The Plant Cell, Vol. 25: 1979–1989, June 2013, www.plantcell.org

RESEARCH ARTICLES

Formation of a Functional Maize Centromere after Loss of Centromeric Sequences and Gain of Ectopic Sequences[C][W]

Bing Zhang,[a,b,1] Zhenling Lv,[a,1] Junling Pang,[b,c,1] Yalin Liu,[a,b] Xiang Guo,[a,b] Shulan Fu,[a] Jun Li,[a] Qianhua Dong,[a,b] Hua-Jun Wu,[c] Zhi Gao,[d] Xiu-Jie Wang,[c,1] and Fangpu Han[a,1,2]

[a] State Key Laboratory of Plant Cell and Chromosome Engineering, Institute of Genetics and Developmental Biology, Chinese Academy of Sciences, Beijing 100101, China

[b] University of Chinese Academy of Sciences, Beijing 100049, China

[c] State Key Laboratory of Plant Genomics, Institute of Genetics and Developmental Biology, Chinese Academy of Sciences, Beijing 100101, China

[d] Division of Biological Sciences, University of Missouri, Columbia, Missouri 65211-7400

ORCID ID: 0000-0001-8393-3575 (F.H.).

The maize (*Zea mays*) B centromere is composed of B centromere–specific repeats (ZmBs), centromere-specific satellite repeats (CentC), and centromeric retrotransposons of maize (CRM). Here we describe a newly formed B centromere in maize, which has lost CentC sequences and has dramatically reduced CRM and ZmBs sequences, but still retains the molecular features of functional centromeres, such as CENH3, H2A phosphorylation at Thr-133, H3 phosphorylation at Ser-10, and Thr-3 immunostaining signals. This new centromere is stable and can be transmitted to offspring through meiosis. Anti-CENH3 chromatin immunoprecipitation sequencing revealed that a 723-kb region from the short arm of chromosome 9 (9S) was involved in the formation of the new centromere. The 723-kb region, which is gene poor and enriched for transposons, contains two abundant DNA motifs. Genes in the new centromere region are still transcribed. The original 723-kb region showed a higher DNA methylation level compared with native centromeres but was not significantly changed when it was involved in new centromere formation. Our results indicate that functional centromeres may be formed without the known centromere-specific sequences, yet the maintenance of a high DNA methylation level seems to be crucial for the proper function of a new centromere.

INTRODUCTION

The centromere is a unique chromosomal region that ensures accurate segregation of chromosomes during mitosis and meiosis in eukaryotes. A normal eukaryotic chromosome only has one centromere, typically located within a heterochromatic region mainly composed of repetitive sequences (Jiang et al., 2003; Ma et al., 2007; Henikoff and Furuyama, 2010). Although the DNA sequences of centromeres are highly variable among species, common specific epigenetic modifications have been found within centromeric regions of all higher eukaryotes. Typical epigenetic marks of functional centromeres include the binding of a conserved variant of conventional histone H3, termed CENP-A in mammals or CENH3 in plants (Henikoff et al., 2001; Cleveland et al., 2003; Birchler et al., 2011), as well as the presence of H2A phosphorylation at Thr-133 (Dong and Han, 2012) and H3 phosphorylation at Ser-10 (Houben et al., 1999).

Newly established centromeres, termed neocentromeres, have been found in several organisms, including humans (Choo, 2001; Marshall et al., 2008; Burrack and Berman, 2012). In humans, over 90 neocentromeres have been reported on 20 different chromosomes (Marshall et al., 2008). In plants, neocentromeres formed in barley (*Hordeum vulgare*) and oat (*Avena sativa*)-maize (*Zea mays*) addition lines can also be stably transmitted to offspring (Nasuda et al., 2005; Topp et al., 2009). Neocentromeres are usually formed at noncentromeric locations and may lack certain centromere-specific DNA sequences. Dicentric chromosomes contain another form of epigenetically determined centromere. Chromosomal rearrangement, or de novo centromere formation, can produce a chromosome containing two centromeric regions. To be stably inherited, one centromere of the dicentric chromosome must be inactivated, as a chromosome with two functional centromeres will suffer chromosomal breakage (Han et al., 2006, 2007; Zhang et al., 2010; Gao et al., 2011). If both centromeres remain active and the chromosome pulls apart, the broken ends of the chromosome will subsequently fuse and undergo the breakage-fusion-bridge (BFB) cycle to form a new dicentric chromosome (McClintock, 1939, 1941). The mechanism of how neocentromeres are formed is unknown, but identification of neocentromeres can shed light on the understanding of centromere structural variation and the mechanism of centromere formation.

Maize has been widely used as a model organism for centromere studies. In addition to the standard A chromosomes,

[1] These authors contributed equally to this work.

[2] Address correspondence to fphan@genetics.ac.cn.

The author responsible for distribution of materials integral to the findings presented in this article in accordance with the policy described in the Instructions for Authors (www.plantcell.org) is: Fangpu Han (fphan@genetics.ac.cn).

[C] Some figures in this article are displayed in color online but in black and white in the print edition.

[W] Online version contains Web-only data.

www.plantcell.org/cgi/doi/10.1105/tpc.113.110015

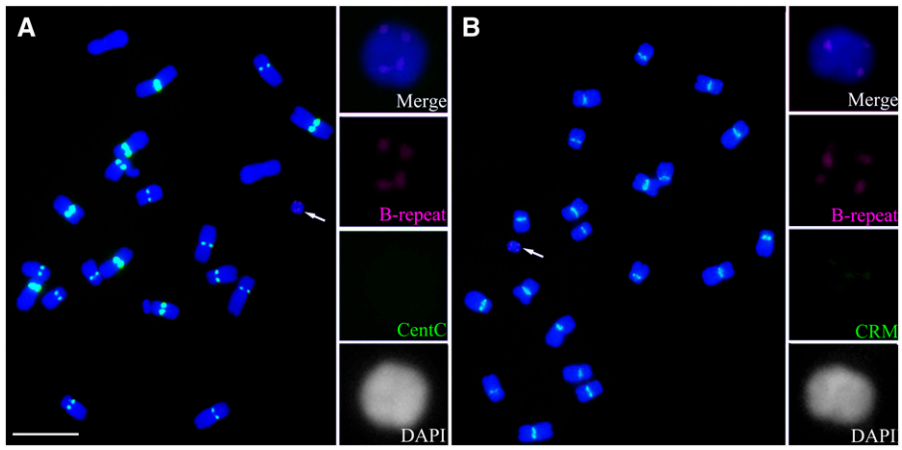

Figure 1. Cytological Analysis of Chromosome sDic-15.

(A) Somatic metaphase chromosomes of a maize line containing sDic-15, probed with the centromere-specific tandem repeat CentC (green) and B chromosome–specific ZmBs (magenta). Bar = 10 μm.
(B) Somatic metaphase chromosomes of maize probed with CRM (green) and ZmBs (magenta).
Arrows indicate the B chromosome, and insets show a higher magnification view of the B chromosome.

some maize lines possess a supernumerary B chromosome. The centromeres of both A and B chromosomes contain two types of centromere-specific sequences: the 156-bp satellite repeat CentC and centromeric retrotransposon of maize (CRM) (Jiang et al., 1996; Ananiev et al., 1998; Nagaki et al., 2003; Sharma et al., 2008; Birchler and Han, 2009). In addition, the centromeres of B chromosomes also contain two megabase blocks of B centromere–specific repeats (ZmBs) that are interspersed with the CentC and CRM elements and form a 700-kb core region (Alfenito and Birchler, 1993; Jin et al., 2005; Lamb et al., 2005). Translocations between the B and A chromosomes (or B-A translocated chromosomes) provide good materials for centromere studies (Birchler et al., 2011).

Previously, we reported a dicentric chromosome named Dic-15 containing a large and a small B centromere generated from the cross of two B-A translocated chromosomes TB-9Sb-Dp9 and T3-5(+) (Han et al., 2009). In the stable dicentric chromosome Dic15, the small centromere is inactivated. However, the inactive centromere can be reactivated when separated from the large active centromere by intrachromosomal recombination. Dic-15 can generate many distinct progenies through this form of recombination. We describe here a progeny of Dic-15 named sDic-15 (based on its centromere coming from the small centromere of Dic15). This newly formed B centromere has no detectable CentC sequences and strongly reduced CRM and B repeat sequences. Using chromatin immunoprecipitation sequencing (ChIP-Seq), we found that the new centromere contains a 723-kb-long CENH3 binding domain with sequence similarity to a region within the short arm of maize chromosome 9 (9S). The 723-kb region has a higher content of repetitive sequences compared with average euchromatic regions and contains only six protein-coding genes, all of which are actively transcribed. This region is also enriched for two kinds of DNA motifs, which are shared with centromeres in maize. Both the 723-kb neocentromere region and its parental sequences on chromosome 9S have a higher DNA methylation level than euchromatin regions, suggesting that the maintenance of a high DNA methylation level, similar to the native centromeres, is important for the formation of a new centromere.

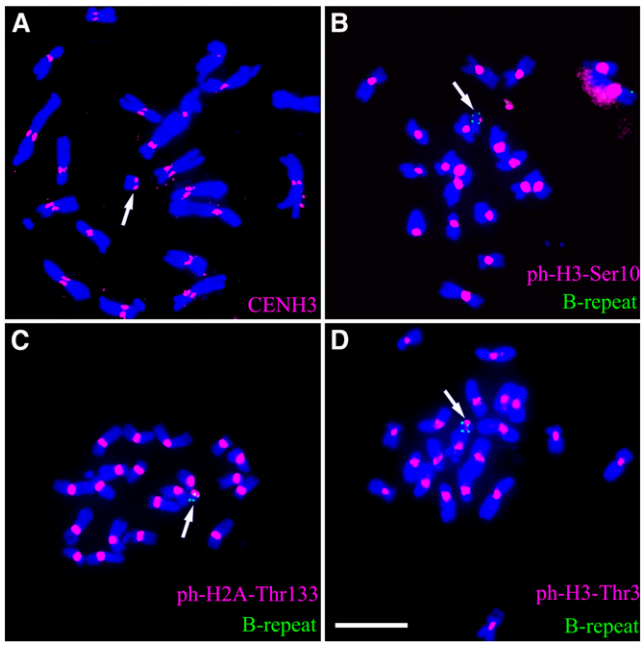

Figure 2. Immunolocalization Analysis of Chromosome sDic-15.

(A) Immunostaining with anti-CENH3 (magenta) antibodies on somatic metaphase chromosomes on sDic-15 (arrow).
(B) Immunostaining with ph-H3-Ser10 (magenta) antibodies and FISH with ZmBs (green) sequences on sDic-15 (arrow).
(C) Immunostaining with ph-H2A-Thr133 (magenta) antibodies and FISH with ZmBs (green) sequences on sDic-15 (arrow).
(D) Immunostaining with ph-H3-Thr3 (magenta) antibodies and FISH with ZmBs (green) sequences on sDic-15 (arrow). Bar = 10 μm.

RESULTS

Identification of A Dicentric Chromosome sDic-15 with a Newly Formed Centromere

Our previous work identified a maize line with a dicentric chromosome, Dic-15, which contains both a large and a small version of the B centromere (Han et al., 2009). Further analysis revealed that Dic-15 contains two copies of the 9S (short arm of chromosome 9), which allowed it to undergo centromere activity state changes either through intrachromosomal recombination or by a centromere reactivation process (Han et al., 2009). From our analyses of Dic-15 progenies, we discovered a new chromosome called sDic-15. The sDic-15 chromosome contained two altered B repeat signals, both of which no longer contained detectable levels of CentC, as determined by fluorescence in situ hybridization (FISH), and had dramatically reduced CRM sequences compared with native centromeres (Figures 1A and 1B). sDic-15 also contained a small amount of B repeat sequence in the centromeres, as detected by FISH (Figures 1A and 1B). Strong telomere signals were detected at the distal end of the sDic-15 chromosome (see Supplemental Figure 1 online), indicating that the newly formed chromosome was a stable linear chromosome.

Functional Analysis of the Newly Formed Dicentric Chromosome sDic-15

FISH analysis of sDic-15 using centromeric-specific sequences as probes indicated that both centromeres of sDic-15 have altered centromeric sequences, including deletion of the CentC sequences. Both of the centromeres appeared to be stable in mitosis, as no breakage of sDic15 was observed in root tip samples from 230 seedlings. We used antibodies against CENH3, H2A phosphorylation at Thr-133 (Dong and Han, 2012), and H3 phosphorylation at Ser-10 and Thr-3 to characterize the centromere activity of sDic-15. Only one centromere exhibited detectable immunostaining signals with the four antibodies (Figure 2). We also detected the CENH3 protein signals in meiosis and confirmed that only one centromere of sDic-15 was active (see Supplemental Figure 2 online).

The sDic-15 Chromosome Can Be Stably Inherited through Meiosis

To test whether sDic-15 can be stably transmitted during meiosis, we collected tassel samples from several greenhouse-grown plants containing sDic-15 and analyzed the meiotic behavior of pollen mother cells. Although the linear sDic-15 chromosome contained a large fragment of 9S, it did not pair with chromosome 9, and there were no crossovers detected (Figure 3A). sDic-15 exhibited behavior typical of a univalent chromosome (Figure 3B). A total of 156 from four plants were examined, and during meiosis I, sDic-15 randomly moved to one pole of the cell (Figures 3C and 3D). The sister chromatids separated at anaphase II (Figure 3E). We then screened the progeny of sDic-15 and found that all examined seedlings that contained a single copy of sDic-15 showed no breakage of the sDic-15 chromosome during mitosis in dividing cells collected from root tips. However, we did observe chromosome fragments at a rather lower frequency (see Supplemental Figure 3 online); nine out of 239 plants contained a monocentric chromosome fragment that was likely to be the result of breakage of anaphase II bridges (Han et al., 2009).

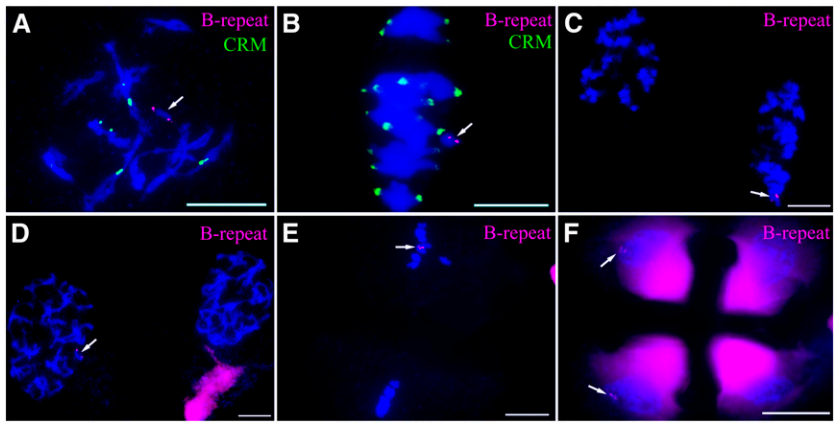

Figure 3. Meiotic Analysis of One Copy of Chromosome sDic-15.

Examination of meiotic chromosome behavior of maize sDic-15 in pollen mother cells. ZmBs sequences are labeled in magenta, and CRM is labeled in green. Arrows point to the sDic-15 chromosome. Bars = 10 μm.

(A) Diakinesis.

(B) Metaphase I. sDic-15 has moved to the plate with the other bivalents.

(C) Anaphase I. sDic-15 has moved to one pole randomly.

(D) Early prophase II. sDic-15 still stays at one pole.

(E) Metaphase II. sDic-15 has moved to the plate with the other chromosome.

(F) Tetrad. sDic-15 sister chromatids are separated to the two poles.

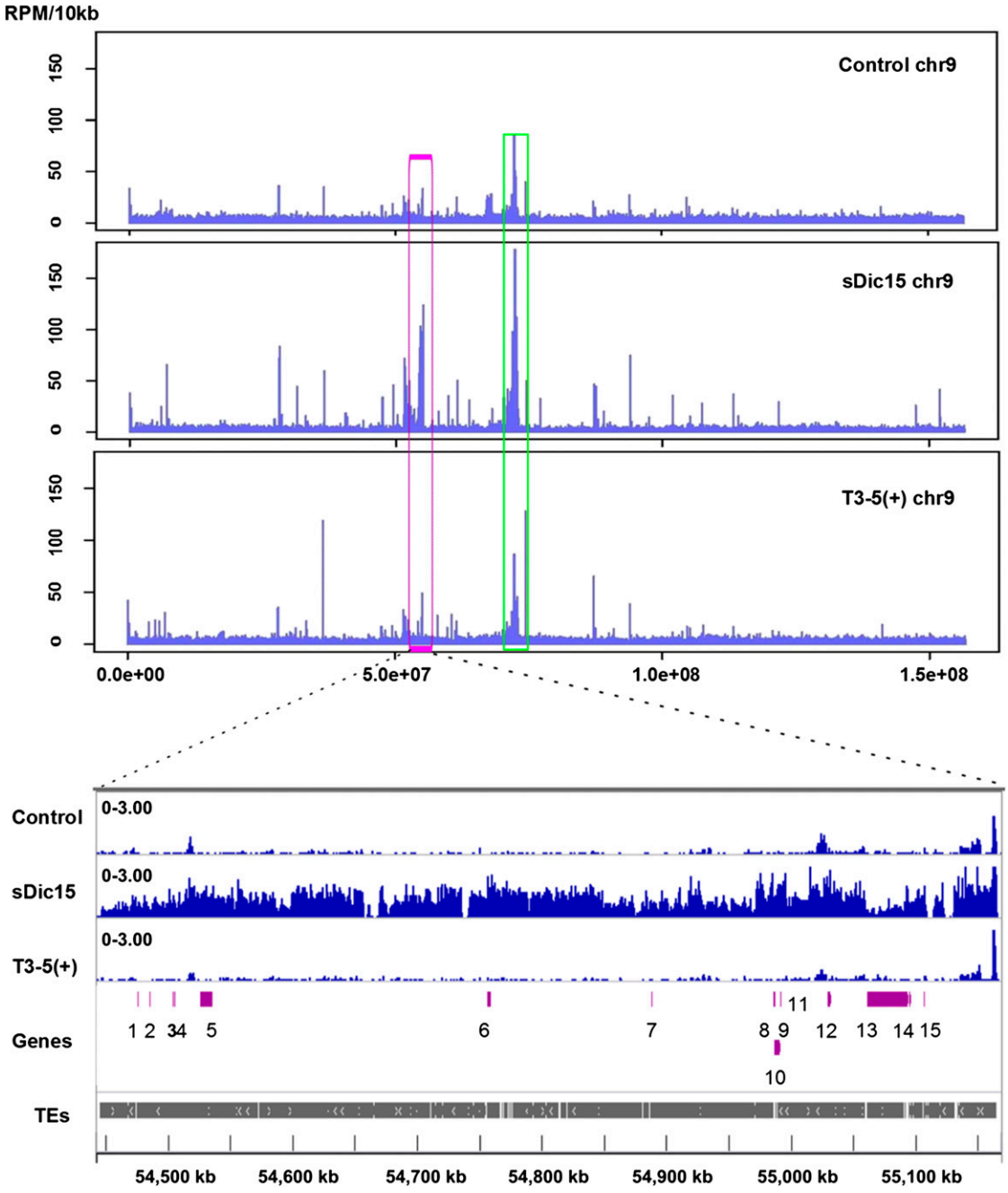

Figure 4. Mapping Reads Resulting from ChIP with CENH3 Antibodies along Maize Chromosome 9.

The mapping results of chromosome 9 in 10-kb windows in three samples are shown in the top half of the figure. The *x* axis marks the position on the chromosome, and the centromere region is indicated by the green box. The *y* axis indicates the number of reads mapping uniquely to each position, normalized as per million mapped reads. There is a significant peak in the sDic-15 sample, which is 723-kb long, spanning from 54,447,000 to 55,170,000 bp, and is marked with a magenta box comparing with the two control lines ["control" and "T3-5(+)"], which do not contain sDic-15. The bottom half of the figure is the detailed layout of the 723-kb region using IGV, a visualization tool. Four tracks represent the ChIP-Seq mapping results of "control" and "sDic-15," and the distribution of genes and transposable elements (TEs), which are represented using gray bars with white arrowheads that indicate the direction of each TE in this region.

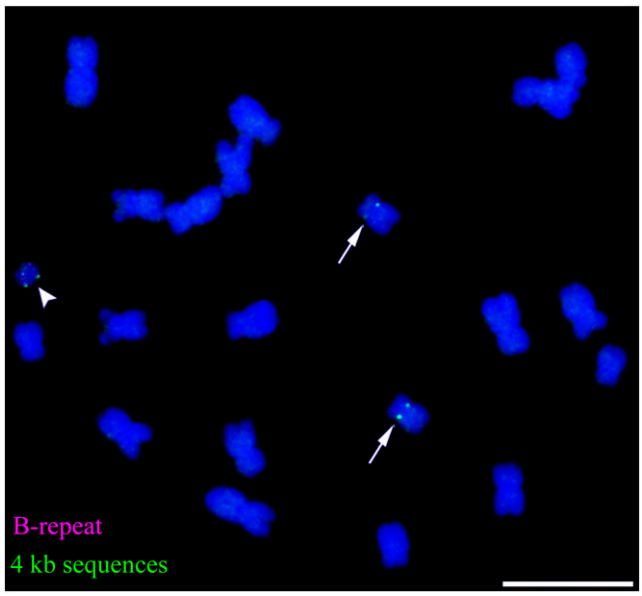

Figure 5. FISH Pattern of sDic-15 Using DNA Sequences from the 723-kb Region.

The ZmBs sequences are labeled in magenta, a 4-kb sequence from GRMZM2G057743, which is from the 723-kb region, is labeled in green. sDic-15 is marked by the arrowhead, and the arrows mark chromosome 9.

A 723-kb Sequence Was Used for New Centromere Formation

Based on the FISH results, we predicted that a nontypical centromeric sequence might be responsible for the maintenance of centromeric functions of sDic-15. In order to identify the DNA sequences associated with CENH3 in sDic-15, we performed a chromatin immunoprecipitation (ChIP) experiment using nuclei isolated from young leaf tissue of maize plants with or without sDic-15. The ChIP experiment was performed using maize-specific CENH3 antibodies. FISH experiments using the ChIPed DNA from sDic-15 plants as probes revealed that majority of the ChIPed sequences were CRM and CentC sequences, and they were highly enriched in the centromeric regions of maize chromosomes. However, compared with the small centromere of Dic-15, the neocentromere of sDic-15 had dramatically reduced CRM and lacked detectable CentC sequences (see Supplemental Figure 4 online).

The high-quality ChIPed DNAs were then sequenced using the Illumina platform, a total of 51.4 and 39.6 million 100-nucleotide-long reads were obtained for the control and sDic-15 samples, respectively. Most of the sequenced reads had perfect matches on the assembled B73 genome (Schnable et al., 2009).

Nonredundant reads with unique genomic mapping loci were used in the comparison of control plants and samples with sDic-15. A 723-kb-long region with significant read enrichment was found in the sDic-15 sample but not in the syntenic genomic region on 9S of the control sample, spanning from 54,447,000 to 55,170,000 bp (Figure 4). By labeling the 4-kb sequence from a single-copy gene GRMZM2G057743 within the 723-kb region as the FISH probe, we detected a signal that colocalized with ZmBs on the sDic-15 chromosome and chromosome 9 (Figure 5). This FISH result confirmed that the 723-kb region originated from chromosome 9 and also demonstrated that the 723-kb region was involved in the formation of the sDic-15 centromere.

As the progenitor of sDic-15, Dic-15 originated from the tug of war between the large centromere of TB-9Sb-Dp9 and the small centromere of T3-5(+). We wondered whether the 723-kb region was involved in centromere formation when T3-5(+) was created. We conducted ChIP-Seq for T3-5(+) using CENH3 antibodies. Mapping of ChIPed DNA sequences associated with CENH3 revealed no significant peak in the T3-5(+) sample (Figure 4). This finding demonstrated that the 723-kb region was not involved in centromere formation of the misdivision derivative T3-5(+). These data indicated that sDic-15 gained sequences from 9S when Dic-15 released a new dicentric chromosome through BFB cycles to rebuild a new CENH3 binding domain and finally formed a new centromere.

Two DNA Motifs Are Enriched in the 723-kb Region and in Known Centromeric Regions

The 723-kb CENH3 binding region of sDic-15 was significantly enriched for transposable elements (88.67% retrotransposons and 2.29% DNA transposons) compared with the average genome transposon content of maize ($P < 1e\text{-}4$, *t* test; see Supplemental Figure 5 online). In addition to transposons, the region also contained six protein-coding genes (see Supplemental Table 1 online).

In order to examine the sequence characteristics of the 723-kb region, we performed motif analysis (see Methods). We found two significantly enriched motifs (E-value < 1e-100) in this region (Figure 6). Aside from this 723-kb region, these motifs were more likely to be enriched in the centromeric region of normal chromosomes, especially motif-1, which was almost exclusively present in the 723-kb region as well as the centromere region of chromosome 9 (Figure 6B). To further confirm the distribution pattern of the two motifs, we screened the motifs on chromosomes 2 and 5, which contain the best-assembled centromere sequences of the maize chromosomes. We found that similar to the observation of chromosome 9, the centromeres of chromosomes 2 and 5 were also enriched for these two motifs (Figures 6C and 6D). These two motifs had no overlap with the CentC and CRM sequences. The enrichment of these motifs in the centromeric regions indicated that they might serve as a new type of characteristic sequences for centromere formation.

The 723-kb Region of sDic-15 Was Highly Methylated yet Maintained Active Gene Transcription

During the process of centromere formation, many factors have been identified to influence neocentromere formation (Choo, 2000). Chemical modifications such as methylation of DNA may be one of the ways to mark a region to form a new centromere. To determine the role that DNA methylation may have played in formation of the new centromere, we examined DNA methylation by bisulfite sequencing of the CENH3-ChIPed DNA sequences (Lister et al., 2009). By comparing the methylation level of the 723-kb region on sDic-15 with its syntenic parental region

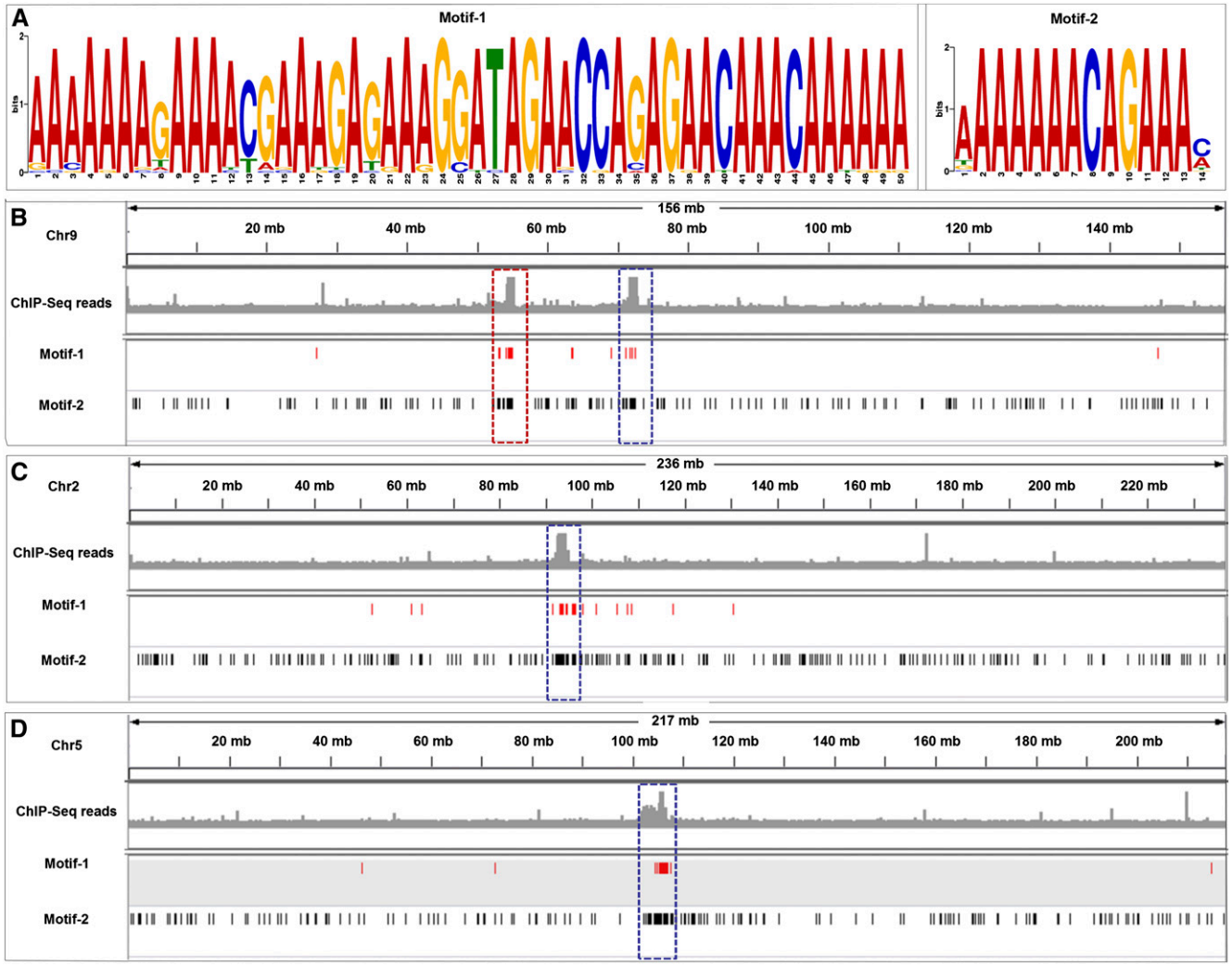

Figure 6. Two Significantly Enriched Motifs Found in the 723-kb Region and in the Centromeric Regions.

(A) The pattern of the two significant motifs found in the 723-kb region.
(B) The distribution of the two motifs on chromosome 9. The red box indicates the 723-kb region, and the blue box shows the suspected centromeric region of the chromosome 9 based on the ChIP-Seq data in the first lane. The red bars in lane 2 represent motif-1, and the black bars in lane 3 represent motif-2.
(C) The distribution of the two motifs on chromosome 2. The blue box indicates the centromeric region.
(D) The distribution of the two motifs on chromosome 5. The blue box indicates the centromeric region.

on 9S using the maize whole-genome bisulfite sequencing results as a reference (SRP011933), we found that the two regions had similar DNA methylation levels (Figure 7). The total amount of CpG methylation sites within the 9S region was 219,193, with a rate of ~0.317 [219193/(219193+473315)], which was nearly identical to the rate of CpG methylation for sDic-15 [154481/(154481+330405)≈0.319]. Then, we compared their methylation levels with known centromeric regions (*Cen2* and *Cen5*) as well as with randomly selected genomic regions (Rand-1000) and found that the methylation level of the 723-kb region, either on sDic-15 or on the normal 9S, was similar to the native centromeric regions and higher than most of the randomly selected genomic regions (Figure 7A). Similar trends were found when we analyzed the CG and CHG methylation contexts, but the CHH-type methylation was low in all sequence categories (Figures 7B to 7D). There was no significant DNA methylation level change within the 723-kb region after neocentromere formation.

Furthermore, we wondered whether the high DNA methylation level could affect the transcription of the genes in this region or whether the status of genes was altered with the process of new centromere formation. By comparing the transcription level of these protein-coding genes by quantitative RT-PCR using seedlings from maize with or without the sDic-15 chromosome, we found that three genes within the 723-kb region were actively expressed (Figure 8; see Supplemental Table 2 online).

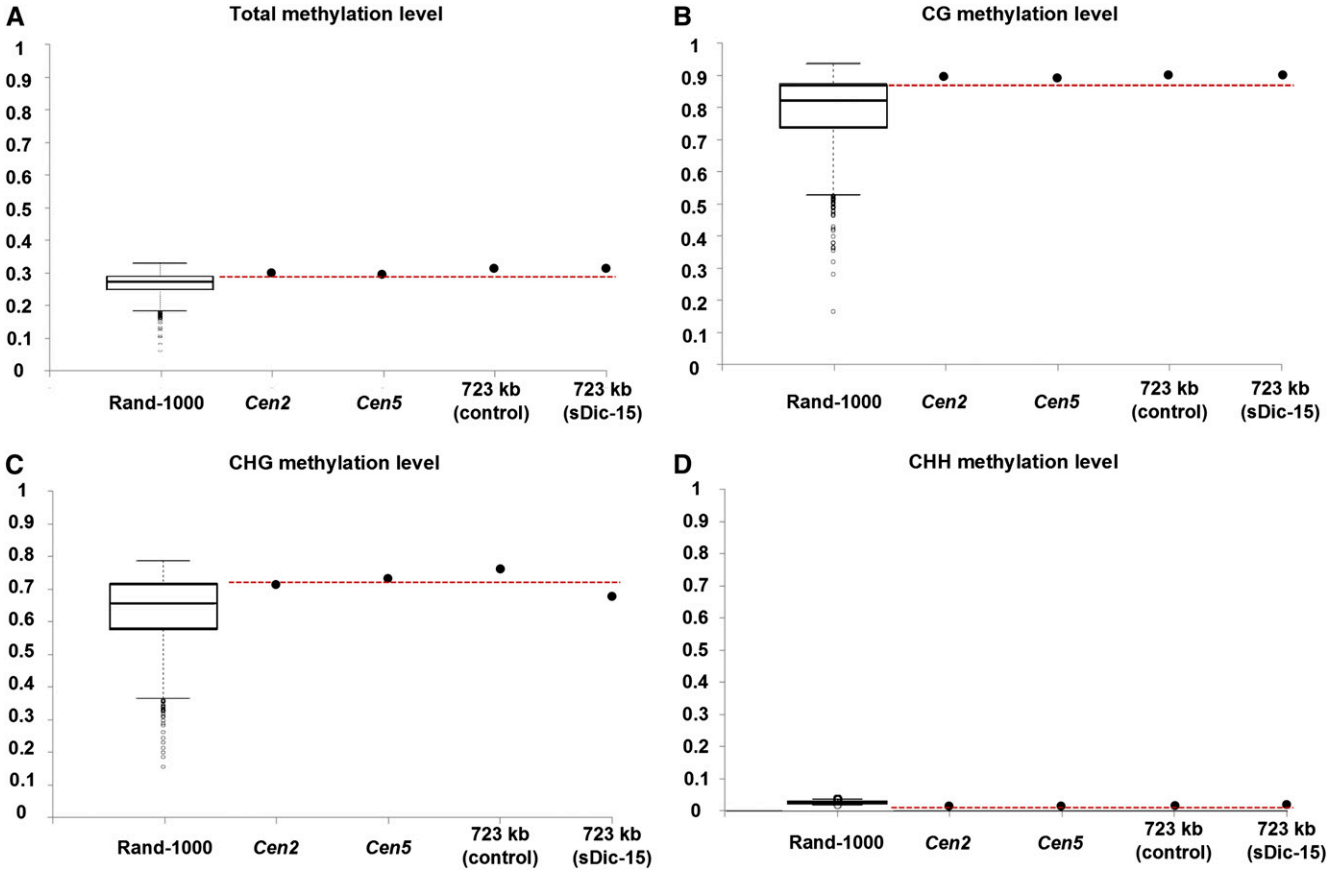

Figure 7. Comparison of DNA Methylation Levels between the 723-kb Region and Other Genomic Regions.

DNA methylation level, with 1 representing complete methylation, of the 723-kb region on sDic-15 and normal chromosome 9 compared with 1000 randomly selected genomic regions (Rand-1000) and two well-assembled centromere sequences, centromeres 2 and 5 (*Cen2* and *Cen5*). The Rand-1000 regions, each 350 kb in length, were selected randomly from chromosomes 1 to 10. The *Cen2* region was defined from 92,200,000 to 95,000,000 bp on chromosome 2, while *Cen5* from 102,000,000 to 106,550,000 bp on chromosome 5. The methylation level includes all CpG contexts of selected regions. The dashed line highlights the 75th percentile of the Rand-1000 regions. The *x* axis represents the different region selected from the chromosome, and the *y* axis represents the DNA methylation level.

(A) Methylation level including all CpG contexts of selected regions. The dashed line highlights the 75th percentile of the Rand-1000 regions.

(B) CG methylation level. The dashed line highlights the 75th percentile of the Rand-1000 regions.

(C) CHG methylation level. The dashed line highlights the 75th percentile of the Rand-1000 regions.

(D) CHH methylation level. The dashed line highlights the 25th percentile of the Rand-1000 regions.

[See online article for color version of this figure.]

DISCUSSION

The centromere is an essential region of a chromosome required for faithful chromosomal separation as it serves as the platform for kinetochore assembly. Alterations in centromere structure, sequence, and organization have been found in different organisms (Stimpson and Sullivan, 2010). In maize and wheat (*Triticum aestivum*), stable dicentric chromosomes have been found with one inactive centromere, even though the inactivated centromeres contain all the DNA elements considered to be necessary for functional centromeres (Han et al., 2006; Gao et al., 2011; Zhang et al., 2011; Fu et al., 2012). In barley, a chromosome that has lost detectable centromeric repeat sequences can be recovered with full centromere function (Nasuda et al., 2005), demonstrating that DNA sequences alone are insufficient for centromere formation in plants.

Thus, the question is raised whether a functional centromere or neocentromeres loses its original centromeric DNA elements or do they gain other sequences to wrap the nucleosome to form centromeric chromatin? In our analysis, the newly formed centromere seemed to be created by centromere breakage and fusion. In the process of its formation, most of the native centromeric repetitive sequences were lost and a unique genomic region (a 723-kb region of chromosome 9S) was recruited for CENH3 loading and seeding of a new centromere. We have shown that this newly formed centromere was from an inactive B centromere in Dic-15, and in the sDic15 it was functional and stably transmitted.

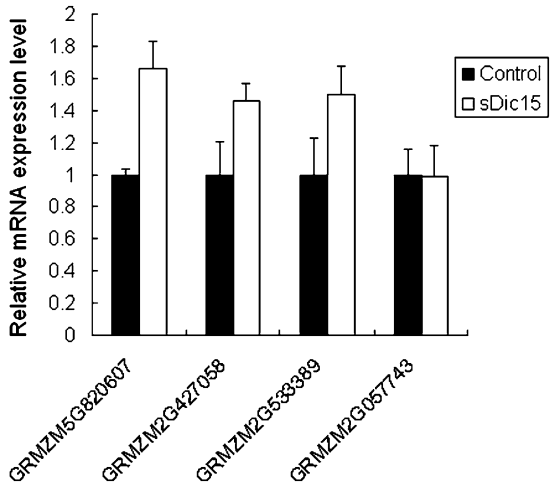

Figure 8. Transcription Analysis of Genes from the 723-kb Region.

Expression levels of GRMZM5G820607, GRMZM2G427058, GRMZM2G533389, and GRMZM2G057743 were examined by quantitative RT-PCR in seedlings without (control, $n = 4$) and with sDic-15 (sDic-15, $n = 4$). The columns and error bars represent the mean relative mRNA expression level and SD from three independent experiments.

When we screened the self-progeny of maize containing one copy of Dic-15, several newly formed centromeres with altered structures were found. Based on previous knowledge (Han et al., 2009), the intrachromosomal recombination of one copy of Dic-15 could form new chromosomes only with large or small centromeres or new dicentric chromosomes containing small centromeres (Han et al., 2009). We predicted that sDic-15 would be produced in a manner similar to a dicentric chromosome with two small centromeres, undergoing similar intrachromosomal recombination, bridge formation, and BFB cycle (see Supplemental Figure 6 online).

Although sDic15 came from the inactive centromere of Dic15, it has recovered centromeric function during formation. As our results show, the active centromere possesses all of the biochemical markers of a functional centromere, and it also can be transmitted to offspring through meiosis. So the newly formed centromere behaves just like a typical centromere.

Our results have shown that the typical centromere sequence is different from the conventional sequences. The 723-kb region has an abundance of transposable elements compared with other genomic sequences. Genes from the centromeric region of both human neocentromeres and plant centromeres can be actively transcribed (Saffery et al., 2003; Nagaki et al., 2004; May et al., 2005; Wong et al., 2006; Alonso et al., 2010; Wu et al., 2011; Gong et al., 2012). Most of the transcripts are associated with epigenetic modifications to maintain centromere identity. Recently, mammalian centromere researchers have suggested that active transcription is important for centromere function (Chan et al., 2012). We also observed active gene transcription within the new centromere region.

Satellite DNA and centromeric retrotransposons are major components of centromeres in plants and animals. However, the failure to identify conserved sequences among species and the emergence of neocentromeres without centromeric sequences raise the question of whether centromere function is sequence dependent (Dawe, 2005). In our analysis, we found that the newly formed centromere contains two significant DNA motifs, which are also enriched in the normal centromere regions in maize. This gives us a hint that these motifs may have a functional role in some aspect of the centromere formation.

In recent years, accumulating evidence has revealed the importance of epigenetic factors, including DNA methylation and histone modifications, in the establishment of centromeric chromatin (Choo, 2001; Schueler and Sullivan, 2006). In the newly formed sDic-15 centromere, the DNA methylation pattern of the parental region in 9S of the 723-kb CENH3 binding sequences had a higher DNA methylation level compared with other euchromatin regions (Figure 7). Such high methylation might confer this region with the ability to form a neocentromere, as a similar DNA methylation level was observed in native centromeres. It is possible that during the formation of sDic-15, the loss of CentC and reduction of CRM sequences in the centromere was caused by intrachromosomal recombination and BFB cycles, and the 723-kb region was subsequently recruited due to its high DNA methylation levels and proximity to the original centromere region.

In summary, we discovered a newly formed centromere in maize, sDic-15, which has almost completely lost CentC sequences and contains significantly reduced CRM and ZmBs sequences compared with canonical maize centromeres. We identified the centromere of sDic-15 as a neocentromere sequence and characterized its sequence characteristics and epigenetic features. Our results indicate that specific sequence characters plus a high DNA methylation level might be necessary for the formation of neocentromeres in maize.

Table 1. Statistics of the Mapping Results of ChIP-Seq Data

	Control (Pair-End 100 Nucleotides)		sDic15 (Pair-End 100 Nucleotides)		T3-5(+) (Pair-End 100 Nucleotides)	
Categories	Reads	Percentage	Reads	Percentage	Reads	Percentage
Total reads	51,453,782	100%	39,616,030	100%	37,775,428	100%
Total mapped	39,326,480	76.43%	23,467,483	59.24%	28,738,416	76.08%
Uniquely and nonredundantly mapped	33,992,185	66.06%	16,569,775	41.83%	24,602,730	65.13%

Table 2. Statistics of the Mapping Results of ChIP-Bisulfite-Seq of Control and sDic-15

	Control (Pair-End 100 Nucleotides)		Control (Single-End 100 Nucleotides)		sDic15 (Pair-End 100 Nucleotides)	
Categories	Reads	Percentage	Reads	Percentage	Reads	Percentage
Total reads	193,955,788	100%	282,553,129	100%	30,204,964	100%
Total mapped	155,258,280	80.05%	262,236,518	92.81%	22,023,510	72.91%
Uniquely and nonredundantly mapped	96,614,812	49.81%	47,415,179	16.78%	7,248,039	24.00%

METHODS

Plant Materials

Dicentric chromosome Dic-15 (Han et al., 2009), telocentric chromosome T3-5(+) (Kaszás and Birchler, 1998), and sDic-15 developed from the progeny of Dic-15 were kindly provide by James A. Birchler (University of Missouri-Columbia). These materials were planted in our greenhouse and were used for ChIP and cytogenetic analyses. sDic-15 seedlings were screened by FISH using probes combining the ZmBs (B chromosome-specific repeat) (Alfenito and Birchler, 1993), CentC (centromeric satellite repeats) (Ananiev et al., 1998), and CRM (Zhong et al., 2002) sequences as probes. Several seedlings (150) containing one copy of sDic-15 were screened. The candidate seedlings were transferred to the greenhouse for further analysis. Male inflorescences at the meiotic stage were fixed in ethanol:acetic acid (3:1, v/v) for 4 h at room temperature and transferred to 70% ethanol and stored at −20°C.

DNA Probe Preparation

For mitotic and meiotic analysis, ZmBs was labeled with Alexa Fluor-488-5-dUTP or Alexa Fluor-594-5-dUTP, and the CentC and CRM sequences were labeled with Alexa Fluor-488-5-dUTP, both by the nick translation method as previously described (Kato et al., 2004). The 4-kb sequence from the 723-kb region was labeled with Alexa Fluor-488-5-dUTP (for primer sequences, see Supplemental Table 3 online).

Immunolocalization in Mitotic and Meiotic Cells

Immunolocalization for mitosis and meiosis were performed as described (Han et al., 2009). The maize (*Zea mays*) CENH3 antibody is a rabbit polyclonal antiserum and was raised against the peptides C-RPGTVALREIRKYQKS STSATPERAAGTGGR-C (GL Biochem). A monoclonal rabbit antibody (04-817) raised against histone H3 phosphorylated at Ser-10 was obtained from Millipore. The phosphorylated H2A antibody was previously described (Dong and Han, 2012). The images were taken as a confocal z-stack (Zeiss LSM 710 NLO), processed with Photoshop CS 3.0.

Meiotic Analysis

Anthers of various meiotic stages were collected and slides were prepared as described (Gao et al., 1999). After hybridization, the slides were washed in 2× SSC (1× SSC is 0.15 M NaCl and 0.015 M sodium citrate) and mounted in mounting medium (containing 1.5 μg/mL 4′,6-diamidino-2-phenylindole; Vector Laboratories). The FISH images were recorded using an epifluorescence Olympus BX61 microscope equipped with a cooled charge-coupled device camera operated with MetaMorph software and processed with Photoshop CS 3.0.

ChIP and ChIP-Seq

ChIP was performed as previously described (Nagaki et al., 2003). Approximately 20 g of fresh leaf tissue was collected from young sDic-15 seedlings containing zero or one copy of sDic15 as well as T3-5(+). The chromatin was digested with micrococcal nuclease (Sigma-N3755) and was used for ChIP experiments using the maize CENH3 antibody. The ChIPed DNA was prepared for Illumina sequencing as described previously (Lister et al., 2009). The enriched DNA sample was sequenced using Illumina's sequencer to generate pair-ended 100-bp sequence reads. Sequences are available at the National Center for Biotechnology Information Gene Expression Omnibus database under accession number GSE43140 (Table 1).

Mapping of ChIP-Seq Reads to the Maize Genome

The reads from ChIP-Seq were mapped to the reference genome with BWA software (Li and Durbin, 2009), allowing at most three mismatches. Reads with unique genomic mapping locus were chosen for further analysis. Identical reads were merged to avoid redundancy caused by PCR reaction. The abundance of reads was calculated by reads per million values with 10-kb windows sliding along the examined genomic regions.

Quantitative RT-PCR

Total RNA was extracted using TRIzol reagent (Invitrogen) and possible contaminating DNA was removed using DNase I (New England Biolabs). cDNA synthesis was performed using Moloney Murine Leukemia Virus Reverse Transcriptase (Promega). Quantitative RT-PCR was performed using SYBR Green PCR Master Mix (Roche) on a Roche 480 LightCycle real-time PCR system (Roche 480) according to the manufacturer's instructions. cDNA equivalent to 20 ng RNA was amplified with 200 nM forward and reverse primers in a 20-μL reaction (for primer sequences, see Supplemental Table 4 online). Dissociation curves were performed to confirm specific amplifications without primer dimer formation. Sequencing experiments were also performed to confirm the identity of each RT-PCR product for the transcription assay of the genes in the 723-kb region. Calculations and statistics in the analysis were as follows.

The comparative C_T (for cycle threshold) method was used for data analysis in transcription detection assay and quantitative RT-PCR analysis (Schmittgen and Livak, 2008). The $\triangle C_T$ value was calculated as (C_T = C_T test gene − C_T internal control). Value of $\triangle\triangle C_T$ was calculated as ($\triangle\triangle C_T$ = $\triangle C_T$ test gene − $\triangle C_T$ experimental control) for fold change analysis of transcripts. Differences between experiment groups were compared by Student's *t* test. Significant differences were considered at $P < 0.05$. Data analysis was performed with SPSS 19.0 software (IBM).

ChIP-Bisulfite-Seq and DNA Methylation Analysis

Anti-CENH3 ChIP was performed using the same method as described above. The ChIPed DNA was then treated with sodium bisulfite before sequencing. Pair-end 100-bp reads were generated using the Illumina HiSeq2000 platform. Sequences are available at the National Center for Biotechnology Information Gene Expression Omnibus database with accession number GSM1057283 (Table 2).

BRAT software (Harris et al., 2010) was used to map the bisulfite sequencing reads to the reference genome. The weighted DNA

methylation level of selected regions was calculated using methods discussed (Schultz et al., 2012), using cytosines that were covered by at least two reads. Basically, the methods can be expressed in the following formula: $\sum_{i=1}^{n} C_i \Big/ \sum_{i=1}^{n} (C_i + T_i)$, where i represents the positions of cytosine meeting the coverage criteria, C represents reads supporting methylated cytosine, and T represents reads supporting unmethylated cytosine. Perl scripts and IGV tools (Robinson et al., 2011) were used in the process of data analysis.

Analysis of Motifs

Self-BLAST of the 723-kb region was conducted using BLASTn with "-F T" parameter. Low-complexity sequences were masked and extracted using perl script. Motifs were searched using MEME software (Bailey and Elkan, 1994) using the sequences of the 723-kb region as data set, with the following parameters: "-dna -minw 6 -maxw 50." The distribution pattern of the motifs on the whole genome was searched using FIMO software (Grant et al., 2011). Only sites with scores greater than zero were kept for further analysis. The final results were shown using IGV tools.

Accession Numbers

Sequence data from this article can be found in the Arabidopsis Genome Initiative or GenBank/EMBL databases under the following accession numbers: GSE43140, GSM1057283, and SRP011933.

Supplemental Data

The following materials are available in the online version of this article.

Supplemental Figure 1. FISH Detection Using Telomeres in sDic-15.

Supplemental Figure 2. CENH3 Detection for sDic-15 Meiotic Cells.

Supplemental Figure 3. FISH Detection the Fragment from the sDic-15.

Supplemental Figure 4. FISH Pattern Using ChIPed DNA as Probes.

Supplemental Figure 5. The TE and LTR Distribution on the 723 kb.

Supplemental Figure 6. The Model of Chromosome sDic-15 Formation.

Supplemental Table 1. Detailed Information on the 15 Genes Located in the 723-kb Region of sDic-15.

Supplemental Table 2. Expression of Six Protein-Coding Genes in the 723-kb Region of sDic-15.

Supplemental Table 3. FISH Probe Primers Used in This Study.

Supplemental Table 4. qRT-PCR Primers Used in This Study.

ACKNOWLEDGMENTS

We thank Jinsheng Lai (Chinese Agriculture of University) for the B73 seedling DNA methylation data. James A Birchler supplied the seeds and Ingo Schubert and Andreas Houben provided critical reading of the article. Ryan Douglas helped edit the revised article. This work is supported by the National Natural Science Foundation of China (31130033 and 31071083), the National Basic Research Program of China (973 Program 2011CB944601), and the Chinese Academy of Sciences (KSCX2-EW-R-0103).

AUTHOR CONTRIBUTIONS

F.H., Z.L., and X.-J.W designed the research. B.Z., Z.L., J.P., Y.L., S.F., X. G., Z.G., H.-J.W., Q.D., and J.L. performed the research. B.Z., Z.L., J.P., X.-J.W., and F.H. analyzed the data. F.H. and X.-J.W. wrote the article.

Received January 25, 2013; revised May 19, 2013; accepted May 23, 2013; published June 14, 2013.

REFERENCES

Alfenito, M.R., and Birchler, J.A. (1993). Molecular characterization of a maize B chromosome centric sequence. Genetics **135:** 589–597.

Alonso, A., Hasson, D., Cheung, F., and Warburton, P.E. (2010). A paucity of heterochromatin at functional human neocentromeres. Epigenetics Chromatin **3:** 6.

Ananiev, E.V., Phillips, R.L., and Rines, H.W. (1998). Chromosome-specific molecular organization of maize (*Zea mays* L.) centromeric regions. Proc. Natl. Acad. Sci. USA **95:** 13073–13078.

Bailey, T.L., and Elkan, C. (1994). Fitting a mixture model by expectation maximization to discover motifs in biopolymers. Proc. Int. Conf. on Intell. Syst. Mol. Biol. **2:** 28–36.

Birchler, J.A., Gao, Z., Sharma, A., Presting, G.G., and Han, F. (2011). Epigenetic aspects of centromere function in plants. Curr. Opin. Plant Biol. **14:** 217–222.

Birchler, J.A., and Han, F. (2009). Maize centromeres: Structure, function, epigenetics. Annu. Rev. Genet. **43:** 287–303.

Burrack, L.S., and Berman, J. (2012). Neocentromeres and epigenetically inherited features of centromeres. Chromosome Res. **20:** 607–619.

Chan, F.L., Marshall, O.J., Saffery, R., Kim, B.W., Earle, E., Choo, K.H.A., and Wong, L.H. (2012). Active transcription and essential role of RNA polymerase II at the centromere during mitosis. Proc. Natl. Acad. Sci. USA **109:** 1979–1984.

Choo, K.H.A. (2000). Centromerization. Trends Cell Biol. **10:** 182–188.

Choo, K.H.A. (2001). Domain organization at the centromere and neocentromere. Dev. Cell **1:** 165–177.

Cleveland, D.W., Mao, Y.H., and Sullivan, K.F. (2003). Centromeres and kinetochores: From epigenetics to mitotic checkpoint signaling. Cell **112:** 407–421.

Dawe, R.K. (2005). Centromere renewal and replacement in the plant kingdom. Proc. Natl. Acad. Sci. USA **102:** 11573–11574.

Dong, Q., and Han, F. (2012). Phosphorylation of histone H2A is associated with centromere function and maintenance in meiosis. Plant J. **71:** 800–809.

Fu, S., Gao, Z., Birchler, J., and Han, F. (2012). Dicentric chromosome formation and epigenetics of centromere formation in plants. J. Genet. Genomics **39:** 125–130.

Gao, Z., Fu, S., Dong, Q., Han, F., and Birchler, J.A. (2011). Inactivation of a centromere during the formation of a translocation in maize. Chromosome Res. **19:** 755–761.

Gao, Z., Han, F.P., He, M.Y., Ma, Y.Z., and Xin, Z.Y. (1999). Characterization of genome and chromosomes in octoploid wheat-wheatgrass amphiploid Zhong 2 using fluorescence in situ hybridization and chromosome pairing analysis. Acta Bot. Sin. **41:** 25–28.

Gong, Z., Wu, Y., Koblízková, A., Torres, G.A., Wang, K., Iovene, M., Neumann, P., Zhang, W., Novák, P., Buell, C.R., Macas, J., and Jiang, J. (2012). Repeatless and repeat-based centromeres in potato: Implications for centromere evolution. Plant Cell **24:** 3559–3574.

Grant, C.E., Bailey, T.L., and Noble, W.S. (2011). FIMO: Scanning for occurrences of a given motif. Bioinformatics **27:** 1017–1018.

Han, F., Gao, Z., and Birchler, J.A. (2009). Reactivation of an inactive centromere reveals epigenetic and structural components for centromere specification in maize. Plant Cell **21:** 1929–1939.

Han, F., Lamb, J.C., Yu, W., Gao, Z., and Birchler, J.A. (2007). Centromere function and nondisjunction are independent components of the maize B chromosome accumulation mechanism. Plant Cell **19:** 524–533.

Han, F.P., Lamb, J.C., and Birchler, J.A. (2006). High frequency of centromere inactivation resulting in stable dicentric chromosomes of maize. Proc. Natl. Acad. Sci. USA **103:** 3238–3243.

Harris, E.Y., Ponts, N., Levchuk, A., Roch, K.L., and Lonardi, S. (2010). BRAT: Bisulfite-treated reads analysis tool. Bioinformatics **26:** 572–573.

Henikoff, S., Ahmad, K., and Malik, H.S. (2001). The centromere paradox: Stable inheritance with rapidly evolving DNA. Science **293:** 1098–1102.

Henikoff, S., and Furuyama, T. (2010). Epigenetic inheritance of centromeres. Cold Spring Harb. Symp. Quant. Biol. **75:** 51–60.

Houben, A., Wako, T., Furushima-Shimogawara, R., Presting, G., Kunzel, G., Schubert, I., and Fukui, K. (1999). Short communication: The cell cycle dependent phosphorylation of histone H3 is correlated with the condensation of plant mitotic chromosomes. Plant J. **18:** 675–679.

Jiang, J.M., Birchler, J.A., Parrott, W.A., and Dawe, R.K. (2003). A molecular view of plant centromeres. Trends Plant Sci. **8:** 570–575.

Jiang, J.M., Nasuda, S., Dong, F.G., Scherrer, C.W., Woo, S.S., Wing, R.A., Gill, B.S., and Ward, D.C. (1996). A conserved repetitive DNA element located in the centromeres of cereal chromosomes. Proc. Natl. Acad. Sci. USA **93:** 14210–14213.

Jin, W.W., Lamb, J.C., Vega, J.M., Dawe, R.K., Birchler, J.A., and Jiang, J. (2005). Molecular and functional dissection of the maize B chromosome centromere. Plant Cell **17:** 1412–1423.

Kaszás, E., and Birchler, J.A. (1998). Meiotic transmission rates correlate with physical features of rearranged centromeres in maize. Genetics **150:** 1683–1692.

Kato, A., Lamb, J.C., and Birchler, J.A. (2004). Chromosome painting using repetitive DNA sequences as probes for somatic chromosome identification in maize. Proc. Natl. Acad. Sci. USA **101:** 13554–13559.

Lamb, J.C., Kato, A., and Birchler, J.A. (2005). Sequences associated with A chromosome centromeres are present throughout the maize B chromosome. Chromosoma **113:** 337–349.

Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics **25:** 1754–1760.

Lister, R., et al. (2009). Human DNA methylomes at base resolution show widespread epigenomic differences. Nature **462:** 315–322.

Ma, J., Wing, R.A., Bennetzen, J.L., and Jackson, S.A. (2007). Plant centromere organization: A dynamic structure with conserved functions. Trends Genet. **23:** 134–139.

Marshall, O.J., Chueh, A.C., Wong, L.H., and Choo, K.H.A. (2008). Neocentromeres: New insights into centromere structure, disease development, and karyotype evolution. Am. J. Hum. Genet. **82:** 261–282.

May, B.P., Lippman, Z.B., Fang, Y.D., Spector, D.L., and Martienssen, R.A. (2005). Differential regulation of strand-specific transcripts from *Arabidopsis* centromeric satellite repeats. PLoS Genet. **1:** e79.

McClintock, B. (1939). The behavior in successive nuclear divisions of a chromosome broken at meiosis. Proc. Natl. Acad. Sci. USA **25:** 405–416.

McClintock, B. (1941). The stability of broken ends of chromosomes in zea mays. Genetics **26:** 234–282.

Nagaki, K., Cheng, Z.K., Ouyang, S., Talbert, P.B., Kim, M., Jones, K.M., Henikoff, S., Buell, C.R., and Jiang, J.M. (2004). Sequencing of a rice centromere uncovers active genes. Nat. Genet. **36:** 138–145.

Nagaki, K., Talbert, P.B., Zhong, C.X., Dawe, R.K., Henikoff, S., and Jiang, J.M. (2003). Chromatin immunoprecipitation reveals that the 180-bp satellite repeat is the key functional DNA element of *Arabidopsis thaliana* centromeres. Genetics **163:** 1221–1225.

Nasuda, S., Hudakova, S., Schubert, I., Houben, A., and Endo, T.R. (2005). Stable barley chromosomes without centromeric repeats. Proc. Natl. Acad. Sci. USA **102:** 9842–9847.

Robinson, J.T., Thorvaldsdóttir, H., Winckler, W., Guttman, M., Lander, E.S., Getz, G., and Mesirov, J.P. (2011). Integrative genomics viewer. Nat. Biotechnol. **29:** 24–26.

Saffery, R., Sumer, H., Hassan, S., Wong, L.H., Craig, J.M., Todokoro, K., Anderson, M., Stafford, A., and Choo, K.H.A. (2003). Transcription within a functional human centromere. Mol. Cell **12:** 509–516.

Schmittgen, T.D., and Livak, K.J. (2008). Analyzing real-time PCR data by the comparative C(T) method. Nat. Protoc. **3:** 1101–1108.

Schnable, P.S., et al. (2009). The B73 maize genome: Complexity, diversity, and dynamics. Science **326:** 1112–1115.

Schueler, M.G., and Sullivan, B.A. (2006). Structural and functional dynamics of human centromeric chromatin. Annu. Rev. Genomics Hum. Genet. **7:** 301–313.

Schultz, M.D., Schmitz, R.J., and Ecker, J.R. (2012). 'Leveling' the playing field for analyses of single-base resolution DNA methylomes. Trends Genet. **28:** 583–585.

Sharma, A., Schneider, K.L., and Presting, G.G. (2008). Sustained retrotransposition is mediated by nucleotide deletions and interelement recombinations. Proc. Natl. Acad. Sci. USA **105:** 15470–15474.

Stimpson, K.M., and Sullivan, B.A. (2010). Epigenomics of centromere assembly and function. Curr. Opin. Cell Biol. **22:** 772–780.

Topp, C.N., Okagaki, R.J., Melo, J.R., Kynast, R.G., Phillips, R.L., and Dawe, R.K. (2009). Identification of a maize neocentromere in an oat-maize addition line. Cytogenet. Genome Res. **124:** 228–238.

Wong, N.C., Wong, L.H., Quach, J.M., Canham, P., Craig, J.M., Song, J.Z., Clark, S.J., and Choo, K.H.A. (2006). Permissive transcriptional activity at the centromere through pockets of DNA hypomethylation. PLoS Genet. **2:** e17.

Wu, Y., Kikuchi, S., Yan, H., Zhang, W., Rosenbaum, H., Iniguez, A. L., and Jiang, J. (2011). Euchromatic subdomains in rice centromeres are associated with genes and transcription. Plant Cell **23:** 4054–4064.

Zhang, M., Zhao, H., Xie, S., Chen, J., Xu, Y., Wang, K., Zhao, H., Guan, H., Hu, X., Jiao, Y., Song, W., and Lai, J. (2011). Extensive, clustered parental imprinting of protein-coding and noncoding RNAs in developing maize endosperm. Proc. Natl. Acad. Sci. USA **108:** 20042–20047.

Zhang, W., Friebe, B., Gill, B.S., and Jiang, J. (2010). Centromere inactivation and epigenetic modifications of a plant chromosome with three functional centromeres. Chromosoma **119:** 553–563.

Zhong, C.X., Marshall, J.B., Topp, C., Mroczek, R., Kato, A., Nagaki, K., Birchler, J.A., Jiang, J.M., and Dawe, R.K. (2002). Centromeric retroelements and satellites interact with maize kinetochore protein CENH3. Plant Cell **14:** 2825–2836.