Ma *et al. Plant Methods (2015) 11:39*
DOI 10.1186/s13007-015-0082-6

PLANT METHODS

METHODOLOGY Open Access

A high-throughput pipeline for detecting locus-specific polymorphism in hexaploid wheat (*Triticum aestivum* L.)

Jian Ma[1,2], Jiri Stiller[2], Zhi Zheng[2,3,4], Ya-Xi Liu[1], Yuming Wei[1], You-Liang Zheng[1] and Chunji Liu[2,3*]

Abstract

Background: Bread wheat (*Triticum aestivum* L., $2n = 6x = 42$) is an allohexaploid with a huge genome. Due to the presence of extensive homoeologs and paralogs, generating locus-specific sequences can be challenging, especially when a large number of sequences are required. Traditional methods of generating locus-specific sequences are rather strenuous and time-consuming if large numbers of sequences are to be handled.

Results: To improve the efficiency of isolating sequences for targeted loci, a time-saving and high-throughput pipeline integrating orthologous sequence alignment, genomic sequence retrieving, and multiple sequence alignment was developed. This pipeline was successfully employed in retrieving and aligning homoeologous sequences and 83% of the primers designed based on the pipeline successfully amplified fragments from the targeted subgenomes.

Conclusions: The high-throughput pipeline developed in this study makes it feasible to efficiently identify locus-specific sequences for large numbers of sequences. It could find applications in all research projects where locus-specific sequences are required. In addition to generating locus-specific markers, the pipeline was also used in our laboratory to identify differentially expressed genes among the three subgenomes of bread wheat. Importantly, the pipeline is not only valuable for research in wheat but should also be applicable to other allopolyploid species.

Keywords: Allopolyploid, Multiple sequence alignment, blastn, Primer design, Genome-specificity, Sequence polymorphism

Background

Reference genome sequences of several major crops have been reported and include rice [1], barley [2], foxtail millet [3], maize [4], sorghum [5], potato [6], tomato [7] and *Brassica napus* [8]. Significant progress has also been made in recent years in generating reference genomes for bread wheat [9] and its progenitor species [10, 11]. These genome sequences have been extensively exploited in the whole spectrum of biological studies ranging from basic understanding of crop evolution to applied breeding. With the rapid development in sequencing capacity, it is anticipated that whole genome sequences should soon become available for multiple genotypes for each of the species of agronomic importance.

Knowing the origins of specific DNA or RNA sequences is essential in numerous applications, such as designing locus-specific markers. Although gene duplication is a common feature of all plant species including *Brachypodium* [12], rice [1], and barley [2], developing locus-specific markers for these diploid species is relatively easy. This suggests that enough variation must exist between the majority of duplicated genes in these species. However, isolating locus-specific sequences for a given subgenome of interest from bread wheat or other polyploid species is still challenging as two or more homoeologous sequences exist in each of these genomes. It can be even more daunting when isolating a gene of interest which belongs to an orthologous gene set or a gene family.

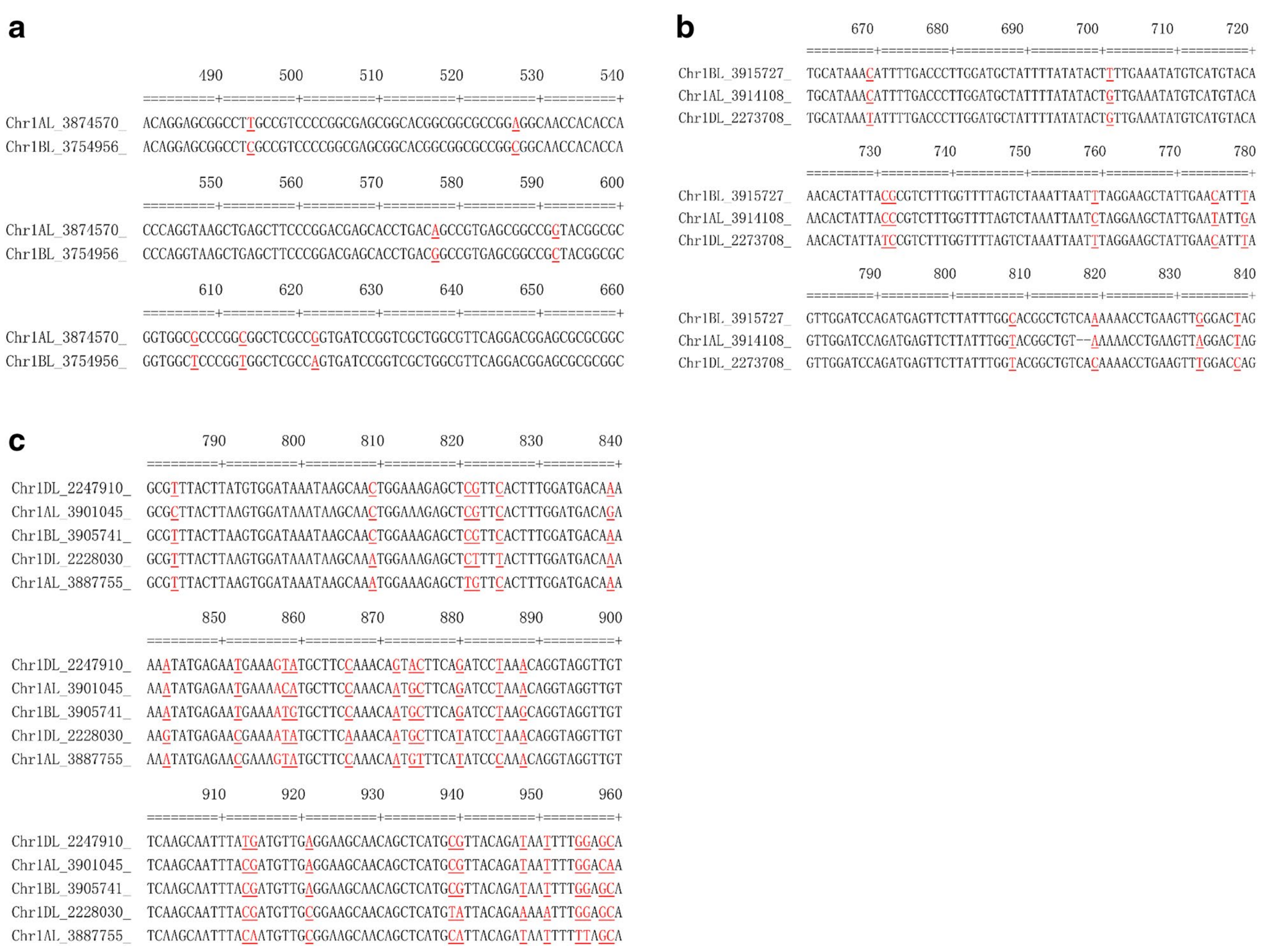
*Correspondence: chunji.liu@csiro.au
[3] School of Plant Biology, The University of Western Australia, Perth, WA 6009, Australia
Full list of author information is available at the end of the article

Currently several steps need to be taken when isolating a specific homoeologous sequence from an allopolyploid species. First, web-based blast servers such as National Center for Biotechnology Information (NCBI) or ViroBLAST in Unité de Recherche Génomique Info (URGI, https://urgi.versailles.inra.fr/blast/blast.php) [13] need to be employed to search for orthologous sequences for a given query sequence. Second, all the orthologous gene sequences for a given species need to be manually retrieved from contigs or scaffold sequences. Third, multiple sequence alignment tools are required to align retrieved orthologous sequences to detect locus-specific sequences. This procedure can be used to manage a limited number of sequences but will become rather strenuous and time-consuming if large numbers of sequences need to be handled.

To improve the efficiency of retrieving sequences from polyploidy species, we have developed a pipeline by streamlining the steps in orthologous sequence alignment, genomic sequence retrieving and multiple

Table 1 Numbers of genes used in blasting against wheat chromosome shotgun sequences (CSSs)

	Brachypodium distachyon	*Triticum urartu*	*Aegilops tauschii*
Total numbers of genes	31,029	34,879	43,150
Genes with hits on CSSs	30,028	34,671	43,126
Genes with generated alignments	27,782	32,378	40,961



```
                          490       500       510       520       530       540
                   =========+=========+=========+=========+=========+=========+
Chr1AL_3874570_    ACAGGAGCGGCCTTGCCGTCCCCGGCGAGCGGCACGGCGGCGCCGGAGGCAACCACACCA
Chr1BL_3754956_    ACAGGAGCGGCCTCGCCGTCCCCGGCGAGCGGCACGGCGGCGCCGGCGGCAACCACACCA

                          550       560       570       580       590       600
                   =========+=========+=========+=========+=========+=========+
Chr1AL_3874570_    CCCAGGTAAGCTGAGCTTCCCGGACGAGCACCTGACAGCCGTGAGCGGCCGGTACGGCGC
Chr1BL_3754956_    CCCAGGTAAGCTGAGCTTCCCGGACGAGCACCTGACGGCCGTGAGCGGCCGCTACGGCGC

                          610       620       630       640       650       660
                   =========+=========+=========+=========+=========+=========+
Chr1AL_3874570_    GGTGGCGCCCGGCGGCTCGCCGGTGATCCGGTCGCTGGCGTTCAGGACGGAGCGCGCGGC
Chr1BL_3754956_    GGTGGCTCCCGGTGGCTCGCCAGTGATCCGGTCGCTGGCGTTCAGGACGGAGCGCGCGGC
```

```
                          670       680       690       700       710       720
                   =========+=========+=========+=========+=========+=========+
Chr1BL_3915727_    TGCATAAACATTTTGACCCTTGGATGCTATTTTATATACTTTTGAAATATGTCATGTACA
Chr1AL_3914108_    TGCATAAACATTTTGACCCTTGGATGCTATTTTATATACTGTTGAAATATGTCATGTACA
Chr1DL_2273708_    TGCATAAATATTTTGACCCTTGGATGCTATTTTATATACTGTTGAAATATGTCATGTACA

                          730       740       750       760       770       780
                   =========+=========+=========+=========+=========+=========+
Chr1BL_3915727_    AACACTATTACGCGTCTTTGGTTTTAGTCTAAATTAATTTAGGAAGCTATTGAACATTTA
Chr1AL_3914108_    AACACTATTACCCGTCTTTGGTTTTAGTCTAAATTAATCTAGGAAGCTATTGAATATTGA
Chr1DL_2273708_    AACACTATTATCCGTCTTTGGTTTTAGTCTAAATTAATTTAGGAAGCTATTGAACATTTA

                          790       800       810       820       830       840
                   =========+=========+=========+=========+=========+=========+
Chr1BL_3915727_    GTTGGATCCAGATGAGTTCTTATTTGGCACGGCTGTCAAAAAACCTGAAGTTGGGACTAG
Chr1AL_3914108_    GTTGGATCCAGATGAGTTCTTATTTGGTACGGCTGT--AAAAACCTGAAGTTAGGACTAG
Chr1DL_2273708_    GTTGGATCCAGATGAGTTCTTATTTGGTACGGCTGTCACAAAACCTGAAGTTTGGACCAG
```

```
                          790       800       810       820       830       840
                   =========+=========+=========+=========+=========+=========+
Chr1DL_2247910_    GCGTTTACTTATGTGGATAAATAAGCAACTGGAAAGAGCTCGTTCACTTTGGATGACAAA
Chr1AL_3901045_    GCGCTTACTTAAGTGGATAAATAAGCAACTGGAAAGAGCTCGTTCACTTTGGATGACAGA
Chr1BL_3905741_    GCGTTTACTTAAGTGGATAAATAAGCAACTGGAAAGAGCTCGTTCACTTTGGATGACAAA
Chr1DL_2228030_    GCGTTTACTTAAGTGGATAAATAAGCAAATGGAAAGAGCTCTTTTACTTTGGATGACAAA
Chr1AL_3887755_    GCGTTTACTTAAGTGGATAAATAAGCAAATGGAAAGAGCTTGTTCACTTTGGATGACAAA

                          850       860       870       880       890       900
                   =========+=========+=========+=========+=========+=========+
Chr1DL_2247910_    AAATATGAGAATGAAAGTATGCTTCCAAACAGTACTTCAGATCCTAAACAGGTAGGTTGT
Chr1AL_3901045_    AAATATGAGAATGAAAACATGCTTCCAAACAATGCTTCAGATCCTAAACAGGTAGGTTGT
Chr1BL_3905741_    AAATATGAGAATGAAAATGTGCTTCCAAACAATGCTTCAGATCCTAAGCAGGTAGGTTGT
Chr1DL_2228030_    AAGTATGAGAACGAAAATATGCTTCAAAACAATGCTTCATATCCTAAACAGGTAGGTTGT
Chr1AL_3887755_    AAATATGAGAACGAAAGTATGCTTCCAAACAATGTTTCATATCCCAAACAGGTAGGTTGT

                          910       920       930       940       950       960
                   =========+=========+=========+=========+=========+=========+
Chr1DL_2247910_    TCAAGCAATTTATGATGTTGAGGAAGCAACAGCTCATGCGTTACAGATAATTTTGGAGCA
Chr1AL_3901045_    TCAAGCAATTTACGATGTTGAGGAAGCAACAGCTCATGCGTTACAGATAATTTTGGACAA
Chr1BL_3905741_    TCAAGCAATTTACGATGTTGAGGAAGCAACAGCTCATGCGTTACAGATAATTTTGGAGCA
Chr1DL_2228030_    TCAAGCAATTTACGATGTTGCGGAAGCAACAGCTCATGTATTACAGAAAAATTTGGAGCA
Chr1AL_3887755_    TCAAGCAATTTACAATGTTGCGGAAGCAACAGCTCATGCATTACAGATAATTTTTAGCA
```

Fig. 1 Examples of multiple alignments for two (**a**), three (**b**), and five (**c**) homoeologous sequences for identifying allele-specific sequences for marker development.

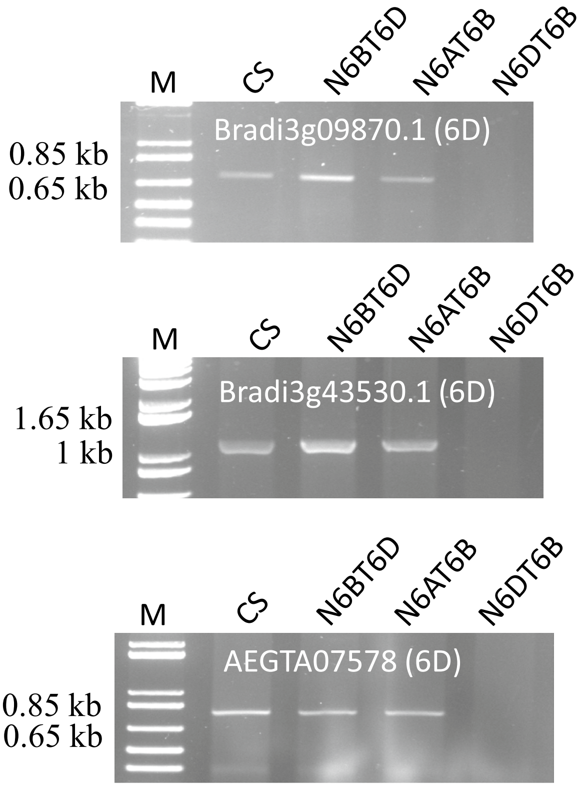


Fig. 2 PCR profiles of the hexaploid wheat genotype 'Chinese Spring' and its nulli-tetrasomic (NT) lines with primers for three genes assessed. Locations of the genes detected from the chromosome shotgun sequences are given in *brackets*. 1 kb plus DNA ladder (M) was used as the size marker.

sequence alignment. This time-saving and high-throughput pipeline significantly simplifies the detection of locus-specific sequences in allopolyploid species. The pipeline has also been successfully used in differentiating expressed genes among the three bread wheat subgenomes.

Results and discussion

The percentages of query gene sequences which detected two or more orthologous sequences from chromosome shotgun sequences (CSSs) were about 90% from *Brachypodium*, 93% from *Ae. tauschii* and 95% from *T. urartu* (Table 1). Examples of these stringent alignments containing two orthologous sequences (e.g. *Bradi2g16370.1*), three homoelogous sequences (e.g. *Bradi2g33190.1*), and more than four homoeologous sequences (e.g. *Bradi2g14840.1*) are shown in Fig. 1. Each of the alignments with the suffix '.htm' is easily readable by any web browser. The alignments generated and described in 'Methods' can be directly used to check possible allele-specific loci for isolating genes in hexaploid wheat.

Given that primers designed based on a single nucleotide polymorphisms (SNP) did not always amplify a specific fragment in our previous studies, primers designed in this study were based on two or more SNPs or indels (Additional file 1: Table S1). Of the 36 primer pairs designed for selected loci, 30 (83%) amplified a product on the expected chromosomes, two failed to amplify any PCR products, and the other four generated locus-specific fragments (Fig. 2 and Additional file 1: Table S1). Eleven of the 30 pairs of primers were further assessed against other bread wheat genotypes (Additional file 1: Table S1). Sequence alignments indicated that, without exception, they all amplified sequences homologous with those from the expected chromosomes as shown in 'Chinese Spring' ('CS') (Additional file 1: Table S1). Four of these primer pairs generated polymorphic fragments between the parents of the mapping population used in this study. The polymorphic sequences were used to develop cleaved amplified polymorphic sequence (CAPS) markers. Each of the four CAPS markers was successfully mapped to the anticipated chromosome as originally detected using 'CS' aneuploids (Fig. 3, Additional file 2: Fig. S1 and Additional file 3: Fig. S2).

A pipeline for generating SNP markers, PolyMarker, was reported recently and it is used to design primers for KASP™ (Kompetitive Allele Specific PCR) assay. KASP™ is a very unique system in that it uses three primers in PCR reactions. Two of them are allele-specific forward primers. Sequences from parental genotypes are required in designing the two forward primers which make accurate bi-allelic discrimination possible [14–18].

Different from the PolyMarker/KASPTM system, the method reported in this paper does not need sequences from parental genotypes. Allele-specific primers are designed based on sequence alignments from all subgenomes of a species. In addition to designing allele-specific markers as shown in the current study, we have also adapted the pipeline to design allele-specific primers for reverse transcription quantitative PCR (RT-qPCR) analysis in wheat. For example, several positions of orthologous sequences of *Bradi1g04060* could be used to design RT-qPCR primers in bread wheat (Fig. 4). We have also successfully used the pipeline to retrieve conserved regions that could be used for differentially expressed analysis in bread wheat (not published). Obviously, this high-throughput pipeline would be applicable to other allopolyploid species such as rapeseed, cotton, or oat.

Conclusion

Here we reported on a high-throughput pipeline which integrates orthologous sequence alignment, genomic sequences retrieval, and multiple sequence alignment. The pipeline can be conveniently used to identify

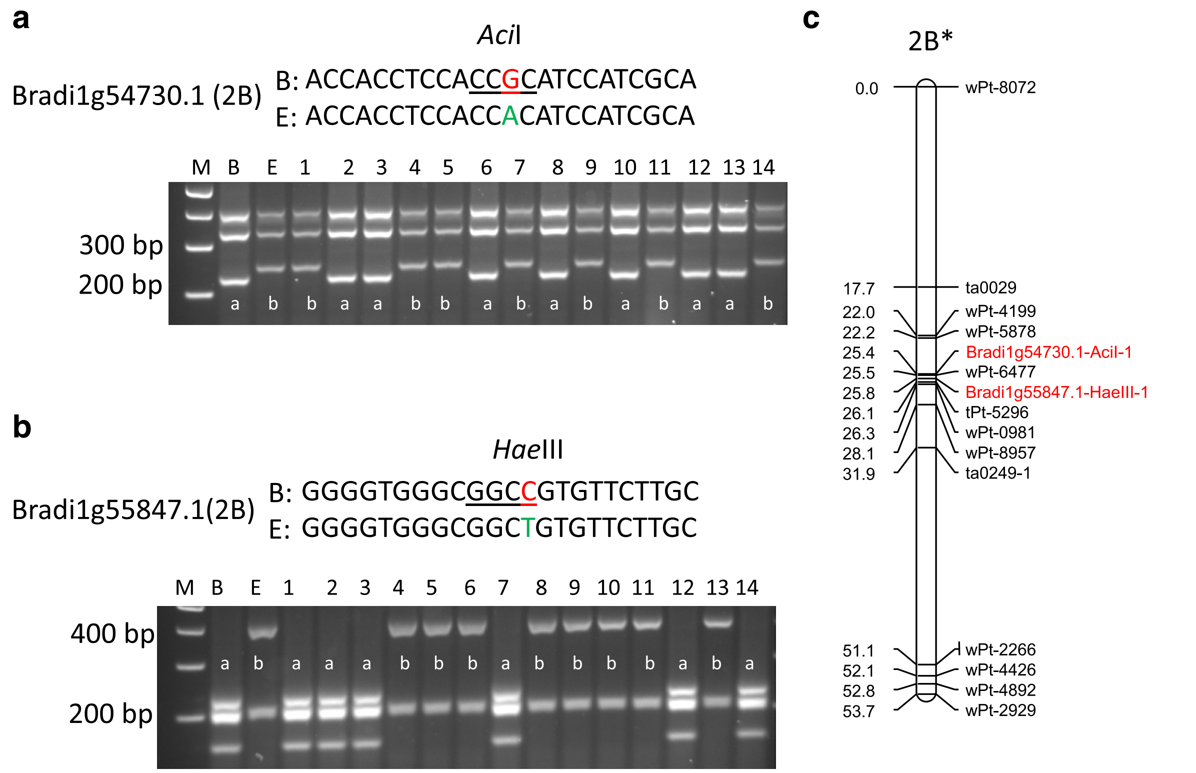

a
AciI
Bradi1g54730.1 (2B)
B: ACCACCTCCACCGCATCCATCGCA
E: ACCACCTCCACCACATCCATCGCA
M B E 1 2 3 4 5 6 7 8 9 10 11 12 13 14
300 bp
200 bp
a b b a a b b a b a b a b a a b
b
HaeIII
Bradi1g55847.1(2B)
B: GGGGTGGGCGGCCGTGTTCTTGC
E: GGGGTGGGCGGCTGTGTTCTTGC
M B E 1 2 3 4 5 6 7 8 9 10 11 12 13 14
400 bp
200 bp
a b a a a b b b a b b b b a b a
c
2B*
0.0 wPt-8072
17.7 ta0029
22.0 wPt-4199
22.2 wPt-5878
25.4 Bradi1g54730.1-AciI-1
25.5 wPt-6477
25.8 Bradi1g55847.1-HaeIII-1
26.1 tPt-5296
26.3 wPt-0981
28.1 wPt-8957
31.9 ta0249-1
51.1 wPt-2266
52.1 wPt-4426
52.8 wPt-4892
53.7 wPt-2929

Fig. 3 Validation of marker locations using a DH (doubled haploid) population. Orthologous sequences of *Bradi1g54730.1* and *Bradi1g55847.1* were amplified from the two parents of the DH population, B ('Batavia') and E ('Ernie'), and sequenced. The single nucleotide polymorphism (in *red* and *green*) and restriction enzyme sites (*underlined*) were identified between B and E for *Bradi1g54730.1* with restriction enzyme *Aci*I (**a**) and *Bradi1g55847.1* with *Hae*III (**b**). The amplified products of the two parents and 14 of the DH lines were digested and separated on agarose gels. Map positions of the two new markers on chromosome 2B (**c**) were calculated based on the linkage map published by Li et al. [21].

locus-specific sequences for marker development and RT-qPCR and transcriptome analyses, especially when large numbers of sequences need to be dealt with. Examples of its application in wheat are given in this publication but the pipeline would also be valuable for similar applications in other allopolyploid species as well.

Methods

Plant materials

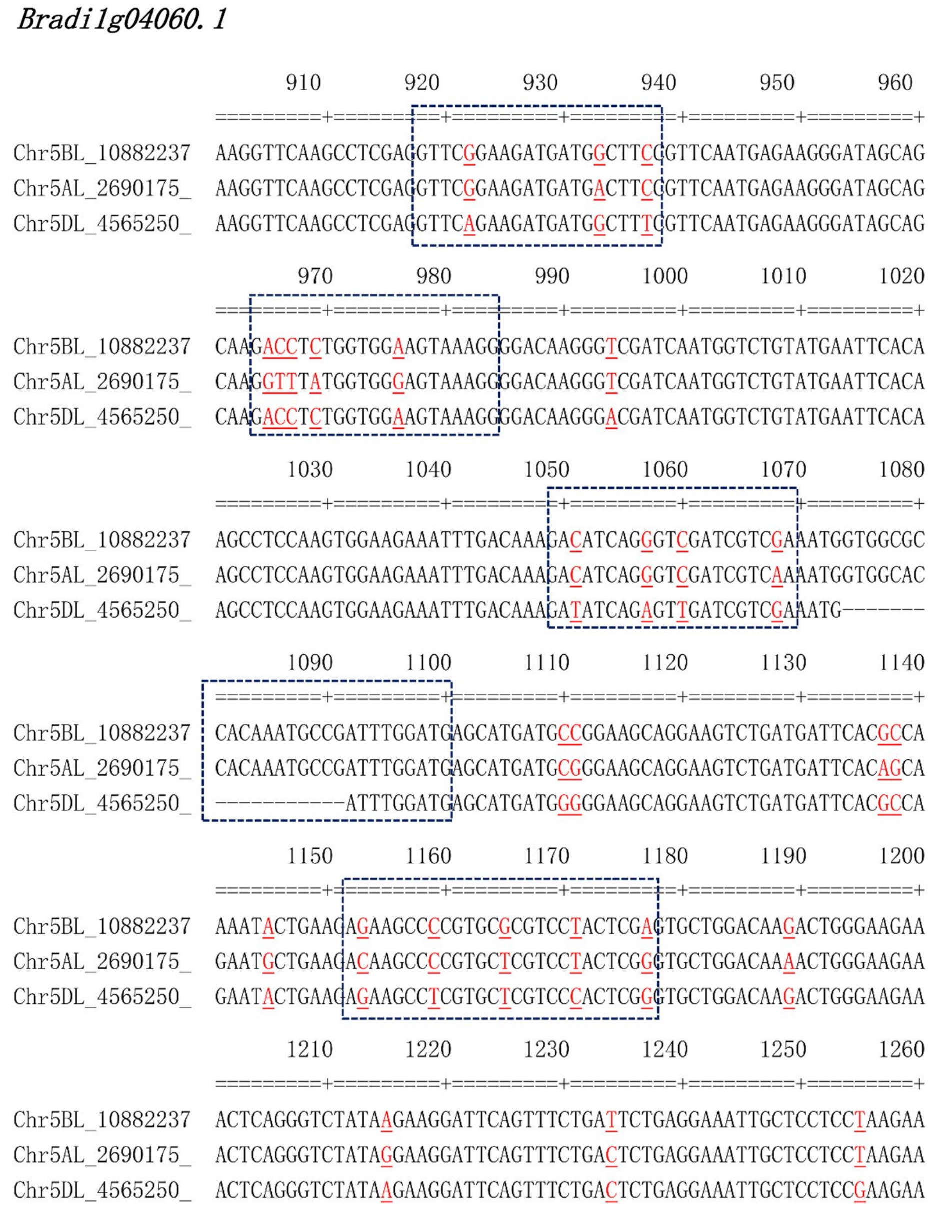
The euploid and selected nullisomic-tetrasomic 'CS' lines [19, 20] were used to locate PCR-amplified fragments to specific chromosomes. Two wheat populations were employed to further validate the location of DNA fragments amplified from primers designed in this study. One is a doubled haploid (DH) population with 153 lines generated from the 'Batavia'/'Ernie' cross [21], and the other one is an F8 population of recombinant inbred lines (RILs) with 92 lines derived from the 'Lang'/'CSCR6' cross [22].

Data collection

Gene-coding sequences (CDS) of *Brachypodium* genome version 1.2 were downloaded from http://www.plantgdb.org/BdGDB [12]. CDS of *Ae. tauschii* (wheatD_final_43150.gff.cds) [10] and *T. urartu* (TRIUR3_120813_filter150_cds) [11] were both downloaded from GIGA_DB (http://gigadb.org/). CSSs of 'CS' were downloaded from https://urgi.versailles.inra.fr/download/iwgsc/Science/ [9].

Generation of multiple sequence alignments and primer design

Alignments of orthologous sequences from *Brachypodium*, *Ae. tauschii*, and *T. urartu* were generated following the steps outlined in Fig. 5. First, gene sequences from *Brachypodium*, *Ae. tauschii*, and *T. urartu* were blasted against CSSs using the BLAST + blastn algorithm with the parameters '-num_descriptions 10 -num_alignments 10 -evalue 0.00001' (i.e. a maximum of 10 hits



Bradi1g04060.1

```
                        910       920       930       940       950       960
                 =========+=========+=========+=========+=========+=========+
Chr5BL_10882237  AAGGTTCAAGCCTCGAGGTTCGGAAGATGATGGCTTCCGTTCAATGAGAAGGGATAGCAG
Chr5AL_2690175_  AAGGTTCAAGCCTCGAGGTTCGGAAGATGATGACTTCCGTTCAATGAGAAGGGATAGCAG
Chr5DL_4565250_  AAGGTTCAAGCCTCGAGGTTCAGAAGATGATGGCTTTCGTTCAATGAGAAGGGATAGCAG

                        970       980       990      1000      1010      1020
                 =========+=========+=========+=========+=========+=========+
Chr5BL_10882237  CAAGACCTCTGGTGGAAGTAAAGGGGACAAGGGTCGATCAATGGTCTGTATGAATTCACA
Chr5AL_2690175_  CAAGGTTTATGGTGGGAGTAAAGGGGACAAGGGTCGATCAATGGTCTGTATGAATTCACA
Chr5DL_4565250_  CAAGACCTCTGGTGGAAGTAAAGGGGACAAGGGACGATCAATGGTCTGTATGAATTCACA

                       1030      1040      1050      1060      1070      1080
                 =========+=========+=========+=========+=========+=========+
Chr5BL_10882237  AGCCTCCAAGTGGAAGAAATTTGACAAAGACATCAGGGTCGATCGTCGAAATGGTGGCGC
Chr5AL_2690175_  AGCCTCCAAGTGGAAGAAATTTGACAAAGACATCAGGGTCGATCGTCAAAATGGTGGCAC
Chr5DL_4565250_  AGCCTCCAAGTGGAAGAAATTTGACAAAGATATCAGAGTTGATCGTCGAAATG-------

                       1090      1100      1110      1120      1130      1140
                 =========+=========+=========+=========+=========+=========+
Chr5BL_10882237  CACAAATGCCGATTTGGATGAGCATGATGCCGGAAGCAGGAAGTCTGATGATTCACGCCA
Chr5AL_2690175_  CACAAATGCCGATTTGGATGAGCATGATGCGGGAAGCAGGAAGTCTGATGATTCACAGCA
Chr5DL_4565250_  -----------ATTTGGATGAGCATGATGGGGGAAGCAGGAAGTCTGATGATTCACGCCA

                       1150      1160      1170      1180      1190      1200
                 =========+=========+=========+=========+=========+=========+
Chr5BL_10882237  AAATACTGAAGAGAAGCCCCGTGCGCGTCCTACTCGAGTGCTGGACAAGACTGGGAAGAA
Chr5AL_2690175_  GAATGCTGAAGACAAGCCCCGTGCTCGTCCTACTCGGGTGCTGGACAAAACTGGGAAGAA
Chr5DL_4565250_  GAATACTGAAGAGAAGCCTCGTGCTCGTCCCACTCGGGTGCTGGACAAGACTGGGAAGAA

                       1210      1220      1230      1240      1250      1260
                 =========+=========+=========+=========+=========+=========+
Chr5BL_10882237  ACTCAGGGTCTATAAGAAGGATTCAGTTTCTGATTCTGAGGAAATTGCTCCTCCTAAGAA
Chr5AL_2690175_  ACTCAGGGTCTATAGGAAGGATTCAGTTTCTGACTCTGAGGAAATTGCTCCTCCTAAGAA
Chr5DL_4565250_  ACTCAGGGTCTATAAGAAGGATTCAGTTTCTGACTCTGAGGAAATTGCTCCTCCGAAGAA
```

Fig. 4 An example of selecting primer sequences for qPCR analysis. *Dotted boxes* represent regions that could be used for primer design.

for each gene query and with E-value threshold of 10^{-5}) [23]. Second, an in-house script was used to retrieve the coordinates of each hit for a given gene query from the blast results. A maximum of 5,000 bp intron and minimum of 200 bp exon were used to limit the retrieved coordinates for a given hit. Third, the 5′ and 3′ flanking regions of 300 bp were isolated from each of the contigs (hits) according to the coordinates obtained. Fourth, the isolated genomic sequences from all the hits for a given query were written to a single file. Finally, a script integrated with Gblocks_0.91ba [24] and Clustal W 2.1 [25] was used to generate the alignments of all the retrieved genomic sequences for a given query (Fig. 5). The alignments and in-house developed scripts are available at http://dx.doi.org/10.6084/m9.figshare.1393103; http://dx.doi.org/10.6084/m9.figshare.1393106; http://dx.doi.org/10.6084/m9.figshare.1393105.

Validation of primers designed from the alignments

For validating the efficiency of the multiple sequence alignments generated, genes that were polymorphic between the parental lines of the mapping populations

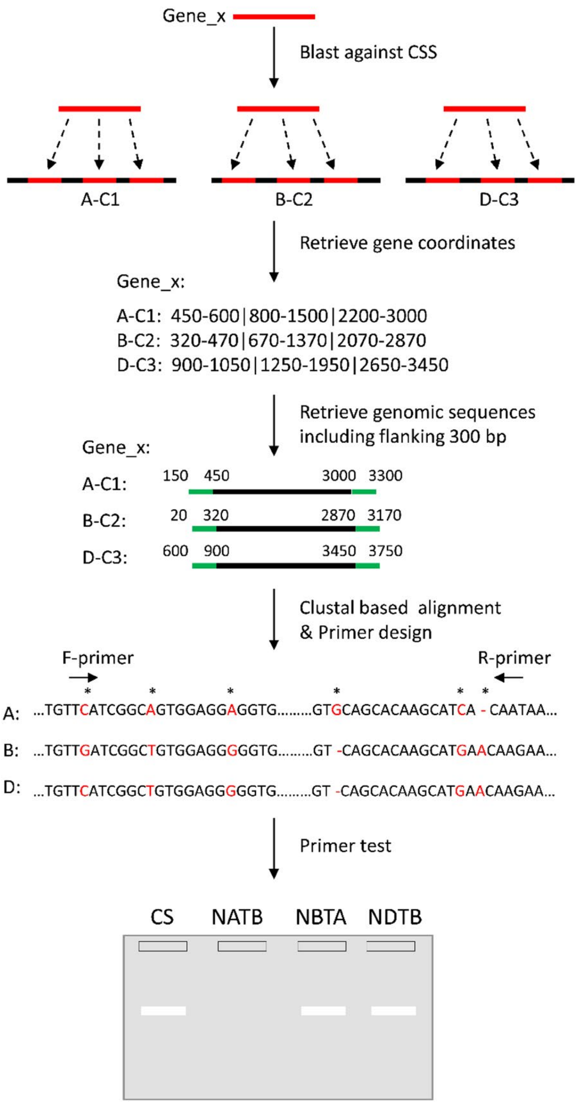

Gene_x
Blast against CSS
A-C1
B-C2
D-C3
Retrieve gene coordinates
Gene_x:
A-C1: 450-600|800-1500|2200-3000
B-C2: 320-470|670-1370|2070-2870
D-C3: 900-1050|1250-1950|2650-3450
Retrieve genomic sequences including flanking 300 bp
Gene_x:
A-C1: 150 450 3000 3300
B-C2: 20 320 2870 3170
D-C3: 600 900 3450 3750
Clustal based alignment & Primer design
F-primer
R-primer
* * * * * *
A: ...TGTTCATCGGCAGTGGAGGAGGTG.........GTGCAGCACAAGCATCA - CAATAA...
B: ...TGTTGATCGGCTGTGGAGGGGGTG.........GT -CAGCACAAGCATGAACAAGAA...
D: ...TGTTCATCGGCTGTGGAGGGGGTG.........GT -CAGCACAAGCATGAACAAGAA...
Primer test
CS
NATB
NBTA
NDTB

Fig. 5 The pipeline of generating multiple sequence alignments in bread wheat. *CSSs* chromosome shotgun sequences; *A*, *B*, and *D* the three subgenomes of bread wheat; *C1*, *C2*, and *C3* the three contigs from the three subgenomes; *'CS'* euploid wheat 'Chinese Spring'; *NATB*, *NBTA*, and *NDTB* 3 nulli-tetrasomic (NT) lines for chromosomes belonging to a given homoeologous group. *Asterisks* representing polymorphic bases targeted in designing primers specific to a subgenome. The lack of a PCR product in the lane of nulli-A and tetra-B (or NATB) indicates that primers are specific to chromosome A.

were assessed. Where possible, sequences differing in more than 1 SNP or indel were used to design primers targeting specific chromosomes for the selected genes (Additional file 1: Table S1).

The euploid and selected nullisomic–tetrasomic lines of 'CS' [19, 20] were analysed. Genomic DNA was extracted from 20-day-old seedlings using the hexadecyltrimethylammonium bromide method (CTAB) [26]. PCR amplification was performed in 10 μl reaction mixtures with 50 ng of genomic DNA, 200 μM of each dNTP, 0.2 μM of each primer, and 0.5 units of *Taq* polymerase. The cycling parameters were 94°C for 5 min to pre-denature, which was followed by 35 cycles of 94°C for 45 s, 40 s at the appropriate annealing temperature (ranging from 50 to 70°C depending on the primers, see Additional file 1: Table S1), 72°C for 1 min, and a final extension at 72°C for 10 min. Amplification products were separated on 1.5% agarose gels.

To further confirm the effectiveness of the primers in the RIL and DH populations, fragments of interest were purified using the QIAquick Extraction Kit (QIAGEN). The recovered PCR fragments were inserted into the pGEM-T easy vector (Promega) and transformed into *Escherichia coli* (Top10). At least three independent clones for each fragment were sequenced in both directions by the Australian Genome Research Facility Ltd. Sequenced fragments were aligned using by the DNAman software package (V5. 2.10; Lynnon Biosoft). To identify whether the sequenced fragments were from the expected chromosomes as found in the 'CS' aneuploids, they were aligned with all of the orthologous sequences from 'CS' for a given gene. SNPs between the parents of a given population were exploited to develop cleaved amplified polymorphic sequence (CAPS) marker using dCAPS Finder 2.0 [27]. PCR products were digested with appropriate enzymes from New England Biolabs (NEB) based on target sequences differences and separated on 3% agarose gels. The genetic linkage map was generated using JoinMap 4 [28].

Additional files

Additional file 1: Table S1. Details of primers tested.

Additional file 2: Figure S1. Validation of marker location of *AEGTA18760* using a RIL (recombination inbred lines) population. Orthologous sequences of *AEGTA18760* were amplified from the two parents of the RIL population, C ('CSCR6') and L ('Lang'), and sequenced. The single nucleotide polymorphism (in red and green) and restriction enzyme sites (underlined) were identified between C and L for *AEGTA18760* with restriction enzyme *NlaIII* (A). The amplified products of the two parents and 13 of the RIL lines were digested and separated on agarose gels. The map position of the new marker on chromosome 3B (B) was calculated based on the linkage map published by Ma et al. [22].

Additional file 3: Figure S2. Validation of marker location of *Bradi1g07500*.1 using a DH (doubled haploid) population. Orthologous sequences of *Bradi1g07500*.1 were amplified from the two parents of the DH population, B ('Batavia') and E ('Ernie'), and sequenced. The single nucleotide polymorphism (in red and green) and restriction enzyme sites (underlined) were identified between B and E for *Bradi1g07500*.1 with restriction enzyme *BtgI* (A). The amplified products of the two parents and 13 of the RIL lines were digested and separated on agarose gels. The map position of the new marker on chromosome 3B (B) was calculated based on the linkage map published by Li et al. [21].

Abbreviations

CS: Chinese spring; DH: doubled haploid; RILs: recombinant inbred lines; CDS: coding sequences; CSSs: chromosome shotgun sequences; CAPS: cleaved amplified polymorphic sequence; CTAB: hexadecyltrimethylammonium bromide; SNP: single nucleotide polymorphism; RT-qPCR: reverse transcription quantitative PCR; NCBI: National Center for Biotechnology Information.

Authors' contributions

Conceived and designed the experiments: CL JM JS WM Y-LZ. Performed the experiments: JM JS ZZ Y-XL. Analysed the data and wrote the paper: JM JS Y-LZ CL. All authors read and approved the final manuscript.

Author details

[1] Triticeae Research Institute, Sichuan Agricultural University, Wenjiang, Chengdu 611130, China. [2] CSIRO Agriculture Flagship, 306 Carmody Road, St Lucia, QLD 4067, Australia. [3] School of Plant Biology, The University of Western Australia, Perth, WA 6009, Australia. [4] National Foxtail Millet Improvement Centre, Institute of Millet Crops, Hebei Academy of Agricultural and Forestry Sciences, Shijiazhuang, China.

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 Agriculture Flagship. The authors are grateful to Drs Evans Lagudah (CSIRO) and Peng Zhang (Plant Breeding Institute, The University of Sydney) for providing the aneuploid stocks of 'Chinese Spring', and to Drs Cushla Metcalfe and Jason Carere for helping to improve the manuscript.

Compliance with ethical guidelines

Competing interests

The authors declare that they have no competing interests.

Received: 15 May 2015 Accepted: 22 July 2015
Published online: 04 August 2015

References

1. Yu J, Hu SN, Wang J, Wong GKS, Li SG, Liu B et al (2002) A draft sequence of the rice genome (*Oryza sativa* L. ssp indica). Science 296(5565):79–92. doi:10.1126/science.1068037
2. International-Barley-Genome-Sequencing-Consortium (2012) A physical, genetic and functional sequence assembly of the barley genome. Nature 491(7426):711–716
3. Zhang G, Liu X, Quan Z, Cheng S, Xu X, Pan S et al (2012) Genome sequence of foxtail millet (*Setaria italica*) provides insights into grass evolution and biofuel potential. Nat Biotechnol 30(6):549–554
4. Schnable PS, Ware D, Fulton RS, Stein JC, Wei F, Pasternak S et al (2009) The B73 maize genome: complexity, diversity, and dynamics. Science 326(5956):1112–1115
5. Paterson AH, Bowers JE, Bruggmann R, Dubchak I, Grimwood J, Gundlach H et al (2009) The *Sorghum bicolor* genome and the diversification of grasses. Nature 457(7229):551–556
6. Consortium PGS (2011) Genome sequence and analysis of the tuber crop potato. Nature 475(7355):189–195
7. Consortium TG (2012) The tomato genome sequence provides insights into fleshy fruit evolution. Nature 485(7400):635–641
8. Chalhoub B, Denoeud F, Liu S, Parkin IA, Tang H, Wang X et al (2014) Early allopolyploid evolution in the post-Neolithic *Brassica napus* oilseed genome. Science 345(6199):950–953
9. International-Wheat-Genome-Sequencing-Consortium (2014) A chromosome-based draft sequence of the hexaploid bread wheat (*Triticum aestivum*) genome. Science 345:1251788
10. Jia J, Zhao S, Kong X, Li Y, Zhao G, He W et al (2013) *Aegilops tauschii* draft genome sequence reveals a gene repertoire for wheat adaptation. Nature 496:91–95
11. Ling H-Q, Zhao S, Liu D, Wang J, Sun H, Zhang C et al (2013) Draft genome of the wheat A-genome progenitor *Triticum urartu*. Nature 496:87–90
12. Vogel JP, Garvin DF, Mockler TC, Schmutz J, Rokhsar D, Bevan MW et al (2010) Genome sequencing and analysis of the model grass *Brachypodium distachyon*. Nature 463(7282):763–768
13. Deng W, Nickle DC, Learn GH, Maust B, Mullins JI (2007) ViroBLAST: a stand-alone BLAST web server for flexible queries of multiple databases and user's datasets. Bioinformatics 23(17):2334–2336. doi:10.1093/bioinformatics/btm331
14. Ramirez-Gonzalez RH, Uauy C, Caccamo M (2015) PolyMarker: a fast polyploid primer design pipeline. Bioinformatics 31(12):2038–2039
15. Ramirez-Gonzalez RH, Segovia V, Bird N, Fenwick P, Holdgate S, Berry S et al (2014) RNA-Seq bulked segregant analysis enables the identification of high-resolution genetic markers for breeding in hexaploid wheat. Plant Biotechnol J 13(5):613–624
16. Genomics L (2013) http://www.lgcgroup.com/services/genotyping. Accessed 30 July 2015
17. Kumpatla SP, Buyyarapu R, Abdurakhmonov IY, Mammadov JA (2012) Genomics-assisted plant breeding in the 21st century: technological advances and Progress. In: Abdurakhmonov I (ed) Plant breeding. InTech publishers
18. Semagn K, Babu R, Hearne S, Olsen M (2014) Single nucleotide polymorphism genotyping using Kompetitive Allele Specific PCR (KASP): overview of the technology and its application in crop improvement. Mol Breed 33(1):1–14
19. Sears E (1954) The aneuploids of common wheat. Univ Mo Agric Exp Stn Res Bull 572:1–58
20. Sears ER, Riley R, Lewis KR (1966) Nullisomic–tetrasomic combinations in hexaploid wheat. In: Riley R, Lewis KR (eds) Chromosome manipulation and plant genetics. Oliver and Boyd, London, pp 29–45
21. Li HB, Xie GQ, Ma J, Liu GR, Wen SM, Ban T et al (2010) Genetic relationships between resistances to Fusarium head blight and crown rot in bread wheat (*Triticum aestivum* L.). Theor Appl Genet 121(5):941–950
22. Ma J, Li H, Zhang C, Yang X, Liu Y, Yan G et al (2010) Identification and validation of a major QTL conferring crown rot resistance in hexaploid wheat. Theor Appl Genet 120(6):1119–1128
23. Altschul SF, Madden TL, Schaffer AA, Zhang JH, Zhang Z, Miller W et al (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 25(17):3389–3402. doi:10.1093/nar/25.17.3389
24. Talavera G, Castresana J (2007) Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. Syst Biol 56(4):564–577
25. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H et al (2007) Clustal W and clustal X version 2.0. Bioinformatics 23(21):2947–2948. doi:10.1093/bioinformatics/btm404
26. Murray M, Thompson WF (1980) Rapid isolation of high molecular weight plant DNA. Nucleic Acids Res 8(19):4321–4326
27. Neff MM, Turk E, Kalishman M (2002) Web-based primer design for single nucleotide polymorphism analysis. Trends Genet 18(12):613–615
28. Van Ooijen J (2006) Software for the calculation of genetic linkage maps (Kyazma BV, Wageningen, The Netherlands) version 4