Aus dem Institut für Agrar- und Ernährungswissenschaften

der Naturwissenschaftlichen Fakultät III

der

Martin-Luther-Universität Halle-Wittenberg

A genetically anchored physical map of barley chromosome 2H

Dissertation

Zur Erlangung des akademischen Grades Doctor agriculturarum (Dr. agr.)

Von

Naser Poursarebani

geb. am: 22/11/1978 in: Minab, Iran

Gutachter:

Prof. Dr. Andreas Graner

Prof. Dr. Klaus Pillen

Prof. Dr. Frank Ordon

Verteidigt am

16/07/2012

Halle/Saale

Acknowledgments

I feel so proud and fortunate to complete my PhD as a member of an international scientific community; the barley genome consortium. My deep and sincere appreciation goes to my supervisor Dr. Nils Stein (Head of Genome Diversity group and PI of barley genome consortium), for giving me this opportunity. I am sincerely indebted for his constructive criticism, constant scientific advice, and inspiration that shaped my scientific career. Furthermore, I wish to thank him for his very excellent scientific ideas and constant guidance for writing this dissertation.

It is my pleasure to convey debt of gratitude to Prof. Dr. Andreas Graner (Director of IPK), for accepting me as PhD student at the Martin-Luther-University (MLU), Halle-Wittenberg. I am very thankful for his guidance, encouragement, excellent advice, and the fruitful suggestions during each progress report. I am grateful to Prof. Dr. Thomas Altman who accepted to be my scientific mentor and for the helpful suggestions and discussions during the progress meetings. Furthermore, I would like to thank Prof. Dr. Klaus Pillen and Prof. Dr. Frank Ordon for agreeing to be my supervisors and evaluate this thesis.

I would like to give many thanks to all my colleagues in GED group for their support and providing a very friendly and scientific environment. I have especial thanks to Dr. Daniela Schulte, Dr. Ruvini Ariyadasa, Dr. Rounan Zhou and Dr. Fahimeh Shahinnia members of barley physical map project for their constant support and valuable scientific discussion and Dr. Axel Himmelbach for valuable scientific discussion and for help in translating the summary of the thesis into German. I would like to thank Dr. Khalil Zeynalinejad and Dr. Uttam Kumar for their comments on writhing of this thesis. I am very grateful to my friends Raj Kishore Pasam and Rajiv Sharma for establishing a valuable friendship and for their scientific support and encouragement.

I would like to deeply appreciate all technicians in GED group especially Jelena Perovic and Jenny Knibbiche for their kindness and perfect technical assistance during the entire time of this study.

I am also thankful to “Bioinformatics group” at IPK especially Dr. Uwe Scholz, Burkhard Steuernagel, Thomas Schmutzer for their assistance in DNA sequence analyses and providing very helpful bioinformatics workshops at the IPK-Gatersleben.

I am very thankful to members of Barlex project - a subproject of barley genome consortium - especially Mihaela Maria Martis and Thomas Nussbaumer to provide me by advice in bioinformatics and to let me have access to their unpublished data.

I also would like to thank Dr. Andreas Houben and Dr. Lu Ma and appreciate their collaboration in performing FISH experiments.

I deeply acknowledge Dr. Britt Leps for her efforts - even far beyond her official obligations – that made our life much convenient by supporting our establishment in Gatersleben.

I finally would like to express my special thanks to my family who gave their kind support in many aspects. I would like to express my profound gratitude to my father and mother who helped me bring my dreams to life and take active steps in making them reality; I have no words to express how grateful I am for having them.

I would like to express my deepest sense of gratitude to my wife ‘Zahra’ for her patience, kindness, incredible support, and great help during this tough career. I am more than happy for the coincidence of submission of my thesis and the birth of my little cute baby girl ‘Sarina’. All the problems during my PhD were depreciated by this dissertation and the pleasure of submitting this work is doubled by the arrival of Sarina into my life.

Dedicated to My Dear Parents, My Lovely Wife and to

My New Born Baby; Sarina

Also

Dedicated to all those who taught me and guided me

towards my milestones

Contents

List of figures

List of tables

List of supplementary tables

1. Introduction

1.1. Barley

Barley (*Hordeum vulgare* L.) is one of the most important cereal crops. It has played a great role in human life in developing agriculture and civilizations because of its food and feed quality. From the time of its domestication, barley uses have gradually changed from a food grain to feed and malting grain (Baik and Ullrich, 2008). However, in some parts of Asia and Northern Africa, barley has still remained as a major food source (Newman and Newman, 2006). Recently, interest in barley for food has increased again mainly due to health claims associated to its soluble fiber content. Beta-glucans may lower blood cholesterol by their impact on blood glucose. The latter made the barley grain a major interest for people suffering from diabetes (Baik and Ullrich, 2008). Barley ranks fifth in crop production worldwide after maize, wheat, rice, and soybean (FAOSTAT; http://www.fao.org/faostat). Besides the high value of barley in agronomy and economy, with regard to its self-pollinated and diploid nature, it has played an important role in providing insight into genetics, physiology, and plant breeding long before *Arabidopsis* emerged as a model for plant scientists (Ullrich, 2010). It is still a valuable model plant for research in other species of the tribe Triticeae like wheat and rye because of its diploid genome and its smaller genome size (DNA content = 5.1 Gbp; Dolezel et al., 1998). So far, a broad spectrum of genetic and genomic resources has been developed in barley that extensively facilitated the barley and the related species genome analysis. Large collection of expressed sequence tags (ESTs), several genetic mapping populations that provided the framework for mapping large number of molecular markers, DNA arrays, TILLING (Targeting Induced Local Lesions In Genomes) populations, mutant collections and several other resources have provided well established platforms for genomic research in barley (Sreenivasulu et al., 2008 and the references therein). These tools/resources have contributed in the relative increase of barley production. For instance, identification of molecular genetic markers linked to the important yield related traits has led to more efficient marker-assisted selection (MAS) and marker-assisted breeding (MAB) programs in barley. Graner et al., (1999) developed Bmac SSR markers closely associated (1cM) with *Rym4/5* locus responsible for resistance to the barley yellow mosaic virus. This genetic marker has been used to

select for virus resistance phenotypes, thus led to highly time and cost-effective breeding programs in winter barley cultivars. The gene was later on positionally cloned (Stein et al., 2005). Another example could be the release of mutant cultivars of barley such as “Diamant” and “Golden Promise” with higher yield and short-height that had a significant impact on brewing industry in Europe. “Diamant” is a short barley cultivar produced by gamma-ray from the original parental cultivar “Valticky” that had an increased grain yield of about 12% (for review; Ahloowalia et al., 2004). Despite such improvements in barley performance, further development of barley genomic resources and their subsequent use in breeding programs will greatly impact the present both grain quantity and quality in barley as well as the other related crop plants.

1.2. Development of advanced genomic resources for barley genome analyses

To overcome the future human life challenges of food shortage resulting from population growth and climate changes the existing yield level needs to be improved. Improving and exploring the full crop plant performance depends on the identification of genetic determinants underlying agronomically important traits and subsequent utilization in plant breeding programs to produce new genotypes with higher productivity. For example, chromosomal locations of many important genes in barley have been identified which yet need to be studied (Druka et al., 2011) in which the most agronomically important ones can be targeted for map based gene isolation. Further characterization and deep understanding of biological pathways that the genes are involved along with application of the respective knowledge in crop breeding would eventually lead to the improvement of crop performance. For that purpose, positional gene/QTL cloning and their detailed functional study are the indispensable gateways towards that goal. Albeit, there are few cases of successful map-based gene cloning in barley (Krattinger et al., 2009), the large size of barley genome and the lack of genome sequence have made the large-scale gene cloning inefficient. Therefore, highly comprehensive genomic tools such as genome sequence information and its key prerequisite physical map – especially for plant with large genomes - are required to expedite biological studies and to improve performance of crop plant such as wheat and barley. These genomic tools are now becoming achievable in wheat and barley with the

availability of fast, low cost and high throughput sequencing approaches (Metzker, 2010) and physical mapping technologies (Meyers et al., 2004).

1.2.1. Genome sequence

The knowledge we have gained over the past two decades in molecular biology of plant including barley is deeply rooted in studies performed on model plants like *Arabidopsis* and rice. These species have been selected as models to study the plant kingdom because of their diploid chromosome set, small genome size, small chromosome number, short life cycle and very well developed and adapted genetic and genomic platforms (Somerville and Koornneef, 2002). Complete access to their genome sequence has contributed significantly to our current understanding of basic biological phenomena within these plants. But more importantly, such knowledge could be translated into other systems like important crop species. One such example is the pathway regulating flowering time. Biological pathways and the genes involved in flowering promotion have been well characterized in *Arabidopsis* (Boss et al., 2004). This information along with conserved synteny among wheat, rice, and *Arabidopsis* eventually helped in understanding the system that regulates the pathway in wheat and in characterizing one of the respective genes in wheat genome (*VRN1* gene) (Yan et al., 2003). The gene is referred to as critical component of environmental adaptation that has divided wheat varieties into the winter and spring market classes. Considering such valuable achievements obtained by exploring the high level of conserved synteny and collinearity among grasses, it was originally proposed that sequencing of small genomes like that of *Arabidopsis* (Meinke et al., 1998), rice (Shimamoto and Kyozuka, 2002) and recently *Brachypodium* (Initiative, 2010) could probably compensate for the lack of genome sequence information in plants with large genome sizes. Nevertheless, numerous exceptions in conserved synteny have also being reported (Bossolini et al., 2007; Wicker et al., 2010), thus indicating the limitation of model plants for barley and other cereal grasses. As an example, detailed investigation towards cloning of *Lr34/Yr18* locus controlling leaf and stripe rust resistance in wheat showed the lack of full collinearity between wheat and the small genomes of rice and *Brachypodium* in the respective genomic region (Spielmeyer et al., 2008). Although the conserved synteny was of great help to narrow down the interval of interest towards the gene, it was shown that the region carrying the

locus is absent in both rice and *Brachypodium* syntenic segment thus, confirming the hypothesis of less conserved micro-collinearity for resistance (*R*) genes among grasses (Leister et al., 1998). Therefore, the limitations in utilizing the conserved synteny between Triticeae and model plant reflect the need of getting access to the crop genome sequence (e.g. barley) itself. Furthermore, the availability of DNA sequence of barley genome has been considered as a perfect complementary component to the existing resources noted above that provides the basic insight into genome organization, genome function, genome evolution and in short will help understanding the biology of barley as well as the related species within the tribe Triticeae (IBSC: http://barleygenome.org; white paper).

There are two main principles of sequencing a genome. The first option is to perform whole genome shotgun (WGS) sequencing and the other is to follow a hierarchical clone-by-clone shotgun sequencing procedure (Figure 1-1) as presented in the following parts.

1.2.1.1. Whole genome shotgun (WGS) sequencing

This WGS sequencing method was initially used as a standard approach for small prokaryotic genomes. In brief, it is performed by fractioning the entire genome of the organism into small pieces, and determining the sequence of each fragment to produce highly redundant sequence-fragments (reads) of the whole genome. Computer tools then will be used to identify and assemble overlapping sequence reads and to deduce a consensus sequence. This method was first used to sequence microbial genome of *Haemophilus influenzae* (Fleischmann et al., 1995). The *Drosophila* genome was the first eukaryote's genome to be sequenced by WGS approach in 1999 (Adams et al., 2000). Implementation of WGS method in rice was the initial report in plants (Yu et al., 2002). Application of this approach to large eukaryotic genomes is an efficient way to provide immediate and quick access to a relatively random and representative amount of genomic DNA sequence (Eversole et al., 2009). This information will facilitate comparative genome analysis (Bouck et al., 2000), detection of polymorphism between individuals (Sachidanandam et al., 2001), and also can be used to learn about the repetitive sequence content of the genome of interest (Wicker et al., 2008). The main shortcomings associated with this strategy include sequence gaps and mis-assemblies that are in part the results of repetitive DNA and genome complexity. However, low sequence quality and inadequate assembly stringency can also

produce mis-assembled sequences. To circumvent the problems of repetitive DNA and genome complexity as an alternative approach so called hierarchical physical map-based sequencing (clone by clone) has been developed as described in the following.

1.2.1.2. Hierarchical physical map-based sequencing

In this approach instead of producing the random sequence reads in a genome-wide fashion, the shotgun sequencing of individual BAC (Bacterial Artificial Chromosome) clones are preferred (Figure 1-1). In brief, the consecutive steps of the process include BAC library development, physical map construction, integration of physical map to genetic maps, and identification of minimally overlapped clones as the input of sequencing phase. The method was first used in human genome sequencing (Consortium, 2001) and has been also referred to as physical map based sequencing strategy. Several plant genome projects considered the advantages of clone-by-clone strategy including *Arabidopsis* (Initiative, 2000), rice (Yu et al., 2002), maize (Schnable et al., 2009) and *Sorghum* (Paterson et al., 2009). For barley, an international sequencing consortium selected the physical-map based sequencing strategy (IBSC: International Barley Genome Sequencing Consortium; http://barleygenome.org).

1.2.2. Physical map

A physical map is a model of a reconstructed chromosome. After sub-cloning the large genomic fragments of the genome of interest in a genomic library, it is the aim to identify manageable, overlapping pieces and reconstruct the individual linkage groups. Current physical map methodologies are based on availability of the genomic BAC libraries and the possibility to detect overlaps between BACs (Figure 1-1, A to C).

There are two main large insert cloning systems including YACs (Yeast Artificial Chromosomes) and BACs (Bacterial Artificial Chromosomes). YACs were introduced in 1987 by (Burke et al., 1987). The system uses *Saccharomyces cerevisiae* as the host and is able to carry large inserts up to 1 Mb. The main disadvantage was the high level of chimerism within the YAC cloning system which often led to the cloning of two or more unlinked DNA fragments in a single molecule (Luo and Wing, 2003). Considering the problem associated with the system, it has been gradually replaced by the BAC

cloning system introduced in 1992 (Shizuya et al., 1992). The *Escherichia coli* F-factor has been used as vector in this system that can carry DNA inserts up to 300 kb. The recombinant vector can be cloned and stably maintained in *E. coli*. The smaller capacity of the BAC system is its advantage over the YAC system, because of its ability in maintaining small DNA fragments (compared to YAC system) it prevents co-insertion of two or more DNA fragments in a particular BAC clone, thus precluding chimerism phenomenon (Luo and Wing, 2003). Once provided with a high quality BAC library (libraries), the process of building the physical map for the genome of interest can be actuated (Figure 1-1, D). In BAC clone based physical mapping, each individual clone enters in the process of so-called fingerprinting in which unique landmarks of the BAC clones will be identified. These landmarks can be Sequence Tagged Site (STS) or enzymatic restriction sites (Green, 2001). Each pair of physically overlapped clones must represent a common set of landmarks due to conserved sequence originally shared between the two. The common method of producing BAC clone fingerprints is by cleaving the BAC DNA into reproducible fragments by help of specific restriction endonucleases (Meyers et al., 2004).

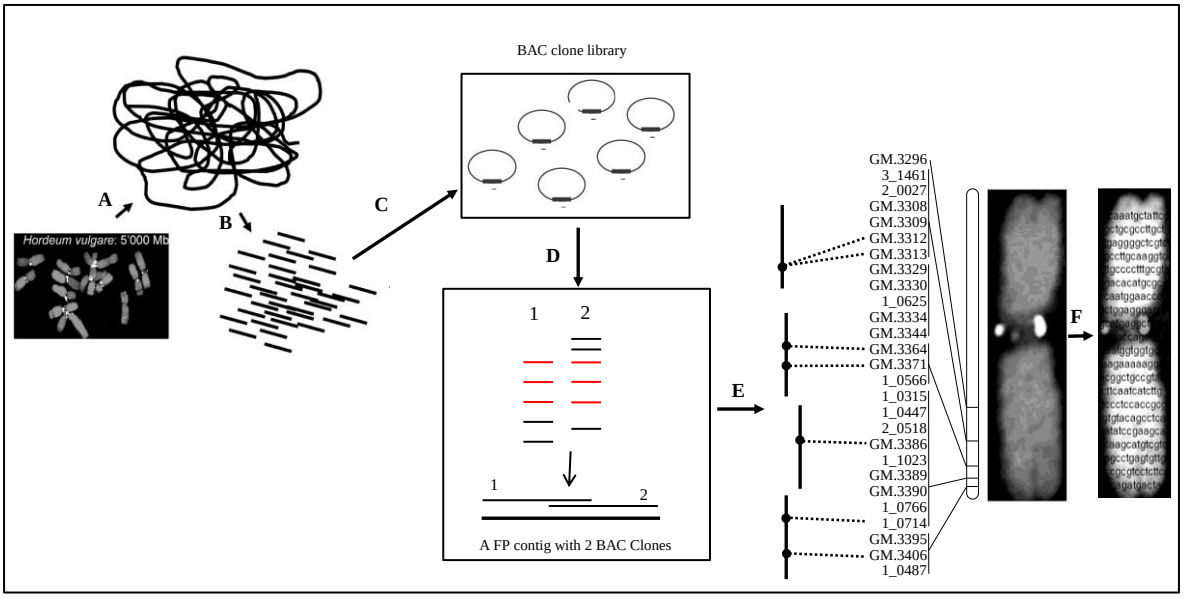


Figure1-1. Schematic presentation of clone by clone sequencing approach: A, B and C) Extracted DNA is digested by a chosen restriction enzyme, size selected fragments (average: 120 Kb) are cloned into a proper vector and stably maintained in *E. coli* . D) Individual BAC clones are subjected to fingerprinting and physical map contig creation in which overlapped BAC

clones are identified by counting the number of fragments representing identical size (red lines). E) Physical map contigs are anchored to their original position in the genome using genetically mapped markers. F) A minimum number of BACs representing the entire length of each contig (MTP= Minimum Tiling Path) enter into the sequencing phase. Figure is modified from Nils Stein (Stein, 2009)

The degree of overlap of two clones will determine the number of shared bands between every two BACs. BAC clones in which the degree of overlap has reached a decided threshold will be assembled as physical map contigs (hereafter FingerPrinted contigs: FPcontigs). Each FPcontig represents a number of overlapping BAC clones from a contiguous segment of the genome. Restriction based fingerprinting can be different in terms of the reaction biochemistry, the separation medium and the information content (the average number of bands per fingerprinted clone). Improvements in the fingerprinting methods were always towards an increase in information content and towards an increase in the band resolution or the space that separates the bands (Meyers et al., 2004). The earlier strategies applied one or two restriction enzymes for clone fragmentation and agarose or acrylamide gel as a separation medium (Coulson et al., 1986; Marra et al., 1997). Several modifications to the basic fingerprinting methods have been proposed including the possibility to increase throughput, sizing accuracy and information content of fingerprinting (Hong, 1997; Zhang and Wing, 1997; Klein et al., 2000; Tao et al., 2001; Luo et al., 2003). The current standard technology 'high information content fingerprinting' uses multi-enzyme restriction, multicolor labeling and separation of fragments utilizing capillary DNA fragment analysis (Luo et al., 2003; Nelson et al., 2007). This method has been applied for physical mapping the barley genome (Schulte et al., 2011; Stein et al., unpublished data).

Despite the pivotal role of the physical map in genome sequencing of large and complex genomes, it can enable for several other applications. Among those, the most important one is the use of physical map in map-based gene isolation even before the availability of whole genome sequence. Establishing a local and region specific physical map by chromosome walking has been applied for map based gene cloning in barley. For instance, Brueggeman et al., (2002) localized a barley stem rust-resistance gene (*Rpg1*) in a genetic interval of 0.21 cM on barley chromosome 7H via high resolution genetic mapping. BAC based physical map of the respective interval resulted in a 330-kb physical contig. Sequencing of the corresponding physical segment and its comparison between the related susceptible and resistance

parental cultivars revealed a candidate gene conferring resistance to stem rust in barley. Numerous genes have been cloned and characterized by similar methodology in the large genomes of cereals (Krattinger et al., 2009). The availability of a whole genome physical map will dramatically expedite and extend such investigations in large cereal genomes.

Physical mapping can also be applied to study the chromosome and genome organization between related species that is known as comparative genomics. Kim et al., (2008) developed a genome-level comparative experimental system for the genus *Oryza.* BAC libraries, BAC fingerprints, BAC-end sequences (BES), and the subsequent physical map frameworks (genome coverage of 77% to 136%) were produced for ten genomes of different *Oryza* species. Kim and co-workers then aligned the respective physical maps to the *O. sativa* reference sequence. Despite the extensive collinearity between individual physical maps and the reference genome, this analysis led to characterization of the repeat content of individual genomes - using BES data - and their possible role in genome variation during evolution. The authors claimed that by providing the phenotypic, genetic, biochemical and physiological information to this comparative framework, fundamental questions in biology and agriculture can be potentially addressed.

In the process of designing a tool for genome synteny analysis Soderlund et al., (2006) aligned the physical map of maize and two species of *Sorghum* against genome sequence of rice. The availability of BAC end sequences and markers with known sequence could provide anchor points of the respective physical maps to rice genome. Using this information, the authors could analyze and visualize the recent and ancient duplications existing among the studied genomes. Another example is reported by Gregory et al., (2002) who aligned the mouse physical map aided by BAC end sequences against human genome. The authors succeeded to improve the physical map of mouse genome and identifying the conserved syntenic blocks between the two genomes in an advance resolution. Moreover they suggested the usefulness of human genome to be used as framework in construction of the other evolutionary related mammalian genome physical maps.

Utilization of a physical map to develop a detailed cytogenetic map of the genome of interest has also been reported. Islam-Faridi et al., (2002) used BAC clones underlying the *Sorghum* physical map to develop the BAC-FISH based cytogenetic map of *Sorghum* chromosome 1. Such resource is extremely useful for determination of chromosomal landmarks like centromere and heterochromatin and studying of recombination frequency across genome. The usefulness of a physical map in structural genome variation, marker development and repeat identification that have been reported for instance by van der Vossen et al., (2000), Kidd et al., (2008) and Cardle et al., (2000), respectively, are also considerable.

All applications introduced and listed above will be feasible only after integration or anchoring of the physical map to a genetic map. Therefore, genetic anchoring plays a major role in the efficiency of the physical map. Various anchoring approaches are briefly introduced in the following parts with focus on barley genome

1.3. Anchoring of physical to genetic map

"Every genome sequence needs a good map" (Lewin et al., 2009). A physical map becomes a good and effective map only after anchoring to a robust genetic map. Only then, the position and relative order of the FPcontigs along each chromosome can be determined and the full potential of a physical map can be explored. A genetic linkage map determines the order and location of genes/markers along each chromosome based on meiotic recombination. Mapped genes/markers can be used to screen BAC libraries underling the physical map to reveal BAC/marker relationships. The corresponding BAC contig will be subsequently placed in the respective marker position on the chromosome and the correct contig orientation may be determined (Figure 1-1, E). Physical/genetic map integration has been performed for several animal and plant genomes, for instance, human (McPherson et al., 2001), mouse (Gregory et al., 2002), rice (Chen et al., 2002), *Sorghum* (Klein et al., 2000), maize (Coe et al., 2002), and wheat (Paux et al., 2008). For barley genome, different genetic maps have been developed before the inception of the current study that renders the potential resource for barley genome physical map anchoring (section 1.3.1). In the following sections, first the barley genetic maps, potential marker resources for physical map anchoring will be described. The approaches by which a physical map can be aligned against the genetic map will also be reviewed.

1.3.1. Genetic map resources of barley – the framework of anchoring of the physical map

Several genetic maps have been developed and reported for barley genome over the last decades. These include different marker sources (from Restriction Fragment Length Polymorphism (RFLP) markers to DArT marker) (Graner et al., 2010) for which several mapping populations have been developed (Table1-1). In this section, the recent high density genetic maps published for the barley genome and available at the starting time of the current study will be reviewed (Table1-1). Our objective was to select among those the genetic maps originating from gene-targeted markers and the maps that fit to the PCR-based anchoring workflow of the present study.

In order to construct a high density genetic map and to genetically map genes responsible for abiotic stress, Rostoks et al., (2005) performed a genome–wide gene-based SNP discovery on barley genome. A set of 1,338 unigenes differentially expressed in response to a variety of abiotic stresses were re-sequenced in eight different barley accessions and were mapped in three different doubled haploid (DH) mapping populations. This analysis generated a consensus map comprising of 1,237 loci, length of 1,211 cM and with 1 locus per cM average resolution. Wenzl et al., (2006) used a hybridization-based technology so called Diversity Arrays Technology (DArT) (Jaccoud et al., 2001) to generate a consensus map using six DH population and three recombinant inbred line (RIL) populations. Analysis of 2,935 markers (2,085 DArT, 850 other loci) produced a map that spanned 1,161cM with an average resolution of 0.7 ± 1.0 cM. The hybridization based technology applied to develop this class of markers and the anonymity of majority of the respective markers preclude the possibility of being directly used in our PCR-based anchoring method. Stein et al., (2007) applied the consensus map method to take the advantages of achieving high level of polymorphism by combining diverse and non-related germplasm- the parents of the mapping populations. The authors reported a gene based consensus map (here in this study called as IPK map) consisting of 1,055 ESTs based markers (total 1,055 loci: 607 RFLP, 190 SSR, and 258 SNP) generated from three different doubled haploid (DH) mapping populations. The number of markers per chromosome ranged from 107 to 179, the total length of map is 1118.3 cM with average marker interval of 0.9 cM. In total 200 common markers were shared between the individual maps. With regards to the application of gene based genetic marker in this map,

Table 1-1. Overview of some genetic maps of barley

Reference [1]	Marker [2]	Majority marker type		Loci	cM	Population [3]	Majority marker system	
		Gene based	Non-genic				PCR based	Hybridization based
Rostoks et al., (2005)	**SNP**,RFLP,AFLP,SSR	×		1237	1121	LH, OWB, SM	×	×
Wenzel et al., (2006)	**DArT**, RFLP, SSR, STS		×	2935	1161	Multiple	×	×
Stein et al., (2007)	EST	×		1055	1118	SM, OWB, IF	×	×
Varshney et al., (2007)	SSR	×		775	1068	OWB, SM, IF	×	
Marcel et al., (2007)	RFLP,**AFLP**,SSR	×	×	3258	1081	Multiple	×	×
Hearnden et al., (2007)	**SSR,** DArT		×	1000	1100	Barque-73 × CPI 71284-48	×	×
Potokina et al., (2008)	TDM	×		1596	1010	SM		×
Sato et al., (2009)	EST	×		2890	?	HNH602	×	
Close et al., (2009)	SNP	×		2493	1099	Multiple	×	

[1)] Listed and introduced in the chapter.
[2)] Marker showed in bold represent markers with the highest number among all marker types of the respective genetic map.
[3)] Abbreviations: LH (Lina x HS92), OWB (Oregon wolf barley recessive x dominant), SM (Steptoe x Morex), IF (Igri x Franka), HNH602 (Haruna Nijo x H602). Multiple represents application of more than three populations for the respective genetic map formation
Table is modified from (Graner et al., 2010)

the PCR-based SSR and SNP markers can be utilized for PCR-based physical map anchoring in the current study. Varshney et al., (2007) has explored a collection of 2,832 non-redundant ESTs, for SSR motif identification in which 3,122 non-redundant SSR-ESTs were identified. Of those, for 754 SSRs a primer set was designed. Only 185 EST-derived SSRs revealed polymorphism between parental genotypes of three DH populations and led to construction of a map with 1,068 cM map length. These markers have also been included in the integrated map constructed by Stein et al., (2007). Marcel et al., (2007) compiled a set of 3,258 markers including RFLP, AFLP, and SSR markers from six different mapping populations to construct a consensus map with the length and resolution of 1,081 cM, and 0.33 cM, respectively. Only 2% of the markers (mainly SSRs) were originated from EST sequence data, hence referred as gene-targeted markers and the rest are from anonymous DNA fragments. In order to avoid the limitation of order accuracy associated with integrated maps Li et al., (2010), and Hearnden et al., (2007) developed a high density barley genetic map composed of 1,000 loci (558 SSRs and 442 DArT markers) using a single mapping population with a map length of 1,100cM. The main disadvantages of the map for the current study are anonymity of markers (only 15% of the markers were gene based markers). In addition almost half of the markers were hybridization based (442 DArT markers). Potokina et al., (2008) developed a set of 1,596 transcript derived markers (TDMs) on a single DH mapping population derived from parental genotypes barley cv. Steptoe and cv. Morex. The polymorphism between the parental genotypes and across the progenies was detected based on signal intensity revealed on Affymetrix microarrays corresponded to the expression level of the respective genes. Considering the hybridization based system that was used for marker development in this analysis, the map has perhaps a high potential to be used in hybridization based library screening and subsequent physical map anchoring (section 1.4.1.2). As a complementary marker resource for IPK map, a high density transcript linkage map of barley derived from a single population developed by Sato et al., (2009) is considered (here in this study called as Okayama map). The authors have developed 10,336 primer sets using 3'-end EST sequences of barley cv. ‘Haruna Nijo’ and an ancestral wild form accession ‘H602’ which were also crossing parents of the mapping population. Using 93 DH individuals, a set of 2,890 PCR-based markers were developed. The total map length was 2136 cM with an average 421 markers per chromosome. As the authors declared,

although the level of missing data and segregation distortion for some loci are considerable, the map quality was not affected as the LOD threshold was 5. Marker information of a high density integrated map (Close et al., 2009) was available for this study. .A total number of 2,943 EST based SNP markers were genotyped in 4 different DH populations using Illumina GoldenGate assay. This map (hereafter Close et al-map) was used as the genetic framework to develop another genomic resource for barley, the so called "barley genome zippers" which will be introduced in section 1.3.2. Regarding the specific primer combination applied in the genotyping technology, the direct use of these markers in our screening platform was not considered. To summarize, although all genetic maps described have been developed to fulfill different goals, they have the potential of being employed for the purpose of physical map anchoring. However, with regards to the physical map anchoring strategy decided for the current study (PCR-based anchoring approach; see section 1.4.1.1 for details), and the priority of anchoring gene containing physical contigs as mentioned above, only two of such maps including EST based barley genetic maps developed by Stein et al., (2007) and Sato et al., (2009) will be considered for physical map anchoring of barley chromosome 2H.

1.3.2. Grass genomes synteny - application in marker development and anchoring

Members of the grass family Poaceae have diverged from a common ancestor 50-70 million years ago (Bolot et al., 2009). This evolutionary period resulted in a substantial divergence in genome organization of grasses including chromosome number and genome size. For example, the genome of bread wheat (1.7 × 1010 bp) is 40 times larger than that of rice (4.3 × 108 bp) (Keller and Feuillet, 2000). Despite this diversity, comparative genetic and genomics have demonstrated that related plant species display extensive conservation in gene content and order (conserved synteny and collinearity) (Paterson et al., 2000; Schmidt, 2002). The term "conserved synteny" reflects co-localization of groups of genes on evolutionary related chromosomal segments in two species, whereas collinearity is more specific form of synteny which indicates preservation of gene order in a syntenic region between species over the time of evolution (Abrouk et al., 2010). Since early studies of comparative genomics (Bonierbale et al., 1988; Berhan et al., 1993; Devos et al., 1993), conserved genomic synteny and collinearity have been employed as useful concepts and tools in plant genomics. This includes map

based gene isolation and gene annotation of big and less-studied genomes such as wheat and barley by taking the advantages of small but well-investigated genome like that of rice. An important application of collinearity in map based gene isolation was the development of genomic tools such as conserved orthologous sequences (COS) markers (Fulton et al., 2002; Liewlaksaneeyanawin et al., 2009; Quraishi et al., 2009).

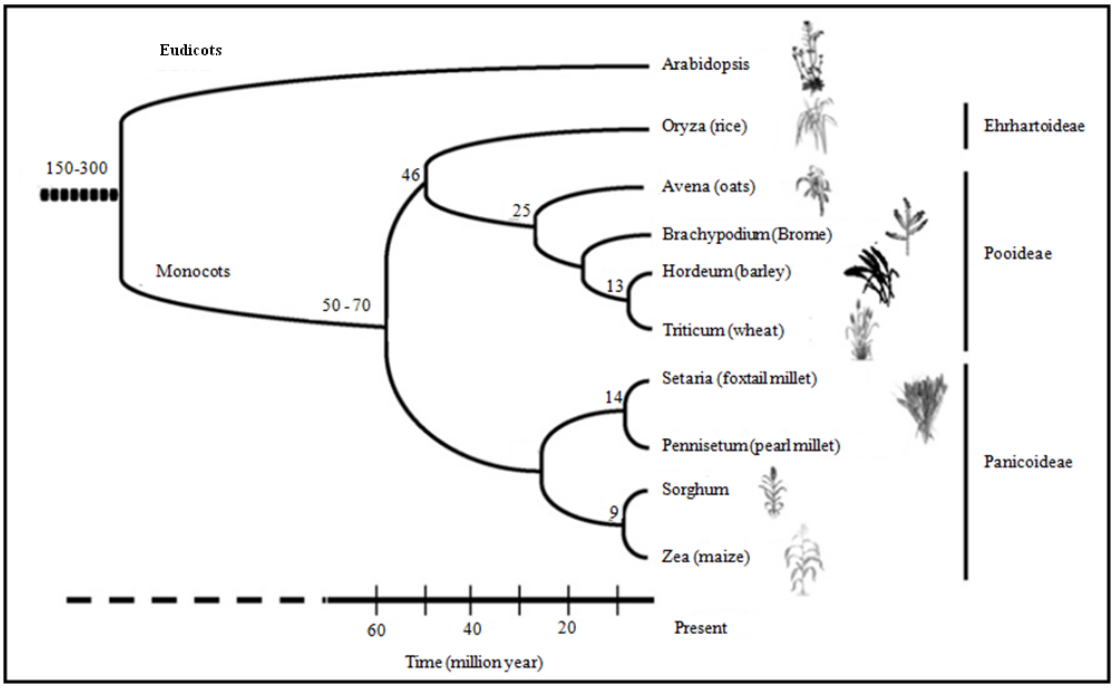
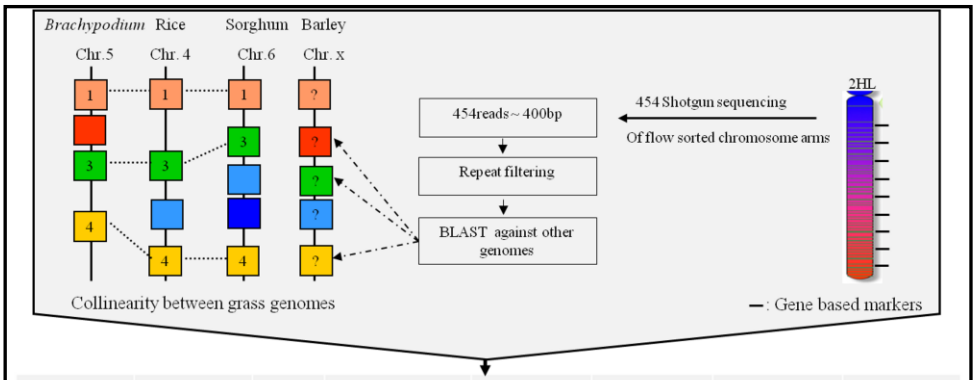


Figure1-2. Phylogenetic relationship among grass species. Divergence times are given in million years on the branches of the phylogenetic tree. Image from Bolot et al., (2009).

In grass, these markers are being developed by availability of rice genome sequence as model genome and large collections of ESTs from other cereal species. The alignment of these ESTs with rice genome sequence can help to predict corresponding location in the genome of interest based on synteny concept. The alignment would also help in identification of intron/exon boundaries in order to provide a possibility for intron-spanning primers design enhancing the chance of SNP detection in the target genome (Quraishi et al., 2009).Considering the genome relation concepts noted above, Mayer et al., (2011) have explored the full potential of conserved synteny and collinearity existing between barley and its relatives to develop a novel synteny based genomic resource for barley. To develop this synteny-based resource, the integrated gene based genetic map developed by Close et al., (2009) was

used as the genetic framework to establish a synteny derived virtual gene order map for barley genome. Such virtual gene order maps or "genome zippers" (Figure1-3) were generated after shotgun 454-sequencing of flow sorted barley chromosome/chromosome arms producing 1- to 2-fold of sequence coverage. After masking repetitive DNA, such chromosome specific sequence data could be compared to the entire genomes of three sequenced grasses including rice, *Sorghum*, and *Brachypodium* in order to detect the homologous genes in the corresponding genomes. Integration of detected homologous genes with gene - based markers of the framework genetic map led to the construction of a virtual linear gene model for each barley chromosome in which a precise genetic location for grass syntenic genes can be predicted in barley chromosomes (Mayer et al., 2011). The number of genes for which their linear order are predicted along each barley chromosome differ from 2,304 genes for chromosome 6H to 3,616 genes for 2H. In case of chromosome 2H on average ~ 22 genes / 1 cM have been order along the chromosome which makes it a valuable resource for marker development and physical map anchoring of chromosome 2H.



Gene Model ID	illumina markers	cM	Brachypodium	Rice	Sorghum	Barley 454 shotgun reads	HarvEST_35
GM3294	3_1461	138.94	-	-	-	R1;R2;R3	5285
GM3295	2_0027	138.94	-	-	-	R100;R5	5674
GM3296	1_0245	139.65	-	-	-	R6;R250	14684
GM3297	3_0914	139.65	-	-	-	R3000;R7	15644
GM3298	3_0352	139.65	Bradi5g23540.2	-	-	R675;R20	19233
GM3299	-	-	-	Os04g0641000	-	R6354;R4	-
GM3300	-	-	Bradi5g23510.1	Os04g0640900	-	R17;R43	-
GM3301	-	-	-	-	Sb06g030300.1	R43;R342	-
GM3302	-	-	-	-	Sb06g030290.1	R325;R11	-
GM3303	-	-	Bradi5g23500.1	Os04g0640800	Sb06g030280.1	R43;R69	-
GM3304	-	-	-	-	Sb06g030275.1	R70;R92	-

Figure1.3. Partial view of the virtual gene order map (genome zipper) of barley chromosome 2H (Mayer et al. 2011): Upper part) Repeat masked low pass shotgun sequencing of individual chromosome/chromosome arms obtained from flow - cytometry were used to identify the homologus genomic regions in the related sequenced grass genomes. Lower part) Detected homologous genes and the associated shotgun reads were integrated with a gene based genetic map (close et al.,

2009) that lead to form a virtual gene order map for each individual barley chromosome. Corresponding gene based sequence information (barley ESTs) were added to the barley gene models as additional supports. R stands for 454 shotgun reads.

1.4. Anchoring of physical to genetic map methodologies

1.4.1. Experimental methods of anchoring a physical map

The experimental methods of library screening/physical map anchoring can be basically categorized into two classes; PCR-based and hybridization based approaches, each of which can be further divided into single *vs* multiplex marker assays (Figure1-4). In all methods the genetic markers will be connected to the respective BAC clones through the process of library screening, thus placing the BACs and the related FPcontigs to the position of the markers in the genome.

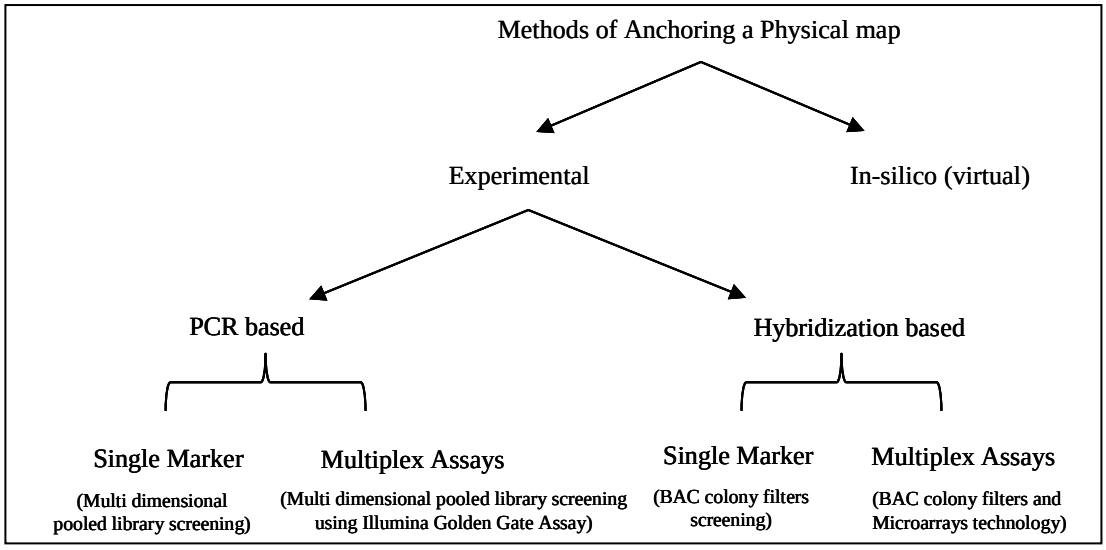


Figure1-4. Current methodologies of physical to genetic map anchoring

1.4.1.1. PCR-based BAC library screening (single and multiplexed assay)

A commonly used method for identification of BAC/marker relations and anchoring physical to genetic map is PCR-based screening of large insert libraries using PCR–based genetic markers. In this approach, rather than carrying out PCR as many as the number of clones present in the library, the library can be condensed into pools in a specific manner to reduce the overall PCR reactions needed to identify a particular BAC clone. Therefore, the efficiency of PCR-based screening can be dramatically improved by constructing Multi dimensional (MD) pools of the

respective BAC libraries. The optimal pooling dimension (D) and the number of pools depends on the redundancy of the library, the number of clones and the desired rate of false positive marker/BAC relations were studied by (Barillot et al., 1991). These authors suggested that in practice, those dimensions greater than five are usually not efficient with PCR-based markers. For example, (Barillot et al., 1991) performed an analysis to identify the best pooling dimension for the CEPH (Centre d'Etude du Polymorphisme Humain) YACs library which contains 72,000 YACs and is 10-fold representative. The optimal pooling dimension was three while the dimension four gives almost equivalent results in terms of the number of false positive marker/BAC relations. In addition, the three dimensions needs less PCR reaction to identify a YAC address. PCR-based anchoring approach has been used in several plant genomes including *Sorghum* (Klein et al., 2000), maize (Yim et al., 2007), soybean (Wu et al., 2008) and for *Aegilops tauschii* (You et al., 2010) physical map anchoring.

The Illumina Golden Gate Assay also known as OPAs (Oligonucleotide Pool Assays) are originally developed as a highly parallel SNP genotyping platform of genomic DNAs (Steemers and Gunderson, 2005). The application of Illumina Golden Gate Assay in BAC library screening was initially reported by Luo et al., (2009) to provide an alternative approach to the single marker library screening described above. The idea was to increase the throughput of the PCR-based library screening by simultaneous screening of up to about 1,500 genetic markers. Luo and co-workers applied this technology to genotype the BAC pools for the presence or absence of the corresponding SNP alleles, thus identifying the BACs carrying the respective SNPs. One year later, an improvement for the method was reported (You et al., 2010). A computational algorithm was integrated into a newly developed software tool, FPCBrowser, for analyzing pooling data and BAC address deconvolution. Although, the strategy is fast, easy and cost efficient as the authors claimed, access to the high number of perfect SNP markers (in OPA format), robust physical map and creation of the specific format of the pooled library are the crucial prerequisites for the strategy.

1.4.1.2. Hybridization based BAC library screening (single and multiplexed assay)

Screening a BAC colony filter (by single or multiple markers) or application of microarrays technology both follow the basic principle of hybridization of two complementary single stranded

nucleic acid molecules. This leads to identifying the genetic markers/BAC clones relations for the subsequent physical map anchoring. BAC filters and the radioactively labeled DNA fragments are the two prerequisites to perform the hybridization based screening. The filters can be screened either by individual or by combination of different labeled probes (multiplex assay). For the sake of throughput, labeled probes can be pooled and then hybridized to colony membranes (Madishetty et al., 2007). Individual BAC clone/marker relations will be identified through the process of de-convolution. This method has been used in a number of organisms including mouse (Cai et al., 1998), rice (Yang et al., 2003), maize (Gardiner et al., 2004) and chicken (Romanov et al., 2003). The microarray technique is also based on hybridization of two complementary single stranded nucleic acid molecules, one of which is immobilized on a matrix (Southern et al., 1999). Microarray technology has been previously used in various biological studies including large scale DNA mapping (Poustka et al., 1986), sequencing (Cantor et al., 1992), and gene expression profiling (Schena et al., 1995). Recently, Liu and co-workers (Liu et al., 2011), used this technology to identify the gene/BAC relationship for subsequent use in genetic anchoring of barley genome physical map. They used Agilent microarrays of barley unigenes by hybridizing to BAC DNA pools originated from a 3D pooling system to identify the marker/BAC relations. Such highly multiplexed BAC screening approach displayed a very time and cost-effective alternative to the conventional BAC/Marker identification procedures.

1.4.2. In-silico (virtual) anchoring of physical map to the genetic map

In-silico anchoring refers to all analyses run via computer tools performing DNA sequence homology search to place the contigs to their initial chromosomal location on the genome. The method uses the available sequence information to improve contig building and anchoring of the evolving physical map of the genome of interest (Virtual library screening). The prerequisite would be the availability of sequence information - the output of genome survey sequencing such as the determination of BAC end sequences - bounded to the FPcontigs with which the respective FPcontigs can be anchored to the genetic markers with known sequence. (Yuan et al., 2000) used the approach for map integration in rice sequencing project. They cleaned and filtered the available EST and BAC end sequences from repetitive sequences and then searched all available rice genetic markers; thereby connected the BACs

(physical map skeleton) to the genetic map. They could anchor 418 markers to a collection of BAC clones, supported by experimental verification. A BAC based physical map of papaya was constructed and integrated with the genetic map and genome sequence (Yu et al., 2009). The entire papaya BAC library of 39,168 BAC clones was either end-sequenced or full length sequenced. Paired ends from 32,397 BAC clones provided anchor points for alignment of the physical map with genome sequences and integration of the genetic and physical maps. Similar approach also was used for physical map orienting in grapevine (Lamoureux et al., 2006), upland cotton (Xu et al., 2008) and soybean (Wu et al., 2008). Consequently, the strategy can be considered as a complementary method for the wet lab library screening in the course of physical map anchoring given the availability of attached sequence information.

1.5. The aims of the study

Chromosome 2H is the biggest among the seven barley chromosomes. During the current study, the physical map anchoring to this barley chromosome will be carried out to complement the efforts of developing a genome-wide genetically anchored physical map of barley. Moreover, chromosome 2H contains loci controlling a variety of agronomically and commercially important traits. These include genes responsible for reproductive development, time to flowering, reproductive frost tolerance, and disease resistance (Costa et al., 2001; Pillen et al., 2004; Reinheimer et al., 2004; Turuspekov et al., 2004; Dilbirligi et al., 2005; Li et al., 2005; Jafary et al., 2006; von Korff et al., 2006; Marcel et al., 2007; Jafary et al., 2008; Vu et al., 2010). Access to an anchored physical map of the chromosome will expedite the detailed study of traits noted above. Therefore, the main objectives of this study are as follows:

1- Establish a first version anchored physical map of chromosome 2H of barley paving the way for map-based cloning of genes located on this chromosome, and to provide the basic information for clone-by-clone sequencing strategy of barley genome.
2- Explore the potential of anchoring the barley genome physical map based on grass genome collinearity.
3- Utilize the genetically anchored physical map to determine the pattern of recombination frequency along chromosome 2H.

2. Material and Methods

2.1. Plant material and DNA samples

To perform genetic mapping of newly developed markers (see section 2.4), a doubled haploid (DH) mapping population originated from a cross between barley cv. Morex and Barke was used. The parent genotypes were used for initial detection of polymorphisms. The population comprised 93 genotypes (SubDate_Table1) and has also been used to construct a consensus map for barley as has been reported before (Close et al., 2009). The map was utilized as framework to develop a virtual gene order map (genome zippers) for barley genome (Mayer et al., 2011). Previously extracted parental and population DNA using the method described by (Graner et al., 1991) were used for initial polymorphism detection and population genotyping.

2.2. Polymerase Chain Reaction (PCR) and gel electrophoresis

Concerning the physical map anchoring approach used in this study (section 2.3), for initial amplification test of each primer set, and for marker development process (section 2.4), PCR experiments were carried out. Except for High Resolution Melting (HRM) curve analysis (section 2.4.3) and for dCAPS markers development (section 2.4.2), a common standard PCR profile was applied. The PCR reagent mixture consisted of 1 µl of genomic DNA (20ng/µl), 1 µl of 10 x PCR buffer, 1 µl of dNTP mixture (2mM each), 1 µl of primers mix (5 pmol/ µl each), 0.05 µl of HotStar Tag DNA polymerase (Qiagen, Hilden, Germany), and 5.95 µl of ddH_2O. All fragments were amplified using the following touch-down PCR profile: an initial denaturing step of 15 min at 95 °C was followed by 40 cycles with denaturation at 94 °C for 30 s and extension at 72 °C for 1 min. The annealing temperature was decreased in 1 °C increments from 65 °C in the first cycle to 60 °C after the 5th cycle and was then kept constant for the remaining 35 cycles (always 30 s). After 40 cycles a final extension step was performed at 72 °C for 7 min. PCR amplifications were carried out using GeneAmp PCR system 9700 (Applied Biosystems). PCR products were checked by 1.5% agarose gel electrophoresis at 80 (V/cm) for 90 min in case of genomic DNA test. For BAC pool DNA analysis,

the electrophoresis time was reduced to 17 min to fit the gel size since for these experiments the Electro-Fast® Gel System; model AB-0826 (http://www.abgene.com) was used.

2.3. PCR-based screening of multi-dimensional DNA pools of a barley BAC library

In order to integrate the barley genome physical map with the genetic map of chromosome 2H, a PCR-based physical map anchoring approach was established by which BAC clones underlying physical map contigs harboring the respective genetic markers were identified. PCR markers originated either from two transcript maps published by (Stein et al., 2007) and (Sato et al., 2009) or were newly developed on the basis of information provided by a virtual linear gene order map ("barley genome zipper", (Mayer et al., 2011) (see section 2.4.1). All primer sets were initially tested for PCR amplification efficiency and specificity on barley cv. Morex genomic DNA applying the standard PCR protocol (section 2.2). Only primer pairs that passed this step of quality check were subjected to PCR-based library screening. The next prerequisite of the PCR-based anchoring approach - a pooled BAC library - was obtained from the barley BAC library HVVMRXALLeA (Schulte et al., 2011) by a commercial service provider (Amplicon Express, Pullman, WA, USA). Multidimensional pooling systems reduce the number of PCR reactions required to screen a complex BAC library for the BAC address harboring the respective marker. The pooling system of Amplicon Express involved 3D pooling schemes of Superpool and Matrix pool design (details: www.amplicon-express.com). In brief, the original library contained 147,840 BAC clones arranged in three hundred eighty-five 384-well plates. All plates were collected into 55 superpools (SPs), each containing 7 consecutive 384-well plates of the library (Figure 2-1). For each superpool, the individual 7 plates, the respective 16 plate rows across all 7 plates and the respective 24 plate columns also across all 7 plates were initially pooled. This created 7 plate pools, 16 row pools and 24 column pools per superpool (Figure2-1). These pools were then further combined to create five Matrix Plate Pools (MPPs), eight Matrix Row Pools (MRPs), and 10 Matrix Column Pools (MCPs), respectively, resulting in 23 Matrix Pools (MPs) for each SP (Figure2-1). The design of Matrix pools resulted in each BAC clone being represented in two different MPs of each MP type (2/5 of MPPs, 2/8 of MRPs, and 2/10 of MCPs). All primer sets yielding a positive superpool (for single copy genes a maximum of 4 hits was expected on average

since the BAC library comprised about 4-fold haploid genome coverage. with the same size as genomic DNA were examined during the second round of library screening (MPs Screening). Figure2-2 summarizes SPs screening, MPs screening, and deconvolution steps of a BAC address for a given marker. The PCR condition for library screening was the same as described in section 2.2

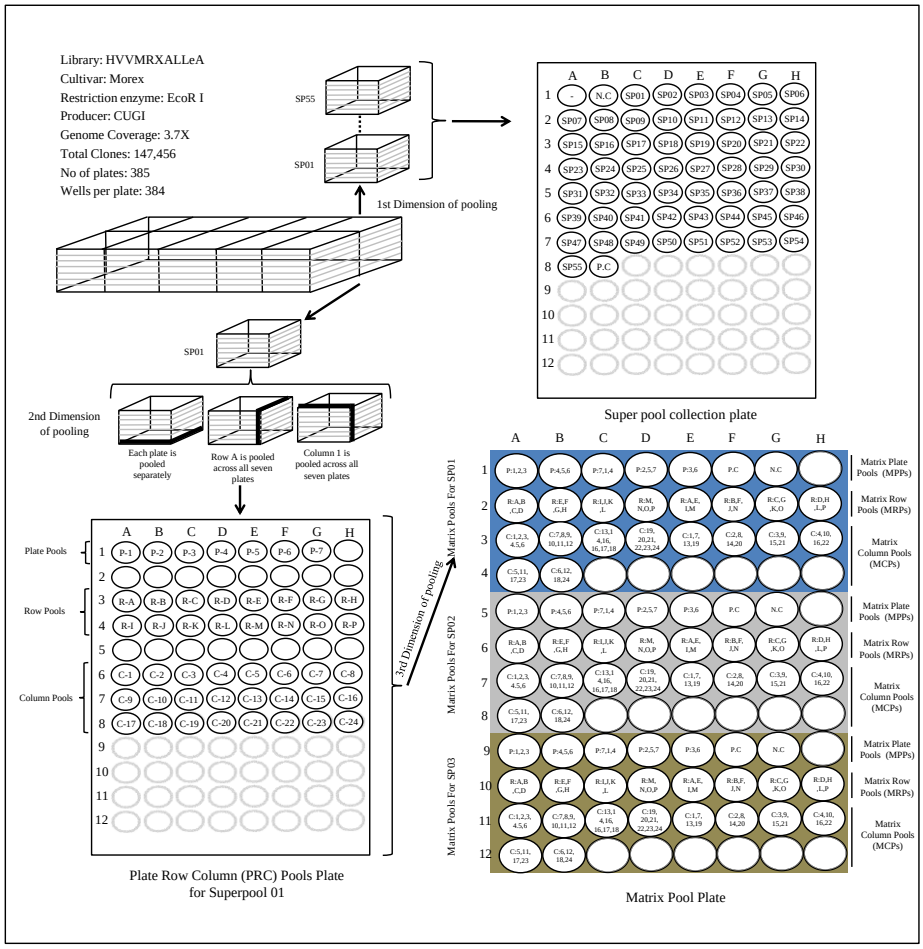


Figure2-1. Steps of three dimensional BAC library pooling constructed and provided by Amplicon Express: Initially, each set of 7 consecutive plates of the library are pooled to form a Superpool (SP). All SPs are arranged in a 96-well plate as superpool collection plate. In the second step, each SP of seven plates is further separated into 7 Plate pools, 16 Row pools, 24 Column pools and arranged in the Plate Row Column (PRC) plate as an intermediate step for the third dimension of pooling. Finally, in matrix pooling step (3rd dimension), each set of pools (Plate, Row and Column) of the PRC plate were then independently further pooled to create 5 Matrix Plate Pools (MPPs), 8 Matrix Row Pools (MRPs), and 10 Matrix Column Pools (MCPs), respectively, in total 23 Matrix Pools (MPs) for each SP. This step of pooling resulted in each BAC being presented in two independent MPs for each type of MP (plate, row, or column). Corresponding Matrix pools of each SP has occupied 1/3 of the Matrix Pool Plate. To identify a BAC address the researcher needs to screen the SP collection plate as Round I PCR and the respective Matrix Pool section as Round II of PCR screening workflow. P.C and N.C stand for Positive and negative control, respectively. For details of screening and deconvolution see Figure2-2.

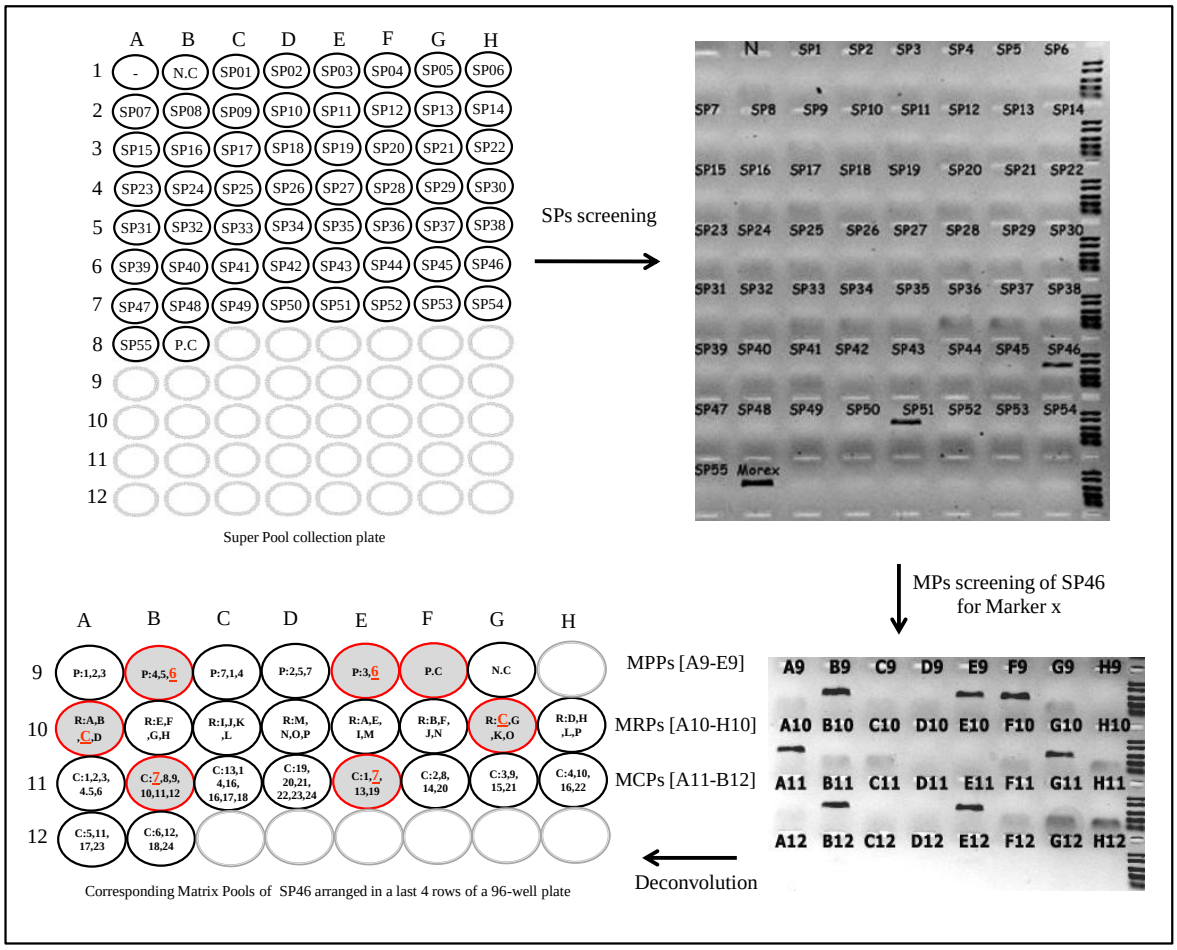


Figure2-2. PCR-based screening of the Matrix 3D pooled library: Superpools (SPs) screening followed by Matrixpools (MPs) screening of marker x is shown. The MPs screening step for only one SP (SP46) is illustrated. In this step, 5 Matrix Plate Pools (A9 to E9), 8 Matrix Row Pools (A10 to H10), and 10 Matrix Column Pools (A11 to B12) are screened and deconvoluted. For instance, to identify the plate in which the BAC clone is located, the plate in common between the two amplified MPPs (B9 and E9) is the initial plate number harboring the BAC clone (plate6; highlighted in red) in the corresponding SP (here SP46). This is because each BAC is presented in two independent MPs for each type of MP (plate, row, or column) to cross-refer each other, see Figure 2-1. The same logic enables deconvolution of respective row letter and column number. Therefore, the corresponding BAC position in SP46 has been identified as row C and column 7 of plate 6. Abbreviation correspondences; P: Plate, R: Row, C: Column, P.C: Positive Control (F9), N.C: Negative Control (G9) and red circles show the corresponding amplified coordinates.

2.4. Marker development

In addition to the above mentioned barley transcript maps a published virtual gene order map of barley chromosome 2H (2H-genome zipper, Mayer et al. 2011) was considered as additional resource for marker development. This dataset has been derived on the basis of a dense transcript map (Close et al.

2009), chromosome-specific sequence datasets and conserved synteny analysis between barley, rice, *Sorghum* and *Brachypodium*.

2.4.1. Genome zipper derived markers

Virtually ordered genes were initially assessed for their potential of being employed in genetic marker development and subsequent physical map anchoring. For this assessment, 139 gene models sequentially ordered in the interval 137.5 cM to 151.4 cM of barley chromosome 2H genome zipper were selected. Roche/454 GSFLX Titanium survey sequence reads (454 reads) of barley genome were associated with gene models. For gene models that more than a single 454 sequence reads was available, the selected 454 read used in primer design considered to acquire highest sequence quality score among all and more than 200 bp length.

Primer design on the basis of individual 454 reads turned out to be inefficient in many cases due to the limited sequence length. In the following experiments, it was therefore the aim to gather more sequence information for primer design. To achieve this goal the complete set of shotgun reads assigned to an individual gene model was assembled to produce "454 read based contigs", if possible. The assembly was performed using Newbler (Roche), version 2.0 applying default parameters (performed by Dr. Rounan Zhou; IPK-Gatersleben, Gatersleben, Germany). All 454 reads associated to the barley genome zippers have been classified either as stringently or not- stringently assigned (Mayer et al., 2011). The former types are the results of Bi-directional Blast Hit (as a stringent blast approach, see below section 2.8.2 for details) performed between each of the model genomes and the 454 reads originated from barley sorted chromosome arms. While, not- stringent class are the first best hits resulted from a single direction blast algorithm applied between the respective sequence dataset (Mayer et al., 2011). Thus, in the current study, genome zipper gene models represented either by 454 read based contigs or stringently assigned singletons (resulted from assembly analysis) were utilized for primer design applying the parameters mentioned in section 2.5 with optimal target size = 500 bp (range 180 to 700bp).

2.4.2. CAPS and dCAPS markers

The majority of genes of the chromosome 2H genome zipper have not been integrated on the basis of genetic mapping information but on the basis of conserved synteny. While exploiting the genome zipper resource for anchoring the physical map of chromosome 2H it was, in the initial pilot phase of the evaluation of the resource, the aim to convert as many as possible genes in genetic markers. Primer sets that could not reveal immediate presence/absence or insertion/deletion (INDEL) polymorphism between parental genotypes were sequenced (ABI PRISM® 3730 DNA Analyzer, Applied Biosystems, CA, USA) for SNP identification. Cycle sequencing experiments and sequence analysis were performed as described by (Shahinnia et al., 2011). Identified SNPs were exploited for the development of either one of the following classes of markers: (A) CAPS markers (Cleaved Amplified Polymorphic Sequence) were developed as described earlier (Thiel et al., 2004; Vu et al., 2010). (B) In case SNP were not leading to the formation or elimination of a restriction site of commonly used and commercially available endonucleases, a dCAPS primer (derived Cleaved Amplified Polymorphic Sequence) (Neff et al., 1998; Komori and Nitta, 2005) was designed using dCAPS Finder (http://helix.wustl.edu/dcaps/dcaps.html). Annealing temperature was adapted for each particular primer pairs. Other conditions for PCR, digestion and electrophoresis were the same as for CAPS markers. (For details of all marker type: SupData_Table1)

2.4.3. High Resolution Melting (HRM) curve analysis

As an alternative method for sequence polymorphism detection the high resolution melting (HRM) curve analysis (Bennett et al., 2003; Herrmann et al., 2007) was applied. By this method differences in DNA sequence are determined by due to differential melting properties of amplicons derived by different alleles. For this analysis, PCR was set up using QIAgility liquid handling instrument (QIAGEN, Germany). The PCR reagent mixture consisted of 1 µl of genomic DNA (20ng/µl), 0.7 µl of primers mix (5 pmol/µl each), 5 µl of HRM mix and 3.3 µl of RNase-free water (last two reagents: Type-it® HRM PCR kit, QIAGEN, Germany). Amplification was achieved by a touchdown PCR protocol: an initial denaturing step of 5 min at 95 °C was followed by 45 cycles with denaturation at 95 °C for 20 s and extension at 72 °C for 20 s. The annealing temperature was decreased in 1 °C

increments from 65 °C in the first cycle to 60 °C after the 5th cycle and was then kept constant for the remaining 40 cycles (always 20 s). PCR amplification was performed on a Rotor-Gene Q real-time PCR Thermocycler (Qiagen, Hilden, Germany). HRM was performed as follows: pre-melt at 75 °C for 90 s, and melt at a ramp of 65 °C to 95 °C at 0.05 °C increments every 2 s. The fluorescence data were acquired at the end of each annealing step during PCR cycles and each of the HRM steps with automatic gain optimization. High resolution melting curve analysis was performed using the HRM module of Rotor-Gene Q realtime PCR Thermocycler (Qiagen, Hilden, Germany). The HRM curve of the respective primer set for each individual was visually scored.

2.5. Primer design

Primers were designed using the public software tool Batchprimer 3 (You et al., 2008). Primer picking parameters were set as follows with the optimal Tm = 62 ºC (range 60 to 64 ºC with difference of maximum 2 ºC between forward and reverse primers) and optimal GC content = 55% (range 45 to 60%). The optimal target size varied depending on input sequence size; 200-500 bp for the sequences used in STS marker development (454 shotgun reads) and ≥ 1300 kb for those used in FISH probe design (see section 2.7) (for example: sequenced BACs). Sequences that did not satisfy these conditions were rejected.

2.6. Genetic mapping

To validate the predicted gene order newly developed genome zipper-derived markers were mapped to the DH population Morex x Barke (Close et al., 2009). Genotyping information was entered into the published genotype information file of this population. Software MAP MANAGER QTX (Manly et al., 2001) was used to fit new marker data into the previously structured linkage groups applying the command “Distribute”. JoinMap V4.0 (Kyazma, The Netherlands) was operated for grouping of markers (LOD score = 4.0) and subsequent marker order determination. The Kosambi mapping function (Kosambi, 1944) was applied for converting recombination units into genetic distances. Graphical genotypes of the resulting chromosome 2H map were visually inspected for consistency.

2.7. Cytogenetic mapping

The resolution provided by barley genetic maps employed for physical map anchoring is limited and does not allow resolving the order of contigs assigned to the genetic centromere and the adjacent pericentromeric region of barley chromosomes which exhibit reduced recombination frequency. These regions are characterized by large numbers of markers clustered to the same genetic position. Cytogenetic placement of gene/contig information to barley chromosomes by Fluorescence in situ hybridization (FISH) might help to add additional resolution for anchoring and ordering of BAC contigs of the physical map. A selection of chromosome 2H BAC contigs was utilized for exploring the potential of cytogenetic anchoring. Two main limitations exist for this method: (i) presence of repetitive DNA in the FISH probe which would lead to innumerous unspecific signals on all barley chromosomes, thus specific allocation is compromised. (ii) Size of the FISH probe – the probe needs to be of sufficient size to allocate sufficient amounts of fluorescent label to one location that still can be detected on the basis of fluorescence microscopy. Barley mitotic chromosome spreads were used as target material.

2.7.1. Probe and chromosome preparation

In-silico -defined unique regions of barley genomic DNA sequence resources [Whole Genome Shotgun sequence contigs (WGS contigs) or sequenced BACs (for repeat masking analyses see section 2.8.3)] were amplified by PCR from genomic DNA of barley cv. Morex [5ng/µl]. The PCR was performed as described in section 2.2. PCR amplicons were purified using QIAquick PCR Purification Kit (QIAGEN, Germany) according to the manufacturer's instruction. The 5S ribosomal probe was generated by PCR as earlier described (Fukui et al., 1994). Mitotic chromosome spreads of barley cv. Morex were prepared using the reported spreading technique (Kato et al., 2006). Probes were directly labeled by nick translation with Texas red-dUTP or Alex-488- dUTP (both from Invitrogen), respectively, as described previously (Kato et al., 2006) (For details of probe information: SupData_Table2).

2.7.2. Fluorescence in situ hybridization (FISH)

For single copy FISH, a previously published procedure has been followed (Ma et al., 2010). In brief, the chromosome slides were treated with 45% acetic acid and pepsin (0.1 mg mL^{-1} in 10 mM HCl) for 10 min separately, post-fixed in 4% formaldehyde in 2×SSC for 10 min, dehydrated in an ethanol series (70, 90 and 96%), and air dried. The hybridization mixture contained 50% deionized formamide, 2×SSC, 1×TE, 50 ng μl^{-1} of each single copy probe and 10× excess of salmon sperm DNA. The hybridization mixture and the treated chromosome slides were denatured together on a heating plate at 80^{o}C for 2 min and incubated in a moist chamber at 37^{o}C overnight (4 to 48h). Post-hybridization washing was done in 2×SSC for 20 min at 55^{o}C. After dehydration in an ethanol series (70, 90 and 96%), the slides were air dried at room temperature (15 to 30 ^{o}C) and counterstained with 4',6-diamidino-2-phenylindole (DAPI) in Vectashield (Vector Laboratories). Images were taken using a cooled CCD camera (Spot 2e, Diagnostic Instruments) on epifluorescence microscope (Axioplan 2, Zeiss) with a Plan Apochromat 63×/1.40. Image process was performed by photoshop 6.0. The images were pseudocolored and merged. Furthermore, brightness and contrast improvement were done on the whole image using command "level". The final resolution was increased to 300 dpi. The chromosome preparation and FISH analysis was provided by Dr. Lu Ma (CSF group, IPK-Gatersleben, Germany) as part of a collaboration for establishing this procedure for FISH mapping of BACs in barley.

2.8. In-silico sequence analysis

All sequence comparison analyses run via computer tools in this study are referred to as *in-silico* experiments. These consisted of different BLAST (Basic Local Alignment Search Tool) types including DNA to DNA (BLASTN) and DNA to protein (TBLASTX) sequence homology search and repetitive sequence identification (Altschul et al., 1990).

2.8.1. BLASTN sequence homology search

A dataset including barley WGS contig linked to the barley FPcontigs (FingerPrinted contig) was kindly provided by Thomas Nussbaumer (Institute for Bioinformatics and Systems Biology, Helmholtz Zentrum Munich, German Research Center for Environmental Health, 85764 Neuherberg, Germany).

This assignment was possible by availability of BACend sequences (BES) and sequenced BACs (sBACs) present in the unpublished physical map assembly of the barley genome (Stein et al. unpublished data).

All BLASTN sequence homology searches were performed at a threshold of E-value ≤ 1E-10. The outputs were parsed using a Perl script (kindly provided by Dr. Mihaela Maria Martis; Institute for Bioinformatics and Systems Biology, Helmholtz Zentrum Munich, German Research Center for Environmental Health, 85764 Neuherberg, Germany) under the parameters of identity ≥ 99% and alignment length ≥ 300bp.

2.8.2. TBLASTX sequence homology search

TBLASTX comparison of barley WGS contigs were performed against *Brachypodium*, rice, and *Sorghum* proteins (*Brachypodium* genome annotation v1.2 [ftp://ftpmips.helmholtzmuenchen.de/plants/*Brachypodium*/v1.2]; rice RAP-DB genome build 4 [http://rapdb.dna.arc.go.jp]; *Sorghum* genome annotation v1.4 [http://genome.jgi-psf.org/Sorbi1/Sorbi1.download.ftp.html]; (Paterson et al., 2009)). In order to identify the genes in model genomes a Bi-directional Best Hit (BBH) blast [In BBH approach gene A (from genome 1) and gene B (from genome 2) are referred as BBHs if there is no gene more similar to A other than B and vice versa (Overbeek et al., 1999)] was performed. This was done on the basis of a barley whole genome shotgun (50x haploid genome coverage, Stein et al. unpublished data) sequence assembly and each individual model genomes. The blast reports were filtered for BBHs according to the following criteria: (1) the best hit display with a similarity ≥ 75% and (2) an alignment length ≥ 30 amino acids using a Perl script kindly provided by Michaela Maria Martis.

2.8.3. Selection of single copy sequences for FISH

Barley genomic sequence information associated to a set of selected FPcontigs (section 2.7) was used for FISH probe development. In order to provide repeat-free and single –copy sequence for probe design, barley genomic sequences where treated by two levels of repeat masking. First the repetitive part was marked after being identified using Vmatch (http://www.vmatch.de) against the MIPS-REdat

Poaceae v8.6.2 repeat library (Mayer et al., 2011) by applying the following parameters: identity ≥70%, 50-bp minimal length, exdrop 5, and e-value 0.001 (This analysis was provided to me by Dr. Heidrun Gundlach (Institute for Bioinformatics and Systems Biology, Helmholtz Zentrum Munich, German Research Center for Environmental Health, 85764 Neuherberg, Germany). Furthermore the repeat masking was followed via Mathematically Defined Repeat (MDR) analysis (Wicker et al., 2008). The selected sequences were virtually fragmented into overlapping 21-mer, and each 21-mer was assessed for its frequency in the barley MDR index generated from an 8x barley whole genome shotgun sequencing dataset (Stein et al., unpublished data). Uniqueness was reached by masking regions that exceeded a 10x repeat threshold (this analysis was kindly performed by Thomas Schmutzer, BIT, IPK-Gatersleben, Germany). Regions that passed both levels of repeat masking and were longer than 1,500 bp were used for primer design. Primer design and PCR experiments were performed as described in section 2.5 and section 2.2, respectively (For details of FISH probe information: SupData_Table2).

2.8.4. Data produced under collaborative efforts and provided to the current study

During the current project, parallel activities were under progress by IBSC members for anchoring the whole genome barley physical map to the individual barley chromosomes. Different data were exchangened among projects including the data produced in this current study that led to complementation of individual activity and to prevent the production of redundant results. Moreover, alternative anchoring methodologies such as FISH were managed to be examined during the current work which was also performed on the basis of cooperation. Therefore, for the purpose of clarification the results produced during the current study itself and the resources provided under collaboration are as listed below:

- All steps of PCR-based anchoring of chromosome 2H; from STS primer development to library screening and BAC address deconvolution: *Naser Poursarebani*
- All steps of *in-silico* anchoring of chromosome 2H physical map to single map of genome zipper and the analyses of recombination frequency: *Naser Poursarebani*
- All genome zipper validation experiments: *Naser Poursarebani*
- Design and the management of FISH experiments along with the FPcontigs selection, corresponding associated sequence identification, primer development and PCR product purification: *Naser Poursarebani*
- Steps of FISH experiments: *Dr. Lu Ma and Dr. Andreas Houben* [1]
- K-mer analysis for repeat masking underlying single copy FISH probe detection: *Thomas Schmutzer* [2]
- *In-silico* assailment of WGS contigs to individual barley genome FPcontigs: *Thomas Nousbammer* [3]
- *In-silico* assignment of barley FPcontigs to the individual barley chromosome arms and Micro array based anchoring: *IBSC members* [4]
- 454 reads assembly of barley genome zippers: *Dr. Rounan Zhou* [5]

[1)] CSF group, IPK-Gatersleben, Germany
[2)] BIT group, IPK-Gatersleben, Germany
[3)] MIPS, Neuherberg, Germany
[4)] International Barley Genome Sequencing Consortium (IBSC; www.barleygenome.org)
[5)] GED group, IPK-Gatersleben, Germany

3. Results

A physical map densely anchored to the genetic map is a basic prerequisite for the hierarchical clone-by-clone based sequencing of the barley genome. By itself, a physical map provides a resource that would greatly facilitate map-based gene isolation in barley. Aspects of comparative genome analysis between barley and related grass species could be addressed at high precision and important characteristics of the barley genome like the ratio of physical map length and distribution of recombination along the chromosome could be analyzed at much higher comprehension. The present work was initiated as a concomitant activity to the international effort of developing a generic physical map of the barley genome. This larger collaborative project aimed at high information content fingerprinting (HICF) (Luo et al., 2003) of several hundred thousand BAC-clones (Schulte et al., 2009; Schulte et al., 2011) in order to generate up to 14-fold haploid genome coverage or even more. This general activity was expected to deliver in the range of ten thousand individual BAC contigs representing the seven barley chromosomes. The assignment of these contigs to the individual chromosomes, however, was not covered by the generic project but was left to be contributed by supporting complementary activities. These were expected to exploit existing public molecular marker maps of barley and generate marker / BAC clone relationships that would allow anchoring of BAC contigs to specific chromosomal regions. Furthermore, there was an urgent need for the exploration of new, alternative sources of marker information since published marker maps would likely be exhausted before all contigs of a chromosome could be anchored successfully. The aim of this study was to utilize PCR-based genetic anchoring of the physical map of barley chromosome 2H. Since in the duration of the project innovative new technological applications as well as unforeseeable genomic sequence datasets became available that provided substantial independent anchoring information for the entire barley genome. This allowed the unique opportunity to this project to combine all information for a by far more comprehensive anchoring of the chromosome 2H physical map. This two-tiered approach is reflected in the following presentation of the results.

3.1. Anchoring of the physical map of chromosome 2H to the genetic map of barley

3.1.1. PCR-based anchoring based on public marker resources

A generic assembly of 571,007 high quality fingerprinted BAC clones comprising 9,435 BAC contigs (FingerPrintedContig, FPcontig) corresponding to 13x genome coverage for the entire barley genome was developed by other efforts (IBSC and Stein et al.; unpublished) in duration of this project and was made available as a resource to this project. A BAC library HVVMRXALLeA (Schulte et al. 2011) used in this effort was used to develop multidimensional BAC DNA pools which could be used in this study for generating marker/gene-BAC clone relationship information.

In the attempt of genetically anchoring BAC contigs of the above mentioned barley physical map to chromosome 2H, published genetic maps - for the convenience called here 'IPK' (Stein et al. 2007) and 'Okayama' (Sato et al. 2009) maps, respectively - were used for screening of BAC library HVVMRXALLeA. These two maps were selected first since they comprised many (IPK) or even exclusively (Okayama) PCR-based STS markers (Sequence Tagged Site). These could be directly employed for PCR screening of multidimensional BAC DNA pools. Fifty (11 GBM and 39 GBS markers) and 492 chromosome 2H markers of the IPK and Okayama maps, respectively, were used (Table3-1). In order to instantly identify marker/BAC relationships, any given marker was examined against a 3-dimensionally (3D) pooled BAC library. For 495 of the above mentioned 542 markers a specific amplicon could be generated from genomic control DNA. Subsequently, these were used for screening of superpools and matrixpools of the pooled BAC library. 414 of these markers were further used for screening of matrixpools. The remaining (81) either revealed no or more than the expected number of signals (considering the library coverage of ~4x; four hits for single copy genes was expected) in the 55 superpools. For a total of 345 markers, matrixpool screening revealed at least one specific BAC address. In several cases more than one address were obtained leading to altogether 532 BAC addresses. After consulting the BAC fingerprint assembly the 345 markers/532 BAC addresses could be assigned to 216 non-redundant FPcontigs for chromosome 2H (For details of Marker/BAC and Marker/FPcontig relations: SupData _ Table 5 and 6). In order to reveal unambiguously a BAC address by matrix pool screening any given marker had to amplify 6 positive PCR signals (2 in Matrix

Plate Pools, 2 in Matrix Row Pools and 2 in Matrix Column Pools; see material and methods Figure2-2). PCR failure in any of these pools would result in failure of BAC identification. Moreover, with eight or more PCR signals BAC address de-convolution would be complicated and could lead to false positives. Therefore, only those markers that produced 6~7 signals in the matrix pools screening step were considered for BAC address deconvolution. For 81 markers matrix pool screening did not lead to coordinates that could be deconvoluted unambiguously to BAC addresses.

Table3-1. Summary of PCR-based genetic anchoring of barley physical map to chromosome 2H

Marker source	Total markers	Amplified on gDNA	Test on Superpool	Test on Matrix pool	Anchored markers	Total Number of BACs identified in the assembly	Total Number of FPcontigs
Zipper[1]	1300	1114	1099	896	759	1042	316
Okayama[2]	492	460	453	380	320	499	212
IPK[3]	50	47	42	34	25	33	17
Total	1842	1621	1594	1310	1104	1574	Non.R=427[4]

[1)] Markers originated from barley virtual gene order map (2H-genome zipper) developed by Mayer et al., (2011).
[2)] Transcript-derived markers developed by Sato et al., (2009) genetically mapped on chromosome 2H.
[3)] Transcript-derived markers developed by Stein et al., (2007) genetically mapped on chromosome 2H.
[4)] Non. R reflects the total non- redundant FPcontigs genetically anchored to chromosome 2H by corresponding marker resources.

3.1.2. PCR-based anchoring based on predicted marker resources: exploring the barley genome zippers

While working on the anchoring with published marker resources a virtual gene order map ("barley genome zippers", Mayer et al., 2011) became available. This provided a novel and comprehensive resource for marker development for anchoring of the physical map. A total of 21,766 barley genes were ordered linearly along the seven barley chromosomes. In case of chromosome 2H the linear order of 3,616 genes was predicted on the basis of shotgun sequencing of sorted chromosomal DNA and integration of syntenic gene content information from rice, *Brachypodium* and *Sorghum*. This outnumbered at least 7-fold marker resources from other genetic maps of barley developed so far (see introduction section 1.3). Nevertheless, it should be kept in mind that these barley "genome zippers" are developed largely on the information of conserved synteny. There are known limitations or interruptions in conserved synteny originating from evolutionary dynamic changes between related

genomes which may have influenced the genome zipper gene order prediction. As an example, in case of chromosome 2H, 87% of the genes virtually positioned in the 2H "genome zipper" have an inferred order/position deduced from conserved synteny. Before the genome zipper information can be used straight ahead for physical map anchoring an estimation of the inherent error was required. Thus an initial validation of the virtual ordering of genes in the 2H genome zipper has been performed (section 3.1.2.1 and section 3.1.2.2) before this dataset could be extensively used for creation of physical map anchoring of chromosome 2H (section 3.2).

3.1.2.1. Evaluation of gene order predicted by barley genome zippers on the basis of comparison to independent gene-based marker maps

The linear gene order of 21,766 gene models has been predicted by barley genome zippers of chromosomes 1H to 7H. This model is based on the framework of an integrated genetic map comprising 2,785 gene based markers (Close et al., 2009) thus the genetic position for 18,981 genes (87%) was mainly inferred from conserved synteny information from sequenced grass genomes (Mayer et al., 2011). All these genes could be related to barley EST unigenes like those represented on Affymetrix Barley1GeneChip (Close et al., 2004). This provided the opportunity to compare the inferred genetic position of a subset of genes of the genome zipper to the genetic position of the respective EST unigenes as available from a transcript derived marker (TDMs) map (Potokina et al., 2008) obtained by transcriptome analysis in a segregating mapping population of barley by help of the Affymetrix Barley1GeneChip. Six-hundred and three of a total of 1,596 TDMs were found in common to genes in the genome zippers. Of these, 324 were present on the basis of an inferred position whereas 279 were among the genetically mapped genes from the framework map (Close et al. 2009). 574 of the genes/markers shared between both maps (95%) were found at the same chromosomal allocation, thus could be used for marker order comparison between the two maps employing the Spearman's Rank-order Correlation (SRC) (Spearman, 2010). Of the common genes (603 genes), 29 genes (5%) were found to be mapped to different chromosomes of the two maps (18 genes positions were inferred from synteny and the remaining 11 genes were mapped by SNP based markers of the genome zippers framework) via visual inspections (Table3-2). The level of SRC ranged from 0.93 (for 7H) to 0.99 (for

5H) with average of 0.96 (Table3-3 and Figure3-1). In other words, on average 96% of all genes that shared *the same chromosomal allocation* between genome zipper and TDMs map exhibited the correctly predicted position.

Table3-2. Chromosomal allocation of Genes/Markers shared between barley genome zippers and transcript derived markers (TDMs) from the map of Potokina et. al (2008)

Chromosome ID	No. of shared genes between the two maps	No. of genes mapped to different chromosomes	Chromosomes
1H	66	3	7H(1),5H(2)
2H	101	9	1H(2),3H(1),4H(1),5H(2),7H(3)
3H	113	7	4H(3),5H(3),6H(1)
4H	70	1	1H(1)
5H	98	6	2H(2),6H(2),7H(2)
6H	66	1	5H(1)
7H	89	2	2H(1),5H(1)
Total	603	29	

[1)] The number in parentheses is the total number of markers showed conflict with the respective chromosome.

Table3-3. Spearman rank order correlation (SRC) between genome zippers and the transcript derived markers (TDMs) map developed by Potokina et. al (2008) for all barley chromosomes

Chromosome ID	No. of shared markers between the two maps	No. of shared markers with the same chromosomal allocation	Spearman Rank order Correlation (SRC)	P-value
1H	66	63	0.95	1.26e-33
2H	101	92	0.98	5.38e-43
3H	113	106	0.96	4.18e-26
4H	70	69	0.99	1.59e-56
5H	98	92	0.99	1.71e-45
6H	66	65	0.95	5.73e-23
7H	89	86	0.93	1.14e-37
Total	603	573	Average:0.96	

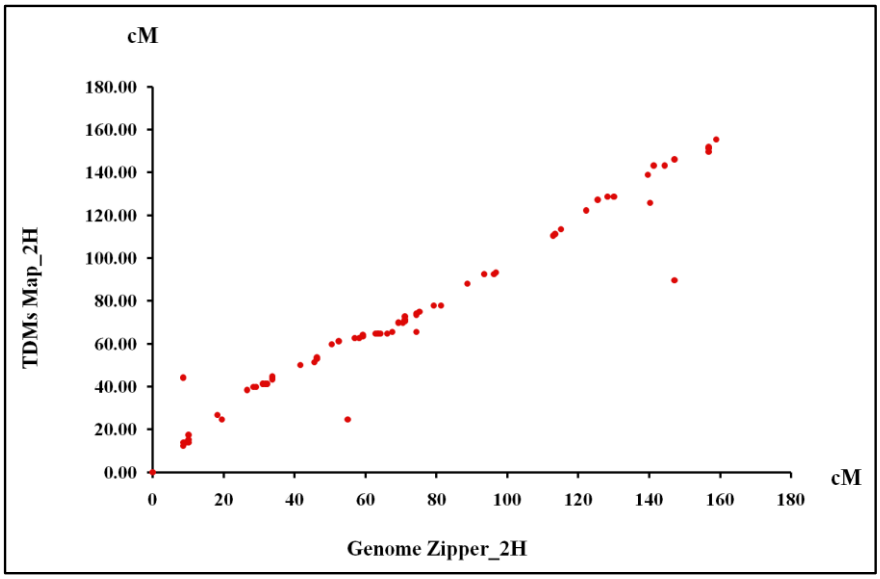


Figure3-1. **Marker order between barley chromosome 2H genome zipper (Mayer et al. 2011) and the genetic map of chromosome 2H from Potokina et al., (2008)**. A Spearman rank-order correlation of 0.98 (P-value = 5.38E-43) was calculated for marker order of chromosome 2H of the two maps.

In total, 91% consistency (the 9% inconsistency included 5% inconsistency in chromosomal allocation + 4% inconsistency revealed by SRC) between barley genome zippers (Mayer et al. 2011) and the TDMs map (Potokina et al. 2008) was observed in this study. Considering that the TDM map bears an inherent 5% prediction inaccuracy (Luo et al., 2007, Potokina et al., 2008 and Prof. Dr. Mike Kearsey; personal communication) the gene order prediction accuracy that eventually could be estimated for barley genome zippers model in this analysis was about 96%.

3.1.2.2. Experimental evaluation of gene order predicted by chromosome 2H genome zipper

3.1.2.2.1. Genetic mapping of genes

The predicted gene order provided by barley chromosome 2H genome zipper was also experimentally evaluated in order to estimate the risk of false genetic assignment of physical map contigs in the case of using genome-zipper based markers for anchoring. One-hundred and thirty-nine genes sequentially ordered in the interval 137.5 cM to 151.4 cM of barley chromosome 2H genome zipper (Mayer et al., 2011) were selected for experimental verification by genetic mapping. PCR primers were designed on the basis of the Roche/454 GSFLX Titanium survey sequence information, which is underlying the

virtual gene order model. All 139 gene models were selected for amplification and sequencing for the determination of sequence polymorphisms between parental genotypes (cultivars ‘Morex’ and ‘Barke’) of a Doubled Haploid (DH) mapping population. For 128 genes (92%) a specific amplicon was obtained and sequenced (Table3-4). Twenty-eight of the 128 amplified gene models were corresponding to genes already genetically mapped by previously established SNP markers (Close et al., 2009). A polymorphism between the cultivars ‘Morex’ and ‘Barke’ was found in 47 out of the 128 genes (36.7%) including 12 genes that had already been mapped previously (Close et al. 2009) (Table3-5). Single nucleotide polymorphisms (SNP) provided the most frequent class of sequence diversity in polymorphic genes (38 out of 47). All could be mapped genetically either by CAPS or dCAPS analysis (Table3-5 and SupData_Table1). Six STS markers were mapped directly in agarose gel electrophoresis due to the presence of a size polymorphism caused by insertion / deletion in one of the parental alleles.

Three STS revealed a presence/absence polymorphism between Morex and Barke (Table3-5) of which one could be genetically mapped. The other two markers GM.3322 and GM.3421 were not placed to any of the barley chromosomes. They exhibited a severe segregation distortion and were skewed towards parental genotype ‘Barke’ with a χ^2-value of 10.33 (P-value = 0.001) and 63.48 (P-value = 0), respectively. Genetic mapping of genes for which before only an inferred position on barley chromosome 2H was known confirmed in most cases the predicted chromosomal location (Figure3-2). Ninety-three percent of the genes (42/45) were mapped to the expected position predicted in the barley chromosome 2H genome zipper (Figure3-2).

Table3-4. Chromosome 2H genome zipper gene models of the 14 cM experimental validation window and their usefulness for anchoring the physical map

Syntenic to	Gene Models selected for primer design	Amplified Gene models	Polymorphic gene models	Allocated to 2HL			
				Genetic Mapping	wheat/barley addition lines[1,2]	FPcontig info	Total (cumulative)
Brachypodium/Rice/*Sorghum*	78	72	27	25	46(10)	8	69
Brachypodium/Rice	3	3	-	-	3(0)	1	3
Brachypodium/*Sorghum*	19	18	5	5	15(2)	4	18
Rice/*Sorghum*	-	-	-	-	-	-	-
Brachypodium	14	11	7	7	7(3)	1	11
Rice	6	6	3	3	4(1)	-	6
Sorghum	9	9	2	2	6(1)	2	7
SNP markers alone	10	9	3	3	8(2)	1	9
Total	139	128	47	45	80(19)[2]	17	123

[1)] Refers to ditelosomic wheat/barley addition lines.
[2)] Numbers in the parentheses indicate number of markers assigned by genetic mapping.

Table3-5. Categories of polymorphism found between parental genotypes (cultivars 'Morex' and 'Barke') PCR amplicons in the 14 cM of chromosome 2H long arm of the genome zipper

Polymorphism type		Polymorphic gene models	Mapped to 2HL[1]	Following predicted position	Provided by the genetic position in the genome zipper [2]
INDEL		6	6	6	2
SNP	CAPS	32	32	29	7
	dCAPS	6	6	6	1
Presence /Absence		3	1	1	2
Total		47	45	42	12

[1)] Two of the presence/absence markers were not mapped to any of the barley chromosomes.
[2)] their order and positions in the genome zipper were supported by the consensus genetic map framework.

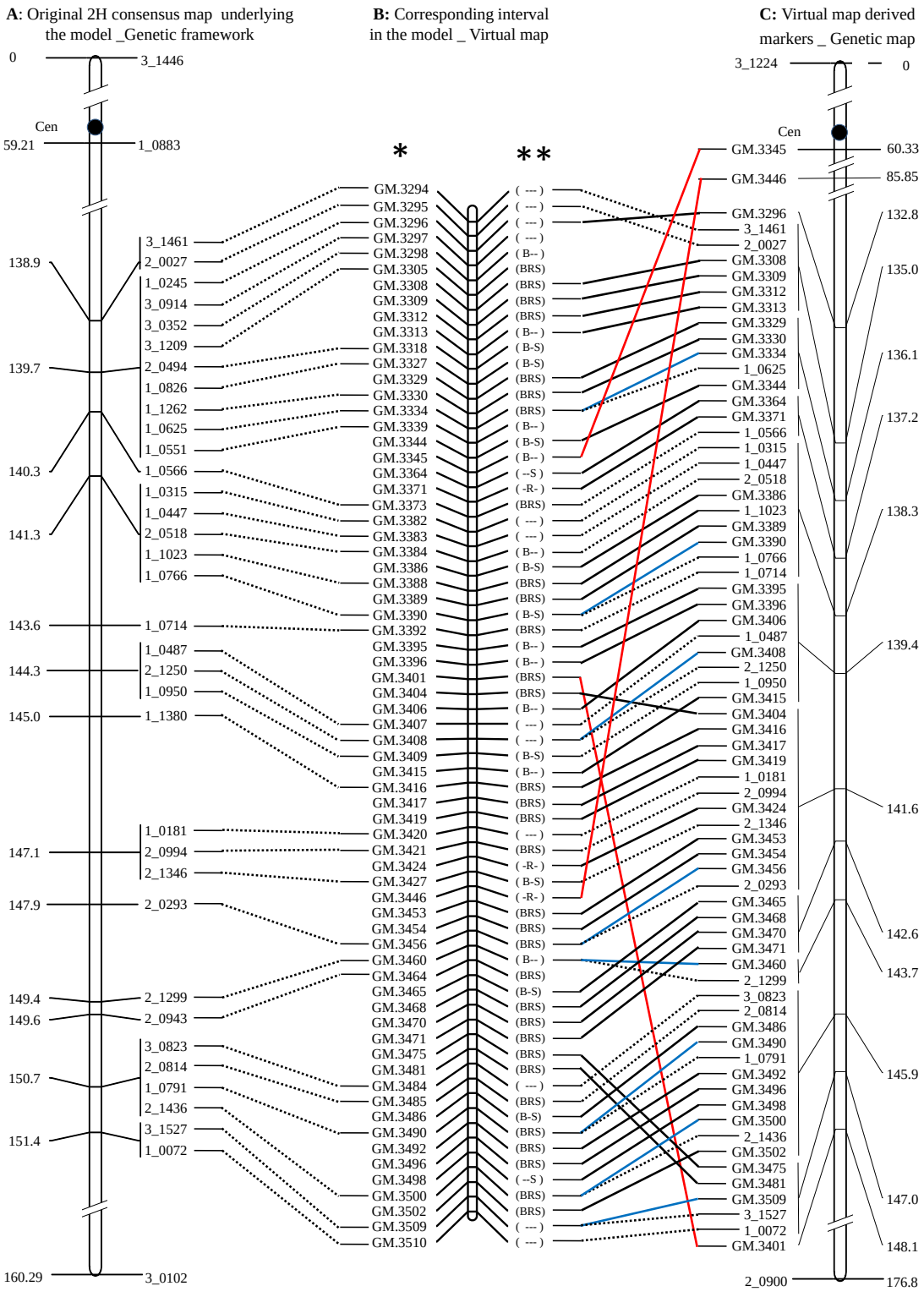


Figure3-2. Integrated virtual map (genome zipper) and genetic map view of a ~14cM region on chromosome 2H long arm: Mapping data of 45 genome zipper derived markers were integrated in a previously established Morex x Barke DH map. (*) Shows barley gene models for which genetically map positions are available (either in the genome zipper, in the Morex x Barke DH map or in both). (**) indicates the presence of orthologous genes from the related sequenced model

genomes for the respective barley gene models. B, R and S stand for support from *Brachypodium*, rice and *Sorghum*, respectively. Dashed lines that connect A to B or B to C indicate the original consensus map markers (Close et al., 2009) underlying the genome zipper. Continuous lines (black, red or blue) that connect B to C indicate the newly developed markers: black lines highlight the markers which follow the proposed order, whereas red lines indicate markers that do not follow the predicted order from the genome zipper. Blue lines pinpoint newly developed markers that have been ordered based on genetic map framework of the genome zipper (positive controls of the analysis).

3.1.2.2.2. Physical mapping of genes by ditelosomic wheat/barley addition lines

Genetic mapping of genes with predicted chromosomal position on the basis of the chromosome 2H genome zipper revealed a high reliability of the predicted map positions. Because of the relatively low level of polymorphism between the parental genotypes of the utilized mapping population, the majority of STS markers selected from the 14 cM target could not be mapped genetically (83 out of 128 genes amplified on genomic DNA). In order to test at least for their chromosomal origin, the aim was to allocate them physically on the basis of PCR analysis of so-called ditelosomic wheat-barley addition lines. Such genetic stocks bear a disomic addition of a specific barley chromosome arm in the background of the entire wheat chromosome complement (Islam et al., 1981). In parallel to genetic mapping, all 128 gene-derived markers from the 14 cM interval of the barley chromosome 2H genome zipper were tested (Table3-4) on a panel of wheat/barley ditelosomic addition lines representing all barley chromosomes except 1H. This chromosome cannot be maintained as single chromosome addition in wheat background. It can, however, be stably inherited together with 6H (Islam and Shepherd, 2000) but a respective confirmed stock was not available for this study. Fifty-two genes could be allocated to chromosome 2HL either because a PCR product was only obtained from barley and the respective ditelosomic addition line and not from wheat (Figure3-3A) or the products of barley and wheat differed sufficiently in fragment size (Figure3-3B).

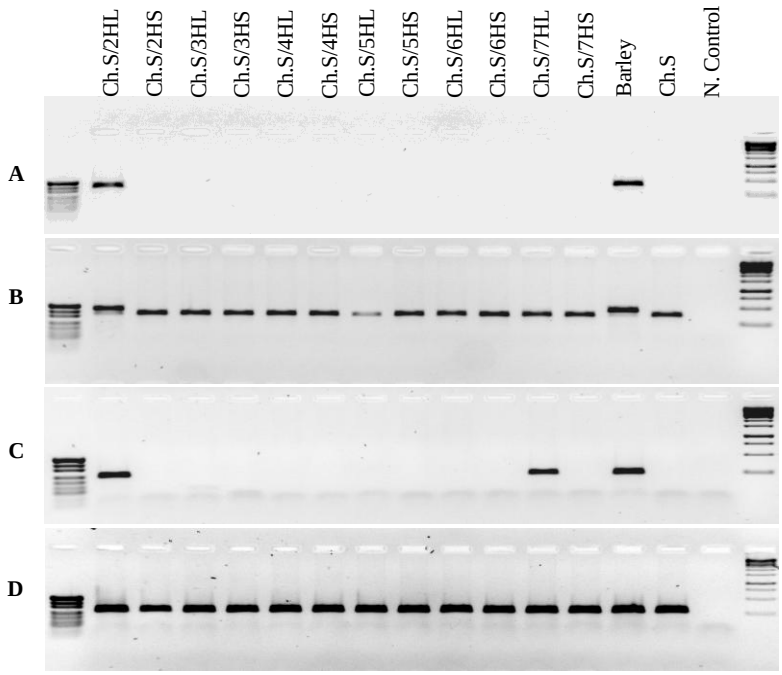


Figure3-3. PCR analysis of ditelosomic wheat/barley addition lines for gene allocation on barley chromosome arms: A and B show assignment of genes/STS markers to the long arm of barley chromosome 2H. C) depicts allocation of a STS marker to 2HL and the long arm of barley chromosome 7H (7HL). D) Shows lack of resolution to identify the chromosomal origin of another STS marker on the agarose gel, such class of markers were further analyzed by High Resolution Melting curve (HRM) analysis for chromosomal origin verification (e.g. Figure3-4). L, S, Ch.S, and N stand for Long arm, Short arm, wheat cv. Chinese Spring, and Negative control, respectively. The barley genomic DNA comes from cv. Betzes. The right and left most lanes of each PCR gel are 1-kb and pUC19 DNA marker ladders, respectively.

Nineteen out of the 52 STS were also mapped genetically (see Table3-4). To this end, 78 genes could be allocated to chromosome 2HL based on genetic mapping (see section 3.1.2.2.1) and physical allocation to chromosome arm addition lines. For the remaining 50 gene models, either a product formed with template DNA from wheat/barley-ditelo 2HL and (an) additional chromosome(s) (Figure3-3C) or the product size obtained from barley and the wheat/ barley addition lines could not be distinguished at the resolution of agarose gels (Figure3-3D). The 50 markers were further characterized either by sequencing of PCR products obtained from the addition lines and comparing the sequence to the original 454 sequence. Or the amplicons obtained from the different addition lines were tested by HRM (high-resolution melting) analysis which is designed to reveal heteroduplex molecules derived from amplicons of mixed alleles or derived from paralogous genes. In 4 cases, sequencing of PCR

products revealed perfect sequence identity between the original 454 sequence and the wheat / barley-ditelo 2HL derived amplicons. In 24 cases the use of high-resolution melting (HRM) curve analysis provided evidence of 2HL specific origin of the respective genes (Figure3-4**)**.

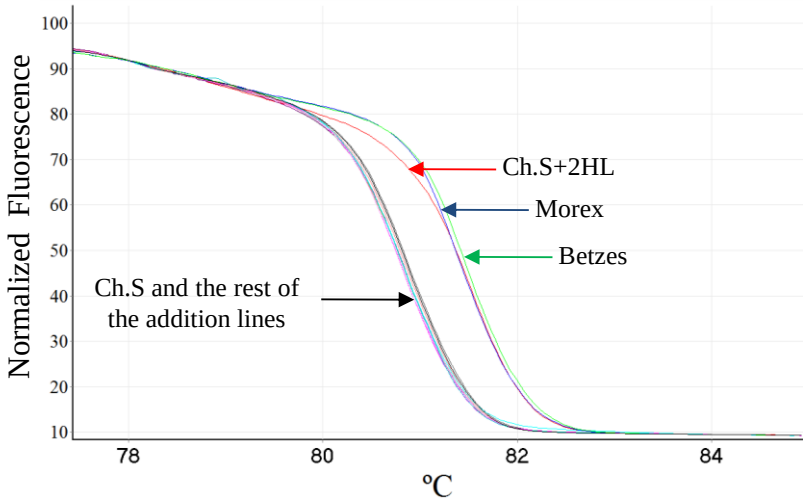


Figure 3-4. A primer set (GM.3432) assigned to barley chromosome 2H long arm by High Resolution Melting (HRM) curve analysis of ditelosomic wheat/barley addition lines. The respective PCR fragments are separated according to their PCR melting (dissociation) curves. CH.S: wheat cv. Chinese Spring.

Altogether, 106 markers (83%) could be confirmed to originate from chromosome 2HL. For the rest of the gene models (22) a chromosomal position could not be determined by genetic or physical mapping. For these genes the respective PCR-based BAC-library screening results and the subsequent anchored physical contig information were considered as described in the next section.

3.1.2.2.3. Physical mapping of genes by assignment to BAC contigs

The genes of the genome zipper validation study were also used for BAC library screening after it had been confirmed by the previous assessments that most of them indeed originated from chromosome 2H. Screening of multi-dimensional BAC DNA pools identified a total of 137 BAC addresses for 96 of the selected gene models (Table3-6).

Table3-6. Summary of the PCR-based library screening for the 14 cM chromosome 2H long arm interval

Number of gene models selected	Amplified gene models	Test on Superpool	Test on Matrix pool	Anchored markers	Number of BACs identified in the assembly	Number of FPcontigs identified
139	128	128	119	96	137	26

These BAC addresses were used to query the genome-wide barley physical map to identify the relevant physical FPcontigs for further analysis. The 137 BACs were distributed in 26 physical contigs. 24 of them could be anchored to the expected zipper interval and followed the predicted ordering (Figure3-5) (For details of Marker/BAC and Marker/FPcontig relations: SupData _ Table8). The order of majority of these physical contigs was further confirmed by the genetic mapping (see section 3.1.2.2.1 for details, Figure3-2). These 24 FPcontigs contained also 17 of the 22 genes for which the chromosomal allocation could not be provided by physical mapping using the wheat-barley addition lines (see section 3.1.2.2.3). Therefore, the total number of genes which could be finally assigned to chromosome 2HL amounted to 123 (106+17) out of 128 genes.

The two FPcontigs 803 (marker GM.3345) and 534 (marker GM.3446) were anchored in the vicinity of the pericentromeric region towards 2HL, thus were located out of the selected target interval (Figure3-5). Localization of 24 out of 26 FPcontigs (i.e. 92%) identified with genes selected for a 14 cM genetic interval from a virtual gene order map could be confirmed to originate from that genetic window of the respective barley chromosome. Therefore, it can be concluded that the genome zipper of barley provides a resource with great potential for anchoring the physical map of the barley genome. Based on these results, entire 2H-genome zippers was considered for STS marker development, library screening and subsequent physical contig anchoring of chromosome 2H. With regards to the highly accurate gene ordering observed in the genome zipper (>94%) the virtual map derived markers also were considered as “genetic” anchors, in this current study.

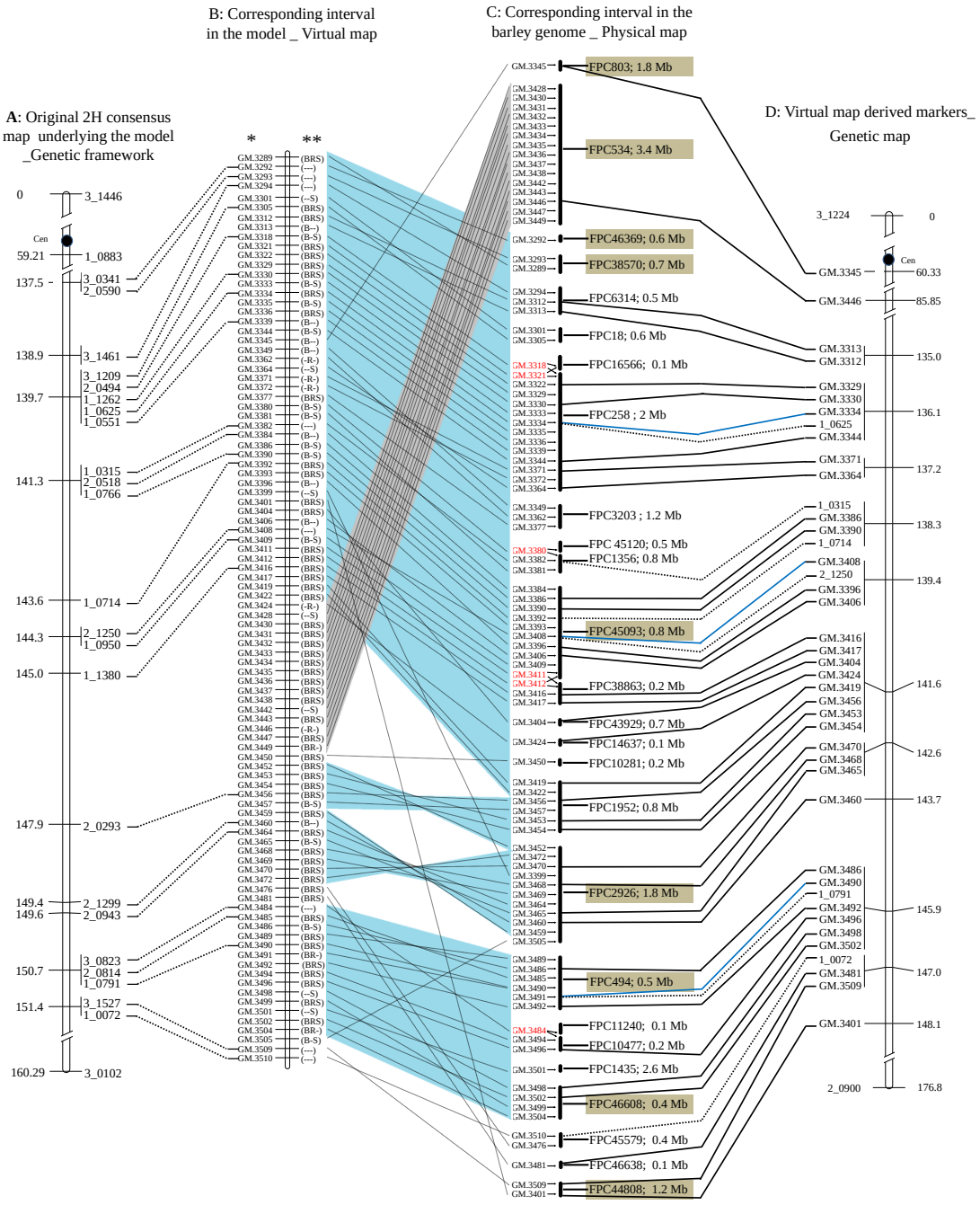


Figure3-5. Comparison of the virtual gene order map (genome zipper), physical map and the new genetic map of the 14cM validation region of chromosome 2HL: (*) shows all gene models (GMs) for which a physical map contig was identified. (**) indicates the presence of orthologous genes from the related sequenced model genomes for the respective barley gene models. B, R and S stand for support from *Brachypodium*, rice and *Sorghum*, respectively. Dashed lines that connect A to B and C to D indicate the original consensus map markers, the genome zipper framework map of Close et al., (2009). Continuous lines (black or blue) connect the newly developed markers to the respective FP (FingerPrinted) contig: black lines are genome zipper derived markers with inferred positions while blue lines pinpoint those genes that have been

ordered based on the genetic map framework of genome zipper (positive controls of the analysis). Markers highlighted in red are shared between the two FPcontigs. FPcontigs highlighted in grey (part C) were selected for FISH analysis (see section 3.4).

3.1.2.3. Large scale anchoring of barley chromosome 2H physical map on the basis of the barley genome zipper

The results obtained from genome-wide validation analysis of genome zipper (section 3.1.2.1) and the pilot study performed for a 14 cM interval of 2HL (section 3.1.2.2) indicated that the position of at least 94% of the genes in the genome zippers has been predicted correctly. Based on these findings, all remaining information of gene order provided by the 2H-genome zipper was considered for anchoring of chromosome 2H physical BAC contigs. All gene models for which 454 read based contigs could be obtained (676 gene models) or stringently assigned 454reads/singletons (710 gene models) were exploited for primer design (1,386 genes in total) (see material and methods section 2.4.1 for 454 read stringency assignment). The stringently assigned 454 reads were preferred since in the pilot study on 2HL, they showed a higher PCR amplification success rate (84%) as compared to non-stringent reads (71.5%). Out of 1,386 genes, primer pairs were designed for 1,141 genes. The remaining 245 genes were omitted either because they were already addressed by other marker resources (171 genes) or could not pass the primer design parameters (74 genes). Including all 159 primer pairs (corresponding to 139 gene models) designed during the genome zipper validation experiment (section 3.1.2.2), a total number of 1,300 primer pairs were designed from the barley 2H-genome zipper and tested on genomic DNA of barley cv. Morex (Table3-1). All primer sets for which a specific amplicon could be obtained from genomic DNA (1114 primer pairs) were directed into the pooled library screening workflow as described earlier. 759 of these assays identified at least a single BAC clone (1042 BAC clone). The remaining primer pairs failed either at the stage of superpool or matrix pool screening. Based on the generic BAC fingerprinting assembly, the 1042 newly identified BAC addresses were located in 316 FPcontigs (Table3-1).

Altogether, the two-tiered approaches of screening multidimensional DNA pools of a BAC library used for physical mapping of the barley genome with 1) markers from published transcript maps (IPK

and Okayama maps) and 2) those derived from the chromosome 2H genome zipper, made it possible to anchor 427 FPcontigs to chromosome 2H. This was corresponding to an estimated total length of 370 Mbp or in other terms 46% of the chromosome (Table3-1 and Figure3-6). Chromosome 2H has been estimated to be 790 Mb long (Suchankova et al., 2006). Using the genome zipper derived information it was possible to double the number of anchored contigs compared to using markers from the two published maps alone (Figure 3-6).

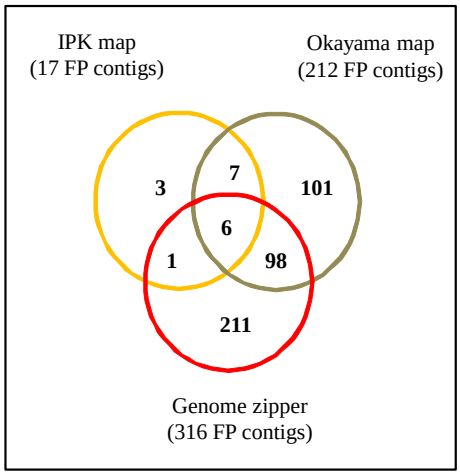


Figure3-6. Contribution of the different genetic maps to PCR-based anchoring of the physical map of chromosome 2H. 427 anchored non-redundant FPcontigs correspond to 46% (370Mb) of the chromosome.

3.2. Chromosome 2H anchoring information from other sources – comparison and integration with this study

3.2.1. Chromosome 2H anchoring information provided by other efforts

427 FPcontigs (370 Mbp, 46% of chromosome 2H) were genetically anchored to chromosome 2H by experimental efforts of the present study. Anchoring of the genome-wide physical map of barley, however, was also a priority for other groups collaborating under the framework of the International Barley Sequencing Consortium (IBSC, http://barleygenome.org). Therefore, it was important to evaluate the specific contribution of chromosome 2H anchoring achieved within this study compared to the general effort (subsequently referred to as IBSC anchoring dataset). An unpublished IBSC dataset of 833 anchored BAC contigs was made available for comparison to the newly generated data.

This IBSC dataset is the result of integrating diverse information (Nussbaumer et al., unpublished results) obtained from 1) BAC end sequencing (Stein, Morgante and Waugh, unpublished data), BAC shotgun sequencing (Stein, Platzer, Scholz and Mayer, unpublished data), 2) gene/marker-BAC relationships obtained from array-based screening of multidimensional BAC DNA pools (Liu et al. 2011), 3) survey-sequence information from sorted barley chromosome arms (Mayer et al. 2011), and 4) Whole Genome Shotgun sequence assembly (WGS-contigs) data of barley (Stein, Platzer, Scholz and Mayer, unpublished data). Of these 833 FPcontigs, 697 (454 Mbp) were assigned only to either short or long arm of chromosome 2H without any genetic allocation whereas 136 contigs (130 Mbp) were anchored by genetic mapping information (Figure3-7A). A comparison of the two datasets of anchored BAC contigs (one set anchored in this study, the other anchored by the IBSC activities) revealed a cumulative number of 867 2H-anchored FPcontigs (referred to as *2H-FP Contigs*) comprising 593 Mbp or 75% of chromosome 2H. The combination of both datasets raised the number of genetically anchored contigs up to 455. Altogether, 386 Mbp of physical BAC contigs where thus anchored to the genetic map and genome zipper of barley chromosome 2H (Figure3-7B). Although the total number of chromosomally anchored contigs increased only by 4% after combining the two datasets (from 833 to 867) the amount of genetically anchored physical map length increased almost three fold.

.

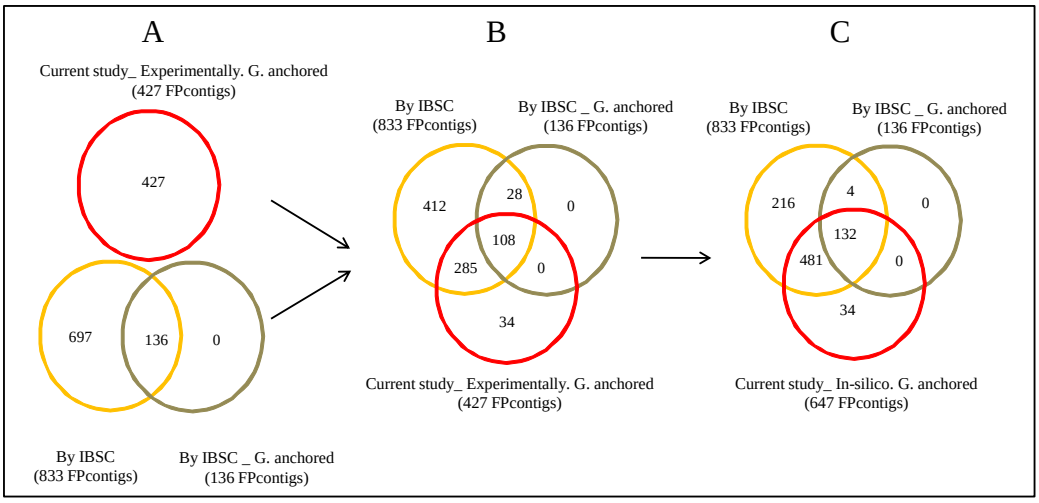


Figure3-7. Genetic anchoring of chromosome *2H-FP Contigs*. A) Shows the anchoring data achieved by IBSC in which out of the total number of anchored contigs, only 136 were genetically anchored. The rest of the contigs was only assigned

either to chromosome long or short arm. B) The experimental anchoring performed in this study increased the number of genetically anchored contigs to 455 [427 (red circle) + 28]. C) The *in-silico* integration using genome zipper and physical map associated sequences resulted in further increased the number of genetically anchored contigs to 651(647+4) in which 647 are arranged along the genome zipper. G. anchored stands for genetically anchored FPcontigs. IBSC refers to International Barley Genome Sequencing Consortium; http://barleygenome.org).

3.2.2. In-silico integration of the chromosome 2H physical map (*2H-FP Contigs)* to the genome zipper

Almost all (98%; 851 FPcontigs) of the above 867 *2H-FP Contigs* were linked to one or the other kind of barley genomic sequence information. For almost 50 % of the BACs included in the FPC assembly, paired end-sequences were generated in an independent project (R Waugh, M Morgante and N Stein, unpublished results). Together with all the gene sequences that were anchored to BAC addresses by the above described efforts this provided anchor points for the integration of further genomic sequence information. Morex genomic DNA was whole genome shotgun (WGS) sequenced to 50-fold coverage and *de-novo* assembled (WGS-contigs) (Stein et al., unpublished data). A total of 29,976 WGS contigs (68.8Mbp) of genomic sequences were linked to 851 *2H-FP Contigs via* BACend and gene sequences (T Nussbaumer et al., unpublished results). Furthermore, assembled sequence information was available for 530 BAC clones from the *2H-FP Contigs* by an independent study (Steuernagel et al., 2009); Stein et al., unpublished results). This sequence information provided the basis for *in-silico* anchoring of the above mentioned 851 FPcontigs by sequence comparison to genes of the 2H genome zipper. The integration of WGS contigs, sequenced BACs, BACend sequences with *2H-FP Contig* information and genetic anchoring data was kindly provided by Thomas Nussbaumer (Helmholtz-Center Munich). On the basis of this integrated sequence information it was the task of this project to query this dataset for the presence of genes and relating this to the genetically anchored physical map and the genome zipper of barley chromosome 2H. The stepwise *in-silico* integration results of the *2H-FP Contigs* to 2H genome zipper consisted of three steps a – c; (a) *In-silico* integration by searching the genome zipper integrated 454 shotgun reads in WGS contigs and sequenced BACs (sBACs) (BLASTN search) to place the corresponding FPcontigs to the genome zipper. A cumulative set of 473

non redundant FPcontigs were assigned to the genome zipper map (471 FPcontigs by WGS contigs and 204 by sBACs) (Table3-9).

Table3-9. *In-silico* integration of the *2H-FP Contigs* to the 2H-genome zipper *via* BLASTN of 454reads against sequence information linked to the *2H-FP Contigs*

BLASTN analysis	Number of 454reads present in 2H-Zipper	Number of 454 reads produced hit in BLASTN [2]	Number of WGS contigs or sBACs involved	Number of identified WGS contigs or BACs which are linked to a 2H-FPcontig	Number of integrated FPcontigs to 2H-genome zipper [3]
454reads *vs.* WGS contig	29,250	8,055	2,420	1,249	471
454reads *vs.* sBACs[1]	29,250	1,791	296	296	204

[1)] Refers to 454 sequenced BACs [2)] BLASTN parameters were identity >= 99% & alignment length >= 300bp. [3)] A cumulative set of **473 non redundant FPcontigs** were integrated to the genome zipper map via this approach.

(b) tBLASTx comparison applying a Bidirectional Best Hit (BBH) strategy of genes from the three model genomes (*Brachypodium, Sorghum* and rice) against WGS contigs was performed. Genes of the model genomes were also considered for *2H-FP Contigs* integration since there were barley gene models for which no 454 read was available (146 genes). Or, even if 454 reads were associated to the gene model, there could be a possibility that they would not provide any connection between model genes and the WGS contigs. This is perhaps because the read and the WGS contig linked to the opposite sides of the respective gene (Figure3-8B). Therefore, performing a sequence homology search against the model genome genes could possibly integrate the WGS contigs – and the corresponding FPcontigs – to the genome zipper (Figure3-8B). Based on that, 997, 1,076, and 1,108 genes were detected from *Brachypodium distachyon*, rice, and *Sorghum*, respectively, corresponding to a cumulative set of 1,459 barley orthologus genes. Of those, 1,059 genes were present in the 2H-genome zipper which led to the assignment of cumulative set of 417 2H-FPcontig to the genome zipper of chromosome 2H (Table3-10).

Table3-10. *In-silico* integration of the *2H-FP Contigs* to the 2H-genome zipper via Bidirectional Best blast Hit (BBH)

Model genomes contributed	Number of genes identified (syntenic and non syntenic)	Number of syntenic genes identified (present in the genome zipper)	Number of FPcontigs integrated to the 2H-genome zipper
Brachypodium	996	723	322
Rice	1,076	654	341
Sorghum	1,108	557	297
Total (Cumulative)	1,469	1,059	417

BLAST parameters were identity ≥ 75% & alignment length ≥ 30 amino acids

(c) The barley unigene set represented in the genome zipper had already been used in physical map anchoring either directly in microarray based analysis (Liu et al. 2011) or in PCR-based approach as genes underlying genetic markers represented in IPK and Okayama maps. Therefore, these unigene/FPcontig relations were also considered for the direct assignment of the related FPcontigs to the genome zipper. A total of 350 FPcontigs could be assigned to the respective genome zipper map (Table3-11).

Table3-11. Direct assignment of the *2H-FP Contigs* into the 2H-genome zipper *via* unigenes/FPcontigs relations

Resource used for Localization	Number of probes present in 2H-genome zipper	Number of probes anchored the FPcontigs	Number of anchored FPcontigs	Number of FPcontigs assigned to the 2H-genome zipper [1]
Unigenes (Array probes)	2,044	978	417	351

[1)] Since not all unigenes were presented in the genome zipper, the rest of the respective FPcontigs (67) could not be assigned.

Considering the three steps described above, a set of 575 *2H-FP Contigs* could be integrated *in-silico* to the 2H-genome zipper (Figure3-8A) providing anchor points for 2,505 chromosome 2H genes.

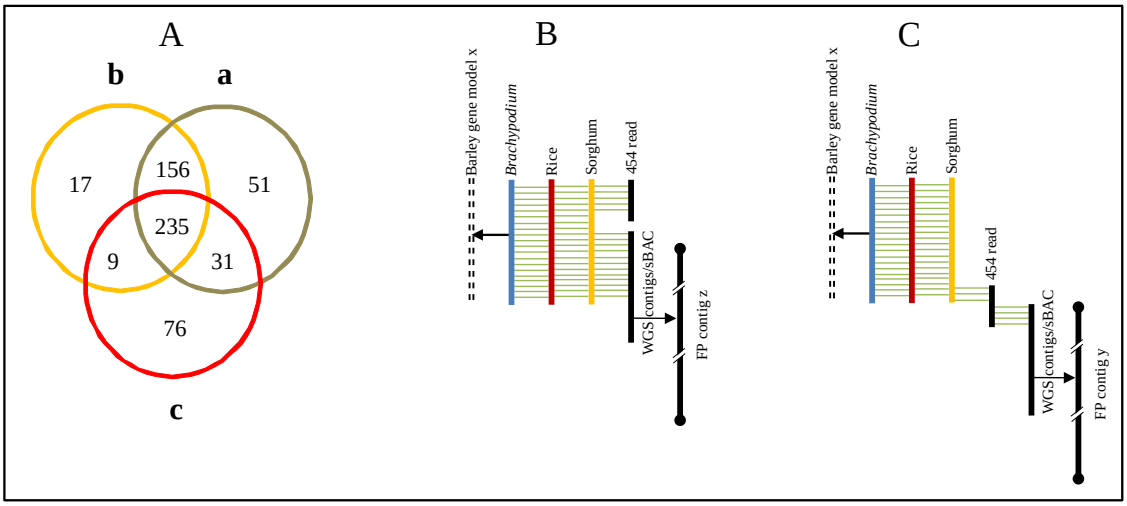


Figure3-8. In-silico integration of the chromosome 2H physical map (2H-FP Contigs) to the genome zipper: A) shows *2H-FP Contigs* integrated *in-silico* to the genome zipper. **a=** BLASTn comparison of 454 reads - associated to genome zipper model genes - against sequence information associated to 2H-FP Contigs . **b=** tBLASTx comparison applying Bidirectional Best Hit (BBH) strategy of model genomes genes against sequence information associated to 2H-FP Contigs . **c=** direct assignment of FP Contigs to the genome zipper using unigenes/FPcontigs relations, unigenes have been assigned to the genome zipper as barley genic sequence support (Mayer et al., 2011). **B)** Assignment of 2H-FP Contigs to barley gene models (genome zipper) via sequence homology (tBLASTx) of the associated sequence information with the model genome gene (s). This assignment was not possible through 454 reads. **C)** Assignment of FPcontigs to barley gene models (genome zipper) via sequence homology (BLASTn) of the associated sequence information with 454 read, this assignment was not possible through model genome gene(s).

Taking into account the 316 FPcontigs anchored *experimentally* by genome zipper based markers alone (see section 3.1.2.3), together with *in-silico* assignment of *2H-FP Contigs* (BBH analysis, BLASTN search of shotgun reads in WGS contigs/sBACs and unigene/FPcontig relations), a cumulative set of 647 non-redundant *2H-FP Contigs* (507 Mbp) could be finally integrated to the genome zipper of chromosome 2H (Table3-12). Thus, *in-silico* anchoring increased the number of *genetically anchored* FPcontigs from 455 to 647 (from 386Mb to 507 Mbp) (Figure3-7C). This corresponds to about 64% of

the chromosome being represented by genetically anchored FPcontigs. The remaining 220 FPcontigs (out of 867) that could not be assigned to 2H-Genome zipper were relatively small contigs (average size = 392 kb) corresponding to a cumulative length of 86 Mbp (Figure3-7C). Moreover, 25% of the chromosome remained unanchored to any chromosome 2H genetic maps utilized.

Table3-12. *2H-FP Contigs in-silico* integrated to the 2H-genome zipper; FPcontigs were assigned to a single or multiple bins.

2H-FP Contigs	Number of FPcontigs	Average size (kb)	Total size (kb)
Assigned to a single bin	603	779	469,608
Assigned to multiple bins	44[1]	850	37,412
Not assigned	220	392	86,207
Total	867	684	593,227

[1)] Eleven of them were allocated into the neighboring bins

3.3. Recombination rate along chromosome 2H

The large proportion of physical contigs that could be genetically anchored to chromosome 2H provided the possibility to analyze the distribution of recombination frequency along this entire barley chromosome. The genetic anchoring of the 2H physical map was based on the linear gene order model of the chromosome 2H genome zipper. For the purpose of studying recombination frequency along this chromosome, the genetic map underlying the genome zipper (Close et al. 2009) was divided into 32 genetic bins, each of a size of 5 cM. The 5 cM interval was selected since the maximum interval between two consecutive genetic markers of this map was 4.92 cM. The resulting 32 bins comprised between nine (Bin_05; in short arm) and 619 genes (Bin_12; containing centromer), respectively. The number of FPcontigs contained in every bin ranged from 1 contig (1.0 Mbp) to 174 contigs (148 Mbp), respectively (Table3-13). In total, 603 *2H-FP Contigs* (out of the 647 *2H-FP Contigs* integrated to 2H-genome zipper) corresponding to 470 Mbp (~60% of the chromosome) could be assigned to single individual bins only. The remaining 44 FPcontigs (37.4 Mbp) were genetically anchored to more than one bin (of those, 11 contigs were assigned to adjacent bins) and, therefore, could not be included in the analysis of recombination frequency (Table3-12). The recombination frequency (cM/Mbp) was

calculated for all bins (Figure3-9A). For the entire chromosome, an average recombination frequency of 1.06 cM / Mbp was determined. Based on this value, three classes of recombination frequency were defined: regions exhibiting (i) low (< 0.5 cM/Mbp), (ii) moderate (0.5–1.06 cM/Mbp) or (iii) high or above average value recombination frequencies (>1.06 cM/Mbp) (Figure3-9A). Areas with low recombination (suppressed in recombination) were mostly confined to bins 10 to 15, (Figure3-9A and Table3-13) thus comprising the centromeric (bin_12) and directly pericentromeric bins. The centromere position of the 2H genome zipper model had been deduced as the transition point from where gene models were associated exclusively with 454 reads from the short to the long chromosome arm (Mayer et al. 2011). In addition to the centromeric and pericentromeric bins two additional regions located at the long arm of the chromosome (bin_19 and bin_32) exhibited a pattern of suppressed recombination.

Table3-13. Dividing the barley chromosome 2H-genome zipper into equally sized genetic bins

Bin number	Start gene model	End gene model	Number of gene model	Number of assigned 2H-FPcontigs	Bin size (Mbp)	Recombination Frequency (cM/Mbp)	Recombination Category[1]
Bin_01	GM.0001	GM.0056	56	8	2.831	1.77	High
Bin_02	GM.0057	GM.0146	90	9	5.887	0.85	Moderate
Bin_03	GM.0147	GM.0245	99	13	9.103	0.55	Moderate
Bin_04	GM.0246	GM.0326	81	5	4.369	1.14	High
Bin_05	GM.0327	GM.0335	9	1	0.953	5.25	High
Bin_06	GM.0336	GM.0389	54	8	6.139	0.81	Moderate
Bin_07	GM.0390	GM.0457	68	8	4.444	1.13	High
Bin_08	GM.0458	GM.0500	43	7	3.215	1.56	High
Bin_09	GM.0501	GM.0547	47	8	3.897	1.28	High
Bin_10	GM.0548	GM.0685	138	27	23.583	0.21	Low
Bin_11	GM.0686	GM.0872	187	34	26.369	0.19	Low
Bin_12	GM.0873	GM.1491	619	174	147.816	0.03	Low
Bin_13	GM.1492	GM.1852	361	76	63.078	0.08	Low
Bin_14	GM.1853	GM.2069	217	27	25.063	0.2	Low
Bin_15	GM.2070	GM.2344	275	46	40.635	0.12	Low
Bin_16	GM.2345	GM.2447	103	13	9.758	0.51	Moderate
Bin_17	GM.2448	GM.2511	64	9	4.831	1.03	Moderate
Bin_18	GM.2512	GM.2590	79	13	6.927	0.72	Moderate
Bin_19	GM.2591	GM.2665	75	9	11.085	0.45	Low
Bin_20	GM.2666	GM.2735	70	8	6.84	0.73	Moderate
Bin_21	GM.2736	GM.2800	65	6	3.995	1.25	High
Bin_22	GM.2801	GM.2868	68	7	3.939	1.27	High
Bin_23	GM.2869	GM.2956	88	8	5.856	0.85	Moderate
Bin_24	GM.2957	GM.3076	120	8	6.127	0.82	Moderate
Bin_25	GM.3077	GM.3116	40	3	1.34	3.73	High
Bin_26	GM.3117	GM.3176	60	10	5.236	0.95	Moderate
Bin_27	GM.3177	GM.3272	96	14	8.576	0.58	Moderate
Bin_28	GM.3273	GM.3372	100	11	7.266	0.69	Moderate
Bin_29	GM.3373	GM.3419	47	4	2.596	1.93	High
Bin_30	GM.3420	GM.3483	64	6	3.488	1.43	High
Bin_31	GM.3484	GM.3521	38	8	4.196	1.19	High
Bin_32	GM.3522	GM.3616	95	15	10.17	0.49	Low
Total	-	-	3616	603	470	Aver= 1.06	-

Bin size was fixed at 5 cM according to the individual largest genetic distance between two markers of the genetic map underlying the genome zipper model.[1)] As compared to the chromosome-wide average of recombination (1.06 cM/Mbp), the bins were classified into low recombination (< 0.5 cM/Mbp), moderate recombination (0.5–1.06 cM/Mbp) and high recombination (>1.06 cM/Mbp) subregions.

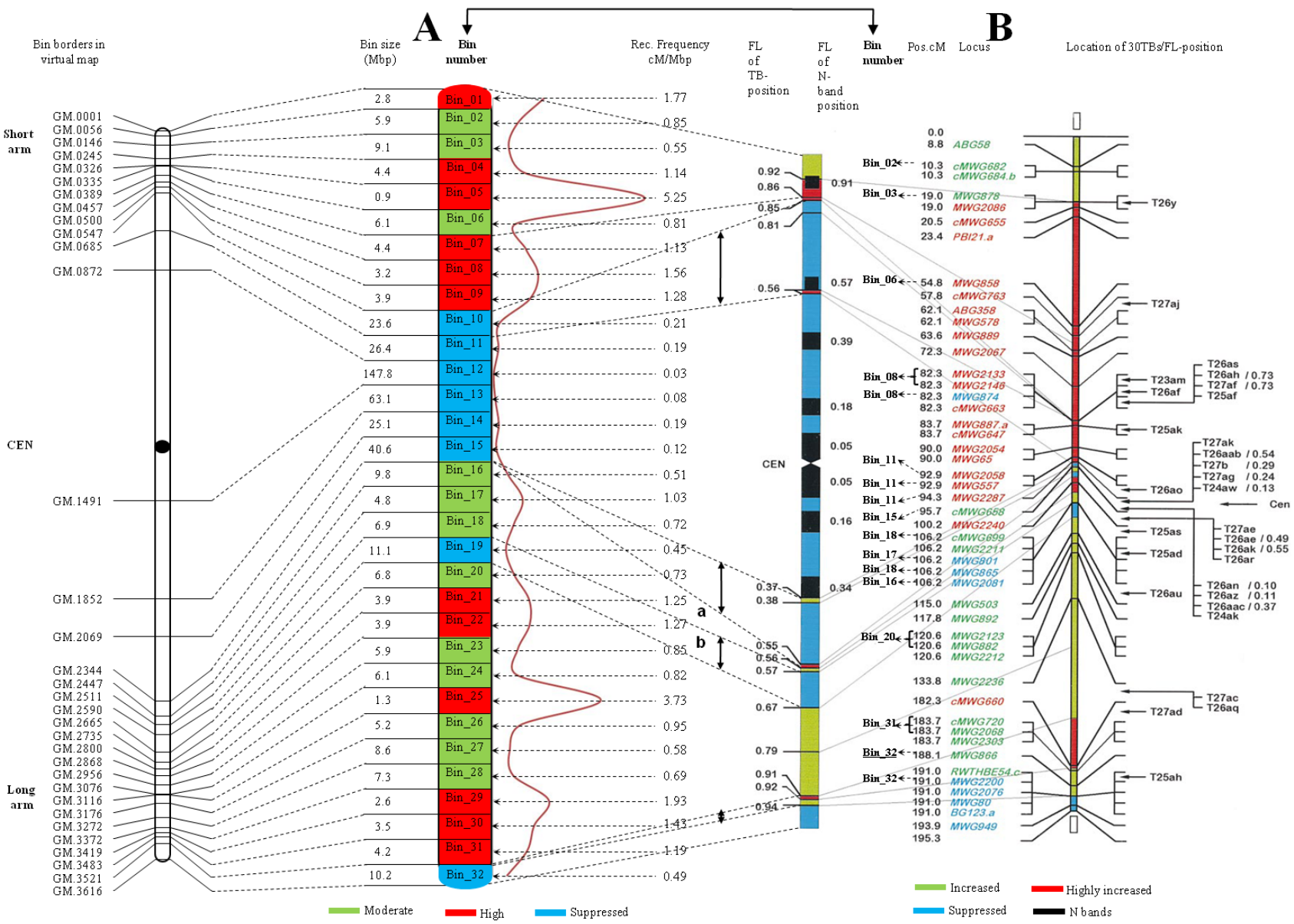

A
B
Bin borders in virtual map
Bin size (Mbp)
Bin number
Rec. Frequency cM/Mbp
FL of TB-position
FL of N-band position
Bin number
Pos.cM
Locus
Location of 30TBs/FL-position
Short arm
Long arm
CEN
Moderate
High
Suppressed
Increased
Highly increased
N bands

Figure3-9: Estimated recombination frequency along barley chromosome 2H. A) In the current study, the chromosome was partitioned into 32 bins each corresponding to 5 cM based on the genetic map underlying 2H-genome zipper. Accumulative physical map length assigned to each bin was used for cM/Mbp ratio calculation. B) Cytogenetically determined recombination rate on chromosome 2H by assigning translocation breakpoints in a RFLP (Restriction Fragment Length Polymorphism) genetic map of barley (Künzel et al., 2000). RFLP markers were compared against sequence information associated with the barley physical map to identify the respective bin location in part A. Bi-directional arrows indicate intervals for which the respective RFLP markers could not be detected in the bin map of the current study. Part B is taken from Kunzel et al., (2000).

In the short arm, three regions including 6 bins (bin_01, bin_04, bin_05, bin_07, bin_08 and bin_09) with recombination frequency higher than average were found. The rate of recombination was observed here to be up to 37 times higher (bin_07) than in the centromeric bin_12. Bin_05, which contained only 9 genes models was connected to a single FPcontig (~ 1Mb). This bin represented the maximum level of recombination with the rate of 5.25 cM/Mbp. Six high recombinogenic bins (21, 22, 25, 29-31) were identified in the long arm. Bin 25 contained the highest recombination frequency which was 45-fold increased compared to the centromeric bin. This bin contained 40 genes in a cumulative physical length of 1.34 Mbp (3 FPcontigs). Overall, only 26 % of the genetically anchored physical map contributed to moderate or high recombinogenic regions of this respective barley chromosome.

The pattern of recombination determined in this study was compared to previously published data that reported recombination frequencies for barley chromosomes on the basis of microdissection of translocation chromosomes and PCR detection of genetic markers (Künzel et al. 2000). This earlier work provided recombination frequency estimates for various regions of different barley chromosomes with a resolution corresponding to a cytogenetic map derived from microscopic specimen (Figure3-9B). Künzel and co-workers employed 31 translocation breakpoints (TB) for barley chromosome 2H. PCR screening of the translocated chromosomal segments with genetic markers led to assignment of the respective segments to the corresponding genetic marker intervals, thus allowed calculation of the recombination gradients.

In order to compare the present datasets and the work published by Künzel et al., (2000), sequence information of 47 published RFLP markers was compared to the physical map associated sequence data. Twenty markers could be anchored to the physical map. The remaining markers either were not detected in the physical map associated sequence data (7 markers) or the respective sequence data were not connected to the physical map (20 markers) (see section 3.2.2). The 20 anchored RFLP markers could be allocated to 14 different bins and therefore, a general comparison of the recombination patterns between the two maps could be achieved (Figure3-9).

RFLP markers cMWG682, MWG878, and MWG858 were situated above Fraction-Length of 0.86 (FL of TB position) in the distal part of the short arm of the cytogenetic map (Künzel et al. 2000) in a region characteristic for high recombination frequency. In the current study, these markers were associated via anchored FPcontigs to bins 02, 03 and 06, respectively, with moderate recombination. Considering the high level of recombination detected in bin_01, both analyses have identified the short arm telomeric regions to be active in recombination. RFLP markers MWG2146, MWG2133 were located in a distal region of short arm corresponding to 0.85 to 0.86 FL with strongly increased level of recombination. These markers were connected to FPcontigs assigned to bin_08 characterized as high recombinogenic areas in the present study.

Chromosome bins 10 to 15 were shown to be suppressed for recombination frequency in our study as they contained the centromeric and pericentromeric zone of the chromosome. According to Künzel and co-workers (2000), a large chromosome segment containing the genetic centromere spanned from FL position 0.56 in short arm to FL of 0.37 in long arm which was characterized by suppressed recombination and absence of markers. In spite of the lack of markers for this region in the cytological physical map of Künzel et al., (2000), a comparison to the integrated map presented here was still feasible by the allocation of flanking markers MWG2287 and cMWG658 to the two almost outermost bins (bin_11 and bin_15) of the centromeric region thus confirming the consistency of assigning the centromeric region as a suppressed recombination zone in both maps.

Markers MWG2058 and MWG557 were allocated to an area with higher recombination frequency just adjacent to a region assigned to the genetic centromere of the cytogenetic map. In the current work, these markers were allocated to the pericentromeric region (Bin_11 and 12) thus reflecting either a contradiction between the two maps or a lack of resolution in the earlier work of Künzel et al., (2000). In case of the long arm, co-segregating markers mapped to 106.2 cM of the RFLP map (cMWG699, MWG801, MWG865 and MWG2081) and were assigned at the border of two adjacent regions on the cytogenetic map. Marker cMWG699 was assigned to a region between 0.56 - 0.57 FL (namely region **a**, Figure3-9) with increased recombination and the other three markers (MWG801, MWG865 and MWG2081) were situated to a segment of 0.57 to 0.67 FL (namely region **b**, Figure3-9) determined as suppressed recombination area of the cytogenetic map. In the present study, all these four markers were assigned to an area spanning from bin_16 to bin_18 with moderate rate of recombination. Considering the localization of MWG801, MWG865 and MWG2081 markers at the border of regions **a** and **b**, the region **b** is likely corresponding to bin_19 with suppressed recombination (Figure3-9).

RFLP markers MWG2123 and MWG882 (located in 0.67-0.79 FL, long arm) and cMWG720 and MWG2068 (from 0.79-0.91 FL, long arm) have been assigned to bin_20 and bin_31, respectively, with same recombination properties.

RFLP markers MWG866 assigned to a small segment (0.92-0.94 FL, long arm) with increased level of recombination in cytogenetic map was assigned to bin_32 with suppressed recombination thus reflecting another inconsistency between the two maps. Nevertheless, assignment of marker MWG2200 to bin_32 confirmed the recombination suppression determination for the very end part of the chromosome 2H long arm in both studies. Overall, except for markers MWG2058, MWG557, MWG866 and MWG874 for which an inconsistency between the two maps was observed, the rest of markers (16) have been localized in areas with similar rate of recombination in this and the previously published study.

3.4. Providing the link between physical and cytogenetic map of long arm of barley chromosome 2H

The results presented above provide the first comprehensive picture of a genetically integrated physical map of barley chromosome 2H. The amount of physical map length anchored to specific genetic bins provides new insights into the correlation of genetic and physical map distances on the barley chromosome. This picture is however still fragmented since not the entire physical length of the chromosome is anchored to the genetic map. Furthermore, several bins contain numerous physical FPcontigs which cannot be arranged in a verified linear order on the basis of the current genetic resolution. Fluorescence in situ Hybridization (FISH) offers a direct way of mapping genomic information to the structure of mitotic or meiotic chromosomes. In order to test the possibility of using cytogenetic mapping as a supportive measure of anchoring (and possibly validating) the genetically anchored physical map, a pilot study was initiated in collaboration with Dr. Lu Ma / Dr. Andreas Houben (CSF group, IPK-Gatersleben) for assessing the usefulness of high-throughput FISH analysis for anchoring BAC contigs to chromosome 2H. The main challenge in this respect was to establish a reliable method of designing suitable FISH probes at high throughput from existing genomic sequence resources integrated to the physical map. The aim then was to map cytogenetically a defined set of contigs and test their genetic and physical order. Because the barley genome contains over 80% of repetitive DNA (Wicker et al., 2009) efficient FISH analysis is often compromised by repetitive fluorescence signals (Ma et al., 2010). Therefore, it was the goal to develop repeat-free sequence probes for the corresponding sequence information associated to each FPcontig. For that purpose, the 14cM genome zipper interval investigated in the pilot study of genome zipper validation (section3.1.2.2.3; Figure3-5) corresponding to an area spanning from bin_28 to bin_31 (Figure3-9A) was considered. Nine different FPcontigs were selected (Figure3-5). This included two FPcontigs that were genetically anchored far from the interval suggested on the basis of the barley chromosome 2H genome zipper interval (FPcontigs 803 and 534; see also section 3.1.2.2.1 and Figure3-2) These two BAC contigs were genetically anchored by two barley gene model derived markers GM.3345 and GM.3446.

The overall genetic order and the respective bin assignment of the selected FPcontigs was as follows: genetic centromere / FPcontig_803(bin_15) / FPcontig_534(bin_19) / FPcontig_46369(bin_28) / FPcontig_38570(bin_28) / FPcontig_45093(bin_29) / FPcontig_2926(bin_30) / FPcontig_494 (bin_31) / FPcontig_46608(bin_31) / FPcontigs_44808(bin_31) / long arm telomere (Figure3-5C).

Two to four primer pairs were designed based on repeat free sequence data for each of the selected FPcontigs. PCR amplicon size varied from 1,333bp to 3,030bp, respectively. PCR products obtained for each FPcontigs were pooled (9 pools corresponding to 9 FPcontigs) thus providing a cumulative FISH probe size of 4,188 bp to 8,742 bp per individual FPcontigs which then were hybridized to mitotic chromosomes of barley cv. Morex (SupData_Table2). A probe for barley 5S rDNA was used to differentiate the seven barley chromosomes, since it is known to produce a chromosome-specific hybridization pattern in the barley genome (Ma et al., 2010).

Eight of the 9 repeat-free probes revealed strong FISH signals specific to chromosome 2HL. FPcontig_46369 was the only contig that produced no clear signal. Based on the genetic mapping analysis, 6 of the 8 FPcontigs (FPcontig_38570 to FPcontig_44808) were closely anchored at the distal part of the long arm (135 cM to 148 cM; Figure3-10) corresponding to bins 28 to 31. This interval was characterized as highly recombinogenic area in the current study (Figure3-9A). The FISH signals (all probes labeled with the same color) obtained for these six closely anchored FPcontigs (genetic distance interval 13 cM) co-localized at the very end of the long arm (Figure3-10) and could not be distinguished at the resolution of mitotic metaphase chromosomes. This supported the observed high recombination rate that could be previously determined for this genetic map interval. FPcontig_803 for which the related genetic marker was mapped to the genetic centromere region produced a clear FISH signal about almost one-sixth (1/6) of the long arm physical length distal to the centromeric constriction of the mitotic chromosome (Figure3-10). This finding underpinned that genes allocated to the genetic centromere of barley may physically be located anywhere between centromere and x% (in this study ~17%) of the chromosome arm. FPcontig_534 co-localized with a 5S rDNA signal of 2HL between the respective signal of the FPcontig_803 and the six co-localizing signals of the remaining FPcontigs (Figure 3-9).

As mentioned above, the two FPcontigs 803 and 534 were expected to be placed at the interval of 137.5 cM to 151.4 cM of chromosome 2H (14 cM validation interval, section 3.1.2.2.3). Genetic mapping shifted them towards the centromere. Cytogenetic mapping allowed testing this observation and basically provided additional evidence for genetic mapping result. Moreover, although the genetic distance between FPcontig_803 and FPcontig_534 was smaller (~26 cM) than the distance between FPcontig_534 and the 6 co-localized FPcontigs (~50cM), the physical distance shown by the respective probe signals was completely opposite (Figure3-10) demonstrating the dynamics in the ratio of physical to genetic distance along barley chromosomes.

In summary, the cytogenetic analysis strongly supported the genetic anchoring results of the physical map especially for FPcontig_803 and FPcontig_534 and secondly showed the efficiency of developing repeat-free FISH probes using an automated approach. For the BAC contigs selected from the distal part of chromosome 2HL, FISH analysis revealed no advantage for physical resolution of FPcontig order. However, large part of the contigs co-localizing genetically in the pericentomeric area would possibly be resolvable in their physical order by FISH.

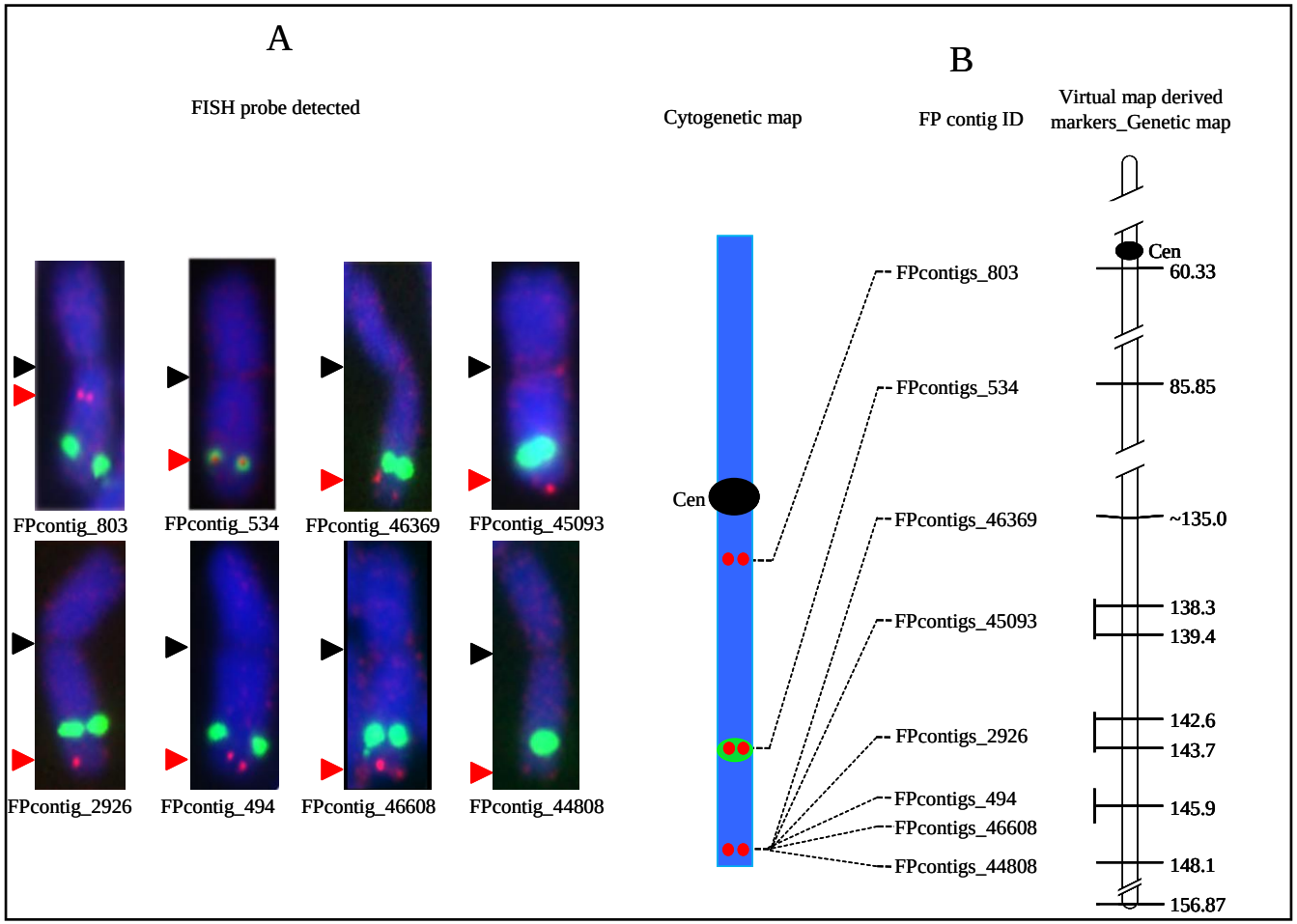


Figure3-10. Cytogenetic mapping of eight genetically anchored barley FPcontigs using FISH on mitotic barley chromosomes: A) shows FISH signals of individual physical contigs/probes (5S rDNA = green signals, FPcontig probes: red signals highlighted by red arrowheads). B) Shows the integrated genetic and cytogenetic map of the selected FPcontigs . Genetic map was derived from a double haploid mapping population consisted of 93 genotypes developed from a cross between barley cv. Morex and Barke (Close et al., 2009). Position of 2H centromere in FISH results is indicated by black arrowheads, 2HS: Chromosome 2H short arm, 2HL: Chromosome 2H long arm.

4. Discussion

The current study presents the first anchored BAC-contig physical map of barley chromosome 2H. 75% of the chromsome could be represented by genetically anchored BAC contigs. Barley chromosome 2H is the biggest among the seven barley chromosomes (Suchankova et al., 2006). This was obtained by different approaches including PCR–based anchoring and *in-silico* strategies. A total number of 1,842 STS (Single Tagged Site) markers were developed and screened against the BAC DNA pools underlying the respective physical map. This approach led to anchoring of 427 physical BAC contigs to the chromosome 2H of barley. This data together with anchoring information provided by IBSC members delivered the first version anchored physical map of this barley chromosome. The map revealed a defined view on distribution of recombination frequency from telomere to telomere. Furthermore, selected anchored BAC contigs could be placed on barley mitotic chromosomes by florescence in situ hybridization (FISH) that allowed insights into the dimension of the genetically anchored physical map of chromosome 2H. However, the 25% unanchored physical map of the chromosome remained to be integrated in the genetic maps by development of the respective suitable genetic markers and anchoring strategies.

Genome-wide physical map of barley that covered up to 14-fold haploid genome coverage had been generated under the framework of the International Barley Sequencing Consortium (IBSC: www.barleygenome.org). Current study was conducted to achieve the full potential of this physical map for the downstream biological studies for which, it needed to be anchored and ordered along the linkage maps of individual chromosomes of barley genome. This accommodates a roadmap towards the gain of several goals such as (1) providing framework for barley genome sequencing, (2) yielding resource for map based gene isolation, (3) serving a basis for analyzing the recombination frequency along the genome, and (4) establishing a framework for comparative genome analysis of barley with other sequenced grass genomes. To complement this larger effort, the current study was focused on anchoring this whole genome barley physical map to chromosome 2H.

4.1. A first version anchored physical map of barley chromosome 2H

In the current study, two independent marker resources including genetic maps developed by Stein et al., (2007) and Sato et al., (2009) were utilized for genetic anchoring of the respective physical map. Among all genetic maps available for barley, these two maps were initially preferred since both were developed exclusively based on genes (ESTs), fit to the anchoring approach (that required PCR-based markers) and were available at the commencement of the current work. Along with them, a virtual gene order map of chromosome 2H (genome zipper) (Mayer et al., 2011) was employed. PCR-based screening of the three maps noted above against BAC DNA pools led to anchoring of 46% (427 FPcontigs; 370 Mbp) of the barley chromosome 2H. In parallel, a dataset of 833 anchored BAC contigs was generated by other IBSC members and made available for comparison to the newly generated data of the current study. Of these 833 FPcontigs, 697 (454 Mbp) had been assigned only to either short or long arm of chromosome 2H without any genetic allocation whereas 136 contigs (130 Mbp) were anchored by genetic mapping information. Compared evaluation of the current study achievement with the anchoring data obtained by IBSC revealed physical map anchoring of 75% (867 FPcontigs called here as *2H-FP Contigs*; 593 Mbp) of barley chromosome 2H. Then it was the task for the present work to *in-silico* integrate/order the *2H-FP Contigs* in a single map – the barley genome zipper of chromosome 2H. This delivered linear ordering of 647 *2H-FP Contigs* (507 Mbp, ~64% of the chromosome) along the chromosome. This was possible by availability of the barley genomic sequences linked to the physical contigs that allowed sequence homology search against the barley genome zipper of chromosome 2H. Since the validation analysis of the genome zipper revealed a high level of accuracy - over 95% - in virtual gene ordering (see section 4.2), the contig ordering on the basis of this virtual map bears a risk of about 5% false anchoring.

Physical maps anchored with different class of markers/genetic maps have been reported for other plants with large and complex genomes like that of wheat chromosome 3B (Paux et al., 2008) and maize (Cone et al., 2002; Wei et al., 2007). BAC based physical map of flow sorted wheat chromosome 3B – the physical map obtained by the same method as barley chromosome 2H - resulted in a final assembly of

1,036 FPcontigs covering 811 Mb (~82%) of the estimated 995 Mb size of chromosome 3B (Paux et al., 2008). By utilizing 1,443 genetic markers originating from different genetic maps, Paux and co-workers could anchor only 61.4% (611Mb) of the chromosome 3B (75% of the 3B physical map). In the absence of genomic resources such as virtual gene order map, Paux and associates employed several other genetic maps/tools to integrate and orient independently the anchored physical map contigs along the chromosome. These included the 3B deletion map, radiation hybrid (RH) map, a composite wheat genetic map, and wheat synteny to rice genome - deletion mapping was performed by utilizing the wheat deletion lines. Each wheat deletion line has lost a specific chromosomal fragment, therefore the lack of any pairs of molecular markers in the respective line reflect the maximal distance between the two markers (Endo and Gill, 1996). Radiation hybrid (RH) mapping is based on radiation-induced chromosome breakage in which retention rate of molecular markers can be analyzed in the respective induced chromosomal segments to develop the RH map (Cox et al., 1990; Hossain et al., 2004) - The largest amount of physical map ordered along the wheat 3B chromosome obtained from two independent map resources including deletion mapping and the wheat genetic map. Deletion mapping resulted in the integration of 599 FPcontigs (556 Mb, ~56% of the chromosome) while only 213 contigs were anchored to the respective genetic map of wheat. In the present study, it was possible to develop a comparable first version anchored and ordered physical map of chromosome 2H (~64% of the chromosome) providing the foundation for genomics informed biological studies on this barley chromosome.

The estimated size of the barley chromosome 2H (790 Mbp) compares to almost two times the size of the entire rice genome (430 Mbp; Chen et al., 2002). Chen et al. could develop the whole rice genome physical map that delivered a total of 458 physical contigs. The barley genome is much bigger and contains a lot more of repetitive DNA (~5Gbp and 80% of repetitive DNA) than the rice genome with 50% of repetitive DNA (Deshpande and Ranjekar, 1980). In this respect to the relation of barley chromosome 2H size with that of rice genome a total number of ~1000 contigs can be expected for this barley chromosome. As of now for 75% of the chromosome 867 FPcontigs have been identified. This can

be already an indication of a comprehensive level of the anchored physical map achieved for this chromosome of barley as compared to the rice genome physical map.

In the current study, 25% of the chromosome remained unanchored to any chromosome 2H genetic map utilized. It can be either because of the lack of marker density in the respective interval of the genetic maps employed or can be the results of small size and highly repetitive nature of the respective contigs. These might have prevented contigs identification - during the screening process - and anchoring by gene based markers utilized during the PCR-based or *in-silico* anchoring. However, these physical contigs most likely can be localized in their original genomic position by performing a reverse anchoring approach in which the sequence information of the respective FPcontigs would be utilized to develop new genetic markers. Subsequent integration of these additional markers to the framework genetic map would anchor the respective physical contigs to the framework map. Different methods to integrate new genetic markers/maps into a standard framework map have been summarized (Li et al., 2010). Of those, construction of so-called "neighbors" map can be considered for barley genome as has been described for maize (Cone et al., 2002) and for wheat (Paux et al., 2008). As the genetic framework map of maize, an intermated map constructed from cross between maize inbred lines B73 and Mo17 was utilized (IBM map). Location of markers from non-IBM maps were extrapolated to their nearest neighbors on the IBM map by considering the shared loci between the two maps that contained the loci of interest (target loci). The distance between shared loci and the target locus on non-IBM map was used to estimate the respective location of the marker in the IBM map. This led to add more than 380 markers to the maize IMB map in the initial experiment (Cone et al., 2002). Development of repeat based markers such as ISBP (insertion site–based polymorphism) markers (Paux et al., 2008) – based on for example BAC end sequence information – and their subsequent integration to the genetic framework map of barley could provide an option to perform the reverse anchoring for barley genome physical map. ISBP can be designed on the basis of the DNA sequence stretches that flank the transposable elements (TE) insertion sites (Paux et al., 2006). ISBP markers developed from BAC end sequences of wheat chromosome 3B

(711 markers) led to anchoring of 472 FPcontigs (452 Mb) in which 296 FPcontigs (221 Mb) were exclusively anchored by this marker type (Paux et al., 2008). The availability of a large set of BAC and BAC end sequence data (Steuernagel et al., 2009; Stein et al., unpublished results) for barley offer the opportunity to apply a comparable strategy in barley genome physical map anchoring. However, in the present work, by application of the large number of gene-based markers most likely the gene containing FPcontigs have been identified and anchored. This has the potential to be considered as starting point to select the gene containing FPcontigs and to initiate the sequencing phase of the chromosome. Moreover, the resulted genetically integrated physical map is of immediate utility to harness the maximum benefits for gene/trait isolation and characterization in barely genome.

4.2. Barley genome zippers is confirmed as a novel resource for synteny based marker development and physical map anchoring

Genome zipper, a virtual high-resolution gene order map, has been developed by combination of chromosome sorting, next generation sequencing and integration with high-resolution synteny data from three grass model genome sequence information (Mayer et al., 2011). This combination of data along with a gene-based genetic map of barley (Close et al., 2009) led to allocation of more than 20,000 genes into the seven chromosomes of barley in a proposed linear order (in case of chromosome 2H on average ~ 22genes/1cM). However, the establishment of the genome zipper is deeply rooted on conserved synteny between related grass genomes. Therefore, in the current study, the map was first validated for the gene order accuracy before being used as a complementary resource of markers for anchoring of chromosome 2H physical map.

It was shown in the present study that the barley virtual gene arrangement formed in the frame of genome zippers for different barley chromosomes was accurate enough - almost 95% accuracy in virtual gene ordering - to be used as a resource for marker development and physical map anchoring. This was assessed by comparing genome zipper gene order against a publicly available transcript derived marker (TDMs) map (Potokina et al., 2008) and by experimental validation of a 14cM genome zipper interval on

the long arm of chromosome 2H. In terms of gene order comparison with TDMs map, a total of 91% consistency (the 9% inconsistency included 5% inconsistency in chromosomal allocation + 4% inconsistency for gene ordering) between barley genome zippers and the TDM map (Potokina et al., 2008) was initially found in this study.

The inconsistencies observed between the two maps could to some extent be the result of inherent inaccuracy attributed with the TDMs map itself. In the process of identifying efficient methods for SFP (single feature polymorphism) prediction in the transcript derived maps (Luo et al., 2007), it was shown that even with applying the most efficient and stringent algorithm only 95% of the gene expression based markers could co-segregate with the genomic SNP markers originated from the same genes. Luo and associates concluded that the remaining 5% of polymorphism in expression data are the result of polymorphism elsewhere in the genome, either trans-acting regulators or duplication of the respective genes (Luo et al., 2007; Potokina et al., 2008). Hence, in the current work, 5% of the total inconsistency between genome zipper and the TDMs map is most likely, though not exclusively, the result of prediction inaccuracy underlying TDMs map itself (Prof. Dr. Mike Kearsey; personal communication). Therefore, the virtual gene order accuracy of the barley genome zippers could amount to almost 96%. Additionally, the accuracy of genome zipper in chromosomal gene allocation was also shown in this study through the course of allocating 128 gene models of the genome zipper to their chromosomal origin by use of wheat/barley ditelosomic addition lines. All 123 genes (out of 128) for which a chromosomal position could be determined by genetic or physical mapping were localized to their proposed origin (chromosome 2HL). Moreover, these results in turn pointed out the efficiency of the flow sorting technique to purify barley chromosome arms from wheat-barley addition lines even bare of wheat genomic segments (Lysak et al., 1999; Suchankova et al., 2006). The high level of gene order uniformity obtained between genome zipper and TDMs map is almost around the level observed in another independent experimental analysis performed in the current work for genome zipper gene order validation on chromosome 2HL. Of the 45 polymorphic genes (between parental genotypes cv. Morex and Barke of a DH population) originated

from a 14 cM interval of the genome zipper, 42 genes (93%) confirmed the proposed order in the barley chromosome 2H genome zipper. The lack of full confirmation perhaps reflects the limitation of synteny between species used for genome zipper construction and barley. However, the gene order/position prediction revealed by the genome zipper is much higher than the general order prediction of 50% to 60 % estimated on the basis of conserved synteny among grasses (Smilde et al., 2001; Gaut, 2002). Gaut reviewed all previous reports on the ability of conserved synteny in gene position prediction and proposed ~50% marker collinearity in a genome-wide scale among grasses. Despite that, synteny based marker development using a single genome such as rice genome sequence information has been continuously explored in local scales with the availability of the large EST sequence collection in barley. For instance, Gottwald et al., (2004), Perovic et al., (2004), and Pourkheirandish et al., (2007) have reported successful utilization - more than 90% marker collinearity for the respective small collinear regions - of grass genome synteny for marker development in the process of map-based cloning of *sdw3*, *Rph16* and *vrs1* genes, respectively. Another example is the application of COS markers (conserved orthologous set among related species) (Fulton et al., 2002; Liewlaksaneeyanawin et al., 2009) as a synteny based marker system. From a set of 31 COS-markers developed from rice genes located on wheat chromosome 7A (Quraishi et al., 2009), 45% were polymorphic between two wheat parental lines. The authors observed that only 86% of the polymorphic markers could be positioned within the interval containing QTL of interest on wheat genome. The remainder (14%) failed to map in the corresponding collinear region. Therefore, range of success rate – in terms of accurate order/position prediction for the respective markers - has been observed using single genome of rice from 56% (Smilde et al., 2001) in macro-collinearity level to more than 90% in micro-collinearity scale (Gottwald et al., 2004; Perovic et al., 2004; Pourkheirandish et al., 2007) to delimit the interval carrying the interested genes in barley. This is because the genome collinearity between grass species is disrupted due to chromosomal rearrangements (inversions, deletions, insertions, translocations) (Bossolini et al., 2007; Faris et al., 2008; Gu et al., 2009). However, despite the aforementioned effects of evolution on synteny, the genome zipper order prediction - in both macro and micro levels - of almost 95% estimated was convincing to consider the

virtual map as a complementary resource for STS marker development and physical map anchoring in barley. It needs to be kept in mind that even for "non-virtual" genetic maps like consensus map, some level of inaccuracy has been reported. There are studies reporting flips of markers between consensus map and the related individual maps which can be the result of significant differences of local recombination frequencies (map length) between populations (Maliepaard et al., 1998; Doligez et al., 2006; Wenzl et al., 2006a; Stein et al., 2007). Therefore, the error rate of gene ordering underlying the virtual gene order map is in the range of other consensus or integrated genetic map available for barley.

4.3. An estimation of recombination frequency along chromosome 2H

In this study, we provided a detailed estimation of recombination rate for barley chromosome 2H by aligning the BAC derived physical map along a genetic consensus map underlying the virtual gene order map of barley. Different rates of recombination along the chromosome were observed from suppression (35 fold below average) in the centromeric region to high rate of recombination (5 fold above average) in the distal region of the short arm. The current finding was compared with the previously reported cytogenetically determined recombination pattern of the chromosome (Kunzel et al., 2000). Kunzel and co-workers utilized 31 translocation breakpoints of barley chromosome 2H and integrated them to a genetic map for recombination rate calculation.

Although, we used different methodology in comparison with the work of Kunzel and associates the general pattern of recombination observed in this study was in excellent agreement with the results presented by these authors. Moreover, the current analysis provides a better resolution by employing a genetic map with larger marker density. The consensus genetic map used in the present analysis was provided by a resolution of 0.37cM, 14 folds more than the resolution of genetic map used in Kunzel et al. analysis (5.2 cM; Graner et al., 1991). However, the minor discrepancies observed between the current study and the previous report of Kunzel and associates could be due to different mapping populations applied to construct the underling genetic maps. Moreover, the ~5% deficiency of the genome zipper in

gene ordering can also have effects on FPcontig mis-ordering and thus causes the contradictions observed.

As expected for plants with large genomes such as maize (Anderson et al., 2003) and wheat (Lukaszewski and Curtis, 1993), it is demonstrated that the recombination has followed a particular pattern with a gradual increase from centromere to the telomeres. For example, in case of maize, Anderson et al., (2003) studied distribution of recombination nodules (RNs) – that are closely related with crossing over - along the maize bivalents to estimate the rate of recombination in the respective chromosomes. Utilizing electron microscopy of synaptonemal complexes (representing pachytene bivalents), the authors showed that all chromosomes shared a common feature of gradual increase of RNs frequency from kinetochores toward distal part of the chromosomal arms, while each chromosome represented its own unique distribution of recombination frequency.

The suppression in crossing over across the centromeric area has been considered as a general characteristic of many species studied until now Drouaud et al., (2006) in *Arabidopsis*, Anderson et al., (2003) in maize, Wu et al., (2003) in rice and Jensen-Seaman et al., (2004) in human. Therefore, we could show that the centromeric region in barley chromosome 2H exhibited the common features of plant chromosomes centromere with being poor in recombination and large in physical size. Suppression of recombination in centromere has been reported to be epigenetically mediated and does not depend on DNA sequence but rather depends on centromere composition (for review see Talbert and Henikoff, 2010). The large size of centromeric zone of plant chromosomes has been reported to be the result of amplification, insertion and duplication of repetitive sequences in the centromeric regions over the time of evolution (Copenhaver et al., 1999; Ma and Jackson, 2006), and most likely is a function of the genome size.

Among the sub regions showing recombination's suppression there was a bin corresponding to the very distal part of the long arm. This decrease of crossing over at the very end of the chromosome 2H was also

observed in the previous work performed in barley chromosome 2H by Kunzel et al., (2000). Moreover, reduction in the rate of recombination at the end of the chromosomes has been also reported in other plants as well (King et al., 2002; Lukaszewski et al., 2004; King et al., 2007; Sourdille et al., 2009). An explanation for that could be the presence of terminal heterochromatin in this region of plant chromosomes, where the formation of crossing over is inhibited (Kumar and Bennetzen, 1999; Gaut et al., 2007). However, this interpretation did not support the suppressed recombination observed in the distal part of barley chromosome 2HL, since the presence of heterochromatin was not confirmed by C-banding pattern of barley chromosome 2H (Kakeda et al., 1991). Moreover, there are genomic sequence features such as CpG motif fraction, GC content and poly(A)/poly(T) fraction (Fullerton et al., 2001; Kong et al., 2002; Muyle et al., 2011) or presence of recombination modifier gene (Ji et al., 1999) that have been shown to correlate with the recombination rate along the genome. Therefore, more investigations and data, e.g. complete genome sequence of barley will be required to study such correlation in various parts of the genome. Access to such information can potentially explain the pattern of recombination rates observed, for example in the distal part of barley chromosome 2HL. In the current study, recombination rates were found to vary within each arm. Peaks of recombination were observed on each of the arms. This in turn shows that the recombination rate does not follow a simple gradient from centromere to the telomeres.

It is very well documented that the recombination and gene density are positively correlated in plant genomes studied so far (for review see Gaut et al., 2007). In wheat, detailed study of recombination gradient and gene distribution for the entire genome of wheat revealed a correlation between these two factors (Erayman et al., 2004). For group two homoeologous chromosomes of wheat - which were of special interest for the current investigation because of synteny with chromosome 2H - the recombinogenic/gene rich areas were shown to be interspersed in the distal portion of the chromosome arms that conform the pattern of recombination frequency distribution observed in this current study.

However, there are studies reported that not all regions with high level of recombination are genes and not all gene rich regions are recombination hot spots (Civardi et al., 1994; Yao et al., 2002).

Several studies have shown the lack of uniform relationship between genetic distance and physical distance in plant and animal genomes (for reviews see Schnable et al., 1998; Nachman, 2002 and Gaut et al., 2007) in that this relationship varies in different parts of the genome, as it was shown in this current study. In spite of this fact, the availability of such knowledge for the genome of interest - in case of the current study chromosome 2H of barley - is a key factor in establishing the profitable plant breeding programs and in map based gene isolation studies. If the region where the gene or QTL of interest is mapped has been now characterized as being suppressed in recombination, the number of F2 population must be large enough to efficiently delimit the interval in the course of fine mapping (Jander et al., 2002). In contrast, for the gene/QTL located in regions highlighted as highly recombinogenic areas lower genetic resolution and subsequently less number of individuals need to be screened in the respective segregating population. Thus having knowledge in recombination pattern along this barley chromosome significantly helps time, effort and cost to be managed more efficiently in the related biological investigations. However, it should also be mentioned that in this current work, the physical map used for the analysis is only representing a portion of chromosome 2H that could be associated to the genetic map underlying the genome zipper (Close et al., 2009; Mayer et al., 2011). Of the 64% of the chromosome length that could be order along the genome zipper 60% were employed for recombination rate calculation. The remaining anchored FPcontigs (4%) have been assigned to more than a single bin, thus were not considered in this analysis. Therefore, fully anchored physical map of the chromosome needs to be established to depict the final pattern and estimates of crossing over rate along the chromosome.

4.4. The perspective of cytogenetic FISH mapping for physical map improvement in the large genome of barley

Integrating the genetic map and BAC contig physical map to the cytogenetic structure of the genome of interest - using florescence in situ hybridization (FISH) - has been employed for direct visualization of

physical contigs. This has enabled not only cross-referencing physical map integration to the linkage map but also has provided further insight into the genome structure and organization. Examples of such genomes investigated include rice (Jiang et al., 2001), human (Furey and Haussler, 2003), sorghum (Kim et al., 2005), maize (Koumbaris and Bass, 2003) and recently *Brachypodium* (Febrer et al., 2010). For instance in case of sorghum (Kim et al., 2005), sets of 18 to 30 BAC clones genetically mapped in regular interval across the linkage map of each of sorghum chromosmes 3-7, 9 and 10 were selected. BAC-FISH cytogenetic mapping was performed on sorghum pachytene bivalents to study chromosomal organization of the genome. The authors were able to confirm BAC genetic anchoring and moreover, could resolve the BAC orders on some linkage map bins - cluster of genetic markers with the same genetic postion - of genetic map. In *Brachypodium,* Febrer et al., (2010) utilized a similar technique on metaphase and pachytene spreads. The authors could assess and validate the contiguity and coverage of physical map and genome sequence contigs of the *Brachypodium* genome. Febrer and co-workers showed that the physical map size is consistent with relative size of each chromosome.

The FISH analysis of the barley genome has been reported based on limited number of probes and in very low-throughput conditions (Lapitan et al., 1997; Stephens et al., 2004; Houben et al., 2007; Phillips et al., 2010; Ma et al., 2010). The major limitations for high-throughput FISH based cytogenetic mapping in barley have been the availability of large genomic sequences, identification of single-copy sequences (Phillips et al., 2010) and low axial resolution of the chromosome type used in cytogenetic studies of barley genome (Valarik et al., 2004). In this current study, the wealth of genomic sequence information integrated to chromosome 2H BAC contigs led to conduct a pilot study to assess the potential of cytogenetic mapping approach using FISH for direct visualization of physical map and to resolve the order of BAC contigs. Resolving of the contigs order is of special importance for regions of the chromosome, e.g. centromeric regions, in which genetic resolution is not sufficient to detect the correct order of the physical contigs.

Nine physical map contigs genetically anchored on chromosome 2H long arm were selected for "FPcontigs-FISH" probe development. Of those probes, ~ 90% revealed strong FISH signals specific to chromosome 2HL on the mitotic metaphase barley chromosomes. The remaining one probe produced no clear signal and was not considered for further optimization. The major bottleneck in detection a specific FISH signal for the respective probe in a large genomes like that of barley after having access to the genomic sequence information is the repetitive nature of the DNA sequences. Because, FISH probes containing repeat sequences would deliver a scattered pattern of hybridization and lead to the failure in the experiment (Islam-Faridi et al., 2002; Stephens et al., 2004). To circumvent this limitation, several techniques have been developed. Several studies in plant with large genomes have applied unlabeled Cot DNA in the respective probe cocktail to block the repetitive DNA of being available in hybridization (Sadder et al., 2000; Sadder and Weber, 2001; Stephens et al., 2004). Recently in barley the technique was used to gather with southern dot blotting using genomic DNA to identify BAC clones with relatively low amounts of repetitive DNA (Phillips et al., 2010). Only 6% of the BACs (7 BACs out of 105) could deliver specific FISH signals on barley mitotic chromosomes and the remainder either landed to multiple loci or did not land at all (Phillips et al., 2010). The application of sequences from small and less repetitive genomes like sorghum as FISH probes in another closely relates species like that of maize with larger genome was considered as a strategy to avoid the cross-hybridization of repetitive DNA (Koumbaris and Bass, 2003). Although, the methods mentioned above can be considered as valuable approaches on a case-by-case basis, the lack of throughput is still the major issue. The most direct and efficient way to develop a specific FISH probe cocktail for organism with large and complex genomes is to assess - if available - its sequence content for uniqueness prior to any wet-lab experiment. Repeat masking algorithms have been applied to mark the repetitive part of the sequences by comparing against repetitive elements library (Wicker et al., 2009). In the current study, two repeat detection approaches including (a) conventional repeat masking using the annotated repetitive element database for Triticeae (Mayer et al., 2011) and (b) mathematically defined repeat (MDR) analysis (Wicker et al., 2008) were employed to detect repetitive sequences in the respective barley sequences. The MDR analysis was

possible by availability of an 8x barley whole genome shotgun sequence database (Stein et al., unpublished data). The selected sequences for probe design were virtually fragmented into overlapping k-mers (21-mers). The resulted fragments were then assessed for their frequency in the barley MDR index generated from the 8x barley whole genome shotgun sequencing dataset. Recently, the similar approach was applied to generate the single-copy sequences for BAC-FISH in barley (Ma et al., 2010). There only 10% (0.1x genome coverage) of the barley genomic sequence was available to generate the MDR index. In the present work, the availability of 8x sequence coverage of the genome for MDR index frequency calculation led to a more accurate repeat masking. Because, first, regions of the genome that were not included in the 0.1x database have been perhaps covered in the 8x sequence information. Therefore, this provides a higher completeness of the repeat representation within the constructed index. Second, k-mer frequencies obtained from the 8x sequence coverage could potentially represent a closer estimation of the natural frequency of the corresponding k-mer in the genome. Overall, application of the k-mer masking can mark the repetitive elements that have not been included in the already existing annotated repeat libraries. These are of major reasons for successful single copy FISH probe development in the current study. The reasons for the only failed probe can be various as described by Kato (2011) including: presence of dust or cell debris, slide scratches, chromosomes swelling, lack of enough fluorescent signal intensity, etc. Considering the results achieved here, it would be possible to combine the process of repeat identification and primer design as a single pipeline to provide a faster method for single-copy probe detection in barley. This will expedite studies related to the chromosomal organization, cytogenetic mapping, and chromosome gene localization in the process of positional gene cloning in barley genome.

Of the eight contigs that provided FISH signals, two contigs (FP contig 803 and FP contig 534) were cytogenetically anchored to centromeric and pericentromeric regions of the chromosome 2H long arm as was shown in genetic anchoring analysis. For the remaining six contigs, although all of them could be landed in the expected regions in which they were genetically anchored, their cytogenetic order could not be resolved.

FPcontig 803 co-segregated with a group of markers that all were mapped to the centromeric region with the same genetic position. Cytogenetic mapping localized this contig far away from centromere by a distance equal to almost one-sixth (1/6) of the chromosome arm. Such observation has been already reported for barley (Stephens et al., 2004). These authors utilized similar approach and attempted to localize 24 cDNA probes on barley metaphase spreads that showed a success rate of only ~ 58% for specific FISH signal detection. Stephens and associates found that a probe genetically mapped close to the centromere was detected to be close to the distal end of the chromosome in the respective cytogenetic map. Therefore, the observations of the current study are in line with the results reported by Stephens et al and suggest that there are still some regions, mainly centromeric areas that have not fully covered by recombination based genetic mapping due to the crossing over suppression. Therefore, the order of genetically anchored physical contigs in such chromosomal regions yet needs to be resolved. Our observation suggests that FISH analysis of the respective contigs on barley metaphase complement can be a suitable tool to resolve contigs order across the corresponding chromosomal segments.

For the remaining six contigs anchored to the distal part of the chromosome 2H long arm, cytogenetic mapping on metaphase chromosomes was not able to reproduce the genetic anchoring order. Therefore, genetic anchoring resolution was much superior over cytogenetic resolution obtained by mitotic metaphase FISH. The cytogenetic marker order clarification for such regions can probably be resolved by application of meiotic pachytene complements and the fiber FISH. These are the other two options for high resolution cytogenetic mapping with spatial resolution of 60-120 kb and 1-5 kp, respectively (Hans de Jong et al., 1999). These techniques can be of special importance to estimate the physical gap size between the respective FPcontigs. The results of the current work showed that the high throughput cytogenetic mapping seems to be feasible in barley by applying a precise repeat masking approach. FISH mapping on mitotic metaphase of barley chromosomes is of great help to resolve the order of physical contigs on pericentromeric and centromeric regions.

4.5. PCR-based pooled BAC library screening employing single marker _ accurate anchoring approach but still laborious and time taking

PCR-based physical map anchoring utilizing single marker assay was the central anchoring approach for the current study that revealed a success rate of 70% for marker/BAC relations identification. The PCR-based approach including the condensed BAC libraries has been applied in several plant genome projects from small genomes like sorghum (Klein et al., 2000), soybean (Wu et al., 2008), grapevine (Lamoureux et al., 2006) to big genomes like that of maize (Yim et al., 2007). The throughput, low cost, and low rate of false positive BAC/marker relation are the advantages of this method over the conventional hybridization based approaches as have been described by (Yim et al., 2007) and (Klein et al., 2000). In this methodology of anchoring, BAC libraries are spotted as high-density colony filters. The radioactively labeled DNA fragments (genetic markers) can then be hybridized to BAC filters to identify the respective BAC clones. Use of hazardous radioactive material and the presence of repeat elements in the labeled probes are the main limitations of the hybridization based screening approaches over the PCR-based screening methods. Although PCR screening is a big improvement if compared to hybridizations approach, it is still time consuming and laborious. In the current study identification of all respective BAC addresses for a given marker required performing 147 PCR reactions [55 superpools and 92 (23×4) matrixpools, since for each marker 4 superpools were considered]. Complete screening for 10 markers required on average the time of three working days. Any approach of multiplexing screening probes would improve screening efficiency. This include application of pooled overgo probes (Madishetty et al., 2007) or recent high throughput multiplex methods such as microarrays (Liu et al., 2011) and illumina golden gate assays (Luo et al., 2009). Agilent microarray technology have been used in barley physical map anchoring (Liu et al., 2011), which significantly reduced the time and efforts needed for BAC library screening. In this technology, BAC DNA pools – the same pooled library was used in the current study - were amplified, labeled by either Cy5 or Cy3 and hybridized to agilent microarrays. The arrays contained 42,302 barley expressed sequence tag (EST) contig sequences in which a 60-mer probe was designed per

selected unigene in 4x44k array format. The data processing scripts were developed and utilized for deconvolution of BAC addresses (Liu et al., 2011). Although the multiplex screening platforms are highly time and cost-efficient, computational programs needed to be developed to analyze and managed the large amount of data produced to facilitate the data handling and rapid BAC address deconvolution. However, it was declared that the recent multiplex screening methodologies can provide significant improvements over traditional single marker-BAC address screening methods (Ariyadasa and Stein, 2012).

5. Outlook

This current work was initiated to integrate the available whole genome BAC based physical map of barley to chromosome 2H in order to deliver the first release anchored physical map of this barley chromosome. To this end, by utilizing different approaches of physical map anchoring and independent resources of gene based molecular markers 75% of the chromosome physical map is now integrated to the respective genetic maps. This information along with a defined view of recombination pattern along the chromosome provides a comprehensive genomic resource for this chromosome so far. The information generated by this study will greatly affect basic and applied research on multiple layers including comparative genomic study, genome structure and evolution, map-based gene isolation and molecular breeding. As with any genome for which a physical map anchoring has been performed, several improvements still need to be accomplished to reach the fully anchored and oriented physical map of the chromosome.

The first task that remains to be performed in future is to target the 25% unanchored physical map of the chromosome to be integrated in the genetic maps. With regards to the repetitive content of >80% of the barley genome (Wicker et al., 2009) and considering that almost all gene bearing physical contigs have been anchored until now, the remaining unanchored ones are small and contain repetitive portion of the chromosome. Therefore, repeat based markers such as ISBP (insertion site–based polymorphism) marker system is of the options to anchor those contigs (Paux et al., 2008). The availability of large set of BAC and BAC end sequence data (Steuernagel et al., 2009; Stein et al., unpublished results) for barley genome

has provided an opportunity to apply such strategy in barley genome physical map anchoring. Otherwise, this data must be generated for FPcontigs that have no such sequence information available.

Second, this study must be continued to better order and orient the FPcontigs along the chromosome especially in regions that lack enough genetic resolution such as centromeric areas. As it was shown in the current study, cytogenetic mapping using FISH can be considered as a proper option. Another possibility is to develop the radiation hybrid (RH) based markers as reported for maize (Riera-Lizarazu et al., 2000) and for wheat genome (Hossain et al., 2004; Paux et al., 2008). Integration of this marker types into the current genetic map framework can improve the physical map anchoring in both recombination-active and recombination-suppressed areas. The availability of wheat-barley addition lines has opened the opportunity for developing such resource for diploid genome of barley. This resource has been produced for chromosome 3H and is being developed for chromosome 2H of barley as well (Dr. S. Kianian, North Dakota University, USA, personal communication).

6. Summary

A BAC (Bacterial Artificial Chromosome) clone - based physical map is the "Reconstructed model" of an individual chromosome by identifying BAC overlaps. This ensures reduction of genome complexity and is essential for sequencing large genomes such as barley. A physical map itself provides a resource that affects basic and applied research at multiple ends. It would greatly facilitate map-based gene isolation, comparative genome analysis, and analysis of genome characteristics, e.g. the distribution of recombination along the chromosome could be analyzed at much higher comprehension. All applications noted above will only be feasible after the integration or anchoring of the physical map to the genetic map, thereby revealing the physical map full potential and efficiency. In the current study it was the aim to develop an anchored physical map of barley chromosome 2H, and herewith complementing the efforts of developing a genome-wide genetically anchored physical map of barley that was being established under the framework work of the International Barley Genome Sequencing Consortium (IBSC).

A combination of two different approaches has been employed, including PCR-based anchoring as the central approach, and an indirect *in-silico* strategy. The former was established by screening DNA pools of BAC clones - underlying the respective physical map - utilizing genetically and accurately, virtually mapped molecular markers. Three independent marker resources, including genetic maps developed by Stein et al., (2007), Sato et al., (2009) along with a virtual gene order map of chromosome 2H (genome zipper) (Mayer et al., 2011) were utilized. The latter marker resource was firstly validated for the precision in virtual gene ordering. The *in-silico* strategy was employed to align anchored physical contigs against a single genetic map. Fluorescence in situ hybridization (FISH) was performed on barley metaphase chromosomes to evaluate its potential for anchoring and ordering of anchored contigs for future barley physical map improvement.

Development/utilization and subsequent PCR screening of 1,842 STS (Sequence Tagged Site) markers against the BAC DNA pools resulted in anchoring of an initial set of 427 physical BAC contigs (FingerPrinted contigs; FPcontigs) to the chromosome 2H of barley. Comparative evaluation of the

current study's achievement with the anchoring data obtained by IBSC revealed that 75% (867 FPcontigs called here as *2H-FP contigs*; 593 Mbp) of barley chromosome 2H had been anchored to the physical map. By performing the *in-silico* integration/ordering of the *2H-FP contigs* in a single genetic map – underlying the barley genome zipper of chromosome 2H - 647 *2H-FP contigs* (507 Mbp, ~64% of the chromosome) were linearly ordered along the chromosome. Using these data a clear view on the pattern of recombination frequency along the chromosome was revealed. Although, this result was in agreement with already published results of recombination frequencies detected in the barley genome (Kunzel et al., 2000), it provided an improved resolution on the distribution of recombination events by employing a genetic map with higher marker density and resolution. A pilot study of FISH mapping of selected physical contigs revealed both a highly improved possibility of single copy FISH probe detection in large genome of barley, and resolving the order of the anchored contigs genetically in areas with suppressed recombination.

Overall, the present study delivered the first release of an anchored physical map of barley chromosome 2H. The resulted genetically integrated physical map can be considered as the starting point to select the gene containing FPcontigs and to initiate the sequencing phase of the chromosome. Moreover, the anchored physical map along with its defined pattern of recombination along the chromosome is of immediate utility to harness the maximum benefits for both gene/trait isolation and molecular plant breeding in barley. However, as with anchoring of physical map of any genome, several improvements still need to be performed to access the fully anchored and oriented physical map of the entire chromosome. Of those tasks that remained to be performed in the future are to target the 25% unanchored physical map of the chromosome to be integrated in the genetic maps by the development of respective suitable genetic markers and anchoring strategies.

7. Zusammenfassung

Eine physikalische Karte basiert auf überlappenden BACs (bakteriellen artifiziellen Chromosomen) und stellt ein „rekonstruiertes Modell“ eines einzelnen Chromosoms dar. Die damit verbundene Verrringerung der Komplexität ist unverzichtbar für die Sequenzierung sehr großer Genome, wie zum Beispiel im Falle der Gerste. Eine physikalische Karte dient in vielerlei Hinsicht als Fundament für Grundlagen- und angewandte Forschung. Sie vereinfacht die Prozedur der kartengestützten Isolation von Genen, die vergleichende Genomanalyse und ermöglicht einen tieferen Einblick in die Verteilung von Rekombinationen entlang eines Chromosoms. Diese Anwendungen sind jedoch nur nach der Integration (Verankerung) von physikalischer und genetischer Karte möglich, welche erst die volle Leistungsfähigkeit einer physikalischen Karte zur Entfaltung kommen lässt. Das Ziel meiner Arbeit war die Entwicklung einer verankerten physikalischen Karte der Gerstenchromosoms 2H. Die vorliegende Studie ergänzt die aktuellen Anstrengungen eine genomweite, genetisch verankerte physikalische Karte der Gerste zu erstellen und wurde im Rahmen des „International Barley Genome Sequencing Consortium“ (IBSC) durchgeführt.

Die PCR-basierte Verankerung wurde experimentell als Hauptansatz eingesetzt und durch eine indirekte *in-silico* Strategie ergänzt. Erstere beruht auf dem Screening von BAC-Pools der physikalischen Karte mit genetisch (experimentell oder *in-silico*) kartierten molekularen Markern. Die Marker stammen aus zwei genetischen Karten (Stein et al., 2007; Sato et al., 2009) und einer virtuellen Anordnung der Gene („genome zipper“) auf dem Chromosom 2H (Mayer et al., 2011). In der vorliegenden Studie wurde die Genauigkeit der virtuellen Genanordung des „genome zipper“ zum ersten Mal validiert. Die *in-silico* Strategie wurde eingesetzt, um verankerte physikalische Contigs an eine einzelne genetische Karte anzupassen. *In situ* Fluoreszenzhybridisierung (FISH) wurde an Metaphasechromosomen der Gerste vorgenommen, um die Leistungsfähigkeit dieser Methode für eine verbesserte Verankerung und Anordnung von verankerten Contigs in der physikalischen Karte der Gerste zu untersuchen.

Für das Screening der BAC-Pools wurden 1842 STS-Marker (Sequence Tagged Site) entwickelt und eingesetzt. Insgesamt wurde ein Satz aus 427 physikalischen BAC Contigs (FingerPrinted contigs; FPcontigs) auf dem Chromosom 2H der Geste verankert. Ein Vergleich der Ergebnisse meiner Arbeit mit den Daten des IBSC zeigte, dass 75% (867 FPcontigs, welche hier „2H-FP contigs“ genannt werden und 593 Mbp umfassen) des Gerstenchromosoms 2H auf der physikalischen Karte verankert wurden. Durch *in-silico* Integration bzw. Anordnung der 2H-FP contigs in einer gemeinsamen genetischen Karte, welche „genome zipper“ des Chromosoms 2H genannt wurde, konnten 647 2H-FP contigs (507 Mbp, ~64% des Chromosoms) linear entlang des Chromosoms angeordnet werden. Basierend auf diesen Ergebnissen wurde ein verbesserter Einblick in die Veränderung der Rekombinationsfrequenz entlang des Chromosoms gewonnen. Diese Daten untermauern vorherige Ergebnisse zur Verteilung von Rekombinationsfrequenzen im Gerstengenom (Künzel et al., 2000) und liefern darüberhinaus durch Einsatz einer genetischen Karte mit höherer Markerdichte eine feinere Auflösung der Verteilung von Rekombinationsereignissen. Die Pilotstudie zur Kartierung ausgewählter physikalischer Contigs mit Hilfe der FISH-Technik lieferte sowohl eine deutlich verbesserte Möglichkeit des Nachweises von Einzel-Kopie FISH-Sonden im komplexen Gerstengenom, als auch eine höhere Auflösung der Reihenfolge von genetisch verankerten Contigs in Bereichen mit unterdrücker Rekombination.

Insgesamt liefert die vorliegende Studie die erste verankerte physikalische Karte des Gerstenchromosoms 2H. Diese genetisch integrierte physikalische Karte liefert die Basis, um gentragende FPcontigs auszuwählen und die Sequenzierung des Chromosoms zu beginnen. Darüberhinaus kann die verankerte physikalische Karte mit ihrem definierten Rekombinationsmuster sofort genutzt werden, um aus Gerste effizienter Gene bzw. Merkmale zu isolieren und die molekulare Gerstenzüchtung zu verbessern. Jede Verankerung einer physikalischen Karte bietet jedoch auch Möglichkeiten zu einer beständigen Verbesserungen, um eine voll verankerte und orientierte, lückenlose physikalische Karte des gesamten Chromosoms zu erlangen. So wird in Zukunft angestrebt, durch Markerentwicklung und geeignete

Verankerungsstrategien, die 25% der noch nicht verankerten physikalischen Karte des Chromosoms in die genetischen Karten zu integrieren.

.

8. References

Abrouk, M., Murat, F., Pont, C., Messing, J., Jackson, S., Faraut, T., Tannier, E., Plomion, C., Cooke, R., Feuillet, C., and Salse, J. (2010). Palaeogenomics of plants: synteny-based modelling of extinct ancestors. Trends in Plant Science 15, 479-487.

Adams, M.D., Celniker, S.E., Holt, R.A., Evans, C.A., Gocayne, J.D., Amanatides, P.G., Scherer, S.E., Li, P.W., Hoskins, R.A., Galle, R.F., George, R.A., Lewis, S.E., Richards, S., Ashburner, M., Henderson, S.N., Sutton, G.G., Wortman, J.R., Yandell, M.D., Zhang, Q., Chen, L.X., Brandon, R.C., Rogers, Y.H., Blazej, R.G., Champe, M., Pfeiffer, B.D., Wan, K.H., Doyle, C., Baxter, E.G., Helt, G., Nelson, C.R., Gabor, G.L., Abril, J.F., Agbayani, A., An, H.J., Andrews-Pfannkoch, C., Baldwin, D., Ballew, R.M., Basu, A., Baxendale, J., Bayraktaroglu, L., Beasley, E.M., Beeson, K.Y., Benos, P.V., Berman, B.P., Bhandari, D., Bolshakov, S., Borkova, D., Botchan, M.R., Bouck, J., Brokstein, P., Brottier, P., Burtis, K.C., Busam, D.A., Butler, H., Cadieu, E., Center, A., Chandra, I., Cherry, J.M., Cawley, S., Dahlke, C., Davenport, L.B., Davies, P., de Pablos, B., Delcher, A., Deng, Z., *Mays*, A.D., Dew, I., Dietz, S.M., Dodson, K., Doup, L.E., Downes, M., Dugan-Rocha, S., Dunkov, B.C., Dunn, P., Durbin, K.J., Evangelista, C.C., Ferraz, C., Ferriera, S., Fleischmann, W., Fosler, C., Gabrielian, A.E., Garg, N.S., Gelbart, W.M., Glasser, K., Glodek, A., Gong, F., Gorrell, J.H., Gu, Z., Guan, P., Harris, M., Harris, N.L., Harvey, D., Heiman, T.J., Hernandez, J.R., Houck, J., Hostin, D., Houston, K.A., Howland, T.J., Wei, M.H., Ibegwam, C., Jalali, M., Kalush, F., Karpen, G.H., Ke, Z., Kennison, J.A., Ketchum, K.A., Kimmel, B.E., Kodira, C.D., Kraft, C., Kravitz, S., Kulp, D., Lai, Z., Lasko, P., Lei, Y., Levitsky, A.A., Li, J., Li, Z., Liang, Y., Lin, X., Liu, X., Mattei, B., McIntosh, T.C., McLeod, M.P., McPherson, D., Merkulov, G., Milshina, N.V., Mobarry, C., Morris, J., Moshrefi, A., Mount, S.M., Moy, M., Murphy, B., Murphy, L., Muzny, D.M., Nelson, D.L., Nelson, D.R., Nelson, K.A., Nixon, K., Nusskern, D.R., Pacleb, J.M., Palazzolo, M., Pittman, G.S., Pan, S., Pollard, J., Puri, V., Reese, M.G., Reinert, K., Remington, K., Saunders, R.D., Scheeler, F., Shen, H., Shue, B.C., Siden-Kiamos, I., Simpson, M., Skupski, M.P., Smith, T., Spier, E., Spradling, A.C., Stapleton, M., Strong, R., Sun, E., Svirskas, R., Tector, C., Turner, R., Venter, E., Wang, A.H., Wang, X., Wang, Z.Y., Wassarman, D.A., Weinstock, G.M., Weissenbach, J., Williams, S.M., WoodageT, Worley, K.C., Wu, D., Yang, S., Yao, Q.A., Ye, J., Yeh, R.F., Zaveri, J.S., Zhan, M., Zhang, G., Zhao, Q., Zheng, L., Zheng, X.H., Zhong, F.N., Zhong, W., Zhou, X., Zhu, S., Zhu, X., Smith, H.O., Gibbs, R.A., Myers, E.W., Rubin, G.M., and Venter, J.C. (2000). The genome sequence of *Drosophila melanogaster*. Science 287, 2185-2195.

Ahloowalia, B.S., Maluszynski, M., and Nichterlein, K. (2004). Global impact of mutation-derived varieties. Euphytica 135, 187-204.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. (1990). Basic Local Alignment Search Tool. Journal of Molecular Biology 215, 403-410.

Anderson, L.K., Doyle, G.G., Brigham, B., Carter, J., Hooker, K.D., Lai, A., Rice, M., and Stack, S.M. (2003). High-resolution crossover maps for each bivalent of *Zea mays* using recombination nodules. Genetics 165, 849-865.

Ariyadasa, R., and Stein, N. (2012). Advances in BAC-based physical mapping and map Integration strategies in plants. Journal of Biomedicine and Biotechnology 2012.

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

Barillot, E., Lacroix, B., and Cohen, D. (1991). Theoretical-analysis of library screening using a N-dimensional pooling strategy. Nucleic Acids Research 19, 6241-6247.

Bennett, C.D., Campbell, M.N., Cook, C.J., Eyre, D.J., Nay, L.M., Nielsen, D.R., Rasmussen, R.P., and Bernard, P.S. (2003). The LightTyper (TM): High-throughput genotyping using fluorescent melting curve analysis. Biotechniques 34, 1288-+.

Berhan, A.M., Hulbert, S.H., Butler, L.G., and Bennetzen, J.L. (1993). Structure and evolution of the genomes of *sorghum-bicolor* and *Zea-Mays*. Theor Appl Genet 86, 598-604.

Bolot, S., Abrouk, M., Masood-Quraishi, U., Stein, N., Messing, J., Feuillet, C., and Salse, J. (2009). The 'inner circle' of the cereal genomes. Current Opinion in Plant Biology 12, 119-125.

Bonierbale, M.W., Plaisted, R.L., and Tanksley, S.D. (1988). RFLP maps based on a common set of clones reveal modes of chromosomal evolution in potato and tomato. Genetics 120, 1095-1103.

Boss, P.K., Bastow, R.M., Mylne, J.S., and Dean, C. (2004). Multiple pathways in the decision to flower: Enabling, promoting, and resetting. Plant Cell 16, S18-S31.

Bossolini, E., Wicker, T., Knobel, P.A., and Keller, B. (2007). Comparison of orthologous loci from small grass genomes *Brachypodium* and rice: implications for wheat genomics and grass genome annotation. Plant J 49, 704-717.

Bouck, J.B., Metzker, M.L., and Gibbs, R.A. (2000). Shotgun sample sequence comparisons between mouse and human genomes. Nat Genet 25, 31-33.

Brueggeman, R., Rostoks, N., Kudrna, D., Kilian, A., Han, F., Chen, J., Druka, A., Steffenson, B., and Kleinhofs, A. (2002). The barley stem rust-resistance gene *Rpg1* is a novel disease-resistance gene with homology to receptor kinases. P Natl Acad Sci USA 99, 9328-9333.

Burke, D.T., Carle, G.F., and Olson, M.V. (1987). Cloning of large segments of exogenous DNA into yeast by means of artificial chromosome vectors. Science 236, 806-812.

Cai, W.W., Reneker, J., Chow, C.W., Vaishnav, M., and Bradley, A. (1998). An anchored framework BAC map of mouse chromosome 11 assembled using multiplex oligonucleotide hybridization. Genomics 54, 387-397.

Cantor, C.R., Mirzabekov, A., and Southern, E. (1992). Report on the Sequencing by Hybridization Workshop. Genomics 13, 1378-1383.

Cardle, L., Ramsay, L., Milbourne, D., Macaulay, M., Marshall, D., and Waugh, R. (2000). Computational and experimental characterization of physically clustered simple sequence repeats in plants. Genetics 156, 847-854.

Chen, M.S., Presting, G., Barbazuk, W.B., Goicoechea, J.L., Blackmon, B., Fang, F.C., Kim, H., Frisch, D., Yu, Y.S., Sun, S.H., Higingbottom, S., Phimphilai, J., Phimphilai, D., Thurmond, S., Gaudette, B., Li, P., Liu, J.D., Hatfield, J., Main, D., Farrar, K., Henderson, C., Barnett, L., Costa, R., Williams, B., Walser, S., Atkins, M., Hall, C., Budiman, M.A., Tomkins, J.P., Luo, M.Z., Bancroft, I., Salse, J., Regad, F., Mohapatra, T., Singh, N.K., Tyagi, A.K., Soderlund, C., Dean, R.A., and Wing, R.A. (2002). An integrated physical and genetic map of the rice genome. Plant Cell 14, 537-545.

Civardi, L., Xia, Y.J., Edwards, K.J., Schnable, P.S., and Nikolau, B.J. (1994). The Relationship between genetic and physical distances in the cloned A1-Sh2 interval of the *Zea-Mays* L genome. P Natl Acad Sci USA 91, 8268-8272.

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.

Coe, E., Cone, K., McMullen, M., Chen, S.S., Davis, G., Gardiner, J., Liscum, E., Polacco, M., Paterson, A., Sanchez-Villeda, H., Soderlund, C., and Wing, R. (2002). Access to the maize genome: An integrated physical and genetic map. Plant Physiol 128, 9-12.

Cone, K.C., McMullen, M.D., Bi, I.V., Davis, G.L., Yim, Y.S., Gardiner, J.M., Polacco, M.L., Sanchez-Villeda, H., Fang, Z.W., Schroeder, S.G., Havermann, S.A., Bowers, J.E., Paterson, A.H., Soderlund, C.A., Engler, F.W., Wing, R.A., and Coe, E.H. (2002). Genetic, physical, and informatics resources for maize on the road to an integrated map. Plant Physiol 130, 1598-1605.

Consortium, I.H.G.S. (2001). Initial sequencing and analysis of the human genome (vol 409, pg 860, 2001). Nature 411, 720-720.

Copenhaver, G.P., Nickel, K., Kuromori, T., Benito, M.I., Kaul, S., Lin, X.Y., Bevan, M., Murphy, G., Harris, B., Parnell, L.D., McCombie, W.R., Martienssen, R.A., Marra, M., and Preuss, D. (1999). Genetic definition and sequence analysis of *Arabidopsis* centromeres. Science 286, 2468-2474.

Costa, J.M., Corey, A., Hayes, P.M., Jobet, C., Kleinhofs, A., Kopisch-Obusch, A., Kramer, S.F., Kudrna, D., Li, M., Riera-Lizarazu, O., Sato, K., Szucs, P., Toojinda, T., Vales, M.I., and Wolfe, R.I. (2001). Molecular mapping of the Oregon Wolfe Barleys: a phenotypically polymorphic doubled-haploid population. Theor Appl Genet 103, 415-424.

Coulson, A., Sulston, J., Brenner, S., and Karn, J. (1986). Toward a physical map of the genome of the nematode *Caenorhabditis-Elegans*. P Natl Acad Sci USA 83, 7821-7825.

Cox, D.R., Burmeister, M., Price, E.R., Kim, S., and Myers, R.M. (1990). Radiation hybrid mapping - a somatic-cell genetic method for constructing high-resolution maps of mammalian chromosomes. Science 250, 245-250.

Deshpande, V.G., and Ranjekar, P.K. (1980). Repetitive DNA in three Gramineae species with low DNA content. Hoppe Seylers Z Physiol Chem 361, 1223-1233.

Devos, K.M., Atkinson, M.D., Chinoy, C.N., Francis, H.A., Harcourt, R.L., Koebner, R.M.D., Liu, C.J., Masojc, P., Xie, D.X., and Gale, M.D. (1993). Chromosomal rearrangements in the rye genome relative to that of wheat. Theor Appl Genet 85, 673-680.

Dilbirligi, M., Erayman, M., and Gill, K.S. (2005). Analysis of recombination and gene distribution in the 2L1.0 region of wheat (*Triticum aestivum L.*) and barley (*Hordeum vulgare* L.). Genomics 86, 47-54.

Dolezel, J., Greilhuber, J., Lucretti, S., Meister, A., Lysak, M.A., Nardi, L., and Obermayer, R. (1998). Plant genome size estimation by flow cytometry: Inter-laboratory comparison. Annals of Botany 82, 17-26.

Doligez, A., Adam-Blondon, A.F., Cipriani, G., Laucou, V., Merdinoglu, D., Meredith, C.P., Riaz, S., Roux, C., This, P., and Di Gaspero, G. (2006). An integrated SSR map of grapevine based on five mapping populations. Theor Appl Genet 113, 369-382.

Drouaud, J., Camilleri, C., Bourguignon, P.Y., Canaguier, A., Berard, A., Vezon, D., Giancola, S., Brunel, D., Colot, V., Prum, B., Quesneville, H., and Mezard, C. (2006). Variation in crossing-over rates across chromosome 4 of *Arabidopsis thaliana* reveals the presence of meiotic recombination "hot spots". Genome Res 16, 106-114.

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.

Endo, T.R., and Gill, B.S. (1996). The deletion stocks of common wheat. J Hered 87, 295-307.

Erayman, M., Sandhu, D., Sidhu, D., Dilbirligi, M., Baenziger, P.S., and Gill, K.S. (2004). Demarcating the gene-rich regions of the wheat genome. Nucleic Acids Research 32, 3546-3565.

Eversole, K., Graner, A., and Stein, N. (2009). Wheat and barley genome sequencing, Genetics and Genomics of the Triticeae, G.J. Muehlbauer and C. Feuillet, eds (Springer US), pp. 713-742.

Faris, J.D., Zhang, Z.C., Fellers, J.P., and Gill, B.S. (2008). Micro-colinearity between rice, *Brachypodium*, and *Triticum monococcum* at the wheat domestication locus Q. Funct Integr Genomic 8, 149-164.

Febrer, M., Goicoechea, J.L., Wright, J., McKenzie, N., Song, X., Lin, J., Collura, K., Wissotski, M., Yu, Y., Ammiraju, J.S., Wolny, E., Idziak, D., Betekhtin, A., Kudrna, D., Hasterok, R., Wing, R.A., and Bevan, M.W. (2010). An integrated physical, genetic and cytogenetic map of *Brachypodium distachyon*, a model system for grass research. Plos One 5, e13461.

Fleischmann, R.D., Adams, M.D., White, O., Clayton, R.A., Kirkness, E.F., Kerlavage, A.R., Bult, C.J., Tomb, J.F., Dougherty, B.A., Merrick, J.M., Mckenney, K., Sutton, G., Fitzhugh, W., Fields, C., Gocayne, J.D., Scott, J., Shirley, R., Liu, L.I., Glodek, A., Kelley, J.M., Weidman, J.F., Phillips, C.A., Spriggs, T., Hedblom, E., Cotton, M.D., Utterback, T.R., Hanna, M.C., Nguyen, D.T., Saudek, D.M., Brandon, R.C., Fine, L.D., Fritchman, J.L., Fuhrmann, J.L., Geoghagen, N.S.M., Gnehm, C.L., Mcdonald, L.A., Small, K.V., Fraser, C.M., Smith, H.O., and Venter, J.C. (1995).

Whole-genome random sequencing and assembly of *Haemophilus-Influenzae* Rd. Science 269, 496-512.

Fukui, K., Kamisugi, Y., and Sakai, F. (1994). Physical Mapping of 5s *Rdna* loci by direct-cloned biotinylated probes in barley chromosomes. Genome 37, 105-111.

Fullerton, S.M., Carvalho, A.B., and Clark, A.G. (2001). Local rates of recombination are positively correlated with GC content in the human genome. Mol Biol Evol 18, 1139-1142.

Fulton, T.M., Van der Hoeven, R., Eannetta, N.T., and Tanksley, S.D. (2002). Identification, analysis, and utilization of conserved ortholog set markers for comparative genomics in higher plants. Plant Cell 14, 1457-1467.

Furey, T.S., and Haussler, D. (2003). Integration of the cytogenetic map with the draft human genome sequence. Human Molecular Genetics 12, 1037-1044.

Gardiner, J., Schroeder, S., Polacco, M.L., Sanchez-Villeda, H., Fang, Z.W., Morgante, M., Landewe, T., Fengler, K., Useche, F., Hanafey, M., Tingey, S., Chou, H., Wing, R., Soderlund, C., and Coe, E.H. (2004). Anchoring 9,371 maize expressed sequence tagged unigenes to the bacterial artificial chromosome contig map by two-dimensional overgo hybridization. Plant Physiol 134, 1317-1326.

Gaut, B.S. (2002). Evolutionary dynamics of grass genomes. New Phytologist 154, 15-28.

Gaut, B.S., Wright, S.I., Rizzon, C., Dvorak, J., and Anderson, L.K. (2007). Opinion - Recombination: an underappreciated factor in the evolution of plant genomes. Nature Reviews Genetics 8, 77-84.

Gottwald, S., Stein, N., Borner, A., Sasaki, T., and Graner, A. (2004). The gibberellic-acid insensitive dwarfing gene *sdw3* of barley is located on chromosome 2HS in a region that shows high colinearity with rice chromosome 7L. Molecular Genetics and Genomics 271, 426-436.

Graner, A., Kilian, A., and Kleinhofs, A. (2010). Barley genome organization, mapping, and synteny. In Barley; Production, Improvement, and Uses (Wiley-Blackwell), pp. 63-84.

Graner, A., Jahoor, A., Schondelmaier, J., Siedler, H., Pillen, K., Fischbeck, G., Wenzel, G., and Herrmann, R.G. (1991). Construction of an RFLP map of barley. Theor Appl Genet 83, 250-256.

Graner, A., Streng, S., Kellermann, A., Schiemann, A., Bauer, E., Waugh, R., Pellio, B., and Ordon, F. (1999). Molecular mapping and genetic fine-structure of the *rym5* locus encoding resistance to different strains of the Barley Yellow Mosaic Virus Complex. Theor Appl Genet 98, 285-290.

Green, E.D. (2001). Strategies for the systematic sequencing of complex genomes. Nature Reviews Genetics 2, 573-583.

Gregory, S.G., Sekhon, M., Schein, J., Zhao, S., Osoegawa, K., Scott, C.E., Evans, R.S., Burridge, P.W., Cox, T.V., Fox, C.A., Hutton, R.D., Mullenger, I.R., Phillips, K.J., Smith, J., Stalker, J., Threadgold, G.J., Birney, E., Wylie, K., Chinwalla, A., Wallis, J., Hillier, L., Carter, J., Gaige, T., Jaeger, S., Kremitzki, C., Layman, D., Maas, J., McGrane, R., Mead, K., Walker, R., Jones, S., Smith, M., Asano, J., Bosdet, I., Chan, S., Chittaranjan, S., Chiu, R., Fjell, C., Fuhrmann, D., Girn, N., Gray, C., Guin, R., Hsiao, L., Krzywinski, M., Kutsche, R., Lee, S.S., Mathewson, C., McLeavy, C., Messervier, S., Ness, S., Pandoh, P., Prabhu, A.L., Saeedi, P., Smailus, D., Spence, L., Stott, J., Taylor, S., Terpstra, W., Tsai, M., Vardy, J., Wye, N., Yang, G., Shatsman, S., Ayodeji, B., Geer, K., Tsegaye, G., Shvartsbeyn, A., Gebregeorgis, E., Krol, M., Russell, D., Overton, L., Malek, J.A., Holmes, M., Heaney, M., Shetty, J., Feldblyum, T., Nierman, W.C., Catanese, J.J., Hubbard, T., Waterston, R.H., Rogers, J., de Jong, P.J., Fraser, C.M., Marra, M., McPherson, J.D., and Bentley, D.R. (2002). A physical map of the mouse genome. Nature 418, 743-750.

Gu, Y.Q., Ma, Y.Q., Huo, N.X., Vogel, J.P., You, F.M., Lazo, G.R., Nelson, W.M., Soderlund, C., Dvorak, J., Anderson, O.D., and Luo, M.C. (2009). A BAC-based physical map of *Brachypodium distachyon* and its comparative analysis with rice and wheat. Bmc Genomics 10, -.

Hans de Jong, J., Fransz, P., and Zabel, P. (1999). High resolution FISH in plants - techniques and applications. Trends in Plant Science 4, 258-263.

Hearnden, P.R., Eckermann, P.J., McMichael, G.L., Hayden, M.J., Eglinton, J.K., and Chalmers, K.J. (2007). A genetic map of 1,000 SSR and DArT markers in a wide barley cross. Theor Appl Genet 115, 383-391.

Herrmann, M.G., Durtschi, J.D., Wittwer, C.T., and Voelkerding, K.V. (2007). Expanded instrument comparison of amplicon DNA melting analysis for mutation scanning and genotyping. Clin Chem 53, 1544-1548.

Hong, G. (1997). A rapid and accurate strategy for rice contig map construction by combination of fingerprinting and hybridization. Plant Molecular Biology 35, 129-133.

Hossain, K.G., Riera-Lizarazu, O., Kalavacharla, V., Vales, M.I., Maan, S.S., and Kianian, S.F. (2004). Radiation hybrid mapping of the species cytoplasm-specific (scsae) gene in wheat. Genetics 168, 415-423.

Houben, A., Schroeder-Reiter, E., Nagaki, K., Nasuda, S., Wanner, G., Murata, M., and Endo, T.R. (2007). CENH3 interacts with the centromeric retrotransposon cereba and GC-rich satellites and locates to centromeric substructures in barley. Chromosoma 116, 275-283.

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

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

Islam-Faridi, M.N., Childs, K.L., Klein, P.E., Hodnett, G., Menz, M.A., Klein, R.R., Rooney, W.L., Mullet, J.E., Stelly, D.M., and Price, H.J. (2002). A molecular cytogenetic map of sorghum chromosome 1. Fluorescence in situ hybridization analysis with mapped bacterial artificial chromosomes. Genetics 161, 345-353.

Islam, A.K.M.R., and Shepherd, K.W. (2000). Isolation of a fertile wheat-barley addition line carrying the entire barley chromosome 1H. Euphytica 111, 145-149.

Islam, A.K.M.R., Shepherd, K.W., and Sparrow, D.H.B. (1981). Isolation and characterization of euplasmic wheat-barley chromosome addition lines. Heredity 46, 161-&.

Jaccoud, D., Peng, K.M., Feinstein, D., and Kilian, A. (2001). Diversity Arrays: a solid state technology for sequence information independent genotyping. Nucleic Acids Research 29.

Jafary, H., Szabo, L.J., and Niks, R.E. (2006). Innate nonhost immunity in barley to different heterologous rust fungi is controlled by sets of resistance genes with different and overlapping specificities (vol 19, pg 1270, 2006). Molecular Plant-Microbe Interactions 19, 1472-1472.

Jafary, H., Albertazzi, G., Marcel, T.C., and Niks, R.E. (2008). High diversity of genes for nonhost resistance of barley to heterologous rust fungi. Genetics 178, 2327-2339.

Jander, G., Norris, S.R., Rounsley, S.D., Bush, D.F., Levin, I.M., and Last, R.L. (2002). *Arabidopsis* map-based cloning in the post-genome era. Plant Physiol 129, 440-450.

Jensen-Seaman, M.I., Furey, T.S., Payseur, B.A., Lu, Y.T., Roskin, K.M., Chen, C.F., Thomas, M.A., Haussler, D., and Jacob, H.J. (2004). Comparative recombination rates in the rat, mouse, and human genomes. Genome Res 14, 528-538.

Ji, Y.F., Stelly, D.M., De Donato, M., Goodman, M.M., and Williams, C.G. (1999). A candidate recombination modifier gene for *Zea mays* L. Genetics 151, 821-830.

Jiang, J.M., Cheng, Z.K., Buell, C.R., Wing, R.A., and Gu, M.H. (2001). Toward a cytological characterization of the rice genome. Genome Res 11, 2133-2141.

Kakeda, K., Fukui, K., and Yamagata, H. (1991). Heterochromatic differentiation in barley chromosomes revealed by C-banding and N-banding techniques. Theor Appl Genet 81, 144-150.

Kato, A. (2011). High-density fluorescence in situ hybridization signal detection on barley (*Hordeum vulgare* L.) chromosomes with improved probe screening and reprobing procedures. Genome 54, 151-159.

Kato, A., Albert, P.S., Vega, J.M., and Birchler, J.A. (2006). Sensitive fluorescence in situ hybridization signal detection in maize using directly labeled probes produced by high concentration DNA polymerase nick translation. Biotechnic & Histochemistry 81, 71-78.

Kidd, J.M., Cooper, G.M., Donahue, W.F., Hayden, H.S., Sampas, N., Graves, T., Hansen, N., Teague, B., Alkan, C., Antonacci, F., Haugen, E., Zerr, T., Yamada, N.A., Tsang, P., Newman, T.L., Tuzun, E., Cheng, Z., Ebling, H.M., Tusneem, N., David, R., Gillett, W., Phelps, K.A., Weaver, M., Saranga, D., Brand, A., Tao, W., Gustafson, E., McKernan, K., Chen, L., Malig, M., Smith, J.D., Korn, J.M., McCarroll, S.A., Altshuler, D.A., Peiffer, D.A., Dorschner, M., Stamatoyannopoulos, J., Schwartz, D., Nickerson, D.A., Mullikin, J.C., Wilson, R.K., Bruhn, L., Olson, M.V., Kaul, R., Smith, D.R., and Eichler, E.E. (2008). Mapping and sequencing of structural variation from eight human genomes. Nature 453, 56-64.

Kim, H., Hurwitz, B., Yu, Y., Collura, K., Gill, N., SanMiguel, P., Mullikin, J.C., Maher, C., Nelson, W., Wissotski, M., Braidotti, M., Kudrna, D., Goicoechea, J.L., Stein, L., Ware, D., Jackson, S.A., Soderlund, C., and Wing, R.A. (2008). Construction, alignment and analysis of twelve framework physical maps that represent the ten genome types of the genus *Oryza*. Genome Biology 9.

Kim, J.S., Islam-Faridi, M.N., Klein, P.E., Stelly, D.M., Price, H.J., Klein, R.R., and Mullet, J.E. (2005). Comprehensive molecular cytogenetic analysis of sorghum genome architecture: Distribution of euchromatin, heterochromatin, genes and recombination in comparison to rice. Genetics 171, 1963-1976.

King, J., Armstead, I.P., Donnison, S.I., Roberts, L.A., Harper, J.A., Skot, K., Elborough, K., and King, I.P. (2007). Comparative analyses between *Lolium/Festuca* introgression lines and rice reveal the major fraction of functionally annotated gene models is located in recombination-poor/very recombination-poor regions of the genome. Genetics 177, 597-606.

King, J., Armstead, I.P., Donnison, I.S., Thomas, H.M., Jones, R.N., Kearsey, M.J., Roberts, L.A., Thomas, A., Morgan, W.G., and King, I.P. (2002). Physical and genetic mapping in the grasses *Lolium perenne* and *Festuca pratensis*. Genetics 161, 315-324.

Klein, P.E., Klein, R.R., Cartinhour, S.W., Ulanch, P.E., Dong, J., Obert, J.A., Morishige, D.T., Schlueter, S.D., Childs, K.L., Ale, M., and Mullet, J.E. (2000). A high-throughput AFLP-based method for constructing integrated genetic and physical maps: progress toward a sorghum genome map. Genome Res 10, 789-807.

Komori, T., and Nitta, N. (2005). Utilization of the CAPS/dCAPS method to convert rice SNPs into PCR-based markers. Breeding Sci 55, 93-98.

Kong, A., Gudbjartsson, D.F., Sainz, J., Jonsdottir, G.M., Gudjonsson, S.A., Richardsson, B., Sigurdardottir, S., Barnard, J., Hallbeck, B., Masson, G., Shlien, A., Palsson, S.T., Frigge, M.L., Thorgeirsson, T.E., Gulcher, J.R., and Stefansson, K. (2002). A high-resolution recombination map of the human genome. Nat Genet 31, 241-247.

Kosambi, D. (1944). The estimation of map distances from recombination values. Ann Eugen 12, 172–175.

Koumbaris, G.L., and Bass, H.W. (2003). A new single-locus cytogenetic mapping system for maize (*Zea Mays* L.): overcoming FISH detection limits with marker-selected sorghum (*S-propinquum* L.) BAC clones. Plant J 35, 647-659.

Krattinger, S., Wicker, T., and Keller, B. (2009). Map-based cloning of genes in triticeae (wheat and barley), Genetics and Genomics of the Triticeae, G.J. Muehlbauer and C. Feuillet, eds (Springer US), pp. 337-357.

Genetics and Genomics of the Triticeae, G.J. Muehlbauer and C. Feuillet, eds (Springer US), pp. 337-357.

Kumar, A., and Bennetzen, J.L. (1999). Plant retrotransposons. Annual Review of Genetics 33, 479-532.

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.

Lamoureux, D., Bernole, A., Le Clainche, I., Tual, S., Thareau, V., Paillard, S., Legeai, F., Dossat, C., Wincker, P., Oswald, M., Merdinoglu, D., Vignault, C., Delrot, S., Caboche, M., Chalhoub, B., and Adam-Blondon, A.F. (2006). Anchoring of a large set of markers onto a BAC library for the development of a draft physical map of the grapevine genome. Theor Appl Genet 113, 344-356.

Lapitan, N.L.V., Brown, S.E., Kennard, W., Stephens, J.L., and Knudson, D.L. (1997). FISH physical mapping with barley BAC clones. Plant J 11, 149-156.

Leister, D., Kurth, J., Laurie, D.A., Yano, M., Sasaki, T., Devos, K., Graner, A., and Schulze-Lefert, P. (1998). Rapid reorganization of resistance gene homologues in cereal genomes. P Natl Acad Sci USA 95, 370-375.

Lewin, H.A., Larkin, D.M., Pontius, J., and O'Brien, S.J. (2009). Every genome sequence needs a good map. Genome Res 19, 1925-1928.

Li, H.B., Kilian, A., Zhou, M.X., Wenzl, P., Huttner, E., Mendham, N., McIntyre, L., and Vaillancourt, R.E. (2010). Construction of a high-density composite map and comparative mapping of segregation distortion regions in barley. Molecular Genetics and Genomics 284, 319-331.

Li, J., Huang, X.Q., Heinrichs, F., Ganal, M.W., and Roder, M.S. (2005). Analysis of QTLs for yield, yield components, and malting quality in a BC3-DH population of spring barley. Theor Appl Genet 110, 356-363.

Liewlaksaneeyanawin, C., Zhuang, J., Tang, M., Farzaneh, N., Lueng, G.L., Cullis, C., Findlay, S., Ritland, C.E., Bohlmann, J., and Ritland, K. (2009). Identification of COS markers in the *Pinaceae*. Tree Genetics & Genomes 5, 247-255.

Liu, H., McNicol, J., Bayer, M., Morris, J.A., Cardle, L., Marshall, D.F., Schulte, D., Stein, N., Shi, B.J., Taudien, S., Waugh, R., and Hedley, P.E. (2011). Highly parallel gene-to-BAC addressing using microarrays. Biotechniques 50, 165.

Lukaszewski, A.J., and Curtis, C.A. (1993). Physical distribution of recombination in B-genome chromosomes of tetraploid wheat. Theor Appl Genet 86, 121-127.

Lukaszewski, A.J., Rybka, K., Korzun, V., Malyshev, S.V., Lapinski, B., and Whitkus, R. (2004). Genetic and physical mapping of homoeologous recombination points involving wheat chromosome 2B and rye chromosome 2R. Genome 47, 36-45.

Luo, M., and Wing, R.A. (2003). An improved method for plant BAC library construction. Methods Mol Biol 236, 3-20.

Luo, M.C., Xu, K., Ma, Y., Deal, K.R., Nicolet, C.M., and Dvorak, J. (2009). A high-throughput strategy for screening of bacterial artificial chromosome libraries and anchoring of clones on a genetic map constructed with single nucleotide polymorphisms. BMC Genomics 10, 28.

Luo, M.C., Thomas, C., You, F.M., Hsiao, J., Shu, O.Y., 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.

Luo, Z.W., Potokina, E., Druka, A., Wise, R., Waugh, R., and Kearsey, M.J. (2007). SFP genotyping from affymetrix arrays is robust but largely detects cis-acting expression regulators. Genetics 176, 789-800.

Lysak, M.A., Cihalikova, J., Kubalakova, M., Simkova, H., Kunzel, G., and Dolezel, J. (1999). Flow karyotyping and sorting of mitotic chromosomes of barley (*Hordeum vulgare* L.). Chromosome Research 7, 431-444.

Ma, J.X., and Jackson, S.A. (2006). Retrotransposon accumulation and satellite amplification mediated by segmental duplication facilitate centromere expansion in rice. Genome Res 16, 251-259.

Ma, L., Vu, G.T.H., Schubert, V., Watanabe, K., Stein, N., Houben, A., and Schubert, I. (2010). Synteny between *Brachypodium distachyon* and *Hordeum vulgare* as revealed by FISH. Chromosome Research 18, 841-850.

Madishetty, K., Condamine, P., Svensson, J.T., Rodriguez, E., and Close, T.J. (2007). An improved method to identify BAC clones using pooled overgos. Nucleic Acids Research 35, -.

Maliepaard, C., Alston, F.H., van Arkel, G., Brown, L.M., Chevreau, E., Dunemann, F., Evans, K.M., Gardiner, S., Guilford, P., van Heusden, A.W., Janse, J., Laurens, F., Lynn, J.R., Manganaris, A.G., den Nijs, A.P.M., Periam, N., Rikkerink, E., Roche, P., Ryder, C., Sansavini, S., Schmidt, H., Tartarini, S., Verhaegh, J.J., Vrielink-van Ginkel, M., and King, G.J. (1998). Aligning male

and female linkage maps of apple (*Malus pumila* Mill.) using multi-allelic markers. Theor Appl Genet 97, 60-73.

Manly, K.F., Cudmore, R.H., and Meer, J.M. (2001). Map manager QTX, cross-platform software for genetic mapping. Mamm Genome 12, 930-932.

Marcel, T.C., Varshney, R.K., Barbieri, M., Jafary, H., de Kock, M.J.D., Graner, A., and Niks, R.E. (2007). A high-density consensus map of barley to compare the distribution of QTLs for partial resistance to *puccinia hordei* and of defence gene homologues. Theor Appl Genet 114, 487-500.

Marra, M.A., Kucaba, T.A., Dietrich, N.L., Green, E.D., Brownstein, B., Wilson, R.K., McDonald, K.M., Hillier, L.W., McPherson, J.D., and Waterston, R.H. (1997). High throughput fingerprint analysis of large-insert clones. Genome Res 7, 1072-1084.

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

McPherson, J.D., Marra, M., Hillier, L., Waterston, R.H., Chinwalla, A., Wallis, J., Sekhon, M., Wylie, K., Mardis, E.R., Wilson, R.K., Fulton, R., Kucaba, T.A., Wagner-McPherson, C., Barbazuk, W.B., Gregory, S.G., Humphray, S.J., French, L., Evans, R.S., Bethel, G., Whittaker, A., Holden, J.L., McCann, O.T., Dunham, A., Soderlund, C., Scott, C.E., Bentley, D.R., Schuler, G., Chen, H.C., Jang, W.H., Green, E.D., Idol, J.R., Maduro, V.V.B., Montgomery, K.T., Lee, E., Miller, A., Emerling, S., Kucherlapati, R., Gibbs, R., Scherer, S., Gorrell, J.H., Sodergren, E., Clerc-Blankenburg, K., Tabor, P., Naylor, S., Garcia, D., de Jong, P.J., Catanese, J.J., Nowak, N., Osoegawa, K., Qin, S.Z., Rowen, L., Madan, A., Dors, M., Hood, L., Trask, B., Friedman, C., Massa, H., Cheung, V.G., Kirsch, I.R., Reid, T., Yonescu, R., Weissenbach, J., Bruls, T., Heilig, R., Branscomb, E., Olsen, A., Doggett, N., Cheng, J.F., Hawkins, T., Myers, R.M., Shang, J., Ramirez, L., Schmutz, J., Velasquez, O., Dixon, K., Stone, N.E., Cox, D.R., Haussler, D., Kent, W.J., Furey, T., Rogic, S., Kennedy, S., Jones, S., Rosenthal, A., Wen, G.P., Schilhabel, M., Gloeckner, G., Nyakatura, G., Siebert, R., Schlegelberger, B., Korenburg, J., Chen, X.N., Fujiyama, A., Hattori, M., Toyoda, A., Yada, T., Park, H.S., Sakaki, Y., Shimizu, N., Asakawa, S., Kawasaki, K., Sasaki, T., Shintani, A., Shimizu, A., Shibuya, K., Kudoh, J., Minoshima, S., Ramser, J., Seranski, P., Hoff, C., Poustka, A., Reinhardt, R., Lehrach, H., and Consortiu, I.H.G.M. (2001). A physical map of the human genome. Nature 409, 934-941.

Meinke, D.W., Cherry, J.M., Dean, C., Rounsley, S.D., and Koornneef, M. (1998). *Arabidopsis thaliana*: A model plant for genome analysis. Science 282, 662.

Metzker, M.L. (2010). Sequencing technologies - the next generation. Nature Reviews Genetics 11, 31-46.

Meyers, B.C., Scalabrin, S., and Morgante, M. (2004). Mapping and sequencing complex genomes: Let's get physical! Nature Reviews Genetics 5, 578-U571.

Muyle, A., Serres-Giardi, L., Ressayre, A., Escobar, J., and Glemin, S. (2011). GC-biased gene conversion and selection affect GC content in the *Oryza* genus (rice). Mol Biol Evol 28, 2695-2706.

Nachman, M.W. (2002). Variation in recombination rate across the genome: evidence and implications. Current Opinion in Genetics & Development 12, 657-663.

Neff, M.M., Neff, J.D., Chory, J., and Pepper, A.E. (1998). dCAPS, a simple technique for the genetic analysis of single nucleotide polymorphisms: experimental applications in *Arabidopsis thaliana* genetics. Plant J 14, 387-392.

Nelson, W.M., Dvorak, J., Luo, M.C., Messing, J., Wing, R.A., and Soderlund, C. (2007). Efficacy of clone fingerprinting methodologies. Genomics 89, 160-165.

Newman, C.W., and Newman, R.K. (2006). A brief history of barley foods. Cereal Foods World 51, 4-7.

Overbeek, R., Fonstein, M., D'Souza, M., Pusch, G.D., and Maltsev, N. (1999). The use of gene clusters to infer functional coupling. P Natl Acad Sci USA 96, 2896-2901.

Paterson, A.H., Bowers, J.E., Burow, M.D., Draye, X., Elsik, C.G., Jiang, C.X., Katsar, C.S., Lan, T.H., Lin, Y.R., Ming, R.G., and Wright, R.J. (2000). Comparative genomics of plant chromosomes. Plant Cell 12, 1523-1539.

Paterson, A.H., Bowers, J.E., Bruggmann, R., Dubchak, I., Grimwood, J., Gundlach, H., Haberer, G., Hellsten, U., Mitros, T., Poliakov, A., Schmutz, J., Spannagl, M., Tang, H.B., Wang, X.Y., Wicker, T., Bharti, A.K., Chapman, J., Feltus, F.A., Gowik, U., Grigoriev, I.V., Lyons, E., Maher, C.A., Martis, M., Narechania, A., Otillar, R.P., Penning, B.W., Salamov, A.A., Wang, Y., Zhang, L.F., Carpita, N.C., Freeling, M., Gingle, A.R., Hash, C.T., Keller, B., Klein, P., Kresovich, S., McCann, M.C., Ming, R., Peterson, D.G., Mehboob-ur-Rahman, Ware, D., Westhoff, P., Mayer, K.F.X., Messing, J., and Rokhsar, D.S. (2009). The Sorghum bicolor genome and the diversification of grasses. Nature 457, 551-556.

Paux, E., Roger, D., Badaeva, E., Gay, G., Bernard, M., Sourdille, P., and Feuillet, C. (2006). Characterizing the composition and evolution of homoeologous genomes in hexaploid wheat through BAC-end sequencing on chromosome 3B. Plant J 48, 463-474.

Paux, E., Sourdille, P., Salse, J., Saintenac, C., Choulet, F., Leroy, P., Korol, A., Michalak, M., Kianian, S., Spielmeyer, W., Lagudah, E., Somers, D., Kilian, A., Alaux, M., Vautrin, S., Berges, H., Eversole, K., Appels, R., Safar, J., Simkova, H., Dolezel, J., Bernard, M., and Feuillet, C. (2008). A physical map of the 1-gigabase bread wheat chromosome 3B. Science 322, 101-104.

Perovic, D., Stein, N., Zhang, H., Drescher, A., Prasad, M., Kota, R., Kopahnke, D., and Graner, A. (2004). An integrated approach for comparative mapping in rice and barley with special reference to the *Rph16* resistance locus. Funct Integr Genomics 4, 74-83.

Phillips, D., Nibau, C., Ramsay, L., Waugh, R., and Jenkins, G. (2010). Development of a molecular cytogenetic recombination assay for barley. Cytogenetic and Genome Research 129, 154-161.

Pillen, K., Zacharias, A., and Leon, J. (2004). Comparative AB-QTL analysis in barley using a single exotic donor of *Hordeum vulgare* ssp *spontaneum*. Theor Appl Genet 108, 1591-1601.

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

Pourkheirandish, M., Wicker, T., Stein, N., Fujimura, T., and Komatsuda, T. (2007). Analysis of the barley chromosome 2 region containing the six-rowed spike gene *vrs1* reveals a breakdown of rice-barley micro collinearity by a transposition. Theor Appl Genet 114, 1357-1365.

Poustka, A., Pohl, T., Barlow, D.P., Zehetner, G., Craig, A., Michiels, F., Ehrich, E., Frischauf, A.M., and Lehrach, H. (1986). Molecular Approaches to mammalian genetics. Cold Spring Harbor Symposia on Quantitative Biology 51, 131-139.

Quraishi, U.M., Abrouk, M., Bolot, S., Pont, C., Throude, M., Guilhot, N., Confolent, C., Bortolini, F., Praud, S., Murigneux, A., Charmet, G., and Salse, J. (2009). Genomics in cereals: from genome-wide conserved orthologous set (COS) sequences to candidate genes for trait dissection. Functional & Integrative Genomics 9, 473-484.

Reinheimer, J.L., Barr, A.R., and Eglinton, J.K. (2004). QTL mapping of chromosomal regions conferring reproductive frost tolerance in barley (*Hordeum vulgare* L.). Theor Appl Genet 109, 1267-1274.

Riera-Lizarazu, O., Vales, M.I., Ananiev, E.V., Rines, H.W., and Phillips, R.L. (2000). Production and characterization of maize chromosome 9 radiation hybrids derived from an oat-maize addition line. Genetics 156, 327-339.

Romanov, M.N., Price, J.A., and Dodgson, J.B. (2003). Integration of animal linkage and BAC contig maps using overgo hybridization. Cytogenetic and Genome Research 102, 277-281.

Rostoks, N., Mudie, S., Cardle, L., Russell, J., Ramsay, L., Booth, A., Svensson, J.T., Wanamaker, S.I., Walia, H., Rodriguez, E.M., Hedley, P.E., Liu, H., Morris, J., Close, T.J., Marshall, D.F., and Waugh, R. (2005). Genome-wide SNP discovery and linkage analysis in barley based on genes responsive to abiotic stress. Molecular Genetics and Genomics 274, 515-527.

Sachidanandam, R., Weissman, D., Schmidt, S.C., Kakol, J.M., Stein, L.D., Marth, G., Sherry, S., Mullikin, J.C., Mortimore, B.J., Willey, D.L., Hunt, S.E., Cole, C.G., Coggill, P.C., Rice, C.M., Ning, Z.M., Rogers, J., Bentley, D.R., Kwok, P.Y., Mardis, E.R., Yeh, R.T., Schultz, B., Cook, L., Davenport, R., Dante, M., Fulton, L., Hillier, L., Waterston, R.H., McPherson, J.D., Gilman, B., Schaffner, S., Van Etten, W.J., Reich, D., Higgins, J., Daly, M.J., Blumenstiel, B., Baldwin, J., Stange-Thomann, N.S., Zody, M.C., Linton, L., Lander, E.S., Altshuler, D., and Grp, I.S.M.W. (2001). A map of human genome sequence variation containing 1.42 million single nucleotide polymorphisms. Nature 409, 928-933.

Sadder, M.T., and Weber, G. (2001). Karyotype of maize (*Zea mays* L.) mitotic metaphase chromosomes as revealed by fluorescence in situ hybridization (FISH) with cytogenetic DNA markers. Plant Molecular Biology Reporter 19, 117-123.

Sadder, M.T., Ponelies, N., Born, U., and Weber, G. (2000). Physical localization of single-copy sequences on pachytene chromosomes in maize (*Zea mays* L.) by chromosome in situ suppression hybridization. Genome 43, 1081-1083.

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.

Schena, M., Shalon, D., Davis, R.W., and Brown, P.O. (1995). Quantitative monitoring of gene-expression patterns with a complementary-DNA microarray. Science 270, 467-470.

Schmidt, R. (2002). Plant genome evolution: lessons from comparative genomics at the DNA level. Plant Molecular Biology 48, 21-37.

Schnable, P.S., Hsia, A.P., and Nikolau, B.J. (1998). Genetic recombination in plants. Current Opinion in Plant Biology 1, 123-129.

Schnable, P.S., Ware, D., Fulton, R.S., Stein, J.C., Wei, F.S., Pasternak, S., Liang, C.Z., Zhang, J.W., Fulton, L., Graves, T.A., Minx, P., Reily, A.D., Courtney, L., Kruchowski, S.S., Tomlinson, C., Strong, C., Delehaunty, K., Fronick, C., Courtney, B., Rock, S.M., Belter, E., Du, F.Y., Kim, K., Abbott, R.M., Cotton, M., Levy, A., Marchetto, P., Ochoa, K., Jackson, S.M., Gillam, B., Chen, W.Z., Yan, L., Higginbotham, J., Cardenas, M., Waligorski, J., Applebaum, E., Phelps, L., Falcone, J., Kanchi, K., Thane, T., Scimone, A., Thane, N., Henke, J., Wang, T., Ruppert, J., Shah, N., Rotter, K., Hodges, J., Ingenthron, E., Cordes, M., Kohlberg, S., Sgro, J., Delgado, B., Mead, K., Chinwalla, A., Leonard, S., Crouse, K., Collura, K., Kudrna, D., Currie, J., He, R.F., Angelova, A., Rajasekar, S., Mueller, T., Lomeli, R., Scara, G., Ko, A., Delaney, K., Wissotski, M., Lopez, G., Campos, D., Braidotti, M., Ashley, E., Golser, W., Kim, H., Lee, S., Lin, J.K., Dujmic, Z., Kim, W., Talag, J., Zuccolo, A., Fan, C., Sebastian, A., Kramer, M., Spiegel, L., Nascimento, L., Zutavern, T., Miller, B., Ambroise, C., Muller, S., Spooner, W., Narechania, A., Ren, L.Y., Wei, S., Kumari, S., Faga, B., Levy, M.J., McMahan, L., Van Buren, P., Vaughn, M.W., Ying, K., Yeh, C.T., Emrich, S.J., Jia, Y., Kalyanaraman, A., Hsia, A.P., Barbazuk, W.B., Baucom, R.S., Brutnell, T.P., Carpita, N.C., Chaparro, C., Chia, J.M., Deragon, J.M., Estill, J.C., Fu, Y., Jeddeloh, J.A., Han, Y.J., Lee, H., Li, P.H., Lisch, D.R., Liu, S.Z., Liu, Z.J., Nagel, D.H., McCann, M.C., SanMiguel, P., Myers, A.M., Nettleton, D., Nguyen, J., Penning, B.W., Ponnala, L., Schneider, K.L., Schwartz, D.C., Sharma, A., Soderlund, C., Springer, N.M., Sun, Q., Wang, H., Waterman, M., Westerman, R., Wolfgruber, T.K., Yang, L.X., Yu, Y., Zhang, L.F., Zhou, S.G., Zhu, Q., Bennetzen, J.L., Dawe, R.K., Jiang, J.M., Jiang, N., Presting, G.G., Wessler, S.R., Aluru, S., Martienssen, R.A., Clifton, S.W., McCombie, W.R., Wing, R.A., and Wilson, R.K. (2009). The B73 maize genome: Complexity, diversity, and dynamics. Science 326, 1112-1115.

Schulte, D., Close, T.J., Graner, A., Langridge, P., Matsumoto, T., Muehlbauer, G., Sato, K., Schulman, A.H., Waugh, R., Wise, R.P., and Stein, N. (2009). The international barley sequencing consortium-At the threshold of efficient access to the barley genome. Plant Physiol 149, 142-147.

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

Shahinnia, F., Druka, A., Franckowiak, J., Morgante, M., Waugh, R., and Stein, N. (2011). High resolution mapping of Dense spike-ar (dsp.ar) to the genetic centromere of barley chromosome 7H. Theor Appl Genet.

Shimamoto, K., and Kyozuka, J. (2002). Rice as a model for comparative genomics of plants. Annual Review of Plant Biology 53, 399-419.

Shizuya, H., Birren, B., Kim, U.J., Mancino, V., Slepak, T., Tachiiri, Y., and Simon, M. (1992). Cloning and Stable Maintenance of 300-Kilobase-Pair Fragments of Human DNA in Escherichia-Coli Using an F-Factor-Based Vector. P Natl Acad Sci USA 89, 8794-8797.

Smilde, W.D., Halukova, J., Sasaki, T., and Graner, A. (2001). New evidence for the synteny of rice chromosome 1 and barley chromosome 3H from rice expressed sequence tags. Genome 44, 361-367.

Soderlund, C., Nelson, W., Shoemaker, A., and Paterson, A. (2006). SyMAP: A system for discovering and viewing syntenic regions of FPC maps. Genome Res 16, 1159-1168.

Somerville, C., and Koornneef, M. (2002). A fortunate choice: the history of *Arabidopsis* as a model plant. Nature Reviews Genetics 3, 883-889.

Sourdille, P., Saintenac, C., Falque, M., Martin, O.C., Paux, E., and Feuillet, C. (2009). Detailed recombination studies along chromosome 3B provide new insights on crossover distribution in wheat (Triticum aestivum L.). Genetics 181, 393-403.

Southern, E., Mir, K., and Shchepinov, M. (1999). Molecular interactions on microarrays. Nat Genet 21, 5-9.

Spearman, C. (2010). The proof and measurement of association between two things. International Journal of Epidemiology 39, 1137-1150.

Spielmeyer, W., Singh, R.P., McFadden, H., Wellings, C.R., Huerta-Espino, J., Kong, X., Appels, R., and Lagudah, E.S. (2008). Fine scale genetic and physical mapping using interstitial deletion mutants of Lr34 /Yr18: a disease resistance locus effective against multiple pathogens in wheat. Theor Appl Genet 116, 481-490.

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

Steemers, F.J., and Gunderson, K.L. (2005). Illumina, Inc. Pharmacogenomics 6, 777-782.

Stein, N. (2009). Physical mapping in the triticeae, Genetics and Genomics of the Triticeae, G.J. Muehlbauer and C. Feuillet, eds (Springer US), pp. 317-335.

Stein, N., Perovic, D., Kumlehn, J., Pellio, B., Stracke, S., Streng, S., Ordon, F., and Graner, A. (2005). The eukaryotic translation initiation factor 4E confers multiallelic recessive Bymovirus resistance in *Hordeum vulgare* (L.). Plant J 42, 912-922.

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.

Stephens, J.L., Brown, S.E., Lapitan, N.L.V., and Knudson, D.L. (2004). Physical mapping of barley genes using an ultrasensitive fluorescence in situ hybridization technique. Genome 47, 179-189.

Steuernagel, B., Taudien, S., Gundlach, H., Seidel, M., Ariyadasa, R., Schulte, D., Petzold, A., Felder, M., Graner, A., Scholz, U., Mayer, K.F.X., Platzer, M., and Stein, N. (2009). De novo 454 sequencing of barcoded BAC pools for comprehensive gene survey and genome analysis in the complex genome of barley. BMC Genomics 10, 547.

Suchankova, P., Kubalakova, M., Kovarova, P., Bartos, J., Cihalikova, J., Molnar-Lang, M., Endo, T.R., and Dolezel, J. (2006). Dissection of the nuclear genome of barley by chromosome flow sorting. Theor Appl Genet 113, 651-659.

Talbert, P.B., and Henikoff, S. (2010). Centromeres convert but don't cross. Plos Biology 8, e1000326.

Tao, Q.Z., Chang, Y.L., Wang, J.Z., Chen, H.M., Islam-Faridi, M.N., Scheuring, C., Wang, B., Stelly, D.M., and Zhang, H.B. (2001). Bacterial artificial chromosome-based physical map of the rice genome constructed by restriction fingerprint analysis. Genetics 158, 1711-1724.

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

Turuspekov, Y., Mano, Y., Honda, I., Kawada, N., Watanabe, Y., and Komatsuda, T. (2004). Identification and mapping of cleistogamy genes in barley. Theor Appl Genet 109, 480-487.

Ullrich, S.E. (2010). Significance, Adaptation, Production, and Trade of Barley. In Barley, Production, Improvement , and Uses (Wiley-Blackwell), pp. 3-13.

Valarik, M., Bartos, J., Kovarova, P., Kubalakova, M., de Jong, J.H., and Dolezel, J. (2004). High-resolution FISH on super-stretched flow-sorted plant chromosomes. Plant J 37, 940-950.

van der Vossen, E.A.G., van der Voort, J.N.A.M.R., Kanyuka, K., Bendahmane, A., Sandbrink, H., Baulcombe, D.C., Bakker, J., Stiekema, W.J., and Klein-Lankhorst, R.M. (2000). Homologues of a single resistance-gene cluster in potato confer resistance to distinct pathogens: a virus and a nematode. Plant J 23, 567-576.

Varshney, R.K., Marcel, T.C., Ramsay, L., Russell, J., Roder, M.S., Stein, N., Waugh, R., Langridge, P., Niks, R.E., and Graner, A. (2007). A high density barley microsatellite consensus map with 775 SSR loci. Theor Appl Genet 114, 1091-1103.

von Korff, M., Wang, H., Leon, J., and Pillen, K. (2006). AB-QTL analysis in spring barley: II. Detection of favourable exotic alleles for agronomic traits introgressed from wild barley (*H-vulgare* ssp *spontaneum*). Theor Appl Genet 112, 1221-1231.

Vu, G.T.H., Wicker, T., Buchmann, J.P., Chandler, P.M., Matsumoto, T., Graner, A., and Stein, N. (2010). Fine mapping and syntenic integration of the semi-dwarfing gene *sdw3* of barley. Functional & Integrative Genomics 10, 509-521.

Wei, F., Coe, E., Nelson, W., Bharti, A.K., Engler, F., Butler, E., Kim, H., Goicoechea, J.L., Chen, M., Lee, S., Fuks, G., Sanchez-Villeda, H., Schroeder, S., Fang, Z., McMullen, M., Davis, G., Bowers, J.E., Paterson, A.H., Schaeffer, M., Gardiner, J., Cone, K., Messing, J., Soderlund, C., and Wing, R.A. (2007). Physical and genetic structure of the maize genome reflects its complex evolutionary history. Plos Genetics 3, 1254-1263.

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

Wicker, T., Buchmann, J.P., and Keller, B. (2010). Patching gaps in plant genomes results in gene movement and erosion of colinearity. Genome Res 20, 1229-1237.

Wicker, T., Taudien, S., Houben, A., Keller, B., Graner, A., Platzer, M., and Stein, N. (2009). A whole-genome snapshot of 454 sequences exposes the composition of the barley genome and provides evidence for parallel evolution of genome size in wheat and barley. Plant J 59, 712-722.

Wicker, T., Narechania, A., Sabot, F., Stein, J., Vu, G.T., Graner, A., Ware, D., and Stein, N. (2008). Low-pass shotgun sequencing of the barley genome facilitates rapid identification of genes, conserved non-coding sequences and novel repeats. BMC Genomics 9, 518.

Wu, J.Z., Mizuno, H., Hayashi-Tsugane, M., Ito, Y., Chiden, Y., Fujisawa, M., Katagiri, S., Saji, S., Yoshiki, S., Karasawa, W., Yoshihara, R., Hayashi, A., Kobayashi, H., Ito, K., Hamada, M., Okamoto, M., Ikeno, M., Ichikawa, Y., Katayose, Y., Yano, M., Matsumoto, T., and Sasaki, T. (2003). Physical maps and recombination frequency of six rice chromosomes. Plant J 36, 720-730.

Wu, X.L., Zhong, G.H., Findley, S.D., Cregan, P., Stacey, G., and Nguyen, H.T. (2008). Genetic marker anchoring by six-dimensional pools for development of a soybean physical map. BMC Genomics 9, 28.

Xu, Z., Kohel, R.J., Song, G., Cho, J., Yu, J., Yu, S., Tomkins, J., and Yu, J.Z. (2008). An integrated genetic and physical map of homoeologous chromosomes 12 and 26 in Upland cotton (G. *hirsutum* L.). BMC Genomics 9, 108.

Yan, L., Loukoianov, A., Tranquilli, G., Helguera, M., Fahima, T., and Dubcovsky, J. (2003). Positional cloning of the wheat vernalization gene *VRN1*. P Natl Acad Sci USA 100, 6263-6268.

Yang, T.J., Yu, Y., Nah, G., Atkins, M., Lee, S., Frisch, D.A., and Wing, R.A. (2003). Construction and utility of 10-kb libraries for efficient clone-gap closure for rice genome sequencing. Theor Appl Genet 107, 652-660.

Yao, H., Zhou, Q., Li, J., Smith, H., Yandeau, M., Nikolau, B.J., and Schnable, P.S. (2002). Molecular characterization of meiotic recombination across the 140-kb multigenic a1-sh2 interval of maize. P Natl Acad Sci USA 99, 6157-6162.

Yim, Y.S., Moak, P., Sanchez-Villeda, H., Musket, T.A., Close, P., Klein, P.E., Mullet, J.E., McMullen, M.D., Fang, Z., Schaeffer, M.L., Gardiner, J.M., Coe, E.H., and Davis, G.L. (2007). A BAC pooling strategy combined with PCR-based screenings in a large, highly repetitive genome enables integration of the maize genetic and physical maps. BMC Genomics 8, 47.

You, F.M., Luo, M.C., Xu, K.N., Deal, K.R., Anderson, O.D., and Dvorak, J. (2010). A new implementation of high-throughput five-dimensional clone pooling strategy for BAC library screening. BMC Genomics 11, 692.

You, F.M., Huo, N.X., Gu, Y.Q., Luo, M.C., Ma, Y.Q., Hane, D., Lazo, G.R., Dvorak, J., and Anderson, O.D. (2008). BatchPrimer3: A high throughput web application for PCR and sequencing primer design. Bmc Bioinformatics 9, -.

Yu, J., Hu, S.N., Wang, J., Wong, G.K.S., Li, S.G., Liu, B., Deng, Y.J., Dai, L., Zhou, Y., Zhang, X.Q., Cao, M.L., Liu, J., Sun, J.D., Tang, J.B., Chen, Y.J., Huang, X.B., Lin, W., Ye, C., Tong, W., Cong, L.J., Geng, J.N., Han, Y.J., Li, L., Li, W., Hu, G.Q., Huang, X.G., Li, W.J., Li, J., Liu, Z.W., Li, L., Liu, J.P., Qi, Q.H., Liu, J.S., Li, L., Li, T., Wang, X.G., Lu, H., Wu, T.T., Zhu, M., Ni, P.X., Han, H., Dong, W., Ren, X.Y., Feng, X.L., Cui, P., Li, X.R., Wang, H., Xu, X., Zhai, W.X., Xu, Z., Zhang, J.S., He, S.J., Zhang, J.G., Xu, J.C., Zhang, K.L., Zheng, X.W., Dong, J.H., Zeng, W.Y., Tao, L., Ye, J., Tan, J., Ren, X.D., Chen, X.W., He, J., Liu, D.F., Tian, W., Tian, C.G., Xia, H.G., Bao, Q.Y., Li, G., Gao, H., Cao, T., Wang, J., Zhao, W.M., Li, P., Chen, W., Wang, X.D., Zhang, Y., Hu, J.F., Wang, J., Liu, S., Yang, J., Zhang, G.Y., Xiong, Y.Q., Li, Z.J., Mao, L., Zhou, C.S., Zhu, Z., Chen, R.S., Hao, B.L., Zheng, W.M., Chen, S.Y., Guo, W., Li, G.J., Liu, S.Q., Tao, M., Wang, J., Zhu, L.H., Yuan, L.P., and Yang, H.M. (2002). A draft sequence of the rice genome (*Oryza sativa* L. ssp *indica*). Science 296, 79-92.

Yu, Q., Tong, E., Skelton, R.L., Bowers, J.E., Jones, M.R., Murray, J.E., Hou, S., Guan, P., Acob, R.A., Luo, M.C., Moore, P.H., Alam, M., Paterson, A.H., and Ming, R. (2009). A physical map of the papaya genome with integrated genetic map and genome sequence. BMC Genomics 10, 371.

Yuan, Q.P., Liang, F., Hsiao, J., Zismann, V., Benito, M.I., Quackenbush, J., Wing, R., and Buell, R. (2000). Anchoring of rice BAC clones to the rice genetic map *in-silico* . Nucleic Acids Research 28, 3636-3641.

Zhang, H.B., and Wing, R.A. (1997). Physical mapping of the rice genome with BACs. Plant Molecular Biology 35, 115-127.

9. Supplementary material

SupData_Table1. Genetically mapped genome zipper based markers

Gene Model ID (Mayer et al., 2011)	Lab Marker ID	Selected barley 454 shotgun read	Forward Primer	Reverse Primer	Annealing TM (Forward-Reverse)	Actual PCR product size(bp)	Polymorphism type	Restriction enzyme
GM.3296	Read3065	FTR0EHA01D1BQB	GTGGGACCAAAGATCCTTATGA	CCATGGCATTTTGCATACCT	60.19-60.72	430	Presence/Absence	-
GM.3334	Read3072	FVG2LPQ01DRMO9	TACGACATGCGCAAGATGCT	GCCGCATACACAAGATCCAA	62.87-61.98	318	CAPS	Cac8I
GM.3330	Read3076	FVG2LPQ02GQ7D6	AATGTTATGGGGGCAGTTGA	CATGGCTTCAAGTTTCCATTG	60.19-60.49	375	CAPS	TagI
GM.3312	Read3087	FVG2LPQ01BHJNH	CCATCGCGGTGAACTTGAT	CGACTGCAAGTAGCAGGAGA	62.03-59.34	223	CAPS	BtsI
GM.3309	Read3090	FTR0EHA01DDH00	ATCTGGTCGTCGGAGAAGAG	CCAGGACTTCGAGACCATCT	59.4-59.25	379	CAPS	AvaII
GM.3345	Read3103	FVG2LPQ02ITLYO	CATGCATCCAGGAAGGAAGT	CATCATCACCTTCAACGGAAA	60.07-60.88	306	CAPS	SfaNI
GM.3364	Read3116	FVG2LPQ02H21XB	GATTGGTTGACACGGTTCAAAA	TTCAAGCCACTGACAATCTCC	62.02-60.25	424	INDEL	-
GM.3371	Read3123	FVG2LPQ01EX1B2	ACTCTTAGCGGCCAAAGGAA	TGGCCAGTGAATCTGTGATG	61.24-60.69	401	CAPS	SMLI
GM.3390	Read3146	FVG2LPQ01AKLBW	CCATCACCTACATGCACGTCTA	ATGGTAAGCACGGCCAGTTA	60.96-60.52	386	CAPS	BbvI
GM.3401	Read3155	FVG2LPQ01C4RA8	AGTGTTCTGCTCTACCCCAAGA	AATCAATCTGTCGTGACTTCGAG	60.3-60.68	406	CAPS	AvaI
GM.3408	Read3157	FTR0EHA01CBJ4H	GTTCACGTTCGTGATGCAGA	AGGAAGTAGGCGACCGAAAT	60.88-60.1	216	CAPS	HgaI
GM.3416	Read3169	FVG2LPQ02GI3E0	GAGTCGGAATCGGAAAGGAT	TTGAAACGTGCCATGAAAGA	60.41-60.23	378	CAPS	BstNI
GM.3456	Read3200	FTR0EHA01AKN8P	TAATGGTCCAGTCGAACGAGTC	GAAGATCATGGCTCGCATCT	61.43-60.33	278	CAPS	AvaII
GM.3453	Read3203	FVG2LPQ01AI2MC	TTGCATTGCCTTATTCTCAGG	TGCTTCTGGTGCTGAACTGA	60.22-60.74	431	CAPS	Cac8I
GM.3471	Read3207	FVG2LPQ01DCP5M	CGGTTTGGATACATCGAGGA	CTTCCAGCACCGTAGAACTGAT	60.85-60.68	422	CAPS	BseYI
GM.3468	Read3210	FTR0EHA01BTGM2	TCCGATGGTGCTCAACTACA	CGGTCATGATCCGTACCTTATT	60.26-60.1	362	CAPS	HaeIII
GM.3460	Read3217	FVG2LPQ01CBHFE	CCCGTCCAGCCATAAACATA	GAGGCGTTACCTCAAACAGCTA	60.71-60.79	389	CAPS	BtgI
GM.3486	Read3230	FVG2LPQ02IPH6O	AAGCTCTGCTTGCAAAATGC	CAAACATAGCTTCGTATGCCCTA	60.67-60.5	334	CAPS	Aci I
GM.3490	Read3233	FVG2LPQ02F29GE	TATATGCCACGACCACGAGA	TGAAAACCACTTGTGCCAAA	60.1-60.13	207	INDEL	-
GM.3498	Read3237	FVG2LPQ01C57SH	GACGCTGAACTTTGGAAGGA	CGTTCCGTTCAACTTGTTCC	60.38-60.53	326	CAPS	BtgI
GM.3496	Read3239	FTR0EHA01DSAQ5	ATTGAGGGTCTGCCTCAACA	TGTACAAGGAATTTCGGAAAGC	60.66-60.47	499	CAPS	ApoI
GM.3509	Read3249	FVG2LPQ01AHVDR	CAACCACAATCGACTCATGC	GCCCTGAGCATCCTAGTGAA	60.12-60.36	244	INDEL	-
GM.3308	2H_Read3084	FTR0EHA01EMLU3	CGTCTCATTGTGAAGCGAAGT	TGAATGCAAGCTACCGTTAGTG	60.45-60.32	358	INDEL	-
GM.3313	2H_Read3089	FTR0EHA01EC1FO	CAGGTCTCGGTTTCGCTTCT	ATAACCACGGGTGTATCAGCA	61.84-60.26	251	CAPS	MseI
GM.3344	2H_Read3116	FVG2LPQ01DZXVE	CTCACAGGGTGCCAAGCA	GAAGCCAGCCGGTTTCAG	62.1-61.89	197	INDEL	-
GM.3386	2H_Read3154	FTR0EHA01ETWLM	AGGCTCGAGAAAGAACTCCAC	ATAGGCAACCTCATGGCTCTT	60.01-60.11	342	CAPS	MscI
GM.3389	2H_Read3156	FVG2LPQ01DWB13	CCTGAGCCAAGAGATGATCC	CGAAACACTTCCCTTCGACTT	59.76-60.65	233	CAPS	HindIII
GM.3396	2H_Read3162	FTR0EHA01BDJ5O	GGAGACACGAGGAGACAGGTTA	TGTACCTCAGGCTTGGATCTG	60.68-60.26	302	CAPS	TagI
GM.3404	2H_Read3170	FVG2LPQ01CFHN4	CTTGGGAGCTGTATGAAGCAG	TTGTTTCACCTAGGGCTGATG	60.02-60.12	355	CAPS	BanI
GM.3406	2H_Read3172	FVG2LPQ02JH9VL	CCTCAGCATTCAGCAATGAAC	TGGGAGAGAAGGAGAAGGTCT	60.79-59.42	200	CAPS	Cac8I
GM.3415	2H_Read3178	FVG2LPQ02HKEQH	CCATCACTAACCTTTCCACGA	ATTATGGGACGGAGGGAGTAG	59.98-59.32	287	INDEL	-
GM.3417	2H_Read3181	FVG2LPQ01EHKI4	AAGGAGCAAAGGATGTCGAA	CTCTAGAATGTGCGTGGGTGT	59.81-60.19	394	CAPS	AvaII
GM.3419	2H_Read3183	FVG2LPQ02FV3VG	TGGCTTATCTCTGAACCCAAAC	GCTCTTTGTGGTGCATCAACT	60.48-60.31	336	CAPS	TagI

SupData_Table1; Continued

Gene Model ID (Mayer et al., 2011)	Lab Marker ID	Selected barley 454 shotgun read	Forward Primer	Reverse Primer	Annealing TM (Forward -Reverse)	Actual PCR product size(bp)	Polymorphism type	Restriction enzyme
GM.3446	2H_Read3207	FTR0EHA01B6W4U	GCCGGACAAGGGAGTTTACTA	TGGCAGTACGCCAGAATCTAC	60.49-60.29	373	CAPS	SspI
GM.3470	2H_Read3212	FTR0EHA01CP1J8	CGAAGAGGTAGAAGGTGAGCA	CCTACTTCAACTACCGCACCA	59.62-60.18	402	CAPS	BseyI
GM.3454	2H_Read3218	FVG2LPQ02FMYJ4	AAACATGTGCTCGCTCCTCT	AGGTCAAGCGGCAATATGAG	60.02-60.24	285	CAPS	Fnu4HI
GM.3475	2H_Read3233	FTR0EHA01CW6DB	GCTCCATCTTCAACGTCTCC	ACGGAGTGGAAGTAGTTGTCG	59.81-59.27	292	CAPS	StyI
GM.3492	2H_Read3247	FVG2LPQ01D9KT5	ATGGCTTGGTAGTGATGTTGG	TCTTTGCTGAGTAGCGACCAT	59.87-60.03	402	CAPS	Aci I
GM.3329	D4-Read3077	FTR0EHA01D8GKV	TACCACGTGCCGAGCTTAAA	TTGGTTAATGAGATATCATTTACTTCG	58.4-60.8	120	dCAPS	TagI
GM.3465	D-Read3215	FVG2LPQ01B5701	TCCACACCAGCCACACTTATAG	TTCATGACTCGCCAATCCTG	62.2-58.4	236	dCAPS	HhaI
GM.3481	D1-Read3227	FVG2LPQ02JSX9H	AGAGAGCTAGTTTCAGATATTACG	GCACACTCAAGTGTCATTGCTT	60.3-60.1	357	dCAPS	BsaI
GM.3500	D1-Read3243	FVG2LPQ01COQD6	TGCCTGTAGTTCATACATCGTA	CACCTGTGGCAAGAGGAAAT	58.4-58.4	323	dCAPS	RsaI
GM.3395	D3-2H_Read3161	FVG2LPQ02JI22U	GGAAGGCTATCATCGAAAGAGTC	CCTGCAAAACTTCAGGACTA	62.9-56.4	111	dCAPS	SpeI
GM.3424	D-2H_Read3188	FVG2LPQ01BLICS	TCCGGTACCTCATAAGGGAAG	CATTACTAGAGATTACCTTATGGTA	61.2-59.5	123	dCAPS	RsaI
GM.3502	2H_Read3256	FVG2LPQ01B9Y70	ATCGGACTGTCGAGCACTTTA	GTTGAGAGCATTGCAGTGAGG	59.89-60.99	285	CAPS	Dde1

SupData_Table2. FISH probes and the related FPcontig/Sequence information

2H FPcontig ID	Sequence ID (WGS contig/sequenced BAC)	Porbe ID	Primer ID_Forward	Primer Seq_Forward	Primer ID_Reverse	Primer Seq_Reverse	amplified size on genomic DNA(cv. Morex)_kbp
803	contig_1567582	1_1	Ctg_1567582_1-F	CTCATCAAGGTC AGATCCAGCTTA	Ctg_1567582_1-R	GGATCCCAATGA GAAGCTACAGA	1483
803	contig_45164	1_2	Ctg_45164_1-F	CACATCTCCCAA GGTTCAATCAG	Ctg_45164_1-R	GGCGATTGGGTA TGAATTTAGC	1435
803	contig_45164	1_3	Ctg_45164_2-F	ATATGATTGCCA GGAACCGAAC	Ctg_45164_2-R	CCTAGCAGCAGT GGGAAGACTAA	1904
803	contig_1567582	1_4	Ctg_1567582_2-F	GAGGGCATCAAC GTCAACAATA	Ctg_1567582_2-R	CTGGTGATCATA TCGCAGAAGC	2199
534	contig_1560534	2_1	Ctg_1560534_2-F	CACTGCTTGCTTG ATATCCTCCT	Ctg_1560534_2-R	CTAGCTTGGGTT GGGTGTTTGT	1975
534	contig_1560534	2_2	Ctg_1560534_3-F	CCTTATACTTTGC AGCGCATGT	Ctg_1560534_3-R	CCGGTGTCTAAT CAGGAAGGAG	2040
534	contig_1560534	2_3	Ctg_1560534_1-F	GAGAGGAGACAC CACCATGATTT	Ctg_1560534_1-R	TTGTGCTGAGGG TCGTAGGTATT	2436
46369	HVVMRXALLhA0182J03_v40_c4	3_1	A0182J03_c4_P005	AGAAGGACCCAG GTCCAAATTA	A0182J03_c4_P005	TCCATGCTGCAG TGATGATGTA	2941
46369	HVVMRXALLhA0182J03_v40_c5	3_2	A0182J03_c5_P025	CCTGTAACTTCTT CCCGTTGCT	A0182J03_c5_P025	TTTGTGTGGAAG ACGATGGAAG	3030
38570	contig_136106	4_1	Ctg_136106_2-F	GATGATAGCGGC GATATTGTGAG	Ctg_136106_2-R	GGAATACCGTAT CCAAACTCGAAG	1445
38570	contig_136106	4_3	Ctg_136106_1-F	GATTAATCAGCC CCACGAAGAT	Ctg_136106_1-R	GGAGTTACAATG TCGTTGCCATC	1439
38570	contig_46527	4_4	Ctg_46527_2-F	TGGTAGACAGAA GACCCGAGTTC	Ctg_46527_2-R	GTCCAGAATGTG CATATGGGTTC	2488
45093	contig_5743	5_1	Ctg_5743_1-F1	TGGTTCAACCTTG TGTACCACTG	Ctg_5743_1-R1	AAGAGGATAGGC AACCTCATGG	2080
45093	contig_5743	5_2	Ctg_5743_2-F	AACTTCCAAAGG TCGTTTCCAC	Ctg_5743_2-R	CCCATCTCGTTTC GATCTGTTT	2719
2926	contig_1568600	6_1	Ctg_1568600_5-F	CAACAGCCATTT CCAGGTACAC	Ctg_1568600_5-R	TTCAGACAATGC AGCTCTCACA	2346
2926	contig_1568600	6_3	Ctg_1568600_3-F	TTGGTTTGACTGG TTGATCATGTA	Ctg_1568600_3-R	GGAAGATCAGCC AACCGTAAATAG	1641
2926	contig_1568600	6_4	Ctg_1568600_2-F	CAGGCTTTATGAT CCTCGTGTG	Ctg_1568600_2-R	AATTCGTACCAC TTTGCAGCAG	2733
2926	contig_1568600	6_5	Ctg_1568600_1-F	ATGGTGTTGAAC GTATGCCAAA	Ctg_1568600_1-R	TGCAACCGAAGT GTAAGAATGG	2022
46608	contig_37676	7_1	Ctg_37676_1-F	TGAGGTACAATA CCTGCTCAACG	Ctg_37676_1-R	GAATAGCGTGGC TCCAATCATAG	1496
46608	contig_2553533	7_2	Ctg_2553533_1-F	CTCTCATCGGTGC TCAGTGG	Ctg_2553533_1-R	CCCAGGTTCCTTC TCAACCAT	1332
46608	contig_2553533	7_3	Ctg_2553533_2-F	AGCTTAGCTGAC TTAGGGCCAGT	Ctg_2553533_2-R	GCATACGCTGAG AGAAGTTACCC	1571
494	contig_2548179	8_1	Ctg_2548179_1-F	GCAGAGAAATGG GTGTAGAACAAG	Ctg_2548179_1-R	GGATTTCTGTCA AGCACCATTG	1451
494	contig_2548179	8_2	Ctg_2548179_2-F	GGATCAGGTACG CAATACAGGAC	Ctg_2548179_2-R	TTACTGTGGAAT CTGTGCCTCTG	1577
494	contig_2548179	8_3	Ctg_2548179_3-F	AGCAACATATCC CCATTTGGTC	Ctg_2548179_3-R	TGGAAGGTGGGC ATTTTCTAGT	1700
44808	contig_49337	9_1	Ctg_49337_2-F	GACTTTGCTGGA ATGGATCCTC	Ctg_49337_2-R	ACCGATGTCATT GAACTGATGC	2000
44808	contig_49337	9_2	Ctg_49337_1-F	CGACAAGTTCCA ACATGTTTCC	Ctg_49337_1-R	CATTGAGACGCT GGATTAGCAG	2188

SupData_Table3. List of enzymes, chemicals and kits utilized

Restriction enzyme	Company	Chemicals	Company	Kits	Company
Aci I	New England Biolabs, Roshe, Switzerland	UltraPure TM Agarose	Invitrogen GmbH, Germany	BigDye® Terminator v3.1 Cycle Sequencing Kit	ABI, CA, USA
Apo I	New England Biolabs, Roshe, Switzerland	Ethylenediaminetetraacetic acid (EDTA)	Roth, Karlsruhe, Germany	Hi-DI TM Formamide	ABI, CA, USA
Ava I	New England Biolabs, Roshe, Switzerland	Ethanol (EtOH)	Roth, Karlsruhe, Germany	QIAquick PCR Purification Kit	QIAGEN, Hilden, Germany
Ava II	New England Biolabs, Roshe, Switzerland	Ethidium bromide (EtBr)	Roth, Karlsruhe, Germany	Type-it® HRM PCR kit	QIAGEN, Hilden, Germany
Ban I	New England Biolabs, Roshe, Switzerland	Sodium acetate, CH3COONa	Roth, Karlsruhe, Germany	NucleoFast TM 96 PCR Purification Kit	MACHEREY-NAGEL GmbHg, Duren, Germany
Bbv I	New England Biolabs, Roshe, Switzerland	Cellulase R10	Duchefa, Haarlem, Netherland		
Bsa I	New England Biolabs, Roshe, Switzerland	Cellulase	CalBioChem, San Diego, USA		
BseY I	New England Biolabs, Roshe, Switzerland	Pectolyase	Sigma, Taufkirchen, Germany		
BstN I	New England Biolabs, Roshe, Switzerland	Cytohelicase	Sigma, Taufkirchen, Germany		
Btg I	New England Biolabs, Roshe, Switzerland	DNA polymerase	Fermentas, Sankt Leon-Rot, Germany		
Bts I	New England Biolabs, Roshe, Switzerland	DNase I	Fermentas, Sankt Leon-Rot, Germany		
Cac8 I	New England Biolabs, Roshe, Switzerland	Texas Red-12-dUTP	Invitrogen, Oregon, USA		
Cac8 I	New England Biolabs, Roshe, Switzerland	Fluor 488-5-dUTP	Invitrogen, Oregon, USA		
Dde I	New England Biolabs, Roshe, Switzerland				
Fnu4H I	New England Biolabs, Roshe, Switzerland				
Hae III	New England Biolabs, Roshe, Switzerland				
Hga I	New England Biolabs, Roshe, Switzerland				
Hha I	New England Biolabs, Roshe, Switzerland				
Hind III	New England Biolabs, Roshe, Switzerland				
Msc I	New England Biolabs, Roshe, Switzerland				
Mse I	New England Biolabs, Roshe, Switzerland				
Rsa I	New England Biolabs, Roshe, Switzerland				
SfaN I	New England Biolabs, Roshe, Switzerland				
SML I	New England Biolabs, Roshe, Switzerland				
Spe I	New England Biolabs, Roshe, Switzerland				
Ssp I	New England Biolabs, Roshe, Switzerland				
Sty I	New England Biolabs, Roshe, Switzerland				
Taq I	Fermentas				
HotStar Taq DNA polymerase	Qiagen, Hilden, Germany				

10. Publications and proceedings related to the submitted thesis

- **Naser Poursarebani**, Mihaela Maria Martis, Ruvini Ariyadasa, Daniela Schulte, Ruonan Zhou, Burkhard Steuernagel, Andreas Graner, Uwe Scholz, Klaus Mayer, Nils Stein, 2012. Virtual gene order map -a novel resource for marker development and physical map anchoring in barley. (In preparation).

- Lu Ma, Thomas Schmutzer, **Naser Poursarebani**, Burkhard Steuernagel, Uwe Scholz, Andreas Graner, Andreas Houben, Nils Stein, 2012. High-throughput single copy FISH probe detection in barley. (In preparation).

- The International Barley Genome Sequencing Consortium (IBSC**), 2012.** Genomic stratification reveals the 5.1 Gb GENE-OME of cultivated barley. (In preparation).

- **Naser Poursarebani**, Lu Ma, Thomas Schmutzer, Ruvini Ariyadasa, Burkhard Steuernagel, Heidrun Gundlach, Uwe Scholz, Klaus Mayer, Andreas Graner, Andreas Houben, Nils Stein. Providing the link between physical and cytogenetic map of barley - a case study on chromosome 2H. Plant & Animal Genomes XX Conference, San Diego, USA. 2012.

- R. Ariyadasa, T. Nussbaumer, D. Schulte, J. Poland, **N. Poursarebani**, R. Zhou, B. Steuernagel, H. Liu, P.E. Hedley, R. Waugh, H. Gundlach, P. Langridge, K. Mayer, U. Scholz, S. Taudien, M. Felder, M. Platzer, Z. Frenkel, A. Korol, K. Madishetty, M. Moscou, T. Close, A. Graner, M. Mascher, J. Keilwagen, N. Stein. The barley physical map – impact of different BAC libraries made of partially digested or randomly sheared genomic DNA. Plant & Animal Genomes XX Conference, San Diego, USA. 2012.

- **Naser Poursarebani**, Mihaela Maria Martis, Ruvini Ariyadasa, Daniela Schulte, Ruonan Zhou, Burkhard Steuernagel, Andreas Graner, Uwe Scholz, Klaus Mayer, Nils Stein. Potential of Grass Genome Synteny for Genetic Anchoring of the Physical Map of Barley (*Hordeum vulgare* L.) - A Case Study On Chromosome 2H. Plant & Animal Genomes XIX Conference, San Diego, USA. 2011.

- Daniela Schulte, Ruvini Ariyadasa, **Naser Poursarebani**, Ruonan Zhou, Tatjana Sretenovic-Rajicic, Peter Langridge, Bu- Jun Shi, Klaus Mayer, Timothy Close, Stephan Weise, Uwe Scholz, Andreas Graner, Nils Stein. Whole Genome Physical Map of Barley (*Hordeum vulgare* L.). Plant & Animal Genomes XVIII Conference, San Diego, USA. 2010.

- Ruvini Ariyadasa, Peter Wenzl, **Naser Poursarebani**, Daniela Schulte, Andrzej Kilian, Andreas Graner, Nils Stein. Steps Towards An Integrated Physical And Genetic Map Of Barley (*Hordeum vulgare* L.). Plant & Animal Genomes XVII Conference, San Diego, USA. 2009.

11. Eidesstattliche erklärung

Herr Naser Poursarebani

Hans-Stubbe Str. 22

06466 Gatersleben

Hiermit erkläre ich, dass mit dieser wissenschaftlichen Arbeit noch keine vergeblichen Promotionsversuche unternommen wurden.

Die eingereichte Dissertation mit dem Thema: *"**A genetically anchored physical map of barley chromosome 2H**"* habe ich selbständig und nur unter Verwendung der angegeben Literatur und Hilfsmittel angefertigt. Des Weiteren erkläre ich, dass keine Strafverfahren gegen mich anhängig sind.

Gatersleben, den …….

Naser Poursarebani

12. Curriculum Vitae

Family name: Poursarebani, **Given name:** Naser
Address: Leibniz Institute of Plant Genetics and Crop Plant Research (IPK), GeneBank Dept. Plant Architecture Group, Corrensstr. 3, 06466, Gatersleben, Germany.
Phone number: +49-39482-5754
E-mail: Poursarebani@ipk-gatersleben.de, Poursarebani@gmail.com
Nationality: Iranian
Date of Birth: 22/11/1978
Place of Birth: Minab / Iran
Gender: male
Marital status: married
Academic Education:

- **Ph.D.** PhD enrolled at Martin Luther University, Halle, Germany (Feb. 2008 – March. 2012), and the thesis was prepared at: Leibniz Institute of Plant Genetics and Crop Plant Research (IPK), Gatersleben, Germany.
- **M.Sc.** Plant breeding, Agricultural College, Mazandaran University, Sari, Iran (Sep.2003 - Sep.2005), Thesis prepared at: Agricultural Biotechnology Research Institute of Iran (ABRII), Karaj, Iran.
- **B.Sc.**; Agronomy and Plant Breeding, Agricultural College, Yasouj University, Yasouj, Iran (Sep.1998 – Jun.2002).
- **Diploma**; Life Sciences, Minab, Iran (1993-1997).

Researches:

- Towards a physical map of wheat chromosome 6A – The German contribution to the International Wheat Genome Sequencing Consortium (Post Doctoral Research position, IPK, Gatersleben, Germany)
- A genetically anchored physical map of barley chromosome 2H (Ph.D. thesis).
- Co-researcher in conducting a research related to analysis of genetic diversity in Iranian wild and cultivated pistachio genotypes using morphological traits and Molecular markers (AFLPs and SSRs), ABRII, Karaj, Iran (Sep2004- Jan2005).
- Determination of Hetrotic Groups between Maize Inbred Lines Using AFLP Markers, ABRII, Karaj, Iran (M.Sc thesis).

Honors:

- Gaining rank 39 in the Plant Breeding Entrance Examination for M.Sc. degree, among approximately 1000 applicants in Iran.(Feb.2003)
- Gaining second price for oral presentation in Plant Science Student Conference (PSSC-2010, IPK, Gatersleben, Germany)

Job Experiences:

- Working as PostDoc employee at the Institute of Plant Genetics and Crop Plant Research (IPK), Gatersleben, Germany (Since Jan.2012).
- Working on Ph.D. thesis as an employee at the Institute of Plant Genetics and Crop Plant Research (IPK), Gatersleben, Germany (Sep.2006-March.2010).
- Working as co- researcher on research projects in Hormozgan Agriculture and Natural Resources Research Center, Minab Station, Minab, Iran (Jan.2006- Sep.2008).
- Teaching a basic course in English (Reading and Sentence Structure) for University Student in Hormozgan Agricultural Education Center, Hajiabad, Iran. (Jan.2007-Dec.2008).
- Teaching Applied Statistics and Plant Breeding Principles for Agricultural Students at Hajiabad Azad University, Hajiabad, Iran (Jan.2007-Dec.2008).

……
Naser Poursarebani