

# Utility of Triti-Map for bulk-segregated mapping of causal genes and regulatory elements in Triticeae

Fei Zhao<sup>1,2</sup>, Shilong Tian<sup>1,2</sup>, Qihong Wu<sup>3</sup>, Zijuan Li<sup>1,2</sup>, Luhuan Ye<sup>1,2</sup>, Yili Zhuang<sup>1,2</sup>, Meiyue Wang<sup>4</sup>, Yilin Xie<sup>1,2</sup>, Shenghao Zou<sup>5,6</sup>, Wan Teng<sup>2,3</sup>, Yiping Tong<sup>2,3</sup>, Dingzhong Tang<sup>3,5,6</sup>, Ajay Kumar Mahato<sup>7</sup>, Moussa Benhamed<sup>8</sup>, Zhiyong Liu<sup>2,3,\*</sup> and Yijing Zhang<sup>4,\*</sup>

<sup>1</sup>National Key Laboratory of Plant Molecular Genetics, CAS Center for Excellence in Molecular Plant Sciences, Shanghai Institute of Plant Physiology and Ecology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, Shanghai 200032, China

<sup>2</sup>University of the Chinese Academy of Sciences, Beijing 100049, China

<sup>3</sup>State Key Laboratory of Plant Cell and Chromosome Engineering, Institute of Genetics and Developmental Biology, The Innovative Academy of Seed Design, Chinese Academy of Sciences, Beijing 100101, China

<sup>4</sup>State Key Laboratory of Genetic Engineering, Collaborative Innovation Center of Genetics and Development, Department of Biochemistry, Institute of Plant Biology, School of Life Sciences, Fudan University, Shanghai 200438, China

<sup>5</sup>State Key Laboratory of Ecological Control of Fujian-Taiwan Crop Pests, Fujian Agriculture and Forestry University, Fuzhou 350002 China

<sup>6</sup>Key Laboratory of Ministry of Education for Genetics, Breeding and Multiple Utilization of Crops, Plant Immunity Center, Fujian Agriculture and Forestry University, Fuzhou 350002 China

<sup>7</sup>Laboratory of Genome Informatics (LGI) In-charge Bioinformatics Wing-A, First Floor Center for DNA Fingerprinting and Diagnostics Inner Ring Road, Uppal, Hyderabad 500039, India

<sup>8</sup>Institute of Plant Sciences Paris-Saclay (IPS2), Université Paris-Saclay, CNRS, INRAE, Univ Evry, Orsay 91405, France

\*Correspondence: Zhiyong Liu ([zyliu@genetics.ac.cn](mailto:zyliu@genetics.ac.cn)), Yijing Zhang ([zhangyijing@fudan.edu.cn](mailto:zhangyijing@fudan.edu.cn))

<https://doi.org/10.1016/j.xplc.2022.100304>

## ABSTRACT

**Triticeae species, including wheat, barley, and rye, are critical for global food security. Mapping agronomically important genes is crucial for elucidating molecular mechanisms and improving crops. However, Triticeae includes many wild relatives with desirable agronomic traits, and frequent introgressions occurred during Triticeae evolution and domestication. Thus, Triticeae genomes are generally large and complex, making the localization of genes or functional elements that control agronomic traits challenging. Here, we developed Triti-Map, which contains a suite of user-friendly computational packages specifically designed and optimized to overcome the obstacles of gene mapping in Triticeae, as well as a web interface integrating multi-omics data from Triticeae for the efficient mining of genes or functional elements that control particular traits. The Triti-Map pipeline accepts both DNA and RNA bulk-segregated sequencing data as well as traditional QTL data as inputs for locating genes and elucidating their functions. We illustrate the usage of Triti-Map with a combination of bulk-segregated ChIP-seq data to detect a wheat disease-resistance gene with its promoter sequence that is absent from the reference genome and clarify its evolutionary process. We hope that Triti-Map will facilitate gene isolation and accelerate Triticeae breeding.**

**Key words:** agronomic gene mapping, Triticeae, wheat, bulk-segregated ChIP-seq, Triti-Map

Zhao F., Tian S., Wu Q., Li Z., Ye L., Zhuang Y., Wang M., Xie Y., Zou S., Teng W., Tong Y., Tang D., Mahato A.K., Benhamed M., Liu Z., and Zhang Y. (2022). Utility of Triti-Map for bulk-segregated mapping of causal genes and regulatory elements in Triticeae. *Plant Comm.* **3**, 100304.

## INTRODUCTION

Triticeae species, including wheat and barley, are among the major food crops. Transposable element bursts before and after the divergence of Triticeae species have contributed to their large genome sizes (Wicker et al., 2018). For example, the worldwide staple common wheat has a genome size of 16 Gb (International Wheat Genome Sequencing Consortium (IWGSC) et al., 2018). In addition, Triticeae species have complex

genomes, and many wild relatives with desirable agronomic traits have been deployed for crop improvement. Distant hybridizations and introgressions have occurred frequently during evolution and domestication of Triticeae, as well as

Published by the Plant Communications Shanghai Editorial Office in association with Cell Press, an imprint of Elsevier Inc., on behalf of CSPB and CEMPS, CAS.

Data type	Sequencing cost	Library construction	Genomic coverage	SNP identification accuracy	Hardware requirements	References
WGS based	high	Simple	whole genome	no obvious bias	high	<a href="#">Abe et al., 2012</a> ; <a href="#">Fekih et al., 2013</a> ; <a href="#">Takagi et al., 2013a, 2013b</a>
RNA-seq based	low	Simple	expressed gene	affected by gene expression level and alternative splicing	low	<a href="#">Liu et al., 2012</a> ; <a href="#">Hill et al., 2013</a> ; <a href="#">Li et al., 2013</a> ; <a href="#">Zhou et al., 2020</a>
Exome capture	low	complicated	exons designed in probe	affected by reference genome, gene annotation, and probe design	low	<a href="#">Ryan et al., 2013</a> ; <a href="#">Mo et al., 2018</a> ; <a href="#">Dong et al., 2020</a>
ChIP-seq based	low	medium	core-genome, including gene and regulatory elements	no obvious bias	medium	<a href="#">Qi et al., 2018</a> ; <a href="#">Wu et al., 2021</a>

**Table 1. Pros and cons of different sequencing strategies for identifying molecular markers**

modern breeding processes ([Feldman and Levy, 2012](#)). The considerable size and complexity of the genomes of these species are substantial obstacles for researchers trying to localize genes or functional elements that control specific agronomic characteristics.

Recently developed mapping methods (e.g., bulked segregant analysis [BSA]) have significantly decreased the cost and labor associated with genetic research ([Zou et al., 2016](#)). Combining BSA with RNA sequencing (RNA-seq) and exome-seq methods has further lowered the total cost of map-based cloning ([Hill et al., 2013](#); [Miller et al., 2013](#)). Several modified strategies based on BSA have enhanced genetic analysis via improved sequence assembly or optimized calculations ([Abe et al., 2012](#); [Takagi et al., 2013a](#); [Fekih et al., 2013](#)). The accuracy and resolution of map-based cloning depend on the accuracy and distribution of genetic markers that are heterogeneous among populations. [Table 1](#) lists the advantages and disadvantages of different sequencing strategies used to obtain molecular markers in terms of overcoming these obstacles. These segregant strategies have facilitated the detection of essential loci that control target traits.

For Triticeae species with extensive and frequent introgression, however, gene mapping is confronted with the following difficulties:

- (1) High computational cost due to the large genome. Specific packages and optimized parameters are needed to overcome the obstacles presented by the large genomes.
- (2) The causal gene may not be present in the reference genome within the candidate region. Two strategies may be used. The first is to collect syntenic regions and genes from other Triticeae genomes. The second is to enrich functional regions by RNA-seq or chromatin immunoprecipitation sequencing (ChIP-seq), followed by sequencing and *de novo* assembly. A recent report demonstrated that a large fraction of genes and regulatory elements could be captured by ChIP-seq without relying on reference genome sequences ([Qi et al., 2018](#)).
- (3) The causal loci may be present in the regulatory region. For example, a recent report on barley revealed that the deletion of one “TA” short tandem repeat in the promoter re-

gion was sufficient to confer the six-rowed trait ([Wang et al., 2021](#)). The large intergenic regions in Triticeae genomes have a substantial abundance of regulatory elements. To identify the regulatory region and the gene region simultaneously, a new strategy involving the use of ChIP-seq technology to capture the core genome was recently proposed ([Qi et al., 2018](#)).

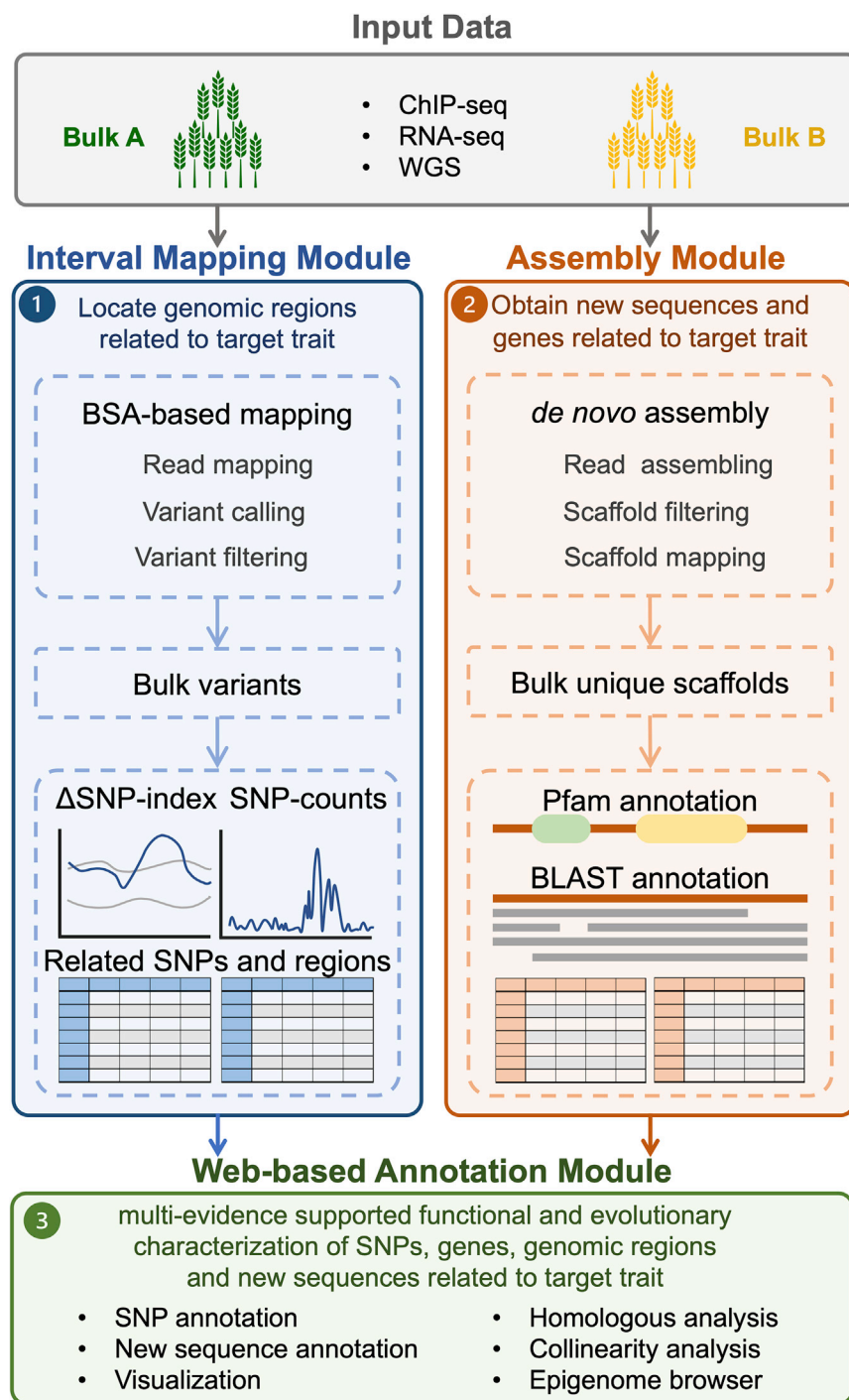
- (4) Because of the generally significant linkage disequilibrium of Triticeae, gene mapping strategies tend to result in large candidate regions, and functional annotation is required to narrow down candidate genes or elements based on multi-omics information. Tissue specificity and specific responses to certain treatments may provide important clues about the causal gene or element, and these require well-processed and organized multi-omics data.

In this study, we introduce Triti-Map, a tool for efficient gene mapping based on bulk-segregated DNA- or RNA-seq data and quantitative trait loci (QTL) data from Triticeae. Triti-Map contains a series of computational packages specifically optimized for Triticeae species, together with a web interface that integrates multi-omics data from Triticeae, to maximize the mining of public data and sequencing results to identify candidate genes. We illustrate how Triti-Map may be used to locate candidate genes that control disease resistance based on the bulked segregant ChIP-seq method, making it an easy-to-use resource for efficient gene mapping in Triticeae species.

## RESULTS

### Computational modules of Triti-Map for locating candidate intervals and detecting specific candidate sequences

The major obstacles to Triticeae gene mapping include their large genomes and the high frequency of introgressed genes that may not be present in reference genomes. Here, we present Triti-Map, which contains a series of computational packages and a web-based interface specifically designed and optimized for Triticeae species to narrow down candidate intervals and detect specific candidate sequences ([Figure 1](#)). [Figure 1](#) illustrates the workflow of the Triti-Map computational modules. Triti-Map uses a pair of bulk-sequencing datasets (another pair of parental

**Figure 1. Triti-Map workflow**

Triti-Map, which accepts raw sequencing data (ChIP-seq, RNA-seq, or WGS data) from bulks with different traits, comprises the interval mapping module (blue) for locating genomic regions associated with a target trait, the *de novo* assembly module (orange) for assembling trait-related sequences, and the web-based annotation module (green) for locating causal variants, candidate genes, or regulatory elements based on integrated multi-omics data and information on Triticeae species.

The assembly module assembles the sequenced reads to identify sequences absent in the reference genome. The assembled sequences specific to the bulk that exhibits the target trait are kept for subsequent analyses. These two computational modules (i.e., the interval mapping module and the *de novo* assembly module) integrate computational steps using Snakemake (Koster and Rahmann, 2012). Users only need to set the basic configuration parameters to complete the analysis and obtain candidate genomic regions and phenotype-associated sequences that are lacking in the reference genome.

#### The web-based annotation module of Triti-Map for locating functional genes and regulatory elements

After detecting a candidate functional interval, candidate genes or regulatory elements must be identified and localized. This is a challenging task, considering the substantial linkage disequilibrium of wheat populations and the frequency of introgressions (Cheng et al., 2019; He et al., 2019; Zhou et al., 2020). Thus, we developed a web-based platform that integrates multi-omics data, sequence data, and functional information to maximize the mining of public data and sequencing results to identify candidate genes (Figure 2A). First, to address the possibility that a causal gene may not be present in the reference genome, a comprehensive collection of collinear regions across Triticeae species, including *Hordeum vulgare*, *Aegilops*

*tauschii*, *Triticum urartu*, *T. dicoccoides*, *T. turgidum*, and *T. aestivum*, is retrieved for any input candidate interval. Lineage-specific and shared genes as well as functional information are listed. Second, for the *de novo* assembled sequences specific to the bulk with the target trait (Figure 2B), functional annotation and phylogenetic analysis of sequences are performed via comparison with all publicly available sequences using the EBI RESTful application programming interface (API) to help narrow down the candidates. Third, for genes within a candidate region, a phylogenetic tree representing the gene

datasets is optional) and a user-set parameter list to reveal candidate regions and sequences. Triti-Map supports the processing of segregated DNA- and RNA-sequencing results and traditional QTL data. The input data types are automatically detected and subjected to appropriate analyses.

The package uses the following analysis pipeline. In brief, the pre-processed reads are analyzed by two modules. The interval mapping module maps the reads to a reference genome, after which a traditional method is used for BSA-based interval detection.

datasets is optional) and a user-set parameter list to reveal candidate regions and sequences. Triti-Map supports the processing of segregated DNA- and RNA-sequencing results and traditional QTL data. The input data types are automatically detected and subjected to appropriate analyses.

evolutionary process is constructed, enabling users to deduce the association between the presence of a gene and a particular trait. Fourth, for SNPs that contribute to interval analyses, the potential functions and feature distributions are presented, which may help to identify functional elements.

Because the enrichment of specific epigenetic markers reflects the presence of active regulatory elements in non-coding regions, we included a search engine and a genome browser for detecting and visualizing epigenetic modifications within candidate regions or regions surrounding candidate genes and SNPs (Figure 2B).

### Optimization to overcome specific challenges in Triticeae gene mapping

The pipeline was optimized in the following ways to address specific challenges in Triticeae gene mapping (Figure 3).

- (1) The software and parameters were optimized to decrease the analysis time. Because of the large genomes and long chromosomes of Triticeae species, commonly used tools for analyzing genomic loci intervals (e.g., bedtools; Quinlan and Hall, 2010) are very slow. We used GIGGLE (Layer et al., 2018), which quickly compares a large number of wheat genomic intervals based on a temporal indexing scheme using a B+ tree to create a single index of the genome intervals, thereby significantly shortening the time required for analysis. To detect variants, we split the genome by chromosomes and used the Genome Analysis Toolkit (GATK) HaplotypeCaller (Van der Auwera et al., 2013) for parallel analysis. The default alignment program for DNA-seq-type sequence alignment is BWA-mem2 (Vasimuddin et al., 2019), which is faster than bwa (Li and Durbin, 2009) because of enhanced cache reuse, simplified algorithms, and the use of single instruction, multiple data (SIMD) wherever applicable. The alignment results produced by this program are identical to those of bwa.
- (2) Given the frequent introgressions and distant hybridizations between Triticeae species, a candidate gene may be absent in the mapped interval of the reference genome. Triti-Map can retrieve new bulk-specific sequences via *de novo* assembly and comparison. Furthermore, it collects Triticeae genomic regions colinear with candidate regions and provides detailed functional annotations to increase the chances of identifying a candidate gene for a particular trait.
- (3) Regarding sequence comparisons and functional annotations, because of the rapid increase in available genome sequences and multi-omics data, rather than using a local database, the pipeline uses the European Molecular Biology Laboratory-European Bioinformatics Institute (EMBL-EBI) RESTful APIs (Madeira et al., 2019) to obtain up-to-date sequence information.

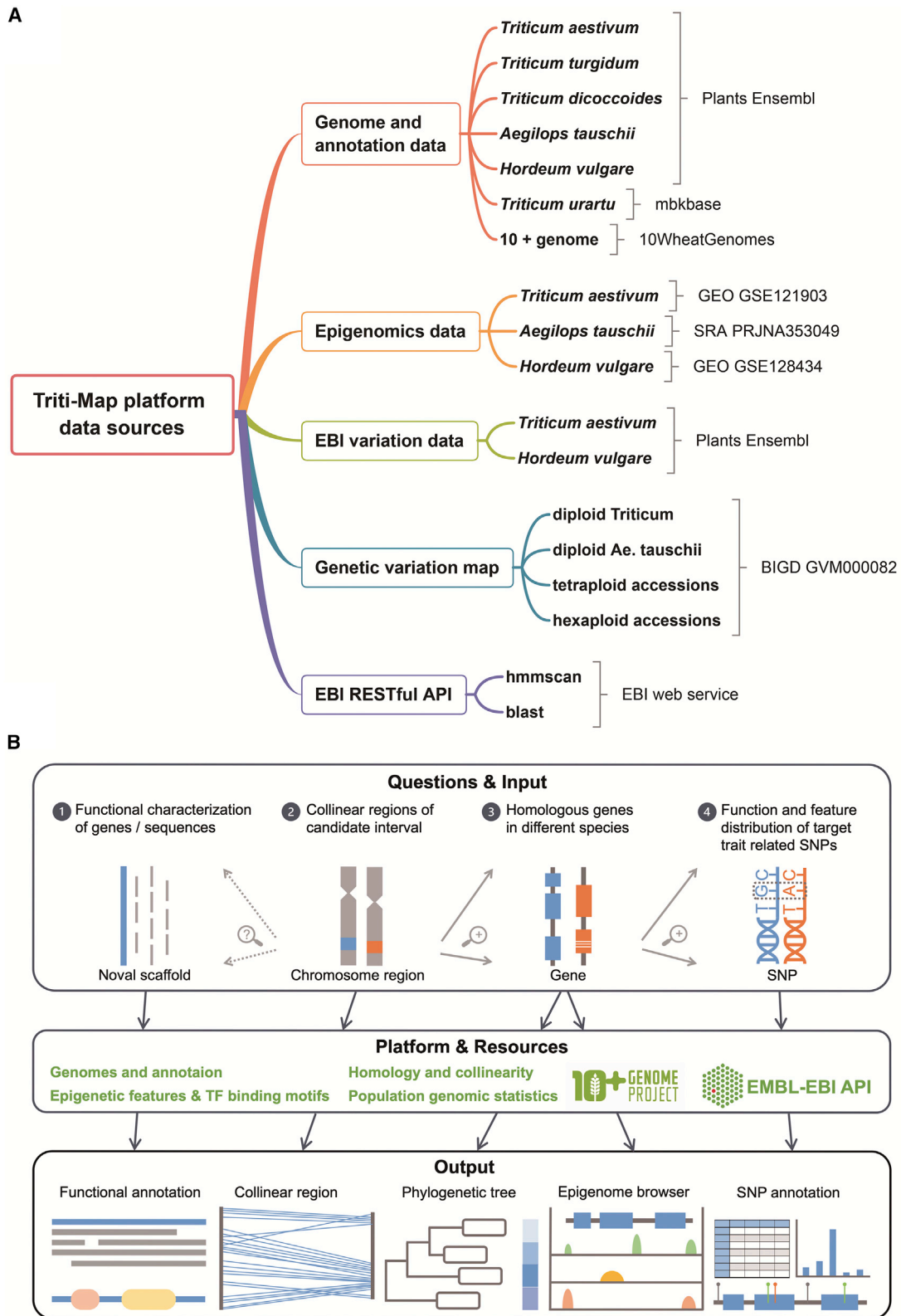
Multiple strategies were used to simplify the use of Triti-Map. The software and environment required for analyses can be quickly deployed through Conda. All of the analysis modules were developed and integrated based on the Snakemake workflow management system. Moreover, the configuration is simple, flexible, and easy to use. Interval mapping (Figure 3, left) and *de novo* sequence assembly (Figure 3, right) can be conducted

separately or together with other modules (Figure 3). The Triti-Map usage document (<https://github.com/fei0810/Triti-Map/wiki>) provides complete instructions and a case study.

### Triti-Map and bulk-segregated ChIP-seq used to identify a disease-resistance gene and its promoter region not present in the reference genome

We next illustrate how Triti-Map and bulk-segregated ChIP-seq help to detect a disease-resistance gene that is not present in the reference genome as recently reported (Wu et al., 2021). Two bulked segregant pools were collected from the F<sub>2</sub> progeny of a cross between Xueza0 (susceptible to powdery mildew) and 3D249 (resistant to powdery mildew). For each pooled sample, a ChIP-seq analysis was performed for three histone marks (H3K4me3, H3K27me3, and H3K36me3) that are closely associated with gene activities. The sequencing data were analyzed using the interval mapping module of the Triti-Map package, resulting in the identification of a 6-Mb region on chromosome 7A (chr7A: 724,111,912–730,119,678) that is highly associated with powdery mildew resistance (Figure 4A, Supplemental Figures 1 and 2). High-quality SNPs within this region were used as input data for the Triti-Map web-based annotation module. Nonsynonymous mutations were detected in two disease resistance-related genes (*TraesCS7A02G551900* and *TraesCS7A02G555200*). In addition, the integrated epigenetic and motif information revealed the presence of the DNase I hypersensitive site (DHS), as well as H3K36me3, H3K4me3, and H3K9ac, in the *TraesCS7A02G551900* promoter, suggesting that the promoter is highly accessible to transcription factors (e.g., autonomously replicating sequence-binding factor 1 [ABF1]) (Figure 4B; Supplemental Table 2). However, the results of an experimental validation indicated that these two genes did not fully segregate with the disease-resistance trait.

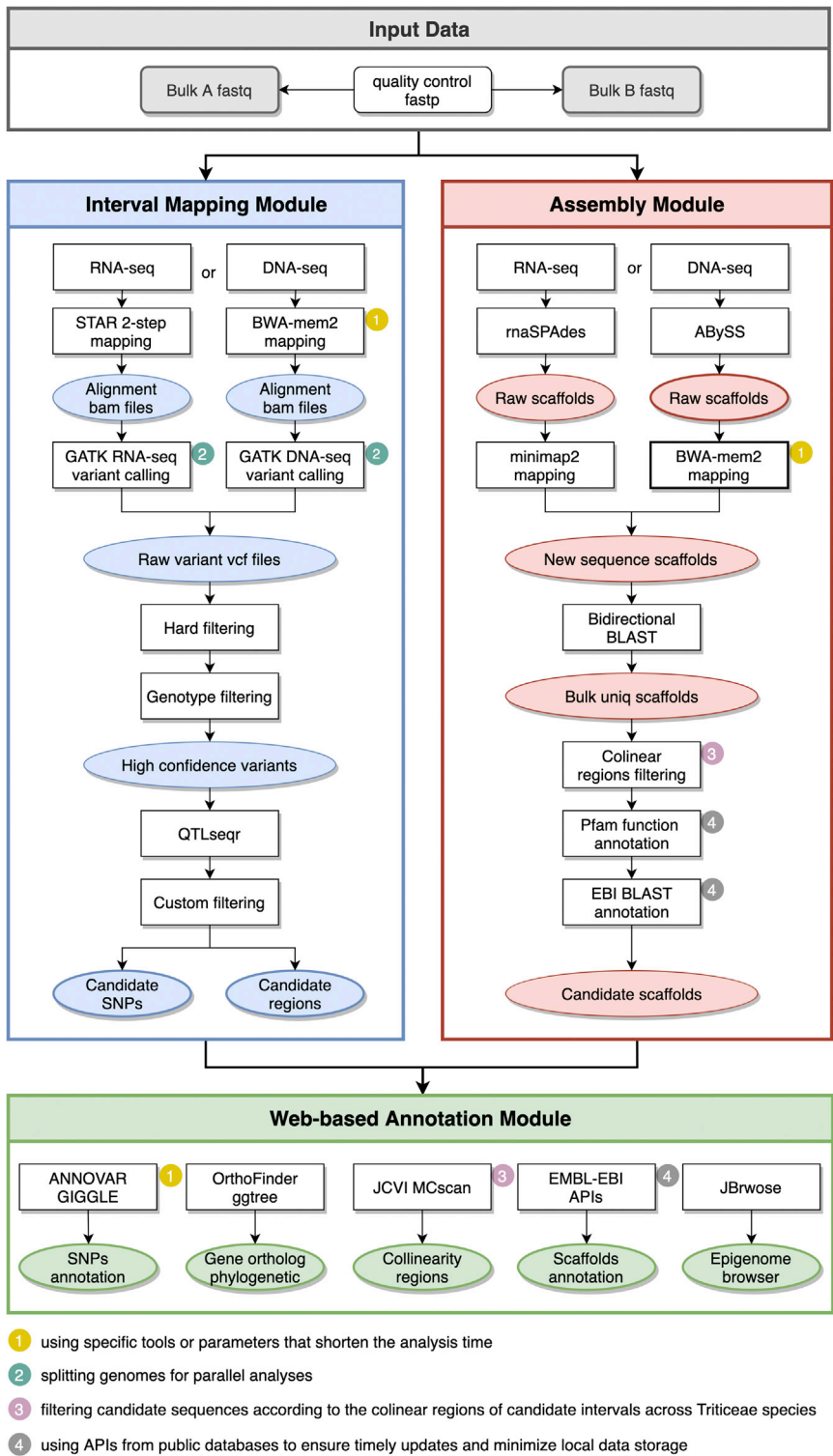
We hypothesized that the candidate gene was present in other species or wheat populations but not in the reference genome. We used the assembly module to detect disease resistance-specific sequences. Among the 10,429 resistant pool-specific scaffolds, 1,704 were partially mapped to Triticeae genomic regions colinear with the candidate region using the collinearity function of the web-based annotation module (Figure 4C). A subsequent functional annotation using the EMBL-EBI hmmscan API detected nine sequences encoding protein domains common among *R* genes. In addition, four sequences encoding the NB-ARC (nucleotide-binding domain shared with apoptotic protease activating factor-1 [APAF-1], various *R* proteins, and CED-4) domain and five sequences encoding the leucine-rich repeat (LRR) domain were detected (Supplemental Table 3). The EMBL-EBI blast API of the assembly module detected sequences encoding one NB-ARC domain and one LRR domain that were highly similar (>99.7% sequence identity) to *Pm60*, which was originally identified as the gene responsible for powdery mildew resistance in the diploid species *T. urartu* (Zou et al., 2018) (Figure 4D; Supplemental Table 4). These two sequences accounted for 69% of the total length of *Pm60*. The *Pm60* sequence was further extended using the sequencing data, resulting in a 235-bp extension at the 5' end and a 203-bp extension at the 3' end (10% of the total length of *Pm60*) (Qi et al., 2018). The 5' end was enriched for both active and repressive marks, indicating that this gene is in a bivalent state (Qi et al., 2018).



**Figure 2. Diagram of the web-based annotation module function**

**(A)** Data integrated with the web-based annotation module.

**(B)** To locate causal variants and candidate genes or regulatory elements, Triti-Map integrates multi-omics data and provides different levels of analysis, including a collinearity analysis of target regions among Triticeae species, as well as a functional and evolutionary characterization of SNPs, genes, or other sequences related to a target trait.



Triti-Map for Triticeae gene mapping

**Figure 3. Optimization to address specific challenges of Triticeae gene mapping and annotation**

The major steps that were optimized are marked by the following numbers: (1) steps that use specific tools or parameters that shorten the analysis time; (2) steps that split genomes for parallel analyses; (3) steps in which candidate sequences are filtered according to the colinear regions of candidate intervals across Triticeae species; (4) steps that use APIs from public databases to ensure timely updates and minimize local data storage. In each module, nodes with a colored background represent important result files, whereas nodes without a colored background represent the main analysis steps and the tools used.

A and hexaploid A subgenomes, implying that *Pm60* originated in the common ancestor of diploid wheat but was lost in the A subgenome progenitor before tetraploidization. The corresponding genes in subgenomes B and D underwent divergent evolution and no longer contributed to powdery mildew resistance (Figure 4D). Considered together, these results demonstrate that bulk-segregated ChIP-seq and Triti-Map enable the rapid identification of a candidate causal gene for a specific phenotype that is not present in the reference genome, facilitating functional and evolutionary analyses.

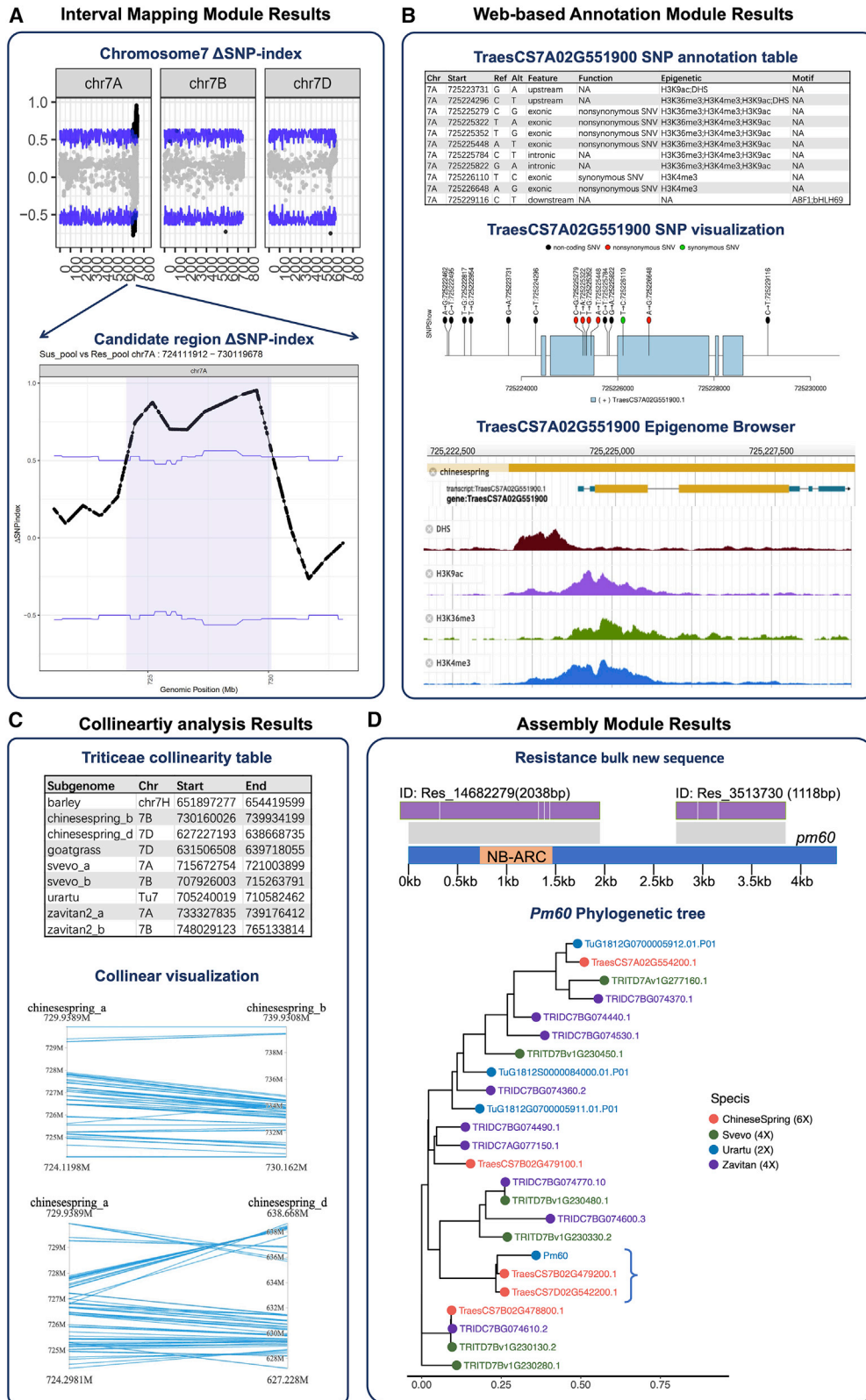
**DISCUSSION**

In this study, we developed Triti-Map, consisting of a suite of scripts and a web-based platform, specifically to address the challenges of Triticeae gene mapping (i.e., large genomes and frequent introgressions), including cases in which the candidate gene is not present in the available reference genome sequences.

First, this pipeline increased the likelihood of identifying agronomic genes absent from the reference genome by integrating information on *de novo* assembled sequences, colinear regions in Triticeae species corresponding to candidate loci, and homologous sequences in public databases. Second, because the ChIP-based strategy

A candidate gene may be introduced from distant relatives or lost during evolution. To assess these two possibilities, we performed collinearity and evolutionary analyses. Collinear regions corresponding to the candidate region of chromosome 7A were detected in both chromosome 7B and chromosome 7D, both of which contain an *R* gene with high sequence homology to *Pm60*. No highly homologous gene was detected in the tetraploid

could help capture core genomic regions, including genes and regulatory elements (Qi et al., 2018), bulk-segregated ChIP-seq facilitated the detection of non-reference genes and elements and decreased the time and labor required for the analysis and enrichment of the core genome, including both gene body and regulatory regions. Third, in addition to ChIP-seq data, the Triti-Map pipeline accepts other types of DNA- and RNA-seq data



**Figure 4. Triti-Map case study results**

(A) Interval mapping module results. Upper panel: causal interval detected using the  $\Delta$ SNP-index method. Lower panel: enlarged candidate region. (B) Web-based annotation module results. From top to bottom: SNP annotation, SNP localization, and epigenome tracks of related regions. (C) Collinearity analysis results. The regions in Triticeae species that are collinear with the detected candidate region are listed. (D) Assembly module results. Top: Two newly assembled sequences (purple) highly similar to *Pm60*. Bottom: Phylogenetic tree presenting the evolutionary distance between *Pm60* and homologous genes in wheat species with a different ploidy level.

and traditional QTL data as the input for locating genes and elucidating their functions. The workflow of this pipeline based on Snakemake is easy to maintain and expand. The ability to add more interval mapping and analysis modules later contributes to the broad utility and flexibility of Triti-Map. Fourth, a comprehensive collection of multi-omics data and a systematic curation of functional and evolutionary information facilitate the functional characterization of sequences. This is useful for precisely and reliably localizing candidate genes and for hypothesis-driven research into specific mechanisms.

There are several publicly available web-based resources for wheat genomic data mining, including resources for visualizing and analyzing BSA results (Zhang et al., 2021) and for identifying collinearity and homology among Triticeae species (Chen et al., 2020). A database of wheat genomic variations and a wheat multi-omics database are also available (Blake et al., 2019; Wang et al., 2020; Chen et al., 2020; Zhang et al., 2021; Ma et al., 2021). Table 2 compares Triti-Map with five wheat genome variation databases and published wheat multi-omics databases. Further mining of the data obtained from bulk-segregated ChIP-seq and Triti-Map using these resources will provide insights into downstream mechanisms.

We anticipate that Triti-Map will improve gene isolation and cloning as well as breeding in Triticeae by decreasing the time and labor required to identify agronomically important genes.

## METHODS

### Data pre-processing

The following data processing pipeline is included in Triti-Map. Triti-Map accepts ChIP-seq and whole-genome sequencing (WGS) (DNA-seq data) or RNA-seq data as the input. The fastp (0.20.1) program (Chen et al., 2018) is used to remove adapter sequences and low-quality sequencing bases from the raw data. The pre-processed DNA-seq reads are mapped to the reference genome using the default parameters of BWA-mem2 (Vasimuddin et al., 2019). The RNA-seq reads are mapped using the two-pass mode of STAR (Dobin et al., 2013). In brief, the reads are mapped to the reference genome, after which the junction information obtained from the mapping results is used to reconstruct the genome index and perform a second round of mapping.

### Variant detection and candidate genomic interval identification

To detect variants from the DNA-seq data, GATK (Van der Auwera et al., 2013) MarkDuplicates is used to remove the duplicated reads resulting from PCR amplification errors during the library preparation step. The RNA-seq data are processed using the pipeline recommended by GATK (<https://gatk.broadinstitute.org/>). Uniquely mapped reads are extracted, and the variant sites are detected using GATK HaplotypeCaller. To detect mutations, the genome is split by chromosomes for parallel computing. After all of the processes are completed, the data for the variants on each chromosome are merged using GATK MergeVcfs to generate a variant call format (VCF) file containing information on all of the mutation sites.

The variant sites are further filtered according to the quality of the variant sites and genotype information using GATK SelectVariants and GATK VariantFiltration. The following combined criteria are used: QualByDepth (QD) >2, FisherStrand (FS) <60, RMSMappingQuality (MQ) >40, StrandOddsRatio (SOR) <10, ReadPosRankSum > -8.0, MQRankSum > -12.5, and QUAL >30. To ensure that SNPs are accurately identified, SNPs within a 35-bp sequence are also filtered. Genotype filtering is performed depend-

ing on the availability of the parental data and information on dominant or recessive alleles. Moreover, mutation sites that satisfy the following conditions are also eliminated: homozygous and identical mutations in two mixed pools; homozygous and identical mutations in the parents; heterozygous mutations in recessive parents or mixed pools; and mutations in recessive mixed pools that are homozygous and different from those in the recessive parent. The filtered VCF file is converted to a matrix format using a self-written bash script for subsequent interval mapping.

The current Triti-Map pipeline supports the use of the allele frequency difference ( $\Delta$ SNP index) method (Takagi et al., 2013b) for mapping candidate intervals. The QTLseql (Mansfeld and Grumet, 2018) package is used to calculate the  $\Delta$ SNP index and identify the trait-associated interval. The results are filtered and visualized using a self-written R script.

### De novo assembly and bulk-specific sequence screening and annotation

The pre-processed DNA-seq data are assembled using ABySS (version 2.0.2) (Jackman et al., 2017), with  $k = 90$ . The transcripts detected by RNA-seq are assembled using the default parameters of rnaSPAdes (version 3.15.0) (Bushmanova et al., 2019).

Assembled sequences longer than 500 bp are retained and mapped to the reference genome using the default parameters of bwa-mem2 (DNA-seq) and minimap2 -ax splice (RNA-seq) (Li, 2018). The assembled sequences of the two mixed pools that are not aligned or partially aligned (i.e., reads containing soft-clipped fragments or with an alternative alignment with cigar strings SA:Z and XA:Z) to the reference genome are selected for subsequent analysis. The screening process involves the SeqKit (Shen et al., 2016) and BEDOPS (version 2.4.36) (Neph et al., 2012) programs, as well as a self-written bash script.

Target trait-related bulk-specific sequences are obtained via a reciprocal Basic Local Alignment Search Tool (BLAST) analysis of sequences from two bulks. To further narrow down the candidate sequences, all regions of Triticeae genomes colinear with the candidate intervals identified using the interval mapping module are identified. Bulk-specific sequences highly similar to these regions are considered to be candidate sequences (BLAST default parameters). The candidate sequences are functionally annotated using the EBI HMMER3 hmmscan API. Their homologous sequences and information on their functions in plants are retrieved from the Ensembl Plants database using the EBI BLAST API. Combining the above information, Triti-Map will generate a table of functional annotations and homologous sequence information for the new candidate sequences and positional information for comparison with the reference genome.

### Triticeae species data collection and processing

Genome sequences and annotation details are obtained from the Ensembl database (Yates et al., 2020) for the following six Triticeae species: *H. vulgare*, *Ae. tauschii*, *T. urartu*, *T. dicoccoides*, *T. turgidum*, and *T. aestivum*. Information is also retrieved for 10 hexaploid wheat genomes (Walkowiak et al., 2020). Epigenomic data from Triticeae species are compiled, including different histone modification ChIP-seq and DNase I hypersensitive site (DHS) data (Figure 2A; Supplemental Table 1). In addition, model-based analysis of ChIP-seq (MACS) (Zhang et al., 2008) is used to identify read-enriched regions, and MotifScan (Sun et al., 2018) is used to locate transcription factor-binding motifs. OrthoFinder (Emms and Kelly, 2019) is used to identify orthologous genes across Triticeae species. JCVI MCscan (Tang et al., 2008) is used to identify gene pairs in colinear regions. EggNOG-mapper (Huerta-Cepas et al., 2017) is used to functionally annotate genes in different Triticeae species, and evolutionary Genotype-Phenotype Systems (eGPS) (Yu et al., 2019) is used to perform a population genetics analysis with a high-density genetic variation map (VMap 1.0) of wheat (Zhou et al., 2020).



	<b>GrainGenes</b> Blake et al. (2019)	<b>Wheat-SnpHub-Portal</b> Wang et al. (2020)	<b>GeneTribe</b> Chen et al., 2020	<b>WheatGmap</b> Zhang et al., 2021	<b>WheatOmics</b> Ma et al., 2021	<b>Triti-Map</b>
Description	an improved resource for the small-grains community	SnpHub, an easy-to-set up web server framework for exploring large-scale genomic variation data in the post-genomic era with applications in wheat	Triticeae GeneTribe, a collinearity-incorporating homology inference strategy for connecting emerging assemblies in the Triticeae Tribe as a pilot practice in the plant pangenomic era	WheatGmap, which integrates multiple BSA mapping models and large amounts of public data to accelerate gene cloning and functional research and facilitate resource sharing	WheatOmics, a platform combining multiple omics data to accelerate functional genomics studies in wheat	Triti-Map is composed of both gene mapping scripts and downstream analysis tools for efficient mapping of both candidate gene and intergenic regulatory elements
Year founded	2005	2020	2020	2020	2018	2021
URL	<a href="https://wheat.pw.usda.gov">https://wheat.pw.usda.gov</a>	<a href="http://guoweilong.github.io/SnpHub">http://guoweilong.github.io/SnpHub</a>	<a href="https://chenym1.github.io/genetribel/">https://chenym1.github.io/genetribel/</a>	<a href="https://www.wheatgmap.org">https://www.wheatgmap.org</a>	<a href="http://wheatomics.sdau.edu.cn/">http://wheatomics.sdau.edu.cn/</a>	<a href="http://bioinfo.cemps.ac.cn/tritimapp">http://bioinfo.cemps.ac.cn/tritimapp</a>
Applicable platform	web	web and Linux	web	web	web	<b>web and Linux</b>
Input data type	/	/	gene, gene list, or fasta file	bulk-sequencing VCF	gene, gene list, or fasta file	<b>fastq raw data</b>
				gene, gene list, or fasta file		bulk-sequencing VCF
						gene, gene list, or fasta file
Data in website	genome, molecular, and phenotypic information for 4 wheat relative species	genomic variation datasets of 7 wheats and their progenitors	homology inference information for 12 Triticeae and 3 outgroup species	high-throughput BSA sequencing datasets of hexaploid wheat (>3,500 groups)	multi-omics data, including genomes, transcriptomes, variomes, and epigenomes of multiple Triticeae species	multi-omics data, including genomes, transcriptomes, genetic variation, and epigenomes of 7 major Triticeae species
Website main function module	blast function	raw variation data and genomic sequence retrieval	collinear information and analysis	gene mapping	multi-omics data information	<b>de novo gene mapping</b>
				gene and SNP annotations	transcriptional analysis	gene and SNP annotations
	transcriptional analysis			regulatory element analysis	collinear information and analysis	
	genome browsers			functional analysis	epigenetic features and transcription factor binding motifs	
	genome browsers			gene identification	<b>population genomic statistics</b>	
		gene identification				
Gene mapping tool	no	No	no	yes	no	<b>yes</b>

**Table 2. Comparison of features between Triti-Map and published databases**

/, no data available. Bold indicates the uniqueness of Triti-Map across multiple feature comparisons.

### Web-based platform construction

A web-based platform was developed using R Shiny, and the front frame was produced using bs4Dash. In this platform, ANNOVAR (Wang et al., 2010) is used to annotate the uploaded mutation information, and GIGGLE (Layer et al., 2018) is used to annotate epigenetic modifications and motifs. The annotations and the results of other analyses are displayed in a formatted table using the R reactable package. The distribution of variant site positions on genes is displayed using the R trackViewer (Ou and Zhu, 2019) package. The distribution of the variant site features, the distribution of chromosomal density, and the results of the collinearity analysis are visualized using echarts4r and plotly.R. The EBI API (Madeira et al., 2019) is used for the functional annotation of sequences and the determination of sequence similarity. The evolutionary tree for homologous genes from different subgenomes is constructed using ggtree (Yu et al., 2017). The genome browser containing data for the apparent modifications in each species is developed based on the JBrowse (Buels et al., 2016) configuration.

### Availability of the Triti-Map package and web-based interface

The Triti-Map package was developed using Snakemake (Koster and Rahmann, 2012) and Conda. The package and manual are available online (<https://github.com/fei0810/Triti-Map>). Triti-Map can be installed from Bioconda (Grüning et al., 2018) and Docker. The web-based Triti-Map interface is also accessible (<http://bioinfo.cemps.ac.cn/tritimap>).

### Sample and sequencing data processing for a case study involving the detection of a disease-resistance gene

The powdery mildew-resistant common wheat cultivar 3D249 is an F<sub>7</sub> wheat-WEW (wild emmer wheat) introgression line developed by Professor Tsomin Yang of China Agricultural University, Beijing (pedigree: Jingshuang 27//Yanda, 1817/WE18/3/Wenmai 4). The common wheat cultivar Xuezaio is highly sensitive to *Blumeria graminis* f. sp. *graminis* (Bgt) #E09. Two-week-old seedlings of 30 F<sub>3</sub> generation homozygous resistant and susceptible materials derived from a Xuezaio × 3D249 hybridization were pooled to construct resistant and susceptible DNA bulks for a ChIP-seq analysis. The ChIP experiments involved antibodies specific for H3K27me3 (Upstate, cat. no. 07-449), H3K4me3 (Abcam, cat. no. Ab8580), and H3K36me3 (Abcam, cat. no. Ab9050). The Illumina HiSeq 2500 system was used for sequencing (150-bp paired-end reads) (Beijing Nuohe Company).

### Data access

The package and manual are available online (<https://github.com/fei0810/Triti-Map>). Triti-Map can be installed from Bioconda and Docker. The web-based Triti-Map interface is also accessible (<http://bioinfo.cemps.ac.cn/tritimap>).

The ChIP-seq data have been deposited in the Sequence Read Archive (SRA) and assigned the identifier accession PRJNA725543 (<https://www.ncbi.nlm.nih.gov/bioproject/PRJNA725543>).

### SUPPLEMENTAL INFORMATION

Supplemental information can be found online at *Plant Communications Online*.

### FUNDING

This study was supported by the National Science Fund for Excellent Young Scholars (32022012).

### AUTHOR CONTRIBUTIONS

Z. Liu and Y. Zhang conceived the project. F.Z. designed the software. F.Z. and S.T. performed the software coding. F.Z., M.W., S.T., and Y.X. analyzed the data. Z. Li, L.Y., Y. Zhuang, and Q.W. performed the experiments. F.Z., A.K.M, M.B. and Y.Z. prepared the figures and wrote the

manuscript. The other co-authors critically reviewed and modified the manuscript.

### ACKNOWLEDGMENTS

We thank Huang Tao for his help in maintaining the high-performance computing server. We thank Zhongeng Xu for insightful communications. We are grateful to Martin Mascher for facilitating access to 10+ Wheat Genomes data. We thank Robbie Lewis, MSc, from Liwen Bianji (Edanz) ([www.liwenbianji.cn/](http://www.liwenbianji.cn/)) for editing a draft of this manuscript. The authors declare that they have no competing interests.

Received: July 5, 2021

Revised: January 13, 2022

Accepted: February 13, 2022

Published: February 18, 2022

### REFERENCES

- Abe, A., Kosugi, S., Yoshida, K., Natsume, S., Takagi, H., Kanzaki, H., Matsumura, H., Yoshida, K., Mitsuoka, C., Tamiru, M., et al. (2012). Genome sequencing reveals agronomically important loci in rice using MutMap. *Nat. Biotechnol.* **30**:174–178.
- Blake, V.C., Woodhouse, M.R., Lazo, G.R., Odell, S.G., Wight, C.P., Tinker, N.A., Wang, Y., Gu, Y.Q., Birkett, C.L., Jannink, J.L., et al. (2019). GrainGenes: centralized small grain resources and digital platform for geneticists and breeders. *Database* **2019**:baz065.
- Buels, R., Yao, E., Diesh, C.M., Hayes, R.D., Munoz-Torres, M., Helt, G., Goodstein, D.M., Elsik, C.G., Lewis, S.E., Stein, L., et al. (2016). JBrowse: a dynamic web platform for genome visualization and analysis. *Genome Biol.* **17**:1–12.
- Bushmanova, E., Antipov, D., Lapidus, A., and Prjibelski, A.D. (2019). rnaSPAdes: a de novo transcriptome assembler and its application to RNA-Seq data. *GigaScience* **8**:giz100.
- Chen, S., Zhou, Y., Chen, Y., and Gu, J. (2018). fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* **34**:i884–i890.
- Chen, Y., Song, W., Xie, X., Wang, Z., Guan, P., Peng, H., Jiao, Y., Ni, Z., Sun, Q., and Guo, W. (2020). A collinearity-incorporating homology inference strategy for connecting emerging assemblies in the Triticeae species as a pilot practice in the plant pangenomic era. *Mol. Plant* **13**:1694–1708.
- Cheng, H., Liu, J., Wen, J., Nie, X., Xu, L., Chen, N., Li, Z., Wang, Q., Zheng, Z., Li, M., et al. (2019). Frequent intra- and inter-species introgression shapes the landscape of genetic variation in bread wheat. *Genome Biol.* **20**:136.
- Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**:15–21.
- Dong, C., Zhang, L., Chen, Z., Xia, C., Gu, Y., Wang, J., Li, D., Xie, Z., Zhang, Q., Zhang, X., et al. (2020). Combining a new exome capture panel with an effective varBScore algorithm accelerates BSA-based gene cloning in wheat. *Front. Plant Sci.* **11**:1249.
- Emms, D.M., and Kelly, S. (2019). OrthoFinder: phylogenetic orthology inference for comparative genomics. *Genome Biol.* **20**:238.
- Fekih, R., Takagi, H., Tamiru, M., Abe, A., Natsume, S., Yaegashi, H., Sharma, S., Sharma, S., Kanzaki, H., Matsumura, H., et al. (2013). MutMap+: genetic mapping and mutant identification without crossing in rice. *PLoS One* **8**, e68529.
- Feldman, M., and Levy, A.A. (2012). Genome evolution due to allopolyploidization in wheat. *Genetics* **192**:763–774.
- Grüning, B., Dale, R., Sjödin, A., Chapman, B.A., Rowe, J., Tomkins-Tinch, C.H., Valieris, R., and Köster, J.; Bioconda Team (2018). Bioconda: sustainable and comprehensive software distribution for the life sciences. *Nat. Methods* **15**:475–476.

- He, F., Pasam, R., Shi, F., Kant, S., Keeble-Gagnere, G., Kay, P., Forrest, K., Fritz, A., Hucl, P., Wiebe, K., et al. (2019). Exome sequencing highlights the role of wild-relative introgression in shaping the adaptive landscape of the wheat genome. *Nat. Genet.* **51**:896–904.
- Hill, J.T., Demarest, B.L., Bisgrove, B.W., Gorski, B., Su, Y.-C., and Yost, H.J. (2013). MMAPPR: mutation mapping analysis pipeline for pooled RNA-seq. *Genome Res.* **23**:687–697.
- Huerta-Cepas, J., Forslund, K., Coelho, L.P., Szklarczyk, D., Jensen, L.J., von Mering, C., and Bork, P. (2017). Fast genome-wide functional annotation through orthology assignment by eggNOG-mapper. *Mol. Biol. Evol.* **34**:2115–2122.
- International Wheat Genome Sequencing Consortium (IWGSC), IWGSC RefSeq principal investigators, Appels, R., Eversole, K., Feuillet, C., Keller, B., Rogers, J., Stein, N., IWGSC whole-genome assembly principal investigators, Pozniak, C.J., et al. (2018). Shifting the limits in wheat research and breeding using a fully annotated reference genome. *Science* **361**, eaar7191.
- Jackman, S.D., Vandervalk, B.P., Mohamadi, H., Chu, J., Yeo, S., Hammond, S.A., Jahesh, G., Khan, H., Coombe, L., Warren, R.L., et al. (2017). ABySS 2.0: resource-efficient assembly of large genomes using a Bloom filter. *Genome Res.* **27**:768–777.
- Koster, J., and Rahmann, S. (2012). Snakemake—a scalable bioinformatics workflow engine. *Bioinformatics* **28**:2520–2522.
- Layer, R.M., Pedersen, B.S., DiSera, T., Marth, G.T., Gertz, J., and Quinlan, A.R. (2018). GIGGLE: a search engine for large-scale integrated genome analysis. *Nat. Methods* **15**:123–126.
- Li, H. (2018). Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics* **34**:3094–3100.
- Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with burrows-wheeler transform. *Bioinformatics* **25**:1754–1760.
- Li, L., Li, D., Liu, S., Ma, X., Dietrich, C.R., Hu, H.-C., Zhang, G., Liu, Z., Zheng, J., Wang, G., et al. (2013). The maize glossy13 gene, cloned via BSR-seq and seq-walking encodes a putative ABC transporter required for the normal accumulation of epicuticular waxes. *PLoS One* **8**, e82333.
- Liu, S., Yeh, C.-T., Tang, H.M., Nettleton, D., and Schnable, P.S. (2012). Gene mapping via bulked segregant RNA-seq (BSR-Seq). *PLoS One* **7**, e36406.
- Ma, S.W., Wang, M., Wu, J., Guo, W., Chen, Y., Li, G., Wang, Y., Shi, W., Xia, G., Fu, D., et al. (2021). WheatOmics: a platform combining multiple omics data to accelerate functional genomics studies in wheat. *Mol. Plant* **14**:1965–1968.
- Madeira, F., Park, Y.M., Lee, J., Buso, N., Gur, T., Madhusoodanan, N., Basutkar, P., Tivey, A.R.N., Potter, S.C., Finn, R.D., et al. (2019). The EMBL-EBI search and sequence analysis tools APIs in 2019. *Nucleic Acids Res.* **47**:W636–W641.
- Mansfeld, B.N., and Grumet, R. (2018). QTLseqr: an R package for bulk segregant analysis with next-generation sequencing. *Plant Genome* **11**.
- Miller, A.C., Obholzer, N.D., Shah, A.N., Megason, S.G., and Moens, C.B. (2013). RNA-seq-based mapping and candidate identification of mutations from forward genetic screens. *Genome Res.* **23**:679–686.
- Mo, Y., Howell, T., Vasquez-Gross, H., de Haro, L.A., Dubcovsky, J., and Pearce, S. (2018). Mapping causal mutations by exome sequencing in a wheat TILLING population: a tall mutant case study. *Mol. Genet. Genomics* **293**:463–477.
- Neph, S., Kuehn, M.S., Reynolds, A.P., Haugen, E., Thurman, R.E., Johnson, A.K., Rynes, E., Maurano, M.T., Vierstra, J., Thomas, S., et al. (2012). BEDOPS: high-performance genomic feature operations. *Bioinformatics* **28**:1919–1920.
- Ou, J., and Zhu, L.J. (2019). trackViewer: a Bioconductor package for interactive and integrative visualization of multi-omics data. *Nat. Methods* **16**:453–454.
- Qi, M., Li, Z., Liu, C., Hu, W., Ye, L., Xie, Y., Zhuang, Y., Zhao, F., Teng, W., Zheng, Q., et al. (2018). CGT-seq: epigenome-guided de novo assembly of the core genome for divergent populations with large genome. *Nucleic Acids Res.* **46**:e107.
- Quinlan, A.R., and Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**:841–842.
- Ryan, S., Willer, J., Marjoram, L., Bagwell, J., Mankiewicz, J., Leshchiner, I., Goessling, W., Bagnat, M., and Katsanis, N. (2013). Rapid identification of kidney cyst mutations by whole exome sequencing in zebrafish. *Development* **140**:4445–4451.
- Shen, W., Le, S., Li, Y., and Hu, F. (2016). SeqKit: a cross-platform and ultrafast toolkit for FASTA/Q file manipulation. *PLoS One* **11**, e0163962.
- Sun, H., Wang, J., Gong, Z., Yao, J., Wang, Y., Xu, J., Yuan, G., Zhang, Y., and Shao, Z. (2018). Quantitative integration of epigenomic variation and transcription factor binding using MAmotif toolkit identifies an important role of IRF2 as transcription activator at gene promoters. *Cell Discov.* **4**:1–4.
- Takagi, H., Uemura, A., Yaegashi, H., Tamiru, M., Abe, A., Mitsuoka, C., Utsushi, H., Natsume, S., Kanzaki, H., Matsumura, H., et al. (2013a). MutMap-Gap: whole-genome resequencing of mutant F2 progeny bulk combined with de novo assembly of gap regions identifies the rice blast resistance gene Pii. *New Phytol.* **200**:276–283.
- Takagi, H., Abe, A., Yoshida, K., Kosugi, S., Natsume, S., Mitsuoka, C., Uemura, A., Utsushi, H., Tamiru, M., Takuno, S., et al. (2013b). QTL-seq: rapid mapping of quantitative trait loci in rice by whole genome resequencing of DNA from two bulked populations. *Plant J.* **74**:174–183.
- Tang, H., Bowers, J.E., Wang, X., Ming, R., Alam, M., and Paterson, A.H. (2008). Synteny and collinearity in plant genomes. *Science* **320**:486–488.
- Van der Auwera, G.A., Carneiro, M.O., Hartl, C., Poplin, R., Angel, G.D., Levy-Moonshine, A., Jordan, T., Shakir, K., Roazen, D., Thibault, J., et al. (2013). From FastQ data to high confidence variant calls: the genome analysis toolkit best practices pipeline. *Curr. Protoc. Bioinformatics* **43**, 11.10.1–11.10.33.
- Vasimuddin, Md., Misra, S., Li, H., and Aluru, S. (2019). Efficient architecture-aware acceleration of BWA-MEM for multicore systems. In 2019 IEEE International Parallel and Distributed Processing Symposium (IPDPS) (IEEE), pp. 314–324.
- Walkowiak, S., Gao, L., Monat, M., Haberer, G., Kassa, M.T., Brinton, J., Ramirez-Gonzalez, R.H., Kolodziej, M.C., Delorean, E., Thambugala, D., et al. (2020). Multiple wheat genomes reveal global variation in modern breeding. *Nature* **588**:277–283.
- Wang, J.H., Xu, Z.-M., Qiu, X.-B., Li, L.-L., Yu, S.-Y., Li, T., Tang, Y.-Y., Pu, X., Zhang, J.-Y., Zhang, H.-L., Liang, J.-J., et al. (2021). Genetic and molecular characterization of determinant of six-rowed spike of barley carrying vrs1.a4. *Theor. Appl. Genet.* **134**:3225–3236.
- Wang, K., Li, M., and Hakonarson, H. (2010). ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res.* **38**:e164.
- Wang, W., Wang, Z., Li, X., et al. (2020). SnpHub: an easy-to-set-up web server framework for exploring large-scale genomic variation data in the post-genomic era with applications in wheat. *Gigascience* **9**:giaa060.
- Wicker, T., Gundlach, H., Spannagl, M., Uauy, C., Borrill, P., Ramirez-Gonzalez, R.H., De Oliveira, R., Mayer, K.F.X., Paux, E., and Choulet, F.; International Wheat Genome Sequencing Consortium (2018). Impact of transposable elements on genome structure and evolution in bread wheat. *Genome Biol.* **19**:103.

- Wu, Q., Zhao, F., Chen, Y., Zhang, P., Zhang, H., Guo, G., Xie, J., Dong, L., Lu, P., Li, M., et al.** (2021). Bulk segregant CGT-Seq-facilitated map-based cloning of a powdery mildew resistance gene originating from wild emmer wheat (*Triticum dicoccoides*). *Plant Biotechnol. J.* **19**:1288–1290.
- Yates, A.D., Achuthan, P., Akanni, W., Allen, J., Allen, J., Alvarez-Jarreta, J., Amode, M.R., Armean, I.M., Azov, A.G., Bennett, R., et al.** (2020). Ensembl 2020. *Nucleic Acids Res.* **48**:D682–D688.
- Yu, D., Dong, L., Yan, F., Mu, H., Tang, B., Yang, X., Zeng, T., Zhou, Q., Gao, F., Wang, Z., et al.** (2019). eGPS 1.0: comprehensive software for multi-omic and evolutionary analyses. *Natl. Sci. Rev.* **6**:867–869.
- Yu, G., Smith, D.K., Zhu, H., Guan, Y., and Lam, T.T.-Y.** (2017). ggtree: an R package for visualization and annotation of phylogenetic trees with their covariates and other associated data. *Methods Ecol. Evol.* **8**:28–36.
- Zhang, L., Dong, C., Chen, Z., Gui, L., Chen, C., Li, D., Xie, Z., Zhang, Q., Zhang, X., Xia, C., et al.** (2021). WheatGmap: a comprehensive platform for wheat gene mapping and genomic studies. *Mol. Plant* **14**:187–190.
- Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., Li, W., et al.** (2008). Model-based analysis of ChIP-seq (MACS). *Genome Biol.* **9**:R137.
- Zhou, Y., Zhao, X., Li, Y., Xu, J., Bi, A., Kang, L., Xu, D., Chen, H., Wang, Y., Wang, Y., et al.** (2020). Triticum population sequencing provides insights into wheat adaptation. *Nat. Genet.* **52**:1412–1422.
- Zou, C., Wang, P., and Xu, Y.** (2016). Bulk sample analysis in genetics, genomics and crop improvement. *Plant Biotechnol. J.* **14**:1941–1955.
- Zou, S., Wang, H., Li, Y., Kong, Z., and Tang, D.** (2018). The NB-LRR gene Pm60 confers powdery mildew resistance in wheat. *New Phytol.* **218**:298–309.