# A high-density, SNP-based consensus map of tetraploid wheat as a bridge to integrate durum and bread wheat genomics and breeding

Marco Maccaferri[1,*], Andrea Ricci[1], Silvio Salvi[1], Sara Giulia Milner[1], Enrico Noli[1], Pier Luigi Martelli[2], Rita Casadio[2], Eduard Akhunov[3], Simone Scalabrin[4,5], Vera Vendramin[4,5], Karim Ammar[6], Antonio Blanco[7], Francesca Desiderio[8], Assaf Distelfeld[9], Jorge Dubcovsky[10,11], Tzion Fahima[12], Justin Faris[13], Abraham Korol[12], Andrea Massi[14], Anna Maria Mastrangelo[15], Michele Morgante[4,5], Curtis Pozniak[16], Amidou N'Diaye[16], Steven Xu[13] and Roberto Tuberosa[1]

[1]*Department of Agricultural Sciences (DipSA), University of Bologna, Bologna, Italy*

[2]*Biocomputing Group, University of Bologna, Bologna, Italy*

[3]*Department of Plant Pathology, Kansas State University, Manhattan, KS, USA*

[4]*Istituto di Genomica Applicata, Udine, Italy*

[5]*Dipartimento di Scienze Agrarie e Ambientali, University of Udine, Udine, Italy*

[6]*CIMMYT Carretera Mexico, Texcoco, Mexico*

[7]*Dipartimento di Biologia e Chimica Agro-forestale ed ambientale, Università di Bari, Aldo Moro, Bari, Italy*

[8]*Consiglio per la ricerca e la sperimentazione in agricoltura, Genomics Research Centre, Fiorenzuola d'Arda, Italy*

[9]*Faculty of Life Sciences, Department of Molecular Biology and Ecology of Plants, Tel Aviv University, Tel Aviv, Israel*

[10]*Department of Plant Sciences, University of California, Davis, CA, USA*

[11]*Howard Hughes Medical Institute, Chevy Chase, MD, USA*

[12]*Department of Evolutionary and Environmental Biology, Institute of Evolution, Faculty of Science and Science Education, University of Haifa, Haifa, Israel*

[13]*USDA-ARS Cereal Crops Research Unit, Fargo, ND, USA*

[14]*Società Produttori Sementi Bologna (PSB), Argelato, Italy*

[15]*Consiglio per la ricerca e la sperimentazione in agricoltura, Cereal Research Centre, Foggia, Italy*

[16]*Crop Development Centre and Department of Plant Sciences, University of Saskatchewan, Saskatoon, SK, Canada*

## Summary

Consensus linkage maps are important tools in crop genomics. We have assembled a high-density tetraploid wheat consensus map by integrating 13 data sets from independent biparental populations involving durum wheat cultivars (*Triticum turgidum* ssp. *durum*), cultivated emmer (*T. turgidum* ssp. *dicoccum*) and their ancestor (wild emmer, *T. turgidum* ssp. *dicoccoides*). The consensus map harboured 30 144 markers (including 26 626 SNPs and 791 SSRs) half of which were present in at least two component maps. The final map spanned 2631 cM of all 14 durum wheat chromosomes and, differently from the individual component maps, all markers fell within the 14 linkage groups. Marker density per genetic distance unit peaked at centromeric regions, likely due to a combination of low recombination rate in the centromeric regions and even gene distribution along the chromosomes. Comparisons with bread wheat indicated fewer regions with recombination suppression, making this consensus map valuable for mapping in the A and B genomes of both durum and bread wheat. Sequence similarity analysis allowed us to relate mapped gene-derived SNPs to chromosome-specific transcripts. Dense patterns of homeologous relationships have been established between the A- and B-genome maps and between nonsyntenic homeologous chromosome regions as well, the latter tracing to ancient translocation events. The gene-based homeologous relationships are valuable to infer the map location of homeologs of target loci/QTLs. Because most SNP and SSR markers were previously mapped in bread wheat, this consensus map will facilitate a more effective integration and exploitation of genes and QTL for wheat breeding purposes.

## Introduction

Durum wheat [*Triticum turgidum* ssp. *durum* (Desf.) Husn.] is grown worldwide on 17 M ha for a production of 35.6 Mt in 2013 (International Grain Council, IGC Grain Market reports). The main durum wheat product, semolina, is at the base of the pasta industrial chain (13.5 MT per year, IPO 2013).

Cultivated durum wheat is tetraploid and its direct ancestor cultivated emmer (*T. turgidum* ssp. *dicoccum* Schrank) was derived from the domestication of a wild tetraploid species (wild emmer, *T. turgidum* ssp. *dicoccoides* Körn. ex Asch. & Graebner, Schweinf.). The tetraploidization event involved a cross between *T. urartu* (genome AA) with an unknown close relative of *Aegilops speltoides* (genome BB) and occurred <0.8 Mya (Marcussen *et al.*, 2014; Mori, 2003; Ozkan *et al.*, 2005; Salamini

et al., 2002). A further hybridization with *Ae. tauschii* (D genome) gave rise to hexaploid wheat (bread wheat, *Triticum aestivum* L., Kihara, 1944; McFadden and Sears, 1946; Dvorak *et al.*, 1998), The genetic constitution of and relationships among tetraploid and hexaploid wheat species were then shaped by domestication, selection bottlenecks, migrations and extensive gene flow (mainly from tetraploid to hexaploid wheat, Nesbitt and Samuel, 1996; Salamini *et al.*, 2002; Ozkan *et al.*, 2005; Dvorak *et al.*, 1998, 2006; Caldwell *et al.*, 2004; Giles and Brown, 2006). Hence, any useful information accrued in durum wheat genetics is likely to be valuable also to bread wheat breeders.

Until now, durum wheat molecular genetics and genomics have largely relied on hexaploid wheat tools (Feuillet *et al.*, 2012; Gupta *et al.*, 2008; Röder *et al.*, 1998; Varshney *et al.*, 2007). Molecular markers (RFLP, SSR, AFLP, DArT®, etc.) developed for hexaploid wheat are generally useful for genetic mapping experiments in tetraploid wheat based on the sharing of the two A and B genomes (Blanco *et al.*, 1998; Korzun *et al.*, 1999; Lotti *et al.* 2000; Mantovani *et al.*, 2008). However, markers' polymorphism information content in durum wheat germplasm differs from that observed in hexaploid wheat (Maccaferri *et al.*, 2003, 2005). Single nucleotide polymorphisms (SNPs) have become the most widely utilized markers and are being extensively developed in crops, including wheat (Allen *et al.*, 2013; Cavanagh *et al.*, 2013; Ganal *et al.*, 2014; Lai *et al.*, 2012; van Poecke *et al.*, 2013; Tiwari *et al.*, 2014; Trebbi *et al.*, 2011). A high-density (90K) wheat SNP array has been recently developed (Wang *et al.*, 2014), mostly based on bread wheat SNPs, but including approx. 8000 SNPs from durum cultivars.

Linkage maps are crucial tools for the identification and cloning of genes and quantitative trait loci (QTL), marker-assisted selection and genome structure analysis. The inherent constraints of linkage maps based on single populations can be overcome by the production of consensus maps of several maps and independent crosses. Consensus maps provide large number of markers, extend the mapped genome portions (including regions identical by descent), overcome local loss of genetic resolution, validate marker order, identify chromosome rearrangements and increase the effectiveness of genome-wide association studies and QTL meta-analyses (Salvi *et al.*, 2010; Swamy *et al.*, 2011).

Integration of mapping information from individual mapping populations has been performed using various approaches (Ronin *et al.*, 2012). Currently, the *genetic merging* method has been used for medium-density (up to 5000 markers) projects (Wang *et al.*, 2011). The *graph*-based method has been used for high-density (of more than 10 000 loci) consensus maps (Gautami *et al.*, 2012; Wang *et al.*, 2014). The *regression* method is also suitable for high-density maps but requires high-quality initial maps (Blenda *et al.*, 2012; Hudson *et al.*, 2012).

In hexaploid wheat, a SSR-based consensus map by Somers *et al.* (2004) has been widely used. In durum wheat, two consensus maps mainly based on SSR and DArT® markers have been developed (Marone *et al.*, 2012; Maccaferri *et al.*,). SNP-based consensus maps are quickly replacing or updating former maps (Bowers *et al.*, 2012; Hyten *et al.*, 2010; Muñoz-Amatriaín *et al.* 2012; Goretti *et al.*, 2014; van Poecke *et al.*, 2013). In bread wheat, six and eight biparental cross-populations were used to build SNP-based consensus maps (Cavanagh *et al.*, 2013; Wang *et al.*, 2014).

In this study, 13 independent tetraploid wheat mapping populations were utilized to assemble a durum wheat consensus map with the highest number of markers (>30 000) to date. These populations included elite × elite, elite × cultivated emmer and elite × wild emmer. Ten of these populations were genotyped with the recently developed Illumina 9K and 90K wheat SNP arrays (Cavanagh *et al.*, 2013; Wang *et al.*, 2014). Informativeness and potential applications of this consensus map were analysed and discussed in terms of genome coverage and marker density, diversity and distribution.

## Results

### Linkage maps from the individual populations (component maps)

The genotypic data of 13 tetraploid wheat (AABB; $2n = 28$) mapping populations were used for assembling the consensus map. In total, the populations included 1928 lines (1773 RILs and 155 DHs, Table 1). The elite × elite populations that were genotyped with the Illumina 90K SNP array (six in total) allowed us to map between 3676 and 6163 SNPs with a mean of 5175 SNPs and a mean map length of 2012 cM. The two durum × cultivated emmer populations allowed us to map 8952 and 10 811 SNPs (3028 and 2363 cM, respectively) while the Svevo × Zavitan (durum × wild emmer) population included 10 911 SNPs (2258 cM).

Despite the high density of polymorphic markers detected using the 90K SNP assay, none of the six elite × elite populations produced component maps with a number of linkage groups equal to the nominal chromosome pair number (14). Rather, linkage groups ranged from 20 to 33 (28 on average) with an average density of 2.2–3.0 SNPs/cM and an average SNP intermarker distance of 0.39 cM/SNP (from 0.33 in Cl × Ld to 0.46 cM/SNP in G9586). In contrast, the component maps derived from the three cultivated emmer and wild emmer-derived maps genotyped with the Illumina 90K SNPs showed 14–15 linkage groups with a more even genome coverage as compared to any map from cultivated durum wheat. However, the increase in marker density of the SNPs mapped in the polymorphic regions (0.23–0.34 cM/SNP for the emmer-derived maps and 0.20 cM/SNP for the wild emmer-derived map) was only slightly higher as compared to the elite × elite populations.

### Consensus map features

The genotypic data of the 1773 RILs were used to construct a framework consensus map comprised of 16 384 common markers (hereafter referred to as anchor markers, 54.3% of the total). In particular, anchors included 15 827 Illumina wheat SNPs, 297 SSRs, 242 DArT® markers and 18 SNPs/STS markers of diverse origin. The anchor markers were grouped into 14 linkage groups, corresponding to the tetraploid wheat chromosomes (Table 2), and were mapped at a mean LOD score = 61.9. The framework map was used to infer the position of the 13 760 markers uniquely mapped in single component populations, leading to a final consensus map including 30 144 markers. This map was based on 5279 unique recombination events (377 per chromosome) and spanned 2631 cM for an average intermarker distance of 0.087 cM, corresponding to 11 markers per cM. The consensus map was thus 30.7 and 17.5% longer than the mean map length of the six component elite × elite durum populations genotyped with the Illumina 90K SNP array and the longest one (Mr × Cd; 2239 cM), respectively. However, the consensus map was very similar in length to those of the three durum × cultivated emmer maps (2631 *vs.* 2639 cM) and slightly longer (14.2%) than the durum × wild emmer Svevo × Zavitan map.

**Table 1** Details of the 13 mapping populations and of the corresponding single component maps

| Mapping populations | | | | | Molecular markers | | | | | Linkage group | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Parents | Acronym | Contributing Institution* | Type DH/RIL | Size no. | Genomic SSR no. | DArT® no. | Illumina SNP† no. | Others‡ no. | Total no. | Linkage group no. | Total length cM | Intermarker distance§ cM/marker |
| *T. durum* × *T. durum* | | | | | | | | | | | | |
| Colosseo × Lloyd¶ | Cl × Ld | UNIBO/UNIUD/PSB | RIL | 176 | 184 | 372 | 6163 | 1227 | 7946 | 20 | 2063.9 | 0.33 |
| Meridiano × Claudio¶ | Mr × Cd | UNIBO/UNIUD/PSB | RIL | 181 | 178 | 608 | 5097 | 87 | 5970 | 27 | 2238.8 | 0.43 |
| Simeto × Levante¶ | Sm × Lv | UNIBO/UNIUD/PSB | RIL | 180 | 142 | 335 | 5315 | 6 | 5798 | 30 | 2184.7 | 0.40 |
| Mohawk × Cocorit69 | Mh × Cr | CIMMYT/USASK | RIL | 81 | – | – | 5554 | – | 5554 | 31 | 2012.7 | 0.36 |
| Svevo × Ciccio¶ | Sv × Cc | UNIBA | RIL | 103 | 16 | 213 | 5246 | 12 | 5487 | 26 | 1887.6 | 0.36 |
| W9292-260D3 × Kofa | G9586 | AAFC/USASK | DH | 155 | 34 | – | 3676 | 2 | 3712 | 33 | 1685.0 | 0.46 |
| Kofa × Svevo¶ | Kf × Sv | UNIBO/UNIUD/PSB | RIL | 249 | 205 | – | – | 38 | 243 | 18 | 1256.2 | – |
| Kofa × UC1113¶ | Kf × UC | UCDAVIS | RIL | 93 | 172 | – | – | 31 | 203 | 24 | 755.1 | – |
| *T. durum* × *T. dicoccum* | | | | | | | | | | | | |
| Ben × PI41025 | Bn × PI_41025 | USDA-ARS | RIL | 200 | 111 | – | 2456 | – | 2567 | 14 | 2526.9 | – |
| Simeto × Molise Colli | Sm × Ml | CRA-Foggia | RIL | 136 | 26 | – | 8926 | – | 8952 | 15 | 3028.4 | 0.34 |
| Latino × MG5323 | Lt × MG_5323 | CRA-Fiorenzuola/UNIBA | RIL | 82 | 216 | – | 10 572 | 23 | 10 811 | 14 | 2363.4 | 0.23 |
| *T. durum* × *T. dicoccoides* | | | | | | | | | | | | |
| Langdon × G18-16¶ | Ln × G18-16 | UHAIFA | RIL | 152 | 120 | 148 | – | – | 268 | 20 | 1577.3 | – |
| Svevo × Zavitan | Sv × Zv | UTELAVIV | RIL | 140 | – | – | 10 911 | – | 10 911 | 14 | 2258.0 | 0.20 |

RIL, recombinant inbred line; DH, double haploid; SSR, simple sequence repeat; DArT®, Diversity Array Technology; SNP, single nucleotide polymorphism.

*UNIBO, University of Bologna; UNIUD, University of Udine; PSB, Produttori Sementi Bologna; CIMMYT, International Maize and Wheat Improvement Center; USASK, University of Saskatchewan; UNIBA, University of Bari; AAFC, Agriculture and Agri-Food Canada; USDA-ARS, Cereal Crop Research Unit, Fargo; CRA, Consiglio per la Ricerca e la Sperimentazione in Agricoltura; UHAIFA, University of Haifa; UTELAVIV, Tel Aviv University.

†Illumina iSelect 90K wheat SNP array used for 9 populations; Illumina iSelect 9K wheat SNP array used for Bn × PI41025.

‡Include sequence tagged sites, morphological and biochemical markers, and 1065 sequence-based genotyping SNPs for Cl × Ld.

§Referred to the Illumina iSelect 90K wheat SNPs only.

¶Details of mapping populations in Materials and Methods.

**Table 2** Features of the consensus framework map summarized by chromosome

| Chromosome | Illumina SNP no. | SSR no. | DArT no. | Others (SNP, STS) no. | Total no. | Common markers no. | Singleton no. | Informative maps no. | Recombination bins no. | LOD LOD units | Length cM | Intermarker distance cM/marker |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1A | 1705 | 44 | 78 | 35 | 1862 | 975 | 887 | 3.0 ± 1.0 | 336 | 78.1 ± 45.3 | 171.6 | 0.09 |
| 1B | 2450 | 69 | 105 | 180 | 2804 | 1481 | 1323 | 2.9 ± 1.0 | 398 | 45.1 ± 32.6 | 178.0 | 0.06 |
| 2A | 1860 | 63 | 72 | 106 | 2101 | 1123 | 978 | 3.2 ± 1.3 | 320 | 67.4 ± 39.1 | 215.5 | 0.10 |
| 2B | 2978 | 85 | 121 | 140 | 3324 | 1789 | 1535 | 3.1 ± 1.1 | 493 | 60.1 ± 38.5 | 199.8 | 0.06 |
| 3A | 1472 | 60 | 74 | 92 | 1698 | 964 | 734 | 3.3 ± 1.3 | 326 | 68.6 ± 47.6 | 184.7 | 0.10 |
| 3B | 2091 | 60 | 150 | 122 | 2423 | 1183 | 1240 | 2.9 ± 1.1 | 470 | 55.2 ± 36.1 | 212.5 | 0.09 |
| 4A | 1346 | 44 | 108 | 65 | 1563 | 813 | 750 | 3.0 ± 1.2 | 320 | 61.7 ± 40.4 | 176.5 | 0.11 |
| 4B | 1254 | 39 | 56 | 77 | 1426 | 924 | 502 | 3.8 ± 1.3 | 307 | 71.7 ± 41.9 | 136.5 | 0.10 |
| 5A | 1442 | 70 | 18 | 78 | 1608 | 860 | 748 | 2.9 ± 1.1 | 377 | 58.1 ± 30.4 | 218.6 | 0.14 |
| 5B | 2140 | 65 | 65 | 53 | 2323 | 1279 | 1044 | 3.4 ± 1.3 | 435 | 72.9 ± 40.7 | 206.2 | 0.09 |
| 6A | 1635 | 27 | 92 | 106 | 1860 | 978 | 882 | 3.0 ± 1.1 | 303 | 63.7 ± 36.1 | 135.1 | 0.07 |
| 6B | 2093 | 46 | 132 | 133 | 2404 | 1337 | 1067 | 3.2 ± 1.2 | 367 | 50.7 ± 42.9 | 156.0 | 0.07 |
| 7A | 2084 | 49 | 108 | 92 | 2333 | 1314 | 1019 | 2.3 ± 1.4 | 418 | 69.8 ± 38.9 | 214.0 | 0.09 |
| 7B | 2076 | 70 | 145 | 124 | 2415 | 1364 | 1051 | 4.0 ± 1.2 | 409 | 43.5 ± 32.9 | 226.3 | 0.09 |
| Mean | 1901 | 57 | 95 | 100 | 2153 | 1170 | 982 | 3.1 ± 1.2 | 377 | 61.9 ± 38.8 | 187.9 | 0.09 |
| Total | 26 626 | 791 | 1324 | 1403 | 30 144 | 16 384 | 13 760 | – | 5279 | – | 2631.3 | – |
| Mean A genome | 1649 | 51 | 79 | 82 | 1861 | 1004 | 857 | – | 343 | – | 188.0 | 0.10 |
| Mean B genome | 2155 | 62 | 111 | 118 | 2446 | 1337 | 1109 | – | 411 | – | 187.9 | 0.08 |
| Total A genome | 11 544 | 357 | 550 | 574 | 13 025 | 7027 | 5998 | – | 2400 | – | 1316 | – |
| Total B genome | 15 082 | 434 | 774 | 829 | 17 119 | 9357 | 7762 | – | 2879 | – | 1315 | – |

SSR, simple sequence repeat; DArT®, Diversity Array Technology; SNP, single nucleotide polymorphism; LOD, logarithm of the odds.
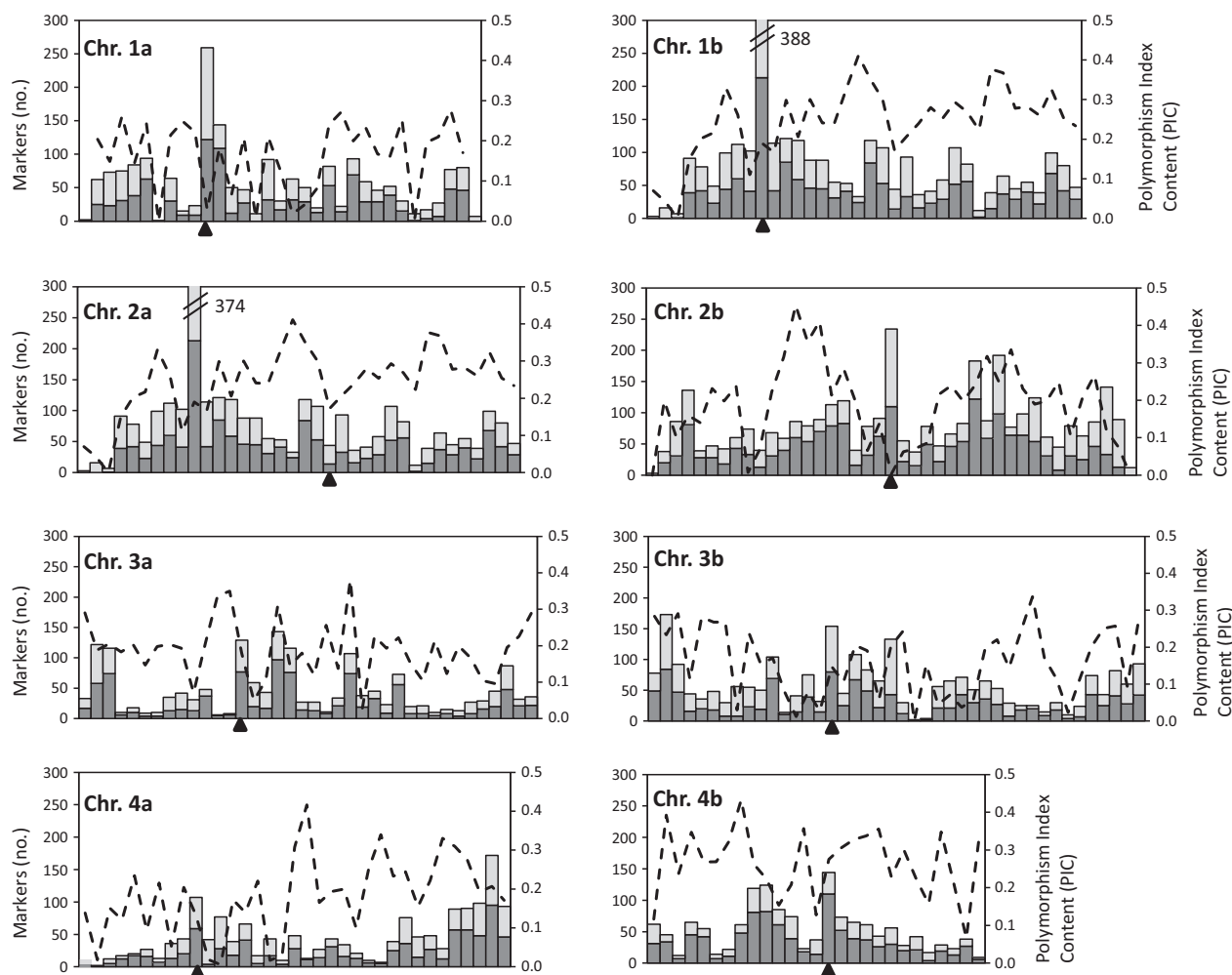
**Figure 1** Distribution of markers on the consensus map. Markers were counted based on 5-cM bins. Anchor and singleton markers are reported separately. Mean Polymorphism Index Content (PIC) is from 55 of elite durum accessions from worldwide. Putative centromere positions are reported.

Considering only materials mapped with the Illumina 90K SNP array, the elite × elite populations allowed us to map 5055 anchor markers and 1064 singletons, whereas the cultivated and wild emmer-derived populations were effective in mapping an additional 7932 anchor SNPs and 2289 singletons. On average, each anchor marker was mapped approximately across 3.0 mapping populations (2.9–3.4). The number of anchor markers and total map length for each chromosome are reported in Table 2. Anchor markers mapped in the elite × elite populations were more informative in terms of the overall number of segregating biparental populations per SNP as compared to the anchors mapped to the interspecific crosses, as expected and most probably due to ascertainment bias effects and divergence of allele frequency between durum and emmer germplasm. Distribution of markers for all maps is reported in details in Tables S1 and S2.

In the consensus map, the A and B genomes had nearly identical map length (1316 and 1315 cM, corresponding to 188.0 and 187.9 cM per chromosome, respectively). However, a greater number of markers mapped to the B genome (17 119) *vs*. the A genome (13 025). This proportion was conserved across all marker classes, and average intermarker distance was equal to 0.079 cM/marker for the B genome and 0.101 cM/marker for the A genome.

For most chromosomes, previously indicated centromere positions (Somers *et al.*, 2004; Sourdille *et al.*, 2004) were coincident with clear spikes in marker density located within a single 5-cM bin (Figure 1), with the exception of chromosome 6A. Conversely, low-marker density regions spanning 10–30 cM were detected in chromosome arms 1AS, 2AS, 3AS, 4AS (proximal regions) and the distal regions of 7AS and 7BS. Although the patterns of PIC value distribution fluctuated widely (Figure 1), the diversity of the Illumina SNPs in the elite durum germplasm could be related to marker distribution and/or chromosome features. A decrease in diversity was observed in the centromeric and pericentromeric regions of several chromosomes (e.g. 1A, 1B, 2A, 2B, 3B, 5B, 6A, 7B) despite the relatively high-marker density of these regions. Conversely, extended regions characterized by low-marker density and diversity were also observed (e.g. in chromosome arms 1BS, 2AS, 4AS, 5AS).

The consensus map contained both transcript-derived SNP markers and nongenic SSR and DArT markers (considered together with SNPs obtained from reduced-representation genomic libraries). The relative distribution of these three marker classes along chromosomes is represented in Figure 2 as standardized relative marker ratio (proportion) per 5-cM bin. Transcript-derived SNP markers showed a tendency towards enrichment in the pericentromeric and centromeric regions (e.g.
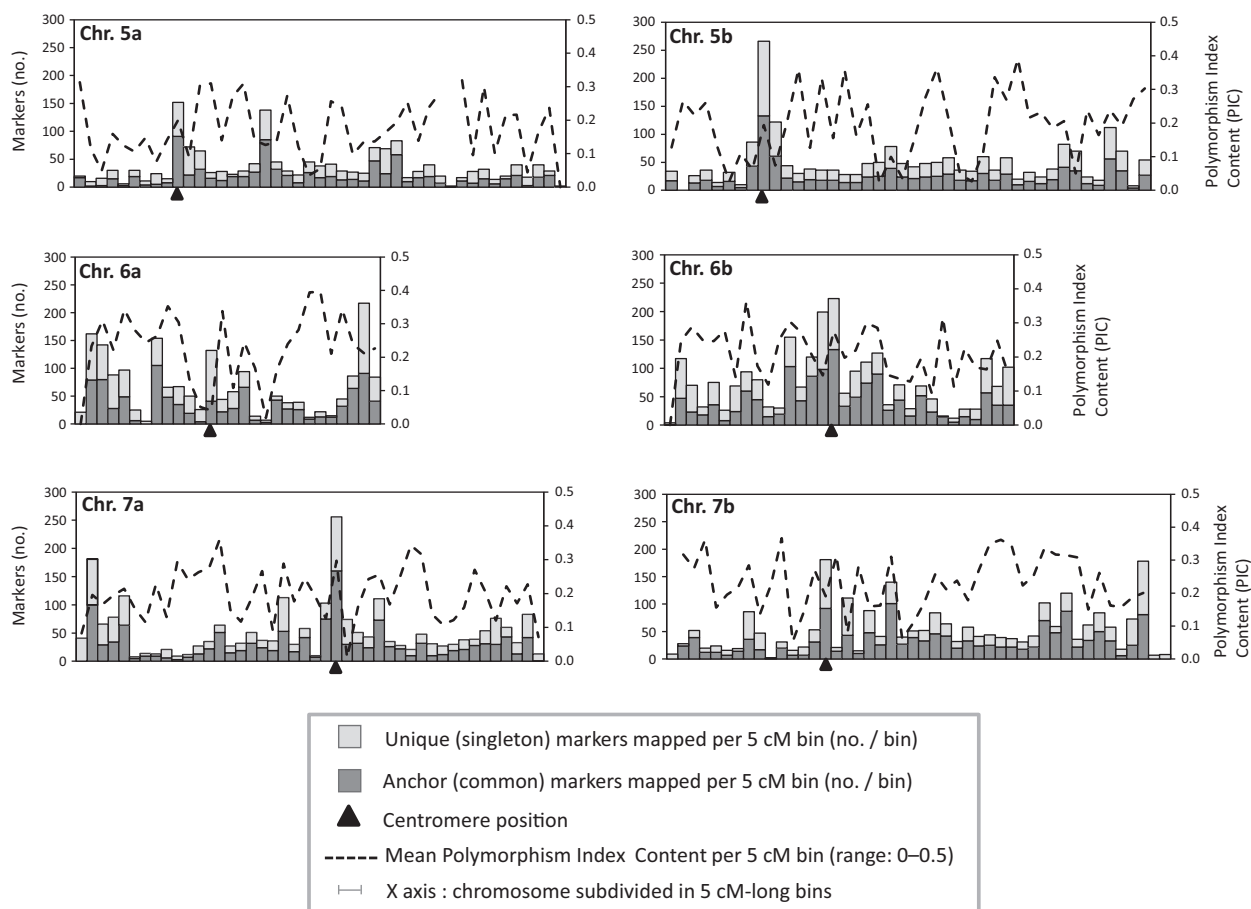
Figure 1 Continued.

in chromosomes 1A, 2A, 3A, 3B, 4A, 6B, 7A). Additionally, their distribution was not regular along chromosomes, with local peaks of high and low density. DArT markers and genomic SNPs obtained from reduced-representation libraries derived from methylation-sensitive restriction enzymes (*Pst*I: KBO and SBG markers) were clearly over-represented in the distal euchromatic regions and under-represented in the centromeric and pericentromeric chromosome regions (enriched in heterochromatin) for most of the chromosomes. The mapped SSR markers, usually chosen based on the previous knowledge on their mapping position, showed a tendency towards even distribution along the chromosomes.

In the elite × elite maps, the marker density differed markedly along the linkage groups, with extensive presence of low-marker density regions and gaps of no marker coverage scattered along chromosomes (Figures S1 and 3, which is magnified in Figure S2). This feature was attenuated in maps involving emmer and wild emmer crosses.

As to the degree of marker order correspondence between the consensus and the component maps, of the 182 pairwise LG comparisons, 147 (80.7%) showed ρ values >0.95 and 56 (30.8%) showed ρ values >0.99 (Figure S1). Conflicts in marker order were mainly limited to a single map among those sharing common markers. Most conflicts involved small distances and the two alternative orders presented nearly equal LOD scores; therefore, the consensus marker order was retained, and the conflicts were not declared. Declared conflicts were detected in 7.4% of markers per chromosome, and the conflict order

information was retained in the consensus map by marker renaming. A small fraction of conflicts involved large structural rearrangements of several linked markers (e.g. on chromosome 3B, data not reported) while the majority was limited to distances <10 cM.

## Genetic-to-physical correspondence between tetraploid and hexaploid chromosome 3B

A total of 1886 gene-derived SNPs genetically mapped on the chromosome 3B tetraploid consensus had significant matches to the 774 Mb pseudomolecule sequence of hexaploid wheat (Choulet *et al.*, 2014). The majority of markers anchored to the pseudomolecule, following a close-to sigmoidal pattern (Figure 4). However, 193 SNPs (10.2%) had significant matches but anchored to noncollinear regions and their positions were scattered throughout the pseudomolecule without any clear pattern. Only 40% of the 193 noncollinear SNPs were univocally positioned on the pseudomolecule, while the remaining 60% showed multiple BLAST hits (from 2 to 10, based on $e^{-10}$ filtering), possibly due to occurrence in duplicated/multiple gene family members. The localization of noncollinear SNPs that were univocally positioned on the hexaploid pseudomolecule supports the occurrence of rearrangements of gene order between the tetraploid wheat parental genotypes and hexaploid wheat.

The genetic-to-physical relationship for chromosome 3B can be described according to subpatterns at five main segmented regions (two distal regions, two proximal regions and the centromeric/pericentromeric region). In the two distal regions
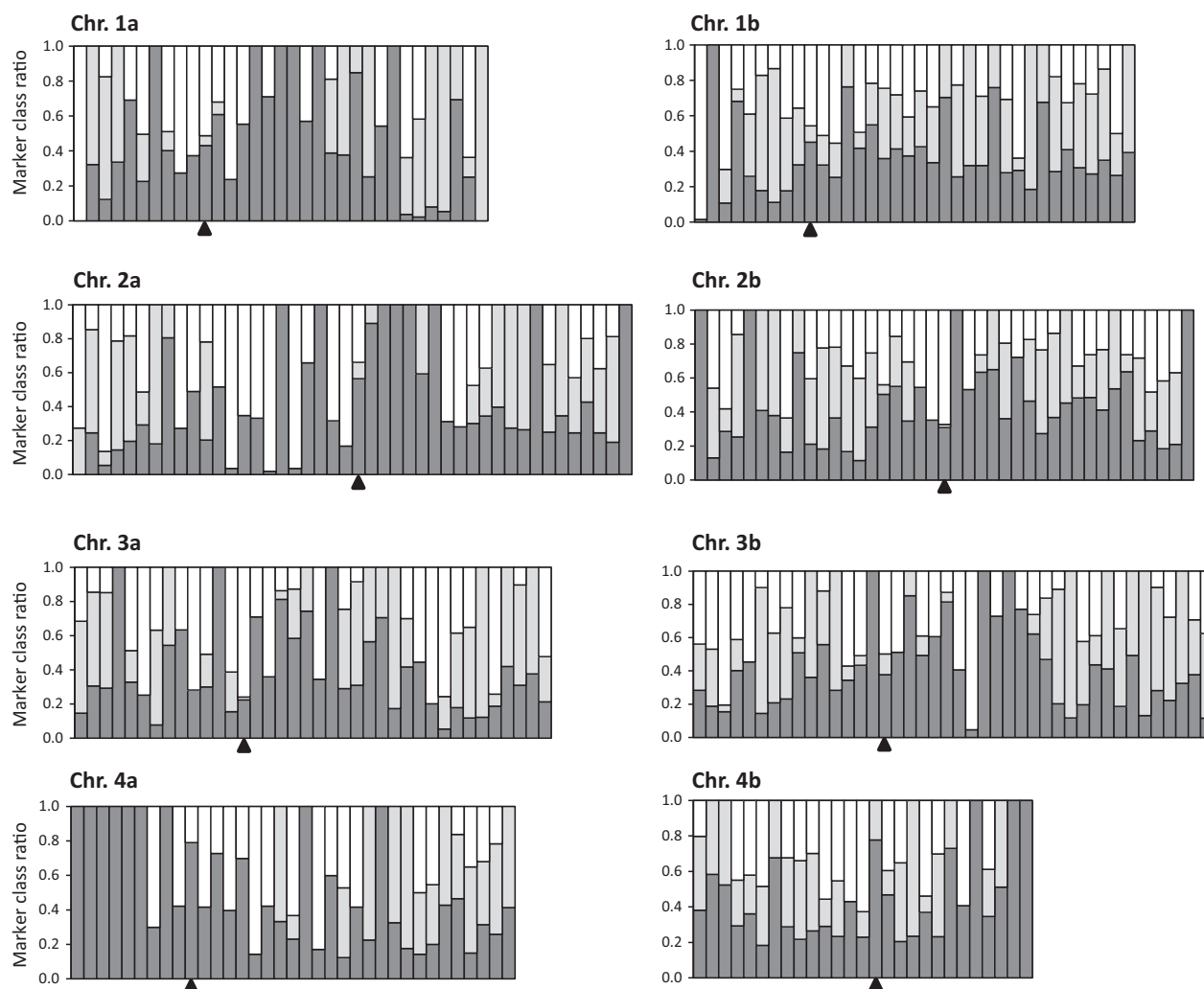
**Figure 2** Relative frequency distribution of genic and nongenic markers on the consensus map. Transcript-associated SNPs, nongenic DArT and SNP markers, nongenic SSR and STS were considered as the main three marker classes. Relative frequencies are reported along chromosomes based on 5-cM bins.

corresponding to 0–61.6 cM in 3BS and 133.4–210.8 cM in 3BL, the physical-to-genetic distance relationship was mostly linear as from the regression results (Figure 4a). These two distal regions extended over ca. 95 (distal 3BS) and 127 Mb (distal 3BL), that accounted for 12.2 and 16.4% of the whole chromosome, and were characterized by an overall recombination rate of 0.64 and 0.61 cM/Mb, respectively. In the two proximal regions, the meiotic recombination rate dropped to 0.15 cM/Mb. The centromeric region was positioned in a 1.4 cM interval (75.2–76.8 cM) extending over 127 Mb (16.4% of the chromosome) and was characterized by strong suppression of recombination. SNP PIC values plotted using a 30-markers sliding window showed higher polymorphism level in the two distal regions than in the rest of the chromosome, apart from localized regions (Figure 4a).

The comparison between the hexaploid (Wang *et al.*, 2014) and the tetraploid consensus maps for chromosome 3B using the SNPs shared between the two maps is reported in Figure 4b. Spearman rank correlation between the two maps (ρ value) was equal to 0.95, thus showing a good collinearity conservation between the two subspecies. The tetraploid wheat map showed a higher homogeneity in recombination event distribution as compared to the hexaploid map, with the latter showing a

markedly lower recombination rate throughout the centromeric–pericentromeric region (as shown in Figure 4b). Despite the overall good correspondence, the collinearity at a finer level of resolution (down to 5 cM) was compromised, mostly due to the low resolution of both maps. When comparing the collinearity (marker order) of the two consensus maps to the 3B pseudo-molecule, the two genetic maps performed similarly based on rank correlation ρ values = 0.89 and 0.84 for the tetraploid and hexaploid map, respectively.

### Homeologous relationships between the A and B genome explored by gene-derived SNPs

To precisely define the homeologous (and paralogous) relationships between the coding sequences tagged by mapped SNPs, the latter were related to the *T. aestivum* genome-specific transcripts. Based on a conserved-sequence similarity search (BLASTN, 1e$^{-10}$), it was possible to relate 21 547 gene-associated SNPs to 29 520 A-, B- and D-genome transcripts with 78 311 hits in total, restricted to 20 045 unique transcripts for the A and B genomes only (53 316 hits). High similarity matches were thus found for 2.66 SNPs per transcript, on average. Unique transcript-SNP relationships (hits) were observed for 9276 cases (46.3% of
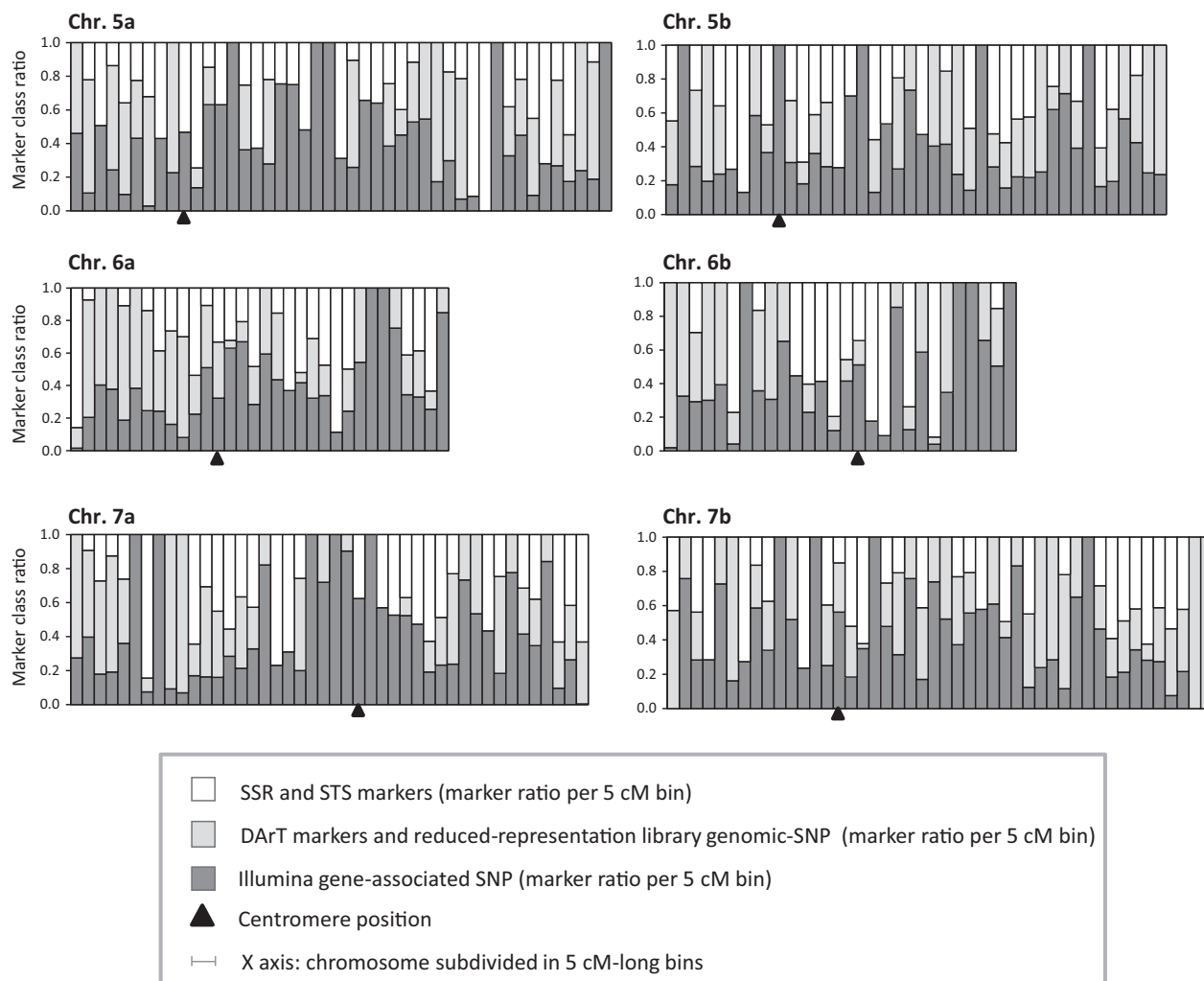
**Figure 2** Continued.

coding sequences) while 8680 (43.3%) transcripts were tagged by 2 to 5 SNP hits (including SNPs from homeologous chromosome regions); the remaining 8.04 and 2.37% were tagged by multiple SNPs (up to 10 and up to 26 tagged SNP, respectively).

The conserved-sequence similarity analysis highlighted the homeologous and paralogous relationships between gene-associated SNPs and transcripts mapped to syntenic regions of homeologous pairs, as well as to other chromosomes. Genome-specific transcripts-SNP hits are reported as Tables S3 and S4. As detailed in Table S3, among the 53 316 genome-specific transcript-SNP hits, 22 579 (42.3%) and 18 372 (34.4%) were traced to direct homologous and indirect homeologous relationships within the seven homeologous groups, respectively (76.7% of the total number of matches). SNP-transcript matches between putative homeologous chromosomes are reported in Table S5. A mean of 558 nonredundant reference Chinese Spring transcripts per homeologous group was tagged by SNP clusters whose components were mapped to homeologous/paralogous positions, for a mean of 1254 SNPs in total (2.25 SNPs per transcript). In homeologous groups 1, 2 and 3, the homeologous relationships were abundant, with 627, 818 and 813 nonredundant transcripts and 1516, 1937 and 1330 nonredundant SNPs, respectively. Chromosome groups 4, 5 and 7 showed the lowest

number of nonredundant transcripts tagged by SNPs in putative homeologous positions, with 130, 426 and 519 reference transcripts and 370, 933 and 1408 SNPs, respectively. This was expected based on the multiple translocation events occurred among these chromosomes. Complex, nonsyntenic homeologous relationships among chromosomes 4A, 5A and 7B, tagging the multiple known translocations involving chromosome arms 4AL, 4AS, 5AL and 7BS (T4A/5A and T7B/4A), were identified. In addition, preliminary evidences of additional putative translocations were obtained. Relationships are detailed in Text S1 and Table S6.

## Comparison between the tetraploid and hexaploid wheat consensus maps

The tetraploid consensus map was compared with the hexaploid Illumina 90K SNP consensus map reported by Wang *et al.* (2014). Some 21 145 SNPs were in common between the two maps (79.4% of the SNP mapped in tetraploid wheat), with 5481 SNPs mapped exclusively in the tetraploid consensus map. Among the common SNPs, 19 202 (90.8%) showed agreement in chromosome assignment between the two maps, whereas 1268 (6.0%) were discordant in assignment to the nominal homeologs and the remaining 675 (3.2%) were assigned to different chromosome
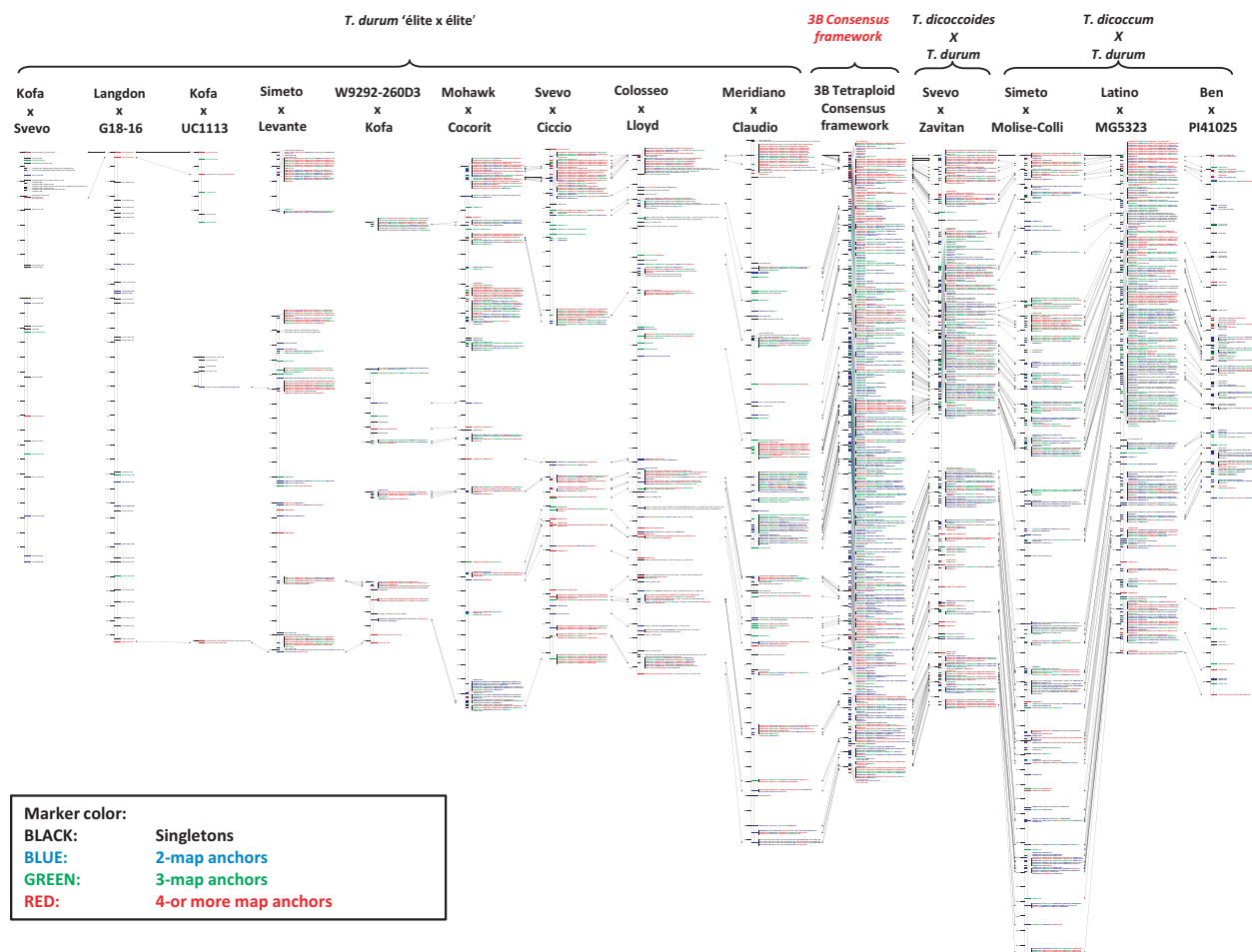
**Figure 3** Single component and consensus map representations for chromosome 3B. The consensus framework map is central. Elite × elite durum maps are reported on the left while durum × cultivated and wild emmer maps on the right. Singleton markers are in black and anchor markers for two, three and four or more maps are reported in blue, green and red, respectively. Connectors join highly informative marker positions. For sake of clarity, unique singleton markers are not reported in the tetraploid consensus.

groups, in balanced frequencies across chromosome groups with the notable exceptions of the few main translocations. One hundred and sixteen SNPs were differentially located to the homeologous 4A-7A and 4A-7D positions in the two maps, thus tagging alternate homeologous sites (tracing to the T4AL/7BS) depending on the maps. Additionally, 65 SNPs showed discordant location to groups 4 and 5 (tracing to T4AL/5AL, with 18 SNPs mapped in the distal 5BL (170.7–206.2 cM interval) on the tetraploid consensus and in the 4AL distal in the hexaploid map, and 13 SNPs mapped in opposite relationship. Additionally, traces of a putative T5A/2B translocation were detected by 15 SNPs differentially mapped in chromosome 2B (tetraploid map, in 75.4–158.3 cM interval)—5B (hexaploid map) or in opposite relationship.

Overall, 9.2% of Illumina SNP assays tagged different homeologous or highly similar loci between durum and bread wheat, a finding already observed by Wang *et al.* (2014) that poses the problem of uniqueness (i.e. adherence to Mendelian inheritance) of the SNP assays. Wang demonstrated that most of these assays detected at least two polymorphisms (target SNPs plus interfering SNPs at the genotyping oligo-annealing site). The precise causes behind this finding remain to be identified. As to the single component populations, it appeared

that most of them held the ancient translocations, with the only notable exception for the 7A-1 linkage group in W9292-260D3 × Kofa, where a notable portion (10.8%) of linked SNPs was mapped in chromosome 4A in other populations and in the consensus map.

The Spearman rank correlations between homologous chromosomes of the two consensus maps ranged from 0.90 to 0.99. Rank-order plots of data pairs from the two consensus (see Figure S3) evidenced extensive regions with completely suppressed recombination in the hexaploid wheat consensus, particularly in the pericentromeric regions of chrs. 1A, 2A, 3A, 4A, 4B, 5A, 6A, 7A, similarly to what we observed for the 3B pericentromeric chromosome region where only the recombination rate of the pericentromeric region was lower as compared to the tetraploid wheat map.

## Discussion

### A tetraploid-specific consensus map as a reference for the A and B genomes of wheat

The consensus mapping data set reported herein builds on a sizeable amount of genotyping data obtained for tetraploid wheat in the last decade. The consensus map integrates data
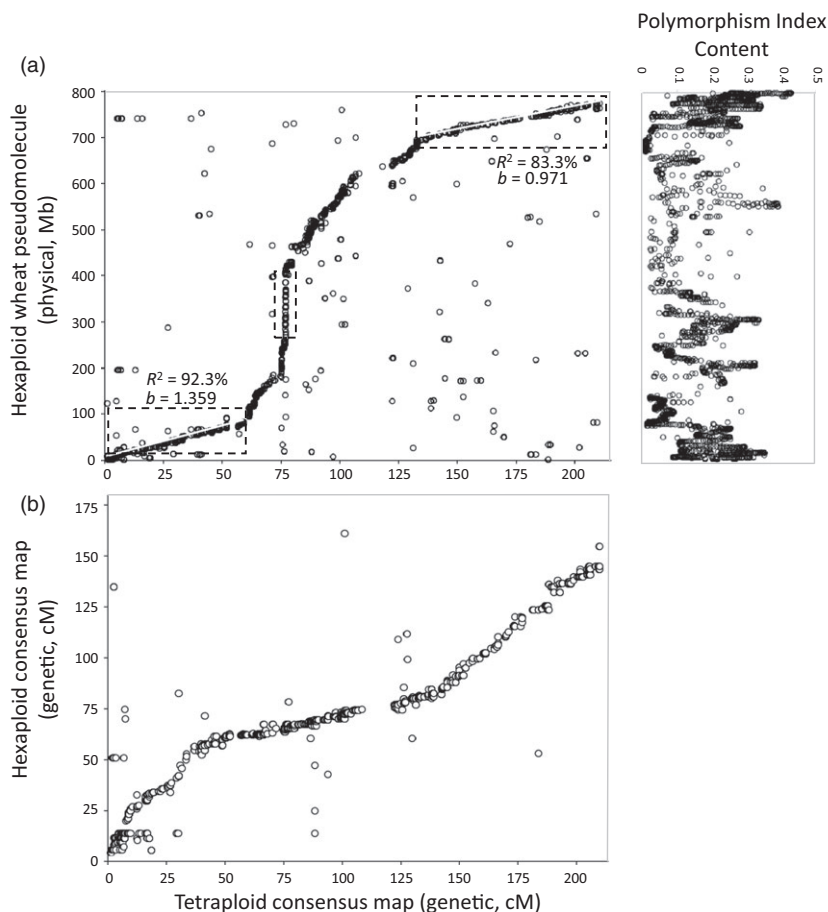
**Figure 4** Physical-to-genetic and genetic-to-genetic relationships between the hexaploid physical map, the tetraploid and the hexaploid consensus genetic maps for chromosome 3B. (a) Plot of the hexaploid physical map ($Y$ axis, Mb) on the tetraploid consensus genetic map ($X$ axis, cM). Correspondence has been established by conserved-sequence similarity search (BLASTN). The two distal chromosome regions with linear relationships between physical and genetic distances and the centromeric/pericentromeric region are indicated by dashed-lined rectangular boxes. Polymorphism Index Content (PIC) of mapped markers from 55 elite durum wheat cultivars is plotted along the physical maps. (b) Plot of the hexaploid consensus genetic map ($Y$ axis, cM) on the tetraploid consensus genetic map ($X$ axis, cM).

from mapping populations obtained from crosses between the cultivated durum varieties and crosses involving the closest ancestors domesticated emmer and wild emmer. This study integrates in a unique framework different types of markers (mostly SSR, DArT® and SNP), hence allowing one to trace the results of genes and QTL mapping studies and genetic diversity analysis in both tetraploid and hexaploid wheat for the A and B genomes. Additionally, our consensus map integrates a large number (26 626) of SNP markers from the recently released wheat 90K SNP array that includes ca. 90 000 gene-associated assays (Wang *et al.*, 2014). This provides a unique opportunity to switch from neutral markers to a large set of functional, physically anchored markers and cross-reference gene/QTL map information and other chromosome and genomic features between bread and durum wheat, thus allows for a highly accurate comparative genotyping in both cereals.

Our results show the utility of the 90K SNP array for map construction providing nearly complete genome coverage in *T. dicoccum*- and *T. dicoccoides*-derived populations compared to elite × elite populations. Gaps in the linkage maps from the elite × elite durum wheat crosses suggest that the genomes of the parental durum wheat genotypes share extensive regions of identity by descent (IBD), difficult to populate with informative markers. The presence of shared long-range haplotypes has been previously demonstrated in elite durum wheat using mapped multi-allelic SSR markers (Maccaferri *et al.*, 2005, 2007). As the IBD regions are randomly scattered across the individual maps, the consensus map approach provided the opportunity to fill linkage gaps and to reconstruct a map with nearly complete

genome coverage. It should be noted, however, that the adoption of an open genotyping technology (such as genotyping by sequencing, Saintenac *et al.*, 2013) could probably partially mitigate the low polymorphism observed in the elite × elite crosses.

### Consensus map construction and map features

To produce a robust consensus map, we used a combination of *genetic merging* for framework mapping, based on common (=anchor) markers, and *interpolation* for singleton markers. Advantages from both mapping methods were thus retained, such as the ability of (i) producing a robust map of anchor markers with nearly complete genome coverage and recombination rate estimates considered as representative for the tetraploid germplasm, and (ii) adding unique markers to a reliable framework using a fast *regression*-based method (Cone *et al.*, 2002; Sosnowski *et al.*, 2012).

*Graph*-based methods are faster alternatives. However, they do not rely on the original set of marker distance matrices (Ronin *et al.*, 2012). A further approach based on the search of best multilocus ordering from consensus analysis of the component data sets was provided by Ronin *et al.* (2012). A more efficient use of anchor loci is expected to speed up such analysis (A. Korol, personal communication).

The *genetic merging* method was chosen based on the availability of genotypic data from 1773 RIL, with 54% of anchor markers, a number sufficiently large to provide sufficiently large to provide genetic distances representative of the distribution of recombination rates in tetraploid wheat. Second, the mapping

populations compared to hexaploid wheat showed a remarkable homogeneity in terms of marker order and recombination rate. In bread wheat, the genomes of modern lines frequently include alien chromosomes transferred from wild relatives that result in extensive suppressions of recombination or aberrations of recombination frequency in mapping populations (Neu *et al.*, 2002; Jayatilake *et al.*, 2013; Xie *et al.*, 2012 supporting materials in Wang *et al.*, 2014). Marker order conservation across maps and relative to the consensus map demonstrated that the marker order was generally well conserved in the tetraploid data set, making it a distinct feature of the tetraploid germplasm as compared to *T. aestivum,* which shows both higher frequency of local rearrangements as well as frequent suppressed recombination events as shown for chromosome 3B and for several other chromosomes as well (Wang *et al.*, 2014). One possible explanation is that, compared to hexaploid wheat, tetraploid wheat germplasm has been less subjected to improvement through introgressions and chromosome engineering from the secondary and tertiary gene pools (Ceoloni *et al.*, 2005). In summary, this consensus map provides an important resource for its resolving power at pericentromeric regions in both durum and bread wheat. On the other hand, the extensive IBD among parents of durum germplasm shown by our study constrain mapping in populations derived from elite × elite crosses. This trend could result from extensive genetic diversity reduction in the elite durum wheat as a whole (Haudry *et al.*, 2007) and consequent hindrance of breeders' effort, or ascertainment bias caused by inadequate sampling of genetic diversity in durum wheat.

Marker density in the consensus map varied widely along chromosomes, even within intervals of 20 cM or less, likely due to both variation in recombination frequency (Erayman *et al.*, 2004; Saintenac *et al.*, 2011) and/or in genetic diversity (ancient genetic drift and selection sweep effects and/or more recent breeder-driven selection). As compared to the hexaploid wheat consensus map and the maps derived from either wild or cultivated emmer (Marone *et al.*, 2012; Maccaferri *et al.*, 2014), at least three large chromosome regions with reduced marker density have been consistently found in chromosomal arms 1AS, 2AS and 3AS of durum wheat across mapping populations and germplasm diversity studies.

Based on the relatively high number of component maps used to produce the consensus and their representativeness of the durum germplasm, these low-marker density regions were considered as a constitutive feature of the cultivated durum wheat genome. The local drop in marker density in these regions was found in all the durum maps but not in the emmer- and wild emmer-derived maps. The projection of the Illumina SNP polymorphism level in elite worldwide durum germplasm (expressed as PIC) to their genetic locations using the consensus map evidenced a marked decline in diversity and recombination rate in the centromeric/pericentromeric regions as well as strong fluctuation in diversity along the chromosomes, as previously shown in wheat and other species (Akhunov *et al.*, 2010; Wang *et al.*, 2011). The distribution of transcript-associated SNPs, with a slight enrichment in the pericentromeric and centromeric regions most probably caused by the lower recombination rates in such regions, supports the hypothesis of relatively even gene distribution along chromosomes, a result that matches the direct sequence observations reported for chromosome 3B (Choulet *et al.*, 2014). Irregular clustering of transcript-associated SNPs may be indicative of presence of gene-rich regions. Conversely,

the skewedness towards distal euchromatic regions observed for DArT markers and genomic SNPs from reduced-representation libraries based on the use of methylation-sensitive restriction enzymes was expected based on previous reports (Akbari *et al.*, 2006; Mantovani *et al.*, 2008; Wenzl *et al.*, 2006), including the absence of DArT coverage in the 4AS chromosome arm (Mantovani *et al.*, 2008).

## Hexaploid physical-to-tetraploid genetic map comparison for chromosome 3B

The graphical representation of the genetic *vs*. physical distances along the chromosome 3B showed a classic sigmoidal pattern that could be split into five regions showing more homogeneous relationships with the physical sequence. The two distal regions showing linear physical-to-genetic distance relationship corresponded closely to the two distal subregions identified by Choulet *et al.* (2014) based on segmentation analysis of the recombination rate and characterized by an elevated recombination rate (of 0.60 and 0.96 cM/Mb) clearly distinct from the low recombination rate (of 0.05 cM/Mb) observed for the large proximal region. These recombination rate values were confirmed in our chromosome 3B tetraploid consensus map. The localization of centromere based on suppressed recombination was also consistent with the physical centromere position in hexaploid wheat, namely 127 Mb in our case and 122 Mb in Choulet *et al.* (2014).

In spite of heterogeneity in the distribution of PIC values along the physical map in the durum wheat germplasm, a strong decline in polymorphism content outside of the two small distal regions—in agreement with observations by Akhunov *et al.* (2010) and Wang *et al.* (2014)—was evident. Meiotic recombination is the likely cause of the variability.

## A database of homeologous relationships for the mapped gene-derived SNPs

In wheat, differential transcriptional regulation of homeologs and duplicated gene copies is a mechanism for fine tuning the phenotypic response (Dubcovsky and Dvorak, 2007), therefore the possibility to easily cross-refer homologous genes/QTLs (and their corresponding tagging SNPs) between their A and B genomes facilitates genetic dissection of traits (Quarrie *et al.*, 2005; Diaz *et al.* 2005; Wilhelm *et al.* 2013; Zheng *et al.* 2013; Avni *et al.* 2014; Nitcher et al. 2014; Hurni *et al.* 2014). To date, homeologous relationships were mainly investigated by RFLPs (Qi *et al.*, 2004) and SSRs (Röder *et al.* 2008) that recognize homeologous copies as distinct patterns in the same genetic profile. Conversely, SNP markers are generally single Mendelian homeologous-specific assays. However, the 90K wheat SNP array has been mostly assembled from wheat transcripts (Wang *et al.*, 2014) and the SNP array includes multiple SNPs per single wheat transcript. These two features were used to establish robust relationships between SNPs mapped to homeologous/paralogous positions.

The mapped SNP sequences were related to the database of chromosome-specific hexaploid transcripts [The International Wheat Sequencing Consortium (IWGSC) (2014)] by conserved-sequence similarity analysis while data sets of SNPs were related to homeologous and paralogous transcript copies have been produced for each homeologous group. These data sets provide homeologous relationship patterns at relatively high density (up to 1937 SNPs per homeologous group for 818 nonredundant transcript sets). Homeologous relationships between

chromosomes harbouring ancestral translocation events (involving chromosome groups 4, 5 and 7; Ma *et al.*, 2013) were also clearly tagged and defined in their extent. In this case, the results were also supported by the absence of the homeologous transcripts in the nominal chromosomes of Chinese Spring (e.g. chrs. 4A-4B, 5A-5B, 7A-7B). Additionally, SNP-transcript relationships between non-nominal homeologous groups were established that could be ascribed to smaller and more fragmented translocation events already noticed based on SSR markers (Marone *et al.*, 2012). In some cases, connections were established between transcript sets of different groups (e.g. chromosome groups 1–3) for which the presence of the nominal homeologous transcripts in the corresponding chromosome-specific database deserves further investigation, possibly ascribed to ancestral genome duplications (Pont *et al.*, 2013).

## Conclusions

By bringing the number of marker loci specifically mapped in tetraploid wheat to a number close to that reached for hexaploid wheat, the present consensus map provides a significant advancement for durum wheat genetics. Most importantly, the majority of mapped SNPs are gene-derived markers that provide valuable anchor points for locus dissection and for bridging durum and bread wheat genetics. The proposed homeologous SNP and transcript database will be useful to define the homeologous relationships between chromosome regions harbouring target genes and QTLs. Additionally, the consensus map provides a framework of unprecedented marker density and genome coverage for association mapping, meta-QTL analysis and positional cloning, thus facilitating the application of molecular breeding approaches to enhance durum wheat productivity and quality. It is worth mentioning that the tetraploid consensus map includes mostly markers common to hexaploid wheat and also tetraploid-specific markers as well, and for the first time, the map integrates SSR, DArT and SNP markers in a single framework. Moreover, the map showed an improved marker order resolution in centromeric and pericentromeric regions as compared to hexaploid wheat map. The tetraploid consensus map will also have potential applications for several aspects of hexaploid wheat research and breeding.

## Experimental procedures

### Mapping populations used for consensus map assembling

Twelve recombinant inbred line (RIL) and one double haploid (DH) mapping populations were derived as follows: (i) eight populations from elite × elite durum wheat crosses including Colosseo × Lloyd (Maccaferri *et al.*, 2008b; Mantovani *et al.*, 2008), Meridiano × Claudio (Maccaferri *et al.*, 2011), Simeto × Levante (Maccaferri *et al.*, 2012), Mohawk × Cocorit69 (K. Ammar, unpublished), Svevo × Ciccio (Gadaleta *et al.*, 2011), W9292-260D3 × Kofa (C. Pozniak, unpublished), Kofa × Svevo (Maccaferri *et al.*, 2008a) and Kofa × UC1113 (Zhang *et al.*, 2008); (ii) three populations from durum × cultivated emmer including Ben × PI_41025 (Faris *et al.*, 2014), Simeto × Molise Colli (A. Mastrangelo, unpublished) and Latino × MG_5323 (F. Desiderio, unpublished); (iii) two populations from durum × wild emmer crosses including Langdon × G18-16 (Peleg *et al.*, 2008) and Svevo × Zavitan (A. Distelfeld, unpublished). Additional details are reported in Table 1.

### Single map analysis

Genotyping was carried out at each contributing institution following standard methods (Mantovani *et al.*, 2008; Peleg *et al.*, 2008; Wang *et al.*, 2014). Segregation data were obtained either as raw or postprocessed data. Low-quality data filtering was carried out according to rules: (i) missing data allowed to a maximum frequency of 0.10 (markers and lines), (ii) data points showing residual heterozygosity were converted to missing data, (iii) marker segregation distortion (departure from the expected 1:1 segregation ratio) allowed up to a threshold $P = 1E^{-04}$, corresponding to a segregation distortion not exceeding 0.7:0.3, above which the markers were eliminated due to residual heterogeneity in the parental lines or excessive distortion. SNPs segregating as dominant markers were retained only upon confirmation in the parental lines, provided that their segregation distortion did not exceed a stringent 0.60:0.40 ratio. After a round of provisional mapping, the distribution in recombination events in each mapping population was tested for departure from the expected Poisson' distribution (Minitab v16, Minitab inc. 2010). This step was critical for uncovering unreliable portions of genotypic data, most likely caused by erroneous DNA/genotyping data handling.

Filtered genotypic data were used to recalculate the linkage maps using Carthagene v. 1.2.3 (De Givry *et al.*, 2005). For the single component maps, nonsegregating SNPs were collapsed (*mrkdouble* function). Robust initial linkage groups (LG) were obtained using the *group* function at recombination frequency = 0.3 and LOD = 6.0. Assignment of LG to wheat chromosomes and their orientation were carried out by cross-checking the position of common SSR loci in the hexaploid wheat consensus map (Ta-SSR-2004, http://wheat.pw.usda.gov/GG2/index.shtml and/or http://www.triticarte.com.au). Once groups were attributed to chromosomes, their genotypic data were pooled on a chromosome-by-chromosome basis and regrouped at a more relaxed threshold of recombination frequency = 0.3 and LOD = 3 or progressively lowered. Mapping was carried using the multipoint maximum likelihood algorithm (*build* command) and the Haldane mapping function based on iterative cycles of: (i) mapping with marker order refinement (*greedy*, *flips* and *polish* commands), (ii) investigation of graphical genotypes for suspicious data points, for example singletons that recombined within ≤5 cM, not supported by other markers and mostly derived from unlikely events, particularly in case of SSR and DArT® markers; (iii) raw data checking and, in case, replacement with missing data. The reconstructed maps are thereafter reported as *component maps*.

### Construction of the consensus framework map and marker interpolation

The consensus framework map was constructed using the markers common to two or more populations (anchor) by collapsing the genotype data sets and using the same procedure described for the single populations. Prior to framework mapping, interpopulation recombination rate heterogeneity in the common chromosome regions was assessed. In the case of strong heterogeneity, a preselection of LGs from populations with relatively homogeneous recombination rates was carried out. Anchors led to obtain a robust framework map with intermarker genetic map distances estimated from mean recombination frequencies across populations. Loci mapped in

single populations only were projected by interpolation from the positions of the two closest flanking common markers (Cone *et al.*, 2002) in Biomercator v.3 (Sosnowski *et al.*, 2012). Conflicts in marker order conservation in single maps compared to the framework or consensus maps were detected in Biomercator (control of monotony step). Re-inspection of the segregation data for single maps carrying conflicts was carried out to assess the multipoint likelihood of the map order previous and after resolving the conflicts. For conflicts that were not alternative to the consensus order, the markers were considered as tagging two potentially different loci and thus renamed using a suffix. As the local conflicting orders across maps were usually two, of which one in low and the other in high frequency, only two potentially conflicting 'marker versions' ('a' and 'b') were used. After marker renaming, one or two cycles of single population and consensus mapping were carried out as necessary. The degree of correspondence and monotony between the single components and consensus maps was inspected by calculating pairwise rho (ρ) Spearman rank marker order correlations, a nonparametric measure of monotonic correlation, that is considering only the direction of marker concatenation.

### Assessment of SNP polymorphism information content in the durum wheat germplasm

A panel of 55 durum wheat accessions (mainly cultivars) including parents of mapping populations and reference cultivars originating from Mediterranean countries (Italy, Morocco, Spain, Syria and Tunisia), North and Southwestern USA, Canada and from CIMMYT and ICARDA breeding programs was used to evaluate the polymorphism information content (PIC) of mapped SNPs and to relate it to the chromosome map location. PIC was computed as follows:

$$PIC = 1 - \sum p_j^2$$

where $p_j$ is the frequency of each of the two alleles across accessions (Powell *et al.*, 1996).

### Investigation of the relationships between the hexaploid physical and tetraploid genetic map of chromosome 3B

The hexaploid wheat chromosome 3B assembled physical sequence (pseudomolecule) of Chinese Spring was downloaded from the URGI repository (http://wheat-urgi.versailles.inra.fr/Seq-Repository/Reference-sequence, traes3bPseudomoleculeV1) (Choulet *et al.*, 2014). The sequences of chromosome 3B-mapped SNPs from the Illumina manifest file (http://triticeaetoolbox.org/wheat/display_genotype.php?trial_code=TCAP90K_SWWpanel) were used to conduct a conserved-sequence search on the chromosome 3B pseudomolecule by BLASTN with a filtering threshold *e*-value <$1e^{-10}$.

### Identification of homeologous relationships between gene-derived SNPs

The mapped Illumina iSelect wheat SNP sequences were used as queries to search the chromosome-specific transcript database of the hexaploid wheat Sequence Survey database (ftp://ftp.ensemblgenomes.org/pub/plants/release22/fasta/triticum_aestivum/cdna/Triticum_aestivum.IWGSP1.22.cdna.all.fa.gz). Based on the recovered hits, SNPs and transcripts were interrelated based on the SNP chromosome and mapping positions and the chromosome-specific origin of transcripts in Excel.

## Conflict of interest

The authors have no conflict of interest to declare.

## Acknowledgements

## References

Akbari, M., Wenzl, P., Vanessa, C., Carling, J., Xia, L., Yang, S., Uszynski, G., Mohler, V., Lehmensiek, A., Kuchel, H., Hayden, M.J., Howes, N., Sharp, P., Rathmell, B., Vaughan, P., Huttner, E. and Kilian, A. (2006) Diversity Arrays Technology (DArT) for high-throughput profiling of the hexaploid wheat genome. *Theor. Appl. Genet.* **113**, 1409–1420.

Akhunov, E.D., Akhunova, A.R., Anderson, O.D., Anderson, J.A., Blake, N., Clegg, M.T., Coleman-Derr, D., Conley, E.J., Crossman, C.C., Deal, K.R., Dubcovsky, J., Gill, B.S., Gu, Y.Q., Hadam, J., Heo, H., Huo, N.X., Lazo, G.R., Luo, M.C., Ma, Y.Q., Matthews, D.E., McGuire, P.E., Morrell, P.L., Qualset, C.O., Renfro, J., Tabanao, D., Talbert, L.E., Tian, C., Toleno, D.M., Warburton, M.L., You, F.M., Zhang, W.J. and Dvorak, J. (2010) Nucleotide diversity maps reveal variation in diversity among wheat genomes and chromosomes. *BMC Genom.* **11**, 702.

Allen, A.M., Barker, G.L.A., Wilkinson, P., Burridge, A., Winfield, M., Coghill, J., Uauy, C., Griffiths, S., Jack, P., Berry, S., Werner, P., Melichar, J.P.E., McDougall, J., Gwilliam, R., Robinson, P. and Edwards, K.J. (2013) Discovery and development of exome-based, co-dominant single nucleotide polymorphism markers in hexaploid wheat (*Triticum aestivum* L.). *Plant Biotechnol. J.* **11**, 279–295.

Avni, R., Zhao, R.R., Pearce, S., Jun, Y., Uauy, C., Tabbita, F., Fahima, T., Slade, A., Dubcovsky, J. and Distelfeld, A. (2014) Functional characterization of *GPC-1* genes in hexaploid wheat. *Planta*, **239**, 313–324.

Blanco, A., Bellomo, M.P., Cenci, A., De Giovanni, C., D'Ovidio, R., Iacono, E., Laddomada, B., Pagnotta, M.A., Porceddu, E., Sciancalepore, A., Simeone, R. and Tanzarella, O.A. (1998) A genetic linkage map of durum wheat. *Theor. Appl. Genet.* **97**, 721–728.

Blenda, A., Fang, D.D., Rami, J.F., Garsmeur, O., Luo, F. and Lacape, J.M. (2012) A high density consensus genetic map of tetraploid cotton that integrates multiple component maps through molecular marker redundancy check. *PLoS One*, **7**, e45739.

Bowers, J.E., Nambeesan, S., Corbi, J., Barker, M.S., Rieseberg, L.H., Knapp, S.J. and Burke, J.M. (2012) Development of an ultra-dense genetic map of the sunflower genome based on single-feature polymorphisms. *PLoS One*, **7**, e51360.

Caldwell, K.S., Dvorak, J., Lagudah, E.S., Akhunov, E., Luo, M.C., Wolters, P. and Powell, W. (2004) Sequence polymorphism in polyploidy wheat and their D-genome diploid ancestor. *Genetics*, **167**, 941–947.

Cavanagh, C.R., Chao, S.M., Wang, S.C., Huang, B.E., Stephen, S., Kiani, S., Forrest, K., Saintenac, C., Brown-Guedira, G.L., Akhunova, A., See, D., Bai, G.H., Pumphrey, M., Tomar, L., Wong, D.B., Kong, S., Reynolds, M., da

Silva, M.L., Bockelman, H., Talbert, L., Anderson, J.A., Dreisigacker, S., Baenziger, S., Carter, A., Korzun, V., Morrell, P.L., Dubcovsky, J., Morell, M.K., Sorrells, M.E., Hayden, M.J. and Akhunov, E. (2013) Genome-wide comparative diversity uncovers multiple targets of selection for improvement in hexaploid wheat landraces and cultivars. *Proc. Natl Acad. Sci. USA*, **110**, 8057–8062.

Ceoloni, C., Forte, P., Gennaro, A., Micali, S., Carozza, R. and Bitti, A. (2005) Recent developments in durum wheat chromosome engineering. *Cytogenet. Genome Res.* **109**, 328–334.

Choulet, F., Alberti, A., Theil, S., Glover, N., Barbe, V., Daron, J., Pingault, L., Sourdille, P., Couloux, A., Paux, E., Leroy, P., Mangenot, S., Guilhot, N., Le Gouis, J., Balfourier, F., Alaux, M., Jamilloux, V., Poulain, J., Durand, C., Bellec, A., Gaspin, C., Safar, J., Dolezel, J., Rogers, J., Vandepoele, K., Aury, J.-M., Mayer, K., Berges, H., Quesneville, H., Wincker, P. and Feuillet, C. (2014) Structural and functional partitioning of bread wheat chromosome 3B. *Science*, **345**, 1249721.

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

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

Diaz, A., Zikhali, M., Turner, A.S., Isaac, P. and Laurie, D.A. (2012) Copy number variation affecting the *photoperiod-B1* and *Vernalization-A1* genes is associated with altered flowering time in Wheat (*Triticum aestivum*). *PLoS One*, **7**, e33234.

Dubcovsky, J. and Dvorak, J. (2007) Genome plasticity a key factor in the success of polyploid wheat under domestication. *Science*, **316**, 1862–1866.

Dvorak, J., Luo, M.C., Yang, Z.L. and Zhang, H.B. (1998) The structure of the *Aegilops tauschii* genepool and the evolution of hexaploid wheat. *Theor. Appl. Genet.* **97**, 657–670.

Dvorak, J., Akhunov, E.D., Akhunov, A.R., Deal, K.R. and Luo, M.C. (2006) Molecular characterization of a diagnostic DNA marker for domesticated tetraploid wheat provides evidence for gene flow from wild tetraploid wheat to hexaploid wheat. *Mol. Biol. Evol.* **23**, 1386–1396.

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

Faris, J.D., Zhang, Q., Chao, S., Zhang, Z. and Xu, S.S. (2014) Analysis of agronomic and domestication traits in a durum × cultivated emmer wheat population using a high-density single nucleotide polymorphism-based linkage map. *Theor. Appl. Genet.* **127**, 2333–2348.

Feuillet, C., Stein, N., Rossini, L., Praud, S., Mayer, K., Schulman, A., Eversole, K. and Appels, R. (2012) Integrating cereal genomics to support innovation in the Triticeae. *Funct. Integr. Genomics*, **12**, 573–583.

Gadaleta, A., Nigro, D., Giancaspro, A. and Blanco, A. (2011) The glutamine synthetase (*GS2*) genes in relation to grain protein content of durum wheat. *Funct. Integr. Genomics*, **11**, 665–670.

Ganal, M., Wieseke, R., Luerssen, H., Durstewitz, G., Graner, E.-M., Plieske, J. and Polley, A. (2014) High-throughput SNP Profiling of Genetic Resources in Crop Plants Using Genotyping Arrays. In *Genomics of Plant Genetic Resources*(Tuberosa, R., Graner, A. and Frison, E., eds), pp. 113–130. Netherlands: Springer.

Gautami, B., Fonceka, D., Pandey, M.K., Moretzsohn, M.C., Sujay, V., Qin, H.D., Hong, Y.B., Faye, I., Chen, X.P., BhanuPrakash, A., Shah, T.M., Gowda, M.V.C., Nigam, S.N., Liang, X.Q., Hoisington, D.A., Guo, B.Z., Bertioli, D.J., Rami, J.F. and Varshney, R.K. (2012) An international reference consensus genetic map with 897 marker loci based on 11 mapping populations for tetraploid groundnut (*Arachis hypogaea* L.). *PLoS One*, **7**, e41213.

Giles, R.J. and Brown, T.A. (2006) GluDy allele variations in *Aegilops tauschii* and *Triticum aestivum*: implications for the origins of hexaploid wheats. *Theor. Appl. Genet.* **112**, 1563–1572.

Goretti, D., Bitocchi, E., Bellucci, E., Rodriguez, M., Rau, D., Gioia, T., Attene, G., McClean, P., Nanni, L. and Papa, R. (2014) Development of single nucleotide polymorphisms in *Phaseolus vulgaris* and related *Phaseolus* spp. *Mol. Breeding*, **33**, 531–544.

Gupta, P.K., Mir, R.R., Mohan, A. and Kumar, J. (2008) Wheat genomics: present status and future prospects. *Int. J. Plant Genomics*, **2008**, 896451.

Haudry, A., Cenci, A., Ravel, C., Bataillon, T., Brunel, D., Poncet, C., Hochu, I., Poirier, S., Santoni, S., Glemin, S. and David, J. (2007) Grinding up wheat: a massive loss of nucleotide diversity since domestication. *Mol. Biol. Evol.* **24**, 1506–1517.

Hudson, C.J., Freeman, J.S., Kullan, A.R.K., Petroli, C.D., Sansaloni, C.P., Kilian, A., Detering, F., Grattapaglia, D., Potts, B.M., Myburg, A.A. and Vaillancourt, R.E. (2012) A reference linkage map for *Eucalyptus*. *BMC Genom.* **13**, 240.

Hurni, S., Brunner, S., Stirnweis, D., Herren, G., Peditto, D., McIntosh, R.A. and Keller, B. (2014) The powdery mildew resistance gene *Pm8* derived from rye is suppressed by its wheat ortholog *Pm3*. *Plant J.* **79**, 904–913. doi:10.1111/tpj. 12593.

Hyten, D.L., Song, Q., Fickus, E.W., Quigley, C.V., Lim, J.-S., Choi, I.-Y., Hwang, E.-Y., Pastor-Corrales, M. and Cregan, P.B. (2010) High-throughput SNP discovery and assay development in common bean. *BMC Genom.* **11**, 475.

IPO International Pasta Organization (2013) *The World Pasta Industry Status Report*. www.internationalpasta.org p. 44.

Jayatilake, D.V., Tucker, E.J., Bariana, H., Kuchel, H., Edwards, J., McKay, A.C., Chalmers, K. and Mather, D.E. (2013) Genetic mapping and marker development for resistance of wheat against the root lesion nematode *Pratylenchus neglectus*. *BMC Plant Biol.* **13**, 230.

Kihara, H. (1944) Discovery of the DD-analyser, one of the ancestors of *Triticum vulgare*. *Agric. Hortic. (Tokyo)*, **19**, 13–14.

Korzun, V., Röder, M.S., Wendehake, K., Pasqualone, A., Lotti, C., Ganal, M.W. and Blanco, A. (1999) Integration of dinucleotide microsatellites from hexaploid bread wheat into a genetic linkage map of durum wheat. *Theor. Appl. Genet.* **98**, 1202–1207.

Lotti, C., Salvi, S., Pasqualone, A., Tuberosa, R. and Blanco, A. (2000) Integration of AFLP markers into an RFLP-based map of durum wheat. *Plant Breeding*, **119**, 393–401.

Lai, K., Duran, C., Berkman, P.J., Lorenc, M.T., Stiller, J., Manoli, S., Hayden, M.J., Forrest, K.L., Fleury, D., Baumann, U., Zander, M., Mason, A.S., Batley, J. and Edwards, D. (2012) Single nucleotide polymorphism discovery from wheat next-generation sequence data. *Plant Biotechnol. J.* **10**, 743–749.

Ma, J., Stiller, J., Berkman, P.J., Wei, Y., Rogers, J., Feuillet, C., Dolezel, J., Mayer, K.F., Eversole, K., Zheng, Y.-L. and Liu, C. (2013) Sequence-based analysis of translocations and inversions in bread wheat (*Triticum aestivum* L.). *PLoS One*, **8**, e79329.

Maccaferri, M., Sanguineti, M.C., Donini, P. and Tuberosa, R. (2003) Microsatellite analysis reveals a progressive widening of the genetic basis in the elite durum wheat germplasm. *Theor. Appl. Genet.* **107**, 783–797.

Maccaferri, M., Sanguineti, M.C., Noli, E. and Tuberosa, R. (2005) Population structure and long-range linkage disequilibrium in a durum wheat elite collection. *Mol. Breeding*, **15**, 271–289.

Maccaferri, M., Stefanelli, S., Rotondo, F., Tuberosa, R. and Sanguineti, M.C. (2007) Relationships among durum wheat accessions. I. Comparative analysis of SSR, AFLP, and phenotypic data. *Genome*, **50**, 373–384.

Maccaferri, M., Sanguineti, M.C., Corneti, S., Ortega, J.L.A., Ben Salem, M., Bort, J., DeAmbrogio, E., Fernando Garcia del Moral, L., Demontis, A., El-Ahmed, A., Maalouf, F., Machlab, H., Martos, V., Moragues, M., Motawaj, J., Nachit, M., Nserallah, N., Ouabbou, H., Royo, C., Slama, A. and Tuberosa, R. (2008a) Quantitative trait loci for grain yield and adaptation of durum wheat (*Triticum durum* Desf.) across a wide range of water availability. *Genetics*, **178**, 489–511.

Maccaferri, M., Mantovani, P., Tuberosa, R., DeAmbrogio, E., Giuliani, S., Demontis, A., Massi, A. and Sanguineti, M.C. (2008b) A major QTL for durable leaf rust resistance widely exploited in durum wheat breeding programs maps on the distal region of chromosome arm 7BL. *Theor. Appl. Genet.* **117**, 1225–1240.

Maccaferri, M., Ratti, C., Rubies-Autonell, C., Vallega, V., Demontis, A., Stefanelli, S., Tuberosa, R. and Sanguineti, M.C. (2011) Resistance to Soil-borne cereal mosaic virus in durum wheat is controlled by a major QTL on chromosome arm 2BS and minor loci. *Theor. Appl. Genet.* **123**, 527–544.

Maccaferri, M., Francia, R., Ratti, C., Rubies-Autonell, C., Colalongo, C., Ferrazzano, G., Tuberosa, R. and Sanguineti, M.C. (2012) Genetic analysis of *Soil-Borne Cereal Mosaic Virus* response in durum wheat: evidence for the role of the major quantitative trait locus QSbm.ubo-2BS and of minor quantitative trait loci. *Mol. Breeding*, **29**, 973–988.

Maccaferri, M., Cane', M.A., Colalongo, C., Massi, A., Clarke, F., Pozniak, C., Korol, A., Fahima, T., Dubcovsky, J., Xu, S., Karsai, I., Knox, R., Clarke, J.M., Salvi, S., Sanguineti, M.C. and Tuberosa, R. (2014) A consensus framework map of durum wheat (Triticum durum Desf.) suitable for linkage disequilibrium and genome-wide association mapping. *BMC Genom.* **15**, 873.

Mantovani, P., Maccaferri, M., Sanguineti, M.C., Tuberosa, R., Catizone, I., Wenzl, P., Thomson, B., Carling, J., Huttner, E., DeAmbrogio, E. and Kilian, A. (2008) An integrated DArT-SSR linkage map of durum wheat. *Mol. Breeding*, **22**, 629–648.

Marcussen, T., Sandve, S.R., Heier, L., Spannagl, M., Pfeifer, M.; The International Wheat Genome Sequencing Consortium, Jakobsen, K.S., Wulff, B.B.H., Steuernagel, B., Mayer, K.F.X. and Olsen, O.A. (2014) Ancient hybridizations among the ancestral genomes of bread wheat. *Science*, **345**, 1250092.

Marone, D., Laido, G., Gadaleta, A., Colasuonno, P., Ficco, D.B.M., Giancaspro, A., Giove, S., Panio, G., Russo, M.A., De Vita, P., Cattivelli, L., Papa, R., Blanco, A. and Mastrangelo, A.M. (2012) A high-density consensus map of A and B wheat genomes. *Theor. Appl. Genet.* **125**, 1619–1638.

McFadden, E.S. and Sears, E.R. (1946) The origin of *Triticum spelta* and its free-threshing hexaploid relatives. *J. Hered.* **37**, 81–89.

Mori, N. (2003) Wheat domestication: when, where and how? *Plant Cell Physiol.* **44**, S2–S2.

Muñoz-Amatriaín, M., Moscou, M.J., Bhat, P.R., Svensson, J.T., Bartos, J., Suchankova, P., Simkova, H., Endo, T.R., Fenton, R.D., Lonardi, S., Castillo, A.M., Chao, S., Cistue, L., Cuesta-Marcos, A., Forrest, K.L., Hayden, M.J., Hayes, P.M., Horsley, R.D., Makoto, K., Moody, D., Sato, K., Valles, M.P., Wulff, B.B.H., Muehlbauer, G.J., Dolezel, J. and Close, T.J. (2011) An improved consensus linkage map of barley based on flow-sorted chromosomes and single nucleotide polymorphism markers. *Plant Genome*, **4**, 238–249.

Nesbitt, M. and Samuel, D. (1996) From stable crop to extinction? The archaeology and history of hulled wheats. In: S. Padulosi, K. Hammer, J. Hellers (Eds) *Hulled wheat*. International Plant Genetic Resources Institute, Rome, pp. 41–100.

Neu, C., Stein, N. and Keller, B. (2002) Genetic mapping of the *Lr20-Pm1* resistance locus reveals suppressed recombination on chromosome arm 7AL in hexaploid wheat. *Genome*, **45**, 737–744.

Nitcher, R., Distelfeld, A., Tan, C., Yan, L. and Dubcovsky, J. (2013) Increased copy number at the *HvFT1* locus is associated with accelerated flowering time in barley. *Mol. Genet. Genomics*, **288**, 261–275.

Ozkan, H., Brandolini, A., Pozzi, C., Effgen, S., Wunder, J. and Salamini, F. (2005) A reconsideration of the domestication geography of tetraploid wheats. *Theor. Appl. Genet.* **110**, 1052–1060.

Peleg, Z., Saranga, Y., Suprunova, T., Ronin, Y., Roeder, M.S., Kilian, A., Korol, A.B. and Fahima, T. (2008) High-density genetic map of durum wheat x wild emmer wheat based on SSR and DArT markers. *Theor. Appl. Genet.* **117**, 103–115.

van Poecke, R.M.P., Maccaferri, M., Tang, J., Truong, H.T., Janssen, A., van Orsouw, N.J., Salvi, S., Sanguineti, M.C., Tuberosa, R. and van der Vossen, E.A.G. (2013) Sequence-based SNP genotyping in durum wheat. *Plant Biotechnol. J.* **11**, 809–817.

Pont, C., Murat, F., Guizard, S., Flores, R., Foucrier, S., Bidet, Y., Quraishi, U.M., Alaux, M., Dolezel, J., Fahima, T., Budak, H., Keller, B., Salvi, S., Maccaferri, M., Steinbach, D., Feuillet, C., Quesneville, H. and Salse, J. (2013) Wheat syntenome unveils new evidences of contrasted evolutionary plasticity between paleo- and neoduplicated subgenomes. *Plant J.* **76**, 1030–1044.

Powell, W., Morgante, M., Andre, C., Hanafey, M., Vogel, J., Tingey, S. and Rafalski, A. (1996) The comparison of RFLP, RAPD, AFLP and SSR (microsatellite) markers for germplasm analysis. *Mol. Breeding*, **2**, 225–238.

Qi, L.L., Echalier, B., Chao, S., Lazo, G.R., Butler, G.E., Anderson, O.D., Akhunov, E.D., Dvorak, J., Linkiewicz, A.M., Ratnasiri, A., Dubcovsky, J.,

Bermudez-Kandianis, C.E., Greene, R.A., Kantety, R., La Rota, C.M., Munkvold, J.D., Sorrells, S.F., Sorrells, M.E., Dilbirligi, M., Sidhu, D., Erayman, M., Randhawa, H.S., Sandhu, D., Bondareva, S.N., Gill, K.S., Mahmoud, A.A., Ma, X.F., Miftahudin, Gustafson, J.P., Conley, E.J., Nduati, V., Gonzalez-Hernandez, J.L., Anderson, J.A., Peng, J.H., Lapitan, N.L.V., Hossain, K.G., Kalavacharla, V., Kianian, S.F., Pathan, M.S., Zhang, D.S., Nguyen, H.T., Choi, D.W., Fenton, R.D., Close, T.J., McGuire, P.E., Qualset, C.O. and Gill, B.S. (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.

Quarrie, S.A., Steed, A., Calestani, C., Semikhodskii, A., Lebreton, C., Chinoy, C., Steele, N., Pljevljakusic, D., Waterman, E., Weyen, J., Schondelmaier, J., Habash, D.Z., Farmer, P., Saker, L., Clarkson, D.T., Abugalieva, A., Yessimbekova, M., Turuspekov, Y., Abugalieva, S., Tuberosa, R., Sanguineti, M.C., Hollington, P.A., Aragues, R., Royo, A. and Dodig, D. (2005) A high-density genetic map of hexaploid wheat (*Triticum aestivum* L.) from the cross Chinese Spring X SQ1 and its use to compare QTLs for grain yield across a range of environments. *Theor. Appl. Genet.* **110**, 865–880.

Röder, M.S., Korzun, V., Wendehake, K., Plaschke, J., Tixier, M.H., Leroy, P. and Ganal, M.W. (1998) A microsatellite map of wheat. *Genetics*, **149**, 2007–2023.

Ronin, Y., Mester, D., Minkov, D., Belotserkovski, R., Jackson, B.N., Schnable, P.S., Aluru, S. and Korol, A. (2012) Two-phase analysis in consensus genetic mapping. *G3-Genes Genomes Genet.* **2**, 537–549.

Saintenac, C., Faure, S., Remay, A., Choulet, F., Ravel, C., Paux, E., Balfourier, F., Feuillet, C. and Sourdille, P. (2011) Variation in crossover rates across a 3-Mb contig of bread wheat (*Triticum aestivum* L.) reveals the presence of a meiotic recombination hotspot. *Chromosoma*, **120**, 185–198.

Saintenac, C., Jiang, D., Wang, S. and Akhunov, E. (2013) Sequence-based mapping of the polyploid wheat genome. *G3-Genes Genomes Genet.* **3**, 1105–1114.

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

Salvi, S., Castelletti, S. and Tuberosa, R. (2010) An updated consensus map for flowering time QTLs in maize. *Maydica*, **54**, 501–512.

Somers, D.J., Isaac, P. and Edwards, K. (2004) A high-density microsatellite consensus map for bread wheat (*Triticum aestivum* L.). *Theor. Appl. Genet.* **109**, 1105–1114.

Sosnowski, O., Charcosset, A. and Joets, J. (2012) BioMercator V3: an upgrade of genetic map compilation and quantitative trait loci meta-analysis algorithms. *Bioinformatics*, **28**, 2082–2083.

Sourdille, P., Singh, S., Cadalen, T., Brown-Guedira, G.L., Gay, G., Qi, L., Gill, B.S., Dufour, P., Murigneux, A. and Bernard, M. (2004) Microsatellite-based deletion bin system for the establishment of genetic-physical map relationships in wheat (*Triticum aestivum* L.). *Funct. Integr. Genomics*, **4**, 12–25.

Swamy, B.P., Vikram, P., Dixit, S., Ahmed, H.U. and Kumar, A. (2011) Meta-analysis of grain yield QTL identified during agricultural drought in grasses showed consensus. *BMC Genom.* **12**, 319.

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

Tiwari, V.K., Wang, S., Sehgal, S., Vrana, J., Friebe, B., Kubalakova, M., Chhuneja, P., Dolezel, J., Akhunov, E., Kalia, B., Sabir, J. and Gill, B.S. (2014) SNP Discovery for mapping alien introgressions in wheat. *BMC Genom.* **15**, 273.

Trebbi, D., Maccaferri, M., de Heer, P., Sorensen, A., Giuliani, S., Salvi, S., Sanguineti, M.C., Massi, A., van der Vossen, E.A.G. and Tuberosa, R. (2011) High-throughput SNP discovery and genotyping in durum wheat (*Triticum durum* Desf.). *Theor. Appl. Genet.* **123**, 555–569.

Varshney, R.K., Langridge, P. and Graner, A. (2007) Application of genomics to molecular breeding of wheat and barley. *Adv. Genet.* **58**, 121–155.

Wang, J., Lydiate, D.J., Parkin, I.A.P., Falentin, C., Delourme, R., Carion, P.W.C. and King, G.J. (2011) Integration of linkage maps for amphiploid Brassica napus and comparative mapping with Arabidopsis and Brassica rapa. *BMC Genom.* **12**, 101.

Wang, S., Wong, D., Forrest, K., Allen, A., Chao, S., Huang, B.E., Maccaferri, M., Salvi, S., Milner, S.G., Cattivelli, L., Mastrangelo, A.M., Whan, A.,

Stephen, S., Barker, G., Wieseke, R., Plieske, J.; International Wheat Genome Sequencing C, Lillemo, M., Mather, D., Appels, R., Dolferus, R., Brown-Guedira, G., Korol, A., Akhunova, A.R., Feuillet, C., Salse, J., Morgante, M., Pozniak, C., Luo, M.-C., Dvorak, J., Morell, M., Dubcovsky, J., Ganal, M., Tuberosa, R., Lawley, C., Mikoulitch, I., Cavanagh, C., Edwards, K.J., Hayden, M. and Akhunov, E. (2014) Characterization of polyploid wheat genomic diversity using a high-density 90 00 single nucleotide polymorphism array. *Plant Biotechnol. J.* **12**, 787–796.

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

Wilhelm, E.P., Boulton, M.I., Al-Kaff, N., Balfourier, F., Bordes, J., Greenland, A.J., Powell, W. and Mackay, I.J. (2013) *Rht-1* and *Ppd-D1* associations with height, GA sensitivity, and days to heading in a worldwide bread wheat collection. *Theor. Appl. Genet.* **126**, 2233–2243.

Xie, W., Ben-David, R., Zeng, B., Dinoor, A., Xie, C., Sun, Q., Roeder, M.S., Fahoum, A. and Fahima, T. (2012) Suppressed recombination rate in 6VS/6AL translocation region carrying the *Pm21* locus introgressed from *Haynaldia villosa* into hexaploid wheat. *Mol. Breeding*, **29**, 399–412.

Zhang, W., Chao, S., Manthey, F., Chicaiza, O., Brevis, J.C., Echenique, V. and Dubcovsky, J. (2008) QTL analysis of pasta quality using a composite microsatellite and SNP map of durum wheat. *Theor. Appl. Genet.* **117**, 1361–1377.

Zheng, B., Biddulph, B., Li, D., Kuchel, H. and Chapman, S. (2013) Quantification of the effects of *VRN1* and *Ppd-D1* to predict spring wheat (*Triticum aestivum*) heading time across diverse environments. *J. Exp. Bot.* **64**, 3747–3761.

## Supporting information

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

**Figure S1** Relationship between marker order and genetic distances of the 13 component maps and the consensus map. Component maps are projected (*Y* values) on the consensus map (*X* values) for all chromosomes. Spearman rank correlations are reported as rho (ρ) values for each linkage group.

**Figure S2** Single component and consensus map representations for chromosome 3B.

**Figure S3** Rank-order plots of the hexaploid consensus map *vs.* the tetraploid wheat consensus for each of the 14 chromosomes. The marker order of tetraploid map is reported in the *X* axisis and the corresponding hexaploid map order is in the *Y* axis.

**Table S1** Single component map features.

**Table S2** Tetraploid consensus map features.

**Table S3** Summary of results of sequence similarity search for mapped Illumina SNPs on the hexaploid wheat Chinese Spring chromosome-sorted survey sequence transcript database.

**Table S4** Results of sequence similarity search for mapped Illumina SNPs on the hexaploid wheat Chinese Spring chromosome-sorted survey sequence transcript database.

**Table S5** Clusters of mapped SNPs tagging homeologous transcripts based on sequence similarity to the hexaploid wheat Chinese Spring chromosome sorted survey sequence transcript database.

**Table S6** Clusters of mapped SNPs tagging homeologous transcripts located on translocated chromosomes based on sequence similarity to the hexaploid wheat Chinese Spring chromosome sorted survey sequence transcript database.

**Text S1** SNP-transcript matches tagging homeologous relationships between chromosome regions in non-nominal homeologous relationships.