Assessing the Barley Genome Zipper and Genomic Resources for Breeding Purposes

Cristina Silvar, Mihaela M. Martis, Thomas Nussbaumer, Nicolai Haag, Ruben Rauser, Jens Keilwagen, Viktor Korzun, Klaus F.X. Mayer, Frank Ordon, and Dragan Perovic*

Abstract
The aim of this study was to estimate the accuracy and convergence of newly developed barley (Hordeum vulgare L.) genomic resources, primarily genome zipper (GZ) and population sequencing (POPSEQ), at the genome-wide level and to assess their usefulness in applied barley breeding by analyzing seven known loci. Comparison of barley GZ and POPSEQ maps to a newly developed consensus genetic map constructed with data from 13 individual linkage maps yielded an accuracy of 97.8% (GZ) and 99.3% (POPSEQ), respectively, regarding the chromosome assignment. The percentage of agreement in marker position indicates that on average only 3.7% GZ and 0.7% POPSEQ positions are not in accordance with their centimorgan coordinates in the consensus map. The fine-scale comparison involved seven genetic regions on chromosomes 1H, 2H, 4H, 6H, and 7H, harboring major genes and quantitative trait loci (QTL) for disease resistance. In total, 179 GZ loci were analyzed and 64 polymorphic markers were developed. Entirely, 89.1% of these were allocated within the targeted intervals and 84.2% followed the predicted order. Forty-four markers showed a match to a POPSEQ-anchored contig, the percentage of collinearity being 93.2%, on average. Forty-four markers allowed the identification of twenty-five fingerprinted contigs (FPCs) and a more clear delimitation of the physical regions containing the traits of interest. Our results demonstrate that an increase in marker density of barley maps by using new genomic data significantly improves the accuracy of GZ. In addition, the combination of different barley genomic resources can be considered as a powerful tool to accelerate barley breeding.

BARLEY was domesticated in the Fertile Crescent about 10,000 yr ago (Badr et al., 2000) and independently in Tibet, as the adaptation to the extreme environmental conditions, about 3500 to 4000 yr ago (Dai et al., 2012). Today, barley is one of the most important cereal crops worldwide, ranking fourth in terms of total production (FAOSTAT, 2012). Such relevance arises from its versatility to adapt to different stress conditions and from its essential use in malting and brewing industries as well as for...
animal feed (Baik and Ullrich, 2008; Ceccarelli et al., 2010; Verstegen et al., 2014). Recent reports on barley’s health benefits have also promoted a renewed interest for this ancient food grain (Brockman et al., 2013; Sullivan et al., 2013). Apart from this key role in agriculture, the diploid and inbreeding nature of barley makes it also a very attractive model species for genetic studies within the Triticeae tribe (Bockelman and Valkoun, 2011). The major impediment for its full exploitation comes from the presence of a large and complex repeat-rich genome of 5.1 Gb (Dolezel et al., 1998). Nevertheless, progress in barley genetics and genomics research has been continuously moving forward (Graener et al., 2011; Kumlehn and Stein, 2014).

From the pioneering work of Sturtevant (1913), who constructed the first genetic map of barley, the mapping of genes, morphological traits, and, now, molecular markers and sequences was one of the most challenging tasks of many generations of geneticists. In the meantime, many tools and strategies for the ordering of markers and sequences were developed, but all of them had some advantageous and disadvantageous features (e.g., in introducing certain level of errors) (Romero et al., 2009; He et al., 2001). With the decline in the costs of next-generation sequencing (NGS) technologies and high-throughput genotyping platforms, permitting the generation of thousands of data points in a very short time, there is a genuine need for new methods for ordering of genetic data and strategies that assess the accuracy of the order. Among the ordering approaches, the barley GZ (Mayer et al., 2011) and POPSEQ (Mascher et al., 2013a) are the most advanced ones in barley genetics. Furthermore, single maps in combination with consensus maps create the basis for ultra-high dense map construction integrating many thousands of markers and NGS data. However, up to now, little is known about the error rate and precision of constructed maps. In last years, high-throughput techniques, that is, Illumina iSelect platform, genotyping-by-sequencing (GBS), along with the flow sorting of chromosomes have revolutionized barley genotyping (Simková et al., 2008; Muñoz-Amatriain et al., 2011, 2014a). For example, an Illumina 9K SNP chip based on sequence polymorphisms in 10 diverse cultivated barley genotypes and a GBS approach for barley have been recently developed (Comadran et al., 2012; Poland et al., 2012). In spite of the enormous progress in barley genomics, these are of limited use without the availability of a draft genome sequence. In 2012, the International Barley Sequencing Consortium (IBSC) generated a densely anchored physical map of the barley genome comprising 9265 fingerprinted bacterial artificial chromosome (BAC) contigs spanning 4.98 Gb (International Barley Genome Sequencing Consortium, 2012). Furthermore, Mayer et al. (2009, 2011) developed another genomic resource, which provided clues on the barley genome composition. They constructed linear ordered virtual gene maps of barley by using the so-called genome zipper approach. The barley GZ assembles 86% of the barley genes in a putative linear order along the individual barley chromosomes by exploiting the high syntenies among three reference grass genomes, namely Brachypodium distachyon (L.) Beauv., rice (Oryza sativa L.), and sorghum [Sorghum bicolor (L.) Moench] (Goff et al., 2002; Yu et al., 2002; Paterson et al., 2009; International Brachypodium Initiative, 2010).

More recently, POPSEQ facilitated the development of genetically ordered contigs from a whole-genome shotgun (WGS) assembly of barley cultivar Morex by genotyping a mapping population with shallow genome coverage (Mascher et al., 2013a). Subsequently, the new information provided by POPSEQ was employed to order and genetically anchor the barley physical map, establishing a minimum tilling path (MTP) that comprises more than 65,000 BAC clones (Ariyadasa et al., 2014). Recently, 15,622 BACs representing the MTP of 72,052 physical-mapped gene-bearing BACs, were identified and sequenced (Muñoz-Amatriain et al., 2015).

Undoubtedly, all these advancements will be extremely beneficial in a wide range of studies both in fundamental and applied barley research (Mascher and Stein, 2014). However, in the breeding context, the use of this information remains mostly unexploited. Only few reports in the last years have been partly focused on the application of the barley GZ for marker saturation of genetic intervals containing interesting traits, such as spike density, resistance to powdery mildew, barley yellow and mild mosaic virus, or barley yellow dwarf disease (Shahinnia et al., 2012; Silvar et al., 2013; Ordon and Perovic, 2013; Yang et al., 2013; Lüpken et al., 2013, 2014). In barley breeding, there is an urgent need for tools mostly directed to the quick and efficient identification of sets of molecular markers that are closely linked to the traits of interest. Such markers may be readily applied to marker-assisted selection (MAS), marker-assisted backcrossing strategies, or so called precision breeding (McCouch, 2004), which enable the selection of traits with greater accuracy and the ability to deploy them cost-effectively into new varieties (Collard and Mackill, 2008). Similarly, those markers would be advantageous in accelerating map-based cloning approaches (Bolger et al., 2014; Yang et al., 2014) followed by allelic mining and exploration of natural genetic variation (Muñoz-Amatriain et al., 2014b). The novel genomic resources of barley are valuable tools in this respect. Nevertheless, to take full advantage of these, it is essential to first evaluate and validate these tools in a breeding context.

In this present work, thirteen linkage maps (Muñoz-Amatriain et al., 2014a; Perovic et al., unpublished data) were used to construct a consensus map, which was compared with the barley GZ, POPSEQ, and the barley physical map. Furthermore, seven well-defined loci were employed to assess the same resources at a microsynteny level to get information on the accuracy and usefulness to accomplish fine mapping schemes and physical delimitation of genomic regions of targeted interest.
MATERIALS AND METHODS

Construction of a Consensus Marker Map
A set of 13 genetic linkage maps was used for the construction of a consensus map. Twelve of them were taken from Muñoz-Amatriain et al. (2014a), while the 13th map was developed on the cross MBR1012 × ‘Scarlett’ (Perovic et al., unpublished data). This latter population of 86 doubled-haploid lines was genotyped with the barley iSelect 9K SNP chip (Comadran et al., 2012). The consensus map was constructed using the R package LPmerge according to Endelman and Plomion (2014). In short, a consensus map was computed independently for each chromosome using a range of possible interval values from 1 to 5. These interval values specify the number of neighboring markers that were used to compute the consensus map. Subsequently, the RMSE values between the corresponding consensus map and the individual maps were computed for each possible interval and chromosome. Based on these RMSE values, the final consensus map was selected for a specific chromosome with the smallest value. All chromosomal maps were manually verified and additional single nucleotide polymorphism (SNP) markers, originating from different Illumina platforms, that is, the 9K iSelect chip (Comadran et al., 2012) and a set of 459 barley oligo pool assay (BOPA) markers (Close et al., 2009) were included.

Comparative Analysis of the Consensus Map to the barley Genome Zipper and the Population Sequencing Map
On the basis of common BOPA markers, the genetic positions between the consensus marker map of this study and the consensus map of Close et al. (2009) were identified and compared. The results of this comparison were visualized for each barley chromosome individually by generating dot plots and statistically evaluated by calculating Spearman correlations using the python packages Matplotlib (Hunter, 2007) and NumPy, respectively.

The consensus map created in this work was used in assessment of the robustness of the GZ and POPSEQ at the genome-wide level. For this purpose, the sequences of 8323 iSelect markers were aligned against the POPSEQ-anchored Morex contigs (Mascher et al., 2013a) and against the gene indices of the GZ for all seven barley chromosomes (Mayer et al., 2011) by using BLASTn (Altschul et al., 1997). The barley zipper dataset consists of anchored barley full-length complementary DNAs (cDNAs), barley markers, and genes from the reference genomes of B. distachyon, rice, and sorghum. Only the first best hits with an alignment length of at least 100 bp and location on the same chromosomes were considered. The quality of the observed overlap between the three maps (consensus map, virtual ordered gene map (GZ), and POPSEQ map) was assessed by dividing each position by the total map length and allowing a 10% difference. The recombination frequency was computed in nonoverlapping bins of 50 GZ loci. All consensus markers in a given bin were considered and the genetic distance between the marker with the highest and lowest position was computed. To filter for wrongly assigned marker, all markers with >10 cM, compared with the median position per bin, were removed. The results were statistically evaluated through a nonparametric measure of correlation (Spearman’s rank correlation coefficient) and visualized by using CIRCOS (Krzywinski et al., 2009).

Microcollinearity: Comparative Analysis of Seven Genetically Mapped Loci to Genome Zipper, Population Sequencing, and Barley Physical Map
Seven loci or QTL (subsequently termed from L1 to L7) located on five different barley chromosomes and genetically mapped in the context of other studies were used for comparative purposes at the microsynteny level: L1 (chromosome 1HS, RphMBR1012; König et al., 2012), L2 (rymy7, 1H centromere; Yang et al., 2013), L3 (2HL, Ryd19-2HL; Perovic et al., 2013a), L4 (ryml1, 4H centromere; Lüpken et al., 2013), L5 (Bq_QTL_6HL, 6HL; Silvar et al., 2011a, 2013), L6 (Bq_QTL_7HS, 7HS; Silvar et al., 2010, 2012, 2013), and L7 (Bq_QTL_7HL, 7HL; Silvar et al., 2010, 2012, 2013). These loci confer resistance to various barley pathogens and some of them (L1, L5, L6, and L7) are under assumption of being less conserved (Leister et al., 1998).

First, the synteny-ordered virtual gene map of barley (GZ) was validated by using this set of seven genetically mapped loci and developing markers from corresponding barley zippers followed by mapping as described in the original publications. For the short arm of chromosome 1H and the long arm of chromosome 2H (unpublished data) comparative analysis to barley zippers, marker design and genotyping was essentially done according to Perovic et al. (2004) and Silvar et al. (2013). Briefly, markers genetically flanking the regions of interest were used to select the target intervals in the virtual linear gene map. Zipper-based markers were used for amplification in both parental lines and amplicons were sequenced on an ABI377XL instrument using BigDye terminator sequencing chemistry (ABI PerkinElmer). Markers, for which polymorphisms were based on presence or absence of polymerase chain reaction (PCR) fragments between parental lines, were directly mapped. In turn, SNPs were transformed to cleaved amplified polymorphic sequence (CAPS) markers (Perovic et al., 2013b) or pyrosequencing markers using a biotin-labeled M13 primer (Silvar et al., 2011b). Linkage analyses were performed with JoinMap 4.0 (Van Ooijen, 2006). Second, 57 markers derived from the target intervals of the seven loci were compared with the POPSEQ-anchored contigs of Morex (Mascher et al., 2013a) by using BLAT (Kent, 2002) requiring an identity of 99% and a minimum match length of 50 nucleotides. The WGS contigs and the physical FPCs were linked to each other by BLAST using stringent criteria and only matches with
at least 99% sequence identity and at least 90% sequence coverage of the smallest sequence (either WGS contig or a physical contig) were considered. The genetic markers from the seven loci were also compared with 15,622 recently sequenced barley clones (Muñoz-Amatriain et al., 2015). Marker sequences were assigned to a clone with help of the Harvest BLAST (http://www.harvest-blast.org/) when clones matched a marker with an e-value of $10^{-5}$ and with a match length of at least 100 nucleotides. Clones were further linked to physical contigs on the same website (http://harvest-web.org/hweb/utilmenu.wc?job=RTRVFORM&db=MOREX_HV3_10.4.1) where a mapping of clones to the physical map of the IBSC is provided.

RESULTS

Construction of a Consensus Genetic Map

The consensus map of barley in this study consists of SNP markers with an average length of 125 nucleotides, and it was constructed by using 13 mapping populations and different Illumina platforms (9K Infinium iSelect high-density custom genotyping bead chip and Illumina BeadXpress Array; Supplemental Table S1). The resulting consensus map holds a total of 6405 markers in 1978 unique positions (bins) (Table 1). The total length of the genetic map is 1120.27 cM, providing a density of theoretically one marker per 0.17 cM and one marker bin every 0.57 cM. Chromosomes 2H and 5H had the highest number of markers and bins and their genetic maps were also the largest in size, whereas chromosomes 1H, 4H, and 6H were the smallest, containing the lowest number of markers and bins (Table 1). Ordering conflicts among the set of linkage maps ranged from zero for chromosome 6H to 18 for chromosome 2H (Table 1). The map displays two gaps of 7.46 and 6 cM in the long arms of chromosomes 2H and 6H, respectively, with the remaining gaps being smaller than 5 cM (Supplemental Table S1; Supplemental Fig. S1).

Comparative Analysis to the Barley Genome Zipper and Population Sequencing at the Genome-Wide Scale

A map comprising 2785 BOPA markers constituted the genome-wide framework along which the barley genes were ordered and positioned for each individual barley chromosome in the GZ (Mayer et al., 2011). A proportion of these BOPA markers is also represented in the iSelect 9K chip and is segregating in the consensus map. First comparison, which also served as a positive control, consisted in corroborating the positions of commonly positioned BOPA markers between both datasets. Entirely, 2447 markers were in common between the consensus map and the map of Close et al. (2009), ranging from 280 markers on chromosome 1H to 450 on chromosome 5H (Supplemental Fig. S2). The agreement in the order of shared BOPA markers between the two maps at the individual chromosome level varied from 0.993 (4H) to 0.999 (2H, 7H), as measured by Spearman’s rank correlation coefficient. On average, the agreement in the intra-chromosomal location for the BOPA markers was 99.6% (Supplemental Fig. S2).

The marker order in the consensus map was employed to evaluate the collinearity of predicted genes in the barley zippers and in their sequence counterparts in the POPSEQ-anchored data. Based on the sequence homology, 689 markers (10.76%) from the consensus map did not find a counterpart, neither in the virtual map nor in the WGS-anchored map (Table 2). Additionally, 1337 (20.9%) markers did not match to any loci inferred by the GZ. Out of 6405 markers, 4379 (68.4%) were represented in the GZ (Table 2). A high percentage (97.8%) of the marker loci is located on the same chromosome in both datasets, and only 2.2% of markers showed a hit to an erroneous chromosome (Table 2). Detailed analysis considering the number of markers per chromosome revealed that chromosome 2H showed the highest amount of disagreements regarding chromosome assignments (3.2%), while chromosome 5H displayed the lowest level of misaligned markers (1.6%). The genetic coordinates of markers in the consensus map and the inferred loci in zippers was consistent at the genome-wide scale, although the resolution of comparison was lower in genetic pericentric regions, were a large amount of genetic markers were cosegregating (Fig. 1). The percentage of agreement in marker position was high (96.24% on average), varying from 94.19 (7H) to 97.63% (4H) (Table 2).

The comparison of the consensus map to POPSEQ revealed that 941 (14.7%) loci did not show any hit to a WGS contig. The remaining markers displayed a match to a POPSEQ contig, 99.30% having congruent chromosome positions (Table 2). Screening of a 5 cM window within both maps indicated that only 0.7% of positions on average disagreed in cM coordinates between the consensus and the POPSEQ, varying from chromosome 5H (0.43%) to chromosome 2H (1.19%) (Fig. 1).

The number of misaligned markers on each chromosome was plotted against their genetic coordinates to evaluate the putative existence of specific regions holding higher frequency of misallocated markers (Supplemental Fig. S3). Misaligned markers between the consensus map and GZ or POPSEQ were evenly distributed along each barley chromosome and a specific pattern of misallocation

<table>
<thead>
<tr>
<th>Chromosome</th>
<th>Map length (cM)</th>
<th>No. markers</th>
<th>No. bins</th>
<th>No. conflicts</th>
</tr>
</thead>
<tbody>
<tr>
<td>1H</td>
<td>146.30</td>
<td>655</td>
<td>226</td>
<td>4</td>
</tr>
<tr>
<td>2H</td>
<td>183.54</td>
<td>1116</td>
<td>357</td>
<td>18</td>
</tr>
<tr>
<td>3H</td>
<td>168.25</td>
<td>1036</td>
<td>355</td>
<td>2</td>
</tr>
<tr>
<td>4H</td>
<td>131.11</td>
<td>671</td>
<td>163</td>
<td>3</td>
</tr>
<tr>
<td>5H</td>
<td>191.15</td>
<td>1245</td>
<td>400</td>
<td>5</td>
</tr>
<tr>
<td>6H</td>
<td>137.68</td>
<td>811</td>
<td>235</td>
<td>0</td>
</tr>
<tr>
<td>7H</td>
<td>162.24</td>
<td>871</td>
<td>242</td>
<td>4</td>
</tr>
<tr>
<td>Total</td>
<td>1120.27</td>
<td>6405</td>
<td>1978</td>
<td>36</td>
</tr>
</tbody>
</table>

Table 1. Statistics of the consensus map.
Seven loci (L1–L7) located on barley chromosomes 1H, 2H, 4H, 6H, and 7H were employed to experimentally test the accuracy in virtually ordered genes of the barley GZ at a fine scale (Fig. 1). Loci L1 and L4 were genetically positioned previously in a high-resolution mapping population, while the others were mapped at a lower resolution. The sequences of flanking markers from all loci were used to survey the data on the barley GZ. The 14 markers matched the corresponding barley unigenes, spanning intervals in the GZ that varied in size from 0.90 (L5) to 7.49 cM (L3). The locus L2, at the centromeric region of chromosome 1H, did not display any interval, since the flanking markers match barley unigenes that are cosegregating (Table 3). The combination of intervals contained a total of 486 loci included in the GZ ranging from 30 (L6) to 198 (L2). Among these, only 62 loci (12.76%) corresponded to originally used BOPA markers, the other 424 were postulated according to the sequence homology to *B. distachyon*, rice, and sorghum genes. In total, 39.7% of targeted GZ loci possessed an orthologous gene in all three reference genomes, while the positions of 16.2, 10.1, and 7.3% of selected genes models was based on their homology to *B. distachyon*, sorghum or rice, respectively. Only one of the postulated barley loci was based on a full-length cDNA without any support from the above-mentioned model grass species (Supplemental Table S2). A set of 179 barley predicted genes, varying from 10 on chromosome 7HL (L7) to 66 on chromosome 1H (L2), putatively located within the target intervals were selected for further work (Table 3). Twenty-five of them (13.9%) did not generate any PCR amplicon, even when different primers pairs were tested at different positions in the gene. The majority of these were identified for the loci L1 (5.02%) and L3 (5.02%) on chromosomes 1HS and 2HL (Table 4). The remaining 154 loci were amplified and sequenced in the corresponding parental lines of the mapping population. Eighty-five markers (55.2%) turned out to be monomorphic. The lowest level of polymorphism was observed in the centromeric region of chromosome 1H, which showed 83.3% of monomorphic loci. On the contrary, those regions on the distal part of chromosomes 6HL and 7HS turned out to be highly polymorphic, the rate being 100% (Table 4). Among the remaining 69 polymorphic markers, seven were genotyped based on the presence or absence of the PCR amplicon in one of the parental lines, whereas 10 markers were mapped according to a size polymorphism. The other 47 loci containing SNPs were converted to CAPS markers (Table 4). Eleven polymorphic markers were identified in the L2 interval, but five of them were not mapped in the original work (Yang et al., 2013), therefore they were not considered in subsequent analyses.

In total, 8, 6, 13, 6, 11, 12, and 8 markers could be genetically mapped to chromosomes 1HS, 1H, 2HL, 4H, 6HL, 7HS, and 7HL, respectively, in the corresponding mapping populations yielding 64 new markers. Seven markers (10.9%) were located outside of the target intervals, mainly on chromosome 7HS (6.3%). Forty-eight (84.2%) markers out of 57 mapped in good collinearity with their estimated positions in the barley zippers (Table 5; Fig. 2), especially on chromosomes 4H and 6HL, where 100% of zipper-based markers are located in the same position as those predicted in the putative barley gene indexes (Table 5; Fig. 2).

The sequences from 57 consensus markers within the selected intervals from all seven target loci were used to survey the POPSEQ (Mascher et al., 2013a), the
International Barley Genome Sequencing Consortium (2012), and the Muñoz-Amatriain et al. (2015) datasets to integrate the genetic maps and the physical map of barley. In total, 44 (77.2%) newly developed markers showed a match to POPSEQ-anchored WGS contigs. The highest number of hits was observed for locus L5 (6HL), where all newly developed markers found a counterpart in POPSEQ. The lowest number of matches (three out of six) was detected for the centromeric region of the chromosome 4H (Table 5; Supplemental Table S3). On average, the percentage of collinearity between new zipper markers and contigs in POPSEQ compared with our consensus map was of 93.2% in the combination of chromosomes (Table 5). The comparison with POPSEQ
allowed verifying the correct ordering for the majority of GZ-based markers in their corresponding genetic map, but also the identification of markers with inconsistent location. Thus, five out of nine inaccurately predicted loci on chromosomes 1H and 7H were mapped at a right position according to POPSEQ (data not shown).

The comparison of the evaluated intervals against the genetically anchored physical map of the International Barley Genome Sequencing Consortium (2012) and the recently 15,622 sequenced barley clones from Muñoz-Amatriaín et al. (2015) allowed identifying the FPCs associated to the GZ-based markers and therefore to delimit the physical regions underlying the seven target loci (Table 5; Fig. 2). In total, 44 markers (77.2%) showed a hit to a FPC allowing the identification of 25 FPCs; two for L6 (7HS) and L7 (7HL); three for L2 (1H); four for L1 (1HS), L4 (4H), L5 (6HL); and six for L3 (2HL) (Table 5; Fig. 2).

**DISCUSSION**

Despite the difficulties resulting from the large size and complexity of the barley genome, relevant achievements in barley genomics have been accomplished during the last years. GZ, POPSEQ, and the resources established in the framework of the IBSC, including various high-density marker maps, have accelerated the exploitation of this Triticeae crop (Graner et al., 2011; Feuillet et al., 2012). However, these resources need further validation to become beneficial for plant breeders, who demand tools that allow not only the expeditious improvement of molecular markers tightly linked to genes or QTL of interest but also the usefulness of these markers across different germplasm resources (Varshney et al., 2006; Kilian and Graner, 2012; Keilwagen et al., 2014).

The construction of robust and highly resolved consensus linkage maps derived from experimental mapping data has been a long-standing challenge in barley genetics. Integrated genetic maps represent a more reliable resource for genetic anchoring of contig-based local or genome-wide physical maps and allow the orientation of scaffolds in genome assemblies (Paux et al., 2008; Alsop et al., 2011). Additionally, the accuracy and density of markers in a consensus map serve as valuable features toward the assessment of newly developed barley genomic resources. Within the barley research community, two integrated consensus maps have been recently published (Muñoz-Amatriaín et al., 2011; 2014a). The consensus map constructed in the present study was intended to be an improved version of those reported above by incorporating additional informative recombination events derived from the mapping population MBR1012 × Scarlett. The inclusion of this population had a relevant impact on the consensus map resolution. Thus, the newly integrated linkage map consists of 6405 markers, which represents an increase of 740 markers over the map developed by Muñoz-Amatriaín et al. (2014a). If only iSelect SNPs are considered, the MBR1012 × Scarlett linkage map contributed 1205 markers to the previous consensus map, representing an improvement of 2438 SNPs with respect to the Morex × ‘Barke’ map developed by Comadran et al. (2012). In general, a comparison of the consensus map generated in this report to the map by Muñoz-Amatriaín et al. (2014a) revealed a good consistence in the locus order, except for chromosome 2H, where a higher number of ordering conflicts were observed. Seventeen out of 18 markers in conflict derived from the MBR1012 × Scarlett population, although they affected intervals ≤2 cM. Considering the numbers of marker bins, the resolution of the current consensus linkage map was particularly improved for chromosomes 2H, 3H, and 5H, each one showing an increase of 5, 18, and 86 unique positions over the consensus map from Muñoz-Amatriaín et al. (2014a). Dense genetic maps will be also valuable for applied barley breeding, to perform precise introgression of improved traits in elite cultivars, as well as for accurate association mapping studies and genomic selection approaches (Heffner et al., 2009; Lorenz et al., 2012).

---

**Table 3. Description of the seven loci employed for the microsyntenic comparisons and number of genome zipper (GZ) loci present in the target intervals.**

<table>
<thead>
<tr>
<th>Locus name</th>
<th>Chromosome type</th>
<th>Population type</th>
<th>Interval in GZ (cM)</th>
<th>No. BOPA markers</th>
<th>No. inferred loci</th>
<th>No. BOPA markers</th>
<th>No. inferred loci</th>
</tr>
</thead>
<tbody>
<tr>
<td>L1</td>
<td>1HS</td>
<td>HR</td>
<td>1.51</td>
<td>10</td>
<td>23</td>
<td>5</td>
<td>19</td>
</tr>
<tr>
<td>L2</td>
<td>1H</td>
<td>LR</td>
<td>0.00</td>
<td>10</td>
<td>188</td>
<td>3</td>
<td>63</td>
</tr>
<tr>
<td>L3</td>
<td>2HL</td>
<td>LR</td>
<td>7.49</td>
<td>20</td>
<td>85</td>
<td>4</td>
<td>22</td>
</tr>
<tr>
<td>L4</td>
<td>4H</td>
<td>HR</td>
<td>1.38</td>
<td>9</td>
<td>30</td>
<td>2</td>
<td>24</td>
</tr>
<tr>
<td>L5</td>
<td>6HL</td>
<td>LR</td>
<td>0.90</td>
<td>4</td>
<td>43</td>
<td>1</td>
<td>11</td>
</tr>
<tr>
<td>L6</td>
<td>7HS</td>
<td>LR</td>
<td>0.64</td>
<td>3</td>
<td>27</td>
<td>1</td>
<td>14</td>
</tr>
<tr>
<td>L7</td>
<td>7HL</td>
<td>LR</td>
<td>4.30</td>
<td>6</td>
<td>28</td>
<td>2</td>
<td>8</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td></td>
<td>62</td>
<td>424</td>
<td>18</td>
<td>161</td>
</tr>
</tbody>
</table>

1 HR, (high resolution) or LR (low resolution) mapping population.

2 Number of GZ loci where position coincides with a barley oligo pool assay (BOPA) marker.

3 Number of GZ loci where position was inferred from their synteny to reference genomes.

---

**Table 4. Description of the genome zipper-based markers developed for the seven target intervals.**

<table>
<thead>
<tr>
<th>Locus name</th>
<th>L1</th>
<th>L2</th>
<th>L3</th>
<th>L4</th>
<th>L5</th>
<th>L6</th>
<th>L7</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chromosome</td>
<td>1HS</td>
<td>1H</td>
<td>2HL</td>
<td>4H</td>
<td>6HL</td>
<td>7HS</td>
<td>7HL</td>
<td></td>
</tr>
<tr>
<td>No. targeted gene models</td>
<td>24</td>
<td>66</td>
<td>26</td>
<td>26</td>
<td>12</td>
<td>15</td>
<td>10</td>
<td>179</td>
</tr>
<tr>
<td>No amplification</td>
<td>9</td>
<td>0</td>
<td>9</td>
<td>2</td>
<td>3</td>
<td>1</td>
<td>3</td>
<td>25</td>
</tr>
<tr>
<td>Monomorphic markers</td>
<td>7</td>
<td>55</td>
<td>4</td>
<td>18</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>85</td>
</tr>
<tr>
<td>No. developed markers</td>
<td>8</td>
<td>11</td>
<td>13</td>
<td>6</td>
<td>11</td>
<td>12</td>
<td>8</td>
<td>69</td>
</tr>
<tr>
<td>Presence or absence</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>3</td>
<td>2</td>
<td>0</td>
<td>7</td>
<td>0</td>
</tr>
<tr>
<td>Size polymorphism</td>
<td>1</td>
<td>2</td>
<td>4</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>10</td>
</tr>
<tr>
<td>CAPS</td>
<td>7</td>
<td>4</td>
<td>7</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>8</td>
<td>47</td>
</tr>
</tbody>
</table>

1 Five of these markers were not mapped previously (Yang et al., 2013), and, accordingly, they were not considered in the present work.

2 CAPS, cleaved amplified polymorphic sequence.
Table 5. Statistics of the comparisons of zipper-based markers to population sequencing (POPSEQ) and International Barley Genome Sequencing Consortium.

<table>
<thead>
<tr>
<th>Locus name</th>
<th>L1</th>
<th>L2</th>
<th>L3</th>
<th>L4</th>
<th>L5</th>
<th>L6</th>
<th>L7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chromosome</td>
<td>1HS</td>
<td>1H</td>
<td>2HL</td>
<td>4H</td>
<td>6HL</td>
<td>7HS</td>
<td>7HL</td>
</tr>
<tr>
<td>Markers outside of target region</td>
<td>0 (0%)</td>
<td>1 (17%)</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>4 (33%)</td>
<td>2 (25%)</td>
<td>7 (10.9%)</td>
</tr>
<tr>
<td>Markers within the target region</td>
<td>8 (100%)</td>
<td>5 (83%)</td>
<td>13 (100%)</td>
<td>6 (100%)</td>
<td>11 (100%)</td>
<td>8 (67%)</td>
<td>57 (89.1%)</td>
</tr>
<tr>
<td>Markers in collinearity with genome zipper</td>
<td>5 (62.5%)</td>
<td>4 (80%)</td>
<td>11 (84.6%)</td>
<td>6 (100%)</td>
<td>11 (100%)</td>
<td>6 (75%)</td>
<td>5 (83.3%)</td>
</tr>
<tr>
<td>Hits to POPSEQ</td>
<td>4 (50%)</td>
<td>4 (80%)</td>
<td>8 (61.5%)</td>
<td>3 (50%)</td>
<td>11 (100%)</td>
<td>8 (100%)</td>
<td>6 (100%)</td>
</tr>
<tr>
<td>Markers in correct collinearity with POPSEQ</td>
<td>4 (100%)</td>
<td>4 (100%)</td>
<td>7 (87.5%)</td>
<td>3 (100%)</td>
<td>11 (100%)</td>
<td>6 (75%)</td>
<td>41 (93.2%)</td>
</tr>
<tr>
<td>Hits to fingerprinted contigs</td>
<td>8 (100%)</td>
<td>3 (60%)</td>
<td>8 (61.5%)</td>
<td>4 (66.6%)</td>
<td>10 (90.9%)</td>
<td>8 (100%)</td>
<td>3 (50%)</td>
</tr>
<tr>
<td>No. identified fingerprinted contigs</td>
<td>4</td>
<td>3</td>
<td>6</td>
<td>4</td>
<td>4</td>
<td>2</td>
<td>25</td>
</tr>
</tbody>
</table>

Figure 2. (continued on next page) Comparison of collinearities among genome-zipper (GZ)-based markers, genome-zipper gene models, and WGS contigs derived from population sequencing at the seven target barley loci. The different colors in the physical map (International Barley Genome Sequencing Consortium [IBSC], 2012) represent the different barley chromosomes. The different shades of gray indicate the genetic position according to the consensus map of this study.
The newly developed consensus map was employed to investigate, in silico, the accuracy and robustness of the barley zippers at the genome-wide scale. A similar approach was performed previously by Poursarebani et al. (2013), but in that report only an individual map containing 1396 transcript derived markers was used for comparative purposes. On the contrary, the use of a high-density consensus map holding 6405 markers is a more appropriate framework for such comparison and validation. Indeed, a higher percentage (68.4%) of shared markers was observed in our work than Poursarebani et al. (2013), who found that only 37.8% of their genetic markers were represented among the GZ gene panels. Additionally, the percentage of markers and gene models that possess the same chromosomal location was also higher in the present study (97.8 vs. 95%). The average percentage of collinearity between both datasets, as measured by the Spearman's coefficient, was similar to that reported by Poursarebani et al. (2013) (96.2 and 96%, respectively). Such results support the greater suitability of consensus maps to validate grass-based comparative genome organization models in barley. Notwithstanding, as proposed by Poursarebani et al. (2013), the predicted chromosomal positions and virtual gene order postulated by genome zippers resulted to be highly precise (~96% accuracy) at the genome-wide level. As expected, the comparison of the consensus map to POPSEQ generated a higher number of matches than those obtained with the GZ, the increase being of 6.2% of the total markers. Population sequencing provides a linear order of WGS contigs genetically positioned along the seven individual barley chromosomes (Mascher et al., 2013a). The power of such methodology has been sufficiently demonstrated in previous reports by using...
different genotyping platforms and mapping populations (Mascher et al., 2013a; Ariyadasa et al., 2014; Chapman et al., 2015). Our results corroborated the robustness of POPSEQ, the genetic coordinates of contigs being coincident with markers positions in the integrated map at a 99.30% on average.

Although the performance of barley zippers at the genome-wide level appeared to be highly reliable, the development of any genomics-based breeding strategy requires the examination of the virtual gene order at a finer scale, when exploited syntenic relationships to *B. distachyon*, rice, and sorghum might be more influenced by misinterpretations (Li and Gill, 2002; Caldwell et al., 2004; Pourkheirandish et al., 2007). With these drawbacks in mind, the predicted linear gene index was investigated with experimental evidences at a low and high genetic resolution level, so called microcollinearity (Keller and Feuillet, 2000). Original data were compiled from earlier studies covering a set of seven loci conferring resistance to various fungal and viral diseases and traced to distinct barley chromosomes (Lüpken et al., 2013; Perovic et al., 2013a; Yang et al., 2013; Silvar et al., 2010, 2012, 2013). A similar procedure was performed by Poursarebani et al. (2013), but in that report, a section in the long arm of chromosome 2H was randomly selected for comparative purposes. On the contrary, the screening of several chromosomal regions spanning resistance genes or QTLs, performed in the present study, will guarantee a more reliable representation of the GZ power at the one-to-one relationship among orthologous genes. Disease resistance loci are particularly unstable, tend to be located in less conserved regions, and are commonly affected by structural variation (Leister et al., 1998; Meyers et al., 2003; Wicker et al., 2009). In total, 179 out of 486 GZ genes were considered for further development of tightly linked markers. Entirely, 13.9% of the loci could not be amplified on parental lines from different mapping populations despite the fact that various primer sets were designed at divergent positions. Such results support previous data by Silvar et al. (2013), who suggested apparent modifications in the gene-space of landrace-derived lines compared with the genome sequences of modern barley cultivars. Additionally, sequencing errors in the reads generated by 454-pyrosequencing could not be discarded as the source of absence of PCR products. Out of 154 primer pairs tested for polymorphism on the five examined chromosomes, 85 (55.2%) generated a monomorphic PCR amplicon. This was due to the presence of two centromeric loci on chromosomes 1H (L2) and 4H (L4), which contributed 85.8% of nonpolymorphic markers. Loci in centromeric and pericentromeric regions are on average less polymorphic than loci located on the rest of the barley chromosomes (Dvorák et al., 1998) as they are commonly organized in haplotype blocks locked into recombination-inert genomic regions (Thiel et al., 2009; Comadran et al., 2010). If these loci were removed, the rate of polymorphisms increases up to 81.3%, which is similar to that found in other reports based on barley expressed sequence tags or unigenes (Liu et al., 2010; Silvar et al., 2012). All 64 newly developed markers were accurately assigned to the corresponding chromosome. However, the genetic maps obtained with those markers were not in complete accordance with the putative linear gene order described in the GZ. Thus, 10.9% were located outside of the initial target intervals defined by the virtual gene index and 15.8% were not genetically mapped according to the gene order expected from the barley zippers. Such absence of collinearity might be attributed to insertions or inversions hypothesized in the virtually ordered gene inventory but not confirmed in our results, as suggested by Silvar et al. (2013). This outcome supports the well-documented existence of multiple inter- or intra-chromosomal rearrangements throughout the evolution among grass genomes (Bossolini et al., 2007; Salse et al., 2008; Bolot et al., 2009; Wicker et al., 2011). Variations in collinearity can also be explained by some gene models that are supported only by their counterpart in one or two reference genomes (Mayer et al., 2011). However, as suggested by Poursarebani et al. (2013), a general rule asserting that more than one model genome will increase the accuracy of the genome zipper should not be established, since a few markers based on single genomes also showed correct collinear positions. Nevertheless, the integrative use of distinctive model genomes might serve to overcome limitations imposed by species-specific regional variations (Mayer et al., 2011). Various reports also demonstrated that collinearity is commonly less conserved at the most distal telomeric regions of chromosomes (Li and Gill, 2002; Caldwell et al., 2004). This was actually the case of target loci on chromosomes 1HS, 2HL, and 7HS (see Fig. 2), but not for that on chromosome 6HL, which showed a 100% of agreement of the order for the zipper-derived markers mapped to the interval of interest. Beyond this, a good performance was observed for the barley zippers, allowing the development of 64 new markers and their mapping with an accuracy of almost 85%. From a breeding point of view, the quality of the order prediction permitted a more precise dissection of the regions containing interesting traits located on five different barley chromosomes.

These new zipper-based markers were employed further for comparisons to the POPSEQ (Mascher et al., 2013a) and barley physical map (International Barley Genome Sequencing Consortium, 2012; Muñoz-Amatriain et al., 2015) to circumscribe the regions in the barley genome conferring resistances. Based on sequence homology, 61.40% of the total markers found a hit to a Morex contig in POPSEQ. The genetic order of these loci in their respective linkage maps coincide with the position of the POPSEQ contigs to a 100%. This output served to verify that 55.5% of the target gene models previously cataloged with erroneous positions according to the GZ, were correctly mapped and allowed to partially resolve in silico blocks of cosegregating markers, which were assigned to distinct positions in the linearly ordered index of WGS contigs. These aspects pave the way toward the valuable
use of POPSEQ in breeding. This tool should be more amenable than the barley zippers for fine-mapping and cloning of agronomically important genes, provided that genetic markers sufficiently close to the loci of interest and adequate resolution in the mapping population are available. The outcome arising from POPSEQ could be anchored in a straightforward manner to the barley physical map, accelerating the identification of BAC contigs and subsequent isolation of genes underlying phenotypic traits (Ariyadasa et al., 2014). Such an approach will be more efficient than ordinary anchoring strategies relying on sequence comparison of flanking markers derived from sequence tags (Mascher and Stein, 2014).

The new markers developed from the barley zippers were also employed to demarcate, based on sequence homology, the physical regions in the barley genome responsible for the traits of interest. Between two and six contigs were identified by sequence comparison for the seven barley loci. Those loci derived from mapping populations with higher resolution permitted the definition of tiling paths holding a lower number of FPCs (data not shown) as long as the target loci are not allocated to centromeric regions (Lüpken et al., 2014). Identified contigs were employed to make a shallow approximation to the lately anchored physical map of barley and the established MTP containing 66,772 overlapping clones (Ariyadasa et al., 2014). Discordant contig placements were only observed on chromosome 7HS and might be explained by the technical and biological inaccuracy inherent to the construction of any genetic map (Wenzl et al., 2006; Wu et al., 2008). In addition, the short arm of chromosome 7H has been described as a “hot spot” of recombination, which might also contribute to the order controversy (Drader et al., 2009). The precise exploration of those disagreements should be carefully considered on the way to positional isolation of resistance genes (Liu et al., 2014). Even though the rough identification of the physical contigs provided extended information about the genomic context and local neighborhood underlying the traits of interest, the low genetic resolution of the majority of assayed mapping populations did not encourage us to speculate about the number or nature of putative candidate genes lying within the delimited genomic areas. That should be the principal task of further projects aimed to the map-based cloning of genes, as it was the case of rym11 (Yang et al., 2014).

The present work clearly demonstrated that recently established barley genomic resources could be efficiently exploited for breeding purposes. In spite of the appearance of few discrepancies, such as zipper-based markers outside the target intervals or erroneously positioned, our data elucidates that GZ and POPSEQ are very powerful tools for marker saturation, chromosome dissection, and physical-map anchoring at high and low resolution, provided that both approaches will also meet some limitations depending on the target chromosomal region. This could be the case for those loci allocated in the proximity of centromeres, which usually show low recombination (International Barley Genome Sequencing Consortium, 2012; Muñoz-Amatriain et al., 2015). Those strategies might be also employed with success in other complex genomic contexts, such as wheat (Triticum aestivum L.) (http://wheat-urgi.versailles.inra.fr) or unsequenced orphan crops of economic importance, like rye (Secale cereale L.) or perennial ryegrass (Lolium perenne L.) (Pfeifer et al., 2013; Martis et al., 2013). As demonstrated in the present work, the combined use of various genomic tools will help plant breeders and geneticists in different manners. First, it will permit the rapid development of markers tightly associated with the gene of interest, which might be further exploited or optimized for molecular MAS or even cataloged as functional markers (Andersen and Lübberstedt, 2003; Palloix and Ordon, 2011). Second, it will facilitate the disclosure of blocks of cosegregating markers, typically associated to low-resolution mapping populations, in a more efficient manner (Silvar et al., 2013). Finally, the combination of those different genomic resources should lead to a more straightforward and faster physical delimitation of promising regions in the barley genome, which constitute the starting point toward map-based cloning strategies (Lüpken et al., 2013; Yang et al., 2014). Some recently published bioinformatics tools, such as EnsemblPlants (http://www.ensemblgenomes.org), chromoWIZ (http://mips.helmholtz-muenchen.de/plant/chromowiz/index.jsp), or BarleyMap (http://floresta.eead.csic.es/barleymap) should ease the integration of different available genomics resources, allowing plant geneticist and breeders to manage these information in a time-saving manner (Kersey et al., 2014; Nussbaumer et al., 2014; Cantalapiedra et al., 2015). To our knowledge, this study is among the first efforts oriented toward the unification of various genomic resources with breeding purposes. Altogether, the fast enrichment of barley genome sequence information and novel techniques such as exome capture (Mascher et al., 2013b) should help to move barley breeding to an unprecedented level of precision and productivity, as foresee by Bevan and Waugh (2007).

Supplementary Material

Figure S1. Root mean square error (RMSE) values between the corresponding consensus map derived by LPmerge and the individual maps for different interval parameters of LPmerge. Among the different potential consensus maps for different interval values, we selected the one yielding the smallest median RMSE and highlighted the corresponding box in blue.

Figure S2. Dot plot comparison of shared BOPA markers between the consensus map and the map developed by Close et al. (2009) and employed as the framework for barley ‘zippers’ anchoring. Numbers at the top of the graphics show the Spearman’s rank correlation coefficient. Numbers at the bottom right hand corner indicate the amount of common markers between both datasets.
Figure S3. Misaligned markers were plotted along the consensus map when the genetic positions were discordant between Consensus/GenomeZipper and Consensus/POPSEQ map.

Table S1. Consensus genetic map

Table S2. Number of GenomeZipper loci comprised in the seven target intervals and their orthology to any of three different reference genomes

Table S3. Markers mapping to the physical contigs of the 15,622 sequenced clones (Muñoz-Amatriain et al., 2015) and information of the respective physical contigs of the International Barley Genome Sequencing Consortium (2012).

Acknowledgments
This work was supported by the German Federal Ministry of Education and Research under the grant number AZ O315702 and by the Spanish Ministry for Science under the grant number EU12009-04075. The authors like to thank Nils Stein and Ping Yang for fruitful discussion about the rym7 locus. CS was supported by a mobility fellowship from Universidade da Coruña.

References
Altschul, S.F., T.L. Madden, A.A. Schaffer, J. Zhang, Z. Zhang, W. Miller, Alsop, A., P. Farre, J.M. Wenzl, M. Wang, X. Zhou, I. Romagosa, A. Kilian, about the rym7 locus. CS was supported by a mobility fellowship from Universidade da Coruña.


Bockelman, H.E., and J. Valkoun. 2011. Barley germplasm conservation and research under the grant number AZ 0315702 and by the Spanish Ministry for Science under the grant number EU12009-04075. The authors like to thank Nils Stein and Ping Yang for fruitful discussion about the rym7 locus. CS was supported by a mobility fellowship from Universidade da Coruña.


