Unlocking the secondary gene pool of barley for breeding and research

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

der

Naturwissenschaftlichen Fakultät III
Agrar- und Ernährungswissenschaften,
Geowissenschaften und Informatik

der Martin-Luther-Universität Halle-Wittenberg

vorgelegt von

Frau M.Sc. agr. Neele Wendler

geb. am 13.05.1987 in Neumünster

Gutachter:
Prof. Dr. Andreas Graner
Prof. Dr. Peter Wehling

Verteidigt am
25/01/2016
Halle/Saale

Table of contents

List of figures

List of supplementary figures Wendler et al. 2014

List of supplementary tables Wendler et al. 2014

List of additional files Wendler et al. 2015

1. General introduction

An estimated 10,000 years ago men started to select and cultivate wild plant species – giving birth to agriculture and domestication of crop plants (Lev-Yadun et al., 2000). Presumably, back in those days, people did not know that they would revolutionize human lifestyle and pave the way for modern human civilization. This milestone in the history of mankind took place in a region called the Fertile Crescent (Gopher et al., 2002; Moore et al., 2000) - spanning modern-day Jordan, Israel, Lebanon, western Syria, and into Turkey, Iraq and Iran. At that time, barley (*Hordeum vulgare* L.) was one of the very first crops to be selected (Salamini et al., 2002) and cultivated, and ever since it remains as one of the most important crop species worldwide. Today, barley is the fourth most important cereal crop in the world in terms of harvested area and tons (http://faostat.fao.org). The largest proportion of the global barley production is utilized for animal feed, followed by the use in malting industries and various food products (Blake et al., 2011). Barley is adapted to a diverse range of environments and it is known to be more stress tolerant than wheat (*Triticum aestivum*) (Nevo et al., 2012). Hence, barley can be cultivated even under unfavorable conditions where alternative food crops are limited and it remains the major source of daily calories for human nutrition especially in some developing countries (Grando and Macpherson, 2005). Furthermore, it is thought to exhibit health benefits as the grain is particularly high in soluble dietary fiber, lowering blood cholesterol, blood pressure and glycaemic index (Ullrich SE, 2008).

Barley belongs to the genus *Hordeum*, which is part of the *Triticeae* tribe and the *Poaceae* (*Gramineae*) family (von Bothmer et al., 1995). The genus *Hordeum* comprises about 33 species (Blattner, 2009) and 45 taxa including diploid, polyploid, annual and perennial types. It shows a wide geographical distribution throughout the world (von Bothmer et al., 1995). Morphologically the genus is defined to have three single-flowered spikelets (triplets) at each rachis node of the inflorescence and it possesses the basic chromosome number of x = 7 (Blattner, 2009). Cultivated barley has been domesticated from *H. vulgare* L. ssp. *spontaneum*, which was first described by the German botanist Carl Koch in Turkey (von Bothmer et al., 1995). Barley is an annual, diploid (2n = 2x = 14) inbreeding species with a genome of > 5 Gbp in size (Dolezel and Bartos, 2005; International Barley Sequencing Consortium, 2012). Even though, more than 80 % of the genome is composed of repetitive DNA (International Barley Sequencing Consortium, 2012), barley has

proven as a valuable reference for other *Triticeae* species with more complex genomes such as rye (*Secale cereale*) and bread wheat (*Triticum aestivum*).

Prehistoric farmers have selected barley for mainly three morphological features that make the crop easier to harvest or consume such as: non-brittleness of the inflorescence – so seeds do not fall apart when being harvested, free threshing naked grains that are no longer covered by the attached glumes, and seed size (Salamini et al., 2002). The millennia of domestication and improvement / adaptation by early farmers and especially the extensive breeding during the last centuries induced the necessary changes to facilitate modern cultivation of barley and laid the foundation for tremendous yield increases (Salamini et al., 2002). Today, breeding has become a highly complex discipline. The major breeding goals for barley and crops in general are to increase yield and quality and to maintain yield and quality at high level by improving the tolerance to biotic and abiotic stress. Efficient and targeted breeding programs seem to be more important than ever, considering the constantly growing world population and rapidly changing environmental conditions (climate change) (Intergovernmental Panel on Climate Change (IPCC), 2013). On the downside however, extensive breeding has often narrowed the genetic diversity of breeding germplasm, resulting in a limited choice of beneficial alleles for further improving elite crops (Feuillet et al., 2008) (Figure 1).

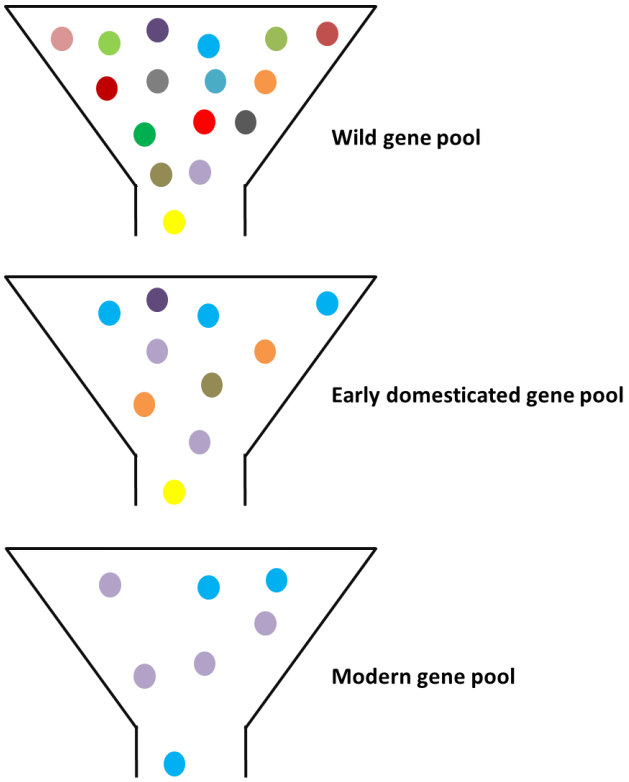

Figure 1. Loss of diversity through domestication and adaptation.

The figure illustrates schematically the loss of diversity through domestication and improvement. Dots represent individual genotypes, where each different color indicates allelic diversity at gene loci leading to different haplotypes. Genetic bottlenecks occur during domestication when a subset of genotypes of the wild are selected during domestication. Consequently cultivated gene pools will be restricted to the diversity from these limited founders. The figure was modified after Tanksley and McCouch (1997).

1.1 Genetic resources

Modern plant breeding often focuses only on a relatively small, improved and adapted sub-population of the available cultivated or on non-cultivated diverse gene pools. This includes strong selection, leading to reduced genetic diversity of the crop as compared to the ancestor (Tanksley and McCouch, 1997). Furthermore, among breeders it is common practice to reuse preferred elite parents to establish new varieties (Kumar et al., 2011). However, a very narrow genetic diversity may pose a significant risk to crop production, as was painfully experienced during the Irish potato famine from 1845 to 1852 as well as during the southern leaf blight epidemic of maize in the US in 1970. Both catastrophes were benefitted due to genetic uniformity

of the cultivated crops (Damania, 2008). The Irish potato production was particularly vulnerable due to a very limited number of varieties grown at that time (Damania, 2008; Hawkes, 1994). In fact a single variety 'Lumper' was extensively cultivated and unfortunately very susceptible to *Phytophthora infestans*, the causal agent of the Irish famine (Damania, 2008; Hawkes, 1994). Likewise, the cause of the leaf blight epidemic in the US was the widespread use of a single genetic male sterility factor, which turned out to be genetically linked to Southern leaf blight (*Bipolaris maydis*) susceptibility (Tanksley and McCouch, 1997). In 1970 this male sterility factor was present in more than 85 % of the US corn cultivation area (Levings, 1990).

For the preservation of genetic resources and to study the origin of cultivated plants Nikolai Ivanovich Vavilov has initiated the collection and conservation of plant genetic resources more than 70 years ago (Vavilov, 1940). Maybe one of the most impressive and important success stories of genetic resources in plant breeding took place in the 1960's. At that time Norman Borlaug incorporated dwarfing genes from a wheat cultivar of Japanese origin and genes conferring durable disease and insect resistance into wheat, promoting the Green Revolution (Rajaram and Hettel, 1994).

It was already in 1972, when the National Research Council, Washington, US, claimed about the urgency and necessity of genetic resources and diversity for a stable food supply (National Research Council, 1972). Since the importance of genetic variation was of global concern, a "plan of action" was published in an international agreement (adopted by 150 countries) in 1996 and updated in 2011 as a "strategic framework for the conservation and sustainable use of the plant genetic diversity" (Food and Agricultural Organization of the United Nations (FAO), 1996; Food and Agricultural Organization of the United Nations (FAO), 2011). Today, plant genetic resources are maintained and conserved in more than 1300 *ex situ* gene banks worldwide (Haussmann et al., 2004).

Genetic resources can be defined as "all materials that are available for improvement of a cultivated plant species" (Becker, 1993). According to the gene pool concept, they can be divided into primary gene pool, secondary gene pool, and tertiary gene pool (Harlan and de Wet, 1971). This classification depends on how closely related they are at the genomic level and based on the degree of sexual compatibility (Harlan and de Wet, 1971). Breeding programs generally consider mainly the primary gene pool, since traits are much easier to utilize and incorporate. Species of

the secondary and tertiary gene pool exhibit crossing barriers to the primary gene pool crop (Harlan and de Wet, 1971). Thus, the incorporation of traits is costlier as compared to the primary gene pool. Furthermore, recombination frequencies between the chromosomes of different gene pool species are usually reduced, slowing down the process of trait incorporation. The larger the genetic distance between a genetic resource and a related crop species, the higher the chance that negative traits are transferred together with the wanted improvement trait (negative linkage drag) (Brinkman and Frey, 1977; Stam and Zeven, 1981). Negative linkage drag may cause yield penalties. Depending on the crop and the genetic resource, it may be a laborious process to remove those negative effects. The importance of genetic resources for crop improvement belonging to any of the three gene pools varies between different crop species. In maize (*Zea mays*) or pearl millet (*Pennisetum glaucum)* the genetic variation within the primary gene pool is high enough to ensure sustainable crop improvement (Simmonds, 1962). This is in strong contrast to rape seed (*Brassica napus*) a recent allopolyploid species with very limited genetic diversity in the primary gene pool. Thus, "novel" traits can be almost exclusively introduced from diploid *Brassica* species of the secondary and tertiary gene pool (Hu et al., 2002). In barley elite germplasm genetic diversity has been narrowed significantly if compared to *H. spontaneum* or early domesticated cultivars (landraces) (Forster et al., 2000; Russell et al., 2000). For instance the analysis of 28 loci over 101 elite barley cultivars and 19 barley landraces revealed that 72 % of the total allele diversity was accounted for by the 19 landraces (Russell et al., 2000). Therefore, barley breeders may benefit by accessing genetic diversity present in less adapted genetic resources.

Since the likelihood that a pathogen can infect two species decreases with their phylogenetic distance (Gilbert and Webb, 2007), the secondary and tertiary gene pools are thought to exhibit a reservoir of more durable non-host resistances compared to those available from the primary gene pool. Thus, despite the substantially higher effort needed to access traits from outside the primary gene pool, it is often worth it. In 1986, it was estimated that the benefit of introgressions from wild relatives into crops accounted for more than 340 million US Dollars per year (Prescott-Allen and Prescott-Allen, 1986). To date, traits from wild relatives have been transferred into at least 13 crops; among them the world's four most important cereals: wheat, rice (*Oryza sativa*), maize (*Zea mays* L.), and barley (Hajjar and Hodgkin, 2007).

1.1.1 Genetic resources for barley improvement

In accordance to the gene pool concept the primary gene pool consists of the crop species itself, landraces, and direct relatives (Harlan and de Wet, 1971). Landraces are locally adapted varieties that have not been undergoing extensive modern line breeding but were selected and adapted as populations by farmers and later also breeders. The members of the primary gene pool can be crossed without obstacles; hybrids are generally fertile with good chromosome pairing; gene segregation is approximately normal and genes can be transferred without barriers (Harlan and de Wet, 1971). The primary gene pool of barley consists of cultivated barley including landraces and its direct ancestor *H. spontaneum* (wild barley). Barley breeding has often taken advantage of the diversity that is provided by landraces (von Bothmer et al., 2003). For instance, broad spectrum resistance against *Blumeria graminis* spp. *hordei* (powdery mildew) was recognized in Ethiopian landraces around 1970 (Jørgensen, 1992a). Subsequently, the resistance gene *mlo-11*, was transferred into elite spring barley and provided durable resistance to powdery mildew in Europe since then (Jørgensen, 1992a). Owing to the emergence of highly yielding elite barley cultivars, the production of barley landraces has diminished constantly. This is because they are often outcompeted regarding yield, especially under technologically advanced, modern plant production practices. Nonetheless, barley landraces still play an important role as major daily calorie source of the human diet in regions with a less developed agriculture such as parts of Central and Southern Asia as well as of Northern Africa and Ethiopia (von Bothmer et al., 2003). In Europe the cultivation has been abandoned except for a few restricted areas, such as Sardinia in Italy (Attene et al., 1996). Landraces make up the largest fraction of barley germplasm in genebanks representing about 23 % of the total collection (Sato, 2014).

Also *H. spontaneum* was found to be a valuable genetic resource for barley breeding. It has a broad geographic distribution, spanning from the eastern Mediterranean and Turkey to Central Asia, and much of Southwestern Asia. The diversity within wild barley was found to be rich, e.g. regarding disease resistance (Abbott et al., 1992; Fetch et al., 2003; Forster et al., 2000; Jana and Nevo, 1991; Moseman et al., 1990), allozymic protein variation (Baek et al., 2003; Nevo, 1998), growth physiology (Van Rijn et al., 2000), agronomic performance (Nevo et al., 1984) and on the level of DNA (Baek et al., 2003; Baum et al., 1997; Turpeinen et al., 2001). Despite of negative linkage drag (Sato and Takeda, 1997), *H. spontaneum* has been used extensively in crosses with *H. vulgare* (Nevo, 1992) conferring new disease resistances and tolerances to abiotic stress

(Eglinton et al., 1999; Ivandic et al., 1998; Lehmann, 1991; Sato and Takeda, 1997) including salt tolerance (Mano and Takeda, 1998), water stress and nitrogen utilization (Gorny, 2001; Honsdorf et al., 2014; Schnaithmann and Pillen, 2013). Also broad-spectrum resistances such as *Rph16*, which confers resistance to all known isolates of barley leaf rust (*Puccinia hordei* Otth.) were gained from *H. spontaneum* (Ivandic et al., 1998). *Mla-6* and *Mla-14*, two genes conferring powdery mildew resistance, were introgressed from *H. spontaneum* (Jørgensen, 1992b). Furthermore, drought tolerant barley cultivars have been obtained from wild barley and were released by the International Center for Agricultural Research, Washington, US (Hajjar and Hodgkin, 2007). *H. spontaneum* is represented by more than 30,000 accessions (Saisho and Takeda, 2011) in 37 gene banks worldwide (Feuillet and Muehlbauer, 2009).

The natural variation of the primary gene pool (or any other gene pool) can be increased artificially by induced mutagenesis (Stadler, 1930). Large mutant collections have been developed for different plant species e.g. Arabidopsis (*Arabidopsis thaliana*) (O'Malley and Ecker, 2010), rice (*Oryza sativa*) (Hirochika et al., 2004) and also barley (Lundquist, 2009; Scholz F, 1962). By now, more than 10,000 barley mutant alleles have been characterized and are stored in the *ex situ* gene bank NordGen, Alnarp, Sweden (Lundquist, 2009). New traits from induced mutations have contributed to more than 300 barley varieties (http://mvgs.iaea.org/Default.aspx). Furthermore, induced barley mutants facilitated the confirmation of candidate genes. For instance, induced knock-out or loss-of-function mutations of candidate genes in independent plant material can confirm the function of candidate genes if an expected phenotype is obtained (Comadran et al., 2012; Mascher et al., 2014).

The secondary gene pool is composed of germplasm that belongs to a different species than the primary pool crop plant. They are more difficult to cross with the (primary gene pool) crop. Progeny may form at lower percentages and may be weak and only partially fertile or even infertile. Chromosome pairing in inter-pool hybrids may be poor or even completely lacking (Harlan and de Wet, 1971). *Hordeum bulbosum* L. (Figure 2) is the only member of the secondary gene pool of barley (von Bothmer et al., 1995). It has been proposed as a valuable genetic resource for barley, especially regarding biotic stress resistances (Pickering et al., 1987; Ruge et al., 2003; Szigat and Szigat, 1991; Walther et al., 2000). Furthermore, *H. bulbosum* is thought to comprise non-host resistances against barley pathogens (Johnston, 2007). A

comprehensive and detailed description of *H. bulbosum* as a genetic resource for barley will be given in the next paragraph.

Figure 2. Inflorescence of *Hordeum bulbosum*

This picture shows a flowering inflorescence (ear, spike) of *Hordeum bulbosum* at full anthesis. (Photograph taken by Heike Ernst, IPK).

Members of the tertiary gene pool are more distantly related to the primary gene pool. Gene transfer to the primary gene pool may only be possible by employing special techniques like embryo culture, protoplast fusion, chromosome doubling or bridge crossing (e.g. make crosses by involving members of the secondary gene pool) (Harlan and de Wet, 1971). Within the genus *Hordeum*, there are about 31 species belonging to the tertiary gene pool (von Bothmer et al., 1995) including diploid, tetraploid or hexaploid, annual, perennial, self-pollinating or out-breeding species (von Bothmer et al., 1995). The geographical distribution is considerably wide with species found in the northern as well as the southern hemisphere such as Central Asia, South Africa, Europe, the Middle East, North and South America (von Bothmer et al., 1995). Thus species of the genus *Hordeum* are endemic in areas from subtropical to arctic climate and from sea level to more than 4500 m above sea level (von Bothmer et al., 1995). Strong crossability barriers and low chromosome pairing occur in hybrids with *H. vulgare* and several attempts to use these species in crosses with *H. vulgare* have been unsuccessful (von Bothmer et al., 1983).

Research on hybrids from protoplast fusion between barley and species of other genera e.g. *Elymus*, *Thinopyrum* and *Pseudoroegneria* (Kim et al., 2008), rice (Kisaka et al., 1998), rye (Fedak and Armstrong, 1981) and even of different plant families like carrot (*Daucus carota)* (Kisaka and Kameya, 1998), soybean (*Glycine max)* (Kao et al., 1974) and tobacco (*Nicotiana tabacum)* (Somers et al., 1986) have been performed. However, obtained hybrids have mostly been sterile and so far they did not play any role for barley improvement.

It was proposed that in modern terminology, genetic resources should also include isolated genes (Becker, 1993). Thus this fourth class of genetic resources could include genes originating from any plant, animal or even microorganism species (Becker, 1993). However, the utilization of this fourth resource requires the implementation of transformation technology for genetic modification. In fact, the fourth gene pool class can play a substantial role for crop improvement as reviewed in Dunwell (2000) and Phipps and Park (2002). Genetic modification is a well-established method for barley (Ji et al., 2013). Nonetheless, genetically modified barley varieties are not in commercial use owing to high regulation costs and a low acceptance by consumers, especially in Europe (Verstegen et al., 2014). However, genetic modification has been successful in barley research mainly for the functional validation of candidate genes as reviewed earlier (Goedeke et al., 2007). The barley gene pool concept is visualized in Figure 3.

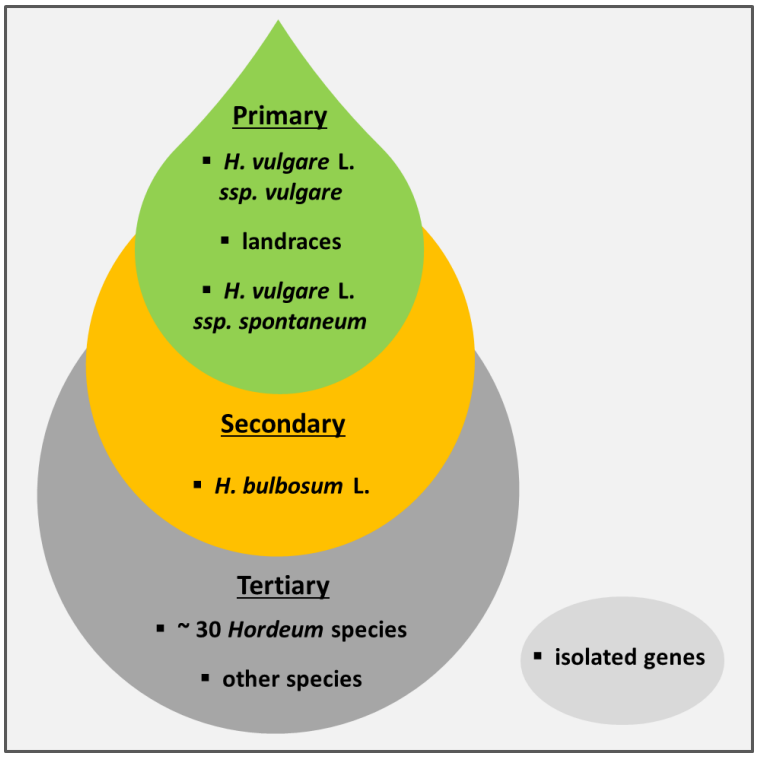

Figure 3. Gene pool concept for barley

This figure illustrates the concept of four gene pools available for barley improvement (Becker, 1993; Harlan and de Wet, 1971). The primary gene pool includes wild barley (*H.vulgare ssp. spontaneum),* barley landraces and cultivars. All members of the primary gene pool mate freely. The secondary gene pool contains *H. bulbosum,* which can be crossed with members of the primary gene pool to produce (partially) fertile hybrids if employing embryo rescue. The tertiary gene pool contains species that are more distantly related to cultivated barley. In nature, hybrids between the third gene pool species and barley do not occur spontaneously. Thus, gene transfer is not possible without applying biotechnology. The forth gene pool represents no living species but includes individual genes that may be transferred to barley by genetic transformation from any species including animals and microorganisms. The figure was modified after Harlan and de Wet (1971) and Becker (1993).

1.1.2 *Hordeum bulbosum* as a resourcefor barley improvement

H. bulbosum is a perennial *Hordeum* species, which has a broad geographical distribution along the Mediterranean basin and the Fertile Crescent (von Bothmer et al., 1995) (Figure 4). *H. vulgare* and *H. bulbosum* were found to constitute the same genome, nowadays generally designated as the 'H' genome (Blattner, 2009). *H. bulbosum* can be propagated sexually via seeds

and also vegetatively via bulbs that form by swelling of the basal internodes giving rise to the species' name of *H. bulbosum*. The out-breeding (cross-fertilizing) species exists as diploid as well as auto-tetraploid cytotypes, and it has a strong, two gene based self-incompatibility system (Lundqvist, 1962). Usually *H. bulbosum* and *H. vulgare* exhibit strong crossability barriers that interfere with the formation of fertile hybrids. If crossed with *H. vulgare*, the *H. bulbosum* chromosomes usually get eliminated during mitosis, leaving behind a haploid barley embryo (Kasha et al., 1970). When treated with colchicine, the chromosomes of the haploid plant can be doubled, resulting in fertile and completely homozygous plants (doubled haploid). Due to this, *H. bulbosum* has been frequently used in barley breeding for the production of doubled haploids (Kasha et al., 1970). However, this approach has recently been largely replaced by anther and microspore culture that were found to be more efficient (Pickering and Devaux, 1992) .

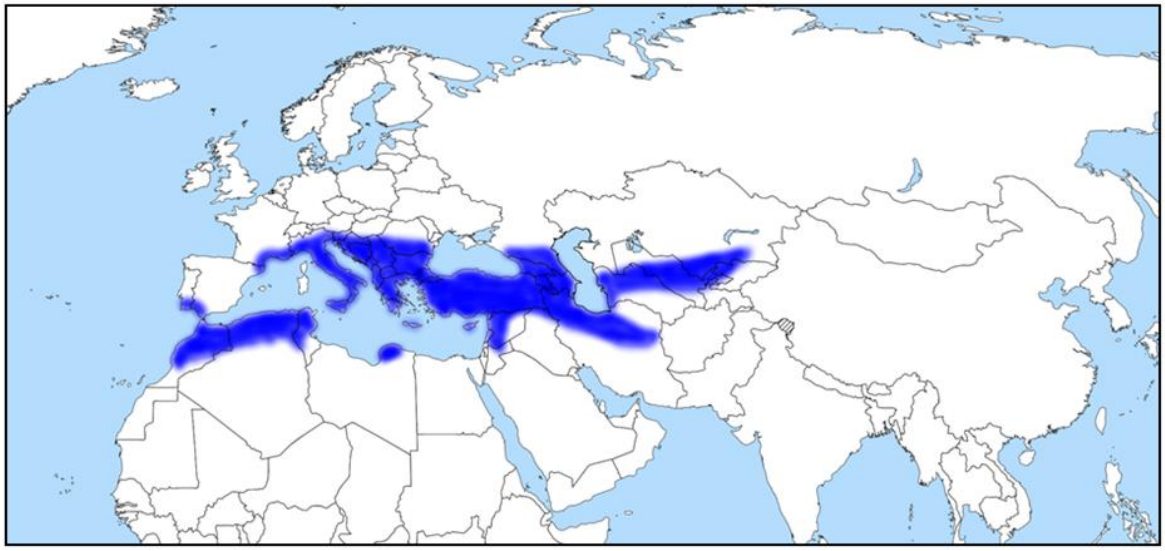

Figure 4. Geographic range of the natural habitat of the species *Hordeum bulbosum*

The map visualizes the geographic range where *H. bulbosum* occurs in natural populations (regions marked in dark blue color). The figure was modified after von Bothmer et al. (1995) and the plain map was freely available and modified from "Wikimedia Commons" (https://upload.wikimedia.org/wikipedia/commons/6/63/A_large_blank_world_map_with_oceans_marked_in_blue.svg).

Besides its importance for doubled haploid plant production, *H. bulbosum* is of interest for barley improvement, mainly as a potential source of genes controlling important agronomic traits such as abiotic and biotic stress resistance or tolerance, perenniality and self-incompatibility

(Pickering and Johnston, 2005). To exploit these traits it is necessary to transfer at least partially the *H. bulbosum* genome to *H. vulgare.* An important step towards this aim was the discovery that the retention of *H. bulbosum* chromosomes can be archived if the genome ploidy ratio favored the *H. bulbosum* parent (Kasha and Sadasiva, 1971). Subsequently if was found that *H. bulbosum* chromosome elimination strongly depends on the parental genotypes and that chromosomes retained at much higher frequency when the plants were crossed and maintained at temperatures below 18°C (Pickering, 1984). As a general feature of wide hybridizations, endosperm degradation occurs gradually after fertilization in crosses between barley and *H. bulbosum* (Pickering and Johnston, 2005). Thus, embryo culture is necessary to regenerate *H. vulgare/H. bulbosum* hybrids. Diploid (VB) hybrids (where V and B denote the haploid genomes from *H. vulgare* and *H. bulbosum*, respectively) were completely sterile, but colchicine treatment allowed regenerating partially fertile tetraploid (VVBB) hybrids (Pickering, 2000). Alternatively, tetraploid hybrids have been obtained by doubling the barley and *H. bulbosum* chromosomes prior to crossing (Szigat and Pohler, 1982). Triploid (VBB) hybrids, produced by crossing diploid barley with tetraploid *H. bulbosum* were partially fertile without colchicine treatment (Pickering, 1991b). Cytogenetic analysis of *H. vulgare/H. bulbosum* hybrids revealed that homeologous pairing between *H. vulgare* and *H. bulbosum* chromosomes occurred during meiosis (Pickering, 1991a). The latter is crucial to obtain recombination between the genomes of *H. vulgare* and *H. bulbosum* within the hybrid. The frequency of homeologous pairing was found to depend on the involved chromosomes (Pickering, 1991a) and on the parental genotypes (Thomas and Pickering, 1985).

After the production of (partially) fertile hybrids, the next step towards the utilization of *H. bulbosum* traits for barley improvement was the transfer of chromosomes and in particular of chromosome segments to barley via recombination and introgression. It was in 1982, when the first evidence was observed that an *H. bulbosum* chromatin introgression was obtained from a hybrid between barley and *H. bulbosum* (Szigat and Pohler, 1982). The transfer of *H. bulbosum* chromatin into the barley genome could be confirmed at the molecular level after the emergence of techniques such as *in situ* hybridization and Southern hybridization (Pickering et al., 1995; Xu and Kasha, 1992). In general, recombination between *H. vulgare* and *H. bulbosum* chromosomes in hybrids is possible, but the frequency is low (Pickering, 1991a). Consequently, "high-pairing" *H. vulgare* and *H. bulbosum* genotypes were selected for crossing to increase the possibility of

homeologous recombination. However, in a subsequent study it was found that recombination frequency was not strongly correlated with homeologous chromosome pairing (Zhang et al., 1999).

Despite all these obstacles, a considerable number of introgression lines (IL)s and substitution lines between barley and *H. bulbosum* could be generated for all barley chromosomes except for a chromosome 1H substitution (Johnston et al., 2009; Pickering et al., 1987; Pickering et al., 1994; Ruge et al., 2003; Szigat and Szigat, 1991; Walther et al., 2000). ILs and substitution lines can be obtained from hybrids by backcrossing with the recurrent barley parent or through selfing (Pickering, 2000; Pickering et al., 1994) (Figure 5). Subsequently, some progeny plants appear to have a diploid *H. vulgare* genome with either introgressed fragments of *H. bulbosum* chromosomes (ILs) or complete, substituted chromosomes (substitution lines). So far, introgressed segments of *H. bulbosum* could only be found at distal positions of the barley chromosomes (Johnston et al., 2009). A large and diverse set of *H. bulbosum* traits have been transferred to barley via those ILs, such as disease resistance against leaf rust (*Puccinia hordei*), barley mild mosaic virus (BaMMV), barley yellow mosaic virus (BaYMV-1,-2), barley yellow dwarf virus (BYDV), stem rust (*Puccinia graminis*), scald (*Rhynchosporium commune*), powdery mildew and septoria speckled leaf blotch (*Septoria passerinii*) (Fetch et al., 2009; Pickering et al., 2004; Ruge et al., 2003; Scholz et al., 2009; Shtaya et al., 2007; Singh et al., 2004; Toubia-Rahme et al., 2003). Primary introgressions may be recognized by a number of morphological characteristics acquired from *H. bulbosum* such as pubescent leaf sheath, overall growth habit, agronomic performance, or by phenotypic tests for disease resistance. Furthermore, the pSc119.1 retrotransposon-based Polymerase Chain Reaction (PCR) assay allowed determining the presence of *H. bulbosum* chromatin since no amplification occured using *H. vulgare* DNA (Johnston and Pickering, 2002; Johnston et al., 2009).

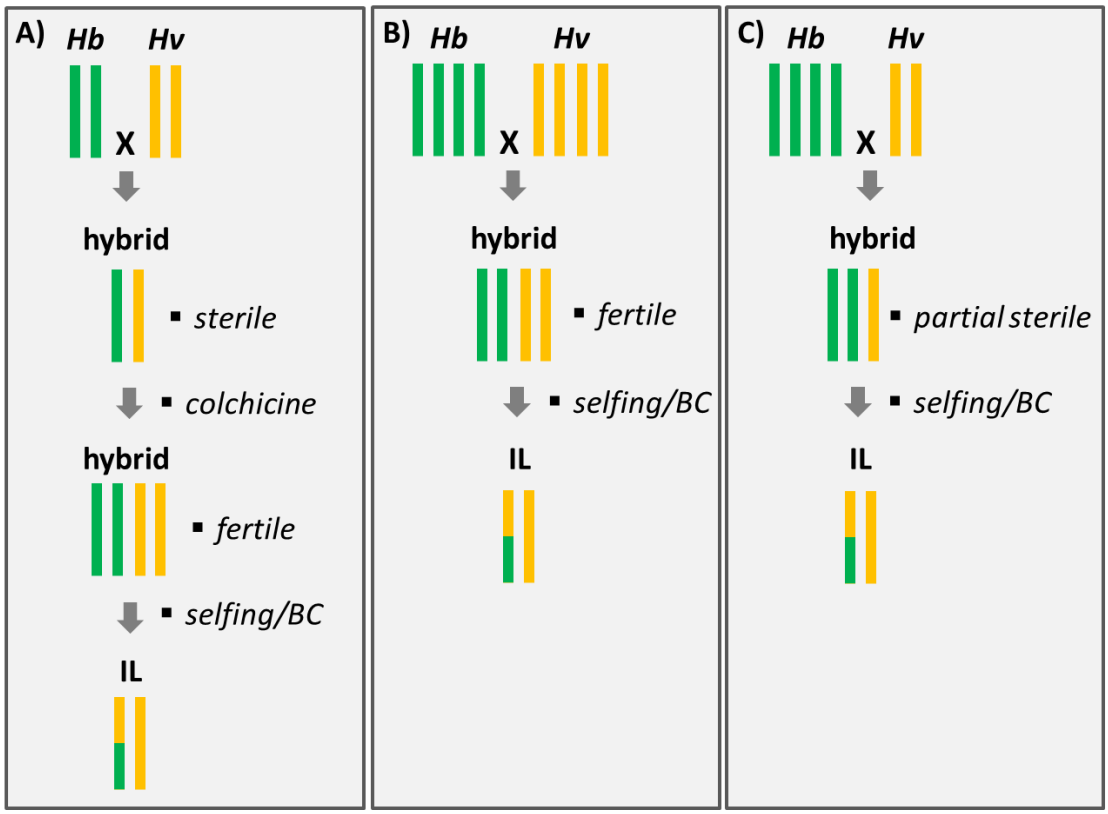

Figure 5. Breeding schema for obtaining introgression lines from *H. vulgare/H. bulbosum* hybrids

This schema illustrates different crossing strategies for the generation of introgression lines between *H. vulgare* (*Hv*) and *H. bulbosum* (*Hb*). Green and orange bars represent single *H. bulbosum* or barley chromosomes, respectively. Thus a single bar illustrates a haploid genome, while two bars of same color represent a diploid genome and so on. A) Diploid barley is crossed with diploid *H. bulbosum*. The obtained hybrid has to be treated with colchicine to induce chromosome doubling and restore fertility. Selfed or back-crossed progeny can then be screened for recombination and introgression events between barley and *H. bulbosum* chromosomes. B) Alternatively, both tetraploid barley and *H. bulbosum* may be crossed. The resulting hybrid will be fertile and can be processed as described above. C) A third option is to cross tetraploid *H. bulbosum* with diploid *H. vulgare*. Hybrids derived from that cross are usually partially fertile and therefore may be either selfed or back-crossed to generate recombinant progeny.

The formation of primary ILs is only representing the initial step towards utilization of *H. bulbosum* characters in *H. vulgare* breeding. Negative linkage drag, that is usually associated with *H. bulbosum* introgressions, has to be eliminated by reducing the size of the initially

transferred *H. bulbosum* segment (Pickering and Johnston, 2005). This requires at least several rounds of selfing or back-crossing. However, since interspecific recombination frequency between *H. vulgare* and *H. bulbosum* chromatin is usually strongly reduced (Johnston et al., 2013; Ruge-Wehling et al., 2006), large populations are required to increase the probability of finding a recombinant individual with appropriately reduced introgression size. This also evokes the need for a detailed genetic characterization. *H. bulbosum* segments must be localized within the barley genome, the precise dimension of each introgression must be determined and no *H. bulbosum* segment should remain undetected. Subsequently, informative markers may be developed for tracing the introgressed segment by marker assisted selection and eventually, mapping and cloning of a gene underlying a favorable trait of *H. bulbosum*. Initial characterization and genomic localization of *H. bulbosum* segments can be performed by *in situ* hybridization (Johnston et al., 2009). This should allow addressing the particular chromosome arm carrying the *H. bulbosum* segment. However, the resolution obtained by this method is usually very low, thus small *H. bulbosum* segments may remain undetected (Lukaszewski et al., 2005). Therefore, molecular markers are an important tool and prerequisite for a detailed characterization of introgressions obtained from *H. bulbosum.* Unfortunately, the development of molecular markers for *H. vulgare/H. bulbosum* ILs has been difficult. Despite the advent of high-throughput sequencing and genotyping technologies, tools for an efficient genetic characterization of *H. vulgare/H. bulbosum* ILs were still lacking.

1.2 Genomic resources and approaches established in *Hordeum vulgare*

While the genomic infrastructure such as sequence information was lacking for *H. bulbosum*, genomic resources are plenty for *H. vulgare* (Close et al., 2009; Comadran et al., 2012; Mascher et al., 2013c; Poland and Rife, 2012; Sato et al., 2009; Stein et al., 2007). In order to allow a more straightforward characterization or exploitation of *H. bulbosum* introgression lines, such genomic resources have to be either developed for *H. bulbosum*, or markers that differentiate between *H. vulgare* and *H. bulbosum* alleles in ILs, can be designed solely based on the *H. vulgare* resources. Furthermore, since the genomes of *H. vulgare* and *H. bulbosum* are expected to be highly conserved and collinear (Jaffe et al., 2000; Salvo-Garrido et al., 2001), the barley genomic infrastructure might also support the development of genomic resources in *H. bulbosum*.

1.2.1 Barley reference maps

An advanced physical map of the barley genome integrated with a high density genetic map and gene-space sequence information was published in 2012 by the International Barley Sequencing Consortium (International Barley Sequencing Consortium, 2012; Schulte et al., 2009). This framework provided a draft reference sequence of the barley genome, which could be utilized to align newly produced re-sequencing data. Thus, genomic sequence information derived from genotypes different than the reference genotype "Morex" can be aligned to the draft reference genome for sequence variant identification.

A physical map of a genome can be built by reconstructing chromosomes from overlapping subgenomic fragments cloned into genomic DNA libraries. In case of barley genomic DNA fragments of more than 100 kbp in size (Schulte et al., 2011), were cloned into bacterial artificial chromosomes (BAC)s. These fragments were produced by either mechanical shearing of or by incomplete enzymatic digestion of high molecular DNA with one out of four different restriction enzymes, resulting in six different BAC libraries (International Barley Sequencing Consortium, 2012; Schulte et al., 2011; Yu et al., 2000). All libraries were produced from the genomic DNA of a single genotype; barley cultivar 'Morex'. Such genomic clones can be analysed by restriction fingerprinting producing a specific sequence-based restriction pattern, which can be utilized to determine physical overlaps between independent genomic clones (Luo et al., 2003) (Schulte et al., 2011; Soderlund et al., 2000). The barley physical map was obtained from high quality fingerprint patterns derived from 571,000 BACs yielding 9255 BAC contigs that represent 95 % (4.98 Gbp) of the barley genome (International Barley Sequencing Consortium, 2012). A subset of 67,000 BAC clones was identified to represent the seven barley chromosomes with minimal possible redundancy (minimum tiling path, MTP) (International Barley Sequencing Consortium, 2012).

To enrich the physical map with sequence information, WGS sequencing was performed for the reference genotype 'Morex' but also other genotypes like cultivars 'Barke' and 'Bowman'. WGS is performed by mechanically shearing and subsequent sequencing of genomic DNA fragments of an average size of 100 – 500 bp. The sequence reads are then assembled based on sequence overlaps between the reads to construct larger WGS sequence contigs. Since most repetitive

DNA cannot be assembled, the barley WGS assembly is gene-focused and it was estimated to contain 86 % of the entire barley gene content (International Barley Sequencing Consortium, 2012). In order to be able to assign WGS contigs to their corresponding physical map contigs, complete or partial sequence information was produced for a subset of BAC clones (International Barley Sequencing Consortium, 2012). Subsequently, 308 Mbp of WGS sequence contigs could be anchored to the BAC-based physical map framework based on sequence homology.

To resolve the order of this sequence-enriched BAC contigs, high-density genetic maps were employed (International Barley Sequencing Consortium, 2012). Information of about 500,000 molecular markers allowed genetically ordering of the physical map. In total 4,556 physical BAC contigs spanning 3.9 Gbp of the barley genome could be assigned to genetic positions along the barley chromosomes (International Barley Sequencing Consortium, 2012).

26,159 genes were annotated for the barley whole genome shotgun sequence assembly based on expressed gene sequence information (International Barley Sequencing Consortium, 2012). About 20,000 of these "high-confidence" genes could be integrated into the physical/genetic framework (International Barley Sequencing Consortium, 2012).

In addition to the above mentioned 500,000 genetic markers used to anchor the BAC contig and WGS sequence based physical map to the genetic map an alternative strategy, POPSEQ (population sequencing) facilitated genetic anchoring of the barley reference WGS contigs (Mascher et al., 2013a). In the POPSEQ approach millions of markers were developed by low coverage (1-2 fold genome coverage) WGS sequencing of two relatively small mapping populations of 90 recombinant inbreed lines and 82 doubled haploid lines. SNP markers were discovered by mapping of the newly developed WGS sequence reads onto the WGS reference sequence assembly of cultivar 'Morex' (International Barley Sequencing Consortium, 2012). This allowed integrating 4.3 million POPSEQ SNP markers into a previously developed high confidence barley map that was based on ~ 5000 molecular markers (Comadran et al., 2012; Mascher et al., 2013a). By this, about 500,000 barley reference WGS contigs with a cumulative length of 927 Mbp could be ordered (Mascher et al., 2013a). This represented nearly 20 % of the barley genome and it allowed to anchor to the physical/genetic barley map more than twice as much WGS data than without POPSEQ (International Barley Sequencing Consortium, 2012;

Mascher et al., 2013a). Thus, the barley POPSEQ map is a valuable additional genomic resource with anchored barley sequence information.

1.2.2 Reduced complexity sequencing

The most straightforward way of developing molecular markers that differentiate between barley and *H. bulbosum* would be based on comparing genome wide sequence information of the genotypes of interest. Today, NGS is principally enabling to produce such data in a matter of weeks. However, even though NGS costs are dropping; it still remains rather expensive to completely sequence large genomes. Furthermore, over 80 % of the barley genome are composed of repetitive DNA sequences, which cannot be efficiently used to build a reference genome assembly (International Barley Sequencing Consortium, 2012). Repetitive sequences tend to collapse during sequence assembly, thus most of them are not represented in the WGS barley reference assembly (International Barley Sequencing Consortium, 2012). Therefore, in most cases, sequencing of whole genomes would be expensive and inefficient, since a majority of the produced sequence may not be utilized. A more efficient way of generating genome wide sequence information would be by targeting only the non-repetitive part of the genome, hence reducing the complexity of the sequencing target. Since genes represent most of the non-repetitive part of the genome, they are in focus of targeted enrichment strategies for genome re-sequencing. Genes evolve relatively slow and are more conserved e.g. between species (Morishige et al., 2002). This feature makes them an ideal resource for comparative genomics and other research purposes, especially beyond species boundaries. For instance about 1,000 markers derived from expressed sequence tags (EST)s in barley were efficiently used to revisit synteny between rice and barley (Stein et al., 2007; Thiel et al., 2009).

By now, several strategies of reduced representation sequencing of genomic DNA have been developed (Davey et al., 2011; Mardis, 2010; Ray and Satya, 2014). There are two basic principles for partial genome representation libraries (Ray and Satya, 2014). Complexity reduction is either achieved by capture assays that are designed for targeted genomic sections (Ray and Satya, 2014), or using digestion by restriction enzymes and sequencing of adjacent regions in sub-genomic fractions (Davey et al., 2011). The former may be applied for a wide or a specific range of genome targets such as exons (Teer and Mullikin, 2010) and other genomic regions such as gene families (Jupe et al., 2013; Teer et al., 2010). Regarding the second type of

strategies to reduce genome complexity for sequencing, various techniques have emerged during the last years including reduced-representation libraries (Gore et al., 2009), complexity reduction of polymorphic sequences (Mammadov et al., 2010), restriction-site associated DNA sequencing (RAD-seq) (Pfender et al., 2011), sequence based polymorphic marker technology (Sahu et al., 2012), multiplexed shotgun genotyping (Andolfatto et al., 2011), and genotyping- by-sequencing (GBS) (Elshire et al., 2011). Recently, exome capture and GBS have been adapted to barley (Elshire et al., 2011; Mascher et al., 2013b; Poland et al., 2012) opening new prospects in barley genomics and genetics.

1.2.2.1 Exome capture

Exome capture is a method for targeted re-sequencing of the gene space (Bamshad et al., 2011) and it has proven successful e.g. for mutation identification in human cancer research (Mardis, 2010). In principle whole genome exome capture should allow to enrich all exon sequences of a genome prior to sequencing. For this, a specific set of oligonucleotide probes (baits) is required that is complementary to the targeted exon regions. During library preparation these baits will hybridize to the corresponding sequences of the DNA sample, which was previously mechanically sheared into shorter fragments. The baits with the hybridized DNA segments can then be filtered and enriched from the remaining sample DNA by e.g. biotin tags (Bainbridge et al., 2010). A key point of this method is that the captured sequences do not need to exhibit 100 % similarity to the assay baits. For example, by reducing the stringency of hybridization conditions it was possible to enrich the sugarcane (*Saccharum officinarum)* exome, based on capture probes designed from sorghum (*Sorghum bicolor)* DNA sequence (Bundock et al., 2012). This feature makes it a valuable platform for research in species lacking genomic sequence information as well as for studies in comparative genomics and for species without a reference genome. Thus, it was possible to predict SNPs between different sugarcane genotypes as well as between sugarcane and sorghum itself, when the exome sequencing reads were mapped onto the original sorghum sequence (Bundock et al., 2012).

A barley whole exome capture was designed based on the annotated genes of the reference draft WGS assembly of barley including also information from full length cDNAs and de novo assembled RNAseq information (Mascher et al., 2013b) resulting in 61.6 Mbp of non-overlapping capture targets. 73.7 % and 40.7 % of the annotated barley high-confidence genes

and low confidence genes were covered, respectively (International Barley Sequencing Consortium, 2012; Mascher et al., 2013b). Based on these templates, millions of oligonucleotide baits were designed along target sequences. The barley exome capture was successfully applied to different barley species like *H. vulgare*, *H. spontaneum*, *H. bulbosum* and *Hordeum pubiflorum* but also to the more distantly related hexaploid species *Triticum aestivum* (bread wheat). Mapping and ordering of exome capture reads along the WGS reference assembly of barley (International Barley Sequencing Consortium, 2012) facilitated the prediction of thousands of potential SNPs within and between species (Mascher et al., 2013b). Therefore, it was hypothesized that "exome capture will become the most widely used approach for re-sequencing studies in barley and its relatives in the near future" (Mascher et al., 2013b). A flow chart highlighting the general procedure of exome capture is shown in Figure 6 A.

1.2.2.2 Genotyping by sequencing

Besides exome capture, also GBS protocols have been readily established for barley (Elshire et al., 2011; Poland et al., 2012). Here, genomic regions flanking restriction sites recognized by site-specific endonucleases are targeted for next generation sequencing. The basic idea is to choose appropriate restriction enzymes to avoid repetitive regions, while targeting low copy regions (Elshire et al., 2011; Gore et al., 2007; Gore et al., 2009). In barley, two protocols have been established, using either *Ape*KI (Elshire et al., 2011) or *Pst*I and *Msp*I in combination (Poland et al., 2012), as methylation sensitive restriction enzymes. Unlike gene space sequences, repetitive DNA in plants is usually highly methylated (Rabinowicz et al., 2005; Rabinowicz et al., 2003). Therefore, methylation sensitive enzymes will cut most dominantly in non-repetitive sequences and thereby facilitate the enrichment of the unique and gene-based genome portions. After digestion of genomic DNA, sequencing adapters are ligated to the resulting DNA fragments. These adapters provide the sequence required to bind the DNA sequencing sample to the sequencing flow cell and they contain the binding sites for primers that are needed during PCR enrichment and sequencing (sequencing primer). In the case of the *Pst*I and *Msp*I double-digestion protocol, the implementation of a "Y" adapter ensures the amplification of DNA fragments only if they contain both of the two different cleavage sites (*Pst*I and *Msp*I) (Poland et al., 2012). Thus, the "Y" adapter contains an overhang to ensure that primers for this respective adapter can only bind after one round of amplification through the reverse primer that binds to the non-"Y" adapter. However, the most important feature is that the adapters contain unique

short sequences (indices or barcodes) for every DNA sample. Such “barcodes” allow pooling of multiple samples prior to sequencing, which is required to fully take advantage of the capacity of NGS platforms. After sequencing, the unique barcode allows tracing back of each sequence read to the corresponding sample. Since the GBS targeted proportion of the genome is considerably small, large numbers of samples can be sequenced together as pools, so sequencing costs per sample are relatively low. Furthermore, the preparation of the GBS library itself is relatively easy and cheap. Elshire et al. (2011) suggested that GBS costs may very soon be dropping below 20 or maybe even 5 US Dollars per sample. The low costs allow performing GBS for large numbers of samples e.g. complete mapping populations. The identified polymorphisms within the genotyped samples can be directly scored as markers, thus no further marker design is needed. Both GBS protocols enabled to reliably detect thousands of SNPs in barley (Elshire et al., 2011; Poland et al., 2012). The double-digestion method further improved the specificity and efficiency of the GBS protocol (Poland et al., 2012). A flow chart of the experimental GBS procedure is given in Figure 6 B.

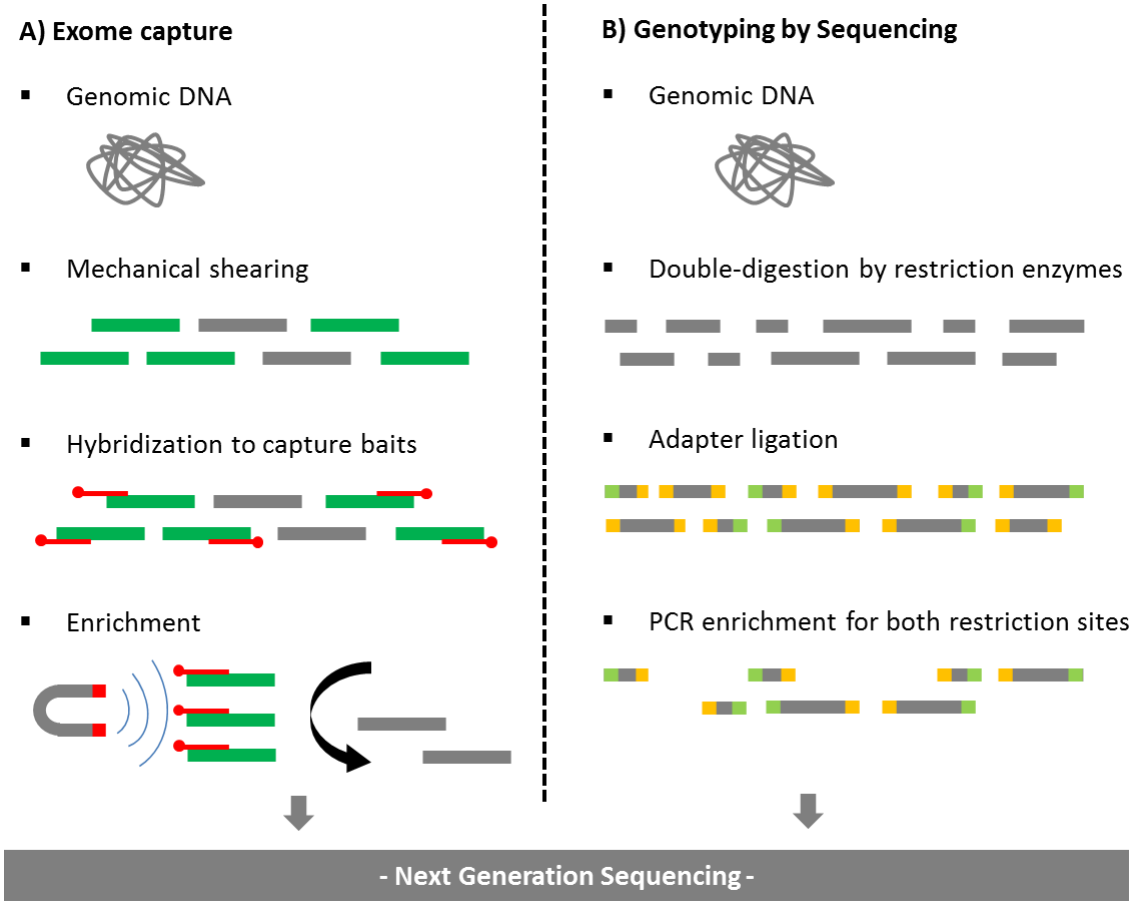

Figure 6. Principles of the two complexity-reducing re-sequencing strategies: exome capture and GBS

Main steps for library preparation involved in exome capture (A) and genotyping-by-sequencing (B). **A)** For exome capture, genomic DNA is first mechanically sheared to smaller fragments by ultrasound. Then capture baits (red lines with red points) – oligonucleotides that are complementary to (barley) exon sequences - will hybridize to the corresponding fragments within the sample. Since the baits contain biotin tags, the hybridization complexes can be enriched via magnetic beads. Simultaneously the non-hybridized portion of the library can be discarded. The image has been modified after (Bamshad et al., 2011). **B)** For genotyping-by-sequencing the genomic DNA is digested by up to two restriction enzymes. Then adapters corresponding to the different restriction site overhangs can be ligated to the generated fragments. If two restriction enzymes are employed, a PCR step will facilitate the enrichment of those fragments that contain two different restriction overhangs only. Both, exome capture and GBS libraries are then sequenced using next generation sequencing platforms.

1.4 Aims of the study

The genetic diversity in *H. bulbosum* populations can be captured and transferred to barley via *H. vulgare/H. bulbosum* ILs and it thus potentially represents a highly valuable resource for barley improvement. Nonetheless, the efficient utilization of *barley/H. bulbosum* ILs has so far been mainly hampered by the lack of suitable molecular tools and infrastructures for their proper genetic characterization. The overall aim of the present PhD study was the development and establishment of approaches and resources for a highly efficient characterization of barley/*H. bulbosom* ILs. The content of this 'cumulative' thesis is essentially based on the aims and results of two peer-reviewed and published scientific articles. The aims of these articles were:

- To establish the necessary molecular methods for the detailed and efficient genetic characterization of *H. vulgare/H. bulbosum* ILs (Wendler et al. 2014);
- To utilize the previously established methods (GBS and exome capture) in order to characterize genotypically a large collection of ILs, and generate a sequence resource that will greatly facilitate future marker development in any barley/*H. bulbosom* IL (Wendler et al. 2015).

Wendler et al. 2014:

Since barley and *H. bulbosum* are closely related (Jakob and Blattner, 2006) and are thought to share rather collinear genomes (Salvo-Garrido et al., 2001), it was postulated that barley genomic resources and tools could be utilized for the characterization of *H. vulgare/H. bulbosum* ILs. The usefulness of two NGS-based, genome-complexity reduction technologies (GBS and exome capture) in combination with the barley genome reference should be evaluated for characterizing *H. vulgare/H. bulbosum* ILs.

Towards this aim, I performed i) exome capture based re-sequencing for three *H. vulgare/H. bulbosum* ILs and their respective barley and *H. bulbosum* donors. Re-sequencing reads were mapped and ordered along the barley draft reference sequence (International Barley Sequencing Consortium, 2012) for sequence variant detection. By that, detected polymorphisms allowed the precise delimitation of introgressed *H. bulbosum* chromatin within the barley genome.
ii) Subsequently, barley/*H. bulbosum* variants were validated by designing a marker assay of 96 SNP markers, which were applied to a population segregating for a *H. bulbosum* segment.

iii) Alternatively, GBS was directly applied to the same mapping population and the respective donor genotypes of barley and *H. bulbosum*. Mapping of re-sequencing reads along the barley reference (International Barley Sequencing Consortium, 2012) allowed the detection of polymorphisms and delimitation of the introgressed *H. bulbosum* segment.
iv) Based on these results, GBS and exome capture were evaluated regarding their general usefulness for effective characterization of barley/*H. bulbosum* ILs.

Wendler et al. 2015:

Over the last decades, a large number of barley/*H. bulbosum* ILs has been generated, but most of these lines had been genotyped only roughly (Johnston et al., 2009). Both, GBS and exome capture proved to be useful methods for characterization of such genetic material (Wendler et al., 2014). Since GBS was of lower cost per individual sample it was chosen for genotyping a larger collection of barley/*H. bulbosum* ILs in a cost-effective manner. GBS sequences and polymorphisms information are made available for future studies as a community resource.

Towards this aim, I genotyped v) 146 ILs and most of the respective donor lines using GBS. Mapping of GBS sequence reads against the barley reference (International Barley Sequencing Consortium, 2012; Mascher et al., 2013a) allowed the precise delimitation of introgressed *H. bulbosum* segments in conjunction with comprehensive sequence polymorphism information for future marker development.

vi) A strategy was proposed to tackle the problem of reduced interspecific recombination frequencies by exploiting the newly derived information about size and location of introgressions in the large panel of ILs. In brief, two ILs with overlapping introgression segments might be crossed to establish a segregating population, which will be based on homologous intraspecific recombination between *H. bulbosum* derived segments. Hence, recombination frequency should restore to normal frequencies since interspecific limitations of recombination will be removed for the two *H. bulbosum* segments. The precise delimitation of introgressed segments as well as the information on polymorphisms for the respective *H. bulbosum* genomic region should support selecting appropriate crossing partners within the panel of ILs.

vii) An integrated sequence resource was developed based on exome capture data of thirteen barley and five *H. bulbosum* accessions. Interspecifically conserved SNPs between barley and *H.*

bulbosum were surveyed and anchored to the barley draft reference genome. These anchored SNPs are likely to be conserved between any barley and *H. bulbosum* accession. Thus, they should be applicable as targets to developed molecular markers in any barley/*H. bulbosum* IL, avoiding the need to produce case specific exome capture data.

2. Unlocking the secondary gene-pool of barley with next-generation sequencing

Neele Wendler[1], Martin Mascher[1], Christiane Nöh[2], Axel Himmelbach[1], Uwe Scholz[1], Brigitte Ruge-Wehling[2] and Nils Stein[1,*]

[1] Leibniz Institute of Plant Genetics and Crop Plant Research (IPK), Gatersleben, Germany
[2] Julius Kühn-Institut (JKI), Institute for Breeding Research on Agricultural Crops, Groß Lüsewitz, Germany

*Correspondence (Tel +49 (0)39482 5522; fax +49 (0)39482 5595; email stein@ipk-gatersleben.de)

Plant Biotechnology Journal (2014), pp. 1–10, doi: 10.1111/pbi.12219

Summary

Crop wild relatives (CWR) provide an important source of allelic diversity for any given crop plant species for counteracting the erosion of genetic diversity caused by domestication and elite breeding bottlenecks. *Hordeum bulbosum* L. is representing the secondary gene pool of the genus *Hordeum*. It has been used as a source of genetic introgressions for improving elite barley germplasm (*Hordeum vulgare* L.). However, genetic introgressions from *H. bulbosum* have yet not been broadly applied, due to a lack of suitable molecular tools for locating, characterizing, and decreasing by recombination and marker-assisted backcrossing the size of introgressed segments. We applied next-generation sequencing (NGS) based strategies for unlocking genetic diversity of three diploid introgression lines of cultivated barley containing chromosomal segments of its close relative *H. bulbosum*. Firstly, exome capture-based (re)-sequencing revealed large numbers of single nucleotide polymorphisms (SNPs) enabling the precise allocation of *H. bulbosum* introgressions. This SNP resource was further exploited by designing a custom multiplex SNP genotyping assay. Secondly, two enzyme- based genotyping-by-sequencing (GBS) was employed to allocate the introgressed *H. bulbosum* segments and to genotype a mapping population. Both methods provided fast and reliable detection and mapping of the introgressed segments and enabled the identification of recombinant plants. Thus, the utilization of *H. bulbosum* as a resource of natural genetic diversity in barley crop improvement will be greatly facilitated by these tools in the future.

Plant Biotechnology Journal

Plant Biotechnology Journal (2014), pp. 1–10 doi: 10.1111/pbi.12219

Unlocking the secondary gene-pool of barley with next-generation sequencing

Neele Wendler[1], Martin Mascher[1], Christiane Nöh[2], Axel Himmelbach[1], Uwe Scholz[1], Brigitte Ruge-Wehling[2] and Nils Stein[1,*]

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

[2]*Julius Kühn-Institut (JKI), Institute for Breeding Research on Agricultural Crops, Groß Lüsewitz, Germany*

Received 11 February 2014;
revised 23 May 2014;
accepted 29 May 2014.
**Correspondence* (Tel +49 (0)39482 5522; fax +49 (0)39482 5595; email stein@ipk-gatersleben.de)

Accession numbers: Accession numbers for sequence data generated in this study are ERP004444 (GBS reads) and ERP004445 (exome capture reads).

Keywords: *Hordeum bulbosum*, crop wild relatives, introgression line, genotyping-by-sequencing, exome capture, next-generation sequencing.

Summary

Crop wild relatives (CWR) provide an important source of allelic diversity for any given crop plant species for counteracting the erosion of genetic diversity caused by domestication and elite breeding bottlenecks. *Hordeum bulbosum* L. is representing the secondary gene pool of the genus *Hordeum*. It has been used as a source of genetic introgressions for improving elite barley germplasm (*Hordeum vulgare* L.). However, genetic introgressions from *H. bulbosum* have yet not been broadly applied, due to a lack of suitable molecular tools for locating, characterizing, and decreasing by recombination and marker-assisted backcrossing the size of introgressed segments. We applied next-generation sequencing (NGS) based strategies for unlocking genetic diversity of three diploid introgression lines of cultivated barley containing chromosomal segments of its close relative *H. bulbosum*. Firstly, exome capture-based (re)-sequencing revealed large numbers of single nucleotide polymorphisms (SNPs) enabling the precise allocation of *H. bulbosum* introgressions. This SNP resource was further exploited by designing a custom multiplex SNP genotyping assay. Secondly, two-enzyme-based genotyping-by-sequencing (GBS) was employed to allocate the introgressed *H. bulbosum* segments and to genotype a mapping population. Both methods provided fast and reliable detection and mapping of the introgressed segments and enabled the identification of recombinant plants. Thus, the utilization of *H. bulbosum* as a resource of natural genetic diversity in barley crop improvement will be greatly facilitated by these tools in the future.

Introduction

Genetic diversity of crop plants has narrowed constantly throughout the process of domestication and breeding. To ensure sustainability or even increase crop yields under changing environmental conditions, permanent breeding for crop improvement is a prerequisite. Crop wild relatives (CWR) are a valued source of allelic diversity for broadening the genetic basis of breeding germplasm (Hajjar and Hodgkin, 2007; Tanksley and McCouch, 1997). Traits were transferred from CWR in at least 13 crops (Hajjar and Hodgkin, 2007) - among them the four world's most important cereals; wheat (*Triticum aestivum* L.), rice (*Oryza sativa* L.), maize (*Zea mays* L.), and barley (*Hordeum vulgare* L.). It was calculated that the benefit from crop-wild introgressions used or imported in North America accounted for more than 340 million US dollar per year (Prescott-Allen and Prescott-Allen, 1986). As limitations due to crossing barriers may be overcome (i.e. Akano *et al.*, 2002; Mallikarjuna, 1999) there is an increasing need for advanced molecular tools that facilitate the efficient utilization of CWR in (pre)-breeding programs.

In barley, three gene pools may be considered as a source of new advantageous alleles in breeding (von Bothmer *et al.*, 1995). The primary gene pool of barley includes domesticated barley (*H. vulgare* L. ssp. *vulgare;* in the following *Hv*) and wild barley (*H. vulgare* L. *ssp. spontaneum;* in the following *Hs*). *Hs* has been repeatedly used to improve *Hv* elite cultivars (Kalladan *et al.*, 2013; Lakew *et al.*, 2013; Nevo, 1992). *H. bulbosum* L. (in the following *Hb*) is the only member of the secondary gene pool of barley. Crossing barriers between *Hb* and *Hv* can be overcome by modifying environmental conditions (Pickering, 1984), use of specific genotypes and biotechnological tools such as embryo rescue (Kasha and Sadasiva, 1971; Pickering, 1983). Thirty wild *Hordeum* species constitute the tertiary gene pool of barley; however, strong crossability barriers and low chromosome pairing in hybrids with *Hv* have interfered with their use in classical barley breeding (von Bothmer *et al.*, 1995). Interspecies hybrids between wild *Hordeum chilense* and domesticated tetraploid durum wheat (*Triticum durum*) were used to form a novel crop species *Tritordeum* with novel nutritional properties (Martin *et al.*, 1999) and *H. chilense* introgressions allowed to introduce abiotic stress tolerances into wheat (Forster *et al.*, 1990).

Hb has been recognized and was used as an important source of advantageous traits for *Hv*, especially with respect to pathogen resistance or tolerance (Pickering *et al.*, 1987; Ruge *et al.*, 2003; Szigat and Szigat, 1991; Walther *et al.*, 2000; Xu and Snape, 1988). The progeny of *Hv*/*Hb* hybrids were surveyed to identify *Hv*

Please cite this article as: Wendler, N., Mascher, M., Nöh, C., Himmelbach, A., Scholz, U., Ruge-Wehling, B. and Stein, N. (2014) Unlocking the secondary gene-pool of barley with next-generation sequencing. *Plant Biotechnol. J.*, doi: 10.1111/pbi.12219

plants carrying preferentially a single *Hb* introgression with a gene of interest (in the following Hv^b). Advantageous traits of *Hb* can often only be transferred at the cost of negative linkage drag that is introduced with the introgressed *Hb* segments—causing a strong penalty on breeding programs. Thus, in order to enable efficient utilization of introgressed *Hb*-loci for barley breeding, marker-assisted reduction of the size of introgressed *Hb* segments is needed. The availability of molecular markers is no longer a limiting factor for the differentiation of barley cultivars (Close *et al.*, 2009; Comadran *et al.*, 2012; Mascher *et al.*, 2013c; Poland *et al.*, 2012; Sato *et al.*, 2009; Stein *et al.*, 2007); however, markers that differentiate between *Hv* and *Hb* alleles are rare. The paucity of *Hb* specific molecular markers has so far been one of the main bottlenecks for the incorporation of *Hb* traits into elite barley germplasm.

The development of molecular markers is greatly facilitated by access to genomic or expressed sequence information of the genotypes of interest (Davey *et al.*, 2011). Even though next-generation sequencing (NGS) costs are continuously decreasing, sequencing of the whole barley genome still remains very costly due to its size (>5 Gbp) and high repetitive DNA content. Targeted sequencing of a sample with reduced complexity would simultaneously decrease sequencing costs and minimize redundancy.

One approach to complexity reduction is a targeted sequence capture, which is used to enrich for defined sequences prior to NGS (Asan *et al.*, 2011; Sulonen *et al.*, 2011). Recently, a barley exome capture assay, targeting ~62 Mb of exon sequence was developed and used to capture genomic sequence information of several barley cultivars and barley CWR (Mascher *et al.*, 2013b). A second approach, genotyping-by-sequencing (GBS), combines complexity reduction, multiplexing of samples, and the use of NGS methods for genotyping of whole mapping populations (Elshire *et al.*, 2011; Poland *et al.*, 2012).

Here, we utilized both, exome capture and GBS, to survey sequence polymorphisms in three independent Hv^b lines. The sequence information obtained by exome sequencing for one introgression on chromosome 2HL was used to design a marker assay to genotype a population segregating for the introgressed segment on chromosome 2HL. Both approaches proved to be highly efficient and are readily applicable for evaluation and characterization of CWR introgressions in other crop species.

Results

Identification of *Hv*/*Hb* SNPs and the *Hb* introgression intervals by exome capture

Three Hv^b lines and the respective donor lines were (re)-sequenced after exome capture, in order to survey sequence polymorphisms between *Hv* and *Hb* and to identify and locate the *Hb* introgression segment in the Hv^b lines. Although all three Hv^b lines were previously characterized, the working hypothesis of the present study was that the new approaches should principally allow the identification of the introgressions *de novo* at much higher precision than by previous attempts.

For the identification and comparison of sequence polymorphisms between *Hv*, *Hb* and Hv^b, sequencing reads of all samples were mapped to the barley reference. This allowed comparing sequence reads of each Hv^b line and the respective *Hv* and *Hb* progenitor genotypes based on their position within the reference sequence. Nearly, 90% and 70% of total *Hv* and *Hb* reads could be mapped to the reference, respectively (Table 1). Sequence or SNP targets between samples can only be compared if they have a given minimum read coverage across all samples. At a minimum 10-fold coverage, the three analysed Hv^b lines shared more than 45 Mb of sequence with both respective donor lines (data not shown).

Table 1 Exome capture and GBS mapping results

Method	Sample	% mapped	% mapped to EC*,‡	% mapped to HC†,‡	% mapped to LC†,‡
EC	A42	69	85	63	17
EC	Borwina	87	76	46	22
EC	Vogelsanger G.	88	76	47	22
EC	Hv^b 2HL	89	79	49	23
EC	Hv^b 5HL	80	63	38	18
EC	Hv^b 6HS	89	75	46	22
GBS	A42	37	50	43	13
GBS	Borwina	82	34	29	14

*EC: Exome capture target region.

†HC/LC: High/low confidence genes of barley (IBSC *et al.*, 2012).

‡The percentage of reads mapped to EC, HC, and LC regions is relative of mapped reads.

To localize the *Hb* introgression segments in the three Hv^b lines, we filtered for (only homozygous) single-nucleotide-polymorphisms (SNPs) between Hv^b and the respective *Hv* donor lines —assuming that sequences of these two genotypes should differ nearly exclusively in the region of the *Hb* introgression segment. To further validate the identified polymorphisms, *Hv*/Hv^b SNPs were only included if they had the same genotype calls in Hv^b (only homozygous) and its respective *Hb* donor lines (heterozygous and homozygous calls allowed). Filtered SNPs that fit all these criteria were considered as *Hb*-specific. The frequency of *Hb*-specific SNPs was then computed in non-overlapping 5 Mb bins along the physical map of barley (IBSC *et al.*, 2012), assuming that SNP frequency should be highest within the introgressed *Hb* segment (Figure 1). This assumption was based on previous findings (Mascher *et al.*, 2013b), that SNP frequency between the barley reference 'Morex' and *Hb* was at least 5 times higher than between the barley reference and cultivars of *Hv*.

Hb-specific SNP frequency revealed clear peaks (>600 to >1000 SNPs per 5 Mb bin) at 2HL, 5HL, and 6HS for the Hv^b lines holding previously allocated *Hb* introgressions at 2HL, 5HL, and 6HS, respectively (Figure 1). SNPs that matched our criteria for *Hb*-specific SNPs could also be indentified on other chromosomes (or chromosome arms) mainly towards the telomeres of the three Hv^b lines, albeit at much lower frequencies (<200 SNPs per 5 Mb bin). Instead of representing SNPs truly originating from a *Hb* introgression, these SNPs likely represent false positives, i.e. sequencing errors, incorrect read mapping, or SNPs between the (re)-sequenced *Hv* genotype and the original *Hv* donor plant that was used to generate the initial hybrid.

In case of the 2HL Hv^b genotype two introgressed *Hb* segments, separated by approximately 22 CentiMorgans (cM) (23 Mb) of *Hv* DNA, could be determined. Based on the barley reference (IBSC *et al.*, 2012), the proximal introgressed segment was located at approximately 110–114 cM while the distal introgression was placed approximately at 136–149 cM in the iSelect genetic map (Comadran *et al.*, 2012), and corresponded to 5 and 16 Mb of physical sequence, respectively. Within these two *Hb* segments, more than 4000 *Hb*-specific SNPs were

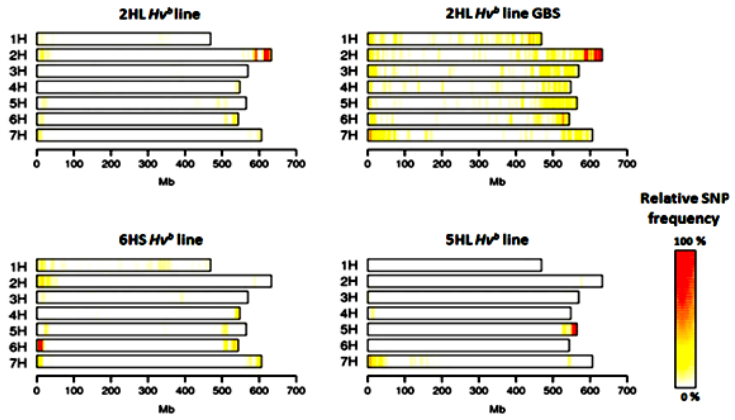

Figure 1 *Hb*-specific SNP frequency distribution in Hv^b lines. *Hb*-specific SNPs were discovered by exome capture-based (re)-sequencing of three independent *Hb* introgression lines, as well as by Genotyping-by-Sequencing (GBS) in case of 2HL Hv^b (upper right panel). The SNP frequency was visualized as a heat map of 5 Mb bins along the physical length of the seven barley chromosomes. The positions of contigs were taken from the barley reference (IBSC *et al.*, 2012). Relative SNP frequency was visualized in 1/1000 colour steps from white (0% SNP frequency) to red (100% SNP frequency) using heat.colors with the R statistical environment (R Core Team, 2012).

detected. The two other Hv^b introgressions on chromosomes 5HL and 6HS each comprised a single *Hb* introgression at approximately 167–169 cM (8 Mb) of chromosome 5HL and 0–10 cM (10 Mb) of chromosome 6HS, respectively. In both cases, more than 900 *Hb*-specific SNPs were assigned to the introgressed segments.

96-plex SNP Genotyping of a 2HL Hv^b mapping population

Out of 4000 SNPs from exome capture (re)-sequencing, nearly 200 fulfilled the 'Golden Gate assay' criteria, and a custom assay targeting 96 independent SNP loci was designed. The aim of this assay was to verify the exome capture SNP information and to identify recombination events in the segment introgressed from *Hb*. SNP markers were selected considering a more or less equidistant genetic distribution across the two introgressed segments according to the physical/genetic map of the barley reference. To analyse the detected pattern of the two (separated) introgressed segments in the 2HL Hv^b line, six SNPs were selected to be located proximal to the proximal *Hb* introgression and nine SNPs to be placed in between the two *Hb* introgressions.

The assay was used to genotype a 2HL Hv^b F_7 mapping population comprising 88 individuals, including three replicates of each donor (*Hv* and *Hb*) genotype. Genotyping resulted in 90% successful SNP markers (88 successful SNP markers), which gave clear and consistent genotype calls (max. 16% missing data) and all successful markers were polymorphic between the donor lines as predicted from the assay design. The 50% GC score of a SNP—a metric to indicate the reliability for the genotype calls (0 = failed, 1 = excellent) revealed that 88% of the loci targeted by the assay (84 SNPs) produced a 50% GC score of 0.5 or higher (Figure S1). Genotyping revealed six recombinant plants and proved the presence of the two detected *Hb* introgression segments as well as the presence of the *Hv* fragment that was identified between those segments. The proximal *Hb* introgression as well as the putative intercalary *Hv* segment was not segregating in the analysed mapping population; thus, orientation of these segments to each other could not be resolved. However, the high success rate and the high quality of the obtained genotype calls revealed the strong reliability of the sequences obtained by exome capture.

Genotyping-by-sequencing of the 2HL Hv^b introgression

Genotyping-by-sequencing (GBS) was applied as an alternative strategy for generating new molecular markers for *Hb* introgressions targeting the same individuals that were genotyped by the Golden Gate assay. Similar to the analysis of the exome capture data, GBS sequence reads were mapped to the barley reference for sequence variant detection. More than 80% of sequence reads of *Hv* but <40% of the *Hb* reads mapped to the barley reference (Table 1), respectively. A shared total of 2 Mb of sequence was represented by a minimum 10-fold sequence read coverage in 95% of all samples of the 2HL Hv^b mapping population, enabling high confidence SNP detection (Figure 2). High quality GBS SNPs between the *Hv* and *Hb* donor genotypes were selected based on genotype call quality, percentage of missing data, sequence read coverage and segregation ratio of parental alleles among the population.

In total, 57 *Hv*/*Hb* SNP markers could be scored on the basis of GBS data. Some SNP markers clustered with tight physical linkage of only several bp distances, thus, were redundant for the purpose of detecting recombination events. In such cases, only SNP markers lying on different reference contigs were selected as independent markers. Overall, 33 non-redundant GBS markers were used for mapping and identification of six recombination events in the introgression interval.

The analysed 2HL Hv^b mapping population was only segregating for the distal *Hb* introgression. Due to the filtering of GBS markers with a 1 : 2 : 1 segregation pattern, no markers were detected for the proximal *Hb* introgression; however, *Hb*-specific GBS SNP frequency indicated the detection of the proximal *Hb* introgression to the same region as identified by exome capture (Figure 1). Thus, GBS is an efficient method to detect and to precisely define even small *Hb* introgressions in an *Hv* background, similarly to exome capture.

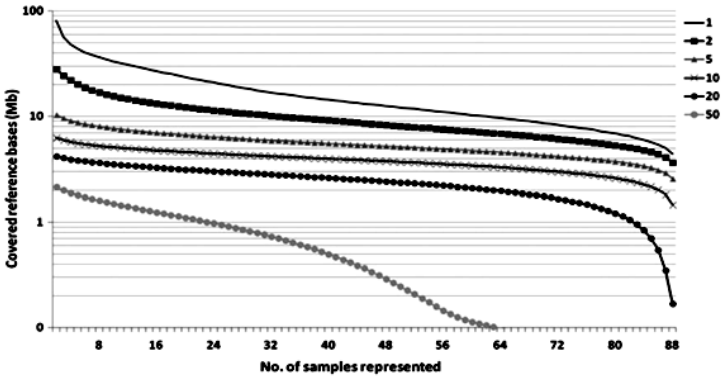

Figure 2 Read coverage across 88 GBS samples of the 2HL Hv^b mapping population. The cumulative size of sequence intervals on the barley reference, covered by at least x sequence reads (x = 1, 2, 5, 10, 20 or 50) that is shared in different numbers of GBS-sequenced samples of the 2HL Hv^b mapping population is visualized.

Genetic map of a 2HL Hv^b population and collinearity with barley

For construction of the genetic maps, a GBS marker that was showing three additional recombinants was excluded, as it was referring to the same barley reference contig with two other GBS markers, which did not support these recombination events. Furthermore, one GBS and one Golden Gate marker were (re)-sequenced using Sanger sequencing in order to revise two unexpected genotypes in one individual, respectively. Subsequently, genotyping results were corrected accordingly for map construction. Genetic maps were constructed only for the distal segment of the 2HL Hv^b line, since the proximal *Hb* introgression segment was not segregating.

When genotyping the 2HL Hv^b mapping population, GBS as well as the Golden Gate assay markers each revealed six recombinant plants. Two of these recombinants were differing between the data sets. GBS detected one recombination event towards the centromere, which was not recognized by GBS. The Golden Gate assay, however, could identify an additional distal recombination event, which in turn was not documented by the Golden Gate assay. Thus, both marker sets were spanning the same genomic regions with slightly differing extensions on either side and each of the maps covered a region that was comprised of the same 24 anchored BAC contigs of the physical map of barley (Ariyadasa *et al.*, 2014) with an estimated cumulative length of >17 Mb (Table S1). Both maps were spanning a distance of 3.4 cM and a consensus map interval of 4.5 cM, respectively. Recombination frequency was greatly reduced as compared to the published *Hv* genetic map of the same genomic region (Mascher *et al.*, 2013a), where the same genetic interval spanned more than 12 cM. When Golden Gate and GBS marker order was compared to the predicted marker position in the barley genomic reference (Mascher *et al.*, 2013a), two and three markers showed a rearranged order, respectively (Figure 3). However, further analysis is needed to determine if this finding is real or attributable to the limited genetic resolution of the barley reference genome.

Discussion

We demonstrated two efficient strategies for generating large numbers of molecular markers for the genetic resource of Hv^b

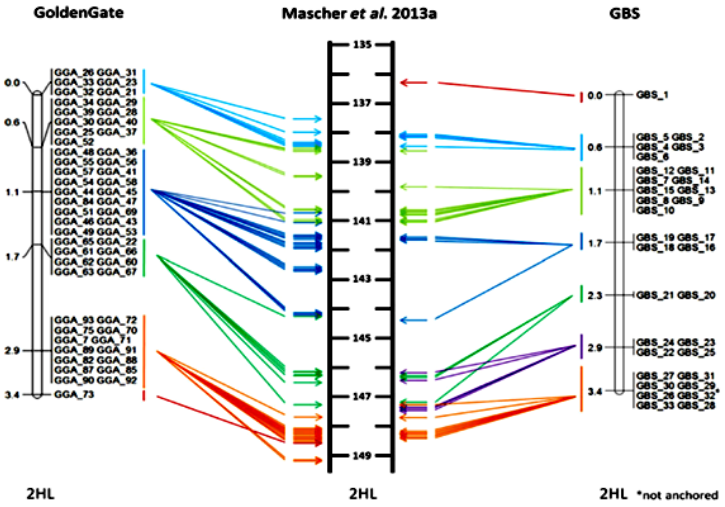

Figure 3 Comparison of marker order. The genetic maps derived by genotyping of the 2HL Hv^b mapping population using either GBS or the Golden Gate assay were compared to the genetic reference map of *Hv* (Mascher *et al.*, 2013a). Distances are given in centiMorgans (cM) and arrows indicate marker positions on the reference map.

lines. Our analysis took advantage of the recently published physical and genetic framework of barley (IBSC *et al.*, 2012; Mascher *et al.*, 2013b) that served both as a reference for computational read mapping and for the chromosomal localization of variants that were discovered.

Both methods allowed to precisely define the locations and sizes of the *Hb* introgressions and to identify recombinant plants. Exome capture based (re)-sequencing as well as GBS of the 2HL Hv^b line revealed an *Hb* introgression located distally at chromosome 2HL. These findings are in line with previous results by Ruge-Wehling *et al.* (2006) who genotyped this Hv^b line using RFLP, STS, and SSR markers. In this study, we identified an additional proximal, smaller *Hb* introgression at 110–114 cM, presumably 22 cM proximal to the larger *Hb* introgression at 136–149 cM on 2HL. Interestingly, this smaller and more proximal *Hb* segment was not detected earlier (Ruge-Wehling *et al.*, 2006), most likely due to a lack of markers. This underpins the strength of both NGS-based marker development strategies applied in the present work. Moreover, the *Hb* segments of two additional Hv^b lines could be identified in the expected chromosomal regions by means of exome capture, corroborating the sensitivity, reproducibility, and reliability of the approach. It is noteworthy that exome capture generates a substantial amount of additional information, such as a large number of SNP calls (i.e. 900–4000 SNPs for the *Hb* segments in the 3 Hv^b lines). Furthermore, the complete sequence of candidate genes may be contained in the exome capture data. In contrast, GBS does not come with this additional information; instead, it would require either deeper sequencing or new GBS libraries using different restriction enzymes, to generate further independent SNP (marker) information.

The power of both methods for producing SNP information was mainly limited by the efficiency of mapping *Hb* sequence reads to the barley reference (IBSC *et al.*, 2012). This limitation was more pronounced in the GBS approach compared to exome capture (re)-sequencing. Both methods allowed mapping of more than 80% of *Hv* sequence reads, while <40–70% of *Hb* sequence reads could be mapped for GBS and exome capture, respectively. Principally, it should be possible to increase mapping rates of *Hb* reads by decreasing stringency of the mapping parameters. Default parameters that were used in this study allow a maximum of 4% sequence polymorphism for a read to be mapped. This permits theoretically 1 sequence polymorphism to appear every 25 bp. As sequence divergence between *Hv* and *Hb* was found to be 1 SNP per 32.4 bp in introns and 1 SNP per 69.6 bp in exons (Johnston *et al.*, 2009), the default parameters are already beyond polymorphism rates. Therefore, decreasing stringencies seems to be unfavourable and would probably only increase miss-mapping of reads within repetitive DNA rather than increasing efficiency. In general, robust mapping of *Hb* reads to the *Hv* reference sequence was mainly restricted to gene regions (HighConfidence genes, LowConfidence genes and exome capture targets, Table 1), as genes appear to be highly conserved between both species. Thus, overall higher success of *Hb* sequence read mapping in exome capture can be attributed to the stronger enrichment for genic regions if compared to GBS.

Exome capture-enriched (re)-sequencing and GBS significantly shorten the time-consuming steps that have been involved to locate and map introgressions from *Hb*. In brief, the rough localization of an *Hb* introgression has typically been carried out using a combination of genomic *in situ* hybridization and fluorescent *in situ* hybridization steps (Pickering *et al.*, 2000). Subsequently, molecular markers had to be found and distributed across the *Hb* introgression. The development of new markers without access to genomic or expressed sequence information of the genotypes of interest was generally time consuming, and transferability of existing non-genic *Hv* markers to Hv^b was limited due to the high degree of sequence divergence between *Hv* and *Hb*. For instance, barley sequence tagged site (STS) markers were found to be unreliable in Hv^b lines. They either failed completely on any *Hb* genotype or they preferentially amplified *Hv* alleles in heterozygotes (Pickering and Johnston, 2005). Johnston *et al.* (2009) developed EST-based markers using the HarvEST database (Wanamaker *et al.*, 2006) to map Hv^b lines. This revealed 92 markers of an original set of 216 amplicons, giving a success rate of <50%. When the 1536-SNP barley BOPA1 set (Close *et al.*, 2009) was applied to localize introgressions of *H. vulgare* ssp. *spontaneum* (wild barley, primary gene pool) in *Hv*, only ~55% of all markers gave useful and polymorphic genotypes (Schmalenbach *et al.* 2011). Considering this low success rate and the high diversity between *Hv* and *Hb* (secondary gene pool), such primer-based, pre-designed marker assays do not seem promising to analyse Hv^b lines. Mascher *et al.* (2013b) reported efficient SNP calling from exome capture data obtained for *Hb* accessions, but transferring existing *Hv*/*Hb* markers or SNP information to new Hv^b lines is difficult, owing to the highly polymorphic nature of the *Hb* genepool.

In contrast, in this study, GBS as well as exome capture based approaches proved useful to precisely delineate *Hb* introgressions without the need for initial screenings. A large set of markers/SNP information and a precise allocation of introgressed segments both were obtained in the same experiment.

By applying exome capture and GBS a dense map of the 2HL Hv^b introgression could be generated. This map, however, was substantially condensed (nearly 4-fold reduction in interspecific recombination frequency of *Hv* and *Hb*) compared to the corresponding region of the barley reference map (Mascher *et al.*, 2013a). Such suppression of recombination frequency was previously observed for introgressed segments from *Hb* (Johnston *et al.*, 2009, 2013; Ruge *et al.*, 2003; Ruge-Wehling *et al.*, 2006), which magnifies the need for inexpensive and easy to handle molecular markers. Large numbers of markers will not compensate for the lack of recombination. However, the possibility to allocate them to the physical/genetic and sequence framework of barley (Ariyadasa *et al.*, 2014; IBSC *et al.*, 2012; Mascher *et al.*, 2013a) helps to more robustly estimate physical distances represented by recombination events which will also provide an estimate of the involved genes of the respective genomic segments.

The number of suitable markers was reported to be one of the major impediment when using CWR (Hajjar and Hodgkin, 2007) in plant breeding. The results of this study strongly suggest that GBS as well as exome capture are two promising tools that will be applicable to other crop-wild introgression systems, as well.

Accuracy and ease-of-use are not the only considerations when selecting a genotyping strategy targeting a specific research aim. Among other factors, the cost in terms of both money and time required by different approaches often play a pivotal role in choosing a method. In this study, the two approaches were compared according to their cost and time requirements (Figure 4 and Table 2). Both methods can be considered as being cost efficient. A caveat for the multiplex SNP assay is that the manufacturer set a minimum order size of 480 genotyping reactions, making this approach only efficient when 480 or more

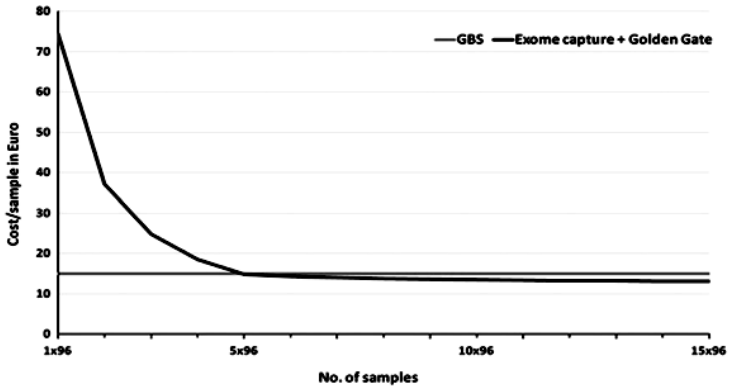

Figure 4 Cost comparison. Per-sample costs for sequencing and genotyping by GBS or exome capture/Golden Gate are compared across different numbers of analysed samples (96-plex format). Costs were calculated based on the experiences during the present study. However, costs may vary considerably due to different manufacturers, reagents, scales, and laboratories. Labour costs have not been considered.

Table 2 Time flowchart of exome capture (EC), Golden Gate assay (GGA) and GBS experiments

	Method		
Exp. step	EC (days)	GGA/96 samples (days)	GBS/96 samples (days)
Library preparation	14*		6
Sequencing	11		7
Analysis	7		7
GGA design		20	
GGA manufacturing		(30–60)**	
GGA genotyping		3	
GGA analysis		1	
Total	56	20	

*Assay assumed to be already in place ** not included.

samples are analysed (Figure 4). Furthermore, costs of the initial exome capture experiment for marker discovery are fixed, favouring a larger amount of subsequently analysed samples. In contrast, per-sample costs of GBS do not change as the number of samples analysed increases. Both strategies rely on NGS and may thus become even cheaper in the near future as sequencing costs are expected to continue to decrease. However, it needs to be considered that to apply both methods, sufficient computational power and bioinformatics knowledge is needed to analyse these fairly large amounts of NGS data. The latter might be a disadvantage for the presented methods as some laboratories or institutes might be lacking the required infrastructure, a limitation, however, that might be easily circumvented by establishing scientific collaborations. GBS was found to be the more flexible approach compared to exome capture/multiplex SNP assay if considering the time required carrying out the experiments (Table 2). This is mainly due to the initial exome capture experiment as well as the time-consuming steps of SNP marker selection for the SNP assay and its manufacturing by the company. However, once the SNP assay is in place, genotyping of 96 samples takes only 2–3 days.

In conclusion, both GBS and exome capture-enriched (re)-sequencing in combination with a multiplexed marker assay are powerful and convenient methods for the precise localization and genotyping of the *Hb* introgression segments in Hv^b lines. This will ultimately pave the way to incorporating advantageous *Hb* traits into elite *Hv* cultivars. Furthermore, we suggest that both methods are highly promising to be successfully implemented in other crop-wild introgression systems as well. An attractive immediate application of the two approaches outlined here would be the characterization of a larger set of Hv^b lines (i.e. Johnston *et al.*, 2009). Such an experiment may further validate putative introgressions and exactly delimit introgression intervals. Analogous to the vast collection of nearly-isogenic barley mutant lines (Druka *et al.*, 2011), a comprehensive, well-characterized catalogue of Hv^b lines would be an excellent community resource and may pave the way for the wide-spread utilization of Hv^b lines in applied breeding and basic research.

Experimental procedures

Plant material and DNA extraction

The plant material of the 2HL introgression (Hv^b) line originated from a single diploid resistant F_2 recombinant plant, which had been obtained from an interspecific tetraploid *Hordeum vulgare* L. *(Hv)* cv. 'Borwina' × *Hordeum bulbosum* L. (*Hb*, accession 'A42') hybrid (Szigat and Szigat, 1991) as described previously (Ruge-Wehling *et al.*, 2006). The resulting F_7 population that was subject to GBS and the Golden Gate assay, segregated for the introgressed *Hb* segment, which was mapped distally to the long arm of chromosome 2H (Ruge-Wehling *et al.*, 2006). To obtain this population, a single plant heterozygous for the *Hb* introgression was self-pollinated. The 2HL Hv^b plant targeted in the exome capture, was a single plant of the ancestral F_6 population, homozygous for the *Hb* introgression. Furthermore, a diploid plant of cultivar 'Borwina' and a tetraploid clonal progeny plant of the 'A42' donor line were used for exome capture and were included in the GBS experiments to capture information on donor genotype specific DNA polymorphisms.

Two additional Hv^b plants known to carry *Hb* introgressions on barley chromosomes 6HS (Ruge *et al.*, 2003) and 5HL (unpublished

results), respectively, were analysed by exome capture. The 6HS Hv^b genotype (F_5) originated from the same hybridization event that gave rise to the above mentioned 2HL Hv^b line. The 5HL Hv^b genotype (F_2) was kindly provided by Dr. U. Walther (BAZ, Aschersleben). It was developed by crossing a tetraploid *Hv* (cv. 'Vogelsanger Gold') with an unknown tetraploid *Hb* parent. Both Hv^b plants (5HL and 6HS) selected for exome capture (re)-sequencing were confirmed before to be homozygous for the *Hb* introgression interval using RFLP, STS, and SSR markers. A diploid plant of cultivar 'Vogelsanger Gold' was (re)-sequenced by exome capture to replace the original donor genotype of the initial hybrid for subsequent variant detection. As the *Hb* donor line of the 5HL introgression was unknown, the 'A42' *Hb* accession was used as a surrogate for the original *Hb* parent.

DNA extraction was performed as described previously (Stein *et al.*, 2001).

Exome capture and sequencing

Construction of exome capture libraries, and sequencing followed previously established procedures (Mascher *et al.*, 2013b). In brief, Illumina TruSeq Paired End libraries (Illumina Part # 15026486) were prepared according to the manufacturer using DNA Adapter Indexes 4, 5, 6 and 12 (Illumina, Inc., San Diego, CA). Genomic DNA (1 μg) was fragmented (size range: 200–300 bp) using Covaris microTUBES and the Covaris S220 Instrument (175W ultrasonic power, 10% duty factor, 200 cycles per burst, 100 s treatment time, Covaris, Inc., Woburn, MA). Adapter ligated DNA products were isolated (size range: 320–420 bp) by excision from an SYBR-Gold stained agarose gel (Life Technologies GmbH, Invitrogen, Darmstadt, Germany) and purified using AmPure XP beads (Beckman Coulter GmbH, Krefeld, Germany). Correctly ligated DNA fragments were enriched using a pre-capture LM-PCR reaction (ligation-mediated PCR) and purified as described (Haun *et al.*, 2011; Mascher *et al.*, 2013b). DNA was quantified with a Qubit 2.0 fluorometer (Life Technologies GmbH, Invitrogen) and analysed using an Agilent 2100 Bioanalyser (Agilent Technology, Santa Clara, Part# 5067-1506) on a DNA 7500 chip (between 250 and 500 bp).

Equimolar amounts of up to four amplified libraries containing different indices were pooled for hybridization. The library capture, liquid array processing, quantification, and sequencing on the Illumina HiSeq2000 device were as described (Mascher *et al.*, 2013a). Equimolar amounts of up to five pooled libraries, each containing different mixes of indices, were pooled for sequencing.

Selection of SNPs for the Golden Gate assay

For the 2HL Hv^b line, a Veracode Golden Gate assay (Illumina, Inc.) of 96 SNP markers was designed based on SNP information gained by exome captured (re)-sequencing of the 2HL Hv^b genotype and the donor genotypes 'Borwina' (*Hv*) and 'A42' (*Hb*), respectively. To improve the selection of high quality SNPs and to avoid false-positive SNP calls, single nucleotide variants between *Hv* and *Hb* with samtools SNP call score <200 and which were located within 200 bp of the end of a reference sequence contig were excluded from further evaluation. Furthermore, SNPs were only selected if they were supported by a minimum of fivefold sequence read coverage in the SNP itself, plus the flanking sequence (50–60 bp) of the (re)-sequenced Hv^b as well as the two donor genotypes.

The selected SNPs along with 50–60 bp flanking sequence to either side of the SNP position and harbouring at most one additional SNP between in each of the flanking sides were provided to the company Illumina Inc., for a pre-assessment using the Array Design Tool (ADT) (Illumina Inc.). ADT provides a feasibility rank score for each SNP ranging from 0 to 1, giving a benchmark about how likely a particular SNP will convert into a successful marker assay. According to recommendations by Illumina Inc., only SNPs with preliminary design evaluation rank scores of 0.6 or higher were selected for the custom assay design (Table S2).

Illumina Golden Gate assay-based genotyping

Genotyping was carried out in a 96-well, 96 SNP format (Fan *et al.*, 2003; Shen *et al.*, 2005) using the Illumina BeadXpress Reader Array platform (Illumina Inc.) according to manufacturer's instructions. Raw data processing and genotype calling were performed using the GenomeStudio software (Illumina Inc.). As recommended by Illumina, the minimum threshold for the 'GenCall' score (GenomeStudio, Illumina Inc.) was set to 0.25. In addition, all generated SNP clusters were visually inspected and manually re-clustered, using the GenomeStudio software (Illumina Inc.), if SNP clusters appeared to be incorrect.

GenCall (GC) scores and 50% GC scores were calculated with the GenomeStudio software (Illumina Inc.). The GC score is a quality value between 0 and 1 (0 = failed, 1 = excellent) to indicate the reliability for each genotype call, while the 50% GC score of a SNP represents the 50th percentile rank for all GenCall scores for that SNP. For visualization the obtained 50% GC score decimals were rounded to the nearest tenth.

GBS library construction and sequencing

Genomic DNA (200 ng) was cleaved over night at 37 °C using 14 units each of the restriction enzymes PstI-HF (CTGCAG, NEB Inc., Ipswich, UK) and MspI (CCGG, NEB Inc.), respectively. The reactions contained 1× NEB buffer 4 and 100 ng/μL BSA (NEB Inc.) in a total volume of 20 μL. After restriction digest, samples were incubated at 65 °C for 20 min to inactivate the restriction enzymes and were subsequently used for adapter ligation without purification. Illumina sequencing libraries were constructed as described with some modifications (Meyer and Kirchner, 2010). Briefly, for the reaction clean-up steps (solid phase reversible immobilization, SPRI) AMPure XP beads (Beckman Coulter GmbH, Krefeld, Germany) were replaced by MagNa beads (Thermo Scientific, Inc. Waltham, MA) for cost reduction (Rohland and Reich, 2012). The initial blunt-end repair step was omitted, as the adapter mix (P57_GBS) was provided with overhangs complementary to the ends obtained from the restriction digest. The P57_GBS adapter mix was prepared by annealing oligonucleotides P5_GBS_PstI (IS1_GBS_PstI: A*C*A*C*TCTTTCCCTACACGACGCTCTTCCGATCT*T*G*C*A and IS3_adapter.P5 + P7: A*G*A*T*CGGAA*G*A*G*C) and oligonucleotides P7_GBS_MspI (I2_adapter.P7: G*T*G*A*CTGGAGTTCAGACGTGTGCTCTTCCG*A*T*C*T and IS3_GBS_MspI: C*G*A*G*ATCGGAA*G*A*G*C; '*' indicates a PTO bond) as described for P5 and P7, respectively (Meyer and Kirchner, 2010). Adapter ligation and adapter fill-in were performed as published (Meyer and Kirchner, 2010). Following adapter fill-in, the DNA was purified and eluted in 20 μL EB (10 mM Tris-HCl, pH 8.0). For indexing PCR, 8 μL of the eluate was used as template DNA. The indexing PCR was performed in 50 μL volume with a final concentration of 1× Phusion HF buffer, 200 μM each dNTP, 200 nM primer IS4_indPCR.P5 (Meyer and Kirchner, 2010), 200 nM indexing primer (Table S3) and 0.02 U/μL Phusion Hot

Start Flex (NEB Inc., Ipswich, UK). After initial incubation at 98 °C for 30 s, the amplification was performed for 16 cycles (98 °C for 10 s, 60 °C for 30 s, 72 °C for 5 s) followed by a final extension (72 °C, 10 min). Products of adapter ligation, indexing PCR and primer positions are shown in Figure S2. The products were purified using SPRI and eluted in 25 μL EB. DNA was then quantified using the Quant-iT PicoGreen dsDNA assay kit (Life Technologies GmbH) and a Synergy HT microplate reader (BioTek, Bad Friedrichshall, Germany). Subsequently, the indexed samples were pooled in equimolar ratios. 500 ng pooled DNA was size-fractionated electrophoretically using a 2% agarose gel (Life Technologies GmbH, Invitrogen) and SYBR Gold (Life Technologies GmbH, Invitrogen) staining. The DNA (size between 150 and 600 bp) was recovered by excision and purified using a MinElute Spin column according to the manufacturer's instructions (Qiagen, Hilden, Germany). The GBS library was eluted from the column in 20 μL EB and analysed electrophoretically with an Agilent 2100 Bioanalyser (Agilent Technology, Santa Clara) using the Agilent High Sensitivity DNA kit (Part# 5067-4626).

Finally, the library was quantified using qPCR essentially as described previously (Mascher *et al.*, 2013b). The concentration was determined based on a standard curve and the average size of the GBS library. The sample was diluted to 10 nM for cluster formation on an Illumina cBot (Illumina, Inc.). In this study, 200 μL custom sequencing primer (Read1_GBS_PstI: CTTTCCCTACAC GACGCTCTTCCGATCTTGCAG; 500 nM in HT1 buffer from Illumina Inc.) were provided per lane for processing on the cBot (Illumina Inc., program: SR_Amp_Lin_Block_TubeStripHyb.v.8.0). Cluster formation and 1 × 100 bp single-end sequencing-by-synthesis using Illumina's HiSeq2000 instrument were performed according to protocols provided by the manufacturer (Illumina Inc.).

NGS read mapping and SNP calling

Sequence data were de-multiplexed using the CASAVA pipeline 1.8 (Illumina, Inc.). Exome capture read data for 'Borwina', 'Vogelsanger Gold' and the *Hb* accession 'A42' were previously published (Mascher *et al.*, 2013b) (accession number PRJEB1810) and were analysed together with sequence data generated in the present study.

Adapter trimming of GBS sequence reads was performed with cutadapt (Martin, 2011) and reads shorter than 30 bp after adapter removal were discarded. Trimmed GBS and exome capture reads were mapped against the barley reference (IBSC *et al.*, 2012) with BWA version 0.6.2 (Li and Durbin, 2009) (commands aln and samse, GBS reads or sampe, exome capture reads). The BWA command aln was called with the parameter '-q 15' for quality trimming, otherwise default parameters were used. Duplicate reads were removed from the exome sequencing mapping files with samtools rmdup. The read mapping files of three GBS replicates of 'A42' and 'Borwina' were merged using samtools merge (Li *et al.*, 2009) before analysis. SNP calling was performed with the SAMtools pipeline (version 0.1.18) using the command samtools aln and bcftools with default parameters (Li *et al.*, 2009). The additional parameter '-D' was used for samtools mpileup to obtain per-sample read depth.

For GBS genotyping, genotype calls were filtered with a custom AWK script and the R statistical environment (R Core Team, 2012), selecting only SNPs matching the following criteria: (i) read depth at least 10 (5, 20) and genotype quality at least 10 (5, 20) for homozygous and heterozygous calls, (ii) at most 5% missing genotype calls (genotype calls with insufficient coverage or quality were set to missing) (iii) a 1 : 2 : 1 segregation ratio in the segregating population.

Exome capture SNP calls were discarded if their SAMtools SNP quality score was below 40, the sequence read coverage was smaller than fivefold or they were positioned on a reference assembly contig shorter than 2000 bp. SNP frequency as determined from GBS and from exome capture data at a minimum 10-fold coverage was visualized along the integrated physical/genetic map of barley (IBSC *et al.*, 2012) using R (R Core Team, 2012). GBS SNP frequency was analysed for minimum SNP call scores of 200. Read depth statistics were performed using SAMtools depth (Li *et al.*, 2009). Only properly paired non-duplicated reads were considered for coverage calculation in the exome sequencing data. The overlap with different genomic intervals such as gene models (IBSC *et al.*, 2012) and exome capture targets (Mascher *et al.*, 2013b) was analysed using bedtools intersect (Quinlan and Hall, 2010) and samtools flagstat (Li *et al.*, 2009).

Primer design and Sanger sequencing

PCR primers were designed using the online software Batch primer 3 (Untergasser *et al.*, 2012) and were obtained from Eurofins (Ebersberg, Germany) (Table S4). PCR amplification of genomic DNA was performed in a final volume of 20 μL, containing 20 ng genomic DNA, 0.1 μM forward and reverse primers, 0.5 U *Hot-star* polymerase (Qiagen), 1× PCR reaction buffer and 0.1 mM dNTPs (Fermentas, Fisher scientific, Schwerte, Germany). PCR amplification was performed with 15 min at 95 °C for initial denaturation, followed by 10 cycles of 30 s at 95 °C, 30 s annealing (60–55 °C, −0.5 °C per cycle), 1 min at 72 °C for extension, and 35 cycles of 30 s at 95 °C, 30 s at 55 °C for extension, 1 min at 72 °C, and a final extension for 7 min at 72 °C.

Sanger cycle-sequencing of PCR products was performed following the manufacturer's instructions. In brief, PCR products were purified using the NucleoFast 96 PCR Kit (Macherey-Nagel, Dueren, Germany). The concentration of purified amplicons was subsequently analysed on an agarose gel (Life Technologies GmbH, Invitrogen). Normalized amplicons (10 ng per 100 bp of template fragment) were used for cycle-sequencing (BigDye®Terminator v3.1, Applied Biosystems, Darmstadt, Germany) by using ABI-3730xl technology. Sequence analysis was performed using Sequencher 4.7 (Gene Codes, Ann Arbor).

Map construction and comparison to barley

Genetic maps based on GBS and/or Golden Gate markers were constructed with JoinMap v4.0 (Van Ooijen, 2006) using the Kosambi function. The resulting marker order of each map was compared to the genetic order by Mascher *et al.* (2013a).

Acknowledgements

We gratefully acknowledge the excellent technical support by Ines Walde, Manuela Knauft and Sandra Driesslein (IPK). We thank Doreen Stengel for raw data submission. The work was financially supported as part of the collaborative project 'TRANSBULB' by a grant (0315966) from the German Federal Ministry of Education and Research (BMBF) to BRW and NS.

References

Akano, A.O., Dixon, A.G.O., Mba, C., Barrera, E. and Fregene, M. (2002) Genetic mapping of a dominant gene conferring resistance to cassava mosaic disease. *Theor. Appl. Genet.* **105**, 521–525.

Ariyadasa, R., Mascher, M., Nussbaumer, T., Schulte, D., Frenkel, Z., Poursarebani, N., Zhou, R.N., 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.X., 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 Physiol.* **164**, 412–423.

Asan, Xu.Y., Jiang, H., Tyler-Smith, C., Xue, Y.L., Jiang, T., Wang, J.W., Wu, M.Z., Liu, X., Tian, G., Wang, J., Wang, J., Yang, H.M. and Zhang, X.Q. (2011) Comprehensive comparison of three commercial human whole-exome capture platforms. *Genome Biol.* **12**, R95.

von Bothmer, R., Jacobsen, N., Baden, C., Jorgensen, R.B. and Linde-Laursen, I. (1995) *An Ecogeographical Study of the Genus Hordeum*. Rome: International Plant Genetic Resources Institute (IPGRI).

Close, T.J., Bhat, P.R., Lonardi, S., Wu, Y.H., Rostoks, N., Ramsay, L., Druka, A., Stein, N., Svensson, J.T., Wanamaker, S., Bozdag, S., Roose, M.L., Moscou, M.J., Chao, S.A.M., 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.

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. *Nat. Genet.* **44**, 1388–1392.

Davey, J.W., Hohenlohe, P.A., Etter, P.D., Boone, J.Q., Catchen, J.M. and Blaxter, M.L. (2011) Genome-wide genetic marker discovery and genotyping using next-generation sequencing. *Nat. Rev. Genet.* **12**, 499–510.

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 Physiol.* **155**, 617–627.

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**, e19379.

Fan, J.B., Oliphant, A., Shen, R., Kermani, B.G., Garcia, F., Gunderson, K.L., Hansen, M., Steemers, F., Butler, S.L., Deloukas, P., Galver, L., Hunt, S., McBride, C., Bibikova, M., Rubano, T., Chen, J., Wickham, E., Doucet, D., Chang, W., Campbell, D., Zhang, B., Kruglyak, S., Bentley, D., Haas, J., Rigault, P., Zhou, L., Stuelpnagel, J. and Chee, M.S. (2003) Highly parallel SNP genotyping. *Cold Spring Harb. Symp. Quant. Biol.* **68**, 69–78.

Forster, B.P., Phillips, M.S., Miller, T.E., Baird, E. and Powell, W. (1990) Chromosome location of genes-controlling tolerance to salt (Nacl) and vigor in *Hordeum-vulgare* and *Hordeum-chilense*. *Heredity*, **65**, 99–107.

Hajjar, R. and Hodgkin, T. (2007) The use of wild relatives in crop improvement: a survey of developments over the last 20 years. *Euphytica*, **156**, 1–13.

Haun, W.J., Hyten, D.L., Xu, W.W., Gerhardt, D.J., Albert, T.J., Richmond, T., Jeddeloh, J.A., Jia, G.F., Springer, N.M., Vance, C.P. and Stupar, R.M. (2011) The composition and origins of genomic variation among individuals of the soybean reference cultivar Williams 82. *Plant Physiol.* **155**, 645–655.

IBSC, Mayer, K.F.X., Waugh, R., Langridge, P., Close, T.J., Wise, R.P., Graner, A., Matsumoto, T., Sato, K., Schulman, A., Muehlbauer, G.J., Stein, N., Ariyadasa, R., Schulte, D., Poursarebani, N., Zhou, R.N., Steuernagel, B., Mascher, M., Scholz, U., Shi, B.J., Langridge, P., Madishetty, K., Svensson, J.T., Bhat, P., Moscou, M., Resnik, J., Close, T.J., Muehlbauer, G.J., Hedley, P., Liu, H., Morris, J., Waugh, R., Frenkel, Z., Korol, A., Berges, H., Graner, A., Stein, N., Steuernagel, B., Taudien, S., Groth, M., Felder, M., Platzer, M., Brown, J.W.S., Schulman, A., Platzer, M., Fincher, G.B., Muehlbauer, G.J., Sato, K., Taudien, S., Sampath, D., Swarbreck, D., Scalabrin, S., Zuccolo, A., Vendramin, V., Morgante, M., Mayer, K.F.X., Schulman, A. and Conso, I.B.G.S. (2012) A physical, genetic and functional sequence assembly of the barley genome. *Nature*, **491**, 711–716.

Johnston, P.A., Timmerman-Vaughan, G.M., Farnden, K.J.F. and Pickering, R. (2009) Marker development and characterisation of *Hordeum bulbosum* introgression lines: a resource for barley improvement. *Theor. Appl. Genet.* **118**, 1429–1437.

Johnston, P.A., Niks, R.E., Meiyalaghan, V., Blanchet, E. and Pickering, R. (2013) Rph22: mapping of a novel leaf rust resistance gene introgressed from the non-host *Hordeum bulbosum* L. into cultivated barley (*Hordeum vulgare* L.). *Theor. Appl. Genet.* **126**, 1613–1625.

Kalladan, R., Worch, S., Rolletschek, H., Harshavardhan, V.T., Kuntze, L., Seiler, C., Sreenivasulu, N. and Roder, M.S. (2013) Identification of quantitative trait loci contributing to yield and seed quality parameters under terminal drought in barley advanced backcross lines. *Mol. Breed.* **32**, 71–90.

Kasha, K.J. and Sadasiva, R.S. (1971) Genome relationship between *Hordeum-Vulgare* L and *H-Bulbosum* L. *Chromosoma*, **35**, 264–287.

Lakew, B., Henry, R.J., Ceccarelli, S., Grando, S., Eglinton, J. and Baum, M. (2013) Genetic analysis and phenotypic associations for drought tolerance in *Hordeum spontaneum* introgression lines using SSR and SNP markers. *Euphytica*, **189**, 9–29.

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

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R. and Proc, G.P.D. (2009) The sequence alignment/map format and SAMtools. *Bioinformatics*, **25**, 2078–2079.

Mallikarjuna, N. (1999) Ovule and embryo culture to obtain hybrids from interspecific incompatible pollinations in chickpea. *Euphytica*, **110**, 1–6.

Martin, M. (2011) Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet. J.* **17**, 10–12.

Martin, A., Alvarez, J.B., Martin, L.M., Barro, F. and Ballesteros, J. (1999) The development of tritordeum: a novel cereal for food processing. *J. Cereal Sci.* **30**, 85–95.

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., Scholz, U., Poland, J.A., Stein, N. and Waugh, R. (2013a) Anchoring and ordering NGS contig assemblies by population sequencing (POPSEQ). *Plant J.* **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. (2013b) Barley whole exome capture: a tool for genomic research in the genus Hordeum and beyond. *Plant J.* **76**, 494–505.

Mascher, M., Shuangye, W., Amand, P.S., Stein, N. and Poland, J. (2013c) Application of genotyping-by-sequencing on semiconductor sequencing platforms: a comparison of genetic and reference-based marker ordering in barley. *PLoS ONE*, **8**, e76925.

Meyer, M. and Kirchner, M. (2010) Illumina Sequencing Library Preparation for Highly Multiplexed Target Capture and Sequencing. *Cold Spring Harb. Protoc.* doi: 10.1101/pdb.prot5448.

Nevo, E. (1992) Origin, evolution, population genetics and resources for breeding of wild barley, *Hordeum spontaneum*, in the fertile crescent. In: *Genetics, Biochemistry, Molecular Biology and Biotechnology* (Shewry, P.R., ed.), pp. 19–43. Oxford: C.A.B. International, The Alden Press.

Pickering, R.A. (1983) The location of a gene for incompatibility between *Hordeum-Vulgare*-L and *H-Bulbosum* L. *Heredity*, **51**, 455–459.

Pickering, R.A. (1984) The influence of genotype and environment on chromosome elimination in crosses between *Hordeum-Vulgare*-L × *Hordeum-Bulbosum* L. *Plant Sci. Lett.* **34**, 153–164.

Pickering, R.A. and Johnston, P.A. (2005) Recent progress in barley improvement using wild species of Hordeum. *Cytogenet. Genome Res.* **109**, 344–349.

Pickering, R.A., Rennie, W.F. and Cromey, M.G. (1987) Disease resistant material available from the wide hybridization programme at DSIR. *Barley Newsl.* **31**, 248–250.

Pickering, R.A., Malyshev, S., Kunzel, G., Johnston, P.A., Korzun, V., Menke, M. and Schubert, I. (2000) Locating introgressions of *Hordeum bulbosum* chromatin within the *H-vulgare* genome. *Theor. Appl. Genet.* **100**, 27–31.

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.

Prescott-Allen, C. and Prescott-Allen, R. (1986) *The First Resource: Wild Species in the North American Economy*. New Haven: Yale University.

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

R Core Team (2012) *Language and Environment for Statistical Computing*. Austria: R Foundation for Statistical Computing. ISBN3-900051-900007-900050.

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

Ruge, B., Linz, A., Pickering, R., Proeseler, G., Greif, P. and Wehling, P. (2003) Mapping of Rym14(Hb), a gene introgressed from *Hordeum bulbosum* and conferring resistance to BaMMV and BaYMV in barley. *Theor. Appl. Genet.* **107**, 965–971.

Ruge-Wehling, B., Linz, A., Habekuss, A. and Wehling, P. (2006) Mapping of Rym16(Hb), the second soil-borne virus-resistance gene introgressed from *Hordeum bulbosum*. *Theor. Appl. Genet.* **113**, 867–873.

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

Schmalenbach, I., March, T.J., Bringezu, T., Waugh, R. and Pillen, K. (2011) High-Resolution Genotyping of Wild Barley Introgression Lines and Fine-Mapping of the Threshability Locus thresh-1 Using the Illumina GoldenGate Assay. *G3-Genes Genomes Genetics*, **1**, 187–196.

Shen, R., Fan, J.B., Campbell, D., Chang, W.H., Chen, J., Doucet, D., Yeakley, J., Bibikova, M., Garcia, E.W., McBride, C., Steemers, F., Garcia, F., Kermani, B.G., Gunderson, K. and Oliphant, A. (2005) High-throughput SNP genotyping on universal bead arrays. *Mutat. Res.* **573**, 70–82.

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 Breed.* **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. *Theor. Appl. Genet.* **114**, 823–839.

Sulonen, A.M., Ellonen, P., Almusa, H., Lepisto, M., Eldfors, S., Hannula, S., Miettinen, T., Tyynismaa, H., Salo, P., Heckman, C., Joensuu, H., Raivio, T., Suomalainen, A. and Saarela, J. (2011) Comparison of solution-based exome capture methods for next generation sequencing. *Genome Biol.* **12**, R94.

Szigat, G. and Szigat, G. (1991) Amphidiploid hybrids between *Hordeum vulgare* and *H. bulbosum*—basis for the development of new initial material for winter barley breeding. In *Vortr Pflanzenzüchtg* (Winkel, A., ed.), pp. 34–39. Wageningen: EUCARPIA.

Tanksley, S.D. and McCouch, S.R. (1997) Seed banks and molecular maps: unlocking genetic potential from the wild. *Science*, **277**, 1063–1066.

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 Res.* **40**, e115.

Van Ooijen, J.W. (2006) *JoinMap® 4, Software for the Calculation of Genetic Linkage Maps in Experimental Populations*. Wageningen, Netherlands: Kyazma BV.

Walther, U., Rapke, H., Proeseler, G. and Szigat, G. (2000) *Hordeum bulbosum* - a new source of disease resistance - transfer of resistance to leaf rust and mosaic viruses from *H-bulbosum* into winter barley. *Plant Breed.* **119**, 215–218.

Wanamaker, S., Close, T., Roose, M. and Lyon, M. (2006) *HarvEST: Barley EST database version 1.44*.

Xu, J. and Snape, J.W. (1988) The cytology of hybrids between *Hordeum vulgare* and *H. bulbosum* revisited. *Genome Biol.* **30**, 486–494.

Supporting information

Additional Supporting information may be found in the online version of this article:

Figure S1 Histogram of the 50% GenCall (GC) Score of the 96 Golden Gate SNP markers, obtained by genotyping in the 2HL Hv^b mapping population.

Figure S2 GBS adapter, primer for indexing PCR and Illumina sequencing.

Table S1 List of BAC contigs.

Table S2 List of Golden Gate SNP markers.

Table S3 Table of GBS indices.

Table S4 Primer sequences for marker verification.

3. *Bulbosum* to go: a toolbox to utilize *Hordeum vulgare* / *bulbosum* introgressions for breeding and beyond

Neele Wendler[1], Martin Mascher[1], Axel Himmelbach[1],Paul Johnston[2],Richard Pickering[2]and Nils Stein[1]

[1]Leibniz Institute of Plant Genetics and Crop Plant Research (IPK), Corrensstr. 3, 06466 Seeland (OT) Gatersleben, Germany
[2]New Zealand Institute for Plant and Food Research Limited, Private Bag 4704, Christchurch 8140, New Zealand

*Corresponding author: Nils Stein (Tel +49 (0)39482 5522; fax +49 (0)39482 5595; email stein@ipk-gatersleben.de)

Molecular Plant (2015), doi: 10.1016/j.molp.2015.05.004

Abstract

Hordeum bulbosum L. a wild relative of barley (*Hordeum vulgare* L.) has been considered as a valuable source of genetic diversity for barley improvement. Since the 1990's a considerable number of barley/*H. bulbosum* introgression lines (IL)s has been generated, with segments introgressed from *H. bulbosum*,harboringa diverse set of desirable traits. However, the efficient utilization of these ILs has been hampered, largely due to the lack of suitable molecular tools for theirgenetic characterization and highly reduced interspecific recombination frequencies in the region of the introgression. In the present study, we utilizedgenotyping-by-sequencingfor the detailed molecular characterization of 145 ILs.Genotypic information allows determining the genetic diversity within the set of ILs and a strategy was outlined to tackle the obstacle of reduced recombination frequencies. Furthermore, wecompiled exome capture re-sequencing information of barleyand *H. bulbosum*and designed an integrated barley*/H.bulbosum* sequence resource with polymorphism information on interspecific and intraspecific sequence variations ofboth species. The integrated sequence will bevaluable for marker development in barley/H. *bulbosum* ILs derived from any barley and *H. bulbosum* donors. This study provides the tools for the wide-spread utilization of barley/*H. bulbosum* ILs in applied barley breeding andacademic research.

Molecular Plant
Research Article

CellPress
PARTNER JOURNAL

Bulbosum to Go: A Toolbox to Utilize *Hordeum vulgare/bulbosum* Introgressions for Breeding and Beyond

Neele Wendler[1], Martin Mascher[1], Axel Himmelbach[1], Paul Johnston[2], Richard Pickering[2] and Nils Stein[1,*]

[1]Leibniz Institute of Plant Genetics and Crop Plant Research (IPK), Corrensstraße 3, 06466 Seeland (OT) Gatersleben, Germany

[2]New Zealand Institute for Plant and Food Research Limited, Private Bag 4704, Christchurch 8140, New Zealand

*Correspondence: Nils Stein (stein@ipk-gatersleben.de)

http://dx.doi.org/10.1016/j.molp.2015.05.004

ABSTRACT

***Hordeum bulbosum* L., a wild relative of barley (*Hordeum vulgare* L.), has been considered as a valuable source of genetic diversity for barley improvement. Since the 1990s, a considerable number of barley/*H. bulbosum* introgression lines (IL)s has been generated, with segments introgressed from *H. bulbosum* harboring a diverse set of desirable traits. However, the efficient utilization of these ILs has been hampered, largely due to the lack of suitable molecular tools for their genetic characterization and highly reduced interspecific recombination frequencies in the region of the introgression. In the present study, we utilized genotyping-by-sequencing for the detailed molecular characterization of 145 ILs. Genotypic information allows the genetic diversity within the set of ILs to be determined and a strategy was outlined to tackle the obstacle of reduced recombination frequencies. Furthermore, we compiled exome capture re-sequencing information of barley and *H. bulbosum* and designed an integrated barley/*H. bulbosum* sequence resource with polymorphism information on interspecific and intraspecific sequence variations of both species. The integrated sequence will be valuable for marker development in barley/*H. bulbosum* ILs derived from any barley and *H. bulbosum* donors. This study provides the tools for the widespread utilization of barley/*H. bulbosum* ILs in applied barley breeding and academic research.**

Key words: *Hordeum vulgare*, *Hordeum bulbosum*, introgression line, genotyping-by-sequencing (GBS), exome capture, NGS

Wendler N., Mascher M., Himmelbach A., Johnston P., Pickering R., and Stein N. (2015). *Bulbosum* to Go: A Toolbox to Utilize *Hordeum vulgare/bulbosum* Introgressions for Breeding and Beyond. Mol. Plant. **8**, 1507–1519.

INTRODUCTION

Barley (*Hordeum vulgare* L.) is one of mankind's oldest cultivated crops. The domestication of this cereal species started about 10 000 years ago with the beginning of agriculture (Lev-Yadun et al., 2000; Salamini et al., 2002). Today, barley is the fourth most important cereal in the world in terms of area grown as well as tons harvested (http://faostat.fao.org) and it has a big impact on the modern human diet in many developing as well as developed countries (Blake et al., 2011). These millennia of domestication and improvement by early farmers as well as extensive breeding in the last centuries fostered significant morphological changes facilitating modern cultivation of barley and leading to a tremendous increase in yield (Szigat and Pohler, 1982; Salamini et al., 2002). On the downside, however, this simultaneously narrowed the genetic diversity of highly yielding elite material, resulting in a restricted choice for new beneficial traits within the elite germplasm (Feuillet et al., 2008). Cultivation of barley and crops in general is constantly being challenged by a range of biotic and abiotic stresses that trigger the inevitable need for permanent genetic improvement. To tackle this problem, crop-wild relatives are an important resource for crop improvement globally to enrich the cultivated gene pools by incorporating beneficial alleles (Tanksley and McCouch, 1997; Hajjar and Hodgkin, 2007). In 1986, the benefit of crop-wild introgressions in North America has been estimated to account for more than US$ 340 million per year (Prescott-Allen and Prescott-Allen, 1986).

Published by the Molecular Plant Shanghai Editorial Office in association with Cell Press, an imprint of Elsevier Inc., on behalf of CSPB and IPPE, SIBS, CAS.

Hordeum bulbosum L. (*Hb*) is the only representative of the secondary gene pool of barley (von Bothmer et al., 1995) and it has been found to be a valuable source of genetic diversity for barley improvement, especially regarding resistance or tolerance to pathogens (Pickering et al., 1987; Jie and Snape, 1988; Szigat and Szigat, 1991; Walther et al., 2000; Ruge et al., 2003). The first report on introgression lines (IL)s between barley and *Hb* dates back to 1982 (Szigat and Pohler, 1982). However, the incorporation of *Hb* chromatin into barley chromosomes could be confirmed only after the development and use of molecular techniques such as *in situ* hybridization and Southern blotting (Xu and Kasha, 1992; Pickering et al., 1995). Since then, considerable numbers of ILs and substitution lines have been generated. The introgressed segments cover different terminal regions of the *H. vulgare* (*Hv*) genome and provide diversity for a set of traits including disease resistances as well as morphological characteristics (Pickering et al., 1987; Szigat and Szigat, 1991; Pickering et al., 1997; Walther et al., 2000; Ruge et al., 2003; Johnston et al., 2009). Considerable effort has been made in various studies for the genetic, cytogenetic, and phenotypic characterization of most of the generated ILs. However, the genetic analysis of *Hv*/*Hb* ILs has been difficult due to the lack of genomic sequence information of *Hb* as a basis for marker development. For instance, characterization with PCR-based marker assays derived from barley has often been inefficient due to a high rate of failed amplification in *Hb* and favored or exclusive amplification of barley alleles in heterozygotes or hybrids (Johnston, 2007). Laborious *in situ* hybridization techniques were required to pre-define the physical chromosomal location of introgressed *Hb* chromatin (Pickering et al., 2000), which subsequently had to be confirmed and addressed by molecular markers.

The transfer of advantageous traits from *Hb* into *Hv* is often accompanied by additional factors that are associated with the introgressed *Hb* segments. These factors can confer a yield penalty in adapted elite material or maybe other unfavorable characteristics (negative linkage drag). Marker-assisted reduction of the size of introgressed *Hb* segments is needed to reduce the chances of negative linkage drag. Therefore, accurate genetic characterization of *Hv*/*Hb* ILs and the availability of molecular markers for this material are a prerequisite to the utilization of *Hb* traits in barley breeding. Genotyping-by-sequencing (GBS) has been established as a cost-efficient, high-throughput genotyping method that combines complexity reduction and the strengths of next-generation sequencing (NGS) (Elshire et al., 2011; Poland et al., 2012). GBS was found to be an efficient tool for the genetic characterization of *Hv*/*Hb* ILs, enabling the precise and rapid detection of all introgressed *Hb* segments of an IL (Wendler et al., 2014). At the same time, it avoids the time-consuming steps of marker development across the previously identified putative region of introgressed *Hb* chromatin (Wendler et al., 2014). Here, we utilized GBS for the detailed characterization of a larger set of ILs. The genotyping data obtained can be used to determine the diversity between *Hb* segments of different ILs. In addition to the lack of *Hb* sequence information, fine mapping and cloning of *Hb* genes is still hampered due to strong reduction of interspecific recombination frequencies within the introgressed *Hb* segments. Based on information on the extent and diversity between *Hb* segments, we suggest a crossing strategy to overcome this obstacle toward a normal recombination frequency. Moreover, a barley exome capture assay has proven to be a useful tool for re-sequencing of *Hb* and subsequent marker development in *Hv*/*Hb* ILs (Mascher et al., 2013b; Wendler et al., 2014). *Hb* and *Hv* exome capture sequence data was used to design an integrated *Hv*/*Hb* sequence resource with defined *Hv*/*Hb* polymorphism information on interspecific and intraspecific sequence variations. This catalogue of precisely characterized ILs in combination with the integrated *Hv*/*Hb* sequence will be a valuable community resource to advance the widespread utilization of *Hv*/*Hb* ILs in applied barley breeding as well as basic research.

RESULTS

Detection of Introgressed *H. bulbosum* Segments by GBS

GBS was applied to 146 *Hv*/*Hb* ILs (Supplemental File 1), as well as the frequently used *Hv* hybrid donors "Vada", "Morex", "Emir", "Golden Promise" and the Hb clones "2920/4", "2032", "2929/1", "A17" (*Hb*) to detect and localize introgressed *Hb* segments and to identify sequence polymorphisms of the analyzed ILs.

The GBS sequence reads of individual samples were mapped to the barley reference for sequence variant detection. This facilitated comparison of sequence reads or recognized polymorphisms between samples based on their position within the reference sequence. For the ILs, GBS yielded on average 2.2 million sequence reads and 242 Mbp of sequence per sample (minimum, 0.2 million reads [27 Mbp]; maximum, 4.3 million reads [461 Mbp]) (Supplemental File 2). Differences in read numbers can arise from deviations during pooling of each sample prior to sequencing. GBS was performed based on single-end instead of paired-end sequencing mainly with respect to costs and data redundancy. With paired-end sequencing, the second read of a pair will be very closely linked to the first read of the pair (e.g. a few hundred base pairs apart). Due to this close linkage, any further SNPs that are identified in this second read of the pair will not add much information with regard to genome coverage. However, paired-end sequencing would almost double sequence costs compared with single-end sequencing. To ensure a high quality of detected sequence variants and thus a high chance of correct allele calling, single nucleotide polymorphisms (SNP)s were only considered for downstream analysis if they were supported by at least five sequence reads. To determine the regions of introgressed *Hb* chromatin in individual ILs, SNPs between ILs and their respective *Hv* and *Hb* parents were compared. An IL SNP was considered to be *Hb* specific if it showed different genotype calls between the IL and the *Hv* donor but identical calls in the IL and the *Hb* donor. These diagnostic *Hb*-specific SNPs should indicate the locations of *Hb* segments, assuming that sequence variations between an IL and its *Hv* donor is present exclusively in the region of the introgressed *Hb* chromatin.

It was known that ILs 125 and 155 have been out-crossed (accidently) or backcrossed (intentionally) to *Hv* cultivars "Emir"(sequenced) and "907-12" (not sequenced), respectively. IL 125 was supposed to be derived from a cross involving cultivar "Morex". Based on the observed SNP profiles of the

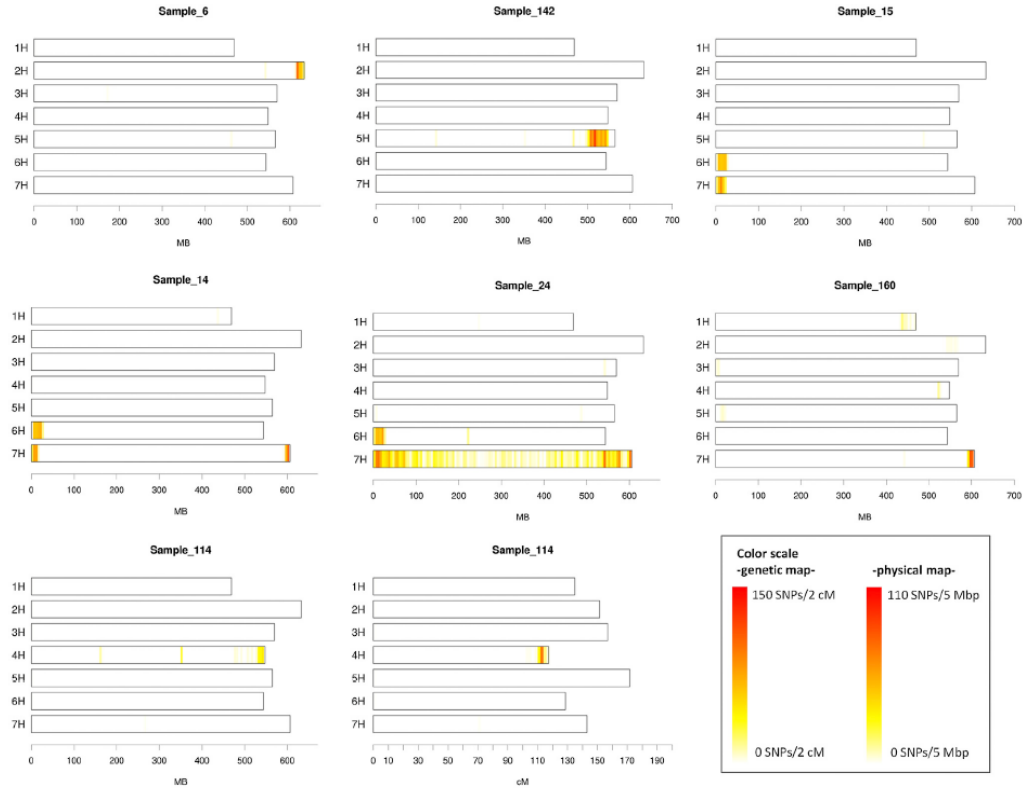

Figure 1. *Hb*-specific SNP Frequency Distribution in a Subset of 7ILs.
Hb-specific SNPs were discovered by genotyping-by-sequencing (GBS). The SNP frequency was visualized as a heat map with resolution of 5 Mbp or 2 cM bins, respectively, along the physical and genetic scale of the seven barley chromosomes (1H–7H). The physical and genetic positions of contigs were taken from the barley reference draft genome information (International Barley Sequencing Consortium, 2012; Mascher et al., 2013a). Relative SNP frequency was visualized in 1/110 (physical map) or 1/150 (genetic map) color steps from white (no detected SNPs) to red (maximum SNP frequency) using heat.colors with the R statistical environment (R Core Team, 2012).

IL and another cultivar "Emir", the latter could be easily identified as the out-crossing parent. *Hb*-specific SNP filtering was altered in these two cases, to prevent the detection of *Hv*/*Hv* SNPs instead of *Hv*/*Hb* polymorphisms. Thus, *Hb*-specific SNPs in these lines were accepted only if they had the same genotype calls in all *Hv* cultivars (IL 155) or *Hv* cultivars "Emir" and "Morex" (IL 125). This corrected for any "Emir"/"Morex" SNPs in IL 125. The data of IL 155 was substantially improved but some *Hv*/*Hv* SNPs still passed our filters (Supplemental Files 2 and 3).

Hb-specific SNPs were filtered allowing only homozygote variant calls. In the case of a heterozygote introgressed *Hb* segment, only heterozygous SNP calls were allowed in the IL, allowing only homozygotes in the *Hv* and *Hb* donors. In cases where the original *Hb* donor was not re-sequenced, GBS data of *Hb* clone "A17" (with the highest amount of raw sequence data) was used as a surrogate for the true hybrid parent. Supplemental File 3 contains the filtered SNPs, their associated genotypes, and their positions on the physical and genetic maps of barley (International Barley Sequencing Consortium, 2012; Mascher et al., 2013a). Supplemental File 4 contains FASTA sequences for each IL, GBS targets with a minimum five-fold sequence read coverage in the IL and the respective donor lines. The FASTA files are based on the barley reference assembly of cultivar "Morex" (International Barley Sequencing Consortium, 2012). Base pair positions with less than five sequence reads are masked as "N" and the filtered *Hb*-specific SNPs are given as IUPAC codes (IUPAC-IUB Commission on Biochemical Nomenclature (CBN), 1970). The frequency of the filtered *Hb*-specific SNPs for each IL were plotted along the barley physical (International Barley Sequencing Consortium, 2012) and genetic (Mascher et al., 2013a) reference maps in 5 Mbp or 2 cM bins, respectively (Figure 1 and Supplemental File 5). A list of SNP frequencies in 5 Mbp bins is given in Supplemental File 2.

Chromosome	IL IPK-ID	Approx. start (cM[a])	Approx. end (cM[a])	No of SNPs (cM[a])	Approx. start (Mbp[b])	Approx. end (Mbp[b])	No of SNPs (Mbp[b])	Estimated size (cM[a])	Estimated size (Mbp[b])
1HS	37	0.1	4.9	24	0.2	4	23	5	4
	64	0.1	4.9	24	0.2	4	23	5	4
	36	0.1	8.4	25	0.2	5.2	29	8	5
	90	0.1	8.6	68	0.2	7	75	9	7
	73	0.1	18.2	89	0.2	15.3	96	18	15
	18	0.2	0.2	6	0.2	0.7	9	0	1

Table 1. Approximate Genetic and Physical Positions of *Hb* Segments on Chromosome 1HS based on Genotyping-By-Sequencing Data.

[a]Mascher et al. (2013a).

[b]International Barley Sequencing Consortium (2012).

The *Hb*-specific SNP frequency plots allowed the detection of introgressed *Hb* segments in all but one of the 146 ILs analyzed (Supplemental File 1). The low number of SNPs may be a result of incorrect read mapping, sequencing errors, or similar technical artifacts. When only homozygous SNPs were considered, it was found that 1–3 SNPs per individual Mbp bin were frequently distributed along the genomes of all samples. Most of these SNPs could be found always at the same position in (almost) all samples, thus they were likely derived by incorrect mapping of reads and they can be easily recognized as being noise and not from true introgressions. Thus, up to three SNPs per bin in homozygous ILs were considered as background noise. As the genotyping error rate is higher for heterozygous positions, the threshold was raised to 10 SNPs per bin for ILs that were presumed to carry their introgressed segments in heterozygous state. The number of SNPs in single 5 Mbp bins ranged from >3 to 109. Detected polymorphic regions were considered to represent introgressed segments only if SNPs were present in at least two different GBS targets flanking different *Pst*I restriction sites within the barley genome. Polymorphisms derived from only a single GBS target may be due to a low fraction of cross-contamination introduced during the indexing PCR step of GBS library construction (Kircher et al., 2012). Due to the possibility of false read mapping and false SNP calling, regions with very few SNPs cannot be claimed with good confidence as representing true introgressions. Three introgressed segments were described based on the presence of not more than five SNPs in the physical and the genetic map (IL 64 3HS, 88 5HL, and 152 7H). Thus, these introgressions should be treated with caution.

Up to three *Hb* segments, including terminal and interstitial introgressions as well as a disomic substitution of chromosome 7H, were detected across the ILs (Figure 1; samples 6, 142, 15, 14, and 24). A questionable pattern was detected for *Hb* introgressions located on 4HL based on the physical reference map of barley (International Barley Sequencing Consortium, 2012), where a single bin of high SNP frequency was found at coordinate 350 Mbp, far from the remaining bins of these introgressed segments. However, when plotting the same ILs along the genetic map (Mascher et al., 2013a), this inconsistency was resolved and all high-frequency bins occurred consecutively at the telomere of chromosome 4HL (i.e. Figure 1, sample 114). Thus, this pattern on 4HL is most likely an artifact due to an error in ordering sequence data in the barley physical map. For 8 ILs (9, 28, 68, 94, 102, 104, 107, and 150) *Hb* chromatin could be detected if heterozygous SNPs were considered. The *Hb* segment of IL 150 appeared to be partly heterozygous and partly homozygous in a single plant. Furthermore, IL 94 showed a homozygous *Hb* segment on 6HS and a heterozygous *Hb* segment on 5HL. Some ILs derived from *Hv* cultivar "Emir" (16, 17, 25, 31, and 160) showed an identical distribution of introgressed chromatin on chromosomes 1HL, 2HL, 3HS, 4HL, and 5HS (ILs 17 and 31 showed this pattern on 1HL only). Comparison of the SNP haplotypes in these regions with *Hv* and *Hb* revealed that they had much higher identity with *Hv* (94%–100%) than *Hb* (0%–6%). Therefore, it seems likely that these segments are actually *Hv* derived; thus an out-crossed or non-pure "Emir" plant containing a mixed haplotype with segments originating from another *Hv* cultivar was most likely used as a donor for these lines. Consequently, these regions were not considered as *Hb* segments for these lines.

Additional introgressed segments that had not been recorded previously (by *in situ* hybridization or molecular markers) were detected in 13 ILs (1, 4, 18, 20, 46, 60, 64, 66, 67, 80, 94, 97, and 99). Furthermore, two ILs had only one introgressed segment detected by GBS compared with the two segments initially recorded (ILs 118 and 150). Another four ILs were found to carry *Hb* segments in different genomic regions from those determined earlier. While *in situ* hybridization of ILs 29, 32, 110, and 133 revealed *Hb* chromatin on 4HL, 5HL, 2HL, and 2HS, GBS detected segments on 2HL, 4HS, 4HL, and 2HL, respectively.

Approximate physical and genetic start and end positions of individual introgressed segments were estimated based on the occurrence of *Hb*-specific SNPs (Table 1 [1HS] and Supplemental File 6 [1H–7H]). Since the barley genome reference is currently in an incomplete draft state (International Barley Sequencing Consortium, 2012), start and end positions of *Hb* segments were determined based on an uninterrupted SNP coverage along the introgressed regions. Thus, if a minority of *Hb*-specific SNPs were allocated some centiMorgans or megabase pairs apart from the bulk of *Hb*-specific SNPs, the former were excluded since their remote allocation may indicate artifacts in the assembly of the barley reference instead of pointing to a second but smaller introgression. This rather conservative approach was chosen to avoid over-estimation of the dimensions of *Hb* segments.

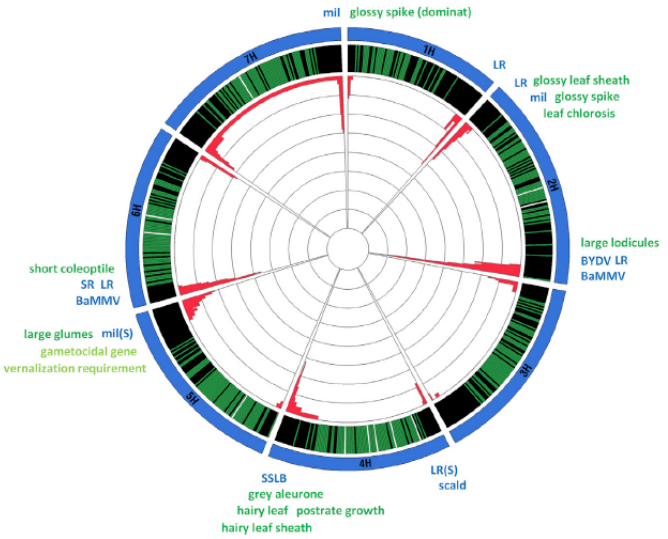

Figure 2. Overview of the Distribution and Frequency of Detected *Hb* Segments and *Hv*/*Hb* SNPs along the Barley Genome.

The figure includes a set of three concentric circles. The outer circle represents the seven barley chromosomes in a clockwise manner from the short to the long arms, scaled in 5 Mbp bins along the physical barley reference map (International Barley Sequencing Consortium, 2012). Moving inwards, the second circle is a heat map of genotyping-by-sequencing *Hv*/*Hb* SNPs frequencies found between *Hv* cultivar "Emir" and *Hb* clone "A17/1" at a minimum five-fold sequence read coverage. The frequency is plotted in 5 Mbp bins along the physical barley reference (International Barley Sequencing Consortium, 2012). The color code of these bins is set as: white, 0 SNPs/bin; dark green, 1–20 SNPs/bin; and black, >20 SNPs/bin. The innermost circle is a histogram of the frequency of *Hb* segments that were detected with genotyping-by-sequencing of the 146 ILs. The frequency is given in 5 Mbp bins along the barley genome and the gridlines of the y-axis are in a scale of 5. The three concentric circles were generated using Circos (Krzywinski et al., 2009). Outside the circles, *Hb* traits are listed that were uniquely attributed to *Hb* segments on individual chromosome arms of barley within the panel of 146 characterized ILs. Disease-associated traits are colored in blue, morphological traits are colored in dark green, and other traits are colored in light green. BaMMV, barley mild mosaic virus; BYDV, barley yellow dwarf virus; LR, leaf rust; mil, powdery mildew; (S), susceptible; scald, *Rynchosporium commune*; SR, stem rust; SSLB, Septoria speckled leaf blotch.

Hb segments were identified on all chromosomal arms (Figure 2) and GBS SNP markers were present in nearly all 5 Mbp bins (Figure 2). Based on their positions, the approximate percentage of the barley genome covered by *Hb* segments (excluding the substitution) was estimated (Table 2). Across the whole genome, about 42% and 13% of the barley genetic (Mascher et al., 2013a) and physical map (International Barley Sequencing Consortium, 2012) were covered by *Hb* segments, respectively.

Diversity in Overlapping *Hb* Segments

GBS data of the ILs was used to identify duplicates and can be exploited to evaluate the diversity of overlapping *Hb* segments. The collection of 146 ILs is the result of transferring chromatin from 12 *Hb* clones into different *Hv* genomic backgrounds. ILs derived from the same *Hb* donor clone have the probability of exhibiting identical *Hb* haplotypes if the *Hb* segments of these lines are located within the same genomic region. However, since *Hb* is a self-incompatible and obligatory outbreeding species, each *Hb* clone is likely to be heterozygous at most genomic loci. Thus, each introgression event tracing back to independent *Hb* gametes may carry a different haplotype at orthologous introgressions even if they result from the same *Hb* clone. In order to evaluate the genetic diversity between orthologous and homozygous *Hb* segments in different ILs, GBS haplotypes can be compared between ILs carrying introgressions at the same *Hv* linkage group.

Potentially overlapping *Hb* introgressions were determined based on their physical positions predicted on the basis of the *Hv* reference genome. A pair of ILs with overlapping *Hb* segments was only considered for further investigation if there were at least 20 SNP loci with sufficient read coverage in the two lines that should be compared (Supplemental File 7). In this context, an SNP locus was defined as an SNP position showing at least one SNP in at least one combination of two samples within the dataset (i.e. a polymorphic SNP between *Hv* and *Hb*). This revealed a total of 1383 pairs of ILs with overlapping *Hb* segments across the whole genome. A large proportion of the overlapping IL pairs (665) were allocated to 2HL, whereas, for 5HS, no comparable overlap was detected.

Identification of Duplicates

In order to identify overlapping *Hb* segments of the same haplotype, homozygous (non-missing) SNPs between each two ILs with overlapping *Hb* segments were counted for overlapping intervals and the genetic diversity within the overlap was measured as the percentage of identical genotype calls between the samples (identity by state) (Supplemental File 7).

The identity at variant positions in overlapping *Hb* fragments ranged from 24% to 100%. Two *Hb* segments were interpreted as representing the same haplotype, only if they had a minimum of 98% of identical SNP calls. In total, 71 *Hb* segments of individual ILs were found to be identical to at least one other *Hb* segment (Figure 3 and Supplemental File 7). All but three of the identical pairs shared the same *Hb* donor. In the remaining three cases, the compared samples were presumed to have originated from different *Hb* donors. One such case was the overlap of introgressions of IL 32 and 58. GBS genotyping of IL 32 had already predicted a mix-up of this sample. The other two cases were overlapping introgressed segments on 2HL between IL 5 and IL 79 and between IL 2 and IL 79. While ILs 2 and 5 are

Chromosome	Total length (cM)	Total length (Mbp)	*Hb* coverage (cM)	*Hb* coverage (Mbp)	% *Hb* (cM)	% *Hb* (Mbp)
1H	130	460	51	47	39	10
2H	145	625	59	72	41	12
3H	150	560	46	52	31	9
4H	115	540	69	130	60	24
5H	165	560	78	87	47	16
6H	125	535	42	33	34	6
7H	140	600	63	91	45	15
Total	970	3880	408	512	42	13

Table 2. Approximate Genetic and Physical *Hv* Genome Coverage of *Hb* Segments in 145 *Hb*/*Hv* ILs Based on Genotyping-By-Sequencing Data.

derived from *Hb* clone "2032," IL 79 was presumed to have originated from "A17/1." However, comparison of the genotypes at the GBS SNP loci between "A17" and "2032," revealed 100% and 95% identity between these clones allowing only homozygous or both homozygous and heterozygous SNPs, respectively. The origin of "A17/1" is not completely resolved, but it is either progeny or a sister plant of "A17"; thus, both should share a substantial part of their genome. The high identity between "2032" and "A17" and consequently also between "2032" and "A17/1" may explain why ILs 2, 5, and 79 have identical haplotypes within the 2HL *Hb* segments. As all of the ILs with recognized identical haplotypes shared the same or a related *Hb* donor, it is likely that these represent true duplicates of the same introgression.

Determination of Diversity

Apart from the foregoing, based on the analysis pipeline highlighted above, it cannot be concluded that *Hb* segments with less than 98% GBS haplotype identity are truly diverse. The determination of the position of *Hb* introgressions is not absolute due to uncertainties in the positioning of reference contigs of the barley draft genome sequence. This ambiguity may lead to a comparison of a longer *Hb* segment in one IL to a shorter *Hb* segment in a second IL. Thus, an *Hb* segment in one IL is actually compared with *Hv* chromatin in the second one. The apparent identity will be decreased as *Hv*/*Hb* SNPs will alter the result. However, it is possible to predict polymorphic *Hb* sites if the potentially diverse (<98% identity) overlapping GBS haplotypes are compared individually. For this, the following should be considered: the GBS haplotypes of the *Hv* and *Hb* donors (if applicable); the distribution and location of the SNPs (*Hv* SNPs are likely to appear as clusters and/or toward the end of the smaller introgression); *Hv* alleles should be found only within the region of the smaller *Hb* segment; the diversity that is expected based on the breeding scheme (i.e. two *Hb* segments originate from the same or different *Hb* clone); if only a few potential *Hb*/*Hb* SNPs are detected, they are likely to be sequencing errors, etc.; however, recombination can occur within the *Hb* donor clone and should also be considered.

For example, ILs 2 and 7 were derived from the same tetraploid *Hv*/*Hb* hybrid, which was produced by crossing diploid *Hv* with diploid *Hb* and subsequent chromosome doubling within the hybrid. Thus, the derived hybrid is known to contain just a single haplotype of the *Hb* genome and ILs from this hybrid with overlapping *Hb* segments must exhibit identical haplotypes within these regions. Based on the automated pipeline described above, ILs 2 and 7 were found to be 96% identical within the overlap of the 2HL *Hb* segment (Supplemental File 7). When looking at the predicted polymorphic SNPs, we found that all 12 SNPs were located on a single barley reference contig (Supplemental File 8, example 1). The SNPs of IL 2, which was proposed to have the smaller *Hb* segment, showed the identical haplotype to the *Hv* donor genotype "Emir". Furthermore, the *Hb* donor "2032" was not heterozygous at these positions. Consequently, those SNPs cannot be derived by different *Hb* haplotypes; instead the size of the *Hb* segment of IL 2 was overestimated and these SNPs are actually polymorphisms between *Hv* and *Hb*. As a second example, ILs 139 and 147 were derived from different hybrids but both of them originated from the *Hb* clone "2920/4"; thus, polymorphisms are only expected at heterozygous sites of the *Hb* donor. Within the 6HS overlap, 14 SNPs were polymorphic between the ILs and heterozygous within the *Hb* segments (Supplemental File 8, example 2) and are probably true *Hb*/*Hb* polymorphisms. A third example is the 2HL overlap between ILs 143 and 92. These ILs are derived from different *Hb* clones "A17" and "2920/4"; thus, they are expected to be polymorphic. Within the overlapping region, 47 polymorphic SNPs between IL 143 and 92 can be compared with the respective *Hb* donors (Supplemental File 8, example 3). The vast majority of 38 SNPs can be clearly attributed to the different *Hb* haplotypes; thus, they likely represent true *Hb*/*Hb* polymorphisms. In nine of the 47 cases (SNPs on morex_contig_2552101, morex_contig_184743, morex_contig_2549473, morex_contig_244682, and morex_contig_511440), the SNP calls of the IL alleles cannot have originated from the *Hb* donor (identity by descent). Thus, these sites might indicate sequencing errors in either *Hb* or the ILs, or polymorphisms that in fact represent *Hv*/*Hb* SNPs. It should be also kept in mind that all of the nine cases are associated with the "A17" *Hb*. This *Hb* is tetraploid; thus, SNP calling in this line will be more erroneous than in diploids.

Number of Possible *Hb* Alleles

The number of possible *Hb* alleles on each chromosomal arm was calculated based on the pedigree information of each IL. Our panel of ILs originated from 12 different *Hb* clones, which contain the genomes of 11 different *Hb* genotypes. For instance, *Hb* clone "2920/4 × Tinos16/1" was produced by crossing *Hb* clones "2920/4" and "Tinos16/1". Thus, this clone does not

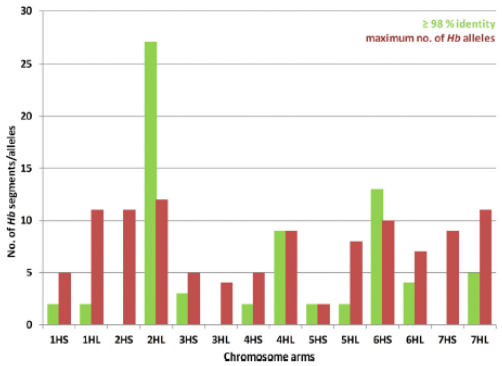

Figure 3. Diversity Analysis of Overlapping *Hb* Segments.
The diversity of overlapping *Hb* segments was analyzed across the 14 chromosome arms of barley (1HS–7HL). The identity between orthologous *Hb* segments in independent ILs was determined based on the genotyping-by-sequencing data (green bar ≥98 identity). Furthermore, the highest potential number of *Hb* alleles represented by *Hb* segments on the different chromosomal arms is given (red bar). The latter was calculated based on the *Hb* donor clones utilized, the *Hb* clone cytotype, and the breeding scheme of the ILs with *Hb* segments on individual chromosome arms.

contain additional diversity. Based on the breeding information, it was possible to calculate the maximum possible number of *Hb* alleles for each chromosome arm (Figure 3). This calculation took into account the *Hb* donor clones utilized, the *Hb* clone cytotype, and the breeding scheme of the ILs with *Hb* segments on individual chromosome arms. This analysis revealed that the highest probability of possible allele diversity existed for introgressions present on chromosome arms 2HL, while this probability was lowest on chromosome 5HS. All GBS genotype calls supported by at least five-fold sequence read coverage and a minimum SAMtools SNP call quality of five for all ILs can be found in Supplemental File 9.

Design of an Integrated *Hv*/*Hb* SNP Sequence Resource

Exome capture sequences of five *Hb* accessions and 13 *Hv* cultivars (Mascher et al., 2013b) were utilized to develop an integrated *Hv*/*Hb* sequence as an extensive resource of *Hv*/*Hb* polymorphism and sequence information to be explored for marker development in ILs of any *Hv* and *Hb* donor.

Within introgressed *Hb* segments, GBS provided about four SNPs per Mbp (International Barley Sequencing Consortium, 2012) and cM (Mascher et al., 2013a). These SNPs and the associated information on GBS sequence reads may be directly utilized for the targeted design of molecular markers within the introgressed fragments. However, this polymorphism resource is not endless; it is limited due to the relatively short GBS sequence reads with about 100 bp of continuous sequence information. Exome capture based re-sequencing of individual ILs would be a good way to increase the number of polymorphisms and the amount of sequence information (Mascher et al., 2013b; Wendler et al., 2014). For instance, exome capture provided about 150 SNPs/Mbp within the introgressions and 45 Mbp of sequence that could be compared between an IL and its two donor lines, at a minimum of 10× sequence read coverage (Wendler et al., 2014). However, exome capture might be not feasible for certain research purposes. Data mining on existing exome capture datasets should provide a good alternative if resources are lacking or other reasons hinder re-sequencing experiments. Therefore, the idea was to design an integrated *Hv*/*Hb* exome capture sequence based on the available exome capture data of several *Hv* and *Hb* accessions.

Toward this aim, exome capture re-sequencing data of five *Hb* accessions "A42", "BCC 2061", "2940/4", "A17", and "A40" and 13*Hv* cultivars "Barke", "Bowman", "Morex", "Steptoe", "Bonus", "Borwina", "Foma", "Gull", "Vogelsanger Gold", "Kindred", "Haruna Nijo", "Igri", and "Harrington" were mapped to the barley draft reference sequence (International Barley Sequencing Consortium, 2012) for data mining of sequence variants. SNP loci displaying consistently one allele (homozygous) in all *Hb* and the other allele (also homozygous) in all *Hv* accessions, respectively, were classified as interspecifically conserved *Hv*/*Hb* SNP loci (Table 3). Those SNP loci displaying different alleles in either *Hb* or *Hv*, or in both species, were considered as sites of intraspecific variation. The information gathered from the entire sequence variation dataset was integrated into a FASTA-format sequence file on the basis of the whole-genome shotgun sequence assembly of cultivar "Morex" (International Barley Sequencing Consortium, 2012) (Supplemental File 10). Base pair positions with less than five-fold sequence coverage in any of the 18 samples were masked as "N" and any detected SNP (interspecific, intraspecific, and low quality) within the set of accessions was highlighted using IUPAC code nomenclature (IUPAC-IUB Commission on Biochemical Nomenclature (CBN), 1970) (Figure 4A). This resulted in almost 36 Mbp of sequence information with 1.1 M IUPAC code masked SNPs as a basis for marker development. In total, 112 847 physically (International Barley Sequencing Consortium, 2012) and/or genetically (Mascher et al., 2013a) anchored polymorphisms were considered as conserved, interspecific *Hv*/*Hb* SNPs with an average of 21 SNPs per Mbp and 105 SNPs per cM. A list of alleles and genetic and physical positions of all filtered conserved *Hv*/*Hb* SNPs is given in Supplemental File 11. As an example, within the ~10 Mbp 2HS *Hb* segment of IL 1, GBS provided 138 anchored SNPs located on 60 different reference contigs with around 6.5 kbp of GBS sequence information (Supplemental File 3). By contrast, in the same genomic regions, the integrated *Hv*/*Hb* sequence provided 996 interspecific conserved SNPs on 195 different reference contigs and ~400 kbp of associated sequence information. Thus, when the genomic position of an IL is known, the SNP information can be accessed in the integrated *Hv*/*Hb* sequence (Figure 4B).

DISCUSSION

Genotyping of Introgression Lines

GBS re-sequencing allowed the precise genomic characterization of all but one of the 146 publicly available *Hv*/*Hb* ILs. *Hb* segments were identified on all chromosomal arms and even very small (<1 MB) introgressed fragments were detected and localized. In fact, 13 ILs displayed additional introgressed segments that had not been identified by previous attempts. Most of these

	Hv-1	*Hv*-2	*Hv*-n	*Hb*-1	*Hb*-2	*Hb*-n	type
Allele	A	A	A	B	B	B	Interspecific
	A	B	A	B	B	B	Intraspecific
	A	A	A	B	A	B	Intraspecific
	A	B	A	B	A	B	Intraspecific

Table 3. Scheme of *Hv*/*Hb* SNP Classification for an Integrated *Hv*/*Hb* Sequence Resource.
The color code of column "type" refers to the nucleotide color code in Figure 4.

fragments were smaller than 5 Mbp. Thus, these segments were either too small to be detected by cytogenetic methods (Lukaszewski et al., 2005) or the molecular markers used had an insufficient density along the genome. Variant detection and delineation of *Hb* segments were performed with the help of the barley physical and genetic reference maps (International Barley Sequencing Consortium, 2012; Mascher et al., 2013a). As almost all introgressions could be detected and localized in the homozygous state (i.e. considering only homozygous variant calls in the IL), they should represent homologous translocation of orthologous and highly collinear regions of the *Hb* and *Hv* genomes. In the case of non-homologous translocation, these fragments should appear in a heterozygous state since the respective syntenic *Hv* region would still be present in the background of the IL genome. Thus, GBS sequences of these *Hv* and *Hb* segments would map to the same region of the *Hv* reference. The same is true for reciprocal translocations between non-homologous chromosomes. It is possible that either scenario applies to the eight ILs where *Hb* segments were detected as heterozygotes. However, it is more likely that these ILs are true heterozygotes as a result of the breeding scheme or other reasons such as lethality factors that have been described in the context of *Hv*/*Hb* ILs (Ruge-Wehling et al., 2006).

Overall, the concordance between the GBS results and previous studies was very high. Genotypic data for most of the ILs characterized here have been published before (e.g. Johnston, 2007; Johnston et al., 2009). However, in most cases, genotyping was based on a few markers and/or *in situ* hybridization only. In two ILs, GBS detected only one of two *Hb* segments that were initially proposed. We cannot exclude the possibility that these fragments are located in a region were *Hv*/*Hb* GBS markers are lacking. However, this is rather unlikely as GBS markers were detected in almost all 5 Mbp bins, especially in non-centromeric regions. Presumably, these *Hb* segments occurred in a heterozygous state in early IL plants and were lost (i.e. fixed for the *Hv* allele) during subsequent breeding and propagation of the material. Another four ILs were found to have *Hb* segments in different genomic regions than detected earlier. While cytogenetic analysis of ILs 29, 32, 110, and 133 revealed *Hb* chromatin on 4HL, 5HL, 2HL, and 2HS, GBS detected segments on 2HL, 4HS, 4HL, and 2HL, respectively. Only the 5HL location of IL 32 was previously also characterized using molecular markers. *In situ* hybridization signals of chromosomes 2H and 4H usually have clearly distinct signals, but the prostrate habit and pubescence of IL 110 suggest a 4HL *Hb* introgression as detected with GBS, as this phenotype has otherwise mainly been associated with *Hb* introgressions on 4HL (R.P., unpublished data). IL 133 was originally classified as carrying an introgression on 2HS; however, 2HS and 2HL can be difficult to discern since fluorescence *in situ* hybridization signals can be fairly similar (Pedersen and Linde-Laursen, 1994). We have no explanation for the inconsistencies in ILs 29 and 32, so a mislabeled seed stock or a mix-up during DNA extraction and library construction are possible explanations. In a single line (IL 137), GBS did not detect any *Hb* segment even though this plant tested positive for the presence of an *Hb* introgression and GBS sequence information was not significantly lower compared with other ILs.

To date, *Hb*/*Hv* ILs are of great interest to a diverse community of plant breeders and researchers. Between 1998 and 2009, there have been about 30 requests per year regarding ILs maintained at the New Zealand Institute for Plant and Food Research Limited, New Zealand (R.P., unpublished results). Despite this obviously high demand, scientific publications with respect to *Hv*/*Hb* ILs are scarce and hard to trace. Furthermore, to date there has not been a single report on the release of barley varieties carrying *Hb* traits. The contradiction between seed request and utilizations has so far been mainly attributed to the paucity of *Hb*-specific molecular markers and *Hb* sequence information, which is a prerequisite for in-depth characterization of the ILs and the reduction of negative linkage drag. For instance, in a large-scale screening of 110 previously confirmed ILs using 46 PCR-based molecular markers, *H. bulbosum* segments were identified only in 88 of the lines (Johnston et al., 2009). Thus, genomic positions of the remaining ILs could not be determined. Recently, GBS and exome capture were established as efficient methods for resequencing, marker development, and characterization of *Hb* ILs (Wendler et al., 2014). The present study provides comprehensive genomic information for a publicly available collection of 145 ILs. GBS markers were found to be fairly evenly distributed across the genome; thus, there remains only a minor risk that any potential *Hb* segment escaped genotyping undetected in any of these 145 ILs. This is of high relevance for future utilization of these ILs, since any *Hb* segment might contribute beneficial as well as detrimental agronomical traits. Information on the location and polymorphisms of introgressed *Hb* segments can be utilized for the development of molecular markers.

Haplotype Analysis

We analyzed the diversity of overlapping *Hb* segments within the panel of 140 ILs with homozygous *Hb* segments. The information on the identity of haplotypes will expedite the exploitation of these lines in later studies and in breeding. Our analysis revealed that in 71 pairwise comparisons of introgressions, the overlapping *Hb* segments carried identical haplotypes and thus contain identical *Hb* alleles of any given gene within the overlapping segments. Nevertheless, it is not possible to determine whether the compared introgressed segments of these pairs are exact duplicates, as the exact physical borders of the introgressions could not be revealed. However, they will share a large proportion of the *Hb* haplotype.

On the contrary, however, this automated pipeline was not suitable for predicting with certainty the diversity within pairwise

Figure 4. Schematic Visualization of the Setup and Utilization of an Integrated *Hv/Hb* Sequence Resource.
The integrated *Hv/Hb* sequence was designed based on exome capture re-sequencing data of five *Hb* accessions and 13 *Hv* cultivars.
(A) A FASTA file example for an imaginary integrated *Hv/Hb* whole-genome shotgun assembly contig of barley cultivar "Morex." The color code of the nucleotides, given as IUPAC codes (IUPAC-IUB Commission on Biochemical Nomenclature (CBN), 1970), refers to the code of interspecific and intraspecific *Hv/Hb* SNPs as listed in Table 3. Furthermore, nucleotides with less than five-fold sequence read coverage in any exome capture sample were masked as "N" and potential low-quality SNPs were incorporated as IUPAC codes as well (IUPAC-IUB Commission on Biochemical Nomenclature (CBN), 1970).
(B) A schematic example of the utilization of the integrated *Hv/Hb* sequence resource. Integrated *Hv/Hb* sequence and SNP information can be exploited for marker development in ILs if the location of the *Hb* segment is known based on the barley physical and/or genetic map (International Barley Sequencing Consortium, 2012; Mascher et al., 2013a). Thus, integrated *Hv/Hb* sequence contigs and associated information on conserved *Hv/Hb* SNPs, which are located within the *Hb* segment, can be extracted based on their position within the barley genome.

comparisons of introgressions with less than 98% identity. The fact that the exact physical limits of *Hb* introgressions could not be determined by our approach may result in a comparison of differently sized *Hb* segments. The detectable identity will deviate from 100% due to the number of polymorphisms originating from *Hv/Hb* comparisons in the region that is only represented by the larger introgression. Consequently, based on this automated pipeline, we cannot exclude the possibility that pairs with less than 98% identity are in fact duplicates if they are derived from the same *Hb* clone.

One option to increase the precision in identifying duplicates in ILs that are derived from the same *Hb* donor would be to consider only SNP positions, which are heterozygous within the *Hb* clone. However, only 3 of the 12 *Hb* donor clones have been re-sequenced in the present study. Therefore, this analysis was not an option and, in addition, the problem of incorporating *Hv/Hb* SNPs into the analysis would remain.

Alternatively, it is possible to judge each pairwise comparison by visually inspecting the graphical genotypes as explained above. This allows the reliable prediction of polymorphic sites between *Hb* segments. It would be particularly advantageous if GBS data of the original *Hb* and *Hv* genotypes were available. While GBS from *Hv* donors of all but one of the ILs is present, only 71 ILs have GBS data of the original *Hb* donor. Thus, for future studies, the missing *Hb* donor clones should be genotyped if available.

Strategy to Avoid Low Interspecific Recombination

Low frequencies of interspecific recombination in *Hb* introgressions complicate genetic fine mapping and map-based cloning of genes underlying favorable traits contributed by *Hb* introgressions (Ruge et al., 2003; Ruge-Wehling et al., 2006; Johnston et al., 2013). In previous studies, an up to 14-fold reduction of recombination frequency was found when mapping traits in *Hv/Hb* ILs (Ruge-Wehling et al., 2006; Johnston et al., 2013). Genetic mapping thus requires the screening of much larger segregating populations to find a sufficient number of informative recombinant plants. The information on the extent and diversity within overlapping *Hb* fragments of different ILs provided in this study might be exploited in a strategy to tackle the problem of reduced recombination frequencies (Figure 5). If

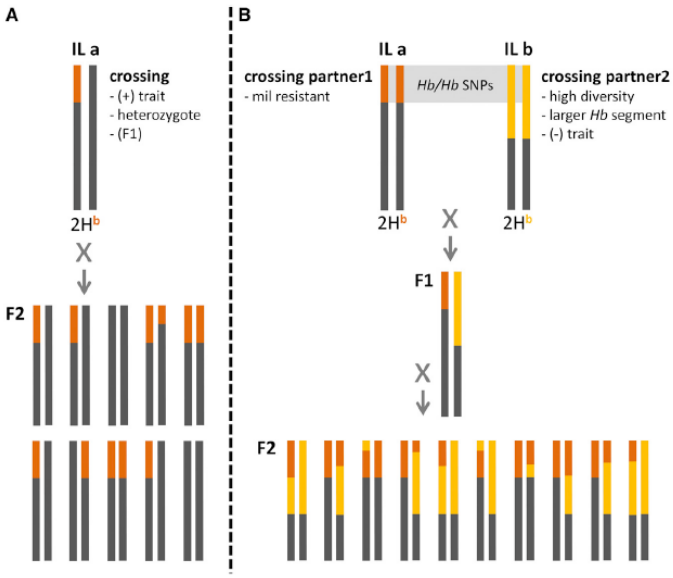

Figure 5. Alternative Crossing Strategies for *Hv/Hb* ILs.
Schema of alternative crossing strategies for the production of populations that segregate for introgressed *Hb* segments.
(A) Usually *Hv/Hb* IL mapping populations are produced by selfing of an IL plant (here with resistance to powdery mildew [mil]), which is known to be heterozygous for the *Hb* segment. Progeny of this cross will segregate like an F_2 population within the *Hb* interval and usually exhibit strongly reduced interspecific recombination frequency.
(B) Alternatively, two ILs with overlapping *Hb* segments might be crossed to produce an F_1 hybrid, which is heterozygous for the two combined *Hb* segments. If the F_1 is selfed, the resulting F_2 population will segregate between the *Hb* segments and in theory should exhibit normal intraspecific recombination frequency. Preferably, the introgressed segment from *Hb* of the respective IL, which is to be mapped (crossing partner1), should be smaller than the *Hb* segment of the crossing partner (crossing partner2). *Hb* segments should exhibit a high number of polymorphisms and, if possible, crossing partner2 should be negative for the trait to be mapped.

two ILs exhibit overlapping *Hb* segments of high diversity (e.g. derived from different *Hb* donors), these ILs might be crossed to obtain progeny that exhibits normal intraspecific recombination rates within the introgressed *Hb* segments. Usually *Hv/Hb* IL mapping populations are produced by selfing an IL plant that is heterozygous for the *Hb* segment. This results in progeny principally segregating like an F_2 population within the genomic region of the introgression (Figure 5A). For example, IL 1 was found to be resistant to leaf rust and powdery mildew associated with the *Hb* segment allocated to 2HS. Genotyping of 82 individuals of an F_2 population using four restriction fragment length polymorphic markers spanning a 49-cM interval on the barley map (Timmerman et al., 1993) revealed no informative recombination event (Pickering et al., 1998). In our proposed alternative strategy, this line may be crossed with another IL with an overlapping 2HS *Hb* segment. The segment of the crossing partner (crossing partner2, Figure 5) should be larger, to ensure the complete coverage of the *Hb* segment that is to be mapped (crossing partner1, Figure 5). The overlapping segments should exhibit a high number of polymorphisms as a source for marker development and preferably both ILs should be derived from different *Hb* donor clones to ensure the highest diversity. In the case of IL 1, IL 88 might be an optimal crossing partner. Both ILs are derived from different *Hb* donors ("S1" and "A17/1") and GBS revealed 58 potential *Hb/Hb* SNPs within the ~10 Mbp overlap (Supplemental File 8, example 4). The 2HS *Hb* segment of IL 88 extends about 9 Mbp beyond the *Hb* segment of IL 1. Furthermore, IL 88 was found to be susceptible to leaf rust and some powdery mildew isolates (Shtaya et al., 2007). When IL 1 and 88 were crossed, the F_1 generation would be heterozygous for the two distinct *Hb* segments. Selfing of the F_1 would then result in an F_2 population segregating within the two *Hb* segments; thus, in theory a normal intraspecific (intra-*Hb*) recombination frequency should be expected (Figure 5B). Even in cases where crossing partner2 did not differ in the phenotype of the target trait, the strategy would still allow one to genetically order markers due to increased intraspecific recombination frequencies. This strategy will be of particular value for chromosomal regions with a high abundance of *Hb* segments, since the possibility of finding an appropriate crossing partner is high, e.g. 2HL. Recently, Johnston et al. (2015) has demonstrated the usefulness of a similar approach using overlapping *Hb* segments to generate ILs with particular small introgression sizes.

Integrated *Hv/Hb* Sequence

To increase the SNP and sequence resource for *Hv/Hb* ILs, we developed an integrated *Hv/Hb* sequence resource providing access to interspecific and intraspecific SNP information based on exome capture re-sequencing data and the whole-genome shotgun assembly of barley (International Barley Sequencing Consortium, 2012).

The interspecific SNPs are fixed differences between *Hv* and *Hb* accessions. These SNPs are likely to be convertible into polymorphic markers for ILs derived from any *Hv* and *Hb* donor. Furthermore, information on intraspecific and low-quality SNPs within the sequence will be important to optimize marker development, for instance, by avoiding primer placement in polymorphic regions.

The present study provides a rich toolkit for the utilization of a large panel of *Hv/Hb* ILs in applied plant breeding, gene isolation, and comparative genomics. Furthermore, the established molecular methods and bioinformatics pipelines can be readily applied

to characterize and utilize further newly developed ILs. It is noteworthy that none of the presented results could have been obtained without the availability of the barley reference genome (International Barley Sequencing Consortium, 2012; Mascher et al., 2013a).

METHODS

Plant Material

Seeds of ILs containing DNA segments of *Hb* in the genetic background of *Hv* as well as seeds of the associated *Hv* donors were kindly provided by the NordGen gene bank (Alnarp, Sweden). In brief, ILs were derived from either triploid or tetraploid hybrids between *Hv* and *Hb*. Those interspecific hybrids were obtained using 12 different *Hb* clones "S1", "3811/5", "2920/4 × 2929/1'' (diploid), "2920/4 × Tinos16/1", "Tinos16/1", "Crete6/2", "PI 204579", "PI 206443/3", "A17", "A17/1" (tetraploid), and "2032", "2920/4'' (used as diploid and tetraploid) in different combination with the five *Hv* cultivars "Morex", "Vada", "Emir", "Golden Promise" and "907-12". For the formation of triploid hybrids (VBB, where V and B stand for haploid genomes of *Hv* and *Hb*, respectively), tetraploid *Hb* (BBBB) clones were crossed as pollinator with diploid *Hv* (VV) as previously described (Pickering, 1991). Tetraploid hybrids were produced by crossing diploid *Hb* (BB) clones as pollinator to diploid *Hv* (VV) and young tillers of the developing hybrids were subsequently treated with colchicine to restore fertility as described (Zhang et al., 2001). Alternatively, diploid *Hb* (and *Hv*) donor genotypes were treated with colchicine previously to crossing to produce triploid (and tetraploid) hybrids (Szigat and Pohler, 1982). To obtain diploid progeny, hybrids were either backcrossed to the recurrent *Hv* parent (triploid hybrids) or allowed to self-fertilize (tetraploid and occasionally triploid hybrids) (Johnston et al., 2009). Details on the breeding scheme and donor genotypes for the individual ILs and some description regarding the *H. bulbosum* donors are given in Supplemental File 1.

Initial characterization of ILs by either *in situ* hybridization or molecular markers was performed as described (if not published otherwise) (Pickering et al., 2000; Johnston et al., 2009). Recombinant plants with verified *Hb* segments were propagated by self-fertilizing or by doubled haploid production (Kasha and Kao, 1970) to obtain (if possible) ILs with homozygous *Hb* segments. If not published otherwise, phenotyping of these lines was done previously at the New Zealand Institute for Plant & Food Research Limited, Christchurch, New Zealand, for obvious morphological traits (i.e. pubescent leaf sheath, growth habit, etc.) as well as in various locations for their resistance to leaf rust (*Puccinia hordei*), barley mild mosaic virus, barley yellow dwarf virus, stem rust (*Puccinia graminis*), scald (*Rhynchosporium commune*), powdery mildew (*Blumeria graminis* spp. *hordei*), and septoria speckled leaf blotch (*Septoria passerinii*) in glasshouse and/or field tests as previously described (Ruge et al., 2003; Toubia-Rahme et al., 2003; Pickering et al., 2004; Singh et al., 2004; Shtaya et al., 2007; Fetch et al., 2009; Scholz et al., 2009). Leaf material of *Hb* donor lines "2032," "2920/4," and "2929/1" was kindly provided by Katrin Kumke (Leibniz Institute of Plant Genetics and Crop Plant Research, Gatersleben, Germany) and leaf material of the *Hb* clones "A17" and "A40'' (tetraploid) were kindly provided by Brigitte Ruge-Wehling (Julius Kühn-Institut, Institute for Breeding Research on Agricultural Crops, Groß Lüsewitz, Germany). The *Hb* clones were maintained vegetatively and prevention of outcrossing was done by regularly removing the emerging ears. However, it remains possible that the re-sequenced *Hb* clones may have out-crossed.

For the re-sequencing experiments, DNA extraction of young leaf tissue was performed essentially as described before (Stein et al., 2001). Extracted DNA of all analyzed ILs tested positive for the pSc119.1 assay (Johnston and Pickering, 2002) to confirm the presence of an *Hb* segment. GBS based re-sequencing was applied to 146 ILs as listed in Supplemental File 1. Furthermore, four diploid *Hv* cultivars "Vada", "Morex", "Emir", and "Golden Promise" and the four *Hb* clones "2920/4", "2032'', "2929/1", and "A17" were genotyped with GBS. Exome capture was performed on the *Hb* clones "A17" and "A40."

NGS Library Construction and Sequencing

GBS library construction using restriction enzymes *MspI* in combination with *PstI* High-Fidelity (NEB, Ipswich, UK) and sequencing was performed following previously established procedures (Wendler et al., 2014). Up to 94 pooled samples with individual indices (Wendler et al., 2014) were sequenced together on a single HiSeq 2000 lane (Illumina, San Diego, CA, USA). The formation of clusters on cBot and 1 × 100 bp single-end sequencing (HiSeq2000, Illumina, San Diego, CA, USA) were performed according to the manufacturer's instructions.

Construction of exome capture libraries and sequencing were performed as previously described (Mascher et al., 2013b; Himmelbach et al., 2014; Wendler et al., 2014). Illumina TruSeq Paired End libraries (Illumina Part #15026486) were prepared using DNA Adapter Indices 5 and 12 (Illumina, San Diego, CA, USA). For cluster formation on the cBot (Illumina, San Diego, CA, USA), equimolar amounts of up to four libraries with different indices were pooled. Library capture, liquid array processing, quantification, and 2 × 100 bp paired-end sequencing on the Illumina HiSeq 2000 device (Illumina, San Diego, CA, USA) were performed as previously described (Mascher et al., 2013b; Himmelbach et al., 2014).

Analysis of Next-Generation Sequencing Data

De-multiplexing of raw GBS and exome capture sequence data was done with the CASAVA pipeline 1.8 (Illumina, San Diego, CA, USA). Adapter sequences were trimmed from GBS sequence reads using cutadapt (Martin, 2011) and reads shorter than 30 bp after adapter removal were discarded. Trimmed GBS reads were mapped (command mem) against the barley reference sequence (International Barley Sequencing Consortium, 2012) using BWA version 0.7.5 (Li and Durbin, 2009). The BWA command mem was called with parameter "-k 10" for a minimum seed length and "-M" to mark shorter split hits as secondary, otherwise default settings were used. Exome capture reads were mapped against the barley reference (International Barley Sequencing Consortium, 2012) with BWA version 0.6.2 (Li and Durbin, 2009) (commands aln and sampe) using the parameter "-q 15" for quality trimming. Duplicate exome capture reads were removed for subsequent analysis with SAMtools rmdup (Li et al., 2009). Raw sequence reads for three *Hb* accessions "A42", "BCC2061", and "2940/4" and 13 *Hv* cultivars "Barke", "Bowman", "Morex", "Steptoe", "Bonus", "Borwina", "Foma", "Gull", "Vogelsanger Gold", "Kindred'', "Haruna Nijo", "Igri", and "Harrington" were downloaded from https://www.ncbi.nlm.nih.gov (accession number PRJEB1810). The sequence read coverage of individual samples was determined with SAMtools depth (Li et al., 2009). SNP calling on exome capture and GBS sequence data was performed with the SAMtools pipeline (version 0.1.19) using the command SAMtools mpileup and bcftools with default settings (Li et al., 2009). The SAMtools command mpileup was called with the parameter "-C 50" for downgrading mapping quality for reads containing excessive mismatches. The additional parameter "-D" was used for SAMtools mpileup to obtain per sample read depth. Subsequently, GBS genotype calls were filtered with a custom AWK script (Mascher et al., 2013c) and the R statistical environment (R Core Team, 2012) (http://www.r-project.org/), selecting only SNPs with a minimum five-fold sequence read depth and genotype quality of at least five for homozygous and heterozygous calls.

GBS SNP frequency was visualized along the integrated physical/genetic map (International Barley Sequencing Consortium, 2012) and the population sequencing (POPSEQ)-based genetic map (Mascher et al., 2013a) of barley using R. The diversity analysis of overlapping *Hb* segments in different ILs was performed using R, considering only SNPs with a maximum of 20% of missing data points within all ILs and which occur at least once in a homozygous state. An integrated *Hv*/*Hb*

sequence FASTA-file was generated based on the whole-genome assembly of cultivar "Morex" (International Barley Sequencing Consortium, 2012) using a custom AWK script. SNPs were visualized as IUPAC codes (IUPAC-IUB Commission on Biochemical Nomenclature (CBN), 1970) within the sequence. Interspecific *Hv*/*Hb* SNP positions were filtered to conduct a minimum SAMtools SNP quality score of 200 and a minimum sequence read coverage of five-fold in all 18 samples. Base pair positions with less than five-fold sequence read coverage within any of the 18 samples were masked as "N". The frequency and distribution of introgressed *Hb* segments and GBS SNP markers was visualized using Circos (Krzywinski et al., 2009).

ACCESSION NUMBERS

The accession numbers for sequence data generated in this study are PRJEB7908 (GBS reads) and PRJEB7909 (exome capture reads). The direct links are: http://www.ebi.ac.uk/ena/data/view/PRJEB7908 and http://www.ebi.ac.uk/ena/data/view/PRJEB7909.

SUPPLEMENTAL INFORMATION

Supplemental Information is available at *Molecular Plant Online*.

FUNDING

The work was financially supported as part of the collaborative project TRANSBULB by a grant (0315966) from the German Federal Ministry of Education and Research (BMBF) to N.S.

ACKNOWLEDGMENTS

We gratefully acknowledge the excellent technical support from Ines Walde, Manuela Knauft, Susanne König, and Sandra Driesslein (IPK). We thank Doreen Stengel for raw data submission and Uwe Scholz and the research group Bioinformatics and Information Technology at IPK for computational resources. We also give our best thanks to Daniel Arend for uploading additional files at *e!DAL*. Furthermore, we greatly acknowledge Katrin Kumke and Brigitte Ruge-Wehling for providing us with leaf material of the *Hordeum bulbosum* accessions "2032", "2920/4", "2929/1", and "A17" and "A40", respectively. No conflict of interest declared.

Received: January 20, 2015
Revised: April 23, 2015
Accepted: May 10, 2015
Published: May 14, 2015

REFERENCES

Blake, T., Blake, V., Bowman, J., and Abdel-Haleem, H. (2011). Barley: Production, Improvement and Uses (Oxford: Wiley-Blackwell).

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**:e19379.

Fetch, T., Johnston, P.A., and Pickering, R. (2009). Chromosomal location and inheritance of stem rust resistance transferred from *Hordeum bulbosum* into cultivated barley (*H. vulgare*). Phytopathology **99**:339–343.

Feuillet, C., Langridge, P., and Waugh, R. (2008). Cereal breeding takes a walk on the wild side. Trends Genet. **24**:24–32.

Hajjar, R., and Hodgkin, T. (2007). The use of wild relatives in crop improvement: a survey of developments over the last 20 years. Euphytica **156**:1–13.

Himmelbach, A., Knauft, M., and Stein, N. (2014). Plant sequence capture optimised for Illumina sequencing. Bio-Protocol **4**:e1166.

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

IUPAC-IUB Commission on Biochemical Nomenclature (CBN). (1970). Abbreviations and symbols for nucleic acids, polynucleotides, and their constituents. Eur. J. Biochem. **15**:203–208.

Jie, X., and Snape, J.W. (1988). The cytology of hybrids between *Hordeum vulgare* and *H. bulbosum* revisited. Genome **30**:486–494.

Johnston, P.A. (2007). Molecular Characterisation of Chromatin Introgressed from *Hordeum bulbosum* L. into *Hordeum vulgare* L (New Zealand: Department of Biochemistry, University of Otago, Dunedin), p. 340.

Johnston, P.A., and Pickering, R.A. (2002). PCR detection of *Hordeum bulbosum* introgressions in an *H. vulgare* background using a retrotransposon-like sequence. Theor. Appl. Genet. **104**:720–726.

Johnston, P.A., Timmerman-Vaughan, G.M., Farnden, K.J.F., and Pickering, R. (2009). Marker development and characterisation of *Hordeum bulbosum* introgression lines: a resource for barley improvement. Theor. Appl. Genet. **118**:1429–1437.

Johnston, P.A., Niks, R.E., Meiyalaghan, V., Blanchet, E., and Pickering, R. (2013). *Rph22*: mapping of a novel leaf rust resistance gene introgressed from the non-host *Hordeum bulbosum* L. into cultivated barley (*Hordeum vulgare* L.). Theor. Appl. Genet. **126**:1613–1625.

Johnston, P., Meiyalaghan, V., Forbes, M., Habekuß, A., Butler, R., and Pickering, R. (2015). Marker assisted separation of resistance genes *Rph22* and *Rym16*^{Hb} from an associated yield penalty in a barley: *Hordeum bulbosum* introgression line. Theor. Appl. Genet. **128**:1137–1149.

Kasha, K.J., and Kao, K.N. (1970). High frequency haploid production in barley (*Hordeum vulgare L.*). Nature **225**:874–876.

Kircher, M., Sawyer, S., and Meyer, M. (2012). Double indexing overcomes inaccuracies in multiplex sequencing on the Illumina platform. Nucleic Acids Res. **40**:e3.

Krzywinski, M., Schein, J., Birol, I., Connors, J., Gascoyne, R., Horsman, D., Jones, S.J., and Marra, M.A. (2009). Circos: an information aesthetic for comparative genomics. Genome Res. **19**:1639–1645.

Lev-Yadun, S., Gopher, A., and Abbo, S. (2000). Archaeology - the cradle of agriculture. Science **288**:1602–1603.

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

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., and 1000 Genome Project Data Processing Subgroup. (2009). The sequence alignment/map format and SAMtools. Bioinformatics **25**:2078–2079.

Lukaszewski, A.J., Lapinski, B., and Rybka, K. (2005). Limitations of in situ hybridization with total genomic DNA in routine screening for alien introgressions in wheat. Cytogenet. Genome Res. **109**:373–377.

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet. J. **17**:10–12.

Mascher, M., Muehlbauer, G.J., Rokhsar, D.S., Chapman, J., Schmutz, J., Barry, K., Muñoz-Amatriain, M., Close, T.J., Wise, R.P., Schulman, A.H., et al. (2013a). Anchoring and ordering NGS contig assemblies by population sequencing (POPSEQ). Plant J. **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., et al. (2013b). Barley whole exome capture: a tool for genomic research in the genus *Hordeum* and beyond. Plant J. **76**:494–505.

Mascher, M., Wu, S., Amand, P., Stein, N., and Poland, J. (2013c). Application of genotyping-by-sequencing on semiconductor sequencing platforms: a comparison of genetic and reference-based marker ordering in Barley. PLoS One **8**:e76925.

Pedersen, C., and Linde-Laursen, I. (1994). Chromosomal locations of four minor rDNA loci and a marker microsatellite sequence in barley. Chromosome Res. **2**:65–71.

Pickering, R.A. (1991). The production of fertile triploid hybrids from crosses between *Hordeum vulgare* L (2n = 4x = 28) and *H. bulbosum* L (2n = 2x = 14). Hereditas **114**:227–236.

Pickering, R.A., Rennie, W.F., and Cromey, M.G. (1987). Disease resistant material available from the wide hybridization programme at DSIR. Barley Newsl. **31**:248–250.

Pickering, R.A., Hill, A.M., Michel, M., and Timmerman-Vaughan, G.M. (1995). The transfer of a powdery mildew resistance gene from *Hordeum bulbosum* L. to barley (*H. vulgare* L.) chromosome 2 (2L). Theor. Appl. Genet. **91**:1288–1292.

Pickering, R.A., Hill, A.M., and Kynast, R.G. (1997). Characterization by RFLP analysis and genomic in situ hybridization of a recombinant and a monosomic substitution plant derived from *Hordeum vulgare* L. x *Hordeum bulbosum* L. crosses. Genome **40**:195–200.

Pickering, R.A., Steffenson, B.J., Hill, A.M., and Borovkova, I. (1998). Association of leaf rust and powdery mildew resistance in a recombinant derived from a *Hordeum vulgare* x *Hordeum bulbosum* hybrid. Plant Breed. **117**:83–84.

Pickering, R.A., Malyshev, S., Kunzel, G., Johnston, P.A., Korzun, V., Menke, M., and Schubert, I. (2000). Locating introgressions of *Hordeum bulbosum* chromatin within the *H. vulgare* genome. Theor. Appl. Genet. **100**:27–31.

Pickering, R., Niks, R.E., Johnston, P.A., and Butler, R.C. (2004). Importance of the secondary genepool in barley genetics and breeding II. Disease resistance, agronomic performance and quality. Czech J. Genet. Plant Breed. **40**:79–85.

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.

Prescott-Allen, C., and Prescott-Allen, R. (1986). The First Resource: Wild Species in the North American Economy (New Haven, CT: Yale University Press).

R Core Team. (2012). R: A Language and Environment for Statistical Computing (Vienna: R Foundation for Statistical Computing).

Ruge, B., Linz, A., Pickering, R., Proeseler, G., Greif, P., and Wehling, P. (2003). Mapping of *Rym14* (Hb), a gene introgressed from *Hordeum bulbosum* and conferring resistance to BaMMV and BaYMV in barley. Theor. Appl. Genet. **107**:965–971.

Ruge-Wehling, B., Linz, A., Habekuss, A., and Wehling, P. (2006). Mapping of *Rym16*(Hb), the second soil-borne virus-resistance gene introgressed from *Hordeum bulbosum*. Theor. Appl. Genet. **113**:867–873.

Salamini, F., Özkan, H., Brandolini, A., Schäfer-Pregl, R., and Martin, W. (2002). Genetics and geography of wild cereal domestication in the Near East. Nat. Rev. Genet. **3**:429–441.

Scholz, M., Ruge-Wehling, B., Habekuss, A., Schrader, O., Pendinen, G., Fischer, K., and Wehling, P. (2009). *Ryd4* (Hb): a novel resistance gene introgressed from *Hordeum bulbosum* into barley and conferring complete and dominant resistance to the barley yellow dwarf virus. Theor. Appl. Genet. **119**:837–849.

Shtaya, M.J.Y., Sillero, J.C., Flath, K., Pickering, R., and Rubiales, D. (2007). The resistance to leaf rust and powdery mildew of recombinant lines of barley (*Hordeum vulgare* L.) derived from *H. vulgare* x *H. bulbosum* crosses. Plant Breed. **126**:259–267.

Singh, A.K., Rossnagel, B.G., Scoles, G.J., and Pickering, R.A. (2004). Identification of a quantitatively inherited source of *Hordeum bulbosum* derived scald resistance from barley line 926K2/11/1/5/1. Can. J. Plant Sci. **84**:935–938.

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 Breed. **120**:354–356.

Szigat, G., and Pohler, W. (1982). *Hordeum bulbosum* x *H. vulgare* hybrids and their backcrosses with cultivated barley. Cereal Res. Commun. **10**:73–78.

Szigat, G., and Szigat, G. (1991). Amphidiploid hybrids between *Hordeum vulgare* and *H. bulbosum* - basis for the development of new initial material for winter barley breeding. In Vortr Pflanzenzüchtg, A. Winkel, ed. (Wageningen: EUCARPIA), pp. 34–39.

Tanksley, S.D., and McCouch, S.R. (1997). Seed banks and molecular maps: unlocking genetic potential from the wild. Science **277**:1063–1066.

Timmerman, G.M., Pickering, R.A., and Melz, G. (1993). Characterization of *Hordeum vulgare* - *Hordeum bulbosum* chromosome substitution lines by Restriction-Fragment-Length-Polymorphism analysis. Genome **36**:507–511.

Toubia-Rahme, H., Johnston, P.A., Pickering, R.A., and Steffenson, B.J. (2003). Inheritance and chromosomal location of *Septoria passerinii* resistance introgressed from *Hordeum bulbosum* into *Hordeum vulgare*. Plant Breed. **122**:405–409.

von Bothmer, R., Jacobsen, N., Baden, C., Jorgensen, R.B., and Linde-Laursen, I. (1995). An Ecogeographical Study of the Genus *Hordeum*, 2nd edn (Rome: International Plant Genetic Resources Institute).

Walther, U., Rapke, H., Proeseler, G., and Szigat, G. (2000). *Hordeum bulbosum* - a new source of disease resistance - transfer of resistance to leaf rust and mosaic viruses from *H. bulbosum* into winter barley. Plant Breed. **119**:215–218.

Wendler, N., Mascher, M., Nöh, C., Himmelbach, A., Scholz, U., Ruge-Wehling, B., and Stein, N. (2014). Unlocking the secondary gene-pool of barley with next-generation sequencing. Plant Biotechnol. J. **12**:1122–1131.

Xu, J., and Kasha, K.J. (1992). Transfer of a dominant gene for powdery mildew resistance and DNA from *Hordeum bulbosum* into cultivated barley (*Hordeum vulgare*). Theor. Appl. Genet. **84**:771–777.

Zhang, L., Pickering, R.A., and Murray, B.G. (2001). *Hordeum vulgare* x *H. bulbosum* tetraploid hybrid provides useful agronomic introgression lines for breeders. New Zeal. J. Crop Hort. **29**:239–246.

4. General discussion

The key objective and overall aim of this study was to provide an infrastructure of advanced genomic resources for the characterization and utilization of *H. bulbosum* traits that are transferred into barley via *H. vulgare/H. bulbosum* introgression lines (IL)s. Barley and *H. bulbosum* are two closely related species that have diverged at proximately 6 million years ago (mya) (Jakob and Blattner, 2006). The two species share conserved and collinear genomes (Salvo-Garrido et al., 2001). Therefore, the hypothesis was that latest molecular methods that have been proven successful in barley as well as the barley genomic resources themselves could be exploited for the characterization of *H. vulgare/H. bulbosum* ILs. Towards this aim we utilized two state-of-the-art molecular technologies (Wendler et al., 2015; Wendler et al., 2014).

In fact, the strategic approach of this thesis had changed from the original plan that was to screen for polymorphisms by the amplification of syntenic genes from both species. This original concept of the project was based on the resource of the barley "genome zipper", a platform that utilized conserved synteny between grasses to assign more than 80 % of the barley genes along individual chromosomes (Mayer et al., 2011). Thus, primers should have been designed based on barley genes and their respective orthologs of *H. bulbosum*, situated in *H. bulbosum* introgressions. Subsequent amplicon sequencing would have been used for variant detection between barley and *H. bulbosum*. If polymorphic, markers could have been developed and applied to characterize the ILs.

In the light of the emergence of an advanced barley reference sequence (International Barley Sequencing Consortium, 2012) and highly efficient NGS-based genotyping technologies such as GBS and exome capture (Mascher et al., 2013b; Poland et al., 2012), the strategy was changed accordingly. Previously, exome capture and GBS had been effective for variant detection and genetic mapping in barley (Mascher et al., 2013b; Poland et al., 2012). Furthermore, the sequence enriched genome framework of barley (International Barley Sequencing Consortium, 2012; Mascher et al., 2013a) allowed to map and order newly obtained re-sequencing data along the barley draft genome. Therefore, these two NGS-based, genome-complexity reduction technologies (GBS and exome capture) in combination with the barley draft reference genome were evaluated regarding their usefulness to characterize *H. vulgare/H. bulbosum* ILs. The outcome of these studies (Wendler et al., 2015; Wendler et al., 2014) as well as further

opportunities and challenges for the here established methods and resources are discussed in the following.

4.1 Advanced genomic resources for barley/*H. bulbosum* ILs

GBS and exome capture were utilized to characterize *H. vulgare/H. bulbosum* ILs. When directly applied to *H. vulgare/H. bulbosum* ILs and the corresponding barley and *H. bulbosum* donor genotypes, both methods enabled for the immediate detection of *H. bulbosum*-specific polymorphisms within the ILs (Wendler et al., 2015; Wendler et al., 2014). Since mapping of the sequencing reads was performed based on the barley reference maps (International Barley Sequencing Consortium, 2012; Mascher et al., 2013a), the identified *H. bulbosum*-specific SNPs could be directly anchored to the barley genome (Wendler et al., 2015; Wendler et al., 2014). Thereby, *H. bulbosum* segments could be immediately detected and delimited, avoiding the previously required time-consuming procedure of empirical marker development and saturation.

In principle, pre-designed marker assays that have been established in barley may be applied to *H. vulgare/H. bulbosum* ILs (Ruge-Wehling et al., 2006; Ruge et al., 2003). However, if pre-designed marker assays are applied to genotypes, which have not been utilized for the marker development, sequence variations that are absent in members of the discovery panel cannot be assessed. This feature is generally known as ascertainment bias (Lachance and Tishkoff, 2013). It was found to be specifically important as more distantly related the genotyped germplasm from the germplasm of the discovery panel. For example, the BOPA1 assay (Close et al., 2009) – a marker assay designed on cultivated barley – was found to significantly underestimate the diversity in barley landraces, while performing fine in the modern cultivated gene pool (Moragues et al., 2010). *H. bulbosum* has an even larger distance to elite cultivars, so *H. vulgare/H. vulgare* variants will be different from polymorphisms between *H. vulgare* and *H. bulbosum.* Furthermore, *H. bulbosum* as a species itself is highly polymorphic (Jaffe et al., 2000). Therefore, pre-designed markers for either barley/barley polymorphisms or a particular barley/*H. bulbosum* IL are highly inefficient when applied to (different) *H. vulgare/H. bulbosum* ILs. The situation is even worse, since pre-designed barley markers often completely fail on *H. bulbosum* alleles (Johnston, 2007; Pickering and Johnston, 2005). Thus, the immediate and case-specific detection of polymorphisms is maybe the major advantage of NGS-based marker technologies and plays a particularly important role in the case of *H. vulgare/H. bulbosum* ILs

Since GBS and exome capture were both efficient for the detection and delimtaion of the *H. bulbosum* segments of ILs, it is reasonable to compare advantages and disatvatges of either method. Exome capture has been designed to enrich for complete barley genes and the captured sequence space is about 60 Mbp in size (Mascher et al., 2013b). By contrast, GBS sequences cover a few Mbp of the genome only (Wendler et al., 2014). Exome capture allowed unique mapping of nearly 90 % and 70 % of barley and *H. bulbosum* sequence reads to the barley draft reference sequence, respectively (Wendler et al., 2014). By contrast, more than 80 % and nearly 40 % from barley and *H. bulbosum* GBS reads could be mapped to the barley draft reference sequence, respectively (Wendler et al., 2014). In a consequence, the number of nucleotide variation sites detected is much higher from exome capture than from GBS. For instance, 4 GBS SNPs/Mbp compared to 150 SNPs/Mbp that were discovered using exome capture (Wendler et al., 2015; Wendler et al., 2014). Furthermore, GBS SNPs will be accompanied by maximum 50 bp (in case of 100 bp single-end sequencing) of adjacent genome sequence, while exome capture provides flanking sequences of up to several hundred or even thousand base pairs. Therefore, exome capture data will be of advantage for downstream applications, e.g. designing primers for individual markers that would assay a selected locus. This feature is not exclusive for exome capture re-sequencing data since marker systems, such as high resolution melting (HRM) markers (Reed and Wittwer, 2004), that are applicable to the information provided by the short GBS reads are also available. However, GBS targets a much smaller proportion of the genome leading to significantly reduced per sample sequencing costs for GBS compared to exome capture (Poland and Rife, 2012; Wendler et al., 2014). Also the GBS library itself is much cheaper than the exome capture library (Wendler et al., 2014). Thus, the additional power of exome capture due to higher amounts of sequence and polymorphism information comes at the expenses of higher per sample costs as compared to GBS (Wendler et al., 2014).

The higher efficiency of mapping *H. bulbosum* exome capture reads to the barley reference assembly (International Barley Sequencing Consortium, 2012) as compared to GBS, also highlights the importance of enriching for low-copy regions. Exome capture based re-sequencing is more strongly restricted to exonic, low-copy genome regions, than GBS (Wendler et al., 2014). Since the alignment of sequence reads depends mainly on sequence identity, mapping of sequence reads of one species to the reference of another species will improve if considering these more conserved regions.

As an alternative for the need to perform exome capture based re-sequencing and/or GBS on any possible individual *H. vulgare/H. bulbosum* IL and their associated donor lines, an integrated *H. vulgare/H. bulbosum* sequence resource was developed taking advantage of exome capture re-sequencing data obtained from 13 *H. vulgare* cultivars and 5 *H. bulbosum* accessions (Wendler et al., 2015). This dataset allowed to identify conserved, interspecific SNPs between *H. vulgare* and *H. bulbosum,* respectively. The species specificity of this set of SNPs implies a high probability for their diagnostic value in many if not in any possible *H. vulgare/H. bulbosum* combinations. However, this hypothesis will require additional testing in the future. A similar approach was implemented previously, with the aim to establish an "informative" set of markers based on conserved SNPs between 4 *H. vulgare* cultivars and 4 *H. bulbosum* accessions (Johnston et al., 2009). However, this study resulted in 46 fully informative markers only. By contrast, the integrated *H. vulgare/H. bulbosum* sequence resource contains information on 112 847 interspecifically conserved polymorphisms anchored to the barley draft reference sequence and genetic maps (International Barley Sequencing Consortium, 2012; Mascher et al., 2013a), together with about 35 Mbp of flanking sequence information (Wendler et al., 2015).

The availability of the barley draft reference (International Barley Sequencing Consortium, 2012; Mascher et al., 2013a) greatly benefited the present study, by allowing to directly map and anchor sequence reads and polymorphisms to the barley genome. Thus, using the barley resources it was possible to deliver highly accurate allocation information for over 145 barley/*H. bulbosum* ILs (Wendler et al., 2015; Wendler et al., 2014). However, it should be considered that their is a chance that the genetic and phiscal positions provided by the barley refrenece in some cases may not reflect the true genomic positions in barley and *H. bulbosum,* respectively.

Even though highly advanced, the WGS sequence assembly as well as the genetic and physical maps are not yet at gold standard quality (International Barley Sequencing Consortium, 2012; Mascher et al., 2013a). The WGS sequence assembly itself is highly fractionated and it is likely that certain regions such as high-copy regions carry assembly errors (International Barley Sequencing Consortium, 2012). The genetic maps utilized to anchor the physical map and WGS sequence assembly consisted of relatively low number of individuals, thus providing limited genetic resolution only (International Barley Sequencing Consortium, 2012; Mascher et al., 2013a). Also the physical barley map will contain errors, which may be due to the laboratory effort of handling this large and complex resource as well as repetitive DNA regions that will

cause miss-assemblies during BAC-fingerprinting as reviewed earlier (Meyers et al., 2004). Thus, the delimitation of *H. bulbosum* introgressed segments by help of draft genome sequence information and genetic maps of barley will be restricted to the accuracy and resolution of these resources. For instance, based on the genetically anchored barley physical map, introgressed *H. bulbosum* segments on 4HL seemed to be fractionated, since a single region of high SNP frequency was constantly found far apart from the remaining region of these introgressed segments (Wendler et al., 2015). However, by comparison to the high density barley genetic map (Mascher et al., 2013a) it was obvious that this pattern was caused by falsely anchored physical map contigs, rather than pointing at a possible presence of an additional *H. bulbosum* segment (Wendler et al., 2015).

Previous studies have suggested good overall collinearity between barley and *H. bulbosum* on the basis of 136 molecular markers (Jaffe et al., 2000; Salvo-Garrido et al., 2001). However, since barley and *H. bulbosum* diverged about 6 mya (Jakob and Blattner, 2006) (smaller) rearrangements between the genomes are likely. Such differences will be missed or cannot be observed since the analysis of data was solely based on considering the barley draft reference genome sequence. The latter should be kept in mind when working with the newly established resources in barley/*H. bulbosum* ILs. Thus, markers can be developed without obstracles based on these resources; however, there is a chance that the given genomic positions based on barley do not represent the absolutely true genomic position in *H. bulbosum* i.e. in the introgressed *H. bulbosum* segment.

4.2 Opportunities and challenges of a barley/*H. bulbosum* IL collection

Considerable numbers of barley/*H. bulbosum* ILs have beend developed over the laste decades (Johnston et al., 2009). However, most of them have been charachterized only rothly by *in situ* hybridization or a few number of mulecular markers (Johnston et al., 2009). To compile a large collection of presicely characterize *H. vulgare/H. bulbosum* ILs we utilize the newly established GBS pipeline to genotype 146 *H. vulgare/H. bulbosum* ILs and some of the corresponding donor lines towards a more educeted and more efficient use of this resource (Wendler et al., 2015). For almost all ILs the previously identified *H. bulbosum* introgressions could be confirmed, however, at much higher precision and resolution than before (Johnston et al., 2009; Pickering et al., 2000; Wendler et al., 2015).

When looking at *H. bulbosum*-specific GBS SNP frequencies, *H. bulbosum* segments could be identified and localized in 145 ILs. Owing to the high GBS marker density, 13 ILs displayed additional introgressed segments, which had not been identified by previous attempts (Wendler et al., 2015). With an average of 4 GBS SNP markers per cM and Mbp within the introgressed segments, the precision of localization is supposed to be above cM and Mbp resolution. This significantly improved the sensitivity compared to previous studies. For instance, Johnston et al. (2009) developed a set of 46 *H. vulgare/H. bulbosum* markers distributed over the barley genome, to facilitate detection and localization of introgressed *H. bulbosum* chromatin. However, when a set of 110 previously confirmed ILs was screened using these markers, *H. bulbosum* segments were identified only in 88 of the lines (Johnston et al., 2009). By contrast, the promis of GBS is the high marker density; however, the distribution of GBS markers is not completely even along the genome, since it depends on the location of restriction sites within the genome. Thus, it is possible that some regions of the genome will produce fewer than 4 SNPs per cM and Mbp. This may explain why a single IL was found to be negative for any *H. bulbosum* introgression (Wendler et al., 2015).

Genetically and phenotypically well characterized crop-wild IL libraries display an important genetic resource in plant science and applied breeding (Zamir, 2001). The benefit is quiet obvious, since exotic libraries represent a permanent genetic resource (Zamir, 2001). The material with different introgressed segments will be directly available for breeders and researchers. Thus, no time-consuming hybridization, introgression and localization of the introgressed segments are necessary. Many resistance loci for some of the most important barley diseases can be found in the *H. vulgare/H. bulbosum* IL collection (Wendler et al., 2015). These may represent monogenetic as well as polygenetic traits. Further phenotyping may facilitate the discovery of additional desirable traits also associated to yield or quality aspects. The *H. bulblosum* fragments of the entire *H. vulgare/H. bulbosum* IL collection covered about 13 % of the physical barley reference genome (International Barley Sequencing Consortium, 2012; Wendler et al., 2015). In principle exotic libraries should contain a set of introgression lines with single defined chromosome segments that as a whole will represent the complete exotic genome (Zamir, 2001). The development of further ILs that will embody the entire *H. bulbosum* genome, hence all possible *H. bulbosum* loci, would be desirable. However, this is mainly hampered by recombination/introgression frequencies and distribution. Possible approaches to alter

recombination formation in ILs and hybrids will be discussed below. It should be considered that some ILs with *H. bulbosum* segments replacing particular regions of the barley genome may be instable e.g. infertile, week or even unviable (Johnston, 2007). This may be caused by negative interactions between the *H. bulbosum* segment and the barley genome or the lacking ability of the *H. bulbosum* chromatin to compensate for the replaced *H. vulgare* segment (Johnston, 2007). For instance, no *H. bulbosum* substitution line for chromosome 1H was generated so far.

In total, the collection of 146 ILs was derived by using 11 different *H. bulbosum* donor genotypes (Wendler et al., 2015). Thus, the *H. vulgare/H. bulbosum* IL library potentially contains redundancy since some genomic regions were contributed by the same *H. bulbosum* donor genotype. The GBS analysis provided experimental proof for this level of redundancy since the same *H. bulbosum* donor haplotype was identified for several ILs for the same genomic region (Wendler et al., 2015). It can be expected that these identical introgressions carry the same *H. bulbosum* alleles and thus, do not add to the diversity of the collection. However, in the remaining non-identical IL segments the 11 *H. bulbosum* donors may have provided different alleles for any given gene (Wendler et al., 2015). This information may be utilized in the future to complement the *H. vulgare/H. bulbosum* ILs collection especially for regions of currently low *H. bulbosum* haplotype diversity by introgressions derived from new/different *H. bulbosum* accessions (Wendler et al., 2015).

4.3 Applicability of GBS and exome capture to other crop-wild introgressions

The present study may also serve as a blueprint for similar projects in other crop-wild introgression systems, especially those addressing the secondary or even tertiary gene pools of crop species. Since NGS is still being improved towards decreasing costs and increasing throughput, NGS based genotyping technologies such as GBS and exome capture are constantly and increasingly adapted to more and more species, especially if possessing large and complex genomes. However, the pipeline to characterize the crop-wild introgressions will differ between species depending mainly on the genome regarding size and ploidy and the available genomic resources. These different factors will influence the utilization of NGS-based complexity reduction based methods such as GBS and exome capture.

For instance, with about 17 Gbp (The International Wheat Genome Sequencing Consortium, 2014) the genome of wheat is more than three times larger than the barley genome. Thus,

genotyping will greatly benefit from genome-complexity reduction methods. Other species, with smaller genomes may not always require complexity-reduction prior to re-sequencing and the usefulness of complexity-reduction will depend on the difference between the size of the genome itself and the library after complexity reduction. Thus, it may not make sense to reduce the complexity of the ~ 900 Mbp genome of tomato (*Solanum lycopersicum*) (The Tomato Genome Consortium, 2012) to the 60 Mbp of exome targets, since the cost saving achieved by the lower amount of data to be sequenced may not exceed the increase of costs due to the additional effort of preparing exome capture libraries. For instance, Menda et al. (2014) analyzed wild-species introgressions in cultivated tomato using WGS sequencing. However, since GBS targets a much smaller proportion of the genome as compared to exome capture and library preparation is easy and cheap, it is a highly valuable method even for species with smaller genomes such as tomato (Labate et al., 2014).

As described above, a reference genome sequence, even if present only at the status of genetically ordered sequence contigs, is of great advantage for the here established methods. However, a reference genome or sequence is not an absolutely necessary prerequisite. For instance, in the absence of a reference genome sequence, GBS was successfully employed to construct a *de novo* genetic map for wheat from a segregating population (Poland et al., 2012). Here, GBS reads were used as an internal sequence reference to map GBS reads (Poland et al., 2012). Thus, GBS can be used without reference genomes when directly applied to a mapping population. Similarly, exome capture reads may be assembled *de novo* directly and could be used for mapping of exome capture (or GBS) reads instead of using a WGS assembly. If no genome anchoring information is available, it would be still possible to filter for exome capture *de novo* contigs with (high frequencies of) wild-species-specific SNPs. SNPs on these contigs could then be used as marker assays to construct a genetic map by the help of a segregating mapping population.

The situation might be different for polyploid species that lack any reference genome or sequence, since mapping of sequence reads and SNP calling will be much more challenged by the different sub-genomes in combination with the introgressed chromatin. Thus, for polyploid species a reference sequence that is at minimum anchored to individual chromosomes, might be a prerequisite for GBS and exome capture based genotyping of crop-wild ILs, especially if derived from the secondary or tertiary gene pool. For instance, in wheat a considerable large number of introgressions from secondary and tertiary gene pool are available. These involve chromatin from

more than 52 species (Wulff and Moscou, 2014). Thus, there is an open field of application of the heare established procedures.

Even though primary gene pool ILs can be usually genotyped using pre-designed and non-NGS based marker systems, the power of case-specific SNP detection from NGS based genotyping has been recognized as a good alternative. For instance, ILs from *H. spontanem* in barley have been developed and characterized by the help of GBS (Honsdorf et al., 2014). In rice two populations of primary gene pool ILs have been generated using GBS (Arbelaez et al., 2015). Recently, exome capture based re-sequencing allowed the prediction of thousands of SNPs between *H. vulgare* and *H. spontaneum* (Mascher et al., 2013b).

4.4 Towards *Hordeum bulbosum* gene isolation and utilization for breeding

The here established genomic resources will be directly applicable for mapping of genes underlying valuable *H. bulbosum* phenotypes. The vast pool of *H. vulgare/H. bulbosum* SNPs and sequence information will enable the targeted development of molecular markers. Flanking markers that are as close as possible to a particular gene/trait of interest may facilitate the incorporation of *H. bulbosum* genes into elite barley breeding material by marker assisted selection. However, the ultimate goal might be to know the underlying target gene in order to establish perfectly linked markers. This would allow selecting of as close as possible recombination events, which are required in order to reduce as much as possible any negative linkage drag.

Towards the efficient utilization of *H. bulbosum* genes/traits in breeding, the first and in most cases inevitable step would be the generation of a high-resolution genetic map for the trait of interest. Interestingly, besides the rich pool of genetic resources, by now, genetic maps have been developed only for six *H. bulbosum* loci in *H. vulgare/H. bulbosum* ILs (Fetch et al., 2009; Korzun et al., 1999; Pickering et al., 2006; Ruge-Wehling et al., 2006; Ruge et al., 2003). These loci have been mapped to different resolution and mapping was often hampered by reduced interspecific recombination frequencies.

Once a reliable and high-resolution genetic map has been generated, candidate genes might be identified via map-based cloning (Stein and Graner, 2005). The basic principle of map-based cloning is to establish the physical map around the target trait, i.e. between the two closest

flanking markers. The physical map can then be sequenced (if not already done) and exploited to predict possible candidate genes based on gene annotations. Physical maps generally require the availability of large insert libraries (e.g. BAC libraries) (Stein and Graner, 2005). No such library is available for *H. bulbosum* yet. Since barley is the closest relative of *H. bulbosum* with an advanced physical map and available BAC libraries, it would be obvious to exploit this resource first. However, it has to be considered that barley and therefore also the barley BAC library may not comprise the gene of interest. Due to their phylogenetic distance (Jakob and Blattner, 2006), barley and *H. bulbosum* may exhibit a considerable amount of rearrangements within the genome, such as translocations, inversion, duplications and deletions (Coghlan et al., 2005). For example cultivated tomato and its close relative *Solanum pennelli* were estimated to have diverged about 2.7 mya (Kamenetzky et al., 2010). By comparative genomics of several genome locations several structural micro-rearrangements have been detected e.g. chromosomal inversion, insertions, deletions and even compressions of *S. pennellii* genomic intervals (Kamenetzky et al., 2010). It is not possible to predict the extent of chromosomal rearrangements, since the degree varies strongly between and within phyla, being inconsistent across species or time (Coghlan et al., 2005; Fischer et al., 2006; Murphy et al., 2005; Scannell et al., 2006). Furthermore, some gene families, e.g. resistance genes, are known to evolve more rapidly leading to a high degree of rearrangements even within the same species (Richter and Ronald, 2000). For instance, when copy number variations (CNV) were assessed in different barley accessions, it was found that 9.5 % of all analyzed genes were prone to CNV (Munoz-Amatriain et al., 2013). Of these, many were annotated as disease resistance proteins and protein kinases (Munoz-Amatriain et al., 2013).

As an alternative it would be reasonable to construct a "reference" BAC library for *H. bulbosum,* or even better a case-specific BAC library of the targeted IL itself. Relying on a BAC library that is derived from a *H. bulbosum* "reference" genotype would still not ensure that any particular gene will be present within this library. *H. bulbosum* is a highly polymorphic species (Jaffe et al., 2000), so *H. bulbosum* genotypes may differ especially regarding particular polymorphism-rich genes as described for barley above. Furthermore, *H. bulbosum* is heterozygous and often polyploid, which will complicate the utilization of the BAC library due to presence of different haplotypes in individual genotypes. Thus, it would be most straight forward to work with a BAC

library that is directly produced from the IL of interest. If possible, the BAC library should be obtained from a homozygous IL, so it will contain a single copy of the target region only.

A different strategy may be provided by a "candidate gene approach" (Pflieger et al., 2001). For that, possible candidate genes are predicted and screened for sequence polymorphisms based on the target trait (Pflieger et al., 2001). For instance, it would be possible to utilize sequences from cloned genes with similar functions (especially if located in syntenic regions) and to exploit these genes for polymorphisms e.g. via sanger-based re-sequencing from the ILs or within exome capture data if available. However, it should be considered that the capture is not complete since it contains only about 74 % and 40 % of annotated barley high-confidence genes and low confidence genes, respectively (International Barley Sequencing Consortium, 2012; Mascher et al., 2013b). Furthermore, it was designed based on barley, so *H. bulbosum* genes that are absent in barley will not be within the assay. Thus, it is possible that a particular candidate gene may not be represented within the data. New methods are constantly emerging and may further speed up the process of such candidate approaches. For instance, RenSeq (Resistance gene enrichment Sequencing) a capture for the enrichment of *NB-LRR* (nucleotide-binding leucine-rich repeat domain-containing) genes allowed the discovery of many previously unknown *NB-LRR* genes in potato (Jupe et al., 2013). Notable, the majority of disease resistance genes in plants encode *NB-LRR* genes (McHale et al., 2006). Thus, in the future such assays may also accelerate candidate identification.

Another strategy would be based on functional genomics and the analysis of differential gene expression between different genotypes in relation to the expression of the desired phenotype. Previously this method relied on measuring gene expression by help of microarrays (Zhu et al., 2001). NGS allowed the implementation of direct RNA sequencing (RNA-seq) and quantifying transcript abundance by counting total number of sequence reads for individual genes of a genome as reviewed elsewhere (Wang et al., 2009). Here the use of near isogenic lines (NILs – lines that differ exclusively within a single interval i.e. the *H. bulbosum* segment) would facilitate faster candidate gene discovery. Differentially expressed and/or polymorphic genes between the samples can then be functionally annotated to identify candidate genes, e.g. resistance-related genes. For this strategy it is important to consider that some genes may not be constitutively expressed. Thus, they may require an induction e.g. via inoculation with the pathogen prior to sequencing.

In any of these candidate gene based and functional genomic approaches a good genetic map will be helpful to reduce the number of possible candidate genes. For the same reason it would also be of benefit if the the *H. bulbosum* segments of the analyzed IL would be reasonably small.

Once a candidate gene has been identified, the gene function has to be confirmed by complementation (via transformation) or mutant analysis (Stein and Graner, 2005). Only after verifying the function, the gene can be declared as "isolated" or "cloned".

Even though most of the so far reported *H. bulbosum* genes have been characterized as dominant traits it is also possible that the desired phenotype is caused by a recessive gene e.g. as described for *rpg6* a recessively inherited *H. bulbosum* trait conferrig resistance agaist steam rust (*Puccinia graminis* f. sp. *tritici*) (Fetch et al., 2009). In such case the *H. bulbosum* phenotype may be derived from a mutation in an orthologous gene of barley, or the corresponding barley gene may be completely absent in *H. bulbosum*. If the gene is absent in *H. bulbosum*, it would substantially complicate the gene isolation process, since it would require to directly identify the gene conferrig suseptibility in barley. Otherwise the recessive *H. bulbosum* gene may be cloned as described above.

After a *H. bulbosum* gene has been isolated from an IL and/or suitable close flanking markers were identified, pre-breeding will follow to prepare the plant material for utilization of the desired trait in barley breeding. This pre-breeding involves mainly minimizing negative linkage drag by reducing the size of the introgressed segments. If a high-resolution mapping population is available, recombinant plants with smaller introgressed segments may be directly selected from this material. Otherwise, markers over the whole segment and in immediate proximity to the gene can be implemented to reduce the size of the flanking *H. bulbosum* chromatin. This step will involve repeated backcrossing or selfing of the IL and it will again be hampered by reduced interspecific recombination frequencies.

Genetic modification can be an alternative way for transferring the *H. bulbosum* trait. Stable gene transformation may be achieved via *Agrobacterium tumefaciens*-mediated transformation, which has been readily established for the barley cultivar 'Golden Promise' (Tingay et al., 1997). In case of a dominant trait, the addition of the *H. bulbosum* gene into the barley genome will facilitate the expression of the desired phenotype, unless the gene is expressed and no negative

interactions between the transgene and the barley genome occur (reviewed in Lorence and Verpoorte, 2004). In case the *H. bulbosum* trait is based on a recessive gene, transformation may also be used to mutate or knock-out the orthologous (dominant) barley gene. Systems for targeted genome mutation ("Genome Editing") (Cheng and Alper, 2014) such as transcription activator-like effector nucleases (TALENs) have been reported for barley (Gurushidze et al., 2014) and may allow to manipulate the target trait in barley. CRISPR/Cas is another promising method for introducing site-specific double strand DNA breaks, but it has not been established in barley yet (reviewed in Belhaj et al. 2013). Other systems for targeted suppression of gene expression, such as RNA-interference (RNAi) in barley are available (Mendiondo et al., 2015) and may also be used. However, RNAi may be less efficient, since it usually does not result in a complete knock-out but only in a knock-down of gene expression resulting in varying reduced expression levels (reviewed in Matzke et al., 2001). In any case, TALEN, CRISPR/Cas or RNAi would also require stable transformation of the particular "knock-down"-construct via *Agrobacterium* at first. The use of transformation has the general potential to boost the process of *H. bulbosum* trait utilization, due to removing the need of removing negative linkage drag on the basis of recombination. However, genetic modification does not replace mapping of the trait and, more importantly, a lack of societal acceptance of the method complicates the regulation of such events for the use in agriculture (Verstegen et al., 2014).

4.5 Managing reduced recombination frequencies in *H. vulgare/H. bulbosum* ILs

Reduced interspecific recombination frequencies within introgressed segments pose an important bottleneck towards an efficient utilization of barley/*H. bulbosum* ILs. It has a strong negative effect on the efficiency of finding the desired recombination events e.g. for mapping of *H. bulbosum* traits or reducing the size of introgressed fragments. Thus, whenever recombination is diminished to a certain level, it will require increasing a population to the same factor, to obtain enough recombination. For instance, assuming that about 2000 individuals of a F_2 mapping population may be necessary for a high-resolution map in region of intermediate or high recombination frequencies in barley e.g. at the telomere (Stein and Graner, 2005). If recombination now would be reduced to half, the number of individuals would have to be increased to 4000 in order to obtain a map of similar resulution as comprated to the situations with "normal" recombination. Furthermore, crossovers between *H. vulgare* and *H. bulbosum* seem to be strongly restricted to distal (and interstitial) positions of the chromosomes (Johnston

et al., 2009). This crossover distribution prevents the construction of an exotic introgressions library that covers the (nearly) complete *H. bulbosum* genome. Thus, *H. bulbosum* trait utilization would highly benefit from methods that allow increasing (homeologous) recombination frequency and distribution.

Elevating the homeologous recombination frequency would be desirable for the utilization of *H. bulbosum* traits. In fact the benefit would be threefold as it would increase the efficiency i) to generate new ILs, ii) to genetically map *H. bulbosum* traits and iii) to reduce the size of *H. bulbosum* segments (i.e. to reduce negative linkage drag). The formation of a crossover in general is a very complex process, involving multiple factors with many still remaining unknown or not well understood (reviewed by Page and Hawley, 2003; Mézard et al., 2015) . Nonetheless, it is known that recombination and crossover depends on a homology search that follows a double-strand break (Ma, 2006). In accordance to this, the ratio of recombination was found to be strongly correlated with sequence identity/diversity (Li et al., 2006). Several proteins are involved in this process to prevent the junction of non-homologous regions (Li et al., 2006; Svetlanov and Cohen, 2004). For instance, in Arabidopsis silencing of *AtMSH2* increased recombination frequencies in divergent chromosome regions, suggesting that the gene is involved in controlling recombination in respect to sequence divergence (Li et al., 2006). Sequences of homeologous chromosomes might differ at varying intensities and may include inversions, translocations or mutations. Thus, the reduced crossover frequency between homeologous chromosomes of barley and *H. bulbosum* may be attributed to the nucleotide divergence between them.

The fact, that it is controlled genetically, may be utilized to increase homeologous recombination. Maybe the best known example, is the gene *Pairing homoeologous1* (*Ph1*) in wheat on 5B, which was found to inhibit pairing and recombination between homeologous chromosomes (Al-Kaff et al., 2008; Griffiths et al., 2006). If *Ph1* is absent, pairing and recombination between homoeologous chromosomes is frequent (Qi et al., 2007). After its discovery in 1958 (Riley and Chapman, 1958; Sears and Okamoto, 1958), *ph1* mutant wheat lines were used repeatedly to introgress chromosome segments of wild relatives into wheat (Islam and Shepherd, 1992; Niu et al., 2011). A recent study revealed that the locus name "*Pairing homoeologous"* was actually not completely correct, since homeologous pairing was not reduced in hybrids between wheat and rye (n=28) when no homologous chromosomes were present (Martin et al., 2014). In these

hybrids *Ph1* affected neither the level of synapsis nor the number of potential recombination sites, which can be recognized by MLH1 sites via immunolocalization (Martin et al., 2014). MLH1 is a mismatch repair protein necessary to process Double Holliday Junctions to crossovers (Ashley et al., 2001; Lhuissier et al., 2007; Martin et al., 2014). Thus, the presence of *Ph1* must have stopped the progression of these MLH1 sites to crossovers in these hybrids (Martin et al., 2014). Therefore, it was concluded that in wheat, *Ph1* promotes pairing of homologs rather than suppressing homeologous pairing and it stabilizes polyploidy by preventing crossovers between homeologs (Martin et al., 2014). It is possible that similar factors are acting during suppression of recombination in barley/*H. bulbosum* ILs. One approach would be to investigate the presence of MLH1 sites in "haploid" barley/*H. bulbosum* hybrids (VB) in a similar approach as described for the wheat-rye hybrid.

In wheat, *Ph1* was delimitated to a cluster of cyclin-dependent kinases (cdk) -like genes by using deletion mutants (Al-Kaff et al., 2008). Expression profiling of these genes in the 5B region and comparison to their homeologs on 5A and 5D, revealed that the deletion of the cdk-like locus on 5B results in activation of transcription of functional cdk-like copies on 5A and 5D (Al-Kaff et al., 2008). In Arabidopsis the closest homolog of the *Ph1* kinase was found to regulate pairing and recombination (Zheng et al., 2014), suggesting that this function it is not restricted to wheat or polyploids. It would be interesting to investigate whether paralogs of the identified cdk-like genes of 5B, 5A and 5D can be found in barley. If yes, they might be studied for instance in barley/*H. bulbosum* ILs with introgressed *Hb* segments, which are located on the same position within the genome, but at the same time, exhibiting different recombination frequencies. Fortunately, such ILs have been described earlier (Johnston et al., 2013).

The study of Al-Kaff et al. (2008) suggests that cdk-like activity is elevated and not degraded if *Ph1* is missing (Knight et al., 2010). Therefore, it was postulated that if cdk-like activity could be increased, homeologous pairing behavior should be elevated even in the presence of *Ph1* (Knight et al., 2010). Okadaic acid was found to be a protein phosphatase inhibitor that activates cdks and induces premature chromosome condensation (Yamashita et al., 1990). When tiller of wheat-rye hybrids were treated with okadaic acid, homeologous chromosome pairing was found to be similar to that of *ph1* mutants (Knight et al., 2010). Thus, it might be interesting to investigate if the latter may also increase homeologous recombination in barley/*H. bulbosum* ILs. However, the

correct concentration and time point of the okadaic acid treatment was critical and needs to be precisely defined (Knight et al., 2010).

Alternatively, it was found that alien genes from other species, such as Ph^I from *Ae. speltoides* can function as inhibitors of *Ph1* and *Ph2* through epistatic interactions (Chen et al., 1994). It would be possible that such genes exist also in barley or *H. bulbosum*. Furthermore, irradiation can be used to promote homeologous recombination. However, generated translocations are often deleterious due to genetic imbalance (Wulff and Moscou, 2014). For instance, when irradiation was used to generate translocations from *Ae. umbellulata* into wheat, only one out of 17 introgressions was not deleterious (Sears, 1956). Thus, irradiation is rather exclusively used to generate initial introgressions but not to increase recombination for genetic mapping experiments.

The genetic resource of well characterized ILs might be a good material to study interspecific recombination in barley as discussed above. Furthermore, GBS and exome capture SNPs/markers could be used in such approaches to trace recombination events.

Nonetheless, since homeologous recombination occurs between barley and *H. bulbosum*, high-resolution mapping should principally be also possible by screening very large populations segregating in the region of the introgression. It should be considered that large introgressed segments may exhibit higher relative recombination frequencies than smaller ones (Canady et al., 2006). In tomato it was found that the size-relative homeologous crossover frequency was positively correlated to the length of the introgressed segment (Canady et al., 2006). In other words, within smaller segments recombination was more strongly reduced than within larger ones (Canady et al., 2006). Consequently, whole chromosome substitutions exhibited the highest relative recombination frequencies (Canady et al., 2006). A similar pattern has been found when comparing *H. bulbosum* introgressions (Johnston et al., 2013). However, recombination will also strongly depend on the donor genotypes themselves and on the genome location (Wijnker and de Jong, 2008). In some genomic regions crossover formation may be completely absent due to larger chromosome rearrangements (Hoffmann et al., 2004).

To circumvent the problem of rare interspecific crossover occurrence, an alternative strategy may be to take advantage of intraspecific recombination between different *H. bulbosum* segments rather than relying on the limitations of interspecific recombination between barley and *H. bulbosum*. Two such strategies have been described specifically addressing i) mapping of *H.*

bulbosum traits (Wendler et al., 2015) and ii) reducing the size of *H. bulbosum* segments after the *H. bulbosum* trait has been mapped (Johnston et al., 2015). The first possible pipeline would be based on the accumulated haplotype diversity information of ILs carrying introgressions on the same chromosome arm (Wendler et al., 2015). F_2 populations derived from ILs with different haplotypes in the introgressed homologous *H. bulbosum* segments will potentially show normal rates of intraspecific recombination frequency. This scenario has been proven successful in tomato (Canady et al., 2006). By crossing an IL to a second line holding a segment that was more closely related to the exotic donor line, recombination frequencies could be significantly increased within the two exotic overlapping fragments (Canady et al., 2006). It is important that the two "crossing partners" carry sufficiently different haplotypes; thus, recombination can be traced easily by molecular markers. For the 145 ILs genotyped by GBS in frame of this thesis, appropriate crossing partners can be selected and the available GBS data can be directly utilized to develop polymorphic markers (Wendler et al., 2015). Preferably the second IL should be negative for the trait of interest to be mapped, thus the resulting progeny will segregate for the respective phenotype. However, even if the selected ILs show the same phenotype, the resulting mapping population can be used to generate a better marker order resolution within the introgressed segments, allowing e.g. to detect rearrangements within the *H. bulbosum* genome as compared to barley.

The second strategy that relies on intraspecific recombination is promising to obtain ILs with particularly small introgressed segments. This strategy has been suggested earlier (Sears, 1981; Wijnker and de Jong, 2008) and it was recently successfully transferred to barley/*H. bulbosum* ILs (Johnston et al., 2015). This strategy invokes crossing of two recombinant lines with larger introgressed segments on different sides of the target trait and that interleave at a very small interval just covering the gene of interest (Sears, 1981; Wijnker and de Jong, 2008). Thus, crossover can take place within the small homologous overlapping interval of the two wild-species segments, which facilitates to generate ILs with particularly small introgression size. The identification of the appropriate crossing partners requires access to a rather large panel of ILs and they may be most likely only identified within the mapping population itself, if available (Johnston et al., 2015).

Finally, the restriction of *H. bulbosum* introgressions to distal (and interstitial) positions may be a cause of the *H. bulbosum* genome structure itself, since strongly reduced recombination

frequencies were also observed in centromeric regions of *H. bulbosum* mapping data (Johnston, 2007; Salvo-Garrido et al., 2001). Compared to barley, the reduction of recombination at the centromeres was found to be even more severe, while recombination frequencies were higher at the telomeres (Kunzel et al., 2000; Salvo-Garrido et al., 2001).

In plants several factors have been shown to influence recombination frequency and distribution, such as presence of B chromosomes (Jones and Rees, 1967) and application of various chemical agents or physical stress, such as temperature shock or UV exposure (Fedak, 1973; Hassan and Jones, 1995; Preuss and Copenhaver). Furthermore, many proteins involved in crossover formation have been identified (Ma, 2006; Mercier and Grelon, 2008). Manipulation of recombination is a desirable goal for barley itself. Mainly for increasing the efficiency of map-based cloning in recombination cold spots of the genome (genomic areas of low recombination, e.g. genetic centromeres), as well as a better introgression of desirable traits in breeding programs (Phillips et al., 2010). Thus, research is underway for a better understanding of barley meiosis, which may in the future also allow manipulation of barley recombination (Barakate et al., 2014; Higgins et al., 2012; Phillips et al., 2012; Phillips et al., 2013). If recombination can be altered in barley, such methods may be also promising to be transferred to *H. bulbosum*, ILs or hybrids and to develop a more comprehensive set of *H. vulgare/H. bulbosum* ILs.

5. Conclusion & Outlook

Owing to the results of the presented thesis work, lack of molecular markers is no longer a limiting factor for the utilization of *H. bulbosum* introgression lines in barley breeding (Wendler et al., 2015; Wendler et al., 2014). The collection of well characterized ILs will provide a valuable community resource for barley breeding and research. By contrast, reduced homeologous recombination frequencies will remain a challenge for the broad utilization of these resources and will still complicate the transfer of advantageous *H. bulbosum* traits. However, several possible pathways have been discussed and suggested that may be suitable to either increase or avoid interspecific recombination. Genetic stocks of the IL library may provide a good source of material to further study homeologous recombination in barley. Nonetheless, interspecific recombination does occur and large populations can facilitate the identification of appropriate recombinants. For the purpose of gene mapping and isolation the possibility of rearranged genome between barley and *H. bulbosum* should always be considered and may

provide challenges. Thus, a high-quality genetic map of the *H. bulbosum* genome and comprehensive comparative genomics analysis between barley and *H. bulbosum* would be highly desirable for the near future. Finally, methods and pipelines that have been established during the present study will be highly promising for other crop-wild introgressions, especially if involving secondary and tertiary gene pool species.

6. References

Abbott, D.C., Brown, A.H.D., and Burdon, J.J. (1992). Genes for Scald Resistance from Wild Barley (*Hordeum vulgare* ssp. *spontaneum*) and Their Linkage to Isozyme Markers. Euphytica 61:225-231.

Al-Kaff, N., Knight, E., Bertin, I., Foote, T., Hart, N., Griffiths, S., and Moore, G. (2008). Detailed dissection of the chromosomal region containing the *Ph1* locus in wheat *Triticum aestivum*: With deletion mutants and expression profiling. Ann Bot-London 101:863-872.

Andolfatto, P., Davison, D., Erezyilmaz, D., Hu, T.T., Mast, J., Sunayama-Morita, T., and Stern, D.L. (2011). Multiplexed shotgun genotyping for rapid and efficient genetic mapping. Genome Res 21:610-617.

Arbelaez, J., Moreno, L., Singh, N., Tung, C.-W., Maron, L., Ospina, Y., Martinez, C., Grenier, C., Lorieux, M., and McCouch, S. (2015). Development and GBS-genotyping of introgression lines (ILs) using two wild species of rice, *O. meridionalis* and *O. rufipogon*, in a common recurrent parent, *O. sativa* cv. Curinga. Mol Breeding 35:1-18.

Ashley, T., Walpita, D., and de Rooij, D.G. (2001). Localization of two mammalian cyclin dependent kinases during mammalian meiosis. J Cell Sci 114:685-693.

Attene, G., Ceccarelli, S., and Papa, R. (1996). The barley (*Hordeum vulgare* L.) of Sardinia, Italy. Genet Resour Crop Evol 43:385-393.

Baek, H.J., Beharav, A., and Nevo, E. (2003). Ecological-genomic diversity of microsatellites in wild barley, *Hordeum spontaneum*, populations in Jordan. Theor Appl Genet 106:397-410.

Bainbridge, M.N., Wang, M., Burgess, D.L., Kovar, C., Rodesch, M.J., D'Ascenzo, M., Kitzman, J., Wu, Y.Q., Newsham, I., Richmond, T.A., et al. (2010). Whole exome capture in solution with 3 Gbp of data. Genome Biol 11:R62.

Bamshad, M.J., Ng, S.B., Bigham, A.W., Tabor, H.K., Emond, M.J., Nickerson, D.A., and Shendure, J. (2011). Exome sequencing as a tool for Mendelian disease gene discovery. Nat Rev Genet 12:745-755.

Barakate, A., Higgins, J.D., Vivera, S., Stephens, J., Perry, R.M., Ramsay, L., Colas, I., Oakey, H., Waugh, R., Franklin, F.C.H., et al. (2014). The Synaptonemal Complex Protein ZYP1 Is Required for Imposition of Meiotic Crossovers in Barley. The Plant Cell 26:729-740.

Baum, B.R., Nevo, E., Johnson, D.A., and Beiles, A. (1997). Genetic diversity in wild barley (*Hordeum spontaneum* C. Koch) in the Near East: A molecular analysis using random amplified polymorphic DNA (RAPD) markers. Genet Resour Crop Ev 44:147-157.

Becker, H. (1993). Pflanzenzüchtung. Stuttgart, Germany: Ulmer Verlag Eugen

Belhaj, K., Chaparro-Garcia, A., Kamoun, S., and Nekrasov, V. (2013). Plant genome editing made easy: targeted mutagenesis in model and crop plants using the CRISPR/Cas system. Plant Methods 9:39.

Blake, T., Blake, V., Bowman, J., and Abdel-Haleem, H. (2011). Barley Feed Uses and Quality Improvement. Oxford, UK: Wiley-Blackwell.

Blattner, F.R. (2009). Progress in phylogenetic analysis and a new infrageneric classification of the barley genus *Hordeum* (*Poaceae: Triticeae*). Breeding Sci 59:471–480.

Brinkman, M.A., and Frey, K.J. (1977). Yield component analysis of oat isolines that produce different grain yields. Crop Sci 17:165-168.

Bundock, P.C., Casu, R.E., and Henry, R.J. (2012). Enrichment of genomic DNA for polymorphism detection in a non-model highly polyploid crop plant. Plant Biotechnology Journal 10:657-667.

Canady, M.A., Ji, Y., and Chetelat, R.T. (2006). Homeologous Recombination in *Solanum lycopersicoides* Introgression Lines of Cultivated Tomato. Genetics 174:1775-1788.

Chen, P.D., Tsujimoto, H., and Gill, B.S. (1994). Transfer of *Ph(I)* Genes Promoting Homoeologous Pairing from *Triticum speltoides* to Common Wheat. Theor Appl Genet 88:97-101.

Cheng, J.K., and Alper, H.S. (2014). The genome editing toolbox: a spectrum of approaches for targeted modification. Current Opinion in Biotechnology 30:87-94.

Close, T.J., Bhat, P.R., Lonardi, S., Wu, Y., Rostoks, N., Ramsay, L., Druka, A., Stein, N., Svensson, J.T., Wanamaker, S., et al. (2009). Development and implementation of high-throughput SNP genotyping in barley. Bmc Genomics 10:582-582.

Coghlan, A., Eichler, E.E., Oliver, S.G., Paterson, A.H., and Stein, L. (2005). Chromosome evolution in eukaryotes: a multi-kingdom perspective. Trends Genet 21:673-682.

Comadran, J., Kilian, B., Russell, J., Ramsay, L., Stein, N., Ganal, M., Shaw, P., Bayer, M., Thomas, W., Marshall, D., et al. (2012). Natural variation in a homolog of *Antirrhinum CENTRORADIALIS* contributed to spring growth habit and environmental adaptation in cultivated barley. Nat Genet 44:1388 - 1392.

Damania, A.B. (2008). History, Achievements, and Current Status of Genetic Resources Conservation. Agron. J. 100:27-39.

Davey, J.W., Hohenlohe, P.A., Etter, P.D., Boone, J.Q., Catchen, J.M., and Blaxter, M.L. (2011). Genome-wide genetic marker discovery and genotyping using next-generation sequencing. Nat Rev Genet 12:499-510.

Dolezel, J., and Bartos, J. (2005). Plant DNA flow cytometry and estimation of nuclear genome size. Ann Bot-London 95:99-110.

Dunwell, J.M. (2000). Transgenic approaches to crop improvement. Journal of Experimental Botany 51:487-496.

Eglinton, J.K., Evans, D.E., Brown, A.H.D., Langridge, P., McDonald, G., Jefferies, S.P., and Barr, A.R. (1999). The use of wild barley (*Hordeum vulgare* ssp. *spontaneum*) in breeding for quality and adaptation. Proceedings of 9th Australian Barley Technical Symposium 1:1–6.

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:e19379.

Fedak, G. (1973). Increased Chiasma Frequency in Desynaptic Barley in Response to Phosphate Treatments. Can J Genet Cytol 15:647-649.

Fedak, G., and Armstrong, K.C. (1981). Cytogenetics of the Trigeneric Hybrid, (*Hordeum vulgare X Triticum aestivum*) X *Secale cereale*. Theor Appl Genet 60:215-219.

Fetch, T., Johnston, P.A., and Pickering, R. (2009). Chromosomal location and inheritance of stem rust resistance transferred from *Hordeum bulbosum* into cultivated barley (*H. vulgare*). Phytopathology 99:339-343.

Fetch, T.G., Steffenson, B.J., and Nevok, E. (2003). Diversity and sources of multiple disease resistance in *Hordeum spontaneum*. Plant Dis 87:1439-1448.

Feuillet, C., Langridge, P., and Waugh, R. (2008). Cereal breeding takes a walk on the wild side. Trends Genet 24:24-32.

Feuillet, C., and Muehlbauer, G.J. (2009). Genetics and Genomics of the Triticeae. Berlin, Germany: Springer Science.

Fischer, G., Rocha, E.P.C., Brunet, F., Vergassola, M., and Dujon, B. (2006). Highly Variable Rates of Genome Rearrangements between Hemiascomycetous Yeast Lineages. PLoS Genet 2:e32.

Food and Agricultural Organization of the United Nations (FAO). (1996). The Global Plan of Action fot the Conservation and Sustainable Use of Plant Genetic Resources for Food And Agriculture. Rome, Italy: Food and Agricultural Organization of the United Nations.

Food and Agricultural Organization of the United Nations (FAO). (2011). The Second Global Plan of Action fot the Conservation and Sustainable Use of Plant Genetic Resources for Food And Agriculture. Rome, Italy: Food and Agriculture Organization of the United Nations

Forster, B.P., Ellis, R.P., Thomas, W.T.B., Newton, A.C., Tuberosa, R., This, D., El-Enein, R.A., Bahri, M.H., and Ben Salem, M. (2000). The development and application of molecular markers for abiotic stress tolerance in barley. Journal of Experimental Botany 51:19-27.

Gilbert, G.S., and Webb, C.O. (2007). Phylogenetic signal in plant pathogen-host range. P Natl Acad Sci USA 104:4979-4983.

Goedeke, S., Hensel, G., Kapusi, E., Gahrtz, M., and Kumlehn, J. (2007). Transgenic Barley in Fundamental Research and Biotechnology. Transgenic Plant Journal 1:104-117.

Gopher, A., Abbo, S., and Lev-Yadun, S. (2002). The 'when', the 'where' and the 'why' of the Neolithic revolution in the Levant. Documenta Praehistorica 28:49–62.

Gore, M., Bradbury, P., Hogers, R., Kirst, M., Verstege, E., van Oeveren, J., Peleman, J., Buckler, E., and van Eijk, M. (2007). Evaluation of Target Preparation Methods for Single-Feature Polymorphism Detection in Large Complex Plant Genomes. Crop Sci. 47:135-148.

Gore, M.A., Chia, J.-M., Elshire, R.J., Sun, Q., Ersoz, E.S., Hurwitz, B.L., Peiffer, J.A., McMullen, M.D., Grills, G.S., Ross-Ibarra, J., et al. (2009). A First-Generation Haplotype Map of Maize. Science 326:1115-1117.

Gorny, A.G. (2001). Variation in utilization efficiency and tolerance to reduced water and nitrogen supply among wild and cultivated barleys. Euphytica 117:59-66.

Grando, S., and Macpherson, H.G. (2005). Food barley: importance, uses and local knowledge. In: Proceedings of the International Workshop on Food Barley Improvement, Hammamet, Tunisia, 14-17 January, 2002 Aleppo, Syria: International Center for Agricultural Research in the Dry Areas (ICARDA).

Griffiths, S., Sharp, R., Foote, T.N., Bertin, I., Wanous, M., Reader, S., Colas, I., and Moore, G. (2006). Molecular characterization of *Ph1* as a major chromosome pairing locus in polyploid wheat. Nature 439:749-752.

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:e92046.

Hajjar, R., and Hodgkin, T. (2007). The use of wild relatives in crop improvement: a survey of developments over the last 20 years. Euphytica 156:1-13.

Harlan, J.R., and de Wet, J.M.J. (1971). Toward a rational classification of cultivated plants. Taxon 20:509–517.

Hassan, L., and Jones, R.N. (1995). Colchicine-Induced Heritable Modification of Chiasma Frequency and Distribution in Perennial Ryegrass (*Lolium perenne*). Genome 38:356-361.

Haussmann, B.I.G., Parzies, H.K., Presterl, T., Susic, Z., and Miedaner, T. (2004). Plant genetic resources in crop improvement. Plant Genetic Resources 2:3–21.

Hawkes, J.G. (1994). Origins of cultivated potatoes and species relationships. In: Potato Genetics--Bradshaw, J.E., and Mackay, G.R., eds. Wallingford, UK: CAB International. 3-42.

Higgins, J.D., Perry, R.M., Barakate, A., Ramsay, L., Waugh, R., Halpin, C., Armstrong, S.J., and Franklin, F.C.H. (2012). Spatiotemporal Asymmetry of the Meiotic Program Underlies the Predominantly Distal Distribution of Meiotic Crossovers in Barley. The Plant Cell 24:4096-4109.

Hirochika, H., Guiderdoni, E., An, G., Hsing, Y.-i., Eun, M., Han, C.-d., Upadhyaya, N., Ramachandran, S., Zhang, Q., Pereira, A., et al. (2004). Rice Mutant Resources for Gene Discovery. Plant Mol Biol 54:325-334.

Hoffmann, A.A., Sgro, C.M., and Weeks, A.R. (2004). Chromosomal inversion polymorphisms and adaptation. Trends Ecol Evol 19:482-488.

Honsdorf, N., March, T., Hecht, A., Eglinton, J., and Pillen, K. (2014). Evaluation of juvenile drought stress tolerance and genotyping by sequencing with wild barley introgression lines. Mol Breeding 34:1475-1495.

Hu, Q., Hansen, L.N., Laursen, J., Dixelius, C., and Andersen, S.B. (2002). Intergeneric hybrids between *Brassica napus* and *Orychophragmus violaceus* containing traits of agronomic importance for oilseed rape breeding. Theor Appl Genet 105:834-840.

Intergovernmental Panel on Climate Change (IPCC). (2013). Climate change 2013: the physical science basis. In: Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change Cambridge, UK: Cambridge University Press.

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

Islam, A.K.M.R., and Shepherd, K.W. (1992). Production of wheat-barley recombinant chromosomes through induced homoeologous pairing. Theor Appl Genet 83:489-494.

IUPAC-IUB Commission on Biochemical Nomenclature (CBN). (1970). Abbreviations and symbols for nucleic acids, polynucleotides, and their constituents. European Journal of Biochemistry 15:203-208.

Ivandic, V., Walther, U., and Graner, A. (1998). Molecular mapping of a new gene in wild barley conferring complete resistance to leaf rust (*Puccinia hordei* Otth). Theor Appl Genet 97:1235-1239.

Jaffe, B., Caligari, P.D.S., and Snape, J.W. (2000). A skeletal linkage map of *Hordeum bulbosum* L. and comparative mapping with barley *(H. vulgare* L.). Euphytica 115:115-120.

Jakob, S.S., and Blattner, F.R. (2006). A chloroplast genealogy of *Hordeum* (*Poaceae*): Long-term persisting haplotypes, incomplete lineage sorting, regional extinction, and the consequences for phylogenetic inference. Molecular Biology and Evolution 23:1602-1612.

Jana, S., and Nevo, E. (1991). Variation in Response to Infection with *Erysiphe graminis hordei* and *Puccinia hordei* in Some Wild Barley Populations in a Center of Diversity. Euphytica 57:133-140.

Ji, Q., Xu, X., and Wang, K. (2013). Genetic transformation of major cereal crops. Int J Dev Biol 57:495-508.

Johnston, P., Meiyalaghan, V., Forbes, M., Habekuß, A., Butler, R., and Pickering, R. (2015). Marker assisted separation of resistance genes *Rph22* and *Rym16^{Hb}* from an associated yield penalty in a barley: *Hordeum bulbosum* introgression line. Theor Appl Genet 128:1137-1149.

Johnston, P.A. (2007). Molecular characterisation of chromatin introgressed from *Hordeum bulbosum* L. into *Hordeum vulgare* L. In: Department of Biochemistry Dunedin, New Zealand: University of Otago.

Johnston, P.A., Niks, R.E., Meiyalaghan, V., Blanchet, E., and Pickering, R. (2013). *Rph22*: mapping of a novel leaf rust resistance gene introgressed from the non-host *Hordeum bulbosum* L. into cultivated barley (*Hordeum vulgare* L.). Theor Appl Genet 126:1613-1625.

Johnston, P.A., and Pickering, R.A. (2002). PCR detection of *Hordeum bulbosum* introgressions in an *H. vulgare* background using a retrotransposon-like sequence. Theor Appl Genet 104:720-726.

Johnston, P.A., Timmerman-Vaughan, G.M., Farnden, K.J.F., and Pickering, R. (2009). Marker development and characterisation of *Hordeum bulbosum* introgression lines: a resource for barley improvement. Theor Appl Genet 118:1429-1437.

Jones, R.N., and Rees, H. (1967). Genotypic Control of Chromosome Behaviour in Rye. 11. Influence of B Chromosomes on Meiosis. Heredity 22:333–347.

Jørgensen, I.H. (1992a). Discovery, characterization and exploitation of *Mlo* powdery mildew resistance in barley. Euphytica 63:141-152.

Jørgensen, J.H. (1992b). Sources and genetics of resistance to fungal pathogens. In: Barley: Genetics, Biochemistry, Molecular Biology and Biotechnology--Shewry, P.R., ed. Oxford, UK: Alden Press. 442–457.

Jupe, F., Witek, K., Verweij, W., Śliwka, J., Pritchard, L., Etherington, G.J., Maclean, D., Cock, P.J., Leggett, R.M., Bryan, G.J., et al. (2013). Resistance gene enrichment sequencing (RenSeq) enables reannotation of the *NB-LRR* gene family from sequenced plant genomes and rapid mapping of resistance loci in segregating populations. The Plant Journal 76:530-544.

Kamenetzky, L., Asís, R., Bassi, S., de Godoy, F., Bermúdez, L., Fernie, A.R., Van Sluys, M.-A., Vrebalov, J., Giovannoni, J.J., Rossi, M., et al. (2010). Genomic Analysis of Wild Tomato Introgressions Determining Metabolism- and Yield-Associated Traits. Plant Physiol 152:1772-1786.

Kao, K.N., Constabe.F, Michaylu.Mr, and Gamborg, O.L. (1974). Plant Protoplast Fusion and Growth of Intergeneric Hybrid Cells. Planta 120:215-227.

Kasha, K.J., Kao, K.N., and Reinberg.E. (1970). Genetic Control over Chromosome Stability in Hybrids from Interspecific *Hordeum* Crosses. Genetics 64:263–275.

Kasha, K.J., and Sadasiva, R. (1971). Genome Relationship between *Hordeum vulgare* L. and *H. bulbosum* L. Chromosoma 35:264-287.

Kim, N.S., Fedak, G., Han, F., and Cao, W. (2008). Cytogenetic analyses of intergeneric hybrids between barley and nine species of *Elymus*. Genome 51:897-904.

Kisaka, H., and Kameya, T. (1998). Cold and salt tolerance of somatic hybrid calli between barley *(Hordeum vulgare* L.) and carrot (*Daucus carota* L.). Breeding Science 48:11-15.

Kisaka, H., Kisaka, M., Kanno, A., and Kameya, T. (1998). Intergeneric somatic hybridization of rice (*Oryza sativa* L) and barley (*Hordeum vulgare* L) by protoplast fusion. Plant Cell Rep 17:362-367.

Knight, E., Greer, E., Draeger, T., Thole, V., Reader, S., Shaw, P., and Moore, G. (2010). Inducing chromosome pairing through premature condensation: analysis of wheat interspecific hybrids. Funct Integr Genomic 10:603-608.

Korzun, V., Malyshev, S., Pickering, R.A., and Borner, A. (1999). RFLP mapping of a gene for hairy leaf sheath using a recombinant line from *Hordeum vulgare* L. x *Hordeum bulbosum* L. cross. Genome 42:960-963.

Kumar, J., Choudhary, A.K., Solanki, R.K., and Pratap, A. (2011). Towards marker-assisted selection in pulses: a review. Plant Breeding 130:297-313.

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.

Labate, J.A., Robertson, L.D., Strickler, S.R., and Mueller, L.A. (2014). Genetic structure of the four wild tomato species in the *Solanum peruvianum* s.l. species complex. Genome 57:169-180.

Lachance, J., and Tishkoff, S.A. (2013). SNP ascertainment bias in population genetic analyses: Why it is important, and how to correct it. BioEssays: news and reviews in molecular, cellular and developmental biology 35:doi:10.1002/bies.201300014.

Lehmann, L. (1991). The use of genetic resources for isolating disease resistance for barley cultivar development. Proceedings 6th Barley Genetics Symposium 1:650–652.

Lev-Yadun, S., Gopher, A., and Abbo, S. (2000). Archaeology - The cradle of agriculture. Science 288:1602-1603.

Levings, C.S. (1990). The Texas Cytoplasm of Maize: Cytoplasmic Male Sterility and Disease Susceptibility. Science 250:942-947.

Lhuissier, F.G.P., Offenberg, H.H., Wittich, P.E., Vischer, N.O.E., and Heyting, C. (2007). The mismatch repair protein MLH1 marks a subset of strongly interfering crossovers in tomato. Plant Cell 19:862-876.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., and 1000 Genome Project Data Processing Subgroup. (2009). The Sequence Alignment/Map format and SAMtools. Bioinformatics 25:2078-2079.

Li, L., Jean, M., and Belzile, F. (2006). The impact of sequence divergence and DNA mismatch repair on homeologous recombination in Arabidopsis. The Plant Journal 45:908-916.

Lorence, A., and Verpoorte, R. (2004). Gene Transfer and Expression in Plants. In: Recombinant Gene Expression--Balbás, P., and Lorence, A., eds.: Humana Press. 329-350.
Lukaszewski, A.J., Lapinski, B., and Rybka, K. (2005). Limitations of in situ hybridization with total genomic DNA in routine screening for alien introgressions in wheat. Cytogenet Genome Res 109:373-377.
Lundquist, U. (2009). Eighty years of scandinavian barley mutation genetics and breeding. In: Induced plant Mutations in the Genomics Era--Shu, Q.Y., ed. Rome, Italy: Food and Agriculture Organization of the United Nations (FAO). 39-43.
Lundqvist, A. (1962). Self-incompatibility in diploid *Hordeum bulbosum* L. Hereditas 48:138-152.
Luo, M.-C., Thomas, C., You, F.M., Hsiao, J., Ouyang, S., Buell, C.R., Malandro, M., McGuire, P.E., Anderson, O.D., and Dvorak, J. (2003). High-throughput fingerprinting of bacterial artificial chromosomes using the snapshot labeling kit and sizing of restriction fragments by capillary electrophoresis. Genomics 82:378-389.
Ma, H. (2006). A Molecular Portrait of Arabidopsis Meiosis. The Arabidopsis Book / American Society of Plant Biologists 4:e0095.
Mammadov, J.A., Chen, W., Ren, R.H., Pai, R., Marchione, W., Yalcin, F., Witsenboer, H., Greene, T.W., Thompson, S.A., and Kumpatla, S.P. (2010). Development of highly polymorphic SNP markers from the complexity reduced portion of maize [*Zea mays* L.] genome for use in marker-assisted breeding. Theor Appl Genet 121:577-588.
Mano, Y., and Takeda, K. (1998). Genetic resources of salt tolerance in wild *Hordeum* species. Euphytica 103:137-141.
Mardis, E. (2010). Cancer genomics identifies determinants of tumor biology. Genome Biol 11:211.
Martin, A.C., Shaw, P., Phillips, D., Reader, S., and Moore, G. (2014). Licensing MLH1 sites for crossover during meiosis. Nat Commun 5:4580.
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 Biol 15:R78.
Mascher, M., Muehlbauer, G.J., Rokhsar, D.S., Chapman, J., Schmutz, J., Barry, K., Muñoz-Amatriain, M., Close, T.J., Wise, R.P., Schulman, A.H., et al. (2013a). Anchoring and ordering NGS contig assemblies by population sequencing (POPSEQ). Plant J 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., et al. (2013b). Barley whole exome capture: a tool for genomic research in the genus *Hordeum* and beyond. Plant J 76:494-505.
Mascher, M., Wu, S., Amand, P., Stein, N., and Poland, J. (2013c). Application of Genotyping-by-Sequencing on Semiconductor Sequencing Platforms: A Comparison of Genetic and Reference-Based Marker Ordering in Barley. Plos One 8:e76925.
Matzke, M.A., Matzke, A.J.M., Pruss, G.J., and Vance, V.B. (2001). RNA-based silencing strategies in plants. Current Opinion in Genetics & Development 11:221-227.
Mayer, K.F.X., Martis, M., Hedley, P.E., Simkova, H., Liu, H., Morris, J.A., Steuernagel, B., Taudien, S., Roessner, S., Gundlach, H., et al. (2011). Unlocking the Barley Genome by Chromosomal and Comparative Genomics. Plant Cell 23:1249-1263.
McHale, L., Tan, X., Koehl, P., and Michelmore, R. (2006). Plant NBS-LRR proteins: adaptable guards. Genome Biol 7:212.
Menda, N., Strickler, S.R., Edwards, J.D., Bombarely, A., Dunham, D.M., Martin, G.B., Mejia, L., Hutton, S.F., Havey, M.J., Maxwell, D.P., et al. (2014). Analysis of wild-species introgressions in tomato inbreds uncovers ancestral origins. Bmc Plant Biology 14:287.
Mendiondo, G.M., Gibbs, D.J., Szurman-Zubrzycka, M., Korn, A., Marquez, J., Szarejko, I., Maluszynski, M., King, J., Axcell, B., Smart, K., et al. (2015). Enhanced waterlogging tolerance in barley by manipulation of expression of the N-end rule pathway E3 ligase PROTEOLYSIS6. Plant Biotechnology Journal:doi:10.1111/pbi.12334.

Mercier, R., and Grelon, M. (2008). Meiosis in plants: ten years of gene discovery. Cytogenet Genome Res 120:281-290.
Meyers, B.C., Scalabrin, S., and Morgante, M. (2004). Mapping and sequencing complex genomes: let's get physical! Nat Rev Genet 5:578-588.
Mézard, C., Tagliaro Jahns, M., and Grelon, M. (2015). Where to cross? New insights into the location of meiotic crossovers. Trends Genet:doi:10.1016/j.tig.2015.1003.1008.
Moore, A.M.T., Hillman, G.C., and Legge, A.J. (2000). Village on the Euphrates, from Foraging to Farming at Abu Hureyra. Oxford, UK: Oxford Univ. Press.
Moragues, M., Comadran, J., Waugh, R., Milne, I., Flavell, A.J., and Russell, J. (2010). Effects of ascertainment bias and marker number on estimations of barley diversity from high-throughput SNP genotype data. Theor Appl Genet 120:1525-1534.
Morishige, D.T., Childs, K.L., Moore, L.D., and Mullet, J.E. (2002). Targeted Analysis of Orthologous Phytochrome A Regions of the Sorghum, Maize, and Rice Genomes using Comparative Gene-Island Sequencing. Plant Physiol 130:1614-1625.
Moseman, J.G., Nevo, E., and Elmorshidy, M.A. (1990). Reactions of *Hordeum spontaneum* to Infection with 2 Cultures of *Puccinia hordei* from Israel and United States. Euphytica 49:169-175.
Munoz-Amatriain, M., Eichten, S., Wicker, T., Richmond, T., Mascher, M., Steuernagel, B., Scholz, U., Ariyadasa, R., Spannagl, M., Nussbaumer, T., et al. (2013). Distribution, functional impact, and origin mechanisms of copy number variation in the barley genome. Genome Biol 14:R58.
Murphy, W.J., Larkin, D.M., der Wind, A.E.-v., Bourque, G., Tesler, G., Auvil, L., Beever, J.E., Chowdhary, B.P., Galibert, F., Gatzke, L., et al. (2005). Dynamics of Mammalian Chromosome Evolution Inferred from Multispecies Comparative Maps. Science 309:613-617.
National Research Council. (1972). Genetic Vulnerability of Major Crops. Washington, US: National Academy of Sciences.
Nevo, E. (1992). Origin, evolution, population genetics and resources for breeding of wild barley, Hordeum spontaneum, in the Fertile Crescent Wallingford, UK: CAB International. 19-43.
Nevo, E. (1998). Genetic diversity in wild cereals: Regional and local studies and their bearing on conservation ex situ and in situ. Genet Resour Crop Ev 45:355-370.
Nevo, E., Beiles, A., Gutterman, Y., Storch, N., and Kaplan, D. (1984). Genetic Resources of Wild Cereals in Israel and Vicinity. 2. Phenotypic Variation within and between Populations of Wild Barley, *Hordeum spontaneum*. Euphytica 33:737-756.
Nevo, E., Fu, Y.B., Pavlicek, T., Khalifa, S., Tavasi, M., and Beiles, A. (2012). Evolution of wild cereals during 28 years of global warming in Israel. P Natl Acad Sci USA 109:3412-3415.
Niu, Z., Klindworth, D.L., Friesen, T.L., Chao, S., Jin, Y., Cai, X., and Xu, S.S. (2011). Targeted Introgression of a Wheat Stem Rust Resistance Gene by DNA Marker-Assisted Chromosome Engineering. Genetics 187:1011-1021.
O'Malley, R.C., and Ecker, J.R. (2010). Linking genotype to phenotype using the Arabidopsis unimutant collection. The Plant Journal 61:928-940.
Page, S.L., and Hawley, R.S. (2003). Chromosome choreography: The meiotic ballet. Science 301:785-789.
Pfender, W.F., Saha, M.C., Johnson, E.A., and Slabaugh, M.B. (2011). Mapping with RAD (restriction-site associated DNA) markers to rapidly identify QTL for stem rust resistance in *Lolium perenne*. Theor Appl Genet 122:1467-1480.
Pflieger, S., Lefebvre, V., and Causse, M. (2001). The candidate gene approach in plant genetics: a review. Mol Breeding 7:275-291.
Phillips, D., Nibau, C., Ramsay, L., Waugh, R., and Jenkins, G. (2010). Development of a Molecular Cytogenetic Recombination Assay for Barley. Cytogenet Genome Res 129:154-161.
Phillips, D., Nibau, C., Wnetrzak, J., and Jenkins, G. (2012). High Resolution Analysis of Meiotic Chromosome Structure and Behaviour in Barley (*Hordeum vulgare* L.). Plos One 7:e39539.

Phillips, D., Wnetrzak, J., Nibau, C., Barakate, A., Ramsay, L., Wright, F., Higgins, J.D., Perry, R.M., and Jenkins, G. (2013). Quantitative high resolution mapping of HvMLH3 foci in barley pachytene nuclei reveals a strong distal bias and weak interference. Journal of Experimental Botany 64:2139-2154.

Phipps, R.H., and Park, J.R. (2002). Environmental Benefits of Genetically Modified Crops: Global and European Perspectives on Their Ability to Reduce Pesticide Use. Journal of Animal and Feed Sciences 11:1-18.

Pickering, R. (2000). Do the wild relatives of cultivated barley have a place in barley improvement? Adelaide, Australia: Barley Genetics Conference.

Pickering, R., and Johnston, P.A. (2005). Recent progress in barley improvement using wild species of *Hordeum*. Cytogenet Genome Res 109:344-349.

Pickering, R., Niks, R.E., Johnston, P.A., and Butler, R.C. (2004). Importance of the Secondary Genepool in Barley Genetics and Breeding II. Disease Resistance, Agronomic Performance and Quality. Czech J. Genet. Plant Breed 40:79–85.

Pickering, R., Ruge-Wehling, B., Johnston, P.A., Schweizer, G., Ackermann, P., and Wehling, P. (2006). The transfer of a gene conferring resistance to scald (*Rhynchosporium secalis*) from *Hordeum bulbosum* into *H. vulgare* chromosome 4HS. Plant Breeding 125:576-579.

Pickering, R.A. (1984). The Influence of Genotype and Environment on Chromosome Elimination in Crosses between *Hordeum vulgare* L. X *Hordeum bulbosum* L. Plant Sci Lett 34:153-164.

Pickering, R.A. (1991a). Comparison of crossover frequencies in barley (*Hordeum vulgare*) and *H. vulgare* x *H. bulbosum* hybrids using a paracentric inversion. Genome Biol 34:666-673.

Pickering, R.A. (1991b). The production of fertile triploid hybrids from crosses between *Hordeum vulgare* L. (2n=4x=28) and *H. bulbosum* L. (2n=2x=14). Hereditas 114:227-236.

Pickering, R.A., and Devaux, P. (1992). Haploid production: approaches and use in plant breeding. Wallingford, UK: CAB International.

Pickering, R.A., Hill, A.M., Michel, M., and Timmerman-Vaughan, G.M. (1995). The transfer of a powdery mildew resistance gene from *Hordeum bulbosum* L. to barley (*H. vulgare* L.) chromosome 2 (2L). Theor Appl Genet 91:1288-1292.

Pickering, R.A., Malyshev, S., Kunzel, G., Johnston, P.A., Korzun, V., Menke, M., and Schubert, I. (2000). Locating introgressions of *Hordeum bulbosum* chromatin within the *H. vulgare* genome. Theor Appl Genet 100:27-31.

Pickering, R.A., Rennie, W.F., and Cromey, M.G. (1987). Disease resistant material available from the wide hybridization programme at DSIR. Barley Newsletter 31:248–250.

Pickering, R.A., Timmerman, G.M., Cromey, M.G., and Melz, G. (1994). Characterization of Progeny from Backcrosses of Triploid Hybrids between *Hordeum vulgare* L (2x) and *H. bulbosum* L. (4x) to *Hordeum vulgare*. Theor Appl Genet 88:460-464.

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.

Poland, J.A., and Rife, T.W. (2012). Genotyping-by-Sequencing for Plant Breeding and Genetics. Plant Gen. 5:92-102.

Prescott-Allen, C., and Prescott-Allen, R. (1986). The first resource: Wild species in the North American economy. New Haven, US: Yale University Press.

Preuss, D., and Copenhaver, G.P. (2000). Chemical and physical treatment that stimulate recombination. University of Chicago:WO/2000/054574.

Qi, L.L., Friebe, B., Zhang, P., and Gill, B.S. (2007). Homoeologous recombination, chromosome engineering and crop improvement. Chromosome Research 15:3-19.

R Core Team. (2012). R: A Language and Environment for Statistical Computing Vienna: R Foundation for Statistical Computing.

Rabinowicz, P.D., Citek, R., Budiman, M.A., Nunberg, A., Bedell, J.A., Lakey, N., O'Shaughnessy, A.L., Nascimento, L.U., McCombie, W.R., and Martienssen, R.A. (2005). Differential methylation of genes and repeats in land plants. Genome Res 15:1431-1440.

Rabinowicz, P.D., Palmer, L.E., May, B.P., Hemann, M.T., Lowe, S.W., McCombie, W.R., and Martienssen, R.A. (2003). Genes and Transposons Are Differentially Methylated in Plants, but Not in Mammals. Genome Res 13:2658-2664.

Rajaram, S., and Hettel, G.P. (1994). Wheat Breeding at CIMMYT: Commemorating 50 Years of Research in Mexico for Global Wheat Improvement. Mexico: CIMMYT.

Ray, S., and Satya, P. (2014). Next generation sequencing technologies for next generation plant breeding. Frontiers in Plant Science 5: 367.

Reed, G.H., and Wittwer, C.T. (2004). Sensitivity and specificity of single-nucleotide polymorphism scanning by high-resolution melting analysis. Clin Chem 50:1748-1754.

Richter, T.E., and Ronald, P.C. (2000). The evolution of disease resistance genes. Plant Mol Biol 42:195-204.

Riley, R., and Chapman, V. (1958). Genetic Control of the Cytologically Diploid Behaviour of Hexaploid Wheat. Nature 182:713-715.

Ruge-Wehling, B., Linz, A., Habekuss, A., and Wehling, P. (2006). Mapping of *Rym16*(Hb), the second soil-borne virus-resistance gene introgressed from *Hordeum bulbosum*. Theor Appl Genet 113:867-873.

Ruge, B., Linz, A., Pickering, R., Proeseler, G., Greif, P., and Wehling, P. (2003). Mapping of *Rym14* (Hb), a gene introgressed from *Hordeum bulbosum* and conferring resistance to BaMMV and BaYMV in barley. Theor Appl Genet 107:965-971.

Russell, J.R., Ellis, R.P., Thomas, W.T.B., Waugh, R., Provan, J., Booth, A., Fuller, J., Lawrence, P., Young, G., and Powell, W. (2000). A retrospective analysis of spring barley germplasm development from 'foundation genotypes' to currently successful cultivars. Mol Breeding 6:553-568.

Sahu, B.B., Sumit, R., Srivastava, S.K., and Bhattacharyya, M.K. (2012). Sequence based polymorphic (SBP) marker technology for targeted genomic regions: its application in generating a molecular map of the *Arabidopsis thaliana* genome. Bmc Genomics 13:20.

Saisho, D., and Takeda, K. (2011). Barley: emergence as a new research material of crop science. Plant and Cell Physiology 52:724–727.

Salamini, F., Özkan, H., Brandolini, A., Schäfer-Pregl, R., and Martin, W. (2002). Genetics and geography of wild cereal domestication in the Near East. Nat Rev Genet 3:429-441.

Salvo-Garrido, H., Laurie, D.A., Jaffé, B., and Snape, J.W. (2001). An RFLP map of diploid *Hordeum bulbosum* L. and comparison with maps of barley (*H. vulgare* L.) and wheat (*Triticum aestivum* L.). Theor Appl Genet 103:869-880.

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

Sato, K., and Takeda, K. (1997). Net blotch resistance in wild species of *Hordeum*. Euphytica 95:179-185.

Sato, K.F., A. Russell J. Börner, A. Valkoun J. (2014). Genetic Diversity and Germplasm Management: Wild Barley, Landraces, Breeding Material. In: Biotechnological Approaches to Barley Improvement Berlin, Germany: Springer.

Scannell, D.R., Byrne, K.P., Gordon, J.L., Wong, S., and Wolfe, K.H. (2006). Multiple rounds of speciation associated with reciprocal gene loss in polyploid yeasts. Nature 440:341-345.

Schnaithmann, F., and Pillen, K. (2013). Detection of exotic QTLs controlling nitrogen stress tolerance among wild barley introgression lines. Euphytica 189:67-88.

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

Scholz, M., Ruge-Wehling, B., Habekuss, A., Schrader, O., Pendinen, G., Fischer, K., and Wehling, P. (2009). *Ryd4* (Hb): a novel resistance gene introgressed from *Hordeum bulbosum* into barley and

conferring complete and dominant resistance to the barley yellow dwarf virus. Theor Appl Genet 119:837-849.
Schulte, D., Ariyadasa, R., Shi, B., Fleury, D., Saski, C., Atkins, M., deJong, P., Wu, C.-C., Graner, A., Langridge, P., et al. (2011). BAC library resources for map-based cloning and physical map construction in barley (*Hordeum vulgare* L.). Bmc Genomics 12:247.
Schulte, D., Close, T.J., Graner, A., Langridge, P., Matsumoto, T., Muehlbauer, G., Sato, K., Schulman, A.H., Waugh, R., Wise, R.P., et al. (2009). The International Barley Sequencing Consortium-At the Threshold of Efficient Access to the Barley Genome. Plant Physiol 149:142-147.
Sears, E.R. (1956). The transfer of leaf rust resistance from *Aegilops umbellulata* to wheat. Brookhaven Symp. Biol. 9:1-21.
Sears, E.R. (1981). Transfer of alien genetic material to wheat. In: Wheat Science – Today and Tomorrow--Evans, L.T., and Peacock, W.J., eds. Cambridge, UK: Cambridge University Press. 75–89.
Sears, E.R., and Okamoto, M. (1958). Intergenomic chromosome relationships in hexaploid wheat. Proc. Int. Cong. Genet. 2:258–259.
Shtaya, M.J.Y., Sillero, J.C., Flath, K., Pickering, R., and Rubiales, D. (2007). The resistance to leaf rust and powdery mildew of recombinant lines of barley (*Hordeum vulgare* L.) derived from *H. vulgare* x *H. bulbosum* crosses. Plant Breeding 126:259-267.
Simmonds, N.W. (1962). Variability in crop plants: its use and conservation. Biological Reviews 37:422-465.
Singh, A.K., Rossnagel, B.G., Scoles, G.J., and Pickering, R.A. (2004). Identification of a quantitatively inherited source of *Hordeum bulbosum* derived scald resistance from barley line 926K2/11/1/5/1. Can J Plant Sci 84:935-938.
Soderlund, C., Humphray, S., Dunham, A., and French, L. (2000). Contigs Built with Fingerprints, Markers, and FPC V4.7. Genome Res 10:1772-1787.
Somers, D.A., Narayanan, K.R., Kleinhofs, A., Cooperbland, S., and Cocking, E.C. (1986). Immunological Evidence for Transfer of the Barley Nitrate Reductase Structural Gene to Nicotiana-Tabacum by Protoplast Fusion. Molecular & General Genetics 204:296-301.
Stadler, L.J. (1930). Some Genetic Effects of X-Rays in Plants. J. Hered 21:3-19.
Stam, P., and Zeven, A.C. (1981). The theoretical proportion of the donor genome in near-isogenic lines of self-fertilizers bred by backcrossing. Euphytica 30:227–238.
Stein, N., and Graner, A. (2005). Map-Based Gene Isolation in Cereal Genomes. In: Cereal Genomics--Gupta, P.K., and Varshney, R.K., eds. Netherlands: Springer Netherlands. 331-360.
Stein, N., Prasad, M., Scholz, U., Thiel, T., Zhang, H.N., Wolf, M., Kota, R., Varshney, R.K., Perovic, D., Grosse, I., et al. (2007). A 1,000-loci transcript map of the barley genome: new anchoring points for integrative grass genomics. Theor Appl Genet 114:823-839.
Svetlanov, A., and Cohen, P.E. (2004). Mismatch repair proteins, meiosis, and mice: understanding the complexities of mammalian meiosis. Experimental Cell Research 296:71-79.
Szigat, G., and Pohler, W. (1982). *Hordeum bulbosum* x *H. vulgare* hybrids and their backcrosses with cultivated barley. Cereal Res Commun 10:73–78.
Szigat, G., and Szigat, G. (1991). Amphidiploid hybrids between *Hordeum vulgare* and *H. bulbosum* - basis for the development of new initial material for winter barley breeding. In: Vortr Pflanzenzüchtg--Winkel, A., ed. Wageningen, Netherlands: EUCARPIA. 34–39.
Tanksley, S.D., and McCouch, S.R. (1997). Seed banks and molecular maps: Unlocking genetic potential from the wild. Science 277:1063-1066.
Teer, J.K., Bonnycastle, L.L., Chines, P.S., Hansen, N.F., Aoyama, N., Swift, A.J., Abaan, H.O., Albert, T.J., Program, N.C.S., Margulies, E.H., et al. (2010). Systematic comparison of three genomic enrichment methods for massively parallel DNA sequencing. Genome Res 20:1420-1431.

Teer, J.K., and Mullikin, J.C. (2010). Exome sequencing: the sweet spot before whole genomes. Human Molecular Genetics 19:R145-R151.

The International Wheat Genome Sequencing Consortium. (2014). A chromosome-based draft sequence of the hexaploid bread wheat (Triticum aestivum) genome. Science 345:doi:10.1126/science.1251788.

The Tomato Genome Consortium. (2012). The tomato genome sequence provides insights into fleshy fruit evolution. Nature 485:635-641.

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 Evol Biol 9:209.

Thomas, H.M., and Pickering, R.A. (1985). The influence of parental genotype on the chromosome behaviour of *Hordeum vulgare × H. bulbosum* diploid hybrids. Theor Appl Genet 71:437–442.

Tingay, S., McElroy, D., Kalla, R., Fieg, S., Wang, M., Thornton, S., and Brettell, R. (1997). *Agrobacterium tumefaciens*-mediated barley transformation. The Plant Journal 11:1369-1376.

Toubia-Rahme, H., Johnston, P.A., Pickering, R.A., and Steffenson, B.J. (2003). Inheritance and chromosomal location of *Septoria passerinii* resistance introgressed from *Hordeum bulbosum* into *Hordeum vulgare*. Plant Breeding 122:405-409.

Turpeinen, T., Tenhola, T., Manninen, O., Nevo, E., and Nissila, E. (2001). Microsatellite diversity associated with ecological factors in *Hordeum spontaneum* populations in Israel. Mol Ecol 10:1577-1591.

Ullrich SE, B., B.K., Quinde-Axtell, Z., Nair, S. (2008). Barley for food: traits and improvements. In: Proceedings of the 10th International Barley Genetics Symposium--Ceccarelli, A., and Grando, S., eds. 563-575.

Van Rijn, C.P.E., Heersche, I., Van Berkel, Y.E.M., Nevo, E., Lambers, H., and Poorter, H. (2000). Growth characteristics in *Hordeum spontaneum* populations from different habitats. New Phytol 146:471-481.

Vavilov, N.I. (1940). The new systematic of cultivated plants. In: The New Systematics--Huxley, J., ed. Oxford, UK: Clareon Press.

Verstegen, H., Köneke, O., Korzun, V., and von Broock, R. (2014). The World Importance of Barley and Challanges to Further Improvements. In: Biotechnological Approaches to Barley Improvement Berlin, Germany: Springer.

von Bothmer, R., Flink, J., Jacoben, N., Kotimäki, M., and Landström, T. (1983). Interspecific hybridization with cultivated barley (*Hordeum vulgare* L.). Hereditas 99:219–244.

von Bothmer, R., Jacobsen, N., Baden, C., Jorgensen, R.B., and Linde-Laursen, I. (1995). An ecogeographical study of the genus *Hordeum* (2nd edition). Rome, Italy: International Plant Genetic Resources Institute.

von Bothmer, R., Sato, R., Knüpffer, H., and van Hintum, T. (2003). Diversification through breeding. In: Diversity in Barley Amsterdam, Netherlands: Elsevier Science. 29-52.

Walther, U., Rapke, H., Proeseler, G., and Szigat, G. (2000). *Hordeum bulbosum* - a new source of disease resistance - transfer of resistance to leaf rust and mosaic viruses from *H. bulbosum* into winter barley. Plant Breeding 119:215-218.

Wang, Z., Gerstein, M., and Snyder, M. (2009). RNA-Seq: a revolutionary tool for transcriptomics. Nat Rev Genet 10:57-63.

Wendler, N., Mascher, M., Himmelbach, A., Johnston, P., Pickering, R., and Stein, N. (2015). *Bulbosum* to go: a toolbox to utilize *Hordeum vulgare/ bulbosum* introgressions for breeding and beyond. Molecular Plant 128:1137-1149.

Wendler, N., Mascher, M., Nöh, C., Himmelbach, A., Scholz, U., Ruge-Wehling, B., and Stein, N. (2014). Unlocking the secondary gene-pool of barley with next-generation sequencing. Plant Biotechnology Journal 12:1122-1131.

Wijnker, E., and de Jong, H. (2008). Managing meiotic recombination in plant breeding. Trends in Plant Science 13:640-646.
Wulff, B.B.H., and Moscou, M.J. (2014). Strategies for transferring resistance into wheat: from wide crosses to GM cassettes. Frontiers in Plant Science 5:doi:10.3389/fpls.2014.00692.
Xu, J., and Kasha, K.J. (1992). Transfer of a dominant gene for powdery mildew resistance and DNA from *Hordeum bulbosum* into cultivated barley (*Hordeum vulgare*). Theor Appl Genet 84:771-777.
Yamashita, K., Yasuda, H., Pines, J., Yasumoto, K., Nishitani, H., Ohtsubo, M., Hunter, T., Sugimura, T., and Nishimoto, T. (1990). Okadaic Acid, a Potent Inhibitor of Type-1 and Type-2a Protein Phosphatases, Activates Cdc2/H1 Kinase and Transiently Induces a Premature Mitosis-Like State in Bhk-21-Cells. Embo J 9:4331-4338.
Yu, Y., Tomkins, J.P., Waugh, R., Frisch, D.A., Kudrna, D., Kleinhofs, A., Brueggeman, R.S., Muehlbauer, G.J., Wise, R.P., and Wing, R.A. (2000). A bacterial artificial chromosome library for barley (Hordeum vulgare L.) and the identification of clones containing putative resistance genes. Theor Appl Genet 101:1093-1099.
Zamir, D. (2001). Improving plant breeding with exotic genetic libraries. Nat Rev Genet 2:983-989.
Zhang, L.T., Pickering, R., and Murray, B. (1999). Direct measurement of recombination frequency in interspecific hybrids between *Hordeum vulgare* and *H. bulbosum* using genomic in situ hybridization. Heredity 83:304-309.
Zheng, T., Nibau, C., Phillips, D.W., Jenkins, G., Armstrong, S.J., and Doonan, J.H. (2014). CDKG1 protein kinase is essential for synapsis and male meiosis at high ambient temperature in Arabidopsis thaliana. P Natl Acad Sci USA 111:2182-2187.
Zhu, T., Chang, H.-S., Schmeits, J., Gil, P., Shi, L., Budworth, P., Zou, G., Chen, X., and Wang, X. (2001). Gene Expression Microarrays: Improvements and Applications towards Agricultural Gene Discovery. Journal of the Association for Laboratory Automation 6:95-98.

7. Summary

Crop wild relatives have been recognized as a source of beneficial traits to a given crop species and to overcome the erosion of genetic diversity resulting from domestication and small effective population sizes in elite breeding programs. *Hordeum bulbosum* L. is the only representative of the secondary gene pool of barley (*Hordeum vulgare* L.) and it has been found as a valuable source of genetic diversity for barley improvement, especially regarding to pathogen resistance or tolerance. Since the 1990s a considerable number of barley/*H. bulbosum* introgression lines (ILs) have been generated, with introgressed *H. bulbosum* segments replacing different regions of the *H. vulgare* genome. These ILs harbor a diverse set of desirable traits. So far, the efficient utilization of such ILs for improving elite barley germplasm has been hampered, largely due to the lack of suitable molecular tools for locating introgressed segments and reduced interspecific recombination frequencies. Furthermore, advantageous traits of *H. bulbosum* can often only be transferred at the cost of transferring also negative traits in addition, which are associated with the introgressed *H. bulbosum* segments. Therefore, for the utilization of *H. bulbosum* traits in barley breeding the accurate genetic characterization of barley/*H. bulbosum* ILs and the availability of molecular markers is the prerequisite. Thus, the focus of the present study was to establish suitable and state-of-the-art molecular tools and resources for the efficient utilization of *H. bulbosum* traits in barley breeding and research.

Towards this aim, I explored the usefulness of two next generation sequencing technologies for developing and scoring molecular markers in diploid barley/*H. bulbosum* ILs. A recently developed whole exome capture assay in combination with a custom Singe-Nucleotide-Polymorphisms (SNP) genotyping assay as well as two-enzyme genotyping-by-sequencing (GBS) were used to allocate the introgression interval of a barley/*H. bulbosum* IL and to genotype progeny segregating for the introgression. Both methods provided fast and reliable detection and mapping of the introgressed segment and enabled the identification of recombinant plants avoiding tedious and iterative steps of marker development.

After prove of concept, GBS was applied to characterize 146 barley/*H. bulbosum* ILs by looking at *H. bulbosum*-specific SNP frequencies along the barley reference. *H. bulbosum* segments could be identified in all but one of the analyzed ILs. Furthermore, *H. bulbosum* introgressions were localized on all chromosomal arms and even very small (< 1 Mbp) introgressed fragments could be detected and localized. A crossing strategy, tackling the problem of reduced interspecific

recombination frequency for mapping of *H. bulbosum* traits in ILs, by exploiting the here generated polymorphisms, information has been proposed and needs to be validated in future studies.

Furthermore, exome capture re-sequencing data of five *H. bulbosum* accessions and 13 *H. vulgare* cultivars was utilized to design an integrated *H. vulgare/H. bulbosum* sequence resource. This resource contains defined polymorphism information on 112 847 proposed interspecifically conserved *H. vulgare/H. bulbosum* sequence variations, all of which have been anchored to the barley reference sequence. This integrated sequence resource will greatly faciltate future marker development in ILs derived from any *H. vulgare* and *H. bulbosum* donors.

The here established methods and pipelines, the comprehensive catalogue of precisely characterized ILs and the integrated *H. vulgare/H. bulbosum* sequence will be an excellent community resource and of major benefit for the wide-spread utilization of barley/*H. bulbosum* ILs in applied barley breeding as well as in basic research.

8. Zusammenfassung

Die genetische Diversität von wilden Verwandten der Kulturpflanzen können genutzt werden um Elitematerial mit vorteilhaften Merkmalen zu bereichern und um dabei dem Verlust an Diversität durch Züchtung entgegenzuwirken. *Hordeum bulbosum* L. repräsentiert als einziges den sekundären Genpool von Kulturgerste (*Hordeum vulgare* L.). Sie birgt daher wertvolles genetisches Potential für die Gerstenzüchtung, besonders im Hinblick auf ihre Widerstandfähigkeit gegenüber wichtigen Krankheitserregern der Kulturgerste. Seit den 1990er Jahren wurde eine beträchtliche Anzahl an Gerste/*H. bulbosum* Introgressionslinen (ILs) generiert, bei denen kleinere Chromosomenabschnitte von *H. bulbosum* an unterschiedlichen Stellen des Gerstengenomes integriert wurden. Diese ILs beinhalten viele wünschenswerte *H. bulbos*um Eigenschaften. Bis jetzt wurde die effiziente Nutzung dieser ILs für die Gerstenzüchtung hauptsächlich durch das Fehlen an geeigneten molekularen Methoden für die genetische Charakterisierung der Introgression behindert. Verminderte interspezifische Rekombinationshäufigkeiten stellen ein zusätzliches Hindernis dar. Des Weiteren sind die vorteilhaften *H. bulbosum* Gene in den ILs meist mit negativen *H. bulbosum* Eigenschaften gekoppelt, die sich auf dem gleichen Introgressionsstück befinden. Eine detaillierte genetische Charakterisierung und das Vorhandensein von molekularen Markern ist daher Voraussetzung für die Nutzbarmachung vorteilhafter Eigenschaften von *H. bulbosum*. Das Hauptziel dieser Arbeit war daher, die Etablierung modernster molekularer Methoden und Ressourcen für eine effiziente Nutzbarmachung von *H. bulbosum* Merkmalen in der Gerstenzüchtung und Forschung.

Zu diesem Zweck wurde zunächst die Anwendbarkeit zweier auf next generation sequencing basierenden Technologien für die Entwicklung von molekularen Markern in diploiden ILs evaluiert. Ein kürzlich entwickeltes Exome Capture Assay in Kombination mit einem selbsterstellten Singe-Nucleotide-Polymorphisms (SNP) Genotypisierungs Assay als auch genotyping-by-sequencing (GBS) basierend auf zwei Restriktionsenzymen wurden hierbei genutzt um die Introgressionsintervalle von ILs zu bestimmen und segregierende Nachkommenschaften dieser ILs zu genotypisieren. Beide Methoden ermöglichten dabei die schnelle und akkurate Lokalisation der integrierten Segmente als auch die Detektion der rekombinanten Nachkommen.

Nach der Etablierung der beiden Methoden, wurde GBS genutzt um 146 bereits vorhandene ILs anhand von *H. bulbosum*-spezifische SNP Frequenzen entlang der Gerstenreferenz zu charakterisieren. Dabei konnte in 145 ILs *H. bulbosum* Introgressionen nachgewiesen werden. *H. bulbosum* Segmente wurden auf allen Gerstenchromosomen Armen detektiert und selbst sehr kleine Segmente (< 1Mbp) konnten dabei lokalisiert werden. Ein möglicher Lösungsansatz für das Problem der reduzierten interspezifischen Rekombinationshäufigkeit bei der Kartierung von *H. bulbosum* Merkmalen, der auf den hier generierten SNP – und Sequenzinformationen basiert, wurde vorgestellt. Der Ansatz muss in zukünftigen Studien getestet werden.

Exome capture Daten von fünf *H. bulbosum* Akzessionen und 13 *H. vulgare* Kultivaren wurden für das Design einer integrierten *H. vulgare/H. bulbosum* Sequenz Ressource genutzt. Diese Ressource beinhaltet Informationen über 112 847 mögliche konservierter interspezifischer *H. vulgare/H. bulbosum* Polymorphismen, von denen alle in der Gerstenreferenz verankert sind. Es ist wahrscheinlich, dass diese integrierte Sequenz eine umfangreiche Ressource für die Entwicklung molekularer Marker in ILs bieten wird, unabhängig von den genutzten *H. vulgare* und *H. bulbosum* Spender-Genotypen.

Die hier etablierten Methoden und Pipelines, der umfangreiche Katalog an charakterisierten ILs und die integrierte *H. vulgare/H. bulbosum* Sequenz werden eine hervorragende gemeinnützige Ressource bilden und dabei die effiziente Nutzbarmachung von *H. vulgare/H. bulbosum* ILs in der Gerstenzüchtung sowie in der Forschung ermöglichen.

9. Acknowledgments

First and foremost, I offer my sincerest gratitude to my supervisor Dr. Nils Stein (PI of Genome Diversity group at Leibniz Institute of Plant Genetic and Crop Plant Research (IPK), Gatersleben) for giving me the opportunity to work in the Plant 2030 - TRANSBULB project (financed by the German Ministry of Education and Research, BMBF). I deeply appreciate his outstanding scientific knowledge and everlasting enthusiasm. His patient way of teaching and continued constructive criticism, while giving me the freedom for self-development have been essential as a foundation for completing my PhD dissertation.

Moreover, I would like to thank my supervisor Prof. Dr. Andreas Graner (Director of IPK, Gatersleben) for his great interest in my work. I am very grateful for his profound discussions, fruitful suggestions and consideration, which have greatly supported the progression of my PhD project. I also acknowledge Dr. Frank Blattner (PI of Experimental Taxonomy group at IPK, Gatersleben) for being the scientific mentor during my PhD study. His excellent discussions and consideration have markedly benefitted the PhD project. I appreciate Prof. Dr. Peter Wehling, Head of the Julius Kühn Institute for Breeding Research on Agricultural Crops for agreeing as my supervisor to evaluate this thesis.

I owe my earnest gratitude to my colleague Dr. Martin Mascher (PI of Domestication Genomics group IPK, Gatersleben), for sharing his great expertise in many helpful discussions and especially for his support and guidance for analyzing the excessive amount of next generation sequencing data that was generated during this project.

I also want to thank all my kind colleagues in the Genome Diversity group for creating a pleasant atmosphere and for many nice discussions and suggestions. I especially thank Dr. Axel Himmelbach for sharing his notable technical expertise and his support for next generation sequencing. I am very grateful to have shared this exciting PhD time with my colleagues and friends Matthias Jost and Mingjiu Li. They have provided me with great support in many scientific and non-scientific aspects. I truly acknowledge Dr. Ping Yang, Dr. Ruonan Zhou and Dr. Andriy Kochevenko for their valuable support and suggestions.

I also would like to take this opportunity to convey my sincere and wholehearted thanks to all our collaborators. I am grateful for the excessive support and discussions and all previous research

work by Dr. Brigitte Ruge-Wehling (Julius Kühn-Institut (JKI), Groß Lüsewitz), Dr. Richard Pickering and Dr. Paul Johnston (New Zealand Institute for Plant and Food Research Limited, Christchurch, New Zealand), which has created the basis for the actual PhD project. I especially thank Christiane Nöh (JKI, Groß Lüsewitz) for the intensive and fruitful collaboration, for handling a major proportion of plant material and for many helpful discussions.

I greatly acknowledge the excellent technical assistance from Manuela Knauft, Susanne Koenig, Ines Walde, Mary Ziems, Jacqueline Pohl, Jelena Perovic and Sandra Driesslein, who had made my project running smoothly. I am also grateful to the gardeners, for carefully looking after my plant material in the greenhouse.

Finally, I want to express my sincere gratitude to my family for their understanding, patience and continuous support.

10. Abbreviations

CA	California
%	Percent
˚C	Celsius
µg	Microgram
µL	Microliter
µM	Micromolar
ADT	Array Design Tool
ATP	Adenosine triphosphate
B	Haploid *H. bulbosum* genome
BAC	Bacterial Artificial Chromosome
BaMMV	barley mild mosaic virus
BaYMV	Barley Yellow Mosaic Virus
BLAST	Basic Local Alignment Search Tool
BMBF	Bundesministerium für Bildung und Forschung
BOPA	Barley Oligonucleotide Pool Assays
bp	Base pairs
BRW	Brigitte Ruge-Wegling
BSA	Bovine Serum Albumin
BWA	Burrows-Wheeler Aligner
CBN	Commission on Biochemical Nomenclature
Cdk	Cyclin-dependent kinases
cDNA	Complementary DNA
cm	Centimeter
cM	Centimorgan
CNV	Copy number variations
CTAB	Cetyltrimethylammonium bromide
cv.	Cultivar
CWR	Crop wild relatives
DH	double haploid
DH	Doubled Haploid
DNA	Deoxyribonucleic acid
dNTP	Desoxyribonukleosidtriphosphate
EB	Elution buffer
EC	Exome capture
EST	Expressed Sequence Tags
F_1	First generation
F_2	Second filial generation
F_6	Sixth filial generation
F_7	Seventh filial generation
FAO	Food and Agricultural Organization of the United Nations
g	Gram

Gbp/Gb	Giga base pairs
GBS	Genotyping by sequencing
GC score	Gen Call score
GGA	Golden Gate assay
GP	Golden Promise
H	*Hordeum*
Hb	*Hordeum bulbosum*
HC	High-confidence
Het	Heterozygous
HF	High Fidelity
HRM	High resolution melting
Hv	*Hordeum vulgare*
Hv^b	*Hordeum vulgare/ H. bulbosum* introgression line
IBSC	International Barley Genome Sequencing Consortium
ID	Identifier
IL	Introgression line
IPCC	Intergovernmental Panel on Climate Change
IPK	Leibniz Institute of Plant Genetics and Crop Plant Research
IUPAC	International Union of Pure and Applied Chemistry
JKI	Julius Kühn-Institut
Kbp/Kb	Kilo base pairs
LC	Low-confidence
LM-PCR	Ligation-mediated PCR
LOD	Logarithm of Odds
LR	leaf rust
m	Meter
M	Molar
MA	Massachusetts
Mbp/Mb	Mega base pairs
mg	Milligram
mil	Powdery mildew
min	Minute(s)
mM	Millimolar
mRNA	Messenger RNA
MTP	Minimal Tiling Path
mya	million years ago
NB-LRR	Nucleotide-binding leucine-rich repeat domain-containing
NCBI	National Center for Biotechnology Information
ng	Nanogram
NGS	Next Generation Sequencing
nm	Nanometer
nM	Nanomolar
NS	Nils Stein

NZ	New Zealand
Oligo	Oligonucleotide
PCR	Polymerase Chain Reaction
Ph1	*Pairing homoeologous1*
POPSEQ	Population Sequencing
PTO bond	Phosphorothioate bond
qPCR	QuantitativePCR
RAD-seq	Restriction-site associated DNA sequencing
RenSeq	Resistance gene enrichment Sequencing
RFLP	Restriction Fragment Length Polymorphisms
RIL	Recombinant Inbred Line
RNA	Ribonucleic acid
RNA-seq	RNA sequencing
s	Second(s)
S	Solanum
S	susceptible
scald	*Rynchosporium commune*
SNP	Single Nucleotide Polymorphism
SPRI	Solid phase reversible immobilization
SR	steam rust
SSLB	Septoria speckled leaf blotch
SSR	Simple Sequence Repeats
STS	Sequence Tagged Site
TALEN	Transcription Activator-like Effector Nuclease
U	Unit
US	United States
V	Haploid *H. vulgare* genome
V	Voltage
W	Watt
WGS	Whole Genome Shotgun

11. Supplementary Data

Wendler at al. 2014

Additional Supporting information may be found in the online version of this article:

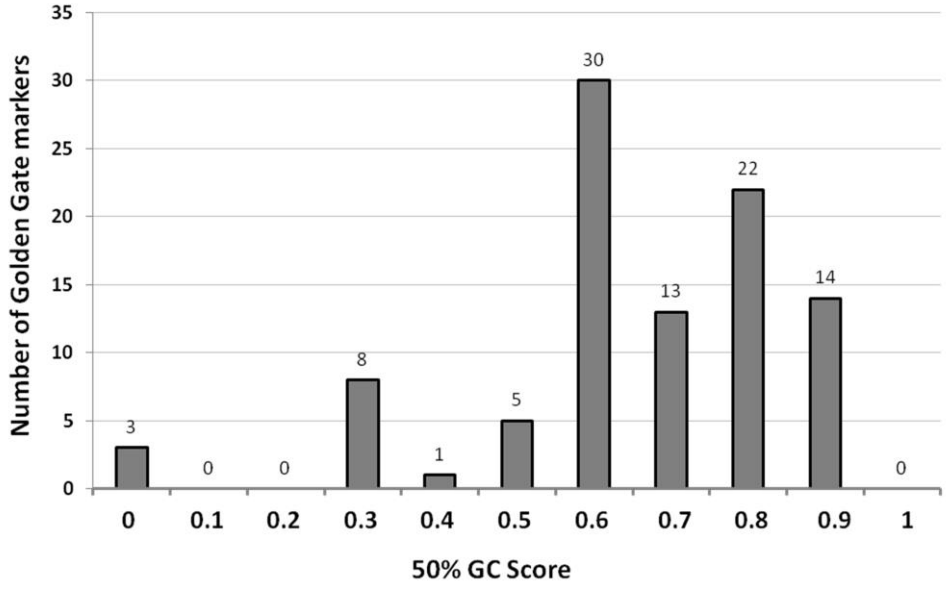

Figure S 1 Histogram of the 50% GenCall (GC) score of the 96 Golden Gate SNP markers, obtained by genotyping in the 2HL *Hv*^{b} mapping population.

Figure S 2 GBS adapter, primer for indexing PCR and Illumina sequencing**.**

Table S 1 List of BAC contigs.

BAC_contig	contig size (Mb)	chromosome	cM
contig_1952	1,7	2	136,0482
contig_14658	0,1	2	138,4561
contig_2926	2,1	2	138,5977
contig_11240	0,13	2	140,7932
contig_9	0,3	2	140,7932
contig_10477	0,3	2	140,7932
contig_46638	0,2	2	141,5014
contig_494	0,6	2	141,5014
contig_46362	0,4	2	141,5014
contig_9745	0,2	2	141,5722
contig_45579	0,5	2	141,7847
contig_44808	1,5	2	142,3513
contig_38793	0,1	2	143,1303
contig_42233	1,2	2	144,2635
contig_39614	0,8	2	146,2465
contig_6372	0,5	2	146,3881
contig_501	0,6	2	146,5297
contig_45091	0,3	2	146,5297
contig_44195	1,9	2	147,3088
contig_44567	0,4	2	147,7337
contig_45937	0,6	2	148,1586
contig_45938	0,6	2	148,1586
contig_45939	0,4	2	148,1586
contig_43474	2,1	2	148,1586

Table S 2 List of Golden Gate SNP markers.

Reference contig	SNP position	Genotype introgression*	Genotype bulbosum*	Genotype borwina*	Reference allele	Alternate allele	Barley chromo-some	cM	MB	Golden Gate marker
morex_contig_40945	18485	0	2	0	T	C	2	93,7	559	GGA_11
morex_contig_135636	2230	0	2	0	G	A	2	94,9	563	GGA_24
morex_contig_1562222	1174	0	2	0	G	A	2	96,1	564	GGA_95
morex_contig_55088	5613	0	2	0	T	C	2	99,3	568	GGA_94
morex_contig_1561819	4162	0	2	0	T	C	2	100,4	569	GGA_96
morex_contig_56924	3161	0	2	0	G	A	2	102,9	571	GGA_14
morex_contig_50835	4191	2	2	0	T	C	2	104,6	572	GGA_7
morex_contig_40851	2774	0	0	2	G	A	2	107,2	576	GGA_64
morex_contig_39371	6164	2	2	0	A	G	2	110,9	585	GGA_1
morex_contig_37648	2495	2	1	0	T	C	2	111,7	585	GGA_2

morex_contig_54147	234	2	2	0	T	G	2	113,5	585	GGA_3
morex_contig_37410	2617	2	1	0	T	G	2	113,9	586	GGA_10
morex_contig_50365	3258	2	2	0	A	G	2	113,9	585	GGA_12
morex_contig_40862	13298	2	1	0	A	G	2	113,9	585	GGA_13
morex_contig_1558026	6732	2	2	0	C	T	2	113,9	588	GGA_4
morex_contig_275706	939	2	1	0	G	A	2	113,9	588	GGA_6
morex_contig_46919	2469	2	1	0	G	C	2	113,9	588	GGA_8
morex_contig_275706	4111	0	0	2	C	T	2	113,9	588	GGA_86
morex_contig_70314	2065	2	2	0	C	G	2	113,9	588	GGA_9
morex_contig_37468	2211	2	2	0	C	T	2	114,1	589	GGA_15
morex_contig_1593008	2995	2	2	0	C	T	2	114,1	588	GGA_16
morex_contig_40876	2333	2	2	0	T	C	2	114,1	588	GGA_17
morex_contig_6807	4275	2	1	0	C	T	2	114,1	588	GGA_18
morex_contig_61715	3255	2	2	0	T	C	2	114,1	588	GGA_5
morex_contig_137011	3660	2	2	0	C	A	2	114,2	589	GGA_19
morex_contig_137075	4084	2	2	0	A	T	2	114,9	589	GGA_20
morex_contig_52733	1212	0	2	0	A	G	2	118,1	589	GGA_35
morex_contig_45476	1765	0	2	0	A	G	2	118,7	591	GGA_83
morex_contig_49247	6423	2	2	0	G	A	2	119,8	592	GGA_22
morex_contig_56633	5780	0	2	0	T	C	2	119,8	592	GGA_76
morex_contig_56583	10275	0	2	0	C	G	2	120,7	594	GGA_77
morex_contig_38530	2951	0	2	0	T	C	2	123,9	598	GGA_78
morex_contig_1564718	2303	0	2	0	A	T	2	125,2	600	GGA_79
morex_contig_57338	2010	0	2	0	T	C	2	127,2	602	GGA_59
morex_contig_44060	2196	0	2	0	C	T	2	129,7	604	GGA_80
morex_contig_1561585	4081	0	2	0	G	A	2	132,6	606	GGA_81
morex_contig_275606	4659	2	2	0	G	A	2	137,0	612	GGA_21
morex_contig_368802	4945	2	2	0	A	G	2	137,6	612	GGA_23
morex_contig_1568600	722	2	2	0	C	A	2	138,6	614	GGA_25
morex_contig_106745	3127	2	2	0	T	C	2	138,6	614	GGA_26
morex_contig_39549	5303	2	2	0	T	C	2	138,6	614	GGA_28
morex_contig_43637	5327	2	2	0	C	T	2	138,6	614	GGA_29
morex_contig_1566005	1647	2	1	0	A	G	2	138,6	612	GGA_31
morex_contig_2522914	310	2	2	0	A	G	2	138,6	612	GGA_32
morex_contig_276884	2246	2	2	0	T	A	2	140,3	614	GGA_33

morex_contig_134768	2847	2	1	0	A	G	2	140,3	614	GGA_34
morex_contig_8868	2280	2	2	0	T	C	2	140,3	614	GGA_36
morex_contig_50394	6190	2	1	0	C	A	2	140,5	614	GGA_27
morex_contig_50394	880	2	2	0	A	G	2	140,5	614	GGA_30
morex_contig_6903	2865	2	2	0	G	T	2	140,7	614	GGA_37
morex_contig_47733	7337	2	2	0	G	A	2	140,8	615	GGA_38
morex_contig_136074	2012	2	2	0	A	T	2	140,8	614	GGA_39
morex_contig_2553533	1032	2	1	0	A	G	2	140,8	614	GGA_52
morex_contig_2548179	13796	2	2	0	C	T	2	140,9	615	GGA_40
morex_contig_37165	4647	2	1	0	G	A	2	140,9	615	GGA_41
morex_contig_135472	1830	2	2	0	T	A	2	141,0	616	GGA_42
morex_contig_2547841	3736	2	2	0	G	A	2	141,0	616	GGA_43
morex_contig_6586	3000	2	2	0	C	G	2	141,0	616	GGA_44
morex_contig_40976	3165	2	1	0	G	T	2	141,1	616	GGA_45
morex_contig_53846	2346	2	2	0	A	T	2	141,6	616	GGA_46
morex_contig_104818	308	2	2	0	G	A	2	142,1	616	GGA_69
morex_contig_1560652	8173	2	2	0	G	C	2	142,3	616	GGA_47
morex_contig_2552114	1345	2	2	0	C	A	2	142,4	618	GGA_48
morex_contig_2552817	5942	2	1	0	G	A	2	142,4	618	GGA_49
morex_contig_274895	2386	2	2	0	A	G	2	142,4	618	GGA_50
morex_contig_47457	3554	2	1	0	C	T	2	142,4	618	GGA_51
morex_contig_51188	2066	2	1	0	T	G	2	142,4	618	GGA_53
morex_contig_44257	9743	2	1	0	G	A	2	142,4	618	GGA_54
morex_contig_43340	4471	2	2	0	T	C	2	142,4	618	GGA_84
morex_contig_43982	4598	2	1	0	G	C	2	144,1	618	GGA_55
morex_contig_1562165	2776	2	2	0	A	G	2	144,3	620	GGA_56
morex_contig_1563321	2446	2	2	0	T	C	2	144,3	620	GGA_57
morex_contig_43759	6854	2	1	0	T	G	2	144,3	620	GGA_58
morex_contig_41725	3650	2	2	0	T	C	2	145,6	620	GGA_60
morex_contig_438554	1092	2	1	0	C	G	2	145,6	620	GGA_61
morex_contig_42008	3889	2	2	0	T	C	2	145,6	621	GGA_62
morex_contig_135277	6524	2	2	0	T	A	2	146,5	621	GGA_63
morex_contig_9839	844	2	2	0	T	C	2	146,5	621	GGA_65
morex_contig_158172	2012	2	2	0	T	C	2	146,9	623	GGA_66
morex_contig_41279	976	2	2	0	C	A	2	146,9	622	GGA_67

morex_contig_37285	2324	2	2	0	A	G	2	148,4	623	GGA_68
morex_contig_45528	1469	2	1	0	A	C	2	149,3	626	GGA_85
morex_contig_44962	3125	2	2	0	A	G	2	149,3	626	GGA_70
morex_contig_453346	8458	2	2	0	G	A	2	149,3	626	GGA_71
morex_contig_45528	5739	2	2	0	T	C	2	149,3	626	GGA_72
morex_contig_1570648	594	2	2	0	G	A	2	149,3	626	GGA_82
morex_contig_41516	11578	2	2	0	G	A	2	149,4	627	GGA_73
morex_contig_1579900	1570	2	2	0	G	A	2	149,4	626	GGA_74
morex_contig_55146	956	2	2	0	A	G	2	149,4	627	GGA_75
morex_contig_44687	2329	2	2	0	T	C	2	149,4	628	GGA_87
morex_contig_38887	11979	2	1	0	G	A	2	149,4	626	GGA_88
morex_contig_1564940	961	2	2	0	T	C	2	149,4	628	GGA_89
morex_contig_158546	1861	2	2	0	A	G	2	149,4	627	GGA_90
morex_contig_511440	1664	2	2	0	C	T	2	149,4	627	GGA_91
morex_contig_55742	1306	2	2	0	G	A	2	149,4	628	GGA_92
morex_contig_7601	6980	2	2	0	T	G	2	149,4	628	GGA_93

* 0 = reference allele; 2 = alternative allele; 1 = heterozygous

Table S 3 Table of GBS indices.

Oligo ID	Oligo sequence (5'-3') (index marked with small letters)	Index sequence (5'-3')
indexing7001	CAAGCAGAAGACGGCATACGAGATcctgcgaGTGACTGGAGTTCAGACGTGT	TCGCAGG
indexing7002	CAAGCAGAAGACGGCATACGAGATtgcagagGTGACTGGAGTTCAGACGTGT	CTCTGCA
indexing7003	CAAGCAGAAGACGGCATACGAGATacctaggGTGACTGGAGTTCAGACGTGT	CCTAGGT
indexing7004	CAAGCAGAAGACGGCATACGAGATttgatccGTGACTGGAGTTCAGACGTGT	GGATCAA
indexing7005	CAAGCAGAAGACGGCATACGAGATatcttgcGTGACTGGAGTTCAGACGTGT	GCAAGAT
indexing7006	CAAGCAGAAGACGGCATACGAGATtctccatGTGACTGGAGTTCAGACGTGT	ATGGAGA
indexing7007	CAAGCAGAAGACGGCATACGAGATcatcgagGTGACTGGAGTTCAGACGTGT	CTCGATG
indexing7008	CAAGCAGAAGACGGCATACGAGATttcgagcGTGACTGGAGTTCAGACGTGT	GCTCGAA
indexing7009	CAAGCAGAAGACGGCATACGAGATagttggtGTGACTGGAGTTCAGACGTGT	ACCAACT
indexing7010	CAAGCAGAAGACGGCATACGAGATgtaccggGTGACTGGAGTTCAGACGTGT	CCGGTAC
indexing7011	CAAGCAGAAGACGGCATACGAGATcggagttGTGACTGGAGTTCAGACGTGT	AACTCCG
indexing7012	CAAGCAGAAGACGGCATACGAGATacttcaaGTGACTGGAGTTCAGACGTGT	TTGAAGT
indexing7013	CAAGCAGAAGACGGCATACGAGATtgatagtGTGACTGGAGTTCAGACGTGT	ACTATCA
indexing7014	CAAGCAGAAGACGGCATACGAGATgatccaaGTGACTGGAGTTCAGACGTGT	TTGGATC
indexing7015	CAAGCAGAAGACGGCATACGAGATcaggtcgGTGACTGGAGTTCAGACGTGT	CGACCTG
indexing7016	CAAGCAGAAGACGGCATACGAGATcgcattaGTGACTGGAGTTCAGACGTGT	TAATGCG
indexing7017	CAAGCAGAAGACGGCATACGAGATggtacctGTGACTGGAGTTCAGACGTGT	AGGTACC
indexing7018	CAAGCAGAAGACGGCATACGAGATggacgcaGTGACTGGAGTTCAGACGTGT	TGCGTCC
indexing7019	CAAGCAGAAGACGGCATACGAGATgagattcGTGACTGGAGTTCAGACGTGT	GAATCTC

indexing7020	CAAGCAGAAGACGGCATACGAGATgagcatgGTGACTGGAGTTCAGACGTGT	CATGCTC
indexing7021	CAAGCAGAAGACGGCATACGAGATgttgcgtGTGACTGGAGTTCAGACGTGT	ACGCAAC
indexing7022	CAAGCAGAAGACGGCATACGAGATccaatgcGTGACTGGAGTTCAGACGTGT	GCATTGG
indexing7023	CAAGCAGAAGACGGCATACGAGATcgagatcGTGACTGGAGTTCAGACGTGT	GATCTCG
indexing7024	CAAGCAGAAGACGGCATACGAGATcatattgGTGACTGGAGTTCAGACGTGT	CAATATG
indexing7025	CAAGCAGAAGACGGCATACGAGATgacgtcaGTGACTGGAGTTCAGACGTGT	TGACGTC
indexing7026	CAAGCAGAAGACGGCATACGAGATtggcatcGTGACTGGAGTTCAGACGTGT	GATGCCA
indexing7027	CAAGCAGAAGACGGCATACGAGATgtaattgGTGACTGGAGTTCAGACGTGT	CAATTAC
indexing7028	CAAGCAGAAGACGGCATACGAGATcctatctGTGACTGGAGTTCAGACGTGT	AGATAGG
indexing7029	CAAGCAGAAGACGGCATACGAGATcaatcggGTGACTGGAGTTCAGACGTGT	CCGATTG
indexing7030	CAAGCAGAAGACGGCATACGAGATgcggcatGTGACTGGAGTTCAGACGTGT	ATGCCGC
indexing7031	CAAGCAGAAGACGGCATACGAGATagtactgGTGACTGGAGTTCAGACGTGT	CAGTACT
indexing7032	CAAGCAGAAGACGGCATACGAGATtactattGTGACTGGAGTTCAGACGTGT	AATAGTA
indexing7033	CAAGCAGAAGACGGCATACGAGATccggatgGTGACTGGAGTTCAGACGTGT	CATCCGG
indexing7034	CAAGCAGAAGACGGCATACGAGATaccatgaGTGACTGGAGTTCAGACGTGT	TCATGGT
indexing7035	CAAGCAGAAGACGGCATACGAGATcggttctGTGACTGGAGTTCAGACGTGT	AGAACCG
indexing7036	CAAGCAGAAGACGGCATACGAGATtattccaGTGACTGGAGTTCAGACGTGT	TGGAATA
indexing7037	CAAGCAGAAGACGGCATACGAGATcctcctgGTGACTGGAGTTCAGACGTGT	CAGGAGG
indexing7038	CAAGCAGAAGACGGCATACGAGATaggtattGTGACTGGAGTTCAGACGTGT	AATACCT
indexing7039	CAAGCAGAAGACGGCATACGAGATgcattcgGTGACTGGAGTTCAGACGTGT	CGAATGC
indexing7040	CAAGCAGAAGACGGCATACGAGATttgcgaaGTGACTGGAGTTCAGACGTGT	TTCGCAA
indexing7041	CAAGCAGAAGACGGCATACGAGATttgaattGTGACTGGAGTTCAGACGTGT	AATTCAA
indexing7042	CAAGCAGAAGACGGCATACGAGATctgcgcgGTGACTGGAGTTCAGACGTGT	CGCGCAG
indexing7043	CAAGCAGAAGACGGCATACGAGATagaccttGTGACTGGAGTTCAGACGTGT	AAGGTCT
indexing7044	CAAGCAGAAGACGGCATACGAGATgtccagtGTGACTGGAGTTCAGACGTGT	ACTGGAC
indexing7045	CAAGCAGAAGACGGCATACGAGATacctgctGTGACTGGAGTTCAGACGTGT	AGCAGGT
indexing7046	CAAGCAGAAGACGGCATACGAGATccggtacGTGACTGGAGTTCAGACGTGT	GTACCGG
indexing7047	CAAGCAGAAGACGGCATACGAGATcttgaccGTGACTGGAGTTCAGACGTGT	GGTCAAG
indexing7048	CAAGCAGAAGACGGCATACGAGATcatcattGTGACTGGAGTTCAGACGTGT	AATGATG
indexing7049	CAAGCAGAAGACGGCATACGAGATtctgactGTGACTGGAGTTCAGACGTGT	AGTCAGA
indexing7050	CAAGCAGAAGACGGCATACGAGATtctagttGTGACTGGAGTTCAGACGTGT	AACTAGA
indexing7051	CAAGCAGAAGACGGCATACGAGATgccatagGTGACTGGAGTTCAGACGTGT	CTATGGC
indexing7052	CAAGCAGAAGACGGCATACGAGATaccgtcgGTGACTGGAGTTCAGACGTGT	CGACGGT
indexing7053	CAAGCAGAAGACGGCATACGAGATcttggttGTGACTGGAGTTCAGACGTGT	AACCAAG
indexing7054	CAAGCAGAAGACGGCATACGAGATtacgccgGTGACTGGAGTTCAGACGTGT	CGGCGTA
indexing7055	CAAGCAGAAGACGGCATACGAGATggactgcGTGACTGGAGTTCAGACGTGT	GCAGTCC
indexing7056	CAAGCAGAAGACGGCATACGAGATgcgcgagGTGACTGGAGTTCAGACGTGT	CTCGCGC
indexing7057	CAAGCAGAAGACGGCATACGAGATgtcgcagGTGACTGGAGTTCAGACGTGT	CTGCGAC
indexing7058	CAAGCAGAAGACGGCATACGAGATcatacgtGTGACTGGAGTTCAGACGTGT	ACGTATG
indexing7059	CAAGCAGAAGACGGCATACGAGATtcagtatGTGACTGGAGTTCAGACGTGT	ATACTGA
indexing7060	CAAGCAGAAGACGGCATACGAGATctaagtaGTGACTGGAGTTCAGACGTGT	TACTTAG
indexing7061	CAAGCAGAAGACGGCATACGAGATttagcttGTGACTGGAGTTCAGACGTGT	AAGCTAA
indexing7062	CAAGCAGAAGACGGCATACGAGATcgccgtcGTGACTGGAGTTCAGACGTGT	GACGGCG
indexing7063	CAAGCAGAAGACGGCATACGAGATgtcttctGTGACTGGAGTTCAGACGTGT	AGAAGAC

indexing7064	CAAGCAGAAGACGGCATACGAGATgccggacGTGACTGGAGTTCAGACGTGT	GTCCGGC
indexing7065	CAAGCAGAAGACGGCATACGAGATaagctgaGTGACTGGAGTTCAGACGTGT	TCAGCTT
indexing7066	CAAGCAGAAGACGGCATACGAGATgcgctctGTGACTGGAGTTCAGACGTGT	AGAGCGC
indexing7067	CAAGCAGAAGACGGCATACGAGATcgtaggcGTGACTGGAGTTCAGACGTGT	GCCTACG
indexing7068	CAAGCAGAAGACGGCATACGAGATatgattaGTGACTGGAGTTCAGACGTGT	TAATCAT
indexing7069	CAAGCAGAAGACGGCATACGAGATgcaggttGTGACTGGAGTTCAGACGTGT	AACCTGC
indexing7070	CAAGCAGAAGACGGCATACGAGATaatcgtcGTGACTGGAGTTCAGACGTGT	GACGATT
indexing7071	CAAGCAGAAGACGGCATACGAGATcggcctaGTGACTGGAGTTCAGACGTGT	TAGGCCG
indexing7072	CAAGCAGAAGACGGCATACGAGATctatgccGTGACTGGAGTTCAGACGTGT	GGCATAG
indexing7073	CAAGCAGAAGACGGCATACGAGATggttgaaGTGACTGGAGTTCAGACGTGT	TTCAACC
indexing7074	CAAGCAGAAGACGGCATACGAGATgagttaaGTGACTGGAGTTCAGACGTGT	TTAACTC
indexing7075	CAAGCAGAAGACGGCATACGAGATtagactaGTGACTGGAGTTCAGACGTGT	TAGTCTA
indexing7076	CAAGCAGAAGACGGCATACGAGATtcatgcaGTGACTGGAGTTCAGACGTGT	TGCATGA
indexing7077	CAAGCAGAAGACGGCATACGAGATgcttattGTGACTGGAGTTCAGACGTGT	AATAAGC
indexing7078	CAAGCAGAAGACGGCATACGAGATcaaggctGTGACTGGAGTTCAGACGTGT	AGCCTTG
indexing7079	CAAGCAGAAGACGGCATACGAGATaggttggGTGACTGGAGTTCAGACGTGT	CCAACCT
indexing7080	CAAGCAGAAGACGGCATACGAGATcttctgcGTGACTGGAGTTCAGACGTGT	GCAGAAG
indexing7081	CAAGCAGAAGACGGCATACGAGATtaattctGTGACTGGAGTTCAGACGTGT	AGAATTA
indexing7082	CAAGCAGAAGACGGCATACGAGATgatgctgGTGACTGGAGTTCAGACGTGT	CAGCATC
indexing7083	CAAGCAGAAGACGGCATACGAGATcctagaaGTGACTGGAGTTCAGACGTGT	TTCTAGG
indexing7084	CAAGCAGAAGACGGCATACGAGATctagaggGTGACTGGAGTTCAGACGTGT	CCTCTAG
indexing7085	CAAGCAGAAGACGGCATACGAGATtatccggGTGACTGGAGTTCAGACGTGT	CCGGATA
indexing7086	CAAGCAGAAGACGGCATACGAGATaggcggcGTGACTGGAGTTCAGACGTGT	GCCGCCT
indexing7087	CAAGCAGAAGACGGCATACGAGATggtcgttGTGACTGGAGTTCAGACGTGT	AACGACC
indexing7088	CAAGCAGAAGACGGCATACGAGATccgctggGTGACTGGAGTTCAGACGTGT	CCAGCGG
indexing7089	CAAGCAGAAGACGGCATACGAGATggaactaGTGACTGGAGTTCAGACGTGT	TAGTTCC
indexing7090	CAAGCAGAAGACGGCATACGAGATattgccaGTGACTGGAGTTCAGACGTGT	TGGCAAT
indexing7091	CAAGCAGAAGACGGCATACGAGATatatacgGTGACTGGAGTTCAGACGTGT	CGTATAT
indexing7092	CAAGCAGAAGACGGCATACGAGATgattagcGTGACTGGAGTTCAGACGTGT	GCTAATC
indexing7093	CAAGCAGAAGACGGCATACGAGATagaagtcGTGACTGGAGTTCAGACGTGT	GACTTCT
indexing7094	CAAGCAGAAGACGGCATACGAGATatagtacGTGACTGGAGTTCAGACGTGT	GTACTAT

Table S 4 Primer sequences for marker verification.

Name	Sequence 5´ to 3´	marker_resequenced
GGA_61_2_F	GTGTGTTGCTGTTTGGTGGA	GGA_61
GGA_61_2_R	CCCTTACGAGCTCACCATGT	GGA_61
Contig_41725_1_F	AAGATGGAGCTGCAGAAGGA	GBS_22
Contig_41725_1_R	CCCGCACCAATAAGCAGTAT	GBS_22

Wendler at al. 2015

Additional files can be downloaded via *e!DAL*(Arend et al., 2014) under:
http://dx.doi.org/10.5447/IPK/2015/3

Additional_file 1 Genotyping and crossing information for 146 re-sequenced ILs.

The first slide (Introgression lines) of this file contains the genotyping information and information on the crossing schemes of the analyzed ILs. An "IPK ID" was provided for each IL to simplify tracking and designating of the lines within the manuscript. The designation under "IL code" and "DH" provides the original naming, which has been commonly used at the New Zealand Institute for Plant & Food Research Limited, Christchurch, New Zealand. "IL code" and "DH" printed in "bold" designate the lines or double haploids ("DH") that were used for re-sequencing. The *Hv* cultivar 'Golden Promise' was abbreviated with "GP" and the description under the column "hybrid breeding" encodes as follows: 1. cross tetraploid *Hv* with tetraploid *Hb*; 2. cross diploid *Hv* with diploid *Hb* then colchicine treatment; 3. cross tetraploid *Hb* with diploid *Hv* than backcross to *Hv*; 4. cross tetraploid *Hb* with diploid *Hv*. The columns "introgression location" indicate the genomic location of detected *Hb* segments using either *in situ* hybridization and/or molecular markers at the New Zealand Institute for Plant & Food Research Limited, Christchurch, New Zealand ("NZ"), or genotyping-by-sequencing ("GBS"). The second slide (Hordeum bulbosum) contains descriptions for *H. bulbosum* clones that have been the donors of ILs.

Additional_file 2 Genotyping-by-sequencing SNP frequencies and raw sequencing data output of 145 ILs.

This file contains the *Hb*-specific SNP frequencies and raw sequence data output for 145 ILs. Genotyping-by-sequences SNPs were filtered to identify *Hb* derived SNPs (*Hb*-specific SNP) within each IL. These SNPs were plotted along the physical reference map of barley (International Barley Sequencing Consortium, 2012) in bins of 5 Mbp ("Mbp bin"). The ILs (Sample_1 to Sample_160) are listed by their IPK ID. In case of ILs with heterozygous *Hb* segments, the designation "_het" was added behind the IL´s ID. The amount of raw sequencing output "Mbp" and "Number of sequence reads" is given for each sample.

Additional_file 3 *Hb*-specific SNPs and anchoring information from genotyping-by-sequencing of 145 ILs.

Genotyping-by-sequencing derived SNPs were filtered to identify *Hb* derived SNPs (*Hb*-specific SNPs) within each IL. A list of filtered *Hb*-specific SNPs is given for each IL. The ILs (Sample_1 to Sample_160) are listed by their IPK ID. In case of ILs with heterozygous *Hb* segments, the designation "_het" was added behind the IL´s ID. *Hb*-specific SNPs were anchored based on the genetically and physical barley reference maps (International Barley Sequencing Consortium, 2012; Mascher et al., 2013a). The column "contig" declares the reference sequence contig of cultivar 'Morex' (International Barley Sequencing Consortium, 2012), on which a particle SNP was detected. The columns "pos", "ref", "alt" and "qual" designate the position of the SNP within the reference contig, the allele of the reference sequence, the alternative allele and the SAMtools SNP quality score (Li et al., 2009), respectively. The genotypes for the ILs and the respective *Hb* and *Hv* donor lines are listed in columns "F", "G" and "H" (0: homozygous the reference allele; 2: homozygous the alternative allele; 1: heterozygous). The anchoring information for each SNP position is given in columns I to N. The columns "chr", "cM", "parent_fpc" and "MB" declare the barley chromosome, CentiMorgans, the BAC (fingerprint) contig and the Mbp position of a SNP locus based on the physical/genetical reference (International Barley Sequencing Consortium, 2012), respectively. Furthermore, columns "chr_popseq" and "popseq_cM" describe the barley chromosome and the CentiMorgans of a SNP locus based on the genetical barley reference map (Mascher et al., 2013a).

Additional_file 4 Genotyping-by-sequencing FASTA-files.

FASTA-files of genotyping-by-sequencing targets with a minimum of 5-fold sequence reads in the IL and the respective *Hb* and *Hv* donor lines are given for each IL. The FASTA-files are based on the whole-genome shotgun assembly of cultivar 'Morex' (International Barley Sequencing Consortium, 2012). Base pair positions with less than 5-fold sequence read coverage are masked as 'N' and filtered *Hb*-specific SNPs (Additional_file3) are visualized as IUPAC codes (IUPAC-IUB Commission on Biochemical Nomenclature (CBN), 1970).

Additional_file 5 Genotyping-by-sequencing SNP heat-plots of 145 ILs.

Hb-specific SNPs were discovered with genotyping-by-sequencing and plotted along the physical (International Barley Sequencing Consortium, 2012) (5 Mbp bins) and genetic (Mascher et al., 2013a) (2 cM bins) reference maps of barley. The frequency of *Hb*-specific SNP was visualized in 1/110 (physical map) and 1/150 (genetical map) color steps from white to red using heat.colors with the R statistical environment (R Core Team, 2012).

Additional_file 6 Approximate physical and genetic locations of detected *Hb* segments.

Approximate physical (Mbp) (International Barley Sequencing Consortium, 2012) and genetic (popseq_cM) (Mascher et al., 2013a) start and end positions of individual introgressed *Hb* segments were estimated based on the occurrence of genotyping-by-sequencing derived *Hb*-specific SNPs. Start and end positions of *Hb* segments were determined based on an uninterrupted SNP coverage along the introgressed regions. The numbers of detected *Hb*-specific SNPs within the potential *Hb* segments are given ("no. of SNPs").

Additional_file 7 Diversity analysis between overlapping *Hb* segments of different ILs.

Genotyping-by-sequencing SNP haplotypes were compared pairwise between ILs with overlapping *Hb* segments. ILs with *Hb* segments that were estimated to physically (International Barley Sequencing Consortium, 2012) overlap were compared (identity-by-state) if at least 20 SNP loci were shared and non-missing ("nshared") within the region of the overlapping *Hb* segment in a pairwise comparison ("Sample1" to "Sample2"). The percentage of identical SNP calls ("nid") between each two compared ILs was calculated for the region of the overlapping *Hb* segment ("percent_id").

Additional_file 8 Four examples for haplotype comparisons of overlapping *Hb* segments.

The genotyping-by-sequencing haplotypes of overlapping *Hb* segments were compared to identify potential polymorphic SNPs between the overlapping fragments and to determine the diversity within the segments. The column "contig", "pos", "ref", "alt", "qual", "chr", "cM", "parent_fpc", "MB", "popseq_chr" and "popseq_cM" are in accordance as described in Additional_file3. The genotypes for the compared ILs and the respective *Hb* and *Hv* donor lines

are listed in columns “F” to “I” (example 1), “F” to “J” (example 2 and 4) or “F” to “K” (example 3) (0: homozygous the reference allele; 2: homozygous the alternative allele; 1: heterozygous).

Additional_file 9 Genotyping-by-sequencing SNP calls for all ILs and donor lines.

Genotyping-by-sequencing derived SNPs were filtered to conduct a minimum sequence read coverage of 5 and a minimum SAMtools SNP call quality score of 5. The column “contig”, “pos”, “ref”, “alt”, “qual”, “chr”, “cM”, “parent_fpc”, “MB”, “popseq_chr” and “popseq_cM” are in accordance as described in Additional_file3. The remaining columns contain the SNP genotype call information of all analyzed ILs as well as the donor *Hv* and *Hb* lines (Sample_xx). The genotype description encodes as follows: 0: homozygous the reference allele; 2: homozygous the alternative allele; 1: heterozygous.

Additional_file 10 Integrated *Hv/Hb* sequence: FASTA-file.

An integrated *Hv/Hb* FASTA-file was designed based on exome capture re-sequencing data of 5 *Hb* accessions, 13 *Hv* cultivars and the whole-genome shotgun assembly of cultivar ‘Morex’ (International Barley Sequencing Consortium, 2012). Base pair positions with less than 5-fold coverage in any sample were marked as ‘N’ and detected SNPs were given in IUPAC codes (IUPAC-IUB Commission on Biochemical Nomenclature (CBN), 1970).

Additional_file 11 Integrated *Hv/Hb* sequence: List of interspecific SNPs.

An integrated *Hv/Hb* polymorphism resource was designed based on exome capture re-sequencing data of 5 *Hb* accessions, 13 *Hv* cultivars and the whole-genome shotgun assembly of cultivar ‘Morex’ (International Barley Sequencing Consortium, 2012). This file contains the filtered interspecific, conserved SNPs and their associated anchoring information. The column “contig”, “pos”, “ref”, “alt”, “qual”, “chr”, “cM”, “parent_fpc”, “MB”, “popseq_chr” and “popseq_cM” are in accordance as described in Additional_file3. The column “which_alt” declares the species, which contains the alternative allele at the particular SNP position.

12. Declaration of author contribution

This present thesis is written as a "cumulative thesis" including two scientific papers. The majority of the data presented in this thesis has been generated and analyzed by the author of the thesis (*Neele Wendler*). However, the two papers are still a product of collaboration between different people and institutes. Therefore, for the purpose of clarification the responsibilities and contributions for the present thesis are as listed below:

Wendler et al. 2014

- the design of the project was performed by *Neele Wendler* and *Nils Stein*
- introgression lines and mapping population used in this study have been generated by *Brigitte Ruge-Wehling* and *Christiane Nöh*
- exome capture experiments were performed in the Genome Diversity group IPK Gatersleben under supervision of *Neele Wendler*
- exome capture data analysis was performed by *Martin Mascher* and *Neele Wendler*
- design of Golden Gate Assay (Illumina, Inc., San Diego, US) was performed by *Neele Wendler*
- Golden Gate Assay genotyping and data analysis was performed by *Neele Wendler*
- genotyping-by-sequencing experiments were performed by *Neele Wendler*
- genotyping-by-sequencing data analysis was performed by *Neele Wendler* and *Martin Mascher*
- data analysis, validation and comparison between exome capture and GBS was done by *Neele Wendler*
- writing of manuscript was done by *Neele Wendler* and Nils Stein with the help of co-authors

Wendler et al. 2015

- the design of the project was performed by *Neele Wendler* and *Nils Stein*
- introgression lines used in this study have been generated by *Richard Pickering* and *Paul Johnston*
- genotyping-by-sequencing experiments were performed by *Neele Wendler*
- genotyping-by-sequencing data analysis was performed by *Neele Wendler* and *Martin Mascher*

- exome capture experiments were performed in the GED group IPK Gatersleben under supervision of *Neele Wendler*
- exome capture data analysis was performed by *Neele Wendler* and *Martin Mascher*
- writing of manuscript was done by *Neele Wendler* and Nils Stein with the help of co-authors

13. Curriculum vitae

Personal details

Date of birth:	13.05.1987
Place of birth:	Neumünster, Germany
Nationality:	German
Marital status:	Unmarried
E-mail:	wendler@ipk-gatersleben.de
Present address:	Leibniz Institute of Plant Genetics and Crop Plant Research (IPK) OT Gatersleben, Corrensstr. 3 D-06466 Stadt Seeland, Germany

Academic Formation

2012 - 2015	PhD thesis at Martin-Luther-University Halle-Wittenberg, Germany and Leibniz Institute of Plant Genetic and Crop Plant Research (IPK), Gatersleben, Germany (Supervisor: Dr. Nils Stein and Prof. Dr. Andreas Graner)
2010 - 2012	Master study: Crop science, agriculture at Christian-Albrechts-University Kiel, Germany
2011 - 2012	Experimental master thesis: "Expression of aluminium tolerance genes in triticale" at CSIRO Plant Industry, Canberra, Australia and Christian-Albrechts-University Kiel, Germany (Supervisor: Prof. Dr. Karl Heinz Mühling, Dr. Emmanuel Delhaize and Dr. Peter R. Ryan)
2007 - 2010	Bachelor study: Crop science, agriculture at Christian-Albrechts-University Kiel, Germany
2010	Erasmus exchange program at University of Aberdeen, UK
2009	Experimental bachelor thesis: „Lokalisation von Magnesium in Blättern der Ackerbohne (*Vivia faba*) nach einer Blattdüngung" at Christian-Albrechts-University Kiel, Germany (Supervisor: Prof. Dr. Karl Heinz Mühling)

Publications – manuscripts

1. Mühling K. H., **Wendler N.**, Zörb C., Geilfus C. M. and Neuhaus C. (2013) Amelioration of the nutritional status and yield of Faba Bean as affected by Mg foliar application. In: XVII. *International Plant Nutrition Colloquium and Boron Satellite Meeting Proceedings Book*, pp 471-472, Sabanci University, Istanbul, ISBN 978-605-4348-62-6.
2. **Wendler N.,** Mascher M., Nöh C., Himmelbach A., Scholz U., Ruge-Wehling B. and Stein N. (2014) Unlocking the secondary gene-pool of barley with next-generation sequencing. *Plant Biotechnology Journal,* 12: 1122–1131.
3. **Wendler N.,** Mascher M., Himmelbach A., Scholz U., Johnston P., Pickering R. and Stein N. (2015) *Bulbosum* to go: a toolbox to utilize *Hordeum vulgare / bulbosum* introgressions for breeding and beyond. *Molecular Plant* doi:10.1016/j.molp.2015.05.004

Publications – talks

1. **Wendler N**., Mascher M., Nöh C., Himmelbach A., Scholz U., Ruge-Wehling B., Stein N. (2014) Ways to unlock barley´s secondary gene pool for breeding by next-generation-sequencing. *PAG XXII, San Diego, US.*
2. **Wendler N**., Mascher M., Nöh C., Himmelbach A., Scholz U., Ruge-Wehling B., Stein N. (2014) Ways to unlock barley´s secondary gene pool for breeding with Next-Generation-Sequencing. *PSSC meeting, Halle, Germany.*
3. **Wendler N**., Mascher M., Nöh C., Himmelbach A., Scholz U., Ruge-Wehling B., Stein N. (2014) Ways to unlock barley´s secondary gene pool for breeding with next-generation-sequencing. *Joint PGR Secure and EUCARPIA International Conference, Cambridge, UK.*
4. **Wendler N**., Mascher M., Nöh C., Himmelbach A., Scholz U., Ruge-Wehling B., Stein N. (2014) Ways to unlock barley´s secondary gene pool for breeding with next-generation-sequencing. Lab meeting, *The Sainsbury Laboratory, Norwich, UK.*
5. **Wendler N**., Nöh C., Ruge-Wehling B. and Stein N. (20014) TRANSBULB: Genomics-based approaches to utilize the secondary gene pool of barley for sustainable breeding. *KWS Cereals IT meeting, Einbeck, Germany.*
6. **Wendler N**., Nöh C., Habekuß A., Westphal A., Herz M., Schweizer G., Korzun V., Großer J., Einfeldt C, Jaiser H., Spiller M, Stiewe G., Greif P., Laubach E., Ruge-Wehling B., Stein N. (2015) TRANSBULB: Unlocking the secondary gene pool for barley breeding. *Plant 2030 status seminar, Potsdam, Germany*

Publications– posters

1. **Wendler, N**., Gerendas J., Mühling K. H. (2010) Influence of Mg foliar application on subcellular cation concentrations of *Vicia faba* leaves. *International Symposium of the German Society of Plant Nutrition, Hannover, Germany.*
2. **Wendler N**., Delhaize E., Ryan P. R., Horst W. J, Mühling K. H. (2012) Expression of aluminum tolerance genes in triticale. *German Society of Plant Nutrition, Bonn, Germany.*
3. **Wendler N.,** Mascher M., Himmelbach A., Scholz U., Ruge-Wehling B., Stein N. (2012) Marker development for barley introgression lines using the linear gene order of the barley genome as an unique and novel platform. *PSSC meeting, Gatersleben, Germany.*

4. **Wendler N.,** Mascher M., Himmelbach A., Scholz U., Ruge-Wehling B., Stein N. (2012) Marker development for barley introgression lines using exome-capture as a novel platform. *German Society for Plant Breeding, Halle, Germany.*
5. **Wendler N**., Mascher M., Himmelbach A., Scholz U., Nöh C., Richmond T., Jeddeloh J., Waugh R., Gerhardt D., Muehlbauer G., Habekuss A., Westphal A., Herz M., Schweizer G., Ruge-Wehling B., Stein N. (2013) TRANSBULB: Utilizing Exome Capture for barley breeding. *Plant 2030 status seminar, Potsdam, Germany.*
6. Nöh C., **Wendler N.**, Mascher M., Scholz U., Himmelbach A., Stein N., Habekuß A., Schweizer G. F., Herz M., Ruge-Wehling B. (2012) Fine mapping of five resistance genes on introgressions of Hordeumbulbosum chromosomal segments in barley with SNP markers. *JKI-Nachwuchswissenschaftlerforum (NWF), Quedlinburg, Germany.*
7. **Wendler N**., Nöh C., Mascher M., Himmelbach A., Scholz U., Richmond T, Jeddeloh J., Waugh R., Gerhardt D., Muehlbauer G., Ruge-Wehling B., Stein N. (2013) Utilizing exome capture to unlock barley's secondary gene pool. *PSSC meeting, Halle, Germany* .
8. Nöh C., **Wendler N**., Stein N., Habekuß A., Schmidt M., Ruge-Wehling B. (2013) TRANSBULB: Fine mapping of the resistance gene *Rym16Hb* introgressed from *H. bulbosum* to barley. *PSSC meeting, Halle, Germany*.
9. Yang P., **Wendler N**., Himmelbach A., Mascher M, Habekuß A, Ordon F., Stein N. (2013) Towards map-based cloning of the resistance gene *rym7* by genotyping-by-sequencing. *The 7th International Triticeae Symposium (7ITS), Chengdu, China.*
10. **Wendler N**., Mascher M., Himmelbach A., Scholz U.,Nöh C.,Ruge-Wehling B., Stein N. (2013) Utilizing next-generation-sequencing to unlock barley's secondary gene pool. *SFB924 meeting, Freising, Germany*.
11. **Wendler N**., Mascher M., Nöh C., Himmelbach A., Scholz U., Ruge-Wehling B., Stein N. (2014) Ways to unlock barley´s secondary gene pool for breeding by NGS. *Plant 2030 status seminar, Potsdam, Germany.*
12. Lüders T., Keilwagen J., **Wendler N**., Himmelbach A., Sharma R., Kilian B., Stein N. and Ordon F. (2014) Genotyping-by-Sequencing of a set of diverse spring barley (*Hordeum vulgare*) accessions. *EUCARPIA Cereals Section – ITMI Joint Conference, Wernigerode, Germany*.
13. **Wendler N**., Mascher M., Nöh C., Himmelbach A., Scholz U., Ruge-Wehling B., Stein N. (2014) Ways to unlock barley´s secondary gene pool for breeding with Next-Generation-Sequencing. *EUCARPIA Cereals Section – ITMI Joint Conference, Wernigerode, Germany.*

14. Declaration under oath

Declaration under oath

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.

_____

Date/ Signature of the applicant (Neele Wendler)