the plant journal

SEB Society for Experimental Biology

*The Plant Journal* (2016) **86,** 195–207 doi: 10.1111/tpj.13153

RESOURCE

# A whole-genome, radiation hybrid mapping resource of hexaploid wheat

**Vijay K. Tiwari[1,*], Adam Heesacker[2], Oscar Riera-Lizarazu[3], Hilary Gunn[2], Shichen Wang[1], Yi Wang[4], Young Q. Gu[4], Etienne Paux[5,6], Dal-Hoe Koo[1], Ajay Kumar[7], Ming-Cheng Luo[8], Gerard Lazo[4], Robert Zemetra[2], Eduard Akhunov[1], Bernd Friebe[1], Jesse Poland[1], Bikram S. Gill[1], Shahryar Kianian[9] and Jeffrey M. Leonard[2,*]**

[1]*Department of Plant Pathology, Wheat Genetics Resource Center, Kansas State University, Manhattan, KS, USA,*
[2]*Department of Crop and Soil Science, Oregon State University, Corvallis, OR, USA,*
[3]*Dow AgroSciences, Pullman, WA, USA,*
[4]*Crop Improvement and Genetics Research Unit, USDA-ARS, Albany, NY, USA,*
[5]*Diversité et Ecophysiologie des Céréales, INRA, UMR 1095 Génétique, 5 chemin de Beaulieu, F-63039, Clermont-Ferrand, France,*
[6]*Diversité et Ecophysiologie des Céréales, UMR 1095 Génétique, Université Blaise Pascal, F-63177, Aubière Cedex, France,*
[7]*Department of Plant Sciences, North Dakota State University, Fargo, ND, USA,*
[8]*Department of Plant Sciences, University of California, Davis, CA, USA, and*
[9]*Cereal Disease Laboratory, University of Minnesota, Saint Paul, MN, USA*

Received 17 November 2015; revised 4 February 2016; accepted 22 February 2016; published online 4 March 2016.
*For correspondence (e-mails vijtiwari@ksu.edu or jeff.leonard@oregonstate.edu)

## SUMMARY

**Generating a contiguous, ordered reference sequence of a complex genome such as hexaploid wheat (2n = 6x = 42; approximately 17 GB) is a challenging task due to its large, highly repetitive, and allopolyploid genome. In wheat, ordering of whole-genome or hierarchical shotgun sequencing contigs is primarily based on recombination and comparative genomics-based approaches. However, comparative genomics approaches are limited to syntenic inference and recombination is suppressed within the pericentromeric regions of wheat chromosomes, thus, precise ordering of physical maps and sequenced contigs across the whole-genome using these approaches is nearly impossible. We developed a whole-genome radiation hybrid (WGRH) resource and tested it by genotyping a set of 115 randomly selected lines on a high-density single nucleotide polymorphism (SNP) array. At the whole-genome level, 26 299 SNP markers were mapped on the RH panel and provided an average mapping resolution of approximately 248 Kb/cR$_{1500}$ with a total map length of 6866 cR$_{1500}$. The 7296 unique mapping bins provided a five- to eight-fold higher resolution than genetic maps used in similar studies. Most strikingly, the RH map had uniform bin resolution across the entire chromosome(s), including pericentromeric regions. Our research provides a valuable and low-cost resource for anchoring and ordering sequenced BAC and next generation sequencing (NGS) contigs. The WGRH developed for reference wheat line Chinese Spring (CS-WGRH), will be useful for anchoring and ordering sequenced BAC and NGS based contigs for assembling a high-quality, reference sequence of hexaploid wheat. Additionally, this study provides an excellent model for developing similar resources for other polyploid species.**

**Keywords: Radiation hybrid mapping, wheat resource, *Triticum aestivum*, physical mapping, genome assembly, gamma radiation.**

## INTRODUCTION

Genomic mapping and sequencing of polyploid crops is difficult in general, and crops with large and highly repetitive genomes, such as bread wheat (*Triticum aestivum* L.; 2n = 6x = 42, AABBDD; approximately 17 Gb), pose an even greater challenge (Choulet *et al.*, 2010). The available technology and abundance of retrotransposons limited the

earliest sequencing approaches to targeting gene regions through sequencing of expressed sequence tags (Lazo *et al.*, 2004). As technology developed, reducing the complexity in wheat was pursued by sequencing BAC contigs developed from the flow-sorted chromosomes (Gill *et al.*, 2004; Feuillet and Eversole, 2007) and whole-genome pools (Luo *et al.*, 2013; Choulet *et al.*, 2014). The advent of NGS technology allowed for the whole-genome, shotgun sequencing of *T. aestivum* (Brenchley *et al.*, 2012), and diploid progenitors *Aegilops tauschii* (Coss.) (Jia *et al.*, 2013) and *Triticum urartu* (Tumanian ex. Gandilyan) (Ling *et al.*, 2013). Most recently, the shotgun sequencing of flow-sorted chromosome arms was used to assemble draft sequences of wheat (IWGSC, 2014).

Irrespective of whether whole-genome or hierarchical shotgun sequencing data are employed, significant obstacles impede assembling a complete genome sequence. The short reads produced by NGS and the abundance of repetitive elements limit the ability to assemble complete genomes. For example, although chromosomes from the recently released wheat genome draft sequence were sequenced to 241x, only 61% of the genome was covered with an L50 of repeat-masked contigs as small as 1.7 kb (IWGSC, 2014). Therefore, the problem of correctly ordering thousands of contigs relies on linking the contigs with map-ordered markers (Lewin *et al.*, 2009).

Genetic maps are used widely in many eukaryotic organisms for anchoring and ordering contigs, but their efficiency is limited because of the widespread phenomenon of suppression of recombination in proximal regions of chromosomes (Faris *et al.*, 2000; Saintenac *et al.*, 2009). In bread wheat, pericentromeric regions encompassing up to 70% of the chromosome flanking the centromeres often show limited or no recombination. Alternate strategies to determine physical order, including GenomeZipper (Mayer *et al.*, 2009) and population sequencing (POPSEQ) (Mascher *et al.*, 2013), ultimately are dependent on genetic maps and subject to the same limitations. These limitations suggest the need for an alternative approach, independent of recombination, for the precise ordering of contigs in a plant with a complex genome such as wheat.

Deletion stocks (Endo and Gill, 1996) have been used for construction of a reference deletion map of the expressed portion of the wheat genome (Qi *et al.*, 2004). However, the number of deletion breaks is limited, isolation of new stocks is time consuming and, hence, this approach is not practical for the large-scale ordering of contigs. RH mapping uses radiation-induced chromosome breaks to induce marker segregation and offers an alternative solution to these problems (Goss and Harris, 1975; Cox *et al.*, 1990). In RH mapping, the physical distance between markers is estimated based on the frequency with which two markers are lost or retained in a chromosome segment following irradiation. RH mapping in plants was first explored by Riera-Lizarazu *et al.* (2000) where an RH map of a maize chromosome was constructed using an oat–maize addition line. Wardrop *et al.* (2002) reported a somatic cell fusion-based approach for construction of a whole-genome RH panel for barley. The utility of these hybrids was limited by technical difficulties in their generation and continuous culture (Wardrop *et al.*, 2002, 2004). The first RH mapping studies in wheat focused on single chromosome RH panels (Hossain *et al.*, 2004; Kalavacharla *et al.*, 2006). Subsequently, Riera-Lizarazu *et al.* (2010) reported an RH panel to map a sub-genome of wheat using D-genome radiation hybrids. However, these panels (Riera-Lizarazu *et al.*, 2010; Kumar *et al.*, 2015) were derived from a seed-irradiation technique, which produced infrequent deletions (approximately 3–4%) in a small number of chromosomes per sample. This rate of deletion is insufficient to accurately order markers. The utility of RH mapping in wheat was advanced when irradiation of wheat pollen was shown to be an effective method of inducing and recovering a sufficient number of deletions for RH mapping (Tiwari *et al.*, 2012).

To counter the low level of D-genome polymorphism in wheat, the RH mapping protocol being used places the D-genome chromosomes in monosomic condition by pollinating tetraploid plants (AABB) with irradiated pollen from a hexaploid plant (AABBDD) producing a quasi-pentaploid (**AABBD**) $F_1$. In this manner, any D-genome marker can be accurately mapped by its absence or presence. Here, we report the generation of a whole-genome radiation hybrid (WGRH) resource for hexaploid wheat and show that WGRHs provide an excellent resource for ongoing wheat genome sequencing to precisely order sequenced contigs throughout the length of any given chromosome.

## RESULTS

### Developing the whole-genome RH panel

A dosage-optimization curve, based on $\gamma$-irradiation treatments ranging from 0 to 50 Gy of the Chinese Spring (CS) pollen samples, showed no viable embryo development beyond 15-Gy treatments (Figure S1). In total, 540 15-Gy $RH_1$ plants were obtained out of approximately 2000 $RH_1$ seeds (approximately 1500 healthy and approximately 500 shriveled), the remaining seeds either did not germinate or did not produce plants that survived beyond the seedling stage. Subsequently, the CS-whole-genome RH (CS-WGRH) panel was also developed using lower doses, including 5-Gy (355 lines) and 10-Gy (192 lines) treatments. In all of the CS-WGRH panels (5-Gy, 10-Gy and 15-Gy panels), each $RH_1$ seed contained unique deletion events generated in a single irradiated pollen grain (Figure S2). For the A- or B-genome chromosomes, the detection of deletions requires markers that are either co-dominant (different alleles in CS and tetraploid Altar 84) or dominant for CS alleles. For the D genome, markers can be mapped on

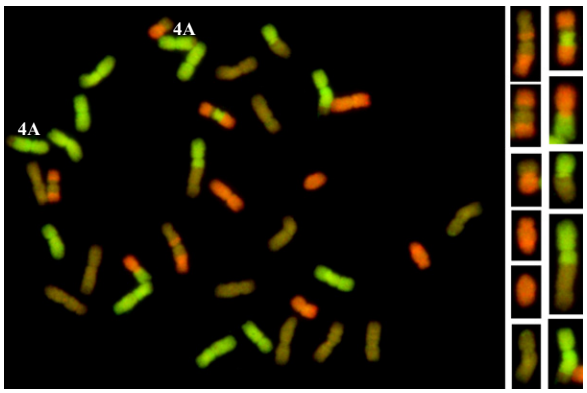


**Figure 1.** Cytological evidence of gamma radiation-induced chromosomal rearrangements in a quasi-pentaploid radiation hybrid of hexaploid wheat. The total chromosome number is 2n = 5x = 35 with 21 chromosomes derived from hexaploid cultivar Chinese Spring and 14 chromosomes derived from the tetraploid cultivar Altar. Although the origin of the A- and B-genome chromosomes cannot be determined, deletions and intergenomic translocation can be easily identified (A-genome chromosomes painted in green, D-genome painted in red, and unlabeled B-genome chromosomes pseudo-colored in brown). The seven D-genome chromosomes are highly fragmented and present in 10 chromosomes in this individual.

this panel as presence–absence variation without a requirement for polymorphism, as all RH1 were monosomic for D-genome chromosomes (Figure S2).

Examination of mitotic metaphases coupled with genomic *in situ* hybridization (GISH) in RHs from the WGRH panel showed extensive chromosome breakage resulting in multiple random deletion and translocation events (Figure 1). In a single plant we observed at least nine major chromosome aberrations. Small terminal and interstitial deletion events within a chromosome or small translocations between chromosomes from the same genome cannot be easily identified by GISH analysis. The D-genome chromosomes were monosomic in the WGRH panel, and the breakpoints were close to the centromeres in four chromosomes, indicating that these chromosomal breaks will allow mapping of markers around the centromeric regions (Figure 1).

### Characterization and high-throughput single nucleotide polymorphism genotyping of the CS-WGRH panel

A subset of 115 randomly selected RH lines from the 15-Gy CS-WGRH panel was used for the initial characterization. To this aim a preliminary screening of these lines with 27 simple sequence repeat (SSR) markers distributed on all 21 chromosomes was performed. The marker retention frequencies were found to range from 50 to 95%, with an average of 85.7% (Figure S3). The same panel was then genotyped on an iSelect single nucleotide polymorphism (SNP) array (Wang *et al.*, 2014) (see the method section) that provided 38 404 markers that could be mapped on RH lines. Marker retention frequencies, based on all 38 404 markers, ranged from 55.9 to 96.5% and with mean marker retention of 82.1%. Analysis of distribution of marker retention bins based on 27 SSR and 38 404 SNPs was performed to find out whether or not a smaller set of markers can help in identification of informative lines (Figure S3), and it suggested that selection of informative lines can be done using three or four markers per chromosome (Figure S3).

### RH mapping

Carthagene (de Givry *et al.*, 2005) was used to group the 30 158 SNP markers, with a threshold distance of 0.3 and a logarithm of the odds (LOD) score set at 12. In total, 53 large RH groups were identified, and these groups contained 26 784 SNPs (see Experimental Procedures). Five small groups with less than 20 markers occupying same loci (considered as singletons) were also found and not included in further mapping. After this step, individual RH groups (53) were used for RH mapping using stringent mapping criteria (threshold distance of 0.3 and LOD ranging from 12 to 15) to break any pseudo-linkages among the markers as described previously (Tiwari *et al.*, 2012). In total, 26 299 SNPs (termed as markers hereafter) were mapped on all 21 wheat chromosomes and comparisons of markers common to the RH maps, and SNP-based consensus maps of all chromosomes (Wang *et al.*, 2014) were used to assign chromosomal locations to the RH groups. Chromosome or chromosome arm assignment was not possible for three RH groups with 25, 27 and 34 markers, and these were not included further in this study. Most chromosomes were represented by two RH marker groups; one per arm, with the exception of a few chromosomes that had three or four RH groups (Table 1). The total number of markers mapped on the A-, B- and D-genome chromosomes was 8470, 11 461 and 6368, respectively. At the whole-genome level, the average marker retention was 78.2% with an average of 78.9, 75.4 and 80.1%, for A-, B- and D-genome chromosomes, respectively. Marker retention frequencies among the homoeologous chromosome groups ranged between 72.5 and 85.5%, and the distribution of marker retention across the homoeologous groups was heterogeneous (Figure 2), indicating chromosome breaks were independent and random events ($P = 0.05$).

The WGRH map comprises 7058 bins with the total map length of approximately 68 660 cR1500. The 26 313 SNP markers were ordered into 2640, 3308 and 1110 bins with map lengths of 25 937.7 cR1500, 34 593.8 cR1500 and 8449.2 cR1500 on A, B, and D genomes, respectively. Considering a genome size of approximately 17 000 Mb for wheat, this whole-genome RH panel provides a resolution of approximately 0.25 Mb/$cR_{1500}$.

### RH maps versus genetic maps

Wang *et al.* (2014) recently presented genetic maps based on eight doubled-haploid (DH) wheat mapping populations

 

**Table 1** Details of the RH maps for all the wheat chromosomes based on SNP markers

| Genome and chromosome | RH groups | Markers mapped | Number of resolved loci | Map length ($cR_{1500}$) | Retention frequency |
|---|---|---|---|---|---|
| 1A | 2 | 1431 | 423 | 4102.7 | 74.5 |
| 2A | 2 | 1690 | 532 | 5954.0 | 75 |
| 3A | 2 | 1216 | 424 | 3765.3 | 72.5 |
| 4A | 3 | 833 | 196 | 1661.1 | 85.5 |
| 5A | 2 | 1261 | 519 | 5143.5 | 81 |
| 6A | 3 | 837 | 233 | 2172.6 | 85 |
| 7A | 3 | 1214 | 321 | 3136.5 | 78.5 |
| *A-Genome* | *17* | *8470* | *2640* | *25 937.7* | *78.9* |
| 1B | 2 | 1467 | 431 | 4582.4 | 75 |
| 2B | 4 | 2321 | 544 | 6365.8 | 72.5 |
| 3B | 3 | 1781 | 470 | 4338.7 | 79.5 |
| 4B | 2 | 1085 | 346 | 3256.6 | 74.5 |
| 5B | 2 | 1612 | 514 | 5237.6 | 77.5 |
| 6B | 2 | 1956 | 513 | 6161.6 | 78 |
| 7B | 2 | 1239 | 490 | 4651.1 | 70.5 |
| *B-Genome* | *17* | *11 461* | *3308* | *34 593.8* | *75.4* |
| 1D | 2 | 1157 | 143 | 1422.3 | 80.5 |
| 2D | 4 | 1156 | 188 | 1152.2 | 82.5 |
| 3D | 2 | 980 | 176 | 1201.1 | 80.5 |
| 4D | 2 | 635 | 177 | 1172.7 | 75 |
| 5D | 2 | 756 | 100 | 767.8 | 84.5 |
| 6D | 2 | 590 | 148 | 719.1 | 76.5 |
| 7D | 2 | 1094 | 178 | 2014 | 81.5 |
| *D-Genome* | *16* | *6368* | *1110* | *8449.2* | *80.1* |
| *Wheat Genome* | *50* | *26 313* | *7058* | *68 659.7* | *78.1* |

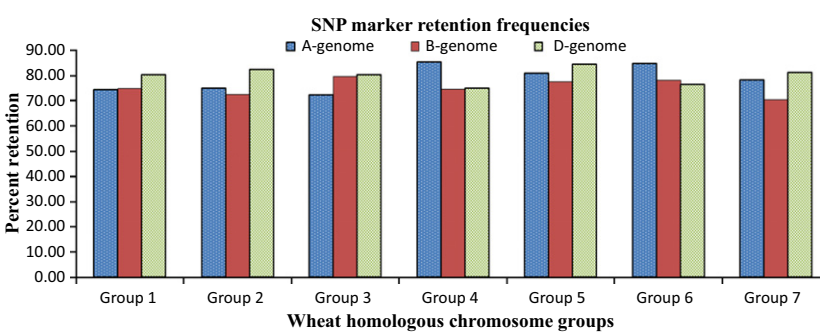


**Figure 2.** SNP marker retention frequencies.
Percentage of SNP markers retained on each of the 21 chromosomes in $F_1$ plants derived from irradiated (15 Gy) Chinese Spring pollen.

along with a consensus map based on six of those populations. As we utilized the same 90 K iSelect SNP array, we were able to assess the efficiency of our calculated RH maps by comparing the number of markers mapped in each of the eight genetic mapping populations against our RH maps. We found that the RH panel allowed mapping of more markers in each genome than any single, DH populations (Figure 3). For the A genome, the number of markers mapped on DH populations (8) ranged from 3980 to 6546, compared to 8470 on the RHs. For the B genome, it ranged from 3529 to 9557 compared to 11 461 on the RHs and 630–2606 for the D genome compared to 6368 on the RHs.

But the biggest difference comes from the D genome. While 630–2606 markers were mapped on genetic maps, our RH map contains 6368. At the chromosome level, the number of markers mapped on the D genome in DH populations ranged between 13 and 693, with an average of 243 markers per chromosome (Figure S4). The RH map allowed anchoring three- to six-fold more markers than the DH populations, with an average of 910 markers per chromosome (Figure S4).

The number of mapping bins provides key information about the resolution of a given mapping panel. At the whole-genome level, the number of resolved genetic bins

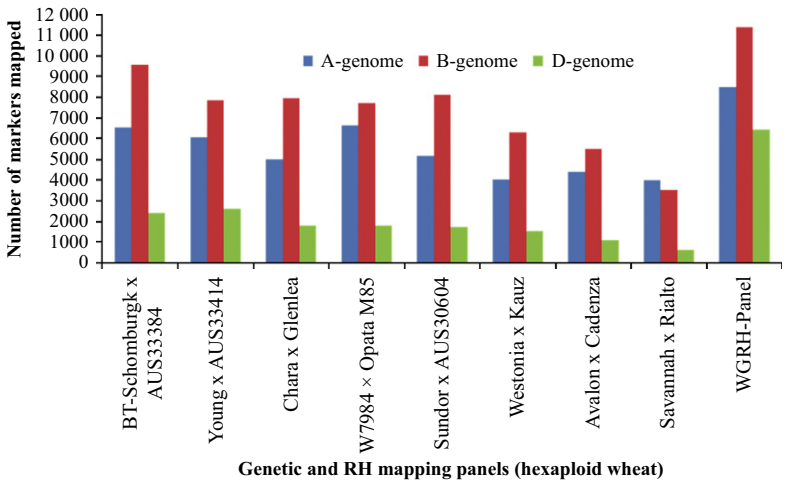


**Figure 3.** Number of mapped markers in DH and RH populations.
Shows number of markers mapped on A-, B- and D-genomes in eight DH populations as well as in a CS-WGRH mapping panel. The information of DH populations in the study was derived from Wang *et al.* (2014).

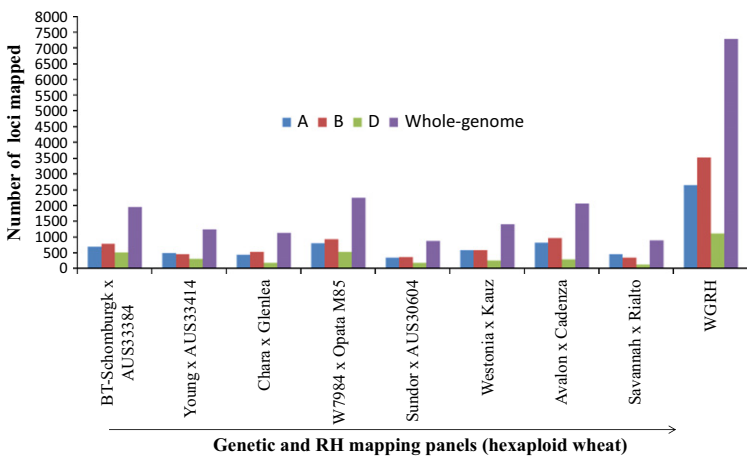


**Figure 4.** Number of mapped well resolved mapping bins.
Shows number of resolved mapping bins on A- B- and D-genome chromosomes in eight genetic populations (genetic bins) and CS-WGRH panel (deletion bins). In genetic population each well resolved locus is separated by at least one recombination event where as in RH maps each mapping bin is separated by a break between two markers.

among the eight DH populations ranged from 873 to 1960 (Figure 4) with the highest number of bins in the B genome followed by the A and D genomes. Average mapping resolution for the D genome ranged from 220 to 550 bins for each of the eight DH maps with an average of 41 mapping bins per chromosome across all eight populations. We observed a three- to eight-fold increase in resolved bins for the RH panel over the DH populations with a total of 7058 total bins (Figure 4). Considering the large genome size of wheat (approximately 17 GB), the 7058 mapping bins provided a mapping bin at every approximately 2.4 Mb.

To validate the marker order and integrity of the RH map, mapping bins also were compared with a recently developed POPSEQ resource (Chapman *et al.*, 2015) where a genetic map of wheat chromosomes was derived from 90 DH individuals from a synthetic W7984 × Opata M85 (SynOpDH) population (Sorrells *et al.*, 2011). Markers for the map were generated using low-coverage, whole-genome sequencing of the population (Chapman *et al.*, 2015). The total number of resolved genetic bins mapped in this population was 1335, which was approximately 20% of the total number of RH bins mapped in the CS-WGRH panel with 115 RH lines.

To assess the validity of marker ordering using the RH approach, we compared the order of markers on the RH map of chromosome 7D to the order of markers on the SNP-based consensus genetic map (Wang *et al.*, 2014).

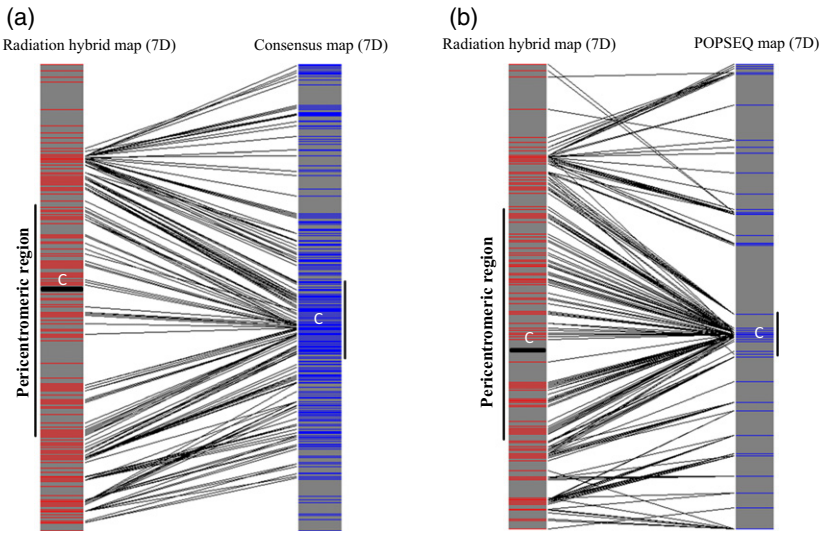


**Figure 5.** Comparative mapping between RH, genetic consensus and POPSEQ maps,
(a) Comparison of RH map (left) of wheat chromosome 7D from reference wheat line 'Chinese Spring' and consensus map (on the right) derived from eight DH population, based on common markers mapped on both of them.
(b) Comparison of CSS contigs (chromosome 7D) mapped and ordered on RH map and on POPSEQ genetic map of chromosome 7D. Black lines connect CSS contigs mapped on RH map as well as POPSEQ map of chromosome 7D. Black vertical bars on left (RH map) and right (consensus and POPSEQ maps) sides in Figure 5(a, b) indicated pericentromeric region on these maps whereas small black box within RH maps indicated putative location of centromere. To identify the putative location of pericentromeric region, BLAST based searches were performed using SNP sequences against deletion bin mapped EST sequences of chromosome 7D.

Chromosome 7D was chosen because it contained the highest number of markers common to both maps; the overall order of markers in both of the maps was largely in agreement (Figure 5a). As expected, the consensus genetic map had better resolution in the telomeric regions of the chromosome, whereas the RH map was highly informative for the centromeric regions. Out of 1094 7D-specific markers compared with consensus DH map, 564 were common on both the maps. Out of these 564 markers, 11 markers showed flips and three small inversions and the rest of the markers were collinear.

A similar comparison for the same chromosome was made with a genetic map of chromosome 7D generated using a POPSEQ approach. The POPSEQ data generated (IWGSC, 2014) was used to order CSS (Chromosome Survey Sequence for hexaploid wheat Chinese Spring) based contigs. We used SNP sequences of the 7D RH map and identified common contigs that were anchored based on POPSEQ (IWGSC, 2014) and RH maps. A comparison of the CSS contigs ordered based on POPSEQ and RH data is presented in Figure 5(b). A visual comparison of the maps (Figure 5b) showed that the POPSEQ map offered better resolution towards the telomeric regions. Resolution of the POPSEQ map, relative to the RH map, appeared uneven over the length of the chromosome and very poor towards the centromere of 7D. In total, 610 markers were common between both the maps. The relative marker order was collinear with the exceptions of 13 markers with local flips and six anomalously placed markers (Figure 5a,b).

### Comparing the RH and reference sequences

The reference sequence of chromosome 3B (Choulet *et al.*, 2014) was used to validate the ordering potential and estimate the mapping resolution of the RH map for that chromosome. The long arm of chromosome 3B was selected because of the high number of well distributed SNP markers on the reference 3B and RH maps. To identify the physical location of 3BL-specific SNPs, we used sequences of markers mapped to the 3BL RH map (Table S1) and ran a BLAST search against the 3B pseudomolecule. A set of 516 markers was selected that mapped to the long arm of pseudomolecule 3B with a 99% similarity over the SNP sequences. The genotyping data for selected markers were used to construct an RH map of chromosome 3BL using Carthagene. The RH map of chromosome 3BL resolved 114 RH bins with a map length of 796.4c$R_{1500}$ (Figure 6 and Table 2) The map spanned 422.45 Mb of the reference sequence including seven 3BL deletion bins, from a centromeric deletion bin (C-3BL2-0.22) to a terminal deletion bin (3BL7-0.63-1.00). The RH map resolution averaged 0.53 Mb/c$R_{1500}$, with an average RH mapping bin approximately every 3.7 Mb. Comparing the order and distance of the deletion bin map, RH map, physical map, and genetic map, we noted a high level of collinearity between the RH

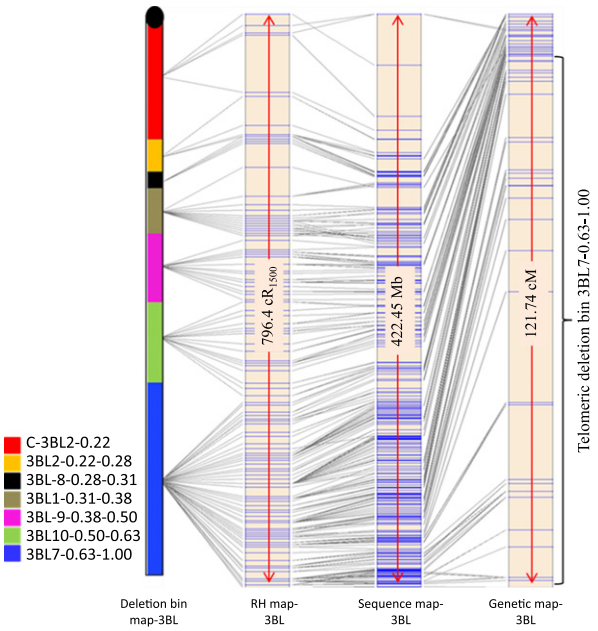


**Figure 6.** Comparison of RH map with a reference physical map.
3BL deletion bin, RH and genetic, maps were compared with long arm of reference chromosome 3B (Choulet *et al.*, 2014). The genetic map used here was derived from the Opata population (Wang *et al.*, 2014). Black circle on the top of deletion bin map represents the centromere.

and physical map; whereas the genetic map did not show a linear relationship with either the physical or the RH map (Figure 6 and Tables 2 and S1). Table 2 presents a detailed comparison of RH map and genetic map with respect to 3BL sequence map (Figure 6 and Table S1). The comparison of RH map and DH map clearly indicated that RH offers many fold higher resolution than the DH or the genetic maps in the proximal region of the chromosome (Table 2).

Dot plots of RH marker $cR_{1500}$ distance against the physical distance on the reference 3BL map showed the strong linear agreement between the RH map and physical positions (Figure 7a). Overall, the RH map distances reflected physical distances more accurately than genetic map distances while providing more uniform coverage (Figure 7a). Similar comparisons were made for 3BL between the reference sequence and genetic map of the SynOpDH population (Figure 7b) and the physical map and the consensus genetic map (Wang *et al.*, 2014) (Figure 7c). There was good agreement in marker order between the physical positions and both the RH and genetic maps. However, while we observed a linear agreement between the RH map distance and physical position on the chromosome arm, the genetic maps were non-linear relative to the physical position, showing the pronounced recombination suppression in a large proximal part of the chromosome.

### Anchoring and ordering NGS contigs

Our results for chromosome arm 3BL indicated that RH maps can be used to anchor and order map scaffolds of either BAC or NGS contigs. To both test and extend the utility of contig anchoring, sequences of RH-mapped markers were used to anchor and order CSS contigs from hexaploid wheat Chinese Spring (IWGSC, 2014). A BLASTN search was made using 8370, 11 461 and 6368 markers mapped on A-, B-, and D-genome chromosomes, respectively, from this study, against the CSS assembly (IWGSC, 2014).

For the CSS assemblies of the reference CS wheat, markers mapped on the WGRH panel were useful to anchor a total of 16 299 NGS contigs with 5426, 8486, and 3060 NGS contigs on A, B and D genomes, respectively. In total, 16 299 NGS contigs were ordered in more than 4400 mapping bins on the RH map. When compared to the CS-WGRH panel, the SynOpDH genetic mapping population

**Table 2** Bin-specific RH map information for chromosome arm 3BL and its comparison with genetic map[a]

| | | WGRH map | | DH map | | Average Mb/mapping bin | |
|---|---|---|---|---|---|---|---|
| Deletion bin | Megabases spanned | Markers mapped | RH bins | Markers mapped | Genetic bins | RH map | DH map |
| C-3BL2-0.22 | 85.59 | 16 | 7 | 2 | 2 | 12.23 | 42.8 |
| 3BL2-0.22-0.28 | 24.84 | 21 | 4 | 5 | 2 | 6.21 | 12.42 |
| 3BL8-0.28-0.31 | 0.99 | 9 | 4 | 4 | – | 0.25 | – |
| 3BL1-0.31-0.38 | 38.85 | 47 | 17 | 13 | 2 | 2.29 | 19.43 |
| 3BL9-0.38-0.50 | 39.12 | 38 | 15 | 3 | 2 | 2.61 | 19.56 |
| 3BL10-0.50-0.63 | 46.33 | 34 | 17 | 19 | 7 | 2.73 | 6.62 |
| 3BL7-0.63-1.00 | 160.50 | 309 | 50 | 93 | 30 | 3.21 | 5.35 |
| 3BL | 422.45[b] | 474[c] | 114[c] | 139[c] | 45[c] | 3.70[d] | 9.39[d] |

[a]DH map corresponds to W7984 × Opata M85 (SynOpDH) population (Sorrells *et al.*, 2011).
[b]Total of the Mb spanned on chromosome 3BL (see Table S1).
[c]Total of column covering seven deletion bins of chromosome 3BL.
[d]Average Mb/mapping bin, calculated by dividing total Mb spanned by total number of mapping bins in RH and genetic DH maps.

 

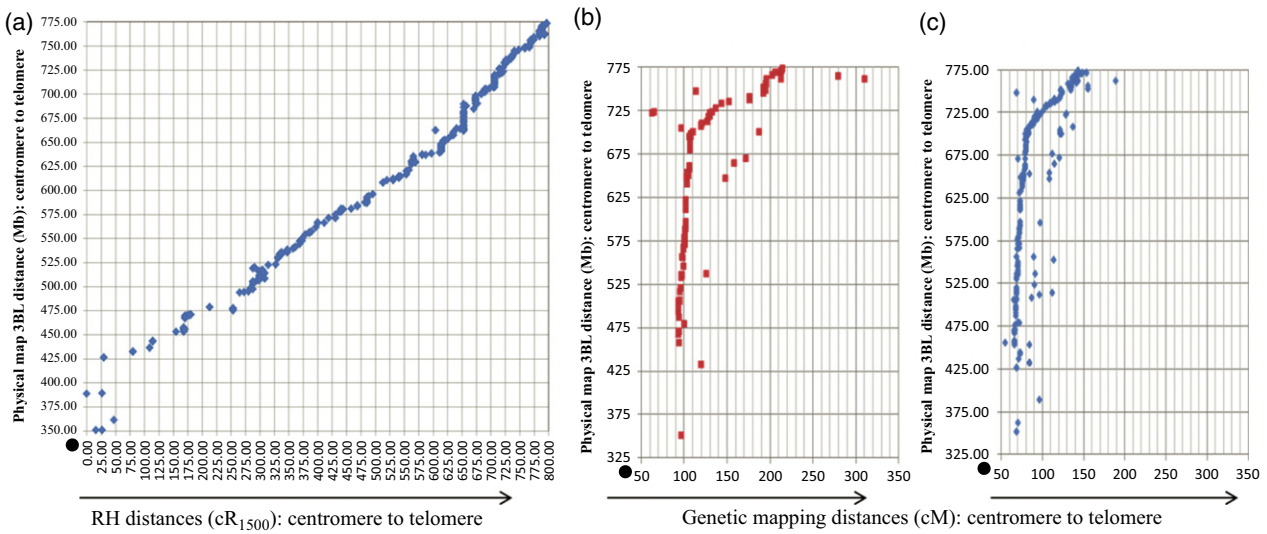


**Figure 7.** RH and genetic marker distances are plotted against their respective positions in the reference physical map of the chromosome arm 3BL (from centromere to telomere).
(a) RH mapping distances ($cR_{1500}$) plotted against their reference physical map position on 3BL.
(b) Genetic mapping distances (in a double haploid SynOpDH population) plotted against their reference physical map on 3BL.
(c) Genetic mapping distances of a consensus map (derived from six double haploid populations) plotted against the reference physical map positions on 3BL. The small black circle indicates the putative location of the centromere. Genetic mapping information of SynOpDH and SNP consensus maps were derived from Wang *et al.* (2014).

had 4366, 5123 and 2366 CSS contigs anchored on A, B, and D genomes, respectively, ordered in approximately 1000 genetic bins.

A similar search was done with RH-mapped A-genome markers against *T. urartu* assemblies (Jia *et al.*, 2013), and RH-mapped D-genome markers against *Ae. tauschii* assemblies (Ling *et al.*, 2013). More than 8000 A-genome markers from the CS-WGRH panel anchored 6513 NGS contigs of *T. urartu*, than the 3100 contigs ordered using the SynOpDH derived map (Table 3). Similarly, 5266 *Ae. tauschii* NGS contigs were anchored with D-genome markers (RH-mapped) while only 2328 were tagged with SynOpDH-mapped markers (Table 3).

**Table 3** Application of WGRH panel in anchoring NGS assemblies for *T. urartu*, *Ae. tauschii* and chromosome survey sequences of hexaploid wheat 'Chinese Spring'

| | | | Number of NGS contigs anchored | |
|---|---|---|---|---|
| S. no | NGS assemblies | Genome 2n | Synthetic × Opata map | WGRH map |
| 1 | Ae. tauschii | $D^tD^t$ | 2328 | 5266 |
| 2 | T. urartu | $A^uA^u$ | 3100 | 6513 |
| 3 | CSS-A-genome | AABBDD | 4366 | 5426 |
| 4 | CSS-B-genome | AABBDD | 5123 | 8486 |
| 5 | CSS-D-genome | AABBDD | 1366 | 3060 |

## DISCUSSION

The difficulty in ordering BAC, sequence contigs, and scaffolds to chromosomes is a major bottleneck in assembling the wheat genome and other large, complex genomes (Paux *et al.*, 2008; Philippe *et al.*, 2013). For example, a large genetic population of approximately 1100 individuals was used to anchor and order BAC contig based *Ae. tauschii* physical maps, yet 36% of contigs were unordered due to limited recombination events or lack of polymorphic markers (Luo *et al.*, 2013). Strategies combining genetic mapping, deletion bin assignment, and synteny with grass genomes were employed to improve ordering of chromosome 1BL BAC contigs along with multi-parent advanced generation inter-cross populations, yet 50% of the contigs containing 37% of the genes remain unordered because of insufficient resolution (Philippe *et al.*, 2013).

This study provides a comprehensive report on developing and utilizing a WGRH resource for wheat. Chinese Spring wheat was chosen because it has been extensively exploited for genetic and genomic studies due to an abundance of genomic resources available, including nullisomic–tetrasomic, ditelosomic, double-ditelosomic and deletion lines (Sears, 1954; Sears and Sears, 1978; Endo and Gill, 1996). Chinese Spring is also the cultivar being sequenced in the ongoing genome sequencing of hexaploid wheat (IWGSC, 2014). The available genetic stocks and genetic maps were pivotal for wheat genome analysis and gene discovery. The RH panel and RH map provide an

additional resource for wheat genome analysis, for efficient anchoring of contigs, and for the assembly of a complete wheat genome sequence.

The mapping information of a RH panel is maximized as the deletion rate approaches 50%. Because of its ability to tolerate high rates of deletion, endosperm derived from irradiated pollen is a superior approach for wheat RH mapping (Tiwari *et al.*, 2012). However, we considered the possibility that marker probe hybridization to multiple copies of homeologs presented by the heptaploid nature of endosperm tissue (AAABBBD) in this system would hinder the ability of the Illumina iSelect system to distinguish marker loss (Akhunov *et al.*, 2009). To reduce genomic complexity and therefore marker noise, we developed a RH panel of $F_1$ plants (AABBD) generated from pollination of tetraploid wheat with the irradiated pollen of hexaploid wheat. The successful identification and mapping of 26 299 markers demonstrate the utility of the iSelect platform to detect marker loss for SNP-based RH mapping in wheat. Because SNPs are now the predominant marker system in wheat and other plants, the RH maps generated here and in the future can be integrated with genetic maps to form integrated consensus maps.

Marker retention rates were similar for the A-, B-, and D-genomes (79, 77 and 80%, respectively) yet 32.2, 43.6, and 24.2% of the markers mapped to the A-, B- and D-genome RH maps respectively. The most significant factor contributing to uneven marker distribution in the RH-mapped markers was likely the distribution of SNPs on the iSelect panel wherein 35, 50, and 15% of the mapped SNPs were located on the A-, B-, and D-genome respectively (Wang *et al.*, 2014).

The most significant difference between the two mapping systems, RH versus genetic, was the large increase in the number of D-genome markers mapped in the RH system. The unique nature of the cross (hexaploid by tetraploid wheat) (Riera-Lizarazu *et al.*, 2010) generates $F_1$s monosomic for D-genome chromosomes (AABBD). Since RH genotyping and mapping rely on detection of marker loss, all detectable D-genome markers become scorable without the requirement for polymorphism inherent in genetic mapping protocols. The lower number of D-genome SNP markers in the genetic populations is expected due to low levels of polymorphism found in this genome (Akhunov *et al.*, 2009). A factor that likely increased the total number of markers mapped in the RH is the ability to detect loss of and map homoeologs of SNPs in the using the Genome Studio Polyploid Clustering algorithm (see method section for SNP genotyping). The A- and B-genome RH maps are longer and have higher number of resolved bins compared to the D-genome map. The high clustering and a smaller map length can be attributed to lower number and poor distribution of D-genome markers.

The marker retention rates we measured were similar for the A-, B-, and D-genomes (79, 77 and 80%, respectively), indicating that the total proportion of the deleted regions is similar within the three sub-genomes. Our results with GISH analysis showed chromosome breaks in centromeric regions of the chromosomes indicated that markers can be ordered in these regions using RH mapping panel (Figure 1). Our previous work (Tiwari *et al.*, 2012) found that most deletions (approximately two-thirds) in irradiated-pollen RH panels were interstitial while the remaining one-third of deletion events was terminal. In that study the number of deletions per chromosome ranged from 0 to 7. Due to the lower radiation dose and lack of line selection in this study, we expected to find approximately only three or four deletions per chromosome. Because each RH line has 21 chromosomes and 115 lines were genotyped, there were 2415 CS chromosomes targeted in this work. Therefore we estimated that the number of radiation-induced deletion bins available for mapping would be between 7245 and 9660. This estimate is in line with our results wherein we mapped 7058 deletion bins (Table 1). Our previous results with an endosperm-derived RH panel (Tiwari *et al.*, 2012) indicated that smaller (<20 Mb) and interstitial deletions are favored and provide an excellent platform to develop high-resolution, overlapping deletion bins for genome mapping.

On average we found two unlinked RH groups per chromosome. Two factors may contribute to this phenomenon. First, the SNPs used in this study were developed from RNA-seq data mapped on reference transcripts (Wang *et al.*, 2014), and the panel developed from that resource is not evenly distributed along the chromosome. Subsequently a lower number of SNPs is expected in the centromeric regions. In support, we observed that the number of markers common to consensus and RH maps was higher towards the telomeric region of the chromosomes. Linkage between markers in the RH maps is established by the frequency of simultaneous loss of markers within a deletion. Because radiation-induced breaks accumulate randomly, the number of overlapping deletions failing to incorporate at least two markers extends to regions of low SNP density, and the ability to extend linkage groups is lost. As genotyping technologies such as insertion site-based polymorphisms (ISBPs) (Paux *et al.*, 2010) and repeat junction markers, which are based on markers distributed more randomly than intergenic SNPs, become available, we expect the capability of RH mapping to be extended. The second factor limiting our ability to generate a single linkage group per chromosome may be the necessity of retaining a functional centromere in $F_1$ plants. The deletions required to create linkage between markers flanking either side of a centromere would create a non-viable plant and are therefore never recovered. In animal sys-

tems, generating an RH panel relies on incorporating chromosome fragments into functional chromosomes of a host cell (Cox *et al.*, 1990). In those systems, a functional centromere from the target species is not required. Wardrop *et al.* (2002) developed a similar system in plants, but the technical challenges of that system made it difficult to implement (Wardrop *et al.*, 2004). The system developed here is simpler to implement than cell-fusion-based techniques but requires that chromosomes be maintained throughout early plant development. However, as more lines are accumulated and markers are developed from non-recombinant regions, each linkage group should approach the functional centromere.

POPSEQ of a 90-member DH hexaploid wheat population generated 112 687 markers, but these markers only define 1335 recombination bins (approximately 64 bins per chromosome) covering 2826 cM (Chapman *et al.*, 2015) which limits ordering of sequenced contigs and physical maps on a scaffold. RH maps at the whole-genome level provided 7058 well resolved mapping bins, with an average of approximately 340 bins per chromosome. The average mapping resolution provided by the POPSEQ approach was approximately 12 Mb whereas the average resolution in our study was approximately 2.4 Mb. Part of this difference can be attributed to the smaller genetic mapping population (90 DHs) versus the 115 RHs described in this report. However it also suggests that deep sequencing of an informative RH panel may be useful for genome mapping studies.

Unlike recombination events in a genetic population, we observed that chromosome breaks are largely random in RH mapping. Resolution in terminal regions of RH maps will increase with sample size, but a combination of genetic and RH approaches may offer the best opportunity to order contigs near the telomeres for wheat genome. Nearer to the centromere, RH maps offer a unique tool to order markers/contigs as it is suggested by our comparison of marker distribution along chromosome 7D (Figure 5). RH map resolution is a function of the number of chromosome breaks, marker distribution, retention patterns, and number of lines used. Doubling the number of RH lines should approximately halve map resolution unlike genetic mapping populations.

RH mapping can potentially provide a solution for constructing a marker scaffold along the length of a chromosome irrespective of recombination patterns (Olivier *et al.*, 2001). This is supported by the co-linearity of the 3BL RH map with the physical map of 3BL (Figure 7). While the 3BL map spans approximately 422 Mb, the unequal distribution of SNP markers on the iSelect panel limited our ability to explore the full mapping potential of the RH panel. For example approximately 85 Mb of the physical map was represented by only seven bins in centromeric deletion bin C-3BL2-0.22 whereas map distances indicated that many more markers would be required for developing a high-resolution map of this region (Table 2).

In a previous study we achieved an average map resolution of approximately 700 Kb by selecting lines with 40–70% marker retention and deletions for multiple chromosomes (Tiwari *et al.*, 2012). In this pilot study, lines were randomly picked and genotyped. Screening $F_1$s and selection of the most informative lines with a small panel of markers prior to SNP genotyping would further optimize mapping resolution. We generated a 540-member 15-Gy CS-WGRH panel. Based on the measured distribution of per-line retention frequencies from our previous work, we estimate that selection of approximately 188 of the most informative lines from the 540 member panel would be sufficient to generate a RH map with approximately 1 Mb resolution.

**Application of the WGRH panel**

Based on our results for chromosome 3BL, we propose an approach for the effective utilization of the CS-WGRH resource for anchoring and ordering sequenced BAC and NGS contigs. Because SNP sequences are short, an *in silico*-based application requires stringent filtering criteria. The approach used in constructing the 3BL map wherein stringent filtering of markers against a database of chromosome-localized sequences is performed prior to mapping is suggested. All genotyping data for 115 RH lines and 81 587 SNP markers scored in this study are publically available (http://probes.pw.usda.gov/ATRJM/RHmaps/). The RH panel has been derived from CS and thus may be beneficial for wheat genome sequencing groups (http://www.wheatgenome.org/Projects) assembling sequence for any given wheat chromosome.

More than 6000 markers mapped on D-genome chromosomes provide an opportunity to compare the D genome of reference wheat line CS with *Ae. tauschii* accession (AL8/78) that has been used for sequencing (Luo *et al.*, 2013). The iselect SNP array includes a large number of D-genome specific SNPs already mapped on *Ae. tauschii* chromosomes (Luo *et al.*, 2013). Our study provides an opportunity to compare both maps using common markers (http://probes.pw.usda.gov/ATRJM/RHmaps/). The D-genome specific markers on the RH map can be used as an additional resource to anchor and order the *Ae. tauschii* BAC contigs in low recombination regions.

We cannot detect small deletions in our system due to the markers utilized and mapping algorithms, but they are expected to be abundant (Tiwari *et al.*, 2012). Various studies have provided high-density maps based on sequence tags from both low- and high-depth sequencing of genetic populations (Elshire *et al.*, 2011; Poland *et al.*, 2012; Mascher *et al.*, 2013; Chapman *et al.*, 2015). Because these datasets were generated by sequencing of genetic populations, however, resolution is limited in the proximal or cen-

tromeric region due to lack of recombination events. Deep sequencing of a selected and informative RH panel would be an important asset for utilizing the enormous data generated by POPSEQ approaches. The present RH panel would be an appropriate candidate for sequencing as it has been characterized by a dense marker system. Selected panels of RH lines developed in this study are available for genotyping on any platform able to detect marker loss. DNA samples and additional information will be made available on request for the entire CS-WGRH panel.

The method described here is a quick and easy approach to develop RH panels in wheat and can be used as a model for developing similar resource for other polyploid crop plants. The CS-WGRH resource and genome mapping data generated in this study will be useful for anchoring and ordering sequenced BAC- and NGS-based contigs for assembling a high-quality, reference sequence of hexaploid wheat.

## EXPERIMENTAL PROCEDURES

### Plant material

Hexaploid wheat line CS, a standard genotype for cytogenetics, genetics, and genomics research, was used as the male parent. The tetraploid cultivar Altar 84 (*Triticum turgidum* L. (2n = 4x = 28; AABB) was used as the female parent.

### Generating pollen irradiation based radiation hybrid panel

Pollen radiation experiments were performed as reported by Tiwari *et al.* (2012). To generate RH mapping panels, dehiscent spikes from greenhouse-grown CS were excised from plants and stems were placed in water. Sets of intact spikes were irradiated with $\gamma$-rays, and pollen from the irradiated spikes was used to pollinate previously emasculated spikes of Altar 84. All radiation experiments used a Gamma Cell 220 irradiator at the Radiation Center of Oregon State University, USA. A radiation dose curve using treatments of 0-, 10-, 15-, 20-, 30-, 40- or 50-Gy irradiated pollen samples and approximately 260 emasculated Altar 82 spikelets per treatment were used. To develop the CS/AL 15-Gy panel (Figure S2), approximately 200 spikes of Altar84 were emasculated and pollinated with 15-Gy irradiated CS pollen. To avoid an embryo-rescue step, we modified the approach and harvested the seeds from the spikes 20 days after pollination and kept them in an oven at 37°C for 30 days; after drying seeds were kept on moist filter paper for germination. Each seed represents an independent RH event (Figure S2). The $F_1$ plants from the crosses between these genotypes were quasi-pentaploid (AABBD).

### Preliminary assessment of the CS-whole-genome RH panel

DNA extraction from 10-Gy and 15-Gy pollen plant lines used the method described by Riera-Lizarazu *et al.* (2010). Initial assessment of 10-Gy and 15-Gy CS-$WGRH_1$s based on 14 SSR markers from D-genome chromosomes (two per chromosome), six A-genome, and seven B-genome-specific SSR markers were tested on the RH panels. Polymerase chain reaction assays and the markers used were as described by Tiwari *et al.* (2012). The 10-Gy panel had significantly higher marker retention than that of 15-Gy panel; therefore, 10-Gy panels were not used for further work.

### High-throughput genotyping using the iSelect 90 K SNP genotyping assay

A randomly selected subset of 115 RH lines from 15-Gy CS-WGRH panel was genotyped on the iSelect 90 K SNP array (Wang *et al.*, 2014). Infinium SNP genotyping assays used the Illumina BeadStation and iScan instruments according to the manufacturer's protocols (Illumina, San Diego, CA, USA). The allele clustering and genotype calling of 81 587 SNP markers was performed using Genome Studio Polyploid Clustering v1.0 software (Illumina). Automated calling of clusters was performed using both optical- and density-based algorithms. All markers that produced two clusters under the optical scan or the density-based scan were utilized for mapping as follows. Because the program arbitrarily assigns cluster numbering, all lines in clusters with frequencies of less than 0.5 were assigned a value of 0 (absence) for that marker. Lines included in the larger cluster were assigned a marker value of 1 (presence). Unclustered data points were scored as missing data. Lines that produced three clusters in the initial automated clustering were assumed to be detecting radiation-induced loss at two loci (homeologs or paralogs). These lines were manually added to the mapping data by sequentially regrouping each of the minor clusters with the dominant cluster. In this manner, each SNP that produced three clusters yielded two data points per line. Markers that produced more than three clusters were excluded from further analysis. In a RH there is the possibility of all homeologs present in a SNP assay. All the RHs carry the same Altar allele for all A-, and B-genome homeologs. The area where each RH is different is where hybrids may have a loss in the A, B, or D genome or in rare cases a loss in two genomes in a single hybrid. As with all SNP genotyping there are only two possibilities, so genotyping this way can pick up both the loss of an A-type and B-type allele. This is the reason why we were able to genotype multiple homeologs from SNPs in the array. The maximum number of loci from a SNP able to be reliably mapped was two.

### Radiation hybrid mapping

To filter out low-quality markers in the first step of RH mapping, only markers with an 88% or lower marker retention were selected for grouping (marker retention is the percentage of lines that retain a specific marker). The purpose of the filtering was to reduce the number of markers that represent bad genotyping data rather than true deletions. In total, 30 158 SNPs were selected for further analysis, and provided high-quality data for RH mapping. Initial grouping and constructing of maps based on RH marker data used the software package Carthagene (de Givry *et al.*, 2005), which utilizes an equal-retention model and computes all two-point distance scores. RH groups were determined using minimum two-point LOD scores of 12.0 and a threshold distance of 0.3. RH groups were assigned to chromosomes by comparing them to a SNP-based consensus genetic map (Wang *et al.*, 2014). After this step, individual RH groups were grouped at LOD scores of 14.0 and a threshold distance of 0.3; the resulting groups were used for constructing maps for individual subgroups. Marker order of initial RH maps was improved using the Carthagene commands 'greedy', 'annealing', 'flip' and 'polish' using default settings. One group identified to have a chimera from 7A and 7D markers was further regrouped at threshold distance of 0.3 and an LOD score of 15, and then maps from two large subgroups were developed for 7A and 7D chromosomes.

 

### The 3BL RH map and mapping of NGS contigs

We performed a BLAST search based on SNP sequences against the 3B pseudomolecule database and sequences with approximately 99% or higher similarity were used to identify markers that may represent a locus or its physical location on the 3BL reference map. If multiple hits for a given sequence were obtained, only the best hit was used for further analysis. SNPs with a known 3BL deletion bin location were selected (519). Because these bins occupied large physical distances and marker coverage was sparse, individual deletion bin-specific maps were made (threshold distance 0.5 and LOD 8). Subsequently, individual groups were merged together followed by map validation commands including annealing, greedy, flips and polish (using default settings) to get a whole chromosome arm-specific map.

The RH data were used to map NGS contigs (Table 3) from CSS (Chromosome Survey Sequences from hexaploid wheat line Chinese Spring), *Ae. tauschii* and *T. urartu* NGS assemblies (IWGSC, 2014; Jia *et al.*, 2013; Ling *et al.*, 2013). We performed BLAST-based searches of SNP sequences against sorted CSS and *T. urartu*, *Ae. tauschii* assemblies, and used a stringent filtering criterion of 99% of the sequence similarity for at least 100 bp of sequences as described by Wang *et al.* (2014) and IWGSC (2014).

Sequences of the markers mapped on D-genome chromosomes were used to BLAST search against extended marker sequences of the *Ae. tauschii* physical maps that are publically available (http://probes.pw.usda.gov/WheatDMarker/). In the case of multiple hits, only the best hits were selected and associated information with markers, including genetic map position and contig assignments of the physical maps, were extracted and comparative map information of *Ae. tauschii* markers against RH maps of the D genome was produced (http://probes.pw.usda.gov/ATRJM/RHmaps/).

### Chromosome painting

The protocols for chromosome painting were essentially as described in Zhang *et al.* (2001) with minor modifications. The genomic DNA of *T. monococcum* and *A. tauschii* were labeled by nick translation with biotin-16-dUTP and dig-11-dUTP, respectively. Genomic DNA of wheat was used as a blocker. Chromosomes were counterstained with 4′,6-diamidino-2-phenylindole (DAPI) in Vectashield antifade solution (Vector Laboratories, Burlingame, CA, USA). The images were captured with a Zeiss Axioplan 2 microscope using a cooled charge coupled device camera CoolSNAP HQ2 (Photometrics, Tucson, AZ, USA) and AxioVision 4.8 software (Zeiss, Jena, Germany). The final contrast of the images was processed using Adobe Photoshop CS5 software (San Jose, CA, USA).

### Availability of supporting data

Whole-genome chromosome maps, raw data files including genotyping data (GenomeStudio file), and RH maps of all wheat chromosomes along with marker sequences are available at http://probes.pw.usda.gov/ATRJM/RHmaps/.

## COMPETING INTERESTS

The authors declare no competing financial interests.

## AUTHORS' CONTRIBUTIONS

VT, JL, OL, YG, RZ, ML, GL, BG, and SK planned the research. AH, EP, DK, AK, JP, YW, HG, EA, and BF provided significant help with data analysis. VT and JL wrote the manuscript.

## ACKNOWLEDGEMENTS

This work was supported by funding from the National Science Foundation, Plant Genome Research Program (NSF-PGRP) grant no. IOS-082210. A part of research was also supported by Oregon Wheat Commission. VKT was supported by WGRC I/UCRC partially funded by an NSF grant contract (IIP-1338897). Computational support for this project was provided by the Beocat High-performance Computing Cluster at Kansas State University, which is funded in part by NSF grants CNS-1006860, EPS-1006860, and EPS-0919443. We thank Jon W. Raupp and Nidhi Rawat, WGRC, for critical reading of the manuscript. We also thank Oregon State University undergraduates Lisa Van Ryssegem, Tianxiao Li, and Leman for their help with SNP genotyping. This work represents contribution number www-J from Kansas Agricultural Experiment Station, Kansas State University, Manhattan KS 66506–502, USA.

## SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article.

**Figure S1.** Dosage optimization curve for the development of CS-WGRH panel.

**Figure S2.** Development of CS-WGRH panel.

**Figure S3.** Distribution of RH lines among different marker retention bins.

**Figure S4.** Number of markers mapped on D-genome chromosomes in recombination and WGRH maps.

**Table S1.** Comparative mapping of chromosome arm 3BL.

## REFERENCES

**Akhunov, E., Nicolet, C. and Dvorak, J.** (2009) Single nucleotide polymorphism genotyping in polyploid wheat with the Illumina GoldenGate assay. *Theor. Appl. Genet.* **119**, 507–517.

**Brenchley, R., Spannagl, M., Pfeifer, M. *et al.*** (2012) Analysis of the bread wheat genome using whole-genome shotgun sequencing. *Nature*, **491**, 705–710.

**Chapman, J.A., Mascher, M., Buluç, A. *et al.*** (2015) A whole-genome shotgun approach for assembling and anchoring the hexaploid bread wheat genome. *Genome Biol.* **16**, 26.

**Choulet, F., Wicker, T., Rustenholz, C. *et al.*** (2010) Megabase level sequencing reveals contrasted organization and evolution patterns of the wheat gene and transposable element spaces. *Plant Cell*, **22**, 1686–1701.

**Choulet, F., Alberti, A., Theil, S. *et al.*** (2014) Structural and functional partitioning of bread wheat chromosome 3B. *Science*, **345**, 1249721.

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

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

**Elshire, R.J., Glaubitz, J.C., Sun, Q., Poland, J.A., Kawamoto, K., Buckler, E.S. and Mitchell, S.E.** (2011) A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLoS ONE*, **6**, e19379. doi:10.1371/journal.pone.0019379.

**Faris, J.D., Haen, K.M. and Gill, B.S.** (2000) Saturation mapping of a gene-rich recombination hot spot region in wheat. *Genetics*, **154**, 823–835.

**Feuillet, C. and Eversole, K.** (2007) Physical mapping of the wheat genome: a coordinated effort to lay the foundation for genome sequencing and develop tools for breeders. *Isr. J. Plant Sci.* **55**, 307–313.

**Gill, B.S., Appels, R., Botha-Oberholster, A.M. *et al.*** (2004) A workshop report on wheat genome sequencing: International Genome Research on Wheat Consortium. *Genetics*, **168**, 1087–1096.

**de Givry, S., Bouchez, M., Chabrier, P., Milan, D. and Schiex, T.** (2005) CARHTA GENE: multipopulation integrated genetic and radiation hybrid mapping. *Bioinformatics*, **21**, 1703–1704.

**Goss, S.J. and Harris, H.** (1975) New method for mapping genes in human chromosomes. *Nature*, **255**, 680–684.

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

**International Wheat Genome Sequencing Consortium (IWGSC)** (2014) A chromosome-based draft sequence of the hexaploid bread wheat (*Triticum aestivum*) genome. *Science*, **345**, 1251788.

**Jia, J., Zhao, S., Kong, X. *et al.*** (2013) *Aegilops tauschii* draft genome sequence reveals a gene repertoire for wheat adaptation. *Nature*, **496**, 91–95.

**Kalavacharla, V., Hossain, K., Gu, Y., Riera-Lizarazu, O., Vales, M.I., Bhamidimarri, S., Gonzalez-Hernandez, J.L., Maan, S.S. and Kianian, S.F.** (2006) High-resolution radiation hybrid map of wheat chromosome 1D. *Genetics*, **173**, 1089–1099.

**Kumar, A., Seetan, R., Mergoum, M. *et al.*** (2015) Radiation hybrid maps of the D-genome of *Aegilops tauschii* and their application in sequence assembly of large and complex plant genomes. *BMC Genom.* **16**, 800.

**Lazo, G.R., Chao, S., Hummel, D.D. *et al.*** (2004) Development of an expressed sequence tag (EST) resource for wheat (*Triticum aestivum* L.): EST generation, unigene analysis, probe selection and bioinformatics for a 16,000-locus bin-delineated map. *Genetics*, **168**, 585–593.

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

**Ling, H.-Q., Zhao, S., Liu, D. *et al.*** (2013) Draft genome of the wheat A-genome progenitor *Triticum urartu*. *Nature*, **496**, 87–90.

**Luo, M.-C., Gu, Y.Q., You, F.M. *et al.*** (2013) A 4-gigabase physical map unlocks the structure and evolution of the complex genome of *Aegilops tauschii*, the wheat D-genome progenitor. *Proc. Natl Acad. Sci. USA*, **110**, 7940–7945.

**Mascher, M., Muehlbauer, G.J., Rokhsar, D.S. *et al.*** (2013) Anchoring and ordering NGS contig assemblies by population sequencing (POPSEQ). *Plant J.* **76**, 718–727.

**Mayer, K.F.X., Taudien, S., Martis, M. *et al.*** (2009) Gene content and virtual gene order of barley chromosome 1H. *Plant Physiol.* **151**, 496–505.

**Olivier, M., Aggarwal, A., Allen, J. *et al.*** (2001) A high-resolution radiation hybrid map of the human genome draft sequence. *Science*, **291**, 1298–1302.

**Paux, E., Sourdille, P., Salse, J. *et al.*** (2008) A physical map of the 1-gigabase bread wheat chromosome 3B. *Science*, **322**, 101–104.

**Paux, E., Faure, S., Choulet, F. *et al.*** (2010) Insertion site-based polymorphism markers open new perspectives for genome saturation and marker-assisted selection in wheat. *Plant Biotechnol. J.* **8**, 196–210.

**Philippe, R., Paux, E., Bertin, I. *et al.*** (2013) A high density physical map of chromosome 1BL supports evolutionary studies, map-based cloning and sequencing in wheat. *Genome Biol.* **14**, R64.

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

**Qi, L.L., Echalier, B., Chao, S. *et al.*** (2004) A chromosome bin map of 16,000 expressed sequence tag loci and distribution of genes among the three genomes of polyploid wheat. *Genetics*, **168**, 701–712.

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

**Riera-Lizarazu, O., Leonard, J.M., Tiwari, V.K. and Kianian, S.F.** (2010) A method to produce radiation hybrids for the D-genome chromosomes of wheat (*Triticum aestivum* L.). *Cytogenet. Genome Res.* **129**, 234–240.

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

**Sears, E.** (1954) The aneuploids of common wheat. *Mo. Agric. Exp. Stn. Res. Bull.* **52**, 1–58.

**Sears, E. and Sears, M.** (1978) The telocentric chromosomes of common wheat. In 23–28 Feb 1978, 5th International Wheat Genetics Symposium. New Delhi, India, pp. 389–407.

**Sorrells, M.E., Gustafson, J.P., Somers, D. *et al.*** (2011) Reconstruction of the Synthetic W7984 × Opata M85 wheat reference population. *Genome*, **54**, 875–882.

**Tiwari, V.K., Riera-Lizarazu, O., Gunn, H.L., Lopez, K., Iqbal, M.J., Kianian, S.F. and Leonard, J.M.** (2012) Endosperm tolerance of paternal aneuploidy allows radiation hybrid mapping of the wheat D-genome and a measure of γ ray-induced chromosome breaks. *PLoS ONE*, **7**, e48815.

**Wang, S., Wong, D., Forrest, K. *et al.*; International Wheat Genome Sequencing Consortium** (2014) Characterization of polyploid wheat genomic diversity using a high-density 90 000 single nucleotide polymorphism array. *Plant Biotechnol. J.* **12**, 787–796.

**Wardrop, J., Snape, J., Powell, W. and Machray, G.C.** (2002) Constructing plant radiation hybrid panels. *Plant J.* **31**, 223–228.

**Wardrop, J., Fuller, J., Powell, W. and Machray, G.C.** (2004) Exploiting plant somatic radiation hybrids for physical mapping of expressed sequence tags. *Theor. Appl. Genet.* **108**, 343–348.

**Zhang, P., Friebe, B., Lukaszewski, A.J. and Gill, B.S.** (2001) The centromere structure in Robertsonian wheat-rye translocation chromosomes indicates that centric breakage-fusion can occur at different positions within the primary constriction. *Chromosoma*, **110**, 335–344.

 