Sequential de novo centromere formation and inactivation on a chromosomal fragment in maize

Yalin Liu*,b,1, Handong Su*a,b,1, Junling Pang*a,b,2, Zhi Gao4, Xiujie Wang3, James A. Birchlerb,d,2, and Fangpu Han*a,2

*State Key Laboratory of Plant Cell and Chromosome Engineering, State Key Laboratory of Plant Genomics, Institute of Genetics and Developmental Biology, Chinese Academy of Sciences, Beijing 100101, China; bDivision of Biological Sciences, University of Missouri, Columbia, MO 65211-7400; and cUniversity of Chinese Academy of Sciences, Beijing 100049, China

Edited by Ronald L. Phillips, University of Minnesota, St. Paul, MN, and approved February 4, 2015 (received for review September 22, 2014)

The ability of centromeres to alternate between active and inactive states indicates significant epigenetic aspects controlling centromere assembly and function. In maize (Zea mays), misdivision of the B chromosome centromere on a translocation with the short arm of chromosome 9 (TB-9Sb) can produce many variants with varying centromere sizes and centromeric DNA sequences. In such derivatives of TB-9Sb, we found a de novo centromere on chromosome derivative 3-3, which has no canonical centromeric repeat sequences. This centromere is derived from a 288-kb region on the short arm of chromosome 9, and is 19 megabases (Mb) removed from the translocation breakpoint of chromosome 9 in TB-9Sb. The functional B centromere in progenitor telo2-2 is deleted from derivative 3-3, but some B-repeat sequences remain. The de novo centromere of derivative 3-3 becomes inactive in three further derivatives with new centromeres being formed elsewhere on each chromosome. Our results suggest that de novo centromere initiation is quite common and can persist on chromosomal fragments without a canonical centromere. However, we hypothesize that when de novo centromeres are initiated in opposition to a larger normal centromere, they are cleared from the chromosome by inactivation, thus maintaining karyotype integrity.

centromere misdivision | epigenetics | de novo centromere | centromere inactivation

The centromere is one of the most complex regions on the chromosome, and complete DNA sequencing through the centromeric region is difficult to obtain due to their highly repetitive nature. Centromere sizes, defined by CENP-A/CENH3 binding, range from 125 bp in Saccharomyces cerevisiae to 500–1,500 kb in humans and mice (11). In plants, centromere sizes can range to several megabases (Mb) with many repetitive transposable elements, which makes it difficult to study centromere structure and function. For example, the sizes of centromere 2 and 5 in maize are roughly 2 and 7 Mb, respectively.

Previous work sought misdivision derivatives of the B chromosome centromere using a translocation between the supernumerary chromosome and the short arm of chromosome 9 (9S) to reduce the size of the centromere for functional studies (13–15). B chromosomes are extra chromosomes that have been found in many plants, animals, and fungi. In maize, a reciprocal translocation between a B chromosome and the short arm of chromosome 9 produced two chromosomes referred to as B-9 and 9-B (13), together referred to as TB-9Sb. Chromosome 9-B contains the long arm of the B chromosome and most of chromosome 9, including its centromere. Correspondingly, chromosome B-9 contains part of the short arm of chromosome 9 and the other part of the B chromosome with the active B centromere. The translocation breakpoint is near WxI, which is located on 9-B (16). The B centromere of B-9 can undergo misdivision during meiosis, producing many derivatives (14). The first misdivision derivative was a pseudoisochromosome, and subsequently, many telocentric chromosomes and isochromosomes were derived by additional misdivisions (13, 17). Misdivision events can be recognized in crosses of TB-9Sb onto a tester via a fusion-breakage cycle recognized by the behavior of the C1 color marker on the B-9 chromosome. The cycle continues during endosperm development to produce a mosaic phenotype.
but is “healed” in the embryo, which when grown and analyzed cytologically will reveal the nature of the new chromosomes formed (13, 17). This type of screen was used to assemble a large collection of misdivisions to examine the structural features of the B centromere (18). Centromere sizes of these derivatives were changed and progressively reduced. In these previous studies, molecular analysis of centromere size relied on studying the B centromere-specific DNA repeat before the maize centromere elements, CentC and CRM, were known. The B-specific repeat allows this centromere to be studied against the background of the other centromeres; it surrounds and is interspersed within the active core of the B centromere (15, 19).

New functional centromeres formed at ectopic locations rather than native centromeric regions on the chromosomes are called de novo centromeres. Many de novo centromeres have been found in human patients and other organisms (20, 21). There are reports of de novo centromeres in plants, such as barley (Hordeum vulgare) (22), oat (Avena sativa)-maize addition lines (23), and maize (24, 25). The conditions for de novo centromere formation remain unclear (26). Recent research revealed that many de novo centromeres prefer to form near native centromeric regions or in the heterochromatic regions, such as the pericentromere and telomere (27, 28). There are also de novo centromeres in human formed far from native centromeres (20). We have previously described two de novo centromeres in maize: one is near the position of the native centromere (25) and the other is distal to the site of the corresponding native centromere (24). Specific chromatin environments may be required for centromere formation, but the major elements are as yet unknown.

DNA sequence alone is insufficient to direct centromere formation, and dicentric chromosomes containing two centromeres are good examples. To be stable, structurally dicentric chromosomes must have one inactive and one active centromere; otherwise, two active centromeres will lead to chromosome breakage during cell division. In maize, many dicentric chromosomes have been reported from B-A translocation chromosome derivatives (29). Dicentric chromosomes can be produced through the process of the chromosome type breakage-fusion-bridge (BFB) cycle, and the inactive centromeres can be reactivated by intrachromosomal recombination (30). The DNA sequences of the active and inactive centromeres of dicentric chromosomes are essentially identical, but the centromere activity states are completely different. We screened several misdivision derivatives using FISH probes specific to maize centromere sequences, CentC and CRM, as well as probes specific to the B centromere repeat sequence (B-repeat) to gain further insight into the nature of the centromeres in this collection. We discovered that one such chromosome, derivative 3-3, lacks detectable CentC and CRM signals, but still has a functional centromere that is not associated with the B-repeat sequence. The results of chromatin immunoprecipitation sequencing (ChIP-seq) using maize CENH3 antibody revealed that a 288-kb region on 9S is involved in the de novo centromere formation. The functional B centromere of progenitor telo2-2 is deleted from derivative 3-3. Further, new derivatives of derivative 3-3 had been selected (31) but there was no change in the B-specific repeat patterns. Here, we found that the de novo centromere of 3-3 has become inactive in all of its derivatives, and in each case a shift to a new de novo centromere position occurred; one of these contains only a 200-kb CENH3 binding region within 9S. The other two are apparently in B chromosome sequences. Thus, sequential de novo centromere formation and exchange of centromere activity occurred in chromosome 3-3 and its derivatives, providing new insight into centromere formation and maintenance.

These results help formulate the nature of de novo centromere formation. In all of the examples now documented in maize, the size range is within a few hundred kilobases. In contrast, normal maize centromeres, as noted above, are typically several megabases. The regular occurrence of de novo centromeres found here and previously (24, 25) indicates that they are capable of being formed regularly on chromosomal fragments that are structurally acentric; however, they do not persist in normal chromosomes. The reason might reside in the previous observation in maize (30) and wheat (32) that in functional dicentrics the smaller centromere becomes inactive in a tug of war between large and small. However, in the absence of a normal centromere, the present work illustrates that de novo centromeres can persist. Thus, in normal chromosomes, if a de novo is initiated, it will be as quickly inactivated in opposition to the much larger preexisting centromere; the chromosome will not be affected, and will seldom change structure over evolutionary time despite such a high rate of de novo formation. This hypothesis also suggests that a selective pressure will be placed on the normal centromeres to expand to a size that can regularly inactivate de novo centromeres based on their initial size at formation.

Results and Discussion

Derivative 3-3 Contains a Functional de Novo Centromere. The B centromere of chromosome B-9 of TB-9Sb undergoes misdivision during meiosis, generating progeny with altered centromere sizes and centromeric DNA sequences (14, 15, 18). Previous work characterized one such derivative, previously called telo3-3 and now called derivative 3-3, which by fiber FISH analysis showed B-repeat signals that, unlike others in the collection, were not associated with CentC and CRM (15). To examine this chromosome further, root tip metaphase spreads were performed but no detectable FISH signals of maize centromere sequences CentC and CRM were found (Fig. 1A and B). There are B-repeat signals at both ends of the chromosome, but these signals do not coincide with CENH3 localization (Fig. 1C). Telomere signals are detected at the distal ends of derivative 3-3 (Fig. S1A), indicating that this chromosome is linear. There are also 180-bp heterochromatic knob signals on both ends of the derivative 3-3 chromosome (Fig. S1B).

To confirm whether derivative 3-3 has a functional centromere, we used antibodies against CENH3 and H2A phosphorylation at Thr-133 (7) to characterize centromere activity. The results of immunostaining show that both of these markers of active centromeres are present (Fig. 1C and D), and the B-repeat signal does not colocalize with CENH3 (Fig. 1C). Thus, derivative 3-3 contains a de novo centromere without classic

![Fig. 1. Cytological analysis of derivative 3-3. (A) Somatic metaphase chromosomes of derivative 3-3, with probes for CentC (green) and B-repeat (red). (B) Somatic metaphase chromosomes of derivative 3-3, with probes for CRM (green) and B-repeat (red). (C) Immuno-FISH of derivative 3-3, with CENH3 antibody (green) and B-repeat probe (red). (D) Immunostaining of derivative 3-3, with antibody against H2A phosphorylation at Thr-133 (red). Arrow indicates chromosome derivative 3-3. (Insets) Higher-magnification view of chromosome derivative 3-3. Blue indicates chromosomes counterstained with DAPI. (Scale bar: 10 μm.)](https://www.pnas.org/content/114/18/4281)
centromeric repeats, and the B-repeat sequence is not associated with the new centromere formation.

Derivative 3-3 is descended from a telocentric chromosome called telo2-2. Telo2-2 has an active B centromere as evidenced by the presence of phosphorylated H2A at the usual terminal position with obvious CentC, CRM, and B-repeat signals at the typical B chromosome centromeric region (Fig. S2). There is no 180-bp knob signal at the B centromere of telo2-2, whereas there is a large block of 180-bp knob on the end of telo2-2 (Fig. S3). To produce derivative 3-3, it is possible that telo2-2 might have first formed a ring chromosome, which regularly occurs from misdivision (18), by fusing the broken B centromere with the terminal knob region to bring them together, because the current structure of derivative 3-3 has this knob/B-repeat juxtaposition at both ends of the chromosome. The progression from telo2-2 to derivative 3-3 deleted the remaining B centromere sequences in the latter.

**De Novo Centromere of Derivative 3-3 Is Derived from a 288-kb Region on 9S.** Because there are no detectable classical maize centromeric repeats on chromosome derivative 3-3 using FISH, we considered that some other DNA sequences were involved in de novo centromere formation. To identify new DNA sequences associated with CENH3, we performed a ChIP-seq experiment using a maize-specific CENH3 antibody. ChIP was conducted using nuclei isolated from young leaves of plants with derivative 3-3, and with TB-9Sb as a control. The quality of ChIP DNA was checked by FISH using probes prepared from ChIP DNA. The results showed that FISH signals are predominantly located in the centromeric regions, indicating that most of the ChIP DNA was obtained from centromeres (Fig. S4).

ChIP DNA from derivative 3-3 and TB-9Sb were sequenced by the Illumina Hiseq2000 platform to generate pair-ended 100-bp reads (Table S1), and then the raw reads were mapped to the

---

Fig. 2. ChIP-seq mapping results of derivative 3-3 with CENH3 antibody along maize chromosome 9. The mapping results of control and derivative 3-3 as well as telo2-2 in 10-kb windows on chromosome 9 are shown (Upper). The x axis represents positions on chromosome 9. The y axis represents numbers of reads mapping on each position uniquely, which is normalized as per-million mapped reads. The green box indicates the region of centromere 9. The red box shows the new centromere region of derivative 3-3 with a significant peak, which is 288 kb in length. There is no peak in telo2-2 in the same region. The 288-kb region spans from 4,300,000 to 4,588,000 bp. (Lower) Detailed layout of the 288-kb region using the Integrative Genomic Viewer (50). The upper three tracks represent ChIP-seq mapping results of control and telo2-2 as well as derivative 3-3. The lower two tracks show the distribution of genes and transposable elements.

maize B73 RefGen v3 using BWA software (33). According to mapping results of ChIP-seq data, a 288-kb peak on 9S was detected in derivative 3-3 compared with the control (Fig. 2), implying that this site could be the CENH3 binding region. The 288-kb region spans from 4,300,000 to 4,588,000 bp on 9S, which is 19 Mb removed from the translocation breakpoint on chromosome 9 of TB-9Sb. To further confirm the ChIP-seq results, we labeled a 5-kb single-copy DNA sequence from the 288-kb region. Somatic metaphase chromosomes of derivative 3-3 (Insets) Higher-magnification view of chromosome derivative 3-3. Blue indicates chromosomes counterstained with DAPI. (Scale bar: 10 μm.)

Sequence analysis of the B73 reference genome revealed that there is no copy of CentC or CRM sequence inside the 288-kb region in which the de novo centromere of derivative 3-3 is derived; although one cannot rule out their presence in derivative 3-3, this comparison reveals there are no fixed invariant centromere elements at this site. The percentage of LTR retrotransposons is 62.53% according to the Repeat Masker (www.repeatmasker.org) results, with 28.52% Ty1-copia and 33.51% Ty3-gypsy retrotransposons. Compared with native centromere 2 and 5, the percentage of LTR retrotransposons is higher, being 84%, and the percentage of Ty3-gypsy retrotransposons is about eight times for Ty1-copia. The GC level of this region and the native centromere is almost the same, with 44.97% in the 288-kb region and 44.85% in centromere 5 (45.92% in centromere 2).

The sequence composition of the 288-kb region is different from native centromeres in B73 (Table S3).

**Derivative 3-3 Behavior in Meiosis.** Previous work showed that derivative 3-3 can be lost during both mitosis and meiosis (18, 31). Albeit at a low rate (13%), derivative 3-3 can transmit through meiosis (18). Because this de novo centromere represents such a small functional centromere, its meiotic behavior is of interest. Examination of meiocytes showed that the derivative 3-3 chromosome does not pair with normal chromosome 9 (Fig. 4 A and B). Sister chromatids separated early in meiotic anaphase I, forming a biorientation (Fig. 4 C and D) as is routinely found with smaller chromosomes in maize (34). Thus, despite a high frequency of loss, the behavior of this de novo centromere is typical of others on small chromosomes.

**Subsequent Derivatives of Chromosome 3-3.** Previous work had sought further misdivision derivatives of chromosome derivative 3-3 by selecting for exceptional kernels with the C1 anthocyanin color marker showing evidence of a fusion-breakage cycle as is used to recover such chromosomes (31). Analysis of the B-repeat sequences showed no detectable change relative to the progenitor derivative 3-3 and, at the time, these chromosomes were considered false positives for misdivision. These subsequent derivatives were reexamined, with the realization that derivative 3-3 has a de novo centromere. In these three further misdivision chromosomes, we found no obvious CentC and CRM signals detected by FISH as expected from their derivation from chromosome 3-3. Derivative 3-3-6 has only a single nearly terminal B-repeat signal in contrast to 3-3 (Fig. 5 A and B). Chromosome derivative 3-3-7 has a B-repeat signal internally on a chromosome arm that is clearly not in the primary constriction (Fig. 5 D and E). Derivative 3-3-11 is a chromosome with B-repeat signals at two locations (Fig. 5 G and H). These three chromosomes were apparently formed following misdivision of the de novo
centromere in derivative 3-3. If misdivision of the de novo centromere occurred, it might have initiated a series of chromosomal rearrangements that gave rise to the diverse structures found in these descendents.

To address centromere activities of these derivatives, we conducted immunostaining using an antibody against H2A phosphorylation at Thr133. The results show that derivative 3-3-6 has the signal of antibody against H2A phosphorylation at Thr-133 near the B-repeat signal (Fig. 5C). In chromosome derivative 3-3-7, the B-repeat signal is clearly distinct from the active centromere region (Fig. 5F). Chromosome derivative 3-3-11 has one B-repeat signal that is near the signal of antibody against H2A phosphorylation at Thr-133 but not coincident (see below).

The De Novo Centromere of Derivative 3-3 Is Inactive in Its Descendents.
To test whether the 288-kb de novo centromere of derivative 3-3 is still active in its own derivatives, we performed FISH experiments using probe 1, described above, to represent the centromeric position of 3-3. Based on the FISH results, the signal of probe 1 is not detectable in the active centromere regions of any of the three subsequent derivatives. In derivative 3-3-6, the signal from probe 1 is on the opposite end of the chromosome from the active centromere (compare Figs. 5C and 6A). There are two sites of probe 1 hybridization signal on one chromosome arm of derivative 3-3-7 (compare Figs. 5F and 6B). There are also two positions of probe 1 hybridization signal on the chromosome arm of derivative 3-3-11, but there are no signals present at the position of the active cen-
The de novo centromere of derivative 3-3 is inactivated in all three of its derivatives. During misdivision, chromosomes experience large rearrangements as well as insertion, deletion, and duplication of some chromosomal fragments. Such chromosomal rearrangements may cause changes in centromere location and centromeric chromatin states that would be related to centromere activity. To determine further the inactivation of the 3-3 centromere in the derivatives, we designed probes flanking the original 288-kb centromeric region of 3-3, referred to as probe L (left side of 3-3 centromere) and probe R (right side of 3-3 centromere). The sequence used to produce probe L was within 1 Mb of sequence to the left side of the 288-kb region (Table S4). For probe R, first we designed several probes according to sequences located within 1, 1.5, 2.5, and 3.5 Mb to the right side of 288 kb, respectively. We found that all these probes have no signals on 3-3 and its derivatives, but are present on the normal chromosome 9. These results indicate that 3-3 has lost the corresponding 3.5 Mb of sequence of the progenitor chromosome 9. Thus, probe R was prepared from sequences 7 Mb distant from the 288-kb centromeric region (Table S4).

The FISH signals of probe L and R are close in 3-3 and derivative 3-3-6 (Fig. S6 A and B), indicating that the location of the 3-3 centromere has no significant rearrangement in 3-3-6. However, in derivative 3-3-7, probe L and probe R are located on different chromosomal arms (Fig. S6C), indicating significant change occurred in 3-3-7. For derivative 3-3-11, there is no detectable signal of probe R (Fig. S6D), which suggests that 3-3-11 is differently rearranged. Despite these rearrangements, probe 1 from within the 3-3 de novo centromere is still present in all three derivatives, suggesting that the de novo 3-3 centromere is inactivated in the derivatives and not deleted.

De Novo Centromere Formation Occurred in the New Derivatives of 3-3.

To identify DNA sequences associated with CENH3 in the subsequent derivatives, we performed anti-CENH3 ChIP-seq using young leaves of these new derivatives, again using TB-9Sb as a control (Table S1). The translocation breakpoint of TB-9Sb is near Wx1 (16), and the position of Wx1 on 9S is from 23,256,000 to 23,260,000 bp. After mapping the sequenced reads to the B73 reference, we sought obvious CENH3 binding peaks along 9S from 0 to 23 M. Compared with the control, a CENH3 binding peak spanning from 20,210,000 to 20,410,000 bp on 9S in derivative 3-3-11 was found (Fig. S7). We then performed FISH using a 7-kb single-copy sequence from this 200-kb peak region as a probe (probe 2; Table S2) to confirm whether this region is consistent with the position of the active centromere of derivative 3-3-11. The FISH results show that probe 2 is located there (Fig. 7). The de novo centromeric region of derivative 3-3-11 is near the translocation breakpoint of TB-9Sb, which is different from 3-3.

According to the ChIP-seq results, we cannot detect an obvious CENH3 binding peak region along 9S in derivatives 3-3-6 and 3-3-7. As noted above, there are no detectable CentC and CRM signals on these chromosomes, so there should be other DNA sequences associated with centromere formation. We performed FISH using probe 2 on derivatives 3-3-6 and 3-3-7 to determine whether this sequence exists in the centromere regions. The results showed that probe 2 is on the chromosomal arm of derivative 3-3-6 (Fig. S8A); the centromere of derivative 3-3-6 may be within the B chromosome-derived sequences that will not be detected in the B73 reference genome. With derivative 3-3-7, the signal of probe 2 is present at two sites, neither of which is consistent with it residing in the active centromere (Fig. S8B). We were also unable to find the sequences associated with CENH3 binding in this derivative, potentially because a de novo centromere was formed over B chromosome sequences. Chromosome patterns and centromere formation of these three new derivatives of 3-3 are quite different, indicating large chromosomal changes took place leading to their selection in the phenotypic genetic screen.
There is also no CentC and CRM sequence in the B73 reference in the de novo centromeric region of derivative 3-3-11 when we performed a BLAST analyzing for CentC and CRM sequences in the 200-kb centromeric region. The percentage of LTR retrotransposon Ty3-gypsy is three times that of Ty1-copia in the 200-kb centromeric region, even though the total percentage of LTR retrotransposons is lower than the level of the whole genome and native centromeres (Table S3).  

Discussion  
From TB-9Sb, misdivision of the B centromere on B-9 produced many variants with different centromere sizes (18). The functional centromere size in different derivatives of TB-9Sb has been examined previously by using the B-repeat sequence via Southern blot analysis and fiber FISH probing (14, 15). That work showed that derivative 3-3 had the smallest quantity of B-repeat sequences among all derivatives of B-9, with evidence that the B-repeat region in derivative 3-3 is less than 280 kb, which is approximately seven times smaller than the size of its progenitor telo2-2 (18). The assumption had been that all of the derivatives possessed a functional B centromere that originated from the initial B centromere of B-9, even though the organization of B-repeat sequence had changed. However, here we showed that the B-repeat sequences are not associated with centromere formation in derivative 3-3 and that the residual B-repeat region is inactive. B-repeat signals are distributed on both ends of derivative 3-3, and CENH3 signal is visible between these signals. Therefore, the centromere of 3-3 is no longer a canonical B centromere.

Structures of centromeres in higher eukaryotes are complex with large centromeric regions and abundant repeat sequences. The sizes of centromere regions vary greatly among different species from a 125-kb single CENH3 containing nucleosome in S. cerevisiae to CENH3 binding domains that span several megabases in humans (11). In plants, the centromere region of Arabidopsis thaliana ranges from 4 to 9 Mb (35–37), whereas sizes of rice centromeres range from 60 kb to 7 Mb (38). In maize, native centromeres range from hundreds of kilobases to several megabases, whereas the de novo centromere region of chromosome Dp3a is only 350 kb (24), and sDic-15 is 723 kb (25). In the present study, derivative 3-3 was found to have a much smaller centromere, which only spans 288 kb. This 288-kb region was detected by using ChIP-seq with a maize-specific CENH3 antibody and was confirmed by FISH using a probe made of the 5-kb single-copy sequence from within the 288-kb CENH3 binding region. In derivative 3-3-11, the 200-kb CENH3 binding domain is the smallest maize centromere found to date. The sizes of human neocentromeres range from 100 to 400 kb (26) and thus are also much smaller than normal centromeres. Collectively, the results revealed that de novo centromere formation is common, but they are much smaller than the normal centromeres and apparently only regularly persist on otherwise structurallyacentric fragments as opposed to a normal chromosome, a point to which we will return below.

There is limited understanding as to why de novo centromeres arise over their underlying sequences (26). The de novo centromere on sDic15 (25) formed near a native centromere, whereas the de novo centromere on Dp3a formed distal from the original (24). Commonalities regarding the positions of formation for de novo centromeres have not been identified to date; the major chromatin environments that control centromere formation remain unknown. During the misdivision of B-9, due to complicated changes in chromosomal structure and chromatin modification, the de novo centromere of derivative 3-3 formed distal from the translocation breakpoint, but the centromere of derivative 3-3-11 is near the translocation breakpoint.

Centromere inactivation and reactivation is an interesting aspect of centromere function. In maize, many structurally dicentric chromosomes with one active and one inactive centromere have been found from B-A translocation chromosomes undergoing chromosome-type BFB cycle (29). The inactive centromere of structurally dicentric chromosomes can be reactivated upon release from the chromosome after intrachromosomal recombination (30). Here we followed a chromosome undergoing rearrangement that exhibits a sequential series of de novo formations and inactivation (Fig. 8). First, derivative 3-3 contains a de novo centromere following the loss of the functional B centromere of telo2-2 with retention of only reduced B-repeat copies. Then, in the subsequent derivatives of 3-3, the de novo centromere was inactivated and another de novo centromere formed in derivative 3-3-11. In
derivatives 3-3-6 and 3-3-11, the centromere region of 3-3 is inactivated and the functional centromere is apparently present in B chromosome-derived sequences. The misdivision screen yielded centromeres of reduced size (15) but can also lead to changes of centromeric DNA sequences, loss of canonical centromeric DNA sequences, and increased or decreased copy number of the flanking B-repeat sequences.

The current maize genome experienced large-scale chromosome evolution from the ancient tetraploid ancestor via translocation, insertion, and fusion (39). The chromosome number and hence number of centromeres has been reduced presumably by centromere inactivation to leave traces of their former presence in the genome (39). The sequences of the de novo centromeres described here are not present at the positions of former centromeres and therefore represent true newly formed sites rather than reactivation events.

The results reported here reveal the regularity with which de novo centromeres can form on a chromosome fragment without a canonical centromere. For this to occur, CENH3 nucleosomes must be present in a critical mass in the chromosome arms. Ordinarily, CENH3 is enriched at the positions of the centromeres but can also be present to some degree elsewhere on the chromosome (40). However, the incorporation into arms is a regular occurrence upon overexpression of CENH3 (41, 42), which can produce ectopic centromeres, suggesting that a continuum of frequency of de novo centromere formation might exist. Under normal conditions, one possibility might be that any de novo centromere formation is blocked in the arms. However, given the findings that de novo centromeres are formed on structurally acentric fragments frequently (this study and refs. 22–25), it is likely that de novo centromeres are indeed organized on chromosome arms but are inactivated as described below.

With these realizations, the following narrative provides a model of how chromosome integrity is maintained. In the absence of a canonical centromere that range to megabases in maize (43), smaller de novo centromeres can become established. Because de novo centromere formation is common on chromosomal fragments in maize (refs. 24 and 25 and this study) and barley (22), it is likely that de novo centromeres are initiated regularly in chromosomal arms, but when opposed in anaphase against the large canonical centromere, they are as quickly inactivated, leaving no trace. Previous work suggests that the smaller centromere in functionally dicentric chromosomes is prone to inactivation in both maize (30) and wheat (32). These findings indicate that if a centromere is not used during a particular anaphase it becomes epigenetically inactivated and inherited in that state. Indeed in budding yeast, the deposition of CENH3 to an artificially generated neocentromere relied on interactions with kinetochore protein complexes (44). Therefore, it is reasonable to suggest that disruption of such interaction might cause centromere inactivation. The regular occurrence of de novo centromeres might help maintain the size of the canonical centromere through selection because the normal centromere must be large enough to be capable of purging any de novo centromeres that are initiated. However, in structurally acentric fragments, de novo centromeres can become established because of the absence of such an inactivating force. In the sequence of events described here, it seems reasonable that the first formed de novo centromere in derivative 3-3 was insufficiently large to predominate over subsequently initiated de novo cases, leading to inactivation and shifts to new functional sites of activity. In a normal chromosome, the canonical centromere position would seldom change because of size predominance, but in chromosomal fragments the sites of activity can be dynamic. Thus, despite regular initiation of small de novo centromeres, they are soon inactivated and chromosome structure is seldom altered over evolutionary time.

Materials and Methods

Plant Materials. The progenitor TB-95b, telocentric chromosomes telo2-2, and derivative 3-3 (18), as well as derivatives 3-3-6, 3-3-7, and 3-3-11, are from the collections of J.A.B. All these materials were screened by FISH using probes for CentC (12), CRM (3), 180-bp knob (45), and B-repeat (46). Suitable materials were planted in the greenhouse or in the field for cytogenetic analysis and ChiP experiments.

DNA Probe and FISH. For mitotic analysis, CentC, CRM, telomere repeat sequence, and 180-bp knob sequences were labeled with Alexa Fluor 488-S-DUTP, and B-repeat sequences were labeled with Texas Red-S-DCTP by nick translation as described (47). Probe 1 was prepared from the de novo centromere 288-kb region and probe 2 from the 200-kb region; they were labeled with Alexa Fluor 488-S-DUTP (primer sequences are listed in Table S2). FISH in mitosis was performed as described (48). Images were taken as a confocal z-stack (Zeiss LSM 710 NLO) and processed with Adobe Photoshop CS 3.0.

Immunolocalization in Mitotic Cells. Immunolocalization for mitosis was performed as described (36). The maize CENH3 antibody (25) and the phosphorylated H2A antibody were previously described (7). The images were taken as described above.

Meiotic Analysis. FISH on anthers of different meiotic stages was conducted as previously described (25). FISH images were recorded by an epifluorescence Olympus BX61 microscope and processed with Adobe Photoshop CS 3.0.

ChiP and ChiP-seq. ChiP was performed as described (49). Roughly 20 g of young leaves from TB-95b, telo2-2, and derivatives 3-3, 3-3-6, 3-3-7, and 3-3-11 were used for ChiP. The maize CENH3 antibody used for ChiP was the same for immunolocalization. ChiP-seq was conducted as previously described (25). The enriched DNA samples were sequenced using Illumina Hiseq2000 platform to generate pair-ended 100-bp sequence reads.

Analysis of ChiP-seq Data. Approximately 300- to 400-Mb ChiP-seq paired-end reads were mapped to maize B73 RefGen_v3 using BWA software as described (33). Only uniquely mapping reads were chosen for further analysis (Table S1).

Then we calculated the abundance of reads by reads-per-million values with 10-kb windows sliding along the genomic regions of interest. Figures were produced with R scripts. The anti-CENH3 ChiP-seq data were deposited in the Gene Expression Omnibus database (accession no. GSE59124).

ACKNOWLEDGMENTS. We thank Dr. Ryan Douglas (University of Missouri) and Nathan D. Han (Washington University in St. Louis) for critically reading the manuscript and valuable suggestions. This work was supported by National Science Foundation of China Grants 31100033 and 31220103912 and National Science Foundation Grant DBI 0927203.


