

Comparative genetic analysis of a wheat seed dormancy QTL with rice and *Brachypodium* identifies candidate genes for ABA perception and calcium signaling

Suthasinee Somyong · Jesse D. Munkvold ·
James Tanaka · David Benscher · Mark E. Sorrells

Received: 18 October 2010 / Revised: 5 March 2011 / Accepted: 12 March 2011 / Published online: 6 April 2011
© Springer-Verlag 2011

Abstract Wheat preharvest sprouting (PHS) occurs when seed germinates on the plant before harvest resulting in reduced grain quality. In wheat, PHS susceptibility is correlated with low levels of seed dormancy. A previous mapping of quantitative trait loci (QTL) revealed a major PHS/seed dormancy QTL, *QPhs.cnl-2B.1*, located on wheat chromosome 2B. A comparative genetic study with the related grass species rice (*Oryza sativa* L.) and *Brachypodium distachyon* at the homologous region to the *QPhs.cnl-2B.1* interval was used to identify the candidate genes for marker development and subsequent fine mapping. Expressed sequence tags and a comparative mapping were used to design 278 primer pairs, of which 22 produced polymorphic amplicons that mapped to the group 2 chromosomes. Fourteen mapped to chromosome 2B, and ten were located in the QTL interval. A comparative analysis revealed good macrocollinearity between the PHS

interval and 3 million base pair (mb) region on rice chromosomes 7 and 3, and a 2.7-mb region on *Brachypodium* Bd1. The comparative intervals in rice were found to contain three previously identified rice seed dormancy QTL. Further analyses of the interval in rice identified genes that are known to play a role in seed dormancy, including a homologue for the putative Arabidopsis ABA receptor ABAR/GUN5. Additional candidate genes involved in calcium signaling were identified and were placed in a functional protein association network that includes additional proteins critical for ABA signaling and germination. This study provides promising candidate genes for seed dormancy in both wheat and rice as well as excellent molecular markers for further comparative and fine mapping.

Keywords Preharvest sprouting · Comparative genomics · ABA · Calcium signaling · Marker development

Electronic supplementary material The online version of this article (doi:10.1007/s10142-011-0219-2) contains supplementary material, which is available to authorized users.

S. Somyong
National Center for Genetic Engineering and
Biotechnology (BIOTEC),
113 Thailand Science Park,
Phaholyothin Road, Klong 1, Klong Luang,
Pathumthani 12120, Thailand

J. D. Munkvold
Boyce Thompson Institute for Plant Research,
Tower Road,
Ithaca, NY 14853, USA

J. Tanaka · D. Benscher · M. E. Sorrells (✉)
Department of Plant Breeding and Genetics, Cornell University,
240 Emerson Hall,
Ithaca, NY 14853, USA
e-mail: mes12@cornell.edu

Introduction

Comparative mapping has become a tool for the comparison of gene order and content across related grass species such as rice, maize, wheat, barley, sorghum, and millets. Comparative mapping can be defined as the comparison of homologous gene content and order in a specific genomic interval between two or more species. A cross mapping of gene sequences by using molecular markers is the first step in studying genome relationships and is useful for identifying candidate genes controlling a trait of interest and for developing markers used in fine mapping. Even though the level of macrocollinearity is conserved among related species, microcollinearity may vary depending on the region (Devos and Gale 2000). A comparative analysis of

4,485 mapped wheat unigenes with the rice genomic sequence revealed numerous chromosomal rearrangements and discontinuities in gene order between these two species (Sorrells et al. 2003). Liu et al. (2006) reported a complexity in the microcollinearity due to inversions and insertions/deletions. However, there have been several successful studies that used orthologous regions from divergent grass species for fine mapping and identification of candidate genes. This approach was used to identify a gene controlling vernalization requirement in winter wheat. The wheat (*Triticum monococcum*) chromosome 5A region containing the vernalization gene, *VRN1*, is collinear with rice chromosome 3 and sorghum BACs (Yan et al. 2003). A perfect microcollinearity was reported for a distance of 2.6 cM on wheat chromosome 6B and rice chromosome 2, at the grain protein locus *Gpc-6B1*. The new markers based on the rice sequence were used to narrow the interval to 0.3 cM and identify five candidate genes within 64 kilobase pairs (kb), based on the rice sequence (Distelfeld et al. 2004). In another example, *Rht1*, which is a gibberellin-insensitive gene that controls plant height and is located on wheat chromosomes 4B and 4D, was found to be orthologous with maize *d8* on chromosome 1 (Peng et al. 1999). Collinearity was also reported for wheat chromosome 3BS, rice chromosome 1S, and barley chromosome 3HS, at the Fusarium head blight locus (*Fhb1*) (Liu et al. 2006). Some studies have shown that certain wheat regions were not collinear with rice, sorghum, and maize but were collinear with barley and rye. The region at shrunken 2 (*Sh2*) and the anthocyaninless1 (*Al*) regions were collinear among rice, sorghum, and maize but not between wheat and barley (Li and Gill 2002).

Wheat preharvest sprouting (PHS) is the germination of grain on the plant before harvesting, resulting in a reduced grain quality and harvestable yield. Preharvest sprouting and seed dormancy (SD) in wheat are expressed as quantitatively inherited traits that are strongly influenced by the environment. Quantitative trait locus (QTL) studies are useful to identify the genome regions that control PHS and SD and their relative importance for further studies such as fine mapping. Fine mapping can narrow down the specific QTL regions, perhaps to the gene level. In a previous work done by Munkvold et al. (2009), PHS QTLs were mapped in a doubled haploid population of soft white winter wheat from a cross between the cultivars Cayuga (PHS-resistant wheat) and Caledonia (PHS-susceptible). The major PHS QTL, *QPhs.cnl-2B.1*, located on chromosome 2B was significant in all 16 environments and explained 5% to 31% of the phenotypic variation. The location of the QTL mapping interval varied somewhat in different environments; however, in all the environments tested, the QTL confidence intervals overlapped. Based on the PHS mean from all environments, the QTL peak was

close to the marker Barc55 and was flanked by markers Gwm429 and Wmc474. The additional QTLs detected on chromosomes 2D, 3D, and 6D were found to be significant in seven, four, and ten out of 16 environments, respectively. The PHS QTL on 2B and 2D may be homoeologous, but more mapping is required for confirmation. In addition, an SD QTL, *QDor.cnl-2B.1*, was located in the same interval as the *QPhs.cnl-2B.1*, which suggests that the variation in PHS contributed by this locus is due to varying levels of seed dormancy (Munkvold et al. 2009). Fine mapping would help narrow down the region containing this QTL and delimit the gene(s) contributing to both PHS and SD.

The objectives of this study were to: (1) develop new markers designed from wheat expressed sequence tags (ESTs), EST contigs, and transcripts in the *QPhs.cnl-2B.1* QTL interval on 2B and (2) use comparative mapping with rice, and *Brachypodium* as a tool to identify candidate genes located in the orthologous region necessary for marker development and fine mapping. By identifying and analyzing the comparative regions between wheat and rice, it was possible to develop improved molecular markers for marker-assisted selection and fine mapping. Additionally, we were able to identify candidate genes with previously characterized roles in ABA perception and calcium signaling, both of which are important biological processes for the establishment and maintenance of seed dormancy.

Materials and methods

Data resources

The *QPhs.cnl-2B.1* QTL interval covering the three simple sequence repeat (SSR) markers, Gwm429, Barc55, and Wmc474, was the target for comparative mapping. The genetic distance of the interval was previously found to be 8.3 cM in a doubled haploid population (Munkvold et al. 2009). However, the physical/genetic relationship between the Gwm429 to Wmc474 marker interval is unknown. Based on the location of the SSR markers, a single deletion bin, 2BS-0.53-0.75, covered the *QPhs.cnl-2B.1* interval. The location and primer information for marker Barc55 was obtained from GrainGenes, a database for Triticeae and *Avena* (<http://wheat.pw.usda.gov/GG2/index.shtml>). Information regarding ESTs previously mapped to the deletion interval 2BS1-0.53-0.75 was obtained from GrainGenes wEST-SQL resources (http://wheat.pw.usda.gov/cgi-bin/westsqldb/map_locus.cgi). The comparisons of the homologous rice genes and wheat ESTs were obtained from the rice genome browser TIGR Release 5.0 (<http://www.modelcrop.org/cgi-bin/gbrowse/rice/>). The comparisons between wheat and *Brachypodium* were obtained from the

Brachy 8× release assembly (<http://www.brachypodium.org>). The rice QTLs that were homologous with the region of the wheat deletion interval 2BS1-0.53-0.75 can be found in Gramene (<http://www.gramene.org/>).

Marker development

For the development of gene-based PCR markers, primers were designed based on wheat EST sequences. To avoid designing primers on the splicing junction between an intron and an exon, the EST sequences were compared with the genomic sequences of the homologous rice genes. Information on all new markers can be found in Supplementary Table 1.

Designing polymorphic markers based on intron variation

Limited genomic sequence is currently available for wheat, but extensive information is available for ESTs, EST contigs, and transcripts. Wheat ESTs and EST contigs were used for marker design. EST contigs were provided by the gene index project, computational biology, and functional genomics laboratories (<http://compbio.dfci.harvard.edu>). The exon sequences were used to design primers for amplifying PCR products across the introns. Each EST and EST contig was compared with rice homologous genes whose sequences were provided by the TIGR rice genomic annotation project (<http://blast.jcvi.org>) and the Gramene database (<http://www.gramene.org/>). The alignment between sequences was compared by NCBI blast (<http://blast.ncbi.nlm.nih.gov>). Sequences were also compared directly by using the Vector NTI-Align X program (Lu and Moriyama 2004) to avoid the splicing junction. The primers were designed using Primer3 (<http://frodo.wi.mit.edu/primer3/>) to exclude putative splice junctions and span introns.

Designing primers based on variation in coding sequences

The variation between the two parents, Cayuga and Caledonia, can be located in introns or in coding regions and may include insertions or deletions (indels), and/or single nucleotide polymorphisms (SNPs). The nucleotide sequences were useful for designing new polymorphic markers, including a cleaved amplified polymorphic sequence (CAPS), EST–SSR, SNP, and indel markers. CAPS markers were designed using the Solanaceae Genome network tool (http://solgenomics.net/tools/caps_designer/caps_input.pl) and SSRs were identified using the Gramene Simple Sequence Repeat Identification Tool (<http://www.gramene.org/db/markers/ssrtool>). Polymorphic primers were first identified in white winter hexaploid wheat (*Triticum aestivum*) cultivars, Cayuga (PHS-resistant) and

Caledonia (PHS-susceptible) and mapped in the Cayuga × Caledonia doubled haploid population as described by Munkvold et al. (2009). PCR products from both parents were cloned into a pGEM-T Easy vector (Promega, Madison, WI) for sequencing. Because wheat contains three homoeologous copies of each gene, multiple clones were sequenced and compared using the Vector NTI-Align X program to assess variations between homoeologues. The 500–1,000 bp sequences were also used to design primers for amplifying short 150–300 bp sequences. The shorter sequences were compared by polyacrylamide gel electrophoresis (PAGE). If the PCR products were monomorphic on PAGE gels, then single-strand conformation polymorphism (SSCP) gels were used to compare products based on differences in mobility resulting from differences in the sequence that alter the secondary or tertiary structure of the DNA molecules (Sunnucks et al. 2000; Liu et al. 1999).

The sequences of select *Myb*-related genes were used to design SSR markers. The primer sequences for the *Myb*-related genes were designed from the sequences in the TIGR plant transcript assemblies (<http://plantta.jcvi.org/index.shtml>). For some genes, such as the calmodulin/calcium-dependent protein kinase (CDPK) (Li et al. 2008), the SSR identification tool (<http://www.gramene.org/db/markers/ssrtool>) and primer3 (<http://frodo.wi.mit.edu/primer3>) were used to design the primers. The *Myb*-related genes and CDPK genes were selected as seed dormancy candidate genes because of both molecular functions and the location of these genes in the homologous rice and *Brachypodium* regions.

Testing for polymorphic markers

Parental DNA was used for testing polymorphism, and the location of the polymorphic markers was determined using a doubled haploid (DH) population of 149 individuals from a cross between the cultivars Cayuga and Caledonia (Munkvold et al. 2009). The set of 149 individuals was used in our study because it was the first set of the DH population with sufficient marker data. The PCR annealing temperature was optimized for each primer, and the PCR products were either separated by size on a 4% PAGE gel or by mobility on a 10% SSCP gel. The locations of the new markers were determined by using the previous map (Munkvold et al. 2009) and MapManager QTXb20 software (Manly et al. 2001).

Identification of candidate genes and functional protein association network analysis

The comparative map between wheat 2B and rice chromosomes 3 and 7 was used to identify candidate genes for seed dormancy and PHS. All wheat ESTs located in the 2BS1-

0.53-0.75 deletion bin with a homologue in the comparative map were compared to the Arabidopsis protein sequences using BLASTX (<http://www.ncbi.nlm.nih.gov>). The best hit annotations were reviewed for known or putative roles in seed dormancy and related biological processes such as ABA and GA signaling. The homologous proteins from Arabidopsis were then used to build functional protein association networks using the STRING 8.3 database (<http://string-db.org>), where the edges were based on co-expression or experimental data with a summary string significance greater than 0.7. The network was first seeded using the Arabidopsis protein homologues for candidate genes identified from the wheat/rice comparative map. Proteins with a first-order connection to the seed proteins were added as input nodes and the network was updated to find additional connected proteins. The network was then visualized and edited using the program Medusa (<http://coot.embl.de/medusa>).

Results

Comparative analysis between wheat deletion bin 2BS1-0.53-0.75 and rice chromosomes 7 and 3

In order to delimit the wheat genomic region for comparative mapping, markers flanking the *QPhs.cnl-2B.1* QTL interval were used to identify the corresponding wheat deletion bin. The SSR marker Barc55, a marker centered in the major PHS resistance QTL interval *QPhs.cnl-2B.1* (Munkvold et al. 2009), was previously mapped to the 2BS1-0.53-0.75 deletion bin (Somers et al. 2004). For comparative analyses of the *QPhs.cnl-2B.1* interval on chromosome 2B, new markers were designed based on

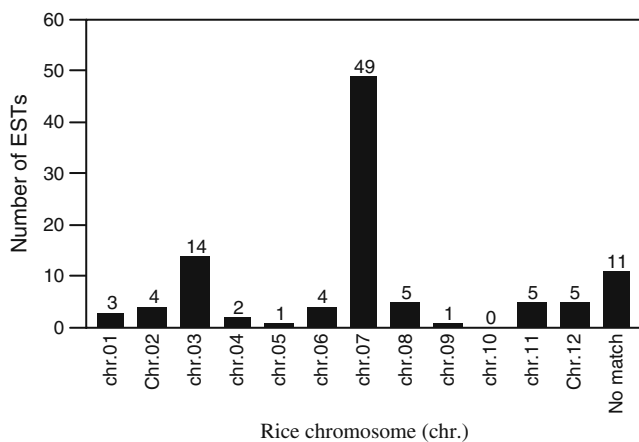
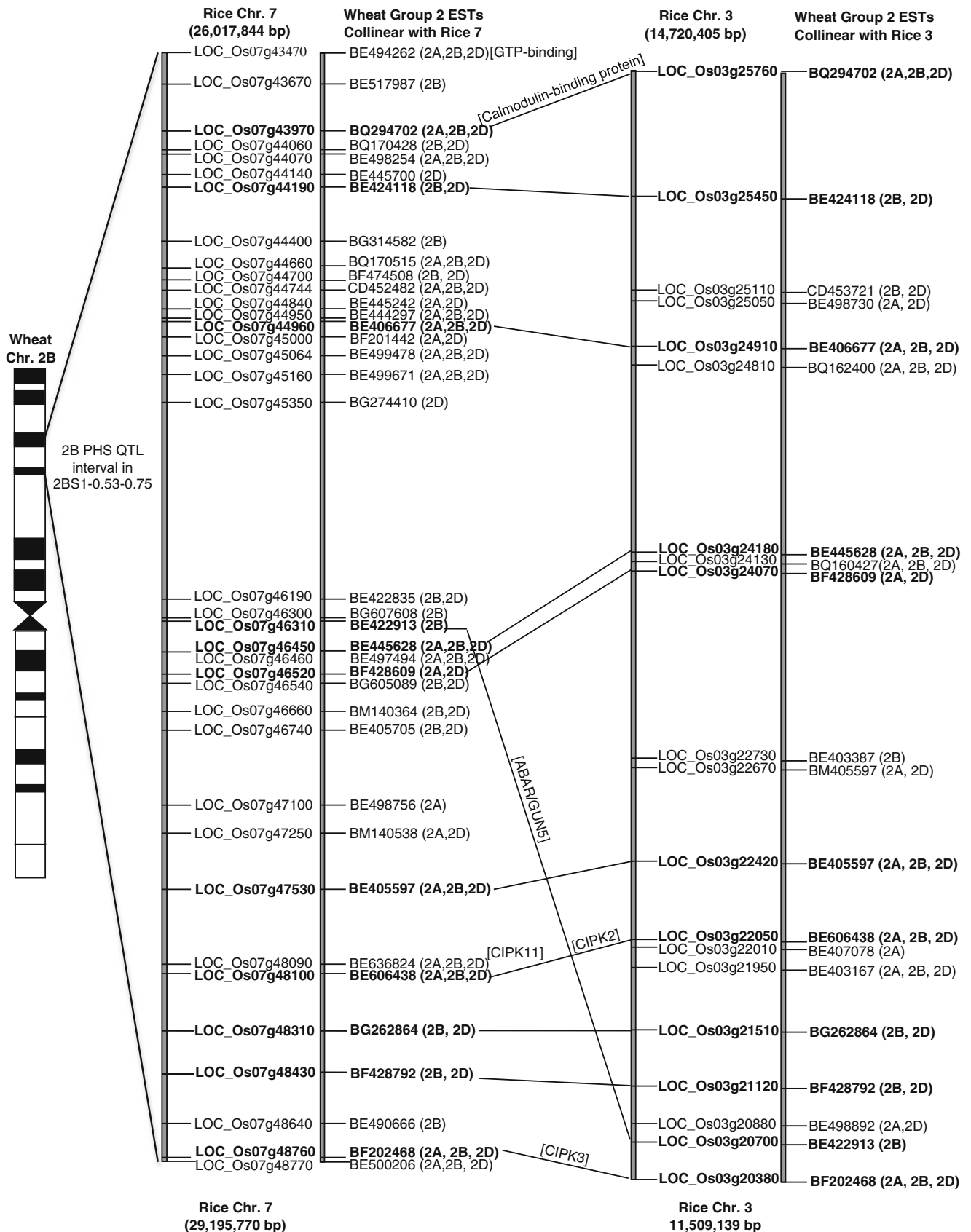


Fig. 1 The number of rice genes homologous to the wheat ESTs in deletion bin 2BS1-0.53-0.75 and their rice chromosome location. Some wheat ESTs were homologous to genes that were located on multiple rice chromosomes, but only the rice chromosome with the most similar homologue was tabulated

Fig. 2 A comparative map representing the collinear regions between wheat ESTs in the *QPhs.cnl-2B.1* interval with rice chromosomes 7 and 3. The comparative map covered the *QPhs.cnl-2B.1* interval between wheat ESTs BE494262 and BE500206. The homologous regions are the distal part of the rice chromosome 7 long arm (~3 mb) from position 26,017,844 bp to position 29,195,770 bp and the middle part of rice chromosome 3 (~3 mb), from position 11,509,139 bp to position 14,720,405 bp. A marker designed from BE494262 was located close to the distal flanking SSR marker Wmc474. A marker designed from BE500206 was located close to the proximal flanking SSR marker Gwm429. The rice gene loci on chromosomes 7 and 3 were placed in order based on the rice pseudomolecule sequence. The homologous wheat ESTs were placed according to the order of rice gene loci. This figure represents only collinearity between the rice genes and the wheat ESTs on wheat group 2 chromosomes (2A, 2B, and 2D). These genes were useful for saturating the PHS QTL region. The genes or ESTs in *bold* represent the collinear genes on both rice chromosomes 7 and 3. The chromosome locations of wheat ESTs are noted in parentheses. The annotation of candidate genes based on BLASTX comparison with Arabidopsis is denoted in *brackets*

ESTs, EST contigs, and known seed dormancy candidate genes previously mapped in the 2BS1-0.53-0.75 deletion interval. A total of 104 ESTs were previously mapped to this bin (<http://wheat.pw.usda.gov/wEST/binmaps/>), of which 93 (89%) showed high homology with rice genes located on all rice chromosomes except chromosome 10. Some ESTs were homologous with genes located on multiple rice chromosomes, but only the ESTs that gave the lowest *E*-value were included (Fig. 1). Rice chromosome 7 had the most genes ($n=49$) homologous to the wheat ESTs while chromosome 3 was second with 14 homologous genes. A comparative mapping of these ESTs was used to delimit the homologous rice region for the PHS QTL *QPhs.cnl-2B.1* on rice chromosomes 7 and 3. To construct the comparative map between the *QPhs.cnl-2B.1* interval and the homologous region in rice, new markers close to the flanking markers Wmc474 and Gwm429 were used to help define the distal and proximal regions (Fig. 2). The comparative mapping revealed macrocollinearity in the *QPhs.cnl-2B.1* interval with rice chromosomes 3 and 7 based on ten wheat ESTs with homologous matches to both regions. The wheat ESTs were ordered based on their putative orthologs on rice chromosomes because of the known macrocollinearity with rice and the unknown order of the wheat ESTs in the deletion bin. All ESTs were also located on one or both of the other group 2 wheat chromosomes. The *QPhs.cnl-2B.1* interval showed homology with the long arm of rice chromosome 7 from the distal position 26,017,844 bp to position 29,195,770 bp (~3.15 million base pair (mb)) and with rice chromosome 3 from a proximal position 11,509,139 bp to position 14,720,405 bp (~3.2 mb).

The region on rice chromosome 7 contained approximately 61 rice genes that matched wheat ESTs. Thirty-seven of these matching genes were previously mapped on wheat chromosomes 2A, 2B, and 2D. Thirty out of these 37



ESTs were mapped in the 2BS1-0.53-0.75 deletion bin. For the homologous region on rice chromosome 3, approximately 58 rice genes matched wheat ESTs. Twenty of these ESTs mapped to wheat group 2 chromosomes and 12 were mapped in the 2BS1-0.53-0.75 bin. Forty-seven percent of the ESTs in the comparative map were located on all three wheat group 2 chromosomes. Thirty-two percent of the ESTs in the comparative map were located on two of the three, group 2 wheat chromosomes. Twenty-one percent of the ESTs were located on only one group 2 chromosome (Sorrells et al. 2003).

The aim of the comparative mapping was to develop all possible markers from ESTs located in the wheat deletion bin 2BS1-0.53-0.75 containing the *QPhs.cnl-2B.1* interval and flanked by markers Wmc474 and Gwm429. A new polymorphic marker, CNL414-BE494262, was found to map near Wmc474. This EST was homologous with the rice gene LOC_Os07g43470, located on rice chromosome 7 but no homologous gene was found on rice chromosome 3. The wheat EST BQ294702 was homologous with both rice chromosome 7 (LOC_Os07g43970) and rice chromosome 3 (LOC_Os03g25760) and was used to design marker CNL413-BQ294702 that mapped at the distal end of the comparative map. On the proximal region of the map, the new polymorphic marker CNL406-BE405569 was mapped near Gwm429 but was found to have homology with a gene on rice chromosome 1 and not chromosomes 3 or 7. The nearby wheat EST BF202468, which was used to design marker, CNL407-BF202468, was located on both rice chromosomes 7 (LOC_Os07g48760) and 3 (LOC_Os03g20380) at the proximal end of the comparative map (Fig. 2). The markers CNL421-BG314234, CNL423-BG274905, and CNL424-BE488865 were monomorphic on chromosome 2B but polymorphic on chromosomes 2A and 2D (Supplementary Table 1).

The comparative map was then used as a tool to identify useful ESTs, transcripts, and candidate genes located in the rice homologous regions. In this step, polymorphic markers, CNL410-BE636824, CNL412-BF201533, and CNL413-BQ294702, were designed based on the variation in coding regions of 22 wheat EST contigs (Supplementary Table 1). In addition to ESTs and their contigs, the candidate genes located in the homologous rice regions were also used for marker development. Eleven genes were selected for mapping based on previous reports that they may play a role in seed dormancy in Arabidopsis. Primers for five genes had already been developed as described above; CNL407-BF202468 was homologous with CBL-interacting protein kinase 3 (CIPK3), CNL409-BE606438 with CBL-interacting serine/threonine kinase 2 (CIPK2), CNL414-BE494262 with a putative GTP binding protein (At4g39520), CNL410-BE636824 with CBL-interacting ser-

ine/threonine kinase 11 (CIPK11), and CNL413-BQ294702 with a calmodulin-binding protein (At3g58480). Additional primers were designed based on CDPK-related kinase 1 (CRK1) (CNL415 and CNL420), a Myb binding protein (CNL417 and CNL426) and a FAD-binding protein (CNL418). Five of these markers were polymorphic: CNL420 on 2A, CNL415, CNL417 and CNL418 on 2B, and CNL426 was unlinked (Supplementary Table 1)

Comparative analysis of rice seed dormancy QTL with the 2BS1-0.53-0.75 bin and the wheat PHS QTL interval

A comparative analysis of the rice SD QTL was used to identify the possible homologous wheat ESTs in both rice and wheat QTL regions. Previous QTL studies reported at least 12 SD QTLs on rice chromosome 7, some of which were probably the same in different reports (Lin et al. 1998; Ishimaru et al. 2001; Cai and Morishima 2000; Miura et al. 2002; Jiang et al. 2003; Gu et al. 2004, 2005a, b; Wan et al. 2005). Seven SD QTLs have been reported on rice chromosome 3 (Ishimaru et al. 2001; Cai and Morishima 2000, 2002; Jiang et al. 2003; Takeuchi et al. 2003) (Fig. 3). Based on the QTL data obtained from Gramene, the position of rice SD QTLs was compared with the annotated Nipponbare sequence 2006. The SD QTLs on rice chromosome 7 started at around position 2 mb and continued to 27 mb nearly spanning the entire rice chromosome. The SD QTLs on rice chromosome 3 started at position 8 mb and continued to 34 mb, again nearly covering the entire rice chromosome. Some QTL regions were small enough to allow identification of the nearest genes within that rice SD QTL interval providing a resource for marker development and comparative mapping. For example, an SD QTL, Gramene QTL accession ID AQCZ006, was located close to a calcium-dependent protein kinase isoform (AK1) gene reported to be involved in seed dormancy (Ok et al. 2005; Woodger et al. 2003; Sheen 1996). These genes may be used for marker development to confirm whether they collocate with the *QPhs.cnl-2B.1* interval. However, there are still numerous genes under those QTL intervals that may be potential candidate genes. Additional agronomic QTLs such as heading date, grain weight, abiotic stress, and biotic stress tolerances were also located on rice chromosome 3 while heading date, seed set percent, amylase content, and leaf nitrogen content were located on rice chromosome 7.

A comparative analysis of the location of rice SD QTLs indicated that there was only one rice SD QTL (Gramene QTL accession ID AQCZ009) on chromosome 7 (Jiang et al. 2003) that was located in the region homologous to the *QPhs.cnl-2B.1* interval in the Cayuga × Caledonia population. That rice map position was estimated to be around 24.29–27.16 mb. The other 11 rice SD QTLs on chromo-

some 7 were located outside the *QPhs.cnl-2B.1* interval. The closest QTLs were all located proximal to the rice centromere from 21 to 25.6 mb (Fig. 3) but in the distal region of wheat chromosome 2BS.

In addition, there were two rice SD QTLs on chromosome 3 that were predicted to lie within the *QPhs.cnl-2B.1* interval. The chromosome 3 regions were Gramene ID CQE17, from position 8.4–21.2 mb (Ishimaru et al. 2001) and AQCZ001, from position 10.35–11.53 mb (Jiang et al. 2003). Five SD QTLs were reported to be outside of the comparative mapping region. However, one of the rice QTL intervals is very large, indicating that the accuracy for that rice QTL interval was limited. In total, there were three rice SD QTLs predicted to lie within the regions homologous to the *QPhs.cnl-2B.1* interval.

Comparative analysis among wheat ESTs, rice, and *Brachypodium*

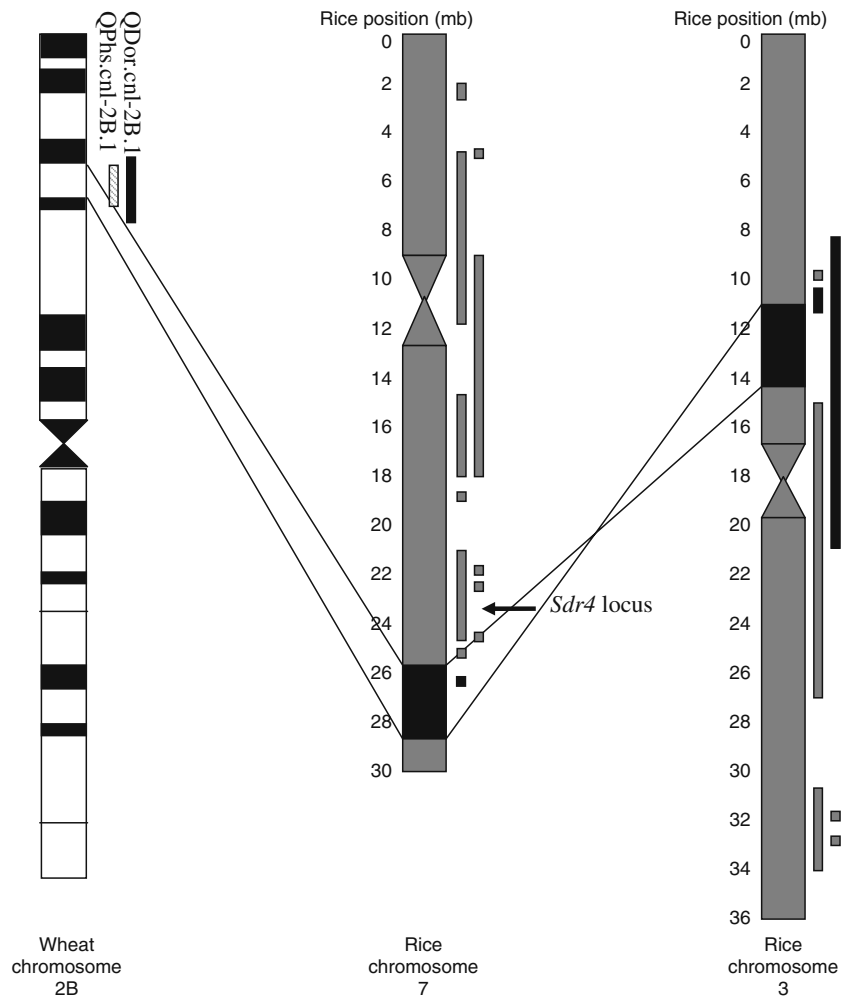
The delimited region on the wheat/rice comparative map was also homologous with *Brachypodium* Bd1. A total of

33 of the 37 wheat ESTs (89%) in the comparative map between wheat group 2 and rice 7 were found in the homologous interval on *Brachypodium* Bd1 (Fig. 4). The gene order in *Brachypodium* was the same as that of rice chromosome 7 and the *QPhs.cnl-2B.1* interval with the exception of two small inversions. Homologs of all six of the candidate genes annotated in Fig. 2 were also present in the homologous region of *Brachypodium* Bd1. Two of the candidate genes from wheat, BE636824 (CIPK11) and BE606438 (CIPK2), were involved in one of the small inversions compared to rice chromosome 7. The comparative analysis between rice, wheat, and *Brachypodium* helped to identify the candidate genes for further analysis.

Identification of seed dormancy candidate genes and functional protein association networks

In order to identify candidate genes for seed dormancy and PHS resistance, we used homologous proteins from *Arabidopsis* for wheat ESTs contained in the rice/wheat/*Brachypodium* comparative interval (Fig. 4). The wheat

Fig. 3 A comparative analysis between the *QPhs.cnl-2B.1* interval and the rice seed dormancy (SD) QTLs on chromosomes 7 and 3 based on Gramene information (<http://www.gramene.org/>). On wheat chromosome 2B, the striped bar represents the *QPhs.cnl-2B.1* interval, whereas the black bar represents the *QDor.cnl-2B.1* interval. On both rice chromosomes 7 and 3, the grey bar represents the rice SD QTL intervals outside the homologous region, whereas the black bars represent the rice SD QTL intervals within the homologous region. The connection lines localized the homologous regions. A gene at the *Sdr4* locus was the first isolated seed dormancy gene reported in cereals by map-based cloning (Sugimoto et al. 2010)



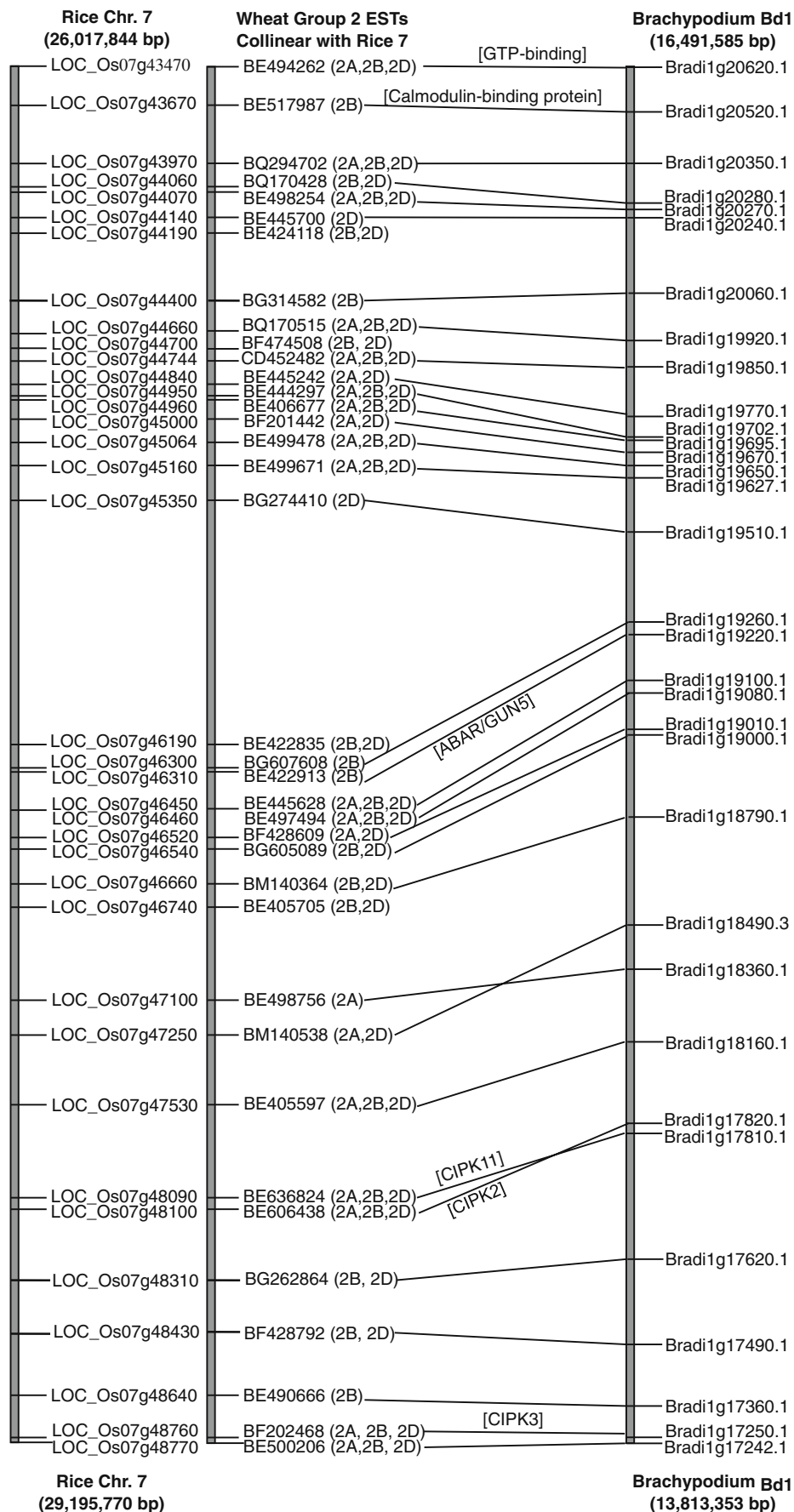


Fig. 4 A comparative map representing the collinear regions between wheat ESTs in the *QPhs.cnl-2B.1* interval with rice chromosome 7 and *Brachypodium*. The comparative map covered the *QPhs.cnl-2B.1* interval on wheat chromosome 2B between ESTs BE494262 and BE500206. The homologous regions are the distal part of rice chromosome 7 (~3 mb) from position 26,017,844 bp to position 29,195,770 bp and *Brachypodium* Bd1 (~2.7 mb), from position 13,813,353 bp to position 16,491,585 bp. A marker designed from BE494262 was located close to the distal flanking SSR marker Wmc474. A marker designed from BE500206 was located close to the proximal flanking SSR marker Gwm429. The rice gene loci on chromosome 7 and gene models of *Brachypodium* were placed in order based on the genome sequence information. The homologous wheat ESTs were placed according to the order of the rice gene loci. This figure represents only collinearity between the rice genes and the wheat ESTs on wheat group 2 chromosomes (2A, 2B, and 2D)

EST BE422913, located near the middle of the comparative map (Fig. 4), was found to be homologous with the Arabidopsis protein ABAR/GUN5 (At5g13630). The ABAR/GUN5 protein is known to play a role in seed dormancy in Arabidopsis and is a possible receptor for the hormone ABA (Shen et al. 2006). Additional candidate genes involved in calcium signaling were identified, including wheat homologues of CIPK2, CIPK3, and CIPK11 from Arabidopsis. The Arabidopsis CIPK3 protein was found to regulate ABA signaling, and CIPK3 mutants were hypersensitive to ABA with regard to seed germination (Kim et al. 2003).

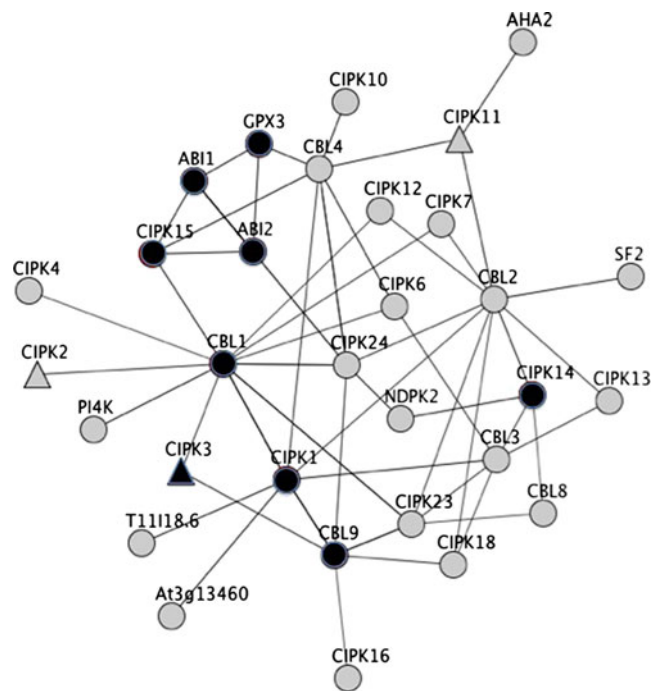
To further characterize the possible role of the CIPK candidates in seed dormancy, a functional protein annotation network was constructed using the STRING 8.3 database (<http://string-db.org>). The network was seeded with the three Arabidopsis proteins CIPK2, CIPK3, and

Fig. 5 A functional protein association network from Arabidopsis developed using the STRING 8.3 database (<http://string-db.org>). Connections are based on co-expression and experimental evidence with a STRING summary score above 0.7. The Arabidopsis homologous proteins for three wheat candidates BE606438 (CIPK2), BF202468 (CIPK3), and BE636824 (CIPK11) are designated by triangles. The nodes for proteins with known roles in ABA signal transduction and seed germination are highlighted in black

CIPK11. All first-order connections were added to the input nodes and the network was redrawn to include additional connected proteins (Fig. 5). The functional protein annotation network included 31 nodes and 62 edges. In addition to CIPK3, eight additional proteins with known roles in seed dormancy and germination were connected in the network, including ABI1, ABI2, CBL1, CBL9, CIPK1, CIPK14, CIPK15, and GPX3. The Arabidopsis protein phosphatase 2C proteins ABI1 and ABI2 are known to be negative regulators of ABA signaling in both seeds and other tissues (Meyer et al. 1994; Rodriguez et al. 1998). Arabidopsis mutants in the Ca²⁺-binding protein CBL1 were found to be hypersensitive to ABA in seed germination. The CBL1 protein was also determined to physically interact with both ABI1 and ABI2 (Guo et al. 2002). Arabidopsis *cbl9* and *cipk1* mutants were found to be impaired in ABA suppression of germination (Pandey et al. 2004; D’Angelo et al. 2006). In contrast, both *cipk14* and *cipk15* mutants were found to be hypersensitive to ABA in relation to seed germination (Guo et al. 2002; Qin et al. 2008). Glutathione peroxidase 3 was shown to interact with ABI1 and ABI2 and identified a possible connection between H₂O₂ and ABA signaling. Mutant *gpx3* plants were also found to be more sensitive to the mannitol suppression of germination (Miao et al. 2006).

Discussion

In this study, comparative mapping was used to identify homologous regions in the rice and *Brachypodium*



genomes for seed dormancy and PHS in wheat. The coding regions in those genomes were used to develop new markers for fine mapping of the *QPhs.cnl-2B.1* QTL in wheat. This approach revealed seed dormancy QTL in the corresponding rice genome region, thus providing additional evidence for the genetic control of this trait in wheat and rice.

Comparative analysis of rice genomic regions homologous to the PHS QTL interval in the wheat 2BS1-0.53-0.75 deletion bin

Comparative analyses between wheat, rice, and *Brachypodium* were used to develop new markers, locate rice SD QTL, identify candidate known function genes, and assess collinearity for the PHS QTL region, *QPhs.cnl-2B.1*. The *QPhs.cnl-2B.1* interval was determined to lie within the deletion bin 2BS1-0.53-0.75 and shares homology with the regions on rice chromosome 7, rice chromosome 3, and with *Brachypodium* Bd1. This 2BS1-0.53-0.75 bin was previously reported to be homologous to rice chromosome 7 (La Rota and Sorrells 2004), but these results suggest that it is also partially homologous to rice chromosome 3. Not all wheat ESTs located in the 2BS1-0.53-0.75 bin were located within the *QPhs.cnl-2B.1* interval flanked by markers Wmc474 and Gwm429. Approximately 70% of all wheat ESTs in the 2BS1-0.53-0.75 bin that were homologous to rice chromosome 7 sequences, were located in the *QPhs.cnl-2B.1* interval, whereas 30% were located distal to Wmc474 or proximal to Gwm429. Similarly, only 50% of all wheat ESTs in the 2BS1-0.53-0.75 bin that were homologous to rice chromosome 3 were located in the *QPhs.cnl-2B.1* interval. Even though the comparative map was used to narrow down the possible wheat ESTs and candidate genes located in the *QPhs.cnl-2B.1* interval, some genes present in the homologous rice regions were not used due to a lack of wheat EST information. Also, this study focused only on the homologous regions on rice chromosomes 7 and 3 that contained 60% matching homologous genes in the comparative map. This may have eliminated wheat genes that are located in the *QPhs.cnl-2B.1* region but are homologous with genes on other rice chromosomes. In the absence of wheat genomic sequence, it not possible to order the wheat ESTs within a given deletion bin. The order of rice homologues provides a useful approximation of the wheat order based on the relatively high general macrocollinearity between the two species. It is possible that rearrangements in microcollinearity have occurred between wheat and rice and could lead to false interpretations of gene locations if examined in isolation. Ishimaru et al. (2001) summarized the location of QTL affecting many traits in rice, including seed dormancy. Their map identified loci on chromosomes 3, 5, 7, and 8. A comparative

mapping between the 2BS1-0.53-0.75 wheat bin and rice revealed a macrocollinearity within a 3-mb segment on the long arm of rice chromosome 7 and the short arm of rice 3 but was unable to assess microcollinearity because of the lack of sufficient polymorphic markers.

Homoeologous relationship among the wheat group 2 chromosomes

The physical map and BAC sequence information available for *Aegilops tauschii* may be useful for comparative genomic studies in hexaploid wheat and other grasses; however, the D-genome may not be a good representative for the *QPhs.cnl-2B.1* QTL region on chromosome 2B. Our results suggest that chromosomes 2A and 2B are more closely related to each other at this locus than they are to 2D with regard to gene order and content because some markers (CNL409 and CNL413) that mapped on 2A and 2B did not map on 2D, possibly because of monomorphism of the third band. Despite these differences, the previous comparative mapping indicated that all three genomes were highly similar in gene content and order (Chao et al. 1989). According to Devos et al. (1993) collinearity for group 2 chromosomes was conserved, except the distal region of the short arm of chromosome 2B, which was likely involved in an interchromosomal translocation. Our results presumably did not involve the region of the translocation, as most of the ESTs located within the QTL interval were also located in a similar location on the other two homoeologous group 2 chromosomes.

New markers were used to compare three linkage groups, 2A.2, 2B.1, and 2D reported by Munkvold et al. (2009). All three linkage groups contained a region homoeologous to deletion bin 2BS1-0.53-0.75. There was a minor PHS QTL mapped on 2A.2 by Munkvold et al. (2009) suggesting that it might be homologous to the PHS QTLs on 2B.1 or 2D; however, our results indicated that the PHS QTL on 2A.2 was not homologous with *QPhs.cnl-2B.1* but seems homologous to a distal region outside Wmc474. Also, the PHS QTL on the 2A.2 linkage group is not homologous with *QPhs.cnl-2D.1* on wheat chromosome 2D. Similarly, *QPhs.cnl-2B.1* on wheat chromosome 2B is not homologous with the *QPhs.cnl-2D.1* on wheat chromosome 2D because all three new markers were located distal to the 2D QTL.

Marker development for polyploid species

The difficulty of working in a polyploid has limited progress of marker development in this project. Polyploidy, autogamy, and the similarity between the parents are the limiting factors for finding the necessary polymorphism for mapping in a population (Tanksley and Nelson 1996).

During this study, 22 polymorphic loci (8%) were found between the two parents, from a total of 278 designed primers. Not only parental relationship, but also the types of markers that are developed, limit polymorphism. The markers in this project were developed from ESTs, EST contigs, and gene transcripts that exclude the 5' untranslated region (UTR) of genes. It is possible that an additional polymorphism could be found in a 5'UTR. The information for the 5'UTR sequence was not available in wheat but was available in rice. However, it is difficult to use 5'UTR information from distantly related species because this region tends to be highly variable even in closely related species. Markers developed from coding regions were more useful than SSR markers because the sequences in coding regions are conserved among species.

The regions in rice and *Brachypodium* homologous to the PHS QTL interval on 2B.1 were generally collinear, and new markers were developed to compare the relationships among the PHS QTLs on homoeologues of wheat group 2. The markers and the comparative map presented here provide a valuable resource for the fine mapping and cloning of genes important for PHS and numerous other traits located in the same region on wheat chromosome 2B.

Seed dormancy QTL and the identification of candidate genes involved in seed dormancy

There were several rice seed dormancy QTLs located in the rice regions homologous with the *QPhs.cnl-2B.1* interval, suggesting that some of the same genes may be involved in both wheat and rice SD. From the comparative mapping between wheat, rice, and *Brachypodium*, four strong candidate genes encoding homologues to the Arabidopsis proteins ABAR/GUN5, CIPK2, CIPK3, and CIPK11, were identified for the PHS and SD traits. The ABAR/GUN5 and CIPK3 protein have already been identified to play a role in ABA signaling and seed germination in Arabidopsis (Shen et al. 2006; Kim et al. 2003). The construction of a functional protein annotation network with CIPK2, CIPK3, and CIPK11 helped to identify associations with proteins known to be involved in ABA signaling and seed dormancy and helped validate the inclusion of CIPK2 and CIPK11 as candidate genes.

ABA regulation is also related to abiotic and biotic stress regulation. For example, a study in wheat by Li et al. (2008) revealed that the response to abiotic and biotic stress involves CDPK. There are QTLs for both abiotic stress and biotic stress tolerance that are located in the homologous regions in rice. Several traits related to biotic stress response, including resistance to diseases such as stem rust (Wu et al. 2009), leaf rust (Leonova et al. 2007), and yellow rust (Mallard et al. 2005) were also controlled by genes on wheat chromosome 2B. The stem rust resistance genes *Sr36*

and *Sr40* are located between the *QPhs.cnl-2B.1* flanking markers Wmc474 and Gwm429 (Wu et al. 2009).

The comparative map, molecular markers, and candidate genes identified in this project revealed genes contributing to PHS and SD and will be useful for the mapping and characterizations of other traits that are located in this region.

Acknowledgments This research was partially supported by Hatch project149402 and a USDA National Research Initiative Competitive grant from the USDA Cooperative State Research, Education, and Extension Service Coordinated Agriculture Project grants 2005-05130 and 2006-55606-16629.

References

- Cai H, Morishima H (2000) Genomic regions affected seed shattering and seed dormancy in rice. *Theor Appl Genet* 100:840–846
- Cai H, Morishima H (2002) QTL clusters reflect character associations in wild and cultivated rice. *Theor Appl Genet* 104:1217–1228
- Chao S, Sharp PJ, Worland AJ, Warham EJ, Koebner RMD, Gale MD (1989) RFLP-based genetic maps of wheat homoeologous group 7 chromosomes. *Theor Appl Genet* 78:495–504
- D'Angelo C, Weinl S, Batistic O, Pandey GK, Cheong YH, Schultze S, Albrecht V, Ehlert B, Schulz B, Harter K, Luan S, Bock R, Kudla J (2006) Alternative complex formation of the Ca²⁺-regulated protein kinase CIPK1 controls abscisic acid-dependent and independent stress responses in Arabidopsis. *Plant J* 48:857–872
- Devos KM, Millan T, Gale MD (1993) Comparative RFLP maps of the homoeologous group-2 chromosomes of wheat, rye and barley. *Theor Appl Genet* 85:784–792
- Devos KM, Gale MD (2000) Genome relationships: the grass model in current research. *Plant Cell* 12:637–646
- Distelfeld A, Uauy C, Olmos S, Schlatter A, Dubcovsky J, Fahima T (2004) Microcollinearity between a 2-cM region encompassing the grain protein content locus *Gpc-6B1* on wheat chromosome 6B and a 350-kb region on rice chromosome 2. *Funct Integr Genomics* 4:59–66
- Gu XY, Kianian SF, Foley ME (2004) Multiple loci and epistases control genetic variation for seed dormancy in weedy rice (*Oryza sativa*). *Genetics* 166:1503–1516
- Gu X, Kianian SF, Foley ME (2005a) Phenotypic selection for dormancy introduced a set of adaptive haplotypes from weedy into cultivated rice. *Genetics* 171:695–704
- Gu X, Kianian SF, Hareland GA, Hoffer BL, Foley ME (2005b) Genetic analysis of adaptive syndromes interrelated with seed dormancy in weedy rice (*Oryza sativa*). *Theor Appl Genet* 110:1108–1118
- Guo Y, Xiong L, Song CP, Gong D, Halfter U, Zhu JK (2002) A calcium sensor and its interacting protein kinase are global regulators of abscisic acid signaling in Arabidopsis. *Dev Cell* 3:233–244
- Ishimaru K, Yano M, Aoki N, Ono K, Hirose T, Lin SY, Monna L, Sasaki T, Ohsugi R (2001) Toward the mapping of physiological and agronomic characters on a rice function map: QTL analysis and comparison between QTLs and expressed sequence tags. *Theor Appl Genet* 102:793–800
- Jiang L, Cao YJ, Wang CM, Zhai HQ, Wan JM, Atsushi Y (2003) Detection and analysis of QTL for seed dormancy in rice (*Oryza*

- sativa* L.) using RIL and CSSL population. *Acta Genetica Sinica* 30:453–458
- Kim K-N, Cheong YH, Grant JJ, Pandey GK, Luan S (2003) *CIPK3*, a calcium sensor-associated protein kinase that regulates abscisic acid and cold signal transduction in *Arabidopsis*. *Plant Cell* 15:411–423
- La Rota M, Sorrells ME (2004) Comparative DNA sequence analysis of mapped wheat ESTs reveals the complexity of genome relationships between rice and wheat. *Funct Integr Genomics* 4:34–46
- Leonova I, Laikova L, Popova O, Unger O, Börner A, Röder M (2007) Detection of quantitative trait loci for leaf rust resistance in wheat—*T. timopheevii*/*T. tauschii* introgression lines. *Euphytica* 155:79–86
- Li A, Wang X, Leseberg CH, Jia J, Mao L (2008) Biotic and abiotic stress responses through calcium-dependent protein kinase (CDPK) signaling in wheat (*Triticum aestivum* L.). *Plant Signal Behav* 3:654–656
- Li W, Gill BS (2002) The collinearity of the Sh2/A1 orthologous region in rice, sorghum and maize is interrupted and accompanied by genome expansion in the triticeae. *Genetics* 160:1153–1162
- Lin SY, Sasaki T, Yano M (1998) Mapping quantitative trait loci controlling seed dormancy and heading date in rice, *Oryza sativa* L., using backcross inbred lines. *Theor Appl Genet* 96:997–1003
- Liu Q, Feng J, Buzin C, Wen C, Nozari G, Mengos A, Nguyen V, Liu J, Crawford L, Fujimura FK, Sommer SS (1999) Detection of virtually all mutations-SSCP (DOVAM-S): a rapid method for mutation scanning with virtually 100% sensitivity. *Biotechniques* 26:932–942
- Liu S, Zhang X, Pumphrey MO, Stack RW, Gill BS, Anderson JA (2006) Complex microcollinearity among wheat, rice, and barley revealed by fine mapping of the genomic region harboring a major QTL for resistance to fusarium head blight in wheat. *Funct Integr Genomics* 6:83–89
- Lu G, Moriyama EN (2004) Vector NTI, a balanced all-in-one sequence analysis suite. *Brief Bioinform* 5:378–388
- Mallard S, Gaudet D, Aldeia A, Abelard C, Besnard AL, Sourdille P, Dedryver F (2005) Genetic analysis of durable resistance to yellow rust in bread wheat. *Theor Appl Genet* 110:1401–1409
- Manly KF, Cudmore RH Jr, Meer JM (2001) Map manager QTX, cross-platform software for genetic mapping. *Mamm Genome* 12:930–932
- Meyer K, Leube MP, Grill E (1994) A protein phosphatase 2C involved in ABA signal transduction in *Arabidopsis thaliana*. *Science* 264:1452–1455
- Miao Y, Lv D, Wang P, Wang X-C, Chen J, Miao C, Song C-P (2006) An *Arabidopsis* glutathione peroxidase functions as both a redox transducer and a scavenger in abscisic acid and drought stress responses. *Plant Cell* 18:2749–2766
- Miura K, Lin Y, Yano M, Nagamine T (2002) Mapping quantitative trait loci controlling seed longevity in rice (*Oryza sativa* L.). *Theor Appl Genet* 104:981–986
- Munkvold JD, Tanaka J, Benschler D, Sorrells ME (2009) Mapping quantitative trait loci for preharvest sprouting resistance in white wheat. *Theor Appl Genet* 119:1223–1235
- Ok SH, Jeong HJ, Bae JM, Shin JS, Luan S, Kim KN (2005) Novel CIPK1-associated proteins in *Arabidopsis* contain an evolutionarily conserved C-terminal region that mediates nuclear localization. *Plant Physiol* 139:138–150
- Pandey GK, Cheong YH, Kim KN, Grant JJ, Li L, Hung W, D'angelo C, Weint S, Kudia J, Luan S (2004) The calcium sensor calcineurin B-like 9 modulates abscisic acid sensitivity and biosynthesis in *Arabidopsis*. *Plant Cell* 16:1912–1924
- Peng J, Richards DE, Hartley NM, Murphy GP, Devos KM, Flintham JE, Beales J, Fish LJ, Worland AJ, Pelica F, Sudhakar D, Christou P, Snape JW, Gale MD, Harberd NP (1999) 'Green revolution' genes encode mutant gibberellin response modulators. *Nature* 400:256–261
- Qin Y, Li X, Guo M, Deng K, Lin J, Tang D, Guo X, Liu X (2008) Regulation of salt and ABA responses by CIPK14, a calcium sensor interacting protein kinase in *Arabidopsis*. *Sci China C Life Sci* 51:391–401
- Rodriguez PL, Benning G, Grill E (1998) ABI2, a second protein phosphatase 2C involved in abscisic acid signal transduction in *Arabidopsis*. *FEBS* 421:185–190
- Sheen J (1996) Ca²⁺-dependent protein kinases and stress signal transduction in plants. *Science* 274:1900–1902
- Shen Y-Y, Wang X-F, Wu F-Q, Du S-Y, Cao Z, Shang Y, Wang X-L, Peng C-C, Yu X-C, Zhu S-Y, Fan R-C, Xu Y-H, Zhang D-P (2006) The Mg-chelatase H subunit is an abscisic acid receptor. *Nature* 443:823–826
- Somers D, Isaac P, Edwards K (2004) A high-density microsatellite consensus map for bread wheat (*Triticum aestivum* L.). *Theor Appl Genet* 109:1105–1114
- Sorrells ME, La Rota M, Bermudez-Kandianis CE, Greene RA, Kantety R, Munkvold JD, Miftahudin MA, Ma X, Gustafson PJ, Qi LL, Echalié B, Gill BS, Matthews DE, Lazo GR, Chao S, Anderson OD, Edwards H, Linkiewicz AM, Dubcovsky J, Akhunov ED, Dvorak J, Zhang D, Nguyen HT, Peng J, Lapitan NL, Gonzalez-Hernandez JL, Anderson JA, Hossain K, Kalavacharla V, Kianian SF, Choi DW, Close TJ, Dilbirli M, Gill KS, Steber C, Walker-Simmons MK, McGuire PE, Qualset CO (2003) Comparative DNA sequence analysis of wheat and rice genomes. *Genome Res* 13:1818–1827
- Sugimoto K, Takeuchi Y, Ebana K, Miyao A, Hirochika H, Hara N, Ishiyama K, Kobayashi M, Ban Y, Hattori T, Yano M (2010) Molecular cloning of *Sdr4*, a regulator involved in seed dormancy and domestication of rice. *Proc Natl Acad Sci USA*. doi:10.1073/pnas.0911965107
- Sunnucks P, Wilson AC, Beheregaray LB, Zenger K, French J, Taylor AC (2000) SSCP is not so difficult: the application and utility of single-stranded conformation polymorphism in evolutionary biology and molecular ecology. *Mol Ecol* 9:1699–1710
- Takeuchi Y, Lin SY, Sasaki T, Yano M (2003) Fine linkage mapping enables dissection of closely linked quantitative trait loci for seed dormancy and heading in rice. *Theor Appl Genet* 107:1174–1180
- Tanksley SD, Nelson JC (1996) Advanced backcross QTL analysis: a method for the simultaneous discovery and transfer of valuable QTLs from unadapted germplasm into elite breeding lines. *Theor Appl Genet* 92:191–203
- Wan JM, Cao YJ, Wang CM, Ikehashi H (2005) Quantitative trait loci associated with seed dormancy in rice. *Crop Sci* 45:712–716
- Woodger F, Millar A, Murray F, Jacobsen J, Gubler F (2003) The role of GAMYB transcription factors in GA-regulated gene expression. *J Plant Growth Regul* 22:176–184
- Wu S, Pumphrey M, Bai G (2009) Molecular mapping of stem-rust-resistance gene *Sr40* in wheat. *Crop Sci* 49:1681–1686
- Yan L, Loukoianov A, Tranquilli G, Helguera M, Fahima T, Dubcovsky J (2003) Positional cloning of the wheat vernalization gene *VRN1*. *Proc Natl Acad Sci USA* 100:6263–6268