**Identification of genes bordering breakpoints of the
pericentric inversions on 2B, 4B and 5A in Bread Wheat
(*Triticum aestivum* L.)**

Journal:	<i>Genome</i>
Manuscript ID	gen-2015-0060.R1
Manuscript Type:	Note
Date Submitted by the Author:	12-Jul-2015
Complete List of Authors:	Ma, Jian; Sichuan Agricultural University, Triticeae Research Institute Gao, Shang; Sichuan Agricultural University, Triticeae Research Institute Stiller, Jiri; CSIRO, Agriculture Flagship Jiang, Qian-Tao; Sichuan Agricultural University, Triticeae Research Institute Lan, Xiu-Jin; Sichuan Agricultural University, Triticeae Research Institute Liu, Ya-Xi; Sichuan Agricultural University, Triticeae Research Institute Pu, Zhi-En; Sichuan Agricultural University, Triticeae Research Institute Wang, Jirui; Sichuan Agricultural University, Triticeae Research Institute Wei, Yu-Ming; Sichuan Agricultural University, Triticeae Research Institute Zheng, You-Liang; Sichuan Agricultural University,
Keyword:	genome evolution, bread wheat, inversion, translocation breakpoint, Brachypodium
Note: The following files were submitted by the author for peer review, but cannot be converted to PDF. You must view these files (e.g. movies) online.	
2015 0060 Supplementary data 10 files.zip	

SCHOLARONE™
Manuscripts

Identification of genes bordering breakpoints of the pericentric inversions on 2B, 4B and 5A in Bread Wheat (*Triticum aestivum* L.)

Jian Ma^{1, #}, Shang Gao^{1, #}, Jiri Stiller², Qian-Tao Jiang^{1, *}, Xiu-Jin Lan¹, Ya-Xi Liu¹, Zhi-En Pu¹, Jirui Wang¹, Yuming Wei¹, You-Liang Zheng^{1, *}

¹*Triticeae Research Institute, Sichuan Agricultural University, Wenjiang, Chengdu 611130, China*

²*CSIRO Agriculture Flagship, 306 Carmody Road, St Lucia, QLD 4067, Australia*

These authors contributed equally to this work.

*Authors for correspondence:

Dr. Qian-Tao Jiang:

E-mail: qiantaojiang@sicau.edu.cn; Tel: +86 28 86290958; Fax: +86 28 82650350

Prof. You-Liang Zheng:

E-mail: ylzheng@sicau.edu.cn; Tel: +86 835 2882007; Fax: +86 835 2883153

Abstract

Chromosome translocation is an important driving force in shaping genomes during evolution. Detailed knowledge of chromosome translocations in a given species and its close relatives should increase the efficiency and precision of chromosome engineering in crop improvement. To identify genes flanking the breakpoints of translocations and inversions as a step toward identifying breakpoints in bread wheat, we systematically analysed genes in the *Brachypodium* genome against wheat survey sequences and bin-mapped ESTs (Expressed Sequence Tags) derived from the hexaploid wheat genotype 'Chinese Spring'. In addition to those well-known translocations between group 4, 5 and 7 chromosomes, this analysis identified genes flanking the 3 pericentric inversions on chromosomes 2B, 4B and 5A. However, numerous chromosomal rearrangements reported in early studies could not be confirmed. The genes flanking the breakpoints reported in this study are valuable for isolating these breakpoints.

Keywords genome evolution, bread wheat, inversion, translocation breakpoint, *Brachypodium*

Introduction

It is widely believed that chromosome translocation is an important driving force in shaping genomes during evolution (Brown and O'Neill 2010; Burt et al. 1999; Colson et al. 2004; Morrow and Cooper 2012; Sankoff and Nadeau 2003). In human (Brown and O'Neill 2010), animals (Burt et al. 1999) and microbes (Colson et al. 2004; Morrow and Cooper 2012), chromosome translocation is usually related to genomic disorders. Some studies also reported that translocation could alter levels of recombination (McKim et al. 1988; Sherizen et al. 2005) which is not only a major source of intra-specific variation but also an important constraint in crop improvement programs. Such programs aim to bring together multiple chromosomal segments containing favourable alleles into single plant lines.

It is well known that the non-homoeologous translocations between chromosomes 4A, 5A and 7B are important translocations in the hexaploid wheat (*Triticum aestivum* L., $2n=6x=42$, genome AABBDD) (Devos et al. 1995; Liu et al. 1992; Naranjo et al. 1987; Nelson et al. 1995). Further, for the first time, by analysing *Brachypodium* genes against survey sequences of sorted wheat chromosome arms and by analysing *Brachypodium* genes against deletion bin-mapped wheat ESTs, we have previously identified *Brachypodium* orthologs bordering several translocation and inversion breakpoints on the modern wheat chromosomes 4A, 5A and 7B (Ma et al. 2013).

In addition to these major translocation events, many other likely chromosomal rearrangements have been reported in the hexaploid wheat: 1) a pericentric inversion in 4B evidenced by cytology (Dvořák et al. 1984; Endo and Gill 1984; Gill et al. 1991), gene-based marker (Mickelson-Young et al. 1995b), and bin-mapped expressed sequence tags (ESTs) (Miftahudin. et al. 2004); 2) a non-homoeologous translocations between chromosomes 2BS and 6BS evidenced by cDNA clone (Devos et al. 1993) and bin-mapped ESTs (Conley

et al. 2004); 3) a pericentric inversion in 2B deduced by bin-mapped ESTs (Conley et al. 2004); 4) a putative pericentric inversion in 5A deduced by bin-mapped ESTs (Linkiewicz et al. 2004); 5) some possible rearrangements in chromosome group 1 and 3 deduced by bin-mapped ESTs (Munkvold et al. 2004; Peng et al. 2004). Further, by analysing homoeoloci for a set of nonredundant genes from *Brachypodium distachyon*, *T. urartu*, and *Aegilops tauschii* in the 'CS' chromosome shotgun sequence obtained from individual chromosome arms flow-sorted from 'CS' aneuploid lines, we systematically assessed chromosomal rearrangements in the hexaploid wheat genotype 'Chinese Spring' ('CS') (Ma et al. 2014; Ma et al. 2015).

Despite of numerous deduced rearrangements being identified, identification of the breakpoints for these translocations and inversions has been so far unsuccessful. Recent development in wheat genome sequencing (International Wheat Genome Sequencing Consortium 2014) offers an excellent opportunity to develop methods that enable rapid and reliable detection of these translocations and inversions. This can be achieved by firstly identifying genomic sequences covering each of these breakpoints and then developing PCR primers that allow the amplification of these sequences. Once such PCR primers become available, cloning and sequencing a PCR product will be all required to detect the presence or absence of any of these translocations and inversions in a given genotype. Thus large numbers of genotypes representing different types of wheats and their close relatives can be assessed, making it possible to evaluate the effects of these translocations in wheat evolution.

Aiming at finding genes flanking the breakpoints of translocations and inversions as a step toward identifying breakpoints in bread wheat, we systematically analysed the structures of the reported translocations and inversions by comparing *Brachypodium* genes against survey sequences of

sorted wheat chromosome arms and validating locations of selected genes using bin-mapped ESTs.

Materials and methods

It is well known that genome segments are highly conserved between wheat and *Brachypodium* although small disruptions of colinearity are not uncommon (Griffiths et al. 2006; Hernandez et al. 2012; Vogel et al. 2010). Analyses conducted in this study focused on those chromosomal rearrangements evidenced by two or more genes with the same pattern of chromosome arm locations. Coding DNA sequences (CDS) from *Brachypodium* genome version 1.2 were downloaded from <http://www.plantgdb.org/BdGDB>. The locations of orthologs to the genes of *Brachypodium* within the survey sequence data of wheat chromosome arms from 'CS' were determined using the BLAST++ facility of the International Wheat Genome Sequencing Consortium (IWGSC) (<http://www.wheatgenome.org/>) hosted by URGI (<http://urgi.versailles.inra.fr/>) (International Wheat Genome Sequencing Consortium 2014). The BLASTN algorithm was applied for all analyses using an E value cut-off of 0.0001. Data based on comparison of *Brachypodium* genes with deletion bin-mapped wheat ESTs were used to determine the relative positions and orientations of orthologs within segments of chromosome arms. Wheat ESTs for individual deletion bins were downloaded from <http://wheat.pw.usda.gov/GG2/index.shtml>.

It is believed that *Ae. tauschii* and *T. urartu* are donors of the D and A genome of hexaploid wheat (McFadden and Sears 1946; Peng et al. 2011). To enhance our analysis, their CDS have been employed to detect the chromosome locations as well. Scaffold sequences of *Ae. tauschii* (Jia et al. 2013) and *T. urartu* (Ling et al. 2013) were retrieved from NCBI's TSA data base (<http://blast.ncbi.nlm.nih.gov>). CDS of *Ae. tauschii* (wheatD_final_43150.gff.cds) and *T. urartu* (TRIUR3_120813_filter150_cds)

were downloaded from GIGA_DB (<http://gigadb.org/>). By comparing CDS against the above retrieved genome sequences, the genes mapped on the given scaffold were retrieved for detecting chromosome locations. The locations of orthologs to the genes of *Ae. tauschii* and *T. urartu* within the survey sequence data of wheat chromosome arms were determined as described before. Unless especially noted, all analyses of sequence comparisons used the BLASTN algorithm with an E value cut-off of 0.00001.

As a consequence of the high degree of rearrangement on what is considered the modern chromosome 4A, its arm ratio has been reversed (Devos et al. 1995; Liu et al. 1992; Mickelson-Young et al. 1995a; Miftahudin. et al. 2004; Nelson et al. 1995). As a result, discussion of the various historical states of this chromosome can become difficult to understand. To alleviate this confusion, this manuscript uses 4AS and 4AL to refer to the arms of the original ancestral version of this chromosome, while the modern chromosome arms are referred to as '4AS' and '4AL'.

Results

Chromosomal locations of genes on the modern 2B

Brachypodium orthologs on chromosome 2BS could be placed into two sets. Genes in Set 1 had homoeologous sequences on 2BS, 2AS and 2DS, respectively. The pattern of these chromosome arm locations shows that they are from the original 2BS and were not involved in any interchromosomal translocations. Genes in this set were orthologous to genes on *Brachypodium* chromosome 5 with *Bradi5g00570* as the most likely gene bordering the breakpoint (Table 1 and S1; Fig.1A).

Genes in Set 2 detected homoeologous sequences on 2BS, 2AL and 2DL, respectively. The pattern of these chromosome arm locations shows that these genes were derived from the original 2BL. Genes in this set were orthologous to genes on *Brachypodium* chromosomes 1 with *Bradi1g55930* bordering the breakpoint (Table 1 and S1; Fig.1A).

Genes on chromosome 2BL detected homoeologous sequences on 2BL, 2AL, and 2DL, respectively. The pattern of these chromosome arm locations shows that they were not involved in any interchromosomal translocations. Genes on this arm were orthologous to genes on two *Brachypodium* chromosomes, 1 and 5, with *Bradi1g55960* bordering the breakpoint (Table 1 and S1; Fig.1A).

Chromosomal locations of genes on the modern 4B

Brachypodium orthologs on chromosome 4BL were placed into two sets. Genes in Set 1 detected homoeologous sequences on 4BS, '4AL' and 4DS, respectively. The pattern of these chromosome arm locations shows that these genes belong to the original 4BS. Genes in this set were orthologous to those on *Brachypodium* chromosome 4. *Bradi4g09050* can be conservatively assigned as the one bordering the breakpoint (Table1 and S2; Fig.1B).

Most of genes in Set 2 detected homoeologous sequences on 4BS, '4AS' and 4DL, respectively. The pattern of these chromosome arm locations shows that these genes were derived from the original 4BL. A few genes in Set 2 had homoeologous sequences on 4BS, '4AL' and 4DL, respectively. It is not difficult to understand such kind of location pattern as a pericentric inversion event was previously reported to happen in 4A, resulting in transferring of a small segment proximal to the centromere on original 4AL to the modern '4AL' (Ma et al. 2013). Thus, this pattern of chromosome arm locations also indicates that these genes were derived from the original 4BL. Genes in Set 2 were orthologous to those on *Brachypodium* chromosome 1, with *Bradi1g62860* boarding the breakpoint (Table1 and S2; Fig.1B).

Most of genes on chromosome 4BL had homoeologous sequences on 4BL, '4AS', and 4DL, respectively. As the 4AL/5AL reciprocal translocation happened in diploid (Devos et al. 1995), a few of genes detected homoeologous sequences on 4BL, 5AL and 4DL. Hence, both of these two patterns of chromosome arm locations shows that they were from the original 4BL and not involved in any interchromosomal translocations. Genes on this arm were orthologous to genes on two *Brachypodium* chromosomes 1 with *Bradi1g62890* bordering the breakpoint (Table 1 and S2; Fig.1B).

Chromosomal locations of genes on the modern 5A

Brachypodium orthologs on the modern chromosome arm 5AS could be placed into two sets based on chromosomal locations of wheat sequences they detect. Genes in Set 1 detected sequences on 5AS, 5BS and 5DS, respectively. This pattern of the chromosome arm locations shows that they were derived from the original 5AS and were not involved in any interchromosomal translocations. Genes in this set have orthologs on *Brachypodium* chromosomes 4 with *Bradi4g06510* as the most likely gene bordering the breakpoint (Table 1 and S3; Fig.1C).

Genes in Set 2 detected homoeologous sequences on 5AS, 5BL, 5BS and 5DL, respectively. The pattern of these chromosome arm locations shows that these genes were derived from the original 5AL. These genes detected sequences on both 5BL and 5BS, indicating that they are duplicated on 5BS. Genes in this set were orthologous to genes on *Brachypodium* chromosomes 4 with *Bradi4g41560* bordering the breakpoint (Table 1 and S3; Fig.1C).

Brachypodium orthologs on the modern chromosome arm 5AL could be placed into two sets as well. Genes in Set 1 detected sequences on 5AL, 5BL and 5DL, respectively. This pattern of the chromosome arm locations shows that they were derived from the original 5AL and were not involved in any interchromosomal translocations. Genes in this set have orthologs on two *Brachypodium* chromosomes, 1 and 4, with *Bradi4g41697* and *Bradi1g03330* boarding the breakpoints proximal and distal to the centromere, respectively (Table 1 and S3; Fig.1C)(Ma et al. 2013).

Genes in Set 2 had homoeologous sequences on 5AL, 4BL and 4DL, respectively, being previously reported to be involved into reciprocal translocation between 5AL/4AL(Ma et al. 2013). Genes in this set have orthologs on three *Brachypodium* chromosomes, 1, 4 and 5, with *Bradi1g75560* boarding the breakpoint (Table S3, Fig. 1C)(Ma et al. 2013).

Discussion

To identify genes flanking the breakpoints of translocations and inversions as a step toward identifying breakpoints in bread wheat, we systematically analysed genes in the *Brachypodium* genome against wheat survey sequences and bin-mapped ESTs derived from the hexaploid wheat genotype 'Chinese Spring'. In addition to those well-known translocations between group 4, 5 and 7 chromosomes, this analysis identified genes flanking the 3 pericentric

inversions on chromosomes 2B, 4B and 5A. However, numerous chromosome rearrangements reported in early studies could not be confirmed in this systematic analysis. Thus attempts to detect genes flanking these untrue chromosome rearrangements are failure. (Fig. 1).

By using a cDNA clone PSR899 resided on 2BS, a 2BS/6BS translocation has been proposed in both hexaploid wheat and tetraploid wheat (*T. turgidum* L.) (Blanco et al. 1998; Devos et al. 1993). This supposed translocation was further verified by the bin-mapped EST, BE604879; and it was proposed to be on the distal region of 2BS collinear with 6AS5-0.65-1.00 and 6DS6-0.99-1.00 (Conley et al. 2004). Sequence comparison between PSR899 and BE604879 indicated that BE604879 actually belongs to partial sequence of PSR899. In this study, *Brachypodium* ortholog of BE604879 cannot be detected; but the *Ae. tauschii* and *T. urartu* orthologs were detected as *AEGTA32343* and *TRIUR3_28537*, respectively. Both *AEGTA32343* and *TRIUR3_28537* had homoeologous sequences on 2BS, 6AS and 6DS. However, attempts to detect one more genes with same pattern of chromosome arm locations (2BS, 6AS and 6DS) have been unsuccessful so far. For example, we cannot detect any orthologs on 2BS, 6AS and 6DS from bin1 on the distal region of chromosome 6DL in *Ae. tauschii* and orthologs from deletion of 6AS5-0.65-1.00 on chromosome 6AS in *T. urartu* (Tables S4 and S5). Further, the pattern of chromosome arm locations, 6BS, 2AS and 2DS, should be detected if the reciprocal translocation event between 2BS and 6BS indeed happened. Nevertheless, our efforts combined reported results failed to detect any genes or/and orthologs on 6BS, 2AS and 2DS so far. These results indicated that the chromosome location pattern of PSR899 (BE604879) is not likely caused by the translocation between 2BS and 6BS, but more like an example of non-collinearity among three different genomes, A, B and D in the hexaploid wheat. It is not uncommon to detect such a non-collinearity.

For example, we detected that *Bradi3g09210* has homoeologous sequences on chromosome 6AS, 6BL and 6DL and *Bradi3g08450* on 6AS, 6DS and 4AL; but the flanking genes did not show the same patterns of chromosome location as these two genes, indicating that they are probably not involved into any chromosome rearrangements (Table S6 and S7). Another example is the possible 3BL/6AS translocation detected by a genomic probe (PSR904) (Devos et al. 1992). However, our analysis of the *Brachypodium* genes failed to detect such patterns of chromosome location as either '3AL, 6AS and 3DL' or '3BL, 6BS and 6DS'. These examples indicated that chromosome rearrangements usually happened with transferring of a segment involved in a set of genes other than a single gene. This is the reason that our analyses in this study only focused on those chromosomal rearrangements evidenced by two or more genes with the same pattern of chromosome arm locations.

Linkiewicz et al (Linkiewicz et al. 2004) proposed that there might be a reciprocal inversion between 5AS and 5AL. The reason is that they had detected one EST (BE403618) being mapped on 5AL and 5BS, although not in the same proximal bin (C-5AL12 and 5BS4); and they suggested that the breakpoint on the 5AL was more distal than that on 5AS (Linkiewicz et al. 2004). BE403618 matched *Brachypodium* gene *Bradi4g39600*. However, *Brachypodium* orthologs on either side of this gene failed to detect a clear pattern of chromosome arm location, i.e. 5AL, 5BS and 5DS (Table S8). Thus the question whether there is a reciprocal inversion between 5AS and 5AL remains unanswered.

Previous analysis of bin-mapped ESTs identified quite a few anomalies in group 1 chromosomes. Of them, one was most probably involved in either a pericentric inversion between long arm and short arm or gene duplication; and it was that four ESTs from the telomeric bin of 1DL2-0.41-1.00 on chromosome 1DL detected homoeologous sequences on 1BL and 1AS (Peng et al. 2004).

However, in our present study, all of these four ESTs together with the genes on either side of their matched *Brachypodium* orthologs failed to detect any homoeologous sequence on 1AS but 1AL, 1BL and 1DL (Table S9), indicating that they are most likely involved in gene duplication not inversion. In group 3 chromosomes, two anomalies were also supposed to be caused by either chromosome rearrangement or duplication. Two ESTs, BE424097 and BE442882, were mapped on 3BS and 3DL (Munkvold et al. 2004). In the current study, BE424097 detected homoeologous sequences on 3AS, 3B and 3DS; and none of the genes on either side of its matched *Brachypodium* ortholog (*Bradi3g19960*) could detect sequence on 3DL (Table S10). The other one EST (BE442882) and a few of in-continuous *Brachypodium* orthologs detected sequences on (3AL), 3B and 3DL (Table S10). Given the unclear matching chromosome pattern, we suggested that these genes are more possibly involved in gene duplication other than chromosome rearrangement.

Two additional ESTs, BE405291 and BF202610, were mapped on 3BL and 3DS (Munkvold et al. 2004). In this study, BE405291 detected sequences on 3B, 3DL and the other 19 chromosome arms. BF202610 detected sequence on 3AL, 3AS, 3DS, 3DL, 3B and other three chromosome arms (Table S10). Although the *Brachypodium* orthologs to neither of these two EST can be detected, the multiple locations of them undoubtedly confirm that they belong to duplications other than chromosome rearrangement.

Acknowledgements

This publication is based upon work supported by a project from the International Science & Cooperation Program of China (No. 2015DFA30600). JM is grateful to the Sichuan Agricultural University and the China Scholarship Council for funding his visit to CSIRO Plant Industry. The authors are also very grateful to Dr. Chunji Liu (CSIRO Agriculture Flagship, Australia) for his critical reading of the manuscript during the preparation of this publication.

References

- Blanco, A., Bellomo, M., Cenci, A., De Giovanni, C., D'ovidio, R., Iacono, E., Laddomada, B., Pagnotta, M., Porceddu, E., and Sciancalepore, A. 1998. A genetic linkage map of durum wheat. *Theor. Appl. Genet.* **97**(5-6): 721-728.
- Brown, J.D., and O'Neill, R.J. 2010. Chromosomes, conflict, and epigenetics: chromosomal speciation revisited. *Annu. Rev. Genomics Hum. Genet.* **11**: 291-316.
- Burt, D.W., Bruley, C., Dunn, I.C., Jones, C.T., Ramage, A., Law, A.S., Morrice, D.R., Paton, I.R., Smith, J., and Windsor, D. 1999. The dynamics of chromosome evolution in birds and mammals. *Nature* **402**(6760): 411-412.
- Colson, I., Delneri, D., and Oliver, S.G. 2004. Effects of reciprocal chromosomal translocations on the fitness of *Saccharomyces cerevisiae*. *EMBO reports* **5**(4): 392-398.
- Conley, E., Nduati, V., Gonzalez-Hernandez, J., Mesfin, A., Trudeau-Spanjers, M., Chao, S., Lazo, G., Hummel, D., Anderson, O., and Qi, L. 2004. A 2600-locus chromosome bin map of wheat homoeologous group 2 reveals interstitial gene-rich islands and colinearity with rice. *Genetics* **168**(2): 625-637.
- Devos, K., Atkinson, M., Chinoy, C., Liu, C., and Gale, M. 1992. RFLP-based genetic map of the homoeologous group 3 chromosomes of wheat and rye. *Theor. Appl. Genet.* **83**(8): 931-939.
- Devos, K., Dubcovsky, J., Dvořák, J., Chinoy, C., and Gale, M. 1995. Structural evolution of wheat chromosomes 4A, 5A, and 7B and its impact on recombination. *Theor. Appl. Genet.* **91**(2): 282-288.
- Devos, K., Millan, T., and Gale, M. 1993. Comparative RFLP maps of the homoeologous group-2 chromosomes of wheat, rye and barley. *Theor. Appl. Genet.* **85**(6-7): 784-792.
- Dvořák, J., McGuire, P.E., and Mendlinger, S. 1984. Inferred chromosome morphology of the ancestral genome of *Triticum*. *Plant Syst. Evol.* **144**(3-4): 209-220.
- Endo, T.R., and Gill, B.S. 1984. Somatic karyotype, heterochromatin distribution, and nature of chromosome differentiation in common wheat, *Triticum aestivum* L. em Thell. *Chromosoma* **89**(5): 361-369.

- Gill, B., Friebe, B., and Endo, T. 1991. Standard karyotype and nomenclature system for description of chromosome bands and structural aberrations in wheat (*Triticum aestivum*). *Genome* **34**(5): 830-839.
- Griffiths, S., Sharp, R., Foote, T.N., Bertin, I., Wanous, M., Reader, S., Colas, I., and Moore, G. 2006. Molecular characterization of *Ph1* as a major chromosome pairing locus in polyploid wheat. *Nature* **439**(7077): 749-752.
- Hernandez, P., Martis, M., Dorado, G., Pfeifer, M., Gálvez, S., Schaaf, S., Jouve, N., Šimková, H., Valárik, M., and Doležel, J. 2012. Next-generation sequencing and syntenic integration of flow-sorted arms of wheat chromosome 4A exposes the chromosome structure and gene content. *Plant J.* **69**: 377-386.
- International Wheat Genome Sequencing Consortium. 2014. A chromosome-based draft sequence of the hexaploid bread wheat (*Triticum aestivum*) genome. *Science* **345**: 1251788.
- Jia, J., Zhao, S., Kong, X., Li, Y., Zhao, G., He, W., Appels, R., Pfeifer, M., Tao, Y., and Zhang, X. 2013. *Aegilops tauschii* draft genome sequence reveals a gene repertoire for wheat adaptation. *Nature* **496**: 91-95.
- Ling, H.-Q., Zhao, S., Liu, D., Wang, J., Sun, H., Zhang, C., Fan, H., Li, D., Dong, L., and Tao, Y. 2013. Draft genome of the wheat A-genome progenitor *Triticum urartu*. *Nature* **496**: 87-90.
- Linkiewicz, A., Qi, L., Gill, B., Ratnasiri, A., Echaliier, B., Chao, S., Lazo, G., Hummel, D., Anderson, O., and Akhunov, E. 2004. A 2500-locus bin map of wheat homoeologous group 5 provides insights on gene distribution and colinearity with rice. *Genetics* **168**(2): 665-676.
- Liu, C., Atkinson, M., Chinoy, C., Devos, K., and Gale, M. 1992. Nonhomoeologous translocations between group 4, 5 and 7 chromosomes within wheat and rye. *Theor. Appl. Genet.* **83**(3): 305-312.
- Ma, J., Stiller, J., Berkman, P.J., Wei, Y., Rogers, J., Feuillet, C., Dolezel, J., Mayer, K.F., Eversole, K., and Zheng, Y.-L. 2013. Sequence-based analysis of translocations and inversions in bread wheat (*Triticum aestivum* L.). *PloS one* **8**(11): e79329.
- Ma, J., Stiller, J., Wei, Y., Zheng, Y.-L., Devos, K.M., Doležel, J., and Liu, C. 2014. Extensive Pericentric Rearrangements in the Bread Wheat (*Triticum aestivum* L.) Genotype “Chinese Spring” Revealed from Chromosome Shotgun Sequence Data. *Genome Biol. Evol.* **6**(11): 3039-3048.
- Ma, J., Stiller, J., Zheng, Z., Wei, Y., Zheng, Y.-L., Yan, G., Doležel, J., and Liu, C. 2015. Putative interchromosomal rearrangements in the hexaploid wheat (*Triticum aestivum* L.) genotype ‘Chinese Spring’ revealed by gene locations on homoeologous chromosomes. *BMC Evol. Biol.* **15**(1): 37.
- McFadden, E., and Sears, E. 1946. The origin of *Triticum spelta* and its free-threshing hexaploid relatives. *J. Hered.* **37**(4): 107-116.
- McKim, K.S., Howell, A.M., and Rose, A.M. 1988. The effects of translocations on recombination frequency in *Caenorhabditis elegans*. *Genetics* **120**(4): 987-1001.
- Mickelson-Young, L., Endo, T., and Gill, B. 1995a. A cytogenetic ladder-map of the wheat homoeologous group-4 chromosomes. *Theor. Appl. Genet.* **90**(7): 1007-1011.

- Mickelson-Young, L., Endo, T., and Gill, B. 1995b. A cytogenetic ladder-map of the wheat homoeologous group-4 chromosomes. *Theor. Appl. Genet.* **90**(7-8): 1007-1011.
- Miftahudin., Ross, K., Ma, X.F., Mahmoud, A., Layton, J., Milla, M.A.R., Chikmawati, T., Ramalingam, J., Feril, O., Pathan, M., and Momirovic, G.S. 2004. Analysis of expressed sequence tag loci on wheat chromosome group 4. *Genetics* **168**(2): 651-663.
- Morrow, J.D., and Cooper, V.S. 2012. Evolutionary effects of translocations in bacterial genomes. *Genome Biol. Evol.* **4**(12): 1256-1262.
- Munkvold, J., Greene, R., Bermudez-Kandianis, C., La Rota, C., Edwards, H., Sorrells, S., Dake, T., Benscher, D., Kantety, R., and Linkiewicz, A. 2004. Group 3 chromosome bin maps of wheat and their relationship to rice chromosome 1. *Genetics* **168**(2): 639-650.
- Naranjo, T., Roca, A., Goicoechea, P., and Giraldez, R. 1987. Arm homoeology of wheat and rye chromosomes. *Genome* **29**(6): 873-882.
- Nelson, J.C., Sorrells, M.E., Van-Deynze, A., Lu, Y.H., Atkinson, M., Bernard, M., Leroy, P., Faris, J.D., and Anderson, J.A. 1995. Molecular mapping of wheat: major genes and rearrangements in homoeologous groups 4, 5, and 7. *Genetics* **141**(2): 721.
- Peng, J., Zadeh, H., Lazo, G., Gustafson, J., Chao, S., Anderson, O., Qi, L., Echaliier, B., Gill, B., and Dilbirli, M. 2004. Chromosome bin map of expressed sequence tags in homoeologous group 1 of hexaploid wheat and homoeology with rice and Arabidopsis. *Genetics* **168**(2): 609-623.
- Peng, J.H., Sun, D., and Nevo, E. 2011. Domestication evolution, genetics and genomics in wheat. *Mol. Breed.* **28**(3): 281-301.
- Sankoff, D., and Nadeau, J.H. 2003. Chromosome rearrangements in evolution: From gene order to genome sequence and back. *Proc. Natl. Acad. Sci. USA* **100**(20): 11188-11189.
- Sherizen, D., Jang, J.K., Bhagat, R., Kato, N., and McKim, K.S. 2005. Meiotic recombination in *Drosophila* females depends on chromosome continuity between genetically defined boundaries. *Genetics* **169**(2): 767-781.
- Vogel, J.P., Garvin, D.F., Mockler, T.C., Schmutz, J., Rokhsar, D., Bevan, M.W., Barry, K., Lucas, S., Harmon-Smith, M., and Lail, K. 2010. Genome sequencing and analysis of the model grass *Brachypodium distachyon*. *Nature* **463**(7282): 763-768.

Figure legend

Fig. 1. The structures of modern chromosomes 2B (**A**), 4B (**B**) and 5A (**C**) and their evolutionary origins.

Pericentric inversion events resulted in these structures are marked as a: 2BL/2BS inversion; b: 4BL/4BS inversion; c: 5AL/5AS inversion. Loci defining these inversion breakpoints and the genes proximal to the telomere are indicated.

Draft

Table 1. *Brachypodium* orthologs bordering inversion breakpoints on chromosome arms 2A, 4B, and 5A

Chromosome arm	Segments#	Border genes##
2BS	2BS <i>BP</i> 2BL	... <i>BP</i> Bradi1g55930 <
2BL	C <i>BP</i> 2BL	C <i>BP</i> Bradi1g55960 >
4BS	4BS <i>BP</i> 4BL	... <i>BP</i> Bradi1g62860 <
4BL	C <i>BP</i> 4 BL	C <i>BP</i> Bradi1g62890 >
5AS	5AS <i>BP</i> 5AL	... <i>BP</i> Bradi4g41560 <
5AL	C <i>BP</i> 5AL	C <i>BP</i> Bradi4g41697 >

#: Definitions of these chromosome segments are provided in Figure 1. Breakpoints are represented by ‘BP’. ## Orientations of genes bordering each of the breakpoints are indicated by ‘>’ (increasing gene IDs) or ‘<’ (decreasing gene IDs). Thus ‘BP Bradi1g55930 <’ indicates that ID numbers for genes moving away from the breakpoint decrease, and ‘BP Bradi1g55960 >’ indicates that ID numbers for genes moving away from the breakpoint increase. ‘...’ indicates that the genes boarding the ‘BP’ cannot be determined. Centromere is represented by ‘C’.

1 **Identification of genes bordering breakpoints of the pericentric**
2 **inversions on 2B, 4B and 5A in Bread Wheat (*Triticum aestivum* L.)**

3 Jian Ma^{1, #}, Shang Gao^{1, #}, Jiri Stiller², Qian-Tao Jiang^{1, *}, Xiu-Jin Lan¹, Ya-Xi
4 Liu¹, Zhi-En Pu¹, Jirui Wang¹, Yuming Wei¹, You-Liang Zheng^{1, *}

5

6

7 ¹*Triticeae Research Institute, Sichuan Agricultural University, Wenjiang,*
8 *Chengdu 611130, China*

9 ²*CSIRO Agriculture Flagship, 306 Carmody Road, St Lucia, QLD 4067,*
10 *Australia*

11

12 # These authors contributed equally to this work.

13

14 *Authors for correspondence:

15 Dr. Qian-Tao Jiang:

16 E-mail: qiantaojiang@sicau.edu.cn; Tel: +86 28 86290958; Fax: +86 28
17 82650350

18 Prof. You-Liang Zheng:

19 E-mail: ylzheng@sicau.edu.cn; Tel: +86 835 2882007; Fax: +86 835 2883153

20

21

1 **Abstract**

2 Chromosome translocation is an important driving force in shaping genomes
3 during evolution. Detailed knowledge of chromosome translocations in a given
4 species and its close relatives should increase the efficiency and precision of
5 chromosome engineering in crop improvement. To identify genes flanking the
6 breakpoints of translocations and inversions as a step toward identifying
7 breakpoints in bread wheat, we systematically analysed genes in the
8 *Brachypodium* genome against wheat survey sequences and bin-mapped
9 ESTs (Expressed Sequence Tags) derived from the hexaploid wheat genotype
10 'Chinese Spring'. In addition to those well-known translocations between group
11 4, 5 and 7 chromosomes, this analysis identified genes flanking the 3
12 pericentric inversions on chromosomes 2B, 4B and 5A. However, numerous
13 chromosomal rearrangements reported in early studies could not be confirmed.
14 The genes flanking the breakpoints reported in this study are valuable for
15 isolating these breakpoints.

16 **Keywords** genome evolution, bread wheat, inversion, translocation breakpoint,
17 *Brachypodium*

18

19

1 Introduction

2 It is widely believed that chromosome translocation is an important driving
3 force in shaping genomes during evolution (Brown and O'Neill 2010; Burt et al.
4 1999; Colson et al. 2004; Morrow and Cooper 2012; Sankoff and Nadeau
5 2003). In human(Brown and O'Neill 2010), animals(Burt et al. 1999) and
6 microbes(Colson et al. 2004; Morrow and Cooper 2012), chromosome
7 translocation is usually related to genomic disorders. Some studies also
8 reported that translocation could alter levels of recombination (McKim et al.
9 1988; Sherizen et al. 2005) which is not only a major source of intra-specific
10 variation but also an important constraint in crop improvement programs. Such
11 programs aim to bring together multiple chromosomal segments containing
12 favourable alleles into single plant lines.

13 It is well known that the non-homoeologous translocations between
14 chromosomes 4A, 5A and 7B are important translocations in the hexaploid
15 wheat (*Triticum aestivum* L., $2n=6x=42$, genome AABBDD) (Devos et al. 1995;
16 Liu et al. 1992; Naranjo et al. 1987; Nelson et al. 1995). Further, for the first
17 time, by analysing *Brachypodium* genes against survey sequences of sorted
18 wheat chromosome arms and by analysing *Brachypodium* genes against
19 deletion bin-mapped wheat ESTs, we have previously identified *Brachypodium*
20 orthologs bordering several translocation and inversion breakpoints on the
21 modern wheat chromosomes 4A, 5A and 7B(Ma et al. 2013).

22 In addition to these major translocation events, many other likely chromosomal
23 rearrangements have been reported in the hexaploid wheat: 1) a pericentric
24 inversion in 4B evidenced by cytology (Dvořák et al. 1984; Endo and Gill 1984;
25 Gill et al. 1991), gene-based marker (Mickelson-Young et al. 1995b), and
26 bin-mapped expressed sequence tags (ESTs) (Miftahudin. et al. 2004); 2) a
27 non-homoeologous translocations between chromosomes 2BS and 6BS
28 evidenced by cDNA clone (Devos et al. 1993) and bin-mapped ESTs (Conley

1 et al. 2004); 3) a pericentric inversion in 2B deduced by bin-mapped ESTs
2 (Conley et al. 2004); 4) a putative pericentric inversion in 5A deduced by
3 bin-mapped ESTs (Linkiewicz et al. 2004); 5) some possible rearrangements
4 in chromosome group 1 and 3 deduced by bin-mapped ESTs (Munkvold et al.
5 2004; Peng et al. 2004). Further, by analysing homoeoloci for a set of
6 nonredundant genes from *Brachypodium distachyon*, *T. urartu*, and *Aegilops*
7 *tauschii* in the 'CS' chromosome shotgun sequence obtained from individual
8 chromosome arms flow-sorted from 'CS' aneuploid lines, we systematically
9 assessed chromosomal rearrangements in the hexaploid wheat genotype
10 'Chinese Spring' ('CS') (Ma et al. 2014; Ma et al. 2015).

11 Despite of numerous deduced rearrangements being identified, identification
12 of the breakpoints for these translocations and inversions has been so far
13 unsuccessful. Recent development in wheat genome sequencing (International
14 Wheat Genome Sequencing Consortium 2014) offers an excellent opportunity
15 to develop methods that enable rapid and reliable detection of these
16 translocations and inversions. This can be achieved by firstly identifying
17 genomic sequences covering each of these breakpoints and then developing
18 PCR primers that allow the amplification of these sequences. Once such PCR
19 primers become available, cloning and sequencing a PCR product will be all
20 required to detect the presence or absence of any of these translocations and
21 inversions in a given genotype. Thus large numbers of genotypes representing
22 different types of wheats and their close relatives can be assessed, making it
23 possible to evaluate the effects of these translocations in wheat evolution.

24 Aiming at finding genes flanking the breakpoints of translocations and
25 inversions as a step toward identifying breakpoints in bread wheat, we
26 systematically analysed the structures of the reported translocations and
27 inversions by comparing *Brachypodium* genes against survey sequences of

1 sorted wheat chromosome arms and validating locations of selected genes
 2 using bin-mapped ESTs.

3 **Materials and methods**

4 It is well known that genome segments are highly conserved between wheat
 5 and *Brachypodium* although small disruptions of colinearity are not uncommon
 6 (Griffiths et al. 2006; Hernandez et al. 2012; Vogel et al. 2010). Analyses
 7 conducted in this study focused on those chromosomal rearrangements
 8 evidenced by two or more genes with the same pattern of chromosome arm
 9 locations. Coding DNA sequences (CDS) from *Brachypodium* genome version
 10 1.2 were downloaded from <http://www.plantgdb.org/BdGDB>. The locations of
 11 orthologs to the genes of *Brachypodium* within the survey sequence data of
 12 wheat chromosome arms from 'CS' were determined using the BLAST++
 13 facility of the International Wheat Genome Sequencing Consortium (IWGSC)
 14 (<http://www.wheatgenome.org/>) hosted by URGI
 15 (<http://urgi.versailles.inra.fr>)([International](http://urgi.versailles.inra.fr) Wheat Genome Sequencing
 16 Consortium 2014). The BLASTN algorithm was applied for all analyses using
 17 an E value cut-off of 0.0001. Data based on comparison of *Brachypodium*
 18 genes with deletion bin-mapped wheat ESTs were used to determine the
 19 relative positions and orientations of orthologs within segments of
 20 chromosome arms. Wheat ESTs for individual deletion bins were downloaded
 21 from <http://wheat.pw.usda.gov/GG2/index.shtml>.

22 It is believed that *Ae. tauschii* and *T. urartu* are donors of the D and A genome
 23 of hexaploid wheat (McFadden and Sears 1946; Peng et al. 2011). To
 24 enhance our analysis, their CDS have been employed to detect the
 25 chromosome locations as well. Scaffold sequences of *Ae. tauschii*(Jia et al.
 26 2013) and *T. urartu* (Ling et al. 2013) were retrieved from NCBI's TSA data
 27 base (<http://blast.ncbi.nlm.nih.gov>). CDS of *Ae. tauschii*
 28 (wheatD_final_43150.gff.cds) and *T. urartu* (TRIUR3_120813_filter150_cds)

1 were downloaded from GIGA_DB (<http://gigadb.org/>). By comparing CDS
2 against the above retrieved genome sequences, the genes mapped on the
3 given scaffold were retrieved for detecting chromosome locations. The
4 locations of orthologs to the genes of *Ae. tauschii* and *T. urartu* within the
5 survey sequence data of wheat chromosome arms were determined as
6 described before. Unless especially noted, all analyses of sequence
7 comparisons used the BLASTN algorithm with an E value cut-off of 0.00001.

8 As a consequence of the high degree of rearrangement on what is considered
9 the modern chromosome 4A, its arm ratio has been reversed (Devos et al.
10 1995; Liu et al. 1992; Mickelson-Young et al. 1995a; Miftahudin. et al. 2004;
11 Nelson et al. 1995). As a result, discussion of the various historical states of
12 this chromosome can become difficult to understand. To alleviate this
13 confusion, this manuscript uses 4AS and 4AL to refer to the arms of the
14 original ancestral version of this chromosome, while the modern chromosome
15 arms are referred to as '4AS' and '4AL'.

16

1 **Results**

2 Chromosomal locations of genes on the modern 2B

3 *Brachypodium* orthologs on chromosome 2BS could be placed into two sets.
4 Genes in Set 1 had homoeologous sequences on 2BS, 2AS and 2DS,
5 respectively. The pattern of these chromosome arm locations shows that they
6 are from the original 2BS and were not involved in any interchromosomal
7 translocations. Genes in this set were orthologous to genes on *Brachypodium*
8 chromosome 5 with *Bradi5g00570* as the most likely gene bordering the
9 breakpoint (Table 1 and S1; Fig.1A).

10 Genes in Set 2 detected homoeologous sequences on 2BS, 2AL and 2DL,
11 respectively. The pattern of these chromosome arm locations shows that these
12 genes were derived from the original 2BL. Genes in this set were orthologous
13 to genes on *Brachypodium* chromosomes 1 with *Bradi1g55930* bordering the
14 breakpoint (Table 1 and S1; Fig.1A).

15 Genes on chromosome 2BL detected homoeologous sequences on 2BL, 2AL,
16 and 2DL, respectively. The pattern of these chromosome arm locations shows
17 that they were not involved in any interchromosomal translocations. Genes on
18 this arm were orthologous to genes on two *Brachypodium* chromosomes, 1
19 and 5, with *Bradi1g55960* bordering the breakpoint (Table 1 and S1; Fig.1A).

20 Chromosomal locations of genes on the modern 4B

21 *Brachypodium* orthologs on chromosome 4BL were placed into two sets.
22 Genes in Set 1 detected homoeologous sequences on 4BS, '4AL' and 4DS,
23 respectively. The pattern of these chromosome arm locations shows that these
24 genes belong to the original 4BS. Genes in this set were orthologous to those
25 on *Brachypodium* chromosome 4. *Bradi4g09050* can be conservatively
26 assigned as the one bordering the breakpoint (Table1 and S2; Fig.1B).

1 Most of genes in Set 2 detected homoeologous sequences on 4BS, '4AS' and
2 4DL, respectively. The pattern of these chromosome arm locations shows that
3 these genes were derived from the original 4BL. A few genes in Set 2 had
4 homoeologous sequences on 4BS, '4AL' and 4DL, respectively. It is not
5 difficult to understand such kind of location pattern as a pericentric inversion
6 event was previously reported to happen in 4A, resulting in transferring of a
7 small segment proximal to the centromere on original 4AL to the modern
8 '4AL'(Ma et al. 2013). Thus, this pattern of chromosome arm locations also
9 indicates that these genes were derived from the original 4BL. Genes in Set 2
10 were orthologous to those on *Brachypodium* chromosome 1, with
11 *Bradi1g62860* boarding the breakpoint (Table1 and S2; Fig.1B).

12 Most of genes on chromosome 4BL had homoeologous sequences on 4BL,
13 '4AS', and 4DL, respectively. As the 4AL/5AL reciprocal translocation
14 happened in diploid (Devos et al. 1995), a few of genes detected
15 homoeologous sequences on 4BL, 5AL and 4DL. Hence, both of these two
16 patterns of chromosome arm locations shows that they were from the original
17 4BL and not involved in any interchromosomal translocations. Genes on this
18 arm were orthologous to genes on two *Brachypodium* chromosomes 1 with
19 *Bradi1g62890* bordering the breakpoint (Table 1 and S2; Fig.1B).

20 Chromosomal locations of genes on the modern 5A

21 *Brachypodium* orthologs on the modern chromosome arm 5AS could be
22 placed into two sets based on chromosomal locations of wheat sequences
23 they detect. Genes in Set 1 detected sequences on 5AS, 5BS and 5DS,
24 respectively. This pattern of the chromosome arm locations shows that they
25 were derived from the original 5AS and were not involved in any
26 interchromosomal translocations. Genes in this set have orthologs on
27 *Brachypodium* chromosomes 4 with *Bradi4g06510* as the most likely gene
28 bordering the breakpoint (Table 1 and S3; Fig.1C).

1 Genes in Set 2 detected homoeologous sequences on 5AS, 5BL, 5BS and
 2 5DL, respectively. The pattern of these chromosome arm locations shows that
 3 these genes were derived from the original 5AL. These genes detected
 4 sequences on both 5BL and 5BS, indicating that they are duplicated on 5BS.
 5 Genes in this set were orthologous to genes on *Brachypodium* chromosomes
 6 4 with *Bradi4g41560* bordering the breakpoint (Table 1 and S3; Fig.1C).

7 *Brachypodium* orthologs on the modern chromosome arm 5AL could be
 8 placed into two sets as well. Genes in Set 1 detected sequences on 5AL, 5BL
 9 and 5DL, respectively. This pattern of the chromosome arm locations shows
 10 that they were derived from the original 5AL and were not involved in any
 11 interchromosomal translocations. Genes in this set have orthologs on two
 12 *Brachypodium* chromosomes, 1 and 4, with *Bradi4g41697* and *Bradi1g03330*
 13 boarding the breakpoints proximal and distal to the centromere, respectively
 14 (Table 1 and S3; Fig.1C)(Ma et al. 2013).

15 Genes in Set 2 had homoeologous sequences on 5AL, 4BL and 4DL,
 16 respectively, being previously reported to be involved into reciprocal
 17 translocation between 5AL/4AL(Ma et al. 2013). Genes in this set have
 18 orthologs on three *Brachypodium* chromosomes, 1, 4 and 5, with
 19 *Bradi1g75560* boarding the breakpoint (Table S3, Fig. 1C)(Ma et al. 2013).

20

21 Discussion

22 To identify genes flanking the breakpoints of translocations and inversions as a
 23 step toward identifying breakpoints in bread wheat, we systematically analysed
 24 genes in the *Brachypodium* genome against wheat survey sequences and
 25 bin-mapped ESTs derived from the hexaploid wheat genotype 'Chinese
 26 Spring'. In addition to those well-known translocations between group 4, 5 and
 27 7 chromosomes, this analysis identified genes flanking the 3 pericentric

1 inversions on chromosomes 2B, 4B and 5A. However, numerous chromosome
2 rearrangements reported in early studies could not be confirmed in this
3 systematic analysis. Thus attempts to detect genes flanking these untrue
4 chromosome rearrangements are failure. (Fig. 1).

5 By using a cDNA clone PSR899 resided on 2BS, a 2BS/6BS translocation has
6 been proposed in both hexaploid wheat and tetraploid wheat (*T. turgidum* L.)
7 (Blanco et al. 1998; Devos et al. 1993). This supposed translocation was
8 further verified by the bin-mapped EST, BE604879; and it was proposed to be
9 on the distal region of 2BS collinear with 6AS5-0.65-1.00 and 6DS6-0.99-1.00
10 (Conley et al. 2004). Sequence comparison between PSR899 and BE604879
11 indicated that BE604879 actually belongs to partial sequence of PSR899. In
12 this study, *Brachypodium* ortholog of BE604879 cannot be detected; but the
13 *Ae. tauschii* and *T. urartu* orthologs were detected as *AEGTA32343* and
14 *TRIUR3_28537*, respectively. Both *AEGTA32343* and *TRIUR3_28537* had
15 homoeologous sequences on 2BS, 6AS and 6DS. However, attempts to detect
16 one more genes with same pattern of chromosome arm locations (2BS, 6AS
17 and 6DS) have been unsuccessful so far. For example, we cannot detect any
18 orthologs on 2BS, 6AS and 6DS from bin1 on the distal region of chromosome
19 6DL in *Ae. tauschii* and orthologs from deletion of 6AS5-0.65-1.00 on
20 chromosome 6AS in *T. urartu* (Tables S4 and S5). Further, the pattern of
21 chromosome arm locations, 6BS, 2AS and 2DS, should be detected if the
22 reciprocal translocation event between 2BS and 6BS indeed happened.
23 Nevertheless, our efforts combined reported results failed to detect any genes
24 or/and orthologs on 6BS, 2AS and 2DS so far. These results indicated that the
25 chromosome location pattern of PSR899 (BE604879) is not likely caused by
26 the translocation between 2BS and 6BS, but more like an example of
27 non-collinearity among three different genomes, A, B and D in the hexaploid
28 wheat. It is not uncommon to detect such a non-collinearity.

1 For example, we detected that *Bradi3g09210* has homoeologous sequences
2 on chromosome 6AS, 6BL and 6DL and *Bradi3g08450* on 6AS, 6DS and 4AL;
3 but the flanking genes did not show the same patterns of chromosome location
4 as these two genes, indicating that they are probably not involved into any
5 chromosome rearrangements (Table S6 and S7). Another example is the
6 possible 3BL/6AS translocation detected by a genomic probe (PSR904)
7 (Devos et al. 1992). However, our analysis of the *Brachypodium* genes failed
8 to detect such patterns of chromosome location as either '3AL, 6AS and 3DL'
9 or '3BL, 6BS and 6DS'. These examples indicated that chromosome
10 rearrangements usually happened with transferring of a segment involved in a
11 set of genes other than a single gene. This is the reason that our analyses in
12 this study only focused on those chromosomal rearrangements evidenced by
13 two or more genes with the same pattern of chromosome arm locations.

14 Linkiewicz et al (Linkiewicz et al. 2004) proposed that there might be a
15 reciprocal inversion between 5AS and 5AL. The reason is that they had
16 detected one EST (BE403618) being mapped on 5AL and 5BS, although not in
17 the same proximal bin (C-5AL12 and 5BS4); and they suggested that the
18 breakpoint on the 5AL was more distal than that on 5AS (Linkiewicz et al.
19 2004). BE403618 matched *Brachypodium* gene *Bradi4g39600*. However,
20 *Brachypodium* orthologs on either side of this gene failed to detect a clear
21 pattern of chromosome arm location, i.e. 5AL, 5BS and 5DS (Table S8). Thus
22 the question whether there is a reciprocal inversion between 5AS and 5AL
23 remains unanswered.

24 Previous analysis of bin-mapped ESTs identified quite a few anomalies in
25 group 1 chromosomes. Of them, one was most probably involved in either a
26 pericentric inversion between long arm and short arm or gene duplication; and
27 it was that four ESTs from the telomeric bin of 1DL2-0.41-1.00 on chromosome
28 1DL detected homoeologous sequences on 1BL and 1AS (Peng et al. 2004).

1 However, in our present study, all of these four ESTs together with the genes
2 on either side of their matched *Brachypodium* orthologs failed to detect any
3 homoeologous sequence on 1AS but 1AL, 1BL and 1DL (Table S9), indicating
4 that they are most likely involved in gene duplication not inversion. In group 3
5 chromosomes, two anomalies were also supposed to be caused by either
6 chromosome rearrangement or duplication. Two ESTs, BE424097 and
7 BE442882, were mapped on 3BS and 3DL (Munkvold et al. 2004). In the
8 current study, BE424097 detected homoeologous sequences on 3AS, 3B and
9 3DS; and none of the genes on either side of its matched *Brachypodium*
10 ortholog (*Bradi3g19960*) could detect sequence on 3DL (Table S10). The
11 other one EST (BE442882) and a few of in-continuous *Brachypodium*
12 orthologs detected sequences on (3AL), 3B and 3DL (Table S10). Given the
13 unclear matching chromosome pattern, we suggested that these genes are
14 more possibly involved in gene duplication other than chromosome
15 rearrangement.

16 Two additional ESTs, BE405291 and BF202610, were mapped on 3BL and
17 3DS (Munkvold et al. 2004). In this study, BE405291 detected sequences on
18 3B, 3DL and the other 19 chromosome arms. BF202610 detected sequence
19 on 3AL, 3AS, 3DS, 3DL, 3B and other three chromosome arms (Table S10).
20 Although the *Brachypodium* orthologs to neither of these two EST can be
21 detected, the multiple locations of them undoubtedly confirm that they belong
22 to duplications other than chromosome rearrangement.

23

1 **Acknowledgements**

2 This publication is based upon work supported by a project from the
3 International Science & Cooperation Program of China (No. 2015DFA30600).
4 JM is grateful to the Sichuan Agricultural University and the China Scholarship
5 Council for funding his visit to CSIRO Plant Industry. The authors are also very
6 grateful to Dr. Chunji Liu (CSIRO Agriculture Flagship, Australia) for his critical
7 reading of the manuscript during the preparation of this publication.

8

9 **References**

10 Blanco, A., Bellomo, M., Cenci, A., De Giovanni, C., D’ovidio, R., Iacono, E., Laddomada, B.,
11 Pagnotta, M., Porceddu, E., and Sciancalepore, A. 1998. A genetic linkage map of
12 durum wheat. *Theor. Appl. Genet.* **97**(5-6): 721-728.
13 Brown, J.D., and O'Neill, R.J. 2010. Chromosomes, conflict, and epigenetics: chromosomal
14 speciation revisited. *Annu. Rev. Genomics Hum. Genet.* **11**: 291-316.
15 Burt, D.W., Bruley, C., Dunn, I.C., Jones, C.T., Ramage, A., Law, A.S., Morrice, D.R., Paton, I.R.,
16 Smith, J., and Windsor, D. 1999. The dynamics of chromosome evolution in birds and
17 mammals. *Nature* **402**(6760): 411-412.
18 Colson, I., Delneri, D., and Oliver, S.G. 2004. Effects of reciprocal chromosomal
19 translocations on the fitness of *Saccharomyces cerevisiae*. *EMBO reports* **5**(4): 392-398.
20 Conley, E., Nduati, V., Gonzalez-Hernandez, J., Mesfin, A., Trudeau-Spanjers, M., Chao, S.,
21 Lazo, G., Hummel, D., Anderson, O., and Qi, L. 2004. A 2600-locus chromosome bin
22 map of wheat homoeologous group 2 reveals interstitial gene-rich islands and
23 colinearity with rice. *Genetics* **168**(2): 625-637.
24 Devos, K., Atkinson, M., Chinoy, C., Liu, C., and Gale, M. 1992. RFLP-based genetic map of the
25 homoeologous group 3 chromosomes of wheat and rye. *Theor. Appl. Genet.* **83**(8):
26 931-939.
27 Devos, K., Dubcovsky, J., Dvořák, J., Chinoy, C., and Gale, M. 1995. Structural evolution of
28 wheat chromosomes 4A, 5A, and 7B and its impact on recombination. *Theor. Appl.*
29 *Genet.* **91**(2): 282-288.
30 Devos, K., Millan, T., and Gale, M. 1993. Comparative RFLP maps of the homoeologous
31 group-2 chromosomes of wheat, rye and barley. *Theor. Appl. Genet.* **85**(6-7): 784-792.
32 Dvořák, J., McGuire, P.E., and Mendlinger, S. 1984. Inferred chromosome morphology of the
33 ancestral genome of *Triticum*. *Plant Syst. Evol.* **144**(3-4): 209-220.
34 Endo, T.R., and Gill, B.S. 1984. Somatic karyotype, heterochromatin distribution, and nature
35 of chromosome differentiation in common wheat, *Triticum aestivum* L. em Thell.
36 *Chromosoma* **89**(5): 361-369.

- 1 Gill, B., Friebe, B., and Endo, T. 1991. Standard karyotype and nomenclature system for
2 description of chromosome bands and structural aberrations in wheat (*Triticum*
3 *aestivum*). *Genome* **34**(5): 830-839.
- 4 Griffiths, S., Sharp, R., Foote, T.N., Bertin, I., Wanous, M., Reader, S., Colas, I., and Moore, G.
5 2006. Molecular characterization of *Ph1* as a major chromosome pairing locus in
6 polyploid wheat. *Nature* **439**(7077): 749-752.
- 7 Hernandez, P., Martis, M., Dorado, G., Pfeifer, M., Gálvez, S., Schaaf, S., Jouve, N., Šimková,
8 H., Valárik, M., and Doležel, J. 2012. Next-generation sequencing and syntenic
9 integration of flow-sorted arms of wheat chromosome 4A exposes the chromosome
10 structure and gene content. *Plant J.* **69**: 377-386.
- 11 International Wheat Genome Sequencing Consortium. 2014. A chromosome-based draft
12 sequence of the hexaploid bread wheat (*Triticum aestivum*) genome. *Science* **345**:
13 1251788.
- 14 Jia, J., Zhao, S., Kong, X., Li, Y., Zhao, G., He, W., Appels, R., Pfeifer, M., Tao, Y., and Zhang, X.
15 2013. *Aegilops tauschii* draft genome sequence reveals a gene repertoire for wheat
16 adaptation. *Nature* **496**: 91-95.
- 17 Ling, H.-Q., Zhao, S., Liu, D., Wang, J., Sun, H., Zhang, C., Fan, H., Li, D., Dong, L., and Tao, Y.
18 2013. Draft genome of the wheat A-genome progenitor *Triticum urartu*. *Nature* **496**:
19 87-90.
- 20 Linkiewicz, A., Qi, L., Gill, B., Ratnasiri, A., Echaliier, B., Chao, S., Lazo, G., Hummel, D.,
21 Anderson, O., and Akhunov, E. 2004. A 2500-locus bin map of wheat homoeologous
22 group 5 provides insights on gene distribution and colinearity with rice. *Genetics*
23 **168**(2): 665-676.
- 24 Liu, C., Atkinson, M., Chinoy, C., Devos, K., and Gale, M. 1992. Nonhomoeologous
25 translocations between group 4, 5 and 7 chromosomes within wheat and rye. *Theor.*
26 *Appl. Genet.* **83**(3): 305-312.
- 27 Ma, J., Stiller, J., Berkman, P.J., Wei, Y., Rogers, J., Feuillet, C., Dolezel, J., Mayer, K.F.,
28 Eversole, K., and Zheng, Y.-L. 2013. Sequence-based analysis of translocations and
29 inversions in bread wheat (*Triticum aestivum* L.). *PloS one* **8**(11): e79329.
- 30 Ma, J., Stiller, J., Wei, Y., Zheng, Y.-L., Devos, K.M., Doležel, J., and Liu, C. 2014. Extensive
31 Pericentric Rearrangements in the Bread Wheat (*Triticum aestivum* L.) Genotype
32 “Chinese Spring” Revealed from Chromosome Shotgun Sequence Data. *Genome Biol.*
33 *Evol.* **6**(11): 3039-3048.
- 34 Ma, J., Stiller, J., Zheng, Z., Wei, Y., Zheng, Y.-L., Yan, G., Doležel, J., and Liu, C. 2015. Putative
35 interchromosomal rearrangements in the hexaploid wheat (*Triticum aestivum* L.)
36 genotype ‘Chinese Spring’ revealed by gene locations on homoeologous chromosomes.
37 *BMC Evol. Biol.* **15**(1): 37.
- 38 McFadden, E., and Sears, E. 1946. The origin of *Triticum spelta* and its free-threshing
39 hexaploid relatives. *J. Hered.* **37**(4): 107-116.
- 40 McKim, K.S., Howell, A.M., and Rose, A.M. 1988. The effects of translocations on
41 recombination frequency in *Caenorhabditis elegans*. *Genetics* **120**(4): 987-1001.
- 42 Mickelson-Young, L., Endo, T., and Gill, B. 1995a. A cytogenetic ladder-map of the wheat
43 homoeologous group-4 chromosomes. *Theor. Appl. Genet.* **90**(7): 1007-1011.

1 Mickelson-Young, L., Endo, T., and Gill, B. 1995b. A cytogenetic ladder-map of the wheat
2 homoeologous group-4 chromosomes. *Theor. Appl. Genet.* **90**(7-8): 1007-1011.

3 Miftahudin., Ross, K., Ma, X.F., Mahmoud, A., Layton, J., Milla, M.A.R., Chikmawati, T.,
4 Ramalingam, J., Feril, O., Pathan, M., and Momirovic, G.S. 2004. Analysis of expressed
5 sequence tag loci on wheat chromosome group 4. *Genetics* **168**(2): 651-663.

6 Morrow, J.D., and Cooper, V.S. 2012. Evolutionary effects of translocations in bacterial
7 genomes. *Genome Biol. Evol.* **4**(12): 1256-1262.

8 Munkvold, J., Greene, R., Bermudez-Kandianis, C., La Rota, C., Edwards, H., Sorrells, S., Dake,
9 T., Benscher, D., Kantety, R., and Linkiewicz, A. 2004. Group 3 chromosome bin maps
10 of wheat and their relationship to rice chromosome 1. *Genetics* **168**(2): 639-650.

11 Naranjo, T., Roca, A., Goicoechea, P., and Giraldez, R. 1987. Arm homoeology of wheat and
12 rye chromosomes. *Genome* **29**(6): 873-882.

13 Nelson, J.C., Sorrells, M.E., Van-Deynze, A., Lu, Y.H., Atkinson, M., Bernard, M., Leroy, P.,
14 Faris, J.D., and Anderson, J.A. 1995. Molecular mapping of wheat: major genes and
15 rearrangements in homoeologous groups 4, 5, and 7. *Genetics* **141**(2): 721.

16 Peng, J., Zadeh, H., Lazo, G., Gustafson, J., Chao, S., Anderson, O., Qi, L., Echallier, B., Gill, B.,
17 and Dilbirligi, M. 2004. Chromosome bin map of expressed sequence tags in
18 homoeologous group 1 of hexaploid wheat and homoeology with rice and Arabidopsis.
19 *Genetics* **168**(2): 609-623.

20 Peng, J.H., Sun, D., and Nevo, E. 2011. Domestication evolution, genetics and genomics in
21 wheat. *Mol. Breed.* **28**(3): 281-301.

22 Sankoff, D., and Nadeau, J.H. 2003. Chromosome rearrangements in evolution: From gene
23 order to genome sequence and back. *Proc. Natl. Acad. Sci. USA* **100**(20): 11188-11189.

24 Sherizen, D., Jang, J.K., Bhagat, R., Kato, N., and McKim, K.S. 2005. Meiotic recombination in
25 *Drosophila* females depends on chromosome continuity between genetically defined
26 boundaries. *Genetics* **169**(2): 767-781.

27 Vogel, J.P., Garvin, D.F., Mockler, T.C., Schmutz, J., Rokhsar, D., Bevan, M.W., Barry, K., Lucas,
28 S., Harmon-Smith, M., and Lail, K. 2010. Genome sequencing and analysis of the model
29 grass *Brachypodium distachyon*. *Nature* **463**(7282): 763-768.

30

31

1 **Figure legend**

2 **Fig. 1.** The structures of modern chromosomes 2B (**A**), 4B (**B**) and 5A (**C**) and
3 their evolutionary origins.

4 Pericentric inversion events resulted in these structures are marked as a:
5 2BL/2BS inversion; b: 4BL/4BS inversion; c: 5AL/5AS inversion. Loci defining
6 these inversion breakpoints and the genes proximal to the telomere are
7 indicated.

Draft

1 Table 1. *Brachypodium* orthologs bordering inversion breakpoints on
2 chromosome arms 2A, 4B, and 5A

Chromosome arm	Segments#	Border genes##
2BS	2BS <i>BP</i> 2BL	... <i>BP</i> Bradi1g55930 <
2BL	C <i>BP</i> 2BL	C <i>BP</i> Bradi1g55960 >
4BS	4BS <i>BP</i> 4BL	... <i>BP</i> Bradi1g62860 <
4BL	C <i>BP</i> 4 BL	C <i>BP</i> Bradi1g62890 >
5AS	5AS <i>BP</i> 5AL	... <i>BP</i> Bradi4g41560 <
5AL	C <i>BP</i> 5AL	C <i>BP</i> Bradi4g41697 >

3 #: Definitions of these chromosome segments are provided in Figure 1.
4 Breakpoints are represented by '*BP*'. ## Orientations of genes bordering each
5 of the breakpoints are indicated by '>' (increasing gene IDs) or '<' (decreasing
6 gene IDs). Thus '*BP* Bradi1g55930 <' indicates that ID numbers for genes
7 moving away from the breakpoint decrease, and '*BP* Bradi1g55960 >'
8 indicates that ID numbers for genes moving away from the breakpoint increase.
9 '...' indicates that the genes boarding the '*BP*' cannot be determined.
10 Centromere is represented by 'C'.
11