at *vrs1* locus in 302 accessions of Intermedium collection, 100 accessions of *spontaneum* and 10 accessions of *agriocrithon*. sp. (*spontaneum*), Int. (Intermedium).

4.4.3 Genotypic status at the *int-c* locus in 'Intermedium' and wild barleys

Sequencing data of the *int-c* locus (**sup. Table 4-3**) revealed that all wild barley and all 'Intermedium' accessions (100%) have one or another of the known alleles and haplotypes at *int-c* reported by Ramsay et al. (2011) and Youssef et al. (2012). Of the 'Intermedium' barley collection studied, 291 accessions (96.4%) carried the six-rowed allele (*Int-c.a*) and only four accessions (HOR3825, HOR10281, HOR14233 and HOR17788) possessed the two-rowed allele (*int-c.b1*). The remaining accessions: HOR5327, HOR5733, HOR5912, HOR6407, HOR6408, HOR9563, HOR10701 carried the *Int-c.a haplotype2* which showed a highly consistent putatively heterozygous position at the *Int-c.a* allele at 437bp. The wild barley ssp. *spontaneum* (100 accessions) sequences data revealed that 30 accessions (30%) carried the six-rowed *Int-c* allele (*Int-c.a*) and a further 22 accessions (22%) the six-rowed *Int-c* allele *HvTB1*.spontc. The other accessions, however, possessed two-rowed alleles *int-c.b1* (43 accessions) or *HvTB1*.spontd (five accessions: FT67, FT222, FT229, FT746 and FT670). These results are in accordance with Ramsay et al. (2011). The sequence data of the 10 wild barley accessions of ssp. *agriocrithon* showed that all carried the six-rowed allele (*Int-c.a*).

4.4.4 Genotypic status at the *vrs4* locus in 'Intermedium' and wild barleys

Sequencing of the *vrs4* locus in the wild barley and 'Intermedium' collections (**sup. Table 4-4**) revealed no new haplotypes when compared to Koppolu et al. (2013). All accessions carried the functional *Vrs4* haplotypes which is present in two- and six-rowed cultivated barleys. The majority of the 'Intermedium' (239 accessions, 78%), ssp. *spontaneum* (63 accessions, 63%) and ssp. *agriocrithon* (10 accessions, 100%) carried *Vrs4_Hap1* which is found in most of cultivated barley such as Morex (six-rowed) and Bonus (two-rowed). Some 48 accessions of the 'Intermedium' collection and the remaining 37 accessions of ssp.

spontaneum possessed *Vrs4.Hap_2*, which is also found in two- and six –rowed cultivated barley such as Bowman and Shimabara. Thirteen 'Intermedium' barley accessions carried the sequence of *Vrs4.Hap_3*, one (HOR6661) that of *Vrs4.Hap_4* and one (HOR13728) that of *Vrs4.Hap_5*.

4.4.5 Haplotype combination/Phenotype relationships

To better understand the phenotype status in the barley accessions studied , the *vrs1* and *int-c* loci, or their haplotypes, were compared with those of *Vrs1.b/int-c.b* or *Vrs1.b/ HvTB1_Hap_spontd* (normal two-rowed), *Vrs1.b/Int-c.a*, or *Vrs1.b2/HvTB1_Hap_spontc* 'Intermedium' and *vrs1.a/Int-c.a* (complete six-rowed) haplotypes. As shown in **Table 4-2**, among 100 accessions of wild barley ssp. *spontaneum*, 52 accessions carried the haplotypes for the 'Intermedium' phenotype. The phenotypic data of lateral spikelet length and width (**Fig. 4-2**) in these accessions clearly showed that lateral spikelets were developed and enlarged but sterile (**Fig. 4-1a**). The other 48 accessions showed normal two-rowed phenotype (**Fig. 4-1b**) and carried *Vrs1.b* (or its haplotypes; *Hap.19 or Hap.20)/int-c.b* or *Vrs1.b/ HvTB1_Hap_spontd*). All accessions from ssp. *agriocrithon* had *vrs1.a1* (or its haplotype; *Hap12*)*/Int-c.a1* and showed complete six-rowed phenotype (**Fig. 4-1c**). Interestingly, we found that in the 'Intermedium' collection, 26 accessions had *Vrs1.b3/Int-c.a* and 257 accessions carried their haplotypes (*Hap_1* to *Hap_18* except *Hap_12 and 14* at *vrs1* locus). In contrast to all wild barley ssp. *spontaneum*, almost all lateral spikelets of the 'Intermedium' accessions were developed and setting (small) seeds. The accessions with *vrs3* phenotype showed *Vrs1.b2_Hap1 / Int-c.a1* haplotype. From the 'Intermedium' collection, four accessions, two carrying *vrs1.a1 / Int-c.a1*, one with *vrs1.a3 / Int-c.a1* and one with *vrs1.a1_Hap14 / Int-c.a_Hap2*, showed a six-rowed phenotype.

Table 4-2: *Vrs1* and *Int-c* allele combinations in all studied wild barley (ssp *spontaneum* and *agriocrithon*) and Intermedium collections and their phenotype.

	Wild barley		Inter.barley	
Haplotypes	***Agrio.***	***Spont.***	**Inter.**	**Phenotype**
Vrs1.b2 / Int-c.a1	-	1	-	intermedium *spontaneum*
Vrs1.b2 /HvTB1_Hap_spontc	-	1	-	intermedium *spontaneum*
Vrs1.b2 / int-c.b1	-	10	-	Two-rowed *spontaneum*
Vrs1.b2 /HvTB1_Hap_spontd	-	1	-	Two-rowed *spontaneum*
Vrs1.b3_Hap19 / HvTB1_Hap_spontc	-	8	-	intermedium *spontaneum*
Vrs1.b3_Hap19 / Int-c.a1	-	6	-	intermedium *spontaneum*
Vrs1.b3_Hap19 / int-c.b1	-	7	-	Two-rowed *spontaneum*
Vrs1.b3_Hap20 / int-c.b1	-	26	-	Two-rowed *spontaneum*
Vrs1.b3_Hap20 / Int-c.a1	-	23	-	intermedium *spontaneum*
Vrs1.b3_Hap20 / HvTB1_Hap_spontc	-	13	-	intermedium *spontaneum*
Vrs1.b3_Hap20 / HvTB1_Hap_spontd	-	4	-	Two-rowed *spontaneum*
Vrs1.b2_Hap1 / Int-c.a1	-	-	181	Intermedium
Vrs1.b2_Hap1 / Int-c.a_Hap2	-	-	1	Intermedium
vrs1.a1_Hap12 / Int-c.a1	5	-	-	Six-rowed
Vrs1.b2_Hap3 / Int-c.a1	-	-	23	Intermedium
Vrs1.b2_Hap4 / Int-c.a1	-	-	4	Intermedium
Vrs1.b2_Hap5 / Int-c.a1	-	-	2	Intermedium
Vrs1.b2_Hap6 / Int-c.a1	-	-	2	Intermedium
Vrs1.b2_Hap2 / Int-c.a1	-	-	39	Intermedium
Vrs1.b2_Hap2 / Int-c.a1_Hap1	-	-	1	Intermedium
Vrs1.b2_Hap2 / Int-c.a1_Hap2	-	-	1	Intermedium
Vrs1.b2_Hap7 / Int-c.a1	-	-	3	Intermedium
Vrs1.b2_Hap8 / Int-c.a1	-	-	1	Intermedium
Vrs1.b2_Hap9 / Int-c.a1	-	-	1	Intermedium
Vrs1.b2_Hap10 / Int-c.a1	-	-	1	Intermedium
Vrs1.b2_Hap11 / Int-c.a1	-	-	1	Intermedium
Vrs1.b3_Hap13 / Int-c.a_Hap2	-	-	1	Intermedium
vrs1.a1_Hap14 / Int-c.a_Hap2	-	-	1	Six-rowed
Vrs1.b3_Hap15 / Int-c.a	-	-	1	Intermedium
Vrs1.b3_Hap16 / Int-c.a1	-	-	1	Intermedium
Vrs1.b3_Hap17 / Int-c.a1	-	-	1	Intermedium
Vrs1.b3_Hap17 / int-c.b1	-	-	1	Two-rowed
Vrs1.b3_Hap18 / int-c.b1	-	-	1	Two-rowed
Vrs1.b3_Hap18 / Int-c.a1			2	Intermedium
vrs1.a1 / Int-c.a1	5	-	2	Six-rowed
vrs1.a1 / int-c.a_Hap2	-	-	1	*labile*
vrs1.a3 / Int-c.a1	-	-	1	Six-rowed
Vrs1.b3 / Int-c.a1	-	-	26	Intermedium
Vrs1.b3 / int-c.b1	-	-	2	Two-rowed
Vrs1.b3 / Int-c.a_Hap2	-	-	2	Intermedium

Another four accessions carrying *Vrs1.b3* or its haplotypes *(Hap-17* or *Hap_18)/ int-c.b1* haplotypes showed a normal two-rowed phenotype. We presume that these last eight accessions with either two- or six-rowed phenotype may have been misclassified as 'Intermedium'.

4.4.6 qRT PCR analysis of *spontaneum* and Intermedium accessions

Results from *Vrs1* and *Int-c* transcript analysis in immature spike meristems at the awn primordium stage of four ssp. *spontaneum* accessions and six selected accessions from the Intermedium collection, the latter with clear variation for lateral spikelet fertility ranging from enlarged to complete six-rowed phenotype (see **Fig. 4-3**), showed that both genes were expressed in all accessions. *Vrs1*was more highly expressed in all accessions compared with *Int-c* (**Fig. 4-5**). In most of the tested Intermedium accessions (HOR7076, HOR7078, HOR804 and HOR7211; **Fig. 4-5**) *Vrs1* expression was clearly lower compared with ssp. *spontaneum* accessions, clearly in line with enlarged and setting seeds in the lateral spikelets in the Intermedium accessions. The tested accessions FT11 and FT144, which classify as ssp. *spontaneum* 'Intermedium', showed slightly lower expression levels of *Int-c* compared to FT237 and FT332 (i.e. ssp. *spontaneum* 'Distichon'; **Fig. 4-5**). While, three accessions from the Intermedium collection (HOR7078, HOR804 and HOR13728; **Fig. 4-5**) showed slightly lower expression of *Int-c*. This lower expression of *Int-c* in these accessions explains the lateral spikelets phenotype; enlarged in ssp *spontaneum* 'Intermedium' and enlarged and setting seeds in the Intermedium collection. The lower expression of *Vrs1* in the Intermedium collection explains partially the increase of lateral spikelet fertility in this collection, which was not the case in wild barley ssp. *spontaneum*. The lateral spikelet fertility phenotype of Intermedium accessions HOR4838 and HOR13728 which showed high *Vrs1* expression (**Fig. 4-5**) suggested that other unknown gene(s) playing the main role of spikelet fertility in this collection.

4.5 Discussion

Based on the fertility of the two lateral spikelets at each rachis node, wild and cultivated barleys are grouped into two- (unfertile lateral spikelets) or six-rowed (fertile lateral spikelets) forms. The spikes of wild barley are of the two-rowed type (*H. vulgare* ssp. *spontaneum*) (Saisho et al. 2009); whereas the six-rowed form is represented by ssp. *agriocrithon*. 'Intermedium' barleys show an intermediate spike phenotype between two- and six-rowed condition in which lateral spikelets may be developed and are enlarged but without seed setting, yet in other cases setting small seeds and showing complete fertility of the spikelet triplet at each rachis node. By re-sequencing the *int-c* locus in 12 accessions of wild barley ssp. *spontaneum* Ramsay et al. (2011) identified that seven had the *int-c.b1* allele, one had an allele identical to *Int-c.a*, three had a novel allele denoted as *HvTB1.spontc* and one had a novel allele denoted as *HvTB1.spontd*. These observations posed an interesting question on the allele/haplotype structure of wild and 'Intermedium' barleys at the known row-type genes *Vrs1*, *Vrs4* and *Int-c*. To answer this question, we analyzed the allele and/or haplotype/phenotype structure of wild and 'Intermedium' barley at these loci.

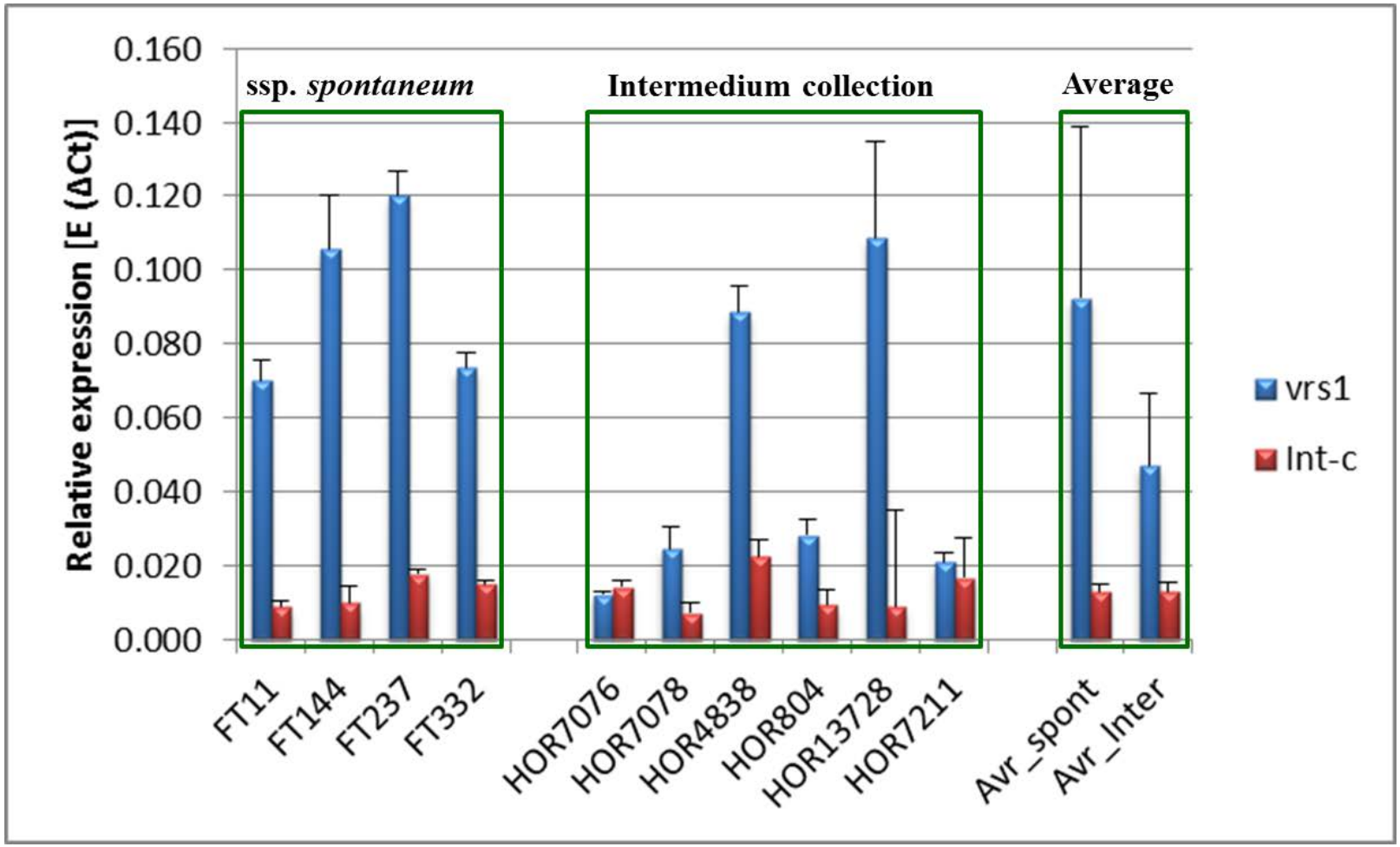


Fig. 4-5: Relative expression levels of *vrs1* and *Int-c* in wild barley ssp. *spontaneum* accessions (FT11, FT144, FT237 and FT332) and Intermedium accessions (HOR7076, HOR7078, HOR4838, HOR804, HOR13728 and HOR7211) as well as the averages of both groups. Relative expression levels of both genes shown are based on delta Ct calculation. Ct values for both genes were normalized to the Ct value of the *HvActin* included in each PCR run. The values were assessed from two biological replicates each of them contains 8 different spike at the awn primordium stage. (Standard errors)

Sequence analysis of 100 wild barley ssp. *spontaneum*, 10 ssp. *agriocrithon* and 302 'Intermedium' accessions at the *vrs1* locus revealed the previously reported alleles *Vrs1.b2*, *Vrs1.b3* (two-rowed alleles), and *vrs1.a1*, *vrs1.a3* (six-rowed alleles) (Komatsuda et al. 2007) as well as 20 new haplotypes. We discovered that 17 accessions of ssp. *spontaneum* and 30 'Intermedium' accessions carried the two-rowed alleles *Vrs1.b2* and *Vrs1.b3*, respectively. The six-rowed alleles *vrs1.a1* and *vrs1.a3* were found in the ssp. *agriocrithon* and 3 'Intermedium' accessions (**sup. Table 4-2**). While the *vrs1.a1* allele is due to a premature stop codon resulting in a truncated and non-functional protein, the *vrs1.a3* has an amino acid substitution of phenylalanine to lysine in the conserved HD domain. These mutations in the *Vrs1* gene presented in these two alleles cause the six-rowed phenotype in ssp. *agriocrithon* and the three 'Intermedium' accessions. All wild and 'Intermedium' barleys tested in our study carried functional *Vrs4* haplotypes (*Hap_1 - 5*). Koppolu et al. (2013) identified *Vrs4* as an important regulator of *Vrs1* in the barley row–type pathway. The two-rowed alleles *Vrs1.b2* and *Vrs1.b3* repress lateral spikelet development in two-rowed types. These two-rowed alleles and their new haplotypes identified here are found in all one hundred ssp. *spontaneum* accessions and 98.3% of the 'Intermedium' accessions, while all ten ssp. *agriocrithon* accessions carry six-rowed alleles.

Ramsay et al. (2011) observed that while two-rowed barleys carried the *int-c.b* allele, almost all six-rowed barleys had the alternative *Int-c.a* allele, irrespective of their particular *vrs1.a* (six-rowed) allele. Also, Cuesta-Marcos et al. (2010) suggested that the *Int-c.a* allele is necessary for the commercial utilization of the six-rowed phenotype; and that the repressing *int-c.b* allele is predominantly present in cultivated two-rowed germplasm. In our study, re-sequencing of *int-c* in a large set of wild and Intermedium barleys revealed all previously reported alleles/haplotypes (Ramsay et al. 2011 and Youssef et al. 2012) and did not show new haplotypes at the *int-c* locus. All ten ssp. *agriocrithon* accessions, 98.7% of 'Intermedium' and 52% of the ssp. *spontaneum* accessions showed a preference for a six-

rowed allele at the *int*-c locus (*Int-c.a*) or its haplotypes. The presence of *int-c.b* in six-rowed cultivars (*vrs1.a*/*int-c.b*) results in the development of smaller lateral spikelets (Lundqvist et al. 1997). This allele combination was not present in the accessions tested in this study. However, Ramsay et al. (2011) observed that the *int-c.b* allele in normal two-rowed (*Vrs1.b*/*int-c.b*) barley represses lateral spikelet development resulting in the two-rowed phenotype. In contrast, *Int-c.a* in two-rowed cultivars (*Vrs1.b*/*Int-c.a*) is associated with partially male fertile lateral spikelets. This intermediate state between two- and six-rowed forms is characteristic of the Intermedium phenotype and by far the most frequent allele combination that we detected in our Intermedium collection.

In the present study the haplotype combinations observed at the *vrs1* and *int-c* loci in wild barley ssp. *spontaneum* indicated that 48 accessions carried the ssp. *spontaneum* 'Distichon' allele combination (i.e. *Vrs1.b*/*int-c.b*) or their corresponding haplotypes, while the other 52 accessions had the *spontaneum* 'Intermedium' allele combination *Vrs1.b*/*Int-c.a*. These results lead to the novel view that wild barley ssp. *spontaneum* comprises two main row-type groups: ssp. *spontaneum* 'Intermedium', i.e. *Vrs1.b*/*Int-c.a*, possibly resulting from a natural mutation of *int-c.b* in the ssp. *spontaneum* 'Distichon' group (i.e. *Vrs1.b*/*int-c.b*). We hypothesize that the *spontaneum* 'Intermedium' group (*Vrs1.b*/*Int-c.a*) may be an intermediate form between the standard six-rowed cultivars (*vrs1.a*/*Int-c.a*) and the wild ancestor of barley ssp. *spontaneum* 'Distichon' (*Vrs1.b*/*int-c.b*). This hypothesis is in agreement with Nevski (1941) who considered that the lateral spikelets of *H. vulgare* ssp. *spontaneum* were too small to develop into forms with the large lateral spikelets of cultivated six-rowed barley. Nevski postulated that the common ancestor of barley had intermediately developed lateral spikelets and a fragile rachis. Furthermore, the normal two-rowed cultivated barley (*Vrs1.b*/*int-c.b*) may be directly derived from the wild barley *spontaneum* 'Distichon' group through human selection for the non-brittle rachis character (**Fig 4-6**).

The six-rowed wild barley ssp. *agriocrithon* always carried the six-rowed allele combination (*vrs1.a*/*Int-c.a*) at *vrs1* and *int*-c loci. Interestingly, we found that all 10 *agriocrithon* accessions tested in this study originated from China and were considered as Tibetan six-rowed wild barley. Feng et al (2006) proposed that the Tibetan six-rowed wild barley is an intermediate form in the transformation from two-rowed wild barley to cultivated barley. In agreement with this we hypothesize that six-rowed wild barley ssp. *agriocrithon* (i.e. *vrs1.a*/*Int-c.a*) might be directly converted from a ssp. *spontaneum* 'Intermedium' form through a single natural mutation of *Vrs1.b* to *vrs1.a*, or alternatively but less likely, from the wild ancestor *spontaneum* 'Distichon' (i.e. *Vrs1.b*/*int-c.b*) through mutations at *Vrs1* and *Int-c*. As such, ssp. *agriocrithon* can be considered as an intermediate form of cultivated six-rowed barleys in Central and East Asia. However, it cannot be completely excluded that ssp. *agriocrithon* may also be the result of a crossing between cultivated six-rowed and *spontaneum* 'Intermedium' types (**Fig 4-6**).

Allele combinations at *vrs1* and *int-c* loci (**Table 4-2**) in the Intermedium collection showed a high preference for two-rowed haplotypes over six-rowed alleles/haplotypes at the *vrs1* locus, and a clear preference of the six-rowed allele (*Int-c.a1*) or its haplotypes over two-rowed alleles/haplotypes at the *int-c* locus. In total, 97% of the Intermedium accessions displayed Intermedium haplotype combinations. In the same time the network analysis at *vrs1* showed that the haplotypes are predominantly derived from the *Vrs1.b2* allele, *Hap_19* and *Hap_20*. These results guided us to hypothesize that this collection was either derived directly from *spontaneum* 'Intermedium' by human selections for the non-brittle rachis character, or was derived from crossing between cultivated two- and six-rowed barleys. Ramsay et al. (2011) noted that the presence of *Int-c.a* in two-rowed cultivars, which carry a functional *Vrs1* allele, causes enlarged, partially male fertile, lateral spikelets. However, in the present study the majority of the accessions displaying an Intermedium row-type, set seeds along the spike similar as in six-rowed forms but with lateral seeds smaller than the central ones. However,

this was clearly not the case in all accessions of wild barley ssp. *spontaneum* 'Intermedium' carrying the identical allele combination at *Vrs1* and *Int-c*, clearly suggesting that detected allelic variation at *Int-c* is not sufficient to provide any seed set in wild barley ssp. *spontaneum* accessions.

The *Vrs1* gene belongs to the HD-ZIP I class of homeobox transcription factors. Low expression of *VRS1* in lateral spikelet primordia enabled complete fertility, suggesting that VRS1 protein suppresses the development of lateral spikelets in barley (Komatsuda et al 2007). The qPCR analysis of *Vrs1* and *Int*-c in *spontaneum* and Intermedium collection accessions showed that both genes are expressed in *spontaneum* as well as in Intermedium accessions. In Intermedium accessions the expression of *Vrs1* was lower, but not significantly, compared to *spontaneum* accessions, which suggests that the increasing lateral spikelet fertility in Intermedium accessions could (partially) be a result of this lower expression of *Vrs1*. *Int-c* expression clearly showed low expression level in all tested accessions either in ssp. *spontaneum* (lateral sterile) or Intermedium collection (lateral fertility varied from enlarged not setting seeds to complete six-rowed phenotype). In agreement with Lundqvist and Lundqvist (1988) the low expression of *Int-c* in all tested accessions suggests that the phenotypic effect of *Vrs1* in the Intermedium collection might be influenced by one or more of the other ten independent *Intermedium* (*int*) genes distributed all over the barley genome. Also, increased fertility in 'Intermedium' barleys appears as a result of naturally derived mutations at other fertility repressing loci. One of which could be for example the *vrs3* locus; however, the low number of detected *vrs3*-like phenotypes (11 accessions) among the 'Intermedium' accessions cannot fully explain the seen phenotypic diversity for increased lateral spikelet fertility.

Taken together, obtained results rather point to the presence of other yet unknown genes playing a major role in lateral spikelet fertility and seed setting either directly or

through regulating other known or unknown genes in ‘Intermedium’ barleys (Lundqvist and Lundqvist 1988). More future work is required to fully elucidate the molecular basis of spikelet/floret fertility and seed setting in ‘Intermedium’ barleys.

4.6 Acknowledgments:

The authors gratefully acknowledge careful maintenance and supply of the Intermedium barley accessions from IPK gene bank, Gatersleben, Germany. We are grateful to Mr. Michael Grau for providing the Gerste_Schlüssel file which helped understanding the ‘Intermedium’ phenotype and classification at the IPK gene bank. We thank Mrs. H. Ibrahim, S. König and I. Walde for excellent technical support. We kindly acknowledge Robbie Waugh (JHI, Dundee) and Gary Muehlbauer (UMN, St. Paul) for providing pre-publication access to exome capture re-sequencing data.

Conflict of Interest

The authors declare that they have no conflict of interest.

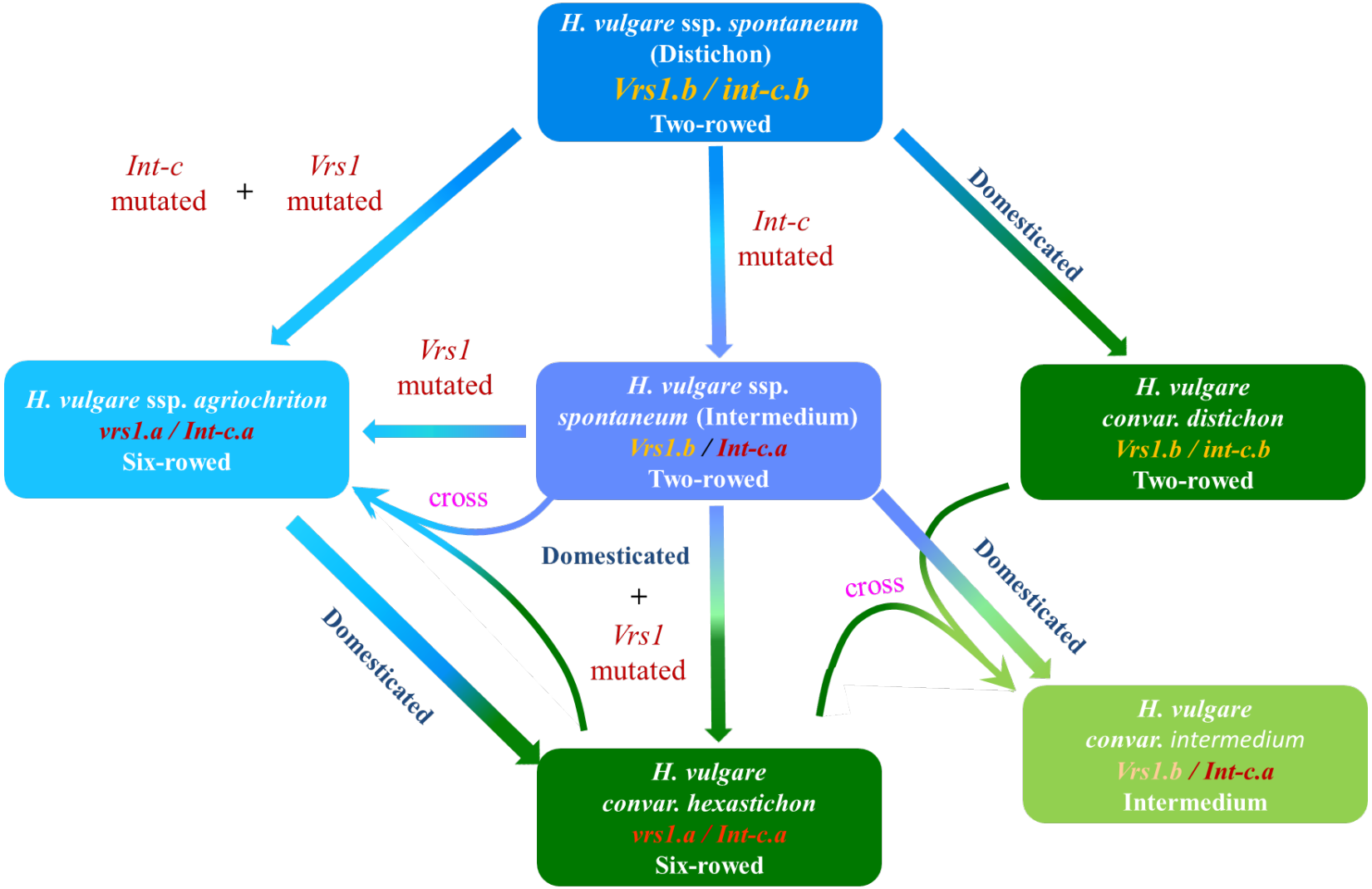


Fig. 4-6: Model of the row-type relationship for the wild and domesticated barleys based on their *Vrs1* and *Int-c* genotypes.

CHAPTER FIVE

General Discussion

5.0 General discussion

Based on the fertility of the two lateral spikelets at each rachis node, wild and cultivated barleys are traditionally separated into two-rowed (having unfertile lateral spikelets) or six-rowed (having fertile lateral spikelets) forms. The spikes of *H. vulgare* ssp. *spontaneum*, which is the ancestor of all domesticated forms, are monomorphic for the two-rowed phenotype (Saisho et al. 2009), whereas the six-rowed form is represented by ssp. *agriocrithon* (six-rowed wild barley). A special case forms, the *labile* and Intermedium barleys which show an intermediate spike phenotype between two-rowed and six-rowed condition. In these barleys the lateral spikelets may develop and are enlarged without setting seeds yet in other cases setting small seeds as in the Intermedium collection or missing at all as in the *labile* collection. Representatives of the Intermedium accessions show a consistent spike form in all plant tillers. In *labile* accessions, however, lateral spikelet fertility can vary from spike to spike on a single plant, and even within a single spike, and from plant to plant within one progeny (Djalali et al. 1970; Takeda and Saito 1988*)*. The development of spike phenotype in both *labile* and Intermedium barleys is very intriguing. Clearly, the relevance of elucidating the genetic background of lateral spikelet fertility will go beyond the barley species itself as it can be expected to also contribute to a better general understanding of spikelet development in related grass species including important crop plants such as wheat.

In the second and fourth chapters of this thesis I describe the allele/haplotype structure of *labile* and Intermedium barleys at the known row-type genes *Vrs1* and *Int-c* in addition to *Vrs4* in Intermedium barley. This is followed by (in the third chapter) generating mapping populations to locate the *lab* locus on the barley chromosome arm using the whole genome mapping techniques.

Komatsuda et al. (2007) and Ramsay et al. (2011) provided new insights in the control of lateral spikelet fertility in barley. These authors showed that when loss-of-function of the

Vrs1 gene (i.e. *vrs1.a*) is complemented by the *Int-c.a* allele, complete fertility of lateral spikelets is achieved resulting in a six-rowed phenotype. About the same time Cuesta-Marcos et al. (2010*)* reported that in two-rowed barley the *Vrs1.b* allele is complemented by the *int-c.b* allele. Thus, lack of the VRS1 protein in lateral spikelet primordia enabled complete fertility, suggesting that VRS1 suppresses the development of lateral spikelets in barley.

5.1 Labile haplotypes at *vrs1* and *int-c* loci

By re-sequencing of the known row-type loci in *labile* it was found that every single *labile* accessions carried six-rowed alleles at *vrs1* and at *Int-c*. For *Vrs1* the results split up in 99,08% of accessions containing *vrs1.a1* and 0.92% *vrs1.a3*, while in case of *Int-c* 53,6% of accessions contained *Int-c.a*, 0,5% *Int-c.a_Hap1* and 45,9% *Int-c.a/Int-c.a_Hap1*. This is strong evidence that *labile*-barleys have a six-rowed genetic background. The *vrs1.a1* allele, which is found in six-rowed barleys from all over the world, and the *vrs1.a2* allele, which is characteristic for barleys from Western Mediterranean regions are due to a premature stop codon resulting in a truncated and non-functional protein. In contrast to this, the *vrs1.a3* allele, which is only found in the regions of East Asia, has an amino acid substitution from phenylalanine to lycine in the conserved HD domain. The present study lead to the discovery of two novel *vrs1.a1 haplotypes, Hap_66* and *Hap_67*. Whereas *Hap_67* is caused by an amino acid substitution (aspartic acid to valine) within the conserved HD domain and has been observed in four of the *labile* accessions, *Hap_66* is due to an SNP in the promoter region and found in nine accessions. These newly identified haplotypes of *vrs1.a1* may be specific to the highlands of Ethiopia and Eretria, which are also the cradle of *labile*-barleys. However, this hypothesis remains tentative since the number of accessions (nine for *Hap_66* and four for *Hap_67*) carrying these alleles is very limited. This might indicate that the evolutionary events leading to these particular haplotypes were rather recent. Further studies with a larger set of *labile* and non-*labile* accessions may provide a deeper insight into the rise

of these two novel haplotypes. An even larger shroud of uncertainty surrounds the *Int-c.a* haplotype (*Hap_1*). This haplotype is due to an amino acid substitution in the conserved TCP domain. It occurred in only one of the accessions and here it was found in the homozygous state. At this stage it is too early to conclude that it is specific to *labile*-barleys.

A remarkable observation was that 101 of the accessions tested contained both the *Int-c.a* and the *Hap_1* allele. This suggests that heterozygosity had been maintained at the *int-c* locus, which is puzzling, given that barley is a highly self-fertilizing species. Hence, another possibility might be that the *int-c* locus has been duplicated in these 101 accessions and that both *int-c* genes differ by only one SNP. A final answer to this question will have to await further work. Although the significance of this allelic combination is not yet clear, the phenotypical data showed that the lateral spikelet fertility within these 101 apparently heterozygous *Int-c.a/Hap_1* accessions did not significantly differ from that found within the 118 homozygous *Int-c.a* accessions. Similar results were obtained when the phenotypic analysis was extended to all allele/haplotypes combinations found in *labile.*

From previous work it was assumed that the *labile* locus interacts with the recessive six-rowed *vrs1* locus to produce irregular spikes with variable abortion of lateral spikelets (Lundqvist and Franckowiak 2003). After crossing *labile*-barley with two-rowed barleys Djalali et al. (1970) reported a recessive mode of gene action for *labile* in F_1 and F_2 plants .

5.2 Labile (lab) gene mapping

To better understand the genetic background of the *labile* phenotype *labile*-barleys were crossed with six-rowed types. Using two mapping populations, we were the first to show that the *lab* locus can be mapped as a distinct Mendelian trait located on the long arm of chromosome 5H at approximately 80 cM to an interval of 4.5–5.8 cM. The segregating F_2 populations showed either the *labile* or the six-rowed spike phenotype. This segregation pattern (1:3) in conjunction with reports by Djalali (1970) confirmed that the reliable

detection of the *labile* character is primarily dependent upon two conditions: (i) the presence of a six-rowed spike phenotype (i.e. genetically constituted as *vrs1.a* + *Int-c.a*), and (ii) the presence of a recessive allele for the *labile* character at the *lab* locus. Furthermore, we show that the newly identified *lab* locus is genetically distinct from all previously identified loci conferring the six-rowed spike phenotype in barley (i.e. *vrs1*, *vrs2*, *vrs3*, *vrs4* and *int-c* on chromosomes 2HL, 5HL, 1HS, 3HS and 4HS, respectively; Pourkheirandish and Komatsuda 2007). Among all row-type mutants, the *labile*-barleys are unique in their spike architecture; i.e. irregular spikelet fertility. The analysis of scanning electron micrographs revealed that up until the lemma primordium stage there were no visible differences in spike development between two-rowed, six-rowed and *labile* spikes. During these early stages, the central floral primordia were always further developed than those of the lateral floral primordia (see also Komatsuda et al. 2007). It was not before late stamen and early awn primordium that first morphological differences in overall spike development became apparent. In the case of six-rowed barleys, central and lateral spikelets/florets displayed complete development, whereas two-rowed barleys consistently displayed a clearly retarded development in all lateral florets. In contrast to two-rowed barleys, infertile lateral florets of *labile*-barleys are completely reduced without any floral development except the glumes, thus rather resembling the *deficiens* phenotype (Mansfeld 1950). An additional remarkable feature of *labile*-barleys is that occasionally also the central spikelets remain reduced to glumes only without any floral meristem development. This unique spikelet fertility phenotype only observed in *labile*-barleys clearly suggests that the *labile* gene is important for floral meristem identity and development. This stresses the importance for the identification of the gene underlying the *lab* locus. The present work underlines that a combination of molecular-genetic and phenotypical analysis of lateral and central spikelet development in *labile*- and other row-types is a promising approach to elucidate the role of the *lab* gene in relation to other floret development and fertility factors in barley. Taken all

observations together, we found that the awn primordium stage is the period when genes controlling the number of spikelet/floret per spike maximal (Kirby and Appleyard 1987). Future work studying hormone contents and gene expression at specific stages may help to better understand which genetic factors are controlling final yield per spike in barley and other grasses. In case of *labile* barley, identification of the *lab* gene and studying its function will primarily help understanding floret development and fertility in all spikelets along the spike.

5.3 Intermedium and wild barleys haplotypes at *vrs1*, *vrs4* and *int-c* loci

The spikes of wild barley (*H. vulgare* ssp. *spontaneum*), the recent ancestor of the domesticated form, are monomorphic for the two-rowed type (Saisho et al. 2009), whereas the six-rowed form is represented by wild ssp. *agriocrithon*. By re-sequencing the *int-c* locus in 12 accessions of wild barley ssp. *spontaneum* Ramsay et al. (2011) identified that seven had the *int-c.b1* allele, one had an allele identical to *Int-c.a*, three had a novel allele denoted as *HvTB1.spontc* and one had a novel allele denoted as *HvTB1.spontd*. These observations threw up an interesting question on the allele/haplotype structure of wild and 'Intermedium' barleys at the known row-type genes *Vrs1*, *Vrs4* and *Int-c*. To answer this question, we analyzed the allele and/or haplotype / phenotype structure of wild and 'Intermedium' barley at these loci. Sequence analysis of 100 wild barley ssp. *spontaneum*, 10 ssp. *agriocrithon* and 302 'Intermedium' accessions at the *vrs1* locus revealed the previously reported two-rowed alleles *Vrs1.b2*, *Vrs1.b3*, and the six-rowed alleles *vrs1.a1*, *vrs1.a3* (Komatsuda et al. 2007) as well as 20 new haplotypes. We discovered that 13 accessions of ssp. *spontaneum* and 30 accessions of Intermedium carried the two-rowed alleles *Vrs1.b2* and *Vrs1.b3*, respectively. The six-rowed alleles *vrs1.a1* and *vrs1.a3* were found in the ssp. *agriocrithon* and in three Intermedium accessions. These mutations in the *Vrs1* gene presented in these two alleles are responsible for the six-rowed phenotype in ssp. *agriocrithon* and the three Intermedium

accessions. All wild and Intermedium barley tested carried functional *vrs4* haplotypes (*Hap_1 - 5*). In this study, re-sequencing of *int-c* in a large set of wild and Intermedium barleys brought to light all previously reported alleles/haplotypes (Ramsay et al. 2011 and Youssef et al. 2012) and did not show any new haplotypes at *int-c* locus. All accessions of ssp. *agriocrithon* and 98.7% of Intermedium as well as 52% of those of ssp. *spontaneum* showed a preference for a six-rowed allele at *Int*-c (*Int-c.a*) or its haplotypes. The haplotype combinations observed at the *vrs1* and *int-c* loci in wild barley ssp. *spontaneum* indicated that 48 accessions carried the ssp. *spontaneum* 'Distichon' allele combination (i.e. *Vrs1.b*/*int-c.b*) or their corresponding haplotypes, while the other 52 accessions had the *spontaneum* 'Intermedium' allele combination *Vrs1.b*/*Int-c.a*.

5.4 Wild and cultivated barley relationship.

From these results we can conclude that wild barley ssp. *spontaneum* comprises two main row-type groups: ssp. *spontaneum* 'Intermedium', i.e. *Vrs1.b*/*Int-c.a*, possibly resulting from a natural mutation of *int-c.b* in the ssp. *spontaneum* 'Distichon' group (i.e. *Vrs1.b*/*int-c.b*). We hypothesize that the *spontaneum* 'Intermedium' group (*Vrs1.b*/*Int-c.a*) is an intermediate form between the standard six-rowed cultivars (*vrs1.a*/*Int-c.a*) and the wild ancestor of barley ssp. *spontaneum* 'Distichon' (*Vrs1.b*/*int-c.b*). This hypothesis is in agreement with Nevski (1941) who considered that the lateral spikelets of *H. vulgare* ssp. *spontaneum* were too small to develop into forms with the large lateral spikelets of cultivated six-rowed barley. Nevski postulated that the common ancestor of barley had intermediately developed lateral spikelets and a fragile rachis. Furthermore, the normal two-rowed cultivated barley (*Vrs1.b*/*int-c.b*) may be directly derived from the wild barley *spontaneum* 'Distichon' group through human selection for the non-brittle rachis character.

The six-rowed wild barley ssp. *agriocrithon* always carried the six-rowed allele combination (*vrs1.a*/*Int-c.a*) at *vrs1* and *int*-c loci. Interestingly, we found that all 10

agriocrithon accessions tested in this study originated from China and were considered as Tibetan six-rowed wild barley. Feng et al (2006) proposed that the Tibetan six-rowed wild barley is an intermediate form in the transformation from two-rowed wild barley to cultivated barley. In agreement with this we hypothesize that ssp. *agriocrithon* (i.e. *vrs1.a*/*Int-c.a*) might be directly converted from a *spontaneum* 'Intermedium' form through a single natural mutation of *Vrs1.b* to *vrs1.a*, or alternatively, from the wild ancestor *spontaneum* 'Distichon' (i.e. *Vrs1.b*/*int-c.b*) through mutations at *Vrs1* and *Int-c*. As such, ssp. *agriocrithon* can be considered as an intermediate form of cultivated six-rowed barleys in Central and East Asia. However, it cannot be completely excluded that ssp. *agriocrithon* may also be the result of a crossing between cultivated six-rowed and *spontaneum* 'Intermedium' types. Studying more *agriocrithon* accessions collected from different areas may help in understanding of moving the wild barley from the main center of domestications (Fertile Crescent) to other centers such as Tibet and will also help in understanding human civilization (Badr and El-Shazly 2012). Allele combinations at *vrs1* and *int-c* loci in the Intermedium collection showed a high preference for two-rowed haplotypes over six-rowed alleles/haplotypes (opposite of *labile* collection) at the *vrs1* locus, and a clear preference of the six-rowed allele (*Int-c.a1*) or its haplotypes over two-rowed alleles/haplotypes at the *int-c* locus. These results in addition to the results from the biological network analysis at *vrs1* locus guide us to hypothesize that this collection was either derived directly from *spontaneum* 'Intermedium' by human selections for the non-brittle rachis character, or was derived from crossing between two- and six-rowed cultivated barleys. In total, 97% of the Intermedium accessions displayed Intermedium haplotype combinations. The majority of the accessions displaying an Intermedium row-type in this study, set seeds along the spike similar as in six-rowed forms but with lateral seeds smaller than the central ones. However, this was clearly not the case in all accessions of wild barley *spontaneum* 'Intermedium' carrying the identical allele combination at *vrs1* and *int-c*, clearly suggesting that detected allelic variation at *int-c* is not sufficient to provide any seed

set in wild barley ssp. *spontaneum* accessions. These results guided us to analyze the *vrs1* and *Int-c* expression in both; *spontaneum* and Intermedium accessions. The relative expression of *Vrs1* and *Int-c* genes showed that both genes are expressed in *spontaneum* as well as in Intermedium accessions**.** In Intermedium accessions the expression of *Vrs1* as expected was lower, but not significantly, compared to *spontaneum* accessions, which suggests that the increasing lateral spikelet fertility in Intermedium accessions could (partially) be a result of this lower expression of *Vrs1*. *Int-c* expression clearly showed low expression level in all tested accessions either in ssp. *spontaneum* or Intermedium collection (lateral spikelet fertility varies from enlarged not setting seeds to complete six-rowed phenotype). In agreement with Lundqvist and Lundqvist (1988) the low expression of *Int-c* in all tested accessions suggests that the phenotypic effect of *Vrs1* in the Intermedium collection might be influenced by one or more of the other ten independent *Intermedium* (*int*) genes distributed all over the barley genome. On the other hand, increased fertility in 'Intermedium' barleys could be appearing as a result of naturally derived mutations at other fertility repressing loci. Taking all in considerations, the obtained results rather point to the presence of other yet unknown genes playing a major role in lateral spikelet fertility and seed setting in 'Intermedium' and in *labile* barleys (Lundqvist and Lundqvist 1988). Theses unknown genes might be playing a direct role of lateral spikelet fertility or regulating/controlling the functions of other known or unknown spike architecture genes. Establishing F2 mapping population from crossings between BW-NIL Intermedium mutants such as *int.f*, *int.h*, *int.i*, *int.k*, *int.l* and *int.m* and two-rowed cultivar Bowman could be used for identifying those genes which may illuminate important spikelet/floret fertility factors for barley and other agronomically important crops such as wheat.

5.5 Concluding remarks

Barley is a valuable model crop to study the different mechanisms conferring spikelet fertility in cereals. Analyzing the genotypic status of the known row-type loci in different barley collections and combining this data with the phenotypic status allowed us to efficiently explore its genetic diversity of spikelet fertility. The approach applied in this thesis involved an extensive study of the haplotype/phenotype relations in a large number of wild, *labile* and Intermedium barleys. The overall picture that emerged suggested that besides the known row-type genes, there must be additional genes involved in the control of spikelet fertility. From the results a clear idea arose about the genotypic status of the *Vrs1* and *Int-c* in *labile*, Intermedium and wild barley collections and the role of these loci in spikelet fertility.

Special attention was paid to *labile* barley, which is considered an intermediate form between two- and six-rowed barley. Investigating *labile* accessions we managed to map the *lab* locus on barley chromosome 5HL. Subsequent studies confirmed that *lab* is controlling floret development and fertility during spike developmental stages. Another remarkable observation was that a group of wild barley ssp. *spontaneum* genotypically carried the Intermedium haplotype at *vrs1* and *Int-c* loci, yet displayed an Intermedium phenotype with enlarged lateral spikelets but not setting any seeds. These latter plants were named ssp. *spontaneum* (Intermedium). It was subsequently discovered that the Intermedium barley collection carried the same genotype/haplotype as the ssp. *spontaneum* "Intermedium" group yet in contrast to the latter showed lateral spikelet fertility, including setting of (small) seeds. This suggests that other additional genes are involved in spikelet fertility in this collection. In conclusion, a new model is introduced on the relation between wild and cultivated barley during domestication based on the re-sequencing of the *Vrs1*, *Vrs4* and *Int-c* genes. I believe that whole genome studies in such different barley collections will confirm the proposed

domestication model presented in this study and that we will thus be able to better understand barley domestication and evolution process.

6.0 Outlook

We still have only a fragmentary knowledge of the genetics underlying inflorescence development in barley. A deficient *Vrs1* gene function in combination with *Int-c* may explain the lateral spikelet fertility in Intermedium and *labile*-barleys. The same combination, however, is insufficient to clarify the random floral sterility seen in lateral and central spikelets in *labile* accessions and the increased fertility in Intermedium barley. Evidently additional genes must be involved in floral development and spike fertility. Looking for a possible candidate we turned our attention to the *lab* locus. To identify the underlying gene, map-based and mutant analysis approaches seems most promising. Molecular-genetic results in combination with the examination of detailed lateral and central spikelet development in *labile*- and other row-types may help elucidate the role of the *lab* gene in relation to other floret development and fertility factors in barley. To understand the genetic effect of the genes controlling spikelet fertility in barley, environmental effects have to be excluded by comparing *Vrs1, Vrs2, Vrs3, Vrs4, Int-c* and *labile* mutants grown in the field with those grown in the greenhouse, something we plan to do in the near future. We found that the awn primordium stage is the stage when the genes controlling the spikelet/floret number per spike showing the maximum number of spikelets at end of awn primordium stage. Future work to study the hormone contents and gene expression of those controlling hormones biosynthesis at these specific stages will help much to understand what is controlling the final yield in barley and other grasses.

The F3 plants from the mapping population (Morex × HOR2573) clearly segregate in traits such as plant height, leaf area, heading date, spike length, production of anthocyanin (only in the spike) and additional spikelet or branching from the central spikelet which I take in consideration and scored them preparing for identifying the loci controlling all of these traits using our VeraCode data for the F2 set of DNA. Studying all of these traits in the same

mapping population might be helpful to understand the relation among all of these traits and the spike development in barley.

Re-sequencing data of wild and Intermedium collections revealed that all of the Intermedium accessions carried *Vrs1.b*/Int-c.*a* and that setting seeds varied from one to three seeds at each rachis node. At the same time a group of wild barley ssp. *spontaneum* 'Intermedium' carrying the same haplotype *Vrs1.b*/*Int-c.a* had its lateral spikelets only enlarged without setting seeds. Crossing between these two genotypes could help to identify the responsible locus for setting seeds in the 'Intermedium' collection. Also, crossing between BW-NIL Intermedium mutants such as *int.f, int.h, int.i, int.k, int.l and int.m* and two-rowed cultivar Bowman to create F2 mapping population which could be used for identifying the genes underling these mutants will help in understanding the barley spike development.

7.0 Zusammenfassung

Gerste repräsentiert ein Modellsystem um Mechanismen der Ährchenfertilität bei Getreiden zu untersuchen. Durch genotypische Untersuchungen mithilfe von bekannten Ährenarchitekturgenen (‘row-type genes‘) in Kombination mit phänotypischen Untersuchungen in verschiedenen Gerstenkollektionen wurden neue Erkenntnisse über die genetische Diversität der Ährchenfertilität gewonnen. Daneben konnten durch das in dieser Studie gezeigte Haplo-/Phänotyp-Verhältnis in den getesteten Wild-, *labile*- und Intermedium-Kollektionen Hinweise erhalten werden, dass neben den bekannten Ährenarchitekturgenen auch andere Gene eine wesentliche Rolle bei der Ausprägung der Ärchenfertilität spielen. Die Ergebnisse zeigen Korrelationen zwischen dem genotypischen Status der *vrs1* und *int-c* Loci und den phänotypischen Ausprägungen in den verschiedenen Gerstenkollektionen, was neue Erkenntnisse über deren Rolle in der Fertilitätsausprägung lieferte. Neben diesen Loci konnte der *lab* Lokus auf dem Gerstenchromosom 5H (langer Arm) lokalisiert werden. Des Weiteren konnte gezeigt werden, dass das *lab*-Gen die Blütchenbildung und -fertilität während der Ährenentwicklung kontrolliert. Ein neues Resultat dieser Studie zeigt, dass eine Gruppe von Wild-Akzessionen (ssp. *Spontaneum*) genotypisch dem Intermedium Haplotyp an den *vrs1* und *int*-c Loci entspricht und dabei auch den entsprechenden Phänotyp mit vergrößerten seitlichen Ährchen, aber ohne zusätzliche Samenbildung zeigt. Diese Gruppe wird als ssp. *spontaneum* (Intermedium)“ bezeichnet. In den Intermedium-Akzessionen konnten aber auch Gruppen gefunden werden, die genotypisch spp. *spontaneum* (Intermedium)“ entsprechen, aber fertile Seitenährchen mit Samenbildung enthalten, woraus hervorgeht, dass zumindest in dieser Kollektion auch andere, unbekannte Gene eine Rolle in der Ährchenfertilität spielen.

Auch für die Domestikation von Gerste wurde eine neue Hypothese aus der Re-sequenzierung der *Vrs1*, *Vrs4* und *Int-c* Gene in Wild- und Kulturakzessionen abgeleitet. Dieses

Domestikationsmodell könnte durch genomweite Untersuchungen (GWAS) verifiziert werden. Dieses würde neuartige Einblicke in die Evolution und Domestikationsprozesse der Gerste liefern.

8.0 References

Abay F, Bjørnstad A (2009) Specific adaptation of barley varieties in different locations in Ethiopia. Euphytica 167:181–195

Åberg E (1938) Hordeum agriocrithon nova sp., a wild six-round barley. The Annals of the Agricultural College of Sweden 6: 159–216.

Åberg E (1940) The taxonomy and phylogeny of *Hordeum* L. sect. Cerealia Ands. With special reference to Thibetan barleys. Symbolae Bot. Upsaliensis 4: 1-156.

Åberg E, Wiebe GA (1945) Ash content of barley awns and kernels as influenced by location, season and variety. J Agronomy 37:583-586

Åberg E, Wiebe GA. (1945) Irregular barley, *Hordeum irregulare*, sp. nov. J Washington Academ Sci 35:161-164

Ariyadasa R, Mascher M, Nussbaumer T, Schulte D, Frenkel Z, Poursarebani N, Zhou R, Steuernagel B, Gundlach H, Taudien S, et al. (2014) A sequence-ready physical map of barley anchored genetically by two million single-nucleotide polymorphisms. Plant Physiol., 164 (2014), pp. 412‑423

Azhaguvel P, Komatsuda TA (2007) A phylogenetic analysis based on nucleotide sequence of a marker linked to the brittle rachis locus indicates a diphyletic origin of barley. Ann Bot 100: 1009–1015.

Badr A, El-Shazly H (2012) Molecular approaches to origin, ancestry and domestication history of crop plants: Barley and clover as examples. Journal of Genetic Engineering and Biotechnology 10:1-12.

Badr A, Mueller K, Schafer-Pregl R, El Rabey H, Effgen S, et al. (2000) On the origin and domestication of barley (Hordeum vulgare). Mol Biol Evol 17: 499–510.

Baik BK, Ullrich SE (2008) Barley for food: Characteristics, improvement, and renewed interest. Journal of Cereal Science 48:233-242

Baktheyev F (1963) Origin and phylogeny of barley. In: Barley Genetics I. Proc. 1st Int. Barley Genet. Symp., Wageningen, The Netherlands, pp. 1-18.

Barski A, Cuddapah S, Cui K, Roh TY, Schones DE, Wang Z, Wei G, Chepelev I, Zhao K (2007) High-resolution profiling of histone methylations in the human genome. Cell 129:823–837

Batley J, and Edwards D (2007) SNP applications in plants. p. 95–102. *In*Oraguzie, N.C. . (ed.) Association mapping in plants. Springer, New York, NY.

Bell GDH (1965) The comparative phylogeny of the temperate cereals. In: J.B. Hutchinson (ed), Essays on Crop Plant Evolution. Cambridge University Press, pp. 70-101.

Bennett MD, Smith JB. (1976) Nuclear DNA amounts in angiosperms. *Philosophical Transactions of the Royal Society of London* B 274: 227–274

Bernatsky R and Tanksley S (1986). Toward a saturated linkage map in tomato based on isozymes and random cDNA sequences. Genetics 112(4): 887-898.

Bjørnstad A, Abay F (2010) Multivariate patterns of diversity in Ethiopian barleys. Crop Sci 50:1579–1586

Blake T, Blake V, Bowman J, Abdel-Haleem H, (2011) *Barley: Production, Improvement and Uses* (ed. S. E. Ullrich) 522–531.

Bonnett OT (1935) A description of the development of the barley spike. Journal of Agricultural Research 51:451-457

Breitenfeld C (1957) Genetische Untersuchungen zur Phylogenie und Systematik der *intermedium*- und *labile*-Gersten. Z Pflanzenzüchtung 38:275-312

Ching A, Caldwell K, Jung M, Dolan M, Smith O, Tingey S, Morgante M, and Rafalski A (2002) SNP frequency, haplotype structure and linkage disequilibrium in elite maize inbred lines. BMC Genet. 3:19.

Close TJ, Bhat PR, Lonardi S, Wu YH, Rostoks N, Ramsay L, Druka A, Stein N, Svensson JT, Wanamaker S, Bozdag S, Roose ML, Moscou MJ, Chao SAM, Varshney RK, Szucs P, Sato K, Hayes PM, Matthews DE, Kleinhofs A, Muehlbauer GJ, DeYoung J, Marshall DF, Madishetty K, Fenton RD, Condamine P, Graner A, Waugh R (2009) Development and implementation of high-throughput SNP genotyping in barley. BMC Genomics 10:582 doi:10.1186/1471-2164-10-582

Cockram J, White J, Zuluaga DL, Smith D, Comadran J, Macaulay M, Luo Z, Kearsey MJ, Werner P, Harrap D, Tapsell C, Liu H, Hedley PE, Stein N, Schulte D, Steuernagel B, Marshall DF, Thomas WTB, Ramsay L, Mackay I, Balding DJ, Consortium TA, Waugh R, O'Sullivan DM (2010) Genome-wide association mapping to candidate polymorphism resolution in the unsequenced barley genome. Proceedings of the National Academy of Sciences 107:21611-21616

Colmsee C, Beier S, Himmelbach A, Schmutzer T, Stein N, Scholz U, Mascher M (2015) BARLEX‐the Barley Draft Genome Explorer. Molecular plant 8: 964-966

Comadran J, Kilian B, Russell J, Ramsay L, Stein N, Ganal M, Shaw P, Bayer M, Thomas W, Marshall D, Hedley P, Tondelli A, Pecchioni N, Francia E, Korzun V, Walther A, Waugh R (2012) Natural variation in a homolog of Antirrhinum CENTRORADIALIS contributed to spring growth habit and environmental adaptation in cultivated barley. Nat Genet 44:1388-1392

Cuesta-Marcos A, Szucs P, Close TJ, Filichkin T, Muehlbauer GJ, Smith KP, Hayes PM (2010) Genome-wide SNPs and re-sequencing of growth habit and inflorescence genes in barley: implications for association mapping in germplasm arrays varying in size and structure. BMC Genomics 11:707

Cutler DJ, Zwick ME, Carrasquillo MM, Yohn CT, Tobin KP, Kashuk C, Mathews DJ, Shah NA, Eichler EE, Warrington JA, Chakravarti A (2001) High-throughput variation detection and genotyping using microarrays. Genome Res., 11: 1913–1925

Dai F, Nevo E, Wu D, Comadran J, Zhou M, et al. (2012) Tibet is one of the centers of domestication of cultivated barley. Proc Natl Acad Sci USA www.pnas.org/cgi/doi/10.1073/pnas.1215265109.

De Candolle A (1959) Origin of Cultivated Plants. 2nd ed. Hafner, New York. (Translated from the 1896 edition).

Dewey DR (1984) The genome system of classification as a guide to hybridization with perennial Triticeae. In: J.P. Gustafson (ed), Gene Manipulation in Plant Taxonomy. Plenum Publ. Corp., New York, pp. 209-279.

Djalali M (1970b) Investigations on expressivity and penetrance of *labile* character of barley (*Hordeum vulgare* L.). Z Pflanzenzüchtung 63:274-322

Djalali M, Hoffman W, Plarre W (1970a) Genetics and varaibility of the *labile*-gene in barley under different environmental conditions. In Barley Genetics II:p.p. 201-207, WSUP, Washington

Doebley J, Stec A, Hubbard L (1997) The evolution of apical dominance in maize. Nature 386:485-488

Doyle JJ, Doyle JL (1990) Isolation of plant DNA from fresh tissue. Focus 12:13-15.

Druka A, Franckowiak J, Lundqvist U, Bonar N, Alexander J, Houston K, Radovic S, Shahinnia F, Vendramin V, Morgante M, Stein N, Waugh R (2011) Genetic dissection of barley morphology and development. Plant Physiol 155:617-627

Duke, J. A. 1983. Handbook of Energy Crops. unpublished. Hordeum vulgare L.

Echt C, Knapp S, Liu BH (1992) Genome mapping with non-inbred crosses using Gmendel 2.0. Maize Genet. Coop. Newslett. 66: 27-29.

Edwards AWF (2012) Reginald Crundall Punnett: First Arthur Balfour Professor of Genetics, Cambridge, 1912. Genetics 192: 3–13

Engledow FL (1924) Inheritance in barley III. Journ. Genetics 14: 57-58, 60-62 illus

Feng ZY, Liu XJ, Zhang YZ, Ling HQ (2006) Genetic diversity analysis of Tibet wild barley using SSR markers. Acta Genet Sin 33: 917–928.

Ferreira A, da Silva MF, da Costa e Silva L, Cruz CD (2006) Estimating the effects of population size and type on the accuracy of genetic maps. Genet. Mol. Biol. 29: 182-192.

Forster BP, Franckowiak JD, Lundqvist U, Lyon J, Pitkethly I, Thomas WTB (2007) The barley phytomer. Annals of Botany 100:725-733

Gawroński P, Schnurbusch T (2012) High-density mapping of the *earliness per se-3A*^{m} (*Eps-3A*^{m}) locus in diploid einkorn wheat and its relation to the syntenic regions in rice and *Brachypodium distachyon* L. Mol Breeding 30:1097–1108

Gilbert W, Maxam A (1973) The Nucleotide Sequence of the lac Operator. PNAS U.S.A. 70 (12): 3581–3584.

Grando S, Macpherson HG (2005) Food Barley: Importance, Uses and Local Knowledge. Proc. Int'l. workshop on Food Barley Improvement. Hammamet. Tunisia. January, 14-17, 2002, ICARDA, Aleppo, Syria.

Graner A, Jahoor A, Schondelmaier J, Siedler H, Pillen K, Fishbeck G, Wenzel G, Herrmann RG (1991) Construction of an RFLP map of barley. Theor. Appl. Genet. 83: 250–256

Greenway H, Munns R (1980) Mechanisms of Salt Tolerance in Non-Halophytes. Annu Rev Plant Phys 31:149-190

Hadado TT, Rau D, Bitocchi E, Papa R (2009) Genetic diversity of barley (*Hordeum vulgare* L.) landraces from the central highlands of Ethiopia: comparison between the Belg and Meher growing seasons using morphological traits. Genet Resour Crop Evol 56:1131–1148

Hadado TT, Rau D, Bitocchi E, Papa R (2010) Adaptation and diversity along an altitudinal gradient in Ethiopian barley (*Hordeum vulgare* L.) landraces revealed by molecular analysis. BMC Plant Biol 10:121

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

Harlan HV (1914) Some distinctions in our cultivated barleys with reference to their use in plant breeding. U.S. Dept. Agr. Bull. 137: 38 pp. illus

Harlan JR (1968) On the origins of barley. In: Barley: Origin, Botany, Culture, Winterhardiness, Genetics, Utilization. USDA Handb. 338. U.S. Gov. Print Office, Washington, DC.

Harlan JR (1992) Crops and Man. American Society of Agronomy, Madison, Wisconsin, 2nd ed.

Harlan JR, Zohary D (1966) Distribution of wild wheat and barley. Science 153: 1074–1080.

Hearnden PR, Eckermann PJ, McMichael GL, Hayden MJ, Eglinton JK, Chalmers KJ (2007) A genetic map of 1,000 SSR and DArT markers in a wide barley cross. Theor Appl Genet 115:383-391

Helbaek H (1959) Domestication of food plants in Old World. Science 130: 365–372.

Helbaek H (1960) Ecological effects of irrigation in ancient Mesopotamia. Iraq 22: 186-196.

Helbaek H (1969) Plant collecting, dry-farming, and irrigation agriculture in prehistoric Deh, Luran. Mem. Mus. Anthrop. Univ. Michigan 1: 383-426.

Heun M, Kennedy AE, Anderson JA, Lapitan NLV, Sorrells ME, Tanksley SD (1991) Construction of a restriction fragment lenght polymorphism map for barley (Hordeum vulgare). Genome 35: 1019–1025.

IBSC (2012) A physical, genetic and functional sequence assembly of the barley genome. Nature 491:711-716

Isidore E, Scherrer B, Bellec A, Budin K, Faivre-Rampant P, Waugh R, Keller B, Caboche M, Feuillet C, Chalhoub B (2005) Direct targeting and rapid isolation of BAC clones spanning a defined chromosome region. Functional & Integrative Genomics 5:97-103

Jacob HJ, Lindpaintner K, Lincoln SE et al (1991) Genetic mapping of a gene causing hypertension in the stroke-prone spontaneously hypertensive rat. Cell 67: 213–224, 1991.

Jiang Guo-Liang (2013) Plant Breeding from Laboratories to Fields. Molecular Markers and Marker-Assisted Breeding in Plants. http://dx.doi.org/10.5772/52583

Jones H, Lister DL, Bower MA, Leigh FJ, Smith LM, Jones MK (2008) Approaches and constraints of using existing landrace and extant plant material to understand agricultural spread in prehistory. Plant Genet Res. 6:98–112.

Kirby E, Appleyard M (1987) Cereal development guide. NAC Cereal Unit, Stoneleigh, UK

Kirby EJM, Riggs TJ (1978) Developmental consequences of two-row and six-row ear type in spring barley: 2. Shoot apex, leaf and tiller development. J Agric Sci Camb 91:207-2016

Kling, J. 2004. An Introduction to Barley - Notes from CSS 330 World Foods Class. Accessed April 18, 2006.

Komatsuda T, Pourkheirandish M, He CF, Azhaguvel P, Kanamori H, Perovic D, Stein N, Graner A, Wicker T, Tagiri A, Lundqvist U, Fujimura T, Matsuoka M, Matsumoto T, Yano M (2007) Six-rowed barley originated from a mutation in a homeodomain-leucine zipper I-class homeobox gene. Proc Natl Acad Sci USA 104:1424-1429

Koppolu R, Anwar N, Sakuma S, Tagiri A, Lundqvist U, Pourkheirandish M, Rutten T, Seiler C, Himmelbach A, Ariyadasa R, Youssef HM, Stein N, Sreenivasulu N, Komatsuda T,

Schnurbusch T (2013) *Six-rowed spike4* (*Vrs4*) controls spikelet determinacy and row-type in barley. Proc Natl Acad Sci USA 110: 13198–13203

Körnicke F (1885) Handbuch des Getreidebaues I. Berlin

Lander ES, Green P, Abrahamson J, Barlow A, Daly MJ, Lincoln SE, Newburg L (1987) MAPMAKER: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. Genomics 1: 174–181.

Li H, Durbin R (2009) Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics*, *25*(14), 1754-1760.

Li JZ, Sjakste TG, Röder MS, Ganal MW (2003) Development and genetic mapping of 127 new microsatellite markers in barley. Theor Appl Genet 107:1021–1027

Llaca V (2012) Sequencing Technologies and Their Use in Plant Biotechnology and Breeding. DNA Sequencing – Methods and Applications. www.intechopen.com.

Lundqvist U, Franckowiak JD (2003) Diversity of barley mutants. In: R von Bothmer, T van Hintum, H Knüpffer and K Sato (eds), Diversity in Barley (*Hordeum vulgare*), pp 77-96 Elsevier Science BV, Amsterdam, The Netherlands

Lundqvist U, Franckowiak JD, Konishi T (1997) New and revised descriptions of barley genes. Barley Genetics Newsletter 26:22-516

Lundqvist U, Lundqvist A (1988) Induced *intermedium* mutants in barley: origin, morphology and inheritance. Hereditas 108, 13–26

Lundqvist U, Lundqvist A (1989) The co-operation between *intermedium* genes and the six-row gene *hex-v* in a six-row variety of barley. Hereditas 110:227-233

Luo ZW, Potokina E, Druka A, Wise R, Waugh R, Kearsey MJ (2007) SFP Genotyping From Affymetrix Arrays Is Robust But Largely Detects Cis-acting Expression Regulators. Genetics 176:789-800

Maas EV, Hoffman GJ (1977) Crop salt tolerance-current assessment. J Irrig Drain 103(IR2):115-134

Maeso I, Roy SW, Irimia M (2012) Widespread recurrent evolution of genomic features. Genome Biol Evol 4: 486–500.

Malcomber ST, Preston JC, Reinheimer R, Kossuth J, Kellogg EA. (2006). Developmental gene evolution and the origin of grass inflorescence diversity. Advances in Botanical Research 44: 425–481

Mammadov J, Aggarwal R, Buyyarapu R, and Kumpatla S (2012) SNP markers and their impact on plant breeding International Journal of Plant Genomics. doi:10.1155/2012/728398

Manikanda BN (2012) Genetic Mapping and Marker Assisted Selection: Basics, Practice and Benefits. Springer, pp 272-273

Manly KF, Cudmore Jr RH, Meer JM (2001). Map Manager QTX, crossplatform software for genetic mapping. Mamm. Genome 12: 930–932.

Mansfeld R (1950) Das morphologische System der Saatgerste, *Hordeum vulgare* L. s. l. Züchter 20:8-24

Marcel TC, Varshney RK, Barbieri M, Jafary H, de Kock MJD, Graner A, Niks RE (2007) A high-density consensus map of barley to compare the distribution of QTLs for partial

resistance to Puccinia hordei and of defence gene homologues. Theor Appl Genet 114:487-500

Margulies M, Egholm M, Altman WE, Attiya S, Bader JS, Bemben LA, Berka J, et al (2005) Genome sequencing in microfabricated high-density picolitre reactors. Nature, 437: 376‑380

Marwat SK, Hashimi M, Usman Khan K, Aslam Khan M, Shoaib M, Rehman F (2012) Barley (Hordeum vulgare L.) a prophetic food mentioned in ahadith and its ethnobotanical importance. American-Eurasian Journal of Agricultural & Environmental Sciences 12:835-841

Mascher M, Richmond TA, Gerhardt DJ, Himmelbach A, Clissold L, Sampath D, Ayling S, Steuernagel B, Pfeifer M, D'Ascenzo M, Akhunov ED, Hedley PE, Gonzales AM, Morrell PL, Kilian B, Blattner FR, Scholz U, Mayer KFX, Flavell AJ, Muehlbauer GJ, Waugh R, Jeddeloh JA, Stein N (2013) Barley whole exome capture: a tool for genomic research in the genus *Hordeum* and beyond. The Plant Journal 76:494-505

Mayer K, Martis M, Hedley P, Simková H, Liu H, Morris J, Steuernagel B, Taudien S, Roessner S, Gundlach H, Kubaláková M, Suchánková P, Murat F, Felder M, Nussbaumer T, Graner A, Salse J, Endo T, Sakai H, Tanaka T, Itoh T, Sato K, Platzer M, Matsumoto T, Scholz U, Dolezel J, Waugh R, Stein N (2011) Unlocking the barley genome by chromosomal and comparative genomics. Plant cell 23: 1249-1263

Mikkelsen TS, Ku M, Jaffe DB, Issac B, Lieberman E, Giannoukos G, Alvarez P, Brockman W, Kim TK, Koche RP, Lee W, Mendenhall E, Donovan AO, Presser A, Russ C, Xie X, Meissner A, Wernig M, Jaenisch R, Nusbaum C, Lander ES, Bernstein BE (2007) Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. Nature 448:553–560

Molina-Cano JL, Russell JR, Moralejo MA, Escacena JL, Arias G, et al. (2005) Chloroplast DNA microsatellite analysis supports a polyphyletic origin for barley. Theor Appl Genet 110: 613–619.

Morrell PL, Clegg MT (2007) Genetic evidence for a second domestication of barley (Hordeum vulgare) east of the Fertile Crescent. Proc Natl Acad Sci USA 104:3289–3294.

Munns R, Tester M (2008) Mechanisms of salinity tolerance. Annu Rev Plant Biol 59:651-681

Nese Sreenivasulu, Andreas Graner, and Ulrich Wobus (2008) Barley Genomics: An Overview. International Journal of Plant Genomics, Article ID 486258, 13 pages

Nevo E (2006) Genome evolution of wild cereal diversity and prospects for crop improvement. Plant Genet Resour 4: 36–46.

Nevo E (2012) Evolution of wild barley and barley improvement. Advance in Barley Sciences, Proceedings of 11th International Barley Genetics Symposium. 1–16.

Nevo E, Beiles A, Gutterman Y, Storch N, Kaplan D (1984) Genetic resources of wild cereals in Israel and vicinity: II. Phenotypic variation within and between populations of wild barley, Hordeum spontaneum. Euphytica 33: 737–756.

Nevo E, Beiles A, Zohary D (1986) Genetic resources of wild barley in the Near East: Structure, evolution and application in breeding. Biol J Linn Soc Lond 27: 355–380.

Nevo E, Chen G (2010) Drought and salt tolerances in wild relatives for wheat and barley improvement. Plant Cell Environ 33:670-685

Nevski SA (1941) Beiträge zur Kenntnis der wildwachsenden Gersten im Zusammenhang mit der Frage über den Ursprung von *Hordeum vulgare* L. und *Hordeum distichon* L. (Versuch einer Monographie der Gattung *Hordeum*). Trudy Bot. Inst. Akad. Nauk SSSR, ser. 1, 5: 64-255.

Newman CW, Newman RK (2006) A brief history of barley foods. Cereal Foods World 51:4-7

Nötzel H (1952) Genetische Untersuchungen an röntgeninduzierten Gerstenmutanten. Kühn-Arch 66:72-132

Nyren P (2007) The history of pyrosequencing. Methods Mol. Biol., 373 (2007), pp. 1‑14

Ossowski S, Schneeberger K, Clark RM , Lanz C, Warthmann N, and Weigel D (2008) Sequencing of natural strains of *Arabidopsis thaliana* with short reads. Genome Res.18:2024–2033.

Pettersson E, Lundeberg J, Ahmadian A (2009) Generations of sequencing technologies. Genomics 93:105–111

Potokina E, Druka A, Luo ZW, Wise R, Waugh R, Kearsey M (2008) Gene expression quantitative trait locus analysis of 16,000 barley genes reveals a complex pattern of genome-wide transcriptional regulation. Plant Journal 53:90-101

Pourkheirandish M, Komatsuda T (2007) The Importance of Barley Genetics and Domestication in a Global Perspective. Annals of Botany 100: 999–1008

Poursarebani N (2012) A genetically anchored physical map of barley chromosome 2H. Agriculture. Martin Luther University of Halle, Halle, p 107

Powell W, Machray GC, Provan J (1996). Polymorphism revealed by simple sequence repeats. Trends Plant Sci. 7: 215–222.

Ramsay L, Comadran J, Druka A, Marshall DF, Thomas WTB, Macaulay M, MacKenzie K, Simpson C, Fuller J, Bonar N, Hayes PM, Lundqvist U, Franckowiak JD, Close TJ, Muehlbauer GJ, Waugh R (2011) *INTERMEDIUM-C*, a modifier of lateral spikelet fertility in barley, is an ortholog of the maize domestication gene *TEOSINTE BRANCHED 1*. Nat Genet 43:169-U125

Rostoks N, Mudie S, Cardle L, Russell J, Ramsay L, Booth A, Svensson JT, Wanamaker SI, Walia H, Rodriguez EM, Hedley PE, Liu H, Morris J, Close TJ, Marshall DF, Waugh R (2005) Genome-wide SNP discovery and linkage analysis in barley based on genes responsive to abiotic stress. Molecular Genetics and Genomics 274:515-527

Rostoks N, Mudie S, Cardle L, Russell J, Ramsay L, Booth A, Svensson JT, Wanamaker SI, Walia H, Rodriguez EM, Hedley PE, Liu H, Morris J, Close TJ,Marshall DF, Waugh R. (2005) Genome-wide SNP discovery and linkage analysis in barley based on genes responsive to abiotic stress. Mol Genet Genomics. 274:515-27.

Rothberg, J.M. Hinz W, Rearick TM, Schultz J, Mileski W, et al (2011) An integrated semiconductor device enabling non-optical genome sequencing. Nature 475(7356): 348-352.

Russell J, Booth A, Fuller J, Harrower B, Hedley P, Machray G, and Powell W (2004) A comparison of sequence-based polymorphism and haplotype content in transcribed and anonymous regions of the barley genome. Genome.47:389–398.

Saisho D, Myoraku E, Kawasaki S, Sato K, Takeda K (2007) Construction and Characterization of a Bacterial Artificial Chromosome (BAC) Library from the Japanese Malting Barley variety 'Haruna Nijo'. Breeding Science 57:29-38

Saisho D, Pourkheirandish M, Kanamori H, Matsumoto T, Komatsuda T (2009) Allelic variation of row type gene *Vrs1* in barley and implication of the functional divergence. Breed Sci 59:621-628

Saisho D, Takeda K (2011) Barley: Emergence as a New Research Material of Crop Science. Plant Cell Physiol 52:724-727

Sakuma S, Pourkheirandish M, Matsumoto T, Koba T, Komatsuda T (2010) Duplication of a well-conserved homeodomain-leucine zipper transcription factor gene in barley generates a copy with more specific functions. Functional & Integrative Genomics 10:123-133

Sakuma S, Salomon B, Komatsuda T (2011) The Domestication Syndrome Genes Responsible for the Major Changes in Plant Form in the Triticeae Crops. Plant Cell Physiol 52:738-749

Salamini F, Özkan H, Brandolini A, Schäfer-Pregl R, Martin W (2002) Genetics and geography of wild cereal domestication in the Near East. Nature Reviews Genetics3:429-441.

Sanger F, Nicklen S, Coulson AR (December 1977). "DNA sequencing with chain-terminating inhibitors". P N A S. U.S.A. 74 (12): 5463–7

Sato K, Nankaku N, Takeda K (2009) A high-density transcript linkage map of barley derived from a single population. Heredity 103:110-117

Schulte D, Ariyadasa R, Shi BJ, Fleury D, Saski C, Atkins M, deJong P, Wu CC, Graner A, Langridge P, Stein N (2011) BAC library resources for map-based cloning and physical map construction in barley (Hordeum vulgare L.). BMC Genomics 12:247

Schulte D, Close TJ, Graner A, Langridge P, Matsumoto T, Muehlbauer G, et al (2009) The international barley sequencing consortium—at the threshold of efficient access to the barley genome. Plant Physiol. 149:142-147.

Schuster WH (1997) Welchen Beitrag leistet die Pflanzenzüchtung zurLeistungssteigerung von Kulturpflanzenarten, Pflanzenbauwissenschaften, vol. 1, pp. 9–18.

Semagn K, Bjørnstad Å and Ndjiondjop MN (2006) Principles, requirements and prospects of genetic mapping in plants. African Journal of Biotechnology 5: 2569-2587.

Shi BJ, Sutton T, Collins NC, Pallotta M, Langridge P (2010) Construction of a barley bacterial artificial chromosome library suitable for cloning genes for boron tolerance, sodium exclusion and high grain zinc content. Plant Breeding 129:291-296

Smith D (2013) The First Genetic-Linkage Map. http://www.caltech.edu/news/first-genetic-linkage-map-38798

Squirrell J, Hollingsworth PM, Woodhead M, Russell J, Lowe AJ, Gibby M, Powell W (2003). How much effort is required to isolate nuclear microsatellites from plants? Mol. Ecol. 12: 1339-1348.

Sreenivasulu N, Graner A, Wobus U (2008) Barley genomics: an overview. Int. J. Plant Genomics. 2008:486258.

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

Stam P (1993a). Construction of integrated genetic linkage maps by means of a new computer package: JoinMap. Plant J. 3: 739-744.

Stein N, Prasad M, Scholz U, Thiel T, Zhang H, Wolf M, Kota R, Varshney RK., Perovic D, Grosse I, Graner A (2007) A 1,000-loci transcript map of the barley genome: new anchoring points for integrative grass genomics. Theor Appl Genet 114:823–839

Suiter KA, Wendel JF, Case JS (1983) Linkage-I: A pascal computer program for the detection and analysis of genetic linkage. J. Hered. 74: 203-204.

Takahashi R, and Tomihisa Y, (1970) Genetic approach to the origin of two wild forms of barley, *lagunculiforme* Bacht. and *proskowetzii* Nabelek (*Hordeum spontaneum* C. Koch emend. Bacht.). In: Barley Genetics II. Proc. 2nd Int. Barley Genet. Symp., Pullman, WA, USA, pp. 51-62.

Takeda K, Saito W (1988) Inheritance of the percentage of missing lateral florets in '*Irregular*' barley. Japan J Breed 38:72-80 (in Japanese with English summary p. 80)

Thompson JD, Linard B, Lecompte O, Poch O (2011) A Comprehensive Benchmark Study of Multiple Sequence Alignment Methods: Current Challenges and Future Perspectives. DOI: 10.1371/journal.pone.0018093

U.S. Grains Council (2006) Barley: Genus Hordeum, family poaceae, a cereal grain.

van Ooijen JW (2006) JoinMap4, software for calculation of genetic linkage maps in experimental populations. Kyazma B.V., Wageningen, Natherlands.

Varshney RK, Grosse I, Hahnel U, Siefken R, Prasad M, Stein N, Langridge P, Altschmied L, Graner A (2006) Genetic mapping and BAC assignment of EST-derived SSR markers shows non-uniform distribution of genes in the barley genome. Theor Appl Genet 113:239-250

Vincze T, Posfai J, Roberts RJ (2003) NEBcutter: a program to cleave DNA with restriction enzymes. Nucleic Acids Res 31:3688-3691

von Bothmer R, Jacobsen N (1986) Interspecific crosses in Hordeum (Poaceae). Pl Syst Evol 153:49-64

von Bothmer R, Jacobsen N, Baden C, Jorgensen RB, Linde-Laursen I (1995) An ecogeographical study of the genus Hordeum. Systematic and ecogeographical studies on crop genepools, 7, 2nd edn. Rome: IPGRI.

von Bothmer R, Jacobsen N, Jorgensen RB (1985) Two new American species of *Hordeum* (Poaceae). Willdenowia 15:85-90

von Bothmer R, Komatsuda T (2011) in Barley: Production, Improvement, and Uses, ed Ullrich SE (Wiley-Blackwell, Chichester, UK), 14–62.

von Bothmer R, Sato K, Komatsuda T, Yasuda S, Fischbeck G (2003) The domestication of cultivated barley. In: von Bothmer R, Hintum Tv, Knu¨pffer H, Sato K, eds. Diversity in barley (*Hordeum vulgare*). Amsterdam: Elsevier, 9–27.

von Bothmer R, Yen C, Yang JL (1990) Does wild, six-rowed barley, *Hordeum agriocrithon* really exist? Plant Genet. Res. Newsl. 77: 17-19.

Vos P, Hogers R, Bleeker M et al (1995) AFLP: a new technique for DNA fingerprinting Nucleic Acids Research 23: 4407–4414

Wang DG, Fan JB, Siao CJ et al (1998) Large-scale identification, mapping, and genotyping of single- nucleotide polymorphisms in the human genome. Science 280:1077–1082

Weiss E, Wetterstrom W, Nadel D, Bar-Yosef O (2004) The broad spectrum revisited: evidence from plant remains. Proceedings of the National Academy of Sciences of the USA. 101:9551-9555.

Welsh J and McClelland M (1990) Fingerprinting genomes using PCR with arbitrary primers. Nucleic Acids Research 18: 7213–7218

Wenzl P, Li HB, Carling J, Zhou MX, Raman H, Paul E, Hearnden P, Maier C, Xia L, Caig V, Ovesna J, Cakir M, Poulsen D, Wang JP, Raman R, Smith KP, Muehlbauer GJ, Chalmers KJ, Kleinhofs A, Huttner E, Kilian A (2006) A high-density consensus map of barley linking DArT markers to SSR, RFLP and STS loci and agricultural traits. BMC Genomics 7:206

White TL, Adams WT, Neale DB (2007) Forest Genetics, CABI, Science - 682 pages

Xu TW (1985) The wild close relatives of cultivated barley in Qinghai-Tibet Plateau. Barley Newslet 29: 105–106.

Xu, Y. 2010. Molecular plant breeding. CAB International.

Yang JL, Yen C (1985) Distribution of weedrace barley in China and the center of origin of cultivated barley. Utsadesforenings Tidskrift 95: 71–79.

Young, B. (2001) Barley: The Versatile Crop. Southern Illinois University, College of Science, Ethnobotanical Leaflets.

Youssef HM, Fazeli A, Koppolu R, Rutten T, Ariyadasa R, Himmelbach A, Stein N, Schnurbusch T (2013) Does *six rowed spike 2* (*vrs2*) play a role in regulating barley lateral spikelet fertility? VIPCA, Plant Genetics and Breeding Technologies, 18-20 February 2013, Vienna, Austria.

Youssef HM, Koppolu R, Rutten T, Korzun V, Schweizer P, Schnurbusch T (2014) Genetic mapping of the labile (lab) gene: a recessive locus causing irregular spikelet fertility in *labile*-barley (*Hordeum vulgare* convar. *labile*). *Theor Appl Genet* 127:1123-1131

Youssef HM, Koppolu R,Schnurbusch T (2012) Re-sequencing of *vrs1* and *int-c* loci shows that labile-barleys (*Hordeum vulgare* convar. *labile*) have a six-rowed genetic background. Genet Resour Crop Evol 59:1319-1328

Yu Y, Tomkins JP, Waugh R, Frisch DA, Kudrna D, Kleinhofs A, Brueggeman RS, Muehlbauer GJ, Wise RP, Wing RA (2000) A bacterial artificial chromosome library for barley (Hordeum vulgare L.) and the identification of clones containing putative resistance genes. Theor Appl Genet 101:1093-1099

Zhang Q, Yang GP, Dai XK, Su JZ (1994) A comparative analysis of genetic polymorphism in wild and cultivated barley from Tibet using isozyme and ribosomal DNA markers. Genome 37: 631–638.

Zohary D, Hopf M (2000) Domestication of plants in the old world: the origin and spread of cultivated plants in West Asia, Europe, and the Nile Valley, 3rd edn. New York, NY: Oxford University Press.

Zohary D, Hopf M, Weiss E (2012) Domestication of Plants in the Old World: The origin and spread of domesticated plants in Southwest Asia, Europe, and the Mediterranean Basin. 4th ed. Oxford: Oxford University Press.

9.0 Supplemental materials

9.1 Supplemental Tables *Chapter two*

Supplementary Table 2-1: 221 Ethiopian barley accessions, gene bank accession number and their corresponding *Vrs1* and *Int-c* alleles, including row-type.

No.	Accessions number	*Vrs1* Allele/Haplotype	*Int-c* Allele/Haplotype	Row-type
1	HOR1560	*vrs1.a1*	*Hap_2*	*labile*
2	HOR1630	*vrs1.a1*	*Hap_2*	*labile*
3	HOR1631	*vrs1.a1*	*Hap_2*	*labile*
4	HOR1632	*vrs1.a1*	*Int-c.a*	*labile*
5	HOR1633	*vrs1.a1*	*Hap_2*	*labile*
6	HOR1634	*vrs1.a1*	*Hap_2*	*labile*
7	HOR1635	*vrs1.a1*	*Hap_2*	*labile*
8	HOR1636	*vrs1.a1*	*Hap_2*	*labile*
9	HOR1637	*vrs1.a1*	*Int-c.a*	*labile*
10	HOR1639	*vrs1.a1*	*Int-c.a*	*labile*
11	HOR1640	*vrs1.a1*	*Hap_2*	*labile*
12	HOR1642	*vrs1.a1*	*Hap_2*	*labile*
13	HOR1643	*vrs1.a1*	*Hap_1*	*labile*
14	HOR2337	*vrs1.a1*	*Int-c.a*	*labile*
15	HOR2380	*vrs1.a1*	*Hap_2*	*labile*
16	HOR2925	*vrs1.a1*	*Hap_2*	*labile*
17	HOR2926	*vrs1.a1*	*Hap_2*	*labile*
18	HOR2927	*vrs1.a1*	*Hap_2*	*labile*
19	HOR2928	*vrs1.a1*	*Hap_2*	*labile*
20	HOR2933	*vrs1.a1*	*Hap_2*	*labile*
21	HOR2935	*vrs1.a1*	*Hap_2*	*labile*
22	HOR3065	*vrs1.a1*	*Int-c.a*	*labile*

23	HOR3209	*vrs1.a1*	*Int-c.a*	*labile*
24	HOR3228	*vrs1.a1*	*Hap_2*	*labile*
25	HOR3279	*vrs1.a1*	*Hap_2*	*labile*
26	HOR3281	*vrs1.a1*	*Int-c.a*	*labile*
27	HOR3284	*vrs1.a1*	*Hap_2*	*labile*
28	HOR3525	*vrs1.a1*	*Hap_2*	*labile*
29	HOR3527	*vrs1.a1*	*Hap_2*	*labile*
30	HOR3528	*vrs1.a1*	*Int-c.a*	*labile*
31	HOR3530	*vrs1.a1*	*Int-c.a*	*labile*
32	HOR3531	*vrs1.a1*	*Hap_2*	*labile*
33	HOR3532	*vrs1.a1*	*Hap_2*	*labile*
34	HOR3533	*vrs1.a1*	*Int-c.a*	*labile*
35	HOR3587	*vrs1.a1*	*Hap_2*	*labile*
36	HOR3588	*vrs1.a1*	*Int-c.a*	*labile*
37	HOR4002	*vrs1.a1*	*Int-c.a*	*labile*
38	HOR4251	*vrs1.a1*	*Int-c.a*	*labile*
39	HOR4651	*vrs1.a1*	*Hap_2*	*labile*
40	HOR5050	*vrs1.a1*	*Int-c.a*	*labile*
41	HOR5065	*vrs1.a1*	*Int-c.a*	*labile*
42	HOR5066	*vrs1.a1*	*Int-c.a*	*labile*
43	HOR5068	*vrs1.a1*	*Int-c.a*	*labile*
44	HOR5075	*vrs1.a1*	*Int-c.a*	*labile*
45	HOR5079	*vrs1.a1*	*Hap_2*	*labile*
46	HOR5086	*vrs1.a1*	*Int-c.a*	*labile*
47	HOR5092	*vrs1.a1*	*Int-c.a*	*labile*
48	HOR5094	*vrs1.a1*	*Int-c.a*	*labile*

49	HOR5097	*vrs1.a1*	*Hap_2*	*labile*
50	HOR5101	*vrs1.a1*	*Hap_2*	*labile*
51	HOR5181	*vrs1.a1*	*Hap_2*	*labile*
52	HOR5203	*vrs1.a1*	*Hap_2*	*labile*
53	HOR5224	*vrs1.a1*	*Int-c.a*	*labile*
54	HOR5284	*vrs1.a1*	*Int-c.a*	*labile*
55	HOR5285	*vrs1.a1*	*Int-c.a*	*labile*
56	HOR5288	*vrs1.a1*	*Int-c.a*	*labile*
57	HOR5289	*vrs1.a1*	*Int-c.a*	*labile*
58	HOR5290	*vrs1.a1*	*Hap_2*	*labile*
59	HOR5334	*vrs1.a1*	*Hap_2*	*labile*
60	HOR5344	*vrs1.a1*	*Int-c.a*	*labile*
61	HOR5397	*vrs1.a1*	*Hap_2*	*labile*
62	HOR5398	*vrs1.a1*	*Hap_2*	*labile*
63	HOR5399	*vrs1.a1*	*Hap_2*	*labile*
64	HOR5465	*vrs1.a1*	*Int-c.a*	*labile*
65	HOR5473	*vrs1.a1*	*Int-c.a*	*labile*
66	HOR5535	*vrs1.a1*	*Hap_2*	*labile*
67	HOR5586	*vrs1.a1*	*Int-c.a*	*labile*
68	HOR5588	*vrs1.a1*	*Hap_2*	*labile*
69	HOR5601	*vrs1.a1*	*Hap_2*	*labile*
70	HOR5610	*vrs1.a1*	*Int-c.a*	*labile*
71	HOR5635	*vrs1.a1*	*Hap_2*	*labile*
72	HOR5649	*vrs1.a1*	*Int-c.a*	*labile*
73	HOR5716	*vrs1.a1*	*Hap_2*	*labile*
74	HOR5751	*vrs1.a1*	*Int-c.a*	*labile*

75	HOR5765	*vrs1.a1*	*Int-c.a*	*labile*
76	HOR5885	*vrs1.a1*	*Int-c.a*	*labile*
77	HOR5930	*vrs1.a1*	*Hap_2*	*labile*
78	HOR5931	*vrs1.a1*	*Hap_2*	*labile*
79	HOR5932	*vrs1.a1*	*Hap_2*	*labile*
80	HOR5936	*vrs1.a1*	*Hap_2*	*labile*
81	HOR5942	*vrs1.a1*	*Int-c.a*	*labile*
82	HOR5953	*vrs1.a1*	*Int-c.a*	*labile*
83	HOR5981	*vrs1.a1*	*Int-c.a*	*labile*
84	HOR5983	*vrs1.a1*	*Int-c.a*	*labile*
85	HOR5985	*vrs1.a1*	*Int-c.a*	*labile*
86	HOR5988	*vrs1.a1*	*Int-c.a*	*labile*
87	HOR5994	*vrs1.a1*	*Int-c.a*	*labile*
88	HOR5995	*vrs1.a1*	*Int-c.a*	*labile*
89	HOR6002	*vrs1.a1*	*Int-c.a*	*labile*
90	HOR6004	*vrs1.a1*	*Int-c.a*	*labile*
91	HOR6008	*vrs1.a1*	*Int-c.a*	*labile*
92	HOR6011	*vrs1.a1*	*Int-c.a*	*labile*
93	HOR6014	*vrs1.a1*	*Int-c.a*	*labile*
94	HOR6017	*vrs1.a1*	*Int-c.a*	*labile*
95	HOR6034	*vrs1.a1*	*Int-c.a*	*labile*
96	HOR6035	*vrs1.a1*	*Int-c.a*	*labile*
97	HOR6066	*vrs1.a1*	*Hap_2*	*labile*
98	HOR6071	*vrs1.a1*	*Hap_2*	*labile*
99	HOR6102	*vrs1.a1*	*Hap_2*	*labile*
100	HOR6103	*vrs1.a1*	*Hap_2*	*labile*

101	HOR6105	*vrs1.a1*	*Hap_2*	*labile*
102	HOR6114	*vrs1.a1*	*Hap_2*	*labile*
103	HOR6117	*vrs1.a1*	*Hap_2*	*labile*
104	HOR6120	*vrs1.a1*	*Int-c.a*	*labile*
105	HOR6122	*vrs1.a1*	*Hap_2*	*labile*
106	HOR6124	*vrs1.a1*	*Hap_2*	*labile*
107	HOR6138	*vrs1.a1*	*Int-c.a*	*labile*
108	HOR6173	*vrs1.a1*	*Hap_2*	*labile*
109	HOR6174	*vrs1.a1*	*Hap_2*	*labile*
110	HOR6201	*vrs1.a1*	*Int-c.a*	*labile*
111	HOR6212	*vrs1.a1*	*Int-c.a*	*labile*
112	HOR6213	*vrs1.a1*	*Int-c.a*	*labile*
113	HOR6214	*vrs1.a1*	*Int-c.a*	*labile*
114	HOR6215	*vrs1.a1*	*Int-c.a*	*labile*
115	HOR6216	*vrs1.a1*	*Int-c.a*	*labile*
116	HOR6220	*vrs1.a1*	*Int-c.a*	*labile*
117	HOR6221	*vrs1.a1*	*Int-c.a*	*labile*
118	HOR6227	*vrs1.a1*	*Int-c.a*	*labile*
119	HOR6234	*vrs1.a1*	*Int-c.a*	*labile*
120	HOR6304	*vrs1.a1*	*Hap_2*	*labile*
121	HOR6307	*vrs1.a1*	*Int-c.a*	*labile*
122	HOR6308	*vrs1.a1*	*Int-c.a*	*labile*
123	HOR6311	*vrs1.a1*	*Hap_2*	*labile*
124	HOR6314	*vrs1.a1*	*Int-c.a*	*labile*
125	HOR6316	*vrs1.a1*	*Int-c.a*	*labile*
126	HOR6319	*vrs1.a1*	*Int-c.a*	*labile*

127	HOR6324	*vrs1.a1*	*Int-c.a*	*labile*
128	HOR6326	*vrs1.a1*	*Int-c.a*	*labile*
129	HOR6331	*vrs1.a1*	*Int-c.a*	*labile*
130	HOR6332	*vrs1.a1*	*Int-c.a*	*labile*
131	HOR6333	*vrs1.a1*	*Int-c.a*	*labile*
132	HOR6423	*vrs1.a1*	*Hap_2*	*labile*
133	HOR6435	*vrs1.a1*	*Hap_2*	*labile*
134	HOR6438	*vrs1.a1*	*Hap_2*	*labile*
135	HOR6472	*vrs1.a1*	*Int-c.a*	*labile*
136	HOR6492	*vrs1.a1*	*Int-c.a*	*labile*
137	HOR6500	*vrs1.a1*	*Int-c.a*	*labile*
138	HOR6504	*vrs1.a1*	*Int-c.a*	*labile*
139	HOR6505	*vrs1.a1*	*Int-c.a*	*labile*
140	HOR6506	*vrs1.a1*	*Int-c.a*	*labile*
141	HOR6514	*vrs1.a1*	*Int-c.a*	*labile*
142	HOR6523	*vrs1.a1*	*Int-c.a*	*labile*
143	HOR6541	*vrs1.a1*	*Int-c.a*	*labile*
144	HOR6545	*vrs1.a1*	*Hap_2*	*labile*
145	HOR6558	*vrs1.a1*	*Hap_2*	*labile*
146	HOR6569	*vrs1.a1*	*Int-c.a*	*labile*
147	HOR6579	*vrs1.a1*	*Int-c.a*	*labile*
148	HOR6581	*vrs1.a1*	*Hap_2*	*labile*
149	HOR6605	*vrs1.a1*	*Int-c.a*	*labile*
150	HOR6607	*vrs1.a1*	*Int-c.a*	*labile*
151	HOR6619	*vrs1.a1*	*Int-c.a*	*labile*
152	HOR6621	*vrs1.a1*	*Int-c.a*	*labile*

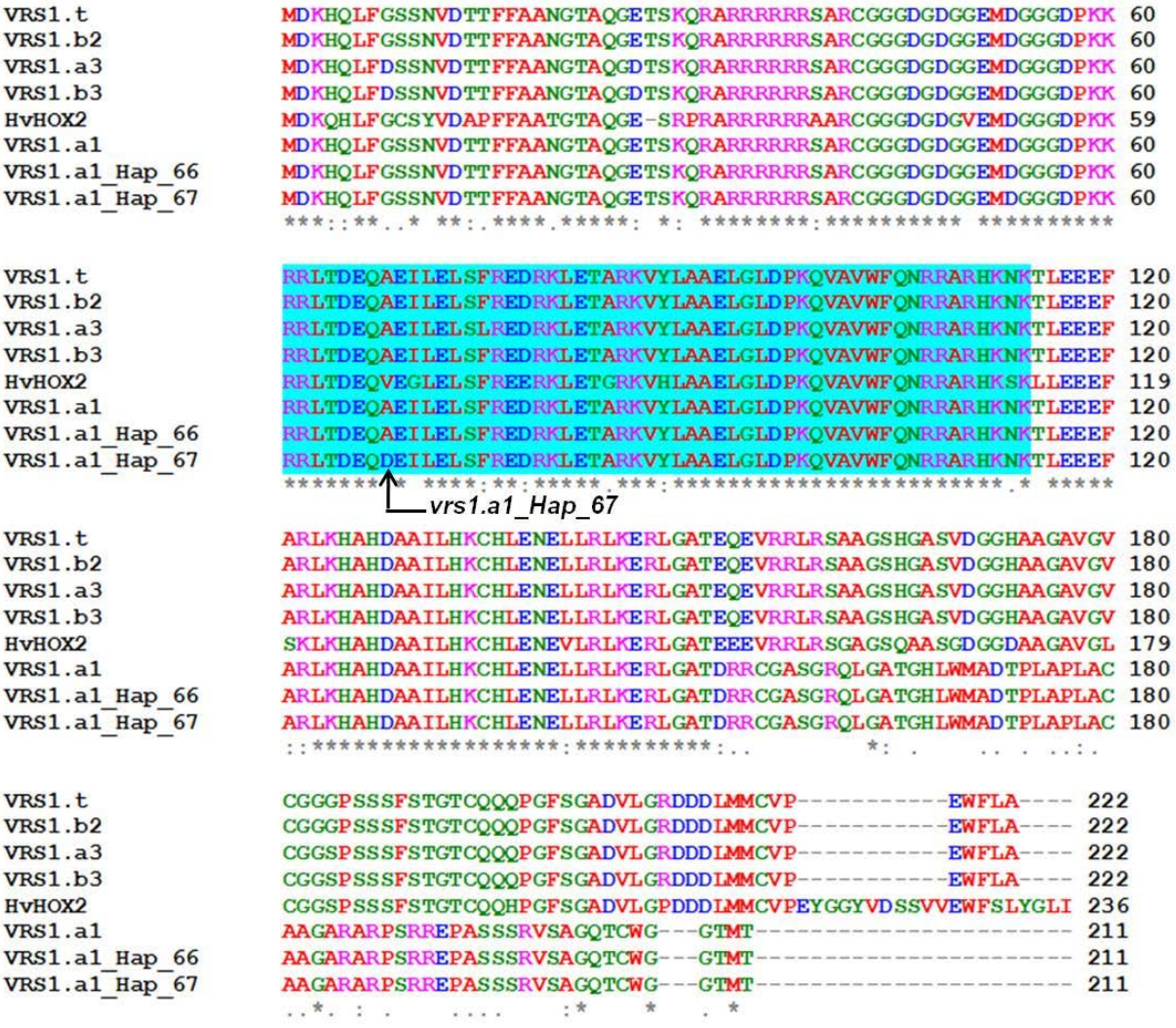
153	HOR6622	*vrs1.a1*	*Int-c.a*	*labile*
154	HOR6627	*vrs1.a1*	*Int-c.a*	*labile*
155	HOR6631	*vrs1.a1*	*Int-c.a*	*labile*
156	HOR6639	*vrs1.a1*	*Int-c.a*	*labile*
157	HOR6724	*vrs1.a1*	*Hap_2*	*labile*
158	HOR6776	*vrs1.a1*	*Int-c.a*	*labile*
159	HOR6852	*vrs1.a1*	*Hap_2*	*labile*
160	HOR7270	*vrs1.a1*	*Int-c.a*	*labile*
161	HOR7720	*vrs1.a1*	*Int-c.a*	*labile*
162	HOR7744	*vrs1.a1*	*Hap_2*	*labile*
163	HOR7746	*vrs1.a1*	*Hap_2*	*labile*
164	HOR7750	*vrs1.a1*	*Hap_2*	*labile*
165	HOR7767	*vrs1.a1*	*Hap_2*	*labile*
166	HOR7786	*vrs1.a1*	*Hap_2*	*labile*
167	HOR7814	*vrs1.a1*	*Int-c.a*	*labile*
168	HOR7828	*vrs1.a1*	*Int-c.a*	*labile*
169	HOR7835	*vrs1.a1*	*Hap_2*	*labile*
170	HOR7843	*vrs1.a1*	*Int-c.a*	*labile*
171	HOR7847	*vrs1.a1*	*Int-c.a*	*labile*
172	HOR7864	*vrs1.a1*	*Int-c.a*	*labile*
173	HOR8992	*vrs1.a1*	*Int-c.a*	*labile*
174	HOR9003	*vrs1.a1*	*Int-c.a*	*labile*
175	HOR9004	*vrs1.a1*	*Hap_2*	*labile*
176	HOR9077	*vrs1.a1*	*Hap_2*	*labile*
177	HOR9105	*vrs1.a1*	*Hap_2*	*labile*
178	HOR9123	*vrs1.a1*	*Int-c.a*	*labile*

179	HOR9124	*vrs1.a1*	*Int-c.a*	*labile*
180	HOR9158	*vrs1.a1*	*Hap_2*	*labile*
181	HOR9193	*vrs1.a1*	*Hap_2*	*labile*
182	HOR9340	*vrs1.a1*	*Int-c.a*	*labile*
183	HOR9365	*vrs1.a1*	*Hap_2*	*labile*
184	HOR9370	*vrs1.a1*	*Int-c.a*	*labile*
185	HOR9401	*vrs1.a1*	*Int-c.a*	*labile*
186	HOR9917	*vrs1.a1*	*Hap_2*	*labile*
187	HOR9935	*vrs1.a1*	*Int-c.a*	*labile*
188	HOR9939	*vrs1.a1*	*Int-c.a*	*labile*
189	HOR9940	*vrs1.a1*	*Int-c.a*	*labile*
190	HOR9948	*vrs1.a1*	*Hap_2*	*labile*
191	HOR9949	*vrs1.a1*	*Int-c.a*	*labile*
192	HOR10271	*vrs1.a1*	*Int-c.a*	*labile*
193	HOR10336	*vrs1.a1*	*Hap_2*	*labile*
194	HOR10455	*vrs1.a1*	*Hap_2*	*labile*
195	HOR10457	*vrs1.a1*	*Hap_2*	*labile*
196	HOR10458	*vrs1.a1*	*Hap_2*	*labile*
197	HOR10491	*vrs1.a1*	*Int-c.a*	*labile*
198	HOR10513	*vrs1.a1*	*Hap_2*	*labile*
199	HOR10515	*vrs1.a1*	*Hap_2*	*labile*
200	HOR10516	*vrs1.a1*	*Hap_2*	*labile*
201	HOR10517	*vrs1.a1*	*Hap_2*	*labile*
202	HOR10522	*vrs1.a1*	*Hap_2*	*labile*
203	HOR10523	*vrs1.a1*	*Hap_2*	*labile*
204	HOR10759	*vrs1.a1*	*Int-c.a*	*labile*

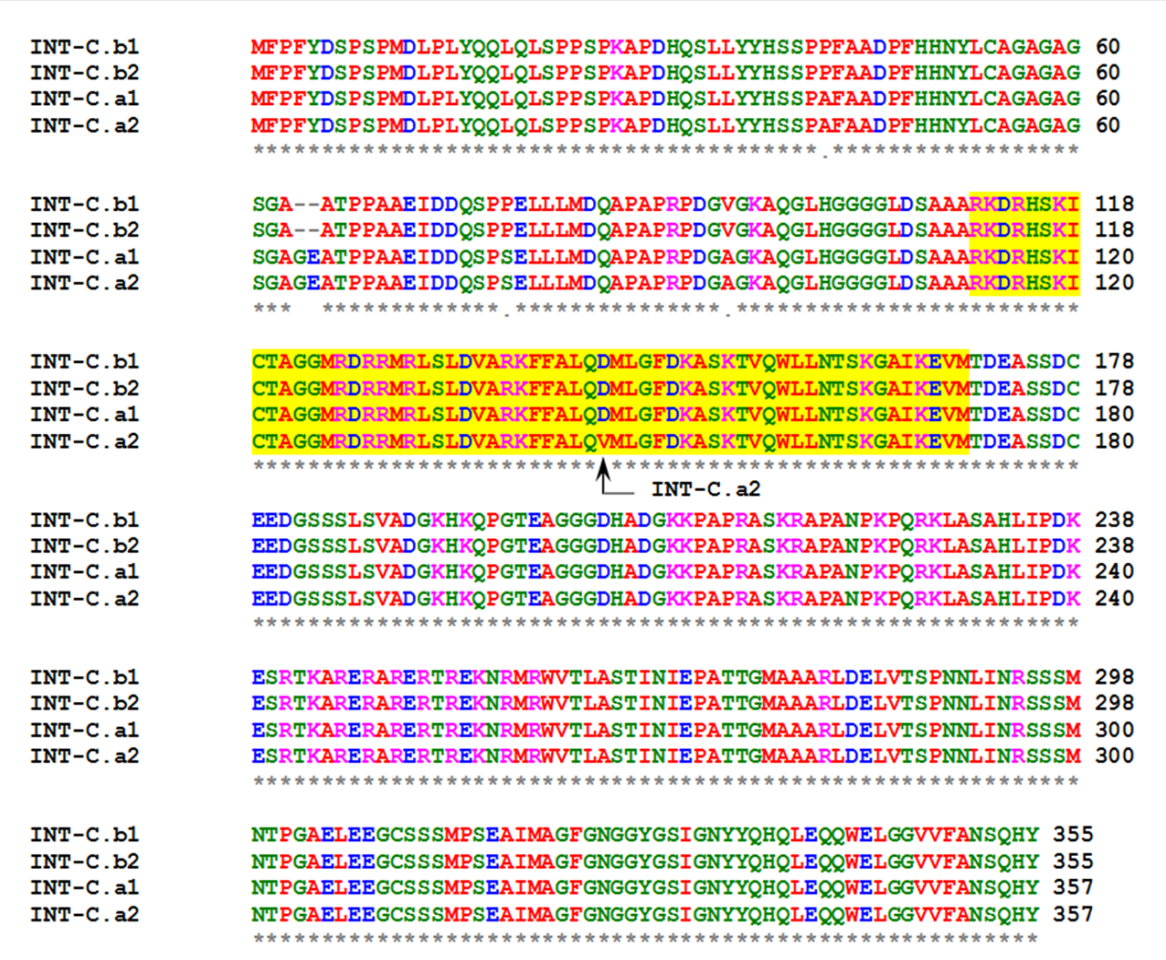
205	HOR3529	*vrs1.a3*	*Int-c.a*	*labile*
206	HOR3587	*vrs1.a3*	*Int-c.a*	*labile*
207	HOR5172	*Hap_67*	*Hap_2*	*labile*
208	HOR6400	*Hap_67*	*Hap_2*	*labile*
209	HOR6420	*Hap_67*	*Hap_2*	*labile*
210	HOR6440	*Hap_67*	*Hap_2*	*labile*
211	HOR6178	*Hap_66*	*Hap_2*	*labile*
212	HOR6179	*Hap_66*	*Hap_2*	*labile*
213	HOR6180	*Hap_66*	*Hap_2*	*labile*
214	HOR6279	*Hap_66*	*Int-c.a*	*labile*
215	HOR7729	*Hap_66*	*Hap_2*	*labile*
216	HOR7734	*Hap_66*	*Hap_2*	*labile*
217	HOR9405	*Hap_66*	*Hap_2*	*labile*
218	HOR10421	*Hap_66*	*Hap_2*	*labile*
219	HOR10490	*Hap_66*	*Hap_2*	*labile*
220	HOR5281*	*Vrs1.b3*	*int-c.b*	*two-rowed Putative mutant in vrs2, 3 or 4?*
221	HOR5471*	*Vrs1.t*	*Int-c.a*	*deficiens*

*Accessions misclassified as *labile* genotypes confirmed by re-sequencing of *vrs1* and *int-c* loci

9.2 Supplemental Figures *Chapter two*

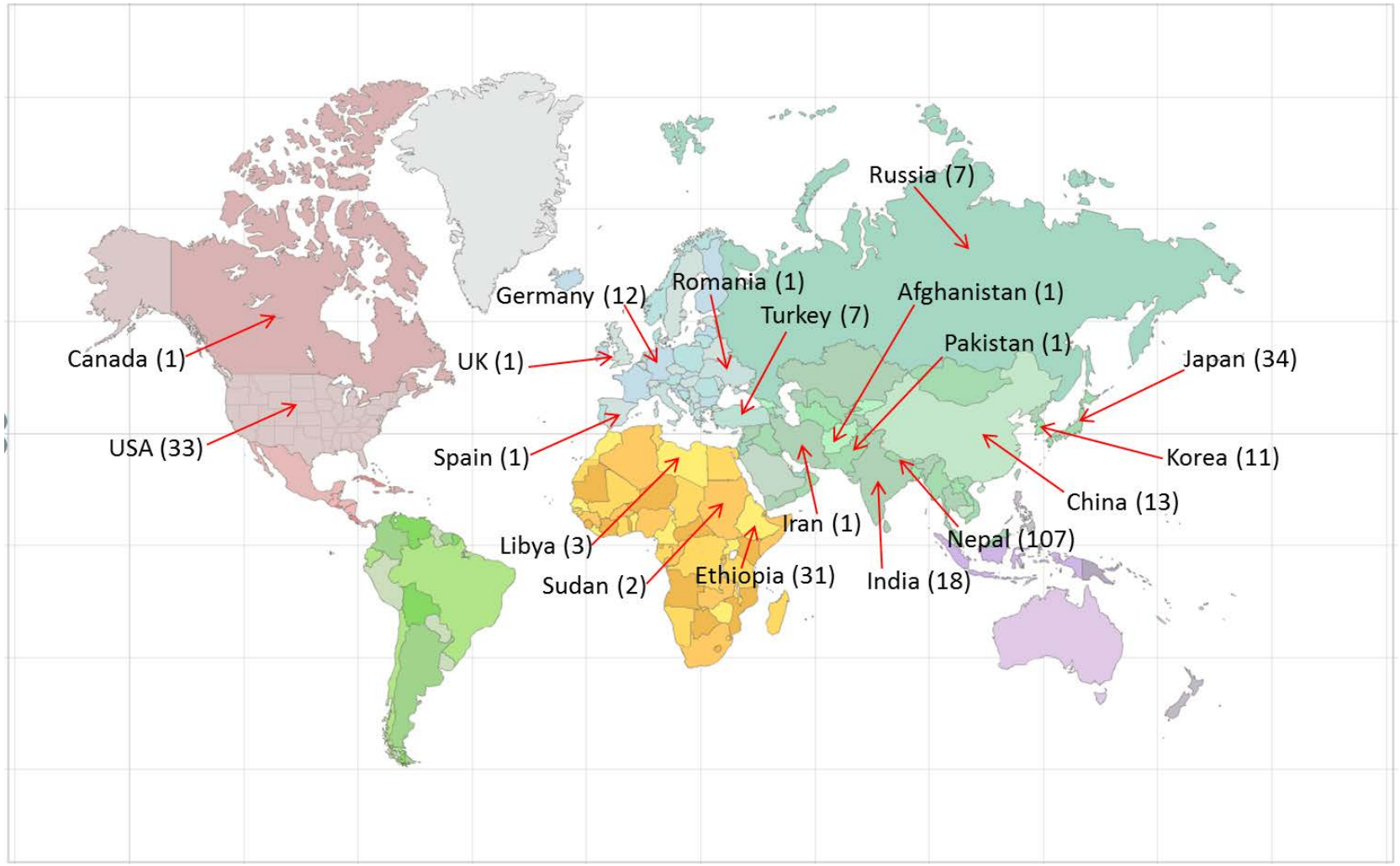


Supplementary Figure 2-1 Amino acid alignments of the novel *vrs1* haplotype (*Hap_67*) in comparison with known *vrs1* alleles and HvHOX2. The HD domain is highlighted in blue. The amino acid substitution due to *Hap_67* of *vrs1.a1* has been indicated with an arrow.



Supplementary Figure 2-2:
Amino acid alignment of *Hap_1* of *Int-c.a* allele with known *int-c* alleles found in barley. The TCP domain is highlighted in yellow. The altered amino acid sequence in *Hap_1* is indicated by the arrow.

9.3 Supplemental Figures *Chapter four*



Sup. Fig 4-1: Distribution of Intermedium collection over the world. Numbers between brackets are the numbers of accessions.

9.1 Supplemental Tables *Chapter four*

Sup. Table 4-1: primers sequence, annealing temperature and fragment length for the primers used in the study

Primer name	Sequence	Annealing Temp.	Fragment length
***Vrs1-1*F**	5′-TATCTAGAGGAACTCGATGAACTTGAG-3′	56	
***Vrs1-1*R**	5′-GTACCATTGGCCGCGAA-3′	56	695 bp
***Vrs1-2*F**	5′-ACACCAACAGGCAACAGAACAACCTA-3′	60	
***Vrs1-2*R**	5′-GGACGCACATCATCAGGTCATCGT-3′	60	980 bp
***Vrs1-3*F**	5′-CAAACATATGGCCAGCTGCT-3′	56	
Vrs1-3R	5′-TGATCTTCAAGAGAGCTGCCA-3′	56	770 bp
***Vrs4-1*F**	5′-CTGTGCGTCCTTGTAGTGAAGTTGAA-3′	60	
Vrs4-1R	5′-AAGCGCATATCATTCTCCACAT-3′	60	1200 bp
***Int-c-1*F**	5′-TCCTTTCTATGATTCCCCAAGCCCC-3′	60	
Int-c-1R	5′-CCACTCCACCGAGCTCCC-3′	60	1074 bp

Sup. Table 4-2: Known alleles and new haplotypes at the *vrs1* locus (2060 bp) in two- and six-rowed cultivars as well as the wild barley (ssp *spontaneum* and *agriocrithon*) and intermedium collections and the number of accessions carrying these haplotypes .

	- 597	- 531	-513	- 429	- 402	-373	-165	-52	23	75	202	243	349	396	564	680	757	775	918	997	Number of accessions		
ID	G/C	T/G	T/C	CGGCGCAAAAC/-----TTTATT	GT/--	T/G	A/T	T/C	A/G	A/C	T/G	T/-	G/C	C/G	T/C	g/-	A/G	G/A	T/C	T/G	Intermedium	agriocrithon	Spontaneum
vrs1.a1	G	G	C	CGGCGCAAAAC	GT	G	T	C	G	C	G	-	C	C	C	-	G	A	T	G	2	5	
vrs1.a2	C	T	T	CGGCGCAAAAC	GT	G	T	C	G	C	G	T	C	C	T	G	G	A	C	T	-		
vrs1.a3	G	G	C	CGGCGCAAAAC	GT	G	T	C	A	C	T	-	G	C	C	G	G	A	C	G	1		
Vrs1.b1	G	G	C	CGGCGCAAAAC	--	G	T	C	G	C	G	-	C	C	C	G	A	A	C	G	-		
Vrs1.b2	C	T	T	CGGCGCAAAAC	GT	G	T	C	G	C	G	-	C	C	T	G	G	A	C	T	-		13
Vrs1.b3	G	G	C	CGGCGCAAAAC	GT	G	T	C	A	C	T	-	C	C	C	G	G	A	C	G	30		
Vrs.t	C	T	T	CGGCGCAAAAC	GT	G	T	C	G	C	G	-	C	C	T	G	G	G	C	T	-		
Hap_1	C	T	T	CGGCGCAAAAC	GT	T	T	C	G	C	G	-	C	C	C	G	G	A	C	T	182		
Hap_2	C	T	T	CGGCGCAAAAC	GT	T	T	C	G	C	G	-	C	C	C	G	G	A	T	G	41		
Hap_3	C	T	T	CGGCGCAAAAC	GT	T	T	C	G	C	G	-	C	C	C	G	G	A	C	G	23		
Hap_4	C	T	T	CGGCGCAAAAC	GT	T	T	C	G	C	G	-	C	C	C	G	G	A	T	T	4		
Hap_5	C	T	T	CGGCGCAAAAC	GT	T	T	C	G	C	G	-	C	C	C	G	G	A	T	G/T	2		
Hap_6	C	T	T	CGGCGCAAAAC	GT	T	T	C	A	C	T	-	C	C	C	G	G	A	C	G	2		
Hap_7	C	T	T	CGGCGCAAAAC	GT	G	T	C	A	C	T	-	C	C	C	G	G	A	C	G	2		
Hap_8	C	T	T	CGGCGCAAAAC	GT	T	T	C	G	C	G	-	C	G	C	G	G	A	C	T	1		
Hap_9	C	T	T	CGGCGCAAAAC	GT	T	T	C	G	C	G/T	-	C	C	C	G	G	A	T	G	1		
Hap_10	C	T	T	CGGCGCAAAAC	GT	T	T	T	G	C	G	-	C	C	C	G	G	A	C	T	1		
Hap_11	C	T	T	CGGCGCAAAAC	GT	G	T	C	G	C	G	-	C	C	FT/RC	G	G	A	C	T	1		
Hap_12	C	T	T	CGGCGCAAAAC	GT	T	T	C	G	C	G	-	C	C	C	-	G	A	C	T	-	5	
Hap_13	G	G	C	CGGCGCAAAAC	GT	G	T	C	A	C	G	-	C	C	C	G	G	A	C	G	1		
Hap_14	G	G	C	CGGCGCAAAAC	GT	G	T	C	A	C	T	-	C	C	C	-	G	A	C	G	1		
Hap_15	G	G	C	CGGCGCAAAAC	GT	G	T	C	G	C	G	-	C	C	C	G	G	A	C	G	1		
Hap_16	G	G	C	CGGCGCAAAAC	GT	G	T	C	G/A	C	G/T	-	C	C	C	G/-	G	A	C	G	1		
Hap_17	G	G	C	CGGCGCAAAAC	GT	G	T	C	A	C	FT/RG	-	C	C	C	G	G	A	C	G	2		
Hap_18	G	G	C	_-----TTTATT	GT	G	T	C	G	C	G	-	C	C	C	G	G	A	C	G	3		
Hap_19	G	G	C	CGGCGCAAAAC	GT	G	A	C	G	A	G	-	C	C	C	G	G	A	C	G	-		21
Hap_20	G	G	T	CGGCGCAAAAC	GT	G	T	C	G	C	G	-	C	C	C	G	G	A	C	G	-		66
Total																					302	10	100

Sup. Table 4-3: Known alleles and haplotypes at the *int-c* locus (1074 bp) in two- and six-rowed cultivars as well as the wild barley (ssp *spontaneum* and *agriocrithon*) and intermedium collections including the amino acids change and the number of accessions carrying these haplotypes .

	124	**189-194**	**235**	**284**	**437**	**498**			
ID	**C/G**	**CGGTGA/------**	**C/T**	**T/C**	**A/T**	**T/C**	**Inter.**	**Agrio.**	**Spont.**
Int-c.a	G	CGGTGA	T	C	A	C	291	10 agrio	34
Int-c.a Haplotype1	G	CGGTGA	T	C	T	C	-		
HvTB1.spontc	G	CGGTGA	T	C	A	C			14
int-c.b1	C	------	C	T	A	C	4		37
int-c.b2	C	------	C	T	A	T	-		
HvTB1.spontd	G	------	T	C	A	C			15
Int-c.a Haplotype2	G	CGGTGA	T	C	W	C	7		
AA Change	Ala to Pro	*	Ser to Pro	Ala to Val	Asp to Val	None	302	10	100

* Position cause the amino acid change from 358 aa in six-rowed allele to 356 aa in two-rowed allele

Sup. Table 4-4: Known haplotypes at the *vrs4* locus (1215 bp) in two- and six-rowed cultivars as well as the wild barley (ssp *spontaneum* and *agriocrithon*) and intermedium collections including the amino acids change and the number of accessions carrying these haplotypes .

	-148 -139	- 47	- 18 - 17	102	141	567			
ID	**TTCCCAGAAC/----------**	**G/A**	**AT/--**	**A/G**	**T/C**	**G/A**	**inter**	**agrio**	**spont**
Hap.1_Morex	TTCCCAGAAC	A	--	G	C	A	239	10	63
Hap.2	TTCCCAGAAC	A	AT	A	T	G	48		37
Hap.4	TTCCCAGAAC	G	/--	G	C	A	1		
Hap.3	_----------	A	/--	A	T	G	13		
Hap.5	_----------	A	/--	G	C	A	1		
AA change				None	None	None	302	10	100

10.0 Curriculum Vita

PERSONAL INFORMATION

Last Name **: Youssef Ibrahim**

First and middle names **: Helmy Mohamed**

Date of Birth **: June 14, 1973**

Gender **: Male**

Citizenship **: Egyptian**

Marital Status **: Married, 4 children**

Home Address (Egypt) **: Arab Elsaha, Elhawamdia, Giza, Egypt**

Present; Finkenweg2, 06466 Gatersleben Germany.

ACADEMIC & PROFESSIONAL INFORMATION

UNIVERSITY	DEGREE/ DIPLOMA	SUBJECTS	YEAR OF PASSING	PERCENTAGE
CAIRO	B.Sc. Ag. I	-AGRI.	1993	V. Good
CAIRO	B.Sc. Ag II	-AGRI.	1994	Good
CAIRO	B.Sc. Ag III	-Agric. Production	1995	V. Good
CAIRO	B.Sc. Ag IIII	- Agric. Production	1996	V. Good
CAIRO	M.Sc. Ag	- Agricultural Science	1998-2002	-
CAIRO	PhD	Stress Physiology and Biotechnology	2003 - 2008	-

PhD Title: Production of wheat plants more tolerant to abiotic stress via modern applied biotechnology techniques.

Employment history

Lecturer: Plant Physiology, Faculty of Agriculture, Cairo University, Giza, Egypt. From 2009 to present**.**

Assistant lecturer: Plant Physiology, Faculty of Agriculture, Cairo University, Giza, Egypt. From 2002 to 2008**.**

Demonstrator: Plant Physiology, Faculty of Agriculture, Cairo University, Giza, Egypt. From 1996 to 2002

LANGUAGE SKILLS

Arabic **Excellent**

English V Good

German Good

DETAILS OF PREVIOUS TRAINING PROGRAMMES ATTENDED

INSTITUTION	PROGRAMME	DURATION	YEAR
-IPK-Gatersleben, Germany	Plant Architecture	-------	May 2010 to present
-UFV Brazil, TWAS/CNPq fellowship	Molecular Plant Physiology	6 Months	1/4/09-1/10/09
- DAAD Germany	Sustainable Development as Strategy and Aim of Higher Education	Two weeks	25-3 to 4-4-09
-Sida, Sweden (2nd phase) in Egypt as a host	Plant Breeding and Seed Production	Two weeks	1-14/3/2008
- Sida, Sweden(1st phase)	Plant Breeding and Seed Production	One month	1-30/9/2007
-Hannover University, German Egyptian year	Biotechnology and Salinity Stress	Two weeks	1-15/7/2007
-In-Went, Germany	Plant Biotech and Biosafety	One year	1-10-2005 to 30-9-2006
-Bio-Center, Halle Uni, Germany.	Theoretical and Practical Basic Course, Development oriented Biotech / Plant Genetic Resources.	One month.	2006
- In-Went, Germany	Improvement of Teaching Quality for Colleges of Agriculture and Forestery.	33 Days	2003
- IAC, The Netherlands	- Training Programme on Biotechnology, Plant Breeding and Seed Sector Development.	38 Days	2003

- IAC, The Netherlands	- International Course on Molecular Markers in Plant Breeding	one week	2003
- IAC, The Netherlands	- International Course on Plant Biotechnology and Biosafety	two weeks	2003
- IAC, The Netherlands	- International Course on Breeding for Resistance	two weeks	2003
- Cairo Univ., Egypt	- DNA fingerprints and molecular markers	8 Days	1999
-Ain Shams Univ., Egypt	- Recombinant DNA technology	one week	1998

Skills:

- DNA and RNA extraction from plants
- PCR and RT-PCR technologies.
- Gel Electrophoresis,
- DNA purification, RNA and DNA Quantitation.
- DNA sequencing and analyses
- Primer design and
- genetic mapping

Teaching Experience: since 1998 to 2010, I was teaching in Plant Physiology Division, Botany Department, Faculty of Agriculture, Cairo University. Courses:

- **Plant physiology**
- **Plant nutrition**
- **Seed physiology**
- **Physiology of flowering in plants.**
- **Physiology of plant under stress**
- **Plant tissue culture.**

Presentations and seminars:

Presentations:

- Teaching system in Faculty of Agriculture, Cairo University (InWent-Germany).

- Training courses and Working System in Plant Biotechnology Research Lab ((InWent- Germany).

Seminars:

- Plant Biotechnology and Biosafety in Egypt (Wageningen – The Netherlands).
- Using Biotechnology Techniques to improvement crops Productions (Faculty of Agriculture –Cairo University –Egypt).
- Internationalen Fachseminar on Moderne Umwelttechnologien 16.12.-.18.12.2005, Berlin, Germany.

Posters in international conferences:

The 22nd International Triticeae Mapping Initiative (ITMI) June 20-24, 2012, Fargo, North Dakota, USA. **Re-sequencing of *vrs1* and *int-c* loci shows that *labile* barleys (*Hordeum vulgare* convar. *labile*) have a six-rowed genetic background. Helmy M. Youssef**, Ravi Koppolu, Thorsten Schnurbusch,

11th International Barley Genetics Symposium (IBGS), April 15 -20, 2012, Hangzhou, China. **Re-sequencing of *vrs1* and *int-c* loci shows that *labile* barleys have a six-rowed genetic background. Helmy M. Youssef**, Ravi Koppolu, Thorsten Schnurbusch.

VIPCA, Plant Genetics and Breeding Technologies, 18-20 February 2013, Vienna, Austria. **Six-rowed spike 4 (Vrs4) mediates spikelet architecture and row-type identity in barley.** Ravi Koppolu, Nadia Anwar, Shun Sakuma, Akemi Tagiri, Udda Lundqvist, Twan Rutten, Christiane Seiler, Mohammad Pourkheirandish, Axel Himmelbach, Ruvini Ariyadasa, **Helmy M. Youssef**, Nils Stein, Nese Sreenivasulu, Takao Komatsuda, Thorsten Schnurbusch.

VIPCA, Plant Genetics and Breeding Technologies, 18-20 February 2013, Vienna, Austria. **Does six rowed spike 2 (vrs2) play a role in regulating barley lateral spikelet fertility? Helmy M. Youssef**, Arash Fazeli, Ravi Koppolu, Twan Rutten, Ruvini Ariyadasa, Axel Himmelbach, Nils Stein, and Thorsten Schnurbusch.

The Plant and Animal Genome (PAG) XXII Conference, January 11-15, 2014 – San Diego, CA, USA. **Genetic control of spikelet determinacy and lateral spikelet fertility in barley.** Ravi Koppolu, Nadia Anwar, Shun Sakuma, Akemi Tagiri, Udda Lundqvist, Twan Rutten, Christiane Seiler, Mohammad Pourkheirandish, Axel Himmelbach, Ruvini Ariyadasa, **Helmy M. Youssef**, Nils Stein, Nese Sreenivasulu, Takao Komatsuda, Thorsten Schnurbusch.

Cereals for Food, Feed and Fuel – Challenge for Global Improvement, Joint **EUCARPIA** Cereal Section & **I•T•M•I** Conference | June 29 - July 4, 2014 | Wernigerode, Germany. **Genetic mapping of the labile (lab) gene: a recessive locus causing irregular spikelet fertility in *labile*-barley (*Hordeum vulgare* convar. *labile*). Youssef, Helmy Mohamed**; Ravi Koppolu; Twan Rutten; Viktor Korzun; Patrick Schweizer and Thorsten. Schnurbusch.

The Plant and Animal Genome (PAG) XXIII Conference, January 10-16, 2015 – San Diego, CA, USA. **Genotypic and phenotypic analyses of wild and Intermedium barleys highlights two distinct row-types in *Hordeum spontaneum*. Helmy M Youssef**, Martin Mascher, Mohammad Ayoub, Benjamin Kilian, Nils Stein and Thorsten Schnurbusch.

Oral presentation in international conferences and meetings:

BarleyGenomeNet (BGN) Meeting 2013, 11-12 Dec., 2013. Faculty of Science, Department of Plant and Environmental Sciences, University of Copenhagen, Copenhagen, Denmark. **Row-type genes in barley: latest analyses and updates.**

List of Publications:

Naser Poursarebani, Tina Seidensticker, Ravi Koppolu, Corinna Trautewig, Piotr Gawroński, Federica Bini, Geetha Govind, Twan Rutten, Shun Sakuma, Akemi Tagiri, Gizaw M.Wolde, **Helmy M. Youssef**, Abdulhamit Battal, Stefano Ciannamea, Tiziana Fusca, Thomas Nussbaumer, Carlo Pozzi, Andreas Börner, Udda Lundqvist, Takao Komatsuda, Silvio Salvi, Roberto Tuberosa, Cristobal Uauy, Nese Sreenivasulu, Laura Rossini, Thorsten Schnurbusch (2015) **The genetic basis of composite spike form in barley and 'Miracle-Wheat'.** Genetics. doi:10.1534/genetics.115.176628

http://www.genetics.org/content/early/recent

Youssef, Helmy Mohamed; Ravi Koppolu; Twan Rutten; Viktor Korzun; Patrick Schweizer and Thorsten. Schnurbusch (2014). **Genetic mapping of the labile (lab) gene: a recessive causing irregular spikelet fertility in *labile*-barley (*Hordeum vulgare* convar. *labile*).** *Theor Appl Genet* 127:1123-1131 ***Theor Appl Genet*** E-pub doi10.1007/s00122-014-2284-0

Ravi Koppolu, Nadia Anwar, Shun Sakuma, Akemi Tagiri, Udda Lundqvist, Mohammad Pourkheirandish, Twan Rutten, Christiane Seiler, Axel Himmelbach, Ruvini Ariyadasa, **Helmy M. Youssef**, Nils Stein, Nese Sreenivasulu, Takao Komatsuda and Thorsten Schnurbusch (2013). **Control of spikelet determinacy and row-type in barley.** ***Proc Natl Acad Sci U S A*** 110: 13198–13203

http://www.pnas.org/content/110/32/13198.abstract

Youssef, Helmy Mohamed; Ravi Koppolu and Thorsten. Schnurbusch (2012) **Re-sequencing of vrs1 and int-c loci shows that labile barleys (*Hordeum vulgare* convar. *labile*) have a six-rowed genetic background.** Genet Resour Crop Evol (2012) 59:1319-1328

http://link.springer.com/article/10.1007%2Fs10722-011-9759-5

Helmy M Youssef ; Mohammad Ayoub; Martin Mascher; Nils Stein and Thorsten Schnurbusch (2014) **Genotypic and phenotypic analyses of wild and Intermedium barleys highlights two distinct row-types in *Hordeum spontaneum*.** (submitted to JXB)

Hereby I declare that all information stated above is true

[ ]
Helmy Mohamed Youssef Ibrahim

11.0 Affirmation/ eidesstattliche Erklärung

Herr Helmy Mohamed Youssef Ibrahim
DFG Heisenberg Plant Architecture Group,
Leibniz-Institute of Plant Genetics and Crop Plant Research (IPK)
Corensstr.3, OT Gatersleben,
D-06466 Stadt Seeland,
Germany.

I hereby declare that the submitted thesis has been completed by me, the undersigned, and that I have not used any other than permitted reference sources or materials or engaged any plagiarism. All the references and the other sources used in the presented work have been appropriately acknowledged in the work. I further declare that the work has not been previously submitted for the purpose of academic examination, either in its original or similar form, anywhere else.

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

Hiermit erkläre ich, dass ich diese eingereichte Dissertation selbständig angefertigt und keine anderen als die angegebenen Quellen und Hilfsmittel verwendet habe. Die den benutzten Hilfsmitteln wörtlich oder inhaltlich entnommenen Stellen habe ich unter Quellenangaben kenntlich gemacht. Die vorliegende Arbeit wurde in gleicher oder ähnlicher Form noch keiner anderen Institution oder Prüfungsbehörde vorgelegt.

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

Gatersleben, den ----------

Helmy Mohamed Youssef Ibrahim