Genome Biology

RESEARCH

Open Access

Genome assembly of the maize B chromosome provides insight into its epigenetic characteristics and effects on the host genome

Qian Liu[1†], Yang Liu[1†], Congyang Yi[1,2†], Zhi Gao[3], Zeyan Zhang[1,2], Congle Zhu[1,2], James A. Birchler[3*] and Fangpu Han[1,2*]

†Qian Liu, Yang Liu and Congyang Yi contributed equally to this work.

*Correspondence: BirchlerJ@missouri.edu; fphan@genetics.ac.cn

[1] State Key Laboratory of Seed Innovation, Institute of Genetics and Developmental Biology, Chinese Academy of Sciences, Beijing 100101, China
[2] University of Chinese Academy of Sciences, Beijing 100049, China
[3] Division of Biological Science, University of Missouri-Columbia, 311 Tucker Hall, Columbia, MO 65211-7400, USA

Abstract

Background: B chromosomes contribute to the genetic variation in numerous eukaryotes. Yet their genetic and epigenetic characteristics, as well as their effects on the host genome remain poorly understood.

Results: Here, we present a comprehensive genome assembly of diploid maize B73 with two copies of B chromosomes using long-read sequencing. We annotate a total of 1124 high-confidence protein-coding genes and 119,579,190 bp repeat elements representing 88.55% of the B chromosome assembly. Using CENH3 ChIP-seq data, we accurately determined the position of the B chromosome centromere, which features a unique monomer-composed satellite array distinct from that found on the chromosome arms. Our research provides detailed genetic and epigenetic maps of the B chromosome, shedding light on its molecular landscape, including DNA sequence composition, DNA methylation patterns, histone modifications, and R-loop distributions across various chromatin regions. Consistent with the cytological morphology of the B chromosome, the less condensed euchromatin regions displayed high levels of H3K4me3, H3K9ac, gene expression, and dense R-loop distributions. DNA methylation on the B chromosome was primarily observed at CG sites. The centromeric region is notably enriched with H3K4me3 and H3K9ac histone modifications and has lower CHG methylation compared to the pericentromeric regions. Moreover, our findings reveal that B chromosome accumulation affects R-loop formation on A chromosomes, and exerts tissue-specific influences on A chromosome gene expression.

Conclusions: The accurate assembly and detailed epigenetic maps of the maize B chromosome will help understand the drive mechanism, reveal its conflict with the host genome, and accelerate the construction of artificial chromosomes.

Keywords: B chromosome, High-quality genome, Centromere, Epigenetic characteristics, R-loop

Background

B chromosomes (Bs), also known as supernumerary chromosomes beyond the basic A chromosome set, are found in numerous individuals across phyla, spanning at least 2951 species [1–3]. Several hypotheses have been proposed for the origin of B chromosomes, though it is widely accepted that Bs arise from standard A chromosomes [4, 5]. Additionally, there are cases where Bs may occasionally form spontaneously in response to new genomic conditions, such as those arising after interspecific hybridization [5] or implicating sex chromosomes for some species [6, 7]. Given this, the de novo formation of Bs is likely a rare event, which likely rationalizes their presence in only some species or inbred lines [8]. Bs are nonessential and selfish constituents of the genome, rarely pairing with A chromosomes in most species. However, recent studies have uncovered genetic activities of Bs, such as improving the recombination frequency of A chromosomes [9, 10], reducing the fitness of *Z. tritici* during host infection [11], and affecting female sex determination in cichlid fish [6]. B chromosomes are often preferentially inherited, following non-Mendelian inheritance, collectively constituting the "chromosome drive". This process involves distinct mechanisms and evolutionary pathways across species, which can occur at pre-meiotic, meiotic, or post-meiotic stages, contributing to variation in B chromosome numbers within species [12–14]. For example, in maize, during the second pollen mitosis, the sister chromatids of the B chromosome undergo nondisjunction at the centromeric region [15, 16], and the sperm with B chromosomes preferentially fertilizes the egg during the double fertilization process [17–20]. In *Nasonia vitripennis*, the B chromosome drive occurs before the first mitotic division of the zygote, manipulating the sex ratio [21]. In *Aegilops speltoides*, Bs exhibit nondisjunction during the first pollen mitosis with root-specific elimination [22–24]. Notably, recent work by Chen et al. utilizing a newly assembled rye B chromosome genome has identified key candidate genes involved in B chromosome drive, including DCR28, a microtubule-associated protein, which is highly expressed during the first pollen mitosis and was also detected in *Aegilops speltoides* [25]. Structural polymorphisms of Bs also have been observed in several species, including *Ae. speltoides* [26], *Brachycome dichromosomatica* [27], and *Scilla autumnalis* [28]. However, in species like maize and rye, B chromosomes maintain a similar molecular and cytological structure at the subspecies level [14].

Maize B chromosomes harbor hundreds of genes and transposable elements (TEs), some of which are shared with the A chromosomes [29]. When their copy numbers are low, the impact of the B chromosome on plant growth and development is not detectable [30]. Therefore, B chromosomes serve as an ideal platform for the insertion and maintenance of exogenous genes, facilitating the construction of artificial mini-chromosomes while minimizing interference with the host genome. Although maize Bs are dispensable, a small portion of genes on Bs are actively transcribed. They can also affect the expression of genes located on the A chromosomes, resulting in noticeable phenotypic effects at high copy numbers, such as reduced plant height and decreased fertility [31]. However, the various epigenetic and transcriptional changes in the host genome caused by B chromosomes remain poorly understood, largely due to the challenges of assembling B chromosome sequences, which possess complex structures and substantial homology with A chromosomes.

The first sequence of the maize B chromosome was completed in 2021 using short-read sequence data , resulting in hundreds of contigs and scaffolds, with the short arm remaining unassembled . Here, we integrated PacBio high-fidelity (HiFi) long-read sequencing and high-throughput chromatin conformation capture (Hi-C) data to assemble maize inbred line B73 containing two copies of the B chromosome (B73 + 2Bs-v2.0). With longer and more accurate sequencing reads, it is now feasible to accurately assemble the B chromosome, including the centromere, the telomeres, and ribosomal DNA repeats. To investigate the genetic and epigenetic characteristics of the maize B chromosome and its effects on the host genome, we conducted a comprehensive analysis utilizing the new B chromosome assembly (ChrB-v2.0), including gene expression levels, histone modification patterns, DNA methylation profiles, and R-loop distributions on the B chromosome, as well as comparisons of transcriptome and R-loop formation in B73 with varying B chromosome copy numbers. Our findings provide valuable genomic resources and reveal novel features of the maize B chromosome.

Results

High-quality genome assembly and annotation of maize inbred line B73 with B chromosomes

Using fluorescence in situ hybridization (FISH) with a B chromosome-specific probe, we screened for plants of the maize inbred line B73 with two copies of the B chromosome (2n = 20 + 2B). This material was subsequently sequenced and assembled. We generated 156.8 Gb (~ 65 × coverage) of PacBio HiFi reads data and 295.7 Gb (~ 123 × coverage) of high-throughput Hi-C data using the PacBio Revio platform and Illumina HiSeq platform, respectively (Additional file 2: Table S1). HiFiasm [32] was used for the primary genome assembly based on HiFi reads, resulting in a draft genome of 2387.3 Mb, containing 1537 contig sequences with an N50 value of 56.40 Mb (Table 1). With the aid of Hi-C reads, the assembled HiFi contigs were further used to refine the ordering and orientation of each cluster to achieve a chromosome-level assembly (Fig. 1a, b; Table S1). The final diploid reference genome of B73 with two B chromosomes comprised 11 chromosomes with a total length of 2288.0 Mb and a scaffold N50 length of 229.22 Mb. Notably, chromosome 10 was devoid of gaps (Table 1; Fig. 1a, b).

Table 1 Comparison of the B73 + 2Bs-v2.0 assembly with the previous genome version

Genomic feature	B73 + 2Bs-v1.0 [15]	B73 + 2Bs-v2.0
Sequencing technology	Hi-C + Bionano + Illumina	PacBio HiFi + Hi-C
Total length (Mb)	2369.4	2387.3
Contig number	NA	1537
Contigs N50 (bp)	NA	56,398,518
Scaffold number	66,269	1494
Scaffolds N50 (bp)	23,788,118	229,222,561
Assigned length (Mb)	NA	2288.0
Gap	57,589	94
Ns (bp)	62,121,592	17,204
BUSCO completeness of annotation (%)	NA	98.8
LAT score	NA	22.99

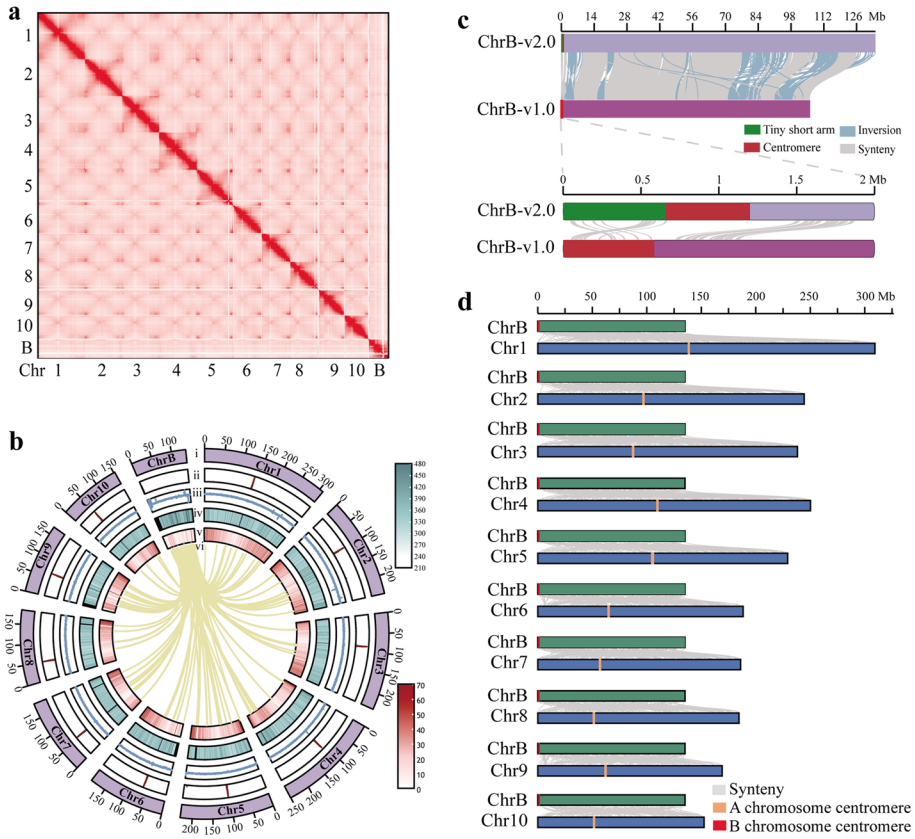

Fig. 1 Overview of the B73 with two B chromosomes reference genome. **a** Hi-C interactions between the ten standard A chromosomes (Chr1-10) and the B chromosome. **b** Circular diagram of major genomic characteristics. Circular tracks from outside to inside indicate (i) chromosome; (ii) CENH3; (iii) GC skew; (iv) repeat density; (v) non-TE gene density; and (vi) syntenic non-TE gene pair identified between the B chromosome and A chromosomes. **c** Alignment of the new B chromosome assembly (ChrB-v2.0) with ChrB-v1.0. The upper panel shows the alignment of the entire chromosome, while the bottom panel provides an enlarged view of the region spanning 0 to 2 Mb. Collinear regions are depicted by gray, and inversion regions are depicted by blue lines, respectively. The tiny short arm and centromere are highlighted in green and red, respectively. **d** Sequence alignments between the A chromosomes and the B chromosome from the B73 + 2Bs-v2.0 assembly. ChrB represents the B chromosome, and Chr1-10 represent the A chromosomes from the B73 + 2Bs-v2.0 assembly. Collinear regions are linked by gray lines. The centromeric regions of the A and B chromosomes are highlighted with orange and red blocks, respectively

The quality of the assembly was assessed using several independent methods. First, genome-wide Hi-C interaction mapping showed a high consistency across all chromosomes and conformed to inter- and intra-chromosome interaction requirements (Fig. 1a). Second, genome completeness was evaluated using benchmarking universal single-copy orthologs (BUSCO), revealing a 98.8% complete score (Additional file 2: Table S2), which is comparable to that of the potato [33] and oat [34] genomes. Third, the high fidelity of the final assembly was also supported by the high mapping rate of 97.2% for RNA-seq reads (Additional file 2: Table S3). Fourth, the annotation of long terminal repeats (LTR) revealed an LTR assembly Index (LAI) score of 22.99, meeting the gold standard for reference genomes, and supporting a high level of genetic continuity (Table 1). Finally, the new B73 with two B chromosomes assembly (B73 + 2Bs-v2.0) exhibited high sequence concordance with the previously assembled B73_v5 [35] and

ChrB-v1.0 genomes , confirming the overall accuracy of the assembly (Fig. 1c, d; Additional file 1: Fig. S1). Taken together, these assessments serve as evidence of the high assembly quality of B73 with 2Bs genome assembly.

Based on a combination of de novo prediction, gene homology-based predictions (GMAP and Gemoma), and RNA-sequencing (RNA-seq) transcript prediction [36, 37], we predicted a total of 42,607 high-confidence protein-coding genes in B73 + 2Bs-v2.0 reference genome (Additional file 2: Table S4). Next, repetitive sequences within the genome were annotated through both ab initio and homolog-based search methods. A total of 2,031,095,324 (85.08%) of repetitive sequences were identified, primarily comprising LTR retrotransposons (73.23%) and DNA transposons (7.13%). Among the LTR retrotransposons, the Gypsy and Copia superfamily accounted for 43.29% and 23.86% of the genome, respectively (Additional file 2: Table S5).

The genetic and epigenetic maps of maize B chromosome

Using this whole-genome assembly as a basis, the B chromosome assembly, named ChrB-v2.0, is nearly complete, spanning 135,035,522 bp. It shows 28.4 Mb of previously uncharacterized regions, including the tiny short arm, which is difficult to observe cytologically, and 27.7 Mb of sequence within the heterochromatic region of the long arm (Fig. 1c, Table 2). Direct sequence comparison between ChrB-v2.0 and ChrB-v1.0 reveals an average sequence identity of 98.39% in the aligned regions (Additional file 2: Table S6). However, there are multiple genomic structural differences, such as inversions and deletions (Fig. 1c; Additional file 2: Table S6). Furthermore, the accuracy of the B chromosome assembly was validated using whole-genome re-sequencing data from various maize B chromosome variants with different B chromosome segments resulting from B-A translocations, B centromere misdivisions, and chromosome breakage events, as reported by Blavet et al. [4] (Additional file 1: Fig. S2). These data correlate cytological

Table 2 Comparison of the ChrB-v2.0 assembly with ChrB-v1.0

	Genomic feature	ChrB-v1.0 [15]	ChrB-v2.0	%Δ
Genome Assembly	Total length (Mb)	125.9	135.0	+ 6.6
	Assigned length (bp)	106,643,106	135,035,522	+ 26.6
	Unlocalized bases (bp)	8,593,788	835,366	− 90.2
	Contig number	NA	18	NA
	Contig N50 (bp)	NA	17,251,596	NA
	Scaffold number	1	1	0
	No. of gaps (bp)	6286	17	− 99.7
	No. of N bases (bp)	9,068,090	3400	− 99.9
	CG content (%)	43.82	44.72	+ 0.21
Annotation	Protein-coding genes	758	1,124	+ 32.5
Repetitive bases	LINE (bp)	27,883	376,510	+ 1142.7
	Gypsy (bp)	39,881,049	73,663,725	+ 84.7
	Copia (bp)	13,875,615	24,966,954	+ 80.0
	Transposon (bp)	542,062	9,446,098	+ 1646
	All repeat classes (kb)	72,322.2	131,935.0	+ 39.4
	TE percentage (%)	57.4	88.55	+ 54.3

%Δ is the percentage change from ChrB-v1.0 to ChrB-v2.0

features of the B chromosome with the sequence. Initially, we annotated 3252 genes. After filtering out low-confidence genes and TE protein-encoding genes, we identified 1124 protein-coding genes, of which 366 were newly anchored (Table 2; Additional file 2: Table S7). The newly assembled sequences significantly increased the percentage of repeats on the B chromosome from 57.4% in ChrB-v1.0 to 88.55% in ChrB-v2.0, corresponding to 72.3 Mb and 131.9 Mb, respectively (Table 2). This high-quality genome assembly of the B chromosome represents a significant improvement over the ChrB-v1.0 reference genome.

As illustrated by the pachytene B chromosome morphology, the short arm and centromere are diminutive (Fig. 2a). The long arm of the B chromosome consists of loosely packed proximal and distal euchromatin (euchromatin 1 and 2), as well as condensed proximal and distal heterochromatin (heterochromatin 1 and 2) (Fig. 2a). Repetitive sequences on the B chromosome comprise 73.52% retrotransposons, 9.62% DNA transposons, and 5.43% satellites, with 2522 full-length LTR-RTs identified (Fig. 2b; Additional file 2: Table S8). Within the short arm, a significant amount of repeats (75%) consist of the B chromosome-specific repeat pZmBs (Fig. 2a; Additional file 2: Table S9). Heterochromatin 1 mainly consists of pZmBs, knob 180 repeat units, and Gypsy elements. Heterochromatin 2, as well as euchromatin 1 and 2 regions, exhibit distinct

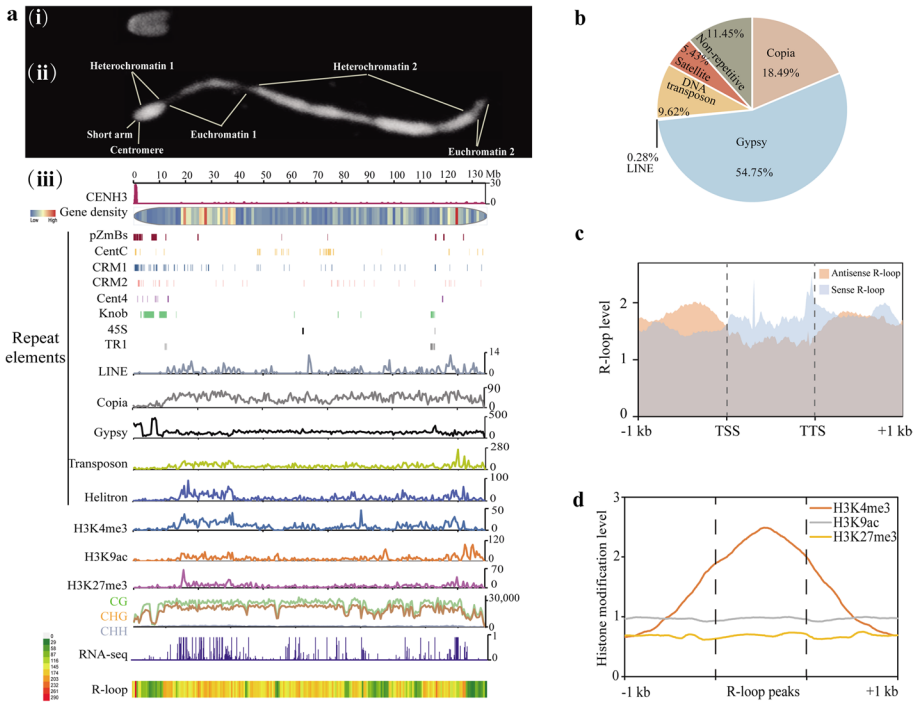

Fig. 2 The genetic and epigenetic maps of the maize B chromosome. **a** Structure and epigenetic characteristics of the maize B chromosome. From top to bottom, mitotic metaphase (i) and pachytene-stage chromosomes (ii); CENH3 enrichment [$\log_2$(ChIP/input)]; density of non-TE protein-coding genes; distribution of pZmBs, CentC, CRM1, CRM2, Cent4, Knob, 45S, and TR1; density of LINE, copia, gypsy, transposon, and helitron; histone modification (H3K4me3, H3K9ac, and H3K27me3); DNA methylation (CG, CHG, and CHH); RNA-seq data; density of R-loop peaks in 500 kb window (iii). **b** Proportion of repetitive sequences in the B chromosome. **c** Metaplots showing sense and antisense R-loop signals on annotated genes of the B chromosome with 1 kb upstream and downstream extensions. **d** Metaplots of H3K4me3, H3K9ac, and H3K27me3 ChIP-seq reads across R-loop peaks on the B chromosome and the up- and downstream regions

repeat compositions. Gypsy elements are predominantly found in the short arm, centromere, and heterochromatin 2 regions. Meanwhile, LINE, copia, DNA transposons, and helitrons are more abundant in the euchromatin 1 and 2 regions (Fig. 2a). The open structure of euchromatin likely facilitates the insertion and accumulation of DNA transposons, contributing to this distinctive distribution. Due to the distribution of pZmBs on the B chromosome, the GC skew peaks co-localize with these repeats, resulting in distinct GC skew patterns that differentiate the B chromosome from A chromosomes (Fig. 1b; Additional file 1: Fig. S3). Genes are unevenly distributed across the B chromosome, with higher densities concentrated in the euchromatin 1 and 2 regions (Fig. 2a). For the identification of significantly expressed genes, we combined the RNA-seq data of B73 with 2Bs from Shi et al. [30] and Huang et al. [38] with RNA-seq data from leaves and spikelets in our study. In total, there are 337 genes with expression levels (FPKM > 1), among which 93 genes are highly expressed (FPKM > 15) (Additional file 2: Table S10). These highly expressed genes are involved in several cellular processes, including but not limited to "structural constituent of chromatin," "microtubule-based process," "pollen germination," and "intracellular protein transport" (Additional file 2: Table S10). The syntenic analyses between the B and A chromosomes indicate that the B chromosome sequence consists of DNA segments contributed by ten A chromosomes, reflecting the complexity and diversity of its composition (Fig. 1b, d; Additional file 1: Fig. S4; Additional file 2: Table S11). Among the B chromosome genes with homologs in the A chromosomes, 58 genes are expressed on the B chromosome (FPKM > 1), while 64 are expressed on the A chromosomes. Of these, 27 genes are expressed on both the A and B chromosomes (Additional file 2: Table S11).

To investigate the epigenetic landscape of the B chromosome, we employed a combination of DNA methylation (whole genome bisulfite sequencing), histone modification data (H3K4me3, H3K9ac, and H3K27me3), and R-loop data (single-strand DNA ligation-based library construction from DNA: RNA hybrid immunoprecipitation, followed by sequencing, ssDRIP-seq) from B73 with 2Bs (Fig. 2a; Additional file 2: Table S3). In the short chromosome arm, centromere, and heterochromatin 1, we observed the lowest levels of histone modification and gene expression, attributable to low gene density in these regions (Fig. 2a). Conversely, both euchromatic regions 1 and 2 of the B chromosome showed higher level of gene expression, and elevated H3K4me3 and H3K9ac signals. Additionally, we observed a predominant enrichment of CG methylation on the B chromosome (Fig. 2a). Compared to the chromosome arms, the centromeric and pericentromeric regions exhibit lower levels of DNA methylation (Fig. 2a). This is in contrast to the high levels of CG and CHG-context methylation typically found in the pericentromeric regions of A chromosomes [39]. We hypothesize that the high methylation levels prevent recombination events in the centromeric regions of normal chromosomes [40]. In contrast, the lower methylation levels observed near the B chromosome centromere might contribute to the unique inheritance behavior of B chromosomes.

R-loops constitute a three-stranded nucleic acid structure with a DNA: RNA hybrid and a single-stranded DNA, playing crucial roles in various cellular processes [41, 42]. We identified a high density of R-loops on the B chromosome, particularly in euchromatin regions, as well as within the centromeric regions (Fig. 2a). Furthermore, an analysis of the sense and antisense R-loop locations revealed that sense R-loop tended to form

within gene bodies, whereas antisense R-loops predominantly localize at promoters on the B chromosome (Fig. 2c). Additionally, H3K4me3 modifications was enriched in the R-loop peaks of the B chromosome (Fig. 2d). These results reveal the complex regulatory landscape of the B chromosome and its unique epigenetic features.

%Δ is the percentage change from ChrB-v1.0 to ChrB-v2.0

Sequence resolution and structure of the B chromosome centromere

The functional region of the centromere is defined by the specific histone H3 variant, CENH3 [43]. To produce a comprehensive map of the B chromosome centromere, we conducted chromatin immunoprecipitation sequencing (ChIP-Seq) using an antibody against maize CENH3. The CENH3 ChIP-seq data were aligned to the B73 + 2Bs-v2.0 assembly. Significant enrichment of CENH3 peaks was observed within the centromeric region, spanning 660,000 to 1,200,000 base pairs (Fig. 3a, b). In this centromeric region, we identified 17 intact long terminal repeat retrotransposons (LTR-RTs) and no genes (Additional file 2: Table S6; Additional file 2: Table S7).

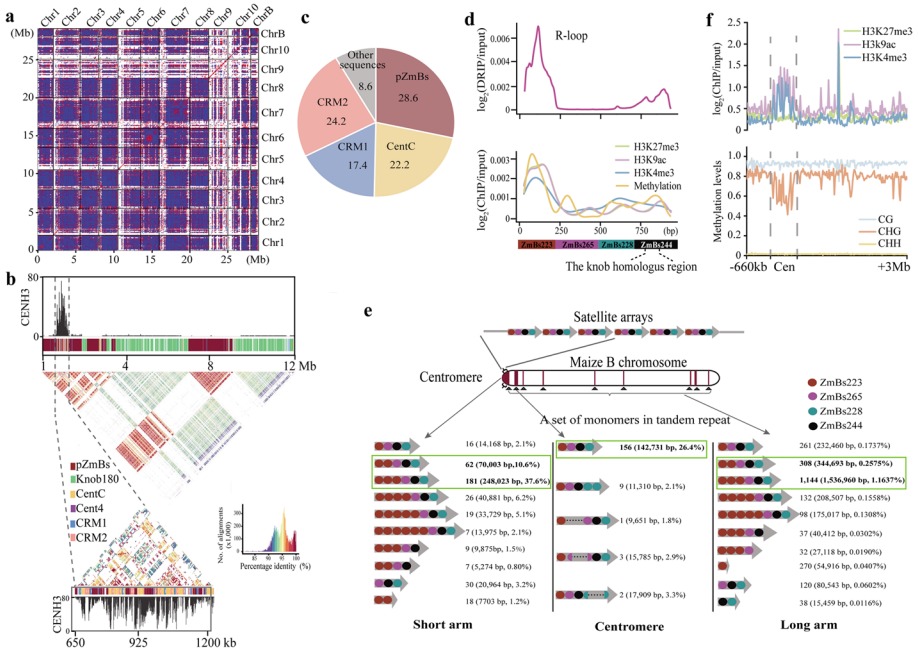

Fig. 3 The structure of maize B chromosome centromere. **a** Dot plots comparing centromeres of A and B chromosomes using a search window of 156 bp. Red and blue indicate forward- and reverse-strand similarity, respectively. **b** Characteristics of the B chromosome centromere. Line 1: CENH3 $\log_2$(ChIP/input) (black) plotted over 0–12 Mb. Line 2: Distribution of various retrotransposons and satellites. Line 3: Heatmap illustrating the pattern of sequence identity. The enlarged view depicts centromere regions of the B chromosome. **c** Composition and proportion of repeat sequences in the B chromosome centromere. **d** R-loop, histone modification, and DNA methylation profiles at four regions of the pZmBs. **e** Organization of the top ten most abundant types of pZmBs on the short arm, centromere, and long arm of the B chromosome. Red vertical lines indicate the distribution of pZmBs on the B chromosome. The four monomers, ZmBs223, ZmBs228, ZmBs244, and ZmBs265, are represented by red, green, dark cyan, and purple dots, respectively. The right side shows the copy number, length (bp), and proportion for each variant. **f** Histone modification ChIP-seq signals (H3K27me3, green; H3K9ac, purple; and H3K4me3, blue) and methylation levels (CG, sky blue; CHG, orange; CHH, yellow) at or in proximity to B chromosome centromere (extending 660 kb to the left and 3 Mb to the right), with a bin size of 20 kb

Furthermore, the centromeres of both A and B chromosomes demonstrate comparable sequence and structural characteristics, except for Cen9, which exhibits lower sequence similarity due to centromere repositioning to a region lacking CentC elements (Fig. 3a; Additional file 1: Fig. S5). Similar to A chromosome centromeres, the functional maize B chromosome centromere is composed of repetitive elements, including CRM1 (17.4%), and CRM2 (24.2%) retrotransposons and CentC satellite sequences (22.2%) (Additional file 2: Table S12). Additionally, the B chromosome centromere also uniquely contains the B chromosome-specific satellite pZmBs (28.6%), which are localized in the region with CENH3 peaks (Fig. 3b, c). Unlike in Arabidopsis, where CENH3 tends to bind to highly conserved repeat sequences [40], CENH3 binding sequences on the maize B chromosome are more variable compared to the flanking regions (Fig. 3b).

When 1001-bp pZmBs sequences previously reported [44] were mapped onto our assembled genome, we observed significant variation in the length of pZmBs fragments, predominantly around ~ 300 bp and ~ 550 bp (Additional file 1: Fig. S6), indicating the variability of pZmBs. To elucidate the sequence characteristics and structure of the B chromosome centromere, we utilized TRASH [45] and RepeatExplorer [46] for tandem repeat annotation. We identified three satellites in the centromeric region, with lengths of 223 bp, 228 bp, and 524 bp. Among these, the 524 bp repeat includes the 223 bp repeat (Additional file 1: Fig. S7). Combining previously reported pZmBs sequences [44], we manually corrected the pZmBs arrays to consist of four monomers nested in tandem: ZmBs223, ZmBs265, ZmBs244, and ZmBs228, named according to their respective lengths, and arranged sequentially (Fig. 3d, e; Additional file 2: Table S9, S13, S14; Additional file 3). Phylogenetic analysis of the four monomers showed that they could be subdivided into four distinct branches, indicating they are different sequences rather than variants of the same repeat (Additional file 1: Fig. S8). Within the ZmBs244 sequence, nucleotides 43–133 bp exhibited homology to the knob 180 tandem repeat, spanning a 90-bp stretch (Additional file 1: Fig. S9), consistent with previous reports [44]. pZmBs arrays in the centromere have 91–94% identity and a major length of approximately 960 bp (Additional file 1: Fig. S8–10; Additional file 3). Among the four sections of pZmBs, ZmBs223 has higher levels of R-loops, H3K4me3, H3K9ac, H3K27me3 modifications, and DNA methylation (Fig. 3d). Knob 180, compared to the knob homologous part of the pZmBs, exhibits lower levels of H3K4me3, H3K27me3, and DNA methylation (Additional file 1: Fig. S11). This distinctive epigenetic landscape likely reflects different functional roles and regulatory mechanisms between knob 180 and the four sections of pZmBs.

The centromere of the B chromosome assembly consists of a 154,676-bp (28.6%) pZmBs satellite array, flanked by pZmBs blocks on the short arm (513,463 bp) and long arm (3,233,306 bp) (Fig. 3b, e). The B chromosome centromeric pZmBs array is primarily composed of four monomers arranged head-to-tail and organized into high-order repeats (HORs) (Fig. 3e). In contrast, the short and long chromosome arms exhibit a greater variety of pZmBs HORs, displaying diverse monomer arrangements (Fig. 3e; Additional file 1: Fig. S12, 13). The most prevalent arrays consist of either five or six monomers, with Zm223 often having one or two additional copies (Fig. 3e). Other arrays include various combinations of monomers, DNA transposon insertions, and monomer deletions or duplications (Fig. 3e; Additional file 1: Fig. S12, 13). In the distal part of

the long arm, two regions with abundant pZmBs distribution were identified (Additional file 1: Fig. S14). The organization of pZmBs around the centromere is similar to that in region 1, while region 2 primarily consists of single 180 bp ZmBs223 elements and fragmented Cent4 sequences (Additional file 1: Fig. S14).

Both the centromeric and pericentromeric regions of the B chromosome contain large amounts of pZmBs. To determine whether centromeres have distinct epigenetic marks compared to the surrounding pericentromeric heterochromatin, we analyzed DNA methylation, histone modifications, and R-loop levels at the centromere, as well as in the pericentromeric regions of the short arm, and long arm of the B chromosome. Our results indicate that CG and CHH methylation levels are comparable between centromeric and pericentromeric regions. However, CHG methylation levels are significantly lower in the centromeric regions (Fig. 3f). Additionally, the centromeric region is notably enriched with H3K4me3 and H3K9ac histone modifications (Fig. 3f).

The influence of B chromosomes on A chromosome gene expression is tissue-specific

In maize, the effects of B chromosomes on the host organism vary significantly depending on their quantity [47]. At low copy numbers, B chromosomes exert subtle influences on the host's metabolism without causing noticeable phenotype changes. However, with five or more B chromosomes, notable phenotypic alterations, such as white longitudinal striping and reduced fertility, are observed (Fig. 4a). To investigate the impact of maize B chromosomes on the transcriptome of A chromosomes, we collected young leaves and tassels from B73 inbred lines with 0B, 2Bs, and 7Bs (displaying the white striping phenotype) for RNA sequencing. High-quality RNA-seq reads obtained from these samples were aligned to our reference genome assembly, achieving over 90% alignment rate for each sample (Additional file 2: Table S3). Comparing transcriptomes at the vegetative stage revealed significant differential expression of genes (DEG) located on

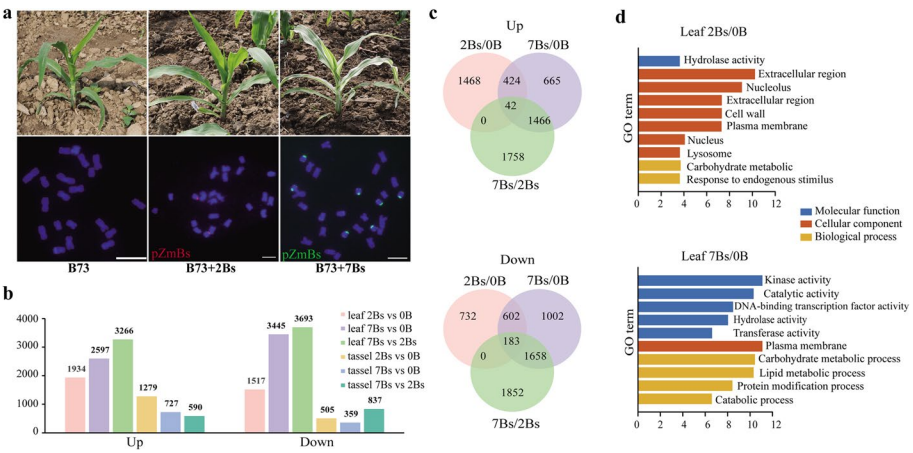

Fig. 4 Influence of different copy numbers of the B chromosome on gene expression in the A chromosomes. **a** Morphological and cytological characteristics of B73 with 0B, 2Bs, and 7Bs. B chromosomes were identified by FISH using the B-specific repeat pZmBs. **b** Number of DEGs located on the A chromosomes. **c** Venn diagram showing the overlap among DEGs that are upregulated (top) and downregulated (bottom) in the leaves of B73 with 0B, 2Bs, and 7Bs. **d** GO functional enrichment analysis for DEGs in leaves of B73 with 0B, 2Bs, and 7Bs

A chromosomes. In young leaves, we found 1934 upregulated and 1517 downregulated DEGs with 2Bs, and 2597 upregulated and 3445 downregulated DEGs with 7Bs (Fig. 4b; Additional file 1: Fig. S15). In tassels, we observed 1279 upregulated and 505 downregulated A-located DEGs with 2Bs, and 727 upregulated and 359 downregulated A chromosome-located DEGs with 7Bs (Fig. 4b; Additional file 1: Fig. S16).

It is important to note that the B73 genetic background and its associated white striping phenotype may contribute to some of the observed gene expression changes. In particular, the white striping phenotype may produce secondary consequences for gene expression that are not directly caused by the presence of B chromosomes. Our findings indicate that the impact of B chromosomes on A chromosome gene expression is more pronounced in young leaves than in tassels. In tassels with B chromosomes and young leaves with 2Bs, the number of upregulated A chromosome-located DEGs is higher than the number of downregulated DEGs (Fig. 4b). This suggests that during the developmental stages with minimal B chromosome influence, A chromosomes tend to exhibit upregulated gene expression. Additionally, comparisons between 7Bs/2Bs and 7Bs/0B samples show the highest number of shared DEGs (Fig. 4c; Additional file 1: Fig. S15, S16), highlighting the complexity and non-linear effects of increasing B chromosome numbers on A chromosome gene expression regulation.

GO term analysis revealed that in young leaves with 2Bs, DEGs were mainly enriched for cellular component terms, including "extracellular region," "nucleolus," "external encapsulating structure," "cell wall," "plasma membrane," "nucleus," and "lysosome" (Fig. 4d). In young leaves with 7Bs, DEGs were mainly enriched in molecular function terms such as "kinase activity," "catalytic activity," "DNA-binding transcription factor activity," and biological process terms like "carbohydrate metabolic process," "lipid metabolic process," and "protein modification process." In contrast, in tassels, DEGs did not show clear preference in enriched biological processes in plants containing either two or seven B chromosomes (Fig. 4d; Additional file 1: Fig. S16). This indicates that in young leaves, an increasing number of B chromosomes has a more complex and specific impact on gene expression, affecting different molecular functions and biological processes. It also highlights the varying impact of B chromosomes on A chromosome-located gene expression in different tissues and developmental stages.

R-loop distributions and properties on maize A chromosomes are influenced by B chromosomes

The R-loop, consisting of a DNA/RNA hybrid and a single-stranded DNA, plays crucial roles in various cellular processes such as DNA replication, gene expression, DNA methylation, and histone modification [48–50]. In maize, R-loops cover approximately 10% of the genome [41]. To investigate the dynamic response of R-loops to varying B chromosome copy numbers during development, we conducted ssDRIP-seq on leaves and tassels of B73 plants with 0B, 2Bs, and 7Bs. To ensure the specificity of the S9.6 signal for R-loops, we treated the negative control sample with RNaseH, which degrades the RNA in RNA: DNA hybrids, and used an untreated sample with S9.6 as the input (Fig. 5a; Additional file 2: Table S3). The ssDRIP-seq data from the replicate experiments demonstrated high reproducibility (Fig. 5a). To maximize R-loop identification, we merged two replicates for further analysis. Snapshots of R-loops showed both shared and distinct

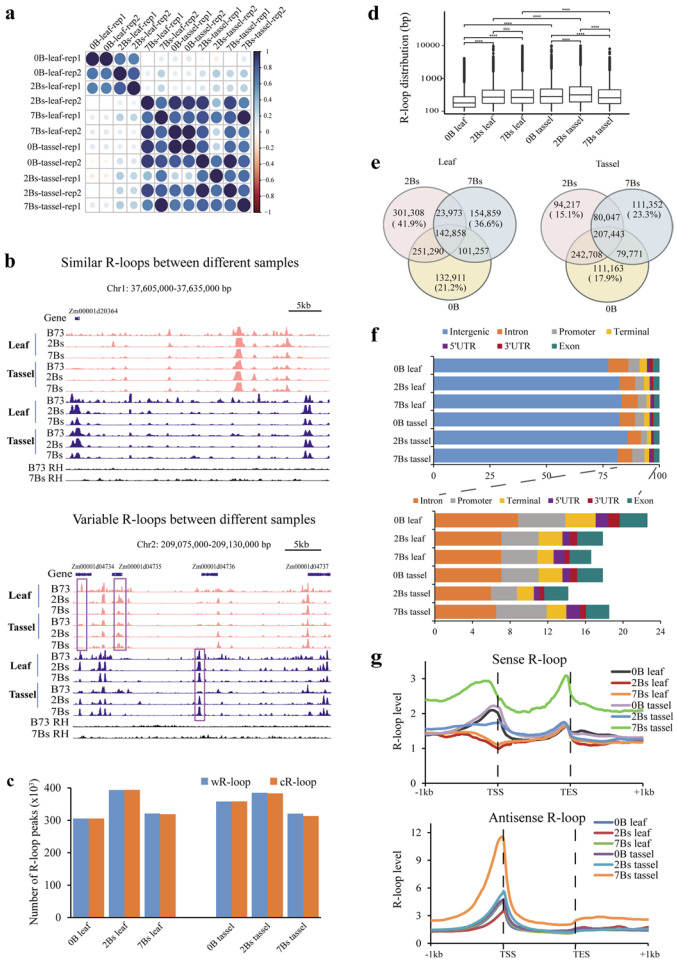

Fig. 5 B chromosomes affect the R-loop landscape on A chromosomes of maize. **a** The correlation between two biological replicates of all samples. **b** Screenshot of representative genomic regions showing similar (left) and variable (right) ssDRIP-seq signals across all samples. ssDRIP-seq data after RNase A and RNase H pre-treatment are shown below. R-loops formed in Waston strand (wR-loops) are indicated in pink, while R-loops formed in Crick strand of DNA (cR-loops) are shown in blue. **c** The numbers of R-loop peaks on the A chromosomes in the leaf (left) and tassel (right) samples of B73 with 0, 2, or 7 B chromosomes. **d** The sizes of R-loop peaks on the A chromosomes in leaf and tassel samples of B73 with 0, 2, or 7 B chromosomes. **e** Shared and unique R-loops on A chromosomes in leaf and tassel samples with varying numbers of B chromosomes. **f** Proportion of R-loop peaks in different genic regions from all samples. **g** Metaplots showing sense (left) and antisense (right) R-loop signals from all samples centered on genes of A chromosomes with 1-kb up- and downstream extensions

R-loops on the A chromosome among samples with 0B, 2Bs, and 7Bs (Fig. 5b). In both leaves and tassels, the number of R-loops on the A chromosomes decreased as the number of B chromosomes increased (Fig. 5c). The size of R-loop peaks on the A chromosomes varied slightly with different B chromosomes copy numbers, ranging from 100 to 500 bp (Fig. 5d). In leaves, we identified 301,308 (41.9%) and 154,859 (36.6%) unique R-loops on the A chromosome in samples with 2Bs and 7Bs, respectively, with about 142,858 R-loops in common to all samples. In tassels, we detected 94,217 (15.1%) and 111,352 (23.3%) unique R-loops in samples with 2Bs and 7Bs, respectively (Fig. 5e). Similar to the RNA-seq results, in leaves, the A chromosomes exhibited more different R-loops compared to tassels. Additionally, R-loops often form at recombination hotspots

(Additional file 1; Fig. S17), suggesting that B chromosomes may influence A chromosome recombination hotspots through R-loop-mediated mechanisms.

We further analyzed the average levels of R-loops on annotated genes of the A chromosomes. R-loop patterns were relatively conserved among most samples, with the greatest enrichment in genes promoter regions. Sense R-loops were mainly enriched in promoters and transcription termination sites (TTS), while antisense R-loops were mainly enriched in the transcription start sites (TSS) (Fig. 5f, g). However, some samples exhibited different levels and proportions of R-loops on genes. For instance, in tassels, the proportion of R-loops forming in promoter regions of A chromosomes decreased with the addition of two Bs but increased with seven Bs. In leaves, neither sense nor antisense R-loops showed enrichment at the TSS in samples with 2Bs and 7Bs. In tassels, B73 with 7Bs significantly increased the levels of both sense and antisense R-loops compared to other samples (Fig. 5g). These results indicate that the presence of B chromosomes can influence the distribution and levels of R-loops on A chromosomes, with a more pronounced impact in leaves.

Discussion

B chromosomes differ in morphology from A chromosomes but exhibit high sequence similarity. Their number and mechanisms vary widely among different species and individuals, adding to the complexity of their study, particularly before the advent of long-read sequencing [3]. Additionally, constructing a comprehensive epigenome of the B chromosome has been challenging due to the absence of a complete reference genome. Here, we present a de novo genome assembly for the maize B chromosome, generated using Hi-C data and deep-sequenced PacBio HiFi long reads with high accuracy. This improved and high-quality genome assembly has enabled us to thoroughly investigate all genomic regions, providing a comprehensive view of the genetic and epigenetic organization of the maize B chromosome, and revealing its unique characteristics. We observed that, consistent with the cytological morphology of the B chromosome, the less condensed euchromatin regions exhibit higher gene density, expression levels, histone modification levels, and R-loop density. Unlike A chromosomes, the centromere of the B chromosome is characterized by unique satellite sequences composed of monomers and lowers level of CHG methylation. while its CG methylation remains comparable to that of the pericentromere. This differs from the centromere of *Arabidopsis* and maize A chromosomes, which typically show high levels of CG methylation and reduced CHG methylation in regions enriched with CENH3. The similar CG methylation levels in the B chromosome centromere may support its unique behaviors, such as nondisjunction during cell division, while reduced CHG methylation could be essential for proper centromere function. Additionally, the B chromosome influences the landscape of R-loops on the A chromosomes. The accumulation of B chromosomes has a stronger impact on A chromosome gene expression in leaves than in tassels. Overall, the high-quality genome assembly and epigenetic insights into the maize B chromosome provide a valuable resource for studying their evolution and drive mechanisms. These findings contribute to our understanding of how B chromosomes influence host genomes and their potential applications in artificial chromosome construction and breeding studies.

Upon examining satellite DNA, we found that an intact pZmBs sequence consists of four monomers: ZmBs223, ZmBs265, ZmBs244, and ZmBs228. Notably, pZmBs showed 139 bp and 90 bp homology to maize Cent4 and knob 180, respectively (Additional file 1: Fig. S9, S10), suggesting a shared ancestor, as discussed [44, 51]. These monomers can be categorized into distinct classes and we identified various types of monomers in HOR across the B chromosome. ZmBs223 monomers typically vary in copy number within a monomer, ranging from 0 to 7 copies. We observed greater variation in monomer composition within chromosome arms compared to the centromere as defined by the CENH3 association. In the centromere, the predominant type consists of an array composed of four monomers, whereas in chromosome arms, arrays composed of six monomers are more prevalent. We hypothesize that the specific binding of centromeric proteins is optimized for the array configuration of four monomers in the centromere, ensuring the proper formation and function of kinetochores. Additionally, the unique satellite DNA array in the centromere likely undergoes specific epigenetic modifications, that contribute to its stability and functionality. Furthermore, the presence of maize A chromosome centromere-specific repetitive sequences, such as CentC and CRM, within B chromosome centromeres adds to the sequence diversity compared to pericentromeric regions. The structural differences between A and B chromosome centromeres, including these distinct compositions of repetitive sequences, likely contribute to the observed variations. Despite the advantages of PacBio HiFi sequencing, such as long read lengths and high accuracy, regions with high sequence similarity between the A and B chromosomes, or the insertion of DNA segments from A chromosomes into the B chromosome, could potentially affect the quality of the B chromosome assembly. In future studies, we will employ chromosome flow sorting in combination with PacBio HiFi and high-throughput Hi-C data to isolate individual B chromosomes for sequencing, which will provide a more accurate and detailed understanding of the B chromosome's structure.

Conclusions

We present a high-quality genome assembly of maize B73 with two B chromosomes, providing a comprehensive view of the genetic and epigenetic landscape of the B chromosome. Our analysis reveals that the B chromosome centromere is characterized by a unique satellite array composed of four monomers which differ from those found on the chromosome arms and exhibit distinct epigenetic modifications on the different monomers. The centromeric region shows enrichment of H3K4me3 and H3K9ac histone modifications with lower CHG methylation compared to the pericentromeric regions. Additionally, we demonstrate that variations in the copy number of the B chromosome influence gene expression and R-loop formation on A chromosomes in a tissue-specific manner. This effect is more pronounced in leaves and is primarily determined by the presence of the B chromosome rather than its copy number. Collectively, the findings improve our understanding of the unique genetic and epigenetic features of the maize B chromosome and its conflict with the host genome.

Methods

Plant materials

To screen for maize (*Zea mays*) inbred line B73 with B chromosomes (2n = 20 + 2Bs), we performed FISH using the B-specific element pZmBs on the root tips of their seeds. B73 plants with 0B, 2Bs, and 7Bs were selected and cultivated in a greenhouse under controlled conditions, with a temperature of 30 °C for 16 h in light and 25 °C for 8 h in the darkness. After 14 days, the plants were transferred to the field.

Library preparation and genome sequencing

For genome sequencing and assembly, young leaves from 14-day-old B73 plants with two B chromosomes were collected, washed with sterile water, immediately frozen in liquid nitrogen, and stored at − 80 °C. Genomic DNA extraction was performed using the cetyltrimethylammonium bromide (CTAB) method, followed by purification using the Qiagen genomic kit (Qiagen, 13,343). The quality and quantity of the DNA were assessed using a NanoDrop spectrophotometer and Qubit fluorometer, respectively.

The extracted high-molecular-weight DNA was utilized to construct both PacBio HiFi sequencing libraries and Hi-C sequencing libraries. The Hi-C and PacBio sequencing libraries were run on the Nanopore PromethION sequencer and PacBio SequeII platform, respectively. These long and highly accurate HiFi and Hi-C reads were assembled using Hifiasm [32] and verkko [52] to generate two versions of the primary genome with default parameter. The Hi-C data were used to correct and validate interaction pairs using Chromap [53], and a non-redundant Hi-C contact matrix was generated using juicer (v 1.6) [54]. Additionally, a homology-based scaffolding approach was employed to ensure the consistency of the primary assembly with B73_v5 and Zm-B73_B-CHROMOSOME using NUCmer [55] to identify homozygous pseudomolecules to chromosome ID. Finally, we filled the gaps with the correct PacBio reads, ultimately resulting in a high-quality genome assembly.

Evaluation of genome assembly quality

For assembly assessment, four main analyses were conducted. (1) The sequences of B73 with 2Bs underwent orthology detection using BUSCO [56], revealing the capture of 98.8% of complete BUSCO. (2) Intact LTR retrotransposons of B73 with 2Bs were identified using LTR_Finder (Version 1.07) [57], LTRharvest (Version 1.6.1) [58], and LTR_retriever (Version 2.9.0) [59], and the LAI value was calculated [59]. (3) RNA-seq reads obtained from leaves and tassels were mapped against our assembled genome using HISAT2 [60] with default settings. (4) Whole-genome re-sequencing data for various maize B chromosome variants were downloaded from the NCBI Gene Expression Omnibus (GEO; https://www.ncbi.nlm.nih.gov/geo/) database under accession numbers SRR11837267 to SRR11837299, respectively. These data were mapped to the B73 + 2Bs-v2.0 reference genome, and the alignment results corresponded to the chromosomal karyotypes.

Genome annotation

Total RNA was extracted from leaf and tassels using TRIzol for sequencing to support genome annotation. Before gene prediction, RepeatMasker [61] was used to mask the genome. The functions of protein-coding genes were predicted by three methods. First, *Ab inito* gene prediction was performed using GeneMark-ET (v4.3) [62] and AUGUSTUS (v3.2.3) [63]. Second, for homology-based gene prediction, protein sequences of B73_v5 were aligned to our genome assembly. Gemoma [64] was then used for gene annotation with default parameters. Third, RNA-seq reads from different tissues were mapped to the reference genome of B73 with 2Bs using BWA and STAR (v2.5.3a) [65]. Gene annotation was then performed using Mikado [66] with the combined evidence. The functional annotation of protein-coding genes and Gene Ontology term annotations were obtained using InterProScan (v.5.55–88.0) [67]. For TE annotation, a de novo repeat library was constructed using RepeatModeler (v2.0.1) (https://github.com/Dfam-consortium/RepeatModeler) and EDTA [59]. The CLARI-TE [64] tool was used to search for TEs using a homology-based search method. FL-LTR-RTs were identified using LTR_retriever (v2.9.0) [59].

Synteny analysis

During the synteny analysis, we compared the A chromosomes of our assembled genome with their corresponding chromosomes in Zm-B73-REFERENCE-NAM-5.0 [68] and the B chromosomes with those in Zm-B73_B-CHROMOSOME . The comparisons were conducted using MUMmer4 [55] with the command "nucmer -t 100" and the parameter "delta-filter -m -i 95 -l 1000." Synteny dotplots were drawn using RectChr (https://github.com/BGI-shenzhen/RectChr) and GenomeSyn [69].

ChIP-seq assay

ChIP was performed following a previously described protocol with some modifications [70]. Tissues were collected and ground in liquid nitrogen, then resuspended in ice-cold chromatin extraction buffer (50 mM HEPES pH7.5, 150 mM NaCl, 1 mM EDTA, 1% Trition X-100, 10% glycerol), and incubated at 4 °C for 20 min with gentle shaking. The solution was filtered through Miracloth and centrifuged at 1250 × g for 15 min at 4 °C. The pellet was washed until the solution color disappeared. The nuclei pellet was resuspended in MNase buffer (50% sucrose, 1 M $MgCl_2$, 1 M Tris–HCl, 0.1 M PMSF) and digested with MNase (New England BioLabs, 0.5 U) at 37℃ for 20 min. Sheared chromatin was centrifuged at 15,000 × g at 4 °C for 10 min, and 10 µl of the supernatant was used as the input control for subsequent sequencing. The remaining chromatin sample was incubated overnight at 4 °C with gentle rotation with 50 µl protein A beads (Invitrogen 10002D). Beads were separated on magnetic racks and washed three times each with low salt wash buffer (50 mM NaCl, 50 mM Tris-HCI, 0.2 mM PMSF, 1 × cOmplete Mini EDTA-free protease inhibitor cocktail, 10 mM EDTA), medium salt wash buffer (100 mM NaCl, 50 mM Tris-HCI, 0.2 mM PMSF, 1 × cOmplete Mini EDTA-free protease inhibitor cocktail, 10 mM EDTA), and high salt wash buffer (150 mM NaCl, 50 mM Tris-HCI, 0.2 mM PMSF, 1 × cOmplete Mini EDTA-free protease inhibitor cocktail, 10 mM EDTA). The beads were then resuspended in elution buffer (50 mM NaCl, 20 mM Tris-HCI, 5 mM EDTA, 2% SDS) and incubated at 65 °C for 55 min.

Sequencing libraries were prepared using the NadPrep DNA library Preparation Kit (1,002,101) and NadPrep UDI Adapter Kit Set C1 for Illumina (1,003,221) according to the standard Illumina protocol. Libraries were sequenced as 150-bp paired-end reads on the Illumina NovaSeq platform.

Identification of centromeric regions

5 g of fresh leaves were collected from B73 plants containing two B chromosomes for nuclei extraction. ChIP-seq was conducted using a maize rabbit polyclonal against anti-CENH3. This antibody targets the peptide RPGTVALREIRKYQKSSTSATPERAAGTG GR, as previously described [71]. Centromeric regions were identified based on the boundaries of the CENH3 ChIP-seq peaks.

ssDRIP-seq assay

ssDRIP-seq was conducted following a previously published protocol [41]. Briefly, nuclei isolation was performed similarly to the ChIP-seq assay described above. Nuclei pellets were resuspended in TE buffer supplemented with 0.1 mg/ml proteinase K and 0.5% SDS, and incubated at 37 °C overnight with gentle rotation. DNA was extracted using phenol/chloroform/isoamyl alcohol (25:24:1), dissolved in TE buffer, and quantified using the Qubit DNA quantification kit (Invitrogen). Approximately 15 μg of DNA was sonicated using a Covaris system with the following parameters: 60 s at a 10% Duty factor, 140 peak incident power, and 200 cycles/burst to yield fragments of approximately 300 bp. 1% of the sample was stored at − 20 °C as input, while half of the sample was treated with RNase A (Takara 2151) and RNaseH (Takara 2151) as a control. Next, both the treated and untreated sample incubated with 10 μg S9.6 antibody (Kerafast ENH001) overnight at 4 °C. Fifty-microliter protein G beads (Invitrogen) were incubated with the mixture for 3 h. The immunoprecipitants in an Eppendorf ThermoMixer were eluted with elution buffer (50 mM Tris pH 8.0, 10 mM EDTA) and proteinase K for 1 h at 55 °C. Single-stranded DNA library preparation was performed using the Accel-NGS Plus DNA Library Kit (Swift Biosciences) following the standard protocol.

ChIP-seq and DRIP-seq raw data processing and alignment

The raw reads of all the ChIP-seq and DRIP-seq data were quality-trimed using Trimmomatic (version 0.36) [72], and the quality-controlled reads were aligned to the reference genome of B73 with 2Bs using BWA with SAMtools (version 1.3.1) [73]. BAM files were processed using Deeptools (version 3.5.1) [74] to generate normalized coverage files in bigwig format with 50 bp bins, facilitating visualization in the Integrative Genomics Viewer (IGV) [75]

RNA-seq

We screened maize plants from the B73 inbred line with 0B, 2Bs, and 7Bs using FISH with the repeat pZmBs probe. Among the plants with 7Bs, those displaying the white striping phenotype were singled out for further analysis. Total RNA was extracted from both young leaves and spikelets in meiosis obtained from the tassels of the screened plants, all grown for the same duration. Each sample type comprised three biological replicates. Total RNA-seq libraries were constructed and sequenced

by AnnoGene (Beijing, China). The clean sequencing reads were mapped to our genome assembly using HISAT2 with default settings. Subsequently, featureCounts was employed to calculate the normalized expression levels (FPKM) of annotated genes. DESeq2 was utilized to identify the DEGs (padj < 0.5, $\log_2$(foldchange ≥ 1). GO enrichment analysis was conducted using R software (version 4.2.1). Overlapping and distinct genes were identified using the jvenn online software [76].

Fluorescence in situ hybridization (FISH)

Root tip cells from different plants were prepared for FISH following previously described protocols [77]. Initially, 1–2 cm root tips were treated with N_2O at 10 atm for 2 h. Subsequently, the tips were fixed in 90% acetic acid for 10 min, washed with water, and the root cap was removed. The remaining root tip was transferred to a tube containing 20 μl enzyme solution (1% w/w pectolyase Y-23, 2% w/w cellulase Onozuka R-10, 10 mM sodium citrate, and 10 mM EDTA), followed by incubation at 37 ℃ for 1 h. After three washes with 70% ethanol, the root cap was carefully broken using a dissecting needle. The cells were then centrifuged, ethanol was removed, and the cells were suspended in 30 μl of 100% acetic acid. Subsequently, 7 μl of the cell suspension was dropped onto a microscope slide in a humid chamber. After drying, the chromatin was crosslinked to the slide by exposure to UV light. Hybridization was carried out using DNA probes labeled with red fluor (5′ TEX 615) or green fluor (5′ 6- FAM). Prior to hybridization, root tip DNA and probe cocktail were denatured by floating in a boiling water bath for 5 min (100 ℃). Finally, the slides were placed in a humid storage container overnight at 55 ℃. Confocal microscopy (Zeiss Cell Observer SD) was employed to capture images, which were subsequently processed using ZEN 2009 Light Edition (Zeiss, http://www.zeiss.com/) and Adobe Photoshop CS 6.0 software.

Identification and sequence identity of centromeric repeats in maize A and B chromosomes

The heatmap depicting tandem repeat structures and their identity was generated using StainedGlass [78, 79]. The tandem repeats on the B chromosome were identified using two methods: TRASH [45] for the genome assembly and RepeatExplore2 [46] for CENH3 ChIP-seq reads. Sequence identity and length of distributions of all pZmBs monomers were derived from BLAST results. Dot plots comparing centromeric sequence similarity within and between chromosomes of our genome assembly were generated using Redotable (https://www.bioinformatics.babraham.ac.uk/projects/download.html#redotable).

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s13059-025-03517-6.

Additional file 1. This file contains Figures S1-S17.

Additional file 2. This file contains Tables S1-S14.

Additional file 3. The FASTA sequences of ZmBs223, ZmBs265,ZmBs244, ZmBs228, and pZmBs.

Additional file 4. Review history.

Acknowledgements
We thank Nathan D. Han from The Edison Family Center for his critical review of the manuscript and valuable feedback. We also thank Lili Dong (Beijing Institute of Genomics, Chinese Academy of Sciences) for her assistance with uploading the sequencing data.

Peer review information
Andrew Cosgrove was the primary editor of this article and managed its editorial process and peer review in collaboration with the rest of the editorial team. The peer-review history is available in the online version of this article.

Review history
The review history is available as Additional file 4.

Authors' contributions
F.P.H., and J.A.B., conceived and designed the project. Q.L., Y.L., Z.Y.Z., and C.L.Z. performed experiments and conducted fieldwork. Q.L., Y.L., C.Y.Y., and Z.G., performed the genome assembly and all other bioinformatic analyses. Q.L. wrote the manuscript with contributions from all authors. All authors read and approved the final manuscript.

Funding
This work was supported by the National Natural Science Foundation of China (31920103006), the National Funded Postdoctoral Researcher Program (GZB20230833), and the National Natural Science Foundation of Youth Fund Project (32400480). Research on this topic in the Birchler lab is supported by a grant from the USA National Science Foundation MCB2214243.

Data availability
The raw sequence data used for de novo whole-genome assembly are available from the NGDC Genome Sequence Archive (GSA; https://bigd.big.ac.cn/gsa/) under accession number CRA020357 [80]. The raw RNA-seq, ChIP-seq, ssDRIP-seq are available under accession numbers CRA020271 [81]. The assembly and gene feature annotation data for B73 with two B chromosomes are available under NCBI BioProject accession PRJNA1191055 [82]. Additionally, we utilized publicly available ChIP-seq datda (SRR1185582 [83], SRR1185584 [84], SRR1652605 [85], and SRR21509778 [86]), RNA-seq data (SRR3329319 [87], SRR3329320 [88], SRR3329321 [89], SRR8646342 [90], SRR8646345 [91], and SRR8646356 [92]), and sequence reads for B-deficiency-carrying maize lines (PRJNA634743 [93]) downloaded from the NCBI Gene Expression Omnibus (GEO; https://www.ncbi.nlm.nih.gov/geo/).

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
Fangpu Han is a guest editor for Genome Biology's article collection on Centromere Structure and Evolution, but was not involved in the editorial assessment or peer review of this article.

Received: 10 October 2024 Accepted: 25 February 2025
Published online: 06 March 2025

References

1. Randolph LF. Genetic characteristics of the B chromosomes in maize. Genetics. 1941;26:608–31.
2. D'Ambrosio U, Alonso-Lifante MP, Barros K, Kovarik A, Mas de Xaxars G, Garcia S. B-chrom: a database on B-chromosomes of plants, animals and fungi. New Phytol. 2017;216:635–642.
3. Ferree PM, Blagojevic J, Houben A, Martins C, Trifonov VA, Vujosevic M. What is a B chromosome? Early definitions revisited. G3 (Bethesda). 2024;16:jkae068.
4. Blavet N, Yang H, Su H, Solansky P, Douglas RN, Karafiatova M, Simkova L, Zhang J, Liu Y, Hou J, et al. Sequence of the supernumerary B chromosome of maize provides insight into its drive mechanism and evolution. Proc Natl Acad Sci U S A. 2021;118: e2104254118.
5. Martis MM, Klemme S, Banaei-Moghaddam AM, Blattner FR, Macas J, Schmutzer T, Scholz U, Gundlach H, Wicker T, Simkova H, et al. Selfish supernumerary chromosome reveals its origin as a mosaic of host genome and organellar sequences. Proc Natl Acad Sci U S A. 2012;109:13343–6.
6. Yoshida K, Terai Y, Mizoiri S, Aibara M, Nishihara H, Watanabe M, Kuroiwa A, Hirai H, Hirai Y, Matsuda Y, Okada N. B chromosomes have a functional effect on female sex determination in Lake Victoria cichlid fishes. PLoS Genet. 2011;7: e1002203.
7. Johnson Pokorna M, Reifova R. Evolution of B chromosomes: from dispensable parasitic chromosomes to essential genomic players. Front Genet. 2021;12: 727570.
8. Carchilan AHaM. Plant B chromosomes: what makes them different? In: Bass, H, Birchler, J (eds). Plant Cytogenet. 2011;4:59–77.
9. Rhoades MM. Studies on the cytological basis of crossing over. In: Peacock WJ, Brock RD, editors. Replication and recombination of genetic material. Canberra: Australian Academy of Science; 1968;229–41.

10. Hong ZJ, Xiao JX, Peng SF, Lin YP, Cheng YM. Novel B-chromosome-specific transcriptionally active sequences are present throughout the maize B chromosome. Mol Genet Genomics. 2020;295:313–25.
11. Habig M, Quade J, Stukenbrock EH. Forward genetics approach reveals host genotype-dependent importance of accessory chromosomes in the fungal wheat pathogen Zymoseptoria tritici. mBio. 2017;8:e01919-01917.
12. Chen J, Birchler JA, Houben A. The non-Mendelian behavior of plant B chromosomes. Chromosome Res. 2022;30:229–39.
13. Banaei-Moghaddam AM, Schubert V, Kumke K, Weibeta O, Klemme S, Nagaki K, Macas J, Gonzalez-Sanchez M, Heredia V, Gomez-Revilla D, et al. Nondisjunction in favor of a chromosome: the mechanism of rye B chromosome drive during pollen mitosis. Plant Cell. 2012;24:4124–34.
14. Houben A, Banaei-Moghaddam AM, Klemme S, Timmis JN. Evolution and biology of supernumerary B chromosomes. Cell Mol Life Sci. 2014;71:467–78.
15. Birchler JA, Yang H. The supernumerary B chromosome of maize: drive and genomic conflict. Open Biol. 2021;11:210197.
16. Roman H. Mitotic nondisjunction in the case of interchanges involving the B-type chromosome in maize. Genetics. 1947;32:391–409.
17. Carlson WR. Locating a site on the maize B chromosome that controls preferential fertilization. Genome. 2007;50:578–87.
18. Han F, Lamb JC, Yu W, Gao Z, Birchler JA. Centromere function and nondisjunction are independent components of the maize B chromosome accumulation mechanism. Plant Cell. 2007;19:524–33.
19. Su H, Liu Y, Liu Y, Birchler JA, Han F. The behavior of the maize B chromosome and centromere. Genes (Basel). 2018;9:476.
20. Roman H. Directed fertilization in maize. Proc Natl Acad Sci U S A. 1948;34:36–42.
21. Werren JH, Nur U, Eickbush D. An extrachromosomal factor causing loss of paternal chromosomes. Nature. 1987;327:75–6.
22. Ruban A, Schmutzer T, Wu DD, Fuchs J, Boudichevskaia A, Rubtsova M, Pistrick K, Melzer M, Himmelbach A, Schubert V, et al. Supernumerary B chromosomes of *Aegilops speltoides* undergo precise elimination in roots early in embryo development. Nat Commun. 2020;11:2764.
23. Ruban A, Fuchs J, Marques A, Schubert V, Soloviev A, Raskina O, Badaeva E, Houben A. B chromosomes of *Aegilops speltoides* are enriched in organelle genome-derived sequences. PLoS One. 2014;9: e90214.
24. Wu D, Ruban A, Fuchs J, Macas J, Novak P, Vaio M, Zhou Y, Houben A. Nondisjunction and unequal spindle organization accompany the drive of *Aegilops speltoides* B chromosomes. New Phytol. 2019;223:1340–52.
25. Chen J, Bartos J, Boudichevskaia A, Voigt A, Rabanus-Wallace MT, Dreissig S, Tulpova Z, Simkova H, Macas J, Kim G, et al. The genetic mechanism of B chromosome drive in rye illuminated by chromosome-scale assembly. Nat Commun. 2024;15:9686.
26. Belyayev A, Raskina O. Chromosome evolution in marginal populations of *Aegilops speltoides*: causes and consequences. Ann Bot. 2013;111:531–8.
27. Houben ATN, Ahne R, Leach CR, Verlin D, Timmis JN. A monophyletic origin of the B chromosomes of Brachycome dichromosomatica (Asteraceae). Plant Syst Evol. 1999;219:127–35.
28. Bakkali MCJ, Lopez-Leon MD, Perfectti F, Camacho JPM. The B chromosome polymorphism of the grasshopper Eyprepocnemis plorans in north Africa. Heredity. 1999;83:428–34.
29. Lamb JC, Kato A, Birchler JA. Sequences associated with A chromosome centromeres are present throughout the maize B chromosome. Chromosoma. 2005;113:337–49.
30. Shi X, Yang H, Chen C, Hou J, Ji T, Cheng J, Birchler JA. Effect of aneuploidy of a non-essential chromosome on gene expression in maize. Plant J. 2022;110:193–211.
31. Jones RN, Viegas W, Houben A. A century of B chromosomes in plants: so what? Ann Bot. 2008;101:767–75.
32. Cheng H, Concepcion GT, Feng X, Zhang H, Li H. Haplotype-resolved *de novo* assembly using phased assembly graphs with hifiasm. Nat Methods. 2021;18:170–5.
33. Sun H, Jiao WB, Krause K, Campoy JA, Goel M, Folz-Donahue K, Kukat C, Huettel B, Schneeberger K. Chromosome-scale and haplotype-resolved genome assembly of a tetraploid potato cultivar. Nat Genet. 2022;54:342–8.
34. Peng Y, Yan H, Guo L, Deng C, Wang C, Wang Y, Kang L, Zhou P, Yu K, Dong X, et al. Reference genome assemblies reveal the origin and evolution of allohexaploid oat. Nat Genet. 2022;54:1248–58.
35. Hufford MB, Seetharam AS, Woodhouse MR, Chougule KM, Ou S, Liu J, Ricci WA, Guo T, Olson A, Qiu Y, et al. *De novo* assembly, annotation, and comparative analysis of 26 diverse maize genomes. Science. 2021;373:655–62.
36. Trapnell C, Roberts A, Goff L, Pertea G, Kim D, Kelley DR, Pimentel H, Salzberg SL, Rinn JL, Pachter L. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. Nat Protoc. 2012;7:562–78.
37. Kovaka S, Zimin AV, Pertea GM, Razaghi R, Salzberg SL, Pertea M. Transcriptome assembly from long-read RNA-seq alignments with StringTie2. Genome Biol. 2019;20:278.
38. Huang W, Du Y, Zhao X, Jin W. B chromosome contains active genes and impacts the transcription of A chromosomes in maize (Zea mays L.). BMC Plant Biol. 2016;16:88.
39. Xu G, Lyu J, Li Q, Liu H, Wang D, Zhang M, Springer NM, Ross-Ibarra J, Yang J. Evolutionary and functional genomics of DNA methylation in maize domestication and improvement. Nat Commun. 2020;11:5539.
40. Naish M, Alonge M, Wlodzimierz P, Tock AJ, Abramson BW, Schmucker A, Mandakova T, Jamge B, Lambing C, Kuo P, et al. The genetic and epigenetic landscape of the Arabidopsis centromeres. Science. 2021;374: eabi7489.
41. Liu Y, Liu Q, Su H, Liu K, Xiao X, Li W, Sun Q, Birchler JA, Han F. Genome-wide mapping reveals R-loops associated with centromeric repeats in maize. Genome Res. 2021;31:1409–18.
42. Liu Y, Su H, Zhang J, Liu Y, Feng C, Han F. Back-spliced RNA from retrotransposon binds to centromere and regulates centromeric chromatin loops in maize. PLoS Biol. 2020;18: e3000582.
43. Henikoff S, Ahmad K, Malik HS. The centromere paradox: stable inheritance with rapidly evolving DNA. Science. 2001;293:1098–102.
44. Alfenito MR, Birchler JA. Molecular characterization of a maize B chromosome centric sequence. Genetics. 1993;135:589–97.

45. Wlodzimierz P, Hong M, Henderson IR. TRASH: tandem repeat annotation and structural hierarchy. Bioinformatics. 2023;39: btad308.
46. Novak P, Neumann P, Pech J, Steinhaisl J, Macas J. RepeatExplorer: a galaxy-based web server for genome-wide characterization of eukaryotic repetitive elements from next-generation sequence reads. Bioinformatics. 2013;29:792–3.
47. Staub RW. Leaf striping correlated with the presence of B chromosomes in maize. J Hered. 1987;78:71–4.
48. Skourti-Stathaki K, Kamieniarz-Gdula K, Proudfoot NJ. R-loops induce repressive chromatin marks over mammalian gene terminators. Nature. 2014;516:436–9.
49. Kabeche L, Nguyen HD, Buisson R, Zou L. A mitosis-specific and R loop-driven ATR pathway promotes faithful chromosome segregation. Science. 2018;359:108–14.
50. Xu Y, Jiao Y, Liu C, Miao R, Liu C, Wang Y, Ma C, Liu J. R-loop and diseases: the cell cycle matters. Mol Cancer. 2024;23:84.
51. Page BT, Wanous MK, Birchler JA. Characterization of a maize chromosome 4 centromeric sequence evidence for an evolutionary relationship with the B chromosome centromere. Genetics. 2001;159:291–302.
52. Rautiainen M, Nurk S, Walenz BP, Logsdon GA, Porubsky D, Rhie A, Eichler EE, Phillippy AM, Koren S. Telomere-to-telomere assembly of diploid chromosomes with Verkko. Nat Biotechnol. 2023;41:1474–82.
53. Zhang H, Song L, Wang X, Cheng H, Wang C, Meyer CA, Liu T, Tang M, Aluru S, Yue F, et al. Fast alignment and preprocessing of chromatin profiles with Chromap. Nat Commun. 2021;12:6566.
54. Durand NC, Shamim MS, Machol I, Rao SS, Huntley MH, Lander ES, Aiden EL. Juicer provides a one-click system for analyzing loop-resolution Hi-C experiments. Cell Syst. 2016;3:95–8.
55. Marcais G, Delcher AL, Phillippy AM, Coston R, Salzberg SL, Zimin A. MUMmer4: a fast and versatile genome alignment system. PLoS Comput Biol. 2018;14:e1005944.
56. Simao FA, Waterhouse RM, Ioannidis P, Kriventseva EV, Zdobnov EM. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. Bioinformatics. 2015;31:3210–2.
57. Xu Z, Wang H. LTR_FINDER: an efficient tool for the prediction of full-length LTR retrotransposons. Nucleic Acids Res. 2007;35:W265-268.
58. Ellinghaus D, Kurtz S, Willhoeft U. LTRharvest, an efficient and flexible software for *de novo* detection of LTR retrotransposons. BMC Bioinformatics. 2008;9:18.
59. Ou S, Chen J, Jiang N. Assessing genome assembly quality using the LTR Assembly Index (LAI). Nucleic Acids Res. 2018;46: e126.
60. Kim D, Langmead B, Salzberg SL. HISAT: a fast spliced aligner with low memory requirements. Nat Methods. 2015;12:357–60.
61. Tarailo-Graovac M, Chen N. Using RepeatMasker to identify repetitive elements in genomic sequences. Curr Protoc Bioinformatics. 2009;25:1–14.
62. Ter-Hovhannisyan V, Lomsadze A, Chernoff YO, Borodovsky M. Gene prediction in novel fungal genomes using an ab initio algorithm with unsupervised training. Genome Res. 2008;18:1979–90.
63. Hoff KJ, Stanke M. Predicting genes in single genomes with AUGUSTUS. Curr Protoc Bioinformatics. 2019;65:e57.
64. Daron J, Glover N, Pingault L, Theil S, Jamilloux V, Paux E, Barbe V, Mangenot S, Alberti A, Wincker P, et al. Organization and evolution of transposable elements along the bread wheat chromosome 3B. Genome Biol. 2014;15:546.
65. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR. STAR: ultrafast universal RNA-seq aligner. Bioinformatics. 2013;29:15–21.
66. Venturini L, Caim S, Kaithakottil GG, Mapleson DL, Swarbreck D. Leveraging multiple transcriptome assembly methods for improved gene structure annotation. Gigascience. 2018;7:giy093.
67. Quevillon E, Silventoinen V, Pillai S, Harte N, Mulder N, Apweiler R, Lopez R. InterProScan: protein domains identifier. Nucleic Acids Res. 2005;33:W116-120.
68. Liu J, Seetharam AS, Chougule K, Ou S, Swentowsky KW, Gent JI, Llaca V, Woodhouse MR, Manchanda N, Presting GG, et al. Gapless assembly of maize chromosomes using long-read technologies. Genome Biol. 2020;21:121.
69. Zhou ZW, Yu ZG, Huang XM, Liu JS, Guo YX, Chen LL, Song JM. GenomeSyn: a bioinformatics tool for visualizing genome synteny and structural variations. J Genet Genomics. 2022;49:1174–6.
70. Liu Q, Yi C, Zhang Z, Su H, Liu C, Huang Y, Li W, Hu X, Liu C, Birchler JA, et al. Non-B-form DNA tends to form in centromeric regions and has undergone changes in polyploid oat subgenomes. Proc Natl Acad Sci U S A. 2023;120:e2211683120.
71. Fu S, Lv Z, Gao Z, Wu H, Pang J, Zhang B, Dong Q, Guo X, Wang XJ, Birchler JA, Han F. *De novo* centromere formation on a chromosome fragment in maize. Proc Natl Acad Sci USA. 2013;110:6033–6.
72. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics. 2014;30:2114–20.
73. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R. The sequence alignment/map format and SAMtools. Bioinformatics. 2009;25:2078–9.
74. Ramirez F, Dundar F, Diehl S, Gruning BA, Manke T. deepTools: a flexible platform for exploring deep-sequencing data. Nucleic Acids Res. 2014;42:187–91.
75. Thorvaldsdottir H, Robinson JT, Mesirov JP. Integrative genomics viewer (IGV): high-performance genomics data visualization and exploration. Brief Bioinform. 2013;14:178–92.
76. Bardou P, Mariette J, Escudie F, Djemiel C, Klopp C. jvenn: an interactive Venn diagram viewer. BMC Bioinformatics. 2014;15:293.
77. Liu Y, Yi C, Fan C, Liu Q, Liu S, Shen L, Zhang K, Huang Y, Liu C, Wang Y, et al. Pan-centromere reveals widespread centromere repositioning of soybean genomes. Proc Natl Acad Sci U S A. 2023;120:e2310177120.
78. Vollger MR, Kerpedjiev P, Phillippy AM, Eichler EE. StainedGlass: interactive visualization of massive tandem repeat structures with identity heatmaps. Bioinformatics. 2022;38:2049–51.
79. Cui J, Zhu C, Shen L, Yi C, Wu R, Sun X, Han F, Li Y, Liu Y. The gap-free genome of Forsythia suspensa illuminates the intricate landscape of centromeres. Hortic Res. 2024;11:uhae185.

80. Liu Q, Liu Y, Yi CY, Gao Z, Zhang ZY, Zhu CL, Birchler JA, Han FP. Genome assembly of the maize B chromosome provides insight into its epigenetic characteristics and effects on the host genome. CRA020357. Genome Sequence Archive. 2025. https://ngdc.cncb.ac.cn/gsa/search?searchTerm=CRA020357.
81. Liu Q, Liu Y, Yi CY, Gao Z, Zhang ZY, Zhu CL, Birchler JA, and Han FP. Genome assembly of the maize B chromosome provides insight into its epigenetic characteristics and effects on the host genome. CRA020271. Genome Sequence Archive. 2025. https://ngdc.cncb.ac.cn/gsa/search?searchTerm=CRA020271.
82. Liu Q, Liu Y, Yi CY, Gao Z, Zhang ZY, Zhu CL, Birchler JA, Han FP. Genome assembly of the maize B chromosome provides insight into its epigenetic characteristics and effects on the host genome. PRJNA1191055. Genome Sequence Archive. 2025. https://www.ncbi.nlm.nih.gov/bioproject/PRJNA1191055.
83. He Y, Wang M, Dukowic-Schulze S, Zhou A, Tiang CL, Shilo S, Sidhu GK, Eichten S, Bradbury P, Springer NM, Buckler ES, Levy AA, Sun Q, Pillardy J, Kianian PMA, Kianian SF, Chen C, Pawlowski WP. Genomic features shaping the landscape of meiotic double-strand-break hotspots in maize. SRR1185582. Sequence Read Archive. 2017. https://www.ncbi.nlm.nih.gov/search/all/?term=SRR1185582.
84. He Y, Wang M, Dukowic-Schulze S, Zhou A, Tiang CL, Shilo S, Sidhu GK, Eichten S, Bradbury P, Springer NM, Buckler ES, Levy AA, Sun Q, Pillardy J, Kianian PMA, Kianian SF, Chen C, Pawlowski WP. Genomic features shaping the landscape of meiotic double-strand-break hotspots in maize. SRR1185584. Sequence Read Archive. 2017. https://www.ncbi.nlm.nih.gov/sra/?term=SRR1185584.
85. Gent JI, Wang K, Jiang J, Dawe RK. Stable patterns of CENH3 occupancy through maize lineages containing genetically similar centromeres. SRR1652605. Sequence Read Archive. 2015. https://www.ncbi.nlm.nih.gov/sra/?term=SRR1652605.
86. Chen J, Wang Z, Tan K, Huang W, Shi J, Li T, Hu J, Wang K, Wang C, Xin B, Zhao H, Song W, Hufford MB, Schnable JC, Jin W, Lai J. A complete telomere-to-telomere assembly of the maize genome. SRR21509778. Sequence Read Archive. 2023. https://www.ncbi.nlm.nih.gov/sra/?term=SRR21509778.
87. Huang W, Du Y, Zhao X, Jin W. B chromosome contains active genes and impacts the transcription of A chromosomes in maize (Zea mays L.). SRR3329319. Sequence Read Archive. 2016. https://www.ncbi.nlm.nih.gov/sra/?term=SRR3329319.
88. Huang W, Du Y, Zhao X, Jin W. B chromosome contains active genes and impacts the transcription of A chromosomes in maize (Zea mays L.). SRR3329320. Sequence Read Archive. 2016. https://www.ncbi.nlm.nih.gov/sra/?term=SRR3329320.
89. Huang W, Du Y, Zhao X, Jin W. B chromosome contains active genes and impacts the transcription of A chromosomes in maize (Zea mays L.). SRR3329321. Sequence Read Archive. 2016. https://www.ncbi.nlm.nih.gov/sra/?term=SRR3329321.
90. Shi X, Yang H, Chen C, Hou J, Ji T, Cheng J, Birchler JA. Effect of aneuploidy of a non-essential chromosome on gene expression in maize. SRR8646342. Sequence Read Archive. 2020. https://www.ncbi.nlm.nih.gov/sra/?term=SRR8646342.
91. Shi X, Yang H, Chen C, Hou J, Ji T, Cheng J, Birchler JA. Effect of aneuploidy of a non-essential chromosome on gene expression in maize. SRR8646345. Sequence Read Archive. 2020. https://www.ncbi.nlm.nih.gov/sra/?term=SRR8646345.
92. Shi X, Yang H, Chen C, Hou J, Ji T, Cheng J, Birchler JA. Effect of aneuploidy of a non-essential chromosome on gene expression in maize. SRR8646356. Sequence Read Archive. 2020. https://www.ncbi.nlm.nih.gov/sra/?term=SRR8646356.
93. Blavet N, Yang H, Su H, Solansky P, Douglas RN, Karafiatova M, Simkova L, Zhang J, Liu Y, Hou J, Shi X, Chen C, El-Walid M, McCaw ME, Albert PS, Gao Z, Zhao C, Ben-Zvi G, Glick L, Kol G, Shi J, Vrana J, Simkova H, Lamb JC, Newton K, Dawe RK, Dolezel J, Ji T, Baruch K, Cheng J, Han F, Birchler JA, Bartos J. Sequence of the supernumerary B chromosome of maize provides insight into its drive mechanism and evolution. PRJNA634743. Sequence Read Archive. 2021. https://www.ncbi.nlm.nih.gov/bioproject/?term=PRJNA634743.

Publisher's Note