Cloning of the plant development regulatory genes *MANY NODED DWARF* (*MND*) and *LAXATUM-A* (*LAX-A*) by taking advantage of an improved barley genomics infrastructure.

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

Herrn Dipl.-Ing. agr. Matthias Jost

Geb. am 29.06.1986 in Rochlitz

Gutachter:
Prof. Dr. Andreas Graner
Prof. Dr. Gary Muehlbauer
Dr. Sarah McKim

Verteidigt am 25.01.2016
Halle / Saale

Publications

Parts of this thesis have been published or are intended to be published in the following articles:

Jost, M.*, Taketa, S.*, Mascher, M., Himmelbach, A., You, T., Shahinnia, F., Rutten, T., Druka, A., Schmutzer, T., Steuernagel, B., Beier, S., Taudien, S., Scholz, U., Morgante, M., Waugh, R., Stein, N.: A homolog of Blade-On-Petiole 1 and 2 (BOP1/2) controls internode length and homeotic changes of the barley inflorescence. **Plant Physiol. (2016).** Epub ahead of print: **dx.doi.org/10.1104/pp.16.00124**

Mascher, M.*, Jost, M.*, Kuon, J.E., Himmelbach, A, Assfalg, A., Beier, S., Scholz, U., Graner, A. & Stein N.: Mapping-by-sequencing accelerates forward genetics in barley. **Genome Biology 15 (2014)** R78.

* Co-first authors

Contents

List of figures

List of tables

Appendix tables

1. Introduction

1.1. General introduction

Barley (*Hordeum vulgare* L.) is the fourth most important cereal in the world after wheat (*Triticum aestivum* L.), rice (*Oryza sativa*) and maize (Zea mays L.) based on the production area. In Germany, barley ranks second after wheat in terms of production with a harvest of 10 million tonnes and 1.6 mio ha cultivation area (FAO 2013, http://faostat.fao.org). Due to its adaptability to diverse climatic conditions and geographic extremes barley has the widest range of production in the world (Bothmer, 2003). In some high altitude areas like the Himalaya or the highlands at the Horn of Africa, barley remains the only cereal which can be cultivated producing reliable yields. In most industrialized countries barley is cultivated as animal feed and as a resource for the malting and brewing industry (Bothmer, 2003). Barley has also high relevance as a food source, especially in the highlands of Tibet, Africa and the Andes. Recently, new interest in barley as a component of healthy food products was generated due to reports on positive dietary effects of barley, such as lowering blood cholesterol, blood pressure and improving the glycaemic index (Ullrich, 2008).

Barley plant architecture (Figure 1) changes during three major growth phases: the (i) vegetative, (ii) reproductive and (iii) grain filling phases (Newman, 2008). The vegetative growth phase starts at seed germination and includes leaf initiation and plant establishment. It is being followed by the reproductive phase. This is characterized by conversion of the vegetative shoot apical meristem into a generative meristem leading to the formation of tillers carrying the inflorescence (Newman, 2008). A tiller (culm) emerges from an axillary bud (axillary meristems) at the proximal plant shoot. It is an elongated stem with leaves initiated at nodes which are separated by extended internodes (Hussien et al., 2014). Barley leaves emerge in an alternating spatial pattern (phyllotaxis) on opposite sides of the rachis spaced by a regular timing (plastochron) of the interval between occurrence of two successive leaves (Hill and Lord, 1990). The basis of each leaf forms a tubular leaf sheath which at about the distance of the next subsequent culm node converts into the leaf blade. At this intersection, the leaf margin forms characteristic outgrowths named auricles whereas at the bottom of the leaf blade a small organ – the ligula, arises. One single inflorescence (ear, spike) is formed at the top of each tiller, which is characterized by a single unbranched main rachis that carries a triplet of sessile single-floreted spikelets at opposing sides of subsequent rachis nodes. This

leads to the characteristic two-rowed (lateral spikelet's infertile) or six-rowed (all three spikelets at a node are fertile) phenotype of the ear of domesticated barley. Each fertile barley floret is composed of a central carpel surrounded by a whorl of three stamen and two lodicules, all enclosed by a leaf-like structure, the palea and the typically long-awned lemma (Figure 1) (Kellogg, 2001).
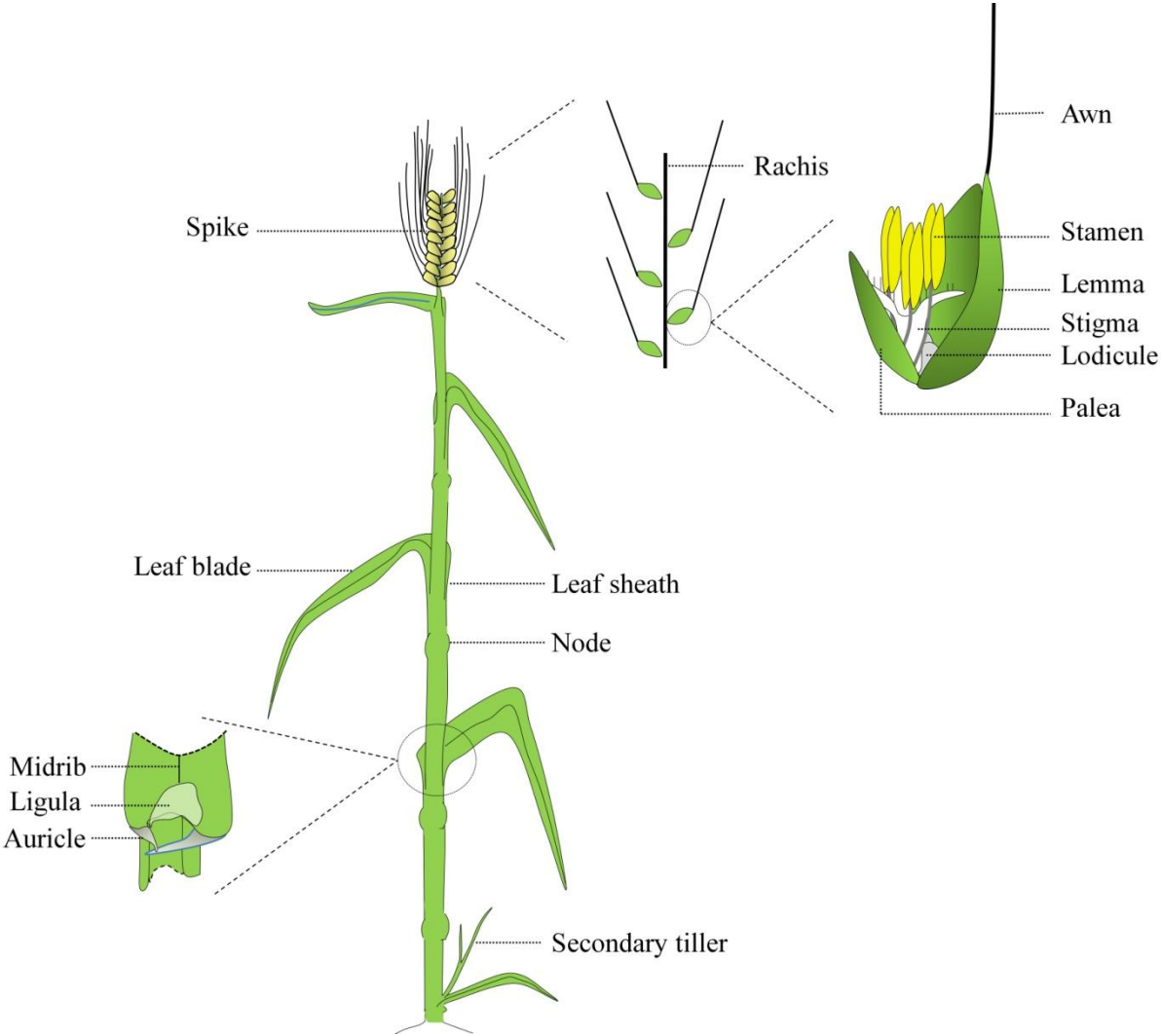
Figure 1 : Schematic overview of barley plant architecture.
For simplicity and visibility reasons the representation of the plant is reduced to a single tiller. Characteristics of leafs, spike architecture and floret composition are shown in detailed views.

Breeding for architectural traits can be important to improve yield. For instance, a major prerequisite for the first 'Green Revolution' was the characterization and introduction of dwarfism genes into varieties to strengthen plants for carrying more heavy and high yielding inflorescences and thus reduce the risk of lodging would lead to yield loss (Hedden, 2003). The first characterized barley dwarfing gene *Uzu1* was identified as member of the Brassinosteroid phytohormone pathway (Chono et al., 2003). This information guided research towards the identification of three additional involved genes (Dockter et al., 2014). The subsequently identified series of mutant alleles can now be utilized to fine-tune the brassinosteroid metabolism to improve lodging resistance in barley (Vriet et al., 2012; Dockter et al., 2014). The gene *Ideal Plant Architecture 1* (*IPA1*) of rice has been reported as important regulatory gene of rice plant architecture with major influence on yield potential (Jiao et al., 2010). A disturbed micro RNA recognition site causes higher transcript accumulation of *IPA1* which leads to plants with a reduced number of tillers in combination with increased plant height and panicle branches. Such growth habit reflects the defined ideal type of plant growth in rice (Khush, 1995). The term ideal plant architecture refers to a plant architecture which enables optimal use of available resources to maximize plant performance. This is mainly defined by an optimal canopy structure to maximize crop photosynthesis, low tiller number for an increased number of grains towards a high grain to straw ratio (harvest index) under repression of unproductive tillers and stable plants against lodging (Donald, 1968; Khush, 1995).

The barley inflorescence carrying the seeds is the harvested part of the plant. Morphology and architecture of the barley spike may impact yield directly. However, the molecular genetic factors underlying barley inflorescence architecture and development have only been revealed for a few characters. Major genes controlling traits like row type [*Vrs1* (Komatsuda et al., 2007), *int-c* (Ramsay et al., 2011), and *Vrs4* (Koppolu et al., 2013)], cleistogamous flowering [*cly1* (Nair et al., 2010)], hulled-ness of the caryopsis [*NUD*, (Taketa et al., 2008)], elongation of awns and pistil morphology [*Lks2* (Yuo et al., 2012)], awn conversion into cap-like structures [*HOODED* (*K*), (Muller et al., 1995)], bract suppression [*Trd1* (Houston et al., 2012)], spike density [*Zeo1b* (Houston et al., 2013)] or spike-branching [COM2, (Poursarebani et al., 2015)] were recently cloned. An extensive collection of seeds from cultivated (*Hordeum vulgare* subsp. *vulgare*) and wild barley (*Hordeum vulgare subsp. spontaneum*) accessions have been maintained by long term seed storage in *ex situ* seed banks and can be used to study natural diversity of genes. This material was used to trace back the

role of genes during barley domestication and revealed the importance of genes regulating architecture traits. For instance, the six-rowed spike type arose by three independent mutation events in the gene *Six-rowed spike 1* (*Vrs1*) of a two-rowed barley (Komatsuda et al., 2007). The today's cleistogamous flowering barley cultivars has two distinct origins which were geographically separated from barley's origin in the 'Fertile Crescent' (Nair et al., 2010). These changes in architecture occurred spontaneously by mutations in genes and were actively selected to maintain the correlated positive traits for barley cultivation. The identification of genes controlling architectural traits is the prerequisite to study the underlying regulating mechanisms and how mutations in these genes affect plant development. This knowledge can be utilized for designing selection schemes in breeding programs towards development of the ideal plant architecture to optimize yield in important crops (Donald, 1968).

This PhD thesis focused on the analysis of two induced architectural mutants of barley obtained from large mutant collections which were generated in the fifties of the last century. First, the *many noded dwarf* (*mnd*) mutant exhibits a higher number of leaves, nodes and tillers in combination with a dwarfed growth habitus. Second, the *laxatum-a* (*lax-a*) mutant is affected in spike length, width of awn basis (leafy awn) and glume development (Harlan, 1922; Larsson, 1985b). In this study, the molecular cloning of both underlying genes is described. The cloning of these genes was supported by applying novel sequencing-based strategies taking advantage of the improved genomic resources of barley. The identified genes were validated by characterizing historical and recent mutant plant resources. Both, the barley genomic infrastructure and plant resources for functional validation are introduced in the following paragraphs.

1.1.1. *Many noded dwarf* (*mnd*) **mutant**

A barley plant with an increased number of nodes was described in detail in 1922 by Harlan and colleagues (Harlan, 1922). They reported a mutant occurring spontaneously in cv. Mesa showing extremely short internodes and an increased number of nodes. Plants carried also an increased number of tillers which were induced by secondary stem branches at lower nodes of the culms. Under field conditions, plants showed a semi-dwarfed growth habitus and developed slightly shorter but fertile spikes. Under greenhouse conditions, plants grew taller with continuously increased number of nodes as compared to wild-type plants. Spikes were

deformed and showed fertility problems. Plants could grow to enormous height due to branching at the basis of the main inflorescence leading to emergence of additional tillers. Spikes formed at such secondary tillers often remained without seeds (Harlan, 1922). This mutant allele was named *mnd1* and was allocated to chromosome 4H (Druka et al., 2011). Other loci affecting many-noded dwarf characteristics were reported as *mnd4* and *mnd6* on chromosome 5H, *mnd3* on 3H and *mnd5* without chromosomal assignment (Walker et al., 1963; Franckowiak, 1995; Pozzi et al., 2003).

1.1.2. *Laxatum* (*lax*) mutant

The term laxatum refers to less compact ('lax') spikes caused by extended length of rachis internodes as if compared to wild-type plants. The *lax-a* mutants display a set of pleiotropic phenotypic characteristics: (i) awns show a very wide (leafy) base, (ii) grains are thin, angular and exposed, (iii) a homeotic conversion of the lodicules into two additional anthers, thus florets contain five instead of three anthers. The gene locus was roughly assigned previously to chromosome 5H by RFLP mapping (Laurie et al., 1996). Two additional *laxatum* loci, *lax-b* and *lax-c*, were assigned to chromosome 6H (Larsson, 1985a). *Lax-c* has been described to have similar spike morphology changes like *lax-a* but misses the transformation of lodicules into stamen. In contrast, *lax-b* was described to have a shrunken endosperm which is already visible at heterozygous stage. In addition, *lax-b* spikes are short and carry low number of grains resulting in a lax spike morphology; all in absence of other *lax-a* described changes (Larsson, 1985b).

1.2. Barley genomic infrastructure

Barley is a diploid species with a genome size of 5.1 Gbp (Dolezel and Bartos, 2005), which is forty times larger than the *Arabidopsis* genome (125 Mb), nineteen times the *Brachypodium* genome (272 Mb), thirteen times the rice genome (389 Mb) and seven times the sorghum (736 Mb) genome (The Arabidopsis Genome Initiative, 2000; Dolezel and Bartos, 2005; Matsumoto et al., 2005; The International Brachypodium Initiative, 2010). The predicted number of genes for barley (26.159) (IBSC, 2012), *Arabidopsis* (25.498), *Brachypodium* (25,532), rice (31.439) and *Sorghum* (27.640) is rather conserved, which is in contrast to the

big differences in genome size. Differences in genome size are mainly attributed to the content of repetitive DNA sequences, which mainly consist of mobile elements or related sequences (retrotransposon, Flavell et al., 1974; IBSC, 2012).

Barley is one out of over 10,000 species of the grass family (*Poaceae* or *Gramineae*) that have diverged from a common ancestor about 55–70 million years ago (Kellogg, 2001). Homology of genomes can be observed within these plant family members although being disturbed by chromosomal rearrangements, inversion, translocations or duplications that occurred during divergence of these species. Nevertheless, comparative genomics among crop species revealed large blocks of conserved collinearity in marker order and gene organization remained in chromosomal segments; information that can be utilized to transfer knowledge between species (Ahn et al., 1993; Moore et al., 1995; Thiel et al., 2009). The general term of homologous chromosomal segments with conserved gene order is called 'synteny' (Freeling, 2001).

Synteny information, high density genetic maps and next generation sequencing technology (NGS) provided the basis for a first gene based reference model for the barley genome. Barley chromosome arms were survey-sequenced by NGS and sequence contigs were used to predict a virtual linear order of 21,766 barley genes by the help of a high density genetic map (Close et al., 2009) and synteny relationships to the sequenced genomes of rice, *Sorghum* and *Brachypodium* (Mayer et al., 2011). To date, still no complete reference sequence is available for barley. However, in 2012, a physical and genetic map framework of barley with integrated whole genome sequence information was published and which can be utilized as first reference genome sequence model (IBSC, 2012). More recently, more than 10 million SNPs derived from population sequencing (POPSEQ) were utilized for an improved genetic anchoring of the sequence resources (Mascher et al., 2013b). Both together promise the possibility for the application of novel approaches facilitating gene isolation in barley. These genomic resources are introduced in greater detail in the following paragraphs.

1.2.1. Genetic reference maps

By following the inheritance of morphological traits in hybrid pea plants, Gregor Mendel discovered the laws of genetic segregation underlying the concept of genetic mapping (Mendel, 1866). After the identification of chromosomes, Thomas Hunt Morgan described first, that traits were inherited as groups, if they were linked to each other on the same

chromosome. He proposed the exchange of segments from homologous chromosome pairs (Morgan, 1915) and such frequencies of exchange during meioses (crossing over) would increase with the distance between genes on the same chromosome. The unit of genetic distance was named centi Morgan (cM) in honor of Thomas Hunt Morgan. One centiMorgan distance between two loci describes 1% of recombination frequency among the total number of analyzed meioses (Tropp, 2011). Based on these early discoveries the first genetic linkage maps could be developed mainly on the basis of different segregating phenotypes caused by variant alleles of genes. Since morphological differences between genotypes that could serve as markers are usually rather limited, genetic mapping remained inefficient until the discovery of molecular markers. A molecular marker is based on changes in the DNA or protein sequence, that do not need to be associated with any measurable phenotypic effect (Griffiths, 2004). First molecular markers were isoenzymes with small changes in their amino acid sequence causing activity changes that could be measured by simple gel electrophoresis (Tanksley, 1983; Schulman, 2006). They were like morphological markers, relatively limited in number and can be influenced by environmental effects. Therefore, they have been quickly replaced after the identification of DNA-based molecular markers (Schulman, 2006). Restriction Fragment Length Polymorphisms (RFLP) were the first DNA-based molecular marker system (Botstein et al., 1980) to be used to establish a genome wide molecular marker map comprising more than 200 loci for the barley genome (Graner et al., 1991). RFLP markers were soon replaced by the much more efficient polymerase chain reaction (PCR) (Saiki et al., 1988) based marker systems: Sequence Tagged Site (STS), Random Amplified Polymorphic DNA (RAPD), Simple Sequence Repeat (SSR) / microsatellites, Amplified Fragment Length Polymorphism (AFLP) and Single Nucleotide Polymorphism (SNP) (Olson et al., 1989; Williams et al., 1990; Weissenbach et al., 1992; Vos et al., 1995; Landegren et al., 1998).

Multiple mapping populations and multiple marker types were used to develop dense gene-based consensus maps. For instance 1,032 expressed sequence tags could be integrated in a genome wide map of a density of less than 1 cM/ marker by different marker types (RFLP, SSR, SNP) (Stein et al., 2007). These gene based markers could be utilized for comparative genomics with other grasses and served as a source for marker saturation in gene cloning projects.

Co-dominant SNP based markers became a standard to establish high density genetic maps. For instance, 3,072 EST-derived SNPs were placed on two Illumina Golden Gate assays

(Illumina Inc., San Diego, CA, USA) BOPA1 and BOPA2 which enabled the development of a consensus map with 2,943 integrated SNPs (Close et al. 2009). Recently, a barley 9K iSelect chip became available comprising 7,864 SNPs on an Illumina Infinium genotyping assay (Illumina Inc., San Diego, CA), including the former BOPA markers. Genotyping 360 recombinant inbreed lines (RIL) of a segregating population (cultivars Morex x Barke) produced a genetic map comprising 3,973 loci (Comadran et al., 2012). Since recently, high throughput sequencing technology - 'Next Generation Sequencing' (NGS) - facilitates genome wide SNP discovery for markers which can be applied to construct maps with more than 30,000 SNPs in a biparental population (Elshire et al., 2011; Poland et al., 2012). These high density genetic maps have a fundamental relevance towards establishing a genome reference of barley which is described in the following paragraph.

1.2.2. Sequence enriched physical map

The size and complexity of the barley genome are the main reasons why there is still a lack of a high quality reference genome sequence of barley. The International Barley Sequencing Consortium decided to establish a physical map consisting of overlapping DNA fragments reproducing the linear chromosomes as a prerequisite for map based cloning (described in part 1.4.1) and complete genome sequencing (Schulte et al., 2009). To construct a physical map of barley, genomic DNA was fragmented and cloned into bacterial artificial chromosome (BAC) libraries. 571,000 barley DNA containing BACs were fingerprinted to identify overlaps between BACs to form overlapping BAC contigs. In total 9255 BAC contigs were identified by this approach representing 4.98 Gb (95 %) of the barley genome (Schulte et al., 2011; IBSC, 2012; Ariyadasa et al., 2014). BAC clones could be assigned to genetic markers from high density genetic maps to provide 4,556 BAC contigs (3.9 Gb) with a genetic position along the barley genome (IBSC, 2012). Whole genome shotgun sequencing (WGS) was performed for cultivars Morex (55-fold sequence read coverage), Bowman (35 fold coverage) and Barke (30-fold coverage), respectively. A large number of relatively short sequence contigs were obtained by assembling the sequence reads. Around 5,000 gene containing BACs were sequenced and additional sequence tags of around 300,000 BACs were generated by Sanger sequencing (Sanger et al., 1977) from the vector into the cloned DNA fragment. Sequence similarity search by BLAST (Altschul et al., 1990) allowed the direct integration of 308 Mbp sequence information to the physical map. RNA sequencing (RNA-seq) data obtained of samples from eight different tissue types were used together with

previously published sequences from multiple full-length cDNA libraries (Matsumoto et al., 2011) to predict genes on WGS contigs. 26,159 high - confidence genes with homology support of at least one sequenced genome of *Brachypodium*, rice and *Sorghum* and *Arabidopsis* could be revealed. 53,220 transcript clusters were assigned as low-confidence genes due to a lack of homology and cluster formation. Among the high-confidence genes, 15,719 could be directly assigned to the physical map frame work; 3,743 positioned by conserved synteny analysis and 4,692 assigned to chromosome arms. Thus around 20,000 high confidence genes were represented within the physical-genetic framework and serve together with anchored sequence information as a draft reference sequence of the barley gene space (IBSC, 2012). Additional genetic markers obtained by skim sequencing of two segregating populations (POPSEQ) provided many more genetic anchor points for integrating more sequence information directly to the genetic map of barley (Mascher et al., 2013b; Ariyadasa et al., 2014).

1.3. Gene identification in Barley

To get insight into regulation of morphological and physiological traits requires the molecular characterization of the respective underlying genes. Two strategies were described to study gene functions: (i) forward genetics and (ii) reverse genetics. The forward genetic approach addresses a specific phenotype and aims to identify the gene which controls the characteristic trait (Hricová, 2010). For instance, seeds were treated with a mutagen (e.g. x-ray radiation) and revealed a plant that differs phenotypically from the original wild-type plant. The forward genetics approach can be used to identify the underlying mutated gene which is responsible for this morphological change. In barley, map-based cloning became a standard forward genetics approach for gene identification (Stein, 2005). Recently, a new innovative method was established for gene identification in model organisms with complete reference genomes by applying next generation sequencing (Schneeberger and Weigel, 2011). The above described achievements of genomic infrastructure improvement in barley should principally also allow adapting these technologies to barley. The reverse genetics methods in barley are gaining importance for validating gene functions to confirm the identified candidate genes of a forward genetic screen. This approach aims to test the impact of a gene on the phenotype by direct modification of the gene, which leads to a reduced or complete loss of function (Hricová, 2010). Those changes in gene function can be introduced by mutagenesis as well by

transgenic approaches. The steps towards isolation and functional validation of genes in barley will be detailed described in the following paragraphs.

1.3.1. Positional cloning

Map-based or positional cloning starts with the development of a small segregating mapping population derived from a cross of genotypes with contrasting phenotypes for the trait of interest. The segregation of the phenotype within the population provides the information if the trait is controlled by one, two or multiple genes. The analysis with genome wide markers allows to identify the linkage group carrying the gene of interest and to select suitable flanking markers for high resolution mapping. This may require the development of further marker in the target interval which can be tested for normal segregation behavior. High resolution mapping in a larger population is required to increase the genetic resolution around the gene of interest. The aim is to identify markers in close linkage that would allow for the physical delimitation in a large DNA insert library. The physical distance behind the genetically defined interval strongly depends on the recombination frequency of the target area (Kuenzel, 2000). Thus, a large number of meiotic events might be necessary for a sufficient genetic resolution to identify recombination events in close proximity to the gene loci located in regions with strongly reduced recombination frequency, like known for the peri-centromeric area (IBSC, 2012). The ideal case, when two flanking markers hit the same DNA fragment of a single BAC clone, is called chromosome landing (Tanksley et al., 1995). If not, chromosome walking is required for stepwise identification of overlapping BAC clones to extend the physical contig. This can be achieved by taking newly generated sequence information (e.g. sequencing of the BAC ends) of the identified BACs to screen again the DNA library to identify the next neighboring overlapping BAC clone (Stein, 2005). However, this can be a time consuming process depending on the distance that needs to be bridged and due to the high amount of repetitive sequence of barley that complicates the development of unique probes / markers from identified BACs in order to allow for specific identification of a BAC clone extending the contig (Tanksley et al., 1995). In barley, the above mentioned genome-wide physical map can be used as a resource of overlapping BAC clones making the general procedure of chromosome walking redundant. However, in recombination poor regions, like the genetic centromere, BAC contigs cannot be ordered due to a lack of genetic resolution (Ariyadasa et al., 2014) and would request flanking markers located on BACs

which belongs to the same BAC contig. The identified BAC clone or overlapping BAC clones then have to be sequenced and annotated for candidate gene identification.

1.3.2. Mapping-by-sequencing

It was shown that NGS can be used to identify a high number of SNPs which allowed a precise mapping of recombinant inbreed lines (RIL) in rice (Huang et al., 2009). A RIL population was generated by a cross of the *indica* and *japonica* rice, whose complete genome sequences are available (Yu et al., 2005). The entire population was sequenced to low genome coverage of around 0.2-fold per individual RIL. The obtained short sequence reads of each RIL were mapped against both reference accessions for SNP calling. The low sequence coverage caused missing data for a number of RILs at a certain SNP position. Thus a sliding window of 15 SNPs was used to screen along the genome to define genotype calls by taking into account how many of them representing the respective parental allele. An *indica*:*japonica* SNP ratio of 11:4 or higher was assigned as homozygous indica (ind/ind). A ratio of 2:13 or lower was called as homozygous *japonica* (jap/jap) and any ratio in between was called heterozygous (ind/jap). The differences in the thresholds related to differences in a SNP error rate indentified by including parental lines as controls in the sequencing. This genotype call was used to calculate a map with average recombination breakpoints every 40 kb and enabled to map a QTL for plant height to a 100 kb interval (Huang et al., 2009).

A simplified application using NGS for mapping was introduced by Schneberger et al. 2009. Their approach, called SHOREmap, described a cloning-by-sequencing strategy for identification of a causal mutated gene within a few working days in Arabidopsis. This software package enabled mapping, candidate gene identification and annotation in a single step (Schneeberger et al., 2009). The principle of this approach goes back to 1991, when the bulked segregant analysis was introduced as a rapid method for identification of markers linked to monogenic recessive traits (Michelmore et al., 1991). A segregating F2 population needed to be divided into two phenotypic groups (bulks) differing for the trait of interest (Figure 2a/b). In case of a recessive Mendelian trait, F2 plants that show a mutant phenotype must be homozygous for the mutant genotype around the gene of interest (Figure 2b). Thus, linked markers differentiate between both phenotypic pools whereas all unlinked markers would show heterozygosity in both bulks (Michelmore et al., 1991). In the SHOREmap approach, they resequenced a pool of 500 mutant plants selected from a segregating F2

population. Instead of neighboring SNPs of individual genotypes, SNPs obtained from sequencing bulks of genotypes were used to score as co dominant marker by deterring the occurrence of reads carrying the mutant or wild-type allele. For this, the short sequence reads needed to be mapped against the wild-type reference sequence for SNP calling. SNP frequency can be determined by counting the number of reads that differ at a certain base pair position compared to the reference sequence in relation to the total of all mapped reads. Mutant genotype specific SNPs closely linked to the mutation are expected to reach a frequency of 100 % in the mutant bulk compared to 25 % in the wild-type pool, originating from sequence reads of heterozygous plants which cannot be phenotypically discriminated from wild-type at F2 stage. SNPs that are unlinked from the mutation were represented by around 50 percent reads from both, wild-type and mutant (Figure 2b). Loosely linked SNPs differ in between according to their distance to the causal mutation (Abe et al., 2012). The accurate SNP frequency determination requires sufficient sequencing depth. Obtained SNPs from sequencing of these bulks can be used to define a target interval by considering detected SNP frequencies in a genome wide physical context (Figure 2c). Genes within the obtained target region can be screened for functional mutations within coding sequences to predict the causal gene of interest {Schneeberger, 2009 #2185}. The principle of mapping-by-sequencing was successfully adapted to the completely sequenced crop plants rice and maize (Abe et al., 2012; Liu et al., 2012).

First studies showed the principle adaptability of such a sequence-based mapping strategy to plants lacking a high quality reference genome sequence, by taking advantage of conserved synteny (Galvao et al., 2012) to closely related and sequenced organisms or comparing variations in *de novo* assembled sequence fragments to each other (Nordstrom et al., 2013). There are two major limiting factors in regard to barley: (i) the lack of a high quality reference sequence; (ii) the enormous genome size would cause high sequencing costs especially in case of higher number of genotypes that need to be analyzed. Nevertheless, the above described anchored sequence assembly of barley (IBSC, 2012; Mascher et al., 2013b) may be utilized as reference for sequence read mapping and variant calling.

Target enrichment based re-sequencing allows generating high informative sequence of the low copy sequences of a genome of interest. Recently, a barley whole exome capture assay was designed representing ~80% and ~40 % of all high and low confidence genes, respectively, which can be used to enrich for genic parts of the genome before sequencing (Mascher et al., 2013c). As an alternative, high throughput cDNA sequencing (RNA-seq) was

reported for complexity reduction by focusing only on transcribed genes (Wang et al., 2009). Thus, taking advantage of cloning-by-sequencing for gene isolation in barley may be feasible.

a) F2 mapping population
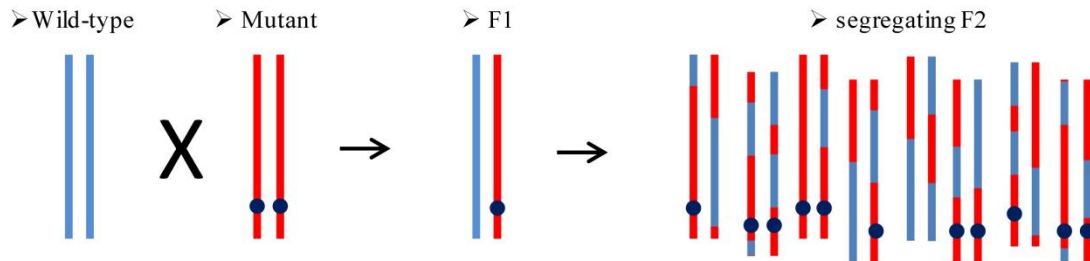
b) Phenotypic bulks, read mapping, SNP detection and SNP frequency determination
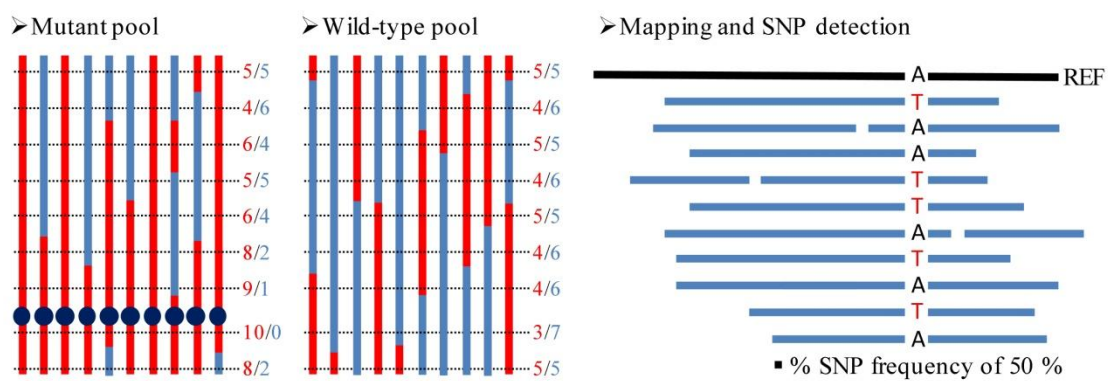
c) Physical distribution of SNP frequencies
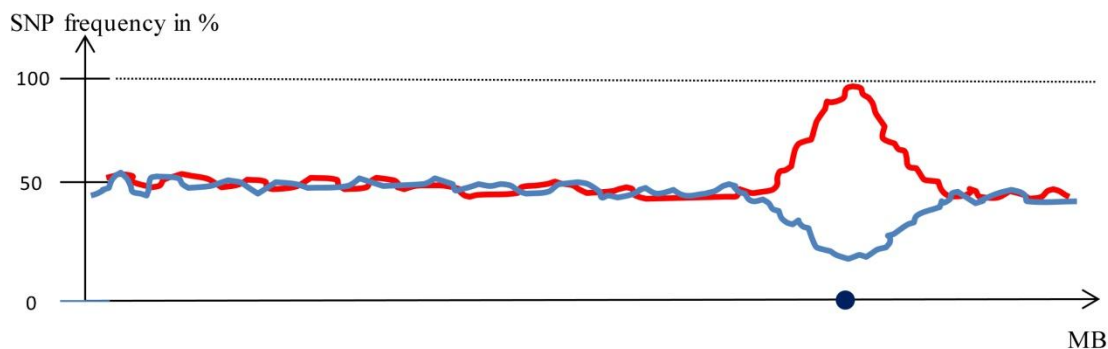
Figure 2: Principle of mapping-by-sequencing
a) The basic approach of mapping-by-sequencing starts with establishing of a segregating F2 population by crossing a wild-type and a mutant plant. b) F2 plants need to be distributed into two contrasting phenotypic pools followed by genotyping (NGS sequencing). Obtained sequence reads can be utilized as genetic markers by determining SNP frequencies in shotgun sequence reads aligned against the reference (WT) sequence. All plants with mutant phenotype will be homozygous for the mutant allele in candidate gene locus. This region can be identified by visualizing SNP frequencies in a physical context (c). In free segregating unlinked regions SNP frequencies pending around 50% for both pools. Within the target region the SNP frequency goes up to 100% mutant allele (all homozygous mutant) in the mutant pool and drops down to 25% mutant allele in the WT pool (homozygous WT and heterozygous WT/mutant). Picture partially modified after Schneeberger and Weigel (2011).

1.3.3. Functional validation of candidate genes

In case, a forward genetics approach led to the identification of a candidate gene, the next experimental step would require functional validation of the gene function. This may be achieved by screening for additional mutant alleles expressing a similar mutant phenotype, providing independent evidence that the identified gene controls the expected phenotypic changes.

Mutant collections are an excellent resource to screen for plants with similar phenotype to the one addressed in the forward genetic screen. The existence of those collections is the result of a long history in barley mutant research and breeding. Mutation breeding history goes back to 1927 when Muller detected that x-ray radiation increases the mutation frequency in fruit fly *Drosophila melanogaster* (Muller, 1927). In 1928, the method was successfully transferred to plants and initiated the establishment of large mutant collections for increasing genetic variability and to identify plants with improved performance in agricultural production or breeding programs. The largest and best known collection of barley mutants was established in Sweden by Nilson-Ehle and Gustafsson by application of different irradiation types; work that was later continued by applying different chemical treatments (Lundquist, 2009). More than 10,000 characterized mutants are stored in Nordic Genetic Resource Center (NordGen, http://www.nordgen.org/) and can be accessed for breeding and research (Lundquist, 2009). In the 1950's, Scholz and Lehman generated at the Zentralinstitut für Genetik, Gatersleben, Germany, [today: Leibniz Institute of Plant Genetics and Crop Plant Research (IPK)] an additional barley mutant collection mainly by x-ray irradiation and later-on by treating seeds with ethyl methansulfonate (EMS). The collection of around 900 accessions is maintained in Gatersleben, Germany (Scholz, 1962). Mutagenesis by irradiation frequently causes large deletions whereas chemical treatment with EMS leads to alkylation of purine bases and subsequently after replication of DNA preferentially to G/C-A/T transitions (Nelson et al., 1994; Serrat et al., 2014). Nevertheless, independent from the type of mutagen, the obtained phenotype and the frequency of induced phenotypic changes are in the center of interest. Thus morphological changes in plant growth or physiology, e.g. flowering time, were used as basis for a classification into main mutant categories (Table 1) (Lundquist, 2009).

However, only specific mutant classes were collected and some alleles might have been lost during propagation of these collections.

Table 1: Main mutant categories (Lundquist, 2009)

#	Category[1]
1	Changes in spike and spikelets
2	Changes in culm length and culm composition
3	Changes in growth type
4	Changes in kernel development formation
5	Physiological mutants
6	Awn changes
7	Changes in leaf blades
8	Changed pigmentation
9	Different chlorophyll development
10	Resistance to powdery mildew

[1] mutant categories (taken from Lundquist, 2009)

In addition, a population of 10,279 mutant plants obtained from EMS mutagenesis of barley cv. Barke seeds provided a resource of high density SNP mutations in the genome (Gottwald et al., 2009). The so called "TILLING" (Targeting Local Lesions IN Genomes) is a reverse genetic approach to screen for the induced single base changes in the genome (Till et al., 2006). PCR amplification with gene-specific primers allows discovering nucleotide changes in the gene of interest. After PCR, a heteroduplex formation at the position of the mismatch is initiated and enables the recognition by a single strand specific nuclease. The resulting digested fragments with different sizes, according to the position of the SNP, can be visualized by gel electrophoreses. A 2D pooling strategy allows to pool the DNA of 8 plants in a single reaction and enables for high throughput screen of the large number of mutated genotypes (Till et al., 2006; Gottwald et al., 2009).

Besides mutagenesis, transformation based approaches can be used to validate gene function. Transgenic complementation by transferring an intact copy of the candidate gene into the mutant background to prove for restoration of wild-type phenotype (Stein, 2005). Another option is to disrupt the gene function in a wild-type genotype. This can be done through knockdown by RNA interference (RNAi). The transcript level of the candidate gene will be down-regulated post-transcriptionally by introducing a specific small interference RNA (siRNA). The siRNA is complementary to the mRNA and leads to the formation of double stranded RNA that will be recognized by the "dicer complex" which cleaves finally the mature mRNA (Matzke et al., 2001). Recently, site directed gene knockout strategies (genome editing) for plants have been reported by introducing double strand breaks in the genomic DNA: e.g. transcription activator-like effector nucleases (TALENs) are proteins with

a specific DNA binding and a *Fok*I endonuclease domain which can introduce double-strand nicks in specific target sequences of the genomic DNA. The DNA repair machinery can introduce various sequence changes at the cleavage site (Joung and Sander, 2013). For instance deletions between 1 to 36 nucleotides could be observed by the first successful studies in barley (Wendt et al., 2013; Gurushidze et al., 2014). Since the DNA binding domain can be designed to target a specific DNA sequence motif, the DSBs can be introduced within the candidate gene to disrupt the functionality.

1.4. The aims of the study

The present PhD project aimed at the isolation and molecular characterization of two induced mutants: *many noded dwarf* (*mnd*) and *laxatum-a* (*lax-a*), both characteristic for fundamental changes in barley plant architecture. The identification of the functional genes is expected to help elucidating the underlying complex regulatory networks of spike and culm development.

The main objectives of this study were:

i) Map-based cloning of two developmental genes in barley.

ii) Exploit the possibilities of mapping-by-sequencing respectively cloning-by-sequencing strategies.

ii) Delimit the target loci physically on the basis of a BAC contig of the barley physical map and identify candidate genes.

iii) Functional test of identified candidate genes by screening for independent mutant alleles by TILLING analysis.

iv) Validate the candidate genes by screening other independent mutant resources.

2. Materials and methods

2.1. Plant material

2.1.1. Plant material utilized for the characterization of *HvMND*

The many noded dwarf mutant accession MHOR474 was obtained by x-ray irradiation of cultivar 'Saale' (Scholz, 1962). Seeds, provided by the IPK Gatersleben genebank were used to produce a segregating F2 population which was derived from a cross of MHOR474 and cv. Barke in advance of the project. Plant material of this F2 population was used for mapping-by-sequencing as well as marker based genetic mapping. For the purpose of *HvMND* candidate gene confirmation, an existing TILLING population (Gottwald et al. 2009) of cv. Barke was analyzed (see 2.9.). Additionally, six Bowman nearly isogenic lines (NIL) (Druka et al., 2011) with introgressed *mnd* alleles and 37 *mnd* mutants from Nordic Genetic Resource Center (NordGen, http://www.nordgen.org) were analyzed (Table 7 and Table 8). The majority of the plant material was cultivated under controlled long day greenhouse conditions (18°C / 16°C day/night). Some of the F3 progenies of the mapping population were also grown under field-like conditions in nursery plots covered by bird-protecting nets during summer season (April to September 2013) for phenotyping.

2.1.2. Plant material utilized for characterization of *HvLAX-A*

The mutant allele *lax-a.8* was induced in 1956 by fast neutron mutagenesis of seeds of cv. Bonus (Franckowiak, 2010). Subsequently the mutant plant was six times backcrossed to cultivar Bowman to generate a nearly isogenic line (NIL) BW457 (Druka et al., 2011). An F2 population was developed by backcrossing of the mutant NIL BW457 with cv. 'Bowman'. Seeds were kindly provided by Dr. Arnis Druka from James Hutton Institute (JHI), Dundee, Scotland. To identify independent mutant alleles for *lax-a*, the same TILLING population as mentioned above (Gottwald et al., 2009) was used. Furthermore, 28 *lax-a* mutant accessions in different genetic background (Table 17) were provided by the Nordic Genetic Resource Center (NordGen, http://www.nordgen.org) and analyzed.

In order to determine natural genetic diversity a worldwide collection of 224 spring barleys, propagated by single seed descent from accessions obtained from the *ex situ* seedbank of IPK Gatersleben (http://www.ipk-gatersleben.de), was analyzed. The collection consisted of 149 improved varieties and 57 landraces plus additional various breeding or research stocks (Haseneyer et al., 2010). Additionally, eighty-three wild barley accessions (*Hordeum spontaneum* L.) from Near East, Europe and Asia were analyzed (Table A1). DNA was kindly provided by Dr. Benjamin Kilian, IPK Gatersleben, Stadt Seeland, Germany (now Bayer CropScience NV Ghent, Belgium). The entire plant material was cultivated under controlled greenhouse conditions as described above (2.1.1.).

2.2. Phenotyping

2.2.1. Phenotypic analysis of a population segregating for *mnd*

Plants of the F2 mapping population were visually scored for number of internodes. Due to the recessive inheritance of the phenotype, plants with more than five internodes at full maturity stage were classified as homozygous mutant (*mnd/mnd*) in contrast to wild-type plants (*Mnd/Mnd* and *Mnd/mnd*). F3 progenies grown under field-like conditions (nursery plots covered by bird-protecting nets) were used to study in addition any other pleiotropic morphological characteristics affecting spike length (five spikes per plant), tiller number and average plant height. The TILLING mutants and the *mnd* mutant accessions from NordGen were characterized under greenhouse conditions solely regarding faster leaf initiation (plastochron) and counting the node/internode number on mature plants.

2.2.2. Phenotypic analysis for *HvLAX-A*

Phenotyping of the *laxatum-a* plant material was exclusively performed under greenhouse conditions. Spikes of plants were visually inspected for width of the lemma awn base and number of anthers after heading stage. The exposure of caryopsis was analyzed at mature spikes. Average rachis internodes length was calculated by dividing overall ear length of mature spikes by number of nodes per spikes. Three spikes per plant were used to calculate the average to estimate a standard deviation and error of the measurement.

2.3. Preparation and quantification of genomic DNA

Plant material for DNA isolation was harvested from greenhouse grown seedlings at three-leaf stage and immediately transferred into liquid nitrogen. The frozen leaf material was homogenized by shaking together with sterilized steel beads (3,175 mm diameter) in 2 ml tubes (Multiply PCR Cups, Sarstedt AG & Co. Nümbrecht, Germany) or plates (96-well racked Collection Microtubes, Qiagen, Hilden, Germany) for one minute at 30 Hz on the 'Retsch MM301' (Retsch, Haan, Germany) instrument. DNA isolation was performed with different protocols described in the following. For obtaining DNA yields higher than 10 µg, DNA was extracted after a modified cetyl-trimethylammonium bromide-based (CTAB) protocol (Stein et al., 2001). In brief, one milliliter of Extraction Buffer (2% (w/v) CTAB, 200 mM Tris/HCl pH 8.0, 20 mM EDTA pH 8.0, 1.4 M NaCl, 1% (w/v) polyvinylpyrrolidone (K30), 1% (v/v) b-mercaptoethanol) was added to the frozen leaf powder in 2 ml reaction tubes (Multiply PCR Cups, Sarstedt AG & Co. Nümbrecht, Germany) and thoroughly shaken. After 30 min incubation at 65 °C, 800 µl -20°C cold Chloroform:Isoamylalcohol (24:1) was added and 15 min incubated at room temperature in a REAX 2 overhead shaker (Heidolph, Schwabach, Germany) followed by centrifugation at 10,000 g (4 °C) for 15 min. The supernatant was transferred to a new 1,5 ml reaction tube and treated with 5 µl RNase A [1,000 U/ml] (Carl Roth GmbH + Co. KG, Karlsruhe, Germany) for 15 min at 37°C. The DNA was precipitated with 570 µl isopropanol (-20 °C) by five times inverting and centrifugation for 15 min (10,000 x g, 4 °C). The liquid phase was removed by carefully inverting the open tube from the pellet which was fixed to the bottom of the tube. 1 ml of Wash Solution I (76% Ethanol, 200 mM Sodium acetate) was added and incubated on ice for 10 min. The Wash Solution I was replaced by Wash Solution II (76% Ethanol, 10 mM Ammonium acetate) followed by a incubation step for 5 min on ice. The wash solution was completely removed and the pellet dried under a fume hood at room temperature for around 15 min. Finally the DNA pellet was eluted in 100 µl TE buffer (10 mM Tris/HCl pH 8.0, 1 mM EDTA) and incubated for 15 min at 37 °C.

Another modified CTAB based DNA isolation protocol (Doyle, 1990) was applied to enable DNA isolation in a 96 well format. In brief, 450 µl of the CTAB Extraction buffer (2% (w/v) CTAB, 200 mM Tris/HCl pH 8.0, 20 mM EDTA pH 8.0, 1.4 M NaCl, 1% (w/v) polyvinylpyrrolidone (K30), 1% (v/v) b-mercaptoethanol) was added to the frozen leaf powder and vigorously shaken followed by a 30 min incubation at 65 °C. 400 µl dichloromethane : isoamylalcohol (24:1) was added and after shaking centrifuged for 10 min

at 5,700 x g (4°C). The supernatant was transferred into a new plate and mixed with 45 µl acetate-mix (10 M NH_4OAc, 3 M NaOAc [pH 5,5]). The DNA was precipitated with 260 µl isopropanol (-20 °C) by inverting the plate five times followed by 30 min centrifugation (5,700 x g, 4 °C). The supernatant was discarded and the pellet washed two times by 500 µl 70 % Ethanol through 10 min centrifugation (5,700 x g, 4 °C). The ethanol was removed carefully and the pellet dried under a fume hood at room temperature. The DNA pellet was eluted in 100 µl TE-buffer (10 mM Tris/HCl pH 8.0, 1 mM EDTA). Two µl RNase A [1000 U/ml] (Carl Roth GmbH + Co. KG, Karlsruhe, Germany) were added to each well and incubated at 37 °C for 15 min.

For high resolution mapping of *HvLAX-A* only small amounts of DNA were required and an automated rapid 96 well plate format DNA isolation on „Biorobot 3000" (Qiagen, Venlo, Nederland) system with MagAttract 96 DNA Plant Core Kit could be used following manufacturer's instructions (Qiagen, Hilden, Germany).

Isolated DNA was analyzed for quality and concentration by gel electrophoresis using 1 % agarose (Invitrogen GmbH, Darmstadt, Germany) in 1xTBE buffer (89 mM Tris, 89 mM Boric acid, 2 mM EDTA pH=8) in comparison to a dilution series (50 –250 ng) of standard λ-DNA (Fermentas GmbH, St. Leon-Rot, Germany). Concentrations were independently measured also by using a NanoDrop spectrophotometer (Thermo Scientific, Wilmington, USA) and/or the Qubit 2.0 Fluorometer (Invitrogen, Carlsbad, CA, USA) according to manufacturer guidelines of the Qubit dsDNA BR Assay Kit which can quantify concentrations between 2 and 1000 ng by measurement with fluorescent dye incorporated in double strand DNA. A Qubit™ working solution was prepared by diluting the Qubit™ dsDNA BR reagent 1:200 in Qubit™ dsDNA BR buffer. Before measurement, 10 µL of each standard (standard 1: 0 ng/µL, standard 2: 100 ng/µl) was mixed with 190 µL of Qubit™ working solution and the DNA samples were diluted 1:200 in the working solution.

2.4. Preparation of RNA

For studying gene expression of *HvLAX-A* in mutant and wild-type plants, 15-18 spike meristems of independent plants were collected and pooled for three developmental stages: (i) glume primordium (Figure 3a) (ii) stamen primordium (Figure 3b) (iii) completely developed spikes at awn primordium (Figure 3c) (Kirby, 1987). Three biological replicates were

sampled for each stage. Spike meristems were checked under stereomicroscope 'Stemi 2000-C (Carl Zeiss MicroImaging GmbH, Jena, Germany) for the exact developmental stage, collected in 1.5 ml tubes and immediately frozen by transferring the tubes (Multiply PCR Cups, Sarstedt AG & Co. Nümbrecht, Germany) to liquid nitrogen.
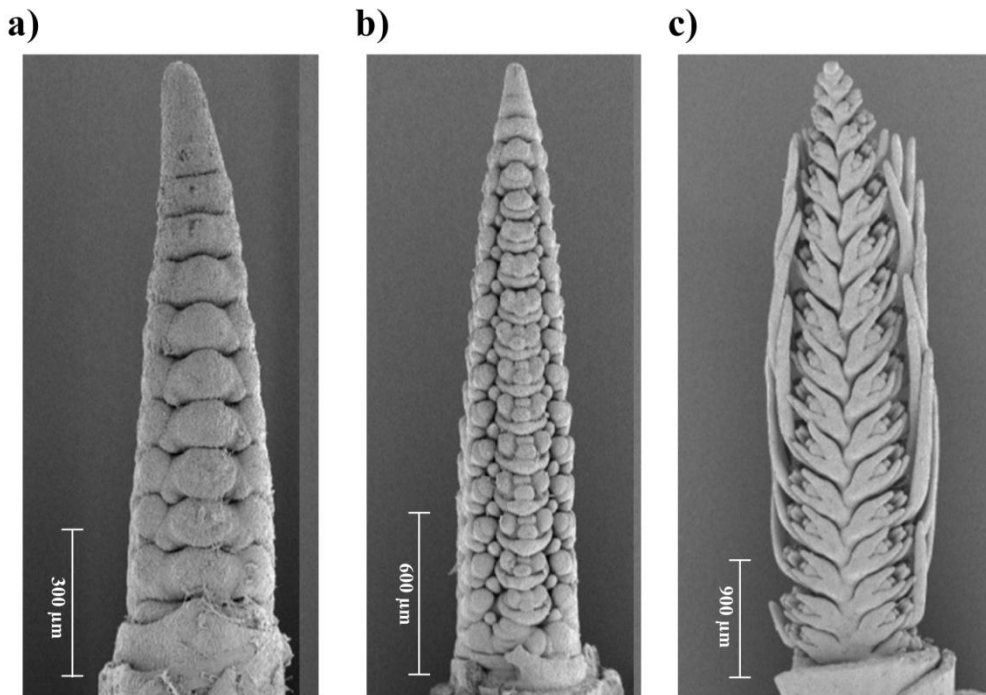
Figure 3: Material harvested for RNA isolation.
Immature spike meristems at three stages were collected from mutant and wild type plants: a) glume primordium b) stamen primordium c) completely developed spike.

Total RNA isolation was performed with TRIzol reagent (Ambion, Austin, TX, USA). In brief, the spike meristems were grinded with a pistel within the liquid nitrogen cooled tubes followed by adding 1 ml of TRIzol reagent and 0.2 ml chloroform. The samples were shaken vigorously and incubated for 3 minutes at room temperature before the phase separation by centrifugation (15 min, 12,000 x g, 4°C) was performed. For precipitation of the RNA, the aqueous phase of each sample was transferred to a new reaction tube, mixed with 0.5 ml of 100% isopropanol, incubated at room temperature for 10 min and finally centrifuged for 10 min (12,000 x g, 4°C). The obtained pellets were washed in 1 ml of 75% ethanol by centrifugation (5 min, 7,500 x g, 4°C). The air dried pellets were eluted in 30 μl Diethylpyrocarbonate (DEPC, AppliChem GmbH, Darmstadt, Germany) treated water (0.1%). Isolation was performed under RNase free conditions including RNase-free filter tips, reaction tubes (Biozym, Hessisch Oldendorf, Germany) and solutions prepared with DEPC

treated water in order to minimize risks for degradation of the RNA. The integrity of isolated RNA was checked by testing the ratio of the 28s and 18s ribosomal RNA bands by a denaturing agarose gel (1,5% agarose, 1xMOPS buffer, 1.5 % Formaldehyde) electrophoresis. Before loading samples on the gel, 2µl RNA was mixed with 25 µl formamide, 5µl 10x MOPS buffer (200 mM MOPS; 50 mM Sodium acetate; 10 mM EDTA), 90 µl 37% formaldehyde and after denaturation for 5 min at 65 °C mixed with 2µl of a 10% (w/v) ethidiumbromid solution. Electrophoresis was performed in 1 x MOPS buffer with 5V/cm power supply for 1 hour. RNA Integrity Number (RIN, (Schroeder et al., 2006)) was determined with RNA 6000 Nano Assay kit on Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA, USA). Only RNA samples with RIN factor higher or equal than eight were considered for RNA-sequencing (see 2.8.4.). Quantification of RNA was achieved by using RNA BR (Broad-Range) Assay Kit on Qubit 2.0 Fluorometer (Invitrogen, Carlsbad, CA,USA) according to manufacturer's instructions described for DNA isolation (part 2.3).

For performing quantitative reverse transcribed PCR (qRT-PCR), RNA samples were purified from potentially contaminating DNA by treating with DNase. One microgram of total RNA was digested in a 10 µl mix of RNase free water, 1 µl DNase I (5U/µl, Fermentas GmbH, St. Leon-Rot, Germany), 1 µl 10× reaction buffer (Fermentas GmbH, St. Leon-Rot, Germany) by incubation at 37°C for 30 min. The samples were provided with one µl 50 mM EDTA and incubated for 10 min at 65 °C to inactivate the DNase I.

The cDNA synthesis was performed by incubating DNase purified RNA (1 µg) with the RNA Invitrogen SuperScript III First-Strand Synthesis Supermix for qRT-PCR (Invitrogen, Carlsbad, CA, USA) according to manufacturer guidelines.

2.5. Primer design and Polymerase Chain Reaction

Uniquely binding oligonucleotides (primers) are required to amplify a target sequence of DNA by polymerase chain reaction (PCR) (Saiki et al., 1988). The primer design was performed with batch primer 3 software tool (Koressaar and Remm, 2007; Untergasser et al., 2012) using default parameters except for primer GC content which was set to values between 50 % and 60 % and the expected product size was adjusted according to the experimental

requirements (e.g. products for Sanger sequencing were designed to be of less than 900 bp to guarantee complete sequence coverage of both DNA strands).

The PCR reactions were performed on GeneAmp PCR Systems 9700 (Applied Biosystems, Carlsbad, USA). The reaction mix with a total volume of 20 μl consisted of 2μl 10x PCR buffer [Tris-CL, KCL, (NH4)2SO4, 15 mM MgCl2] (Qiagen, Hilden, Germany), 2μl dNTP Mix [2 mM of each dNTP] (Fermentas, Fermentas, St. Leon-Rot, Germany), 1 μl of each Primer [10 mM], 0,1 μl Hot star Taq polymerase [5 units/μl] (Qiagen, Hilden, Germany) and 1 μl DNA template [20 ng/μl]. In some reactions 4 μl Q-solution was added which is supposed to improve the amplification of 'difficult' templates (e.g. high probability of secondary structures or GC-rich templates) according to manufacturer's information's (Qiagen, Hilden, Germany). The reaction mix was filled up with double-distilled water $_{dd}H_2O$ to 20 μl total volume. A standard touchdown (TD-) PCR profile was used for all PCR analyses containing two cycling steps: initial denaturation for 15 min at 95 °C, followed by ten cycles of denaturation at 95 °C / 30 sec; annealing at 60°C / 30 sec (decreasing by 0.5°C per cycle) followed by extension at 72°C / 60 s); then 35 cycles denaturation at 95 °C / 30 sec, annealing at 55°C / 30sec, and extension at 72°C / 60 sec followed by a final extension step at 72°C / 7 min. Extension time was modified according to the length of the PCR target (1 min / 1 kbp). PCR products were resolved by agarose gel electrophoresis (5V/cm) using 1.5 % agarose gel (Invitrogen GmbH, Darmstadt, Germany) in 1xTBE buffer.

2.6. Marker development, genotyping and genetic map construction

Markers were developed by exploiting previously known SNP resources available from the Barley Oligo Pool Assay (BOPA) SNP map (Close et al., 2009) or predicted SNPs obtained by sequence comparison after Next Generation Sequencing (NGS) of mutant and wild-type samples. Thus, specific primers were designed and used to amplify products surrounding SNPs located on WGS contigs. Predicted SNPs were confirmed by Sanger sequencing (see part 2.8.1) and checked for applicability to the Cleaved Amplified Polymorphic Sequences (CAPS) marker system by SNP2CAPS software tool (Konieczny and Ausubel, 1993; Thiel et al., 2004). Restriction digests were performed according to manufacturer guidelines. In general, for enzymes from New England Biolabs (Ipswich, England) 10 μl PCR product was digested by 2 units enzyme, 1,5 μl of the respective NEB buffer and 0,15 μl BSA (*bovine serum albumin,* if required according manufacturer) adjusted with ddH20 to a total volume of

15 µl. For restriction enzymes from Fermentas GmbH (St. Leon-Rot, Germany) 10 units enzyme were used in a 30 µl total volume including 2 µl of the corresponding enzyme buffer to digest 10 µl PCR product. The digest was performed according manufacturer suggested temperature of each enzyme for 2,5 h on GeneAmp PCR System 9700 (Applied Biosystems, Carlsbad, USA).

Some SNP markers were converted to an 8-plex SNaPshot (Applied Biosystems, Foster City, CA, USA) marker assay to screen an extended mapping population for high resolution mapping. The SNP detection method is based on an allele specific oligo extension. For this, a primer designed to hybridize directly next to the SNP position was extended specifically into the allele-specific bp-position by introducing a labeled didesoxy-nucleotide. The extension oligos are differentiating in size between 30 and 74 bp (Table A2) which allows the multiplexing of multiple markers in a single reaction, since they can easily be differentiated by fragment size after capillary electrophoresis. First, amplification of 92 to 227 bp fragments was performed by a multiplex PCR (two reactions, Table A2) at standard PCR conditions (see part 2.5.). PCR reaction cleanup (removal of unincorporated dNTPs, primers) was achieved by incubation with shrimp alkaline phosphatase (5u) and exonuclease 1 (2u) (Affymetrix, Santa Clara, USA). Reaction conditions as well as subsequent steps of sample preparation were performed according to the supplier`s protocol of the ABI PRISM SNaPshot™ Multipex Kit (Applied Biosystems, Foster City, CA, USA). Capillary electrophoresis was performed on Applied Biosystems 3730xl DNA Analyzer equipped with 50 cm capillaries, POP-7™ Polymer matrix and Data Collection Software 3.0 (Applied Biosystems, Foster City, CA, USA). The system was calibrated with Matrix Standard Set DS-02 (Set E5) according to the user bulletin (Applied Biosystems, Foster City, CA, USA). Peak histogram analysis for genotyping was performed with GenMapper4.0 (Applied Biosystems, Foster City, CA, USA) software.

A genetic linkage map was constructed by JoinMap version 4.0 (Kyazma B.V., Wageningen, The Netherlands) by applying the regression mapping algorithm (LOD score = 4.0 for grouping) and Kosambi mapping function under default settings.

2.7. Physical mapping

The available physical map of barley (IBSC, 2012) allows the physical characterization of genomic target regions without the need for performing chromosome walking. Based on the

sequence resources integrated to the physical/ genetic genome reference map of barley a large number of markers and genes had been assigned a physical map position already by in silico sequence comparison (IBSC, 2012). . In case the marker sequence was not assigned to the physical map, the in silico anchoring of genes or markers was repeated by BLASTN (minimum hit length of 700 bp, >99% identity) (Altschul et al., 1990) of the corresponding WGS contigs against sequenced BACs or BACend sequences of the physical map (IBSC, 2012). In case no sequence homology based assignment was achieved, a two step PCR-based screening of multidimensional BAC pools (customized arrangement by Amplicon Express, Pullman, WA, USA) of the library HVVMRXALLeA (Schulte et al., 2011) was performed. The library of 147,456 BAC clones, representing a 3,7 fold haploid genome coverage of barley, was separated in three hundred and eighty-five 384-well plates. Seven of these plates were pooled resulting in 55 super pools. For each super pool 23 matrix pools were generated and arranged in a matrix to form plate, row, and column pools which allowed the deconvolution of a single BAC address (Ariyadasa and Stein, 2012; Poursarebani et al., 2013). To assign the gene *HvLAX-A* to a BAC contig of the physical map (IBSC, 2012), PCR amplification of two gene fragments (HvLAX_F/R1 and HvLAX_F/R3) was performed. A BAC harboring the second gene, *HvMND*, was identified by screening two-dimensional DNA pools comprising only the BAC clones representing the minimum-tiling path of the genome-wide physical map of barley (Ariyadasa et al., 2014) with a single gene fragment (HvMND_F/R4, see Table A3).

2.8. Sequencing and data processing

2.8.1. Sanger sequencing

PCR products were purified with NucleoFast 96 ultra filtration kit (MACHEREY-NAGEL GmbH & Co. KG, Düren, Germany). PCR products were brought to a total volume of 100 µl by adding $_{dd}H_2O$ and then transferred to filter plates equipped with a membrane which is permeable for products smaller than 150 bp. Filtration was achieved by applying vacuum (1000 mbar) after placing the plates to the QIAvac 96 (Qiagen, Venlo, Netherlands) device. Samples were washed twice by adding 100 µl $_{dd}H_2O$ followed by filtration under vacuum. The products were finally eluted in 30 µl $_{dd}H_2O$ by shaking plates for 15 minutes on Titramax 100 plate shaker (Heidolph Instruments GmbH & Co. Kg, Schwabach, Deutschland) at room

temperature. Two µl of the purified product were checked on a 1,5 % Agarose gel together with λDNA (Fermentas, Germany) with the known concentration (50 to 250 ng/µl) to adjust the amount of PCR product to the suggested ten ng per 100 bp PCR product. Sequencing was performed using the ABI PRISM® dGTP BigDye™ Terminator v3.0 Ready Reaction Cycle Sequencing Kit (Applied Biosystems, Carlsbad, USA). For this, a reaction mix consisting of 4 µl BigDye Premix, 1 µl Primer [5µM] and the PCR product was finally adjusted to a total reaction volume of 10 µl with $_{dd}H_2O$. The cycle sequencing was performed on GeneAmp PCR System 9700 (Applied Biosystems, Carlsbad, USA) according to the following protocol: Initial denaturation by 96°C for 1 min, 25 cycles: denaturation (96°C, 10 sec), annealing (50°C, 5 sec), extension (60°C, 4 min) and a finale cooling to 4°C. The sequencing was performed on „3730xl DNA Analyzer" (Applied Biosystems, Carlsbad, USA). PHRED 20 quality (Ewing and Green, 1998; Ewing et al., 1998) trimmed sequence analysis was done with "Sequencher 4" software (Genecodes Corporation, USA).

2.8.2. Whole genome shotgun sequencing

The Bowman NIL BW457, which was used to establish the bi-parental mapping population for the gene *lax-a* by crossing to Bowman, was whole genome shotgun sequenced to about eight-fold haploid genome coverage by. Illumina-Paired-End library preparation was performed according to previously published protocols (IBSC, 2012). In brief, genomic DNA was fragmented mechanically by nebulization to around 500 bp fragments which was controlled by Agilent High Sensitivity DNA Assay on Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA, USA) before sample preparation according Illumina TruSeq DNA Sample Preparation Kit (Illumina Inc., San Diego, CA, USA). In brief, first step was the end repair which was done by mixing 50 µl fragmented DNA with 10 µl End Repair Control and 40 µl End Repair Mix followed by incubation for 30 min at 30 °C in a thermocycler. The control reagent includes a spiked small amount of dsDNA fragments which reports the success of the step and can be used for troubleshooting after sequencing. This solution is optional and can be replaced by resuspension buffer. The total reaction (100 µl) was transferred to a 1,5 µl tube and mixed thoroughly with 160 µl AMPure XP Beads (Beckman Coulter GmbH, Krefeld, Germany) and placed on a DynaMag-2 device (Life Technologies, Invitrogen, Carlsbad, CA, USA). After 15 min separation at room temperature the aqueous phase was discarded and replaced by 200 µl of 80 % ethanol. After 30 sec incbubation the

ethanol was renewed and again after 15 min incubation discarded completely. The beads were dried for 15 min at room temperature and finally eluted in 16,5 µl RSB by thoroughly mixing and 2 min incubation. The tube was placed again to the magnet to remove the beads from the solution. 15 µl of the sample was processed for A-Tailing. A single nucleotide was added to the 3 prime end of the blunt end fragments to avoid self ligation. The reaction comprised of 15 µl DNA, 2,5 µl A-Tailing Control and 12,5 µl A-Tailing Mix was incubated for 30 min at 37 °C. Complementary T overhanging adapter were adjusted in the next step for sequence adapter ligation. The 30 µl A-tailed DNA fragments were mixed with 2,5 µl Ligation Control, 2,5 µl Adapter Index (specific barcode for each sequencing sample) and 2,5 µl Ligation Mix followed by incubation at 30 °C for 10 min. The ligation was stopped by adding 5 µl of the Stop Ligation Buffer. The reaction was cleaned by mixing with 42,5 µl AMPure XP beads and incubation for 15 min at room temperature. The reaction tube was then transferred again to the magnetic device to separate the beads from the liquid. The supernatant was discarded and the beads washed two times with 200 µl of 80 % Ethanol (like described above) and finally completely removed. The beads were dried for 15 min before eluation in 52 µl Resuspension Buffer by thoroughly mixing and incubation for 1 min. The suspension was transferred into a new reaction tube and mixed again with 50 µl AMPure beads. The washing procedure with 80 % ethanol was repeated and the samples finally eluted in 22 µl Resuspension Buffer. The beads were removed by magnetic separation and the DNA sample was processed for gel-purification to select fragments with an appropriately size for sequencing. For this, a 2 % agarose gel was prepared with 1x TAE (40mM Tris, 20mM acetic acid, 1mM EDTA) and 10 % SYBR-Gold solution (Life Technologies, Darmstadt, Germany). The library (20 µl) was mixed with 5 µl 6x loading dye (Thermo Fisher Scientific, Fermentas, *St. Leon-Rot, Germany)* before loading to the gel. As size standard 250 ng of the GeneRuler 50 bp DNA ladder (Thermo Fisher Scientific, Fermentas, St. Leon-Rot, Germany) was mixed with 4 µl of 6 x loading dye (Thermo Fisher Scientific, Fermentas, St. Leon-Rot, Germany) and adjusted with Resuspension buffer to a 20 µl total volume. The electrophoresis was performed in 1x TAE for 2h with 4 V/cm protected against light. The SYBR-Gold dye was visualized under blue light from a "Dark reader" transilluminator (Clare Chemical Research, Dolores, CO, USA) and the region of 420-520 bp was cut out and transferred in a clean 2 ml reaction tube. The DNA was eluted from the gel using the MinElute kit (Qiagen, Hilden, Germany) as described by the manufacturer. The library preparation ends with a PCR based DNA fragment enrichment step. For this, 20 µl of the gel purified DNA was mixed with 5 µl PCR Primer Cocktail and 25 µl PCR Master Mix. The cycling was performed by initial

denaturation for 30 s, 10 cycles of denaturation 98 °C (10 sec), annealing 60 °C (30 sec), extension 72 °C (30 sec) followed by a final extension of 72 °C for 5 min and a cooling to 10 °C. PCR product cleanup was conducted like described above by adding 50 µl AmPure XP beads and with a final elution in 32.5 µl Resuspension Buffer. Finally 30 µl of library was separated from the beads by the magnet. The size profile of the whole genome shotgun library was determined with Agilent High Sensitivity DNA Kit on the Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA, USA).

The Quantification of the sequence library was done by Real-Time PCR on a 7900 HT Fast Real-Time PCR system (Applied Biosystems, Carlsbad, USA) according described procedure (Mascher et al., 2013c). In brief, the reaction mix with a total volume of 15 µl contained 1 µl diluted library, 7.5 µl SYBR Green PCR Master Mix (Qiagen, Hilden, Germany) and 100 nM primer specific for the Illumina TruSeq libraries. The PCR cycling was performed by initial denaturation at 95°C for 14 minutes and 15 seconds, followed by 30 cycles of 94°C denaturation (15 sec), 54°C annealing (30 sec), extension 73°C (45 sec). To avoid errors each measurement was done with three technical replicates of the diluted sequencing library, a non-template control and a dilution series of a former TruSeq library covering the range between 10 pM and 1 fM as reference (standard curve). The concentration was determined based on the standard curve using the ABI SDS 2.2.2 Realtime-PCR Software (Applied Biosystems, Carlsbad, USA). The two times 100 cycles paired end sequencing-by-synthesis was performed in a single compartment (lane) on the Illumina HiSeq2000 instrument (Illumina Inc., San Diego, CA, USA).

Sequence reads were mapped against the 'Bowman' WGS assembly (IBSC, 2012) with BWA version 0.5.9 (Li and Durbin, 2009). SNP calling was performed with SAMtools version 0.1.17 (Li, 2011) using default parameters.

2.8.3. Exome Sequencing

Exome capture re-sequencing was performed on single individuals but also multi sample pools of genomic DNA. For *HvMND*, DNA of plants with the same phenotype were assigned in two DNA pools (wild-type and mutant). For *HvLAX-A*, however, previously existing genotyping information was considered to select groups of recombinant plants sharing the same type of recombination/phenotype combination. DNA of recombinant plants that shared the same marker haplotype and phenotype were pooled together and further processed as

DNA pools. The pooling was done with equimolar amounts of DNA for each genotype, which required a precise quantification of the DNA concentration like described above.

The protocol started with size fragmentation of genomic DNA. For this, 1 µg DNA was mechanically sheared to 200 – 300 bp fragments by ultra-sonication in Covaris microTUBES on CovarisTM S220 (Covaris inc., Woburn, MA, USA) instrument with the following settings: 175W Peak Incident Power, Duty Factor 10%, 200 cycles per burst and 100 seconds 'Treatment Time'. Size profile was controlled with Agilent High Sensitivity DNA Assay on Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA, USA). Sequencing library preparation was done with Illumina TruSeq DNA Sample Preparation Kit (Illumina, San Diego, CA, USA) based on the manufacturer's instructions. Details described for WGS sequencing library preparation (Part 2.8.2). A range of 320 to 430 bp fragments were cut from the gel and eluted with MinElute kit (Qiagen, Hilden, Germany) as described by the manufacturer for further processing of the sequence capture procedure. The sequence enrichment process was performed according the published protocol (Himmelbach, 2014).

First step was the pre-capture LM-PCR (ligation-mediated PCR). For this, 20 µl of the sequencing library was mixed with 50 µl 2x Phusion High-Fidelity PCR Master Mix (New England Biolabs, Ipswich, United Kingdom), 26 µl $_{dd}H_2O$ and 2 µl TS-PCR Oligo 1 (100 µM) and TS-PCR Oligo 2 (100 µM) respectively. PCR cycling was performed by initial denaturation of 98 °C for 30 sec, 8 cycles of denaturation 98 °C (10 sec), annealing 60 °C (30 sec), elongation 72 °C (30 sec) followed by a final extension at 72 °C (5 min) and adjusting a final temperature of 8 °C. The amplified samples were mixed with 500 µl Qiagen buffer PBI (Qiagen, Hilden, Germany) and transferred to a QIAquick column (Qiagen, Hilden, Germany) placed in a collection tube and centrifuge for 1 min at 16,000 x g. *After* discarding the flow-through 740 µl PE wash buffer (Qiagen, Hilden, Germany) was added and centrifuged at 16,000 x g (1 min). The MinElute column (Qiagen, Hilden, Germany) was placed in a new 1.5 ml tube to elute the DNA by adding 50 µl $_{dd}H_2O$ and 1 min incubation. The pre-capture LM-PCR amplified library was collected after 1 min incubation at room temperature by centrifugation for 1 min (16,000 x g). Quality check was performed by Qbit measurement and Agilent 2100 Bioanalyzer with the Agilent DNA 7500 Kit (Agilent Technologies, Santa Clara, CA, USA). The hybridization of sample to exome library was only continued if the amount of the library was higher than 1 µg and the obtained fragment size in a range between 250 and 500 bp.

One microgram of the pre-capture Ligation mediated PCR (LM-PCR) amplified library was mixed with 10 µl of the Sequence Capture Developer Reagent (Material # 06684335001, Roche, Indianapolis, IN, USA). One microlitre of the TruSeq HE Universal Oligo 1 (1 mM) was added to block the universal sequence of the adapters and 1 µl (1 mM) of the TruSeq DNA INV-HE Index Oligo to block the index sequence of the TruSeq library adapters were added to the reaction mix.

The mixture was dried down in a SpeedVac at 60 °C with closed lid punctured by a needle. The lid was than replaced by an intact one and 7.5 µl 2x Sequence Capture (SC) Hybridization Buffer and 3 µl Hybridization Component A was added to the dried material and denatured in a heating block (95 °C, 10 min). The sample was mixed in a new reaction tube with 4.5 µl Exome Library (liquid array) by gently pipetting. The hybridization was done for 64-72 h in a thermocycler at 47 °C (lid heated to 57 °C).

Affinity purification was performed to enrich for hybridized DNA fragments by a reversible streptavidin-biotin binding of the biotin labeled capture probes to streptavidin coated magnetic beads (Life Technologies, Invitrogen, Carlsbad, CA, USA). For preparation of the Streptavidin Dynabeads, 50 µl was aliquoted in a 1.5 ml reaction tube and placed for 2 min in a DynaMag-2 magnet (Life Technologies, Invitrogen, Carlsbad, CA, USA). The clear liquid was discarded and replaced by 100 µl of the 1 x Bead Wash Buffer. The tube was removed from the magnet for a 10 sec vortexing step. Than the tube was transferred again to the magnetic tool for 2 min and the liquid phase was replaced again by 100 µl 1 x Bead Wash Buffer. The sample was then vortexed (10 sec) and briefly centrifuged (10 sec at 16,000 x g). The beads were again separated by the magnet (2 min) and the liquid phase replaced by 50 µl of the 1 x Bead Wash Buffer. The sample was transferred after vortexing (10 sec) and a short spin down (5 sec 16,000 x g) into a PCR plate for further processing. The 96 well plate was moved to DynaMag-96 Side Skirted Magnetic Particle Concentrator (MPC96, Life Technologies, Invitrogen, Carlsbad, CA, USA) for 2 min to discard the liquid phase. The hybridization sample (15 µl) was then added and mixed with the beads by gently pipetting up and down. The biotinylated captured sample was bound to the (streptavidin- coated) Dynabeads by incubation at 47 °C for 45 min in a thermocycler with heated lid (57 °C). In 15 min intervals the plates were vortexed for 3 sec followed by a short centrifugation to pull down the liquid. After hybridization of the DNA to the Dynabeads 100 µl 1X Wash Buffer I (pre-heated to 47 °C) was added and mixed by vortexing for 10 sec. The suspension was transferred to 1,5 ml tubes and placed in the DynaMag-2 device to remove the supernatant after 1 min. After this, 200 µl pre-heated (47 °C) 1x Stringent Wash Buffer was added to the

sample and incubated for 5 min at 47 °C. The washing with Stringent Wash buffer was repeated once before proceeding with Wash Buffer I (200 µl), Wash Buffer II and Wash Buffer III following the same procedure with phase separation by magnetic device like described for Stringent Wash buffer I. Finally the Dynabead-bound captured library was eluted in 50 µl $_{dd}H_2O$.

After affinity purification a second LM-PCR called Post-capture library amplification was done. For this, the Dynabead-bound library (50 µl) was mixed with 100 µl 2x Phusion High-Fidelity PCR Master Mix, 42 µl $_{dd}H_2O$, 4 µl TS-PCR Oligo 1 (100 µM) and 4 µl TS-PCR Oligo 2 (100 µM). The total volume of 200 µl was separated in two 0.2 ml PCR tubes. PCR was performed by initial denaturation of 98 °C for 30 sec, 18 cycles of denaturation 98 °C (10 sec), annealing 60 °C (30 sec), elongation 72 °C (30 sec) followed by a final extension at 72 °C (5 min) and adjusting a final temperature of 8 °C. The two reactions were combined again after PCR for the following cleanup procedure. The sample was mixed with 1 ml Qiagen buffer PBI and transferred to QIAquick column placed in a collection tube to filter by centrifugation (16,000 x g) for 1 min (Qiagen, Hilden, Germany). After this, 740 µl PE wash buffer (Qiagen, Hilden, Germany) was added and centrifuged at 16,000 x g (1 min). The captured DNA was eluated by adding 50 µl pre-heated (50 °C) EB (Qiagen, Hilden, Germany) from the column matrix by 1 min centrifugation at 16,000 x g.

The size profile of the exome capture library was determined with Agilent 2100 Bioanalyzer Agilent DNA 7500 Kit (Agilent Technologies, Santa Clara, CA, USA). The Quantification of the exome capture library was done by Real-Time PCR as described previously (Mascher et al., 2013c).

Sequencing was performed on the HiSeq2000 according to protocols provided by the manufacturer (Illumina Inc., San Diego, CA, USA). Two times 100 bp paired end sequencing was performed on a single lane for HvLAX-A and HvMND cloning projects, respectively.

Sequence reads of the recombinant genotypes for *HvLAX-A* were mapped against the whole genome shotgun assembly of barley cultivar 'Bowman' with BWA version 0.6.2 (Li and Durbin, 2009). The sequence reads obtained from phenotypic pools of *HvMND* were mapped against the WGS assembly of 'Barke' with BWA version 0.5.9 (Li and Durbin, 2009) The BWA parameter "-q 15" was used for quality trimming. Multisample SNP calling was performed with SAMtools 0.1.19 (Li, 2011) using the command "mpileup –q 10 –C50". SNP calls were filtered for coverage and genotype score with a custom script adapted from (Mascher et al., 2013a). A sufficient read coverage is required to avoid calling of false

positive variants induced by sequencing errors or putatively wrong interpretation of detected SNP frequencies. The read depth in capture targets was calculated with BEDtools (Quinlan and Hall, 2010). Genotype calls and coverage values were loaded into the R statistical environment (http://www.r-project.org) and queried according to the biological questions.

2.8.4. BAC sequencing, annotation and deletion detection

Since the barley genome sequencing is still ongoing, BACs clones of interest obtained by physical anchoring (see part 2.7) still needed to be sequenced individually. A minimum number of overlapping BAC clones representing the complete extension of the selected BAC contigs (IBSC, 2012; Ariyadasa et al., 2014) were selected for sequencing. BAC clones from BAC libraries of cv. 'Morex' (Schulte et al., 2011) were used for DNA isolation, indexing and sequencing library preparation according to published protocols (IBSC, 2012).

The DNA isolation was performed in a 96 well plate format. For this, BAC clones were cultivated in 1.2 ml liquid growth medium 2xYT-Medium (Chloramphenicol 12.5 µg/ml) in deep well plates (Riplate 96, Ritter GmbH, Schwabmünchen, Germany) overnight at 37 °C . The Bacteria cells were pelleted by centrifugation (3500 x g) for 20 min and the supernatant carefully discarded. The cells were than frozen for 30 min at -20 °C and subsequently eluted with 200 µl resuspension solution (50 mM Tris Cl, 10 mM EDTA pH 8.0) containing 2 µl RNAseA [50 mg/ml] and incubated for 15 minutes on Titramax 100 plate shaker (Heidolph Instruments GmbH & Co. Kg, Schwabach, Deutschland). After this, 200 µl lysis buffer (0.2 M NaOH, 1% SDS) was added and mixed by slow inverting of the plate followed by incubation in ice water for 5 min. The protein precipitation was done by adding 200 µl solution (3M potassium acetate 11.5 % (v/v) acetic acid in 100 ml), five times inverting, 10 min incubation in a ice bath, and 30 min centrifugation (3700 x g, 20 °C). Two times 260 µl of the supernatant was transferred in a new deep well plate for DNA precipitation with 350 µl Isopropanol by 3 times inverting and centrifugation at 30 min (3600 x g, 20 °C). The supernatant was removed and the pellet washed with 400 µl 70 % EtOH through 15 min centrifugation (3600 x g, 20 °C). After removing the ethanol the pellet was dried at 37 °C (45 min) and eluted in 100 µl TE

The sequencing library production was done according to a published protocol for DNA Illumina multiplex libraries (Meyer and Kircher, 2010). They have described alternative strategy for library preparation which allows avoiding expensive components and commercial

kits. A second feature is given trough more flexible index strategy. The index will be introduced by an additional PCR step (index PCR) after sequencing adapter ligation which allows to use up to 672 custom designed indices - a prerequisite for deep pooling to sequence large number of BAC clones in a cost efficient way on the Illumina HiSeq2000 instrument (Illumina, San Diego, CA, USA).

First step was the blunt end repair by adding a reaction mix of 7.12 µl $_{dd}H_2O$, 7 µl 10x Tango Buffer (Thermo Scientific , Wilmington, USA), 0.28 µl dNTPs (25 nM of each dNTP Fermentas, St. Leon-Rot, Germany), 0.7 µl ATP [100 mM] (Thermo Fisher Scientific, Wilmington, USA), 3.5 µl T4 polynucleotide kinase [10 U/µl] and T4 DNA polymerase [5 U/µl] (Thermo Scientific , Wilmington, USA) to the 50 µl fragmented DNA. Samples were mixed gently with a pipette and incubated for 15 min at 25 °C, 5 min at 12 °C and cooled down to 4 °C final temperature. Blunt end DNA fragments were purified by adding 126 µl MagNA beads. The MagNA beads were used instead of the Illumina recommended AMPure XP beads (Beckman Coulter GmbH, Krefeld, Germany) to reduce costs. They were used according previously described protocol (Rohland and Reich, 2012) with the following composition: 2 % carboxyl-modified Sera-Mag Magnetic Speed-beads (v/v, Thermo Scientific, Wilmington, MA, USA), 18 % PEG-8000 (w/v), 1 M NaCl, 10 mM Tris-HCl (pH 8.0), 1 mM EDTA (pH 8.0), 0.05 % Tween 20 (v/v, Bio-Rad Laboratories, Hercules, CA, USA). After homogeneity mixing by pipetting and incubation at room temperature (5 min) samples were transferred to DynaMag-96 Side Skirted Magnetic Particle Concentrator (MPC96, Life Technologies, Invitrogen, Carlsbad, CA, USA) to separate the beads with the hybridized DNA from the solution. The supernatant was removed and two times washed with 70 % ethanol by incubation for 1 minute. After removal of the ethanol solution, beads were dried for ~30 minutes under a fume hood at room temperature. The plate was removed from the magnetic carrier for resuspending in 22 µl EBT [10 mM Tris-Cl (pH 8.0), 0.05 % (v/v) Tween 20]. The eluted samples were moved back to the magnetic tool to remove the beads from the purified sample.

The sequencing adapters were ligated to the blunt end DNA fragment. For this, the sequencing adapter mix (combined hybridization mix for adapter P5 [200 µM] and adapter P7 [200 µM]) was prepared according protocol (Meyer and Kircher, 2010) and ligated by adding 10 µl $_{dd}H_2O$, 4 µl T4 DNA ligase buffer, 4 µl PEG 4000 (50%), 1 µl of the Adapter mix P57 (100 µM of each) and 1 µl T4 DNA ligase (5 U/µl) to the 20 µl purified DNA samples by incubation for 30 min at 22 °C. After ligation 72 µl MagNA beads were added to

perform a purification step like described above. 20 µl of the purified sample was mixed with 14.1 µl ddH20, 4 µl 10x ThermoPol reaction buffer (New England Biolabs, Ipswich, United Kingdom), 0.4 µl dNTPs (25 mM each) and 1.5 µl large fragment Bst polymerase [8 U/µl] (New England Biolabs, Ipswich, United Kingdom) for adapter fill-in by incubation for 20 min at 37 °C in a PCR-cycler. The sample was than mixed again with 72 µl MagNa beads for reaction clean up like described above. The concentration of the adapter ligated DNA was determined with Qubit® dsDNA HS (High Sensitivity) Assay (Invitrogen, Carlsbad, CA, USA) according manufactory protocol (described in 2.3).

To provide the libraries with an individual index, 8 µl adapter ligated DNA containing 20 - 100 ng was used for as template for the indexing PCR by adding 1.4 µl of the appropriate (index) primer (10 µM). The total reaction mix included in addition 28.7 µl H2O, 10 µl 5x Phusion HF buffer, 0.4 µl dNTPs (25 mM each) Primer IS4 (10 µM), 0.5 µl Phusion Hot Start Flex DNA polymerase [2U/µl] (New England Biolabs, Ipswich, United Kingdom). The PCR was conducted using the protocol: initial denaturation 98 °C for 30 sec, 12 cycles of denaturation 98 °C (10 sec), annealing 60 °C (20 sec), elongation 72 °C (20 sec) followed by a final extension at 72 °C (10 min) and hold temperature at 8 °C. The sample clean-up was done by adding 90 µl MagNA beads to the sample and placing them for 5 min DynaMag-96 Side Skirted Magnetic Particle Concentrator (MPC$_{96}$) to separate the beads carrying the DNA from the solution. Washing was done 2 times by adding 150 µl 70 % ethanol and incubation for 1 min. After air drying for 30 min, the probes were eluted with in 27 µl EBT. Last step of library preparation is gel-purification which was conducted by the gel-purification of the library like described in WGS sequencing (see part 2.8.2).

For *HvLAX-A*, BACs sequenced on the Illumina MiSeq instrument by 2 x 250 bp paired end sequencing (Table A4) and sequences were assembled with CLC assembly cell version 4.0.6 (www.clcbio.com).

Twenty-six BACs for *HvMND* were sequenced on a single lane on Illumina HiSeq2000 system and raw data was assembled after quality clipping by CLC assembly cell version 4.0.6 (www.clcbio.com). Twelve BACs were sequenced in advance on the Roche/454 Titanium platform according to described procedures (IBSC, 2012) and assembled with MIRA (Chevreux, 1999). A list of sequenced BACs is summarized in Table A4.

Overlaps between sequenced BACs were detected by all against-all sequence comparisons with BLAST (Altschul et al., 1990) considering only BLAST hits longer than 2 kb and 99 % sequence identity. For the gene identification, BAC sequence contigs were first subjected to repeat masking by k-mer statistics using the Kmasker pipeline (Schmutzer et al., 2014) to exclude the repetitive parts of the sequence contigs. The software divides sequence contigs in overlapping 21-mer fragments to test with a mathematical approach how often they occur in a eight-fold Whole Genome Shotgun dataset (Schmutzer et al., 2014). All detected repetitive sequence parts that exceeded a repeat threshold of 20 over stretches longer than 100 bp were masked. The non-masked regions of the BAC sequence were used for homology search by BLAST against barley high- and low-confidence genes (IBSC, 2012) as well as ORFs of the sequenced grass genomes of rice (TIGR_Os_V6.1), Brachypodium (CDS_V1.2) and Sorghum (CDS_Sbi1.4).

In addition, automated gene prediction was performed with software tool AUGUSTUS (Stanke et al., 2004) using the maize gene models. Predicted protein sequences were functionally annotated with the AHRD pipeline (https://github.com/groupschoof/AHRD) which parses the description of BLASTP hits against the TAIR (Lamesch et al., 2012) Uniprot/trEMBL, and Uniprot/SwissProt (Apweiler et al., 2013) databases.

BAC sequence contigs were used to estimate the size of the deleted region at the *lax-a* locus within BAC FPcontig_2862. In total, 64 Bowman WGS contigs with a cumulative length of 210,643 bp could be assigned to FPcontig_2862 by BLASTN (alignment length ≥ 500; identity ≥ 99.5 %) and used as a reference to determine the mean read coverage of WGS data of BW457 and wild-type Bowman (IBSC, 2012). The position of Bowman WGS contigs in FPcontig_2862 that showed no read coverage from BW457 data were used to estimate the size of the deletion. Read depth was calculated with BEDtools (Quinlan and Hall, 2010).

2.8.5. Transcriptome sequencing (RNA-seq)

Quantitative transcriptome sequencing (RNAseq) was performed to determine differential gene expression between mutant *lax-a* and wild-type plants. Four micrograms of total RNA (see part 2.4.) in 50 µl total volume was applied for library preparation with IlluminaTruSeq RNA Sample Preparation Kit (Illumina, San Diego, CA, USA). In brief, the procedure starts with a poly A selection for mRNA which was performed by mixing the total RNA with 50 µl

RNA Purification Beads by incubation at 65 °C for 5 min and cooling to 4 °C final temperature. A phase separation was performed by adding the samples to the DynaMag-2 (Life Technologies, Invitrogen, Carlsbad, CA, USA) magnetic device for 5 min. The liquid phase was removed and exchanged by 200 µl Bead Washing Buffer and again separated for 5 min on the magnet. The aqueous phase was discarded, the sample removed from the magnet and the mRNA eluted with 50 µl Elution Buffer by heating to 80 °C for 2 min. The poly A selection step was repeated again and samples finally eluted in 19.5 µl Elute, Prime, Fragment Mix (EPF) by denaturation at 94°C for 8 min. Samples were placed 5 min on the magnetic device for phase separation and 17 µl of the liquid transferred to a new well for first strand cDNA synthesis. A mix in the ratio of 1 µl Superscript II to 9 µl First Strand Master Mix was prepared and 8 µl of them was added to the 17 µl purified RNA. The synthesis was performed in the termo cycler with the following settings: 25°C (10 min), 42°C (50 min), 70°C (15 min) and final temperature of 4°C. Second strand synthesis was done by adding 25 µl of the Second Strand Master Mix and incubation at 16 °C for 60 min. The sample was then mixed with 90µl AMPure XP beads (see part 2.2.1), 15 min incubated at room temperature and placed for further 5 min on the magnetic device. 135 µl of the supernatant was replaced by 200 µl 80 % Ethanol. After 30 sec, the Ethanol was renewed before the solution was completely removed. Samples were dried for 30 min at room temperature, removed from the magnetic device and eluted in 52.5 µl Resuspension Buffer. After 2 min incubation at room temperature, the samples were placed again on the magnet to clean the sample from the magnetic beads. Finally 50 µl of the cDNA was used to process with end repair step and the following steps of library preparation following the same principle like described above for the TruSeq DNA Sample Preparation Guide (Illumina, San Diego, CA, USA). The enrichment step by PCR was done before the gel purification. For this, 20 µl of the adapter ligated fragments were mixed with 5 µl PCR Primer Cocktail and 25 µl PCR Master Mix. The cycling was performed by initial denaturation for 30 s, 15 cycles of denaturation 98 °C (10 sec), annealing 60 °C (30 sec), extension 72 °C (30 sec) followed by a final extension of 72 °C for 5 min and a cooling to 10 °C. PCR cleaning was performed with MinElute Reaction Cleanup Kit (Qiagen, Venlo, Netherlands) by mixing PCR product with 125µl PB-buffer (Qiagen, Venlo, Netherlands) and pipetting to the membrane. The samples were washed with PE buffer (Qiagen, Venlo, Netherlands) and finally eluted with 21 µl EB buffer (10 mM Tris·Cl, pH 8.5) by 1 min centrifugation (16000 x g) respectively. Final step was the gel-purification like described above. For size fragmentation a region of 420-520 bp was cut out from the gel.

Paired end sequencing (2x100bp) was performed on Illumnia HiSeq2000 (Illumina, San Diego, CA). Four sequencing libraries were sequenced together per HiSeq2000 lane. Sequence reads were mapped against the WGS assembly of barley cultivar 'Morex' (IBSC, 2012) with Tophat version 2.0.9 (Kim et al., 2013). Read count data was processed with HTseq framework version 0.5.4 (Anders et al., 2014) using the reference annotation of the Morex WGS assembly (IBSC, 2012). The minimum mapping quality of a sequence read was set to Q10. Differential expression analysis was performed with DEseq (Anders and Huber, 2010). DEseq takes different read counts per sample into account and normalizes raw counts by applying a correction factor. Multiple testing correction was performed with the Benjamini-Hochberg method (Benjamini and Hochberg, 1995).

2.9. TILLING analysis

To identify novel mutant alleles of *HvMND*, *HvLAX-A* and *HvCUL4*, DNA of 7,979 M2 plants of an EMS-treated TILLING (Targeting Local Lesions In Genomes) population of cv. 'Barke' (Gottwald et al., 2009) was screened with two primer combinations per gene to analyze the entire ORF for mutations (HvMND_EX1, HvMND_EX2, Table A3; HvLAX_EX1, HvLAX_Ex2, Table A5; HvCUL4_Ex1, HvCUL4_Ex2, Table A6). A standard PCR reaction was used for amplification which ends with a heteroduplex step through slow cooling (Table 2). After denaturation at 99 C° 10 min a cooling procedure by 25 cycles of 70 °C, 69.7 °C and 69.4 °C for 20 sec, respectively. After each cycle all three annealing temperatures decrease by -0.9 °C per cycle. Enzymatic digestion to cleave the mismatched heteroduplex was processed by dsDNA Cleavage Kit (Advanced Analytical, Iowa, USA). Two microliters of the PCR product were mixed with 2 μl of the 1:125 in T-Digest Buffer diluted dsDNA Cleavage Enzyme (Advanced Analytical, Iowa, USA) and incubated at 60 °C for 15 to 30 min (Table 2).

Table 2: TILLING reaction conditions

	HvMND_EX1	HvMND_EX2	HvLAX_EX1	HvLAX_Ex2	HvCUL4_Ex1	HvCUL4_Ex2
PCR reaction mix						
H2O bidest.	6 µl	8 µl	8 µl	8 µl	6 µl	8 µl
Q-Solution	2 µl	/	/	/	2 µl	/
PCR-Puffer 10x	1 µl	1 µl	1 µl	1µl	1 µl	1 µl
dNTPs [5mM]	0.4 µl	0.4 µl	0.4 µl	0.4 µl	0.4 µl	0.4 µl
Primer F* [10 µM]	0.3 µl	0.3 µl	0.3 µl	0.3 µl	0.3 µl	0.3 µl
Primer R* [10 µM]	0.3 µl	0.3 µl	0.3 µl	0.3 µl	0.3 µl	0.3 µl
Taq [5 U/µl]	0.053 µl	0.053 µl	0.053 µl	0.053 µl	0.053 µl	0.053 µl
DNA	1 µl	1 µl	1 µl	1 µl	1 µl	1 µl
Annealing temperature	58 °C	60 °C	60 °C	60 °C	58 °C	58 °C
Digest temperature	60 °C	60 °C	60 °C	60 °C	60 °C	60 °C
Digest time	15 min	15 min	30 min	30 min	20 min	20 min

* F = forward, R=reverse

TILLING analyses was done by using the Mutation Discovery 910 Gel Kit (Advanced Analytical, Iowa, USA) on the AdvanCE FS96 system according manufacturer's instructions (Advanced Analytical, Iowa, USA). In brief, the digested samples were filled up with TE-buffer to a final volume of 30 µl before the run. In addition, a marker plate was prepared with 30 µl per well of a mix consisting of 3 ml TE-buffer and 5µl 100 bp and 2000 bp NoLimits DNA Fragments (Fermentas GmbH, St. Leon-Rot, Germany), respectively. Five µl Intercalating Dye were mixed with 100 ml of the ds DNA Gel solution (1:20000 dilution) for gel matrix preparation (Advanced Analytical, Iowa, USA). 1200 µl of 1 x 930 dsDNA Inlet Buffer (Advanced Analytical, Iowa, USA) was separated in a deep well plate. Finally the 5 x Capillary Conditional Solution was diluted to 1 x with $_{dd}H_2O$ for the run of the instrument. Capillarelectrophoreses was performed according suggested settings of the Mutation Discovery Kit User Manual (Advanced Analytical, Iowa, USA) with few changes: (i) Marker plate injection time 20 sec (ii) Sample injection at 7.5 kv for 45 sec. (iii) Perform CE Separation for 90 sec. Data analyses was performed with PROSize software (Advanced Analytical, Iowa, USA). The data was normalized for the upper and lower DNA markers followed by visual screening of digested fragments. Identified SNPs were confirmed by Sanger sequencing (see 2.8.1.). M3 progeny of all mutants (M2) carrying non-synonymous SNPs were propagated and their phenotype analyzed under greenhouse conditions.

2.10. Haplotype analysis

In order to study natural genetic diversity, identified candidate genes were amplified from 303 barley genotypes (see 2.1.2) and Sanger sequencing applied to both strands of PCR products (forward and reverse reaction). A haplotype analysis was performed based on re-sequencing data of the complete ORF. Sequence alignments were performed with ClustalW2 (http://www.ebi.ac.uk/Tools/msa/clustalw2/). DnaSP software package (Librado and Rozas, 2009) was used for haplotype detection under consideration of gaps. DNA ALIGNMENT 1.3.1.1 (http://www.fluxus-engineering.com) was used to produce the .rtf input file for the NETWORK 4.6.1.2 software (http://www.fluxus-engineering.com) to generate a Median-Joining network (Bandelt et al., 1999).

2.11. Phylogenetic analysis

To identify closely related paralogous genes or gene family members in barley, a protein sequence homology search by BLASTP (Altschul et al., 1990) was performed against the barley high- and low-confidence gene set (IBSC, 2012). To identify all members of the Blade-on-petiole-like (BOP-like) genes in the plant kingdom, the NCBI protein and Phytozome (http://phytozome.jgi.doe.gov) database were queried. A sequence similarity based conserved sequence search on NCBI (http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi) was applied to all indentified proteins as control for presence of both conserved domains (BTB and ANK domains). Multiple protein sequence alignment was performed using MUSCLE (Edgar, 2004a, b). The phylogenetic analysis was performed using MEGA6 software (Tamura et al., 2013) following the published protocol (Hall, 2013). Maximum Likelihood (ML) tree was constructed using JTT model with discrete Gamma distribution and Nearest Neighbor Interchange (NNI) by applying 1000 Bootstrap replicates.

2.12. Data generated in collaborative efforts

This thesis includes data which resulted from scientific collaboration. For clarification which data resources were generated and provided by the PhD candidate or by collaborative partners, please see the list below. The collaborating efforts are listed separately for the characterization of the *many noded dwarf* and *laxatum-a* mutant to keep clarity:

Molecular characterization of the *many noded dwarf* mutant

- Experimental design and managing work of the master students Axel Aßfalg[1] and Joel E. Kuon[1]: Matthias Jost
- Synteny based physical anchoring and selection of BACs for sequencing: Matthias Jost
- BAC annotation: Matthias Jost
- Mutant analysis by TILLING and resequencing of *mnd* mutants from Nordic Genbank: Matthias Jost
- DNA preparation, quantification and pooling for the mapping-by-sequencing experiment: Axel Aßfalg[1]
- Preparation and quantification of sequencing libraries; sequencing: Dr. Axel Himmelbach[1]
- Next generation data analysis: read mapping, SNP frequency detection, read coverage analysis, graphical visualization: Dr. Martin Mascher[2]
- CAPS marker design: Dr. Martin Mascher[2]
- Genetic mapping and phenotyping: Joel E. Kuon[1]
- PCR based physical anchoring of HvMND: Joel E. Kuon[1]
- BAC assembly: Sebastian Beier[2]

Molecular characterization of the *laxatum-a* mutant

- Marker development and high resolution mapping genetic mapping; identification of recombinant plants for mapping-by-sequencing by exome capture: Matthias Jost
- RNA, DNA isolation and quantification for next generation sequencing approaches: Matthias Jost
- Evaluation of candidate targets obtained from NGS data analysis: Matthias Jost

- Functional validation of the candidate gene by TILLING: <u>Matthias Jost</u>

- Allelism test: <u>Matthias Jost</u>

- Natural diversity analysis: <u>Matthias Jost</u>

- BAC annotation: <u>Matthias Jost</u>

- BAC assembly: <u>Sebastian Beier</u>[2]

- Mutant analysis of independent *lax-a* mutant alleles from Nordic Genbank: <u>Prof. Shin Taketa</u>[3]

- Preparation and quantification of sequencing libraries; sequencing: <u>Dr. Axel Himmelbach</u>[1]

- Next Generation Sequencing data analysis; read mapping, SNP frequency detection, read coverage analysis, differential expression analysis, graphical visualization: <u>Dr. Martin Mascher</u>[2]

- WGS sequencing of the mutant NIL BW457: <u>Dr. Stefan Taudien</u>[4], <u>Dr. Matthias Platzer</u>[4]

- Sequence assembly and SNP determining between BW457 and Bowman: <u>Thomas Schmutzer</u>[2] and <u>Dr. Burkhard Steuernagel</u>[2]

[1] Genome diversity group, IPK Gatersleben, Stadt Seeland, Germany

[2] Bioinformatics and Information Technology group, IPK Gatersleben, Stadt Seeland, Germany

[3] Institute of Plant Science and Resources, Okayama University, Kurashiki, Japan

[4] Leibniz Institute for Age Research – Fritz-Lipmann-Institute, Jena, Germany

3. Results

3.1. Cloning of the gene *MANY NODED DWARF (MND)*

3.1.1. Phenotyping

In a previous study seeds of barley cv. 'Saale' were treated with X-ray radiation to induce mutations (Scholz, 1962) and revealed a mutant (MHOR474) with altered phenotype called *many noded dwarf* (*mnd*), which showed faster leaf initiation (shortened plastochron), higher number of nodes and a dwarfed growth habit (Figure 4a-c). In order to investigate the underlying mutated gene a segregating population was established by crossing the mutant MHOR474 to cv. Barke and self pollination of the F1 plants. 100 F2 plants were screened under greenhouse conditions for phenotypic segregation: 81 wild-type and 19 *mnd* mutant plants were observed resulting in a 3:1 segregation ($\chi2 = 1.92$, P value = 0.165) of a single recessive gene. Twenty F3 family members per F2 genotype were grown to confirm the phenotype scored in the F2 generation. The previously reported pleiotropic morphological effects of the *mnd* mutation were tested by growing plants under field conditions and scoring tiller number, plant height and spike length. All mutant plants showed about two times higher number of nodes and increased number of tillers in comparison to wild-type plants (Figure 4e). The dwarfed phenotype was only observed under field conditions. Plants grown in the greenhouse showed no dwarfism compared to wild-type plants (Figure 4d).

3.1.2. Delimiting a target interval by a mapping-by-sequencing approach

To identify the genomic location of the mutated locus, which is responsible for the pleiotropic growth changes of *mnd* mutant plants, a bulked segregant analysis (Michelmore et al., 1991) in combination with next generation sequencing (NGS) was performed. Leaf material of all F2 plants was sampled for DNA extraction to initiate an attempt of mapping-by-sequencing (see part 1.3.2). DNA of 18 mutant plants and 30 randomly selected wild-type plants were separately collected for two independent DNA pools, respectively. For complexity reduced re-sequencing these pools were applied to exome capture (Mascher et al., 2013c). The liquid array allows to enrich for 61.5 Mbp target sequence preferentially consisting of protein coding genes allowing to re-sequence around 80% of all barley genes (Mascher et al., 2013c).
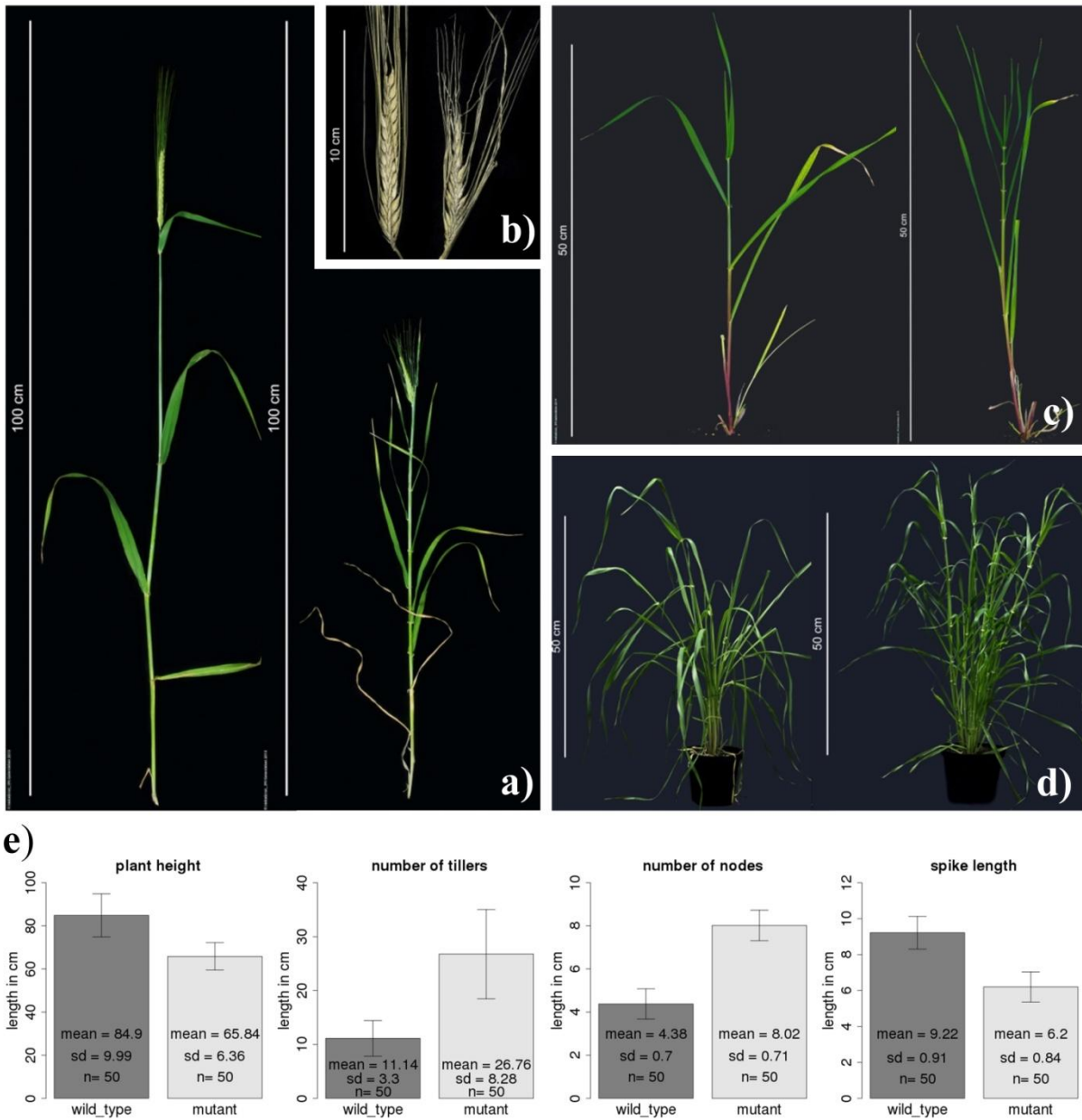
Figure 4: Phenotypic characteristics of *mnd* plants.

(a) Mutant plants (right panel) have a significantly higher number of nodes compared to the wild-type plant (left panel) and showed a semi-dwarf growth habit. (b) Spike lengths of mutants (right panel) were reduced in comparison to wild-type plants (left panel). (c) Faster leaf formation could be observed already in early developmental stages for mutants (right) compared to the wild-type (left). (d) Mutant plants (right) have more internodes without a dwarfing phenotype when they were grown under greenhouse conditions. e) Comparison of average number of nodes, plant height, number of tillers, and spike length (left to the right) of wild-type (left) and mutants (right) scored for 50 plants respectively under field-like conditions. Pictures taken from (Mascher et al., 2014).

Both DNA pools were used to prepare a separate sequencing library which was provided with an individual index, respectively. This allowed combining them for capturing the gene space in a single reaction and proceeding with the sequencing on a single lane of the Illumina HiSeq2000 sequencing system. In total 82 and 70 million sequence read pairs (2x 100 bp) were obtained for mutant and wild-type sample, respectively. To detect sequence variants sequence reads were mapped against the barley whole-genome-sequencing (WGS) assembly of cv. Barke (IBSC, 2012), the wild-type parent of the mapping population. SNP frequencies were determined at each polymorphic position by calculating the ratio of reference wild-type allele and distinct alternative allele of all mapped reads. The causal mutation was expected to be located in a position where the mutant genotype-specific SNP frequency in the mutant pools reaches 100 % (homozygote mutant reads) and in the wild-type pool the mutant allele frequency decreases to 25 %. To identify this region, the physical, genetic and functional sequence assembly of barley (IBSC, 2012) was used as a reference to visualize the obtained SNP frequencies in a physical distribution along the seven barley chromosomes (Figure 5a). Only a single peak on chromosome 5H showed the expected SNP frequency reaching up to ~97 % mutant allele for the mutant pool and as low as ~35 % mutant allele in the wild-type pool. The SNP frequency reached a maximum of 97% in the mutant pool at around 97 cM of the physical/genetic map framework (IBSC, 2012). The portion of wild-type reads found in the mutant pool originated from one falsely phenotyped heterozygous F2 plant, as could be confirmed by phenotyping the respective F3 progeny.

3.1.3. Genetic mapping

Mapping-by-sequencing of the *MND* locus revealed a 30 cM target interval on the long arm of chromosome 5H. The exome capture sequencing data of the two contrasting pools delivered directly a large amount of sequence polymorphism information (SNPs). In order to confirm the initial mapping-by-sequencing result, new markers were developed by taking advantage of physically evenly distributed SNPs of the target interval as determined on the basis of their position in the barley physical/genetic reference (IBSC 2012). Polymorphic SNPs were converted to ten single PCR-based marker assays for the interval between 80 and 110 cM of chromosome 5H (Table 3, Table A7). Genotyping was performed in the same population of 100 F2 plants that were used before to select plants for the pools of the mapping-by-sequencing approach. The *MND* locus was mapped to a 3.5 cM interval in the

resulting linkage map (Figure 5b), thus confirming genetic mapping based on SNP frequencies distribution obtained from bulked segregant analysis.
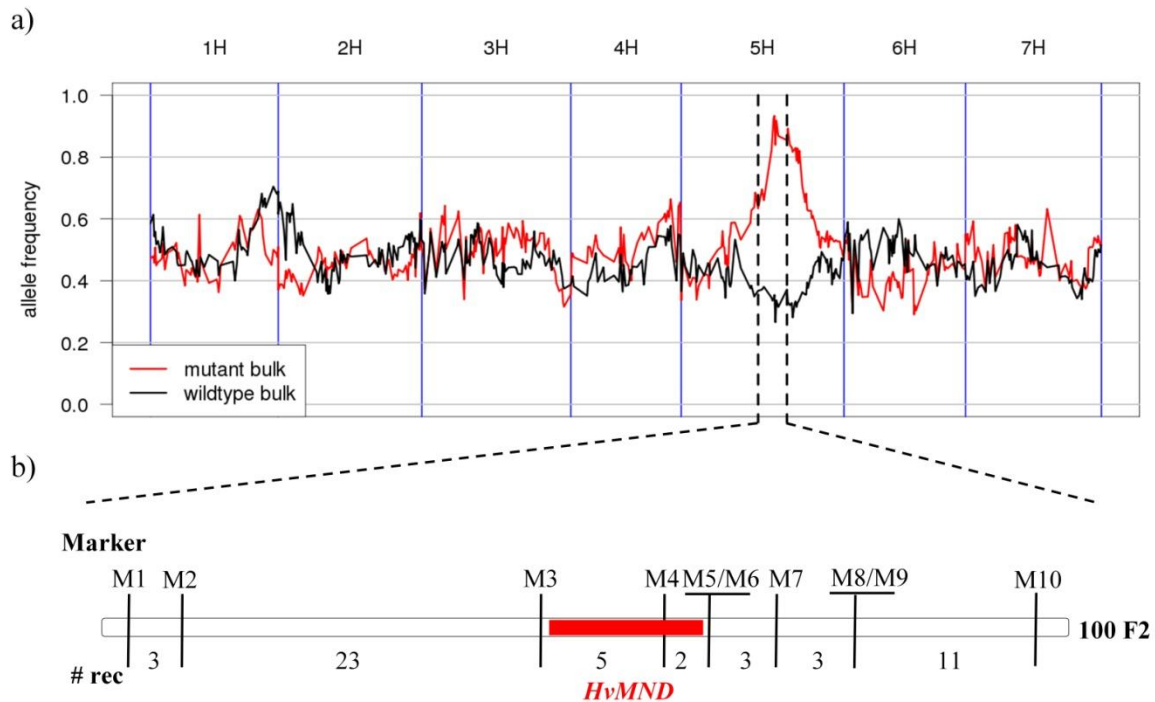
Figure 5: Defined mapping interval for *HvMND*.
A genetic interval for *HvMND* was defined by analyzing SNP frequencies obtained from sequencing of phenotypic pools in a genomic context. a) Allele frequency of mutant pool (red) and wild-type pool visualized along the integrated physical genetic map (IBSC, 2012) along the seven chromosomes (separated by blue lines). The dashed lines delimiting the 30 cM confidence interval which was predicted according frequency distribution of mutant and wild-type pool. b) Ten SNPs, distinct between mutant and wild-type pool, were extracted and converted to CAPS markers for genotyping 100 F2 plants of the same population that was used for mapping-by-sequencing. Markers (black lines) were schematically displayed based on the number of observed recombination events (indicated below by numbers). The gene could be mapped to a 3.5 cM interval, flanked by the proximal marker M3 (five recombination events) and distal markers M5/M6 (two recombination events). Marker M4 cosegregated with the phenotype. Picture modified from Mascher et al. (2014) (Mascher et al., 2014).

Table 3: Detected frequencies of mapped SNPs and positions in the barley reference sequence

Marker	Barke_WGS	Morex_WGS	FPC[1]	cM[1]	cM[2]	SNP frequency* mutant	WT
M1	contig_1794173	contig_1563519	43818	80.34	80.34	0.65	0.35
M2	contig_274977	contig_1582922	no FPC	83.95	84.37	0.73	0.38
M3	contig_1781303	contig_37730	1020	95.12	95.48	0.95	0.35
M4	contig_265820	contig_67963	45097	96.59	96.59	0.97	0.35
M5	contig_1783058	contig_134574	46058	97.29	97.29	0.95	0.35
M6	contig_267961	contig_117772	46058	97.29	97.7	0.96	0.35
M7	contig_2786329	contig_1567878	NA	NA	NA	0.95	0.32
M8	contig_269740	contig_42302	no FPC	98.8	NA	0.91	0.33
M9	contig_481335	contig_49640	48127	98.88	99.93	0.94	0.29
M10	contig_55926	contig_41272	39059	109.65	109.65	0.92	0.36

[1](IBSC, 2012); [2 (Mascher et al., 2013b)]; * determined SNP frequencies in sequence reads of mutant and wild- type (WT) pool

3.1.4. Read coverage analysis leads to identification of a single candidate gene

The type of mutagenesis used to generate the *mnd* mutation provided an alternative option to query the mapping-by-sequencing dataset for potential candidate genes. Irradiation with X-ray is known to often cause larger deletions in the genome (Nelson et al., 1994) including the deletion of an entire gene or several genes. Sequencing reads were mapped against the 'Morex' whole genome sequence assembly to indentify putatively deleted exome capture targets in the mutant pool. Read coverage was calculated for each bp position of the WGS sequence contigs for both, mutant and wild-type pool. Filtering for targets that, over stretches of at least 150 bp, exhibited a coverage of zero in mutant pool and more than 5-fold coverage in the wild-type pool was performed to identify putatively deleted regions. Due to the wrongly phenotyped heterozygous plant, that was accidentally included to the mutant pool (see 3.1.2), a certain number of reads was still expected even in truly deleted targets of the mutant. Thus also targets were considered as putative candidates that showed a four times higher coverage in wild-type pool than in the mutant pool, while their coverage in the mutant pool was smaller than two and higher than five in the wild-type pool. In total, 435 potentially deleted targets on WGS contigs were identified by this query for the entire genome. Eighteen of these targets, including six annotated genes (Table 4), were assigned to the defined *MND* locus confidence interval (80-110 cM) (Mascher et al., 2013b) (Table A8). Among them MLOC_64838.2, which was the only gene located within the narrowed genetic interval (Table

4) characterized by the SNP derived markers (see 3.1.3). Conserved domain search identified MLOC_64838.2 as a member of the cytochrome P450 class 78a (CYP78A) gene family. One member of this gene family was previously characterized in a rice mutant as exhibiting a similar phenotype to barley *mnd*. Mutation in the gene *plastochron 1A* (*pla1*) leads to changes of the plastochron resulting in faster leaf initiation and higher number of leaves in rice (Miyoshi et al., 2004). However, the rice *PLA1* is not the best homolog of MLOC_64838.2. The barley gene has higher sequence similarity with rice gene Os09g35940.1, which is also consistent to the conserved syntenic location of both homologs on barley chromosome 5H and rice chromosome Os09, respectively (Thiel et al., 2009). The *PLA1* gene of rice has no direct homolog in barley but showed equally good DNA sequence similarity to other CYP78A gene family members of the barley genome. Three additional members of cytochrome P450 class 78a (CYP78A) genes could be identified within the barley reference gene set (IBSC, 2012) based on sequence similarity. MLOC_64838.2 has a distinct expression pattern considering published expression data obtained from eight different tissues (IBSC, 2012). In contrast to the other members, MLOC_64838.2 reached highest expression in developing inflorescences whereas the remaining members reached their maximum expression in 4 days old embryo and roots from young seedlings. This may hint at distinct functionalities in comparison to *HvMND* (Figure 6).

Thus, MLOC_64838.2 was considered a candidate gene for *mnd.* In the following the gene name *HvMND* will be used instead of the gene locus identifier MLOC_64838.2. As a result of the *in silico* read coverage analysis, it was indicated that the gene *HvMND* was deleted in the mutant line MHOR474. To test this result, three PCR primer pairs (HvMND_F/R1, HvMND_F/R2, HvMND_F/R4; Table A3) were designed based on the Morex reference sequence information in order to amplify the entire open reading frame (ORF) of the gene including introns from mutant and wild-type plants. While all primer combinations were suited to amplify the gene from wild-type Barke no amplification was observed when mutant MHOR474 was used as a template.

Table 4: Putatively deleted genes with low sequence read coverage in the mutant pool within confidence interval

Morex WGS contig ID[1]	Start	End	Coverage*		barley high confidence gene[1]	Annotation**	Genetic position [cM][2]
			MT	WT			
contig_49382	2107	2455	0.4	24.4	MLOC_64838.2	Cytochrome, P450	96.6
contig_49382	3871	4101	1.1	8.8	MLOC_64838.2	Cytochrome, P450	96.6
contig_159829	785	996	0.9	8.2	MLOC_21734.1	CC-NBS-LRR	99.9
contig_159829	4501	4709	0.4	12.4	MLOC_21734.1	CC-NBS-LRR	99.9
contig_1558349	1236	1477	1.5	11.5	MLOC_10070.1	MATE efflux	99.9
contig_45126	14800	14961	0.8	7.3	AK365660	Tumor susceptibility	101
contig_2547452	4712	4913	0.1	10.5	AK357178	3-5 exoribonuclease CSL4	108.1
contig_58347	7926	8186	1	14.5	MLOC_70680.1	DNA repair protein	109.4

*average read coverage of mapped sequence reads from the mutant and wild-type pool to target regions on the Morex contigs (defined by start and end coordinates); **available predicted functional annotation information (IBSC, 2012); [1](IBSC, 2012); [2](Mascher et al., 2013b)
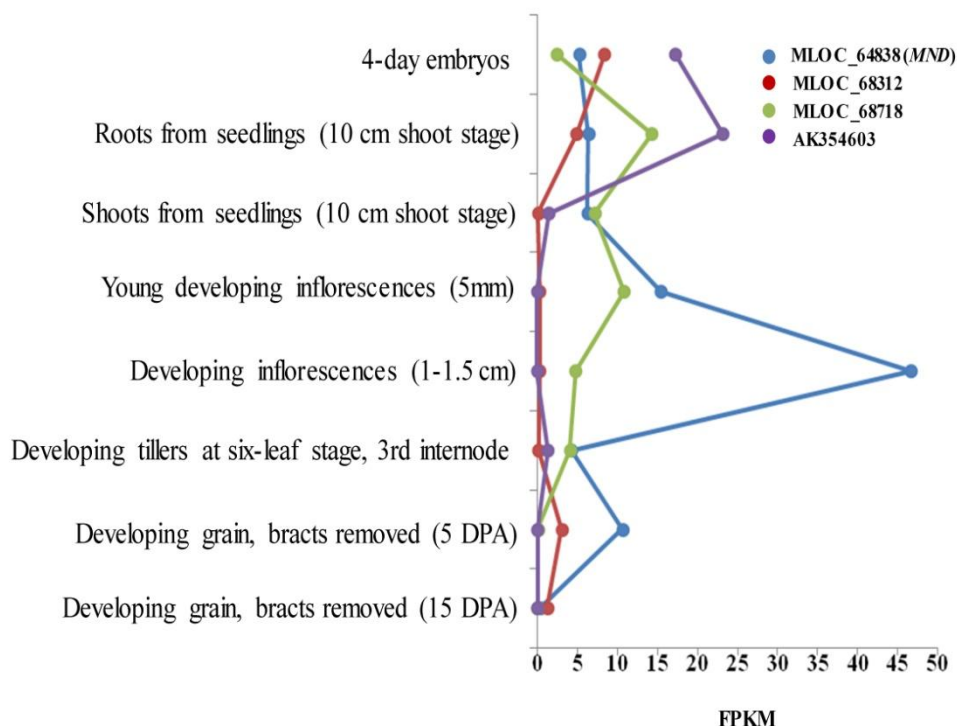
Figure 6: Expression pattern of the CYP78A gene family of barley.
Expression values (x-axis) from of eight different tissues (y-axis) were plotted for the four CYP78A family members of barley. Picture modified from Mascher et al. (2014)

3.1.5. A physical contig carrying the *HvMND* candidate gene

The identified candidate gene was completely deleted in the mutant genotype. To address the question if the candidate gene was the only deleted gene, and thus most likely responsible for controlling the phenotype, it was aimed to identify the flanking genes. The candidate gene *HvMND* was annotated based on WGS contigs of Morex, Barke and Bowman which were not yet integrated in the physical map (IBSC, 2012). Thus, the current status of the barley reference sequence did not allow identification of physically linked genes in the direct proximity to the candidate gene. The putatively flanking genes were deduced from synteny information of the model organism Brachypodium, whose genome reference sequence is available (The International Brachypodium Initiative, 2010). Seven barley genes, homologous to two downstream and five upstream Brachypodium gene models were tested by PCR amplification for presence or absence due to complete deletion (Table 5, Table A9).

Table 5: Syntenic Brachypodium interval flanking the homolog of the *HvMND* candidate gene

Brachypodium ID[1]	Barley homologous gene ID[2]	FPcontig[2]
Bradi4g35867	AK368025	45097
Bradi4g35880	AK355945	45097
Bradi4g35890	**MLOC_64838.2**	**45097***
Bradi4g35900	MLOC_54310.2	46058
Bradi4g35910	AK376953	n.a.**
Bradi4g35930	MLOC_52879.5	46058
Bradi4g35940	MLOC_22687.1	46058
Bradi4g35950	AK374133	46058

[1]Brachypodium CDS v1.2; [2] (IBSC, 2012);
* anchored by PCR based library screening. **n.a. = no similarity to a physical anchored sequence

All seven genes could be amplified in mutant parent MHOR474 and wild-type Barke. The identified candidate gene remained the only one that could not be amplified in the mutant parent but in Barke, provided first evidence for a single gene deletion. Nevertheless, a lack of collinearity and the principle possibility of additional non-syntenic genes being present at the barley *mnd* locus could not be ruled out by this analysis. The ultimate proof of a single gene deletion required to reveal the complete sequence of the target locus flanked by the closest markers in barley. In contrast to the candidate gene, the predicted neighboring genes could be assigned to the physical map and identified two physical BAC contigs (FPcontig_45097 and FPcontig 46058) (Table 5). A single BAC clone HVVMRXALLhB0080C03 was identified

by PCR to carry *HvMND*. The identified clone was belonging to FPcontig_45097, which was already identified with flanking genes AK355945 and AK368025 (Table 5). The distally flanking marker M5 was located on FPcontig_46058 together with four of the predicted neighboring genes. The co-segregating Marker M4 was found to be on FPcontig_45097 together with the candidate gene *HvMND*. The proximally flanking marker M3 was separated by five recombination events from *HvMND* and was allocated to FPcontig_1020 (Figure 7a). Since the SNP of Marker M3 is located in the gene MLOC_52707.1, which is homologous to Bradi4g35710.1 and Marker M4 (AK355945) homologous to the Brachypodium gene model Bradi4g35880.2, 16 Brachypodium gene models were expected to be located within the gap in case of conserved collinearity. No further anchored physical FPcontig of this region in the physical map harbored a gene which would fit to the collinear synteny order of Brachypodium genes. Thus, it was assumed to close the gap with the entire sequence information of the FPcontigs. BAC clones representing the FPcontig_45097 and FPcontig_46058 were sequenced. The sequence of FPcontig_1020 was already generated in advance in frame of the barley genome sequencing project and the sequences were kindly provided for this analysis. The obtained sequences were used to detect overlaps between BACs by sequence comparison (BLAST, Altschul et al., 1990) with a minimum hit length of 2000 bp and a identity higher than 99,5 %. Like assumed, overlapping BAC clones of FPcontig_45097 and FPcontig_46058 could be detected (Figure 7a, Table A10 and Table A11). Unfortunately, no overlapping sequences could be identified between FPcontig_1020 and BAC_contig_45097 (Table A12). Ten out of sixteen predicted genes were not found in the BAC sequences. Two of the missing Brachypodium gene models, Bradi4g35840 (homologous to MLOC_70099.11) and Bradi4g35800 (homologous to MLOC_62256.1), could be assigned, based on their respective morex_contig_57438 and morex_contig_46356, to FP_contig_45219 (HVVMRXALLeA0202J12) and FP_contig_45903 (HVVMRXALLhA0201C20) by BLAST to BACend sequences (IBSC, 2012). Both FPcontigs were not anchored within the physical map framework (IBSC, 2012). These relatively short physical contigs (~227 kb and ~236 kb) contigs were not sequenced and just used to extrapolate the potential physical distance of the target interval. Including these two small physical contigs a physical to genetic ratio of ~320 kb / cM could be estimated for the flanking marker M3 and marker M5.

No complete contig of overlapping BAC clones could be identified between the gene flanking markers. The available sequence information, however, was used for gene prediction by BLAST. Twenty-six previously annotated barley genes (IBSC, 2012) could be identified within the genetic mapping interval by sequence similarity search. All of them confirmed an
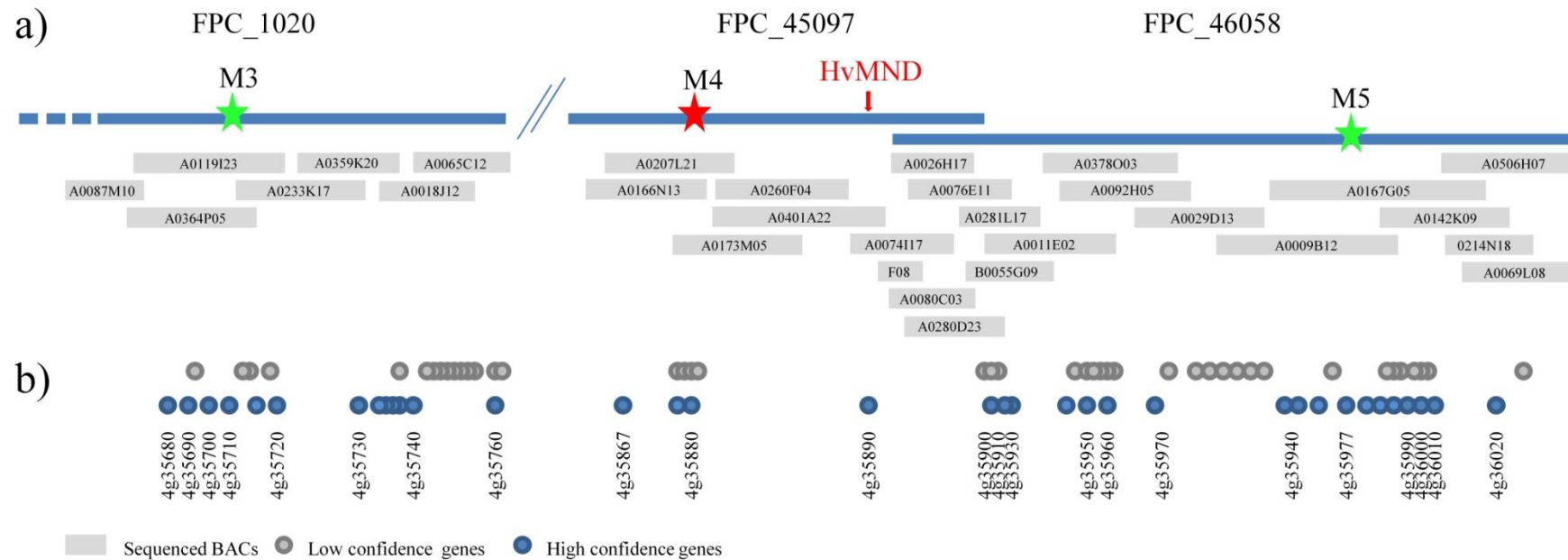
Figure 7: Physical BAC contig of the *MND* locus.

Contigs of overlapping BAC clones carrying flanking markers (green stars), one co-segregating marker (red star) and the gene *HvMND* (labeled by red arrow). Sequenced BACs were positioned according to finger printing coordinates of the physical map (IBSC, 2012). Details of BAC clone overlaps are summarized in Tables A10, A11 and A12. b) Genes annotated based on obtained sequences from BAC sequencing are visualized according their position. Blue dots represent high confidence genes whereas grey dots indicate the positions of low confidence genes from the barley gene set (IBSC, 2012), respectively (Table A14 and A15). Vertical labels are indicating orthologous Brachypodium gene models – their numbering indicates complete collinearity between barley and Brachypodium.

overall conserved synteny to Brachypodium (Figure 7b, Table A13). No gene in direct proximity of the candidate gene was deleted. Thus, the functional validation was initiated to confirm the candidate gene as causal gene for the *mnd* phenotype.

3.1.6. Functional validation of the *mnd c*andidate gene

The final step of a gene cloning project is the functional validation of the candidate gene. Additional mutant alleles in independent genetic background can be used to study the correlation of the gene with the expected phenotype. If those mutants show phenotypic similarity to the analyzed *mnd* mutant it can be concluded that the gene controls all the observed pleiotropic changes. Thus a set of available barley plant resources were evaluated for additional mutant alleles within the gene *HvMND* (MLOC_64838.2). For this, two strategies were used: (i) TILLING for the identification of new alleles for the candidate gene. (ii) Resequencing the gene from available accessions described as many noded dwarf plants.

A population of 7,979 EMS-mutagenized plants of c.v. Barke (Gottwald et al., 2009) was screened by two PCR fragments (HvMND_EX1_F/R, HvMND_EX2_F/R; Table A3) for mutations in the ORF of the *HvMND* candidate gene. Twenty synonymous, seventeen non-synonymous SNPs and one polymorphism leading to a premature stop codon, possibly resulting in a truncated protein upon translation, were identified (Figure 8a, Table 8). M3 progeny of all mutants carrying non-synonymous mutations as well as carrying the premature stop codon mutation were phenotyped under greenhouse conditions. The M3 progeny of the plants 4098_1 and 13490_1 segregated for the phenotype of *mnd* and genotyping of *HvMND* always showed perfect correlation of homozygous mutant genotype with phenotype (Table 6). Plants homozygous for the premature stop codon caused by SNP G261A (W78*, 4098_1) or for the non-synonymous SNP G1651A (G508E, 13490_1) showed the *mnd* characteristic phenotype (Figure 8b). In both cases, only genotypes carrying the homozygous mutant allele showed the *mnd* characteristic phenotype (Figure 8d). The dwarfism could not be observed, probably due to cultivation under greenhouse conditions, similar to the original *mnd* mutant as described above. Mutant plants reached enormous height through secondary tillers emerging from branches at the basal node of the spikes (Figure 8c).
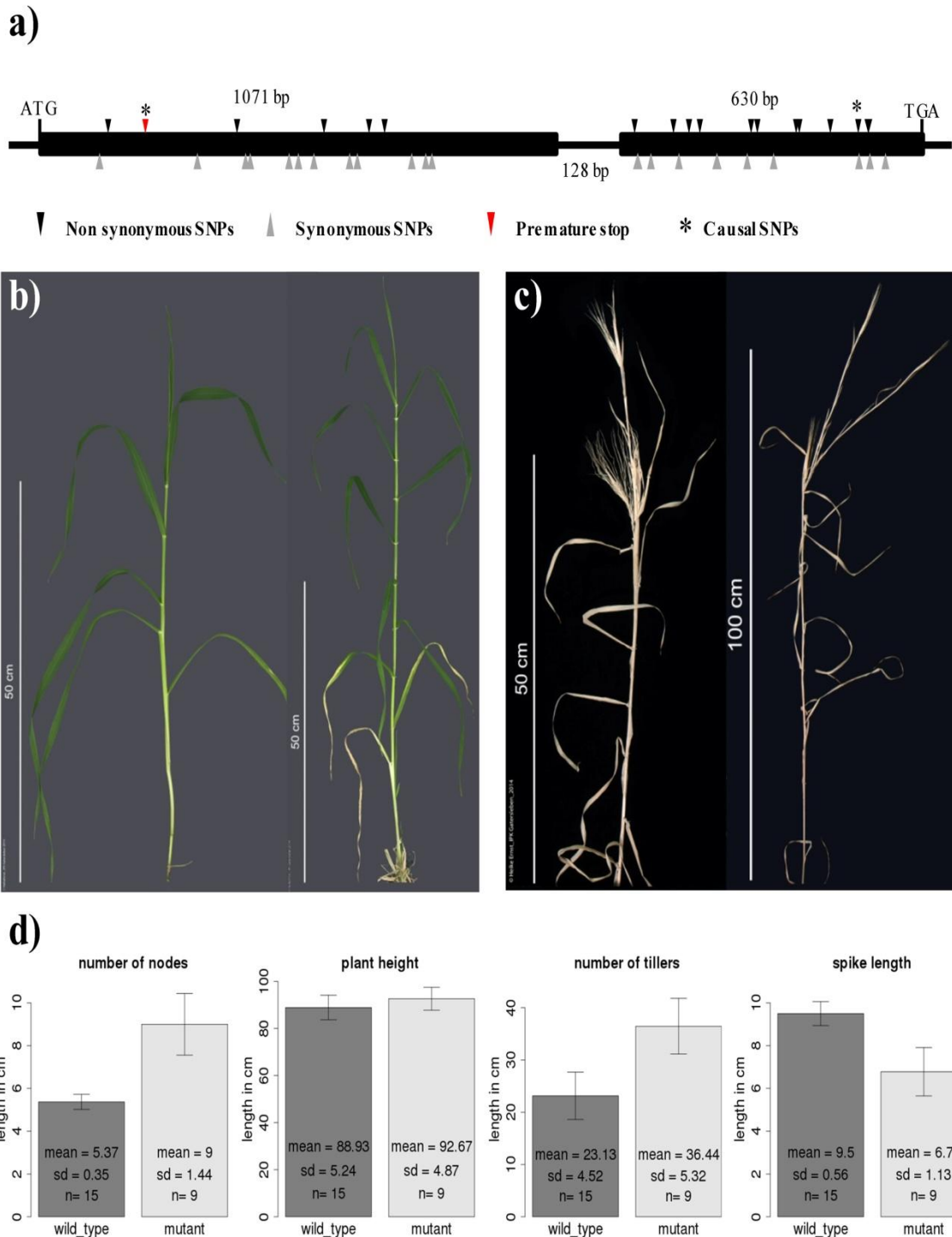
Figure 8: Independent mutant alleles for *HvMND* could be detected by TILLING.
Distribution of non-synonymous (black) and synonymous (grey) SNPs within the two exons of *HvMND*. The SNP causing a pre-mature translation stop within the first exon is highlighted by a red triangle. b) Plants homozygous for the mutant allele (right) showed faster leaf initiation compared to wild-type. c) Upper part of two mutant plants to show the branching that occurred mostly on a single tiller of mutant plants. d) Average number of nodes, plant height, number of tillers, and spike length (left to the right) of wild-type (left) and mutants (right) was compared by analyzing M3 families of TILLING mutant line 4098_1 (carrying the premature stop codon).

Table 6: Segregation of mnd TILLING alleles

Plant	SNP	Effect	Wild-type	Heterozygote	Mutant
4098_1	G261A	W87*	4	15	5
13490_1	G1651A	G508E	1	2	3

As a second resource, the nearly isogenic lines (NIL) collection harboring various morphological mutant alleles introgressed into genomic background of cv. Bowman (Druka et al., 2011) was inspected for *mnd* alleles. Among these lines, seven are described harboring a mutant allele causing an *mnd* phenotype (Table 7). The gene *HvMND* was amplified (primers: HvMND_EX1_F/R, HvMND_F/R2, HvMND_F/R3, Table A3) from all seven lines and amplicons covering the entire ORF were re-sequenced. A complete deletion allele (BW522) and one non-synonymous SNP in the CYP450 domain (BW520) could be identified. The other four BW NILs showed wild-type genotype, which was consistent to their previous mapping information (Druka et al., 2011). Only BW520 and BW522 were described as carrying introgressions of mutated alleles on chromosome 5H (Table 7).

Table 7: Sequence variation of *HvMND* in Bowman nearly isogenic lines

Name	Allele	Chr	Mutagenesis	original cultivar	Backcross	Mutation	Effect
BW516	mnd.f	NA	Spontaneous	UT1713	BC3	/	
BW517	mnd.h	NA	Fast neutron	Steptoe	BC1	/	
BW518	mnd1.a	4HL	Spontaneous	Mesa	BC8	/	
BW519	mnd3.d	3H	Gamma-ray	Montcalm	BC6	/	
BW520	mnd4.e	5HL	ethyl methansulfonate	Akashinriki[1]	BC6	G1642A	R50K
BW521	mnd5.g	NA	Spontaneous	Logan/ND15053	BC3	/	
BW522	mnd6.6	5HL	ethylene imine	Bonus	BC6	complete deletion	

[1] No reference DNA of Akashinriki available, we excluded all SNPs that are shared with cv. Morex and Barke

Thirty-seven accessions available from the Nordic Genbank (NordGen, Alnarp, Sweden) are classified as *mnd* mutants. Although no information about the underlying *mnd* loci was available, the large number of mutants indicated to be a useful resource to identify further alleles for the addressed 5H locus. In all 37 accessions independent mutations of the following categories were found: six complete deletions of the gene, 26 premature stop codons, four with non-synonymous SNPs leading to amino acid changes, one splice site disruption and one 107 bp deletion within the second exon (Table 9). All plants were grown in the greenhouse to confirm the phenotype as provided by the description of Nordic

Genebank. All showed mnd characteristic plant growth but lacked like the other mutants the dwarfism growth due to the greenhouse cultivation.

In summary, two new functional alleles from TILLING were identified which showed, in a homozygous state, the *mnd* characteristic growth habit. In addition, a large spectrum of various functional mutation events was identified by analyzing an *mnd* mutant collection from Nordic Genbank (NordGen, Alnarp, Sweden). The identified mutant alleles resulted in different genomic backgrounds but all showed phenotypic similarity in expressing the *mnd* characteristic phenotype. Thus, independent mutant alleles confirmed that MLOC_64838.2 is the gene underlying the 5H *mnd* phenotype.

Table 8: TILLING mutants of *HvMND*

Accession	SNP[1]	M3 status[2]	Effect
13339_1	C160T	HET	P54S
6869_1	C180T	HOM	L60=
4098_1	G261A	HET	W87*
3895_1	G306A	HOM	R102=
4642_1	G435A	HET	M145I
6631_1	G465A	HET	R155=
4555_1	G468A	HET	E156=
9906_1	C609T	HET	S203=
13573_1	G639A	HOM	Q213=
7036_1	C657T	HET	S219=
16177_1	C671T	HOM	A224V
4610_1	G691A	HET	G231S
3158_1	G699A	HOM	E233=
6465_1	G736A	HET	V246M
9172_1	G755A	HET	G252D
2983_1	G846A	HOM	Q282=
16204_1	C861T	HET	D287=
13675_1	C864T	HOM	H288=
14226_1	G1221A	HET	D365N
15112_1	C1251T	HOM	L375=
10971_1	G1289A	HET	R387=
11186_1	G1314A	HOM	V396M
11295_1	G1319A	HET	G397=
15186_1	G1332A	HET	V402I
10869_1	C1365T	HOM	L413F
10498_1	C1389T	HOM	L421=
8989_1	C1460T	HET	G444=
10895_1	C1464T	HET	L446F
4786_1	C1471T	HET	P448L
15867_1	C1511T	HET	T461=
7736_1	G1566A	HET	A480T
13507_1	G1569A	HOM	G481R
15012_1	G1642A	HOM	S505N
13490_1	G1651A	HET	G508E
13955_1	C1678T	HET	A517V
14646_1	C1707T	HET	L527=
14033_1	C1724T	HOM	S532=
10243_1	C1748T	HET	N540=

[1] genomic DNA coordinates of Barke starting from A of ATG translation start; [2] HOM = HOM, HET = HET

Table 9: Sequence variation of the gene *HvMND* as found by resequencing of independent *many noded dwarf* mutants from Nordic Gene Bank

Accession	Mutagen[1]	Parent cv.	Mutation**	Effect
NGB114513.1*	unknown	BONUS	/	/
NGB114514.1*	EO	BONUS	G231A	stop
NGB114515.1*	ethylene imine	BONUS	/	/
NGB114516.1	Neutrons	FOMA	/	/
NGB114517.1	ethylene imine	FOMA	G255A	stop
NGB114518.1	ethylene imine	FOMA	107 bp deletion (1304-1410 bp)	
NGB114519.1	ethylene imine	FOMA	A833G	Tyr/Cys
NGB114520.1	epichlorhydrine	FOMA	G1688A	stop
NGB114521.1	X-ray	FOMA	C1368-	frameshift
NGB114522.1	EMS	FOMA	G231A	stop
NGB114523.1	EMS	FOMA	/	/
NGB114525.1	ethylene imine	FOMA	/	/
NGB114526.1	EMS	FOMA	G255A	stop
NGB114527.1	ethylene imine	KRISTINA	A1387T	Glu/Val
NGB114528.1	EMS	FOMA	G231A	stop
NGB114529.1	N-methyl-N-nitrosourea	KRISTINA	G264A	stop
NGB114530.1	Fast Neutron	KRISTINA	C1367-	frame shift
NGB114531.1	EMS	KRISTINA	C1555G	Pro/Arg
NGB114532.1	EMS	KRISTINA	G618A	stop
NGB114533.1	EMS	KRISTINA	G876A	stop
NGB114534.1	methyl methansulfonate	KRISTINA	G1531A	stop
NGB114535.1	X-ray	KRISTINA	G876A	stop
NGB114536.1	EMS	KRISTINA	G876A	stop
NGB114537.1	EMS	KRISTINA	/	/
NGB114538.1	EMS	KRISTINA	T747C	stop
NGB114539.1	EMS	KRISTINA	T661C	stop
NGB114540.1	iso-propyl methanesulfonate	KRISTINA	/	/
NGB114541.1	iso-propyl methanesulfonate	KRISTINA	T1801A	stop
NGB114542.1	EMS	KRISTINA	G858A	stop
NGB114543.1	iso-propyl methanesulfonate	Sv 71120	/	/
NGB114544.1*	X-ray	BONUS	/	/
NGB114545.1	Fast neutrons	BONUS	G876A	stop
NGB114546.1	NaAz	BONUS	C619T	stop
NGB114547.1	EMS	BONUS	/	/
NGB114548.1	NaAz	BONUS	G231A	stop
NGB114550.1	NaAz	BONUS	G500A	Arg/His
NGB119361.1	EMS	LINA	G1202A	splice site

* not possible to amplify first primer combination

** positions based on genomic DNA starting on ATG

[1] EMS = ethyl methansulfonate; EO = ethylene oxide; NaAz = sodium azide

3.2. Cloning of the gene *LAXATUM-A*

3.2.1. Phenotyping

The *laxatum-a.8* (*lax-a.8*) mutant was derived by fast neutron mutagenesis from cultivar 'Bonus' (Franckowiak, 2010). For genetic studies the mutation was previously introgressed into cultivar 'Bowman' to produce the nearly-isogenic line (NIL) BW457 (Druka et al., 2011). This homozygous NIL was back-crossed once more to the recurrent parent to obtain a segregating F2 population for the introgressed segment (Arnis Druka, unpublished data). The phenotype segregated in 100 F2 plants with the ratio of 78:22 (expected: 3 (WT) : 1 (MT), $\chi2$ =0.48, P-value =0.488) as expected for a single recessive gene. Plants showing the characteristic *lax-a* morphology traits: five anthers, exposed grains, and leafy awn basis, were scored as homozygous mutant plants and were clearly distinguishable from wild-type (Figure 9).

Figure 9: *Laxatum-a* phenotype.
a) The *laxatum* phenotype (*lax-a*) is characterized by an increased spike length caused by 15% longer rachis internodes compared to wild-type (see also Figure 10). Furthermore, *lax-a* mutants showed exposed caryopses (b) because of reduced size of lemma and palea (c). Awns showed a wide (leafy) base (e) and florets contained five anthers (e), two of them developed from transformed lodicules (d). Dissected florets were shown from front and back respectively and transformed organs were highlighted by white arrows.

Rachis internode lengths were determined for three spikes per plant by dividing total spike length by number of rachis nodes (number of grains). Rachis internodes of mutant spikes were on average 15 % longer than wild-type spikes (Figure 10).
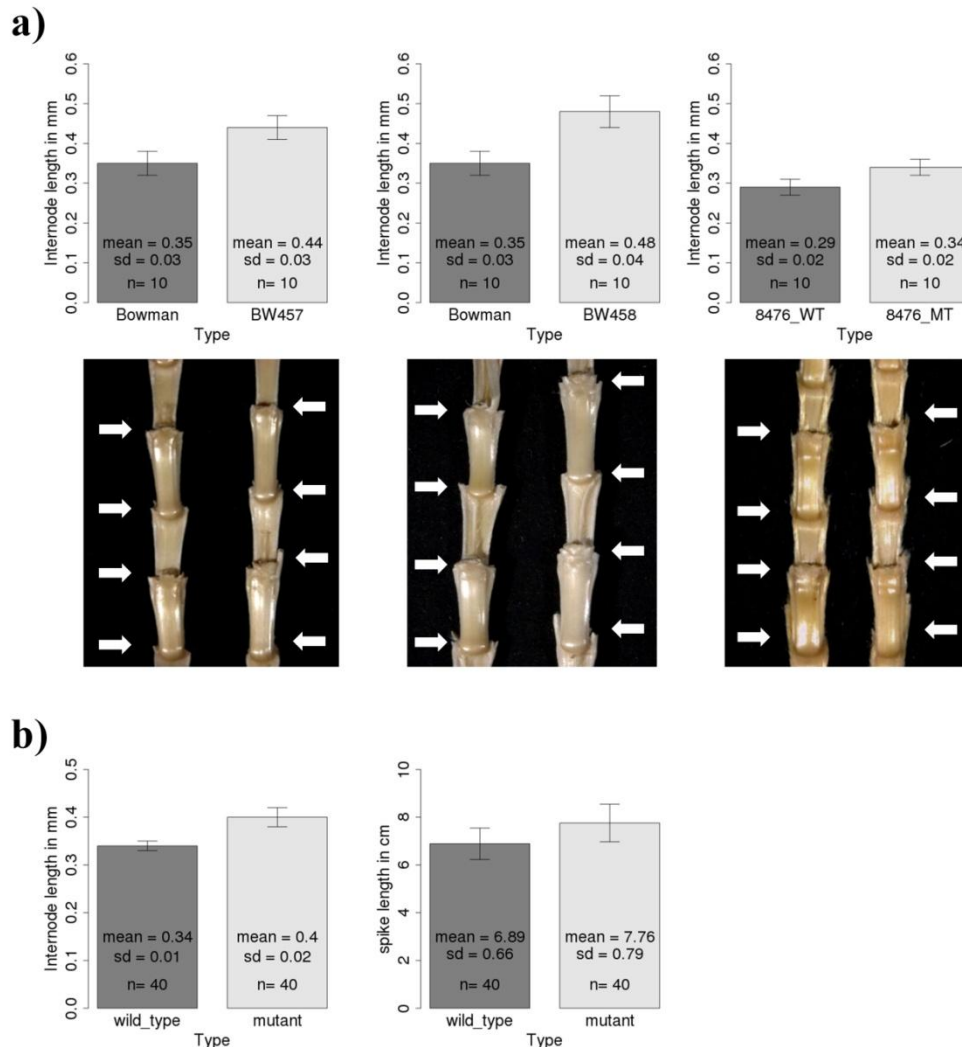
Figure 10: Rachis internode length in *lax-a* mutant and wild type plants.
a) The average rachis internode length for Bowman spikes compared to BW457 and BW458, respectively, as well as for the TILLING mutant 8476 with wild type allele against homozygote mutant plants (upper panel) revealed extended rachis internode lengths for all mutants. The lower panel shows a detailed view from part of the rachis for each wild type (left) and mutant (right), respectively, presented in the upper panel a). Rachis nodes are indicated by white arrows. b) 40 wild-type plants and 40 mutant plants were taken from of the segregating mapping population of the cross between Bowman and BW457 to confirm co-segregating of the trait within the population. While the rachis internode length showed a significant difference the length of the spikes are not significantly different between wild type and mutant plants.

3.2.2. High resolution mapping of *HvLAX-A*

A single introgressed segment of 38.5 cM spanning the centromere of 5H and harboring the *LAX-A* locus was previously detected by genome wide SNP genotyping of the NIL BW457 allele (Druka et al., 2011) (Figure 11a). Polymorphic SNPs of the target interval were utilized to develop PCR based markers. In total ten Cleaved Amplified Polymorphic Sequences (CAPS) markers were developed and genotyped in 100 F2 plants (Table A16). The *lax-a* phenotype co-segregated in a 1.5 cM interval with a cluster of four markers in the genetic centromere of chromosome 5H. These markers were located on both opposite chromosome arms of 5H according to sorted chromosome arm sequencing data (Mayer et al., 2011) (Figure 11b). To increase the genetic resolution, eight SNP loci of the ten initially developed CAPS markers, were combined in a multiplex SNP genotyping assay (SNaPshot technology) to screen an enlarged population of 1,970 F2 plants. The cluster of centromeric markers was still co-segregating in this enlarged population. Two newly identified recombination events between phenotype and the cluster of markers placed the *LAX-A* locus to the long arm of 5H (Figure 11c). A set of about 49,000 high-confidence single nucleotide polymorphisms (SNPs) was already available at the beginning of the study (Stein et al., unpublished data). It was obtained by whole genome shotgun sequencing of genomic DNA of NIL BW475 to about eight-fold haploid genome coverage. Sequence data was mapped to a WGS assembly of cv. Bowman and sequence polymorphisms were automatically data-mined. The resource was used to saturate the genetic map by three additional markers (Table A16). Eventually, the gene locus could be delimited to a 0.2 cM interval on the long arm of chromosome 5H (Figure 11d).

The aim of genetic mapping within the extended population was to reach a sufficiently high genetic resolution which would allow for delimiting the *lax-a* gene in the physical map of barley. Thus, the assigned position of each marker in the physical map of barley was considered to proof the possibility to narrow a physical interval of overlapping FPcontigs (Table 10). The closest markers as well as the cluster of markers co-segregating with *lax-a* could be allocated to the same genetic bin at 42.23 cM of the genetical/physical map of barley, to which 207 physical contigs with a cumulative length of ~176 MBp in size were assigned (IBSC, 2012). Due to a lack of genetic resolution within this region of the physical map, all of these FPcontigs were presented in the physical map at random order (IBSC, 2012). Thus, no neighboring FPcontigs could be selected based on the available map, and still chromosome walking over an unknown distance would be required. To get an impression of

the genes present in the genetic 0.2 cM interval, the virtual gene order map of barley (Mayer et al. 2011) was consulted. This map was constructed of WGS sequencing data of sorted chromosomes of barley. The sequencing data was arranged in an order by syntenic relationships between the complete sequenced grass genomes of *Brachypodium*, rice and sorghum (Mayer et al., 2011). Around 200 potential gene models were assigned to the defined target interval between marker 2_0524 and HvBradi4g43680.

Table 10: Positions of markers used for genetic mapping of HvLAX-A within the physical map of barley

Marker ID	barley high confidence gene	Bowman WGS contig ID	Genetic position [cM]	physical map contig	Physical position [Mbp]
1_0974	AK248898.1	contig_201113	17.92	FPcontig_1748	11244320
1_0580	MLOC_68535.2	contig_844634	31.25	n.a.	19657720
1_1198	MLOC_3013.1	contig_12040	43.95	n.a.	70569640
1_0157	AK251398.1	contig_884009	42.23	n.a.	153040800
1_1469	AK358200	contig_895954	42.23	FPcontig_44251	199813600
1_0481	MLOC_22626.1	contig_143316	42.23	n.a.	253551480
2_0524	AK356301	contig_221162	42.23	n.a.	248503440
HvLAX-A	MLOC_61451.6	contig_68343*	42.23	FPcontig_2862	203930400
Bd4g43680	MLOC_73193.3	contig_68744	42.23	FPcontig_3522	259218280
1_1260	MLOC_4554.1	contig_881005	44.51	n.a.	283803760
2_0239	MLOC_63089.10	contig_10702	42.23	n.a.	264640800
2_1148	MLOC_64817.1	contig_1786585**	45.9	n.a.	298526280
2_0713	MLOC_68788.2	contig_268627	51.45	n.a.	386809320

* contig was anchored by PCR based BAC-library screening; **Barke WGS; n.a. = not anchored to physical map
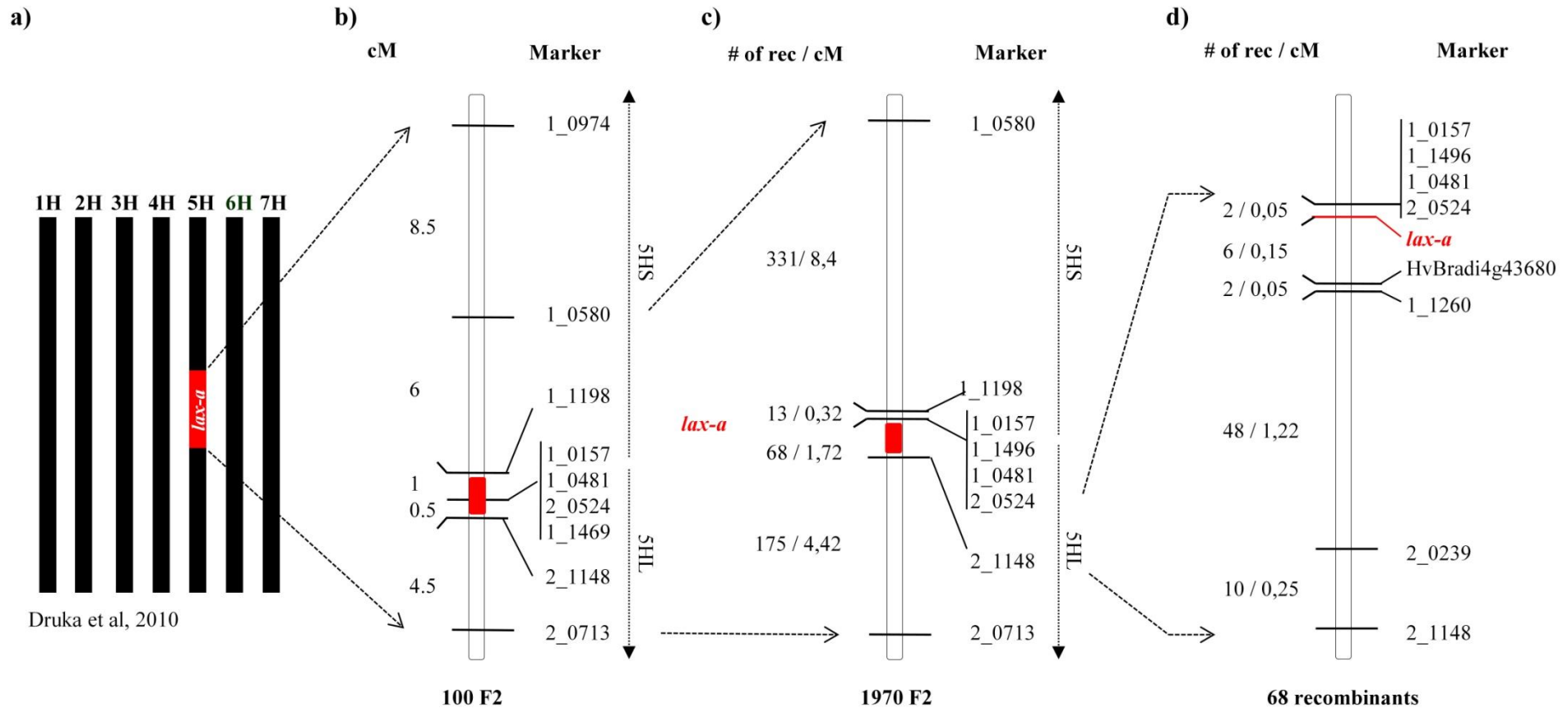
Figure 11: Genetic mapping of *lax-a*.

a) Schematic drawing of the previously localized 38,5 cM introgressed segment (red box) of the original *lax-a.8* mutant on chromosome 5H in the genetic background of Bowman (Druka et al., 2011). b) Low resolution mapping in 100 F2 plants narrowed 1.5 cM interval around the *lax-a* locus. The vertical arrows indicating the chromosomal assignment of markers based on sorted chromosome arm sequencing data (Mayer et al., 2011) c) High resolution mapping in 1,970 F2 plants allocated the gene to a 1.87 cM interval on long arm of 5H. d) The interval around the gene locus was narrowed by further markers to a 0.2 cM interval.

3.2.3. Candidate gene identification by high throughput sequencing of recombinant plants

The genetic resolution was not sufficient to narrow down a physical interval for the target gene. A step wise narrowing of the interval by placing further markers was complicated by the non linear arrangement of the physical map in this region. The stepwise identification of neighboring BAC contigs by chromosome walking is principally possible; however, given the genomic region it was considered a very laborious and risky task. The 200 gene models assigned to the target region based on synteny indicated a possibly large physical distance to be analyzed. Thus, an NGS based approach was aimed to explore the remaining interval by mapping and candidate gene identification in a single experiment. For this, the eight most tightly linked recombinants defining the 0.2 cM target interval of the *LAX-A* locus after high resolution mapping were selected for sequencing (Figure 11d, Figure 12a). DNA of recombinants with the same marker haplotype and plant phenotype was multiplexed leading to small pools of two to three genotypes, respectively (Figure 12a). In three cases a given marker haplotype / phenotype combination was represented by single recombinants only, which were still called 'pools' in the following for simplicity of terminology. Each DNA pool was tagged with sequencing adaptors carrying unique sequence indexes (barcode) and all five samples were combined in a single liquid phase barley whole exome capture assay hybridization (Mascher et al., 2013c) reaction. Paired end sequencing (2x100 bp) was performed on a single lane of the HiSeq2000 instrument and yielded ~200 mio total read pairs after removing duplicates. Sequence reads were assigned to individual pools on the basis of the respective unique barcode sequence tags. The number of read pairs per pool differed between ~32 and ~56 Mio per pool. The sequence reads were mapped against the WGS assembly of cv. Bowman (IBSC, 2012), followed by variant calling to identify SNPs.

For each SNP position, the ratio of reads carrying the alternative allele (mutant) versus reads with the reference allele was calculated through all of the individual pools. Obtained SNP frequencies were plotted for each sample against the reference sequence. Clear shifts in SNP frequencies could be observed for all five samples consistent with the predicted marker haplotypes from genetic mapping (Figure 12b). Since the physical map in this region was not arranged in a robust linear order, a graphical delimitation of a target interval was not possible to detect causal SNPs in targets. To overcome this problem, the raw SNP frequency data was queried for targets of interest. These are SNPs that co-segregate with the phenotype in all five
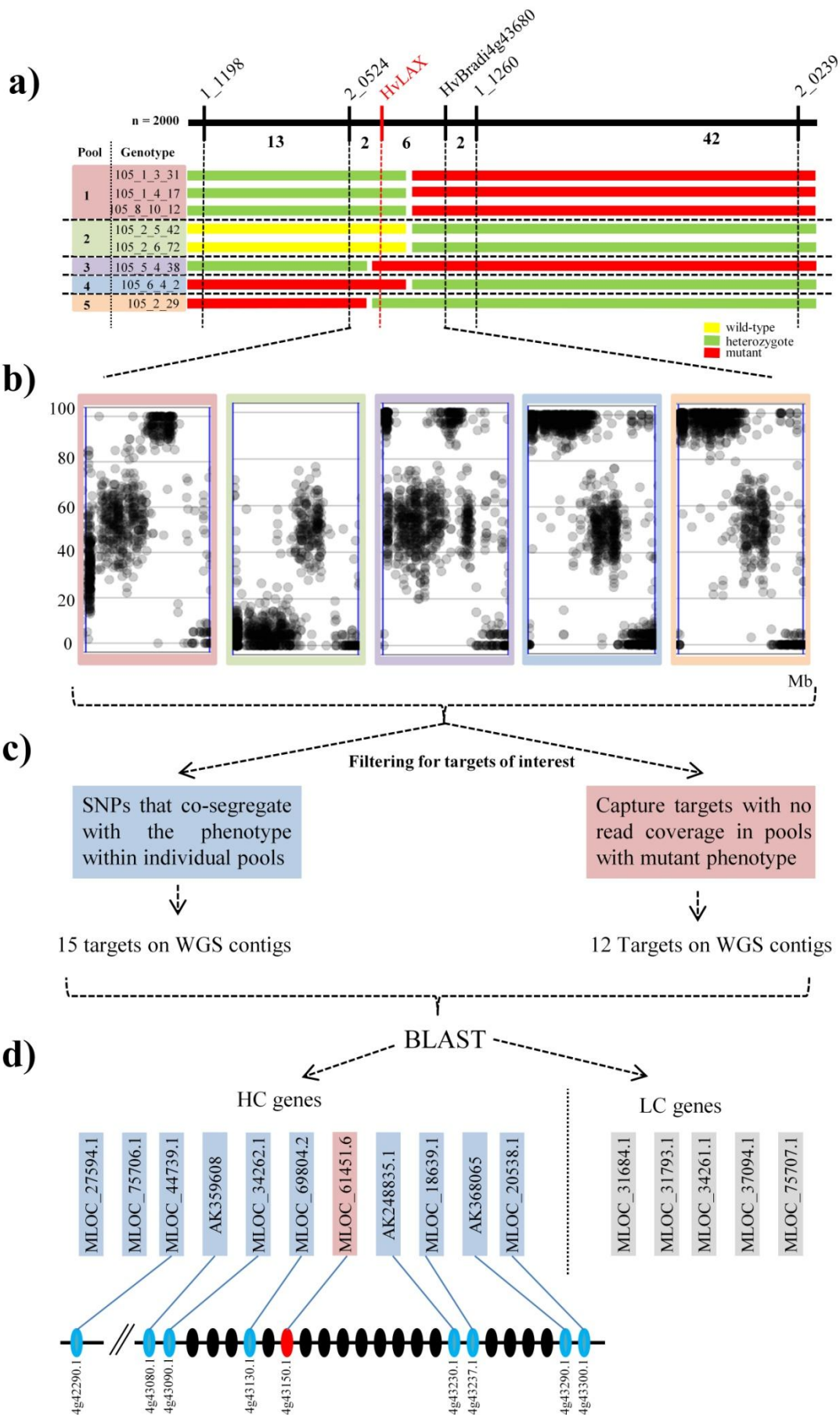
Figure legend on the following page

Figure 12: Strategy for *LAX-candidate* gene identification.
A sequence assisted evaluation of a predefined mapping interval was used to indentify candidate genes underlying the *lax-a* phenotype. a) Eight recombinant plants delimited a 0.2 cM mapping interval on 5H for *lax-a* after genotyping 1,970 F2 plants. Marker score and phenotype score for each genotype is represented by a color code for simplification (yellow = wild type, green = heterozygote, red = mutant). Genotypes with identical marker haplotype / phenotype combination were pooled together before continuing with exome capture (Mascher et al., 2013c) followed by sequencing. b) Obtained sequence reads of the individual pools were mapped to the reference of Bowman (IBSC, 2012) for SNP discovery and determining SNP frequencies. SNP frequency plots restricted to chromosome 5H for each individual pool. The X-axis shows the physical expansion of 5H (IBSC, 2012), the y-axis SNP frequencies from 0 to 100 % and SNPs are visualized as dots. c) Filtering for candidate targets delivered 27 targets on WGS sequence contigs. d) Identified genes on WGS contigs were used for homology analysis and revealed that most of the genes cluster in a small syntenic interval of Brachypodium.

pools. The SNP frequencies can be utilized to score variants as co-dominant marker. In principle, a SNP with a frequency of 100 % reflects homozygote mutant genotype score (*lax-a/lax-a*), 50 % represent a heterozygote genotype call (*LAX-A/lax-a*) and wild-type for 0 % (*LAX-A/LAX-A*). Therefore, SNPs that occur at a frequency of 50 % for pool one, 0 % for pool two, 100 % for pool three and four and 50 % for pool five were selected. These targets with SNPs should be either the causal gene itself or targets in direct proximity. The accurate SNP calling requires a sufficient read coverage. Not all targets are captured with the same efficiency. Additional fraction of wrongly mapped reads can be observed and could complicate the frequency determination. Therefore, an accuracy threshold for SNP frequencies was introduced to compensate erroneously induced changes in SNP frequencies. Frequencies larger than 80 % were called as mutant genotype, below 20 % were called as wild-type genotype and everything in between (20-80 %) as heterozygote genotype. Fifteen targets with SNPs mapped to the WGS contigs of Bowman were identified which showed the requested criteria of co-segregation with the phenotype (Table 11). In total, ten high-confidence genes and five low-confidence genes were identified by sequence comparison of the identified Bowman WGS contigs against the barley gene set (IBSC, 2012).

The *lax-a* mutation was induced by fast neutrons. It was shown that fast neutrons cause predominantly large deletions up to 12 kb (Li et al., 2001). Therefore, it was assumed that the mutation in NIL BW457 may have caused partial or complete deletion of the gene. The exome capture assay is designed to capture exon sequences, however, only ~74 % of the annotated barley genes are represented in the capture target (Mascher et al., 2013c). If the *lax-*

a gene is part of the capture target it can be expected that this part of the target should have zero read coverage in the sequencing data of all pools with mutant phenotype. All targets were filtered for average read coverage of less than two-fold for pool three and four (mutant phenotype) and more than five-fold coverage in all other pools (wild-type phenotype) (Figure 12a). The relaxed filtering criteria by allowing up to 2-fold read coverage in mutant pools was implemented to reduce the risk of missing important targets because a certain percentage of falsely mapped reads to short repetitive parts of the reference sequence or conserved sequences within gene families can be expected. Furthermore, false-assignments of sequences to samples by misidentification of the respective index after multiplex sequencing of samples is a known contamination source of Illumina sequencing (Kircher et al., 2012). Twelve putatively deleted capture targets mapped on Bowman WGS contigs were obtained from read depth analysis (Table 12). Only a single putatively deleted gene (MLOC_61451.6), present on bowman_contig_68383 was identified among the twelve putative targets, by sequence comparison (BLAST) to the barley gene sets (Figure 12d). Details for SNP frequency distribution and coverage values within the five pools are summarized in Table A17 and Table A18.

The genetic resolution was not sufficient enough to delimit an interval in the physical map of barley for the identified candidate genes. The known highly conserved synteny of genes on barley chromosome 5H and Brachypodium chromosome Bd4 was consulted to explore their relationship in a physical context (Mayer et al., 2011; IBSC, 2012). Homologous genes for the identified candidate genes were predicted by BLAST to Brachypodium gene models. The majority of the identified Brachypodium genes were located in a small collinear interval. Interestingly, the homolog of the identified putative deleted gene was located in the center of the interval (Fig 12d). Within the defined interval in Brachypodium, additional gene models were allocated. The Brachypodium gene models were used to search for homologous genes in barley to proof whether additional genes were affected by deletion at the *lax-a* locus. All of the identified barley genes were successfully captured and none of them showed a reduced coverage in the mutant compared to the wild-type pools (Table 13).

The reduced complexity resequencing of the closest recombinants delivered a set of cosegregating candidate genes. Considering their homologous Brachypodium genes, a small interval could be identified by the help of the advanced genome reference available for Brachypodium. Based on the synteny information, only a single deleted gene remained as most prominent candidate gene.

Table 11: Exome capture targets with SNPs which cosegregate with *HvLAX-A* phenotype

Capture targets mapped to WGS contigs of Bowman		Gene information for WGS contigs		
WGS_contig[1]	length	HC_genes[1]	LC_genes[1]	Brachypodium*
contig_166251	1241bp	/	MLOC_31793.1	/
contig_230144	763 bp	/	/	/
contig_245492	677 bp	/	/	/
contig_582766	205 bp	/	/	/
contig_876951	4598 bp	/	/	/
contig_859419	3668 bp		MLOC_37094.1	Bradi4g43080.1
contig_921063	2228 bp	/	/	/
contig_13430	9835 bp	MLOC_20538.1	/	Bradi4g43300.1
		AK368065	/	Bradi4g43290.1
contig_108294	3369 bp	MLOC_27594.1	MLOC_31684.1	/
contig_221364	13089 bp	MLOC_69804.2	/	Bradi4g43130
contig_366956	4997 bp	MLOC_34262.1	MLOC_34261.1	Bradi4g43090.1
contig_871803	3230 bp	AK248835.1	/	Bradi4g43230.1
contig_879642	4375 bp	MLOC_18639.1	/	Bradi4g43237.1
contig_1989744	3111 bp	MLOC_44739.1	/	Bradi4g42290.1
contig_1998524	4042 bp	MLOC_75706.1	MLOC_75707.1	no homolog

1(IBSC, 2012), * Homologous Brachypodium gene models (v1.2) predicted by BLAST

Table 12: Targets with low coverage in captured pools with mutant phenotype

Capture targets mapped to WGS contigs of Bowman		Gene information for WGS contigs		
WGS_contig[1]	length	HC_genes[1]	LC_genes[1]	Brachypodium*
contig_1387056	882 bp	/	/	/
contig_1498435	470 bp	/	/	/
contig_1532510	797 bp	/	/	/
contig_1534971	630 bp	/	/	/
contig_1764579	240 bp	/	/	/
contig_212370	1415 bp	/	/	/
contig_335772	224 bp	/	/	/
contig_350085	316 bp	/	/	/
contig_380230	686 bp	/	/	/
contig_523676	220 bp	/	/	/
contig_872010	5225 bp	/	/	/
contig_68343	10094 bp	MLOC_61451.6	/	Bradi4g43150.1

1(IBSC, 2012), * Homologous Brachypodium gene models (v.1.2) predicted by BLAST

Table 13: Syntenic block in Brachypodium defined by sequence homology of identified candidate capture targets from mapping-by-sequencing of HvLAX-A

Brachypodium[1]	HC_genes[2]	Annotation[2]	Bowman WGS contig[2]	lax_1	lax_2	lax_3	lax_4	lax_5
Bradi4g43080*	AK359608	Eukaryotic rpb5 RNA polymerase subunit	contig_129504	15.37	20.58	18.41	29.37	17.57
Bradi4g43090*	MLOC_34262.1	Response regulator 6	contig_366956	20.28	19.58	20.52	30.54	22.96
Bradi4g43100	/	/	/	/	/	/	/	/
Bradi4g43110	MLOC_10658.1	Cytochrome P450	contig_859290	9.06	10.05	9.13	17.98	11.48
Bradi4g43117	/	/	/	/	/	/	/	/
Bradi4g43130*	MLOC_69804.2	2-isopropylmalate synthase B	contig_221364	5.82	6.42	6.96	11.08	8.70
Bradi4g43137	AK373675	Strictosidine synthase family protein	contig_1990163	11.49	13.67	13.14	19.75	15.57
	AK373675		contig_862830	7.73	9.55	10.10	15.18	9.47
Bradi4g43150**	MLOC_61451.6	NPR1 protein	contig_68343	4.66	12.17	0.42	1.02	6.89
Bradi4g43160	/	/	/	/	/	/	/	/
Bradi4g43170	MLOC_34640.1	70 kDa heat shock protein	contig_92989	18.23	17.11	19.48	29.21	23.89
Bradi4g43180	MLOC_6357.1	Glycosyltransferase	contig_846283	11.46	13.54	12.03	25.38	15.09
Bradi4g43190	/	/	contig_92989	18.23	17.11	19.48	29.21	23.89
Bradi4g43200	MLOC_7637.1	GRAS family transcription factor	contig_189650	3.72	5.42	7.21	10.18	7.76
	MLOC_7637.1		contig_187504	4.05	2.91	4.29	5.84	3.72
Bradi4g43210	MLOC_65272.1	Calmodulin-like protein 1	contig_81200	5.63	7.88	7.23	12.76	6.99
Bradi4g43220	AK358874	FAR1-related sequence 7	contig_328569	15.87	17.87	17.37	26.22	19.74
	AK358874		contig_34063	14.86	14.25	8.96	24.83	12.26
Bradi4g43230*	AK248835.1	PP2A regulatory subunit TAP46	contig_871803	14.28	15.98	15.20	25.62	17.76
	AK248835.1		contig_940211	6.60	6.72	9.30	13.18	9.30
Bradi4g43237*	MLOC_18639.1	unknown protein	contig_879642	10.86	9.86	11.46	19.25	12.53
Bradi4g43250	AK364870	Cytochrome c biogenesis protein	contig_13687	11.45	10.89	11.81	17.48	14.10
Bradi4g43257	MLOC_53116.5	Saccharopine dehydrogenase family protein	contig_844139	19.39	23.95	23.39	32.75	25.43
Bradi4g43270	MLOC_53117.1	NC domain-containing protein	contig_844139	/	/	/	/	/
Bradi4g43280	MLOC_53116.5	Saccharopine dehydrogenase family protein	contig_844139	/	/	/	/	/
Bradi4g43290*	AK368065	Calpain-B	contig_13430	34.59	36.05	38.76	53.48	42.91
Bradi4g43300*	MLOC_20538.1	Calcium-transporting ATPase 1	contig_13430	/	/	/	/	/

[1] Brachypodium gene models of the syntenic block; [2] High confidence gene models with predicted annotation (IBSC, 2012)

* Filtered targets by SNP frequencies request; ** Filtered targets for coverage change; green: targets with >5x coverage, red: targets <2x coverage according filtering criteria's

3.2.4. Physical anchoring of the *LAX-A* candidate gene

The read coverage analyses revealed a single gene which was not covered by sequence reads in pools with mutant phenotype. Thus, the gene was putatively deleted within the mutant NIL BW457. The gene deletion was confirmed by PCR amplification in DNA of wild-type parent Bowman and mutant NIL BW457 using multiple fragments covering the complete gene from translation start to translation stop, including introns (HvLAX_F/R1, HvLAX_F/R2, HvLAX_F/R3, HvLAX_F/R4, Table A5). The size of the deletion comprised at least the entire gene, respectively the area sampled by the employed primer pairs. To explore the entire size of deletion it was aimed to increase the sequence information around the candidate gene by anchoring the gene to a physical contig. The gene itself was not assigned to the available physical map of barley (IBSC, 2012). Thus, a PCR-based screening was performed on multidimensional BAC pools which led to the identification of four BAC clones (HVVMRXALLeA0046E04, HVVMRXALLeA0122D17, HVVMRXALLeA0209O12, HVVMRXALLeA0379H14) which were all belonging to the same FPcontig_2862 with an estimated total length of approximately 2.2 Mbp.

A minimum number of overlapping clones, spanning the complete BAC contig, were shotgun sequenced and assembled. Sequence overlaps were confirmed by all against all sequence comparisons (BLASTN, identity >99% and minimum hit length of 2000 bp) to generate a contiguous sequence scaffold (Figure 13a, Table A19). The sequencing resulted in a total cumulative non redundant sequence length of 2.3 Mbp.

The mutant NIL BW457 was sequenced to an average eight-fold genome coverage. The recurrent genotype Bowman was previously sequenced to about 34-fold genome coverage in frame of the development of the sequence enriched barley physical map (IBSC, 2012). These sequence resources were used to indentify parts of the genome that are under-represented in the WGS data of the mutant compared to wild-type Bowman. The fact that the sequenced BAC clones were generated from cultivar Morex required for a less stringent read mapping which allowed a small percentage of mismatches. Since the obtained read lengths of NGS sequencing are only 100 bp in maximum, an overrepresentation of highly similar short sequence reads can be expected by mapping an entire genome against a short BAC contig, which would most likely lead to distorted read coverage. Therefore, the available 1.8 Gbp sequence assembly of the parental genotype Bowman (IBSC, 2012) was used as reference for mapping. The average read coverage for each Bowman WGS contig was calculated for the

respective mapping of mutant and wild-type reads. Sixty-four sequence contigs of this WGS assembly of cv. Bowman, representing a cumulative length of 211 kbp, were assigned along the sequenced FPcontig by BLAST. Seventeen of these WGS contigs, with a cumulative length of 53 kb, showed no sequence coverage by reads from the WGS assembly of the mutant NIL BW457 but the expected coverage of ~30-fold from the wild-type Bowman reads. All of these contigs were allocated to the center of the sequenced FPcontig_2862 and allowed a prediction of a deletion size between 436 and 554 kb (min/max) (Figure 13c).

The assembled BAC sequence contigs were used for gene identification. Before, the sequences were analyzed by k-mer statistics in order to mask the present repetitive elements. Low copy sequences were used to search for genes in the barley gene set by BLAST (IBSC, 2012). In total four HC_genes (Table 14) and two LC_genes were present on the FPcontig_2862 (Figure 13b). Only the identified candidate gene locus MLOC_61451.6 was allocated to the deletion (Figure 13b). The barley LC and HC gene sets were predicted based on an incomplete WGS shotgun assembly (IBSC, 2012), thus are likely to be incomplete and not fully representative. To minimize the risk of overlooking additional putative barley genes present on the physical contigs, two additional strategies for gene identification were applied: (i) First, available sequence resources of the sequenced genomes of the grass species Brachypodium, rice and Sorghum were included in the sequence comparison analysis and showed no indication for additional genes. (ii) Second, a *de novo* gene prediction (Stanke et al., 2004) was applied to identify putative open reading frames (ORF) of unknown genes. While the sequenced grass genomes confirmed complete synteny to the four predicted barley genes, the *de novo* gene prediction revealed 27 putative ORFs within the deletion. None of these ORFs showed any sequence homology to genes or conserved domains of other plants or human genomes. They are rather short and single exon ORFs. Thus, they might represent pseudo-genes.

Table 14: Annotated genes located on BAC_contig_2862:

HC_genes[1]	Annotation[1]	Brachypodium ID*
MLOC_61451.6	NPR1 protein	Bradi4g43150.1
AK373675	Strictosidine synthase family protein	Bradi4g43137.1
MLOC_69804.2	2-isopropylmalate synthase B, putative	Bradi4g43130.1
MLOC_10658.1	Cytochrome P450	Bradi4g43110.1

[1](IBSC, 2012), Homologous genes in Brachypodium predicted by sequence similarity (BLAST)
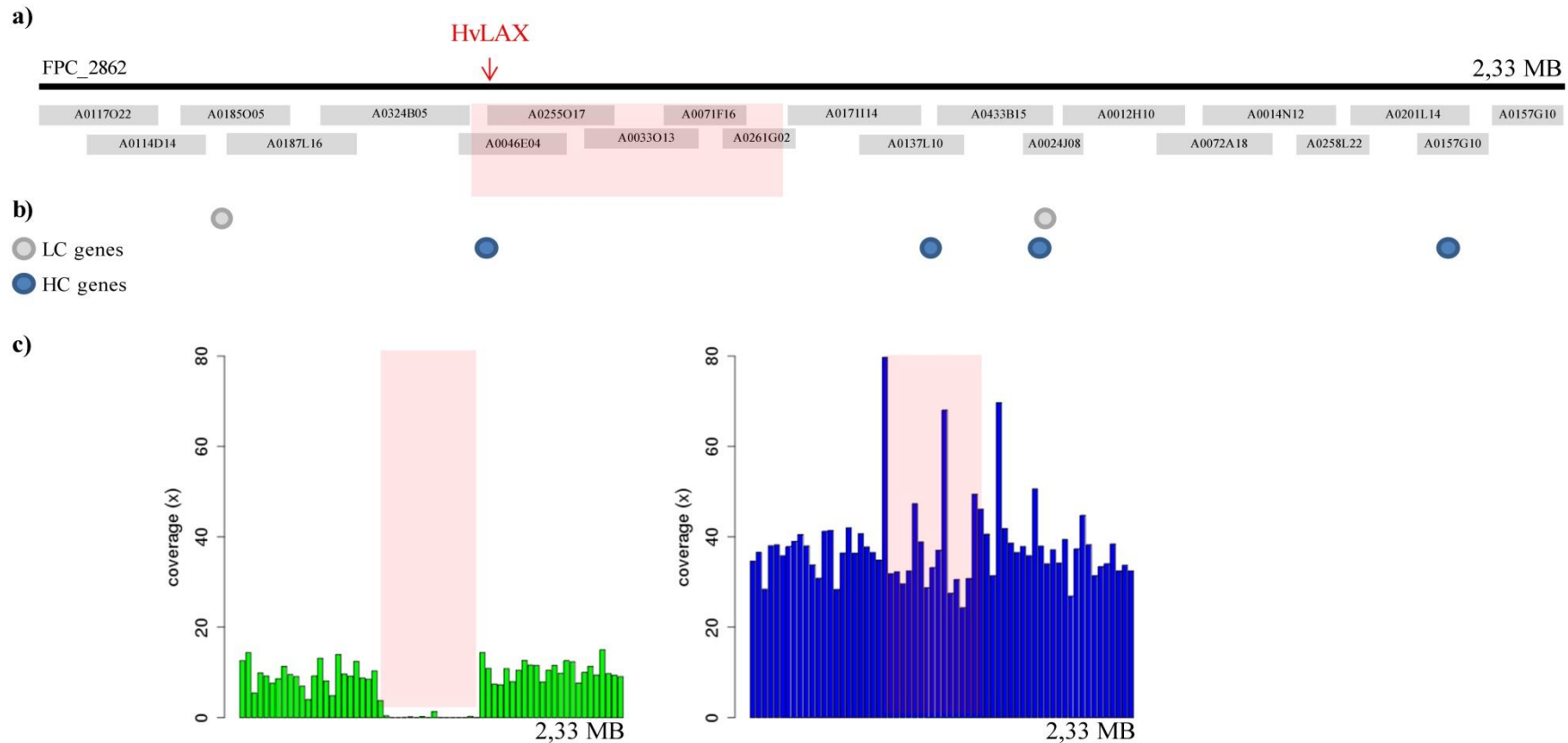
Figure 13: Sequence analysis of FPcontig_2862 containing the *HvLax-A* candidate gene
a) Sequencing a contig of overlapping BAC clones carrying the *lax-a* candidate gene revealed 2.33 Mbp cumulative sequence (Table A19). Sequences were used for gene prediction of high confidence genes (blue) and low confidence genes (grey) of the barley gene set (Table A20). c) Graphical visualization of the average read coverage of Bowman WGS contigs anchored to the BAC sequence contigs of FPcontig_2862. The x-axis represents anchored Bowman WGS contigs arranged along the FPcontig. Each bar represents one Bowman WGS contig. The mean read coverage of raw reads from WGS B457 (left) and Bowman (right) of each anchored WGS Bowman contig is given on the y-axis. In the center of the contig no read coverage was detected for the WGS data of the mutant indicating a large deletion of around 500 kb (marked in red in a and c).

3.2.5. Functional validation of *the LAX-A* candidate gene by mutant analysis

The genetic mapping resolution did not allow establishing a continuous sequence scaffold between both gene flanking markers. A large deletion within an FPcontig was indentified but it was not possible to exclude the possibility of further undetected but linked deletions. In addition, the large deletion included still additional candidate ORFs which might have an impact on expressing the *lax-a* mutant phenotype. It was aimed to analyze the effect of the candidate gene by various mutations in independent plant material. The same strategies like described for the *HvMND* gene were applied. (i) TILLING for the identification of new alleles for the candidate gene. (ii) Resequencing the candidate gene in an existing series of *lax-a* plants obtained from the gene bank.

TILLING analysis was performed on the same 7,979 EMS induced plants of c.v. Barke (Gottwald et al., 2009) like described in part 3.1.6. The TILLING analysis was done for two gene fragments spanning the entire ORF (HvLAX_E1_F/R; HvLAX_EX2_F/R, Table A5). A series of twenty synonymous and twenty-one non-synonymous SNPs was identified (Figure 15a, Table 16). No knockout mutation, e.g. truncation of splice site or introduction of a premature translation stop codon, could be identified. Thus, M3 progeny of all families with non-synonymous variations were grown to analyze for phenotypic variation. All M3 plants were genotyped in order to confirm the presence and zygosity of the detected SNP, since several mutations were found in heterozygous state in the original M2 samples (Table 16). A single mutation, C127T (L43F), revealed the typical *lax-a* phenotype with extended rachis internode lengths (Figure 10), five anthers, broadened base of the lemma awn and uncovered seeds (Figure 14 a-f). The respective M3 family was segregating for the mutation and phenotype since the M2 plant was heterozygous for the C127T mutation. Among the eight M3 plants only the two homozygous mutant plants showed the *lax-a* phenotype, which is consistent with the recessive nature of the mutated gene.

Among the Bowman near isogenic lines, one additional line (BW458) was described to exhibit the *lax-a* characteristic phenotype including the five stamen. The phenotype was confirmed by growing BW458 plants in the greenhouse (Fig. 14 g-l). Part of the first exon could not be amplified and indicated a partial gene deletion (Fig. 15b).

TILLING mutant 8476

BW458
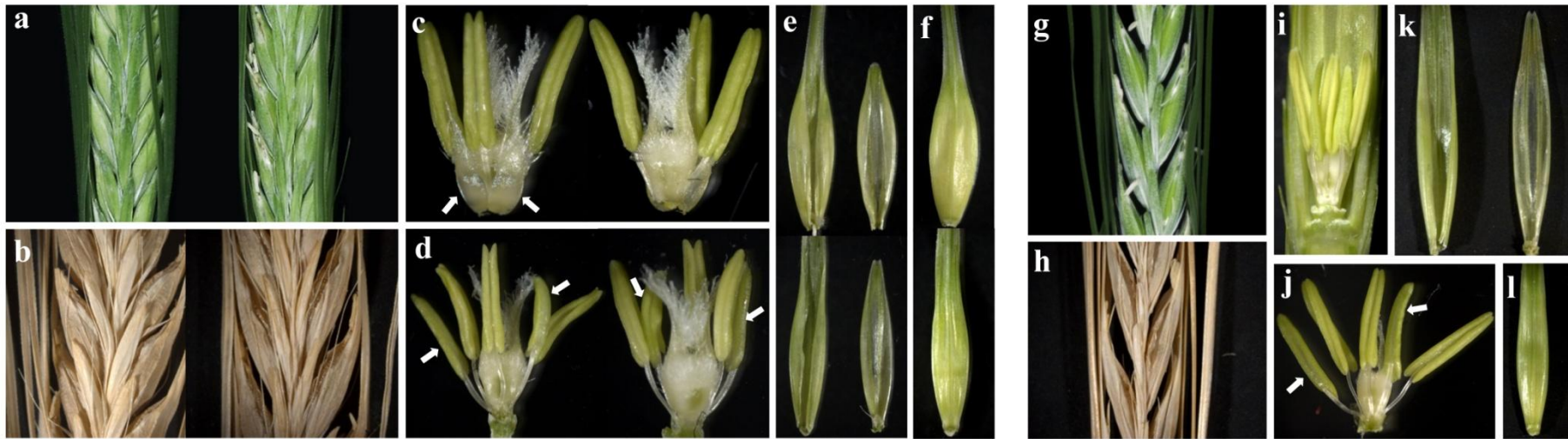
F1 of BW457xTILLING mutant 8476

F1 of BW457xBW458

*Figure legend on the following page

Figure 14: Independent *lax-a* mutant alleles and allelism test.
The figure summarizes the phenotype information for TILLING mutant 8476 (a-f) and NIL BW458 (g-l) and their respective F1 crosses to the original NIL BW457 which was used for cloning of the gene. Genotypes of TILLING mutant family 8476 with wild-type allele and homozygote mutant allele compared for *lax-a* characteristic traits: exposed caryopsis of the mutant (right) to the covered grains of the wild type (a+b), five anthers (d) instead of three for wild-type plants (c), reduced growth of mutant (lower one) lemma and palea (e) as well as the broadened awn base (f) compared to wild-type (upper one). BW458 carries exposed grains (g+h), five anthers (j), lax-a typical lemma and palea (k) and the wide awn base (i). Both crosses resulting in lax-a mutant phenotype: exposed caryopsis (m+n and r+s) five anthers (o and t), broad awn base (p,u) and changed palea and lemma growth (q,v). Arrows labeling the homeotic replaced organs in the barley florets with removed lemma and palea

The gene *HvLAX-A* was not mapped to a resolution which allowed the physical delimitation of a continuously annotated stretch of sequence information around the gene. Since the analyzed mutant alleles originated by artificial mutagenesis, which is known to introduce multiple random mutations in the genome, a further complementation test was aimed to rule out the influence of a background mutation on the phenotype. The test for allelism was performed between the mutant line used for cloning (BW457), the TILLING mutant 8476_1_1 and BW458. The test of allelism refers to a cross of two genotypes with variants in a gene of interest. If both variants are the cause of a shared recessive phenotype, hence they represent independent alleles of the same causal gene, F1 hybrids derived from a combination of both genotypes must express the mutant phenotype. In case the mutation is caused by two independent loci, due to the recessiveness of the trait, the F1 hybrid will show a wild-type phenotype. The Bowman NIL 457 was used as female parent and crossed with TILLING mutant 8476 and NIL BW458, respectively. A codominant SNP marker test for heterozygosity was performed to confirm the success of the cross. Since *HvLAX-A* was identified to be located within a large deletion, two flanking polymorphic SNPs, located on FPcontig_2862 outside of the deletion, were used for genotyping the cross of BW457 and BW458 (Table 15). Two primer combinations (Table A5), were used to amplify these polymorphic SNPs (Table 15) and all analyzed F1 plants proofed to be heterozygous for both markers. All sixteen F1 plants displayed a *lax-a* characteristic phenotype (Figure 14 r-v). For the second cross, between BW457 and the TILLING mutant 8476, only the SNPs on Bowman_contig_129575 could be utilized for co-dominant genotyping (Table 15). To generate a second independent confirmation, the first exon was amplified (HvLAX_F/R4,

Table A5) in all F1 plants as dominant marker and the obtained fragment was sequenced. All fragments carried the homozygous mutant allele (T) at position 127 bp, transferred from the male parent TILLING Mutant 8476. In total six F1 plants were analyzed for the cross between BW457 and TILLING mutant 8476. All were heterozygous and showed the laxatum-characteristic phenotype (Figure 14 m-q).
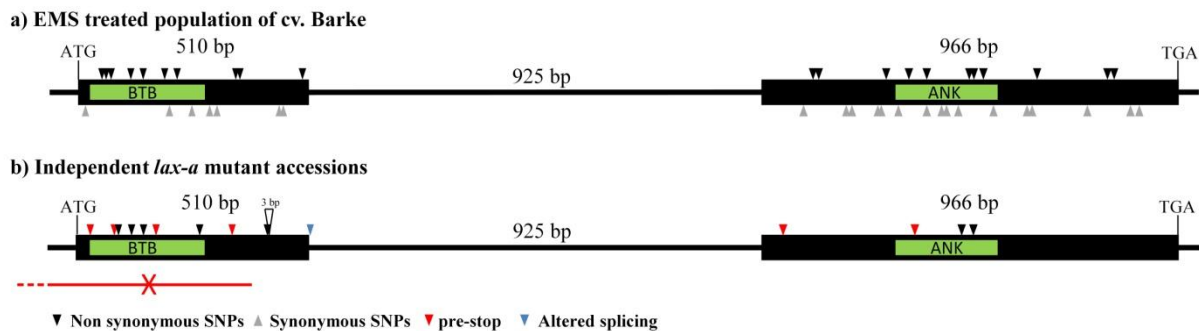
Figure 15: Analysis and characterization of induced mutations in the *LAX-A* candidate gene.
A gene model of *HvLAX*, consisting of two exons (black boxes) spaced by a single intron (black line), is visualized. Green boxes indicate conserved domains of the protein. Distribution of mutant alleles (triangles color coded according functionality) along the *HvLAX-A* gene model for TILLING analysis (a) and *lax-mutant* accessions obtained from gene bank (NordGen, http://www.nordgen.org/). The red line indicates the partial deletion present in Bowman NIL BW458 (b).

The original mutant *lax-a.8*, which was used in this study for gene cloning, was a member of a large collection of morphological and physiological mutants originated by large scale mutagenesis experiments in the mid 20th century in Sweden (Lundquist, 2009). Among this collections, 28 *lax-a* mutant accessions were described and seeds were obtained from Nordic Genetic Resource Center (NordGen, http://www.nordgen.org) to resequence the candidate gene. The available set of 28 accessions was induced by various radiation types and chemical treatments in genotypes 'Bonus', 'Foma' and 'Kristina'. Thus, a number of further mutant alleles was expected. In thirteen accessions the candidate gene could not be amplified which indicated a complete deletion of the gene. Fourteen accessions carried point mutations of which seven led to a pre-mature stop codon, six to non-synonymous amino acid substitutions and one resulted in an altered splicing site (Figure 15b, Table 17). One mutant accession

carried two mutations: An amino acid substitution and a three base-pair deletion leading to a frame shift and causing a premature translation stop codon. Based on this analysis it can be concluded that mutations in the gene MLOC_61451.6 are causal for the expression of the *lax-a* phenotype in barley. The gene name *HvLAX-A* will be used instead of the gene locus identifier MLOC_64838.2.

Table 15: SNP marker used for F1 test

Contig	length	SNP position	Parent	Genotype
Bowman_contig_71318	7031 bp	C5024T	BW458	C
HVVMRXALLeA0185O05*			BW457	T
~350 kb			8476_1_1	/
Bowman_contig_129575	6517 bp	G3377A	BW458	G
HVVMRX83KHA0157G10 *			8476	G
~2275 kb			BW457	A
		T3735G	BW457	T
			BW458	G
			8476	G

* Corresponding BAC, ** position within the FPcontig_2862

Table 16: Identified TILLING mutants within *HvLax-A*

Plant	SNP	Status M3[1]	effect	phenotype
6574_1	C28T	HOM	L10=	/
4164_1	A73C	HOM	S25R	wild-type
14134_1	G89A	HET	S30N	wild-type
10389_1	G94A	HET	E32K	wild-type
8476_1	C127T	HET	L43F	mutant
2905_1	G152A	HET	R51K	wild-type
3105_1	C202T	HET	L68F	wild-type
10899_1	C228T	HOM	G76=	/
4782_1	G256A	HET	G86S	wild-type
14423_1	C309T	HET	I103=	/
12687_1	G378A	HET	K126=	/
9017_1	G387A	HOM	P129=	/
12029_1	G436A	HET	A146T	wild-type
9605_1	G439A	HOM	V147I	wild-type
4125_1	C453A	HET	L151=	/
9122_1	C465T	HOM	A155=	/
10251_1	G505A	HET	V169I	no seeds
4607_1	G1636T	HET	S237=	/
3374_1	C1697T	HOM	P258S	wild-type
15929_1	G1710A	HET	G262E	no seeds
16023_1	G1714A	HET	P263=	/
10074_1	G1741A	HET	K272=	/
3428_1	C1744T	HET	I273=	/
11801_1	C1750T	HET	R275=	/
9346_1	C1773T	HOM	A283V	wild-type
6661_1	C1813T	HOM	G296=	/
6201_1	C1832T	HET	L303F	wild-type
11039_1	C1906T	HET	D327=	/
2573_1	C1916T	HET	R331C	wild-type
10884_1	G1936A	HOM	K337=	/
12210_1	G1936A	HOM	K337=	/
7969_1	G1960A	HET	E345=	/
10241_1	C1991T	HET	L356F	wild-type
4094_1	A1998T	HET	H358L	wild-type
6679_1	C2024T	HET	L367F	no seeds
6835_1	G2056A	HOM	R377=	/
10920_1	C2107T	HET	H394=	/
15189_1	C2110T	HET	I395=	/
11578_1	C2114T	HET	P397S	wild-type
3265_1	T2206A	HET	A427=	/
3265_1	G2207T	HET	G428W	wild-type
3859_1	G2238A	HOM	G438D	wild-type
9661_1	G2258A	HET	V445M	/
7969_1	C2268T	HET	T448I	/

[1] HOM = homozygote, HET = heterozygote

Table 17: Mutant alleles of *HvLAX-A* in independent *lax-a* accessions

NGB Nr.	Mutant	Mutation	original cultivar	Mutagen
116334	*lax-a.01*	complete deletion	Bonus	gamma-rays
116338	*lax-a.4*	C1983T (S348F)	Bonus	Cumarine
116342	*lax-a.8*	complete deletion	Bonus	Neutrons
116354	*lax-a.20*	complete deletion	Bonus	Neutrons
116372	*lax-a.37*	complete deletion	Bonus	Neutrons
116374	*lax-a.39*	complete deletion	Bonus	X-rays
116388	*lax-a.54*	T149C (F50S)	Bonus	ethylene imine
116426	*lax-a.92*	complete deletion	Bonus	X-rays
116436	*lax-a.208*	complete deletion	Foma	X-rays
116446	*lax-a.218*	A347T (Y116F)	Foma	ethylene imine
116450	*lax-a.222*	T125A (V42E)	Foma	ethylene imine
116458	*lax-a.229*	C29A (S9*)	Foma	Glyidol
116483	*lax-a.256*	A1505T (K194*)	Foma	iso-propyl methanesulfonate
116503	*lax-a.278*	C127T (L43F)	Foma	ethyl methanesulfonate (EMS)
116510	*lax-a.286*	C1853T (Q305*)	Foma	ethyl methanesulfonate (EMS)
116560	*lax-a.353*	complete deletion	Kristina	Neutrons
116579	*lax-a.369*	T512A (altered splicing)	Kristina	ethylene imine
116608	*lax-a.398*	T461G (L154R) 3-bp deletion 469 to 471 bp (A157*)	Kristina	gamma-rays
116613	*lax-a.405*	complete deletion	Kristina	X-rays
116614	*lax-a.406*	complete deletion	Kristina	Neutrons
116622	*lax-a.413*	T123A (C41*)	Bonus	ethyl methanesulfonate (EMS)
116647	*lax-a.434*	complete deletion	Bonus	Neutrons
116650	*lax-a.437*	T1992A (L331H)	Bonus	iso-propyl methanesulfonate
116659	*lax-a.444*	complete deletion	Bonus	Neutrons
116664	*lax-a.448*	C184T (Q62*)	Bonus	ethyl methanesulfonate (EMS)
116668	*lax-a.450*	G417A (W139*)	Bonus	sodium azide
116675	*lax-a.455*	G417A (W139*)	Bonus	sodium azide
116695	*lax-a.472*	complete deletion	Bonus	X-rays

3.2.6. Mutant analysis of the paralog of *HvLAX-A* in barley

Sequence homology identified the gene *HvLAX-A* as a homolog of *Arabidopsis thaliana BLADE-ON-PETIOLE 1 and 2* (*BOP1/2*) genes which are known to be involved in regulating leaf morphogenesis, floral organ abscission and flower asymmetry (Hepworth et al., 2005; Norberg et al., 2005). The Arabidopsis proteins carrying a BTB/POZ (broad complex, tramtrack, and brick à brack/poxviruses and zinc finger) and an ANK (ankyrin repeats) domain, a combination that occurs in members of a small gene family. The family includes four homologs of the plant defense related *NONEXPRESSOR OF PR GENES1* (*NPR1*) and the *BLADE-ON-PETIOLE* genes *1* and *2* (*BOP1* and *BOP2*) (Hepworth et al., 2005). A sequence based conserved domain search (Marchler-Bauer et al., 2011) revealed similarity of the sequence of *HvLAX-A* to both domains. In order to check if barley, like Arabidopsis, contains multiple genes containing the BTB/POZ and ANK domain combination, a sequence similarity search was performed for the barley gene set (IBSC, 2012) and revealed five additional barley genes. To check if they can be assigned in similar groups like in Arabidopsis a combined phylogenetic analysis was performed with barley and Arabidopsis genes. Four genes clustered together with the *NPR1-like* genes while *HvLAX-A* as well as a second gene AK360734 grouped within the *BOP-like* glade (Figure 16). The identified *HvLAX-A gene*, together with the paralog on chromosome 3H (AK36074), thus very likely represent the barley orthologs of *BOP1* and *BOP2* in Arabidopsis.

Loss of function mutants of single *BOP* genes were causing no obvious or only weak phenotypes in Arabidopsis (Hepworth et al., 2005; Norberg et al., 2005). A small ectopic outgrowth on leaf petiole could be observed for some *bop1* mutants (Ha et al., 2004; Norberg et al., 2005). Double mutants have severe changes in leaf growth and inflorescence development (Hepworth et al., 2005; Norberg et al., 2005). Over expression lines of *BOP1* and *BOP2* showed strong changes in growth of shoots, leaves and flowers with similarity to the phenotype observed in double mutants (Norberg et al., 2005). Thus, it has been supposed that both genes function redundantly. *AtBOP1* and *AtBOP2* are expressed in similar tissues, while the expression of *AtBOP2* is always higher than *AtBOP1*. To test if similar expression pattern of the barley *BOP-like* genes could be observed, the published quantitative transcriptome sequencing data of eight independent tissues of barley was considered (IBSC, 2012). The highest expression of *HvLAX-A* gene was detected in early inflorescence development whereas the highest transcript accumulation of the paralogous gene AK360734

was observed in development of tillers at six leaf stage (Figure17). The fact that no obvious change in leaf morphology could be observed in *lax-a* mutant plants and the distinct expression pattern indicated that the function of the two genes in barley might be differential than in Arabidopsis.
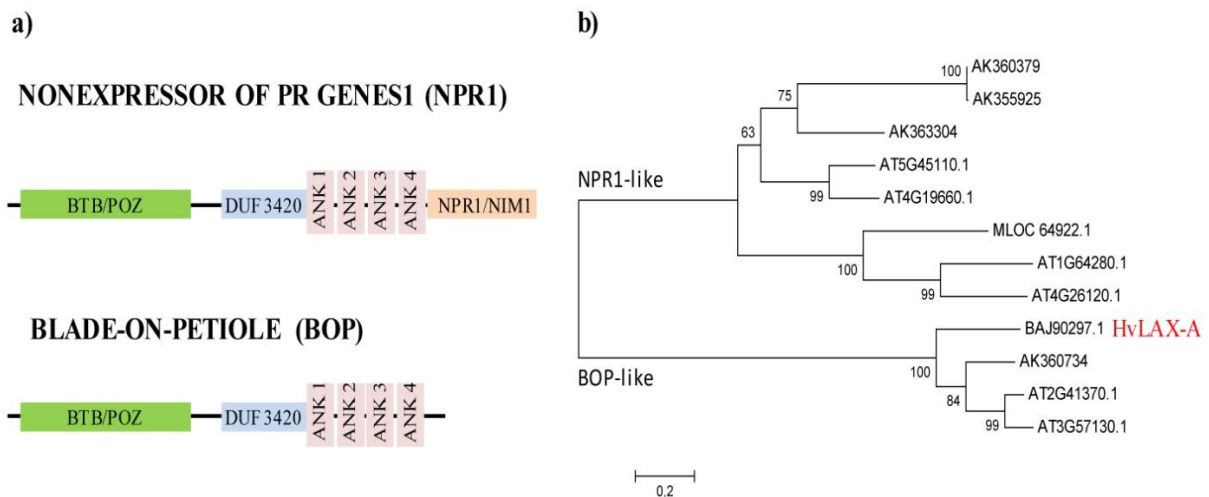
Figure 16: Conserved domain structure and gene family analysis.
The *HvLAX-A* candidate gene showed sequence similarity to a small gene family of Arabidopsis which is specified by a combination of BTB/POZ and ANK repeats conserved domains. This family comprises of the plant defense related genes NPR-like (nonexpressor of PR Genes) and the plant meristem boundary identity genes *BOP-like* (Blade-on-petiole). NPR1-like genes have a specific additional domain at the C-terminal end in contrast to *BOP-like* genes (schematically visualized in a). Figure is modified according to (Khan et al., 2014) b) A combined phylogenetic analysis of barley and Arabidopsis *NPR1-like* and *BOP-like* genes revealed the same grouping of barley and Arabidopsis (gene ID starts with AT) in a phylogenetic tree. The paralogous gene AK360734 is closer related to the Arabidopsis *BOP1/2* genes than *HvLAX-A* (BAJ90297.1).
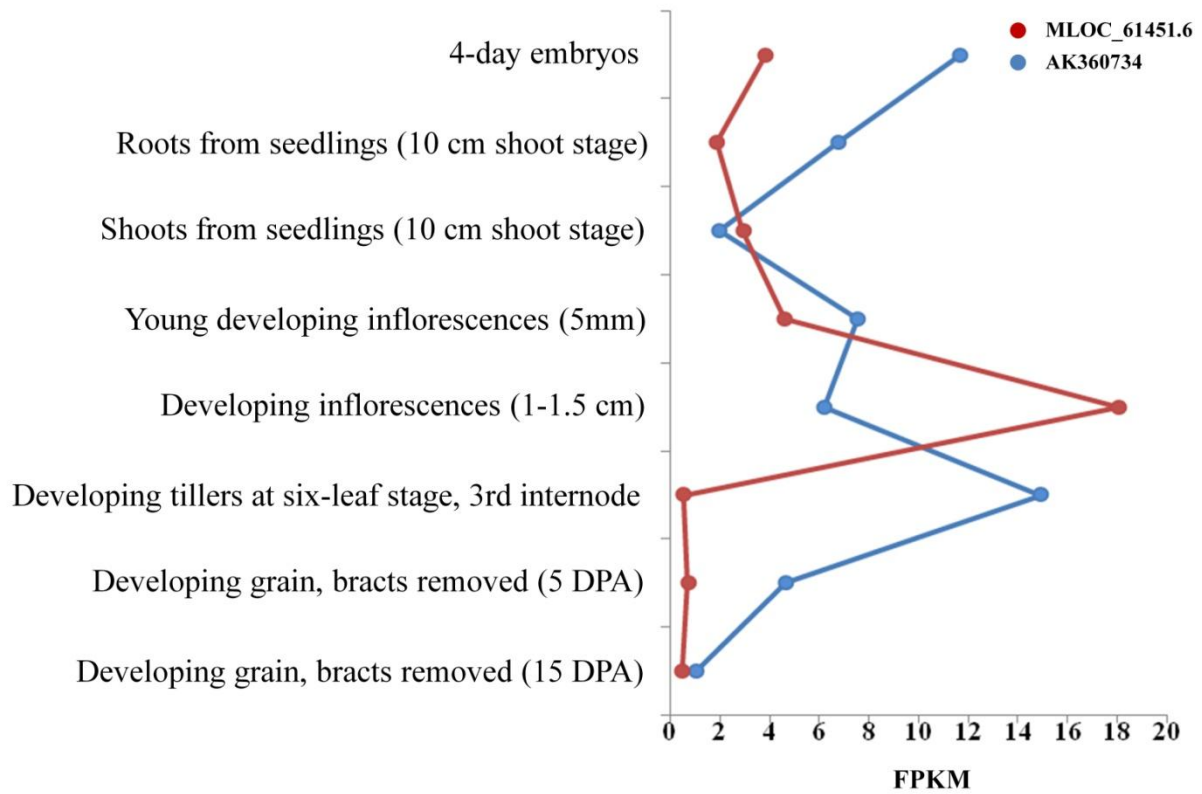
Figure 17: Expression of barley *BOP-like* genes during barley development.
Expression levels of the *lax-a* candidate MLOC_61451.6 (red) and 3H paralog AK360734 (blue) of available expression data (IBSC, 2012) across eight different tissues of development revealed distinct expression maxima. Transcript level is given as fragments per kilo base of exon per million read mapped (FPKM).

In order to test whether the gene function is redundant and leading to similar phenotypes in barley, the paralogous gene AK360734 was analyzed by TILLING. A series of 25 non-synonymous and 21 synonymous SNPs, one premature stop codon (9425_1) and one splice site mutation (13391_1) were identified by screening 7,979 M2 plants (Figure 18a, Table 19). None of the plants with non-synonymous mutation showed any obvious phenotypic effect. Mutants affected by a premature stop codon or a splice site mutation, however, exhibited drastic changes in plant development. Both mutants exhibited a liguleless phenotype and irregular outgrowth of auricles (Figure 18 d-f) and produced only less than three tillers (Figure 18 b-c). Both identified mutants were heterozygous in M2. Thus, segregating families were analyzed in M3, while only homozygous mutant plants showed the described phenotypic alterations (Table 18). Strong differences in plant performance were observed for homozygous mutant plants for both mutant families. Some of the mutant plants showed stunted growth and either died after strong curling in the first weeks (5x 9425_1) or stopped growing before flowering (1x 9425_1, 3x 13391_1). Homozygous mutant plants which grew to maturity (4x 9425_1; 1x 13391_1) showed none of the *lax-a* characteristic changes in spike morphology compared to wild-types of the respective TILLING family. The described reduced tiller number and liguleless plant phenotype was previously described for a mutation called *uniculme4*. By personal communication to the group of Dr. Laura Rossini, University of Milano, Italy, it was confirmed that this *lax-a* paralog is the causal gene of the *uniculme4* mutant in barley. This gene was cloned independently from this study and published recently (Tavakol et al., 2015). Thus, the gene AK360734 was renamed to *HvCUL4* in the following parts of the thesis.

Table 18: TILLING mutants of the *HvLAX-A* paralog gene on 3H

Plant	SNP	Effect	Wild-type	Heterozygote	Mutant
9425_1	G453A	W151*	11	9	10
13391_1	G547A	Splice Junction	2	4	4

* pre stop
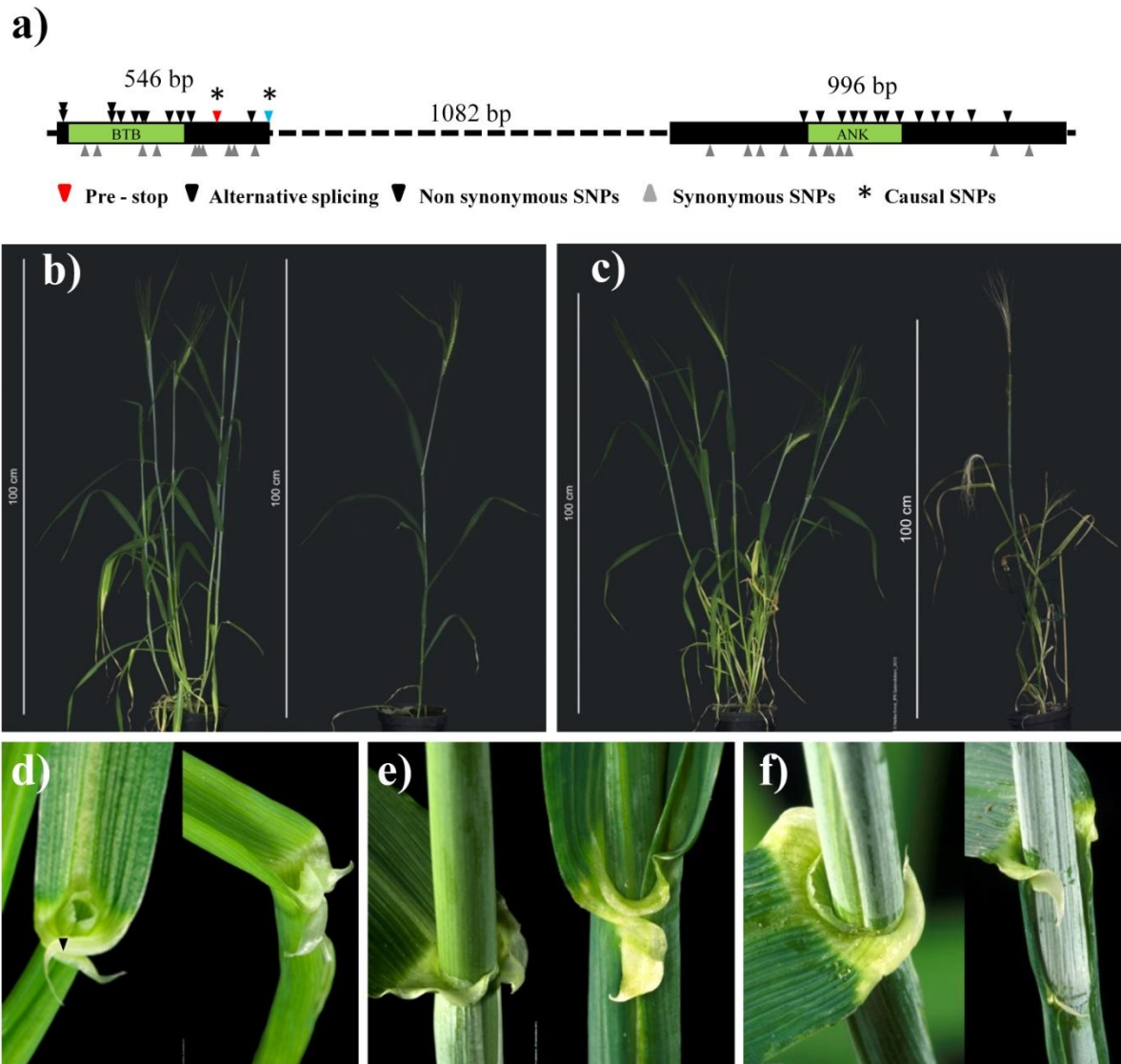
Figure 18: TILLING analysis of *HvCul4* revealed two mutants with *cul4* characteristic phenotypes.

a) Gene model of the gene *HvCul4* represented by two exons (black boxes) spaced by a single intron (black dashed line). SNPs identified by TILLING are displayed as triangles (colored according to classification) along the gene model. a) An example of a mutant plant with splice site mutation (13391_1) and TILLING mutant 9425_1 with premature stop codons (b) developed only few tillers (right) compared to wild-type. b) Plants in the gene (right). c-e) All mutant plants lacked ligules (ligule-less) but showed irregular outgrowth of auricles at the leaf sheath / leaf blade boundary.

Table 19: Identified TILLING mutants for *HvCul4*

plant	SNP	status M3	effect	phenotype
10131_1	G4A	HOM	D2N	wild-type
7902_1	G4A	HET	D2N	wild-type
14598_1	C117T	HOM	C39=	/
9985_1	C147T	HET	R49=	/
14030_1	C217T	HET	L73F	wild-type
8853_1	C217T	HET	L73F	wild-type
15147_1	T241A	HOM	S81T	wild-type
4515_1	G250A	HET	G84S	wild-type
9673_2	C300T	HET	S100=	/
3511_1	G301A	HET	A101T	wild-type
9741_1	G319A	HET	G107R	wild-type
6887_1	G343A	HET	V115M	no seeds
10441_1	C364T	HET	L122=	/
11076_1	A385T	HET	S129C	wild-type
14628_1	A410G	HOM	Q137R	wild-type
7366_1	G411A	HET	Q137=	/
14645_1	G414A	HET	K138=	/
4128_1	G417A	HOM	G139=	/
9425_1	G453A	HET	W151*	mutant
11239_1	C462T	HOM	H154=	/
11494_1	C481T	HET	L161=	/
6539_1	C494T	HET	T165I	no seeds
10561_1	C495T	HET	T165=	/
13391_1	G547A	HET	Splice Junction	mutant
3616_1	C1814T	HET	D244=	/
14799_1	C1847T	HOM	S255=	/
3208_1	G1868A	HET	Q262=	/
6800_1	G1949A	HET	R289=	/
7393_1	G1952A	HOM	M290I	wild-type
3667_1	C1970T	HOM	S296=	/
4419_1	G2007A	HET	E309K	wild-type
13286_1	G2036A	HOM	A318=	/
10095_1	C2037T	HET	L319=	/
14188_1	G2054A	HET	E324=	/
6525_1	G2065A	HET	R328Q	wild-type
15334_1	G2069A	HOM	E329=	/
4087_1	T2071A	HET	V330E	wild-type
6869_1	G2122A	HET	G347E	wild-type
13142_1	C2169T	HOM	P363S	wild-type
14634_1	G2172A	HET	D364N	wild-type
10496_1	T2242A	HOM	L387H	wild-type
10893_1	C2259T	HET	L393F	wild-type
10916_1	C2266T	HET	S395F	wild-type
14600_1	C2274T	HET	L398F	wild-type
11976_1	G2296A	HOM	G405E	wild-type
9922_1	C2417T	HET	H445=	/
4652_1	C2421T	HOM	P447S	wild-type
3387_1	C2480T	HOM	N466=	/

[1] HOM = homozygote, HET = heterozygote

3.2.7. Natural diversity of *HvLAX-A* and *HvCUL4*

The two genes *HvLAX-A* and its paralog *HvCUL4* are involved in regulating the two major agronomical relevant traits of tiller number and spike morphology. It was aimed to analyze natural diversity to test if both genes were involved in a natural selection process during barley cultivation. A statistical test can be applied to prove whether the pattern of polymorphisms observed in a population is affected by a selection process or if it is of random origin (Fu and Li, 1993).
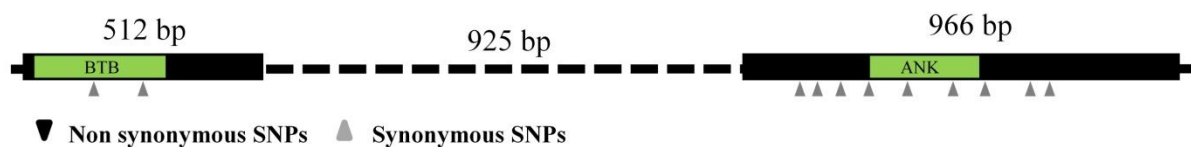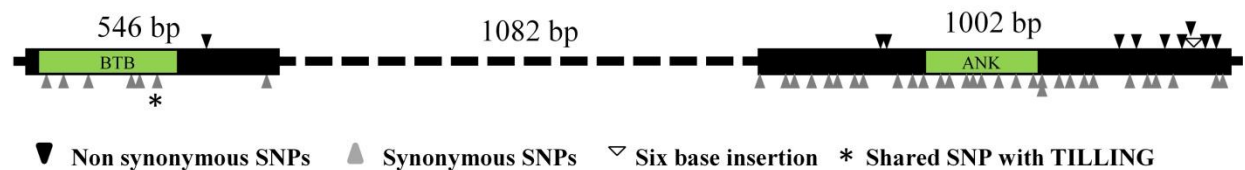
Figure 19: Distribution of indentified polymorphic sites within the ORF of *HvLAX-A* and *HvCUL4*.
Resequencing the ORF of both genes in 305 barley accessions revealed a set of sequence variations which are positioned along the gene models of *HvLAX-A* (a) and *HvCUL4* (b).

A set of 83 wild (*Hordeum spontaneum*) and 222 domesticated barley lines was used as plant material. The 222 adapted lines were taken from a pre-existing worldwide spring barley collection consisting of 150 advanced cultivars, 55 landraces, 17 accessions of breeding or research material (Haseneyer et al., 2010). To identify variation within the genes, the complete ORF was amplified by three primer combinations for *HvLAX-A* (HvLAX_F/R4, HvLAX_F/R5, HvLAX_F/R6) and *HvCUL4* (HvCUL4_F/R1, HvCUL4_F/R2, HvCUL4_F/R3), respectively, and subsequently sequenced (Table A5 and Table A6). In total, 11 polymorphic sites (S) were identified within the coding sequence of *HvLAX-A* consisting of nine synonymous changes and two non-synonymous changes (Figure 19a). A higher

number of polymorphic sites was identified for the paralogous gene *HvCUL4*. The total number of 48 sites was represented by 38 synonymous SNPs and 10 non-synonymous SNPs. Additionally to the SNP variations, some accessions carried a 3 bp (adapted material) or 6 bp (wild material) insertion within the second exon of *HvCUL4* (Figure 19b). Interestingly, a single SNP within *HvCUL4* was shared with the mutations identified from the TILLING analysis.

The obtained re-sequencing data for both genes was not only analyzed as combined dataset of 305 accessions. The accessions were also separately analyzed for wild and adapted barley accessions to study changes of diversity between them. Then a statistical test was applied to test whether these changes followed a random process or were induced by non random selection process. First step was to assign accessions with identical sequences to haplotypes. The sequence polymorphisms detected for *HvLAX-A* distinguished 10 exon haplotypes with a haplotype diversity (hd) of 0.307. While all 10 haplotypes were represented in the wild material, only two of them remained considering the adapted accessions. This led to a reduction of haplotype diversity from hd=0.644 to hd=0.0781 (Table 20). Like already indicated by the increased number of polymorphic sites for *HvCUL4*, a much higher haplotype diversity was observed compared to *HvLAX-A*. In total 31 exon haplotypes (hd=0,8582) were indentified. Among the wild material, 26 haplotypes were detected while the adapted material represented only 13 haplotypes. Of those, 11 were shared between wild and domesticated accessions. The haplotype diversity between wild and adapted material was reduced by about 18.7 % from hd=0.9412 to hd=0.765 (Table 20).

Table 20: Statistics of the diversity analysis of *HvLAX-A* and *HvCUL4*

Parameter	*HvLAX-A*			*HvCUL4*		
	wild	adapted	all	wild	adapted	all
Number of haplotypes	26	13	31	10	2	10
Haplotype diversity (hd)	0.9412	0.765	0.8582	0.6644	0.0781	0.307
Number of sites	1548	1545	1551	1476	1476	1476
Number of sites*	1542	1542	1542	1476	1476	1476
Nucleotide diversity (π)	0.00456	0.00493	0.00499	0.0009	0.00005	0.00034
Polymorphic sites (S)	44	27	48	11	1	11
synonymous	34	24	38	9	1	9
non-synonymous	10	3	10	2	/	2
Tajima's D	-0.70735	1.72692	-0.03471	-1.06982	-0.53007	-1.61662

* gaps excluded

Based on their similarity, haplotypes were connected to each other forming a network (Figure 20). The network of *HvLAX-A* was defined by one major haplotype (Hap-2) from which minor haplotypes differentiated by one or two polymorphisms. The network of *HvCUL4* revealed two groups separated by 10 SNPs between haplotype 11 (Hap-11) and haplotype 2 (Hap-2). In order to test if these two groups can provide a hint for regional adaptation, the geographic origin of each accession was considered. Hap-11 and Hap-2 consisted, like all other major haplotypes, from accessions originating from Eastern Asia, America, Western/Northern Asia as well as Europe (Table A1). Thus, there was no indication that polymorphisms between the groups were caused by regional adaptation concerning latitude. It has previously been reported that the world wild collection of adapted material used in this study showed population structure into two groups according two-rowed and six-rowed barley (Haseneyer et al., 2010). Hence, the accessions of all haplotypes were analyzed for row type. Similar like for the geographic origin, no row type specific structure could be observed for the detected haplotypes.

The gene *HvLAX-A* exhibited highly reduced nucleotide diversity in comparison to the gene *HvCUL4*. The low nucleotide diversity of $\pi = 0.0009$ for *HvLAX-A* in wild barley was further decreased in domesticated barley to $\pi = 0.0005$. A contrasting pattern was observed for *HvCUL4* in terms of nucleotide diversity. A slightly (7.5%) increased nucleotide diversity between wild (0.00456) and adapted (0.00493) material was detected for gene *HvCUL4* (Table 20).

The analysis revealed reduction of haplotype diversity between wild to adapted barley accessions for both genes. The nucleotide diversity was reduced as well for *HvLAX-A* but in contrary it was slightly increased for *HvCUL4* between wild and adapted material. The Tajima test (Tajima, 1989) for neutrality of DNA polymorphisms was performed to test if these observed changes could have been the result of either non-random shift or due to natural selection. None of the calculated Tajima D values reached a significant value (p<0.1). Thus, the observed changes in diversity may not be attributed to selection.
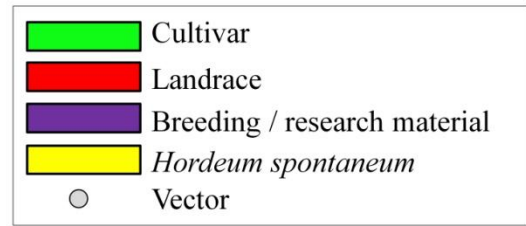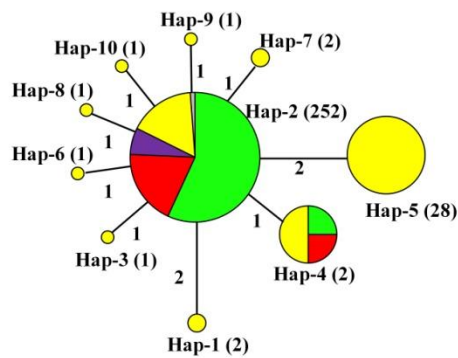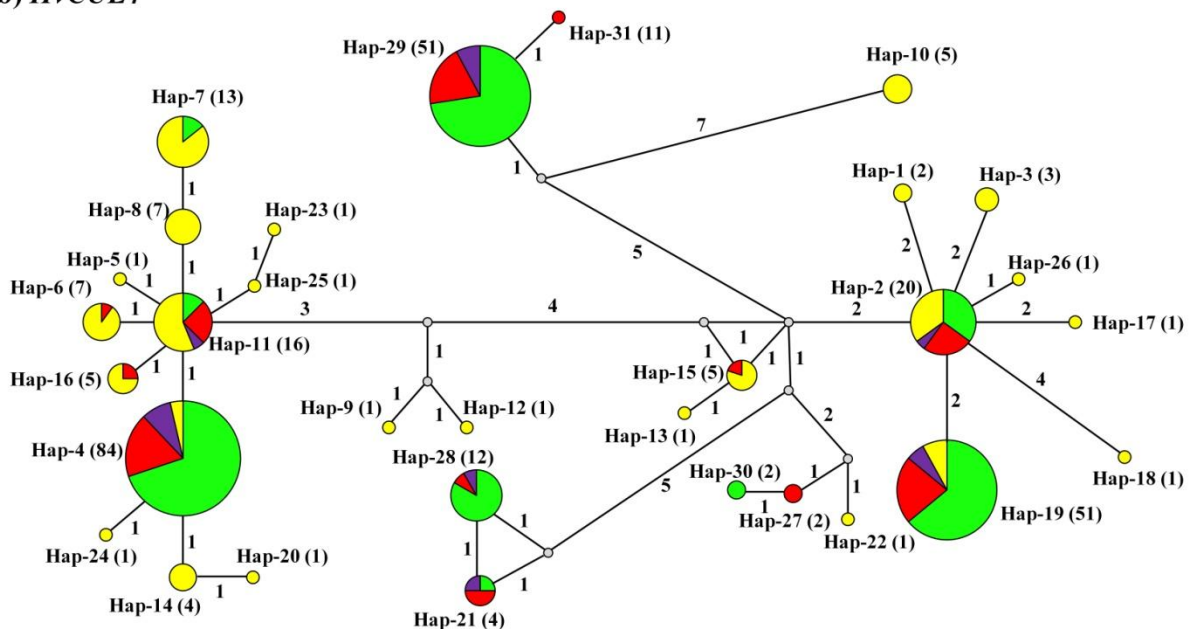
Figure 20: Diversity analysis of *HvLAX-A* and *HvCUL4*.

Median-joining network derived from haplotypes identified by re-sequencing the ORF of *HvLAX-A* (a) and *HvCUL4* (b) in 83 *Hordeum spontaneum* accessions (yellow), 55 barley landraces (red), 150 barley cultivars (green), 17 accessions of breeding / research material (purple). Haplotypes were labeled with haplotype ID and number of accessions sharing the respective haplotype (brackets). Circle sizes are proportional to numbers of accessions per haplotype. Length of connector lines refers to number of nucleotide substitutions between haplotypes (indicated by numbers on connecting lines = number of mutations).

3.2.8. Global analysis of gene expression

The Arabidopsis genes *BOP1/2* are described to contain two protein – protein binding domains and working as transcriptional co-activators in Arabidopsis (reviewed in Khan et al., 2014). Since the two conserved domains were present as well in the barley *HvLAX-A* gene, a similar regulating function may be postulated. To further understand the role of *HvLAX-A* in barley inflorescence development a global transcription profiling experiment was performed by quantitative high-throughput cDNA sequencing (RNA-seq). Available transcription data (IBSC, 2012) revealed the highest transcript accumulation of *HvLAX-A* in developing inflorescences and *lax-a* mutants exhibited significant changes in the inflorescence organ identity. Thus, young developing inflorescence meristems were chosen as tissue for RNA isolation to study the function of *HvLAX-A* on expression of other genes. Three developmental stages were selected as tissues (Figure 3a), covering the complete time frame of inflorescence organ initiation: (i) early development glume primordium stage (ii) stamen primordium stage (iii) completely developed spikes (30 days after germination). Bowman was taken as wild-type and compared with the *lax-a* mutant NIL BW457. Comparison of expression levels of genes between wild-type and mutant was used to identify gene regulations that might be attributed to loss of function of *HvLAX-A*.

Four samples were sequenced together per lane on the Illumina HiSeq2000 instrument which delivered a total of 1.4 billion paired reads after quality filtering. 60 to 100 million read pairs per sample were included for the quantitative expression analysis. Differential expression analysis was performed with DEseq (Anders and Huber, 2010). DEseq takes different read counts per sample into account and normalizes raw counts by applying a correction factor based on a statistical binomial distribution (Anders and Huber, 2010). Obtained mean read values were multiplied with a correction factor between 0.77 und 1.27, determined by the software algorithm, to compensate minor differences in number of mapped reads per sample across the replicates. Significantly differential gene expression was tested for false discovery rate (FDR) by the Benjamini-Hochberg method (Benjamini and Hochberg, 1995). A quality check was introduced to check the reliability of expression values among the biological replicates. For this, a correlation heatmap of gene expression values between samples was drawn in R with the function "heatmap.2" of the "gplots" package. Replicates within a single developmental stage were expected to show highest correlations and should lead to formation of separated clusters according to wild-type and mutant samples. This clustering was observed only for the last stage of development. Unfortunately, the first two developmental

stages did not cluster as expected. The different tissue stages as well as the mutant and wild-type samples did not clustere separately from each other (Figure 21a). Considering only significantly differentially expressed genes leads to the expected genotype specific clustering for all samples of BW457 and Bowman, respectively. Only the sample laxatum_st2 and laxatum_st3 showed a higher correlation to the samples of the first (glume primordium) developmental stage (Figure 22b). However, the overall correlation of all replicates was very high. The introgression of Bowman NIL457 was only restricted to chromosome 5H in a Bowman background. Since expression values were compared to Bowman as wild-type, an overall high amount of similar gene expression can be expected. The samples of the first two developmental stages are very close together in developmental time. Since multiple spike meristems per sample are harvested to have sufficient tissue for RNA isolation, small developmental differences might introduce 'noise' into the analysis. However, as expected for an entirely deleted gene no reads of the gene *HvLAX-A* were detected in all samples derived from BW457. Thus an error in sampling could be excluded and therefore the analysis was conducted with all replicates and analyzed separately for each developmental stage.
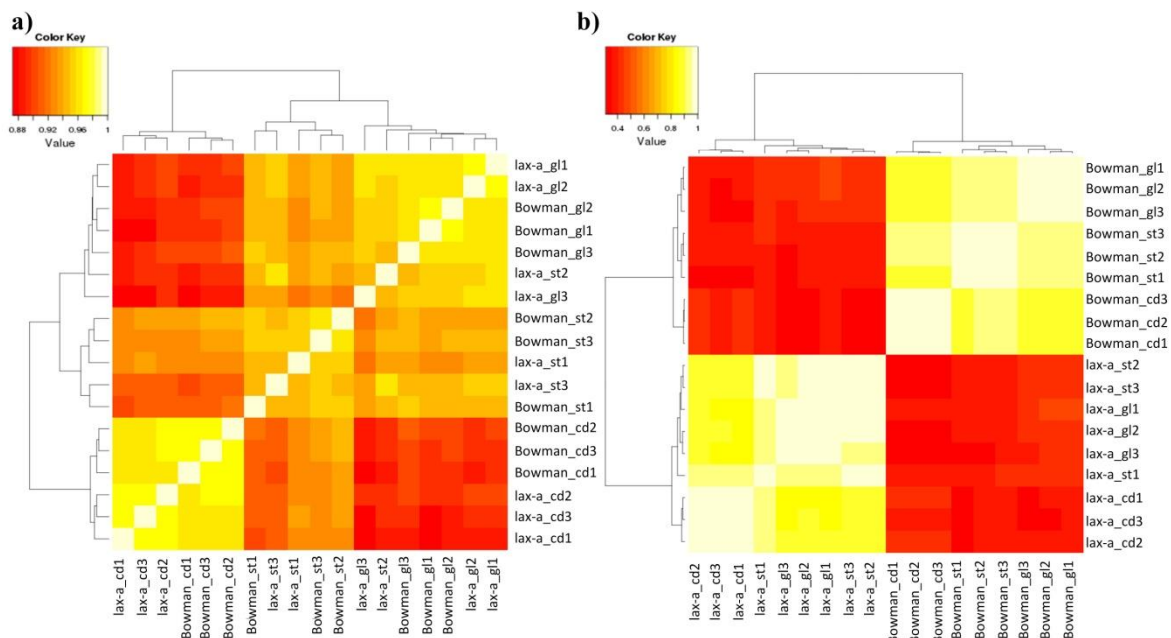
Figure 21: Correlation heat map of all samples and replicates of the RNA-seq expression analysis.
The heatmaps show the correlation of gene expression values between samples and are colored by a red (low correlation) to white (high correlation) scale. a) The correlation matrix was plotted by all genes (a) and only for the differentially expressed genes (b).

In total, 144 high confidence genes (IBSC, 2012) showed a significant (p < 0.05) differential expression between Bowman and BW457 samples in at least one of the addressed tissues. Among these, 20 down regulated and 19 upregulated genes showed high differential expression (log2-fold change >2) between Bowman and BW457 (Table 22). A summary of all differential gene expression statistics (normalized mean read counts per gene and sample, log 2 fold change, significance level and their functional annotation (IBSC, 2012) is provided in Table 22 (log2 fold > 2) and Table A21 (log2 fold < 2).

Only 33 genes were significantly differentially regulated between Bowman and BW457 among all three developmental stages. This included three 'phytohormone associated' genes, six 'transcription related' genes with protein-protein interaction domains as well as 11 unknown proteins according to the published annotation information (IBSC, 2012). Since it is not known at which of the three developmental stages *HvLAX-A* might act for establishing a wild-type barley spike, the three time points of development were inspected separately for putative candidate genes that showed differential regulation between Bowman and BW457. Three down-regulated homeobox leucine zipper transcription factors (TFs) were identified in the glume primordium stage in BW457. In the 'stamen primordium stage two' one ectopically expressed MADS-box TF gene (MLOC_76418.1) was observed, whereas in stage of 'completely developed immature spikes' five transcription factor related genes were up-regulated in *lax-a* samples compared to Bowman.

The paralogous gene *HvCUL4* (AK360734) was expressed at similar levels in both mutant and wild-type genotypes. Thus deletion of *HvLAX-A* did not induce any compensating over-expression of the paralogous gene *HvCUL4* in the analyzed tissues (Figure 22).

The genes *BOP1/2* in Arabidopsis were in the focus of intensive research over the last 10 years and a number of genes were described to be regulated by both genes. The orthologs of those genes known to be involved in regulating pathways were inspected for conserved gene regulation between barley and Arabidopsis. A summary table of known interaction partners is given in Table 21. The nucleotide sequences of Arabidopsis genes (http://www.arabidopsis.org) were used to identify homologous barley genes (Table 21). For some genes the entire gene family was listed since the sequence conservation did not allow a prediction of the putative orthologs. Among these genes, *HvLAX-A* remained the only candidate with high difference in expression between Bowman and BW457 samples. A second gene, a LOB domain containing gene which showed similarity to *AS2* of Arabidopsis reached the significance level (p<0,05) of the DE test in the last sampled tissue stage of

completely developed spikes (Table 23). This indicated that the regulation of *BOP-like* genes may not be well conserved between Arabidopsis and barley in the analyzed tissues. However, some of the described regulations from Arabidopsis are known from studies of morphological changes in leaf development and thus might be not comparable to the analyzed spike meristem tissue in the present study.
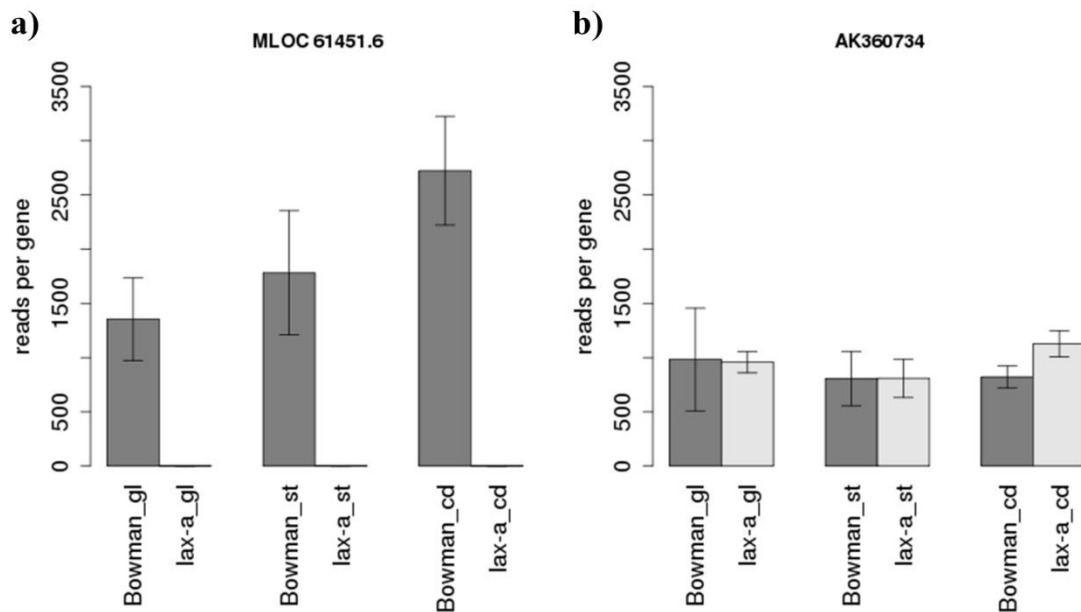
Figure 22: Expression analysis of the genes *HvLAX-A* and *HvCUL4*.
The expression values obtained from RNAseq compared between Bowman and BW457 of the three developmental stages (glume primordium, stamen primordium, completely developed spike) for *HvLAX-A* (a) and *HvCUL4* (b). No reads were detected for *HvLAX-A* in samples of the BW457 mutant line whereas the gene was highly expressed in wild-type tissue. The expression of the paralogous gene *HvCUL4* was not affected thus no differences in transcript level could be observed between wild-type and mutant plants.

Table 21: Arabidopsis genes involved in regulatory pathways of *BOP1/2*

Name	Short	GeneID	comments
BLADE ON PETIOLE 1	*BOP1*	AT3G57130.1	
BLADE ON PETIOLE 2	*BOP2*	AT2G41370.1	
ASYMMETRIC LEAVES 1	*AS1*	AT2G37630.1	MYB-domain protein
ASYMMETRIC LEAVES 2	*AS2*	AT1G65620.1	Lateral organ boundaries (LOB)
JAGGED	*JAG*	AT1G68480.1	zinc finger transcription factor
NUBBIN	*NUB*	AT1G13400.1	zinc finger transcription factor
LEAFY	*LFY*	AT5G61850.1	
PUCHI	*PUCHI*	AT5G18560.1	no homolog in barley
PERIANTHIA	*PAN*	AT1G68640.1	TGA8 TF
APETALA1	*AP1*	AT1G69120.1	MADS-box protein
AGAMOUS-LIKE 24	*AGL24*	AT4G24540.1	MADS-box protein
KNOTTED1-LIKE - HOMEOBOX GENE 6	*KNAT6*	AT1G23380.1	class I knotted1-like homeobox
BREVIPEDICELLUS	*BP*	AT4G08150.1	class I knotted1-like homeobox
PENNYWISE	*PNY*	AT5G02030.1	BEL1-LIKE Homeodomain
HOMEOBOX GENE 1	*ATH1*	AT4G32980.1	

Table 22: Differentially expressed (log2_fold >2) genes between Bowman and BW457

HC_gene[1]	normalized mean wild-type[2]			normalized mean mutant[2]			log2FC_gl	log2FC_st	log2FC_cd	adj p=<0.05			annotation[1]
	bowman_gl	bowman_st	bowman_cd	lax_gl	lax_st	lax_cd				gl	st	cd	
MLOC_61451.6	1320.00	1782.55	2824.00	0.30	1.20	0.29	-12.09	-10.53	-13.24	X	X	X	NPR1 protein = *HvLAX-A*
MLOC_16761.1	1082.49	995.52	876.52	0.30	0.34	0.89	-11.80	-11.50	-9.94	X	X	X	BTB/POZ domain containing protein
MLOC_50578.1	489.55	304.34	390.32	1.89	0.97	1.80	-8.02	-8.29	-7.76	X	X	X	Unknown protein
MLOC_16.1	28.18	23.03	32.90	0.30	0.34	0.31	-6.55	-6.08	-6.74	X	X	X	Unknown protein
AK361956	781.34	611.33	623.28	3.42	5.94	9.16	-7.83	-6.69	-6.09	X	X	X	Unknown protein
AK252827.1	1.15	57.96	970.84	0.30	0.34	571.73	-1.92	-7.41	-0.76		X		Polyamine oxidase, putative
AK361946	4085.77	3157.46	3777.40	20.46	24.34	34.33	-7.64	-7.02	-6.78	X	X	X	Unknown protein
MLOC_80940.1	51.72	80.04	51.64	0.61	2.39	0.94	-6.41	-5.06	-5.78	X	X	X	Unknown protein
MLOC_37360.1	18.74	20.03	8.63	0.34	0.34	0.62	-5.80	-5.88	-3.81	X	X	X	Unknown protein
MLOC_45680.1	13.36	9.33	15.31	0.30	0.34	0.31	-5.46	-4.78	-5.63	X	X	X	Unknown protein
MLOC_19415.1	24.41	29.77	50.24	1.61	0.61	4.11	-3.92	-5.60	-3.61			X	Gibberellin receptor GID1L2
MLOC_78355.1	53.47	46.48	66.19	4.09	3.50	0.93	-3.71	-3.73	-6.16			X	Unknown protein
MLOC_48386.1	515.41	507.85	349.89	54.34	41.26	31.24	-3.25	-3.62	-3.49	X	X	X	Protein kinase family protein
MLOC_5849.1	7.99	21.13	90.95	0.93	1.31	34.12	-3.10	-4.01	-1.41	X	X	X	myb domain protein 36
AK358964	235.98	287.59	176.03	32.44	42.36	26.79	-2.86	-2.76	-2.72	X	X	X	IAA-amino acid hydrolase ILR1
MLOC_31142.2	12.58	10.47	13.70	1.91	0.54	0.61	-2.72	-4.28	-4.50		X	X	B12D protein
AK358685	17.43	17.57	19.34	3.31	16.43	36.32	-2.40	-0.10	0.91	X			GDSL esterase/lipase
MLOC_56666.1	1238.93	1683.87	1590.02	299.26	393.62	487.73	-2.05	-2.10	-1.70	X	X	X	Ubiquitin-like protein 5
MLOC_68776.2	56.15	44.37	98.25	13.78	23.26	70.49	-2.03	-0.93	-0.48	X			VQ motif family protein
AK377028	28.84	40.50	51.98	7.23	36.43	38.78	-2.00	-0.15	-0.42	X			Cytochrome P450
MLOC_58320.2	340.84	202.63	148.80	889.26	824.27	585.08	1.38	2.02	1.98	X	X	X	Unknown protein
AK248367.1	2170.11	503.73	2726.45	1161.62	2086.70	1185.63	-0.90	2.05	-1.20		X		Histone H2A
MLOC_50612.1	15.03	17.53	17.37	74.66	43.43	37.57	2.31	1.31	1.11	X			S-receptor kinase-like
MLOC_33768.7	3.93	8.78	11.47	7.87	18.30	55.30	1.00	1.06	2.27			X	Stress responsive alpha-beta barrel domain protein
MLOC_73626.1	27.24	25.63	23.89	15.69	8.52	121.97	-0.80	-1.59	2.35			X	BZIP transcription factor
MLOC_64580.5	25.88	31.92	23.90	67.08	70.56	121.53	1.37	1.14	2.35	X			Protein kinase superfamily protein
MLOC_76418.1	91.29	38.41	419.66	202.81	211.34	425.58	1.15	2.46	0.02		X		MADS-box transcription factor 8
MLOC_69030.3	8.21	15.74	4.05	47.51	39.02	6.74	2.53	1.31	0.73	X			Serine-glyoxylate aminotransferase
MLOC_15093.1	7.12	4.43	5.98	43.87	33.11	59.73	2.62	2.90	3.32	X	X	X	Kinase, putative
MLOC_12869.1	0.38	0.73	1.73	5.31	15.37	16.50	3.81	4.39	3.25			X	Exostosin family protein, expressed
MLOC_74917.2	1.91	1.47	2.14	53.83	52.98	27.53	4.82	5.17	3.68	X	X	X	F-box domain containing protein
AK377105	3.70	12.25	33.91	189.00	214.54	171.01	5.67	4.13	2.33	X	X	X	Ureide permease 2, putative, expressed
MLOC_48640.1	0.26	0.33	0.72	11.33	11.11	7.97	5.44	5.08	3.47	X	X		Receptor-like kinase
MLOC_29207.2	9.42	14.66	71.36	511.50	676.10	608.67	5.76	5.53	3.09	X	X	X	Gag/pol polyprotein
MLOC_53891.1	0.26	0.63	4.29	27.17	37.16	48.03	6.71	5.89	3.49	X	X	X	Unknown protein
AK372632	0.26	0.33	0.31	74.77	50.32	66.91	8.17	7.25	7.75	X	X	X	Unknown protein
MLOC_31926.3	2.69	0.73	0.31	207.12	210.92	128.97	6.26	8.17	8.70	X	X	X	Ferredoxin-thioredoxin reductase catalytic chain
AK248907.1	0.26	0.33	0.31	432.94	360.48	449.81	10.70	10.09	10.50	X	X	X	Chaperone protein dnaJ, putative
MLOC_2686.1	0.64	0.40	0.31	1667.15	1658.06	1646.66	11.35	12.00	12.37	X	X	X	F-box protein

[1] High confidence gene model with predicted annotation (IBSC, 2012);

[2] normalized read numbers of DEseq analysis for the developmental stages of gl = glumen primordia, st = stamen primordia and cd = complete dveloped spikes

less than 50 reads/gene

color scale   -13.00   0.00   13.00
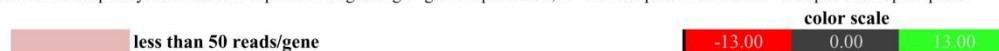
Table 23: Differences in expression values (log2-fold change) between Bowman and BW457 for homologous barley genes of known *AtBOP1/2* regulated genes.

Class[1]	Gene[2]	log2FC_gl	log2FC_st	log2FC_cd
BOP1	AK360734.1	-0.03	0.04	0.26
BOP2	MLOC_61451.6	**-12.09***	**-10.53***	**-13.24***
KNOX1-like	MLOC_57232.2	0.11	-0.05	0.37
	MLOC_11678.1	0.28	0.07	0.11
	AK369536	-0.98	0.97	0.95
	AK376780	0.09	-0.01	0.01
	AK373181	-0.04	0.14	0.13
	AK358211	-0.16	-0.01	-0.19
Ap1-like	AK360697	-0.04	0.10	0.30
	MLOC_61901.1	-0.34	-0.14	0.11
	AK361227	-0.07	0.28	0.20
LFY	MLOC_14305.1	-0.23	0.19	0.16
AS1 –like	AK372839	-0.61	0.55	0.23
	AK372889	-0.44	-0.03	0.33
LOB- like genes	AK373607	-0.32	-0.13	-0.49
	MLOC_61156.1	-0.94	-0.29	**-0.57***
	MLOC_55239.1	-0.67	0.43	-0.43
	MLOC_52276.7	0.28	-0.05	-0.11
	MLOC_51325.1	0.37	1.60	-0.38
	MLOC_58304.1	-0.77	0.03	0.21
	AK373051	-1.09	0.34	0.20
	MLOC_11838.1	-0.20	1.26	0.35
yabby like	MLOC_60897.1	-0.21	0.03	0.42
	MLOC_18466.1	-0.40	0.02	-0.05
	MLOC_70653.1	-0.19	-0.27	-0.16
	MLOC_66260.1	-1.42	-0.44	-0.32
	MLOC_70676.2	-0.12	0.16	0.02
JAG/NUB	MLOC_4391.1	-0.56	0.21	0.00
AGL24-like	AK374424	-0.70	-0.04	0.44
	AK370732	-0.31	0.11	-0.02
	AK250344.1	-0.47	-0.34	-0.13
	AK374272	-0.38	-0.15	0.09
	AK360697	-0.04	0.10	0.30
PNY-like	MLOC_66099.1	-0.13	0.42	0.01
	MLOC_3638.2	-0.03	-0.06	0.25
ATH1-like(bel1-like)	MLOC_53364.1	0.13	-0.11	0.07
	MLOC_5414.1	0.00	0.09	0.27
	MLOC_68213.2	-0.13	0.35	0.97
	AK367579	1.56	3.42**	-1.11
	AK358615	-0.30	-0.08	-0.21
	AK360144	-0.22	-0.39	-0.34
	MLOC_3638.2	-0.03	-0.06	0.25
PAN-like	AK362993	-0.12	-0.29	-0.08
	AK357239	-0.06	0.14	-0.26
	AK359391	-0.01	0.23	0.13
	MLOC_64008.1	-0.03	0.11	0.00
	AK363931	0.27	0.15	0.10
	MLOC_14596.3	-0.22	0.15	0.15

[1] Genes and gene classes reported to be regulated by *BOP1/2* in Arabidopsis

[2] Homologous barley gene models (IBSC, 2012) with changes in expression levels for the 3 different analyzed tissues between Bowman and BW457. Genes with less than 10 reads/gene were excluded and considered at not expressed

* significant differentially expressed (p<0,05); ** 19 reads mapped in single sample. Most likely artifact

3.2.9. Phylogenetic analysis of BOP-like genes within the plant kingdom

In a previous study, it was shown by a global analysis of *BOP-like* genes across the plant kingdom that most species were carrying two or three copies of *BOP-like* genes (Khan et al., 2014). Since *BOP1/2* in Arabidopsis showed redundant function while the Barley *BOP-like* genes indicated an independent function, it was aimed to study the evolutionary conservation of the *BOP-like* genes in the plant kingdom. For this, the protein sequences of *HvLAX-A* and *HvCUL4* were taken as query for screening the protein databases of NCBI and Phytozyome for homologous proteins in a number of distinct plant species. A maximum likelihood-based phylogenetic tree was constructed, which showed a plant family specific separation of the sequences (Figure 23). Two independent clades of gene sequences were found for all analyzed species. HvLAX-A and CUL4 distributed, like BOP1 and BOP2 of Arabidopsis, to distinct clades within the respective gene family. There was no hint for any difference in gene evolution between barley and *Arabidopsis*. No clear separation of monocots and dicots could be observed which may indicate a coevolution of these genes. Interestingly, a third copy present in several species clustered always together in the same subclade of the respective plant families. For instance, the clade containing *HvCUL4* contained only a single copy in all analyzed monocots. The subclade containing *HvLAX-A* included always two copies of *BOP-like* genes except for barley and *Brachypodium*. This pattern could be also observed for the *Brassicaceae* family including Arabidopsis. The only exception is Brassica that exhibit four copies in total, two in every subclade, which might be reasoned by its allotetraploid genome.
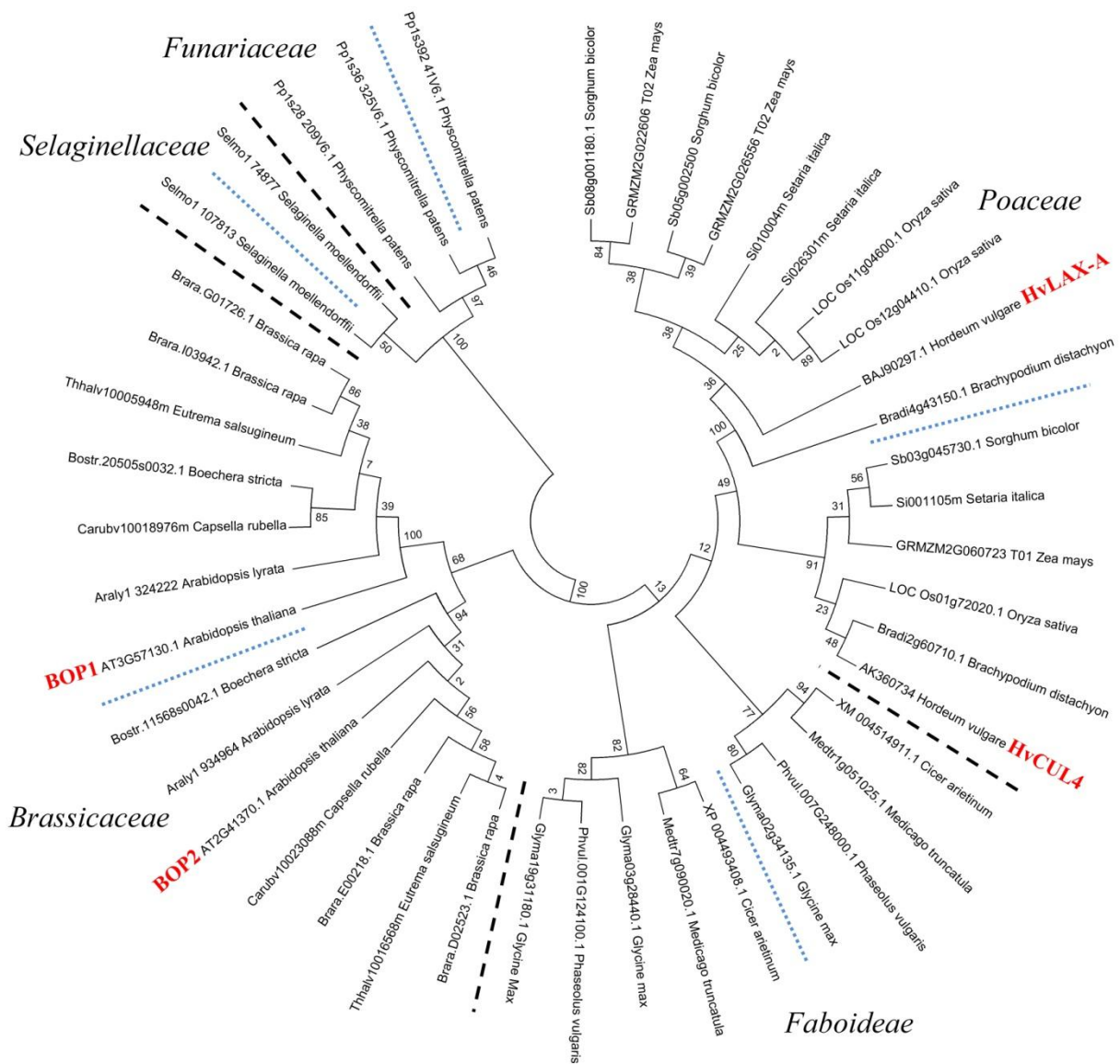
Figure 23: The *BOP-like* gene family organization among the plant kingdom.
A maximum likelihood phylogenetic tree was constructed based on 50 protein sequences of
BOP-like genes within the plant kingdom. All genes separated in clades according to their
respective plant families (indicated by dashed lines). Within each plant family two branches
of BOP genes can be (separated by the blue pointed line). BOP1 and BOP2 of Arabidopsis
and HvLAX-A and HvCUL4 were separated by these clades. Each protein was labeled by the
protein ID as well the plant species.

4. Discussion

This study aimed at the molecular cloning of two barley genes: *MANY NODED DWARF* (*MND*) and *LAXATUM-A* (*LAX-A*), both affecting important aspects of plant architecture. The mutant gene *mnd* influenced the 'plastochron', the frequency of leaf initiation, resulting in plants carrying many more phytomers (plant architecture unit consisting of node, internode, and leave) at a given time after germination compared to wild-type plants. Mutations in the second gene *LAX-A* mainly affected rachis internode length in the inflorescence, but also caused homoeotic changes in the barley flower by converting lodicules into additional stamens. Both genes were allocated to genomic regions that strongly differed in recombination frequency, and thus in their potential of being identified by classical map-based cloning. In both cases, the innovative strategy of mapping (cloning)-by-sequencing resulted in high resolution mapping of the mutant gene loci and the identification of candidate genes. Independent mutant alleles of the two genes could be found in historical mutant collections and by TILLING (targeting local lesions in genomes), which showed the same phenotype as the original *mnd* and *lax-a* mutants. Thus the identified candidate genes were confirmed to represent the causal genes underlying the mutant phenotype in both cases. Sequence comparison and conserved domain structure analysis uncovered the *HvMND* as member of the Cytochrome P450 class 78A gene family. The *LAX-A* gene was identified to be homologous to the *Arabidopsis thaliana BLADE-ON-PETIOLE-1/2* (*BOP1/2*) genes, which are transcriptional regulators that contain ankyrin repeats and a BTB/POZ domain and control together leaf morphogenesis and floral organ abscission in *Arabidopsis*. Like in *Arabidopsis*, a second closely related paralog of *LAX-A* could be identified on barley chromosome 3H. Mutant analysis revealed the influence of this gene on barley shoot branching since two mutants from a TILLING screen showed a low tiller phenotype. Based on personal communication with the group of Dr. Laura Rossini, University of Milano, Italy, it was confirmed that this *lax-a* paralog is the causal gene of the *uniculme4* mutant in barley, exhibiting strongly reduced tiller number and a liguleless plant phenotype (Tavakol et al., 2015). Analysis of natural diversity of both genes indicated distinct patterns of diversity, however, whether these were a consequence of selection or drift or if they were just a result of their different positions in the genome will require further genome-wide studies. Quantitative gene expression analysis of three independent stages of early spike development was performed to reveal candidate genes involved in *lax-a* specific regulatory gene networks leading to the observed phenotype. A number of strong candidate genes could be identified

and will be discussed here as basis for future studies to uncover the underlying regulatory pathways.

4.1. Mapping-by-sequencing accelerates gene cloning

In this study, the cloning of the two genes *HvMND* and *HvLAX-A* demonstrated the applicability of mapping-by-sequencing as an approach for rapid gene isolation in barley. Both genes were allocated to genomic regions that strongly differed in recombination frequency and thus in their potential of being identified by classical map-based cloning. The innovative strategy of mapping-by-sequencing resulted in high resolution mapping of the mutant gene loci and in identification of candidate genes.

The implementation of next-generation sequencing (NGS) technologies enabled the production of large volumes of sequence data at moderate costs (Metzker, 2010). This fundamental technological innovation made genome sequencing at broad scale possible for principally all plants (Michael and Jackson, 2013). The technology enables as well new fast-forward genetic approaches for mutation identification (reviewed in Schneeberger and Weigel, 2011). Performing a bulk segregant analysis (BSA, Michelmore et al., 1991) by sequencing phenotypic pools of a segregating population was shown to be a powerful tool to identify recessive mutations without any prior mapping information. The first report of a successful gene cloning by utilizing a combination of BSA and NGS in plants was published in 2009 for the model plant Arabidopsis. Schneeberger and colleagues developed an analysis pipeline called SHOREmap, which enabled *de novo* marker identification, mapping, candidate gene identification and annotation in a single step (Schneeberger et al., 2009). A mutant plant, which was EMS induced in Columbia (Col-0), was crossed to Landsberg erecta (Ler-1) to establish a segregating population. They resequenced a mutant pool of >500 genotypes to about 22x average genome coverage. A high quality SNP marker map between Col-0 and Ler-1 is available (Clark et al., 2007). The software algorithm calculated for each marker the base frequencies of the Col-0 and Ler-1 alleles which results in a peak were the population is homozygous for the Col-0 background. A causal SNP was pinpointed 4 kb away from the highest SNP frequency peak. The second closest non-synonymous SNP was located 200 kb away.

Meanwhile, this method was also successfully implemented for gene cloning in crop plants with completely sequenced genomes. A pipeline called MutMap was applied for characterization of seven agronomical traits in rice (Abe et al., 2012). Pools of 20 mutants of a segregating population were obtained from a backcross of an EMS induced mutant to the non-mutagenized progenitor followed by selfing. Target intervals were defined by SNP indexes larger than 0.9, which represents homozygous mutant allele in at least 90 % of the mapped reads at this position. In all except one case, an average mapping interval of 2.1 Mb with at most 4 non-synonymous changes in protein coding genes was found (Abe et al., 2012). The approach strongly benefited from the reduction of background mutations by backcrossing the artificial mutant to the non-mutagenized parental line whose reference genome sequence was available. Likewise, in maize, BSA sequencing of 32 mutant and 31 wild-type plants obtained from a biparental cross, narrowed down a 2 Mb target interval including 48 genes (Liu et al., 2012). Since they have used an F1 population of an induced mutant in a non-reference background, a further experiment was required to identify the causal gene. The underling gene of the epicuticular wax mutant *glossy 3* (*gl3*) was identified by additional mutant alleles from a transposon tagging forward genetic approach (Liu et al., 2012).

In the present study, a similar approach was applied to identify the causal gene that controls the *mnd* phenotype. The induced mutant (genetic background cv. 'Saale') was crossed with cv. 'Barke' to establish an F2 segregating population. Sequencing of a DNA pool of 18 *mnd* mutant plants from this population was sufficient to narrow the gene locus to chromosome 5H to a confidence interval of around 10 cM (36 Mb). This broad interval was a conservative interpretation based on SNP frequency distribution in order to minimize the risk missing the gene of interest. The observed maximum frequency peak reached only 97 %. The percentage of wild-type reads originated from a wild-type (heterozygous) plant accidentally included in the mutant pool. Even with this problem, it could be shown how precise SNP frequencies represented recombination events within the bulked genotypes. This was validated by converting some of the SNPs into genetic markers. Two SNPs with frequencies of 95% were genetically mapped which allowed to delimit the target interval further to a region of roughly estimated 1.1 Mbp including eleven annotated genes (IBSC, 2012). Another successful application in barley demonstrated that mapping-by-sequencing works even with quantitative traits, such as flowering time (Pankin et al., 2014). A pool of 208 early flowering plants was sequenced from a mapping population, obtained from a cross of a NIL with introgressed *eam.5* mutant allele in Bowman background and Bowman. Although only a maximum of 75

% frequency could be reached, caused by a closely linked second flowering time gene, an 8 Mb interval could be narrowed including around 210 genes. This formed an excellent basis for candidate gene prediction and marker development which enabled finally the gene identification of a gene *HvPHYTOCHROM C* for the *early maturity 5* locus (Pankin et al., 2014).

The success of a mapping-by-sequencing approach is highly correlated with the size of the narrowed target interval and the number of included candidate genes. In theory, a pool of 100 F2 plants contains the information for mapping towards the resolution of 1 cM. The size of the underlying physical distance depends strongly on genome size and the genomic location of the target interval (Mokry et al., 2011). The genetic to physical distance is not evenly distributed along the barley genome. High numbers of recombination events occur only in few small areas within the barley genome which are spaced by large segments where recombination is suppressed to various degree (Kunzel et al., 2000). The ratio of genetic to physical distance is significantly higher towards the barley telomeres (IBSC, 2012). Especially the peri-centromeric regions are known for recombination suppression. According to recent work of the International Barley Sequencing Consortium, around 49% (1.9 Gb) of the barley genome is allocated to the peri-centromeric (+/-5 cM of the genetic centromer) region in the physical map harboring nearly 3500 high confidence genes. The reduction of recombination frequency might impede the exploitation of those genes for genetic analysis (IBSC, 2012). In this particular region, a SNP frequency distribution of a small population would show a broad plateau of SNPs, which would not allow delimiting a target interval to inspect a certain number of candidate genes. This phenomenon had been recently reported in wheat by BSA-seq sequencing bulks of resistant and susceptible plants of a population of ~200 F2 plants segregating for a resistance trait (Ramirez-Gonzalez et al., 2015). It was not possible to define a smooth curve with a peak by visualizing SNP frequencies along the genome. They observed clusters of SNPs with high frequency ratios between resistant and susceptible pools from both chromosome arms spanning the centromere, which can be utilized for marker-assisted breeding or to use them to narrow the resistance gene by further mapping in a larger population (Ramirez-Gonzalez et al., 2015).

Allocation of *HvMND* in a genomic area with a suitable recombination frequency enabled the precise mapping of *HvMND*. A ratio of physical to genetic distance of ~320 kb / cM was observed for the target interval. In contrast, the second gene of this study, *HvLAX-A*, was allocated to a recombination-poor region, previously. The gene had been assigned to the peri-

centromeric region of chromosome 5H RFLP by linkage mapping and introgression line genotyping (Larsson, 1985a; Druka et al., 2011). Since, polymorphic marker information for the *lax-a* mapping population was available, a low resolution mapping was conducted to localize the gene within the defined introgressed segment (Close et al., 2009; Druka et al., 2011). The genetic mapping in a small sized population of 100 F2 plants resulted in co-segregating markers from both chromosome arms spanning the centromere of chromosome 5H. Therefore, the genetic resolution was increased by genetic mapping of 1970 F2 plants. The final target interval of 0.2 cM was defined by eight recombination events. The flanking markers were assigned to the same genetic bin in the physical map. This position has 207 assigned physical BAC contigs which are so far not arranged in a continuous linear order due to the lack of recombination (IBSC, 2012). Thus, further marker development to narrow down the candidate interval would be a laborious task to identify the target flanking contigs from the 207 BAC contigs. For instance, the closest flanking markers were assigned to 248 Mb and 259 Mb on the physical map coordinates. The finally identified *HvLAX-A* harboring BAC_contig_2862 was assigned to position of 203 Mb in the published physical map of barley (Table 7). The individual re-sequencing of recombinants allowed investigating most of the genes within the target interval in a single step. This mapping-by-sequencing of recombinants enabled mapping of variations based on SNP frequencies of all SNPs within the region independently and in a single step. The identification of SNPs which were co-segregating with the phenotype was independent from the current status of the physical map. Therefore, deep resequencing of preselected recombinants is an innovative approach to access genes located in regions with suppressed recombination frequency because it provides additional resolution from the dedicated genetic populations which is lacking from the physical maps resources, which were anchored by small populations only. In this study, known polymorphic markers of the introgression segment of BW457 and the wild-type Bowman could be utilized for genotyping a large population to identify suitable recombinants for sequencing. Alternatively, a low density genome wide mapping implemented by Genotyping-by-Sequencing (GBS, Elshire et al., 2011; Poland et al., 2012) or SNP assays (BOPA or iSelect Druka et al., 2011; Comadran et al., 2012) could be used to select recombinants from a larger population for deep resequencing by exome capture (Mascher et al., 2013c).

Published cloning-by-sequencing studies strongly benefited from the availability of high-quality reference genome sequences for read mapping and variant calling. The rice MutMap and the SHOREmap approach focused only on the mutant bulks, which were induced in a

well-characterized reference background. Variants obtained by read mapping could be traced back to mutations introduced by the mutagen treatment. In case of barley, a comparable high quality reference sequence with large sequence scaffolds is still missing. The current reference is based on a sequence assembly obtained from WGS sequencing which is partially incomplete and is disrupted into small fragments due to the repetitive nature of the genome (IBSC, 2012). In addition to their small size, the sequence contigs of the barley WGS assembly are error prone towards their ends due to low sequence quality of Illumina short read technology and low sequence coverage towards ends of contigs caused by the assembly procedure (IBSC, 2012). Not all genes are represented in the assembly, and highly conserved gene family members may not be resolved in the assembly and be represented in collapsed contigs. Thus, the incompleteness and potential mis-scored variants due to assembly errors can negatively influence the success of candidate identification. Thus, sequencing the bulk of wild-type DNA has a higher importance in barley to exclude variants due to errors in the reference assembly. For both mutants of this study, mutagenesis caused a complete gene deletion, which facilitated the identification of single candidate genes by sequence depth comparisons between mutant and wild-type pools. Since both genes were reported before to regulate similar or related function in other plants, all remaining SNPs which were linked to the target region had not necessarily been considered as potential candidates. Anyway, for *HvMND* a clear target interval was identified for straightforward candidate gene validation. The linkage information of gene loci was a prerequisite for predicting the best candidate gene. For *HvLAX-A*, we identified a number of co-segregating SNPs located in genes. Synteny to the sequenced model genome of Brachypodium was utilized to predict a candidate interval. Assigning these co-segregating SNPs to an improved linearly arranged complete physical map would also allow the direct prediction of a candidate interval in absence of synteny. Systematic evaluation of neighboring genes or regulatory elements would lead to the identification of the causal gene of interest independent from type of mutagenesis.

When bulk sequencing does not allow direct candidate gene prediction due to limited number of genotypes included or lack of sufficient coverage, target interval enrichment followed by deep sequencing was successfully applied in Arabidopsis (Mokry et al., 2011). Two strategies were described: (i) Genomic sequence enrichment for the entire candidate interval (Mokry et al., 2011), (ii) Deep candidate re-sequencing (dCARE) based on the assumption that the causal change occurs with the highest frequency within the bulk (Hartwig et al., 2012). However, both approaches require a robust genome reference sequence. The sequence enrichment for full genomic intervals will be restricted in barley due to the repetitive nature of

the genome. The knowledge of all annotated genes on continuous sequence scaffolds may allow for target gene enrichment for fine mapping of genes within deep DNA bulks. Sequencing deep pools including large numbers of genotypes to moderate sequencing depths to predict a suitable candidate region followed by dCARE of included genes might be a future perspective as soon the complete barley reference sequence will be available.

A case study for reference-free mutation identification was tested in *Arabis alpina* and rice (Nordstrom et al., 2013). The strategy was applied as well only on induced mutations from EMS treatment. To reduce the number of background mutations, a single backcross was applied followed by selfing. The obtained segregating families were used to select homozygous mutant plants to prepare DNA pools for sequencing. Mutant pools of 97 F2 (*floral defective 1, fde1*) and 86 F2 (*flowering 1-1, pep1-1*) plants of *Arabis alpina* were sequenced to 67- and 158-fold genome coverage, respectively. Sequence reads were *de novo* assembled and directly analyzed for homozygous induced variants based on a pair-wise comparison of low copy sequence fragments. A relatively low number (16 and 13) of EMS induced non-synonymous SNPs were identified as candidates for further confirmation experiments (Nordstrom et al., 2013). Since, depending of the generation time of a given species, backcrossing to reduce the number of background mutations can be a time-consuming process, they also proposed that extending the pipeline to additional allelic mutations would be an advantage for mutation identification. Only genes with variations within all allelic mutants would be considered as candidates. (Schneeberger and Weigel, 2011; Nordstrom et al., 2013). Re-analyzing of the data from MutMap study in rice (Abe et al., 2012) delivered three to five candidate EMS mutations for six of the seven included experiments (Nordstrom et al., 2013). Nonetheless, efficient sequence coverage would be challenging for the large barley genome.

Most of the successful mapping-by-sequencing applications were implemented by whole genome shotgun sequencing of relatively small-sized genomes like Arabidopsis and rice. Considering the achieved sequence coverages by sequencing of the pools within the published experiments it is obvious that they are not sufficient to represent each individual genotype by sufficient coverage. For instance, in the Arabidopsis SHOREmap study, a bulk of 500 genotypes was sequenced to a 22-fold coverage (Schneeberger et al., 2009). In rice bulks of 20 mutant plants were resequenced to an average coverage between 12.5 and 24.1 fold for the MutMap approach (Abe et al., 2012). Simulation studies in Arabidopsis highlighted that the coverage alone in contrast to the pool size has only a small effect on delimiting target

intervals (James et al., 2013). Increasing sequencing coverage from 15- to 200-fold for bulks containing 100 plants leads only to a reduction of the final target interval from 500 (+- 310) kb to 419 (+-298) kb. Nevertheless, the large genome size of barley would require forty times (Arabidopsis) or thirteen times (rice) more sequencing data to reach comparable coverages to the successful reported mapping by sequencing studies. Thus, complexity reduction is required to perform NGS BSA in barley and this complexity reduction can be achieved by exome capture resequencing or RNAseq.

In this study, the barley whole genome exome capture for enrichment of 61.6 MB of gene coding sequence was successfully applied for reducing complexity. Sequencing a single capture on a complete single compartment (lane) of the Illumina HiSeq2000 instrument, which generates >30 Gb of sequence information, results in 95% of all targets with at least 10x coverage. But multiplexing only slightly affected the sensitivity (Mascher et al., 2013c). Forty-eight genotypes for *HvMND and* eight genotypes genotypes of the *HvLAX-A* mapping population were re-sequenced in a single lane of the Illumina HiSeq2000 respectively, which resulted in 30 Gb of sequence information for each project. So there might be the risk that some targets of the exome capture are not sufficiently covered by reads. However, this did not hamper the identification of target regions by a mapping-by-sequencing approach.

RNA-seq is dependent on gene expression and might be an alternative approach of genome complexity reduction. However, if the gene of interest is not expressed in the sampled tissue, it cannot be directly indentified as a candidate. However, RNA-seq of phenotypic pools has delivered sequencing data for 75 % of gene models in maize (Liu et al., 2012). The International Barley Sequencing Consortium has detected a common expression of 72–84% of high-confidence genes across samples from eight different tissues (IBSC, 2012). Thus, compared to the exome capture targets, which represents around 75 % of the barley high confidence genes (Mascher et al., 2013c), no direct beneficial advantage or disadvantage can be assumed for one of the two strategies. RNA-seq can be utilized to prioritize candidates in a delimited target interval in case the causal mutation affects gene expression (Liu et al., 2012). Such regulatory mutations, especially if they were located outside of gene coding regions, might be missed by gene space capture (Mokry et al., 2011). Changes in expression levels of genes might require higher sequencing coverage to access also genes with low expression levels. The exome capture assay is balanced by increasing number of capture probes for genes that are captured with lower intensity (Mascher et al., 2013c). A combination of both datasets would deliver the most comprehensive view towards direct candidate identification.

4.2. *HvLAX-A* is involved in the definition of flower whorl identity

The present study reports the map based cloning of the gene underlying the *laxatum-a* phenotype. In *lax-a* florets, a homeotic replacement of the two lodicules by two additional stamenoid organs was observed. Flower organization and such interchange of organ types in flowers is well characterized in model plants *Arabidopsis thaliana* and *Anthirrhinum majus* by the so called ABC model. The model categorizes organs to four concentric whorls of sepal, petal, stamen and carpel (Figure 24a). The floral organ identities of each whorl are specified through regulation by A-, B-, and C-class genes. The identity of sepals is controlled by A-class, of petals by A- and B-class, of stamen by B- and C-class genes and of carpels by C-class genes alone (Coen and Meyerowitz, 1991). This model was extended to the ABCDE model by including the E-class function encoded by *SEPALLATA (SEP)* - like genes which are important for the specification of all floral organs and *AGAMOUS-LIKE 6* (*AGL6*) - like which are important for the ovule formation (Krizek and Fletcher, 2005). The model has also been partially adapted to rice as reviewed by Yoshida & Nagota (Yoshida and Nagato, 2011). Instead of petals and sepals, grass flowers contain lodicules and two bract like organs: the palea and lemma, which enclose the flower (Kellogg, 2001). It is assumed that lemma and palea are modified sepals and that lodicules are homologous to petals (reviewed in Yoshida, 2012; Lombardo and Yoshida, 2015). Changes in barley *lax-a* flowers are restricted to the outer whorls, that is the sepal-like glumes, lemma and palea and the petal-like lodicules, hence their modified appearance could be affected by an A-class gene.

Arabidopsis *bop1/bop2* double mutants displayed an altered pattern of organ formation in flowers (Hepworth et al., 2005). While wild-type Arabidopsis flowers produce flowers with four sepals, four petals, and six stamens, the *bop1/bop2* double mutants showed five organs in the sepal whorl, four in the petal whorl and six or seven stamen organs (Hepworth et al., 2005) (Figure 24). Mutant *lax-a* plants show no supernumerary number of organs. But lodicules, which are supposed to be homologous to petals (reviewed in Yoshida, 2012), were always transformed into stamenoid organs. The additional organs of *bop1/2* double mutants were always located on the abaxial side of the flower causing changes in the positioning (symmetry) of organs, while the adaxial appears like wild-type (Hepworth et al., 2005). The ectopic organs grew more slowly compared to the rest. This was also observed for the stamenoid lodicules of the *lax-a* mutants. The abaxial organs of *bop1/2* mutants in Arabidopsis grew outwards like wings and sometimes fused together (Hepworth et al., 2005). This ectopic growth of the abaxial organs in *bop1/bop2* mutants might be comparable with the

outgrowth of the lemma on the basis of the awn of *lax-a* mutants. Single loss of function *bop2* mutants were described as showing no obvious phenotype (Hepworth et al., 2005; Norberg et al., 2005). One mutant line with a tDNA insertion within the second exon showed severe changes in the inflorescence development. Inflorescences were fused, sometimes fasciated and multiple flowers developed from the same node (Ha et al., 2007). Based on these observations, it has been supposed that *BOP2* plays a key role in inflorescence development, while *BOP1* is more important for leaf development (Khan et al., 2014).
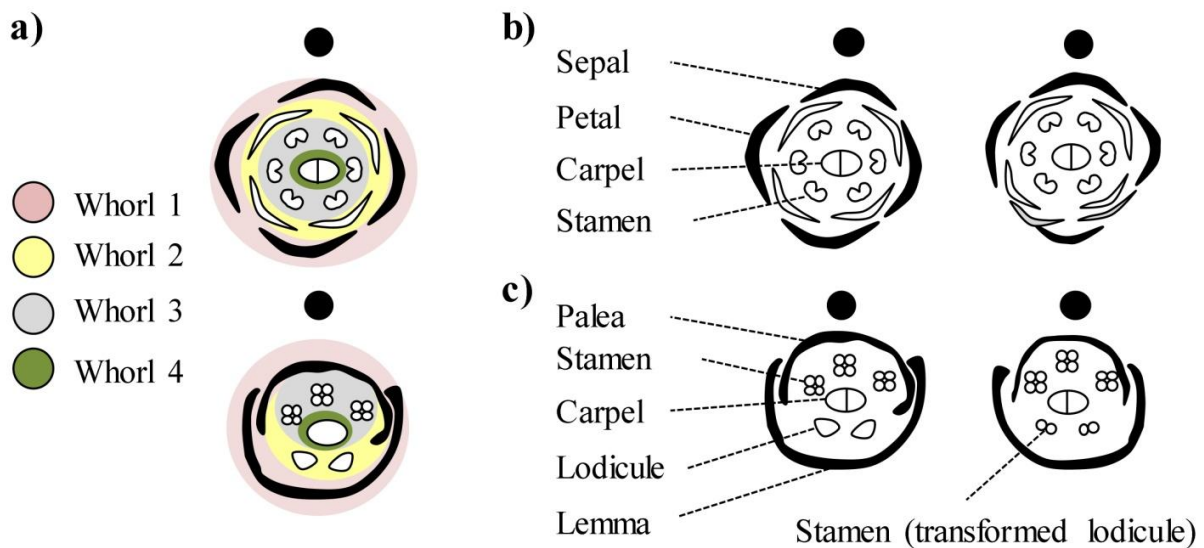
Figure 24: The Arabidopsis genes *BOP1/2* and barley gene *HvLAX* controlling floral whorl identity.
The floral organs are arranged in four concentric whorls according the specific organ type. A schematical drawing of the whorl organisisation for Arabidopsis (upper panel) and barley (lower panel) was provided in a). b) Arabidopsis *bop1/bop2* double mutants (right) exhibit changes in positioning and number of organs on the abaxial site of flowers compared to wild-type (left) flowers (modified picture according to (modified picture according to Hepworth et al., 2005). c) Mutant *lax-a* flowers (right) show a homeotic transformation of the lodicules to stamen. The two additional stamens contain only two migrosporangia in contrast to the four of standard stamens. Black circles marking the stem or floral rachis and indicating the adaxial site.

Independently of this study, the function of the paralogous gene of *HvLAX-A* was recently uncovered by cloning of the gene *Uniculme4* (Tavakol et al., 2015). The *Uniculme4* gene controls the number of tillers, ligule development and proximal-distal leaf patterning (Tavakol et al., 2015). This showed similarities to the observed phenotypes of *bop1/2* mutants in Arabidopsis which also affected proximal regions of leaves. Double mutants showed ectopic leafy outgrowth on petioles (Hepworth et al., 2005; Norberg et al., 2005).

The described phenotypic alterations of the *uniculme4* mutants, i.e. reduction of tiller number and leaf sheath composition, were all confirmed in the analyzed TILLING mutants. Thus, two additional independent mutant alleles were identified in addition to the three alleles reported in the recent cloning study (Tavakol et al., 2015). The difference of importance of the homologous genes in Arabidopsis and barley for regulating inflorescence development and leaf development can be clearly observed in barley mutants. Single *lax-a* mutants showed no obvious leaf defect like observed for *uniculme4* mutants, and vice versa *uniculme4* mutants developed normal spikes carrying florets with three anthers. These observations indicate that in barley both genes might not work in a redundant fashion as implied for Arabidopsis *BOP1/2* genes. However, this hypothesis needs to be confirmed by analyzing the phenotypic characteristics of *lax-a/cul4* double mutants.

4.3. Global gene expression analysis revealed candidate genes putatively regulated by *HvLAX-A*

RNA-sequencing of three developmental stages of immature spike meristems was performed to identify differentially expressed genes between the "wild-type" genotype Bowman and the *lax-a* mutant NIL BW457. A set of 144 differentially expressed high confidence genes were identified. These genes reached a significance level ($p < 0.05$) for differential expression between mutant and wild-type in a least one of the three analyzed tissues. Since several phenotypic similarities were observed between *laxatum-a* mutants and *bop1/bop2* mutants in Arabidopsis, a survey for conserved gene regulation between barley and Arabidopsis was undertaken.

Changes in barley *lax-a* flowers and bop1/2 were restricted to outer two whorls, which might be caused by an A-class gene defect (as described above). APETALA1 *(AP1)* and *LEAFY (LYF)* are key genes for flower initiation and *AP1*, representing an A-class gene, is important for outer whorl organ identity in flowers (reviewed in Krizek and Fletcher, 2005). The role of

AtBOP1/2 in inflorescence meristems is described as a direct and indirect stimulator of *AP1* expression. An interaction with the bZIP transcription factors *TGA8/PERIANTHIA* (*PAN*) allowed *BOP1/2* to directly bind to the promoter of *AP1* (Xu et al., 2010). The indirect activation of *AP1* takes place by promoting *LEAFY* (*LYF*) expression by an interaction with an *EREBP/AP2-type* transcription factor named *PUCHI* (Karim et al., 2009). No indication for a conservation of this regulatory pathway in barley was observed within the expression data of Bowman compared to BW457 (Table 23). Except for the *AP2/ERF* transcription factor *APETALA 2* (A-class gene) all characterized homeotic genes belong to the class of MADS domain (MADS Box) transcription factors (Krizek and Fletcher, 2005). *Lax-a* flowers showed a homeotic conversion of lodicules into stamens. In the stamen primordium stage, a member of the MADS Box transcription factor family of barley showed significant differential expression between Bowman and the *lax-a* NIL (Table 24). In rice, transformation of lodicules into stamens can be induced by ectopic expression of *OsMADS3*, the ortholog of *AGAMOUS (AG)* (Kyozuka and Shimamoto, 2002). However, the homolog of *OsMADS3* in barley did not exhibit changes in transcript level in our expression analysis. The MADS box gene (MLOC_76418.1) identified in the barley expression analysis is a homolog of the so far uncharacterized gene *OsMADS37* showing in BW457 ectopic expression in the stamen primordium stage.

Some of the differentially regulated genes identified here belong to gene classes that have been reported to have an impact during inflorescence differentiation. Three down-regulated homeodomain TFs belonged to the homeodomain-leucine zipper I-class homeobox gene family (Jain et al., 2008). The row-type regulating gene *Vrs1* is a member of this class of genes, demonstrating the potential importance of such genes for regulating spike architecture or plant architecture related traits in general (Komatsuda et al., 2007). Arabidopsis *bop1/2* mutants showed ectopic growth and changes in organ number on the abaxial side of the inflorescence (Hepworth et al., 2005). *YABBY* TFs are important for adaxial - abaxial asymmetry in Arabidopsis (Siegfried et al., 1999). In *lax*-mutants an outgrowth of the lemma and a transformation of petal-like organs into stamens can be observed. This could be affected by an abaxialization fate caused by mis-expression of the *YABBY-like* transcription factor MLOC_66260.1 (Table 24). Next to Arabidopsis, the functional impact of *YABBY* TFs for plant architecture traits has been demonstrated also in the monocot plant rice. Mutations in the *YABBY* Gene *DROOPING LEAF* in rice caused a homeotic transformation of carpels into stamens and failure in midrib development on leaves (Yamaguchi et al., 2004). A further *YABBY* TF member is required for the development of ligulae, auricles and lamina joints (Dai

et al., 2007). It might be possible that the mutation of *HvLAX-A* and *HvCUL4* genes could control expression of two independent members of the *YABBY* family.

The *Auxin response factor* (MLOC_65945.1), the *IAA-amino acid hydrolase* (AK358964) and the *YUCCA-like Flavin monooxygenase* (MLOC_2703.3) were found to be significantly differentially regulated between Bowman and BW457. The latter two genes are key genes involved in the tryptophan-dependent auxin biosynthesis pathway (Zhao, 2008). The Arabidopsis gene *TOPLESS* (*TPL*) directly interacts with *IAA12 amino acid hydrolase* and represses the *auxin response factor 5* (Szemenyei et al., 2008). Studies on the orthologous genes in rice and maize have shown their effects on spike architecture in response to auxin deficiency. The *ABERRANT SPIKELET AND PANICLE1* (*ASP1*) is involved in the regulation of meristem fate (Yoshida et al., 2012). The *asp1* mutants showed elongation of lemma and palea organs, irregular elongation of other organs in the florets as well as compromised auxin signaling (Yoshida et al., 2012). The maize ortholog *RAMOSA ENHANCER LOCUS 2 (REL2)* controls branching of the tassel (Gallavotti et al., 2010). Recently, also the ortholog of *REL2* in barley (*VRS4*) was reported to control row type in barley (Koppolu et al., 2013). Thus, auxin is an important factor for inflorescence development. The *KNOX* gene *KN1* in maize reduces the amount of gibberellic acid by activation of the gene *GA2OX1* (Bolduc and Hake, 2009). Maize *kn1* loss of function mutants have smaller tassels with reduced branching (Bolduc et al., 2012). The expression study in barley revealed a down-regulated *gibberellin 2-oxidase* (AK364775) which may result in higher cytokinin levels in *lax-a* mutants.

Stem cell proliferation and differentiation to organ boundaries is a prerequisite for the development of physically separated organs (Zadnikova and Simon, 2014). Differentiation processes are controlled by hormonal crosstalk and a unique set of transcription factor proteins which locally repress cell proliferation. Loss of function mutants thus may result in organ fusions as well as defects in organ development and phyllotactic patterning (Zadnikova and Simon, 2014).

The transcription factors *KNOTTED1 like homeobox class 1 (KNOX1)* are key factors for balancing between organ initiation and meristem renewal (Bolduc et al., 2012). A direct or indirect interaction with *KNOX class 1 genes* of *BOP-like* genes has been shown in Arabidopsis. *AtBOP2* directly interacted with a complex of *ATH1* to promote expression of the *KNOX1* member *KNOTTED1-LIKE FROM ARABIDOPSIS THALIANA6* (*KNAT6*) to regulate inflorescence elongation (Khan et al., 2012). *AtBOP1/2* were shown to promote the expression of the lateral organ boundary (LOB) domain gene *ASYMMETRIC LEAVES2* (*AS2*)

which form a complex with *ASYMMETRIC LEAVES1* (*AS1*) to suppress *KNOX1* gene expression in leaves (Jun et al., 2010). One *AS2-like* member (MLOC_61156.1) is moderately significant down-regulated in the first analyzed stage of development (Table 24).

The rice homolog of the Arabidopsis LOB domain factor *JAGGED LATERAL ORGANS* (*JAG*) gene called *OPEN BEAK (OPB) / STAMENLESS 1 (SL1)* (Horigome et al., 2009; Xiao et al., 2009) controls, similar to *lax-a,* pleiotropic changes in floral organ identity, reduction in the size of palea/lemma organs and an increase in panicle length. *OPB* functions similar to AtBOP1/2 as a suppressor of *KNOX1-like* genes in organ boundaries. Loss of function of *OPB* leads to up-regulation of *OSH1*, *OSH6*, *OSH15*. However, only *OSH6* shows major changes in expression of young panicle at the spikelet forming stage (Horigome et al., 2009). No significant differentially expressed *KNOX1* was detected in the dataset presented here. The barley homolog of OSH6 is MLOC_57232.2 which is the only *KNOX1-like* gene in our datasets that shows a slightly but not significant up-regulation in the mutant sample of latest developmental stage (Table 24).

The identified candidates must be approached with caution due to the experimental set up and potential technical pitfalls. The correlation analysis of replicates indicated some unexpected clustering of samples which are most likely caused by growing conditions and the mixing of multiple spike meristems per reaction to get sufficient tissue for RNA isolation. Plants were grown under greenhouse conditions, which may be subject to changes in environmental conditions, especially light intensity. The optimal experimental set-up would require plant cultivation in a climate chamber with fully controlled growth conditions. The differential expression analysis was done only by comparing Bowman as wild-type with the NIL BW457 as mutant line. The latter is an introgression line carrying a larger segment of the original donor mutant line which was induced in the genetic background of Bonus. Thus, it cannot be excluded that the observed changes in gene expression were caused by genotype-specific regulation introduced by the Bonus background. Among the 144 differentially expressed genes, only 39 showed a log2 fold change higher than two in one of the three analyzed tissue stages. No log2 cutoff was imposed to minimize the risk of overlooking putative candidate genes. Especially, the expression and function of meristem identity genes are restricted to small organ boundaries on the flanking sites of meristems. A unique set of genes is able to initiate differentiation of organs, like leaves, from meristem cells by temporary regulation of their transcript levels (reviewed in Zadnikova and Simon, 2014). Complete immature spike meristems were taken as plant material, which could result in only small changes of

expression levels by distortion of higher or lower expression levels outside of meristematic active cells. Studying changes in transcript level of genes involved in organ boundary formation would most likely require working with mircodisected meristems in combination with few or single cell RNA-seq analysis (Wu et al., 2014).

The genes with a low number of reads per gene should be treated with caution. Small changes in read number of replicates could cause large changes in log2-fold change (Anders and Huber, 2010). A higher sequencing coverage would be required to confidently estimate the expression of these low abundant genes. As an alternative approach, the expression of these genes could be measured by quantitative RT-PCR in relation to a constitutively expressed gene. The high number of observed weakly expressed genes within the set of differentially expressed genes is a result of using DEseq (Anders and Huber, 2010) software package for the test of differential expression. This method works based on normalized mean values of sequence count data of mapped reads per gene. This method replaces formerly used normalization based on reads per kilo base of transcript per million mapped reads (RPKM) which normalizes for gene length and total mapped reads within a sample (Mortazavi et al., 2008; Garber et al., 2011). The representation of a gene is dependent of the expression level of all other genes. Reads of highly expressed genes mask especially counts of lowly expressed genes (Rapaport et al., 2013).

Table 24: Overview of candidate genes putatively regulated by *HvLAX-A*

| HC_gene[1] | normalized read numbers wild-type[2] | | | normalized read numbers mutant[2] | | | log2FC_gl | log2FC_st | log2FC_cd | adj p=<0.05 | | | annotation[1] |
	mean_bowman_gl	mean_bowman_st	mean_bowman_cd	mean_lax_gl	mean_lax_st	mean_lax_fd				gl	st	cd	
MLOC_19415.1	24.41	29.77	50.24	1.61	0.61	4.11	-3.92	-5.60	-3.61			X	Gibberellin receptor GID1L2
AK358964	235.98	287.59	176.03	32.44	42.36	26.79	-2.86	-2.76	-2.72	X	X	X	IAA-amino acid hydrolase ILR1
MLOC_61156.1	142.97	147.24	192.40	74.40	120.37	129.65	-0.94	-0.29	-0.57	X			LOB domain-containing protein
MLOC_65945.1	61.01	134.48	359.32	32.62	58.20	198.16	-0.90	-1.21	-0.86		X		Auxin response factor
MLOC_74058.1	11.69	13.21	50.81	11.75	9.72	13.09	0.01	-0.44	-1.96			X	auxin response factor 4
MLOC_2703.3	25.56	54.79	213.89	23.40	16.49	153.11	-0.13	-1.73	-0.48		X		Flavin monooxygenase-like protein
AK364775	71.50	271.84	797.45	31.15	186.31	559.07	-1.20	-0.55	-0.51	X			Gibberellin 2-oxidase
MLOC_76418.1	91.29	38.41	419.66	202.81	211.34	425.58	1.15	2.46	0.02		X		MADS-box transcription factor 8
MLOC_66260.1	192.30	282.24	664.11	72.03	207.60	530.95	-1.42	-0.44	-0.32	X			YABBY protein
AK354063	499.36	431.26	1133.36	227.34	649.67	905.87	-1.14	0.59	-0.32	X			Homeobox-leucine zipper protein
MLOC_57232.2	4672.89	4619.75	2550.90	5027.52	4455.86	3293.46	0.11	-0.05	0.37	X			Homeobox protein knotted-1

[1] High confidence gene model with predicted annotation (IBSC, 2012); [2] normalized read numbers of DEseq analysis for the developmental stages of gl = glumen primordia, st = stamen primordia and cd = complete dveloped spikes

Color scale

| -3.00 | 0.00 | 3.00 |

4.4. *BOP-like* gene family organization in Barley

Two paralogous genes with high similarity in gene structure, but pronounced differences in functionality were addressed in this study. Whereas the gene *HvLAX-A* is involved in spike development, a second closely related gene *HvCUL4* regulates tillering and leaf sheath development in barley. Gene duplication was an important driver for morphological evolution of plants (reviewed in Rensing, 2014). In the present case, if the multiple copies of the *BOP-like* genes originated by a gene duplication event, the distinct functions of the paralogous genes can be a result of neofunctionalization or subfunctionalisation. Neofunctionalization occurs if the paralogous gene fulfills a novel beneficial function, which was preserved by natural selection (Lynch and Conery, 2000). The two other categories are subfunctionalisation and nonfunctionalization. Subfunctionalisation describes that both copies divide the original function and nonfunctionalisation refers to the loss of one copy after gene duplication by a silencing mutation (Lynch and Conery, 2000). Since Arabidopsis BOP1/2 regulate leaf and inflorescence development, the modified regulation of traits in barley indicates a subfunctionalization of both genes in barley. However, analyzing *lax-a/cul4* double mutants would be required to completely exclude functional redundancy. Tavakol and coworkers proposed, based on higher sequence conservation between *BOP1* and *BOP2* in Arabidopsis (80%) and Soybean (82%) compared to *HvLAX-A* and *HvCUL4* (58%), that the dicot *BOP* genes might have been derived from a more recent duplication event and thus show functional redundancy (Tavakol et al., 2015). More distinct highly supported clades were observed for monocots within a phylogenetic analysis. Thus, they postulated a separate evolution of the monocot BOP genes from a more ancient duplication event resulting in functional divergence (Tavakol et al., 2015). This was partially supported by the dicot pea *coch* and *Medicago truncatula noot* mutants which exhibit changes in both, leaf and flower development (Couzigou et al., 2012). Mutants show changes in flower symmetry connected with additional floral organ formation and a reduced stipule development (Couzigou et al., 2012). However, so far there is only one copy known for pea, which might be affected by the lack of sequence data. M. *truncatula* has two BOP-like genes, but lacks in knowledge about the function of the second copy as well the analysis of double mutants.

In this study, a more comprehensive phylogenetic analysis was performed. All analyzed BOP-like genes clustered in two separate clades in the phylogenetic tree for monocots as well as for dicots. The clustering in one clade of the reported study in the *Cul4* cloning manuscript is a

result of the smaller number of analyzed members of dicot taxa. Based on the analysis of the present study, there is no indication for a difference in evolution of monocot and dicot *BOP-like* genes which would allow explaining functional redundancy or subfunctionalisation.

Barley contains two members of *BOP-like* genes. Other cereals like maize, rice, *Sorghum* contain three copies of *BOP-like* genes. The two or three gene copy pattern was present as well for dicot families. Interestingly, if multiple members of a plant family have three copies, the third copy of the different species cluster always together in the same subclade. For instance, the clade containing *HvCUL4* contains only a single copy in all analyzed monocots. The single copies of *HvCUL4-like* genes in other monocots might be good candidates to study similar function in leaf development and tiller formation. The subclade containing *HvLAX-A* includes always two copies of BOP-like genes except for barley and *Brachypodium*. The additional highly related *HvLAX-A-like* copies in rice, sorghum and maize genes might have a redundant function. It has been proposed that the pairing of BTB-ankyrin domains occurred within a charophyte algal lineage during the transition to land plants (Lewis and McCourt, 2004; Khan et al., 2014). It was shown by a global analysis of BOP genes among the plant kingdom that chlorophyte green algae *Volvox carteri* and *Chlamydomonas reinhardtii* have no *BOP-1ike* genes, whereas the mosses Physcomitrella patens (3 copies) and Selaginella moellendorffii (2 copies) have multiple copies (Khan et al., 2014). Since multiple copies of the *BOP-like* genes were already present in the first land plants, it is not possible to conclude if the multiple gene copies observed in higher land plants are a result of later gene duplication events or are just differentiated ancient *BOP-like* copies.

Support for the latter hypothesis can be obtained by considering the evolutionary relationship of the barley and rice genome. In rice, more than 60 % of the genome could be assigned to duplicated genome segments which originate from an ancient whole-genome duplication before the divergence of grasses (Yu et al., 2005). The high level of synteny to rice has been used to predict putative duplicated segments of barley (Stein et al., 2007; Thiel et al., 2009). The *HvCUL4* ortholog in rice LOC_Os01g72020.1 is located on chromosome Os1. *HvLAX-A* shows high sequence similarity to LOC_Os12g04410.1 and LOC_Os11g04600.1. The two copies in rice might be resulting from a known recent duplication event in rice between Os11 and Os12 (Yu et al., 2005). Nevertheless, chromosome Os01 shows no duplicated segments to chromosome Os11 or Os12 in rice (Thiel et al., 2009). Thus, it is more likely that *HvLAX-A* and *HvCUL4* are derived versions of already existing *BOP-1ike* genes already available during the transition to land plants.

4.5. Natural diversity analysis revealed distinct haplotype structure

In addition to the obvious functional differences, a distinct haplotype structure between *HvLAX-A* and *HvCUL4* was detected by analyzing the natural diversity in wild and domesticated barley lines. *HvCul4* gene showed a diverse haplotype structure, while the locus *HvLAX-A* gene had a very low level of diversity even in the wild material. None of the identified polymorphisms within the coding sequence of *HvLAX-A* and *HvCUL4* led to a complete truncation of the protein through premature translation stop codons or splice site alterations. In addition, most of the detected non-synonymous SNPs were located outside of the functional domains and the majority of mutated sites were restricted to the C-terminal end of the protein. Considering the high number of tolerated amino acid exchanges without phenotypic alterations in both genes that was observed in the TILLING analysis, it is rather unlikely that these non-synonymous changes have a functional impact on the phenotype.

Mutations with significant positive effects on phenotypic traits can quickly become fixed by positive selection. The other way around, genes that might be highly sensitive for protein changes were fixed to ensure the functionality (negative selection) (Nielsen, 2005). Loss of functional mutants of *HvLAX-A* did not significantly influence the plant fitness. Thus, it is rather unlikely that the low natural diversity of *HvLAX-A* is affected through a selective process. *HvLAX-A* is located within a region of suppressed recombination close to the centromere. Thus, the reduced diversity of *HvLAX-A* was maybe caused by linkage to another important gene with a high biological relevance which led to fixation already in the wild barley genepool. In addition, it was detected that nucleotide diversity was correlated with recombination rate and thus is reduced in the centromeric region (Begun and Aquadro, 1992). However, the flanking genomic region were not analyzed which might be required to confirm one of the hypothesizes.

Under greenhouse cultivation, the changes in leaf sheath composition and missing ligulae have a strong effect on fitness of the TILLING mutant plants for *HvCUL4*. For instance, only five out of ten identified homozygous TILLING mutants with premature stop codons were able to develop a single tiller with seeds. The others struggled with emergence of new leaves which led to an extremely curly growth and consequently the termination of plant development. This was also recently described for the *Uniculme4* cloning project (Tavakol et al., 2015) and thus may not indicate any effect of background mutations introduced by the high mutation density of EMS treatment of the TILLING population. This might be the

reason that non-synonymous SNPs within the ORF are dominantly located towards the C-terminal end, which might not impact functionality of the gene.

The Tajima's D test (Tajima, 1989) was performed to check for neutrality of DNA polymorphism in order to test whether the genes were involved in a selection process during barley cultivation. Tajima's D test relies on a pair wise comparison of the observed number of variations against an estimated number of expected mutations based on analyzed segregating sites within a population of a certain size (Tajima, 1989; Biswas and Akey, 2006). The neutral mutation hypothesis is zero which refers to observed nucleotide diversity is equal to the expected estimated mutation rate. All except one of the observed Tajima values for *HvLAX* and *HvCUL4* have a negative value. Negative values of Tajima's D are observed by segregating sites with low frequencies (rare alleles) within the population and can be a result of population expansion or positive selection (Biswas and Akey, 2006). The Tajima test for the adapted accessions *HvCUL4* resulted in a positive Tajima's D value in contrast to a negative value for wild material. Positive values are a result of intermediate frequency alleles which can be caused by a population bottleneck or balancing selection (Biswas and Akey, 2006). Nucleotide diversity is kept even or even slightly increased within the adapted material for *HvCUL4* compared to the wild barley material. Thus the major haplotypes were maintained, while the majority of rare alleles were lost in domesticated material. The Tajima's D value of both genes did not reach a significance level and thus it is rather likely that the differences at nucleotide diversity level are caused by random bottleneck effects of the population.

4.6. Loss of function of Cytochrome P450 protein controls plastochron in barley *mnd* mutants

The ideotype of the barley crop plant has been defined to develop a limited number of small and erect leaves (Donald, 1968). Larger leaves have a higher affinity to be floppy and might overlit the lower leaves. Thus, the canopy structure has a direct impact on crop photosynthesis and is in addition to the leaf size mainly defined by the spatial distribution (phyllotaxis) and the regular timing (plastochron) of leaf initiation. In this study, a CYTOCHROME P450 SUBFAMILY 78A (CYP78A) protein encoded by a gene allocated on chromosome 5H was identified to be responsible for regulating the plastochron and overall plant architecture in barley.

The nearly isogenic lines of cv. Bowman indicate the presence of multiple mutant loci controlling *mnd* mutant phenotypes. This trait was assigned to chromosomes 4HL and 3HL in the introgression lines BW518 and BW519, respectively, and another two mnd loci remained without chromosomal assignment (Druka et al., 2011). In Arabidopsis, six CYP78A proteins are present (Mizutani and Ohta, 2010). Two of these have been reported to control the plastochron. Mutant *cyp78a5* plants showed a higher leaf initiation rate and a reduced size of leaves, sepals and petals (Zondlo and Irish, 1999). Wang and colleagues (2008) have proposed a redundant function of one paralog *CYP78A7* in regulating leaf initiation. While single mutants did not show any obvious phenotype, the cyp78a5/cyp78a7 double mutants had compact rosette leaves and showed increased leaf initiation (Wang et al., 2008). The CYP78A enzymes have been described to be involved in controlling several plant growth and physiological pathways like endosperm development (Nagasawa et al., 2013), seed size (Adamski et al., 2009), organ size (Anastasiou et al., 2007), reproductive development (Sotelo-Silveira et al., 2013) or fatty acid hydroxylation (Kai et al., 2009). Other members of the *CYP78A5* family of barley had their most abundant transcript levels in different tissues than *HvMND* according to eight analyzed tissues (IBSC, 2012). Hence, they might be involved in other pathways than *HvMND*. In addition to the CYP78A family, two further protein classes have been reported causing shortened plastochron when mutated in plants. The maize TERMINAL EAR 1 (TE1, Veit et al., 1998) and its rice ortholog PLASTOCHRON 2 (PLA2, Kawakatsu et al., 2006) belong to the MEI2-like RNA binding proteins. The rice genes *PLASTOCHRON3* (PLA3, Kawakatsu et al., 2009) and *ALTERED MERISTEM PROGRAMM1* (AMP1, Helliwell et al., 2001) in Arabidopsis encode a GLUTAMATE CARBOXYPEPTIDASE. The distinct classes of proteins involved in plastochron definition are indicative of a complexity of underlying regulatory pathways (Kawakatsu et al., 2006). While the *amp1* mutant shows up-regulation of *CYP78A5* in Arabidopsis, its ortholog in rice (*pla3*) does not affect expression levels of other plastochron genes. No changes in transcript accumulation was detected as well for the other *PLA* genes in *pla1* and *pla2* mutants (Kawakatsu et al., 2006; Kawakatsu et al., 2009). Analyzing *pla1/pla2* double mutants revealed more severe phenotypes compared to single mutants. The combination of *pla1/pla3* and *pla2/pla3* mutants showed a strongly retarded growth and a defective SAM maintenance (Kawakatsu et al., 2009). Thus, it was proposed that the three proteins are regulated independently, but may function redundantly (Kawakatsu et al., 2006; Kawakatsu et al., 2009). Members of these proteins are good candidates to investigate for the identification of further *mnd* alleles in barley. Among the 37 *mnd* accessions from NordGen, seven accessions

did not show any alteration of gene *HvMND*, thus remain as alternative alleles for any of the other unlinked plastochron regulating loci. No functional validation test was performed for the mutants with amino acid changes in *HvMND*. It is not possible to conclude that the *mnd* phenotype is caused by these changes in the protein sequence or a result of a mutation in another locus. A test of allelism to another mutant would be required to show if these changes in the protein sequence have influence on the gene function. Thus, those plants remain as putative candidates for other *mnd* regulating loci.

The barley *mnd* mutants show a strong phenotypic similarity to the *plastochron mutants* (*pla1, pla2, pla3*) mutants in rice. Mutant *mnd* plants exhibit smaller leaves compared to the wild-type. The reduction in length and width of the leaf blades of rice *pla* mutants results from a reduced number of cells caused by precocious maturation of leaves (Kawakatsu et al., 2006; Kawakatsu et al., 2009). A higher number of cell division in the shoot apical meristem could be observed in the rice *pla* mutants which might be caused by the rapid production of leaf primordia (Kawakatsu et al., 2006; Kawakatsu et al., 2009). Differences in the extent of internode elongation were observed between the three rice plastochron mutants. The *pla1* mutant plants were much taller than *pla2* and *pla3* mutant plants (Kawakatsu et al., 2006; Kawakatsu et al., 2009). The different mutant alleles within the Bowman NILs also exhibited different levels of internode elongation, which reinforces the hypothesis of independent mutant alleles from distinct protein classes (Figure 25). Mutant *mnd* plants showed frequently ectopic outgrowth of secondary tillers from the basis of the spikes. Inflorescences of the rice plastochron mutants produce several ectopic shoots instead of branches in the panicle (Miyoshi et al., 2004; Kawakatsu et al., 2006; Kawakatsu et al., 2009). Faster leaf initiation and reduced plant organ size was observed among all plastochron mutants. The ectopic shoot branching was exclusively observed in the rice mutants and is thus most likely monocot specific (Kawakatsu et al., 2009). Nonetheless, all plastochron mutants showed pleiotropic growth defects to different extent on various morphological traits which might refer to the distinct specified plant architecture among the plant kingdom.
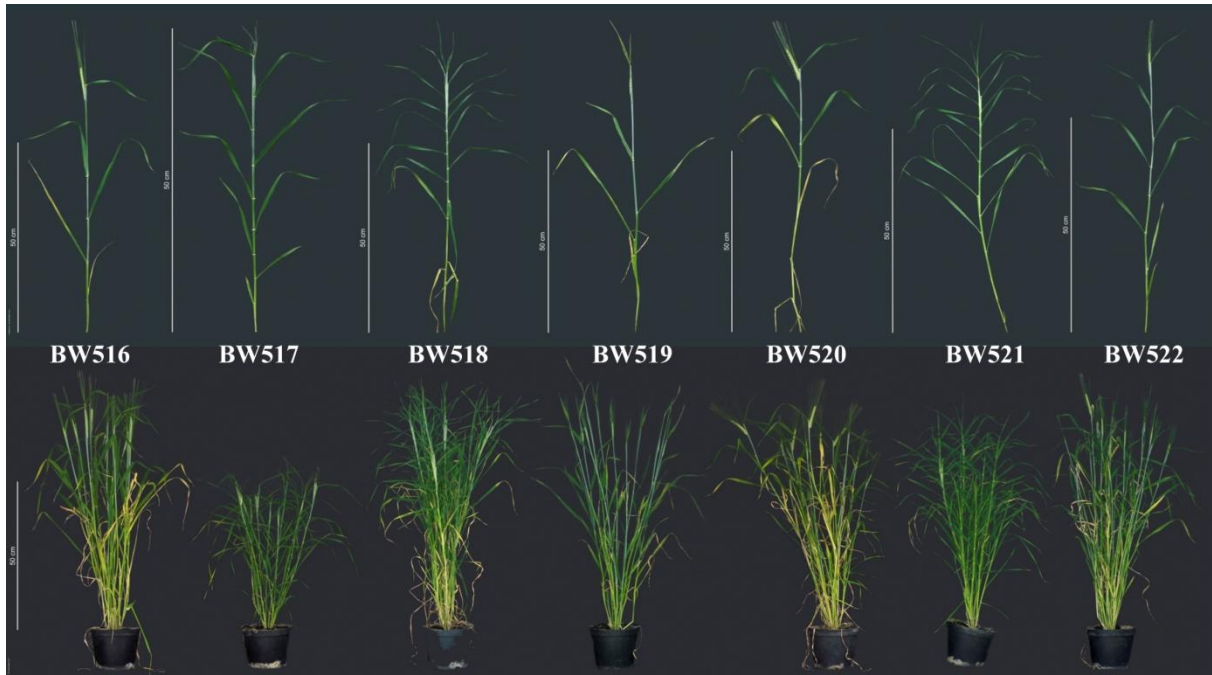
Figure 25: Plant growth of Bowman near isogenic lines with introgressed *mnd* alleles.
Differences in the overall plant growth (lower level) and single tiller view highlights differences of internode elongation between the Bowman NILs.

Little is known about the regulating pathways of the identified proteins. The Arabidopsis *CYP78A5* gene has been shown to be expressed on flanking sites of the shoot apical and inflorescence meristem (Zondlo and Irish, 1999). Alterations in expression levels of *CYP78A5* was observed in *shoot meristemless* (*stm*), *zwille* (*zll*) and *argonaute* (*ago*) mutants - genes that have been reported as key factors of shoot apical meristem specification (Zondlo and Irish, 1999). Thus, *CYP78A5* might act downstream of these genes involved in organ meristem differentiation.

Plant hormones play a significant role in the regulation of growth and development. Since all plastochron mutants showed pleiotropic growth changes, it is likely that the hormonal balance might be changed in these mutants. The Arabidopsis *amp1* mutant shows six-fold increased levels of cytokinin (Chaudhury et al., 1993). The level of cytokinin was also increased in *pla1* and *pla3* but not to the same extent (Kawakatsu et al., 2009). ABA content is decreased significantly in all rice *pla* mutants and the indole-3-acetic acid (IAA) reduced to about 2/3 (Kawakatsu et al., 2009). Thus, the hormonal homeostasis is disturbed in all three rice mutants. Gibberellins (GA) are known to have a fundamental effect on plant height and organ size. The *pla1* and *pla2* mutants show no response to GA treatment. The expression levels of *PLA1* and *PLA2* genes in wild-type show strong response to GA treatment (Mimura et al.,

2012). Thus, it was proposed that *PLA1* and *PLA2* functions downstream of the GA signaling pathway (Mimura et al., 2012; Mimura and Itoh, 2014).

5. Outlook

This study revealed the two barley genes *MANY NODED DWARF* (*MND*) and *LAXATUM-A* (*LAX-A*), which are important regulators for plant architecture development. The mutant gene *mnd* influenced the 'plastochron' resulting in higher number of leaves, nodes and internodes compared to wild-type plants. Mutations in the second gene *LAX-A* caused a homoeotic change in the barley flower by transforming the two lodicules in two stamenoid organs. In addition *lax-a* mutants exhibited pleitropically a broadened awn base, growth change of lemma and palea as well as an extended rachis of the inflorescence. The identified genes could be confirmed by independent mutant alleles in historical mutant collections and by TILLING, showing the same phenotype as the original *mnd* and *lax-a* mutants.

The *LAX-A* gene was identified to be homologous to the *Arabidopsis thaliana BLADE-ON-PETIOLE-1/2* (*BOP1/2*) genes, which are transcriptional regulators containing a BTB/POZ domain and ankyrin repeats and working as transcriptional regulators controlling leaf morphogenesis and flower organization in *Arabidopsis.* Like in Arabidopsis, a second closely related paralog of *LAX-A* could be identified on barley chromosome 3H. Mutant analysis revealed the influence of this gene on barley shoot branching since two mutants from a TILLING screen showed low tiller number phenotype, which turned out to be the *uniculme4* mutation which was already cloned independently of this study. Analysis of natural diversity of both genes indicated distinct patterns of diversity of *HvLAX-A* and *HvCUL4*. However, no statistical support allowed concluding if these differences were caused by selection or just by drift. This might require resequencing more plants than in the current analysis. In addition, analyzing further gene loci around *HvLAX-A* and *HvCUL4* or a genome wide diversity analysis would be required to test whether these differences are caused by different genomic locations of both genes.

The global expression analysis of developing inflorescence meristems delivered a number of putative candidates with different expression levels between wild-type Bowman and *lax-a* mutant NIL BW457. The identified candidate genes could be involved in spike architecture regulation but this requires further experiments. An analysis of independent *lax-a* mutants could confirm that the regulation was introduced by the gene itself or mediated by the genetic

background of the remained Bonus segment within the introgression line used for cloning. Performing RNA-seq with an independent mutant allele would allow ruling out background mutations introduced by the Bonus background of the NIL. The alternative approach would be the systematic investigation of expression levels of the identified differentially expressed genes by quantitative PCR between Bowman and BW457 in reference to a constitutively expressed gene. Confirmed candidates could be screened by TILLING to indentify mutations that would allow studying the function of the gene. Since the *LAX-A* gene contained two protein-protein binding domains a Chromatin Immunoprecipitation Sequencing (CHIP-seq) experiment could be used to identify interacting partners and binding sites to regulatory elements of genes. These genes can be evaluated for changes of expression pattern with the established expression data. The success of this approach was demonstrated for the rice gene *IDEAL PLANT ARCHITECTURE 1* (*IPA1*) which was shown to regulate directly the expression of *DENSE AND ERECT PANICLE 1* (*DEP1*) (Lu et al., 2013). Performing an mRNA in situ hybridization experiment is required to confirm the meristematic activity in the organ boundaries of inflorescence meristems like known from Arabidopsis. This could be extended to promising candidate genes from the expression analysis to confirm co-expression as hint for putative interacting activity.

The innovative sequence assisted mapping of *HvMND* and *HvLAX-A* led to the direct identification of promising candidate genes. The allocation in different genomic areas with highly distinct recombination frequencies provided insights into required experimental setups for further gene cloning projects in barley. A strategy to access genes in regions with suppressed recombination frequency was demonstrated by deep resequencing of preselected recombinants as a prerequisite under the currently available barley genomic infrastructure. The gathered experiences in mapping-by-sequencing in barley can be applied to systematic investigation of further plant architecture traits from mutant collections. A large percentage of the stored mutants within the Nordic Genetic Research Center (NordGen) or at IPK are induced by neutron or radiation treatment which is known to frequently leading to larger lesions in the genome (Nelson et al., 1994; Li et al., 2001). For instance, among the Bowman near isogenic lines, another 280 mutants alleles induced by x-ray, gamma-ray or neutrons are backcrossed in Bowman isogenic background. Most of these NILs are backcrossed for several generations which would allow to systematically perform mapping-by-sequencing approaches according to the genomic location. These mutants belong to 60 different morphological traits. The identification of the causative mutations will deliver first insights into regulating pathways.

6. Summary

Optimizing plant architecture is one major target towards increasing yield potential of crop plants. To identify underlying genes regulating plant growth is one major requirement to understand and intervene towards breeding plants with optimal architecture that allows the maximum usage of the available agriculture production area. Induced mutants are excellent tools for studying function of genes. This study focused on the identification of two barley genes *MANY NODED DWARF* (*MND*) and *LAXATUM-A* (*LAX-A*) which are important regulators for plant architecture development. Mutant gene *mnd* plants showed faster leaf initiation (shortened plastochron), smaller and erect leafs as well as a dwarfism growth habit under field conditions compared to wild-type plants. The *lax-a* mutant plants exhibited pleiotropic changes in inflorescence architecture compared to the wild-type plants. Spike length was increased due to longer rachis internode lengths, grains were exposed due to changes in growth of lemma and palea, a broader awn base and a homeotic transformation of the lodicules to stamenoid organs. The map based cloning approach became standard technique to indentify underlying genes with the causal mutation leading to the observed morphological or physiological traits. However, this approach can be a laborious task depending on the recombination frequency around the target loci. An innovative approach of mapping-by-sequencing was described for model plants which allowed identifying causal mutated genes within few working days. This approach combines a deep sequencing of phenotypic pools of mutant and wild-type plants with Next Generation Sequencing (Bulked segregant analysis by sequencing, BSA-seq) and is enabled by the small genome size and complete genome reference. Recently, reduced complexity sequencing by capturing gene space with the whole barley exome capture enabled cost efficient sequencing and a physically map with assigned sequence information became available for barley. Thus, the study aimed to adapt this mapping-by-sequencing strategy to barley. The addressed genes were allocated to genomic areas with highly distinct recombination frequencies and thus required different experimental setups. The gene *HvMND* was identified within a small population of 100 F2 plants enabled through a location in a recombination hotspot. The gene *HvLAX-A* was previously assigned to the centromere of Chromosome 5H, an area that is known for recombination suppression. Therefore, the gene was genetically mapped in 1970 F2 plants to a 0,2 cM interval to increase the genetic resolution. The interval still included approximately 200 gene models based on synteny information to other grasses and targets a region in the physical map which was not resolved in a linear order of contigs. The Closest linked

recombinant plants, which defined the small genetic interval, were applied for exome capture based mapping-by-sequencing. For both genes, promising candidate genes were directly identified and confirmed by a series of independent mutant alleles, found in historical mutant collections and by TILLING (targeting local lesions in genomes), which showed the same phenotype as the original *mnd* and *lax-a* mutants. Thus the identified genes could be confirmed to represent the causal genes underlying the mutant phenotype in both cases. The gathered information of mapping-by-sequencing can be used in the future to plan fast forward screens of mutant collections. The increased number of known regulating genes would be an excellent starting point for unraveling complex regulations in plant development. The gene *HvMND* was uncovered by sequence similarity search to represent a member of the Cytochrome P450 class 78A gene family. The *HvLAX-A* gene was identified to be homologous to the *Arabidopsis thaliana BLADE-ON-PETIOLE-1/2* (*BOP1/2*) genes, which are transcriptional regulators that contain a BTB/POZ and ankyrin repeats domain and control together leaf morphogenesis and floral organ abscission in *Arabidopsis*. Like in *Arabidopsis*, a second closely related paralog of *LAX-A* could be identified on barley chromosome 3H. Mutant analysis revealed the influence of this gene on barley shoot branching, since two mutants from a TILLING screen showed a low tiller phenotype. In addition, mutant plants showed ectopic outgrowth of the auricles and were missing the ligulae organ. Based on personal communication with the group of Dr. Laura Rossini, University of Milano, Italy, it was confirmed that this *lax-a* paralog is the causal gene of the *uniculme4* mutant in barley, which is known to exhibit the observed traits. Analysis of natural diversity of both genes indicated distinct patterns of diversity after resequencing the coding sequence of each gene in 305 barley accessions. *HvLAX-A* showed a very low diversity even in wild barley accessions whereas *HvCUL4* showed a strong diversity pattern. However, whether these were a consequence of selection or drift or if they just a result of their different positions in the genome will require further genome-wide studies. Based on the conserved domain structure of two protein-protein binding domains and the known role of as transcriptional co-activators in Arabidopsis it was aimed to identify genes which were putatively regulated by *HvLAX-A*. A quantitative global gene expression analysis applied by RNA sequencing was performed on three independent stages of early spike development to reveal candidate genes involved in *lax-a* specific regulatory gene networks leading to the observed phenotype. Eleven strong candidate genes with differential expression between wild-type Bowman and mutant NIL BW457 were identified and discussed as basis for further studies to uncover the underlying regulatory pathways.

7. Zusammenfassung

Die Optimierung der Pflanzenarchitektur ist ein Ansatz um zukünftig höhere Erträge von Kulturpflanzen zu erzielen. Die Identifikation von Genen, welche für die Merkmalsausprägung des Pflanzenwachstums verantwortlich sind, ist eine Grundvoraussetzung um eine gezielte Züchtung von Pflanzen und somit eine optimale Nutzung der landwirtschaftlich verfügbaren Fläche zu ermöglichen. Induzierte Veränderungen durch Mutagenisierung sind eine exzellente Möglichkeit Genfunktionen zu analysieren. Diese Arbeit beschreibt die Identifikation der Gerstengene *MANY NODED DWARF* (*MND*) und *LAXATUM-A* (*LAX-A*), welche wichtige Regulatoren für die Pflanzenarchitektur sind. Mutanten *mnd* Pflanzen zeigen, im Gegensatz zu herkömmlichen Wild-typ Gersten Pflanzen eine schnellere Blattanlage (verkürztes „Plastochron"), haben schmalere aufrecht stehende Blätter und weisen ein verzwergtes Wachstum unter Feldbedingungen auf. Die *lax-a* Mutanten haben hingegen pleiotrophe Veränderungen in der Ährenentwicklung im Vergleich zu Wild-typ Pflanzen. Die Mutanten haben eine weniger dichte Ähre als Resultat einer Streckung der Internodienlänge der Spindelachsen (Rachis), freiliegen Karyopsen aufgrund einer veränderten Struktur von Palea und Lemma, eine verbreiterte Grannenbasis sowie eine homeotischen Austausch der Lodiculae zu zusätzlichen Antheren. Die Kartengestützte Isolierung („Map based cloning") wurde erfolgreich als Standartmethode zur Genidentifikation in Gerste etabliert. Jedoch kann sich dieser Ansatz je nach Rekombinationsrate in der Zielregion zu einer arbeitsaufwendigen Prozedur entwickeln. Ein innovativer Ansatz namens „mapping (cloning)-by-sequencing" ermöglicht innerhalb kürzester Zeit die Identifikation von Genen, welche für die Merkmalsausprägung verantwortlich sind. Dieses Verfahren basiert auf einer tiefen Sequenzierung von phänotypischen DNA Pools mit der „Next Generation Sequencing" Technologie (Bulk Segregant Analyse durch Sequenzierung, BSA-seq) und profitiert dabei von der vergleichsweise geringen Genomgröße und verfügbaren Genom Referenzsequenz.

Ein Verfahren zur Reduktion der Genomkomplexität durch eine Anreicherung gencodierender Abschnitte der DNA („barley whole genome exome caputre") vor der Sequenzierung wurde erfolgreich für Gerste etabliert. Die kürzlich veröffentlichte physikalische Karte liefert erstmals entlang der Gerstenchromosomen verankerte Sequenzinformationen und kann daher als Referenz genutzt werden. Auf Grundlage dessen wurde im Rahmen dieser Arbeit angestrebt, die Kartierung durch Sequenzierung (mapping-by-sequencing) in Gerste zu

etablieren. Beide Gene, welche in dieser Arbeit adressiert wurden, sind in Regionen mit stark unterschiedlichen Rekombinationsraten lokalisiert und verlangten verschiedene experimentelle Ansätze. Das *HvMND* Gen konnte in einer kleinen Population von 100 F2 Pflanzen identifiziert werden, da es in einer Region mit erhöhter Rekombinationshäufigkeit („Hotspot") liegt. Das Gen *HvLAX-A* war bereits im Vorfeld in der Zentromerregion auf Chromosom 5H kartiert, welche für eine stark verminderte Rekombinationshäufigkeit bekannt ist. Infolge dessen wurde der Genort zunächst in 1970 Pflanzen in einem 0,2 cM Intervall genetisch kartiert. Dieses beinhaltete, auf der Basis von Syntänie Information von anderen bereits vollständig sequenzierten Gräsern, immer noch ca. 200 Gene. Zudem war die zugehörige physische Region nicht eingrenzbar und befand sich in einer Region in der, aufgrund der geringen Rekombinationsrate, bis heute keine lineare Anordnung physischer Contigs möglich ist. Die Rekombinanten Pflanzen, welche das kleine genetische Intervall begrenzen, wurden für einen Mapping-by-sequencing Ansatz unter Anwendung von Exome-capture selektiert.

In beiden Experimenten konnten direkt vielversprechende Kandidatengene identifiziert werden und durch eine Reihe unabhängiger Allele aus historischen Mutantensammlungen sowie TILLING (targeting local lesions in genomes)-Analyse bestätigt werden, welche die gleichen Merkmalsausprägungen wie *mnd* bzw. *lax-a* Mutanten aufwiesen. Die erlangten Erkenntnisse können zukünftig genutzt werden um Mutantenkollektionen zu analysieren. Die dabei identifizierten regulierenden Gene wären ein vielversprechender Ausgangspunkt um komplexe Regulationsmechanismen des Pflanzenwachstums zu entschlüsseln. Das *HvMND* Gen konnte auf Basis von Sequenzübereinstimmungen der Proteinfamilie „Cytochrome P450 class 78A" zugeordnet werden. Das Gen *HvLAX-A* zeigte Homologie zu den *Arabidopsis thaliana* Genen *BLADE-ON-PETIOLE-1/2* (*BOP1/2*), welche mithilfe ihrer konservierten Proteindomänen „BTB/POZ" und „ankyrin repeats" als Transkriptionsregulatoren fungieren und dabei das Blattwachstum sowie die Organabgrenzung in den Blüten steuern. Genau wie in Arabidopsis konnte in Gerste auf Chromosom 3H ein paraloges Gen zu *HvLAX-A* identifiziert werden. Die Analyse der Genfunktion durch TILLING zeigte, dass dieses Gen eine Rolle in der Bestockung spielt, da Mutanten nur einen bis wenige Triebe bildeten. Zusätzlich konnte neben dem Auswuchs der Blattöhrchen (Auriculae) das Fehlen des Blatthäutchens (Ligula) beobachtet werden. Eine persönliche Rücksprache mit der Arbeitsgruppe von Dr. Laura Rossini (Universität Mailand, Italien) bestätigte, dass es sich bei dem paralogen Gen um das mutierte Gen der *uniculme4* Mutante handelt, welches die beobachteten Merkmalsausprägungen verursacht. Eine Analyse auf natürliche Diversität der

beiden Gene durch die Re-sequenzierung des Gencodierenden Bereiches in 305 Gersten Akzessionen zeigte sehr unterschiedliche Ausprägung von natürlicher Diversität. So zeigte *HvLAX-A* eine sehr geringe Diversität bereits in den Wildgerste Akzessionen. Im Vergleich dazu konnte für *HvCUL4* eine generell hohe Diversität identifiziert werden. Ob diese Unterschiede in der Diversität einem Selektionsprozess, natürlicher Populationsänderungen (Drift) oder dem Ergebnis der unterschiedlichen Genom Regionen zugrunde liegen, bedarf jedoch einer Genomweiten Ausdehnung der Analyse. Basierend auf der bereits gewonnenen Kenntnisse der Funktionsweise als Transkriptionsregulatoren in Arabidopsis, wurde versucht mögliche Gene zu identifizieren die durch *HvLAX-A* reguliert werden. Eine quantitative Transkriptionsanalyse mittels RNA-Sequenzierung (RNA-seq) wurde an drei unterschiedlichen, frühzeitigen Ährenentwicklungsstadien (Vegetationskegel) durchgeführt um Kandidatengene zu identifizieren, welche in *lax-a* spezifischen, regulierenden Gennetzwerken verantwortlich für die phänotypischen Unterschiede sind. Dabei wurde eine Reihe interessanter Kandidatengene identifiziert und in diesem Zusammenhang als Grundlage weiterer notwendiger Arbeiten zur Entschlüsselung regulatorischer Gennetzwerke diskutiert.

8. Appendix Tables

Table A1 : Information of plant material used to identify sequence haplotypes of *HvLAX-A* and *HvCUL4*

→ Accessions sorted according *HvCUL4* haplotypes

accession	material	region	country	HvCUL4	HvLAX-A
FT011	*Hordeum spontaneum*	WANA	ISR	Hap_1	Hap_2
FT031	*Hordeum spontaneum*	WANA	ISR	Hap_1	Hap_2
FT067	*Hordeum spontaneum*	WANA	ISR	Hap_2	Hap_2
FT144	*Hordeum spontaneum*	WANA	ISR	Hap_2	Hap_2
FT248	*Hordeum spontaneum*	WANA	IRN	Hap_2	Hap_4
FT276	*Hordeum spontaneum*	WANA	JOR	Hap_2	Hap_4
FT361	*Hordeum spontaneum*	WANA	GRC	Hap_2	Hap_2
FT64	*Hordeum spontaneum*	WANA	ISR	Hap_2	Hap_10
FT643	*Hordeum spontaneum*	WANA	ISR	Hap_2	Hap_2
BCC524	cultivar	EA	IND	Hap_2	Hap_2
BCC526	breeding / research material	EA	IND	Hap_2	Hap_2
BCC533	cultivar	EA	IND	Hap_2	Hap_4
BCC538	cultivar	EA	IND	Hap_2	Hap_2
BCC577	cultivar	EA	IND	Hap_2	Hap_2
BCC579	cultivar	EA	IND	Hap_2	Hap_2
BCC581	cultivar	EA	IND	Hap_2	Hap_2
BCC732	landrace	EA	NPL	Hap_2	Hap_2
BCC745	cultivar	EA	NPL	Hap_2	Hap_2
BCC761	landrace	EA	NPL	Hap_2	Hap_2
BCC766	landrace	EA	NPL	Hap_2	Hap_2
BCC768	landrace	EA	NPL	Hap_2	Hap_2
HOR11403	landrace	EA	IND	Hap_2	Hap_4
FT218	*Hordeum spontaneum*	WANA	ISR	Hap_3	Hap_2
FT340	*Hordeum spontaneum*	WANA	ISR	Hap_3	Hap_8
FT569	*Hordeum spontaneum*	WANA	IRN	Hap_3	Hap_2
FT037	*Hordeum spontaneum*	WANA	ISR	Hap_4	Hap_2
FT147	*Hordeum spontaneum*	WANA	ISR	Hap_4	Hap_2
FT219	/	/	/	Hap_4	Hap_2
BCC003	landrace	WANA	AFG	Hap_4	Hap_2
BCC093	/	/	/	Hap_4	Hap_2
BCC1367	cultivar	EU	DEU	Hap_4	Hap_2
BCC1372	cultivar	EU	POL	Hap_4	Hap_2
BCC1376	cultivar	EU	DNK	Hap_4	Hap_2
BCC1377	cultivar	EU	FRA	Hap_4	Hap_2
BCC1378	cultivar	EU	GBR	Hap_4	Hap_2
BCC1381	cultivar	EU	GBR	Hap_4	Hap_2
BCC1382	cultivar	EU	GBR	Hap_4	Hap_2
BCC1384	cultivar	EU	DEU	Hap_4	Hap_2
BCC1386	cultivar	EU	DEU	Hap_4	Hap_2

BCC1387	cultivar	EU	NLD	Hap_4	Hap_2
BCC1388	/	/	/	Hap_4	Hap_2
BCC1389	cultivar	EU	IRL	Hap_4	Hap_2
BCC1392	cultivar	EU	DNK	Hap_4	Hap_2
BCC1394	cultivar	EU	NLD	Hap_4	Hap_2
BCC1395	cultivar	EU	NLD	Hap_4	Hap_2
BCC1396	cultivar	EU	SWE	Hap_4	Hap_2
BCC1397	landrace	EU	HUN	Hap_4	Hap_2
BCC1398	breeding / research material	EU	HUN	Hap_4	Hap_2
BCC1399	cultivar	EU	SWE	Hap_4	Hap_2
BCC1400	cultivar	EU	FRA	Hap_4	Hap_2
BCC1401	cultivar	EU	DEU	Hap_4	Hap_2
BCC1402	cultivar	EU	SWE	Hap_4	Hap_2
BCC1404	cultivar	EU	GBR	Hap_4	Hap_2
BCC1405	cultivar	EU	GBR	Hap_4	Hap_2
BCC1407	cultivar	EU	AUT	Hap_4	Hap_2
BCC1408	cultivar	EU	GBR	Hap_4	Hap_2
BCC1409	cultivar	EU	AUT	Hap_4	Hap_2
BCC1410	cultivar	EU	SWE	Hap_4	Hap_2
BCC1412	cultivar	EU	SWE	Hap_4	Hap_2
BCC1415	cultivar	EU	GBR	Hap_4	Hap_2
BCC1418	cultivar	EU	DNK	Hap_4	Hap_2
BCC1419	cultivar	EU	DEU	Hap_4	Hap_2
BCC1422	cultivar	EU	NLD	Hap_4	Hap_2
BCC1425	cultivar	EU	DEU	Hap_4	Hap_2
BCC1428	cultivar	EU	GBR	Hap_4	Hap_2
BCC1439	cultivar	EU	CZE	Hap_4	Hap_2
BCC1444	cultivar	EU	CZE	Hap_4	Hap_2
BCC1453	cultivar	EU	FIN	Hap_4	Hap_2
BCC1456	cultivar	EU	RUS	Hap_4	Hap_2
BCC1458	cultivar	EU	RUS	Hap_4	Hap_2
BCC1465	cultivar	EU	UKR	Hap_4	Hap_2
BCC1466	cultivar	EU	UKR	Hap_4	Hap_2
BCC1467	cultivar	EU	BLR	Hap_4	Hap_2
BCC1472	cultivar	EU	LTU	Hap_4	Hap_2
BCC1479	landrace	EU	RUS	Hap_4	Hap_2
BCC1482	cultivar	EU	RUS	Hap_4	Hap_2
BCC1484	landrace	EU	RUS	Hap_4	Hap_2
BCC1485	cultivar	EU	RUS	Hap_4	Hap_2
BCC1491	landrace	EU	RUS	Hap_4	Hap_2
BCC1494	landrace	WANA	KAZ	Hap_4	Hap_2
BCC1498	landrace	WANA	UZB	Hap_4	Hap_2
BCC1505	cultivar	EU	UKR	Hap_4	Hap_2
BCC1524	cultivar	EU	DEU	Hap_4	Hap_2
BCC1529	cultivar	EU	AUT	Hap_4	Hap_2
BCC1541	cultivar	EU	YUG	Hap_4	Hap_2
BCC1565	landrace	EU	ALB	Hap_4	Hap_2

BCC161	landrace	WANA	MAR	Hap_4	Hap_2
BCC167	landrace	WANA	OMN	Hap_4	Hap_2
BCC427	breeding / research material	EA	CHN	Hap_4	Hap_2
BCC432	breeding / research material	EA	CHN	Hap_4	Hap_2
BCC666	cultivar	EA	KOR	Hap_4	Hap_2
BCC667	cultivar	EA	KOR	Hap_4	Hap_2
BCC675	cultivar	EA	KOR	Hap_4	Hap_2
BCC807	cultivar	AM	URY	Hap_4	Hap_2
BCC814	breeding / research material	AM	USA	Hap_4	Hap_2
BCC817	cultivar	AM	USA	Hap_4	Hap_2
BCC818	cultivar	AM	USA	Hap_4	Hap_2
BCC847	cultivar	AM	USA	Hap_4	Hap_2
BCC852	cultivar	AM	CAN	Hap_4	Hap_2
BCC860	breeding / research material	AM	URY	Hap_4	Hap_2
BCC861	breeding / research material	AM	URY	Hap_4	Hap_2
BCC929	cultivar	AM	CAN	Hap_4	Hap_2
BCC942	cultivar	AM	USA	Hap_4	Hap_2
HOR11374	breeding / research material	WANA	ISR	Hap_4	Hap_2
HOR1962	landrace	na	na	Hap_4	Hap_2
HOR4727	landrace	WANA	TUR	Hap_4	Hap_2
HOR7985	landrace	WANA	TUR	Hap_4	Hap_2
HOR8006	landrace	WANA	TUR	Hap_4	Hap_2
HOR930	landrace	WANA	TUR	Hap_4	Hap_2
FT222	*Hordeum spontaneum*	WANA	ISR	Hap_5	Hap_2
FT232	*Hordeum spontaneum*	WANA	IRQ	Hap_6	Hap_2
FT268	*Hordeum spontaneum*	WANA	SYR	Hap_6	Hap_5
FT278	*Hordeum spontaneum*	WANA	AFG	Hap_6	Hap_5
FT279	*Hordeum spontaneum*	WANA	AFG	Hap_6	Hap_5
FT376	*Hordeum spontaneum*	WANA	TKM	Hap_6	Hap_5
FT469	*Hordeum spontaneum*	WANA	TUR	Hap_6	Hap_5
BCC190	landrace	WANA	SYR	Hap_6	Hap_2
FT237	*Hordeum spontaneum*	WANA	IRN	Hap_7	Hap_5
FT252	*Hordeum spontaneum*	WANA	TUR	Hap_7	Hap_5
FT462	*Hordeum spontaneum*	WANA	TUR	Hap_7	Hap_2
FT507	*Hordeum spontaneum*	WANA	IRQ	Hap_7	Hap_5
FT578	*Hordeum spontaneum*	WANA	TUR	Hap_7	Hap_5
FT581	*Hordeum spontaneum*	WANA	TUR	Hap_7	Hap_5
FT582	*Hordeum spontaneum*	WANA	SYR	Hap_7	Hap_5
FT660	*Hordeum spontaneum*	WANA	IRN	Hap_7	Hap_5
FT661	*Hordeum spontaneum*	WANA	IRN	Hap_7	Hap_2
FT730	*Hordeum spontaneum*	WANA	TUR	Hap_7	Hap_2
FT731	*Hordeum spontaneum*	WANA	TUR	Hap_7	Hap_5
FT878	*Hordeum spontaneum*	WANA	IRN	Hap_7	Hap_5
BCC1442	cultivar	EU	GBR	Hap_7	Hap_2
FT272	*Hordeum spontaneum*	WANA	JOR	Hap_8	Hap_6
FT333	*Hordeum spontaneum*	WANA	IRN	Hap_8	Hap_2

FT589	*Hordeum spontaneum*	WANA	SYR	Hap_8	Hap_2
FT604	*Hordeum spontaneum*	WANA	IRQ	Hap_8	Hap_5
FT875	*Hordeum spontaneum*	WANA	UZB	Hap_8	Hap_5
FT879	*Hordeum spontaneum*	WANA	IRN	Hap_8	Hap_2
FT885	*Hordeum spontaneum*	WANA	TJK	Hap_8	Hap_5
FT286	*Hordeum spontaneum*	WANA	IRN	Hap_9	Hap_7
FT332	*Hordeum spontaneum*	WANA	IRN	Hap_10	Hap_5
FT379	*Hordeum spontaneum*	WANA	EGY	Hap_10	Hap_5
FT613	*Hordeum spontaneum*	WANA	IRN	Hap_10	Hap_4
FT747	*Hordeum spontaneum*	WANA	TUR	Hap_10	Hap_2
FT880	*Hordeum spontaneum*	WANA	IRN	Hap_10	Hap_5
FT262	*Hordeum spontaneum*	WANA	TUR	Hap_11	Hap_2
FT338	*Hordeum spontaneum*	WANA	ISR	Hap_11	Hap_2
FT473	*Hordeum spontaneum*	WANA	TUR	Hap_11	Hap_2
FT566	*Hordeum spontaneum*	WANA	TKM	Hap_11	Hap_4
FT567	*Hordeum spontaneum*	WANA	TKM	Hap_11	Hap_4
FT568	*Hordeum spontaneum*	WANA	AFG	Hap_11	Hap_4
FT572	*Hordeum spontaneum*	WANA	JOR	Hap_11	Hap_2
FT627	*Hordeum spontaneum*	WANA	JOR	Hap_11	Hap_2
FT658	*Hordeum spontaneum*	WANA	IRQ	Hap_11	Hap_9
BCC131	landrace	WANA	MAR	Hap_11	Hap_2
BCC195	landrace	WANA	SYR	Hap_11	Hap_2
BCC197	landrace	WANA	SYR	Hap_11	Hap_2
BCC527	breeding / research material	EA	IND	Hap_11	Hap_2
BCC535	cultivar	EA	IND	Hap_11	Hap_2
BCC695	cultivar	EA	KOR	Hap_11	Hap_4
HOR12830	landrace	WANA	SYR	Hap_11	Hap_2
FT357	*Hordeum spontaneum*	WANA	LBY	Hap_12	Hap_2
FT590	*Hordeum spontaneum*	WANA	SYR	Hap_13	Hap_2
FT595	*Hordeum spontaneum*	WANA	IRQ	Hap_14	Hap_5
FT670	*Hordeum spontaneum*	WANA	TUR	Hap_14	Hap_2
FT735	*Hordeum spontaneum*	WANA	TUR	Hap_14	Hap_2
FT748	*Hordeum spontaneum*	WANA	TUR	Hap_14	Hap_2
FT363	*Hordeum spontaneum*	WANA	CYP	Hap_15	Hap_2
FT380	*Hordeum spontaneum*	WANA	EGY	Hap_15	Hap_2
FT616	*Hordeum spontaneum*	WANA	IRN	Hap_15	Hap_2
FT734	*Hordeum spontaneum*	WANA	TUR	Hap_15	Hap_2
HOR2828	landrace	WANA	IRN	Hap_15	Hap_2
FT470	*Hordeum spontaneum*	WANA	TUR	Hap_16	Hap_5
FT630	*Hordeum spontaneum*	WANA	LBN	Hap_16	Hap_2
FT671	*Hordeum spontaneum*	WANA	LBN	Hap_16	Hap_2
FT754	*Hordeum spontaneum*	WANA	TUR	Hap_16	Hap_5
FT741	*Hordeum spontaneum*	WANA	TUR	Hap_17	Hap_5
FT042	*Hordeum spontaneum*	WANA	ISR	Hap_18	Hap_7
FT359	*Hordeum spontaneum*	WANA	MAR	Hap_19	Hap_2

FT365	*Hordeum spontaneum*	WANA	CHN	Hap_19	Hap_2
FT873	*Hordeum spontaneum*	WANA	UZB	Hap_19	Hap_5
FT886	*Hordeum spontaneum*	WANA	TJK	Hap_19	Hap_5
BCC118	landrace	WANA	LBY	Hap_19	Hap_2
BCC126	landrace	WANA	MAR	Hap_19	Hap_2
BCC129	landrace	WANA	MAR	Hap_19	Hap_2
BCC1368	cultivar	EU	NLD	Hap_19	Hap_2
BCC139196				Hap_19	Hap_2
BCC1455	cultivar	EU	RUS	Hap_19	Hap_2
BCC1457	cultivar	EU	RUS	Hap_19	Hap_2
BCC1474	cultivar	EU	UKR	Hap_19	Hap_2
BCC1476	cultivar	WANA	UZB	Hap_19	Hap_2
BCC149	landrace	WANA	MAR	Hap_19	Hap_2
BCC1493	cultivar	EU	UKR	Hap_19	Hap_2
BCC1503	landrace	WANA	TKM	Hap_19	Hap_2
BCC1504	cultivar	EU	RUS	Hap_19	Hap_2
BCC1548	cultivar	EU	CYP (GRC)	Hap_19	Hap_2
BCC173	landrace	WANA	PAK	Hap_19	Hap_4
BCC182	landrace	WANA	PAK	Hap_19	Hap_2
BCC218	landrace	WANA	TJK	Hap_19	Hap_4
BCC421	cultivar	EA	CHN	Hap_19	Hap_2
BCC423	cultivar	EA	CHN	Hap_19	Hap_2
BCC434	breeding / research material	EA	CHN	Hap_19	Hap_2
BCC438	cultivar	EA	CHN	Hap_19	Hap_2
BCC439	cultivar	EA	CHN	Hap_19	Hap_2
BCC502	cultivar	EA	CHN	Hap_19	Hap_2
BCC532	cultivar	EA	IND	Hap_19	Hap_2
BCC551	cultivar	EA	IND	Hap_19	Hap_5
BCC625	cultivar	EA	JPN	Hap_19	Hap_2
BCC642	cultivar	EA	JPN	Hap_19	Hap_4
BCC719	cultivar	EA	KOR	Hap_19	Hap_2
BCC729	cultivar	EA	KOR	Hap_19	Hap_2
BCC759	cultivar	EA	NPL	Hap_19	Hap_2
BCC801	landrace	EA	NPL	Hap_19	Hap_2
BCC806	cultivar	AM	MEX	Hap_19	Hap_2
BCC844	cultivar	AM	COL	Hap_19	Hap_2
BCC846	breeding / research material	AM	USA	Hap_19	Hap_2
BCC868	breeding / research material	AM	MEX	Hap_19	Hap_2
BCC869	cultivar	AM	MEX	Hap_19	Hap_2
BCC875	cultivar	AM	USA	Hap_19	Hap_2
BCC888	cultivar	AM	CAN	Hap_19	Hap_2
BCC892	cultivar	AM	BOL	Hap_19	Hap_2
BCC893	cultivar	AM	USA	Hap_19	Hap_2
BCC899	cultivar	AM	CHL	Hap_19	Hap_2
BCC900	cultivar	AM	MEX	Hap_19	Hap_2
BCC903	cultivar	AM	CAN	Hap_19	Hap_2

BCC907	cultivar	AM	USA	Hap_19	Hap_2
BCC921	cultivar	AM	COL	Hap_19	Hap_2
HOR1804	landrace	WANA	AFG	Hap_19	Hap_4
HOR1842	landrace	WANA	AFG	Hap_19	Hap_4
FT584	*Hordeum spontaneum*	WANA	SYR	Hap_20	Hap_5
FT746	*Hordeum spontaneum*	WANA	TUR	Hap_21	Hap_2
BCC1447	cultivar	EU	FRA	Hap_21	Hap_2
BCC1500	landrace	WANA	TJK	Hap_21	Hap_2
BCC219	landrace	WANA	TJK	Hap_21	Hap_2
FT18	*Hordeum spontaneum*	WANA	ISR	Hap_22	Hap_2
FT56	*Hordeum spontaneum*	WANA	ISR	Hap_23	Hap_4
FT75	*Hordeum spontaneum*	WANA	ISR	Hap_24	Hap_2
FT592	*Hordeum spontaneum*	WANA	SYR	Hap_25	Hap_5
FT628	*Hordeum spontaneum*	WANA	SYR	Hap_26	Hap_2
BCC192	landrace	WANA	SYR	Hap_27	Hap_2
HOR2800	landrace	WANA	IRN	Hap_27	Hap_2
BCC1452	cultivar	EU	NLD	Hap_28	Hap_2
BCC1461	cultivar	EU	RUS	Hap_28	Hap_2
BCC1468	cultivar	WANA	KAZ	Hap_28	Hap_2
BCC1469	cultivar	WANA	KAZ	Hap_28	Hap_2
BCC1490	breeding / research material	EU	RUS	Hap_28	Hap_2
BCC1561	landrace	EU	BGR	Hap_28	Hap_2
BCC436	cultivar	EA	CHN	Hap_28	Hap_2
BCC445	cultivar	EA	CHN	Hap_28	Hap_2
BCC446	cultivar	EA	CHN	Hap_28	Hap_2
BCC447	cultivar	EA	CHN	Hap_28	Hap_2
BCC718	cultivar	EA	KOR	Hap_28	Hap_2
BCC881	cultivar	AM	CAN	Hap_28	Hap_2
BCC1370	cultivar	EU	FRA	Hap_29	Hap_2
BCC1371	cultivar	EU	FRA	Hap_29	Hap_2
BCC1373	cultivar	EU	GBR	Hap_29	Hap_2
BCC1374	cultivar	EU	NLD	Hap_29	Hap_2
BCC1379	cultivar	EU	CZE	Hap_29	Hap_2
BCC1380	cultivar	EU	FRA	Hap_29	Hap_2
BCC1383	cultivar	EU	GBR	Hap_29	Hap_2
BCC1385	cultivar	EU	POL	Hap_29	Hap_2
BCC1390	cultivar	EU	SWE	Hap_29	Hap_2
BCC1391	cultivar	EU	DEU	Hap_29	Hap_2
BCC1403	cultivar	EU	DEU	Hap_29	Hap_2
BCC1411	cultivar	EU	DEU	Hap_29	Hap_2
BCC1413	cultivar	EU	DEU	Hap_29	Hap_2
BCC1414	cultivar	EU	CZE	Hap_29	Hap_2
BCC1416	cultivar	EU	SWE	Hap_29	Hap_2
BCC1417	cultivar	EU	DEU	Hap_29	Hap_2
BCC1420	cultivar	EU	NLD	Hap_29	Hap_2
BCC1421	cultivar	EU	CZE	Hap_29	Hap_2

BCC1423	cultivar	EU	FRA	Hap_29	Hap_2
BCC1424	cultivar	EU	DEU	Hap_29	Hap_2
BCC1430	cultivar	EU	FRA	Hap_29	Hap_2
BCC1431	cultivar	EU	AUT	Hap_29	Hap_2
BCC1432	cultivar	EU	CZE	Hap_29	Hap_2
BCC1433	cultivar	EU	DEU	Hap_29	Hap_2
BCC1440	cultivar	EU	AUT	Hap_29	Hap_2
BCC1441	cultivar	EU	DEU	Hap_29	Hap_2
BCC1443	cultivar	EU	DEU	Hap_29	Hap_2
BCC1445	breeding / research material	EU	FRA	Hap_29	Hap_2
BCC1450	cultivar	EU	FIN	Hap_29	Hap_2
BCC1459	cultivar	EU	RUS	Hap_29	Hap_2
BCC1463	cultivar	EU	RUS	Hap_29	Hap_2
BCC1470	cultivar	WANA	UZB	Hap_29	Hap_2
BCC1471	cultivar	EU	ARM	Hap_29	Hap_2
BCC1480	landrace	EU	RUS	Hap_29	Hap_2
BCC1481	landrace	EU	RUS	Hap_29	Hap_2
BCC1483	landrace	EU	RUS	Hap_29	Hap_2
BCC1487	landrace	EU	RUS	Hap_29	Hap_2
BCC1497	landrace	WANA	KGZ	Hap_29	Hap_2
BCC1506	cultivar	EU	UKR	Hap_29	Hap_2
BCC1566	landrace	EU	GRC	Hap_29	Hap_2
BCC1589	landrace	EU	ITA	Hap_29	Hap_2
BCC8050	cultivar	AM	CAN	Hap_29	Hap_2
BCC812	cultivar	AM	MEX	Hap_29	Hap_2
BCC857	cultivar	AM	MEX	Hap_29	Hap_2
BCC913	cultivar	AM	USA	Hap_29	Hap_2
HOR11370	breeding / research material	WANA	ISR	Hap_29	Hap_2
HOR11371	breeding / research material	WANA	ISR	Hap_29	Hap_2
HOR11373	breeding / research material	WANA	ISR	Hap_29	Hap_2
HOR2829	landrace	WANA	IRN	Hap_29	Hap_2
HOR2835	landrace	WANA	IRN	Hap_29	Hap_2
HOR8113	landrace	WANA	TUR	Hap_29	Hap_2
BCC1448	cultivar	EU	FIN	Hap_30	Hap_2
BCC927	cultivar	AM	PER	Hap_30	Hap_2
HOR8160	landrace	WANA	TUR	Hap_31	Hap_2

Table A2: Multiplex SNaPshot assay of *HvLAX-A*

Name	PCR pool	Sequence	size	Prod.size
1_0580_F	1	TGAGCAAGTTATGCATGTGCT	21 bp	92 bp
1_0580_R	1	CAGTGACCGGTTAATTTCACG	21 bp	
1_0580_S		d(T)$_7$GATGGTCAGATTTGCTCTCCTAC	30 bp	
1_1198_F	1	TTGACCTCGACGACGATATG	20 bp	38 bp
1_1198_R	1	AGGTGGAAGATCGCAAACAC	20 bp	
1_1198_S		d(T)$_{17}$CGATATCGACATGGACAGCAG	38 bp	
1_0157_F*	1	ACCACGCCCGAGAAAATAGT	20 bp	140 bp
1_0157_R*	1	GCAGATCAAGGATGGGTTTG	20 bp	
1_0157_S*		d(T)$_{26}$CCAATGTCCTTGAGCGCATT	46 bp	
2_0713_F	1	CGTGAATCAAGGATGTCTGC	20 bp	136 bp
2_0713_R	1	ACTCATTCTTGGGCACCAGT	20 bp	
2_0713_S		d(T)$_{33}$ATAGATGCCGCTGATTCAAGG	54 bp	
1_0481_F	2	CCTCTTGCCTCACATTGTTC	20 bp	128 bp
1_0481_R	2	GGCTGATGAGATGGAGCATT	20 bp	
1_0481_S		d(T)$_{37}$TCTGAGTTGAGACTCTTGAGTA	59 bp	
2_1148_F	2	GCCTGCTCGACTCAAGAGAA	20 bp	158 bp
2_1148_R	2	CGATCTCGAGCAAGGAAAAC	20 bp	
2_1148_S		d(T)$_{44}$GGCAAGTGAAGAGGAGAGGC	64 bp	
2_0524_F	2	GGCGGAGCAGTATCAGAAGA	20 bp	164 bp
2_0524_R	2	GCAGCTCCAGGGAACAAATA	20 bp	
2_0524_S		d(T)$_{47}$TTGCTTTGGAAGTGAGGCCTTC	69 bp	
1_1469_F	2	CGATTCGGTGAGAACTGGAT	20 bp	227 bp
1_1469_R	2	CCGCCTACAAACATGAACCT	20 bp	
1_1469_S		d(T)$_{51}$CCACCTTTTTATATATTCTGTTC	74 bp	

*: not the Illumina SNP but SNP in the same gene
F = forward; R = reverse; S = sequence extension

Table A3: Oligonucleotides of *HvMND*

Name	forward sequence (5' - 3')	reverse sequence (5' - 3')
HvMND_F1/R1	CCCTACTCCTCGATGACGAT	ACACGGACCACATCACGTT
HvMND_F2/R2	CACCATCTCCTGCAACACAC	CCAGGGAGATCCTCAACAGC
HvMND_F3/R3	CATCGCTTGCTAGTTGCTCA	ACATGATCGCGGTGCTCT
HvMND_EX1_F1/R1	GCCTTTTCCTGCGGCTATAA	CAGAGAATACCCCCAACGAA
HvMND_EX2_F2/R2	GTCCTGCTCTCCCTCCAAG	CGCTTGCTAGTTGCTCATCA

Table A4: Overview of sequenced BAC clones

BAC contig	BAC clone[1]
FPcontig_1020:	HVVMRX83KhA0173C12
	HVVMRXALLeA0300M11
	HVVMRXALLeA0380K08
	HVVMRXALLmA0117G20
	HVVMRXALLmA0119I23
	HVVMRXALLmA0158E02
	HVVMRXALLmA0233K17
	HVVMRXALLmA0286G06
	HVVMRXALLmA0359K20
	HVVMRXALLmA0364P05
	HVVMRXALLmA0429F04
	HVVMRXALLrA0018J12
	HVVMRXALLrA0065C12
FPcontig_45097:	HVVMRXALLeA0166N13
	HVVMRXALLeA0173M05
	HVVMRXALLhA0014F08
	HVVMRXALLhB0080C03
	HVVMRXALLmA0207L21
	HVVMRXALLmA0260F04
	HVVMRXALLrA0026H17
	HVVMRXALLrA0074I17
	HVVMRXALLrA0401A22
FPcontig_46058:	HVVMRX83KhA0214N18
	HVVMRXALLeA0009B12
	HVVMRXALLeA0029D13
	HVVMRXALLeA0069L08
	HVVMRXALLeA0076E11
	HVVMRXALLeA0092H05
	HVVMRXALLeA0140P18
	HVVMRXALLeA0280D23
	HVVMRXALLeA0281L17
	HVVMRXALLhB0167G05
	HVVMRXALLhC0055G09
	HVVMRXALLmA0011E02
	HVVMRXALLmA0142K09
	HVVMRXALLmA0506H07
	HVVMRXALLrA0378O03
FPcontig_2862:	HVVMRXALLmA0117O22
	HVVMRXALLmA0012B04
	HVVMRXALLmA0114D14
	HVVMRXALLeA0185O05
	HVVMRXALLrA0187L16
	HVVMRXALLeA0324B05
	HVVMRXALLmA0029K21
	HVVMRXALLeA0046E04
	HVVMRXALLmA0255O17
	HVVMRXALLmA0315H03
	HVVMRXALLmA0033O13
	HVVMRXALLeA0016M20
	HVVMRXALLhA0071F16
	HVVMRXALLrA0261G02
	HVVMRXALLeA0048H07
	HVVMRXALLeA0171I14
	HVVMRXALLeA0137L10
	HVVMRXALLhC0076L02
	HVVMRXALLmA0433B15
	HVVMRXALLmA0029A21

HVVMRXALLmA0012H10
HVVMRXALLeA0074F19
HVVMRXALLmA0072A18
HVVMRXALLmA0014N12
HVVMRXALLmA0187F06
HVVMRXALLeA0258L22
HVVMRXALLmA0201L14
HVVMRXALLhB0169O19
HVVMRXALLHB0067H22
HVVMRXALLEA0228K11
HVVMRXALLHC0099I01
HVVMRXALLMA0216M07
HVVMRXALLRA0067J14
HVVMRX83KHA0024J08
HVVMRX83KHA0157G10
HVVMRXALLRA0259K16
HVVMRXALLEA0228K11
HVVMRXALLHB0099C18
HVVMRXALLMA0474B14
HVVMRXALLEA0298O21
HVVMRXALLRA0392E09
HVVMRXALLEA0366F22
HVVMRXALLRA0127J23
HVVMRXALLRA0403L22
HVVMRXALLRA0061G22
HVVMRXALLEA0122D17
HVVMRXALLEA0209O12
HVVMRXALLEA0379H14
HVVMRXALLEA0084I23
HVVMRXALLEA0046E04
HVVMRXALLHB0082E03
HVVMRXALLMA0062N24
HVVMRXALLEA0041O05
HVVMRXALLMA0096J10
HVVMRXALLMA0094L06
HVVMRXALLMA0106E21
HVVMRXALLHB0148A07
HVVMRXALLRA0164F10

[1] BACs of FPcontig_1020 were sequenced by Roche 454 Titanium

Table A5: Oligonucleotides of *HvLAX-A*

Name	forward sequence (5' - 3')	reverse sequence (5' - 3')
HvLAX_Ex1	ACCCCTCCCACTTTTACCAC	TGGCAGGCCTAACTACCATC
HvLAX_Ex2	ACTGATCCAGACCCCAACAG	GGATCGAGAGCGAGATATGC
HvLAX_1	GCTGGTGTTGCAGTTCCTCT	GAAGCGACAGACACAAACGA
HvLAX_2	GGTGCAGTGTGTACTGAACGA	CATGACCATGAGCTTGACGA
HvLAX_3	CCACTACCTGCCAATCAACA	CATGTCCAACGATGATGCAG
HvLAX_4	CCCTAACCTTTCCGTCTTCC	CGACTGGAGATCGAGTTAGC
HvLAX_5	ATGCATTGCTGTGCGGTA	AGGAGGACGGACACCATATC
HvLAX_6	AGCTCGTCAAGCTCATGGTC	CGATACAGCACCGTCGTATG
contig_71318	CCGTCAGCCTCAGTTTTTGT	AGCAGAGTAAAAGGGCGACA
contig_129575	GGATCCCCTAGATGCAAGGA	GGGAGATGGAGGGGTATATG

Table A6: Oligonucleotides of *HvCUL4*

Name	forward sequence (5' - 3')	reverse sequence (5' - 3')
HvCUL4_Ex1	GATTGGAGGGCAACAACAAG	TCGACGACATACTGCGCTAC
HvCUL4_Ex2	CGGTCTCTCCATGCCATATT	CATTCTCGTCGACCGATCTC
HvCUL4_1	GCTCTAGCTCGATCTGATTCATCC	CCATCTTGATCCCCACAATC
HvCUL4_2	CCTCTAGTAAGCTCTCTCGGTCTC	ATGTCCGGGCAGACCATCT
HvCUL4_3	CTCGACGACCACAACAAGAT	CAGCCCAAAAGGATCATCTG

Table A7: CAPS markers used for genetic mapping

Marker	Barke WGS contig	forward sequence (5' - 3')	reverse sequence (5' - 3')	Enzyme
M1	contig_1794173	TCCTGGTCTAGGAGGATGTTCTG	CACTAGCTCCCAGGAGGTACAAA	EcoRI
M2	contig_274977	AGGATGATCTTCTTGGCCTCCT	GCAGGACAGCTTGAATGTTGAC	HaeII
M3	contig_1781303	CATTATCCTCTGTTGCCACCTG	CTCGGAGACCCATAAGATCGAG	EcoRI
M4*	contig_265820	CGACGAAGATGACTCTGGAAG	GAACAGCACCACCACCATC	/
M5	contig_1783058	GATTACGGGGTGGATGCTC	ACCCTCACCTCCAGCCATA	HpyCH4III
M6	contig_267961	CCACTTCATAGAGTCGCTCCTGT	CGAGAGGATCTATTCGTGTCTGG	HaeII
M7	contig_2786329	AAGCAGGAGAAGGGTATCAGCTC	GTCAACCGGCAAGCCTTAGATA	HaeII
M8	contig_269740	GTCTTCCTTCGAGGTGAGCACT	GTGCCGTAGCTGCAACAATAAG	EcoRI
M9	contig_481335	ATGTATACTGGAGGCTGGAGCTG	AACAGAGCCTGCAAACAAGACC	HindIII
M10	contig_55926	CTCTCCATCCTCAACTCCTTCCT	GGACCTTGACTATAAGGCTCCAC	HaeII

*mapped by Sanger sequencing

Table A8: Capture targets with low coverage in the *mnd* mutant pool

Morex_contig	hc_gene	start	end	length	chr[1]	cM[1]	ibsc_cM	ibsc_fpc	morex_contig _length	hc_gene_annotation	lc_gene
contig_172677	NA	649	904	256	5	91.16	NA	NA	1065	NA	NA
contig_82118	NA	3474	3629	156	5	95.00	93.92	contig_44536	5116	NA	MLOC_78840.1
contig_273931	NA	304	585	282	5	96.60	NA	NA	4254	NA	NA
contig_49382	MLOC_64838.2	2107	2455	349	5	96.60	95.57	no_fpc	8874	Cytochrome P450	NA
contig_49382	MLOC_64838.2	3871	4101	231	5	96.60	95.57	no_fpc	8874	Cytochrome P450	NA
contig_209576	NA	1118	1286	169	5	98.72	96.25	contig_44691	2053	NA	NA
contig_94037	NA	405	666	262	5	99.93	98.71	contig_46725	8771	NA	MLOC_81017.1
contig_1558349	MLOC_10070.1	1236	1477	242	5	99.93	NA	NA	4654	MATE efflux family protein	MLOC_10071.1
contig_1558349	MLOC_10070.1	1236	1477	242	5	99.93	NA	NA	4654	MATE efflux family protein	MLOC_10069.1
contig_159829	MLOC_21734.1	785	996	212	5	99.93	NA	NA	6336	CC-NBS-LRR class	NA
contig_159829	MLOC_21734.1	4501	4709	209	5	99.93	NA	NA	6336	CC-NBS-LRR class	NA
contig_1582387	NA	139	339	201	5	99.93	NA	NA	1826	NA	NA
contig_45126	AK365521	14800	14961	162	5	107.08	97.74	contig_10120	15662	Phytoene synthase 3	NA
contig_45126	AK365660	14800	14961	162	5	107.08	97.74	contig_10120	15662	Tumor susceptibility 101	NA
contig_45126	MLOC_61115.1	14800	14961	162	5	107.08	97.74	contig_10120	15662	MORN repeat protein	NA
contig_2547452	AK357178	4712	4913	202	5	108.14	102.77	no_fpc	5951	3-5 exoribonuclease CSL4	MLOC_36881.1
contig_2222314	NA	207	362	156	5	108.14	NA	NA	579	NA	NA
contig_58347	MLOC_70680.1	7926	8186	261	5	109.38	101.18	no_fpc	13413	DNA repair protein-like	MLOC_70681.1
contig_1612436	NA	1072	1271	200	5	109.38	NA	NA	1629	NA	NA
contig_38917	NA	5097	5392	296	5	109.65	NA	NA	6934	NA	MLOC_54137.1
contig_109704	NA	3139	3357	219	5	109.65	109.65	contig_39059	4478	NA	NA

[1] POPseq; [2] IBSC

Table A9: Oligonucleotides to test for complete deletions of neighboring genes

Brachypodium gene ID	Barley gene ID[1]	FPCcontig ID[1]	forward sequence (5' - 3')	reverse sequence (5' - 3')
Bradi4g35867	AK368025	45097	GTCAGGAGCCTCCGTTCC	CTCTCGTCCACCCAGAAGAC
Bradi4g35880	AK355945	45097	CGGCCTCTTCTTCGTCCT	AGGTCACCAGCAAGGTCCTA
Bradi4g35890	*MLOC_64838.2*	*45097*	*HvMND_F3/R3*	*HvMND_F3/R3*
Bradi4g35900	MLOC_54310.2	46058	ATCTTCTCCCCCTCCCACT	CTAACATTGGCGCACTTTGA
Bradi4g35910	AK376953	no_fpc	GACGGGGACTGGATGATG	CTCGAGTGGAACGGGAAC
Bradi4g35930	MLOC_52879.5	46058	GCGCTCTGTGTCCGTCTT	TTTAGTTCACCTCCGGATCG
Bradi4g35940	MLOC_22687.1	46058	GCTCCATCAACGTCTCCTTC	GCAAGACGTCCAAGAAGTCC
Bradi4g35950	AK374133	46058	AAGCTTGCTAGTCGGCATGT	GTGGTTCACCATTTGCTCCT

[1] (IBSC, 2012)

Table A10: Overlapping BAC clones of FPC_45097

FP_Contig	BAC_clone[1]	cb start[2]	cb_end[2]	cummulative sequence	overlapping BAC clones with >2000 bp alignment length and >99% identity
FPC_45097	HVVMRXALLMA0207L21	0	133	154899	HVVMRXALLeA0166N13; HVVMRXALLeA0173M05; HVVMRXALLmA0260F04;
FPC_45097	HVVMRXALLeA0166N13	28	104	128912	HVVMRXALLmA0207L21; HVVMRXALLeA0173M05
FPC_45097	HVVMRXALLeA0173M05	60	225	159543	HVVMRXALLeA0166N13; HVVMRXALLmA0207L21; HVVMRXALLmA0260F04; HVVMRXALLrA0401A22
FPC_45097	HVVMRXALLmA0260F04	103	233	161408	HVVMRXALLrA0401A22; HVVMRXALLeA0173M05: HVVMRXALLmA0207L21;
FPC_45097	HVVMRXALLrA0401A22	116	279	204615	HVVMRXALLmA0260F04; HVVMRXALLrA0074I17; HVVMRXALLeA0173M05; HVVMRXALLhA0014F08
FPC_45097	HVVMRXALLrA0074I17	224	337	124685	HVVMRXALLhA0014F08; HVVMRXALLrA0026H17; HVVMRXALLrA0401A22; HVVMRXALLeA0076E11; HVVMRXALLhB0080C03; HVVMRXALLeA0280D23
FPC_45097	HVVMRXALLhB0080C03	307	386	98612	HVVMRXALLrA0074I17; HVVMRXALLeA0076E11; HVVMRXALLeA0280D23; HVVMRXALLhA0014F08; HVVMRXALLrA0026H17; HVVMRXALLeA0281L17
FPC_45097	HVVMRXALLrA0026H17	312	374	83779	HVVMRXALLeA0281L17; HVVMRXALLhC0055G09; HVVMRXALLhA0014F08; HVVMRXALLhB0080C03; HVVMRXALLeA0076E11; HVVMRXALLeA0280D23; HVVMRXALLrA0074I17
FPC_45097	HVVMRXALLHA0014F08	289	350	66412	HVVMRXALLeA0076E11; HVVMRXALLrA0026H17; HVVMRXALLrA0074I17 HVVMRXALLhB0080C03; HVVMRXALLrA0401A22; HVVMRXALLeA0280D23

[1] BACs sequenced by Illumina HiSeq2000, [2] Consensus bands of High Information Content Fingerprinting (HICF, IBSC, 2012; Ariyadasa et al., 2014)

Table A11: Overlapping BAC clones of FPC_46058

FP_Contig	BAC_clone[1]	cb start[2]	cb_end[2]	cummulative sequence	overlapping BAC clones with >2000 bp alignment length and >99% identity
FPC_46058	HVVMRXALLMA0506H07	0	139	157629	HVVMRXALLeA0069L08; HVVMRX83KhA0214N18; HVVMRXALLmA0142K09; HVVMRXALLhB0167G05
FPC_46058	HVVMRXALLeA0069L08	21	140	124573	HVVMRXALLmA0506H07; HVVMRX83KhA0214N18; HVVMRXALLmA0142K09
FPC_46058	HVVMRX83KhA0214N18	70	150	115080	HVVMRXALLeA0069L08; HVVMRXALLmA0506H07; HVVMRXALLhB0167G05; HVVMRXALLmA0142K09
FPC_46058	HVVMRXALLmA0142K09	107	224	158416	HVVMRXALLmA0506H07; HVVMRXALLhB0167G05; HVVMRX83KhA0214N18; HVVMRXALLeA0069L08; HVVMRXALLeA0009B12
FPC_46058	HVVMRXALLhB0167G05	141	297	250074	HVVMRX83KhA0214N18; HVVMRXALLmA0142K09; HVVMRXALLeA0009B12; HVVMRXALLmA0506H07
FPC_46058	HVVMRXALLeA0009B12	197	340	207976	HVVMRXALLeA0029D13; HVVMRXALLhB0167G05; HVVMRXALLmA0142K09
FPC_46058	HVVMRXALLeA0029D13	289	388	147011	HVVMRXALLrA0378O03; HVVMRXALLeA0092H05; HVVMRXALLeA0009B12
FPC_46058	HVVMRXALLeA0092H05	339	435	165047	HVVMRXALLrA0378O03; HVVMRXALLmA0011E02; HVVMRXALLeA0029D13
FPC_46058	HVVMRXALLrA0378O03	363	456	172214	HVVMRXALLeA0092H05; HVVMRXALLmA0011E02; HVVMRXALLhC0055G09; HVVMRXALLeA0029D13
FPC_46058	HVVMRXALLmA0011E02	402	525	164652	HVVMRXALLhC0055G09; HVVMRXALLeA0076E11; HVVMRXALLeA0281L17; HVVMRXALLeA0280D23; HVVMRXALLeA0092H05; HVVMRXALLrA0378O03
FPC_46058	HVVMRXALLeA0281L17	474	551	97682	HVVMRXALLeA0076E11; HVVMRXALLeA0280D23; HVVMRXALLhC0055G09; HVVMRXALLmA0011E02; HVVMRXALLrA0026H17; HVVMRXALLhB0080C03
FPC_46058	HVVMRXALLhC0055G09	488	532	110947	HVVMRXALLmA0011E02; HVVMRXALLrA0378O03; HVVMRXALLrA0026H17; HVVMRXALLeA0281L17; HVVMRXALLeA0280D23; HVVMRXALLeA0076E11
FPC_46058	HVVMRXALLeA0076E11	490	559	136887	HVVMRXALLhB0080C03; HVVMRXALLeA0280D23; HVVMRXALLeA0281L17; HVVMRXALLrA0026H17; HVVMRXALLrA0074I17; HVVMRXALLhA0014F08; HVVMRXALLhC0055G09; HVVMRXALLmA0011E02
FPC_46058	HVVMRXALLEA0280D23	500	586	112017	HVVMRXALLeA0281L17; HVVMRXALLeA0076E11; HVVMRXALLhC0055G09; HVVMRXALLrA0074I17; HVVMRXALLhA0014F08; HVVMRXALLhB0080C03; HVVMRXALLrA0026H17; HVVMRXALLmA0011E02

[1] BACs sequenced by Illumina HiSeq2000, [2] Consensus bands of High Information Content Fingerprinting (HICF, IBSC, 2012; Ariyadasa et al., 2014)

Table A12: Overlapping BAC clones of FPcontig_1020

BAC contig	BAC_clone [1]	cb start [2]	cb_end [2]	cumulative sequence	overlapping BAC clones with >2000 bp alignment length and >99% identity
FPC_1020	*HVVMRXALLHA0377F15*	0	64	99653 bp	HVVMRX83KhA0091K10
FPC_1020	HVVMRX83KHA0091K10	36	122	115055 bp	HVVMRXALLeA0380K08; HVVMRXALLhA0377F15
FPC_1020	HVVMRXALLEA0380K08	74	187	165565 bp	HVVMRXALLmA0117G20; HVVMRX83KhA0091K10
FPC_1020	HVVMRXALLMA0117G20	179	258	152254 bp	HVVMRXALLeA0380K08; HVVMRXALLhC0200D12
FPC_1020	*HVVMRXALLHC0200D12*	254	362	142387 bp	HVVMRXALLmA0117G20; HVVMRXALLhA0424E16; HVVMRXALLeA0300M11; HVVMRX83KhA0173C12
FPC_1020	*HVVMRXALLHA0424E16*	275	360	104595 bp	HVVMRXALLhC0200D12; HVVMRXALLeA0300M11; HVVMRX83KhA0173C12
FPC_1020	HVVMRX83KHA0173C12	295	401	112274 bp	HVVMRXALLhC0200D12; HVVMRXALLhA0424E16; HVVMRXALLeA0300M11
FPC_1020	HVVMRXALLEA0300M11	322	480	181481 bp	HVVMRXALLmA0429F04; HVVMRXALLhC0200D12; HVVMRXALLhA0424E16; HVVMRX83KhA0173C12
FPC_1020	HVVMRXALLMA0429F04	385	516	160603 bp	HVVMRXALLeA0300M11; HVVMRXALLmA0158E02
FPC_1020	HVVMRXALLMA0158E02	471	592	153871 bp	HVVMRXALLmA0286G06; HVVMRXALLmA0429F04
FPC_1020	*HVVMRXALLMA0286G06*	546	704	159061 bp	HVVMRX83khA0087M10; HVVMRXALLmA0158E02
FPC_1020	HVVMRXALLMA0364P05	654	811	168099 bp	HVVMRX83khA0087M10; HVVMRXALLmA0119I23; HVVMRXALLmA0233K17
FPC_1020	*HVVMRX83KHA0087M10*	661	765	102809 bp	HVVMRXALLmA0286G06; HVVMRXALLmA0119I23; HVVMRXALLmA0364P05
FPC_1020	HVVMRXALLMA0119I23	765	915	196053 bp	HVVMRXALLmA0364P05; HVVMRXALLmA0233K17; HVVMRX83khA0087M10
FPC_1020	HVVMRXALLMA0233K17	814	977	168059 bp	HVVMRXALLmA0359K20; HVVMRXALLmA0119I23; HVVMRXALLmA0364P05
FPC_1020	HVVMRXALLMA0359K20	896	1001	133652 bp	HVVMRXALLrA0018J12; HVVMRXALLmA0233K17
FPC_1020	HVVMRXALLRA0018J12	961	1083	123614 bp	HVVMRXALLmA0359K20; HVVMRXALLrA0065C12
FPC_1020	HVVMRXALLRA0065C12	1016	1135	126639 bp	HVVMRXALLrA0018J12

[1] BACs sequenced on Roche 454, Italic BACs published by IBSC, 2012, [2] Consensus bands of High Information Content Fingerprinting (HICF, IBSC, 2012; Ariyadasa et al., 2014)

Table A13: Genes and synteny information for genes in the physical interval of *HvMND*

HC_gene	Annotation IBSC[1]	Homoglog[2]
MLOC_4956.1	C-repeat binding factor 3-like protein	Bradi4g35630.1
AK374663	C-repeat binding factor 6	Bradi4g35650.1
MLOC_67798.1	MATE efflux family protein	Bradi4g35660.1
MLOC_53578.1	BAG family molecular chaperone regulator 3	Bradi4g35680.1
AK374095	rRNA biogenesis protein rrp36	Bradi4g35690.1
MLOC_52708.1	F-box family protein	Bradi4g35700.1
MLOC_52707.1	RING/U-box superfamily protein	Bradi4g35710.3
MLOC_8011.1	Protein of unknown function (DUF579)	*Bradi4g40400.1*
MLOC_67041.1	UBX domain-containing protein	Bradi4g35720.1
MLOC_6781.1	Generative cell specific-1	Bradi4g35730.1
MLOC_16893.1	Transferase family protein	*Bradi1g28210.1*
MLOC_51157.1	Transferase family protein	*Bradi4g20790.1*
MLOC_72200.1	Transferase family protein	*Bradi1g28210.1*
MLOC_72201.1	Transferase family protein	*Bradi4g20790.1*
AK368631	Rhomboid family protein	Bradi4g35740.1
AK365794	homeobox-leucine zipper protein 3	Bradi4g35760.1
AK355696	Transposase	*Bradi3g04500.1*
AK368025	6-phosphogluconolactonase	Bradi4g35867.1
MLOC_49081.1	Unknown protein	/
AK355945	GTPase activating-like protein	Bradi4g35880.2
MLOC_64838.2	**Cytochrome P450**	**Bradi4g35890.1**
MLOC_54310.2	Mediator of RNA polymerase II transcription subunit	Bradi4g35900.1
AK376953	Homeobox leucine zipper protein	Bradi4g35910.1
MLOC_52879.5	Kinesin-related protein	Bradi4g35930.1
AK357642	Cytochrome P450	Bradi4g35940.1
AK374000	Cytochrome P450	Bradi4g35940.1
MLOC_66893.1	U-box domain-containing protein 73	*Bradi3g43630.1*
AK374133	Zinc finger protein CONSTANS-like protein	Bradi4g35950.1
MLOC_62610.1	Auxin responsive protein	Bradi4g35960.1
AK373680	Endo-1,4-beta-glucanase	Bradi4g35970.1
AK358124	UDP-Glycosyltransferase superfamily	*Bradi5g11950.1*
MLOC_4008.2	Zinc finger CCCH domain-containing	Bradi4g35977.1
AK360526	NEDD8-activating enzyme E1 regulatory subunit	*Bradi1g03800.1*
MLOC_1996.2	Undecaprenyl-phosphate 4-deoxy-4-formamido-L-arabinose transferase	*Bradi2g19800.1*
MLOC_61767.1	FAR1-related sequence 3	/
AK354518	unknown protein	Bradi4g35990.1
MLOC_55045.1	Subtilase family protein	Bradi4g36000.1
MLOC_55046.1	Oxidoreductase-like protein	Bradi4g36010.1
MLOC_10511.4	Mei2-like protein	Bradi4g36020.1

[1] (IBSC, 2012) [2] gene in italic models are not collinear in Brachypodium
HvMND in bold letters; marker labeled by grey background

Table A14: High confidence genes on sequenced BACs of the physical interval for
HvMND

HC_Gene_ID	BAC sequences containing gene
MLOC_4956.1	HVVMRXALLhC0200D12; HVVMRXALLmA0117G20
AK374663	HVVMRXALLhA0424E16; HVVMRXALLhC0200D12
MLOC_67798.1	HVVMRXALLmA0158E02
MLOC_53578.1	HVVMRX83khA0087M10; HVVMRXALLmA0119I23; HVVMRXALLmA0364P05
AK374095	HVVMRXALLmA0119I23; HVVMRXALLmA0364P05
MLOC_52708.1	HVVMRXALLmA0119I23; HVVMRXALLmA0364P05
MLOC_52707.1	HVVMRXALLmA0119I23; HVVMRXALLmA0364P05
MLOC_8011.1	HVVMRXALLmA0119I23; HVVMRXALLmA0233K17; HVVMRXALLmA0364P05
MLOC_67041.1	HVVMRXALLmA0119I23; HVVMRXALLmA0233K17
MLOC_6781.1	HVVMRXALLmA0233K17; HVVMRXALLmA0359K20
MLOC_16893.1	HVVMRXALLmA0359K20; HVVMRXALLrA0018J12
MLOC_51157.1	HVVMRXALLmA0359K20; HVVMRXALLrA0018J12
MLOC_72200.1	HVVMRXALLrA0018J12
MLOC_72201.1	HVVMRXALLrA0018J12
AK368631	HVVMRXALLrA0018J12
AK365794	HVVMRXALLrA0065C12
AK355696	HVVMRXALLrA0065C12
AK368025	HVVMRXALLeA0166N13; HVVMRXALLmA0207L21
MLOC_49081.1	HVVMRXALLmA0207L21; HVVMRXALLeA0166N13; HVVMRXALLeA0173M05
AK355945	HVVMRXALLeA0173M05; HVVMRXALLmA0207L21; HVVMRXALLeA0166N13
MLOC_64838.2	HVVMRXALLrA0074I17; HVVMRXALLrA0401A22
MLOC_54310.2	HVVMRXALLeA0076E11; HVVMRXALLeA0280D23; HVVMRXALLeA0281L17; HVVMRXALLhC0055G09; HVVMRXALLmA0011E02
AK376953	HVVMRXALLhC0055G09; HVVMRXALLmA0011E02; HVVMRXALLeA0076E11; HVVMRXALLeA0280D23; HVVMRXALLeA0281L17
MLOC_52879.5	HVVMRXALLmA0011E02; HVVMRXALLeA0076E11; HVVMRXALLeA0280D23; HVVMRXALLeA0281L17; HVVMRXALLhC0055G09
AK357642	HVVMRXALLeA0009B12; HVVMRXALLeA0029D13
AK374000	HVVMRXALLeA0009B12; HVVMRXALLeA0029D13
MLOC_66893.1	HVVMRXALLrA0378O03; HVVMRXALLeA0092H05; HVVMRXALLmA0011E02
AK374133	HVVMRXALLeA0092H05; HVVMRXALLmA0011E02; HVVMRXALLrA0378O03
MLOC_62610.1	HVVMRXALLmA0011E02; HVVMRXALLrA0378O03; HVVMRXALLeA0092H05
AK373680	HVVMRXALLeA0029D13; HVVMRXALLeA0092H05; HVVMRXALLrA0378O03
AK358124	HVVMRXALLhB0167G05; HVVMRXALLeA0009B12
MLOC_4008.2	HVVMRXALLhB0167G05; HVVMRXALLeA0009B12
AK360526	HVVMRXALLhB0167G05; HVVMRXALLmA0142K09; HVVMRXALLeA0009B12
MLOC_1996.2	HVVMRXALLhB0167G05; HVVMRXALLeA0009B12; HVVMRXALLmA0142K09
MLOC_61767.1	HVVMRXALLhB0167G05; HVVMRXALLmA0142K09
AK354518	HVVMRXALLmA0142K09; HVVMRXALLhB0167G05

MLOC_55045.1	HVVMRXALLmA0142K09;　HVVMRXALLhB0167G05
MLOC_55046.1	HVVMRXALLmA0142K09;　HVVMRXALLhB0167G05
MLOC_10511.4	HVVMRXALLmA0142K09;　HVVMRXALLmA0506H07;　HVVMRX83KhA0214N18; HVVMRXALLeA0069L08

Table A15: Low confidence genes on Sequenced BACs of the physical interval for
HvMND

LC_gene_ID	BAC sequences containing gene
MLOC_76105.1	HVVMRX83KhA0091K10;　HVVMRXALLhA0377F15
MLOC_81791.1	HVVMRX83KhA0091K10;　HVVMRXALLhA0377F15
MLOC_26918.1	HVVMRX83KhA0091K10;　HVVMRXALLhA0377F15
MLOC_54400.1	HVVMRXALLeA0380K08
MLOC_66767.1	HVVMRXALLhC0200D12;　HVVMRXALLmA0117G20
MLOC_34861.1	HVVMRXALLmA0158E02
AK357794	HVVMRXALLhA0424E16;　HVVMRXALLhC0200D12
MLOC_18170.2	HVVMRXALLmA0119I23;　HVVMRXALLmA0364P05
MLOC_1861.1	HVVMRXALLmA0119I23;　HVVMRXALLmA0233K17
MLOC_24698.1	HVVMRXALLmA0119I23;　HVVMRXALLmA0233K17; HVVMRXALLmA0364P05
MLOC_74421.1	HVVMRXALLmA0119I23;　HVVMRXALLmA0233K17; HVVMRXALLmA0364P05
MLOC_33705.1	HVVMRXALLmA0359K20;　HVVMRXALLrA0018J12
MLOC_28665.1	HVVMRXALLrA0018J12;　HVVMRXALLrA0065C12
MLOC_35114.1	HVVMRXALLrA0018J12;　HVVMRXALLrA0065C12
MLOC_41679.1	HVVMRXALLrA0018J12;　HVVMRXALLrA0065C12
MLOC_52261.1	HVVMRXALLrA0018J12;　HVVMRXALLrA0065C12
MLOC_64495.4	HVVMRXALLrA0018J12;　HVVMRXALLrA0065C12
MLOC_64496.1	HVVMRXALLrA0018J12;　HVVMRXALLrA0065C12
MLOC_64497.1	HVVMRXALLrA0018J12;　HVVMRXALLrA0065C12
MLOC_27836.1	HVVMRXALLrA0018J12;　HVVMRXALLrA0065C12
MLOC_44914.1	HVVMRXALLrA0065C12
MLOC_12487.1	HVVMRXALLrA0065C12
MLOC_14358.1	HVVMRXALLeA0166N13;　HVVMRXALLeA0173M05; HVVMRXALLmA0207L21
MLOC_23913.1	HVVMRXALLeA0166N13;　HVVMRXALLeA0173M05; HVVMRXALLmA0207L21
MLOC_46809.1	HVVMRXALLeA0166N13;　HVVMRXALLeA0173M05; HVVMRXALLmA0207L21
MLOC_32752.1	HVVMRXALLeA0166N13;　HVVMRXALLeA0173M05; HVVMRXALLmA0207L21
MLOC_69709.1	HVVMRXALLeA0076E11;　HVVMRXALLeA0280D23; HVVMRXALLeA0281L17; HVVMRXALLhC0055G09;　HVVMRXALLmA0011E02
MLOC_69711.1	HVVMRXALLmA0011E02;　HVVMRXALLeA0280D23; HVVMRXALLeA0281L17; HVVMRXALLhC0055G09;　HVVMRXALLeA0076E11
MLOC_54309.1	HVVMRXALLeA0076E11;　HVVMRXALLeA0280D23; HVVMRXALLeA0281L17; HVVMRXALLhC0055G09
MLOC_50477.1	HVVMRXALLeA0092H05;　HVVMRXALLmA0011E02; HVVMRXALLrA0378O03
MLOC_48482.1	HVVMRXALLeA0092H05;　HVVMRXALLmA0011E02; HVVMRXALLrA0378O03

MLOC_30901.2	HVVMRXALLeA0092H05;	HVVMRXALLmA0011E02; HVVMRXALLrA0378O03
MLOC_27967.1	HVVMRXALLeA0092H05;	HVVMRXALLmA0011E02; HVVMRXALLrA0378O03
MLOC_30902.2	HVVMRXALLeA0092H05;	HVVMRXALLmA0011E02; HVVMRXALLrA0378O03
MLOC_48460.1	HVVMRXALLeA0092H05;	HVVMRXALLmA0011E02; HVVMRXALLrA0378O03
MLOC_74476.1	HVVMRXALLeA0029D13;	HVVMRXALLeA0092H05; HVVMRXALLrA0378O03
MLOC_2411.1	HVVMRXALLeA0009B12;	HVVMRXALLeA0029D13
MLOC_23382.1	HVVMRXALLeA0009B12;	HVVMRXALLeA0029D13
MLOC_50262.1	HVVMRXALLeA0009B12;	HVVMRXALLeA0029D13
MLOC_20154.2	HVVMRXALLeA0009B12;	HVVMRXALLeA0029D13
MLOC_21287.1	HVVMRXALLeA0009B12;	HVVMRXALLeA0029D13
MLOC_22687.1	HVVMRXALLeA0009B12;	HVVMRXALLeA0029D13
MLOC_784.1	HVVMRXALLeA0009B12;	HVVMRXALLhB0167G05
AK248639.1	HVVMRXALLeA0009B12;	HVVMRXALLhB0167G05; HVVMRXALLmA0142K09
MLOC_27290.1	HVVMRXALLhB0167G05;	HVVMRXALLmA0142K09; HVVMRXALLeA0009B12
MLOC_46593.1	HVVMRXALLhB0167G05;	HVVMRXALLmA0142K09
MLOC_55047.1	HVVMRXALLhB0167G05;	HVVMRXALLmA0142K09
MLOC_18386.1	HVVMRXALLhB0167G05;	HVVMRXALLmA0142K09
MLOC_41937.1	HVVMRX83KhA0214N18;	HVVMRXALLeA0069L08; HVVMRXALLmA0506H07

Table A16: Single marker assays *HvLAX-A*

Marker	Type	forward sequence (5' - 3')	reverse sequence (5' - 3')	enzyme
1_0974*	CAPS	TTGCTGGTTGGGATCTCTTC	ACACTGGCTGACACATGGAA	*Sau3A I*
1_0580	CAPS	CATACCGAAAGTACAGAGCACA	TCACCGAACAGACTCTCAAGG	*Aci I*
1_1198	CAPS	AACCTCCCAGCTGACCTCAT	ACACAATCAGCAGCAACACC	*Fau I*
1_0157	CAPS	ACACCAAAGATGACCTCAAGC	CCTGAAAAGATGCACGGAAT	*HpyCH4 III*
1_0481	CAPS	AGGCTAGACGGCAACGATACT	GACTCCAACAACATCGCAGAG	*Aci I*
2_0524	CAPS	TTCTCTGCCGCATCAACAG	CCGCATGGATACTCACTTCCT	*Taq I*
1_1469*	CAPS	GTGGTGTTCTGGGCAACTTT	ACGGTCATGTGACCCTATCA	*HpyCH4 IV*
2_1148	seq**	ACTACATGTCCGGCTTCCAC	CCTACGTACGACTAGCAAAGCA	/
2_0306	seq**	GCTTCTGCACTTGCCAATTC	GCACAAAGCAAACAGTGCTC	/
2_0713	CAPS	AGCACTCCTAGTGTCCGTGAG	CGGCTCCGCTTCTTATCTTT	*HpyCH4 IV*
Bradi4g43680	CAPS	GAGACAGGACAGAGGCTTGC	ATCCTTCCTTTCCACCCTTG	*BfuA1*
1_1260*	seq*	CTCGGAGCGCCACTTCTTAG	CTGGGTAGATGGCATCCATTG	/
2_0239*	CAPS	TGTAGTCGGGGTGTTGTGTG	AGTGGGTTCTGAGGAAAGCA	*Sau96 I*

*: not the Illumina SNP but located on the same unigene, **sequenced by Sanger sequencing

Table A17: Filter for cosegregating targets with expected SNP frequency within captured pools

Bowman WGS_contig[1]	pos	ref	alt	ratio_lax_1	ratio_lax_2	ratio_lax_3	ratio_lax_4	ratio_lax_5	depth_lax_1	depth_lax_2	depth_lax_3	depth_lax_4	depth_lax_5
contig_221364	6495	C	G	0.57	0	1	0.95	0.59	7	19	5	19	17
contig_859419	2590	C	G	0.55	0	1	0.96	0.57	11	8	9	23	7
contig_13430	1449	C	T	0.43	0	1	1	0.42	7	8	17	11	19
contig_879642	2867	T	C	0.53	0	0.86	0.94	0.47	17	17	14	32	17
contig_1998524	3326	T	G	0.55	0.17	1	0.93	0.58	11	12	6	15	12
contig_108294	2503	A	G	0.6	0	1	0.94	0.45	10	12	8	18	11
contig_108294	2512	T	C	0.55	0	1	1	0.5	11	11	9	18	8
contig_166251	562	T	C	0.61	0.13	0.95	0.97	0.62	18	23	19	35	39
contig_1989744	1326	C	G	0.57	0	1	0.97	0.41	14	13	19	31	17
contig_230144	457	C	A	0.67	0	1	1	0.5	6	7	4	10	8
contig_245492	475	C	T	0.6	0	1	1	0.36	10	8	12	14	14
contig_366956	1131	G	T	0.55	0.05	0.95	0.98	0.49	55	60	60	104	67
contig_366956	1132	C	G	0.54	0.03	0.93	0.94	0.49	56	61	61	108	67
contig_582766	131	G	A	0.5	0	1	1	0.67	4	1	2	6	3
contig_871803	1453	T	C	0.4	0	1	1	0.36	10	11	12	11	14
contig_876951	1295	C	G	0.5	0	1	0.98	0.59	20	19	15	41	27
contig_876951	1995	C	T	0.57	0.12	0.83	1	0.46	7	17	6	13	13
contig_876951	1996	A	G	0.57	0.12	1	1	0.54	7	17	6	13	13
contig_921063	560	A	G	0.57	0	1	0.83	0.5	14	8	10	6	8

ratio = number of reads with alternative base (mutant allele)

depth = read coverage at SNP position

[1] (IBSC, 2012)

Table A18: Filtered candidate targets of read coverage analysis within captured pools

Bowman WGS_contig[1]	contig_length	start	end	length	lax_1	lax_2	lax_3	lax_4	lax_5
contig_1387056	882	412	526	115	5.49	8.07	6.00	10.88	7.50
contig_1387056	882	626	692	67	2.45	1.62	2.12	1.30	7
contig_1387056	882	779	842	64	7.04	6.19	1.79	0.96	5.30
contig_1498435	470	354	421	68	5.65	6.97	1.83	1.38	5.41
contig_1532510	797	337	445	109	6.46	6.60	0.35	1.96	6.57
contig_1532510	797	456	520	65	6.37	2.73	1.85	3.78	3.46
contig_1532510	797	593	670	78	1.53	1.77	2.89	6.23	0.57
contig_1534971	630	550	617	68	5.89	5.79	0.79	1.68	5.02
contig_1764579	240	160	227	68	5.47	5.26	1.58	1.94	5.20
contig_212370	1415	827	903	77	0.34	5.55	1.48	6.71	0.57
contig_212370	1415	1283	1399	117	5.20	5.16	1.92	1.84	6.41
contig_335772	224	41	118	78	5.01	6.22	1.79	1.72	7.25
contig_350085	316	135	237	103	5.33	6.32	1.11	1.11	6.18
contig_380230	686	88	198	111	5.69	5.98	0.55	0.94	6.86
contig_380230	686	316	391	76	5.04	3.30	3.53	4.86	6.13
contig_380230	686	415	676	262	7.26	4.32	5.62	6.49	5.83
contig_523676	220	31	127	97	6.13	6.05	1.89	1.41	7.95
contig_68343	10094	112	203	92	0.23	1.83	0.65	5.86	1.86
contig_68343	10094	3410	3506	97	0.37	5.80	0	0	1.69
contig_68343	10094	3523	3942	420	4.39	14.03	0.38	0.64	7.76
contig_68343	10094	3968	4033	66	0.46	6.13	0	1	0
contig_68343	10094	5216	5285	70	0.85	5.57	1.56	0	2.62
contig_68343	10094	5286	6083	798	2.79	10.35	0.65	0.40	5.13
contig_68343	10094	6103	6714	612	9.53	17.21	0.06	1.63	11.38
contig_872010	5225	4	66	63	7.69	6.61	1.91	0.91	5.61

[1] (IBSC, 2012)

Table A19: Overlapping BACs of FPcontig_2862

BAC_clone	cb_start[1]	cb_end[1]	cumulative sequence	overlapping BAC clones with >2000 bp alignment length and >99% identity
HVVMRXALLmA0117O22	0	82	171357 bp	HVVMRXALLmA0114D14
HVVMRXALLmA0114D14	108	197	171200 bp	HVVMRXALLeA0185O05; HVVMRXALLmA0117O22
HVVMRXALLeA0185O05	157	301	157155 bp	HVVMRXALLmA0114D14; HVVMRXALLrA0187L16
HVVMRXALLrA0187L16	242	427	185401 bp	HVVMRXALLeA0185O05; HVVMRXALLeA0324B05
HVVMRXALLeA0324B05	389	526	213911 bp	HVVMRXALLrA0187L16; HVVMRXALLeA0046E04; HVVMRXALLmA0029K21
HVVMRXALLmA0029K21	446	576	171180 bp	HVVMRXALLeA0324B05; HVVMRXALLmA0255O17; HVVMRXALLeA0046E04
HVVMRXALLeA0046E04	487	634	155472 bp	HVVMRXALLmA0029K21; HVVMRXALLeA0324B05; HVVMRXALLmA0255O17
HVVMRXALLmA0255O17	557	677	181434 bp	HVVMRXALLmA0033O13; HVVMRXALLmA0029K21; HVVMRXALLeA0046E04
HVVMRXALLmA0033O13	663	780	159185 bp	HVVMRXALLhA0071F16; HVVMRXALLmA0255O17
HVVMRXALLhA0071F16	733	867	117817 bp	HVVMRXALLmA0033O13; HVVMRXALLrA0261G02
HVVMRXALLrA0261G02	823	932	102098 bp	HVVMRXALLhA0071F16; HVVMRXALLeA0171I14
HVVMRXALLeA0171I14	921	1042	189691 bp	HVVMRXALLeA0137L10; HVVMRXALLrA0261G02
HVVMRXALLeA0137L10	974	1114	147541 bp	HVVMRXALLmA0433B15; HVVMRXALLeA0171I14
HVVMRXALLmA0433B15	1092	1192	165267 bp	HVVMRXALLeA0137L10; HVVMRX83KhA0024J08
HVVMRX83KHA0024J08	1170	1242	83924 bp	HVVMRXALLmA0012H10; HVVMRXALLmA0433B15
HVVMRXALLmA0012H10	1240	1386	174749 bp	HVVMRX83KhA0024J08; HVVMRXALLmA0072A18
HVVMRXALLmA0072A18	1371	1477	166702 bp	HVVMRXALLmA0014N12; HVVMRXALLmA0012H10
HVVMRXALLmA0014N12	1401	1539	188810 bp	HVVMRXALLeA0258L22; HVVMRXALLmA0072A18
HVVMRXALLeA0258L22	1503	1615	104452 bp	HVVMRXALLmA0201L14; HVVMRXALLmA0014N12
HVVMRXALLmA0201L14	1549	1719	167994 bp	HVVMRXALLeA0258L22; HVVMRX83KhA0157G10
HVVMRX83KHA0157G10	1659	1762	101195 bp	HVVMRXALLmA0201L14;
HVVMRXALLhB0169O19	*1727*	*1793*	*98379 bp*	*no overlap*

[1] consensus bands of High Information Content Fingerprinting (HICF, IBSC, 2012; Ariyadasa et al., 2014)

Table A20: Genes on sequenced BACs of FPC_2862

BAC_clone	HC_genes	LC_genes
HVVMRXALLmA0117O22	/	/
HVVMRXALLmA0114D14	/	/
HVVMRXALLeA0185O05	/	MLOC_8443.1
HVVMRXALLrA0187L16	/	/
HVVMRXALLeA0324B05	/	/
HVVMRXALLmA0029K21	MLOC_61451.6	/
HVVMRXALLeA0046E04	MLOC_61451.6	/
HVVMRXALLmA0255O17	/	/
HVVMRXALLmA0033O13	/	/
HVVMRXALLhA0071F16	/	/
HVVMRXALLrA0261G02	/	/
HVVMRXALLeA0171I14	/	/
HVVMRXALLeA0137L10	AK373675	/
HVVMRXALLmA0433B15	MLOC_69804.2	MLOC_20145.1
HVVMRX83KHA0024J08	MLOC_69804.2	MLOC_20145.1
HVVMRXALLmA0012H10	/	/
HVVMRXALLmA0072A18	/	/
HVVMRXALLmA0014N12	/	/
HVVMRXALLeA0258L22	/	/
HVVMRXALLmA0201L14	MLOC_10658.1	/
HVVMRX83KHA0157G10	MLOC_10658.1	/
HVVMRXALLhB0169O19	/	MLOC_34165.1

Table A21: Differentially expressed genes (log2 fold change <= 2 between Bowman and BW457

HC_gene[1]	normalized mean wild-type[2]			normalized mean mutant[2]			log2FC_gl	log2FC_st	log2FC_cd	adj p=<0.05			annotation[1]
	bowman_gl	bowman_st	bowman_fd	lax_gl	lax_st	lax_cd				gl	st	cd	
MLOC_11814.1	41.91	43.38	33.59	10.53	19.22	40.14	-1.99	-1.17	0.26	X			CsAtPR5
AK361975	40.63	27.96	25.95	11.01	21.26	4.24	-1.88	-0.39	-2.61			X	UDP-Glycosyltransferase superfamily protein LENGTH=480
AK363765	75.25	85.08	113.41	21.48	24.91	45.22	-1.81	-1.77	-1.33	X	X	X	Protein BFR2, putative
AK374864	86.11	67.17	60.63	25.56	33.98	48.27	-1.75	-0.98	-0.33	X			Leucine-rich repeat receptor-like protein kinase
MLOC_81021.1	226.22	247.57	351.36	69.61	97.43	159.98	-1.70	-1.35	-1.14	X	X	X	Mitochondrial transcription termination factor-like
MLOC_76480.1	74.59	79.78	107.03	24.00	73.90	90.64	-1.64	-0.11	-0.24	X			UDP-Glycosyltransferase superfamily protein LENGTH=489
MLOC_70515.0	108.55	128.00	343.59	35.39	107.65	273.96	-1.62	-0.25	-0.33	X			Wound-induced protein 1, putative, expressed
MLOC_77488.1	84.01	220.06	324.35	27.70	175.60	245.58	-1.60	-0.33	-0.40	X			Homeobox-leucine zipper-like protein
MLOC_28382.1	25.77	11.13	22.29	8.54	4.37	4.27	-1.59	-1.35	-2.38			X	Cyst nematode resistance protein-like protein
MLOC_52969.1	53.95	38.89	54.90	19.51	21.72	18.00	-1.47	-0.84	-1.61			X	histidine kinase 5 LENGTH=922
MLOC_5859.1	73.12	97.31	175.52	26.76	55.73	144.93	-1.45	-0.80	-0.28	X			Isoflavone reductase, putative, expressed
MLOC_66260.1	192.30	282.24	664.11	72.03	207.60	530.95	-1.42	-0.44	-0.32	X			YABBY protein
MLOC_27470.1	59.50	21.61	61.58	23.38	7.64	19.54	-1.35	-1.50	-1.66	X			Unknown protein
AK364433	409.78	416.66	351.76	162.42	442.16	295.10	-1.34	0.09	-0.25	X			Sugar phosphate exchanger 2
MLOC_43518.1	44.14	47.68	92.89	17.87	40.62	84.29	-1.30	-0.23	-0.14	X			Calmodulin-like protein
AK375184	707.49	839.28	409.97	289.21	516.28	194.15	-1.29	-0.70	-1.08	X			Poly(RC)-binding protein, putative
MLOC_52345.1	39.83	38.16	44.60	16.29	18.77	14.89	-1.29	-1.02	-1.58			X	histidine kinase 5 LENGTH=922
MLOC_70020.2	122.72	145.04	141.03	50.64	192.05	119.73	-1.28	0.41	-0.24	X			Epoxide hydrolase, putative, expressed
AK376953	172.98	190.25	145.48	72.11	209.38	176.25	-1.26	0.14	0.28	X			Homeobox leucine zipper protein
MLOC_35121.1	176.79	150.76	203.25	73.74	87.74	69.68	-1.26	-0.78	-1.54	X			histidine kinase 5 LENGTH=922
MLOC_42368.1	56.74	54.16	45.72	23.80	25.56	17.04	-1.25	-1.08	-1.42				Histidine kinase cytokinin receptor
MLOC_57248.1	567.54	743.44	625.86	241.88	243.41	263.59	-1.23	-1.61	-1.25	X	X	X	FBD-associated F-box protein
MLOC_41692.1	117.67	106.81	115.05	50.74	51.33	47.09	-1.21	-1.06	-1.29	X	X	X	Histidine kinase cytokinin receptor
AK364775	71.50	271.84	797.45	31.15	186.31	559.07	-1.20	-0.55	-0.51	X			Gibberellin 2-oxidase
MLOC_50854.1	40.15	48.08	56.60	17.51	29.25	19.67	-1.20	-0.72	-1.52			X	Histidine kinase cytokinin receptor
AK354063	499.36	431.26	1133.36	227.34	649.67	905.87	-1.14	0.59	-0.32	X			Homeobox-leucine zipper protein
MLOC_37146.1	396.27	388.78	181.45	182.04	213.06	104.49	-1.12	-0.87	-0.80	X			2-oxoglutarate (2OG) and Fe(II)-dependent oxygenase-like protein
AK370114	86.41	67.05	118.74	39.71	148.96	82.75	-1.12	1.15	-0.52	X	X		Unknown protein
AK361626	3091.51	2863.39	1898.08	1441.98	1247.25	1027.61	-1.10	-1.20	-0.89	X	X	X	Unknown protein
AK353900	535.53	146.68	550.70	250.09	343.57	239.34	-1.10	1.23	-1.20	X	X	X	Histone H3
MLOC_2297.1	90.28	34.19	83.05	42.62	23.93	35.27	-1.08	-0.51	-1.24	X	X	X	ATP-dependent DNA helicase PIF1
MLOC_62152.1	101.07	98.97	134.25	47.85	115.33	123.64	-1.08	0.22	-0.12	X			SPX domain gene 3 LENGTH=245
MLOC_37138.2	629.87	1285.44	6143.10	298.47	647.54	3991.10	-1.08	-0.99	-0.62	X	X	X	Cyclopropane-fatty-acyl-phospholipid synthase
AK367599	67.69	58.20	57.69	32.13	74.68	20.71	-1.07	0.36	-1.48	X			Molybdopterin synthase catalytic subunit
MLOC_71650.1	353.33	402.25	771.40	169.19	445.11	562.02	-1.06	0.15	-0.46	X			unknown protein
MLOC_72621.2	1389.92	1451.84	1665.93	677.33	707.73	902.04	-1.04	-1.04	-0.89	X	X	X	Non-structural maintenance of chromosomes element 1 homolog
MLOC_11367.1	114.02	133.85	143.07	55.73	59.94	104.58	-1.03	-1.16	-0.45	X			Leucine-rich repeat receptor-like protein kinase
MLOC_3822.1	3410.62	2369.02	2066.79	1674.04	2954.86	2117.11	-1.03	0.32	0.03	X			Non-specific lipid-transfer protein
MLOC_35554.2	97.11	111.07	387.67	48.54	124.52	238.72	-1.00	0.16	-0.70	X			Unknown protein
MLOC_65935.3	18671.38	20814.31	12711.54	9414.65	11598.16	8319.32	-0.99	-0.84	-0.61	X	X		Eukaryotic translation initiation factor 3 subunit 9, putative
AK364168	991.95	1011.14	1114.04	502.09	605.01	707.72	-0.98	-0.74	-0.65	X			Bifunctional folylpolyglutamate/dihydrofolate synthasesynthase (FolC)
AK250965.1	216.50	146.28	297.69	110.22	282.30	169.39	-0.97	0.95	-0.81	X			Histone H4
MLOC_52924.1	256.50	189.47	282.41	131.26	236.15	196.48	-0.97	0.32	-0.52	X			P-loop containing nucleoside triphosphate hydrolases superfamily protein
MLOC_16760.1	149.62	92.57	143.37	76.86	126.52	88.04	-0.96	0.45	-0.70	X			Unknown protein
MLOC_52603.1	737.94	729.35	1369.76	380.38	942.56	986.00	-0.96	0.37	-0.47	X			unknown protein
MLOC_61156.1	142.97	147.24	192.40	74.40	120.37	129.65	-0.94	-0.29	-0.57	X			LOB domain-containing protein
AK365916	74.63	65.84	95.93	39.58	42.95	44.17	-0.92	-0.62	-1.12			X	Cyclin-dependent protein kinase-like protein
MLOC_11318.2	3259.29	3143.71	3895.02	1735.21	1596.11	2284.26	-0.91	-0.98	-0.77	X	X	X	GPI ethanolamine phosphate transferase 3
MLOC_6843.1	547.38	193.27	98.25	291.54	480.40	80.77	-0.91	1.31	-0.28		X		Ethylene responsive transcription factor 12
MLOC_65945.1	61.01	134.48	359.32	32.62	58.20	198.16	-0.90	-1.21	-0.86		X		Auxin response factor
MLOC_53792.1	392.89	448.34	394.62	213.00	375.24	366.77	-0.88	-0.26	-0.11	X			Lysine decarboxylase-like protein
AK356252	219.78	250.44	252.14	119.86	104.49	203.78	-0.87	-1.26	-0.31		X		Methyltransferase
MLOC_52879.5	460.05	372.64	632.31	254.88	415.16	778.20	-0.85	0.16	0.30	X			Kinesin-related protein
MLOC_72108.1	189.27	250.04	192.88	105.19	126.99	139.75	-0.85	-0.98	-0.46	X			Unknown protein
AK251852.1	200.23	76.74	265.24	112.06	223.64	134.43	-0.84	1.54	-0.98		X		Histone H2A
MLOC_18499.1	3329.02	4451.45	4006.48	1866.90	3600.94	3378.27	-0.83	-0.31	-0.25			X	Xyloglucan endotransglucosylase hydrolase 2
MLOC_74122.3	2863.69	3123.32	3331.08	1621.19	2010.13	2100.59	-0.82	-0.64	-0.67	X			Tubulin-specific chaperone A
MLOC_59412.1	76.58	34.13	32.90	43.56	78.03	33.91	-0.81	1.19	0.04		X		Protein of unknown function, DUF584 LENGTH=203
AK365885	798.67	440.83	611.58	478.55	780.24	368.73	-0.74	0.82	-0.73		X		Ring finger protein, putative
MLOC_62894.2	512.73	473.29	552.54	316.87	264.51	271.01	-0.69	-0.84	-1.03			X	HAT family dimerisation domain containing protein
AK374582	134.95	178.68	133.91	86.47	72.57	74.97	-0.64	-1.30	-0.84		X		CsAtPR5
MLOC_70988.1	857.72	461.55	690.04	557.19	908.40	538.70	-0.62	0.98	-0.36		X		alpha/beta-Hydrolases superfamily protein LENGTH=330
AK249692.1	1255.24	1402.12	1286.23	840.20	922.98	729.80	-0.58	-0.60	-0.82			X	Unknown protein
MLOC_70352.1	410.02	500.82	440.61	277.78	264.94	320.87	-0.56	-0.92	-0.46		X		ARM repeat-containing protein
MLOC_71824.1	303.26	134.16	339.11	208.66	354.67	172.90	-0.54	1.40	-0.97		X		unknown protein
AK362713	449.82	332.85	793.70	311.19	720.00	626.61	-0.53	1.11	-0.34		X		1-aminocyclopropane-1-carboxylate oxidase-like protein
AK370873	270.62	218.49	277.92	187.35	152.38	144.92	-0.53	-0.52	-0.94			X	5-methylthioadenosine/S-adenosylhomocysteine deaminase
MLOC_51823.1	586.05	418.53	296.95	410.18	745.68	270.98	-0.51	0.83	-0.13		X		F-box protein
MLOC_7524.2	1086.30	715.96	749.96	769.14	1307.33	573.56	-0.50	0.87	-0.39		X		UPA24
MLOC_58853.1	217.39	114.72	54.40	156.39	230.82	49.63	-0.48	1.01	-0.13		X		Ovate protein
MLOC_72040.1	646.75	749.59	619.67	478.10	398.53	359.97	-0.44	-0.91	-0.78		X		Pentatricopeptide repeat-containing protein
AK369281	537.94	429.16	557.87	398.71	758.52	503.27	-0.43	0.82	-0.15		X		Long chain fatty alcohol oxidase, putative
MLOC_17022.2	3995.67	4971.39	5392.59	3023.45	2201.11	4436.19	-0.40	-1.18	-0.28		X		Kinesin-4
MLOC_64667.1	591.96	382.62	601.05	461.42	847.92	433.44	-0.36	1.15	-0.47		X		HXXXD-type acyl-transferase family protein LENGTH=426
AK361731	84.53	67.61	138.53	67.59	57.88	66.15	-0.32	-0.22	-1.07			X	RNA polymerase sigma factor 2
MLOC_67515.1	358.48	301.33	455.10	287.21	165.47	337.44	-0.32	-0.86	-0.43		X		CST complex subunit CTC1
MLOC_38753.1	424.11	329.97	718.27	349.24	170.23	581.38	-0.28	-0.95	-0.31		X		DNA polymerase epsilon subunit 2
AK366398	158.84	84.52	61.20	133.20	120.42	143.47	-0.25	0.51	1.23			X	Transcription factor, putative
MLOC_71159.1	467.36	592.56	865.33	393.09	347.79	690.50	-0.25	-0.77	-0.33		X		Receptor-like kinase
MLOC_65526.2	405.02	297.82	614.25	343.83	163.57	453.54	-0.24	-0.86	-0.44		X		Kinesin-related protein
MLOC_60773.1	4729.13	3918.74	3321.07	4063.70	6728.15	3159.84	-0.22	0.78	-0.07		X		unknown proteinLENGTH=116
MLOC_16485.1	86.12	131.76	547.99	75.46	63.67	209.04	-0.19	-1.05	-1.39		X	X	Unknown protein
AK248373.1	144.89	83.22	69.85	127.46	245.49	58.94	-0.18	1.56	-0.25		X		Calcium-dependent protein kinase 2
MLOC_2703.3	25.56	54.79	213.89	23.40	16.49	153.11	-0.13	-1.73	-0.48		X		Flavin monooxygenase-like protein
MLOC_73503.1	4657.03	5650.75	8328.70	4408.35	3260.36	7385.60	-0.08	-0.79	-0.08		X		Mini-chromosome maintenance complex-binding protein
MLOC_60135.1	993.27	792.20	357.30	941.85	895.34	806.55	-0.08	0.18	1.17			X	BZIP transcription factor
MLOC_74058.1	11.69	13.21	50.81	11.75	9.72	13.09	0.01	-0.44	-1.96			X	auxin response factor 4 LENGTH=788
MLOC_15452.1	202.89	614.94	1007.07	216.39	315.56	1129.61	0.09	-0.96	0.17		X		Rhamnogalacturonate lyase
MLOC_9998.3	745.95	559.71	333.10	874.98	824.16	609.56	0.23	0.56	0.87				Zinc finger protein CONSTANS-like protein
MLOC_2492.1	58.09	43.68	21.18	70.02	59.25	50.41	0.27	0.44	1.25			X	TTF-type zinc finger protein with HAT dimerization domain
MLOC_61857.1	27.06	17.56	18.06	44.54	60.01	35.75	0.72	1.77	0.99		X	X	Malectin/receptor-like protein kinase family protein LENGTH=873
AK367749	0.38	17.29	17.26	0.63	6.54	60.56	0.73	-1.40	1.81				Germin-like protein 4d
MLOC_15119.2	818.67	783.82	1012.21	1437.25	1087.42	1155.88	0.81	0.47	0.19	X			MORC family CW-type zinc finger protein
MLOC_55987.4	467.21	384.44	276.42	893.40	526.43	472.70	0.94	0.45	0.77	X			DNA binding protein, putative
MLOC_63436.1	778.05	1354.51	855.47	1488.45	1359.97	1021.35	0.94	0.01	0.26	X			BZIP transcription factor
AK249203.1	477.89	504.72	679.07	924.10	440.63	650.54	0.95	-0.20	-0.06	X			Ribulose bisphosphate carboxylase/oxygenase activase
AK369747	1522.57	1946.07	1019.96	3321.08	3716.80	1946.50	1.13	0.93	0.93	X			unknown protein
AK371281	61.39	80.11	74.56	134.73	153.89	140.63	1.13	0.94	0.92	X			cDNA clone:J013058P10, full insert sequence
MLOC_72738.1	59.74	60.26	63.75	133.84	89.41	104.51	1.16	0.57	0.71	X			Pentatricopeptide repeat-containing protein, putative
MLOC_51950.1	663.31	595.14	416.70	1669.26	1691.73	1321.40	1.33	1.51	1.66	X	X	X	NBS-LRR disease resistance protein homologue
MLOC_7060.1	74.62	65.85	64.54	188.31	142.11	185.73	1.34	1.11	1.53	X	X	X	Fructose-6-phosphate 1-phosphotransferase
MLOC_52105.1	68.02	51.88	79.83	172.65	91.60	165.54	1.34	0.82	1.05	X	X	X	Zinc finger RAD18 domain-containing protein C1orf124-like
MLOC_73607.2	569.82	603.32	590.08	1537.80	1537.55	1960.52	1.43	1.35	1.73	X	X	X	Dentin matrix acidic phosphoprotein 1, putative
MLOC_34478.1	108.81	123.42	74.77	296.43	146.66	126.16	1.45	0.25	0.75	X			Octicosapeptide/Phox/Bem1p and tetratricopeptide repeat domain-containing
MLOC_12726.1	27.58	36.46	20.33	95.09	52.83	26.92	1.79	0.53	0.41	X			FBD-associated F-box protein LENGTH=459

[1] High confidence gene model with predicted annotation (IBSC, 2012);

[2] normalized read numbers of DESeq analysis for the developmental stages of gl = glumen primordia, st = stamen primordia and cd = complete dveloped spikes

less than 50 reads/gene

9. References

Abe, A., Kosugi, S., Yoshida, K., Natsume, S., Takagi, H., Kanzaki, H., Matsumura, H., Yoshida, K., Mitsuoka, C., Tamiru, M., Innan, H., Cano, L., Kamoun, S., and Terauchi, R. (2012). Genome sequencing reveals agronomically important loci in rice using MutMap. Nature biotechnology **30,** 174-178.

Adamski, N.M., Anastasiou, E., Eriksson, S., O'Neill, C.M., and Lenhard, M. (2009). Local maternal control of seed size by KLUH/CYP78A5-dependent growth signaling. Proceedings of the National Academy of Sciences of the United States of America **106,** 20115-20120.

Ahn, S., Anderson, J.A., Sorrells, M.E., and Tanksley, S.D. (1993). Homoeologous relationships of rice, wheat and maize chromosomes. Molecular & general genetics : MGG **241,** 483-490.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. (1990). Basic Local Alignment Search Tool. J Mol Biol **215,** 403-410.

Anastasiou, E., Kenz, S., Gerstung, M., MacLean, D., Timmer, J., Fleck, C., and Lenhard, M. (2007). Control of plant organ size by KLUH/CYP78A5-dependent intercellular signaling. Developmental cell **13,** 843-856.

Anders, S., and Huber, W. (2010). Differential expression analysis for sequence count data. Genome biology **11**.

Anders, S., Pyl, P.T., and Huber, W. (2014). HTSeq-a Python framework to work with high-throughput sequencing data. Bioinformatics.

Apweiler, R., Martin, M.J., O'Donovan, C., Magrane, M., Alam-Faruque, Y., Alpi, E., Antunes, R., Arganiska, J., Casanova, E.B., Bely, B., Bingley, M., Bonilla, C., Britto, R., Bursteinas, B., Chan, W.M., Chavali, G., Cibrian-Uhalte, E., Da Silva, A., De Giorgi, M., Dimmer, E., Fazzini, F., Gane, P., Fedotov, A., Castro, L.G., Garmiri, P., Hatton-Ellis, E., Hieta, R., Huntley, R., Jacobsen, J., Jones, R., Legge, D., Liu, W.D., Luo, J., MacDougall, A., Mutowo, P., Nightingale, A., Orchard, S., Patient, S., Pichler, K., Poggioli, D., Pundir, S., Pureza, L., Qi, G.Y., Rosanoff, S., Sawford, T., Sehra, H., Turner, E., Volynkin, V., Wardell, T., Watkins, X., Zellner, H., Corbett, M., Donnelly, M., van Rensburg, P., Goujon, M., McWilliam, H., Lopez, R., Xenarios, I., Bougueleret, L., Bridge, A., Poux, S., Redaschi, N., Auchincloss, A., Axelsen, K., Bansal, P., Baratin, D., Binz, P.A., Blatter, M.C., Boeckmann, B., Bolleman, J., Boutet, E., Breuza, L., de Castro, E., Cerutti, L., Coudert, E., Cuche, B., Doche, M., Dornevil, D., Duvaud, S., Estreicher, A., Famiglietti, L., Feuermann, M., Gasteiger, E., Gehant, S., Gerritsen, V., Gos, A., Gruaz-Gumowski, N., Hinz, U., Hulo, C., James, J., Jungo, F., Keller, G., Lara, V., Lemercier, P., Lew, J., Lieberherr, D., Martin, X., Masson, P., Morgat, A., Neto, T., Paesano, S., Pedruzzi, I., Pilbout, S., Pozzato, M., Pruess, M., Rivoire, C., Roechert, B., Schneider, M., Sigrist, C., Sonesson, K., Staehli, S., Stutz, A., Sundaram, S., Tognolli, M., Verbregue, L., Veuthey, A.L., Zerara, M., Wu, C.H., Arighi, C.N., Arminski, L., Chen, C.M., Chen, Y.X., Huang, H.Z., Kukreja, A., Laiho, K., McGarvey, P., Natale, D.A., Natarajan, T.G., Roberts, N.V., Suzek, B.E., Vinayaka, C.R., Wang, Q.H., Wang, Y.Q., Yeh, L.S., Yerramalla, M.S., Zhang, J., and Consortium, U. (2013). Update on activities at the Universal Protein Resource (UniProt) in 2013. Nucleic acids research **41,** D43-D47.

Ariyadasa, R., and Stein, N. (2012). Advances in BAC-Based Physical Mapping and Map Integration Strategies in Plants. Journal of Biomedicine and Biotechnology.

Ariyadasa, R., Mascher, M., Nussbaumer, T., Schulte, D., Frenkel, Z., Poursarebani, N., Zhou, R., Steuernagel, B., Gundlach, H., Taudien, S., Felder, M., Platzer, M., Himmelbach, A., Schmutzer, T., Hedley, P.E., Muehlbauer, G.J., Scholz, U., Korol, A., Mayer, K.F., Waugh, R., Langridge, P., Graner, A., and Stein, N. (2014). A sequence-ready physical map of barley anchored genetically by two million single-nucleotide polymorphisms. Plant physiology **164,** 412-423.

Bandelt, H.J., Forster, P., and Rohl, A. (1999). Median-joining networks for inferring intraspecific phylogenies. Molecular biology and evolution **16,** 37-48.

Begun, D.J., and Aquadro, C.F. (1992). Levels of Naturally-Occurring DNA Polymorphism Correlate with Recombination Rates in Drosophila-Melanogaster. Nature **356,** 519-520.

Benjamini, Y., and Hochberg, Y. (1995). Controlling the False Discovery Rate - a Practical and Powerful Approach to Multiple Testing. J Roy Stat Soc B Met **57,** 289-300.

Biswas, S., and Akey, J.M. (2006). Genomic insights into positive selection. Trends Genet **22,** 437-446.

Bolduc, N., and Hake, S. (2009). The Maize Transcription Factor KNOTTED1 Directly Regulates the Gibberellin Catabolism Gene ga2ox1. The Plant cell **21,** 1647-1658.

Bolduc, N., Yilmaz, A., Mejia-Guerra, M.K., Morohashi, K., O'Connor, D., Grotewold, E., and Hake, S. (2012). Unraveling the KNOTTED1 regulatory network in maize meristems. Genes & development **26,** 1685-1690.

Bothmer, R.v., Sato, K., Komatsuda, T., Yasuda, S., Fischbeck. (2003). "Chapter 2: The domestication of cultivated barley". in Diversity in Barley (Hordeum vulgare). Developments in Plant Genetics and Breeding **7,** 9-27.

Botstein, D., White, R.L., Skolnick, M., and Davis, R.W. (1980). Construction of a Genetic-Linkage Map in Man Using Restriction Fragment Length Polymorphisms. American journal of human genetics **32,** 314-331.

Chaudhury, A.M., Letham, S., Craig, S., and Dennis, E.S. (1993). Amp1 - a Mutant with High Cytokinin Levels and Altered Embryonic Pattern, Faster Vegetative Growth, Constitutive Photomorphogenesis and Precocious Flowering. Plant Journal **4,** 907-916.

Chevreux, B., Wetter, T., Suhai, S. (1999). Genome Sequence Assembly Using Trace Signals and Additional Sequence Information. Computer Science and Biology: Proceedings of the German Conference on Bioinformatics (GCB) **99**.

Chono, M., Honda, I., Zeniya, H., Yoneyama, K., Saisho, D., Takeda, K., Takatsuto, S., Hoshino, T., and Watanabe, Y. (2003). A semidwarf phenotype of barley uzu results from a nucleotide substitution in the gene encoding a putative brassinosteroid receptor. Plant physiology **133,** 1209-1219.

Clark, R.M., Schweikert, G., Toomajian, C., Ossowski, S., Zeller, G., Shinn, P., Warthmann, N., Hu, T.T., Fu, G., Hinds, D.A., Chen, H., Frazer, K.A., Huson, D.H., Scholkopf, B., Nordborg, M., Ratsch, G., Ecker, J.R., and Weigel, D. (2007). Common sequence polymorphisms shaping genetic diversity in Arabidopsis thaliana. Science **317,** 338-342.

Close, T.J., Bhat, P.R., Lonardi, S., Wu, Y., Rostoks, N., Ramsay, L., Druka, A., Stein, N., Svensson, J.T., Wanamaker, S., Bozdag, S., Roose, M.L., Moscou, M.J., Chao, S., Varshney, R.K., Szucs, P., Sato, K., Hayes, P.M., Matthews, D.E., Kleinhofs, A., Muehlbauer, G.J., DeYoung, J., Marshall, D.F., Madishetty, K., Fenton, R.D., Condamine, P., Graner, A., and Waugh, R. (2009). Development and implementation of high-throughput SNP genotyping in barley. BMC genomics **10,** 582.

Coen, E.S., and Meyerowitz, E.M. (1991). The war of the whorls: genetic interactions controlling flower development. Nature **353,** 31-37.

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., and Waugh, R. (2012). Natural variation in a homolog of Antirrhinum CENTRORADIALIS contributed to spring growth habit and environmental adaptation in cultivated barley. Nature genetics **44,** 1388-1392.

Couzigou, J.M., Zhukov, V., Mondy, S., Abu el Heba, G., Cosson, V., Ellis, T.H.N., Ambrose, M., Wen, J.Q., Tadege, M., Tikhonovich, I., Mysore, K.S., Putterill, J., Hofer, J., Borisov, A.Y., and Ratet, P. (2012). NODULE ROOT and COCHLEATA Maintain Nodule Development and Are Legume Orthologs of Arabidopsis BLADE-ON-PETIOLE Genes. The Plant cell **24,** 4498-4510.

Dai, M., Hu, Y., Zhao, Y., Liu, H., and Zhou, D.X. (2007). A WUSCHEL-LIKE HOMEOBOX gene represses a YABBY gene expression required for rice leaf development. Plant physiology **144,** 380-390.

Dockter, C., Gruszka, D., Braumann, I., Druka, A., Druka, I., Franckowiak, J., Gough, S.P., Janeczko, A., Kurowska, M., Lundqvist, J., Lundqvist, U., Marzec, M., Matyszczak, I., Muller, A.H.,

Oklestkova, J., Schulz, B., Zakhrabekova, S., and Hansson, M. (2014). Induced Variations in Brassinosteroid Genes Define Barley Height and Sturdiness, and Expand the "Green Revolution" Genetic Toolkit. Plant physiology.

Dolezel, J., and Bartos, J. (2005). Plant DNA flow cytometry and estimation of nuclear genome size. Annals of botany **95,** 99-110.

Donald, C.M. (1968). Breeding of Crop Ideotypes. Euphytica **17,** 385-&.

Doyle, J.J., Doyle, J.L. (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., and Waugh, R. (2011). Genetic dissection of barley morphology and development. Plant physiology **155,** 617-627.

Edgar, R.C. (2004a). MUSCLE: a multiple sequence alignment method with reduced time and space complexity. BMC bioinformatics **5,** 1-19.

Edgar, R.C. (2004b). MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic acids research **32,** 1792-1797.

Elshire, R.J., Glaubitz, J.C., Sun, Q., Poland, J.A., Kawamoto, K., Buckler, E.S., and Mitchell, S.E. (2011). A Robust, Simple Genotyping-by-Sequencing (GBS) Approach for High Diversity Species. PloS one **6**.

Ewing, B., and Green, P. (1998). Base-calling of automated sequencer traces using phred. II. Error probabilities. Genome research **8,** 186-194.

Ewing, B., Hillier, L., Wendl, M.C., and Green, P. (1998). Base-calling of automated sequencer traces using phred. I. Accuracy assessment. Genome research **8,** 175-185.

Flavell, R.B., Bennett, M.D., Smith, J.B., and Smith, D.B. (1974). Genome size and the proportion of repeated nucleotide sequence DNA in plants. Biochemical genetics **12,** 257-269.

Franckowiak, J.D. (1995). Notes on linkage drag in Bowman backcross derived lines of spring barley. Barley Genetics Newsletter **24,** 24-63.

Franckowiak, J.D.L., U. (2010). Descriptions of barley genetic stocks for 2010. Genetics Newsletter **40,** 45-177.

Freeling, M. (2001). Grasses as a single genetic system. Reassessment 2001. Plant physiology **125,** 1191-1197.

Fu, Y.X., and Li, W.H. (1993). Statistical tests of neutrality of mutations. Genetics **133,** 693-709.

Gallavotti, A., Long, J.A., Stanfield, S., Yang, X., Jackson, D., Vollbrecht, E., and Schmidt, R.J. (2010). The control of axillary meristem fate in the maize ramosa pathway. Development **137,** 2849-2856.

Galvao, V.C., Nordstrom, K.J., Lanz, C., Sulz, P., Mathieu, J., Pose, D., Schmid, M., Weigel, D., and Schneeberger, K. (2012). Synteny-based mapping-by-sequencing enabled by targeted enrichment. The Plant journal : for cell and molecular biology **71,** 517-526.

Garber, M., Grabherr, M.G., Guttman, M., and Trapnell, C. (2011). Computational methods for transcriptome annotation and quantification using RNA-seq. Nature methods **8,** 469-477.

Gottwald, S., Bauer, P., Komatsuda, T., Lundqvist, U., and Stein, N. (2009). TILLING in the two-rowed barley cultivar 'Barke' reveals preferred sites of functional diversity in the gene HvHox1. BMC research notes **2,** 258.

Graner, A., Jahoor, A., Schondelmaier, J., Siedler, H., Pillen, K., Fischbeck, G., Wenzel, G., and Herrmann, R.G. (1991). Construction of an Rflp Map of Barley. Theoretical and Applied Genetics **83,** 250-256.

Griffiths, A.J.F., Wessler, S.R., Lewontin, Gelbart, W.M., Suzuki, D.T., Miller, J.H. . (2004). Chapter 4: Eukaryote Chromosome Mapping by Recombination. in: An Introduction to Genetic Analysis. W.H.Freeman & Co Ltd **8,** 115-150.

Gurushidze, M., Hensel, G., Hiekel, S., Schedel, S., Valkov, V., and Kumlehn, J. (2014). True-Breeding Targeted Gene Knock-Out in Barley Using Designer TALE-Nuclease in Haploid Cells. PloS one **9**.

Ha, C.M., Jun, J.H., Nam, H.G., and Fletcher, J.C. (2004). BLADE-ON-PETIOLE1 encodes a BTB/POZ domain protein required for leaf morphogenesis in Arabidopsis thaliana. Plant and Cell Physiology **45,** 1361-1370.

Ha, C.M., Jun, J.H., Nam, H.G., and Fletcher, J.C. (2007). BLADE-ON-PETIOLE1 and 2 control Arabidopsis lateral organ fate through regulation of LOB domain and adaxial-abaxial polarity genes. The Plant cell **19,** 1809-1825.

Hall, B.G. (2013). Building Phylogenetic Trees from Molecular Data with MEGA. Molecular biology and evolution **30,** 1229-1235.

Harlan, H.V., Pope, M.N. (1922). Many-noded dwarf barley. J Hered **13, no. 6,** 269-273.

Hartwig, B., James, G.V., Konrad, K., Schneeberger, K., and Turck, F. (2012). Fast isogenic mapping-by-sequencing of ethyl methanesulfonate-induced mutant bulks. Plant physiology **160,** 591-600.

Haseneyer, G., Stracke, S., Paul, C., Einfeldt, C., Broda, A., Piepho, H.P., Graner, A., and Geiger, H.H. (2010). Population structure and phenotypic variation of a spring barley world collection set up for association studies. Plant Breeding **129,** 271-279.

Hedden, P. (2003). The genes of the Green Revolution. Trends Genet **19,** 5-9.

Helliwell, C.A., Chin-Atkins, A.N., Wilson, I.W., Chapple, R., Dennis, E.S., and Chaudhury, A. (2001). The Arabidopsis AMP1 gene encodes a putative glutamate carboxypeptidase. The Plant cell **13,** 2115-2125.

Hepworth, S.R., Zhang, Y.L., McKim, S., Li, X., and Haughn, G. (2005). BLADE-ON-PETIOLE-dependent signaling controls leaf and floral patterning in Arabidopsis. The Plant cell **17,** 1434-1448.

Hill, J.P., and Lord, E.M. (1990). A Method for Determining Plastochron Indexes during Heteroblastic Shoot Growth. Am J Bot **77,** 1491-1497.

Himmelbach, A., Knauft, M., Stein, N. (2014). Plant Sequence Capture Optimised for Illumina Sequencing. Bio-protocol **4,** Bio-protocol.

Horigome, A., Nagasawa, N., Ikeda, K., Ito, M., Itoh, J.I., and Nagato, Y. (2009). Rice OPEN BEAK is a negative regulator of class 1 knox genes and a positive regulator of class B floral homeotic gene. Plant Journal **58,** 724-736.

Houston, K., Druka, A., Bonar, N., Macaulay, M., Lundqvist, U., Franckowiak, J., Morgante, M., Stein, N., and Waugh, R. (2012). Analysis of the barley bract suppression gene Trd1. Theoretical and Applied Genetics **125,** 33-45.

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

Hricová, A., Robles, P., Quesada, V. (2010). "Chapter 18: Unravelling Gene Function Trough Mutatgenesis". in: Molecular Techniques in Crop Improvement. 2nd Edition. Springer Science**,** 437-467.

Huang, X.H., Feng, Q., Qian, Q., Zhao, Q., Wang, L., Wang, A.H., Guan, J.P., Fan, D.L., Weng, Q.J., Huang, T., Dong, G.J., Sang, T., and Han, B. (2009). High-throughput genotyping by whole-genome resequencing. Genome research **19,** 1068-1076.

Hussien, A., Tavakol, E., Horner, D.S., Munoz-Amatriain, M., Muehlbauer, G.J., and Rossini, L. (2014). Genetics of Tillering in Rice and Barley. Plant Genome-Us **7**.

IBSC, International Barley Genome Sequencing Consortium. (2012). A physical, genetic and functional sequence assembly of the barley genome. Nature **491l,** 711-716.

Jain, M., Tyagi, A.K., and Khurana, J.P. (2008). Genome-wide identification, classification, evolutionary expansion and expression analyses of homeobox genes in rice. Febs J **275,** 2845-2861.

James, G.V., Patel, V., Nordstrom, K.J., Klasen, J.R., Salome, P.A., Weigel, D., and Schneeberger, K. (2013). User guide for mapping-by-sequencing in Arabidopsis. Genome biology **14,** R61.

Jiao, Y., Wang, Y., Xue, D., Wang, J., Yan, M., Liu, G., Dong, G., Zeng, D., Lu, Z., Zhu, X., Qian, Q., and Li, J. (2010). Regulation of OsSPL14 by OsmiR156 defines ideal plant architecture in rice. Nature genetics **42,** 541-544.

Joung, J.K., and Sander, J.D. (2013). INNOVATION TALENs: a widely applicable technology for targeted genome editing. Nat Rev Mol Cell Bio **14,** 49-55.

Jun, J.H., Ha, C.M., and Fletcher, J.C. (2010). BLADE-ON-PETIOLE1 coordinates organ determinacy and axial polarity in arabidopsis by directly activating ASYMMETRIC LEAVES2. The Plant cell **22,** 62-76.

Kai, K., Hashidzume, H., Yoshimura, K., Suzuki, H., Sakurai, N., Shibata, D., and Ohta, D. (2009). Metabolomics for the characterization of cytochromes P450-dependent fatty acid hydroxylation reactions in Arabidopsis. Plant Biotechnol **26,** 175-182.

Karim, M.R., Hirota, A., Kwiatkowska, D., Tasaka, M., and Aida, M. (2009). A role for Arabidopsis PUCHI in floral meristem identity and bract suppression. The Plant cell **21,** 1360-1372.

Kawakatsu, T., Itoh, J., Miyoshi, K., Kurata, N., Alvarez, N., Veit, B., and Nagato, Y. (2006). PLASTOCHRON2 regulates leaf initiation and maturation in rice. The Plant cell **18,** 612-625.

Kawakatsu, T., Taramino, G., Itoh, J.I., Allen, J., Sato, Y., Hong, S.K., Yule, R., Nagasawa, N., Kojima, M., Kusaba, M., Sakakibara, H., Sakai, H., and Nagato, Y. (2009). PLASTOCHRON3/GOLIATH encodes a glutamate carboxypeptidase required for proper development in rice. Plant Journal **58,** 1028-1040.

Kellogg, E.A. (2001). Evolutionary history of the grasses. Plant physiology **125,** 1198-1205.

Khan, M., Xu, H.S., and Hepworth, S.R. (2014). BLADE-ON-PETIOLE genes: Setting boundaries in development and defense. Plant Sci **215,** 157-171.

Khan, M., Xu, M.L., Murmu, J., Tabb, P., Liu, Y.Y., Storey, K., McKim, S.M., Douglas, C.J., and Hepworth, S.R. (2012). Antagonistic Interaction of BLADE- ON-PETIOLE1 and 2 with BREVIPEDICELLUS and PENNYWISE Regulates Arabidopsis Inflorescence Architecture. Plant physiology **158,** 946-960.

Khush, G.S. (1995). Breaking the Yield Frontier of Rice. Geo Journal **35.3,** 329-332.

Kim, D., Pertea, G., Trapnell, C., Pimentel, H., Kelley, R., and Salzberg, S.L. (2013). TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. Genome biology **14,** R36.

Kirby, E.J.M., Margaret Appleyard. (1987). Cereal development guide. Kenilworth, England : Arable Unit, National Agricultural Centre **2nd ed.**

Kircher, M., Sawyer, S., and Meyer, M. (2012). Double indexing overcomes inaccuracies in multiplex sequencing on the Illumina platform. Nucleic acids research **40**.

Komatsuda, T., Pourkheirandish, M., He, C.F., Azhaguvel, P., Kanamori, H., Perovic, D., Stein, N., Graner, A., Wicker, T., Tagiri, A., Lundqvist, U., Fujimura, T., Matsuoka, M., Matsumoto, T., and Yano, M. (2007). Six-rowed barley originated from a mutation in a homeodomain-leucine zipper I-class homeobox gene. Proceedings of the National Academy of Sciences of the United States of America **104,** 1424-1429.

Konieczny, A., and Ausubel, F.M. (1993). A Procedure for Mapping Arabidopsis Mutations Using Codominant Ecotype-Specific Pcr-Based Markers. Plant Journal **4,** 403-410.

Koppolu, R., Anwar, N., Sakuma, S., Tagiri, A., Lundqvist, U., Pourkheirandish, M., Rutten, T., Seiler, C., Himmelbach, A., Ariyadasa, R., Youssef, H.M., Stein, N., Sreenivasulu, N., Komatsuda, T., and Schnurbusch, T. (2013). Six-rowed spike4 (Vrs4) controls spikelet determinacy and row-type in barley. Proceedings of the National Academy of Sciences of the United States of America **110,** 13198-13203.

Koressaar, T., and Remm, M. (2007). Enhancements and modifications of primer design program Primer3. Bioinformatics **23,** 1289-1291.

Krizek, B.A., and Fletcher, J.C. (2005). Molecular mechanisms of flower development: an armchair guide. Nature reviews. Genetics **6,** 688-698.

Kunzel, G., Korzun, L., and Meister, A. (2000). Cytologically integrated physical restriction fragment length polymorphism maps for the barley genome based on translocation breakpoints. Genetics **154,** 397-412.

Kyozuka, J., and Shimamoto, K. (2002). Ectopic expression of OsMADS3, a rice ortholog of AGAMOUS, caused a homeotic transformation of lodicules to stamens in transgenic rice plants. Plant & cell physiology **43,** 130-135.

Lamesch, P., Berardini, T.Z., Li, D.H., Swarbreck, D., Wilks, C., Sasidharan, R., Muller, R., Dreher, K., Alexander, D.L., Garcia-Hernandez, M., Karthikeyan, A.S., Lee, C.H., Nelson, W.D., Ploetz, L., Singh, S., Wensel, A., and Huala, E. (2012). The Arabidopsis Information Resource (TAIR): improved gene annotation and new tools. Nucleic acids research **40,** D1202-D1210.

Landegren, U., Nilsson, M., and Kwok, P.Y. (1998). Reading bits of genetic information: methods for single-nucleotide polymorphism analysis. Genome research **8,** 769-776.

Larsson, H.E.B. (1985a). Linkage Studies with Genetic-Markers and Some Laxatum Barley Mutants. Hereditas **103,** 269-279.

Larsson, H.E.B. (1985b). Morphological Analysis of Laxatum Barley Mutants. Hereditas **103,** 239-253.

Laurie, D.A., Pratchett, N., Allen, R.L., and Hantke, S.S. (1996). RFLP mapping of the barley homeotic mutant lax-a. TAG. Theoretical and applied genetics. Theoretische und angewandte Genetik **93,** 81-85.

Lewis, L.A., and McCourt, R.M. (2004). Green algae and the origin of land plants. Am J Bot **91,** 1535-1556.

Li, H. (2011). A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. Bioinformatics **27,** 2987-2993.

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

Li, X., Song, Y., Century, K., Straight, S., Ronald, P., Dong, X., Lassner, M., and Zhang, Y. (2001). A fast neutron deletion mutagenesis-based reverse genetics system for plants. The Plant journal : for cell and molecular biology **27,** 235-242.

Librado, P., and Rozas, J. (2009). DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics **25,** 1451-1452.

Liu, S.Z., Yeh, C.T., Tang, H.M., Nettleton, D., and Schnable, P.S. (2012). Gene Mapping via Bulked Segregant RNA-Seq (BSR-Seq). PloS one **7**.

Lombardo, F., and Yoshida, H. (2015). Interpreting lemma and palea homologies: a point of view from rice floral mutants. Front Plant Sci **6**.

Lu, Z.F., Yu, H., Xiong, G.S., Wang, J., Jiao, Y.Q., Liu, G.F., Jing, Y.H., Meng, X.B., Hu, X.M., Qian, Q., Fu, X.D., Wang, Y.H., and Li, J.Y. (2013). Genome-Wide Binding Analysis of the Transcription Activator IDEAL PLANT ARCHITECTURE1 Reveals a Complex Network Regulating Rice Plant Architecture. The Plant cell **25,** 3743-3759.

Lundquist, U. (2009). Eighty years of scandinavian barley mutation genetics and breeding. in. In Induced plant Mutations in the Genomics Era, Food and Agriculture Organization of the United Nations., 39-43.

Lynch, M., and Conery, J.S. (2000). The evolutionary fate and consequences of duplicate genes. Science **290,** 1151-1155.

Marchler-Bauer, A., Lu, S., Anderson, J.B., Chitsaz, F., Derbyshire, M.K., DeWeese-Scott, C., Fong, J.H., Geer, L.Y., Geer, R.C., Gonzales, N.R., Gwadz, M., Hurwitz, D.I., Jackson, J.D., Ke, Z., Lanczycki, C.J., Lu, F., Marchler, G.H., Mullokandov, M., Omelchenko, M.V., Robertson, C.L., Song, J.S., Thanki, N., Yamashita, R.A., Zhang, D., Zhang, N., Zheng, C., and Bryant, S.H. (2011). CDD: a Conserved Domain Database for the functional annotation of proteins. Nucleic acids research **39,** D225-229.

Mascher, M., Wu, S.Y., St Amand, P., Stein, N., and Poland, J. (2013a). Application of Genotyping-by-Sequencing on Semiconductor Sequencing Platforms: A Comparison of Genetic and Reference-Based Marker Ordering in Barley. PloS one **8**.

Mascher, M., Jost, M., Kuon, J.E., Himmelbach, A., Assfalg, A., Beier, S., Scholz, U., Graner, A., and Stein, N. (2014). Mapping-by-sequencing accelerates forward genetics in barley. Genome biology **15**.

Mascher, M., Muehlbauer, G.J., Rokhsar, D.S., Chapman, J., Schmutz, J., Barry, K., Munoz-Amatriain, M., Close, T.J., Wise, R.P., Schulman, A.H., Himmelbach, A., Mayer, K.F.X., Scholz, U., Poland, J.A., Stein, N., and Waugh, R. (2013b). Anchoring and ordering NGS contig assemblies by population sequencing (POPSEQ). Plant Journal **76,** 718-727.

Mascher, M., Richmond, T.A., Gerhardt, D.J., Himmelbach, A., Clissold, L., Sampath, D., Ayling, S., Steuernagel, B., Pfeifer, M., D'Ascenzo, M., Akhunov, E.D., Hedley, P.E., Gonzales, A.M., Morrell, P.L., Kilian, B., Blattner, F.R., Scholz, U., Mayer, K.F., Flavell, A.J., Muehlbauer, G.J., Waugh, R., Jeddeloh, J.A., and Stein, N. (2013c). Barley whole exome capture: a tool for genomic research in the genus Hordeum and beyond. The Plant journal : for cell and molecular biology **76,** 494-505.

Matsumoto, T., Tanaka, T., Sakai, H., Amano, N., Kanamori, H., Kurita, K., Kikuta, A., Kamiya, K., Yamamoto, M., Ikawa, H., Fujii, N., Hori, K., Itoh, T., and Sato, K. (2011). Comprehensive sequence analysis of 24,783 barley full-length cDNAs derived from 12 clone libraries. Plant physiology **156,** 20-28.

**Matsumoto, T., Wu, J.Z., Kanamori, H., Katayose, Y., Fujisawa, M., Namiki, N., Mizuno, H., Yamamoto, K., Antonio, B.A., Baba, T., Sakata, K., Nagamura, Y., Aoki, H., Arikawa, K., Arita, K., Bito, T., Chiden, Y., Fujitsuka, N., Fukunaka, R., Hamada, M., Harada, C., Hayashi, A., Hijishita, S., Honda, M., Hosokawa, S., Ichikawa, Y., Idonuma, A., Iijima, M., Ikeda, M., Ikeno, M., Ito, K., Ito, S., Ito, T., Ito, Y., Ito, Y., Iwabuchi, A., Kamiya, K., Karasawa, W., Kurita, K., Katagiri, S., Kikuta, A., Kobayashi, H., Kobayashi, N., Machita, K., Maehara, T., Masukawa, M., Mizubayashi, T., Mukai, Y., Nagasaki, H., Nagata, Y., Naito, S., Nakashima, M., Nakama, Y., Nakamichi, Y., Nakamura, M., Meguro, A., Negishi, M., Ohta, I., Ohta, T., Okamoto, M., Ono, N., Saji, S., Sakaguchi, M., Sakai, K., Shibata, M., Shimokawa, T., Song, J.Y., Takazaki, Y., Terasawa, K., Tsugane, M., Tsuji, K., Ueda, S., Waki, K., Yamagata, H., Yamamoto, M., Yamamoto, S., Yamane, H., Yoshiki, S., Yoshihara, R., Yukawa, K., Zhong, H.S., Yano, M., Sasaki, T., Yuan, Q.P., Shu, O.T., Liu, J., Jones, K.M., Gansberger, K., Moffat, K., Hill, J., Bera, J., Fadrosh, D., Jin, S.H., Johri, S., Kim, M., Overton, L., Reardon, M., Tsitrin, T., Vuong, H., Weaver, B., Ciecko, A., Tallon, L., Jackson, J., Pai, G., Van Aken, S., Utterback, T., Reidmuller, S., Feldblyum, T., Hsiao, J., Zismann, V., Iobst, S., de Vazeille, A.R., Buell, C.R., Ying, K., Li, Y., Lu, T.T., Huang, Y.C., Zhao, Q., Feng, Q., Zhang, L., Zhu, J.J., Weng, Q.J., Mu, J., Lu, Y.Q., Fan, D.L., Liu, Y.L., Guan, J.P., Zhang, Y.J., Yu, S.L., Liu, X.H., Zhang, Y., Hong, G.F., Han, B., Choisne, N., Demange, N., Orjeda, G., Samain, S., Cattolico, L., Pelletier, E., Couloux, A., Segurens, B., Wincker, P., D'Hont, A., Scarpelli, C., Weissenbach, J., Salanoubat, M., Quetier, F., Yu, Y., Kim, H.R., Rambo, T., Currie, J., Collura, K., Luo, M.Z., Yang, T.J., Ammiraju, J.S.S., Engler, F., Soderlund, C., Wing, R.A., Palmer, L.E., de la Bastide, M., Spiegel, L., Nascimento, L., Zutavern, T., O'Shaughnessy, A., Dike, S., Dedhia, N., Preston, R., Balija, V., McCombie, W.R., Chow, T.Y., Chen, H.H., Chung, M.C., Chen, C.S., Shaw, J.F., Wu, H.P., Hsiao, K.J., Chao, Y.T., Chu, M.K., Cheng, C.H., Hour, A.L., Lee, P.F., Lin, S.J., Lin, Y.C., Liou, J.Y., Liu, S.M., Hsing, Y.I., Raghuvanshi, S., Mohanty, A., Bharti, A.K., Gaur, A., Gupta, V., Kumar, D., Ravi, V., Vij, S., Kapur, A., Khurana, P., Khurana, P., Khurana, J.P., Tyagi, A.K., Gaikwad, K., Singh, A., Dalal, V., Srivastava, S., Dixit, A., Pal, A.K., Ghazi, I.A., Yadav, M., Pandit, A., Bhargava, A., Sureshbabu, K., Batra, K., Sharma, T.R., Mohapatra, T., Singh, N.K., Messing, J., Nelson, A.B., Fuks, G., Kavchok, S., Keizer, G., Llaca, E.L.V., Song, R.T., Tanyolac, B., Young, S., Il, K.H., Hahn, J.H., Sangsakoo, G., Vanavichit, A., de Mattos, L.A.T., Zimmer, P.D., Malone, G., Dellagostin, O., de Oliveira, A.C., Bevan, M., Bancroft, I., Minx, P., Cordum, H., Wilson, R., Cheng, Z.K., Jin, W.W., Jiang, J.M., Leong, S.A., Iwama, H., Gojobori, T., Itoh, T., Niimura, Y., Fujii, Y., Habara, T., Sakai, H., Sato, Y., Wilson, G., Kumar, K., McCouch, S., Juretic, N., Hoen, D., Wright, S., Bruskiewich, R., Bureau, T.,

Miyao, A., Hirochika, H., Nishikawa, T., Kadowaki, K., Sugiura, M., and Project, I.R.G.S. (2005). The map-based sequence of the rice genome. Nature **436,** 793-800.

Matzke, M.A., Matzke, A.J.M., Pruss, G.J., and Vance, V.B. (2001). RNA-based silencing strategies in plants. Curr Opin Genet Dev **11,** 221-227.

Mayer, K.F., Martis, M., Hedley, P.E., Simkova, H., Liu, H., Morris, J.A., Steuernagel, B., Taudien, S., Roessner, S., Gundlach, H., Kubalakova, M., Suchankova, 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., and Stein, N. (2011). Unlocking the barley genome by chromosomal and comparative genomics. The Plant cell **23,** 1249-1263.

Mendel, G. (1866). Versuche über Pflanzen-Hybriden. In: Verhandlungen des Naturforschenden Vereines in Brünn, Bd. IV für das Jahr 1865**,** 3-47.

Metzker, M.L. (2010). Applications of Next-Generation Sequencing Sequencing Technologies - the Next Generation. Nature Reviews Genetics **11,** 31-46.

Meyer, M., and Kircher, M. (2010). Illumina sequencing library preparation for highly multiplexed target capture and sequencing. Cold Spring Harbor protocols **2010,** pdb prot5448.

Michael, T.P., and Jackson, S. (2013). The First 50 Plant Genomes. Plant Genome-Us **6**.

Michelmore, R.W., Paran, I., and Kesseli, R.V. (1991). Identification of markers linked to disease-resistance genes by bulked segregant analysis: a rapid method to detect markers in specific genomic regions by using segregating populations. Proceedings of the National Academy of Sciences of the United States of America **88,** 9828-9832.

Mimura, M., and Itoh, J.I. (2014). Genetic interaction between rice PLASTOCHRON genes and the gibberellin pathway in leaf development. Rice **7**.

Mimura, M., Nagato, Y., and Itoh, J.I. (2012). Rice PLASTOCHRON genes regulate leaf maturation downstream of the gibberellin signal transduction pathway. Planta **235,** 1081-1089.

Miyoshi, K., Ahn, B.O., Kawakatsu, T., Ito, Y., Itoh, J.I., Nagato, Y., and Kurata, N. (2004). PLASTOCHRON1, a timekeeper of leaf initiation in rice, encodes cytochrome P450. Proceedings of the National Academy of Sciences of the United States of America **101,** 875-880.

Mizutani, M., and Ohta, D. (2010). Diversification of P450 Genes During Land Plant Evolution. Annu Rev Plant Biol **61,** 291-315.

Mokry, M., Nijman, I.J., van Dijken, A., Benjamins, R., Heidstra, R., Scheres, B., and Cuppen, E. (2011). Identification of factors required for meristem function in Arabidopsis using a novel next generation sequencing fast forward genetics approach. BMC genomics **12**.

Moore, G., Devos, K.M., Wang, Z., and Gale, M.D. (1995). Cereal genome evolution. Grasses, line up and form a circle. Current biology : CB **5,** 737-739.

Morgan, T.H., Sturtevant, A.H., Muller, H.J., Bridges, C.B. (1915). The Mechanism of. Mendelian Heredity. Chapter I: Mendelian segregation and the chromosomes. Henry Holt and Company**,** 1-26.

Mortazavi, A., Williams, B.A., Mccue, K., Schaeffer, L., and Wold, B. (2008). Mapping and quantifying mammalian transcriptomes by RNA-Seq. Nature methods **5,** 621-628.

Muller, H.J. (1927). Artificial Transmutation of the Gene. Science **66,** 84-87.

Muller, K.J., Romano, N., Gerstner, O., Garcia-Maroto, F., Pozzi, C., Salamini, F., and Rohde, W. (1995). The barley Hooded mutation caused by a duplication in a homeobox gene intron. Nature **374,** 727-730.

Nagasawa, N., Hibara, K.I., Heppard, E.P., Vander Velden, K.A., Luck, S., Beatty, M., Nagato, Y., and Sakai, H. (2013). GIANT EMBRYO encodes CYP78A13, required for proper size balance between embryo and endosperm in rice. Plant Journal **75,** 592-605.

Nair, S.K., Wang, N., Turuspekov, Y., Pourkheirandish, M., Sinsuwongwat, S., Chen, G.X., Sameri, M., Tagiri, A., Honda, I., Watanabe, Y., Kanamori, H., Wicker, T., Stein, N., Nagamura, Y., Matsumoto, T., and Komatsuda, T. (2010). Cleistogamous flowering in barley arises from the

suppression of microRNA-guided HvAP2 mRNA cleavage. Proceedings of the National Academy of Sciences of the United States of America **107,** 490-495.

Nelson, S.L., Giver, C.R., and Grosovsky, A.J. (1994). Spectrum of X-Ray-Induced Mutations in the Human Hprt Gene. Carcinogenesis **15,** 495-502.

Newman, R.K., Newman, C.W. (2008). 2. Barley: Taxonomy, Morphology, and Anatomy. in: Barley for Food and Health: Science, Technology, and Products. Wiley**,** 21-25.

Nielsen, R. (2005). Molecular signatures of natural selection. Annu Rev Genet **39,** 197-218.

Norberg, M., Holmlund, M., and Nilsson, O. (2005). The BLADE ON PETIOLE genes act redundantly to control the growth and development of lateral organs. Development **132,** 2203-2213.

Nordstrom, K.J., Albani, M.C., James, G.V., Gutjahr, C., Hartwig, B., Turck, F., Paszkowski, U., Coupland, G., and Schneeberger, K. (2013). Mutation identification by direct comparison of whole-genome sequencing data from mutant and wild-type individuals using k-mers. Nature biotechnology **31,** 325-330.

Olson, M., Hood, L., Cantor, C., and Botstein, D. (1989). A Common Language for Physical Mapping of the Human Genome. Science **245,** 1434-1435.

Pankin, A., Campoli, C., Dong, X., Kilian, B., Sharma, R., Himmelbach, A., Saini, R., Davis, S.J., Stein, N., Schneeberger, K., and von Korff, M. (2014). Mapping-by-sequencing identifies HvPHYTOCHROME C as a candidate gene for the early maturity 5 locus modulating the circadian clock and photoperiodic flowering in barley. Genetics **198,** 383-396.

Poland, J.A., Brown, P.J., Sorrells, M.E., and Jannink, J.L. (2012). Development of high-density genetic maps for barley and wheat using a novel two-enzyme genotyping-by-sequencing approach. PloS one **7,** e32253.

Poursarebani, N., Ariyadasa, R., Zhou, R., Schulte, D., Steuernagel, B., Martis, M.M., Graner, A., Schweizer, P., Scholz, U., Mayer, K., and Stein, N. (2013). Conserved synteny-based anchoring of the barley genome physical map. Functional & integrative genomics **13,** 339-350.

Poursarebani, N., Seidensticker, T., Koppolu, R., Trautewig, C., Gawronski, P., Bini, F., Govind, G., Rutten, T., Sakuma, S., Tagiri, A., Wolde, G.M., Youssef, H.M., Battal, A., Ciannamea, S., Fusca, T., Nussbaumer, T., Pozzi, C., Borner, A., Lundqvist, U., Komatsuda, T., Salvi, S., Tuberosa, R., Uauy, C., Sreenivasulu, N., Rossini, L., and Schnurbusch, T. (2015). The Genetic Basis of Composite Spike Form in Barley and "Miracle-Wheat". Genetics.

Pozzi, C., di Pietro, D., Halas, G., Roig, C., and Salamini, F. (2003). Integration of a barley (Hordeum vulgare) molecular linkage, map with the position of genetic loci hosting 29 developmental mutants. Heredity **90,** 390-396.

Quinlan, A.R., and Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics **26,** 841-842.

Ramirez-Gonzalez, R.H., Segovia, V., Bird, N., Fenwick, P., Holdgate, S., Berry, S., Jack, P., Caccamo, M., and Uauy, C. (2015). RNA-Seq bulked segregant analysis enables the identification of high-resolution genetic markers for breeding in hexaploid wheat. Plant biotechnology journal **13,** 613-624.

Ramsay, L., Comadran, J., Druka, A., Marshall, D.F., Thomas, W.T., Macaulay, M., MacKenzie, K., Simpson, C., Fuller, J., Bonar, N., Hayes, P.M., Lundqvist, U., Franckowiak, J.D., Close, T.J., Muehlbauer, G.J., and Waugh, R. (2011). INTERMEDIUM-C, a modifier of lateral spikelet fertility in barley, is an ortholog of the maize domestication gene TEOSINTE BRANCHED 1. Nature genetics **43,** 169-172.

Rapaport, F., Khanin, R., Liang, Y.P., Pirun, M., Krek, A., Zumbo, P., Mason, C.E., Socci, N.D., and Betel, D. (2013). Comprehensive evaluation of differential gene expression analysis methods for RNA-seq data. Genome biology **14**.

Rensing, S.A. (2014). Gene duplication as a driver of plant morphogenetic evolution. Current opinion in plant biology **17,** 43-48.

Rohland, N., and Reich, D. (2012). Cost-effective, high-throughput DNA sequencing libraries for multiplexed target capture. Genome research **22,** 939-946.

Saiki, R.K., Gelfand, D.H., Stoffel, S., Scharf, S.J., Higuchi, R., Horn, G.T., Mullis, K.B., and Erlich, H.A. (1988). Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. Science **239,** 487-491.

Sanger, F., Nicklen, S., and Coulson, A.R. (1977). DNA Sequencing with Chain-Terminating Inhibitors. Proceedings of the National Academy of Sciences of the United States of America **74,** 5463-5467.

Schmutzer, T., Ma, L., Pousarebani, N., Bull, F., Stein, N., Houben, A., and Scholz, U. (2014). Kmasker - A Tool for in silico Prediction of Single-Copy FISH Probes for the Large-Genome Species Hordeum vulgare. Cytogenetic and genome research **142,** 66-78.

Schneeberger, K., and Weigel, D. (2011). Fast-forward genetics enabled by new sequencing technologies. Trends in plant science **16,** 282-288.

Schneeberger, K., Ossowski, S., Lanz, C., Juul, T., Petersen, A.H., Nielsen, K.L., Jorgensen, J.E., Weigel, D., and Andersen, S.U. (2009). SHOREmap: simultaneous mapping and mutation identification by deep sequencing. Nature methods **6,** 550-551.

Scholz, F., Lehmann, C.O. (1962). Die Gaterslebener Mutanten der Saatgerste in Beziehung zur Formenmannigfaltigkeit der Art Hordeum vulgare L. IV. Die Kulturpflanze **10,** 312-334.

Schroeder, A., Mueller, O., Stocker, S., Salowsky, R., Leiber, M., Gassmann, M., Lightfoot, S., Menzel, W., Granzow, M., and Ragg, T. (2006). The RIN: an RNA integrity number for assigning integrity values to RNA measurements. BMC molecular biology **7,** 3.

Schulman, A.H. (2006). Molecular markers to assess genetic diversity. Euphytica **158,** 313–321.

Schulte, D., Close, T.J., Graner, A., Langridge, P., Matsumoto, T., Muehlbauer, G., Sato, K., Schulman, A.H., Waugh, R., Wise, R.P., and Stein, N. (2009). The International Barley Sequencing Consortium-At the Threshold of Efficient Access to the Barley Genome. Plant physiology **149,** 142-147.

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

Serrat, X., Esteban, R., Guibourt, N., Moysset, L., Nogues, S., and Lalanne, E. (2014). EMS mutagenesis in mature seed-derived rice calli as a new method for rapidly obtaining TILLING mutant populations. Plant methods **10.**

Siegfried, K.R., Eshed, Y., Baum, S.F., Otsuga, D., Drews, G.N., and Bowman, J.L. (1999). Members of the YABBY gene family specify abaxial cell fate in Arabidopsis. Development **126,** 4117-4128.

Sotelo-Silveira, M., Cucinotta, M., Chauvin, A.L., Montes, R.A.C., Colombo, L., Marsch-Martinez, N., and de Folter, S. (2013). Cytochrome P450 CYP78A9 Is Involved in Arabidopsis Reproductive Development. Plant physiology **162,** 779-799.

Stanke, M., Steinkamp, R., Waack, S., and Morgenstern, B. (2004). AUGUSTUS: a web server for gene finding in eukaryotes. Nucleic acids research **32,** W309-W312.

Stein, N., Herren, G., and Keller, B. (2001). A new DNA extraction method for high-throughput marker analysis in a large-genome species such as Triticum aestivum. Plant Breeding **120,** 354-356.

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

Stein, N., Graner, A. (2005). Map-Based Gene Isolation in Cereal Genomes in: Cereal Genomics; Gupta, P. K., Varshney, R.K. . Kluwer Academic Publishers**,** 331-360.

Szemenyei, H., Hannon, M., and Long, J.A. (2008). TOPLESS mediates auxin-dependent transcriptional repression during Arabidopsis embryogenesis. Science **319,** 1384-1386.

Tajima, F. (1989). Statistical-Method for Testing the Neutral Mutation Hypothesis by DNA Polymorphism. Genetics **123,** 585-595.

Taketa, S., Amano, S., Tsujino, Y., Sato, T., Saisho, D., Kakeda, K., Nomura, M., Suzuki, T., Matsumoto, T., Sato, K., Kanamori, H., Kawasaki, S., and Takeda, K. (2008). Barley grain

with adhering hulls is controlled by an ERF family transcription factor gene regulating a lipid biosynthesis pathway. Proceedings of the National Academy of Sciences of the United States of America **105,** 4062-4067.

Tamura, K., Stecher, G., Peterson, D., Filipski, A., and Kumar, S. (2013). MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. Molecular biology and evolution **30,** 2725-2729.

Tanksley, S.D. (1983). Molecular markers in plant breeding Plant Mol Biol Rep **1,** 3-8.

Tanksley, S.D., Ganal, M.W., and Martin, G.B. (1995). Chromosome Landing - a Paradigm for Map-Based Gene Cloning in Plants with Large Genomes. Trends Genet **11,** 63-68.

Tavakol, E., Okagaki, R., Verderio, G., Shariati, J.V., Hussien, A., Bilgic, H., Scanlon, M.J., Todt, N.R., Close, T.J., Druka, A., Waugh, R., Steuernagel, B., Ariyadasa, R., Himmelbach, A., Stein, N., Muehlbauer, G.J., and Rossini, L. (2015). The Barley Uniculme4 Gene Encodes a BLADE-ON-PETIOLE-Like Protein That Controls Tillering and Leaf Patterning. Plant physiology **168,** 164-174.

The Arabidopsis Genome Initiative. (2000). Analysis of the genome sequence of the flowering plant Arabidopsis thaliana. Nature **408,** 796-815.

The International Brachypodium Initiative. (2010). Genome sequencing and analysis of the model grass Brachypodium distachyon. Nature **463,** 763-768.

Thiel, T., Kota, R., Grosse, I., Stein, N., and Graner, A. (2004). SNP2CAPS: a SNP and INDEL analysis tool for CAPS marker development. Nucleic acids research **32**.

Thiel, T., Graner, A., Waugh, R., Grosse, I., Close, T.J., and Stein, N. (2009). Evidence and evolutionary analysis of ancient whole-genome duplication in barley predating the divergence from rice. BMC evolutionary biology **9,** 209.

Till, B.J., Zerr, T., Comai, L., and Henikoff, S. (2006). A protocol for TILLING and Ecotilling in plants and animals. Nature protocols **1,** 2465-2477.

Tropp, B.E. (2011). 13.1 Introduction to Homologous Recombination. in Molecular Biology - Genes to Proteins. Jones & Bartlett Learning **4,** 512-517.

Ullrich, S.E., Baik, B.K., Quinde-Axtell, Z., Nair, S. (2008). Barley for food: traits and improvements. in Proceedings of the 10th International Barley Genetics Symposium. **,** 563-575.

Untergasser, A., Cutcutache, I., Koressaar, T., Ye, J., Faircloth, B.C., Remm, M., and Rozen, S.G. (2012). Primer3-new capabilities and interfaces. Nucleic acids research **40**.

Veit, B., Briggs, S.P., Schmidt, R.J., Yanofsky, M.F., and Hake, S. (1998). Regulation of leaf initiation by the terminal ear 1 gene of maize. Nature **393,** 166-168.

Vos, P., Hogers, R., Bleeker, M., Reijans, M., Vandelee, T., Hornes, M., Frijters, A., Pot, J., Peleman, J., Kuiper, M., and Zabeau, M. (1995). Aflp - a New Technique for DNA-Fingerprinting. Nucleic acids research **23,** 4407-4414.

Vriet, C., Russinova, E., and Reuzeau, C. (2012). Boosting crop yields with plant steroids. The Plant cell **24,** 842-857.

Walker, G.W.R., Miller, R., Dietrich, J., and Kasha, K. (1963). Recent Barley Mutants and Their Linkages .2. Genetic Data for Further Mutants. Can J Genet Cytol **5,** 200-&.

Wang, J.W., Schwab, R., Czech, B., Mica, E., and Weigel, D. (2008). Dual effects of miR156-targeted SPL genes and CYP78A5/KLUH on plastochron length and organ size in Arabidopsis thaliana. The Plant cell **20,** 1231-1243.

Wang, Z., Gerstein, M., and Snyder, M. (2009). RNA-Seq: a revolutionary tool for transcriptomics. Nature reviews. Genetics **10,** 57-63.

Weissenbach, J., Gyapay, G., Dib, C., Vignal, A., Morissette, J., Millasseau, P., Vaysseix, G., and Lathrop, M. (1992). A second-generation linkage map of the human genome. Nature **359,** 794-801.

Wendt, T., Holm, P.B., Starker, C.G., Christian, M., Voytas, D.F., Brinch-Pedersen, H., and Holme, I.B. (2013). TAL effector nucleases induce mutations at a pre-selected location in the genome of primary barley transformants. Plant Mol Biol **83,** 279-285.

Williams, J.G.K., Kubelik, A.R., Livak, K.J., Rafalski, J.A., and Tingey, S.V. (1990). DNA Polymorphisms Amplified by Arbitrary Primers Are Useful as Genetic-Markers. Nucleic acids research **18,** 6531-6535.

Wu, A.R., Neff, N.F., Kalisky, T., Dalerba, P., Treutlein, B., Rothenberg, M.E., Mburu, F.M., Mantalas, G.L., Sim, S., Clarke, M.F., and Quake, S.R. (2014). Quantitative assessment of single-cell RNA-sequencing methods. Nature methods **11,** 41-46.

Xiao, H., Tang, J.F., Li, Y.F., Wang, W.M., Li, X.B., Jin, L., Xie, R., Luo, H.F., Zhao, X.F., Meng, Z., He, G.H., and Zhu, L.H. (2009). STAMENLESS 1, encoding a single C2H2 zinc finger protein, regulates floral organ identity in rice. Plant Journal **59,** 789-801.

Xu, M., Hu, T., McKim, S.M., Murmu, J., Haughn, G.W., and Hepworth, S.R. (2010). Arabidopsis BLADE-ON-PETIOLE1 and 2 promote floral meristem fate and determinacy in a previously undefined pathway targeting APETALA1 and AGAMOUS-LIKE24. The Plant journal : for cell and molecular biology **63,** 974-989.

Yamaguchi, T., Nagasawa, N., Kawasaki, S., Matsuoka, M., Nagato, Y., and Hirano, H.Y. (2004). The YABBY gene DROOPING LEAF regulates carpel specification and midrib development in Oryza sativa. The Plant cell **16,** 500-509.

Yoshida, A., Ohmori, Y., Kitano, H., Taguchi-Shiobara, F., and Hirano, H.Y. (2012). ABERRANT SPIKELET AND PANICLE1, encoding a TOPLESS-related transcriptional co-repressor, is involved in the regulation of meristem fate in rice. Plant Journal **70,** 327-339.

Yoshida, H. (2012). Is the lodicule a petal: Molecular evidence? Plant Sci **184,** 121-128.

Yoshida, H., and Nagato, Y. (2011). Flower development in rice. Journal of experimental botany **62,** 4719-4730.

Yu, J., Wang, J., Lin, W., Li, S.G., Li, H., Zhou, J., Ni, P.X., Dong, W., Hu, S.N., Zeng, C.Q., Zhang, J.G., Zhang, Y., Li, R.Q., Xu, Z.Y., Li, S.T., Li, X.R., Zheng, H.K., Cong, L.J., Lin, L., Yin, J.N., Geng, J.N., Li, G.Y., Shi, J.P., Liu, J., Lv, H., Li, J., Wang, J., Deng, Y.J., Ran, L.H., Shi, X.L., Wang, X.Y., Wu, Q.F., Li, C.F., Ren, X.Y., Wang, J.Q., Wang, X.L., Li, D.W., Liu, D.Y., Zhang, X.W., Ji, Z.D., Zhao, W.M., Sun, Y.Q., Zhang, Z.P., Bao, J.Y., Han, Y.J., Dong, L.L., Ji, J., Chen, P., Wu, S.M., Liu, J.S., Xiao, Y., Bu, D.B., Tan, J.L., Yang, L., Ye, C., Zhang, J.F., Xu, J.Y., Zhou, Y., Yu, Y.P., Zhang, B., Zhuang, S.L., Wei, H.B., Liu, B., Lei, M., Yu, H., Li, Y.Z., Xu, H., Wei, S.L., He, X.M., Fang, L.J., Zhang, Z.J., Zhang, Y.Z., Huang, X.G., Su, Z.X., Tong, W., Li, J.H., Tong, Z.Z., Li, S.L., Ye, J., Wang, L.S., Fang, L., Lei, T.T., Chen, C., Chen, H., Xu, Z., Li, H.H., Huang, H.Y., Zhang, F., Xu, H.Y., Li, N., Zhao, C.F., Li, S.T., Dong, L.J., Huang, Y.Q., Li, L., Xi, Y., Qi, Q.H., Li, W.J., Zhang, B., Hu, W., Zhang, Y.L., Tian, X.J., Jiao, Y.Z., Liang, X.H., Jin, J.A., Gao, L., Zheng, W.M., Hao, B.L., Liu, S.Q., Wang, W., Yuan, L.P., Cao, M.L., McDermott, J., Samudrala, R., Wang, J., Wong, G.K.S., and Yang, H.M. (2005). The Genomes of Oryza sativa: A history of duplications. Plos Biol **3,** 266-281.

Yuo, T., Yamashita, Y., Kanamori, H., Matsumoto, T., Lundqvist, U., Sato, K., Ichii, M., Jobling, S.A., and Taketa, S. (2012). A SHORT INTERNODES (SHI) family transcription factor gene regulates awn elongation and pistil morphology in barley. Journal of experimental botany **63,** 5223-5232.

Zadnikova, P., and Simon, R. (2014). How boundaries control plant development. Current opinion in plant biology **17,** 116-125.

Zhao, Y.D. (2008). The role of local biosynthesis of auxin and cytokinin in plant development. Current opinion in plant biology **11,** 16-22.

Zondlo, S.C., and Irish, V.F. (1999). CYP78A5 encodes a cytochrome P450 that marks the shoot apical meristem boundary in Arabidopsis. Plant Journal **19,** 259-268.

10. Abbreviations

%	Percent
˚C	Celsius
μg	Microgram
μl	Microliter
μM	Micromolar
AA	Amino acid
ABA	Abscisic acid
AFLP	Amplified Fragment Length Polymorphism
ANK	Ankyrin repeats
ATP	Adenosine triphosphate
BAC	Bacterial Artificial Chromosome
BLAST	Basic Local Alignment Search Tool
BLASTN	Nucleotide BLAST
BLASTP	Protein BLAST
BOP	*BLADE-ON-PETIOLE*
BOPA	Barley Oligo Pool Assay
bp	Base pairs
BR	Broad-Range
BSA	*bovine serum albumin*
BSA	Bulk segregant analysis
BTB/	Broad complex, tramtrack, and brick à brack
BW	Bowman Near Isogenic Line
CAPS	Cleaved Amplified Polymorphic Sequences
cd	Completely developed
cDNA	Complementary DNA
CDS	*Coding sequence*
CHIP-seq	Chromatin Immunoprecipitation Sequencing
cm	Centimeter
cM	Centimorgan
cm	Centimeter
Col-0	Columbia
CTAB	Cetyl-trimethylammonium bromide
cv.	Cultivar
CYP78A	Cytochrome P450 class 78a
$_{dd}H_2O$	Double distilled water
DEPC	Diethylpyrocarbonate
DNA	Deoxyribonucleic acid
dNTP	Deoxynucleotide
DSB	Double strand break
dsDNA	double-stranded DNA
EDTA	Ethylenediaminetetraacetic acid

EMS	Ethyl methanesulfonate
EPF	Elute, Prime, Fragment Mix
F1	First filial generation
F2	Second filial generation
F3	Third filial generation
FDR	False discovery rate
FLcDNA	Full-length cDNA
FPcontig	Finger-Printed contigs
g	Gram
GA	Gibberellins
Gbp	*giga base pairs*
GBS	Genotyping-by-Sequencing
gl	Glume primordium
H. spontaneum	*Hordeum vulgare subsp. spontaneum*
Hap	Haplotype
HCL	hydrogen chloride
hd	Haplotype diversity
HS	High Sensitivity
IAA	Indole-3-acetic acid
IBSC	International Barley Genome Sequencing Consortium
ID	Identifier
ind	*indica*
IPK	Leibniz Institute of Plant Genetics and Crop Plant Research
jap	*japonica*
JHI	James Hutton Institute
kb	Kilobase pairs
Kbp	Kilobase pairs
KCL	potassium chloride
KNOX1	*KNOTTED1 like homeobox class 1*
lax-a	*laxatum-a*
LM-PCR	Ligation mediated PCR
LOD	Logarithm of Odds
log2	Binary logarithm
M	Molar
M_2	Second mutant generation
M_3	Third mutant generation
max	Maximum
Mb	Megabase pairs
mbar	Millibar
Mbp	Megabase pairs
mg	Milligram
MgCl2	Magnesium chloride
min	Minute

ml	Milliliter
mM	Milimolar
mnd	*many noded dwarf*
MOPS	3-(N-morpholino) propanesulfonic acid
MPC96	Magnetic Particle Concentrator
mRNA	Messenger RNA
MT	Mutant
MTP	Minimal tiling path
NaOAc	Sodium acetate
NaOH	Sodium hydroxide
NCBI	National Center for Biotechnology Information
ng	Nanogram
NGS	Next Generation Sequencing
(NH4)2SO4	Ammonium sulfate
NH_4OAc	Ammonium Acetate
NIL	Nearly isogenic lines
NordGen	Nordic Genetic Resource Center
NPR1	*NONEXPRESSOR OF PR GENES1*
ORF	Open reading frame
PCR	Polymerase Chain Reaction
POPSEQ	Population sequencing
POZ	Poxviruses and zinc finger
p-value	probability-value
qRT-PCR	Quantitative reverse transcribed PCR
RAPD	Random Amplified Polymorphic DNA
RFLP	Restriction Fragment Length Polymorphisms
RIL	Recombinant inbreed lines
RIN	RNA Integrity Number
RNA	Ribonucleic acid
RNAi	RNA interference
RNA-seq	RNA sequencing
RPKM	Reads per kilo base of transcript per million mapped reads
S	Polymorphic/segregating sites
SC	Sequence Capture
SDS	Sodium dodecyl sulfate
sec	Seconds
siRNA	Small inference RNA
SNP	Single nucleotide polymorphism
SSR	Simple sequence repeat
st	stamen primordium
STS	Sequence Tagged Site
TAE	Tris-Acetate-EDTA
TALENs	Transcription activator-like effector nucleases
TD-PCR	Touchdown PCR

TFs	Transcription factors
TILLING	Targeting Local Lesions IN Genomes
Tm	DNA melting temperature
Tris	*Tris*(hydroxymethyl)aminomethane
U	Unit
V	Voltage
v/v	volume/volume
w/v	weight/volume
WGS	Whole genome shotgun sequencing
WT	Wild-type
π	Nucleotide diversity
$\chi 2$	Chi-square

11. Eidesstattliche erklärung / *Declaration under Oath*

Herr Matthias Jost

Röpziger Straße 17

06110 Halle (Saale)

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

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

Datum / *Date*   Unterschrift des Antragstellers / *Signature of the applicant* (Matthias Jost)

12. Curriculum Vitae

Personal Details

Date of birth: 29.06.1986
Place of birth: Rochlitz
Nationality: German
Marital status: Single
E-mail: jost@ipk-gatersleben.de; matthias.jost2@gmail.com
Present address: Dieskauer Straße 18
                        D-06112 Halle (Saale)

Academic Education

01/2016 to present: Postdoctoral Researcher in the group Genomics of Genetic Resources at the Leibniz Institute of Plant Genetics and Crop Plant Research (IPK, Gatersleben)

07/2011 to 12/2015: PhD Student at Martin-Luther-University Halle-Wittenberg

- Work and PhD thesis was done in the Genome Diversity group at the Leibniz Institute of Plant Genetics and Crop Plant Research (IPK, Gatersleben)

10/2005 – 06/2011: Study Agricultural sciences at Martin-Luther-University Halle-Wittenberg (MLU)

Experimental diploma thesis: "Genetic mapping of the spike morphology mutant *laxatum-a* in barley" (written in German)

Scientific supervisor: Dr. Nils Stein (IPK Gatersleben)
Reviewer: Prof. Dr. Andreas Graner (University Halle and IPK Gatersleben)
Prof. Dr. Klaus Pillen (University Halle)

Honors and awards:

Poster-Award for young scientists awarded from the scientific advisory board „Pflanzenbiotechnologie der Zukunft" in frame of the PLANT 2030 Status Seminar 06.-08.03.2013

Best Poster at the „10th Plant Science Student Conference" (PSSC), 02.-05.06.2014 in Gatersleben

"Oohara Foundation Scholarship" for short term visit of the "Institute of Plant Science and Resources" (IPSR), Okayama University, Kurashiki Japan, 19-31.08.2015

Publications

Jost, M.*, Taketa, S.*, Mascher, M., Himmelbach, A., You, T., Shahinnia, F., Rutten, T., Druka, A., Schmutzer, T., Steuernagel, B., Beier, S., Taudien, S., Scholz, U., Morgante, M., Waugh, R., Stein, N. A homolog of Blade-On-Petiole 1 and 2 (BOP1/2) controls internode length and homeotic changes of the barley inflorescence. **Plant Physiol. (2016).** Epub ahead of print: **dx.doi.org/10.1104/pp.16.00124**

Mascher, M.*, **Jost, M**.*, Kuon, J.E., Himmelbach, A, Assfalg, A., Beier, S., Scholz, U., Graner, A. & Stein N.: Mapping-by-sequencing accelerates forward genetics in barley. **Genome Biology 15 (2014)** R78.

* Co-first authors

Conferences - oral presentation:

Jost, M.: Cloning of the homeotic gene *laxatum* (*lax.a*) - prospects from an improving barley, **9th Plant Science Student Conference** (**PSSC**)**,** Halle, 29.05.2013

Jost, M.: Cloning of the homeotic gene *laxatum* (*lax.a*) - prospects from an improving barley genomics infrastructure**. 27. Tagung Molekularbiologie der Pflanzen**, Tagungshotel Maria in der Aue, Wermelskirchen, 26.02.2014

Jost, M.: Cloning of the homeotic gene *laxatum* (*lax.a*) - prospects from an improving barley genomics infrastructure. **1st International Workshop on Barley Mutant Research**. IPK Gatersleben, Stadt Seeland, 27.06.2014

Conferences - Posters

Matthias Jost[1], Martin Mascher[1], Burkhard Steuernagel[1], Thomas Schmutzer[1], Axel Himmelbach[1], Uwe Scholz[1], Arnis Druka[2], Robbie Waugh[2], Shin Taketa[3] and Nils Stein[1].
„Cloning of the homeotic gene *Laxatum.a* (*lax.a*) - prospects from an improving barley genomics infrastructure"

[1] Leibniz Institute of Plant Genetics and Crop Plant Research (IPK), Gatersleben, 06466 Stadt Seeland, Germany
[2] James Hutton Institute, Invergowrie, Dundee, Scotland DD2 5DA, UK
[3] Institute of Plant Science and Resources, Okayama University, 2-20-1 Chuo, 710-0046, Japan

- **Genome Research Working Group Conference of the GPZ,** 18.-20.09.2012, Martin-Luther-University Halle-Wittenberg

- **PLANT 2030 Status Seminar**, Potsdam, 06.-08.03.2012
- **8th Plant Science Student Conference (PSSC)**, Gatersleben, 04.-07.06.2012
- **PLANT 2030 Status Seminar**, Potsdam, 06.-08.03.2013
- **"Plant Biology for the Next Generation" - the International Symposium of the SFB924 "Molecular mechanisms regulating yield and yield stability in plants,** Freising, 18-20.09.2013
- **The Plant and Animal Genome Conference (PAG),** San Diego, CA. USA, 11.-15.01.2014
- **10th Plant Science Student Conference (PSSC)**, Gatersleben, 02.-05.06.2014
- **EUCARPIA Cereals Section – ITMI Joint Conference**, Wernigerode, 29.06.-06.07.2014

_____   _____

Datum / Date   Matthias Jost