



Article

# RNA Sequencing-Based Bulk Segregant Analysis Facilitates Efficient D-genome Marker Development for a Specific Chromosomal Region of Synthetic Hexaploid Wheat

Ryo Nishijima<sup>1</sup>, Kentaro Yoshida<sup>1,\*</sup> , Kohei Sakaguchi<sup>1</sup>, Shin-ichi Yoshimura<sup>1</sup>, Kazuhiro Sato<sup>2</sup> and Shigeo Takumi<sup>1,\*</sup>

<sup>1</sup> Graduate School of Agricultural Science, Kobe University, Rokkodai 1-1, Nada, Kobe 657-8501, Japan; ryosnishijima@gmail.com (R.N.); k.sakaguchi@brown.plala.or.jp (K.S.); shin.82353@gmail.com (S.-i.Y.)

<sup>2</sup> Institute of Plant Science and Resources, Okayama University, Kurashiki, Okayama 710-0046, Japan; kzsato@rib.okayama-u.ac.jp

\* Correspondence: kentaro.yoshida@port.kobe-u.ac.jp (K.Y.); takumi@kobe-u.ac.jp (S.T.); Tel.: +81-78-803-5858 (K.Y.); +81-78-803-5860 (S.T.)

Received: 2 November 2018; Accepted: 22 November 2018; Published: 26 November 2018



**Abstract:** Common wheat originated from interspecific hybridization between cultivated tetraploid wheat and its wild diploid relative *Aegilops tauschii* followed by amphidiploidization. This evolutionary process can be reproduced artificially, resulting in synthetic hexaploid wheat lines. Here we performed RNA sequencing (RNA-seq)-based bulk segregant analysis (BSA) using a bi-parental mapping population of two synthetic hexaploid wheat lines that shared identical A and B genomes but included with D-genomes of distinct origins. This analysis permitted identification of D-genome-specific polymorphisms around the *Net2* gene, a causative locus to hybrid necrosis. The resulting single nucleotide polymorphisms (SNPs) were classified into homoeologous polymorphisms and D-genome allelic variations, based on the RNA-seq results of a parental tetraploid and two *Ae. tauschii* accessions. The difference in allele frequency at the D-genome-specific SNP sites between the contrasting bulks ( $\Delta$ SNP-index) was higher on the target chromosome than on the other chromosomes. Several SNPs with the highest  $\Delta$ SNP-indices were converted into molecular markers and assigned to the *Net2* chromosomal region. These results indicated that RNA-seq-based BSA can be applied efficiently to a synthetic hexaploid wheat population to permit molecular marker development in a specific chromosomal region of the D genome.

**Keywords:** allohexaploid; homoeolog; hybrid necrosis; molecular marker; wheat

## 1. Introduction

Wild relatives of common wheat (*Triticum aestivum* L.), including *Aegilops* species, constitute important genetic resources for common wheat breeding. Notably, *Aegilops tauschii* Coss. is the D-genome progenitor of common wheat, and a limited number of *Ae. tauschii* strains appear to have contributed to the speciation of common wheat through interspecific crossing to the cultivated tetraploid wheat (*Triticum turgidum* L. ssp. *durum*) strains carrying the A and B genomes and subsequent amphidiploidization about 8000 years ago in the Fertile Crescent [1,2]. Artificial replication of this evolutionary process can be achieved by generation of synthetic hexaploid wheat derived from the interspecific crosses between cultivated tetraploid wheat and *Ae. tauschii* [3]. Thus, agriculturally useful alleles of *Ae. tauschii* are directly available for common wheat breeding via the transmission from synthetic hexaploid wheat to common wheat [4–10]. In common wheat, the D genome is known

to harbor much lower genetic diversity than do the A and B genomes, whereas introgression of the *Ae. tauschii* natural variation via synthetic hexaploid wheat has been used successfully to enlarge D-genome diversity [11–14].

*Ae. tauschii* is widely distributed, and is present from northern Syria and southeastern Turkey to western China [15,16]. Recent studies on the population structure of *Ae. tauschii* showed that only one (TauL1) of three major lineages has contributed to the wide species range [2,17,18]. The other two major lineages, TauL2 and TauL3, are restricted to the Transcaucasus/Middle East region. TauL2 includes both subspecies *tauschii* and subspecies *strangulata*, and the TauL3 accessions are limited in Georgia [18]. The *Ae. tauschii* strains associated with the origin of common wheat are assumed to be the TauL2 lineage [2,17], and only limited reproductive barriers are thought to exist between tetraploid wheat and many of the TauL2 accessions [19]. TauL1 accessions frequently carry a hybrid incompatibility gene, designated *Net2*, that triggers low-temperature-induced necrotic cell death upon interspecific hybridization with tetraploid wheat, impeding the generation of synthetic hexaploid wheat [20,21]. Thus, *Net2* is a major reproductive barrier for breeding use of the TauL1 accessions, and the development of markers closely linked to *Net2* is needed for efficient use of the TauL1 gene pool. A fine map for the *Net2* chromosomal region already has been constructed on the short arm of chromosome 2D [22].

Genomic approaches using next-generation sequencing (NGS) techniques have been applied to analysis of the genomes of the wild relatives of domesticated crops, expanding the genetic resources available for crop improvement [23]. In common wheat, the D-genome markers remain much less developed than those of the A and B genomes, whereas recent progress using the NGS technique has facilitated an increase in the number of D-genome markers [24]. RNA sequencing (RNA-seq) of the *Ae. tauschii* accessions has generated a huge number of genome-wide polymorphisms, including single nucleotide polymorphisms (SNPs) and insertions/deletions (indels); the genome-wide SNPs and indels can be efficiently anchored to the chromosomes of barley and *Ae. tauschii* [25–27]. The SNP- and indel-based markers are available for construction of linkage maps in the target chromosomal regions of not only *Ae. tauschii* but also the D genome of hexaploid wheat including synthetic wheat [26,27].

Bulked segregant analysis (BSA) combined with NGS allows efficient development of molecular markers linked to a genomic region associated with the target phenotype in cereals [28–31]. In maize, for example, an RNA-seq-based BSA approach has been used to construct high-density linkage maps and to screen among candidate genes for a target locus [32–34]. Hexaploid wheat has three closely related genomes (designated A, B, and D), each of which carries a set of highly similar genes (homoeology). Due to the genome complexity via allopolyploidy and the large proportion of repetitive DNA in polyploid wheat, whole-genome resequencing is still unviable and reduced-representation methods of NGS data have been employed in this species [35]. Recently, RNA-seq-based BSA was employed successfully for the development of molecular markers closely linked to target chromosomal genes such as a grain protein content gene (*GPC-B1*), a yellow rust resistance gene (*Yr15*), and a powdery mildew resistance gene (*Pm4b*) in tetraploid and common wheat [36–38]. However, a limited amount of information has been reported for the RNA-seq-based BSA approach in polyploid wheat. Here, we employed the RNA-seq-based BSA method to develop a molecular marker closely linked to *Net2*. This process used a mapping population generated from a cross of two synthetic hexaploid wheat lines that shared identical A and B genomes but contained diverse D genomes originating from two distinct pollen parents.

## 2. Results and Discussion

Two synthetic hexaploid wheat lines were derived from interspecies crosses of tetraploid wheat cultivar Langdon (Ldn) and two *Ae. tauschii* accessions (KU-2075 and KU-2025). Ldn/KU-2075 and Ldn/KU-2025, respectively, showed normal phenotype (wild-type) and type-II necrosis phenotype [20]. To obtain novel molecular markers tightly linked to *Net2*, RNA sequencing of four bulks of synthetic hexaploid wheat was performed (Figure 1). Each bulk was composed of

ten *Net2*-homozygous individuals or non-carriers (*net2*-homozygous) selected from two F<sub>5</sub> populations between Ldn/KU-2075 and Ldn/KU-2025 (hereafter referred to as Segregating Populations SP1 and SP2). The two bulks of each SP (*Net2*-SP1, non-carrier-SP1, *Net2*-SP2, and non-carrier-SP2) were sequenced twice; the bulks were designated as follows: *Net2*-SP1-1st, *Net2*-SP1-2nd, non-carrier-SP1-1st, non-carrier-SP1-2nd, *Net2*-SP2-1st, *Net2*-SP2-2nd, non-carrier-SP2-1st, and non-carrier-SP2-2nd (Table 1). In each experiment, 4.06 million to 5.22 million paired-end reads were generated. After quality filtering, 2.80 million to 3.71 million high-quality reads were acquired. These reads were aligned with transcripts of the two parental synthetic hexaploid *Ae. tauschii* accessions, KU-2075 and KU-2025, transcriptomes that had been de novo assembled in our previous study [23]. In each experiment, 2.02 million to 2.76 million and 1.93 million to 2.68 million reads were aligned to the KU-2075 and KU-2025 transcriptomes, respectively. The alignment output files of the first and second runs were merged for each alignment pair of bulk and transcript. On average, 292,678 and 278,690 SNPs were detected on KU-2075 and KU-2025 transcripts, respectively (Table 2). Of these, 290,712 (99.33%) and 276,700 (99.29%) SNPs on the KU-2075 and KU-2025 transcripts (respectively) were anchored to the *Ae. tauschii* genome. SNP sites in the eight alignment pairs of bulk and transcript (e.g., *Net2*-SP1 vs. KU-2075), as shown in Table 3, were integrated with the genome sequence, yielding a total of 319,808 non-redundant (NR) SNP sites on the seven chromosomes (Table 3). This section may be divided by subheadings. It should provide a concise and precise description of the experimental results, their interpretation as well as the experimental conclusions that can be drawn.

To distinguish D-genome-specific variations from homoeologous polymorphisms between the tetraploid AB and diploid D genomes, the RNA-seq raw read data of Ldn, the tetraploid parental accession of the synthetic wheat lines, were processed as described above. Out of the total of 6.32 million read pairs obtained, 4.37 million read pairs passed the quality filtering (Table 1). To KU-2075 and KU-2025 transcripts, 2.97 million and 2.66 million reads were aligned, respectively.

**Table 1.** Summary of RNA sequencing results for four pairs of bulks of synthetic hexaploid wheat and for tetraploid wheat cv. Langdon.

| Samples              | Total Read Pairs | Filtered Read Pairs (%) <sup>a</sup> | Aligned to the <i>Ae. tauschii</i> Transcripts <sup>b</sup> |                      |
|----------------------|------------------|--------------------------------------|---|----------------------|
|                      |                  |                                      | (%) <sup>c</sup>  |                      |
|                      |                  |                                      | KU-2075   | KU-2025              |
| Synthetic hexaploids |                  |                                      |   |                      |
| non-carrier-SP1-1st  | 4,202,114        | 2,799,202 (66.61%)                   | 2,020,405 (72.18%)  | 1,930,400.5 (68.96%) |
| non-carrier-SP1-2nd  | 4,059,840        | 2,858,956 (70.42%)                   | 2,062,291 (72.13%)  | 1,971,110.5 (68.95%) |
| non-carrier-SP2-1st  | 4,492,358        | 2,953,088 (65.74%)                   | 2,164,656 (73.3%)   | 2,083,120 (70.54%)   |
| non-carrier-SP2-2nd  | 4,115,352        | 2,864,271 (69.60%)                   | 2,098,383 (73.26%)  | 2,020,752.5 (70.55%) |
| <i>Net2</i> -SP1-1st | 4,710,499        | 3,148,652 (66.84%)                   | 2,208,392 (70.14%)  | 2,110,452.5 (67.03%) |
| <i>Net2</i> -SP1-2nd | 4,403,630        | 3,108,568 (70.59%)                   | 2,178,403 (70.08%)  | 2,082,004 (66.98%)   |
| <i>Net2</i> -SP2-1st | 4,828,182        | 3,249,596 (67.30%)                   | 2,420,056 (74.47%)  | 2,348,814 (72.28%)   |
| <i>Net2</i> -SP2-2nd | 5,216,082        | 3,709,478 (71.12%)                   | 2,763,471 (74.5%)   | 2,684,019 (72.36%)   |
| Tetraploid wheat     |                  |                                      |   |                      |
| cv. Langdon          | 6,316,174        | 4,372,660 (69.23%)                   | 2,974,277 (68.02%)  | 2,661,487 (60.87%)   |

<sup>a</sup> The ratio of the filtered read pairs to the total read pairs. <sup>b</sup> Nishijima et al. [27]. <sup>c</sup> The ratio of the aligned reads to the filtered read pairs.

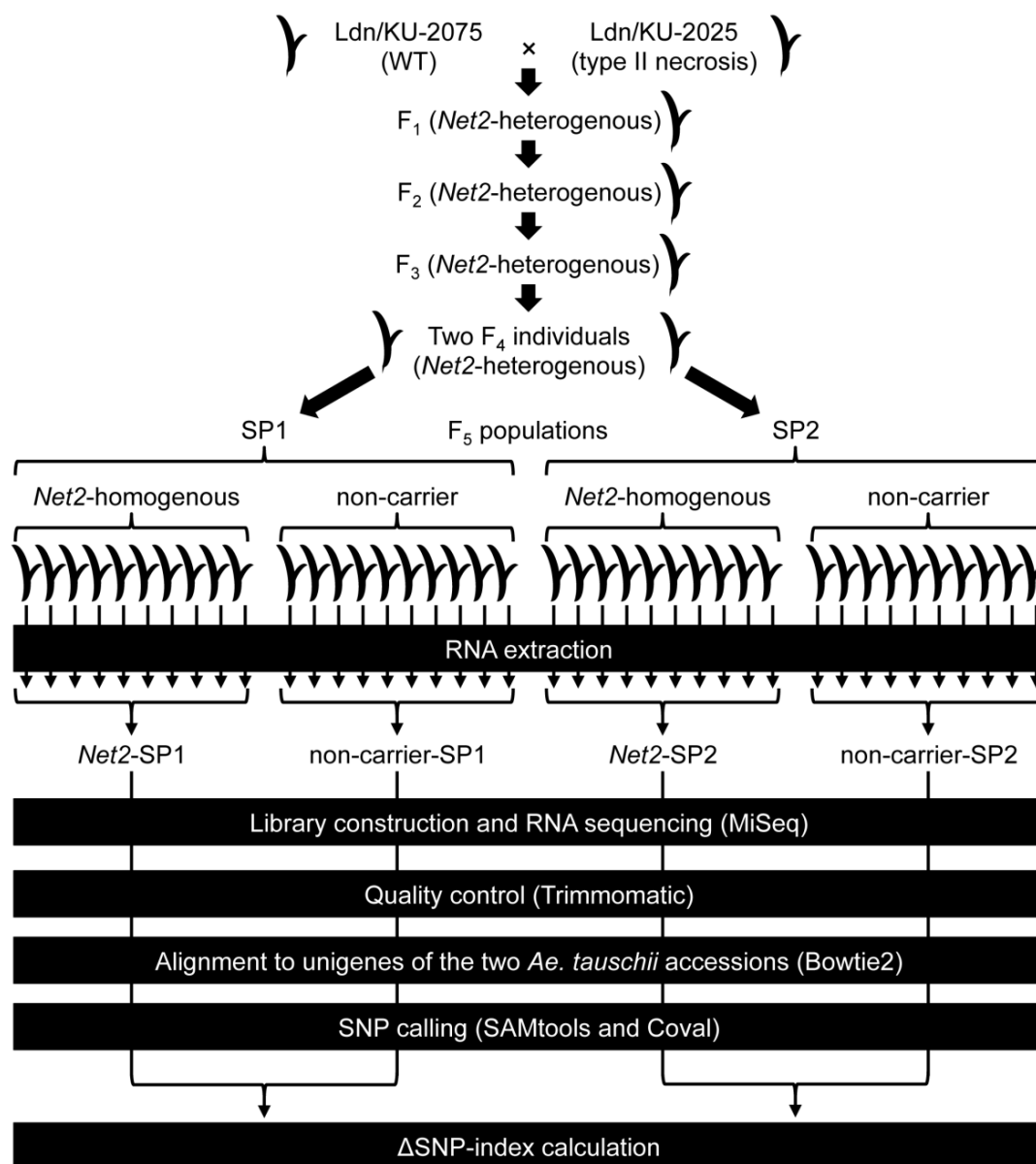


Figure 1. Workflow of the RNA sequencing analysis in this study.

Table 2. The number of single nucleotide polymorphisms (SNPs) detected in four bulks of synthetic hexaploid wheat and tetraploid wheat cv. Langdon compared to the two parental *Ae. tauschii* transcriptomes.

| Transcripts <sup>a</sup> |         | KU-2075                                 |         | KU-2025                                 |  |
|--------------------------|---------|---|---------|---|--|
| The Number of SNP        | Total   | Anchored to the Genome <sup>b</sup> (%) | Total   | Anchored to the Genome <sup>b</sup> (%) |  |
| Synthetic hexaploids     |         |   |         |   |  |
| non-carrier-SP1          | 277,605 | 275,799 (99.35%)                        | 262,966 | 261,128 (99.30%)                        |  |
| non-carrier-SP2          | 276,564 | 274,772 (99.35%)                        | 269,175 | 267,249 (99.28%)                        |  |
| <i>Net2</i> -SP1         | 318,046 | 315,859 (99.31%)                        | 296,819 | 294,739 (99.30%)                        |  |
| <i>Net2</i> -SP2         | 298,496 | 296,419 (99.30%)                        | 285,798 | 283,684 (99.26%)                        |  |
| Tetraploid wheat         |         |   |         |   |  |
| cv. Langdon              | 429,346 | 421,957 (98.28%)                        | 350,871 | 345,657 (98.51%)                        |  |

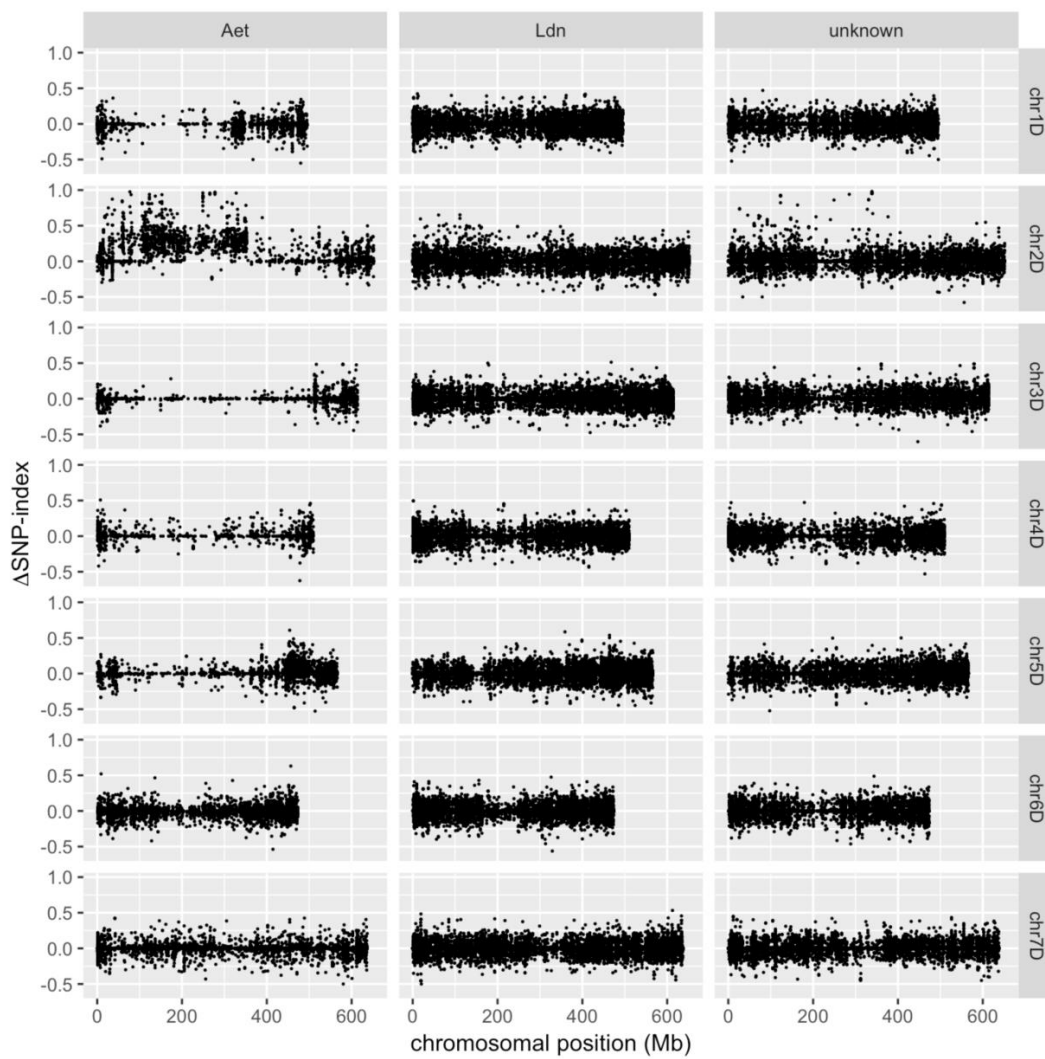
<sup>a</sup> Nishijima et al. [27]. <sup>b</sup> Luo et al. [39].

**Table 3.** The number of SNPs classified into three categories, including D-genome-specific allelic variations, homoeologous polymorphisms, and unclassified (those falling into neither of the other two classes).

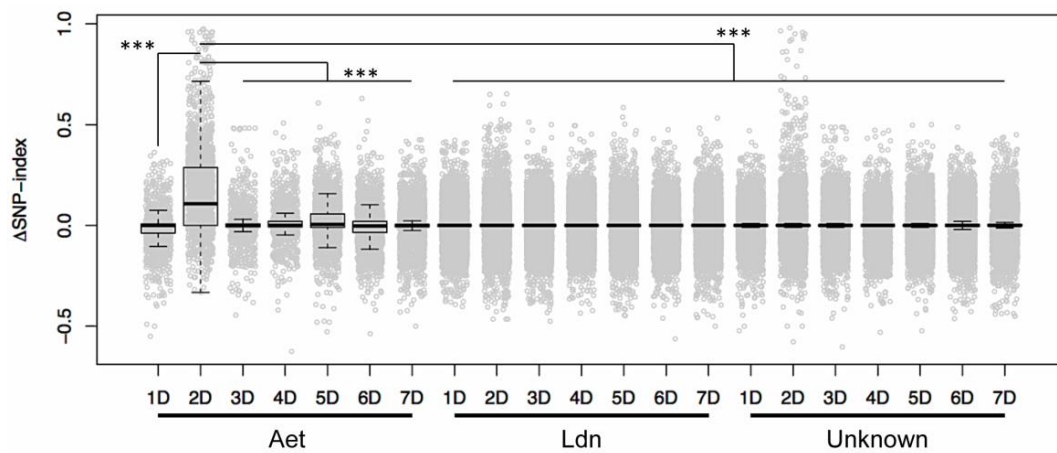
| Chr.  | D-genome-Specific | Homoeologous | Unclassified | Total   |
|-------|-------------------|--------------|--------------|---------|
| 1D    | 1674              | 29,307       | 11,760       | 42,741  |
| 2D    | 2295              | 33,822       | 13,975       | 50,092  |
| 3D    | 1611              | 31,932       | 15,532       | 49,075  |
| 4D    | 1698              | 28,781       | 11,776       | 42,255  |
| 5D    | 2936              | 34,961       | 14,336       | 52,233  |
| 6D    | 3730              | 24,534       | 10,966       | 39,230  |
| 7D    | 3983              | 28,593       | 11,606       | 44,182  |
| Total | 17,927            | 211,930      | 89,951       | 319,808 |

Out of the 429,346 and 350,871 SNPs identified on KU-2075 and KU-2025 transcripts (respectively), 421,957 (98.28%) and 345,657 SNPs (98.51%) (respectively) were anchored to the *Ae. tauschii* genome (Table 2). Based on these SNPs derived from tetraploid wheat reads, and on the SNPs between the two *Ae. tauschii* accessions reported in our previous study [27], NR SNPs identified in synthetic hexaploid bulks were classified into homoeologous polymorphisms and D-genome-specific allelic variation. Out of the 319,808 SNPs, 17,927 were assigned to the allelic variation on the D genome, and 211,930 were grouped into the homoeologous polymorphisms, while the remained 89,951 did not fall into either of these two classes (“unclassified”; Table 3).

SNP-index values were computed for each of the eight alignment pairs, and then  $\Delta$ SNP-index values were calculated for the comparisons between *Net2*-homogenous bulks and non-carrier bulks, yielding indices for *Net2*-SP1 minus non-carrier-SP1 and *Net2*-SP2 minus non-carrier-SP2. The average of the four  $\Delta$ SNP-index values at each SNP site was mapped to the *Ae. tauschii* genome (Figure 2). The  $\Delta$ SNP-index distribution on the chromosomes indicated that SNPs with high  $\Delta$ SNP-index values tended to be located on the short arm of chromosome 2D, where the *Net2* locus resides (Figure 2). Comparison of  $\Delta$ SNP-index values among D-genome-specific SNPs showed that the D-genome-specific SNPs on chromosome 2D possessed significantly higher (Steel-Dwass test,  $p < 0.001$ )  $\Delta$ SNP-index values than those on the other six chromosomes (Figure 3). Theoretically, homozygous chromosomal regions account for 87.5% of the total genome in an  $F_4$  individual derived from two parents. D-genome-specific SNP sites with low  $\Delta$ SNP-index values appeared to already have been fixed as homozygous for the KU-2075 or KU-2025 allele by the  $F_4$  generation of the synthetic hexaploid populations. The  $\Delta$ SNP-index values of D-genome-specific SNPs on chromosome 2D also were significantly higher (Steel-Dwass test,  $p < 0.001$ ) than those of homoeologous and unclassified SNPs (Figure 3). Moreover, most of the homoeologous and unclassified polymorphisms showed  $\Delta$ SNP-index values of approximately zero (Figure 3), indicating that homoeologous polymorphisms had been efficiently removed from candidate SNPs for development of molecular markers, since such candidate SNPs should have possessed high  $\Delta$ SNP-index values. Low  $\Delta$ SNP-index values at homoeologous polymorphic sites were likely due to the similar allele frequency of A- and B-genome-derived reads between the *Net2*-homogenous and non-carrier bulks, which thereby offset the SNP-index values of each other.



**Figure 2.** Distribution of  $\Delta$ SNP-index values along the *Ae. tauschii* chromosomes. The three categories of SNPs (D-genome-specific allelic variations, homoeologous polymorphisms, and unclassified SNPs) are designated as “Aet”, “Ldn”, and “Unknown”, respectively.



**Figure 3.** Box and dot plots for  $\Delta$ SNP-index values based on the seven chromosomes and on the SNP classifications in Figure 2. D-genome-specific SNPs on chromosome 2D had significantly higher  $\Delta$ SNP-index values than SNPs on the other six chromosomes, homoeologous SNPs, and unclassified SNPs. \*\*\*  $p < 0.001$  by the Steel-Dwass test.

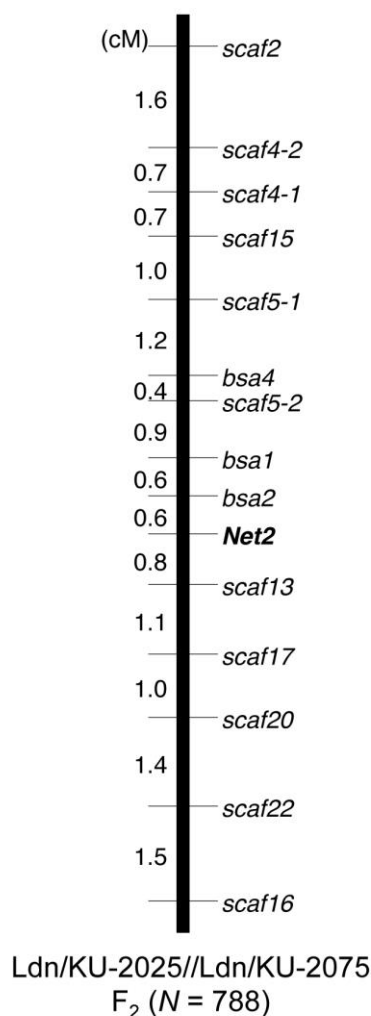
To assess whether SNPs with high  $\Delta$ SNP-index values can be used as molecular markers for the *Net2* locus, derived cleaved amplified polymorphic sequence (dCAPS) markers were designed from these SNPs. Within the physical region from 81.8 Mb to 83.3 Mb on chromosome 2D that was defined by the two markers noted in our previous study [22], 208 SNPs were detected in the present study, including 10 D-genome specific, 144 homoeologous, and 54 unclassified SNPs. The  $\Delta$ SNP-index values of this subset of SNPs ranged from  $-0.4025$  to  $0.5425$ . Four of the D-genome-specific SNPs with  $\Delta$ SNP-index values higher than  $0.38$  were converted into dCAPS markers (Table 4), and three of these markers (designated *bsa1*, *bsa2*, and *bsa4*) were successfully genotyped and mapped in the Ldn/KU-2025//Ldn/KU-2075 population (Figure 4). All three of these markers were anchored within 2.5 cM distal to the *Net2* locus, and two of these markers were closer to *Net2* than to *scaf52*, the most closely linked markers in our previous study [22]. This result indicates that SNPs derived from RNA-seq-based BSA of synthetic hexaploid wheat can be used for molecular marker development in a specific chromosomal region of the D genome.

**Table 4.** List of the derived cleaved amplified polymorphic sequence (dCAPS) markers developed in this study.

| Marker Name | Primer Sequence (5' to 3')                       | Restriction Enzyme |
|-------------|--|--------------------|
| <i>bsa1</i> | TCATGACCTGCTGGTTTGTT<br>GATCCAATGTTATTCTGAACCCT  | <i>StyI</i>        |
| <i>bsa2</i> | TCACAACATTCGCAGGTCAT<br>TGGTCTGTGATCTCACTGCC     | <i>HpaII</i>       |
| <i>bsa4</i> | ACAAGTCGGATATCGCCAAA<br>CAGCTAAAAACTGTTTGCTTGAGA | <i>HinfI</i>       |

The distribution of  $\Delta$ SNP-index values was not continuous (Figure 2), even in a single transcript. The discontinuous pattern was due to the homoeologous and nonpolymorphic RNA-seq reads, which could be derived from homoeologous regions in the A and B genomes and overlap with the reads containing the D-genome-specific SNPs. For complete masking of the overlapping reads from the A and B genomes, precise alignment to all three genomes of hexaploid wheat would be needed. In the present study, although transcripts of the parental diploid species were used as the reference sequences, our strategy succeeded in efficiently detecting candidate SNPs on a specific chromosomal region of synthetic hexaploid wheat. Therefore, this strategy of RNA-seq-based BSA is expected to be applicable to other synthetic polyploids derived from crosses among tetraploid and/or diploid wheat relatives.

Taken together, these data demonstrate that RNA-seq-based BSA can be applied to synthetic hexaploid wheat for the development of molecular markers in specific chromosomal regions. The use of a single tetraploid wheat cultivar as a parental line for synthetic hexaploid wheat effectively cancelled the increase in  $\Delta$ SNP-index values at homoeologous polymorphic sites. To date, a large number of synthetic wheat hexaploids have been produced as part of efforts to enlarge the D-genome genetic diversity [5,6,40]. Epistatic interactions can occur between the parental genomes in the newly synthesized allopolyploids [41], and several phenotypes specific to the synthetic wheat hexaploids have been reported [20,42]. RNA-seq and subsequent de novo transcriptome assembly can be performed even in wild diploid wheat relatives with no reference sequences. This strategy would facilitate molecular marker development in diverse wheat synthetics with various genome constitutions. Thus, RNA-seq-based BSA is expected to serve as a rapid and efficient approach for genetic evaluation of target traits from wild wheat relatives with allopolyploid backgrounds.



**Figure 4.** Genetic map of the short arm of chromosome 2D in the Ldn/KU-2025//Ldn/KU-2075 population. The three newly developed markers are designated with “*bsa*” prefixes.

### 3. Materials and Methods

#### 3.1. Plant Materials

Two synthetic hexaploid wheat lines derived from interspecies crosses of tetraploid wheat cultivar Langdon (Ldn) and two *Ae. tauschii* accessions (KU-2075 and KU-2025) were used in this study. These lines were produced in our previous study [21]. A synthetic hexaploid line (Ldn/KU-2075) showed normal growth, whereas another synthetic hexaploid line (Ldn/KU-2025) exhibited the type-II necrosis phenotype when grown at low temperature [21]. An F<sub>2</sub> mapping population was generated through a cross of the two synthetic hexaploid lines [22]. Subsequently, the F<sub>3</sub> and F<sub>4</sub> generations were developed from selfing F<sub>2</sub> and F<sub>3</sub> individuals (respectively) heterozygous for the chromosomal region of *Net2*, a type-II necrosis-causing gene [22]. Two F<sub>5</sub> populations (SP1 and SP2) were generated through selfing of two independent F<sub>4</sub> individuals heterozygous for the *Net2* region (Figure 1).

#### 3.2. Library Construction and RNA Sequencing

Ten *Net2*-homozygous individuals and ten non-carriers (*net2*-homozygous) each were selected from the SP1 and SP2 populations. Total RNA was extracted from leaves of each of the forty selected plants at the seedling stage using a Plant Total RNA Extraction Miniprep System (Viogene, Taipei Hsien, Taiwan, ROC). The resulting RNA preparations were pooled independently into four bulk preparations according to the genotype of the *Net2* region and to the SP source population (*Net2*-SP1, non-carrier-SP1,



*Net2*-SP2, and non-carrier-SP2). A total of 8 µg of each bulk preparation of RNA was used to construct paired-end libraries using TruSeq RNA Library Preparation Kit v2 (Illumina, San Diego, CA, USA). The libraries were sequenced twice for 300 cycles × 2 on an Illumina MiSeq sequencer with 300-bp paired-end reads. The read data was deposited to the DDBJ Sequence Read Archive under the accession number DRA007501. RNA sequencing data of Ldn, which was obtained in our previous report [43] and deposited under the accession number DRA007097, also was used for the subsequent analysis.

### 3.3. Alignment of RNA-seq Reads to de Novo Assembled Transcripts of the Parental *Ae. tauschii* Accessions

Low-quality bases (average Phred quality score per 4 bp < 30), adapter sequences, and reads < 100 bp were removed from the short reads using the Trimmomatic version 0.32 tool [44]. RNA-seq analysis of KU-2075 and KU-2025 for de novo transcriptome assembly was described in our previous study [27], and the corresponding sequencing data was obtained from the DDBJ Sequence Read Archive DRA004604. The quality-filtered reads of each bulk and Ldn were aligned to the transcripts of KU-2075 and KU-2025 using the Bowtie2 tool [45] with local alignment, obtaining bam outputs.

### 3.4. Identification of D-genome Specific SNPs and Calculation of ΔSNP-Index

After merging the two bam files of each bulk library using SAMtools software (version 1.9; La Jolla, CA, USA) with the command “samtools merge”, SNPs were called from the alignment files using SAMtools and Coval software (version 1.5; Iwate, Japan) with option “-freq 0.1 -m 1000000 -n 10” [46,47]. The bam files of the synthetic hexaploid bulks were supposed to hold homoeologous polymorphisms and allelic variations on the D genome, whereas the alignment outputs of Ldn would contain homoeologous polymorphisms only. We assumed that, of the SNPs derived from the bulks, SNPs also detected in the pairwise comparison of KU-2075 and KU-2025 would represent D-genome specific polymorphisms, and SNPs also found in the alignment files of Ldn would represent homoeologous polymorphisms. The transcript sequences of the two *Ae. tauschii* accessions were mapped to the reference genome of *Ae. tauschii* [39] using GMAP version 2013-03-31 software (South San Francisco, CA, USA) [48,49], and the identified SNPs were anchored to the *Ae. tauschii* genome based on the GMAP outputs. The allele frequency at the SNP sites, a value designated as the SNP-index [28,29], was calculated in each of the eight respective alignment pairs, and the difference in the SNP-index values of the contrasting bulks (*Net2*-homogenous minus non-carrier bulk alignment, e.g., *Net2*-SP1-KU-2075 minus non-carrier-SP1-KU-2075) was defined as the ΔSNP-index by analogy to the work of Takagi et al. [29]. RStudio ver. 0.99.902 [50] with R software ver. 3.3.1 [51] was used for statistical analyses and ΔSNP-index plotting of the *Ae. tauschii* chromosomes.

### 3.5. Molecular Marker Development and Genotyping

In our previous study [22], the *Net2* locus was fine-mapped within a 0.6-cM region on the short arm of chromosome 2D. The two most closely linked markers were assigned to the *Ae. tauschii* genome [39] using the BLASTN search function of the BLAST+ software (Bethesda, MD, USA) [52] to define the physical region where the *Net2* locus was assumed to reside. D-genome-specific SNPs of high ΔSNP-index values within this chromosomal region were converted into derived cleaved amplified polymorphic sequence (dCAPS) markers (Table 4), and these novel markers were used to genotype 788 individuals of the Ldn/KU-2025//Ldn/KU-2075 population, as described in our previous report [22]. The genotyped markers were assigned to the genetic map around *Net2* using the MAPMAKER/EXP version 3.0 package (Cambridge, MA, USA) [53].

**Author Contributions:** Conceptualization, S.T. and K.Y.; methodology, R.N. and K.Y.; formal analysis, R.N., K.S., S.-i.Y. and K.S.; resources, S.T.; writing—original draft preparation, R.N. and S.T.; writing—review and editing, K.Y., K.S. and S.T.; project administration, S.T.; funding acquisition, S.T., R.N. and K.S.

**Funding:** This work was supported by grants from the Ministry of Education, Culture, Sports, Science and Technology of Japan (MEXT; Grant-in-Aid for Scientific Research (B) No. 16H04862 and Grant-in-Aid for Scientific Research on Innovative Areas No. 17H05842) to ST; by a Research Fellowship from the Japan Society for the Promotion of Science for Young Scientists (No. 16J03477) to RN; and by MEXT as part of a Joint Research Program implemented at the Institute of Plant Science and Resources, Okayama University, Japan. KY was supported by the Japan Science and Technology Agency, PRESTO (No. JPMJPR15QB).

**Acknowledgments:** The *Ae. tauschii* seeds used in this study were supplied by the National BioResource Project-Wheat (<http://shigen.nig.ac.jp/wheat/komugi/>).

**Conflicts of Interest:** The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

## Abbreviations

|         |  |
|---------|--|
| RNA-seq | RNA sequencing                                 |
| NGS     | next-generation sequencing                     |
| BSA     | bulked segregant analysis                      |
| SNP     | single nucleotide polymorphism                 |
| Indel   | insertion/deletion                             |
| NR      | non-redundant                                  |
| SP      | Segregating Population                         |
| dCAPS   | derived cleaved amplified polymorphic sequence |
| Ldn     | Langdon  |

## References

1. Matsuoka, Y. Evolution of polyploid *Triticum* wheats under cultivation: The role of domestication, natural hybridization and allopolyploid speciation in their diversification. *Plant Cell Physiol.* **2011**, *52*, 750–764. [[CrossRef](#)] [[PubMed](#)]
2. Wang, J.; Luo, M.C.; Chen, Z.; You, F.M.; Wei, Y.; Zheng, Y.; Dvorak, J. *Aegilops tauschii* single nucleotide polymorphisms shed light on the origins of wheat D-genome genetic diversity and pinpoint the geographic origin of hexaploid wheat. *New Phytol.* **2013**, *198*, 925–937. [[CrossRef](#)] [[PubMed](#)]
3. Matsuoka, Y.; Nasuda, S. Durum wheat as a candidate for the unknown female progenitor of bread wheat: An empirical study with a highly fertile F<sub>1</sub> hybrid with *Aegilops tauschii* Coss. *Theor. Appl. Genet.* **2004**, *109*, 1710–1717. [[CrossRef](#)] [[PubMed](#)]
4. Zohary, D.; Harlan, J.R.; Vardi, A. The wild diploid progenitors of wheat and their breeding value. *Euphytica* **1969**, *18*, 58–65. [[CrossRef](#)]
5. Mujeeb-Kazi, A.; Rosas, V.; Roldan, S. Conservation of the genetic variation of *Triticum tauschii* (Coss.) Schmalh. (*Aegilops squarrosa* auct. non L.) in synthetic hexaploid wheats (*T. turgidum* L. s.lat. x *T. tauschii*; 2n=6x=42, AABBDD) and its potential utilization for wheat improvement. *Genet. Resour. Crop Evol.* **1996**, *43*, 129–134. [[CrossRef](#)]
6. Jones, H.; Gosman, N.; Horsnell, R.; Rose, G.A.; Everst, L.A.; Bentley, A.R.; Tha, S.; Uauy, C.; Kowalski, A.; Novoselovic, D.; et al. Strategy for exploiting exotic germplasm using genetic, morphological, and environmental diversity: The *Aegilops tauschii* Coss. example. *Theor. Appl. Genet.* **2013**, *126*, 1793–1808. [[CrossRef](#)] [[PubMed](#)]
7. Bhatta, M.; Morgounov, A.; Belamkar, V.; Yorgancilar, A.; Baenziger, P.S. Genome-wide association study reveals favorable alleles associated with common bunt resistance in synthetic hexaploid wheat. *Euphytica* **2018**, *214*, 200. [[CrossRef](#)]
8. Bhatta, M.; Morgounov, A.; Belamkar, V.; Baenziger, P.S. Genome-wide association study reveals novel genomic regions for grain yield and yield-related traits in drought-stressed synthetic hexaploid wheat. *Int. J. Mol. Sci.* **2018**, *19*, 3011. [[CrossRef](#)] [[PubMed](#)]
9. Bhatta, M.; Baenziger, P.S.; Waters, B.M.; Poudel, R.; Belamkar, V.; Poland, J.; Morgounov, A. Genome-wide association study reveals novel genomic regions associated with 10 grain minerals in synthetic hexaploid wheat. *Int. J. Mol. Sci.* **2018**, *19*, 3237. [[CrossRef](#)] [[PubMed](#)]

10. Gorafi, Y.S.A.; Kim, J.S.; Elbashir, A.A.E.; Tsujimoto, H. A population of wheat multiple synthetic derivatives: An effective platform to explore, harness and utilize genetic diversity of *Aegilops tauschii* for wheat improvement. *Theor. Appl. Genet.* **2018**, *131*, 1615–1626. [[CrossRef](#)] [[PubMed](#)]
11. Zhang, P.; Dreisigacker, S.; Melchinger, A.E.; Reif, J.C.; Mujeeb-Kazi, A.; Van Ginkel, M.; Hoisington, D.; Warburton, M.L. Quantifying novel sequence variation and selective advantage in synthetic hexaploid wheats and their backcross-derived lines using SSR markers. *Mol. Breed.* **2005**, *15*, 1–10. [[CrossRef](#)]
12. Jafarzadeh, J.; Bonnett, D.; Jannink, J.L.; Akdemir, D.; Dreisigacker, S.; Sorrells, M.E. Breeding value of primary synthetic wheat genotypes for grain yield. *PLoS ONE* **2016**, *11*, e0162860. [[CrossRef](#)] [[PubMed](#)]
13. Bhatta, M.; Morgounov, A.; Belamkar, V.; Poland, J.; Baenziger, P.S. Unlocking the novel genetic diversity and population structure of synthetic hexaploid wheat. *BMC Genom.* **2018**, *19*, 591. [[CrossRef](#)] [[PubMed](#)]
14. Rasheed, A.; Mujeeb-Kazi, A.; Ogbonnaya, F.C.; He, Z.; Rajaram, S. Wheat genetic resources in the post-genomics era: Promise and challenges. *Ann. Bot.* **2018**, *121*, 603–616. [[CrossRef](#)] [[PubMed](#)]
15. Van Slageren, M.W. *Wild Wheats: A Monograph of Aegilops L. and Amblyopyrum (Jaub. & Spach) Eig (Poaceae)*; Wageningen Agricultural University: Wageningen, The Netherlands, 1994; pp. 326–344, ISBN 90-6754-377-2.
16. Matsuoka, Y.; Takumi, S.; Kawahara, T. Flowering time diversification and dispersal in central Eurasian wild wheat *Aegilops tauschii* Coss.: Genealogical and ecological framework. *PLoS ONE* **2008**, *3*, e3138. [[CrossRef](#)] [[PubMed](#)]
17. Mizuno, N.; Yamasaki, M.; Matsuoka, Y.; Kawahara, T.; Takumi, S. Population structure of wild wheat D-genome progenitor *Aegilops tauschii* Coss.: Implications for intraspecific lineage diversification and evolution of common wheat. *Mol. Ecol.* **2010**, *19*, 999–1013. [[CrossRef](#)] [[PubMed](#)]
18. Matsuoka, Y.; Kawahara, T.; Takumi, S. Intraspecific lineage divergence and its association with reproductive trait change during species range expansion in central Eurasian wild wheat *Aegilops tauschii* Coss. (Poaceae). *BMC Evol. Biol.* **2015**, *15*, 213. [[CrossRef](#)] [[PubMed](#)]
19. Matsuoka, Y.; Takumi, S. The role of reproductive isolation in allohexaploid speciation pattern: Empirical insights from the progenitors of common wheat. *Sci. Rep.* **2017**, *7*, 16004. [[CrossRef](#)] [[PubMed](#)]
20. Mizuno, N.; Hosogi, N.; Park, P.; Takumi, S. Hypersensitive response-like reaction is associated with hybrid necrosis in interspecific crosses between tetraploid wheat and *Aegilops tauschii* Coss. *PLoS ONE* **2010**, *5*, e11326. [[CrossRef](#)] [[PubMed](#)]
21. Mizuno, N.; Shitsukawa, N.; Hosogi, N.; Park, P.; Takumi, S. Autoimmune response and repression of mitotic cell division occur in inter-specific crosses between tetraploid wheat and *Aegilops tauschii* Coss. that show low temperature-induced hybrid necrosis. *Plant J.* **2011**, *68*, 114–128. [[CrossRef](#)] [[PubMed](#)]
22. Sakaguchi, K.; Nishijima, R.; Iehisa, J.C.M.; Takumi, S. Fine mapping and genetic association analysis of *Net2*, the causative D-genome locus of low temperature-induced hybrid necrosis in interspecific crosses between tetraploid wheat and *Aegilops tauschii*. *Genetica* **2016**, *144*, 523–533. [[CrossRef](#)] [[PubMed](#)]
23. Brozynska, M.; Furtado, A.; Henry, R.J. Genomics of crop wild relatives: Expanding the gene pool for crop improvement. *Plant Biotechnol. J.* **2016**, *14*, 1070–1085. [[CrossRef](#)] [[PubMed](#)]
24. Ishikawa, G.; Saito, M.; Tanaka, T.; Katayose, Y.; Kanamori, H.; Kurita, K.; Nakamura, T. An efficient approach for the development of genome-specific markers in allohexaploid wheat (*Triticum aestivum* L.) and its application in the construction of high-density linkage maps of the D genome. *DNA Res.* **2018**, *25*, 317–326. [[CrossRef](#)] [[PubMed](#)]
25. Iehisa, J.C.M.; Shimizu, A.; Sato, K.; Nasuda, S.; Takumi, S. Discovery of high-confidence single nucleotide polymorphisms from large-scale de novo analysis of leaf transcripts of *Aegilops tauschii*, a wild wheat progenitor. *DNA Res.* **2012**, *19*, 487–497. [[CrossRef](#)] [[PubMed](#)]
26. Iehisa, J.C.M.; Shimizu, A.; Sato, K.; Nishijima, R.; Sakaguchi, K.; Matsuda, R.; Nasuda, S.; Takumi, S. Genome-wide marker development for the wheat D genome based on single nucleotide polymorphisms identified from transcripts in the wild wheat progenitor *Aegilops tauschii*. *Theor. Appl. Genet.* **2014**, *127*, 261–271. [[CrossRef](#)] [[PubMed](#)]
27. Nishijima, R.; Yoshida, K.; Motoi, Y.; Sato, K.; Takumi, S. Genome-wide identification of novel genetic markers from RNA sequencing assembly of diverse *Aegilops tauschii* accessions. *Mol. Genet. Genom.* **2016**, *291*, 1681–1694. [[CrossRef](#)] [[PubMed](#)]
28. Abe, A.; Kosugi, S.; Yoshida, K.; Natsume, S.; Takagi, H.; Kanzaki, H.; Matsumura, H.; Yoshida, K.; Mitsuoka, C.; Tamiru, M.; et al. Genome sequencing reveals agronomically important loci in rice using MutMap. *Nat. Biotechnol.* **2012**, *30*, 174–178. [[CrossRef](#)] [[PubMed](#)]

29. Takagi, H.; Abe, A.; Yoshida, K.; Kosugi, S.; Natsume, S.; Mitsuoka, C.; Uemura, A.; Utsushi, H.; Tamiru, M.; Takuno, S.; et al. QTL-seq: Rapid mapping of quantitative trait loci in rice by whole genome resequencing of DNA from two bulked populations. *Plant J.* **2013**, *74*, 174–183. [[CrossRef](#)] [[PubMed](#)]
30. Schneeberge, K. Using next-generation sequencing to isolate mutant genes from forward genetic screens. *Nat. Rev. Genet.* **2014**, *15*, 662–676. [[CrossRef](#)] [[PubMed](#)]
31. Zou, C.; Wang, P.; Xu, Y. Bulk sample analysis in genetics, genomics and crop improvement. *Plant Biotechnol. J.* **2016**, *14*, 1941–1955. [[CrossRef](#)] [[PubMed](#)]
32. Liu, S.; Yeh, C.T.; Tang, H.M.; Nettleton, D.; Schnable, P.S. Gene mapping via bulked segregant RNA-Seq (BSR-Seq). *PLoS ONE* **2012**, *7*, e36406. [[CrossRef](#)] [[PubMed](#)]
33. Su, A.; Song, W.; Xing, J.; Zhao, Y.; Zhang, R.; Li, C.; Duan, M.; Luo, M.; Shi, Z.; Zhao, J. Identification of genes potentially associated with the fertility instability of S-type cytoplasmic male sterility in maize via bulked segregant RNA-Seq. *PLoS ONE* **2016**, *11*, e0163489. [[CrossRef](#)] [[PubMed](#)]
34. Du, H.; Zhu, J.; Su, H.; Huang, M.; Wang, H.; Ding, S.; Zhang, B.; Luo, A.; Wei, S.; Tian, X.; et al. Bulk segregant RNA-seq revealed differential expression and SNPs of candidate genes associated with waterlogging tolerance in maize. *Front. Plant Sci.* **2017**, *8*, 1022. [[CrossRef](#)] [[PubMed](#)]
35. Borrill, P.; Adamski, N.; Uauy, C. Genomics as the key to unlocking the polyploid potential of wheat. *New Phytol.* **2015**, *208*, 1008–1022. [[CrossRef](#)] [[PubMed](#)]
36. Trick, M.; Adamski, N.M.; Mugford, S.G.; Jiang, C.C.; Febrer, M.; Uauy, C. Combining SNP discovery from next-generation sequencing data with bulked segregant analysis (BSA) to fine-map genes in polyploid wheat. *BMC Plant Biol.* **2012**, *12*, 14. [[CrossRef](#)] [[PubMed](#)]
37. Ramirez-Gonzalez, R.H.; Segovia, V.; Bird, N.; Fenwick, P.; Holdgate, S.; Bery, S.; Jack, P.; Caccamo, M.; Uauy, C. RNA-Seq bulked segregant analysis enables the identification of high-resolution genetic markers for breeding in hexaploid wheat. *Plant Biotechnol. J.* **2015**, *13*, 613–624. [[CrossRef](#)] [[PubMed](#)]
38. Wu, P.; Xie, J.; Hu, J.; Qiu, D.; Liu, Z.; Li, J.; Li, M.; Zhang, H.; Yang, L.; Liu, H.; et al. Development of molecular markers linked to powdery mildew resistance gene *Pm4b* by combining SNP discovery from transcriptome sequencing data with bulked segregant analysis (BSR-Seq) in wheat. *Front. Plant Sci.* **2018**, *9*, 95. [[CrossRef](#)] [[PubMed](#)]
39. Luo, M.-C.; Gu, Y.Q.; Puiu, D.; Wang, H.; Ywardziok, S.O.; Deal, K.R.; Huo, N.; Zhu, T.; Wang, L.; Wang, Y.; et al. Genome sequence of the progenitor of the wheat D genome *Aegilops tauschii*. *Nature* **2017**, *551*, 498–502. [[CrossRef](#)] [[PubMed](#)]
40. Comai, L. The advantages and disadvantages of being polyploid. *Nat. Rev. Genet.* **2005**, *6*, 836–846. [[CrossRef](#)] [[PubMed](#)]
41. Okada, M.; Yoshida, K.; Takumi, S. Hybrid incompatibilities in interspecific crosses between tetraploid wheat and its wild relative *Aegilops umbellulata*. *Plant Mol. Biol.* **2017**, *95*, 625–645. [[CrossRef](#)] [[PubMed](#)]
42. Miki, Y.; Yoshida, K.; Mizuno, N.; Nasuda, S.; Sato, K.; Takumi, S. Origin of the wheat B-genome chromosomes conferred by RNA sequencing analysis of leaf transcripts in the section *Sitopsis* species of *Aegilops*. *DNA Res.* **2018**, under review.
43. Bolger, A.M.; Lohse, M.; Usadel, B. Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* **2014**, *30*, 2114–2120. [[CrossRef](#)] [[PubMed](#)]
44. Langmead, B.; Salzberg, S.L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **2012**, *9*, 357–359. [[CrossRef](#)] [[PubMed](#)]
45. Li, H.; Handsaker, B.; Wysoker, A.; Fennell, T.; Ruan, J.; Homer, N.; Marth, G.; Abecasis, G.; Durbin, R. Subgroup 1000 Genome Project Data Processing. The sequence alignment/map format and SAMtools. *Bioinformatics* **2009**, *25*, 2078–2079. [[CrossRef](#)] [[PubMed](#)]
46. Kosugi, S.; Natsume, S.; Yoshida, K.; MacLean, D.; Cano, L.; Kamoun, S.; Terauchi, R. Coval: Improving alignment quality and variant calling accuracy for next-generation sequencing data. *PLoS ONE* **2013**, *8*, e75402. [[CrossRef](#)] [[PubMed](#)]
47. Kajimura, T.; Murai, K.; Takumi, S. Distinct genetic regulation of flowering time and grain-filling period based on empirical study of D genome diversity in synthetic hexaploid wheat lines. *Breed. Sci.* **2011**, *61*, 130–141. [[CrossRef](#)]
48. Wu, T.D.; Watanabe, C.K. GMAP: A genomic mapping and alignment program for mRNA and EST sequences. *Bioinformatics* **2005**, *21*, 1859–1875. [[CrossRef](#)] [[PubMed](#)]

49. Quinlan, A.R.; Hall, I.M. BEDTools: A flexible suite of utilities for comparing genomic features. *Bioinformatics* **2010**, *26*, 841–842. [[CrossRef](#)] [[PubMed](#)]
50. RStudio Team. *RStudio: Integrated Development for R*; RStudio, Inc.: Boston, MA, USA, 2016; Available online: <http://www.rstudio.com/> (accessed on 1 November 2016).
51. R Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria; Available online: <https://www.R-project.org/> (accessed on 1 November 2016).
52. Altschul, S.F.; Gish, W.; Miller, W.; Myers, E.W.; Lipman, D.J. Basic local alignment search tool. *J. Mol. Biol.* **1990**, *215*, 403–410. [[CrossRef](#)]
53. Lander, E.S.; Green, P.; Abrahamson, J.; Barlow, A.; Daly, M.J.; Lincoln, S.E.; Newburg, L. MAPMAKER: An interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. *Genomics* **1987**, *1*, 174–181. [[CrossRef](#)]



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).