CENTROMERE FUNCTION AND EVOLUTION IN MAIZE (*ZEA MAYS*)

A Dissertation
Presented to
the Faculty of the Graduate School
University of Missouri – Columbia

In Partial Fulfillment
of the Requirements for the Degree

Doctor of Philosophy

By

JONATHAN CARL LAMB

Dr. James A. Birchler, Dissertation Supervisor

DECEMBER 2006

The undersigned, appointed by the dean of the Graduate School, have examined the dissertation entitled

CENTROMERE FUNCTION AND EVOLUTION IN MAIZE (*ZEA MAYS*)

presented by Jonathan Carl Lamb,

a candidate for the degree of doctor of philosophy of Biological Sciences,

and hereby certify that, in their opinion, it is worthy of acceptance.

Professor James A. Birchler

Professor Kathleen J. Newton

Professor Karen C. Cone

Professor Emmanuel Liscum

Professor Perry Gustafson

ACKNOWLEDGEMENTS

My mentor, Dr. James A. Birchler, has provided an environment in which creativity and collaboration can thrive. The opportunities and freedom that have been presented to me in Dr. Birchler's laboratory have truly been exceptional. At one point he told me that publications were the currency of the academic world. Continuing that metaphor, I do not believe that the pay could have been better in any other laboratory or institution in the world. Nor could I have grown and improved my abilities at the same rate in any other situation. He is truly a kind and wise friend and I will be proud to acknowledge his influence throughout the rest of my scientific career.

The bulk of the research presented in this dissertation was conducted in collaboration with others. I have benefited greatly through these interactions. The community of maize researchers embodies the ideals of collegiality. The willingness of others, from undergraduates to senior scientists, to discuss their ideas and listen to mine has let me feel included in the great scientific endeavor. Of course, I have worked more closely with some. Accordingly, I am grateful to the coauthors on the various publications I have contributed to, especially Drs. Akio Kato and Fangpu Han. Dr. Kato developed the technique I used most and taught it to me thoroughly. Dr. Han is endlessly finding more interesting things to study and has shared it all with me.

TABLE OF CONTENTS

CENTROMERE FUNCTION AND EVOLUTION IN MAIZE

Jonathan Carl Lamb

Dr. James A. Birchler, Dissertation Supervisor

ABSTRACT

A dispensable supernumerary chromosome present in maize, the B chromosome, was the focus of many centromere studies. First, I examine the variation in copy number of centromeric elements and other repeats among different maize lines. Then, I demonstrate that centromeric elements are present away from the centromere on the B chromosome indicating that centromeric elements are not sufficient for centromere function in maize. I demonstrate that the B centromere can be inactivated in dicentric chromosomes to produce stable, functionally monocentric chromosomes. Next, I examine the rate of divergence for centromeric elements in maize and its relatives in relation to other repetitive elements in the genome. I examined the genomic distribution of repetitive elements showing that certain families of retrotransposons are enriched in the heterochromatic regions flanking the centromere. Finally, I describe a novel chromosomal rearrangement, an inversion with a breakpoint in the centromere that splits the tract of centromere repeats creating a chromosome with two distinct sites of centromere elements.

CHAPTERS

1. INTRODUCTION

During my time as a graduate student, I had the opportunity to write two review articles on centromeres. The first, **The role of DNA sequence in centromere formation**, discusses the evidence for and against a role for specific sequences in determining centromere identity. The second, **What's in a centromere?**, written a year later, describes additional research on centromeres that occurred during the intervening year. It also discusses possible roles for different chromatin states and modifications that could influence centromere identity.

These two articles were published in *Genome Biology* and are included to provide a general background on the state of knowledge regarding centromeres at the beginning of my research.

LAMB, J. C., and J. A. BIRCHLER, 2003 The role of DNA sequence in centromere formation. Genome Biol **4:** 214.

LAMB, J. C., J. THEURI and J. A. BIRCHLER, 2004 What's in a centromere? Genome Biol **5:** 239.

The role of DNA sequence in centromere formation

Jonathan C. Lamb and James A. Birchler

Abstract

Centromeres are key to the correct segregation and inheritance of genetic information. Eukaryotic centromeres, which are located in large blocks of highly repetitive DNA, have been notoriously difficult to sequence. Several groups have recently succeeded in analyzing centromeric sequences in human, Drosophila and Arabidopsis, providing new insights into the importance of DNA sequence for centromere function.

Published: 29 April 2003

Genome Biology 2003, 4:214

The electronic version of this article can be found online at:

http://genomebiology.com/2003/4/5/214

Centromeres are essential for the proper segregation of chromosomes during cell division in eukaryotes. They are characterized by highly repetitive DNA regions and bound kinetochore proteins, which are required for the attachment of microtubules to the chromosomes during mitosis. Centromeres are a paradox in that their basic function is highly conserved across eukaryotes but their sequences are divergent, even between closely related species [1]. Several investigators have therefore suggested that the DNA sequence may not be essential in centromere formation [2]. It has been difficult to

address this issue because of a lack of complete sequence for any higher eukaryotic centromere. Sequencing efforts have been confounded because centromeres are located in regions of highly repetitive DNA. Several groups [3-7] have recently developed novel methods to overcome these difficulties and report extensive centromeric sequence data from human, Drosophila and Arabidopsis.

Centromere sequences in different species

Deletion of large regions of the human Y chromosome has shown that centromere activity is associated with a block of tandemly repeated 171 base-pair (bp) units, termed α-satellite DNA [8]. Further work has demonstrated that every human centromere is associated with arrays of this α -satellite DNA that can be several megabases (Mb) in size. These massive arrays are imbedded between blocks of pericentric heterochromatin containing highly repetitive DNA [9]. In situ hybridization with α -satellite and immunolabeling using antibodies against kinetochore proteins also confirms that centromeres are located in these regions [10].

Schueler et al. [3] used variation among the 171 bp repeats of α -satellite DNA in the human centromere to design PCR markers. The markers were used for constructing a 500 kilobase (kb) contig of bacterial artificial chromosomes (BACs) that covers a region that is immediately adjacent to, and including part of, a 3 Mb array of α -satellite located at the centromere of the human X chromosome. Shotgun and BAC end sequencing gave a sampling of this region that consisted of approximately 62% diverged α -satellite DNA, about 24% other satellite repeats, and about 16% LINE-type retroelements, as well as other sequences. The 3 Mb array of α -satellite DNA consists of nearly identical copies of

the 171 bp unit that have more than 99% sequence identity and are all oriented in the same direction. At the edge of the array is approximately 40 kb of α -satellite DNA that becomes more divergent with distance from the center of the 3 Mb array, moving from 98% to 70% identity at the edge.

Arabidopsis centromeres include a 178 bp satellite repeat, which is organized in tandem arrays that range in size from 0.4 Mb to 1.4 Mb on different chromosomes and are located between regions enriched for various satellites and other repetitive elements [6,11]. The clusters of α-satellite DNA in human and the 178 bp centromeric element in Arabidopsis are organized in similar ways, although their primary sequences are completely unrelated. Interestingly, centromeres of other plants have also been shown to contain DNA elements of similar length, and this may reflect a common requirement for centromere function (see, for example, [12]).

To overcome difficulties in sequencing repetitive DNA from Drosophila centromeres, a novel approach [5] was used involving the Drosophila minichromosome Dp1187, which is derived from the X chromosome and retains a fully functional centromere. Several deletion derivatives of this minichromosome were recovered after irradiation and were used to map the centromere to a 420 kb region. One derivative chromosome of 620 kb was isolated by electrophoresis and gel extraction. Its DNA was fractionated and cloned and bacterial transposons were inserted into the cloned DNA [5]. Previous work [13] had demonstrated that the centromere of the Drosophila X chromosome is composed of arrays of two types of simple 5 bp satellites, AATAT and AAGAG, that are interrupted by five retrotransposons and an ‘island’ of complex DNA. Using primers specific to the inserted bacterial transposons or tagged primers that

consisted of satellite sequence attached to non-homologous sequence, Sun et al. [5] were able to sample 31 kb of the AATAT and AAGAG satellites. This study [5] and previous work [13] showed that the arrays in the Drosophila centromere are highly similar - the AATAT sequence had 2.2% variation and AAGAG had only 0.3% variation in sequence - and that the repeats in each satellite are in the same orientation. Whereas transposon-like sequences previously found in Drosophila heterochromatin often consisted of scrambled clusters of different elements [5], the retrotransposons in the centromere of the X chromosome were intact. This suggested that they had recently been inserted into the genome or that their sequence is functionally conserved. The island of complex DNA was shown to be 39 kb long, including 16.2 kb of AT-rich sequence and retrotransposon-like elements that are arranged in blocks in different orientations. The beginning and end of this island contain a similar sequence, but are oriented in opposite directions - an arrangement analogous to fission yeast centromeres [5].

All of the elements identified by Sun et al. [5] are also found at non-centromeric locations in the Drosophila genome; the AATAT and AAGAG satellites are present in other but not all centromeres. Indeed, in Drosophila there are no DNA sequences that are located at every centromere, suggesting that primary centromeric sequence alone is neither sufficient nor necessary for centromere formation. The arrays identified in the X chromosome may therefore be merely permissive for centromere organization.

Insights from aberrant centromeres

Drosophila centromeres are unusual in being composed of sequences that are abundant elsewhere in the genome whereas in plants or mammals this is not the case

under normal circumstances. But there are some cases, in which the usual human centromeric sequences can be found at other chromosomal locations, where they display no detectable centromeric activity. For example, Robertsonian translocations, which are whole-arm rearrangements between acrocentric chromosomes can link two centromeres and yet the resulting chromosome is stably transmitted through mitosis and meiosis. Furthermore, in situ analysis using antibodies against essential kinetochore proteins, such as CENP-C, an essential component of the inner kinetochore plate, and CENP-A, the centromere-specific variant of histone H3 in human, has shown that only one of the two centromeric locations retains function [10].

Also in humans, rearranged chromosomes have been found that lack the region in which the centromere is usually present, and in these cases a new location has acquired centromeric activity. The new site ('neocentromere') has the usual hallmarks of a centromere - it forms a cytologically discernible constriction on the centromere and has kinetochore proteins bound [10,14]. The DNA sequences that gave rise to two of these neocentromeres were determined by immunoprecipitation of chromatin with antibodies against the centromeric histone H3 protein CENP-A. Analysis of the isolated DNA region showed that there are no elements in common between the two neocentromeres and normal centromeres [15,16].

Human artificial chromosomes can be generated by introducing α-satellite DNA arrays into cells [17], but not by introducing the DNA sequences of neocentromeres in a similar fashion [18]. Nevertheless, when the chromosome arms surrounding the neocentromere are truncated by insertion of telomere sequences, the resulting minichromosomes composed of the neocentromere DNA can be perpetuated through cell

divisions [18]. This indicates that the satellite array of normal centromeres can direct de novo centromere formation,whereas the neocentromere DNA cannot. Nevertheless, the chromatin structure of the neocentromere appears to be stably maintained throughout the cell cycle. Because the primary sequences are not similar between neocentromeres and usual centromeres, the presence of neocentromeres suggests that centromere function may be regulated on an epigenetic level independent of DNA sequence.

Models of centromere determination

The importance of chromatin structure for centromere function is supported by the presence of species-specific variants of histone H3 found in the centromeric chromatin of all eukaryotes. The variants interact with the other core histone proteins, H2a, H2b and H4, to form a type of nucleosome that is present only at functional centromeres. It has been suggested that nucleosomes containing centromeric histone H3 are indispensable for centromere function and likely to serve as anchors for kinetochore formation. A model proposing that correct spacing of centromeric and normal nucleosomes is required for centromere function is supported by recent data from Drosophila and human cells showing that stretched chromatin from centromeres is organized into blocks of centromeric nucleosomes interspersed between blocks of nucleosomes containing the normal core histone H3 [19]. This spacing may be facilitated by the satellites present at centromeres. Centromeric satellites from mammals and plants are approximately the length required to wrap around a nucleosome, and even in Drosophila multiples of the 5 bp satellites could add up to a unit of nucleosomal length.

Analysis of centromeric histone H3 in related species of mammals, flies, and plants has shown that the variants are highly similar to core histone H3 proteins in the regions that interact with the other histone proteins [20-22]. But in the region that is likely to contact the DNA strand centromeric histone H3 proteins appear to be under adaptive selection. Because the DNA sequence elements that are in contact with the centromeric H3 histones are divergent between species, it has been suggested that the centromeric histone H3 protein and the DNA are coevolving. Meiotic drive (a distortion of chromosome segregation) resulting from preferential positioning of 'stronger' centromeres to the egg during female meiosis might be the mechanism for this coevolution [20,21].

Many models for centromere determination predict that centromere function is independent of the underlying sequence. Such models are formulated to explain how nucleosomes containing centromeric histone H3 are maintained at all functional centromeres regardless of the DNA sequence with which they are associated. Spatial or temporal sequestration of the centromeres within nuclear compartments coupled to the availability of centromeric nucleosomes within these compartments or time phases has been suggested as a mechanism. Another model predicts that extant nucleosomes containing centromeric histone H3 are distributed to each strand during replication and subsequently used in post-replication recruitment of additional centromeric nucleosomes (for further discussion see [2]).

Models for centromere formation that do not rely on sequence must account for certain elements, such as the human α -satellite DNA and the Arabidopsis 178 bp repeat, that are present at every centromere in a normal karyotype within a given species. It

seems that there must be mechanisms that homogenize repetitive elements such as centromeric repeats. For example, unequal crossing-over has been postulated to explain homogenization of α -satellite DNA within a chromosome [3], but there must also be a process that homogenizes the repeats between nonhomologous chromosomes. Unless the homogenization mechanism imposes constraints on the substrate sequence, changes to centromeric elements that become fixed in different populations would become randomly distributed in the absence of selection for sequence content. The analysis of Arabidopsis and human centromeric satellites identified regions that were conserved among the various iterations, as well as regions that were more variable than average, implying that selection pressures act on the sequence of centromeric elements [7]. The observed non-random distribution of centromeric satellite DNA is not consistent with a model proposing complete irrelevance of sequence.

Some investigators [23,24] have raised the possibility that secondary structure or even higher order DNA structure could be a factor in determining centromere position and function. This idea may reconcile data showing irrelevance of primary sequence on the one hand with data that show conservation of DNA elements on the other. Conservation of DNA secondary structure allows for large variation in sequence, but does not exclude fine-tuning of the primary sequence, perhaps through coevolution with the domain of the centromeric histone H3 that associates with DNA. Similarly, epigenetic models of centromere formation, proposing regulation at the chromatin level, would not exclude fine-tuning of primary sequence. In either model, formation of a centromere with a new sequence would be allowed as long as the region permitted the proper higher order DNA organization.

Data from neocentromere analysis do provide support for the idea that centromeres self-perpetuate without the need for a specific underlying sequence. In contrast, conservation of human and Arabidopsis centromeric repeat sequences suggests specific requirements at this level. Extreme models advocating a specific DNA element at centromeres versus no requirement at all will probably require a new synthesis. The means by which the position of the centromere on the chromosome is determined has yet to be resolved, but the recent elucidation of DNA sequence from the centromeres of various species is valuable information for making new predictive models. To determine the importance of various DNA elements found in or near the centromere, the mechanisms that drive evolution of centromeric DNA need to be clarified. For example, the lack of any centromeric elements common to all centromeres in Drosophila may be the result of a homogenization mechanism that is fundamentally different from the one that seems to function in mammals and plants. As additional centromeric sequences continue to become available from many different species, insights into the homogenization of sequences and their involvement in centromere formation will grow.

References:

1. Henikoff S, Ahmad K, Malik HS: **The centromere paradox: stable inheritance with rapidly evolving DNA.** *Science* 2001,**293:**1098-1102.
2. Sullivan BA, Blower MD, Karpen GH: **Determining centromere identity: cyclical stories and forking paths.** *Nat Rev Genet* 2001, **2:**584-596.
3. Schueler MG, Higgins AW, Rudd MK, Gustashaw K, Willard HF: **Genomic and genetic definition of a functional human centromere.** *Science* 2001, **294:**109-115.
4. Guy J, Hearn T, Crosier M, Mudge J, Viggiano L, Koczan D, Thiesen H, Bailey JA, Horvath JE, Eichler EE, *et al.*: **Genomic sequence and transcriptional profile of the boundary between pericentromeric satellites and genes on human chromosome arm 10p.** *Genome Res* 2003, **13:**159-172.
5. Sun X, Le HD, Wahlstrom JM, Karpen GH: **Sequence analysis of a functional *Drosophila* centromere.** *Genome Res* 2003, **13:**182-194. 6. Copenhaver GP, Nickel K, Kuromori T, Benito M, Kaul S, Lin X, Bevan M, Murphy G, Harris B, Parnell LD, *et al.*: **Genetic definition and sequence analysis of *Arabidopsis* centromeres.** *Science* 1999, **286:**2468-474.
7. Hall SE, Kettler G, Preuss D: **Centromere satellites from *Arabidopsis* populations: maintenance of conserved and variable domains.** *Genome Res* 2003, **13:**195-205.
8. Heller R, Brown KE, Burgtorf C, Brown WR: **Mini-chromosomes derived from the human Y chromosome by telomere directed chromosome breakage.** *Proc Natl Acad Sci USA* 1996, **93:**7125-7130.
9. Murphy TD, Karpen GH: **Centromeres take flight: alpha satellite and the quest for the human centromere.** *Cell* 1998, **93:**317-320.
10. Warburton PE, Cooke CA, Bourassa S, Vafa O, Sullivan BA, Stetten G, Gimelli G, Warburton D, Tyler-Smith C, Sullivan KF, *et al.*: **Immunolocalization of CENP-A suggests a distinct nucleosome structure at the inner kinetochore plate of active centromeres.** *Curr Biol* 1997, **7:**901-904.
11. Haupt W, Fischer TC, Winderl S, Fransz P, Torres-Ruiz RA: **The CENTROMERE1 (CEN1) region of *Arabidopsis thaliana*: architecture and functional impact of chromatin.** *Plant J* 2001, **27:**285-296.
12. Zhong CX, Marshall JB, Topp C, Mroczek R, Kato A, Nagaki K, Birchler JA, Jiang J, Dawe RK: **Centromeric retroelements and satellites interact with maize kinetochore protein CENH3.** *Plant Cell* 2002, **14:**2825-2836.
13. Sun X, Wahlstrom J, Karpen GH: **Molecular structure of a functional *Drosophila* centromere.** *Cell* 1997, **91:**1007-1019.
14. Saffery R, Irvine DV, Griffiths B, Kalitsis P, Wordeman L, Choo KH: **Human centromeres and neocentromeres show identical distribution patterns of >20 functionally important kinetochore-associated proteins.** *Hum Mol Gen* 2000, **9:**175-185.
15. Lo AW, Craig JM, Saffery R, Kalitsis P, Irvine DV, Earle E, Magliano DJ, Choo KH: **A 330 kb CENP-A binding domain and altered replication timing at a human neocentromere.** *EMBO J* 2001, **20:**2087-2096.

16. Lo AW, Magliano DJ, Sibson MC, Kalitsis P, Craig JM, Choo KH: **A novel chromatin immunoprecipitation and array (CIA) analysis identifies a 460-kb CENP-A-binding neocentromere DNA.** *Genome Res* 2001, **11:**448-457.
17. Grimes BR, Rhoades AA, Willard HF: **Alpha-satellite DNA and vector composition influences rates of human artificial chromosome formation.** *Mol Ther* 2002, **5:**798-805.
18. Saffery R, Wong LH, Irvine DV, Bateman MA, Griffiths B, Cutts SM, Cancilla MR, Cendron AC, Stafford AJ, Choo KH: **Construction of neocentromere-based human minichromosomes by telomere- associated chromosomal truncation.** *Proc Natl Acad Sci USA* 2001, **98:**5705-5710.
19. Blower MD, Sullivan BA, Karpen GH: **Conserved organization of centromeric chromatin in flies and humans.** *Dev Cell* 2002, **2:**319-330.
20. Malik HS, Vermaak D, Henikoff S: **Recurrent evolution of DNA binding motifs in the *Drosophila* centromeric histone.** *ProcNatl Acad Sci USA* 2002, **99:**1449-1454.
21. Henikoff S, Malik HS: **Selfish drivers.** *Nature* 2002, **417:**227.
22. Yoda K, Ando S, Morishita S, Houmura K, Hashimoto K, Takeyasu K, Okazaki T: **Human centromere protein A (CENP-A) can replace histone H3 in nucleosome reconstitution *in vitro*.** *Proc Natl Acad Sci USA* 2000, **97:**7266-7271.
23. Koch J: **Neocentromeres and alpha satellite: a proposed structural code for functional human centromere DNA.** *Hum Mol Gen* 2000, **9:**149-154.
24. Grady DL, Ratliff RL, Robinson DL, McCanlies EC, Meyne J, Moyzis RK: **Highly conserved repetitive DNA sequences are present at human centromeres.** *Proc Natl Acad Sci USA* 1992, **89:**1695-1699.

What's in a centromere?

Jonathan C Lamb, James Theuri and James A Birchler

Abstract

The complete sequence of rice centromere 8 reveals a small amount of centromere-specific satellite sequence in blocks interrupted by retrotransposons and other repetitive DNA, in an arrangement that is strikingly similar in overall size and content to other centromeres of multicellular eukaryotes.

Published: 17 August 2004

Genome Biology 2004, 5:239

The electronic version of this article can be found online at:

http://genomebiology.com/2004/5/9/239

Shakespeare's Juliet posed the question "What's in a name?" to explore the connotations that a single word can hold. The name 'centromere' conjures many ideas from classical biology, but genome projects have had a difficult time defining exactly what is present at the portion of the chromosome responsible for microtubule association and segregation at mitosis and meiosis. In humans [1], Arabidopsis thaliana [2], and other model organisms, centromeres appear to contain a core of megabase-sized arrays of a single element (or, in flies, several arrays of a small number of different microsatellite elements [3]). Near the center of this core the repeated elements are arranged in a nearly perfect array, while near the edges the uniformity decreases and the arrays are

interspersed by various repetitive elements. Because of the size and uniformity of the cores, they have been impossible to sequence with standard techniques and so have remained as gaping holes of unsequenced DNA in the otherwise well-defined model-organism genomes obtained by various international efforts.

As in other model organisms, each centromere of members of the grass family (including rice and maize) contains large tandem arrays of a species-specific centromeric repeat (CentO in rice [4]; CentC in maize [5]). Fluorescent in situ hybridization (FISH) using centromere-specific satellite sequence as a probe reveals that their copy number among different rice and maize centromeres varies considerably - almost 30-fold in rice. Because the copy number of the centromeric satellite in rice chromosome 8 is very low, two groups - Nagaki et al. [6] and Wu et al. [7] - were able to sequence the entire centromeric region using standard techniques involving bacterial artificial chromosomes (BACs). The two groups screened BAC libraries, created as part of the ongoing effort to sequence the rice genome, with centromere-specific elements as probes, and then 'walked' from BAC to adjacent BAC, by virtue of overlapping sequence at their ends, so as to form a minimal tiling path, or contig, spanning the genetically defined centromeric region. Their work has resulted in the first complete sequence of a normal centromere from a multicellular organism.

Because CentO is found as a tandem array of repeats and such repetitive DNA tends to be unstable when maintained in Escherichia coli (which is used to replicate BACs), Nagaki et al. [6] used cytological approaches to confirm the location and completeness of their centromere-containing contig. First, they used BACs that flanked the CentO region from the minimal contig of centromere 8 as FISH probes on spreads of

rice pachytene chromosomes, to confirm that the contig included the entire CentO-containing region. Next, they performed 'fiber FISH', probing the same chromosomes in the form of stretched DNA fibers, again using the BACs from the minimal contig as probes and with CentO as a probe, to show that the predicted tiling path reflected the correct physical arrangement of the BACs around the centromere. This procedure also showed that the complete cytologically detectable CentO-containing region was contained in one BAC. Measuring the length of the CentO array in parallel on stretched genomic DNA and on stretched BAC fibers then confirmed that the CentO array contained in the BAC was intact. Nagaki et al. [6] then sequenced 12 BACs containing 1.65 Mb in total, spanning the CentO tract and extending into both the long and short arms of the chromosome. Wu et al. [7] independently obtained 1.97 Mb of sequence from the same centromeric region that includes the 1.65 Mb from the Nagaki et al. [6] study. They sequenced multiple BACs covering the CentO tracts to confirm the size and integrity of the CentO arrays.

In contrast to human [1] and Arabidopsis [2] centromeres, each of which has a large core of nearly homogeneous satellite sequence, the tandem arrays of centromeric satellite in rice chromosome 8 are frequently interrupted by insertions of a particular family of retroelements of the long terminal repeat (LTR) type, called CRR in rice. Using FISH, retroelements of this type can be seen only at the centromere in cytological preparations from numerous grass species [8]. Nagaki et al. [6] report that rice centromere 8 contains only 41 kilobases of CentO sequence, arranged as a cluster of three arrays of CentO separated by full and partial CRR elements. One of the arrays is oriented in the opposite direction to the other two. There is also approximately 2.8 kb of CentO

that is separated from the main site by over 700 kb of sequence that includes repetitive elements and active genes. Analyzing yet another rice centromere, Zhang et al. [9] defined a BAC contig that spans rice centromere 4 and reported sequencing efforts from the single BAC that hybridizes to the CentO element of this centromere. This BAC contained a 124 kb 'core' region made up of 379 copies of CentO arranged in 18 tracts in different orientations interrupted by various repetitive sequences, including CRR elements and other LTR retroelements and repeats not specific to centromeres.

Because many repetitive elements, including the centromeric unit, are highly divergent between maize and oat, it is possible to use FISH to distinguish the centromeres of maize chromosomes that have been artificially transferred to an oat background. Using this type of material, Jin et al. [10] examined the DNA arrangement along stretched chromatin fibers from individual maize centromeres and found that tracts of the maize centromere repeat element CentC were interspersed with CRM, the maize homolog of CRR, and unknown sequences. This pattern is consistent with the results of the sequencing efforts for rice centromeres 4 and 8 as well as other rice [4] and maize [11] BACs that contain centromeric satellite sequence. Taken together, these results suggest a consistent pattern of DNA organization at grass centromeres consisting of tracts of centromeric satellite interspersed with various repetitive elements, especially centromere-specific retrotransposons.

Centromeric chromatin structure

Centromeric chromatin includes a centromere-specific histone H3 variant (CenH3) that is incorporated into nucleosomes underlying the kinetochore. These

nucleosomes remain a part of the chromatin throughout the cell cycle and are essential to both meiotic and mitotic cell divisions [12]. Although it has not been established that CenH3 alone determines centromere identity, the sequence of a complete centromere should at the least include the entire region that is wound around nucleosomes containing CenH3. Nagaki et al. [6] used anti-CenH3 antibodies to immunoprecipitate chromatin (ChIP) comprising DNA bound to CenH3-containing nucleosomes, confirming that CenH3 is associated with both the CentO repeats and the CRR family of retrotransposons. Primer pairs were designed that would amplify sequences scattered along the length of the centromere 8 contig, and these were used to sample the immunoprecipitated DNA using a process called ChIP-PCR, showing that the CenH3-containing region is approximately 750 kb and does not include the small 2.8 kb cluster of CentO that is separated from the three main arrays. Although the region immediately around the CentO tracts for both centromeres 4 and 8 consists entirely of repetitive elements, the 750 kb CenH3-binding domain of rice centromere 8 included 14 putative non-retroelement open reading frames (ORFs), including 4 that were shown to be expressed by reverse-transcriptase-coupled PCR [6]. This observation is reminiscent of human neocentromeres - chromosomal regions that have newly acquired centromere activity. Neocentromeres have also been shown to harbor expressed genes [13], and the rice finding shows that the chromatin structure of both plant and mammalian CenH3-binding domains is open and accessible to the transcriptional machinery.

In addition to binding microtubules, centromeres have other functions, including sister chromatid cohesion and preventing microtubules from both poles attaching to the same chromatid. These other functions may be located in domains with distinct

chromatin structures [14,15]. To examine the chromatin structure of rice centromere 8, Nagaki et al. [6] used ChIP-PCR with antibodies against two different covalent modifications of the canonical H3 histone protein (rather than the centromere-specific CenH3): dimethylation on lysine 9 (dimethyl-K9), which has been shown to be enriched in heterochromatic regions, and dimethyl-K4, which is present in euchromatic portions of the chromosome. The region associated with dimethyl-K9 H3 spans approximately 1.2 Mb and includes all of the CentO arrays. Because this region covers the entire CenH3-binding region (around 750 kb), the authors [6] postulated that CenH3- containing and dimethyl-K9 H3-containing nucleosomes are interspersed and that the position of these nucleosomes is dynamic, so that a population of cells may have the same DNA sequence interacting with both types of nucleosome. Indeed, the interspersion of these two types of nucleosome has been observed on stretched chromatin fibers of both Drosophila [16] and maize [10]. Immunoprecipitation with antibodies against dimethyl-K4 H3 was limited to the edges of the contig flanking the dimethyl-K9 H3 region [6].

Nagaki et al. [6] and Wu et al. [7] chose the rice centromere with the fewest copies of CentO for their sequencing efforts. Although this approach allowed an achievement not otherwise possible, the sequence obtained may not be representative of centromeres of other rice chromosomes and of some other model organisms, because of its unusually small size. Despite the reduced copy number of CentO, however, it should not be concluded that the functional domain of rice centromere 8 is smaller than other centromeres. In humans [1,15] and Arabidopsis [17], which have centromeres made up of numerous copies of satellite sequences, the CenH3- binding region covers only a portion of the central core of the centromeric satellite array. In rice and maize, ChIP analysis

shows that the majority of centromeric satellite is not associated with CenH3 [6,18]. Cytological observation of maize chromosomes shows that while the amount of centromeric satellite varies extensively among centromeres, the amount of CenH3 remains relatively constant [18]. Although it is difficult to determine the precise sizes of centromeres (because they are composed of large arrays of satellite), observations of fragmented centromeres arising from rare events [19,20] have allowed the lengths of some centromeres to be estimated. The rice centromere 8 CenH3-binding domain is consistent with the reported minimal sizes of other centromeres including the maize B chromosome (around 500 kb) [19], the human Y chromosome (not more than 500 kb) [20] and a Drosophila minichromosome (around 420 kb) [3], suggesting a common size requirement. Additional requirements for effective passage through meiosis may necessitate additional chromatin configurations and could explain the excess sequences that are present at many centromeres and whose function is not yet apparent. For example, Drosophila minichromosomes that lack sequences adjacent to the essential core show reduced meiotic transmission [21].

Because human neocentromeres are not composed of repetitive DNA, immunoprecipitation analysis is possible and a direct comparison of chromatin states between neocentromeres and rice centromere 8 is revealing (Figure 1). Human neocentromere 10q25.3 contains a 330 kb CenH3- binding region within a 700 kb domain that can be precipitated by an immune serum containing antibodies against numerous centromeric proteins [22]. These domains are flanked by regions that replicate late in the cell cycle. In total, the region altered by adoption of centromere identity is approximately 1.4-2 Mb, similar in size to the dimethyl-K9 H3-bound region of rice centromere 8.

Although dimethyl- K9 H3 antibodies were not used in the study by Lo et al. [22], the delayed replication of this region probably reflects the presence of dimethyl-K9 H3 or a similar heterochromatic structure. The similarities in chromatin domain size and arrangement between rice centromere 8 and the human neocentromere (Figure 1) suggest that rice and human have similar chromatin requirements for functional centromeres, including a requirement for flanking heterochromatin that is shared with Drosophila [21]. Additional chromatin domains have been identified within the human neocentromere, including a domain that binds the centromere protein CenPH and another enriched for chromosomal scaffold/matrix attachment regions [13]. With the availability of the complete sequence for rice centromere 8, similar analysis can now be performed for this centromere and the findings compared to the human neocentromere results.

Centromere evolution

Taking their cue from the analysis of human neocentromeres, Nagaki et al. [6] suggest that the presence of active genes indicates that rice centromere 8 is relatively 'young', evolutionarily, and may have arisen from a neocentromerization event. In humans, neocentromerization is usually initiated by a significant chromosomal rearrangement, such as a translocation that produces an acentric fragment, but neocentromeres can also arise spontaneously in an intact karyotype within a single generation [23]. Consistent with the hypothesis that rice centromere 8 is a relatively new centromere, the amount of CentO it contains is small and sequence analysis of the LTRs of the CRR-class retroelements shows that they have recently inserted into the region. But because the CenH3-binding domain has not been determined for other rice

centromeres, the possibility that active genes and frequent retrotransposon insertions are a common feature of grass centromeres cannot yet be ruled out. Also, certain maize centromeres in some lines have virtually undetectable amounts of CentC [5] while homologous centromeres of other lines contain numerous copies of the centromeric satellite and are present at the same genetic location [24]. This suggests that aside from neocentromere formation, mechanisms that reduce satellite copy number could account for the small amount of CentO at rice centromere 8. An example of such a reduction is seen in a study of human cells in which centromere 21 spontaneously lost a specific portion of the centromeric repeat array at a measurable frequency [25].

Although rice centromeres 4 and 8 do not contain massive arrays of CentO, other rice centromeres do (for example, centromeres 1 and 11 [4]), indicating that forces that expand centromeric DNA elements are active in rice. Despite the involvement of epigenetic factors that determine centromere identity, certain DNA sequences seem more suited to life in a centromere than others [26]. In chromosomes that contain very few copies of centromeric satellite, flanking sequences, including genes, will be incorporated into the centromere and forced to conform to local centromeric chromatin requirements. Introduction and subsequent expansion of more suitable sequences would push these sequences away from the active centromere core. Such changes would be strongly selected for, especially if the misexpression of genes incorporated into centromeric regions is detrimental to individual fitness and regular expression could be restored by the expansion of centromere repeats. This type of selection pressure on new centromeres to expand would complement other forces that could drive centromere satellite expansion, such as competition among centromeres during female meiosis [27].

The two rice centromere 8 sequences derived from Nipponbare varieties by Nagaki et al. [6] and Wu et al. [7] are essentially identical to each other except for the size of the CentO arrays: 38.2 kb versus 68.5 kb of CentO contained in the major cluster for Nagaki et al. [6] and Wu et al. [7], respectively. Despite the large differences in satellite copy number, the relative orientation of the tandem arrays is the same for the two groups' sequencing efforts, and the CRR elements that separate the three arrays are identical. Because both groups took steps to confirm that the size of their tracts was accurate, it is unlikely that rearrangements resulting from the cloning process account for the differences between the two groups' findings. Instead, the sequencing efforts probably captured ongoing changes in centromeric satellite copy number and underscore how rapidly such change can occur.

In humans, L1 retroelement insertions are scarce in the heart of the centromeric satellite arrays but are more common in the divergent repeat units found on the periphery. Insertions located at some distance from each other are found to be either present or absent as a group, a phenomenon that can be explained by intra-chromosomal recombination between L1 elements simultaneously removing several elements and the intervening satellites [28]. The presence of a centromere-specific LTR retroelement has thus far only been observed in the grasses and, in contrast to human L1 retroelements, the grass centromeric retroelements show a preference for, and frequent insertion into, centromeric regions including satellite arrays. Thus, an accelerated process of continual transposition and subsequent rearrangements coupled with satellite expansion may explain the differences between human and grass centromeres, the latter of which contain

clusters of centromeric satellite organized in fragmented arrays with different orientations and abundant solo LTR elements.

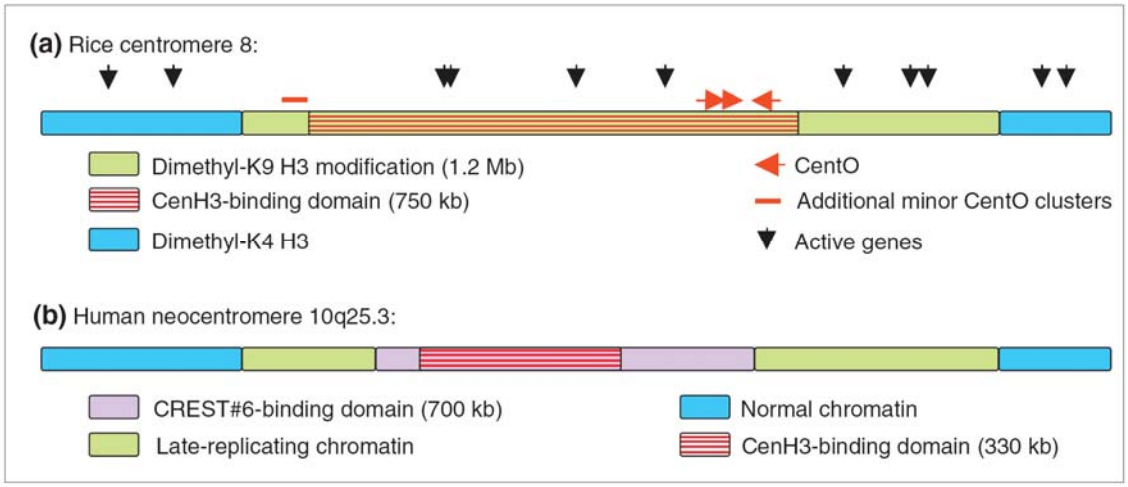
In conclusion, the completion of the first sequences of a centromere from a multicellular eukaryote thus indicates that the necessary regions span hundreds of kilobases and contain a specific repeat. Some of this region is organized around nucleosomes containing CenH3 or histone H3 dimethylated at lysine 9. As other sequences become available, further generalizations will emerge to answer the question from 'Juliet of the genome', "What's in a centromere?"

References

1. Schueler MG, Higgins AW, Rudd MK, Gustashaw K, Willard HF: **Genomic and genetic definition of a functional human centromere.** *Science* 2001, **294:**109-115.
2. Copenhaver GP, Nickel K, Kuromori T, Benito MI, Kaul S, Lin X, Bevan M, Murphy G, Harris B, Parnell LD, *et al.*: **Genetic definition and sequence analysis of *Arabidopsis* centromeres.** *Science* 1999, **286:**2468-2474.
3. Sun X, Le HD, Wahlstrom JM, Karpen GH: **Sequence analysis of a functional *Drosophila* centromere.** *Genome Res* 2003, **13:**182-194.
4. Cheng Z, Dong F, Langdon T, Ouyang S, Buell CR, Gu M, Blattner FR, Jiang J: **Functional rice centromeres are marked by a satellite repeat and a centromere-specific retrotransposon.** *Plant Cell* 2002, **14:**1691-1704.
5. Ananiev EV, Phillips RL, Rines HW: **Chromosome-specific molecular organization of maize (*Zea mays* L.) centromeric regions.** *Proc Natl Acad Sci USA* 1998, **95:**13073-13078.
6. Nagaki K, Cheng Z, Ouyang S, Talbert PB, Kim M, Jones KM, Henikoff S, Buell CR, Jiang J: **Sequencing of a rice centromere uncovers active genes.** *Nat Genet* 2004, **36:**138-145.
7. Wu J, Yamagata H, Hayashi-Tsugane M, Hijishita S, Fujisawa M, Shibata M, Ito Y, Nakamura M, Sakaguchi M, Yoshihara R, *et al.*: **Composition and structure of the centromeric region of rice chromosome 8.** *Plant Cell* 2004, **16:**967-976.
8. Jiang J, Nasuda S, Dong F, Scherrer CW, Woo SS, Wing RA, Gill BS, Ward DC: **A conserved repetitive DNA element located in the centromeres of cereal chromosomes.** *Proc Natl Acad Sci* USA 1996, **93:**14210-14213.
9. Zhang Y, Huang Y, Zhang L, Li Y, Lu T, Lu Y, Feng Q, Zhao Q, Cheng Z, Xue Y, *et al.*: **Structural features of the rice chromosome 4 centromere.** *Nucleic Acids Res* 2004, **32:**2023-2030.
10. Jin W, Melo JR, Nagaki K, Talbert PB, Henikoff S, Dawe RK, Jiang J: **Maize centromeres: organization and functional adaptation in the genetic background of oat.** *Plant Cell* 2004, **16:**571-581.
11. Nagaki K, Song J, Stupar RM, Parokonny AS, Yuan Q, Ouyang S, Liu J, Hsiao J, Jones KM, Dawe RK, *et al.*: **Molecular and cytological analyses of large tracks of centromeric DNA reveal the structure and evolutionary dynamics of maize centromeres.** *Genetics* 2003, **163:**759-770.
12. Sullivan BA, Blower MD, Karpen GH: **Determining centromere identity: cyclical stories and forking paths.** *Nat Rev Genet* 2001, **2:**584-596.
13. Saffery R, Sumer H, Hassan S, Wong LH, Craig JM, Todokoro K, Anderson M, Stafford A, Choo KH: **Transcription within a functional human centromere.** *Mol Cell* 2003, **12:**509-516.
14. Bjerling P, Ekwall K: **Centromere domain organization and histone modifications.** *Braz J Med Biol Res* 2002, **35:**499-507.
15. Spence JM, Critcher R, Ebersole TA, Valdivia MM, Earnshaw WC, Fukagawa T, Farr CJ: **Co-localization of centromere activity, proteins and topoisomerase II within a subdomain of the major human X alpha-satellite array.** *EMBO J* 2002, **21:**5269-5280.

16. Blower MD, Sullivan BA, Karpen GH: **Conserved organization of centromeric chromatin in flies and humans.** *Dev Cell* 2002, **2:**319-330.
17. Nagaki K, Talbert PB, Zhong CX, Dawe RK, Henikoff S, Jiang J: **Chromatin immunoprecipitation reveals that the 180-bp satellite repeat is the key functional DNA element of *Arabidopsis thaliana* centromeres.** *Genetics* 2003, **163:**1221-1225.
18. Zhong CX, Marshall JB, Topp C, Mroczek R, Kato A, Nagaki K, Birchler JA, Jiang J, Dawe RK: **Centromeric retroelements and satellites interact with maize kinetochore protein CENH3.** *Plant Cell* 2002, **14:**2825-2836.
19. Kaszas E, Birchler JA: **Meiotic transmission rates correlate with physical features of rearranged centromeres in maize.** *Genetics* 1998, **150:**1683-1692.
20. Tyler-Smith C, Oakey RJ, Larin Z, Fisher RB, Crocker M, Affara NA, Ferguson-Smith MA, Muenke M, Zuffardi O, Jobling MA: **Localization of DNA sequences required for human centromere function through an analysis of rearranged Y chromosomes.** *Nat Genet* 1993, **5:**368-375.
21. Murphy TD, Karpen GH: **Localization of centromere function in a *Drosophila* minichromosome.** *Cell* 1995, **82:**599-609.
22. Lo AW, Craig JM, Saffery R, Kalitsis P, Irvine DV, Earle E, Magliano DJ, Choo KH: **A 330 kb CENP-A binding domain and altered replication timing at a human neocentromere.** *EMBO J* 2001, **20:**2087-2096.
23. Amor DJ, Bentley K, Ryan J, Perry J, Wong L, Slater H, Choo KH: **Human centromere repositioning "in progress".** *Proc Natl Acad Sci USA* 2004, **101:**6542-6547.
24. Kato A, Lamb JC, Birchler JA: **Chromosome painting in maize using repetitive DNA sequences as probes for somatic chromosome identification.** *Proc Natl Acad Sci USA*, in press.
25. Lo AW, Liao GC, Rocchi M, Choo KH: **Extreme reduction of chromosome-specific alpha-satellite array is unusually common in human chromosome 21.** *Genome Res* 1999, **9:**895-908.
26. Lamb JC, Birchler JA: **The role of DNA sequence in centromere formation.** *Genome Biol* 2003, **4:**214.
27. Henikoff S, Malik HS: **Centromeres: selfish drivers.** *Nature* 2002, **417:**227.
28. Laurent AM, Puechberty J, Roizes G: **Hypothesis: for the worst and for the best, L1Hs retrotransposons actively participate in the evolution of the human centromeric alphoid sequences.** *Chromosome Res* 1999, **7:**305-317.

Figure 1 Similarities between a rice centromere and a human neocentromere
(a) Rice centromere 8 contains an approximately 750 kb CenH3-binding domain that is positioned off-center inside an approximately 1.2 Mb domain where H3 is dimethylated at the lysine that is residue 9 (dimethyl-K9 H3). Active genes are found in and around the CenH3-binding domain. Rice-specific centromeric repeats (CentO) are indicated. **(b)** Human neocentromere 10q25.3 contains an approximately 330 kb CenH3-binding domain contained in an approximately 700 kb region that can be precipitated with CREST#6 antibodies and is flanked by late-replicating regions. Shading is used to indicate potentially analogous regions, and the sizes shown are approximate.



2. CHROMOSOME PAINTING USING REPETITIVE DNA SEQUENCES AS PROBES FOR SOMATIC CHROMOSOME IDENTIFICATION IN MAIZE

Because centromeric elements are common to all chromosomes, it is difficult to study an individual chromosome. This chapter consists of an article that I contributed to that describes a fluorescence in situ hybridization (FISH) technique which allows each of the maize chromosomes to be identified facilitating characterization of cytological features on specific chromosomes.

When I arrived at the Birchler lab, Dr. Akio Kato had been working for some years screening random DNA fragments for individual DNA elements that would produce a distinctive pattern when used as a FISH probe. I worked with Dr. Kato to develop approaches to maximize the efficiency of his screening work. I suggested that subtractive hybridization of a collection of fragments that had already been tested would effectively eliminate those elements leaving a high percentage of untested genomic fragments. Furthermore, I extended the application of the repetitive element FISH probes to meiotic chromosome analysis allowing identification of chromosomes at stages, e.g. diakinesis and anaphase, where chromosome identification was not previously possible.

Relevant to my studies on centromeres, the intensity of hybridization by the centromere specific satellite, CentC, was found to be extremely variable among different inbred maize lines. The extent of variation was not expected because most other species have more uniform sizes of centromere element tracts.

This article was published in the Proceeding of the National Academy of Science, USA, volume 101, pages 13554-9. This chapter retains the reference style of that journal. This article was one of the top 100 downloaded articles the year it was published. The

KATO, A., J. C. LAMB and J. A. BIRCHLER, 2004 Chromosome painting using repetitive DNA sequences as probes for somatic chromosome identification in maize. Proc Natl Acad Sci U S A **101:** 13554-13559.

Chromosome painting using repetitive DNA sequences as probes for somatic chromosome identification in maize

Akio Kato, Jonathan C. Lamb, and James A. Birchler*

Division of Biological Sciences, Tucker Hall, University of Missouri, Columbia, MO 65211-7400

Edited by Susan R. Wessler, University of Georgia, Athens, GA, and approved August 4, 2004 (received for review May 24, 2004)

This paper was submitted directly (Track II) to the PNAS office.

Abbreviations: FISH, fluorescence *in situ* hybridization; BAC, bacterial artificial chromosome; Cent4, centromeric satellite 4; CentC, centromeric satellite C; NOR, nucleolusorganizing region; TR-1, tandemly repeated DNA sequence 1. Data deposition: The sequences reported in this paper have been deposited in the GenBank database (accession nos. CL569181 and CL569243).

*To whom correspondence should be addressed. E-mail: birchlerj@missouri.edu.

Abstract

Study of the maize (*Zea mays* L.) somatic chromosomes ($2n = 20$) has been difficult because of a lack of distinguishing characteristics. To identify all maize chromosomes, a multicolor fluorescence *in situ* hybridization procedure was developed. The procedure uses tandemly repeated DNA sequences to generate a distinctive banding

pattern for each of the 10 chromosomes. Fluorescence *in situ* hybridization screening trials of nonsubtracted or subtracted PCR libraries resulted in the isolation of microsatellite 1-26-2, subtelomeric 4-12-1, and 5S rRNA 2-3-3 clones. These three probes, plus centromeric satellite 4 (Cent4), centromeric satellite C (CentC), knob, nucleolus-organizing region (NOR), pMTY9ER telomere-associated sequence, and tandemly repeated DNA sequence 1 (TR-1) were used as a mixture for hybridization to root-tip chromosomes. All 10 chromosomes were identified by the banding and color patterns in the 14 examined lines. There was significant quantitative variation among lines for the knob, microsatellite, TR-1, and CentC signals. The same probe mixture identifies meiotic pachytene, late prophase I, and metaphase I chromosomes. The procedure could facilitate the study of chromosomal structure and behavior and be adapted for other plant species.

Introduction

In maize (*Zea mays* L., $2n = 20$) pachytene chromosomes have been used extensively for karyotyping and cytogenetic analyses. The first procedure to identify maize meiotic chromosomes was developed by McClintock (1), and the method was refined and detailed by Longley (2) and Rhoades (3) in the middle of the 20th century. Pachytene-stage karyotyping has contributed to maize genetics in numerous ways [i.e., constructing chromosome maps (4), examining the structure and behavior of chromosomal aberrations (5), discovering transposable elements (6), and developing A–A and B–A translocation series (7, 8)]. A detailed morphological pachytene chromosome map is available (4). However, the pachytene stage is a relatively short period, and only a

small percentage of anthers carry this stage in the tassel as a whole. In this sense, the procedure is limited by the availability of the appropriate cell type. Thus, there would be advantages for the study of maize chromosomes if each of the 10 members of the karyotype could be identified in somatic cells. Such a system would permit the screening of many individuals in a short period. A root tip contains many dividing cells and the tissue is readily available. However, the identification of somatic chromosomes has been difficult because the highly condensed chromatin structure conceals the fine details that are used for chromosome identification at the pachytene stage, such as cytologically observable knobs, heterochromatic regions, arm ratios, and total chromosome length (5).

Recent development of fluorescence *in situ* hybridization (FISH) technology has provided improved karyotyping on both meiotic (9) and mitotic (10) cells in maize. However, because of the paucity of landmarks and the polymorphism of knobs among varieties (11), these procedures are effective only for specific lines. To build on this procedure for general use with different maize varieties, an increase in the number of probes was necessary. To this end, we attempted bacterial artificial chromosome (BAC) detection on maize chromosomes. However, without the precise level of blocking with unlabeled repetitive DNA, the FISH procedure tends to label all chromosomes nonspecifically because of the presence of retrotransposons in the probe. Comparative genomic hybridization (CGH) has proved to be effective in distinguishing mammalian chromosomes (12); however, our CGH trials using chromosome 2 and 4 trisomics failed, even after computerized data processing of the images. Of the maize genome, ~85% consists of repeated sequences (13), making BAC and CGH detection difficult because of the low unique-gene density. Therefore, we sought repetitive sequences that are located at

specific chromosomal regions that could be used for karyotyping. Screening was carried out on random PCR libraries to recover sequences that were used as FISH probes to identify the somatic chromosomes. These probes, coupled with changes in the chromosomepreparation procedure that improve fluorescent signal detection, allowed the development of a FISH karyotyping method that is effective on all tested maize lines

Materials and Methods

Construction of a Random PCR Library.

The FISH screen using a random PCR library.

Genomic DNA was extracted from immature ears of maize inbred line Mo17 by using a urea-based extraction protocol. DNA was partially digested by DNase I (catalog no. 104–132, Roche Applied Science), and fragments in the size range of 0.5–1.5 kb were collected by using a repeated gel-shift procedure. An adapter sequence (EBH1F, 5'-AGAATTCGGATCCAAGCTTCTGGTTTGT- 3'; and EBH1R+p, 5'-pACAAACCAGAAGCTTGGATCCGAA- 3') was ligated to the fragments, and the DNA was gel purified again to eliminate the low-molecular-weight DNA and adapter dimers. We suspended 1 µl of the DNA solution in 499 µl of molecular-grade DMSO (stored at room temperature; D-4818; Sigma), and the number of PCR amplifiable DNA fragments was determined by using the EBH1F primer. After dilution, an average of 48 PCR-amplifiable fragments were added to a 300-µl PCR solution (Qiagen, Valencia, CA; with use of nuclease-free water, Ambion, Austin, TX), and 3 µl of this solution was added to 96 PCR tubes (average of 0.5 fragments per tube). The fragments were then amplified by using a GeneAmp 9700 PCR machine (Applied Biosystems), and the PCR

products that showed a single band on gel electrophoresis were reamplified by adding 20 µl of PCR solution to the respective PCR tube. The PCR product was then ethanol-precipitated, and the DNA was labeled with biotin-14- dATP (Invitrogen) by the nick-translation procedure (14). Without further purification, the probes were hybridized to maize Oh43 root-tip chromosome spreads, and FISH signals were detected by dichlorotriazinyl aminofluorescein (DTAF) – streptavidin system (Jackson ImmunoResearch) with or without the tyramide amplification (PerkinElmer Life Science), according to the manufacturer's instructions.

The FISH screen using a subtracted PCR library.

During the initial FISH screening, we found that most of the PCR products that showed signals were homologous to previously identified highly repetitive sequences, such as knob (15), centromeric satellite C (CentC) (16), nucleolus-organizing region (NOR) (17), and retroelements heavily distributed over the maize genome. To eliminate these repeated sequences, the mixture of PCR products of the first screening and B repeat (18) were biotin-labeled by nick translation. Then, a 10-times-larger amount of biotin labeled DNA was added to the partially digested and adapter ligated B73 genomic DNA (containing four B chromosomes). A different adapter was used to avoid amplification of contaminants from the first screening (i.e., BEH2F, 5'-AGGATCCGAATTCAAGCTTGTCTTTG-3'; and BEH2R_p, 5'-pCAAAGACAAGCTTGAATTCGGA3-3'). After denaturing and annealing, the DNA fragments were passed through a Vectrex AvidinD column (Vector Laboratories) according to the manufacturer's instructions. The column selectively binds DNA

sequences that are annealed to the biotin labeled repeated DNA sequences. This subtractive process was repeated, and the second PCR library was constructed by using the same procedures as described above. The DNA sequences were screened by the tyramide- amplified FISH procedure on root-tip chromosome spreads of a B73 X Mo17 hybrid (containing B chromosomes). Among the DNA sequences exhibiting FISH signals, 50 fragments were selected for cloning into the pGem-T vector (Promega) based on potential usefulness for karyotyping and future work.

Preparation of Chromosome Spreads.

For analysis, we used 12 commonly used maize inbred lines (A188, A632, B37, B55, B73, KYS, M14, Mo17, Oh43, stock6, W22, and W23) and two maize varieties [Black Mexican Sweet and an abnormal chromosome 10 line (K10) (19)]. Kernels were germinated at 30°C for 2–3 days. Excised root tips were treated with nitrous oxide gas (20) for 2 h. Treated root tips were fixed in ice-cold 90% acetic acid for 10 min and stored in 70% ethanol at -20°C until use. After washing in water on ice, the section containing dividing cells was dissected and digested in 1% pectolyase Y23 (ICN) and 2% cellulase Onozuka R-10 (Yakult Pharmaceutical, Tokyo) solution for 65 min at 37C (one section per tube with 20 μl of enzyme solution). After digestion, the root sections were washed in ice-cold distilled water and then washed in 100% ethanol two times briefly. The root sections were carefully broken by using a needle and vortexed at maximum speed in 100% ethanol for 30 sec at room temperature to separate cells from one another. The cells were collected at the bottom of the tube by centrifugation and resuspended in

acetic acid/ethanol (9:1 dilution) solution. The cell suspension was dropped onto glass slides in a box lined with wet paper towels and dried slowly.

For meiotic chromosome preparations, immature tassels of inbred line Oh43 were fixed in ethanol/acetic acid (3:1 dilution) and stored at -20°C in 70% ethanol. Anthers at the pachytene, late prophase I, or metaphase I stages were selected under light microscopy by examining one anther stained with iron acetocarmine. The remaining anthers were digested in the enzymatic solution, and air-drying was performed in the same manner as described above for root-tip slides.

Probe-Mixture Constitution.

The following nine repeated DNA sequences were used for karyotyping: coumarin-5-dUTP-labeled knob 180-bp sequence (40 ng/µl) (15), Oregon green-488–5-dUTP-labeled NOR-173 clone (0.2 ng/µl), fluorescein-12-dATP, fluorescein-12-dGTP, fluorescein-12-dCTP, fluorescein-12- dUTP-labeled subtelomeric 4-12-1 clone (40 ng/µl) (Table 1; GenBank accession no. CL569186), fluorescein-12-dUTP-labeled CentC (2 ng/µl) (16), fluorescein-12-dUTP and Texas red-5- dUTP-labeled 5S 2-3-3 clone (Table 1; GenBank accession no. CL569181) (18 ng/µl), Texas red-5-dUTP-labeled microsatellite 1-26-2 clone (Table 1) (6 ng/µl), Texas red-5-dUTP labeled centromeric satellite 4 (Cent4) (10 ng/µl) (21), Texas red-5-dATP, Texas red-5-dGTP, Texas red-5-dCTP, Texas red-5-dUTP-labeled pMTY9ER telomere-associated sequence (40 ng/µl) (22), and Cy5-dUTP-labeled tandemly repeatedDNAsequence 1 (TR-1) (20 ng/µl) (23). These sequences were all labeled by the nicktranslation procedure. The NOR-173 clone was obtained from a separate plasmid-based random cloning of maize genomic DNA

(A.K., unpublished data, GenBank accession no. CL569243); sequencing results indicate that this clone is a 17S maize rRNA sequence. The TR-1 clone was obtained from TA cloning (pGem-T vector; Promega) of PCR products amplified by MR77-specific primers (24). The sequence MR77 (GenBank accession no. AF020266) reported by Chen *et al.* (24) is a TR-1 sequence (88% homology; GenBank accession no. AF071123) that was reported by Ananiev *et al.* (23).

After labeling, these probes were purified by column chromatography (BioGel P-60; Bio-Rad) to eliminate unincorporated dNTPs and then coprecipitated with autoclaved (20 min) salmon-sperm DNA (50 μg) and dried. The pellets were resuspended in 2x SSC (containing 1 mM EDTA; 1X SSC = 0.15 M sodium chloride/0.015 M sodium citrate, pH 7) solution and stored at -20°C. Each probe concentration for the karyotyping mixture was adjusted by increasing the amount of probes that show weaker signals or decreasing the probes that show stronger signals to capture the weakest signals with the more intense one without interference.

FISH Procedure.

After the cell spreads were dried on slides, they were UV-crosslinked for 2 min (total energy, 120 mJ_cm2) and fixed in 10% formaldehyde solution for 5 min. Slides were washed sequentially in water and 100% ethanol and then dried. At the center of the cell spreads, 3 μl of 2X SSC solution containing autoclaved salmon-sperm DNA (1 μg/μl) were dropped. After application of a mineral-oil-coated plastic coverslip, the slide preparation was denatured by being placed on awet paper towel in an aluminum tray floating in boiling water (100°C) for 5 min. The slides were cooled immediately on a

metal plate placed on ice. After removing the plastic coverslip, denatured (100°C, 5 min) and rapidly cooled probe mixture (5 µl, in 2X SSC/1 mM EDTA) was applied. After reapplication of the plastic coverslip, the slides were incubated at 55°C overnight in a humidity chamber containing 2X SSC soaked paper toweling. Slides were washed in 2X SSC for 20 min at 55°C. After a brief wash with PI buffer (0.2 M NaH2PO4, pH 7.8/0.1% Igepal CA-630; I-3021; Sigma), the slides were mounted with Vectashield mounting medium (Vector Laboratories) without counter stain.

Image Capture and Data Processing of FISH Images.

Chromosome spreads were identified by using an oil lens (X25 magnification) and a triple band-pass filter of a Universal microscope (Zeiss). FISH images were captured by an Optronics MagnaFire chargecoupled device (CCD) camera and plan apo oil lenses (X100 objective for mitotic cells, and X63 objective for meiotic cells). Four single channel (blue, green, red, and infrared) images were captured in 8-bit depth black and white and were later superimposed in PHOTOSHOP 7.0 (Adobe Systems, San Jose, CA).

Assignments for color of each channel after being superimposed were the same as the fluorochrome color used (blue, coumarin; green, fluorescein or Oregon green; and red, Texas red), except for Cy5, for which a white color was assigned. Computerized background subtraction and image-feature intensification were conducted by using the ‘‘curve’’ and ‘‘layer overlay’’ command in PHOTOSHOP 7.0 to make the weak signals recognizable. Aligned and paired-chromosome figures were generated by using the ‘‘cut and paste’’ and ‘‘rotation’’ commands in PHOTOSHOP 7.0.

Results

A total of 1,000 DNA fragments in the initial PCR library were examined by the FISH procedure applied to root tips of the inbred line Oh43. Among them, 10% of the fragments showed signals at knobs, 2% showed signals at CentC regions (16), 2% showed signals at the NOR on chromosome 6 (17), 1.6% showed signals at positions on chromosomes 1, 2, and 4 (a microsatellite cluster), and 0.5% showed centromere-specific retrotransposon of maize-like (25) signals. Of the fragments, 75% labeled most of the length of all chromosomes after tyramide amplification or had a distinct enhancement in the regions around the centromeres in what we term a ‘‘centromere-diffuse’’ pattern. The remaining 10% did not show any signal. For the subtracted PCR library, 1,100 fragments were analyzed on the B73 X Mo17 hybrid (with one to two B chromosomes). Among them, 3.5% were present at knob positions, 1.6% showed signals at the NOR, 0.7% showed signals at CentC regions, 0.4% showed centromere patterns different from CentC, 0.3% were present at microsatellite positions, and 0.3% corresponded to TR-1 sites (chromosomes 4 and 6). In addition, one fragment was specific to chromosome 2 (2-3-3), one fragment showed subtelomeric signals (4-12-1), two fragments were more intense on the B than A chromosomes, and two fragments showed labeling along the length of the A chromosomes but with a diminished hybridization on the B; 65% showed nonspecific or centromere diffuse signals, and ~30% were not detectable.

Among these FISH-positive fragments, 50 were cloned and sequenced for the potential use in karyotyping (Table 1) (Gen- Bank accession nos. CL569181–CL569242). The four clones that showed chromosome 1-, 2-, and 4-specific patterns proved to be

difficult to sequence; however, the partial results indicated that they are (TAG)n simple sequence repeats. Sequencing results of clone 2-3-3, which showed a chromosome 2-specific signal, determined this fragment to be part of a 5S rRNA gene (26). Most centromere signals that showed patterns different from those of CentC were homologous to centromere-specific retrotransposon of maize (25, 27) or Cinful1 retrotransposons. The 32 cloned centromere-diffuse sequences were related to various DNA sequences, including segments of maize BACs or retrotransposable elements. The chromosome-specific subtelomere sequence 4-12-1 was homologous to the maize subtelomere sequence (GenBank accession no. S46925) (28) or the pMTY7SC2 maize telomere-associated sequence (GenBank accession no. U39641) (22). We selected the microsatellite 1-26-2 clone, the 5S 2-3-3 clone, and the subtelomeric 4-12-1 clone for FISH karyotyping in combination with the following other repeated sequences: knob, TR-1, CentC, pMTY9ER telomereassociated sequence (22), Cent4, and NOR-173.

Distribution of Each Sequence.

The chromosomes of the inbred lines Oh43 (29) and KYS (9, 10) have been characterized for the position of heterochromatic knobs, 5S sequence, size, arm ratio, and the presence of the secondary constriction (chromosome 6). Based on these and other articles that describe the distribution of the sequences of Cent4 (21) and TR-1 (23), the FISH signal distributions on Oh43 of the repeated sequences used were determined as follows (Fig. 1). Knob 180-bp repeat: 1S (small), 2L, 4L, 5L, 6SL, 7L, 8S (very small), and 9S; TR-1: 2L, 4L, and 6S; microsatellite 1-26-2 clone: 1L, 2SL, and 4S; CentC: present at all centromeres, and chromosomes 7 and 8 have the strongest signals; 5S 2-3-3

clone: 2L; Cent4: chromosome 4 primary constriction; NOR-173 clone: 6S; subtelomeric 4-12-1 clone: 2SL, 4SL, 5S, and 8L; and pMTY9ER telomere-associated sequence: 2SL, 3SL, 4SL, 5S, 7S, and 8L.

All somatic chromosomes showed distinctive staining patterns, and chromosome numbers were identified (Figs. 1 and 2). In meiotic cells of Oh43, all 10 chromosomes are identifiable by using the same hybridization mixture, which helped confirm the location of the hybridization sites. In the late prophase I stage, the chromosomes are well separated and all 10 chromosome pairs can be recognized (Fig. 3). At metaphase I, identification of all chromosome pairs is possible (Fig. 3).

To test the applicability of this system for identifying each chromosome in other varieties, many of the commonly used inbred lines were examined. In the root-tip chromosome spreads of 14 tested lines, distinguishing all 10 chromosomes was possible by using this multicolor FISH procedure (Fig. 4). The features of the each chromosome are as follows.

Chromosome 1. The large microsatellite signals are at the middle of the long arm (red in Fig. 4). Very small knob signals (blue in Fig. 4) are at the tip of the short arm. It is the longest chromosome.

Chromosome 2. The 5S signals (yellow in Fig. 4) are at the long-arm tip. Microsatellite signals (red in Fig. 4) are present at the short-arm tip and middle of the long arm. TR-1 (white in Fig. 4) or knob (blue in Fig. 4) signals are sometimes present in the long arm.

Chromosome 3. The telomere-associated pMTY9ER signal (red in Fig. 4) is present at the ends of both chromosome arms, and 4-12-1 subtelomeric signals are absent.

Chromosome 4. The Cent4 signals (red in Fig. 4) are present at the primary constriction. Microsatellite signals (red in Fig. 4) are present in the short arm. Some lines carry knob (blue in Fig. 4) or TR-1 signals (white in Fig. 4) in the long arm.

Chromosome 5. The subtelomeric 4-12-1 signals (green in Fig. 4) are present on the tip of the short arm and knob (blue in Fig. 4) in the long arm. The arm ratio is 1:1.

Chromosome 6. The NOR signal (green in Fig. 4) is present in the short arm, and knob signals (blue in Fig. 4) are present at both chromosome ends, which are sometimes obscured by adjacent signals. TR-1 signals (white in Fig. 4) are present at the tip of short arm.

Chromosome 7. The telomere-associated pMTY9ER signals (red in Fig. 4) are present at the tip of the short arm. This chromosome tends to have larger CentC signals (green in Fig. 4) and large knob signals (blue in Fig. 4) in the long arm.

Chromosome 8. The 4-12-1 subtelomeric signals (green in Fig. 4) are at the tip of the long arm. This signal is invariant in the lines examined. The arm ratio is 1:3.

Chromosome 9. Knob signals (blue in Fig. 4) are always present at the tip of the short arm but vary in size.

Chromosome 10. This chromosome is the smallest. There are no landmarks on this chromosome, except for abnormal 10 in the K10 line.

There is significant variation for the presence and size of the repeated sequences among lines. The presence of knobs (blue in Fig. 4) is variable on chromosomes 2, 4, 5, and 8 among lines. TR-1 signals (white in Fig. 4) are variable on chromosomes 2, 4, and 6L. Chromosome 9 in stock 6 has a large TR-1 and knob hybridization in the short arm. Microsatellite signals (red in Fig. 4) are variable in the chromosome 6 centromere region

and on the long or short arms of chromosome 5. The size of the CentC signals (green in Fig. 4) is different among chromosomes as well as among lines. All chromosomes of the 14 lines tested have a chromosome 9 short-arm knob, although the size of this knob in B37, BMS, K10, Mo17, and W23 is small. Chromosome 1 short-arm knobs are found consistently, except in BMS and Mo17. The 4-12-1 subtelomeric signals on the chromosome 8 long-arm tip are detected consistently in all of the tested lines. The telomere-associated pMTY9ER signals in the short arm of chromosome 7 are weak and variable. In A188, A632, B37, B73, W22, and W23, the 7S subtelomeric signals are undetectable. Also, these signals on chromosome 3 are weak and variable. Despite these variations, all somatic chromosomes are identifiable in the examined lines based on the conserved signals by using the current multicolor FISH technique.

Discussion

The FISH screening of the random PCR library proved to be effective for separating repeated sequences from the maize genome. Most of the major repeated sequences, knob, CentC, NOR, 5S, TR-1, centromere-specific retrotransposon of maize, microsatellite, and 4-12-1 subtelomeric repeats were separated during the screening. However, Cent4, telomere (30), and pMTY9ER telomere-associated sequences were not found. This result implies that there might be other remaining unknown repeated sequences that could be useful for maize karyotyping. Subtraction was effective in reducing the known repeated sequences from the population of the original PCR library. Judging from the number of PCR fragments that showed knob signals, the process eliminated 70–80% of these sequences.

The current system contains a sufficient number of probes to distinguish all chromosomes in commonly used lines of maize. It provides a baseline to which additional probes can be added. One possibility involves BACs, which are used as probes in other organisms, such as *Arabidopsis thaliana* (31), rice (*Oryza sativa* L.) (32), and sorghum [*Sorghum bicolor* (L.) Moench] (33), as well as sorghum BACs onto maize chromosomes (34). However, detection of maize BACs on chromosome spreads is difficult because of the low gene density (13). Other repeated sequences or tandem gene arrays may provide additional features that could aid in distinguishing chromosomes for the further refinement of the probe collection. The procedure could be adapted to other species by the isolation of a respective collection of repetitive DNA sequences.

The morphologies of chromosomes 2–4 resemble each other, as do chromosomes 7–9 (Fig. 4). Distinguishing these chromosomes cannot be achieved reliably without the use of unique banding patterns. Changes in chromosome arm length can be affected 5–20% by the polymorphism of knobs (Fig. 4) as well as chromosome-preparation procedures. The multicolor FISH procedure described here permits the distinction of these chromosomes.

Examination of the various inbred lines revealed significant variation for many of the repetitive gene arrays examined. The copy number of the respective sequences is apparently subject to change by mechanisms that are currently unknown. The probe collection described here provides the tools to examine this variability throughout other maize lines and the mechanism by which this variation arises.

For detailed structural analysis of chromosomes, the pachytene stage of meiosis is unparalleled because of the less condensed chromatin that permits high resolution of

chromosomal features; however, the ease of observation at this stage is genotype-dependent (5). The technique described here is effective for identifying all maize somatic chromosomes in a wide variety of lines and allows such analysis in a genotypeindependent manner. Somatic chromosomes are typically better spread, permitting greater ease of identification. Also, it is reasonable to design experiments involving hundreds of individuals that can be analyzed in a relatively short period. The use of root tips for karyotyping has great advantages because studies can be completed shortly after germination and individuals of interest can be retained and grown to maturity for future analysis. This system can be used for detection of chromosomal aberrations, determination of specific chromosomes involved in aneuploidy, detection of variation of repetitive sequences in the genome, analysis of chromosomal behavior in mitosis and meiosis, localization of large transgenes to chromosomal region, and many other applications for which chromosomal identification is useful.

Acknowledgements

We thank R. L. Phillips (University of Minnesota, Saint Paul) for providing the CentC clone; L. Dennis (CSIRO Plant Industry, Canberra, Australia) for the knob clone; J. Gardiner (University of Arizona, Tucson) for the pMTY9ER telomere-associated sequence. We also thank J. M. Vega (University of Missouri, Columbia) for providing B73 _ 4 B plant tissue; E. H. Coe (University of Missouri, Columbia) for stock 6; S. Melia-Hancock (University of Missouri, Columbia) for inbred lines A632, B37, and W23; and J. Eta-ndu (University of Minnesota, Saint Paul) for A188; T. L. Phelps-Durr (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY) and J. L. Cooper (University

of Washington, Seattle) for comments on the cloning procedure; T. Wako (National Institute of Agrobiological Sciences, Tsukuba, Japan) for information about the tyramide amplification system; and D. L. Auger (South Dakota State University, Brookings) and T. Ream (Washington University, Saint Louis) for discussions. This work was supported by the Monsanto Company and in part by National Science Foundation Grant DBI 9975827.

References.

1. McClintock, B. (1930) *Proc. Natl. Acad. Sci. USA* **16,** 791–796.
2. Longley, A. E. (1939) *J. Agric. Res.* **59,** 475–490.
3. Rhoades, M. M. (1950) *J. Hered.* **41,** 58–67.
4. Neuffer, G. M., Coe, E. H. & Wessler, S. (1997) in *Mutants of Maize* (Cold Spring Harbor Lab. Press, Plainview, NY), pp. 32–53.
5. Carlson, W. R. (1988) in *Corn and Corn Improvement*, eds. Sprague, G. F. & Dudley, J. W. (Am. Soc. Agronomy, Crop Science Soc. America, and Soil Science Soc. America, Madison, WI), 3rd Ed., pp. 259–331.
6. McClintock, B. (1950) *Proc. Natl. Acad. Sci. USA* **36,** 344–355.
7. Coe, E. H. (1993) in *The Maize Handbook*, eds. Freeling, M. & Walbot, V. (Springer, New York), pp. 364–376.
8. Beckett, J. B. (1978) *J. Hered.* **69,** 27–36.
9. Chen, C. C., Chen, C. M., Hsu, F. C., Wang, C. J., Yang, J. T. & Kao, Y. Y. (2000) *Theor. Appl. Genet.* **101,** 30–36.
10. Sadder, M. T. & Weber, G. (2001) *Plant Mol. Biol. Rep.* **19,** 117–123.
11. Longley, A. E. & Kato, Y. T. A. (1965) *Res. Bull. CIMMYT (Chapingo, Mexico)* **1,** 1–112.
12. Wells, D. & Levy, B. (2003) *BioEssays* **25,** 289–300.
13. Meyers, B. C., Tingey, S. V. & Morgante, M. (2001) *Genome Res.* **11,** 1660–1676.
14. Wiegant, J., Verwoerd, N., Mascheretti, S., Bolk, M., Tanke, H. J. & Raap, A. K. (1996) *J. Histochem. Cytochem.* **44,** 525–529.
15. Peacock, W. J., Dennis, E. S., Rhoades, M. M. & Pryor, A. J. (1981) *Proc. Natl. Acad. Sci. USA* **78,** 4490–4494.
16. Ananiev, E. V., Phillips, R. L. & Rines, H. W. (1998) *Proc. Natl. Acad. Sci. USA* **95,** 13073–13078.
17. Stein, N., Ponelies, N., Musket, T., McMullen, M. & Weber, G. (1998) *Plant J.* **13,** 281–289.
18. Alfenito, M. R. & Birchler, J. A. (1993) *Genetics* **135,** 589–597.
19. Rhoades, M. M. & Dempsey, E. (1966) *Genetics* **53,** 989–1020.
20. Kato, A. (1999) *Biotech. Histochem.* **74,** 160–166.
21. Page, B. T., Wanous, M. K. & Birchler, J. A. (2001) *Genetics* **159,** 291–302.
22. Gardiner, J. M., Coe, E. H. & Chao, S. (1996) *Genome* **39,** 736–748.
23. Ananiev, E. V., Phillips, R. L. & Rines, H. W. (1998) *Proc. Natl. Acad. Sci. USA* **95,** 10785–10790.
24. Chen, C., Yan, H., Zhai, W., Zhu, L. & Sun, J. (2000) *Genome* **43,** 181–184.
25. Zhong, C. X., Marshall, J. B., Topp, C., Mroczek, R., Kato, A., Nagaki, K., Birchler, J. A., Jiang, J. M. & Dawe, R. K. (2002) *Plant Cell* **14,** 2825–2836.
26. Pan, Y. B., Burner, D. M. & Legendre, B. L. (2000) *Genetica* **108,** 285–295.
27. Nagaki, K., Song, J., Stupar, R. M., Parokonny, A. S., Yuan, Q., Ouyang, S., Liu, J., Hsiao, J., Jones, K. M., Dawe, R. K., Buell, C. R. & Jiang, J. (2003) *Genetics* **163,** 759–770.
28. Burr, B., Burr, F. A., Matz, E. C. & Romero-Severson, J. (1992) *Plant Cell* **4,** 953–960.

29. Kakeda, K., Yamagata, H., Fukui, K., Ohno, M., Fukui, K., Wei, Z. Z. & Zhu, F. S. (1990) *Theor. Appl. Genet.* **80,** 265–272. 30. Cox, A. V., Bennett, S. T., Parokonny, A. S., Kenton, A., Callimassia, M. A. & Bennett, M. D. (1993) *Ann. Bot.* **72,** 239–247.
31. Lysak, M. A., Pecinka, A. & Schubert, I. (2003) *Chromosome Res.* **11,** 195–204.
32. Jiang, J., Gill, B. S., Wang, G. L., Ronald, P. C.&Ward, D. C. (1995) *Proc. Natl. Acad. Sci. USA* **92,** 4487–4491.
33. Islam-Faridi, M. N., Childs, K. L., Klein, P. E., Hodnett, G., Menz, M. A., Klein, R. R., Rooney, W. L., Mullet, J. E., Stelly, D. M. & Price, H. J. (2002) *Genetics* **161,** 345–353.
34. Koumbaris, G. L. & Bass, H. W. (2003) *Plant J.* **35,** 647–659. Kato *et al*. PNAS

Table 1. Cloned FISH positive DNA sequences separated by root-tip screening

FISH signals	Clone no.	Homology to known sequences
Chromosomes 1, 2, and 4	1–26-2, 4–12-4, 4–12-6, and 4–12-12	Microsatellite TAG repeat
Chromosome 2	2–3-3	5S
Chromosomes 2, 4, and 6	4–12-7	TR-1
Centromere	1–26-68, 4–12-16, 4–12-22, 6–9-26, 6–9-34, 12–29-1, and 12–29-6	CRM or Cinful1
Centromere diffuse	4–12-18, 4–12-19, 6–9-1, 6–9-2, 6–9-3, 6–9-4, 6–9-5, 6–9-6, 6–9-7, 6–9-9, 6–9-10, 6–9-11, 6–9-12, 6–9-13, 6–9-14, 6–9-16, 6–9-17, 6–9-18, 6–9-19, 6–9-20, 6–9-21, 6–9-22, 6–9-23, 6–9-24, 6–9-25, 6–9-27, 6–9-28, 6–9-29, 6–9-30, 6–9-31, 6–9-36, and 12–2-7	Various sequences, most are homologous to partial sequences of known maize BACs
Centromere diffuse, B chromosome long arm enriched	4–12-20 and 4–12-21	21 4–12-20 is homologous to the maize BAC ZM16H10. 4–12-21 shows no homology to known sequences.
Uniform in A chromosome and less in B chromosome	4–12-9 and 4–12-10	Rire-2 and Huck-2
Chromosome-specific subtelomeric	4–12-1	Gardiner's telomere associated sequence (22), Burr's subtelomere (28)
Clones that showed signals at the positions of 180-bp knob, CentC, and NOR are not included		

Figures

Figure 1. Probe localization on maizeOh43root-tip chromosomes.
(*a*) Microsatellite 1-26-2 clone (red), CentC (green), TR-1 (white), and knob 180-bp (blue) signals.
(*b*) The 5S 2-3-3 clone (yellow), Cent4 (red), NOR-173 clone (green), and knob180-bp (blue) signals. (*c*)pMTY9ERtelomere-associated sequence (red), subtelomeric 4-12-1 clone (green), and knob 180-bp (blue) signals. (*d*) The signals of all nine probes. (Scale bar, 10 μm.)

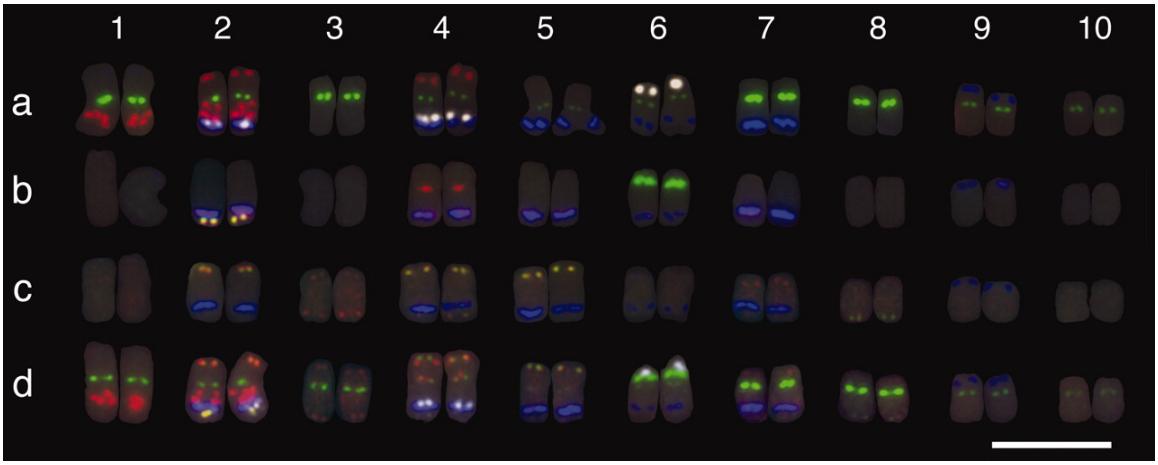


Figure 2. Somatic-chromosome identification in four maize inbred lines probed with the FISH mixture described in the text. (Scale bar, 10 μm.)

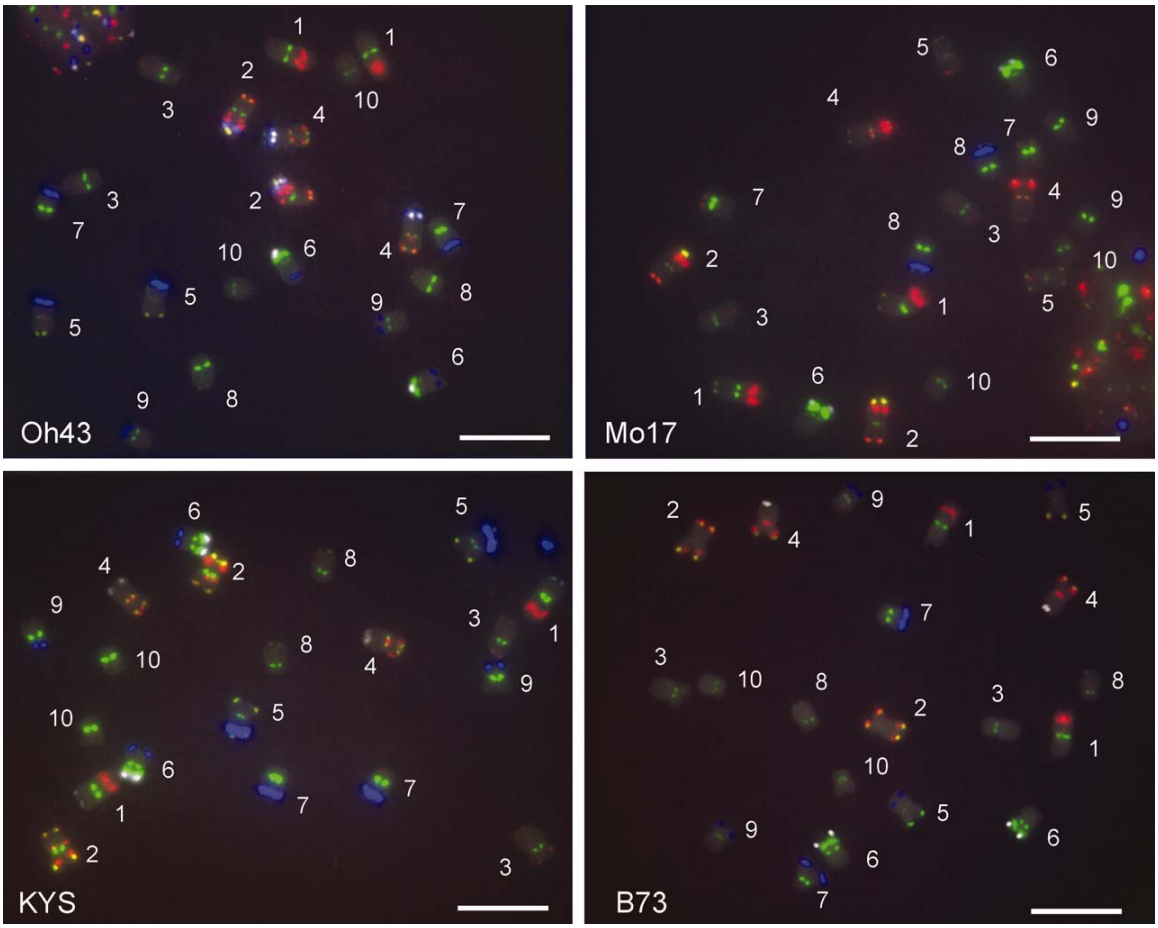


Figure 3. FISH signals on maize Oh43 meiotic cells. (*Upper*) Late prophase I. All 10 chromosome pairs are identifiable. (*Lower*) Metaphase I. Chromosomes are identifiable from the signal combinations. (Scale bar, 10 μm.)

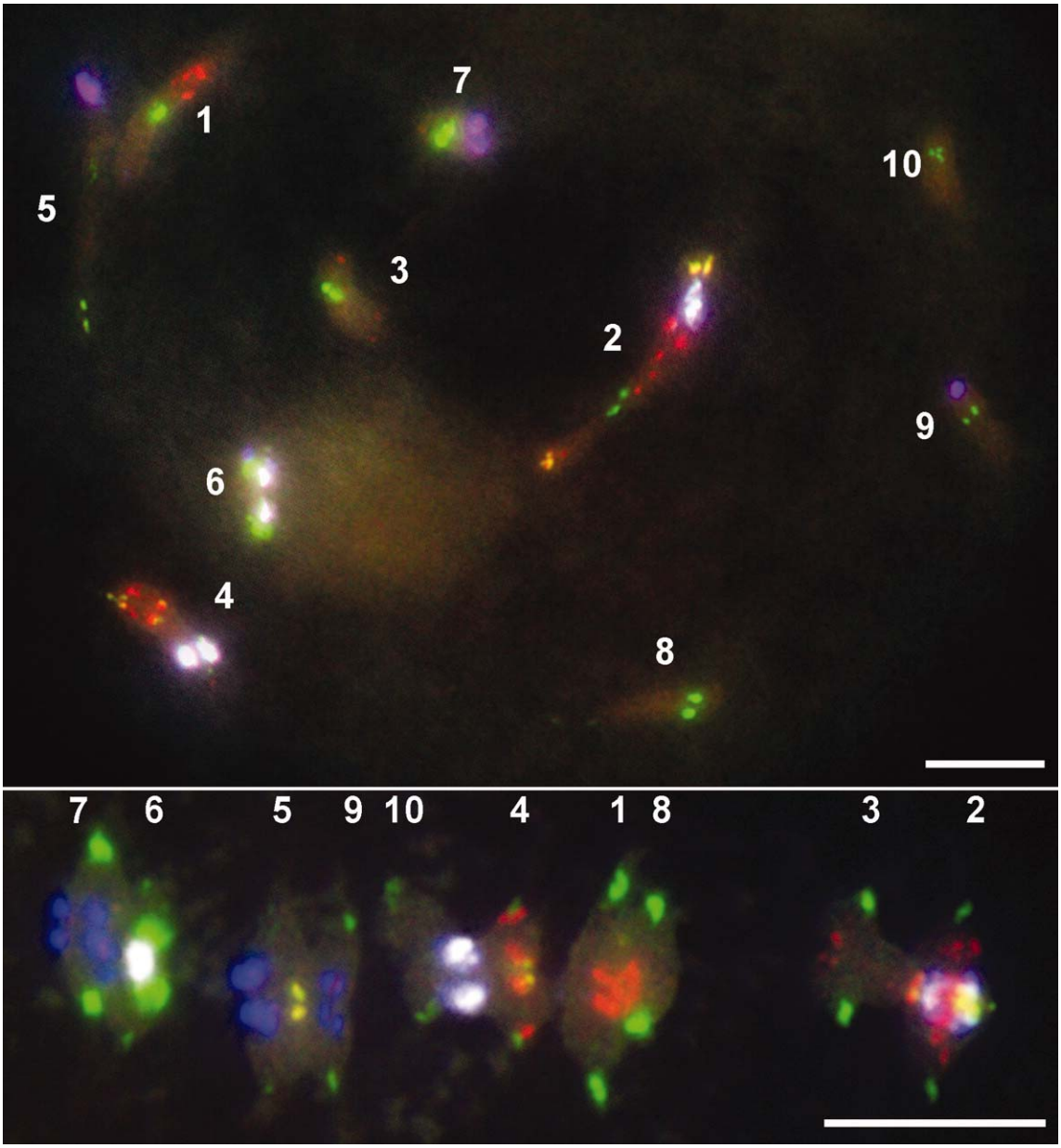


Figure 4. Somatic chromosome karyotyping of 14 maize lines probed with the FISH mixture. Knob180-bp repeat (blue), 5S 2-3-3 (yellow, 2L), NOR-173 clone (green, 6S), CentC (green), subtelomeric 4-12-1 clone (green), Cent4 (red, 4C), microsatellite 1-26-2 clone (red), pMTY9ER telomere-associated sequence (red), and TR-1 (white). (Scale bar, 10 μm.)

3. SEQUENCES ASSOCIATED WITH A CHROMOSOME CENTROMERES ARE PRESENT THROUGHOUT THE MAIZE B CHROMOSOME

This article was published in the journal Chromosoma and the reference style of that journal was retained.

LAMB, J. C., A. KATO and J. A. BIRCHLER, 2005 Sequences associated with A chromosome centromeres are present throughout the maize B chromosome. Chromosoma **113:** 337-349.

Sequences associated with A chromosome centromeres are present throughout the maize B chromosome

Jonathan C. Lamb[1], Akio Kato[1], James A. Birchler[1,2]

[1]Division of Biological Sciences, University of Missouri - Columbia

[2]Communicating Author:

James A. Birchler

Professor of Biological Sciences

University of Missouri

117 Tucker Hall

Columbia, MO 65211

(573) 882-4905 (voice)

(573) 882-0123 (FAX)

BirchlerJ@Missouri.edu

Key Words: B chromosome, centromere, repetitive elements, FISH, maize

Abstract:

Maize chromosome spreads containing the supernumerary B chromosome were hybridized with probes from various repetitive elements including CentC, CRM and CentA, which have been localized to centromeric regions on the A chromosomes. Repetitive elements that are enriched or found exclusively near the centromeres of A chromosomes hybridized to many sites distinct from the centromere on the B chromosome. To examine if these

elements recruit kinetochore proteins at locations other than the canonical B centromere, cells were labeled with antibodies against CENH3, a key kinetochore protein. No labeling was detected outside the normal centromere and no evidence of B chromosome holocentromeric activity was observed. This finding suggests that, as in other higher eukaryotes, DNA sequence alone is insufficient to dictate kinetochore location in plants. Additionally, examination of the B centromere region in pachytene chromosomes revealed that the B-specific element ZmBs hybridizes to a much larger region than the site of hybridization of CentC, CRM, and CentA and the labeling by anti-CENH3 antibodies.

Introduction

The largest of the selfish DNA elements takes the form of dispensable chromosomes, called B chromosomes, that in many species are found in addition to the required complement (reviewed in Jones and Rees 1982). Because B chromosomes do not confer advantages to organisms that harbor them, they may be thought of as parasitic chromosomes that persist in populations by making use of cellular machinery required for normal chromosome maintenance and passage (reviewed in Puertas 2002; Jones and Houben 2003). Therefore, the means by which B chromosomes arise and the mechanisms that allow them to persist have been studied in an effort to understand how regular chromosomes are maintained (e.g. Berdnikov et al. 2003; Cheng et al 2000; Dhar et al. 2002; Kaszas and Birchler 1998).

Although the maize B chromosome was the first B chromosome to be found (Longley 1927), its origin, as well as the molecular basis of the mechanisms that allow it to survive in maize populations, is still not clear. Efforts to understand the nature and origin of the maize B chromosome have led researchers to isolate repetitive DNA that is unique to the B chromosome (Alfenito and Birchler 1993; Cheng and Lin 2004; Stark et al. 1996) potentially suggesting an origin for this chromosome outside the current maize genome. However, it is apparent that most of the DNA on the B chromosome is similar to that of the A chromosomes (Chilton and McCarthy, 1973; Alfenito and Birchler 1993; Cheng and Lin 2003; Stark et al. 1996). For example, using DNA from inbred lines lacking the B chromosome as a probe in GISH preparations reveals hybridization along most of the B chromosome (Stark et al. 1996). The finding that the majority of the DNA on the B chromosome is common to DNA of the A chromosomes supports the hypothesis that the B chromosome may be a degenerated form of an A chromosome or at least that the B chromosome has been present in the maize genome for multiple rounds of repetitive element expansion.

In previous work, a collection of DNA sequences from repetitive elements that can be used as probes in FISH preparations to identify each of the chromosomes in the maize karyotype was developed (Kato et al., 2004a). In this study, we have used this collection to expand on the previous GISH approach by hybridizing individual DNA elements to chromosome spreads containing B chromosomes. The rationale was that if the B chromosome

originated from a degenerated A chromosome, the hybridization pattern of repetitive elements on the B chromosome could reflect the chromosome(s) of origin, depending on the complexity of rearrangements and alterations that took place during the formation of the B chromosome.

As these patterns of hybridization were determined, it was observed that DNA elements normally found exclusively at centromeric locations were abundant on the B chromosome in locations distinct from the B centromere. The presence of centromeric elements outside of the canonical B centromere allowed us to examine the role of these DNA elements in chromosome movement and centromere formation.

In mammals and some other eukaryotes, translocations that result in a chromosome with two centromeres have been observed (reviewed in Sullivan et al. 2001). This situation creates instability as the two centromeres can move to opposite poles during cell division ripping the chromosome in the process. In mammals, chromosomes with two centromeres can be stabilized by inactivation of one of them. Additionally in mammals, de novo centromere formation can occur at chromosomal locations that contain no sequences in common with other centromeres. In plants, however, neither inactive centromeres nor true *de novo* centromere formation has been reported, despite the creation of unstable dicentric chromosomes and acentric fragments (Zheng et al. 1999; Kato et al. 2004b). Also, sequence analysis of Arabidopsis centromeric repeats shows regions of conservation across ecotypes suggesting selection pressure acts on the primary sequence (Hall et al. 2003b). Taken

together, these observations raise the possibility that plant centromere determination is more reliant on primary DNA sequence than in mammals or other eukaryotes.

Because of the high concentration of centromeric sequences at many locations on the B chromosome, we tested these sequences for association with proteins and activities normally found at centromeres. No centromeric function was found for the centromeric sequences present at sites other than the commonly recognized position of the centromere. This finding indicates that in plants, as in other higher eukaryotes, DNA sequence of specific elements alone is insufficient to organize centromere function.

Materials and Methods

B73 lines containing B chromosomes were used as the tissue source for all experiments. Meiotic images were taken from material that contained 2 B chromosomes. Preparation of labeled DNA probes, meiotic and mitotic chromosome spreads and hybridization for FISH analysis were performed essentially as in Kato et al. (2004a). The pZmBs probe was made in the same manner as other probes in Kato el al. (2004a) using the sequence, "ZmBs (Knob-)," that was used in Hsu et al. (2003). This plasmid was derived by Tara Phelps-Durr from the ZmBs sequence by removal of the major region of homology to the 180bp knob repeat. Probes against the LTRs of Huck and Prem1 were derived from plasmids provided by J. Bennetzen (University of Georgia). The probe against CRM was from a plasmid provided by Kelly

Dawe (Zhong et al. 2002) and the probe against CentA was from the LTR region on a plasmid containing CentA (Ananiev et al. 1998). Subclones ZMABC19 and ZMABC91 (Nagaki et al. 2003) were provided by Jiming Jiang (University of Wisconsin).

Immunolabeling involving meiotic chromosomes was performed as described in Kaszas and Cande (2000) except that after the secondary antibody was washed, 10% formaldehyde was applied for 10 min then rinsed several times with phosphate buffered saline (PBS) prior to denaturing and application of labeled DNA probe. The anti-CENH3 antibody (Zhong et al. 2002) was provided by K. Dawe (University of Georgia).

Root tip chromosome spreads for immunolabeling were prepared by placing root tips that were 1.5 cm in length into a pressure chamber containing nitrous oxide at 10 atm for 3 hours then fixing in ice cold 4% paraformaldehyde in one of two solutions, either PBS or Buffer A (Zhong et al. 2002) followed by three 5 minute washes in the same solution without paraformaldehyde. Both PBS and Buffer A worked equally well. Fixed root tips were stored in 100% methanol at –20C. Root tips were then soaked in PBS and the meristematic portion cut out with a scalpel and digested in 0.5% pectinase, 1% cellulase in citrate pH 5, for 45 minutes at 37C. The digestion mixture was removed and the root tip suspended in a small amount (100 ul) of PBS and the root tip was broken apart with a probe. The cells were then thoroughly separated from one another by pipetting up and down several times. Finally, the cells were dropped onto polylysine coated slides and spun

at 150g for 4 minutes in a swinging bucket style centrifuge. Immunolabeling was as described in Manzanero et al. (2000) using rabbit antibodies against CENH3 or phosphorylated H3-Ser10 (polyclonal rabbit from Upstate Biotechnology). After the secondary antibody (goat anti-rabbit) was washed, 75ul of 10% formaldehyde was applied for 10 minutes followed by a brief immersion in distilled water. Slides were then allowed to air dry for at least 1 hour in the dark before DNA probes were applied in the same manner as above.

For images involving only DNA probes, FISH images were captured either by an Optronics MagnaFIRE CCD camera mounted on a Zeiss Universal microscope using 100X or 63X plan apo oil immersion lenses or by a Sensys CCD camera using a 60X plan apo oil immersion lens on an Olympus IX70 inverted microscope with a motorized focus controlled by ImagePro Plus software with a Vaytek Volume Scan plug in. The images were processed as described in Kato et al. (2004a). In cases for which relative positions of probes were to be determined, a few representative images were captured using the Olympus microscope and system to confirm locations because no manual aligning of colors is required on this system. Images involving antibodies were all obtained with the Olympus microscope by capturing a Z series that was processed using Metamorph. Stacks were examined for colocalization of CENH3 labeling and DNA hybridization followed by flattening for presentation.

All results reported represent 5-50 good chromosome spreads examined from at least two individuals. Results from CentC, CRM, ZmBs, and antibody labeling of CENH3 represent at least 40 chromosome spreads obtained from at least four different individuals. Examination of meiotic anaphase/telophase spreads for "holocentromeric" activity involved at least 40 chromosome spreads.

Results:

Hybridization of repetitive elements on the maize B chromosome.

To aid in describing the hybridization patterns of the various elements used as FISH probes, a typical B chromosome and a diagram showing the hybridization patterns of various repetitive elements on the B is shown in Figure 1. The B chromosome consists of a diminutive short arm, a centromere, a proximal block of heterochromatin, a stretch of euchromatin, called the proximal euchromatin, followed by four blocks of heterochromatin, and the distal-most euchromatin (see Figure 1).

The following is a list of elements that were used as probes on spreads containing the B chromosome and a description of their hybridization pattern. These data are summarized in Table 1 and presented graphically in Figure 1.

The ZmBs B specific repeat:

Using a probe made from the ZmBs repeat lacking the region of strong homology to the 180bp knob repeat (Alfenito and Birchler 1993), six distinct regions of ZmBs hybridization were observed: two areas of intense

hybridization flanking the CENH3 binding region (discussed below), the distal half of the proximal heterochromatic block, the fourth block of distal heterochromatin, a minor site in the middle of the proximal euchromatin and a minor site in the distal euchromatin.

CentC:

The DNA element CentC is ~156bp in length and is found in megabase sized tandem arrays that are visualized at the primary constriction of maize chromosomes in FISH preparations (Ananiev et al. 1998). CentC hybridizes to multiple regions along the length of the B chromosome including the CENH3 binding region between two regions of intense ZmBs hybridization, the second half of the proximal heterochromatin, throughout the first block of heterochromatin, at a distinct, intense site in the second distal block, throughout the third block, and in a distinct site near the end of the B chromosome, distal to the ZmBs cluster of the fourth heterochromatic region (Figure 2 and Figure 3). The intensity of hybridization in each heterochromatic region is greater than the intensity of the A centromeres with weak signal

CRM and CentA:

The maize retroelements CentA (Ananiev et al. 1998) and CRM (Zhong et al. 2002) are members of a class of Ty3/Gypsy retroelements that are present at the centromeres of grass species (Miller et al. 1998). CRM is located at the CENH3 binding domain between two regions of intense ZmBs hybridization and the second half of the proximal heterochromatin, co-

localizing with CentC in double labeled preparations (Figure 2 and Figure 3). A minor site of hybridization can be detected on the B chromosome distal to the ZmBs signal of the fourth block of heterochromatin and proximal to the most distal CentC signal (Figure 2). Detection of signal at this distal location is dependent on the CRM probe used and upon the detection method. Using a probe labeled with all four fluorescent nucleotides to detect the CRM LTR and part of the UTR, the distal B chromosome signal and sites of hybridization on the A chromosomes in specific non-centromeric locations are consistently observed. Confirmation of these extracentromeric signals was obtained by comparing CRM to sequences in Genbank using the BLAST program (Altschul et al. 1990). Matches are observed that are from DNA sequences located at positions other than the centromere (accessions #AF090446, #AC144717).

CentA hybridizes to the functional centromere and the proximal heterochromatin in the same approximate location as CentC and CRM. In general, the CentA probe produces high background and hybridization in a pattern similar to the "Centromere Diffuse" elements (discussed below) when stringency conditions are sub-optimal (data not shown).

"Centromere Diffuse" elements:

In addition to elements that are located predominantly at the primary constriction, we examined the location of DNA sequences that exhibit a "Centromere Diffuse" pattern of hybridization (Kato et al 2004a). These elements hybridize along most of the length of the A chromosomes but show a

distinct pattern of enrichment in the pericentromeric regions. They include retroelements (such as Cinful1), and DNA elements of unknown origin (Kato et al. 2004a; Nagaki et al 2003).

Signals from the "Centromere Diffuse" elements from Kato et al. (2004a) ("4-12-18," "4-12-19," "6-9-11," "6-9-12," "6-9-13," "6-9-19," "6-9-21," "6-9-25," "6-9-30") as well as the centromeric BAC subclones ZMABC19 and ZMABC91 from Nagaki et al. (2003) are located at the proximal heterochromatin and in the four distal blocks of heterochromatin. Between the blocks of heterochromatin, the intensity of these signals is diminished and the euchromatic regions are not well labeled. One representative element, "6-9-11", is shown in Figure 4.

180bp knob repeat:

The beginning of the B chromosome proximal heterochromatin contains an area that hybridizes only to the 180bp knob repeat. The distal half of the proximal heterochromatin hybridizes to the 180bp knob repeat as well as other probes (Figure 3). Additionally, there is a region of minor knob hybridization near the ZmBs signal of the fourth block of heterochromatin (Hsu et al. 2003; present results). The arrangement of the 180bp knob repeats on the B differs from the A chromosomes in two ways: on A chromosomes, the 180bp knob repeat is not located near centromeres (McClintock et al. 1981; Kato et al. 2004a) and is generally found as a tandem array with few interruptions except TR1 (Mroczek and Dawe 2003).

TR1:

TR1 was previously reported to hybridize weakly to the proximal heterochromatin in a similar location as the 180bp knob repeat (Hsu et al., 2003). We did not observe any TR1 hybridization on the B chromosome. However, it should be noted that despite amplification, the TR1 signal reported by Hsu et al. (2003) was very weak.

The "AGT" microsatellite:

On the A chromosomes, there are several distinct sites of "AGT" microsatellite hybridization, where the sequence "AGT" is present in hundreds or thousands of copies (Kato et al, 2004a). On the B chromosome, the "AGT" microsatellite hybridizes strongly to a site at the beginning of, and moderately to a site at the end of, the third block of distal heterochromatin (Figure 4).

Cent4:

Cent4 is an element that is found near the primary constriction of chromosome 4 and has homology to the ZmBs and the 180bp knob repeats (Page et al. 2001). Consequently, Cent4 hybridizes weakly and in a diffuse manner to the proximal heterochromatin, as well as the knob sites that contain the 180bp repeat on the A chromosomes. In addition to the hybridization to the B centromeric regions, there is a small site of hybridization that partially but not entirely overlaps the ZmBs site of the fourth heterochromatin.

Telomere repeat:

The maize telomere, (TTTAGGG)(n), consists of an Arabidopsis type telomere repeat (Gardiner et al. 1996; Fuchs et al. 1995) and has homology to

portions of the ZmBs probe (Alfenito and Birchler 1993). Hybridization is observed to colocalize with ZmBs at all sites described above for ZmBs. The intensity of hybridization at these sites varies between these two probes and the colocalization does not always precisely overlap (Figure 2).

LTR retroelements:

The distribution of three members of the Ty3/Gypsy class of retroelements that are abundant in the maize genome, Huck, Prem1, and Cinful-1 (Meyers et al. 2001) was examined. Huck hybridized more strongly to the middle of the chromosome arms while Prem1 hybridization was enriched near the centromeres of A chromosomes, although it was depleted immediately at the primary constriction of some chromosomes (Figure 5). Cinful has a distinct "centromere diffuse" pattern of hybridization. On the B chromosome, Prem1 and Cinful hybridization was more intense than the A chromosomes and was distributed in a similar pattern as the centromere diffuse elements. In contrast, Huck hybridized weakly to the B chromosome and is located primarily in the euchromatic region.

5S and 25S ribosomal genes:

The 5S and 25S ribosomal genes did not give any signal when hybridized to the B chromosome.

Because the observed hybridization of centromeric elements on the long arm of the B chromosome was unexpected, we considered the possibility that these signals result from cross-hybridization to various unknown sequences specific to the B chromosome. Under conditions of increased

stringency (washing at 70ºC), the long arm CentC signals on the B chromosome decrease in intensity, as do the signals on A centromeres, but are not eliminated. For CentA, CRM and the "Centromere Diffuse" elements, conditions were used to prepare slides that resulted in the previously observed patterns of hybridization on the A chromosomes. For example, when the centromere diffuse elements were used as probes, they show enrichment on the B (as described above) as well as the previously observed "centromere diffuse" pattern on A chromosomes. This finding serves as an internal control to the experiment showing that sequences on the B have strong homology to the sequence of the respective probes.

The presence of CentC on the B long arm led us to address whether CentC is present at low levels on other chromosome arms. In numerous mitotic and meiotic chromosomes spreads hybridized with the CentC probe, no CentC hybridization was ever observed away from the centromere. Maize genome sequences were examined for CentC homology by performing a BLAST search. No sequences were found with stretches of homology greater than 20bp outside of regions known to reside at the primary constrictions. This result is consistent with the cytological assays described above.

On The B Chromosome, CENH3 Is Located Only At The Centromere.

CENH3 is a variant of the H3 histone and is found at centromeres during all stages of the cell cycle. CENH3 is thought to serve as the attachment point of the kinetochore proteins to the chromosome (Sullivan et al.

2001). In regular maize chromosomes, nucleosomes containing the centromeric H3-histone variant, CENH3, are located over the centromeric DNA elements CentC, CRM, and CentA (Zhong et al. 2002) although not all copies of these centromeric DNA elements are involved (Jin et al. 2004). Because CentC, CRM, and CentA are located at sites distinct from the centromere on the B chromosome, meiotic and mitotic cells containing B chromosomes were labeled with antibodies against CENH3 to determine if these centromeric elements recruit kinetochore specific nucleosomes. Metaphase and interphase mitotic cells from root tips (Figure 6) and meiotic cells (Figure 7) were labeled with antibodies against CENH3. The cells have also been hybridized with ZmBs to identify the B chromosomes. CENH3 labeling is observed only at the functional centromere. In interphase cells it is not possible to follow along the B chromosome from end to end. However, ZmBs labels both the centromere as well as the distal tip of the B long arm. No trail of CENH3 labeling is observed between the ZmBs signals as would be expected if CENH3 were present along the entire arm.

In examining these preparations, a difference between the amount of labeling of the CENH3 at the centromeres of the A chromosomes and the B chromosome was observed. The B centromere shows a consistently weaker CENH3 signal than the A centromeres for both meiotic and mitotic cells (Figure 7).

Organization of Repetitive Elements at the B Chromosome Centromere

As B chromosomes were examined for the presence of CENH3 outside of the centromere, the organization of repetitive DNA elements at the B centromere and their relation to CENH3 was also determined. Figure 3 shows the location of ZmBs, CentC, CRM, CentA, and the 180bp knob repeat in relation to each other and CENH3 at pachynema. CENH3 labeling colocalizes with hybridization of CentC, CRM, CentA, and ZmBs. On either side of the CENH3 binding domain, there are two regions of intense hybridization to the ZmBs repeat, one of which encompasses the diminutive short arm. These regions of hybridization to ZmBs appear to be large arrays of ZmBs as there is little hybridization to other elements (aside from the telomeric repeat which has homology to ZmBs). After the second region of hybridization to ZmBs, there is a stretch of 180bp knob repeat that is followed by a mixture of ZmBs, CentC, CRM, CentA, the 180bp knob repeat and various "centromere diffuse" elements.

Holocentromeric Activity Is Not Observed On The B Long Arm.

In maize, clusters of specific DNA elements located away from centromeres have been shown to interact with spindle fibers during meiosis I and move toward the poles independently from the centromeres (Rhoades and Vilkomerson, 1942; Hiatt et al. 2002). We examined metaphase I and anaphase I spreads containing B chromosomes for this type of movement of the B long arm. In the majority of cases, the B long arm was clearly lagging behind the centromere (Figure 8). While in some cases the B long arm was

positioned nearly parallel to the metaphase plate, it was never observed to lead the centromere.

Additionally, metaphase and anaphase spreads were probed with the 180bp knob repeat, CentC and CRM to determine whether these sequences on the B chromosome exhibit neocentromeric activity of the type described by Hiatt and coworkers (2002). Such activity would be evidenced by a trail of DNA moving ahead of the centromere that would hybridize with the respective probe. Chromosomes displaying this pattern of hybridization were not observed (Figure 8).

Histone H3-Ser 10 phosphorylation is not observed at positions other than the centromere on the B chromosome.

During mitosis in plants, H3-histone proteins near the centromere are phosphorylated at the Serine10 residue (Kaszas and Cande 2000). Antibodies that recognize H3-pSer10 residues were used to label root tip metaphase spreads containing B chromosomes to examine the distribution of this centromere associated histone modification. The chromosomes were also hybridized with "6-9-11," ZmBs, or CRM.

H3 Ser10 phosphorylation on the B chromosome is only observed near the centromere. The region of intense hybridization of "6-9-11" extends beyond the region labeled by H3-pSer10 antibodies and H3-pSer10 signal was observed between two sites of CRM hybridization (Figure 9).

Discussion

The B Chromosome Contains Elements in Common with A Chromosomes But in Different Arrangements

It is not surprising that some repetitive DNA elements would be abundant on the B chromosome as accumulation there would be unlikely to interrupt any vital gene function. However, while many kinds of repetitive elements are present on the B chromosome, there is a striking pattern of enrichment for DNA elements that are normally found at or near centromeres. Most notably, the centromeric satellite, CentC, is found scattered throughout the long arm of the B (Figure 2), but it is found only at the centromeres of the A chromosomes. Also, DNA elements that show enrichment near centromeres, the "Centromere Diffuse" elements, also show strong hybridization in the heterochromatic regions of the B chromosome (Figure 4). In contrast, Huck, one of the most highly represented retroelements in the maize genome, is not enriched near the centromeres of A chromosomes and does not hybridize strongly to the B chromosome.

In the few cases where nascent B chromosomes have been observed to form, chromosomal aberrations have played a major role (e.g. Berdnikov et al. 2003; Cheng et al 2000; Dhar et al. 2002). For example, in a trisomic of *Plantago lagopus*, the extra A chromosome underwent a series of complex rearrangements including the formation of a ring chromosome and finally an isochromosome that experienced “massive amplification” of specific DNA sequences ultimately leading to development of properties of a B chromosome

over only a few generations (Dhar et al. 2002). In the process that degenerates an A chromosome to form a B, genes that are detrimental to the fitness of the individual when present in greater than 2n copies, must be lost or silenced as part of the B chromosome formation. Zheng et al. (1999) and Kato et al. (2004b) observed that dicentric chromosomes in maize rapidly undergo massive rearrangements via the chromosome type breakage-fusion-bridge cycle generally resulting in deletions of portions of the original chromosome. Therefore, an event that results in a multicentric chromosome could give rise to a small chromosome with few genes from which a B chromosome could form. Furthermore, if one of the centromeres on such a rearranged multicentric chromosome were to become inactivated, the resulting nascent B chromosome would be rich in centromeric elements at a location distinct from the remaining functional centromere. Subsequent expansion of the inactive centromere region by invasion of repetitive elements, unequal crossovers, and other DNA amplification mechanisms that affect repetitive DNA elements could give rise to a chromosome with the DNA arrangement observed in the current maize B chromosome.

If the observed repetitive elements were not present on the nascent B chromosome, they must have transposed to it from the A chromosomes. Mroczek and Dawe (2003) examined many of the LTR retroelements abundant in maize for their chromosomal distribution, including the three elements used in this study. Of those examined, Huck was the only retroelement rarely found in knobs while Cinful-1 hybridization was 2.2 times

higher in knobs composed of TR-1 repeat when compared to euchromatin. The present study shows that Huck is depleted while Cinful-1 and Prem1 are enriched in the heterochromatin around the centromeres of A chromosomes. Therefore, the hybridization patterns of these elements on the B chromosome may be explained by targeted insertion into different types of chromatin.

Although insertion preference may explain the pattern of LTR retroelement hybridization on the B for Huck, Prem1, and Cinful-1, the centromeric repeat, CentC, is not known to transpose. CRM and CentA are LTR retrotransposons that are actively transposing but accumulate only at the centromeres of A chromosomes (Nagaki et al. 2003). Additionally, the CentC and CRM element families evolve rapidly, but closely related sequences are present at all of the nonhomologous centromeres suggesting that these sequences evolve in concert. The presence of CentC along the length of the B chromosome may indicate a hypothetical mechanism which homogenizes centromeric elements and may target the whole chromosome rather than just the centromeric region. Yet another possibility is that CRM or CentC clusters are mobilized at random in the genome but are selected against in the A chromosome arms. Because the B chromosome is gene poor, or because it contains a chromatin type that is unusual at locations other than the centromere, centromeric elements may be able to accumulate on it.

Cheng and Lin (2004) used the B specific element, CL1, to obtain several lambda phages containing DNA from the B chromosome. Sequencing of these phages showed that the CL1 element is organized in tandem arrays

that have been interrupted by insertion of transposable elements. Also, one of the phages shows an insertion into a CL1 array of 7 copies of CentC with a sequence similar to the subclone ZMABC91 taken from an A centromeric BAC. This example of an insertion of centromeric DNA into a B chromosome specific sequence supports the idea that centromeric elements can be moved from the A centromeres to the B chromosome. However, because little sequence is available from the B chromosome, determining how CentC was scattered along the long arm of the B is not possible at this point.

Role of Specific DNA Sequences in Centromere Formation and Chromosome Movement

Although the reason for the concentration of centromeric elements at many locations in the B chromosome is not clear, their presence has allowed us to examine the role they play in centromere formation. In A chromosomes of maize, the satellite CentC and the CRM family of retroelements have been shown to underlie the chromatin domain that contains the modified H3 histone (CENH3) which is required for centromere function. However, many of the copies of CentC and CRM lie outside the region associated with CENH3 suggesting that, at most, these elements nucleate the formation of the CENH3 binding domain and do not define its boundaries (Zhong et al. 2002; Jin et al. 2004). In the B chromosome, the element CentC is present in high copy number throughout the heterochromatin of the B chromosome, including some locations that appear to be large arrays of the repeat, and CRM is found

interspersed with CentC in the proximal heterochromatin. Therefore, the B chromosome is an example of a stable plant chromosome that contains large clusters of centromeric DNA elements away from the functional centromere.

The stability of the B chromosome is a strong argument against the formation of fully functional ectopic centromeres along its length. However, we examined the possibility that the long arm centromeric DNA elements were recruiting centromeric components at a low level that was insufficient to result in chromosome breakage. Using antibodies against CENH3, we labeled chromosome spreads containing the B chromosome from cells in interphase, mitosis, meiosis I, and meiosis II. In no case did we observe labeling of CENH3 at locations other than the previously recognized centromeres (Figures 3 and 4).

In maize, blocks of the 180bp and TR1 knob repeat become mobile in meiosis in the presence of factors on abnormal chromosome 10 and move ahead of the other chromosomes during meiotic anaphase, stretching the attached chromosome arm toward the spindle pole (Rhoades and Vilkomerson 1942). Because this type of conditional chromosome movement involves specific DNA sequences that are present throughout the genome but inactive except under special circumstances (Hiatt et al. 2002) and that do not associate with essential kinetochore proteins (Dawe et al. 1999), we considered the possibility that the B chromosome makes use of similar movement mechanisms involving sequences along its length, such as ZmBs, CL1, CentC, or other unidentified B specific sequences, to promote its

passage through meiosis. Support for this idea came from a report by Carlson and Roseman (1992) in which they examined meiotic passage of translocation chromosomes involving a reciprocal exchange between the short arm of chromosome 9 and the B chromosome (TB-9Sb). They report that when this B-9 translocation chromosome is present as a univalent, it moves towards the pole ahead of the remaining chromosomes. This precocious movement is reminiscent of the movement described above for knob sequences.

Although many chromosome spreads were examined for this type of chromosomal movement during meiosis I, it was never observed. During telophase, only the CENH3 bound portion of the B chromosome, containing overlapping CRM and CentC signals, moves ahead toward the poles. No other part of the chromosome is stretched toward the poles as would be observed in the non-kinetochore mediated chromosome movement analogous to that conditioned by abnormal 10. Although the observed movement of intact B chromosomes in this study differs from the movements of B-A translocation chromosomes observed by Carlson and Roseman (1992), it should be noted that the B chromosomes in this study were present as a bivalent pair, not as univalents. The difference in movement of univalent B-A chromosomes versus paired bivalents B chromosomes could be due to mechanisms that are only activated by a univalent situation. In any case, our results are not consistent with constitutively active "holocentromeric" activity of the centromeric elements on the B chromosome outside the normally recognized centromere.

Plant mitotic and meiosis II metaphase chromosomes show phosphorylation of the serine10 residue of the canonical H3-histone protein in the region immediately around the centromere (Houben et al. 1999). Because of its location, the region of Ser10 p-H3 modification is thought to play a role in sister chromatid cohesion (Kaszas and Cande 2000; Shibata and Murata 2004) although the modification itself appears to be dispensable for proper chromosome transmission (Manzanero et al. 2002; Gernand et al. 2003). Although the role that Ser10 p-H3 plays in chromosome transmission is not certain, the modification likely reflects a difference in chromatin structure from the surrounding non-phosphorylated regions. We examined the distribution of Ser10 phosphorylated H3-histone in relation to the DNA elements enriched near centromeres and on the B chromosome. The "6-9-11" hybridization signal was observed to extend well beyond the region labeled by Ser10 p-H3 in A chromosomes while CRM and CentC are visualized as two spots with p-H3 labeling between and somewhat overlapping them (Figure 9). On the B chromosome, Ser10 p-H3 is located only around the functional centromere and is not found at the distal heterochromatin. These findings indicate that while the same DNA elements are enriched on the B chromosome long arm and the regions immediately surrounding A centromeres, their chromatin structures are different at mitotic metaphase.

There Is Less CENH3 present at the B centromere than the A centromeres.

Examination of numerous labeled cell preparations revealed that the intensity of labeling of the CENH3 proteins at the B centromere is consistently less than in A centromeres. It has been proposed that the centromeres of homologous A chromosomes are in competition with each other for inclusion into the female gamete leading to rapid evolution of centromeres (Malik and Henikoff 2002). The B chromosome, in contrast to A chromosomes, is frequently unpaired during meiosis and the B chromosome fails to disjoin at the second microspore division (Roman 1947). The potentially different requirements for successful transmission of A and B centromeres may involve different selection pressures and could explain why there is a difference in the amount of CENH3 labeling.

Conclusion

The above findings indicate that in maize, as in other model organisms, DNA sequence alone is insufficient to cause centromere formation. Instead, centromere identity may be determined by epigenetic factors including chromatin structure. Because it is phenotypically neutral and therefore may be manipulated without detrimental effects to the plant, the maize B chromosome could provide a tool for investigating the specific chromatin requirements for successful chromosome transmission. The collection of FISH probes described above will facilitate the cytological examination of the B chromosome.

Acknowledgements

This work was supported by grant DBI 9975827 from the National Science Foundation Plant Genome Initiative.

References

Alfenito MR, Birchler JA (1993) Molecular characterization of a maize B chromosome centric sequence. Genetics 135:589-597

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J Mol Biol 215:403-410

Ananiev EV, Phillips RL, Rines HW (1998) Chromosome-specific molecular organization of maize (*Zea mays* L.) centromeric regions. PNAS 95:13073-13078

Berdnikov VA, Gorel FL, Kosterin OE, Bogdanova VS (2003) Tertiary trisomics in the garden pea as a model of B chromosome evolution in plants. Heredity 91:577-583

Carlson W, Roseman RR (1992) A new property of the maize B chromosome. Genetics 131:211-223

Cheng Y, Lin B (2003) Cloning and characterization of maize B chromosome sequence derived from microdissection. Genetics 164:299-310

Cheng Y, Lin B (2004) Molecular organization of large fragments in the maize B chromosome: Indication of a novel repeat. Genetics 166:1947-1961

Cheng ZK, Yu HX, Yan HH, Gu MH, Zhu LH (2000) B chromosome in a rice aneuploid variation. Theor Appl Genet 101:564-568

Chilton MD, McCarthy BJ (1973) DNA from maize with and without B chromosomes: a comparative study. Genetics 74:605-614

Dawe RK, Reed LM, Yu HG, Muszynski MG, Hiatt EN (1999) A maize homolog of mammalian CENPC is a constitutive component of the inner kinetochore. Plant Cell 11:1227-1238

Dhar MK, Friebe B, Koul AK, Gill BS (2002) Origin of an apparent B chromosome by mutation, chromosome fragmentation and specific DNA sequence amplification. Chromosoma 111:332-340

Fuchs J, Brandes A, Schubert I (1995) Telomere sequence localization and karyotype evolution in higher-plants. Plant Systematics and Evolution 196:227-241

Gardiner J. M., Coe E. H., & Chao S (1996) Cloning maize telomeres by complementation in *Saccharomyces cerevisiae*. Genome 39:736-748

Gernand D, Demidov D, Houben A (2003) "The temporal and spatial pattern of histone H3 phosphorylation at serine 28 and serine 10 is similar in plants but differs between mono- and polycentric chromosomes." Cytogenet Genome Res 101:172-176.

Hall SE, Kettler G, Preuss D (2003) Centromere satellites from *Arabidopsis* populations: Maintenance of conserved and variable domains. Genome Res 13:195-205

Hiatt EN, Kentner EK, Dawe RK (2002) Independently regulated neocentromere activity of two classes of tandem repeat arrays. Plant Cell 14:407-420

Hsu FC, Wang CJ, Chen CM, Hu HY, Chen CC (2003) Molecular characterization of a family of tandemly repeated DNA sequences, TR-1, in heterochromatic knobs of maize and its relatives. Genetics 164:1087-1097

Houben A, Wako T, Furushima-Shimogawara R, Presting G, Kunzel G, Schubert I, 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, Birchler JA, Parrott WA, Dawe RK (2003) A molecular view of plant centromeres. Trends Plant Sci 8:570-575

Jin WW, Melo J.R, Nagaki K, Talbert PB, Henikoff S, Dawe RK, and Jiang J (2004) Maize centromeres: Organization and functional adaptation in the genetic background of oat. Plant Cell 16: 571-581

Jones RN, Houben A (2003) B chromosomes in plants: escapees from the A chromosome genome? Trends Plant Sci 8:417-423

Jones RN, Rees H (1982) B Chromosomes, 1st edn. Academic Press, London

Kaszas E, Birchler JA (1998) Meiotic transmission rates correlate with physical features of rearranged centromeres in maize. Genetics 150:1683-1692

Kaszas E, Cande WZ (2000) Phosphorylation of histone H3 is correlated with changes in the maintenace of sister chromatid cohesion during meiosis in maize, rather than the condensation of the chromatin. J Cell Sci 113:3217-3226

Kato A, Lamb JC, Birchler JA (2004a) Chromosome painting in maize using repetitive DNA sequences as probes for somatic chromosome identification. Proc Natl Acad Sci, USA 101:13554-13559.

Kato A, Zheng YZ, Auger DL, Phelps-Durr T, Bauer MJ, Lamb JC, Birchler JA (2004b) Minichromosomes derived from the B chromosome of maize. Cytogenet Genome Res, in press.

Lamb JC, Birchler JA (2003) The role of DNA sequence in centromere formation Genome Biol 4: 214.

Longley AE (1927) Supernumerary chromosomes in *Zea mays*. J Agric Res 35:796

Malik HS, Henikoff S (2002) Conflict begets complexity: the evolution of centromeres. Curr Opin Genet Dev 12:711

Manzanero S, Arana P, Puertas MJ, Houben A (2000) The chromosomal distribution of phosphorylated histone H3 differs between plants and animals at meiosis. Chromosoma 109:308-317

McClintock B, Kato TA, Blumenshein A (1981) Chromosome constitution of races of maize. Colegio de Postgraduados, Chapingo, Mexico.

Meyers BC, Tingey SV, Morgante M (2001) Abundance, distribution, and transcriptional activity of repetitive elements in the maize genome. Genome Res. 11:1660-1676

Miller JT, Dong F, Jackson SA, Song J, Jiang J (1998) Retrotransposon-related DNA sequences in the centromeres of grass chromosomes. Genetics 150:1615-1623

Mroczek RJ, Dawe RK (2003) Distribution of retroelements in centromeres and neocentromeres of maize. Genetics 165:809-19

Nagaki K, Song J, Stupar RM, Parokonny AS, Yuan Q, Ouyang S, Liu J, Hsiao J, Jones KM, Dawe RK, Buell CR, Jiang J (2003) Molecular and cytological analysis of large tracks of centromeric DNA reveal the structure and evolutionary dynamics of maize centromeres. Genetics 163:759-770

Page BT, Wanous MK, Birchler JA (2001) Characterization of a maize chromosome 4 centromeric sequence: Evidence for an evolutionary relationship with the B chromosome centromere. Genetics 159:291-302

Puertas MJ (2002) Nature and evolution of B chromosomes in plants: A non-coding but information-rich part of plant genomes. Cytogenet Genome Res 96:198-205

Rhoades MM, Vilkomerson H (1942) On the anaphase movement of chromosomes. PNAS 28:433-443

Roman H (1947) Mitotic Nondisjunction in the Case of Interchanges Involving the B-Type Chromosome in Maize. Genetics 32:391

Shibata F, Murata M (2004) Differential localization of the centromere-specific proteins in the major centromeric satellite of *Arabidopsis thaliana*. J Cell Sci 117: 2963-70

Stark EA, Connerton I, Bennett ST, Barnes SR, Parker JS, Forster JW (1996) Molecular analysis of the structure of the maize B-chromosome. Chrom Res 4:15-23

Sullivan BA, Blower MD, Karpen GH (2001) Determining centromere identity: Cyclical stories and forking paths. Nature Rev Gen 2:584-596

Zheng YZ, Roseman RR, Carlson WR (1999) Time course study of the chromosome-type breakage-fusion-bridge cycle in maize. Genetics 153:1435-1444

Zhong CX, Marshall JB, Topp C, Mroczek R, Kato A, Nagaki K, Birchler JA, Jiang J, Dawe RK (2002) Centromeric retroelements and satellites interact with maize kinetochore protein CENH3. Plant Cell 14:2825-36

Table

Table 1 Elements used as probes and their pattern of hybridization on the B chromosome (C Centromere, DE distal euchromatin, DH distal heterochromatin (1–4), PH proximal euchromatin, PK proximal heterochromatin, S short arm)

Probe examined	Hybridization on the B chromosome
ZmBs-B specific repeat	S, C, PH, DE, DH4, DE
CentC	C, PH, DH1-3, DE
CL1-B specific repeat	PH, DH1-4
CRM	C, DH4
CentA	C, DH4 (DH1-4 at low stringency)
180 bp knob	PE, DH4
"AGT" microsatellite	DH3 (two distinct locations)
Cent4	PH (possible cross-hybridization to 180 bp knob), DH4
Telomere repeat	S, C, PH, DE, DH4, DE
Gardiner's 1.1 kb subtelomeric repeat	S, C, PH, DE, DH4, DE
4-12-1 subtelomeric repeat	S, C, PH, DE, DH4, DE
"Centromere Diffuse" elements	PH, DH1-4
ZMABC19, ZMABC91	PH, DH1-4
Prem1	PH, DH1-4
Huck	No hybridization except for PE
TR1	PH (as reported by Hsu et al. *2003*)
5S Ribosomal gene cluster	No signal
25S Ribosomal gene cluster	No signal

Figures

Figure 1. The maize B chromosome at pachynema. The image on the left is a DAPI stained maize B chromosome at pachynema. The image on the right is an illustration indicating the hybridization patterns of the various repetitive elements that have been located on the B chromosome.

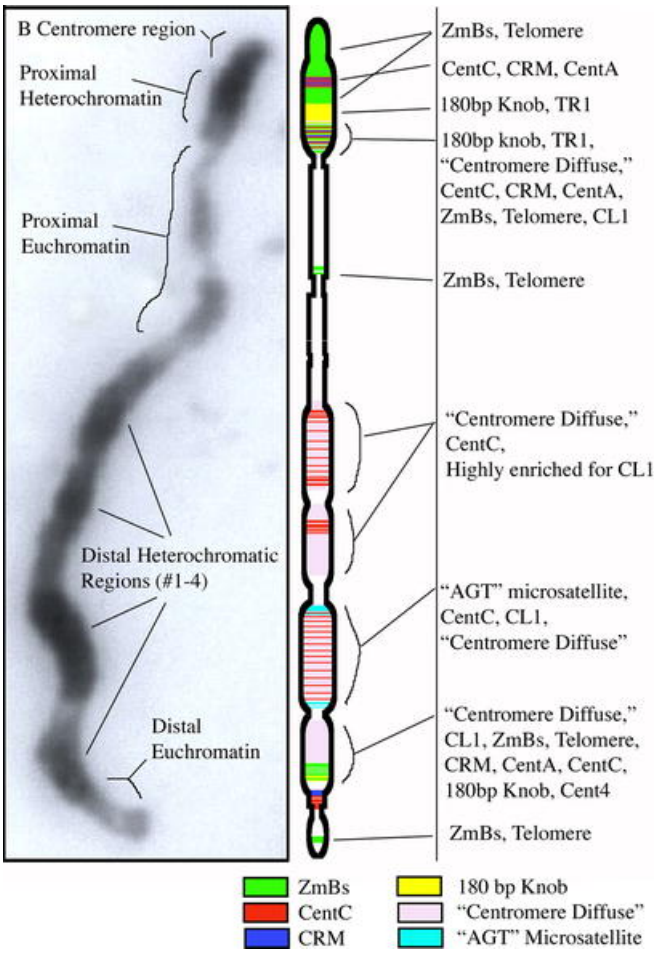


Figure 2. Centromeric elements on the B chromosome.
ZmBs (A), CentC (B), CRM (C), and the telomere repeat (D) were used as FISH probes on pachytene chromosome spreads containing the B chromosome. The white arrows indicate the location of the B centromere. The red arrows indicate regions of hybridization of the respective elements. In the top panels the probe signal and the DAPI counter-stain are merged. In the middle panels, only the probe is shown and the bottom panels shows the chromosomes stained with DAPI.

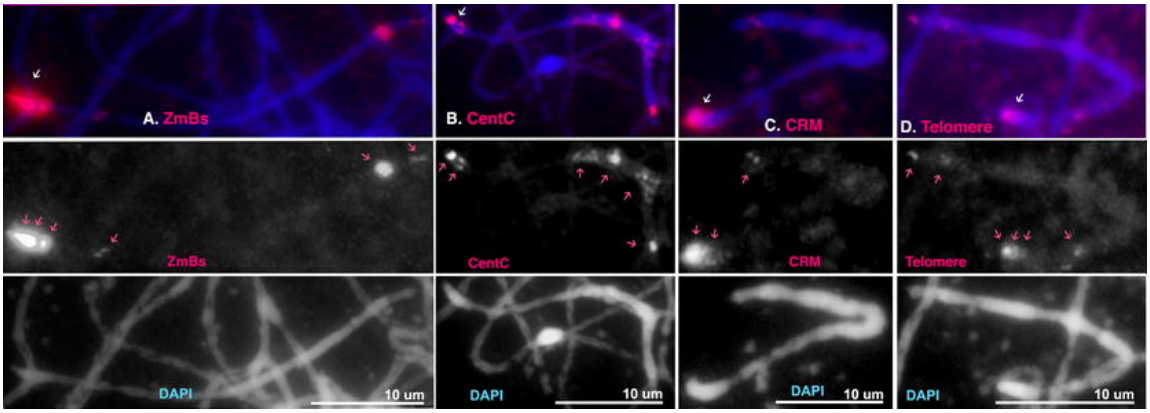


Figure 3. Magnified views of the B centromere region.
Centromeres of B chromosomes are shown that have been labeled with ZmBs and CENH3 (A), ZmBs and CentC (B), ZmBs and CRM (D), ZmBs and 180bp knob repeat (D), CRM and 180bp knob repeat (E), and ZmBs and CentC (F). The chromosomes are at the pachytene stage except in (F), where a pair of B chromosomes are in Anaphase I. The first column shows merged views, including DAPI counter-stain, while the second two columns show individual probe signals.

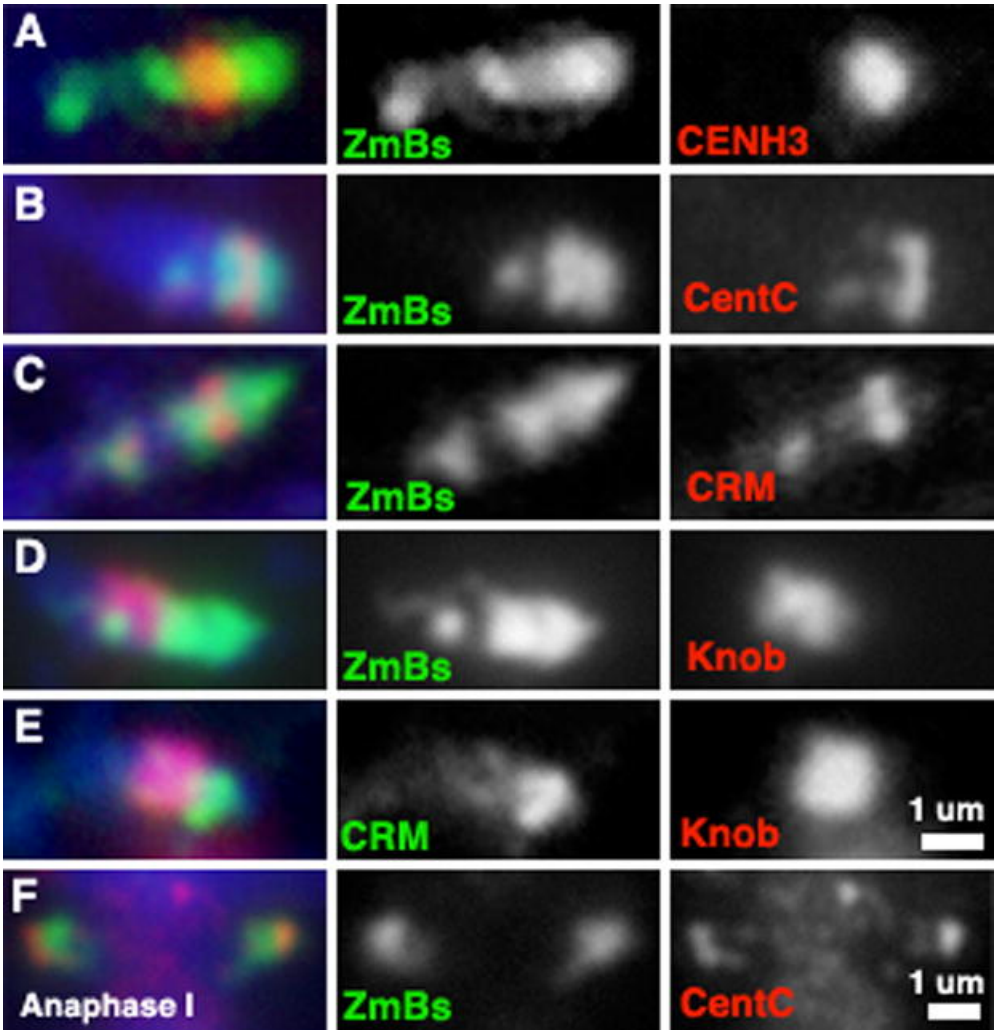


Figure 4. "Centromere Diffuse" elements on the B chromosome. Centromere Diffuse element "6-9-11" was used as a FISH probe on mitotic (A) or meiotic (B) chromosome spreads containing B chromosomes. In (A), the arrows indicate the B chromosomes. In (B), the arrows indicate the centromere, the locations of the "AGT" microsatellite, and the locations of "6-9-11."

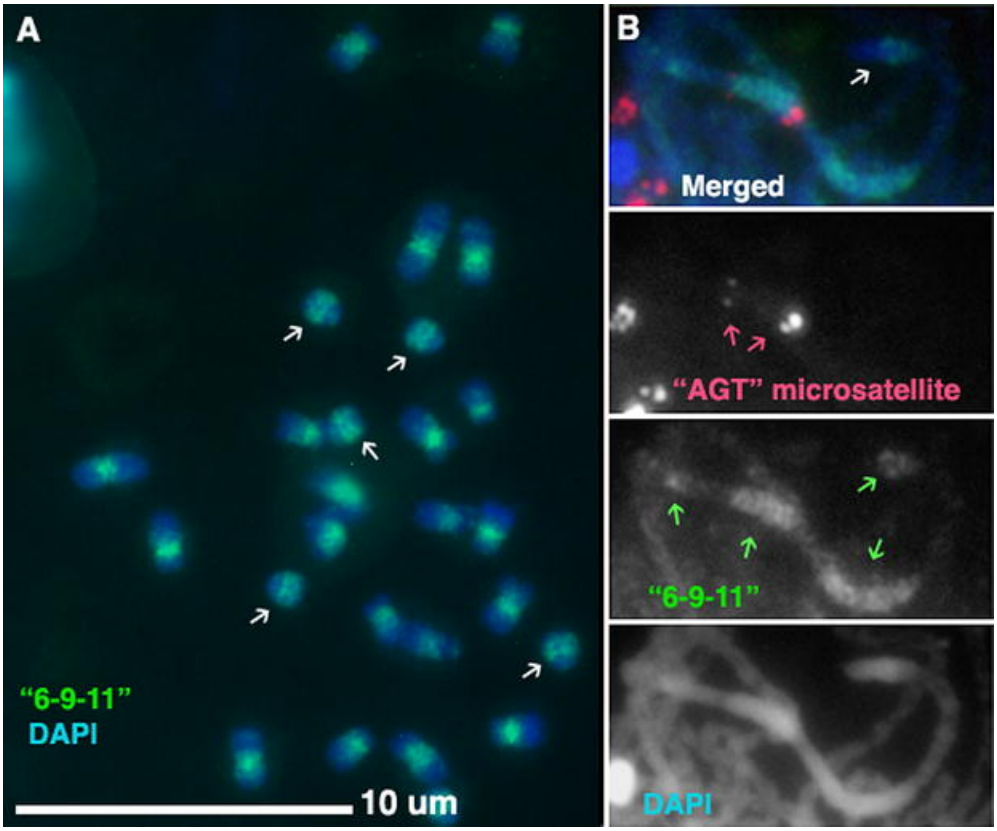


Figure 5. Retroelements on the B chromosome. Huck (in green) and Prem1 (in red) LTRs were used as probes on a mitotic chromosome spread. The arrows indicate the B centromere.

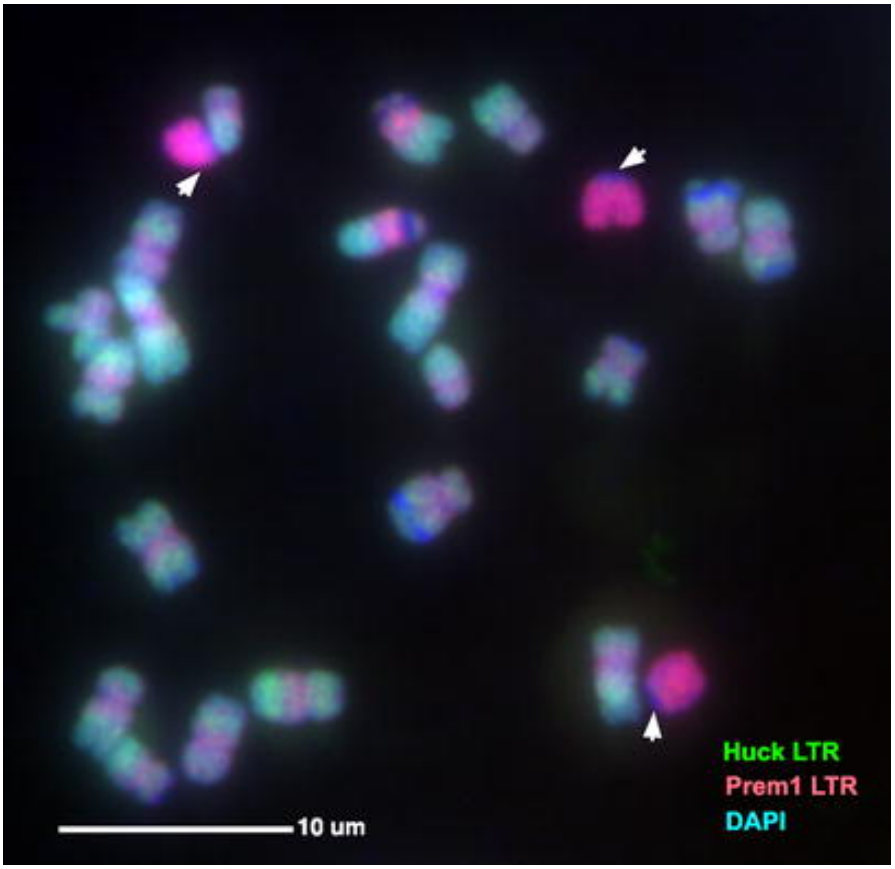


Figure 6. CENH3 labeling of mitotic cells containing B chromosomes. Antibodies against CENH3 (green) were used to label mitotic chromosome spreads containing the B chromosome, which were then hybridized to ZmBs (red) to identify the B chromosome. A root tip cell in interphase (A). A metaphase cell spread (B). (A') and (B') show only the gray value image of CENH3 labeling. The arrows indicate the site of minor hybridization to ZmBs.

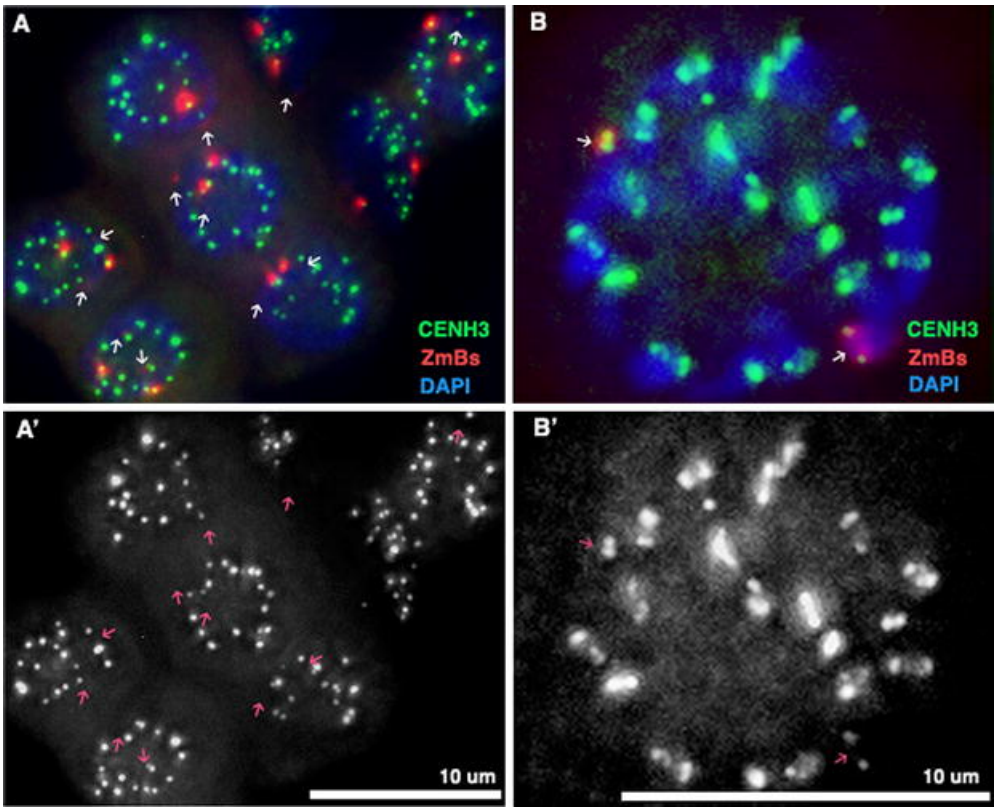


Figure 7. CENH3 labeling of meiotic cells containing the B chromosome. Antibodies against CENH3 (green) were used to label meiotic chromosome spreads containing the B chromosome, which were then hybridized to ZmBs (red) to identify the B chromosome. CENH3 labeling is observed at the centromere of the B chromosome but not at locations on the long arm. In (A) and (B) arrows indicate the locations of the centromeric or distal locations of ZmBs hybridization. In (B), the B chromosome is separated from the remainder of the nucleus. (A') and (B') show only the CENH3 signal. (C) Two cells in meiosis II labeled with antiCENH3 (green) and hybridized to ZmBs (red). (C') CENH3 labeling of the preparation shown in C.

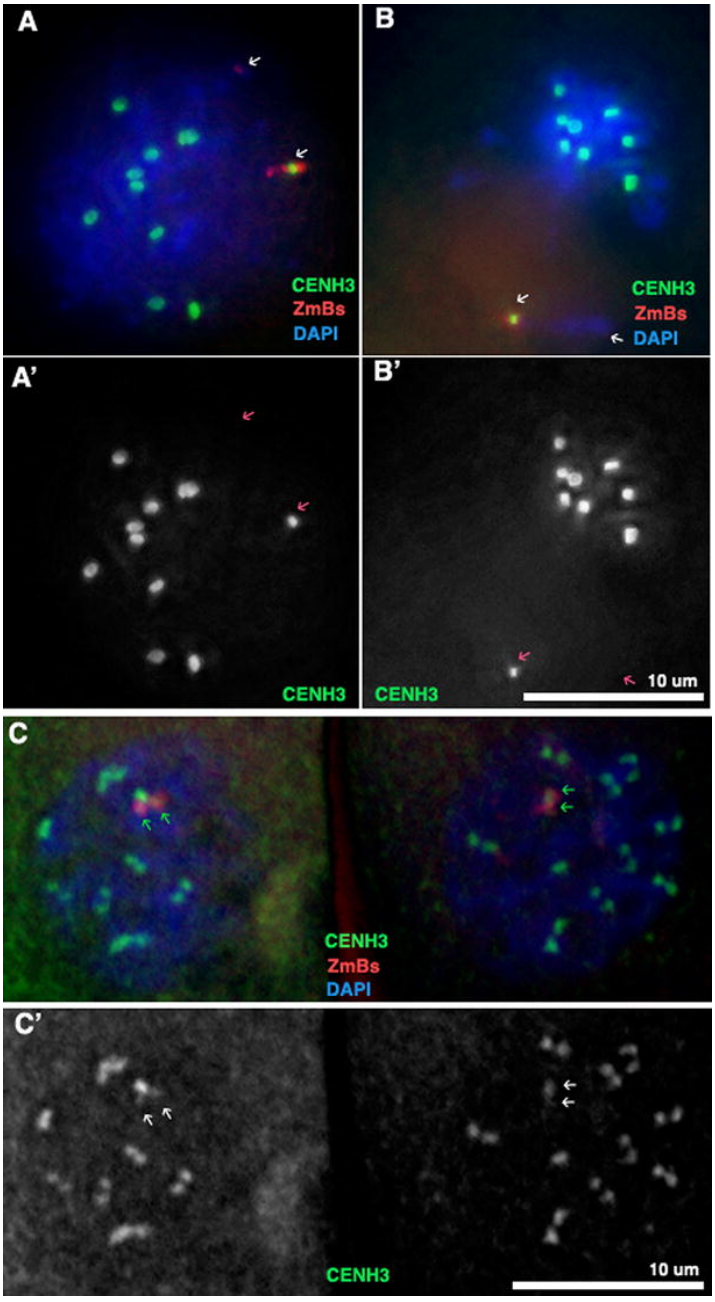


Figure 8. Paired B chromosomes at anaphase/telophase. Representative anaphase I (A) or telophase I (B) spreads are shown. ZmBs signal overlaps (yellow) or is slightly behind (green) the CentC signal (red). The long arm of the B chromosome is not observed to lead the centromere. Arrows indicate the sites of minor hybridization of ZmBs on the B long arm.

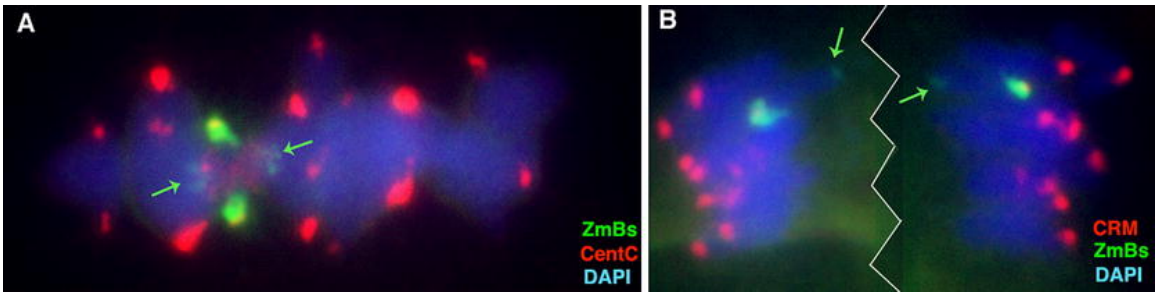
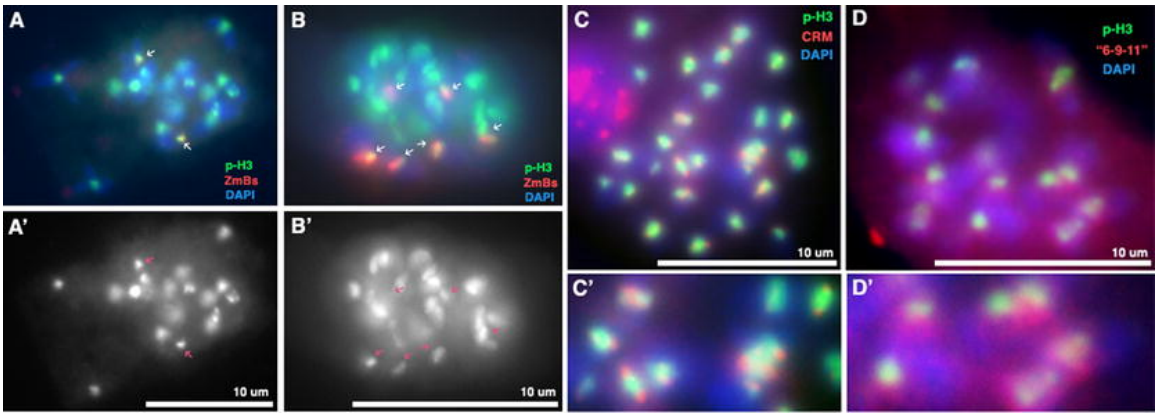


Figure 9. Phosphorylated H3 histone on the B chromosome. Chromosome spreads labeled with antibodies against p-Ser10 of the H3 histone protein and hybridized with ZmBs (A and B), CRM (C), or centromere diffuse element "6-9-11" (D). (A') and (B') show only the p-H3 labeling with arrows indicating the B centromere. (C') and (D') show a higher magnification of the labeled chromosomes. The image in (D) is from a single plane of focus while the other images are compressed stacks.



4. MOLECULAR AND FUNCTIONAL DISSECTION OF THE MAIZE B CHROMOSOME CENTROMERE

The findings described in the previous chapter were presented in a talk at the Plant and Animal Genome conference. Among the attendees at that conference were Dr. Jiming Jiang, a professor at the University of Wisconsin, and a postdoctoral researcher in his lab, Dr. Weiwei Jin. Dr. Jiang is a co-PI with Dr. Birchler and others on a NSF grant to study plant centromeres. After my presentation, Drs. Jiang and Jin discussed the findings and I showed them additional data that indicated that the organization of the B centromere included two large tracts of the ZmBs element flanking a central core composed of CentC and CRM. I described additional experiments that I felt would be valuable to determine the relationship between this central core and the amount of CenH3 that is recruited to the B centromere. Dr. Birchler had also included my proposed experiments in the renewal information for the grant that he, Dr. Jiang, and others had submitted to the NSF. Dr. Jin had been perfecting the fiber-FISH technique and was applying it to a collection of centromere misdivision derivatives that were previously developed in the Birchler lab. He combined this powerful approach to study the DNA structure at the B centromere with my proposed experiments to produce the following article.

I made the initial observation that the amount of CenH3 correlated with the amount of centromeric elements remaining in misdivided centromeres. Dr. Jin examined many cells and produced the data used in the tables of this article. I examined the position of the "core domain" in meiotic samples and produced all images involving meiotic chromosomes. Furthermore, I helped design the experiments, selected material to

examine and assisted in writing the manuscript. The article was published in The Plant Cell and the reference style of that journal has been retained. The American Society of Plant Biology is the copyright holder for this article. It is used in accordance with their policy.

JIN, W., J. C. LAMB, J. M. VEGA, R. K. DAWE, J. A. BIRCHLER *et al.*, 2005 Molecular and Functional Dissection of the Maize B Chromosome Centromere. Plant Cell **17:** 1412-1423.

Molecular and Functional Dissection of the Maize B Chromosome Centromere

Weiwei Jin[a], Jonathan C. Lamb[b], Juan M. Vega[b], R. Kelly Dawe[c], James A. Birchler[b,1], and Jiming Jiang[a,1]

[a] Department of Horticulture, University of Wisconsin–Madison, Madison, Wisconsin 53706

[b] Division of Biological Science, University of Missouri, Columbia, Missouri 65211

[c] Department of Plant Biology, University of Georgia, Athens, Georgia 30602

[1] To whom correspondence should be addressed. E-mail birchlerj @missouri.edu or jjiang1@wisc.edu; fax 573-882-0123 or 608-262-4743.

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: James A. Birchler (birchlerj@missouri.edu).

Article, publication date, and citation information can be found at www.plantcell.org/cgi/doi/10.1105/tpc.104.030643.

Abstract

The centromere of the maize (*Zea mays*) B chromosome contains several megabases of a B-specific repeat (ZmBs), a 156-bp satellite repeat (CentC), and centromere-specific retrotransposons (CRM elements). Here, we demonstrate that only a small fraction of the ZmBs repeats interacts with CENH3, the histone H3 variant specific to centromeres. CentC, which marks the CENH3-associated chromatin in maize A centromeres, is restricted to an ~700-kb domain within the larger context of the ZmBs repeats. The breakpoints of five B centromere misdivision derivatives are mapped within

this domain. In addition, the fraction of this domain remaining after misdivision correlates well with the quantity of CENH3 on the centromere. Thus, the functional boundaries of the B centromere are mapped to a relatively small CentC- and CRM-rich region that is embedded within multimegabase arrays of the ZmBs repeat. Our results demonstrate that the amount of CENH3 at the B centromere can be varied, but with decreasing amounts, the function of the centromere becomes impaired.

Introduction

The chromosomes of most higher eukaryotes contain a single centromere, the region that serves as the attachment site for spindle fibers and sister chromatid cohesion during cell division. The proteins that are involved in centromere function are highly conserved among all eukaryotes. Despite the similarities of form and function, the primary DNA sequence that underlies centromeres has no discernable conservation among various model organisms, underscoring the present lack of understanding of factors that determine centromere identity.

Although there is no obvious relationship among sequences found at centromeres, there are commonalities in organization. In most complex eukaryotes studied so far, including Drosophila melanogaster, humans, mice, maize (*Zea mays*), rice (*Oryza sativa*), and *Arabidopsis thaliana*, the centromeres are embedded within long tracks of highly repetitive DNA sequences consisting primarily of satellite repeats and transposons (Henikoff et al., 2001; Jiang et al., 2003). Human centromeres are the best studied among multicellular eukaryotes and appear to be a representative model. The most abundant DNA sequence in human centromeres is the ~171-bp α-satellite repeat (Willard, 1998).

The amount of α-satellite DNA varies from ~250 kb to >4 Mb in different centromeres (Wevrick and Willard, 1989; Oakey and Tyler-Smith, 1990). Human artificial chromosomes were successfully assembled using either synthetic or cloned α-satellite DNA (Harrington et al., 1997; Ikeno et al., 1998), suggesting that long stretches of α-satellite DNA could act as a functional human centromere. Structural and functional analyses of DNA in the X chromosome centromere revealed that the a-satellite repeats are more diverged in the flanks of the centromere and become increasingly homogenized toward a functional core (Schueler et al., 2001). These data suggest that the centromeres evolve by selecting new repeats at their functional core, pushing older repeats to the flanking regions (Henikoff, 2002).

In contrast with the lack of conservation of centromeric DNA, several proteins specific to the centromere/kinetochore complex are highly conserved (Henikoff et al., 2001; Sullivan et al., 2001; Houben and Schubert, 2003). A centromere-specific histone H3-like protein, referred to as CENP-A or CENH3, has been found to underlie the kinetochore (Henikoff et al., 2001; Sullivan et al., 2001). CENH3 has been found in all model eukaryotes, including several plant species (Talbert et al., 2002; Zhong et al., 2002; Nagaki et al., 2004). There are numerous lines of evidence that CENH3 plays the key role in the establishment and function of kinetochores in various organisms (Henikoff et al., 2001; Sullivan et al., 2001). In maize, CENH3-associated chromatin is exclusively associated with the centromeric satellite and the CRM retroelements from the centromeric retrotransposon (CR) family in grasses (Zhong et al., 2002; Jin et al., 2004; Topp et al., 2004).

Because the stretches of repetitive DNA in centromeres are highly homogeneous and are often megabases in length, they are left as gaps in most sequencing projects of complex eukaryotic genomes (Henikoff, 2002). Currently, nearly complete sequences are available only from two native centromeres of rice (Nagaki et al., 2004; Wu et al., 2004; Zhang et al., 2004) and human neocentromeres (Saffery et al., 2003), which contain minimal or no satellite repeats. Therefore, methods other than sequencing must be used to study the relationship between the centromeric protein structure and the underlying DNA sequences. The centromere of the maize B chromosome provides an excellent model to study centromeres that contain an excess amount of satellite repeats. This supernumerary chromosome is present in some but not all maize varieties and is maintained by an accumulation mechanism involving nondisjunction at the second pollen mitosis followed by preferential fertilization of the egg by the B-containing sperm in the process of double fertilization (Roman, 1947). The maize B centromere region contains a repetitive element ZmBs that is not present on the A chromosomes (Alfenito and Birchler, 1993). The specificity of the ZmBs repeat to the B chromosome allows molecular and cytological studies of a single centromere among others in the genome (Kaszas and Birchler, 1996, 1998). In addition, B centromere misdivision derivatives can be readily developed from specific cytogenetic stocks, and the transmission of these derivatives can be easily tracked (Kaszas and Birchler, 1998). These materials provide a unique system to study a centromere deletion series to establish the molecular correlates of centromere function.

In this study, we report the fine structure and DNA composition of the centromeres from the normal B chromosome and a set of B centromere misdivision derivatives. We describe an ~700-kb domain that consists of all three previously

described repeats (ZmBs repeat, CentC, and CRM). Five misdivision breakpoints are mapped within this domain and the amount of the B centromere–bound CENH3 is correlated with the amount of sequence from this domain. These results provide support for the view that CentC/CRM is present at all sites in the maize genome that specify a centromere (Ananiev et al., 1998; Jin et al., 2004). They also demonstrate that the amount of CENH3 can be altered and correlates with centromere function. These data demonstrate the minimum requirements for a representative centromere in a multicellular eukaryote.

Results

DNA Repeats and Their Organization in the Intact Maize B Centromere

It was previously demonstrated that the centromeres of maize A chromosomes contain a 156-bp centromere-specific satellite repeat CentC (Ananiev et al., 1998) and retrotransposons from the CR family, including CRM1, CRM2 (full-size CR elements), and CentA (a truncated CR element) (Ananiev et al., 1998; Zhong et al., 2002; Nagaki et al., 2003b; Jin et al., 2004; Lamb et al., 2005). In addition to the elements found in A centromeres, the centromere of the maize B chromosome contains a B-specific repeat ZmBs (Alfenito and Birchler, 1993; Kaszas and Birchler, 1996). High-resolution fluorescence in situ hybridization (FISH) mapping on pachytene B chromosomes revealed that the B chromosome contains two major hybridization sites of the ZmBs repeat: one at the centromeric region and another at the subtelomeric region on the long arm (Lamb et al., 2005) (Figure 1). Major CRM signals were only detected in the centromeric regions, although minor CRMsignals were also observed outside of the

centromere (Lamb et al., 2005). The CentC signals, which are specific to the centromeres of the A chromosomes, were dispersed throughout the B chromosome (Lamb et al., 2005) (Figure 1).

We used a fiber-FISH approach to analyze the organization of CentC, CRM, and the ZmBs repeat within the B centromere. We found that the longest array of the ZmBs repeat was ~1150 mm (~3.7 Mb) in length. Because it is technically difficult to obtain DNA fibers >2 Mb, we estimate that the ZmBs array in the centromeric region is at least 3.7 Mb and is potentially longer. The fiber-FISH–based estimate excludes the ZmBs arrays that are near the centromere but separated from the major cluster as well as those in the distal long arm tip (Figure 1).

Two of the longest fiber-FISH signals are shown in Figure 2. Figure 2A shows signals for ZmBs (green) and CentC (red), and Figure 2C shows signals for ZmBs (green) and CRM (red). The CentC signals were restricted to an~700-kb domain that is also rich in CRM sequences. The CentC and CRM sequences in this domain are intermingled (Figures 2B, 2D, 2E, and 2F), similar to the organization of these two sequences in the A centromeres (Jin et al., 2004). Unlike A centromeres, however, the intermingled CentC/CRM sequences are disrupted by five ZmBs arrays (Figures 2E and 2F). This ~700-kb CentC/CRM/ZmBs repeat domain is flanked on either side by long ZmBs arrays that contain short CRM signals (Figure 2D). The ZmBs array flanking on one side (arbitrarily called ‘‘left’’) is longer than the array on the ‘‘right’’ side, with lengths measuring ~2.2 and 0.8 Mb, respectively (Figures 2B and 2D). The fiber-FISH signal patterns in the flanking domains are remarkably uniform, suggesting a high concentration

of the ZmBs repeats in these regions. A diagram of the fiber- FISH–based pattern map of the CentC-rich domain is depicted in Figure 2F.

Mapping the Breakpoints of Misdivided B Centromeres

Early cytogenetic analysis of univalent chromosomes in meiosis demonstrated that the centromere is a compound structure consisting of multiple functional units (Darlington, 1939; Sears, 1952). A misdivision of a centromere will result in two functional derivatives. Therefore, analysis of misdivision breakpoints is a powerful approach to dissect the functional domains of centromeres. Numerous B centromere misdivision derivatives were previously isolated from progenies of a TB-9Sb translocation chromosome (Carlson, 1970; Carlson and Chou, 1981; Kaszas and Birchler, 1996, 1998). A set of seven B misdivision derivatives (Figure 3) was selected for fine mapping of the B centromeric regions. The centromeres of these selected misdivision derivatives have significantly different sizes based on our previous pulsed field gel electrophoresis (PFGE) mapping data (Kaszas and Birchler, 1998).

Fiber-FISH was used to analyze the organization of CentC, CRM, and the ZmBs repeat in the misdivision derivatives. Four lines, Telo4-11(-), Ring4-8(-), Telo4-4(-), and Ring4-12(-), were derived from the same progenitor, Iso3(-) (Figure 3A). The sizes of the fiber-FISH signals from Telo4-11(-) and Ring4-8(-) were 825 ± 55 kb and 440 ± 42 kb, respectively (Figures 4A to 4D), in reasonable agreement with the 490- and 500-kb estimates made previously from PFGE (Kaszas and Birchler, 1998) (Table 1). Both Telo4-11(-) and Ring4-8(-) contain blocks of CentC/CRM repeat arrays that align with the patterns obtained from the normal B centromere (Figures 4A to 4D and 5). Telo4–

11(–) contains the first four of the five short ZmBs arrays within the CentC-rich domain, whereas Ring4-8(-) contains only the first ZmBs cluster (Figures 4A to 4D and 5).

The size and pattern of the fiber-FISH signals derived from Telo4-4(-) is almost identical to those from Telo4-11(-) (Figures 4E and 5). PFGE analysis, however, suggested that Telo4-4(-) contains >2 Mb of the ZmBs repeat (Kaszas and Birchler, 1998) (Table 1). In fiber-FISH analysis, we only collected and analyzed the signals associated with CentC and/or with a significant amount of CRM. Additional fiber-FISH signals consisting of solely ZmBs repeats cannot be reliably identified because these signals may be derived from broken DNA fibers. PFGE analysis, by contrast, identifies all DNA fragments associated with the ZmBs repeat. In addition, analysis of progenies from a single individual containing Telo4-4(-) revealed considerable differences in the amount of the ZmBs repeat (J. Lamb and J.A. Birchler, unpublished data), suggesting that this line may not be stable. Taken together, differences in what the two techniques detect and possibly an instability of Telo4-4(-) might explain the discrepancy of the amount of the ZmBs repeats estimated by fiber-FISH and PFGE.

The final Iso3(-) derivative characterized was Ring4-12(-), which contains only 93 ± 10 kb of CRM and the ZmBs repeat and no detectable CentC. This short fragment also aligns with a small part of the CentC-rich domain that includes the first of the five short ZmBs arrays. The first ZmBs array in this region can be identified based on the presence of a distinctive CRM insertion (Figures 4F and 5).

Line Telo3-3(-) and its parental line Telo2-2(-) have a different lineage from line Iso3(-) (Figure 3A). Two DNA fragments containing the ZmBs repeat were detected by fiber-FISH in Telo2- 2(-). The first fragment is ~800 kb long and consists of all three

elements: CentC, CRM, and ZmBs. The organization of the three repeats in Telo2-2(-) is nearly identical to that of Telo4-11(-) and Telo4-4(-) (Figures 4G and 5). The second fragment is ~240 kb in size and consists exclusively of the ZmBs repeat (Figure 4H). We were not able to detect CentC or CRM signals within this fragment using fiber-FISH and could not align it to the CentC-rich domain. Furthermore, we could not find any unambiguous connection between these two centromeric DNA fragments, suggesting that the two fragments are separated by several hundred kilobases (Jackson et al., 1998).

Pachytene FISH analysis of Telo2-2(-) revealed a larger ZmBs repeat signal, which colocalizes with CentC and CRM, and a separate, nearby but smaller ZmBs repeat signal, which does not colocalize with other centromeric elements (Figure 6). This smaller FISH site is also present in the normal B chromosome (Lamb et al., 2005). Thus, the ~240-kb ZmBs array is likely to coincide with this centromere distal ZmBs site. A single DNA fragment containing ZmBs was found in Telo3-3(-), which appeared identical to the ~240-kb ZmBs fragment in Telo2-2(-) (Figure 4I). Both fiber-FISH and pachytene FISH did not detect CentC and CRM signals in Telo3-3(-). Interestingly, Telo3-3(-) is the most unstable B centromere derivative recovered to date (Kaszas and Birchler, 1998). Telo3-3(-) is frequently lost both mitotically and meiotically and routinely shows developmental loss of genetic markers present on the chromosome (Kaszas and Birchler, 1998). It is also unstable structurally, being found alternately in our stocks as a telocentric or isochromosome.

The ZmBs Repeat Is Associated with Maize CENH3

Transmission studies of misdivision derivatives suggest that the ZmBs repeat is present in the functional region of the B centromere (Kaszas and Birchler, 1998). We conducted chromatin immunoprecipitation (ChIP) experiments using an antibody against maize CENH3. Data from three independent experiments (Figure 7A) showed that, on average, the percentage of immunoprecipitation (%IP) of the ZmBs repeat was 7.1% (SE = ± 0.9%, n = 3), whereas only 1.6% (SE = ± 1.4%) was detected with the rDNA control, a statistically significant difference (Student's t test, P ¼ 0.0022). By contrast, the average %IP for the CentC repeat was 27.8% in the same experiments, similar to our previous results (Zhong et al., 2002). The association of the ZmBs repeat with CENH3 was also visualized by sequential detection of CENH3 and the ZmBs repeat on stretched B centromeres. On stretched chromatin fibers, the CENH3 domain appeared to colocalize with the second quarter of the ZmBs repeat array (Figures 7B to 7D), a similar position as the CentC-rich domain in the B centromere from fiber-FISH analysis (Figures 2A and 2C). On labeled pachytene chromosomes, the location of CENH3 appeared to be near the middle of the major ZmBs array and colocalized with the CentC-containing domain (Lamb et al., 2005). At metaphase I and anaphase I of meiosis, FISH signals derived from CentC were consistently located at the most poleward position on B chromosome bivalents as well as bivalents of Telo2-2(-) (Figure 7E). Taken together, these results provide further support that the CentC-enriched domain within the ZmBs array is the region associated with kinetochore formation.

Relationship between the Amount of CENH3 and the Size of Misdivided B Centromeres

Immunostaining analysis revealed that CENH3 signals are only located at the B centromeres and not at any of the B chromosomal regions containing CentC or the ZmBs repeat (Figures 7F and 7G). Notably, the CENH3 signals in the B centromeres were often weaker than those in the A centromeres (Figure 7F). To investigate the relationship between the amount of centromeric DNA and CENH3, we analyzed several misdivision lines using a combination of FISH and immunolocalization procedures. The CENH3 immunostaining signals derived from the misdivided B centromeres can be unambiguously identified based on their colocalization with the ZmBs repeat (Figure 8). To analyze the results quantitatively, the immunostaining signals were divided into five classes: I, the intensity of the immunostaining signal was similar to those in A centromeres; II, signals weaker than those in A centromeres; III, signals significantly weaker than those in A centromeres; IV, the signal could not be observed without contrast adjustment by computer software; V, signal could not be observed even after contrast adjustment.

We scored 14% class I, 68% class II, 18% class III, but no class IV and V signals from the intact B centromere (Table 2). Telo4-11(-), which contains ~545 kb of the ~700-kb CentC-rich domain, showed a similar amount of CENH3 as intact B centromeres (Table 2). Telo4-4(-), however, with slightly less of this domain (495 kb), contained no class I signals and an increase in the number of class IV and class V signals (Table 2). Ring 4-8(-), with only ~240 kb from the CentC-rich domain, recruited much less CENH3 than the other derivatives. The vast majority of the Ring 4-8(-) signals fell into the class IV and V categories. Finally, Telo3-3(-), which has undetectable amounts of CentC and CRM in the centromeric region, recruited so little CENH3 that it was only weakly

detected in 12% of the cells (class IV) and undetectable (class V) in the remaining cells (Table 2). The very low association of CENH3 with Telo3-3(-) explains the highly unstable nature of this chromosome as noted above.

Discussion

Dissection of the Functional Domain of Maize B Chromosome Centromere

Centromeres in complex eukaryotes often contain arrays of a single class of repetitive DNA elements up to several megabases. However, generally only a portion of such long arrays is associated with CENH3. The repetitive DNA outside of the CENH3-associated domain presumably plays a role in other centromeric functions, such as sister chromatid cohesion and chromosomal condensation (Grewal and Moazed, 2003). For example, the centromere of the human X chromosome contains ~3 Mb of the α-satellite but only a portion of this a-satellite array is associated with centromeric proteins (Schueler et al., 2001; Spence et al., 2002). Association of CENH3 with a subdomain of the centromeric satellite DNA has also been demonstrated in maize (Jin et al., 2004) and Arabidopsis (Shibata and Murata, 2004).

Previous studies demonstrated that the maize B chromosome contains up to 9 Mb of the ZmBs repeat (Kaszas and Birchler, 1996). This ZmBs repeat was considered to be represented in the functional region of the B centromere because all of the B centromere misdivision derivatives retained a portion of the ZmBs repeat and rearranged the restriction pattern. There is a strong correlation between the size of the retained ZmBs repeat array and meiotic transmission (Kaszas and Birchler, 1998). However, it was not known if the ZmBs repeat was the sole DNA element located within the functional B

centromere. It was also found that some misdivided B centromeres retained up to 2.5 Mb of the ZmBs repeat but still showed poor meiotic transmission (Kaszas and Birchler, 1998).

We provide several lines of evidence that an ~700-kb CentC-rich domain represents the primary CENH3 binding and functional kinetochore of the B centromere. (1) The breakpoints of all of the misdivision derivatives analyzed, except for Telo3-3(-), could be recognized within the ~700-kb domain (Figure 5). (2) The relative position of the CENH3 binding domain on stretched B centromeres (Figures 7B to 7D) is similar to the position of the ~700-kb domain on mapped DNA fibers (Figures 2A and 2C). The leading cytological location of the CentC repeats on the B chromosome bivalent at metaphase I also suggests that they are bound within the kinetochore (Figure 7E). (3) The amount of CENH3 in misdivided B centromeres is correlated with the size of the DNA fragment derived from the ~700-kb domain (Table 2). Because the ZmBs repeat is present within this CentC-rich domain (Figure 5), this interpretation is also consistent with prior findings (Kaszas and Birchler, 1996, 1998).

An Estimate of the Normative Size of a Functional Plant Centromere

The size of a functional centromere has been estimated in different species using different approaches. Deletion mapping revealed that the minimum size of a human minichromosome centromere is ~100 kb (Yang et al., 2000). Similarly, deletion mapping placed the minimum centromere of a Drosophila minichromosome within a 420-kb region (Murphy and Karpen, 1995; Sun et al., 1997). The CENP-A binding domains of several human neocentromeres have been determined using ChIP analysis. The chromatin

domain associated with CENP-A in these neocentromeres ranges from ~130 to 460 kb (Lo et al., 2001a, 2001b; Alonso et al., 2003).

In plants, the centromere of rice chromosome 8 (CEN8) contains an ~750-kb region associated with rice CENH3 (Nagaki et al., 2004). Here, we demonstrate that the CENH3 binding region in the maize B centromere is ~700-kb, close in size to the CENH3 binding domain in rice CEN8. We have also previously shown that the centromeres of five of seven maize A chromosomes analyzed contain 300 to 700 kb of intermingled CentC/ CRM sequences that interact with CENH3 (Jin et al., 2004). Our deletion analysis supports the view that these domains correspond to functional centromeric regions. Thus, the functional centromeres of maize A and B chromosomes have a similar size and contain ~300 to 700 kb DNA in the CENH3-associated chromatin.

The B Centromere Is Defined by Both Genetic and Epigenetic Mechanisms

Extensive research in several model eukaryotes has revealed that centromere formation can be controlled by both genetic and epigenetic mechanisms. In humans, fully functional artificial centromeres can be assembled from the α-satellite alone (Harrington et al., 1997; Ikeno et al., 1998), and minor manipulations of the sequence alter this capacity (Ohzeki et al., 2002). However, many human neocentromeres contain no detectable α-satellite DNA, yet associate with the same proteins as normal human centromeres (du Sart et al., 1997; Saffery et al., 2000) and are stable in both mitosis and meiosis (Tyler-Smith et al., 1999; Amor et al., 2004).

The ~700-kb domain in the B centromere contains arrays of three repeats, the ZmBs repeat, CentC, and CRM. These three repeats, however, are not specific to the B centromere and show multiple locations throughout the B chromosome (Lamb et al., 2005). Nevertheless, CENH3 is only associated with the CentC rich domain in the B centromere. Therefore, the repetitive DNA within this domain is possibly marked epigenetically for CENH3 recognition. The CentC repeat is highly specific to the centromeres of maize A chromosomes (Ananiev et al., 1998). It is interesting to note that although the CentC repeat is distributed throughout the B chromosomes, the CentC sequence within the B centromere is restricted to the ~700-kb domain. We previously also demonstrated that maize CENH3 is always associated with intermingled CentC/CRM sequences in A centromeres, although not all of the CentC/CRM sequences are associated with CENH3 (Jin et al., 2004). The CentC/CRM sequences in the ~700-kb region of the B centromere appear to serve a similar role in CENH3 recognition. It has recently been demonstrated in a human neocentromere that the CENH3-associated chromatin is divided into seven blocks by H3-associated chromatin within a 330-kb region (Chueh et al., 2005).

Based on these data, the CentC-containing region within the B centromere appears to be critical for association with CENH3 because removal of CentC within the ~700-kb domain diminishes CENH3 staining. However, when CentC is reduced below detection, as on Telo3-3(-), some CENH3 remains. The small amount of CENH3 remaining on the Telo3-3(-) derivative may be binding to a small amount of epigenetically marked ZmBs repeat derived from the ~700-kb domain. Alternatively, it is possible that sequences adjacent to the CentC-enriched domain are recruited to the

centromere after misdivision, in an analogous fashion to neocentromere formation in Drosophila (Platero et al., 1999; Maggert and Karpen, 2001).

Taken together, the results suggest that the functional centromere of the B chromosome is a small CentC- and CRM-rich domain that is embedded within a much larger array of the ZmBs repeat. The amount of CENH3 found at the B centromere is correlated with the size of this domain and the stability of the chromosome. These findings bolster the view that CentC and CRM are key elements of maize centromeres (Zhong et al., 2002; Jin et al., 2004) and provide genetic support for the idea that CENH3 is necessary for proper kinetochore formation.

Materials and Methods

Materials

A maize (*Zea* mays) line B73 containing multiple B chromosomes was used for cytological analysis of the intact B chromosomes. B centromere misdivision derivatives were produced from the A-B translocation line TB-9Sb (Kaszas and Birchler, 1996, 1998) (Figure 3A). These derivatives were classified by their chromosome type (telocentric chromosome, isochromosome, or ring chromosome) and presence (+) or absence (-) of a knob adjacent to the centromeric region (Kaszas and Birchler, 1996). The centromeric DNA probes, including CentC, CRM, and the ZmBs repeat, were described previously (Alfenito and Birchler, 1993; Nagaki et al., 2003b).

FISH and Fiber-FISH

The FISH and fiber-FISH were performed according to published protocols (Jiang et al., 1995; Jackson et al., 1998). DNA probes were labeled with biotin-dUTP or digoxigenin-dUTP (Roche, Indianapolis, IN). Chromosomes were counterstained by 4',6-diamidino-2-phenylindole in an antifade solution Vectashield (Vector Laboratories, Burlingame, CA). Images were captured digitally using a Sensys CCD camera (Roper Scientific, Tucson, AZ) attached to an Olympus BX60 epifluorescence microscope (Tokyo, Japan). The camera control and imaging analysis were performed using IPLab Spectrum v3.1 software (Signal Analytics, Vienna, VA).

Sequential Detection of CENH3 and the ZmBs Repeat on Chromosomes and Stretched Centromeres

Preparation of somatic chromosomes and stretched centromeres for immunostaining was performed according to Jin et al. (2004) with only minor modifications. Maize nuclei were isolated from young maize kernels rather than callus. The preparations were fixed in 4% paraformaldehyde in PBS for 10 min. Approximately 100 mL rabbit anti-CENH3 antibody (1:200 in TNB buffer [0.1 M Tris-HCl, pH 7.5, 0.15 M NaCl, and 0.5% blocking reagent]) was added to the slides. After incubation in a humid chamber at 378C for 3 h, the slides were washed three times in PBS before adding 100 mL of rhodamine anti-rabbit secondary antibody (1:50 in TNB buffer). The slides were incubated and washed again and fixed in 4% paraformaldehyde. The slides were then probed with the ZmBs repeat. Chromosomes and interphase nuclei were counterstained with 49,6-diamidino-2-phenylindole. Labeling and detection of DNA

repeats and CENH3 on maize pachytene chromosomes were described previously (Lamb et al., 2005).

ChIP

ChIP was performed as described previously (Nagaki et al., 2003a). The nuclei were isolated from leaf tissue and digested with micrococcal nuclease. The resultant nucleosomes were incubated with the maize anti- CENH3 antibody (Zhong et al., 2002). The immune complexes were precipitated and separated into unbound (Sup, for supernant) and bound (Pel, for pellet) fractions. Equal amounts of the Sup and Pel fractions were blotted on membranes and hybridized with ^{32}P-labeled ZmBs repeat. The hybridization was quantified using a phosphor-imager. Mock experiments using preimmunized rabbit serum served as nonspecific binding control for each ChIP assay. The %IP, defined as pel/(pel + sup) of the mock experiments, was subtracted in each case from %IP of the anti-CENH3 treatments. We used 18S-26S rRNA genes (rDNA) and CentC as negative and positive controls, respectively.

Acknowledgements

This research was supported by National Science Foundation Grant 9975827 to R.K.D., J.A.B., and J.J. J.L. was supported by a University of Missouri Life Sciences Graduate Fellowship.

References

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

Alonso, A., Mahmood, R., Li, S., Cheung, F., Yoda, K., and Warburton, P.E. (2003). Genomic microarray analysis reveals distinct locations for the CENP-A binding domains in three human chromosome 13q32 neocentromeres. Hum. Mol. Genet. 12, 2711–2721.

Amor, D.J., Bentley, K., Ryan, J., Perry, J., Wong, L., Slater, H., and Choo, K.H.A. (2004). Human centromere repositioning ‘‘in progress’’. Proc. Natl. Acad. Sci. USA 101, 6542–6547.

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

Carlson, W.R. (1970). Nondisjunction and isochromosome formation in the B chromosome of maize. Chromosoma 30, 356–365.

Carlson, W.R., and Chou, T.-S. (1981). B chromosome nondisjunction in corn: Control by factors near the centromere. Genetics 97, 379–389.

Cheng, Z.K., Buell, C.R., Wing, R.A., and Jiang, J. (2002). Resolution of fluorescence in-situ hybridization mapping on rice mitotic prometaphase chromosomes, meiotic pachytene chromosomes and extended DNA fibers. Chromosome Res. 10, 379–387.

Chueh, A.C., Wong, L.H., Wong, N., and Choo, K.H.A. (2005). Variable and hierarchical size distribution of L1-retroelement-enriched CENP-A clusters within a functional human neocentromere. Hum. Mol. Genet. 14, 85–93.

Darlington, C.D. (1939). Misdivision and the genetics of the centromere. J. Genet. 37, 341–364.

du Sart, D., Cancilla, M.R., Earle, E., Mao, J.I., Saffery, R., Tainton, K.M., Kalitsis, P., Martyn, J., Barry, A.E., and Choo, K.H.A. (1997). A functional neo-centromere formed through activation of a latent human centromere and consisting of non-alpha-satellite DNA. Nat. Genet. 16, 144–153.

Grewal, S.I.S., and Moazed, D. (2003). Heterochromatin and epigenetic control of gene expression. Science 301, 798–802.

Harrington, J.J., Bokkelen, G.V., Mays, R.W., Gustashaw, K., and Willard, H.F. (1997). Formation of de novo centromeres and construction of first-generation human artificial microchromosomes. Nat. Genet. 15, 345–355.

Henikoff, S. (2002). Near the edge of a chromosome’s ‘black hole’. Trends Genet. 18, 165–167.

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

Houben, A., and Schubert, I. (2003). DNA and proteins of plant centromeres. Curr. Opin. Plant Biol. 6, 554–560.

Ikeno, M., Grimes, B., Okazaki, T., Nakano, M., Saitoh, K., Hoshino, H., McGill, N.I., Cooke, H., and Masumoto, H. (1998). Construction of YAC-based mammalian artificial chromosomes. Nat. Biotechnol. 16, 431–439.

Jackson, S.A., Wang, M.L., Goodman, H.M., and Jiang, J. (1998). Application of fiber-FISH in physical mapping of Arabidopsis thaliana. Genome 41, 566–572.
Jiang, J., Birchler, J.B., Parrott, W.A., and Dawe, R.K. (2003). A molecular view of plant centromeres. Trends Plant Sci. 8, 570–575.
Jiang, J., Gill, B.S., Wang, G.L., Ronald, P.C., and Ward, D.C. (1995). Metaphase and interphase fluorescence in situ hybridization mapping of the rice genome with bacterial artificial chromosomes. Proc. Natl. Acad. Sci. USA 92, 4487–4491.
Jin, W.W., Melo, J.R., Nagaki, K., Talbert, P.B., Henikoff, S., Dawe, R.K., and Jiang, J. (2004). Maize centromeres: Organization and functional adaptation in the genetic background of oat. Plant Cell 16, 571–581.
Kaszas, E., and Birchler, J.A. (1996). Misdivision analysis of centromere structure in maize. EMBO J. 15, 5246–5255. Kaszas, E., and Birchler, J.A. (1998). Meiotic transmission rates correlate with physical features of rearranged centromeres in maize. Genetics 150, 1683–1692.
Lamb, J.C., Kato, A., and Birchler, J.A. (2005). Sequences associated A chromosome centromeres are present throughout the maize B chromosome. Chromosoma 113, 337–349.
Lo, A.W., Craig, J.M., Saffery, R., Kalitsis, P., Irvine, D.V., Earle, E., Magliano, D.J., and Choo, K.H. (2001a). A 330 kb CENP-A binding domain and altered replication timing at a human neocentromere. EMBO J. 20, 2087–2096.
Lo, A.W.I., Magliano, D.J., Sibson, M.C., Kalitsis, P., Craig, J.M., and Choo, K.H.A. (2001b). A novel chromatin immunoprecipitation and array (CIA) analysis identifies a 460-kb CENP-A-binding neocentromere DNA. Genome Res. 11, 448–457.
Maggert, K.A., and Karpen, G.H. (2001). The activation of a neocentromere in Drosophila requires proximity to an endogenous centromere. Genetics 158, 1615–1628.
Murphy, T.D., and Karpen, G.H. (1995). Localization of centromere function in a Drosophila minichromosome. Cell 82, 599–609.
Nagaki, K., Cheng, Z.K., Ouyang, S., Talbert, P.B., Kim, M., Jones, K.M., Henikoff, S., Buell, C.R., and Jiang, J. (2004). Sequencing of a rice centromere uncovers active genes. Nat. Genet. 36, 138–145.
Nagaki, K., Song, J., Stupar, S.M., Parokonny, A.S., Yuan, Q., Ouyang, S., Liu, J., Hsiao, J., Jones, K.M., Dawe, R.K., Buell, C.R., and Jiang, J. (2003b). Molecular and cytological analyses of large tracks of centromeric DNA reveal the structure and evolutionary dynamics of maize centromeres. Genetics 163, 759–770.
Nagaki, K., Talbert, P.B., Zhong, C.X., Dawe, R.K., Henikoff, S., and Jiang, J.M. (2003a). Chromatin immunoprecipitation reveals that the 180-bp satellite repeat is the key functional DNA element of Arabidopsis thaliana centromeres. Genetics 163, 1221–1225.
Oakey, R., and Tyler-Smith, C. (1990). Y chromosome DNA haplotyping suggests that most European and Asian men are descended from one of two males. Genomics 7, 325–330.
Ohzeki, J., Nakano, M., Okada, T., and Masumoto, H. (2002). CENP-B box is required for de novo centromere chromatin assembly on human alphoid DNA. J. Cell Biol. 159, 765–775.

Platero, J.S., Ahmad, K., and Henikoff, S. (1999). A distal heterochromatic block displays centromeric activity when detached from a natural centromere. Mol. Cell 4, 995–1004. Roman, H. (1947). Mitotic non-disjunction in the case of interchanges involving the B-type chromosome in maize. Genetics 32, 391–409.
Saffery, R., Irvine, D.V., Griffiths, B., Kalitsis, P., Wordeman, L., and Choo, K.H.A. (2000). Human centromeres and neocentromeres show identical distribution patterns of >20 functionally important kinetochore-associated proteins. Hum. Mol. Genet. 9, 175–185.
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.
Schueler, M.G., Higgins, A.W., Rudd, M.K., Gustashaw, K., and Willard, H.F. (2001). Genomic and genetic definition of a functional human centromere. Science 294, 109–115.
Sears, E.R. (1952). Misdivision of univalents in common wheat. Chromosoma 4, 535–550.
Shibata, F., and Murata, M. (2004). Differential localization of the centromere-specific proteins in the major centromeric satellite of Arabidopsis thaliana. J. Cell Sci. 117, 2963–2970.
Spence, J.M., Critcher, R., Ebersole, T.A., Valdivia, M.M., Earnshaw, W.C., Fukagawa, T., and Farr, C.J. (2002). Co-localization of centromere activity, proteins and topoisomerase II within a subdomain of the major human X alpha satellite array. EMBO J. 21, 5269– 5280.
Sullivan, B.A., Blower, M.D., and Karpen, G.H. (2001). Determining centromere identity: Cyclical stories and forking paths. Nat. Rev. Genet. 2, 584–596.
Sun, X.P., Wahlstrom, J., and Karpen, G. (1997). Molecular structure of a functional Drosophila centromere. Cell 91, 1007–1019. Talbert, P.B., Masuelli, R., Tyagi, A.P., Comai, L., and Henikoff, S. (2002). Centromeric localization and adaptive evolution of an Arabidopsis histone H3 variant. Plant Cell 14, 1053–1066.
Topp, C.N., Zhong, C.X., and Dawe, R.K. (2004). Centromere-encoded RNAs are integral components of the maize kinetochore. Proc. Natl. Acad. Sci. USA 101, 15986–15991.
Tyler-Smith, C., Gimelli, G., Giglio, S., Floridia, C., Pandya, A., Terzoli, G., Warburton, P.E., Earnshaw, W.C., and Zuffardi, O. (1999). Transmission of a fully functional human neocentromere through three generations. Am. J. Hum. Genet. 64, 1440–1444.
Wevrick, R., and Willard, H.F. (1989). Long-range organization of tandem arrays of a satellite DNA at the centromeres of human chromosomes: High frequency array-length polymorphism and meiotic stability. Proc. Natl. Acad. Sci. USA 86, 9394–9398.
Willard, H.F. (1998). Centromeres: The missing link in the development of human artificial chromosomes. Curr. Opin. Genet. Dev. 8, 219–225.
Wu, J.Z., et al. (2004). Composition and structure of the centromeric region of rice chromosome 8. Plant Cell 16, 967–976.

Yang, J.W., Pendon, C., Yang, J., Haywood, N., Chand, A., and Brown, W.R. (2000). Human mini-chromosomes with minimal centromeres. Hum. Mol. Genet. 9, 1891–1902.

Zhang, Y., Huang, Y.C., Zhang, L., Li, Y., Lu, T.T., Lu, Y.Q., Feng, Q., Zhao, Q., Cheng, Z.K., Xue, Y.B., Wing, R.A., and Han, B. (2004). Structural features of the rice chromosome 4 centromere. Nucleic Acids Res. 32, 2023–2030.

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.

Tables

Table 1. Distribution of CentC, CRM, and the ZmBs Repeat among Normal and Misdivision Derivatives of Maize B Centromeres

Lines	n	Full Size (µm)	Full Size (kb)[a]	ZmBs Repeats (kb)	ZmBs Repeats (kb)[b]	Presence of CentC	Presence of CRM	Transmission[c]
Intact B[d]	10	>1150	>3690	>3000	>9000	Yes	Yes	-
Core domain of B	10	220+/-14	705+/-45	~245	ND[e]	Yes	Yes	-
Ring4-8(-)	17	137+/-13	440+/-42	~220	~500	Yes	Yes	ND
Telo4-11(-)	21	257+/-17	825+/-55	~500	~2360	Yes	Yes	42%
Telo4-4(-)	12	228+/-19	732+/-62	~500	~490	Yes	Yes	10%
Ring4-12(-)	12	29+/-3	93+/-10	~60	ND	No	Yes	ND
Telo2-2(-) 1	8	225+/-15	722+/-48	~330	~2150	Yes	Yes	43%
Telo2-2(-) 2	7	95+/-14	305+/-45	~200	ND	No	No	-
Telo3-3(-)	14	87+/-6	280+/-19	~185	~280	No	No	13%

a The length of the fiber-FISH signals (mm) was converted to kilobases using a 3.21 kb/mm conversion rate (Cheng et al., 2002).
b The sum of fragments containing the ZmBs repeat arrays was estimated by PFGE (Kaszas and Birchler, 1998).
c Data from Kaszas and Birchler (1998).
dWe used the longest fiber-FISH signals to represent the minimum size.
e ND, not determined.

Table 2. Signal Intensities of CENH3 Immunostaining in Normal and Misdivision Derivatives of the Maize B Chromosome

Lines	Size (kb)	Core Domain (kb)	n	I [a]	II [b]	III [c]	IV [d]	V [e]
Intact B	>3690	~705	89	13 (14.6%)	59 (66.3%)	17 (19.1%)	0 (0.0%)	0 (0.0%)
Telo4-11(-)	825	~545	98	17 (17.3%)	65 (66.3%)	15 (15.3%)	1 (1.0%)	0 (0.0%)
Telo4-4(-)	732	~495	87	0 (0.0%)	43 (49.4%)	36 (41.4%)	6 (6.9%)	5 (5.7%)
Telo4-8(-)	440	~240	65	0 (0.0%)	0 (0.0%)	2 (3.1%)	26 (40.0%)	37 (56.9%)
Telo3-3(-)	280	0	34	0 (0.0%)	0 (0.0%)	0 (0.0%)	4 (11.8%)	30 (88.2%)

a The intensity of CENH3 signals on the B centromere is similar to those of A centromeres.
b The intensity of CENH3 signals on the B centromere is weaker than those of A centromeres.
c The intensity of CENH3 signals on the B centromere is significantly weaker than those of A centromeres.
d The signal can not be observed without contrast adjustment using computer software.
e The signal can not be observed even with contrast adjustment.

Figures

Figure 1. Locations of the CentC, CRM, ZmBs, and Knob Repeats on the Maize Pachytene B Chromosome.
Minor sites of these repeats on the B chromosome are not shown. The four distinctive heterochromatin blocks on the long arm are illustrated. The diagram is based on Lamb et al. (2005).

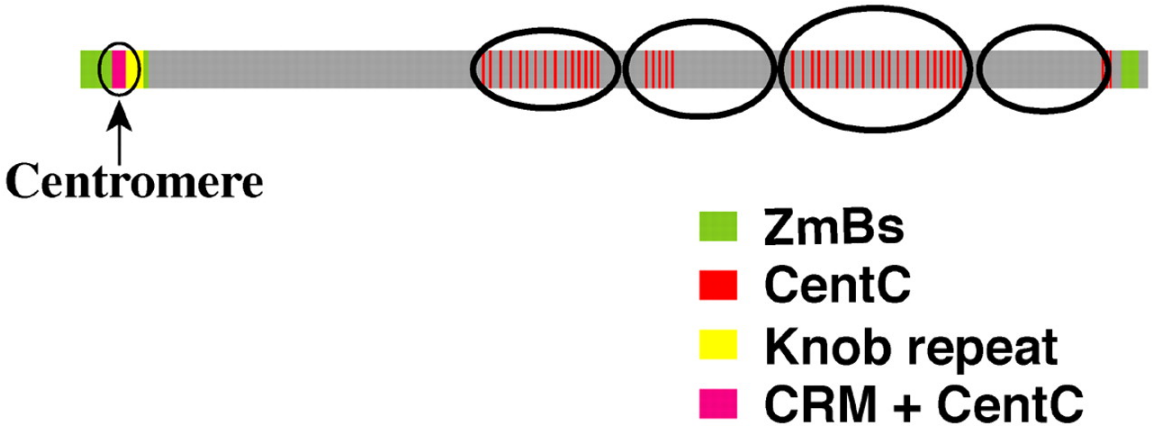


Figure 2. Organization of CentC, CRM, and the ZmBs Repeat in the Centromere of Maize B Chromosomes.
(A) A fiber-FISH signal derived the ZmBs repeat (green) and CentC (red). Arrows mark the border of the CentC-enriched domain.
(B) A diagram based on the fiber-FISH signal pattern in (A). Unambiguous signals from CentC were only observed between the two arrows.
(C) A fiber-FISH signal derived the ZmBs repeat (green) and CRM (red). Arrows mark the border of the CentC-enriched domain.
(D) A diagram based on the fiber-FISH signal pattern in (C). Additional CRM signals located outside of the CentC-enriched domain are marked by arrowheads.
(E) An expanded image of the domain in (A). The five ZmBs repeat arrays are marked.
(F) An expanded image of the domain in (C). The five ZmBs repeat arrays are marked.
(G) A diagram depicting the distribution patterns of the major blocks of CentC, CRM, and ZmBs. Note: the first ZmBs repeat array contains a distinctive CRM insertion in the middle. Bars = 100 μm in (A) and (C) and 20 μm in (E) and (F).

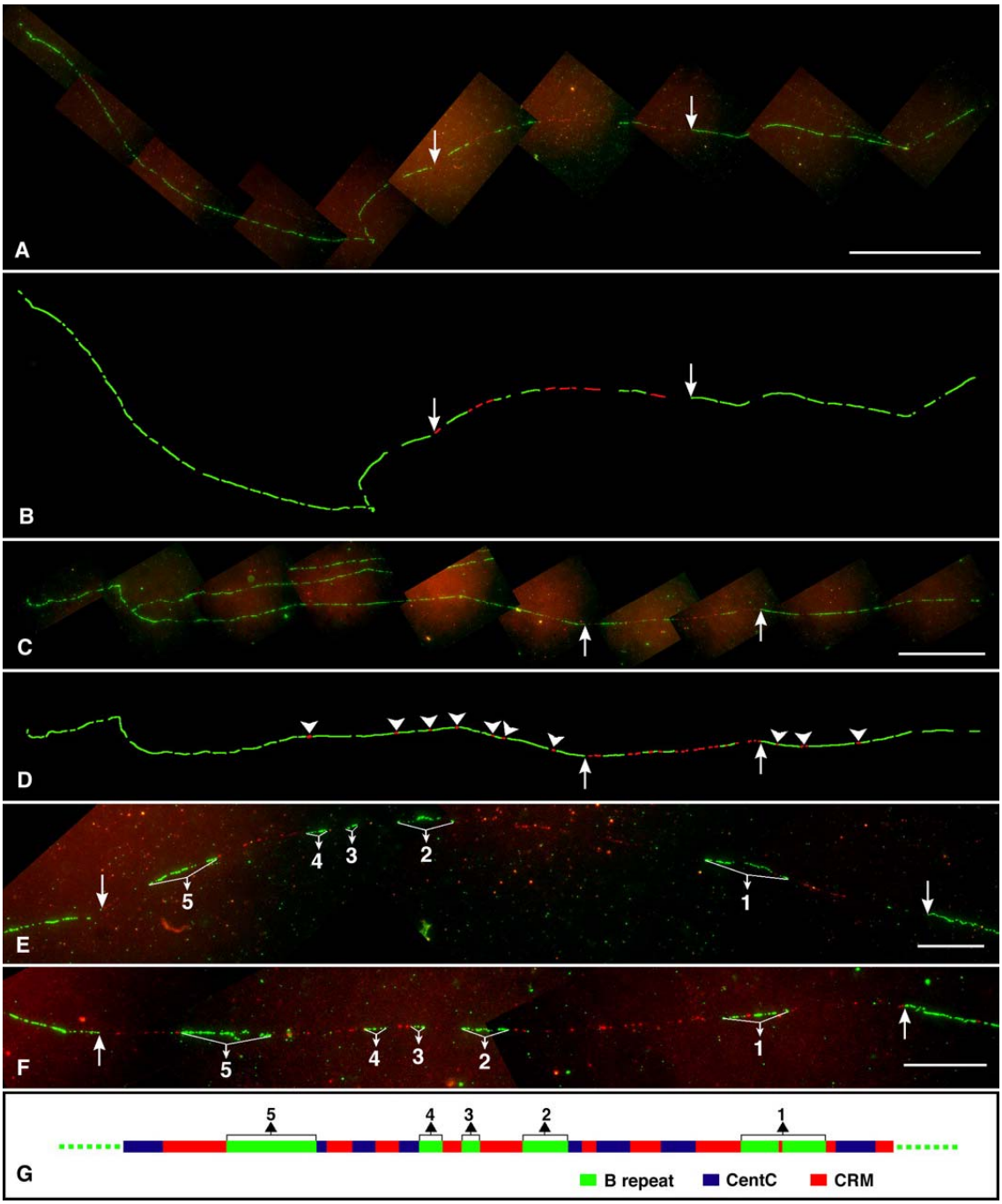

A
B
C
D
E
4
3
2
5
1
F
5
4
3
2
1
G
5
4
3
2
1
B repeat
CentC
CRM

Figure 3. The Pedigree and Structure of the B Centromere Misdivision Derivatives.
(A) The pedigree of B centromere misdivision derivatives used in this study.
(B) Graphical representation of the centromere size of B chromosome misdivision derivatives based on Kaszas and Birchler (1998). The closed boxes represent the centromeric region.

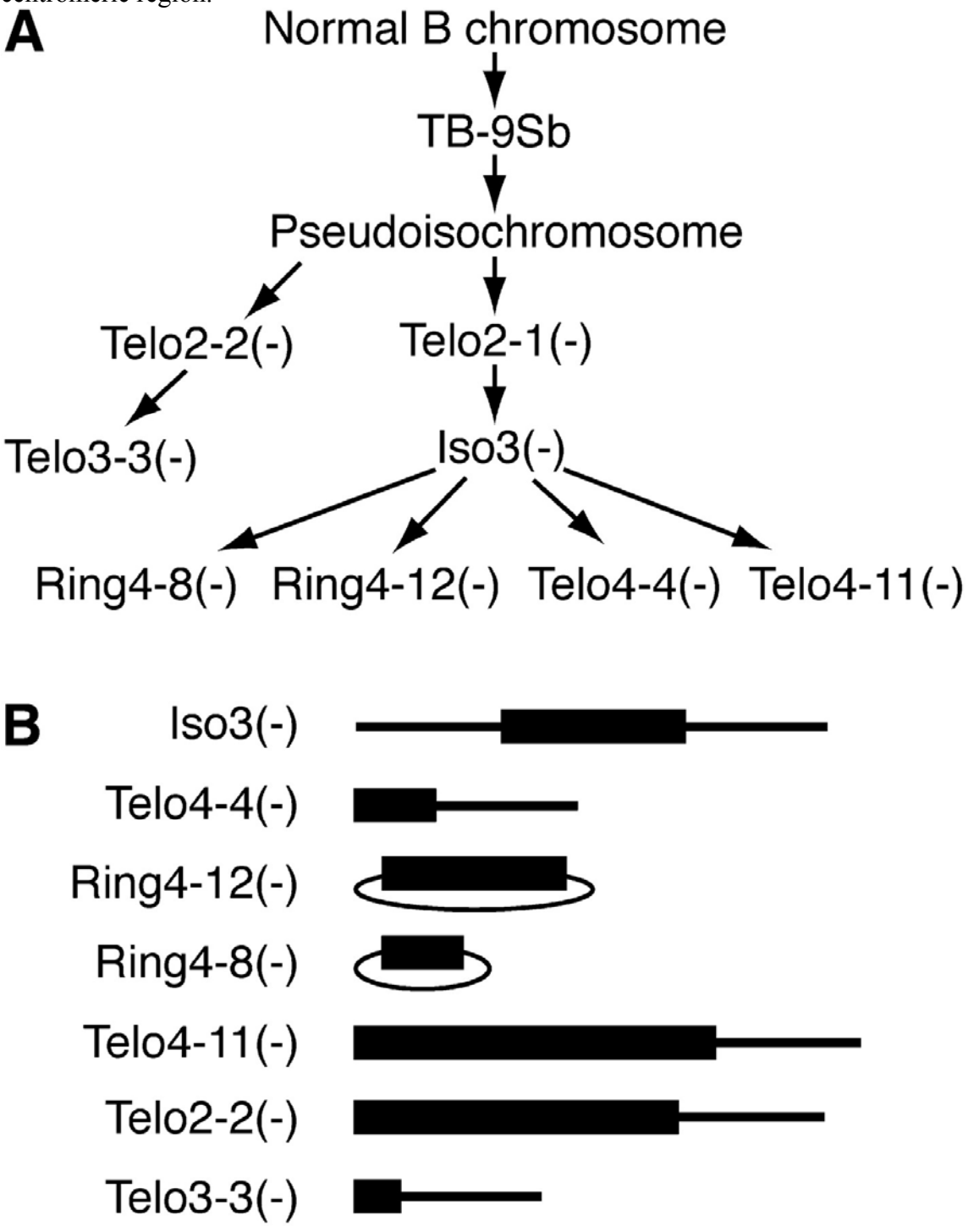


Figure 4. Distribution Patterns of CentC, CRM, and the ZmBs Repeat in the Centromeres of B Centromere Misdivision Derivatives.
The ZmBs repeat arrays derived from the CentC-enriched domain are marked. Bars = 30 mm in (E), 25 mm in (A) and (B), 20 mm in (G), 10 μm in (C), (D), (H), and (I), and 5 μm in (F).
(A) A fiber-FISH signal of ZmBs (green) and CRM (red) from the centromere of Telo4-11(-).
(B) A fiber-FISH signal of ZmBs (green) and CentC (red) from the centromere of Telo4-11(-).
(C) A fiber-FISH signal of ZmBs (green) and CentC (red) from the centromere of Ring4-8(-).
(D) A fiber-FISH signal of ZmBs (green) and CRM (red) from the centromere of Ring4-8(-).
(E) A fiber-FISH signal of ZmBs (green) and CRM (red) from the centromere of Telo4-4(-).
(F) A fiber-FISH signal of ZmBs (green) and CRM (red) from the centromere of Ring4-12(-). Note: the ZmBs array is disrupted by a CRM insertion (arrowhead), which is characteristic of the first ZmBs array in the domain.
(G) A fiber-FISH signal from ZmBs (green) and a mixed probe of CRM and CentC (red) from the centromere of Telo2-2(-) (fragment 1).
(H) A fiber-FISH signal of ZmBs (green) from the centromere of Telo2-2(-) (fragment 2). No CRM and CentC signals were detected within this fragment.
(I) A fiber-FISH signal of ZmBs (green) from the centromere of Telo3-3(-). No CRM and CentC signals were detected within this fragment.

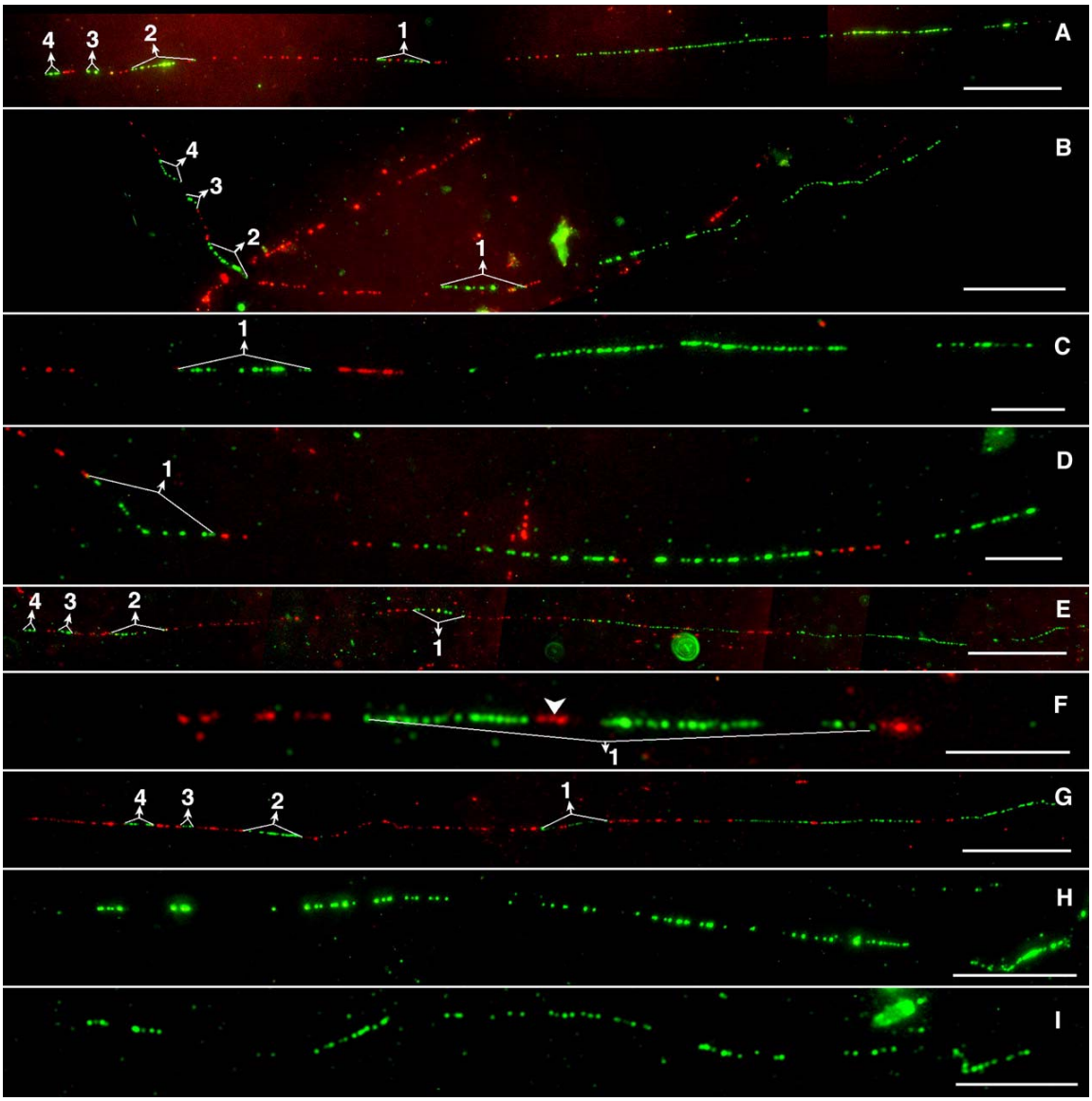

A
4
3
2
1
B
4
3
2
1
C
1
D
1
E
4
3
2
1
F
1
G
4
3
2
1
H
I

Figure 5. Diagrams of DNA Organization of Five B Centromere Misdivision Derivatives.
All five lines, Telo4-11(-), Telo4-4(-), Ring4-8(-), Ring4-12(-), and Telo2-2(-), contain a DNA fragment derived from the CentC-enriched domain of the original B centromere. The misdivision breakpoints of these five derivatives are all located in the middle of this domain. Vertical arrows point to the borders of the ~700-kb CentC-enriched domain.

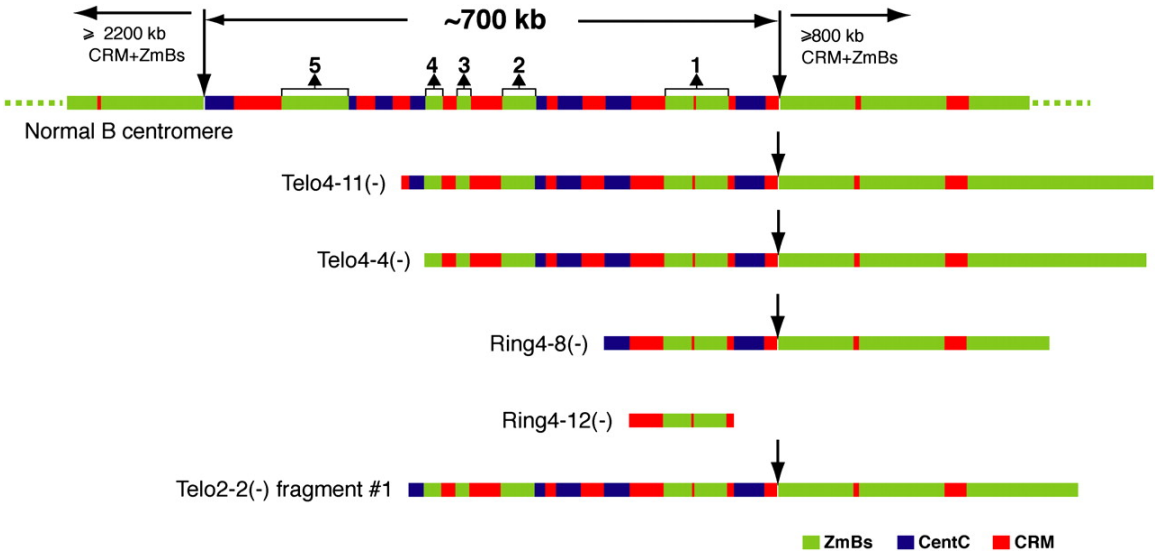


Figure 6. Pachytene Localization of the ZmBs and CENH3.
(A) Pachytene spreads from Telo2-2(-), Telo4-4(-), and Telo3-3(-) (as an isochromosome in this line) were immunolabeled with CENH3 (in green), hybridized with ZmBs (in red), and counterstained with 4',6-diamidino-2- phenylindole.
(B) Only the ZmBs hybridization signal is shown. The arrow indicates a minor hybridization signal distinct from the main ZmBs hybridization site. The arrowhead indicates a cluster of ZmBs on the distal portion of the reciprocal A-B translocation chromosome in this line.
(C) Only the CENH3 signal is shown. Arrows point to the CENH3 signals associated with the misdivided B centromeres. Complete or partial colocalization of ZmBs and CENH3 is always observed for Telo2-2(-) and Telo4-4(-), whereas Telo3-3(-) shows no colocalization in some spreads.
(D) A Telo2-2(-) pachytene chromosome labeled with ZmBs (in red) and CentC (in green). Arrow indicates a minor hybridization signal distinct from the main ZmBs hybridization site. Arrowhead indicates the location of CentC hybridization to the B chromosome heterochromatin of the B-A translocation.

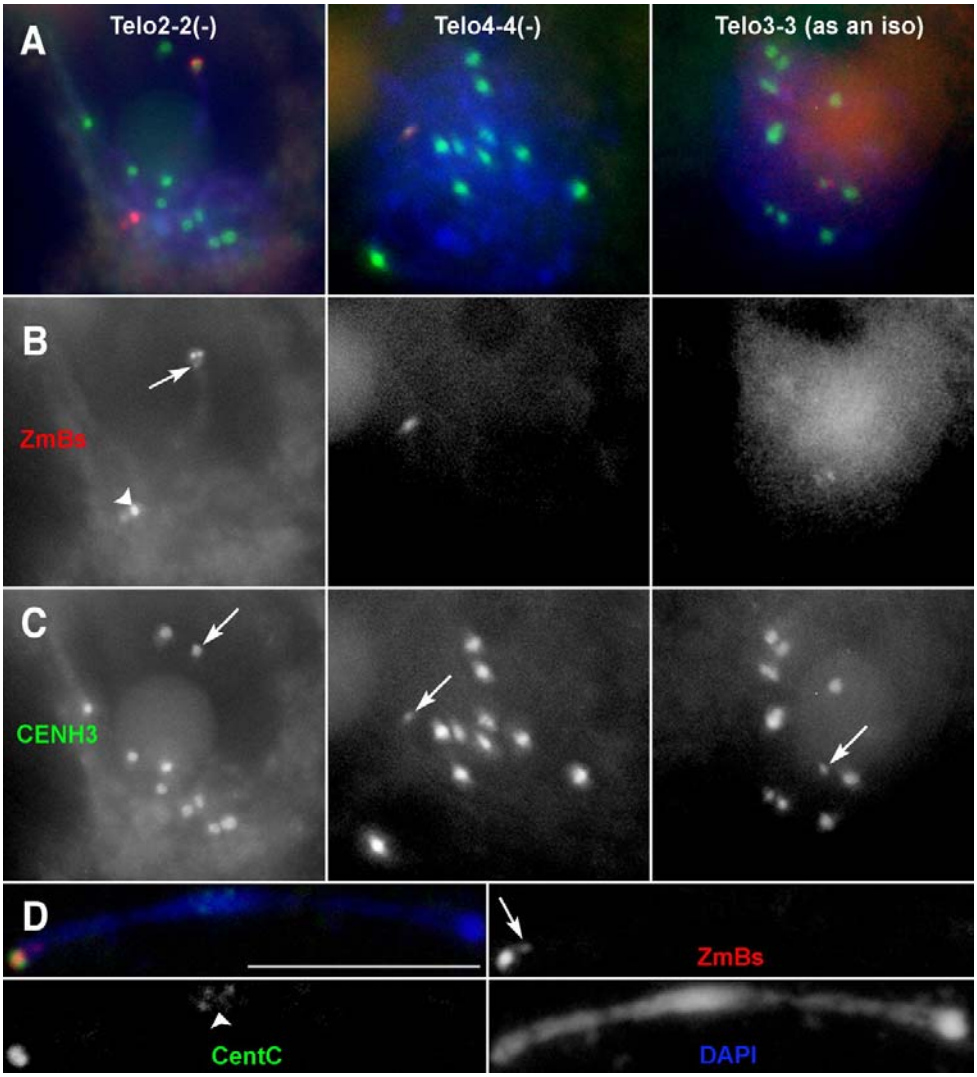


Figure 7. ChIP Analysis and Cytological Mapping of CENH3 on Stretched Chromatin Fibers and Meiotic and Mitotic Chromosomes.
(A) ChIP analysis of CentC and the ZmBs repeat in maize line B73+3B. The ZmBs repeat was coimmunoprecipitated with the anti-CENH3 antibody in all three experiments. The %IP of ZmBs repeat was significantly different from that of the negative control rDNA.
(B) A stretched B centromere is stained by the maize anti-CENH3 antibody. The arrows mark the borders of the CENH3 domain.
(C) FISH mapping on the same stretched B centromere using the ZmBs repeat.
(D) A merged image of the immunostaining and FISH signals. Bar = 10 μm.
(E) FISH mapping of CentC (red) and the ZmBs repeat (green) on metaphase I chromosomes of maize B73 line containing a pair of B chromosomes. The Cent repeats (arrows) are located at the most poleward position (ahead of the ZmBs repeats) on the B chromosome bivalent.
(F) Immunostaining signals from the anti-CENH3 antibody on the somatic metaphase chromosomes of maize line B73+3B. Arrows point to the signals on the B chromosomes, which are weaker than those on the A centromeres.
(G) A merged image of the immunostaining signals with chromosomes, which are stained with 4',6-diamidino-2-phenylindole and are pseudocolored in red. Bar = 5 μm.

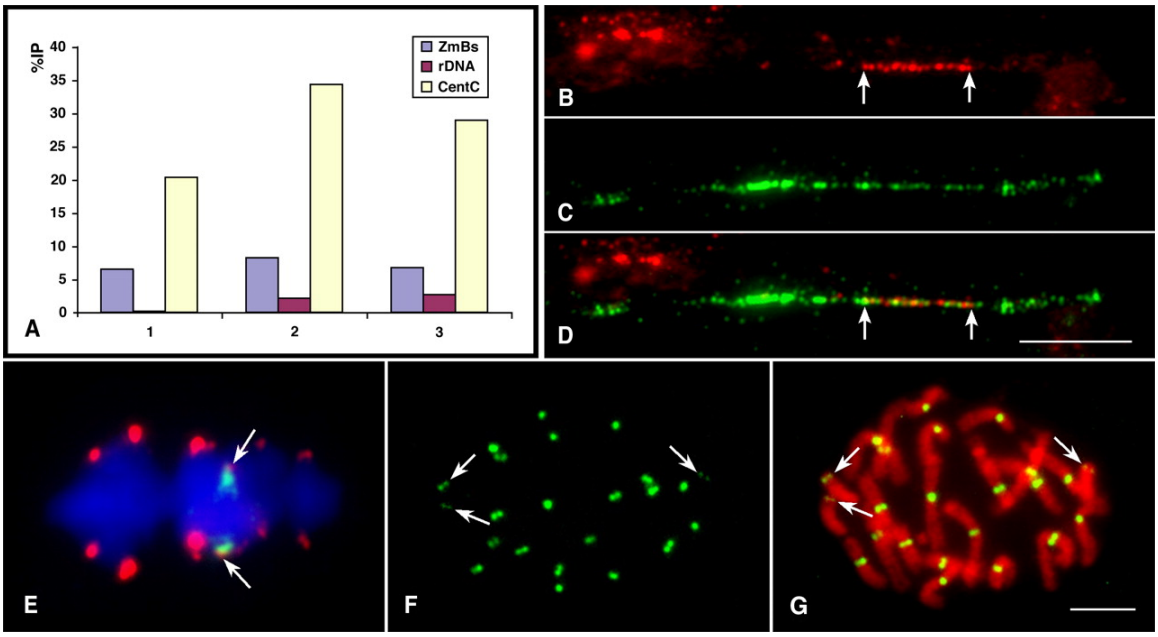
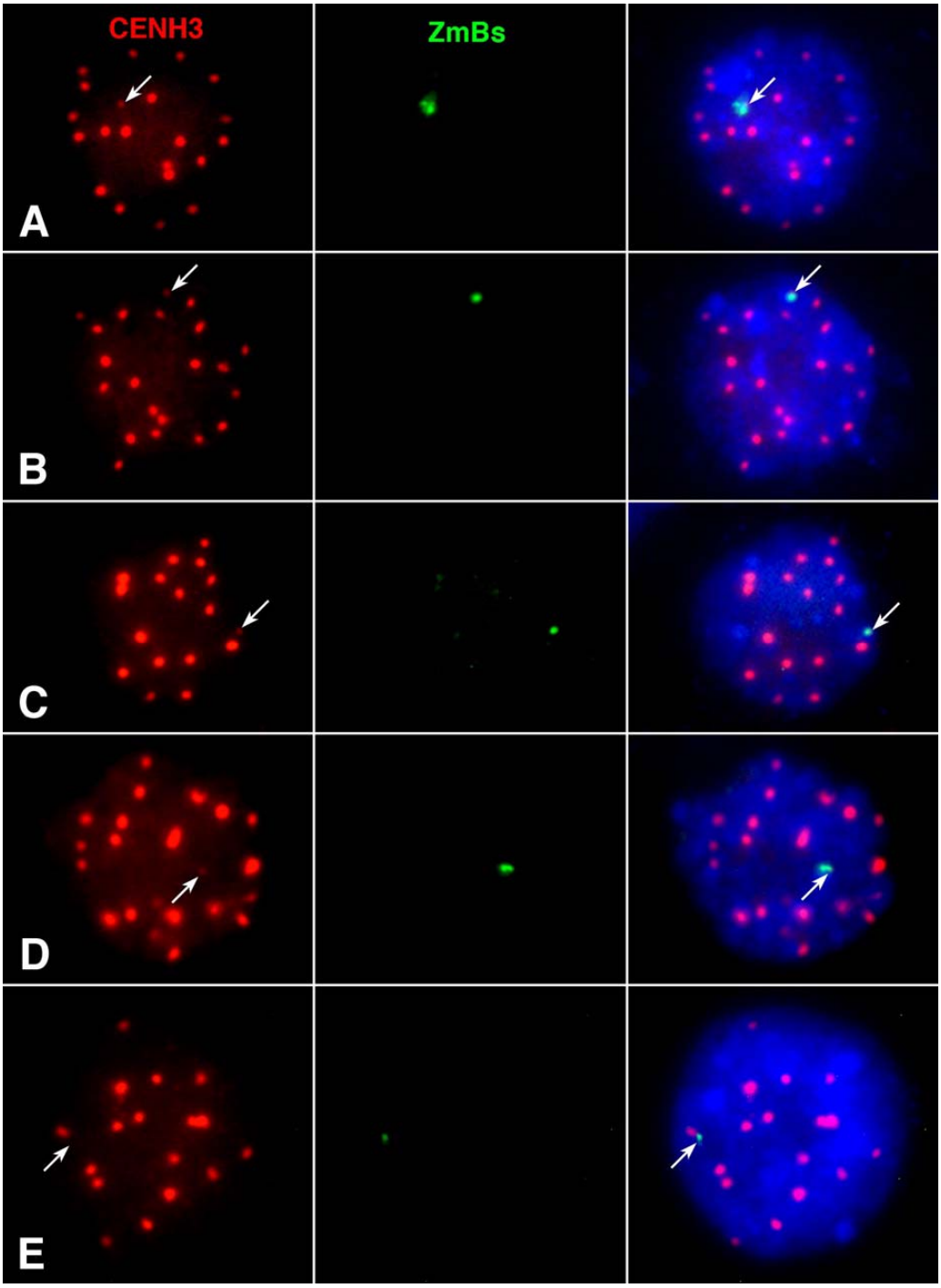


Figure 8. Interphase Detection of CENH3 and the ZmBs Repeat in a Maize B73 Line with a Single B Chromosome, Telo4-11(-), Telo4-4(-), Ring4-8(-), and Telo3-3(-). (A) Maize B73 line with a single B chromosome. (B) Telo4-11(-). (C) Telo4-4(-). (D) Ring4-8(-). (E) Telo3-3(-). The signals associated with normal or the misdivided centromeres are indicated by arrows. Note: the ZmBs site in (E) is not associated with any CENH3 staining.



5. HIGH FREQUENCY OF CENTROMERE INACTIVATION RESULTING IN STABLE DICENTRIC CHROMOSOMES OF MAIZE

Dr. Kato had initially begun cytological characterization of a collection of minichromosomes before he left the lab for a faculty position in Japan. Dr. Kato had noted that some of the minichromosomes appeared to have two sites of centromeric elements. Dr. Fangpu Han began work in Dr. Birchler's lab during my third year of study and continued the work with the minichromosomes. Because of my involvement in the initial characterization of the minichromosomes and familiarity with the manner of their production, Dr. Han solicited my involvement in his work. Dr. Han performed all crosses and conducted the initial screen of root tips from the minichromosome collection. I performed all the antibody labeling reactions of the CenH3 protein. For the remainder of the experiments, it is not possible to separate our individual involvement because we worked together so intimately. In many instances, one of us would begin a procedure and the other would finish it. In other cases, Dr. Han would perform a labeling procedure and I would capture the images. I wrote the manuscript with input from Drs. Han and Birchler.

The article was published in PNAS with Dr. Han and me as co-first authors. The reference style of that journal has been retained.

HAN, F., J. C. LAMB and J. A. BIRCHLER, 2006 High frequency of centromere inactivation resulting in stable dicentric chromosomes of maize. Proc Natl Acad Sci U S A **103:** 3238-3243.

High frequency of centromere inactivation resulting in stable dicentric chromosomes of maize

Fangpu Han[*†], Jonathan C. Lamb[*†] and James A. Birchler[*‡]

[*] Division of Biological Sciences, University of Missouri-Columbia

117 Tucker Hall, Columbia, MO, 65211-7400, U S A

[†]These authors contributed equally to this work.

[‡] To whom correspondence should be addressed.

E-mail: BirchlerJ@missouri.edu

Tel : (573) 882-4905 Fax : (573) 882-0123

Abstract

Somatic chromosome spreads from maize (*Zea mays* L.) plants containing B-A translocation chromosomes undergoing the chromosome type breakage-fusion-bridge (BFB) cycle were examined by FISH. The size and type of extra chromosomes varied among cells of the same individual. A collection of minichromosomes derived from the chromosome type BFB cycle was examined for the presence of stable dicentric chromosomes. Six out of twenty-three chromosomes in the collection contained two regions with DNA sequences typical of centromeres. Functional analysis as well as immunolabeling of CENH3, the centromere-specific histone H3 variant, revealed only one functional centromere per chromosome despite the duplicate centromere sequences. One plant was found with an inactive B centromere that had been translocated to the short arm of chromosome 9. The translocated centromere region appeared identical to that of a normal B chromosome. The inactivation of the centromeres was stable for at

least four generations. By using dicentrics from dispensable chromosomes, centromere inactivation was found to be quite common under these circumstances.

Introduction

Chromosomal rearrangements or *de novo* centromere formation can produce two linked centromeres, which will migrate separately to newly forming daughter cells, forming a chromatin bridge that will break. The broken ends may subsequently fuse reforming a dicentric chromosome, albeit with some of the intervening chromatin missing. This process is called the breakage-fusion-bridge (BFB) cycle (1, 2). When stable dicentric chromosomes have been recovered, they are functionally monocentric. In dicentric chromosomes with well-separated centromeres, stabilization occurs by the poorly understood phenomenon of centromere inactivation (3). Stability can also be achieved if the centromeres are very close together and form only one heterochromatic block (3). To date, no examples of an inactivated plant centromere have been reported.

Extra chromosomes, called B chromosomes, have been identified in diverse taxa (4) including maize (5). The presence of these chromosomes has little obvious effect on the phenotype of a plant that harbors them, yet they persist by taking advantage of the cellular mechanisms responsible for faithful chromosome replication and transmission. Because these chromosomes are entirely dispensable but contain the essential components required for efficient transmission through mitosis and meiosis, they provide an excellent model to study centromeres.

Previous reports of the chromosome type BFB cycle describe the fate of a translocation chromosome involving the B chromosome and a variant of chromosome 9,

which contains an inverted duplication of its short arm (6, 7). The duplicated section of 9S can fold back and recombine with itself during meiosis I creating a chromosome that will be cleaved during anaphase II because it contains two centromeres. The B9-Dp9 chromosome undergoes non-disjunction at the second pollen mitosis and two broken chromosomes can be delivered to the zygote initiating the chromosome type BFB cycle (6). Because this chromosome is dispensable and moves independently from the intact chromosome 9, it provides a method to track the progress of a dicentric chromosome throughout the lifecycle and to recover the resulting chromosomes.

We describe six independent stable dicentric chromosomes resulting from this process including a chromosome in which the B centromere has been transferred to the short arm of chromosome 9. In all cases, only one centromere is functional. By using a dispensable chromosome, it was possible to demonstrate that centromere inactivation can be quite common and can occur in plants.

Results

Cytological examination of somatic metaphase chromosomes in root tips undergoing the BFB cycle

A previous study of the BFB cycle involving the B9-Dp9 chromosome examined cells during telophase for double bridges, which indicate that the chromosome type BFB cycle is active (6). In about one third of the plants, minichromosomes were observed in meiotic samples (6). Subsequent work involved assembling a collection of these minichromosomes, which varied in size, transmission rate and type of DNA elements that

were present (7). Figure 1 illustrates the hypothesized events leading to minichromosome formation.

Fluorescence in situ hybridization (FISH) using the B chromosome specific element, ZmBs, and the 180bp knob heterochromatin repeat probes was performed on somatic chromosome spreads from root tips resulting from a hybrid between a tester stock and a male containing one copy of the B9-Dp9 and two copies of the reciprocal 9-B chromosome in order to visualize the effects of the BFB cycle. The ZmBs probe allows the centromere of the B chromosome to be identified and the 180bp knob repeat hybridizes to a location near the B centromere as well as to the knob normally found near the tip of chromosome 9 (8-10).

Some of the progeny contained intact B9-Dp9 chromosomes that have passed through meiosis without experiencing a crossover in the duplicated region (Figure 2A). All chromosome spreads examined from this type of root tip contained two B9-Dp9 chromosomes. In other root tips, the number and appearance of chromosomes varied from cell to cell indicative of the BFB cycle (Figure 2B-D). Large and small dicentric chromosomes were observed as well as telocentric chromosomes and small chromatin fragments. Up to four chromosomes with ZmBs signals were present in a single cell, all being telocentric. In some cells, two distinct dicentric chromosomes were observed (Figure 2C). In root tips collected from three week old seedlings, no dicentric chromosomes were found, only very small fragments and larger telocentric chromosomes.

Dicentric minichromosomes resulting from the BFB cycle

Because the BFB cycle can be stopped by inactivation of one centromere, the collection of minichromosomes (7) was screened for stable dicentrics. Chromosomes at the pachytene stage of meiosis are much more extended than at somatic metaphase; therefore, meiotic samples were collected and hybridized with FISH probes against the centromeric elements ZmBs, CentC and CRM as well as the 180bp knob repeat. Of the twenty-three chromosomes examined (including nine previously unreported cases), five were identified that contained two sites where ZmBs signal colocalized with CentC and CRM (minichromosomes #2, 3, 5, 10, and 13, Figure 3; Supplementary Figure 1). An additional stable dicentric chromosome was identified that contained a translocated B centromere on chromosome 9S (discussed below).

By comparing the FISH signal patterns of the minichromosomes to those of an intact B centromeric region (Supplementary Figure 2), some of the re-arrangements leading to their formation can be determined. The centromeric region of the B chromosome contains two blocks of ZmBs that are megabases in length and that flank a 700kb "core" consisting of interspersed CentC, CRM, and ZmBs elements. This core contains the functional kinetochore binding domain and is the site of CENH3 localization (9, 11). Adjacent to the block of ZmBs on the long arm, there is a stretch of 180bp knob repeat followed by another block of ZmBs intermixed with CentC and CRM (8, 9). The short arm of chromosome 9 has a site of strong hybridization to the 180bp knob repeat, which appears as two large signals on the B9-Dp9 chromosome (Figure 2A).

At the cytological level minichromosomes #2, 3, and 13 contained two identical B centromere segments (Figure 3). This type of structure is one predicted outcome of the

BFB process (Figure 1) and might result when the breaks in both chromatids of the double bridge occur in the same location. ZmBs, CentC, CRM and the 180bp knob repeat hybridize to multiple sites in minichromosome #10, indicating that a complex series of rearrangements formed this minichromosome.

The ZmBs hybridization signals on minichromosome #5 were well separated and both colocalized with CentC and CRM signals (Figure 3). Only one of the two regions of centromeric element hybridization also contained 180bp knob repeat signals and the intensity of the CentC/CRM signal was greater at this site than the other. The lack of 180bp knob repeat at that centromere suggests that as two active centromeres pulled the chromatin, a break occurred very near to one of the B centromeres removing the block of 180bp knob repeat, two blocks of ZmBs, and a portion of the 700kb "core" domain. The resulting fragment contained little more than the centromere. This small fragment was healed by attachment to another chromatid fragment that resulted from a break further out on the chromosome arm.

Minichromosomes with two centromeric regions contain a single functional centromere

Antibodies against CenH3, the centromeric H3 histone variant, were used to label cytological preparations containing the stable dicentric minichromosomes to confirm that only one centromere remained active. In each case, a single site of localization was observed (Figure 4). For minichromosomes #2, 3, and 13, which contain two identical centromere regions, it was not possible to distinguish which of the centromeric regions contained the functional centromere. In minichromosome #5, the region with no 180bp knob labeling and less hybridization signal to the centromeric element probes is the

region that was labeled by anti-CenH3 antibodies (Figure 4). During anaphase, the smaller region of centromeric elements was stretched toward the two poles and leads the remainder of the chromosome (Figure 4). The smaller site of centromeric elements was the sole site of CenH3 recruitment in five different plants observed over two generations, which indicated that the location of the active centromere is stable. In minichromosome 10, the site of CenH3 labeling is also distinct from the larger site of ZmBs hybridization (Figure 4).

In minichromosomes #3, 10 and 13, FISH showed a distinct site of CentC and CRM located between two blocks of ZmBs hybridization (Figure 3) similar to the distribution seen on an intact B centromere (9, 11). The CenH3 labeling also appeared at this position in the minichromosomes and the intact B centromere (Figure 4). In minichromosome #5, the larger region of centromeric element hybridization shows a similar pattern of element distribution as the intact B centromere but this region does not function as a centromere (Figure 4A and B). Instead, the CenH3 labeling is immediately adjacent to the other, smaller area of ZmBs (Figure 4A). CentC and CRM elements also hybridize to this site (Figure 3E). Thus CenH3 appears to be associated with the same interspersion of CentC, CRM, and ZmBs as occurs in the intact B centromere. Therefore, inactivation of one centromere did not affect the positioning of the remaining active centromere.

Origin of an A-B dicentric translocation chromosome in maize

During a screen for additional minichromosomes, a large chromosome was discovered that contained strong ZmBs and intermediate 180bp knob repeat signals at the

tip of its short arm. Application of a cocktail of FISH probes that allows the maize karyotype to be identified (10) indicated that this chromosome contained two centromeric regions, one from a B chromosome, and the other from chromosome 9. Hereafter, this chromosome is referred to as 9-B inactive centromere-1 (9-Bic-1). CentC and CRM signals colocalized with the ZmBs signal at the tip of 9-Bic-1 (Figure 5A and B, Supplementary Figure 3).

Self pollination of plants carrying this chromosome resulted in progeny containing two copies of 9-Bic-1. Plants homozygous for this chromosome were albino and died at the seedling stage, most probably because of removal of the very distal part of chromosome 9S including the *white deficiency* (*wd*) gene. Kernels on the ears resulting from self-pollination were all recessive *bronze1* (*bz1)* and dominant *Colored1* (*C1)* (Supplementary Figure 4). Because the dominant *C1* allele was present in all kernels (including homozygotes for 9-Bic-1), it must be present on the 9-Bic-1 chromosome. This information allowed the breakpoint of the translocation to be placed on the genetic map between *wd* and *c1*. The original B-9Dp9 chromosome contained the dominant *Bz1* allele, suggesting that 9-Bic-1 was formed by placement of a B centromeric region onto the recessive *bz1* tester chromosome. This could have occurred via a meiotic recombination between the *bz1* tester chromosome and a dicentric minichromosome (Supplementary Figure 5). Pairing between 9-Bic-1 and chromosome 9 occurred along most of the chromosomes during meiosis (Figure 5C and D). The missing portion on 9S is not critical for gametophyte function as 9-Bic-1 was transmitted through both the male and female flowers.

Pachytene FISH analysis of the translocated B centromere of 9-Bic-1

The 9-Bic-1 chromosome was labeled with FISH probes for CentC, CRM, ZmBs, and the 180bp knob repeat to compare the patterns of hybridization to the intact B chromosome. The ZmBs probe hybridized to three blocks, two located immediately adjacent to each other (and that appeared as a single large block in most spreads) and the third was separated by a short distance. The CentC and CRM signal appeared at a site located between the closely adjoined ZmBs signals with an additional weak signal colocalizing with the third block of ZmBs in some spreads (Figure 5C and D). This pattern is the same as an intact B centromeric region (9, 11). Because of the close proximity of the large knob on the normal chromosome 9 that was paired with 9-Bic-1, it was not possible to observe the 180bp knob repeat signal on the translocated B region of 9-Bic-1. However, the 180bp knob signal is revealed in somatic preparations in close proximity to the ZmBs signal (Figure 5B). Thus, the translocated B centromere contained all the same elements as the progenitor B centromere and at the cytological level appeared unaltered.

The translocated B centromere is inactive

Interphase root tip nuclei and meiotic samples containing 9-Bic-1 were labeled with antibodies against CENH3 and hybridized with ZmBs. CENH3 signal did not colocalize with ZmBs (Figure 5E and F). At anaphase, the ZmBs signal lagged behind the centromere from chromosome 9 (Figure 5G and H) indicating that it did not form a functional kinetochore or bound microtubules. In addition to the site of microtubule attachment, centromeric regions serve as the site of sister chromatid cohesion. During

mitotic metaphase, the portions of the chromosomes that remain bound are marked by phosphorylation of the serine 10 residue of the canonical H3 histone protein (12, 13). Mitotic chromosome spreads containing 9-Bic-1 were labeled using antibodies against this chromatin modification (pSer10-H3). Only the chromosome 9 centromeric region was labeled (Supplementary Figure 6). Additionally, ZmBs signals on sister chromatids of 9-Bic-1 separated at mitotic metaphase while the centromeric regions from chromosome 9 remain attached (Figure 5A and B, Supplementary Figure 3). The 9-Bic-1 chromosome has been maintained for more than four generations and over one hundred root tips have been examined as kernels were scored for the 9-Bic-1 chromosome during routine classification. No evidence of chromosome instability that would arise from reactivation of the translocated B centromere has been observed. Taken together, this evidence indicated that the translocated B centromere region located on 9S is stably inactive for kinetochore formation and sister chromatid cohesion typical of centromeres at metaphase.

Discussion

Many studies have been conducted in plants to examine the passage of dicentric chromosomes through cell division, but there are no previous reports of inactivated plant centromeres. In other species, inactivation of one centromere has allowed the recovery of dicentric chromosomes leading to the conclusion that the mere presence of particular DNA elements at centromeres is not sufficient to organize a centromere (14). The previous inability to observe inactive plant centromeres may have been due to the difficulty in recovering chromosomes with large deficiencies, which would be selected

against in the gametophytic generation. The approach described in this study, namely, to use the maize B chromosome to study dicentrics, takes advantage of the dispensable nature of the maize B chromosome. Any chromosomes that result, including the stabilized dicentric chromosomes, segregate independently of the other chromosomes and can be recovered for study.

One third of the plants containing a B9-Dp9 chromosome undergoing the BFB have a minichromosome observed in meiosis (6). In the remaining plants, the B9-Dp9 chromosome is completely lost during development. This loss could occur as dicentric chromosomes lag behind other chromosomes and are excluded from both daughter nuclei. Non-disjunction at any cell division could also result in loss of the additional chromosome in one cell lineage and the gain of a chromosome in another. Some metaphase cells of root tips from B9-Dp9 progeny undergoing the BFB cycle were observed with two dicentric chromosomes or with three or four telocentric chromosomes (Figure 2), suggesting that non-disjunction has occurred during at least two mitotic divisions subsequent to dicentric formation in meiosis. Normal maize B chromosomes regularly fail to disjoin at the second pollen division placing two copies of the B chromosome into one sperm and none into the other (15). Also, non-disjunction occurs at a low level in the first pollen division (16), in tapetal cells (17) and in endosperm cells (18). Non-disjunction of a rye B chromosome results when the chromosome lags behind the other chromosomes at the metaphase plate (19). Frequent non-disjunction of the dicentric B9-Dp9 chromosome during development may result from the tendency of functionally dicentric chromosomes to lag or from the action of the B chromosome non-disjunction mechanism.

Every round of the BFB cycle presents an opportunity for the chromosome to be stabilized by healing the ends of the chromosome or by inactivating a centromere. Six out of the 23 recovered minichromosomes contained inactivated centromeres. Thus, while centromere inactivation is not as frequent as end healing in stabilizing the dicentric chromosomes, the mechanism of centromere inactivation cannot be considered an extremely rare event. Except for 9-Bic-1, the dicentric chromosomes were very small indicating that centromere inactivation occurred after most of the chromatin from the B chromosomes and the duplicated portion of 9S had already been lost, suggesting that smaller dicentric chromosomes are more prone to centromere inactivation. Such inactivation might result by nondisjunction in which one centromere fails to attach to the spindle and does not recover activity in subsequent mitoses.

Previously, a collection of translocation chromosomes containing broken or misdivided maize B centromeres was created and analyzed (11, 20, 21). The misdivisions occur primarily in the 700 kb "core" and some of the resulting centromere derivatives retain only a small portion of their functional chromatin. Despite losing most of the centromeric sequences, these derivatives are able to form functional centromeres demonstrating that even a small portion of the regular DNA elements of the B centromere is sufficient for centromere function.

Both the active and inactive centromeres of the minichromosomes as well as the translocated portion of chromosome 9-Bic-1 all appear to retain at least a portion of the functional core of the B centromere and in many cases the core region appears intact cytologically. Therefore, the inactivation of the B centromere of 9-Bic-1 and the inactive centromeres on the minichromosomes are not due to a lack of suitable sequences to form

a centromere. When the two centromeric regions could be distinguished, as in minichromosome #5 and chromosome 9-Bic-1, the inactivation of an intact B centromere was shown to be stable over several generations.

Because the inactivated centromere contains all DNA elements found in a functional centromere, it can be concluded that primary sequence alone is not sufficient for centromere maintenance in plants. Recovery of barley chromosomes lacking any detectable centromeric elements is complementary evidence that DNA sequence is not necessary for centromere formation (22). Similarities in structure and a shared lack of sequence dependence in plants and animals suggest a common mechanism for centromere identity in the two kingdoms.

Materials and Methods

Plant Materials

Lines containing minichromosomes #2, #3, #5, #10, #13 and #23 were selected from the progeny of plants that were hemizygous for the B9-Dp9 chromosome together with two 9-B chromosomes (6). Individual plants from each line were scored cytologically for the presence of a B chromosome specific element, ZmBs (23), and then grown in the greenhouse or field. Immature tassels were fixed in ethanol: acetic-acid (3:1, v/v) on ice for 2 hours, and transferred to 70% ethanol and stored at -20°C.

Fluorescence In Situ Hybridization and immunostaining

Probes were prepared as described previously (10) except that 10 units of polymerase I was used in the nick translation reaction. Chromosome preparation, FISH,

image capturing and image processing were performed as previously described (10) with the following modifications. Digested root tips were washed and broken in 70% ethanol. The cells were rinsed in 100% ethanol and resuspended in 100% acetic acid before application to slides. The probe mixture (4 ng/ul of each probe in 2X SSC and 1X TE buffer) was heated for five minutes at 95°C then cooled on ice before applying to slides. Probe and chromosomes were denatured together by heating at 100°C for five minutes. Tissue preparation and immunolabeling was performed as described previously (9, 12).

Acknowledgements

This work was supported by the National Science Foundation (DBI0421671 and DBI0423898), the USDA (2002-01280), and a MU Life Sciences Fellowship to JL. We thank M. Sivaguru and B. Hardiman at the Molecular Cytology Core, Life Sciences Center, for their help with image acquisition.

References

1. McClintock, B. (1941) *Genetics* **26,** 234-282.
2. McClintock, B. (1939) *Proc Natl Acad Sci U S A* **25,** 405-416.
3. Sullivan, B. A. & Willard, H. F. (1998) *Nat Genet* **20,** 227-8.
4. Jones, G. H. & Rees, H. (1982) *B Chromosomes* (Academic Press, London, UK).
5. Longley, A. E. (1927) *Journal of Agricultural Research* **35,** 796.
6. Zheng, Y. Z., Roseman, R. R. & Carlson, W. R. (1999) *Genetics* **153,** 1435-44.
7. Kato, A., Zheng, Y. Z., Auger, D. L., Phelps-Durr, T., Bauer, M. J., Lamb, J. C. & Birchler, J. A. (2005) *Cytogenet Genome Res* **109,** 156-65.
8. Hsu, F. C., Wang, C. J., Chen, C. M., Hu, H. Y. & Chen, C. C. (2003) *Genetics* **164,** 1087-97.
9. Lamb, J. C., Kato, A. & Birchler, J. A. (2005) *Chromosoma* **113,** 337-49.
10. Kato, A., Lamb, J. C. & Birchler, J. A. (2004) *Proc Natl Acad Sci U S A* **101,** 13554-9.
11. Jin, W., Lamb, J. C., Vega, J. M., Dawe, R. K., Birchler, J. A. & Jiang, J. (2005) *Plant Cell* **17,** 1412-23.
12. Kaszas, E. & Cande, W. Z. (2000) *J Cell Sci* **113 (Pt 18),** 3217-26.
13. Schroeder-Reiter, E., Houben, A. & Wanner, G. (2003) *Chromosome Res* **11,** 585-96.
14. Sullivan, B. A., Blower, M. D. & Karpen, G. H. (2001) *Nat Rev Genet* **2,** 584-96.
15. Roman, H. (1947) *Genetics* **32,** 391-409.
16. Rusche, M. L., Mogensen, H. L., Shi, L., Keim, P., Rougier, M., Chaboud, A. & Dumas, C. (1997) *Genetics* **147,** 1915-21.
17. Chiavarino, A. M., Rosato, M., Manzanero, S., Jimenez, G., Gonzalez-Sanchez, M. & Puertas, M. J. (2000) *Genetics* **155,** 889-97.
18. Alfenito, M. R. & Birchler, J. A. (1990) *Maydica* **35,** 359-366.
19. Jones, N. & Houben, A. (2003) *Trends Plant Sci* **8,** 417-23.
20. Kaszas, E. & Birchler, J. A. (1996) *EMBO J* **15,** 5246-55.
21. Kaszas, E. & Birchler, J. A. (1998) *Genetics* **150,** 1683-92.
22. Nasuda, S., Hudakova, S., Schubert, I., Houben, A. & Endo, T. R. (2005) *Proc Natl Acad Sci U S A* **102,** 9842-7.
23. Alfenito, M. R. & Birchler, J. A. (1993) *Genetics* **135,** 589-97.

Figures

Figure 1. The B9-Dp9 chromosome initiates the BFB cycle after crossing over with itself.
(1) The duplicated portion of B9-Dp9 pairs with itself during meiosis I and a crossover creates a dicentric chromosome and an acentric fragment. (2) During meiosis II, the sister centromeres migrate to opposite poles and the resulting chromatin bridge is broken. (3) A chromosome with a broken end is delivered to the microsporocyte. (4) In the microsporocyte, the broken chromosome is replicated and fuses with itself. The centromeres migrate to opposite poles during the first pollen mitotic division breaking the chromosome. (5) After the first pollen division, the broken chromosome is replicated in the generative cell. Because the B centromere usually undergoes non-disjunction during the second pollen division, both centromeres move to one pole. (6) In the subsequent somatic division after fertilization, the dicentric chromosome is replicated and the centromeres could move to the poles independently from one another. If two centromeres on the same chromatid move to opposite poles, then a double bridge is formed that might break at different locations. The broken ends of the chromosomes in each daughter cell can fuse leading to another round of the chromosome type BFB cycle. As this process continues, chromosomes can be stabilized by inactivation of one of the centromeres or by healing of a broken end by addition of telomeres leading to a variety of minichromosomes.

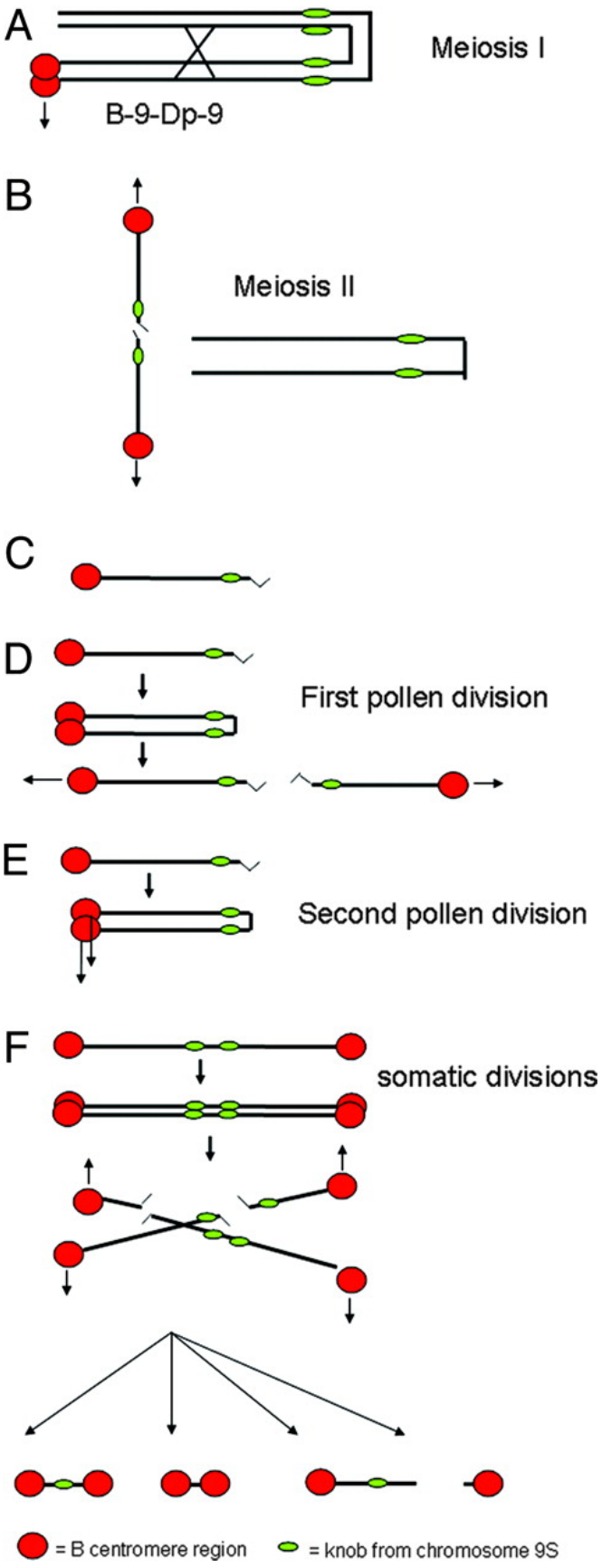

A
Meiosis I
B-9-Dp-9
B
Meiosis II
C
D
First pollen division
E
Second pollen division
F
somatic divisions
= B centromere region
= knob from chromosome 9S

Figure 2. Breakage-fusion-bridge cycle in root tips.
A plant containing one B9-Dp9 and two 9-B translocation chromosomes was crossed as a male to a *yg2*, *bz1* tester line. The resulting kernels were germinated, chromosome spreads prepared from the root tips, and FISH performed using the B chromosome specific ZmBs probe (green) and the 180bp knob heterochromatin probe (red). Lack of crossing over in the duplicated region followed by non-disjunction during the second pollen mitosis results in two intact B9-Dp9 chromosomes as in (A). In (B-D), different cells from a single root tip vary in the size and number of extra chromosomes. The arrowhead indicates the ZmBs signal at the tip of the 9-B chromosome and arrows indicate the ZmBs signal of the centromeric region.

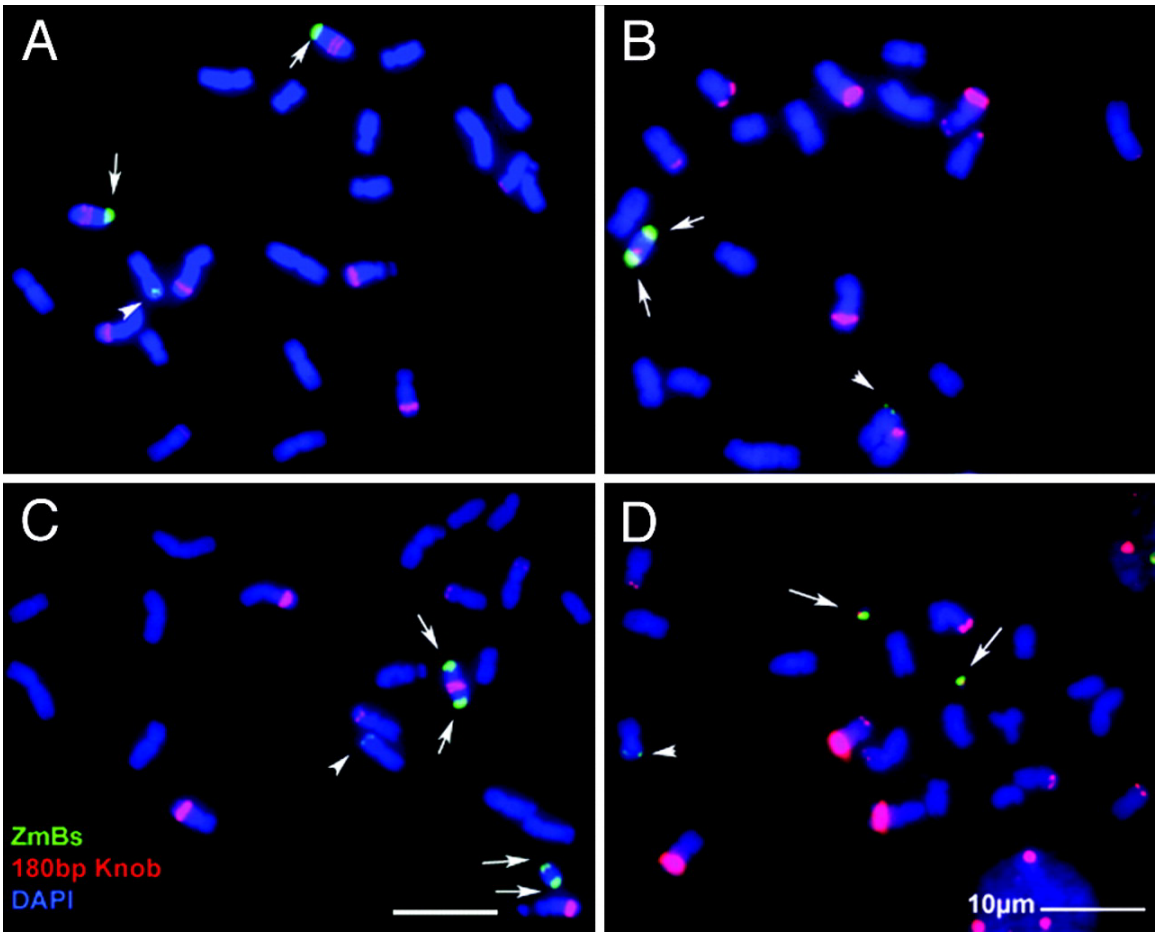


Figure 3. Pachytene FISH of dicentric minichromosomes.
(A) A pachytene spread with an intact B chromosome (indicated by the arrow) and minichromosome #5 hybridized with ZmBs (green) and CRM (red). The arrowhead indicates the active centromere of minichromosome #5. (B) A pachytene spread with two unpaired copies of minichromosome #10 hybridized with ZmBs (green) and CRM (red) illustrates the relative size of the minichromosomes. In (C) through (G) minichromosomes #2, 3, 5, 10, and 13, respectively, are depicted from pachytene spreads. Chromosomes were hybridized with ZmBs (green) and CentC, CRM, or the 180bp knob repeat (red). The gray values for the probes are also displayed; first ZmBs and second the indicated repetitive probe (CentC, CRM, or the 180bp knob repeat).

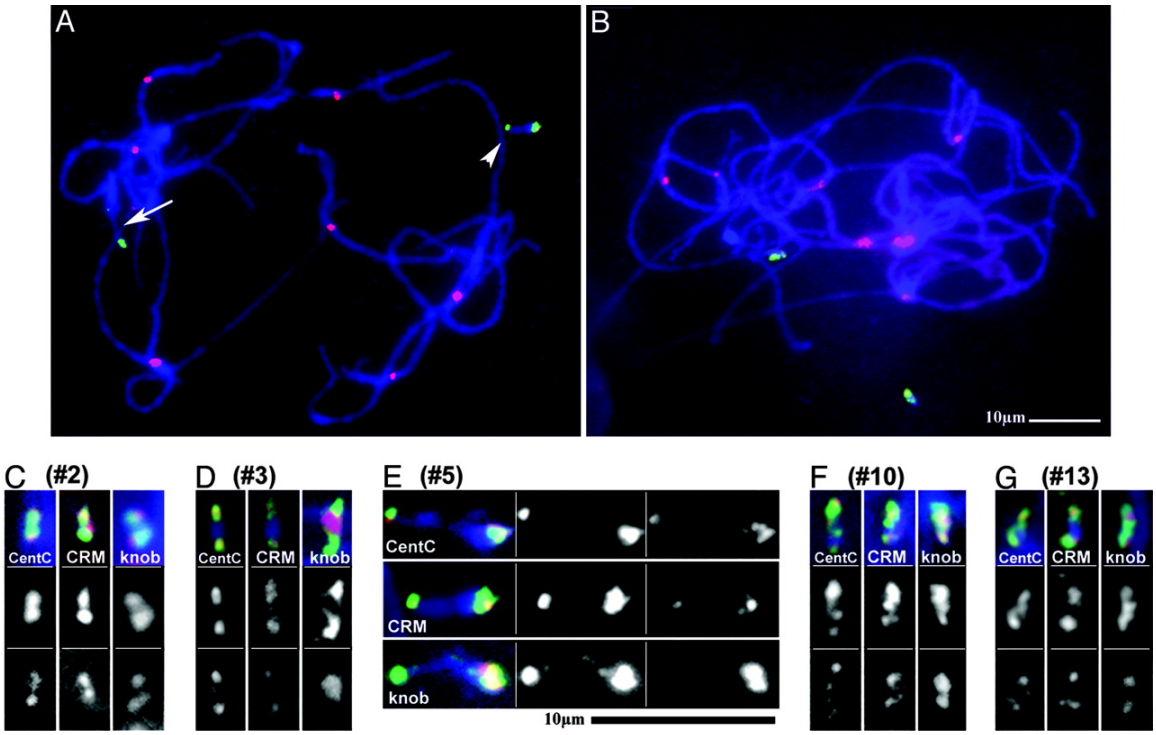


Figure 4. Localization of the active centromeres of the minichromosomes.
(A) Meiotic samples containing minichromosomes #5 (indicated with an arrowhead) and an intact B chromosome (indicated by an arrow) with immunolabled CENH3 (red) hybridized with ZmBs (green). (B) Anaphase I cell spread containing two unpaired copies of minichromsome #5 hybridized with ZmBs (green) and the 180bp knob repeat (red). The smaller sites of ZmBs hybridization (indicated with arrowheads in one homolog) move toward the poles. The figures below show CENH3 immunolabeling (red) and ZmBs (green) hybridization of the minichromosomes and an intact B chromosome. Only one site of CENH3 labeling is observed per chromosome. The gray values for ZmBs (middle) and CENH3 (bottom) are presented below the merged image.

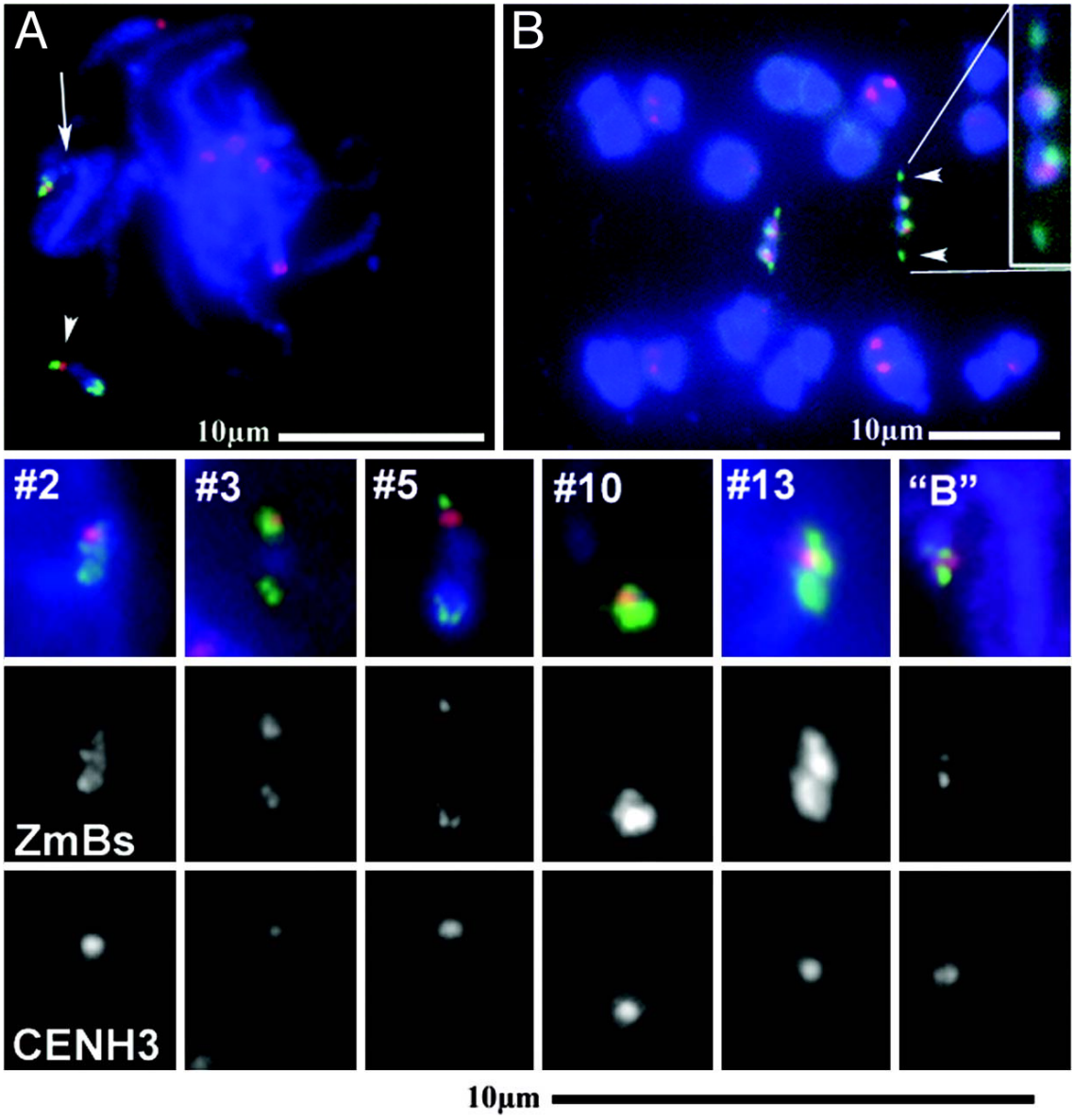
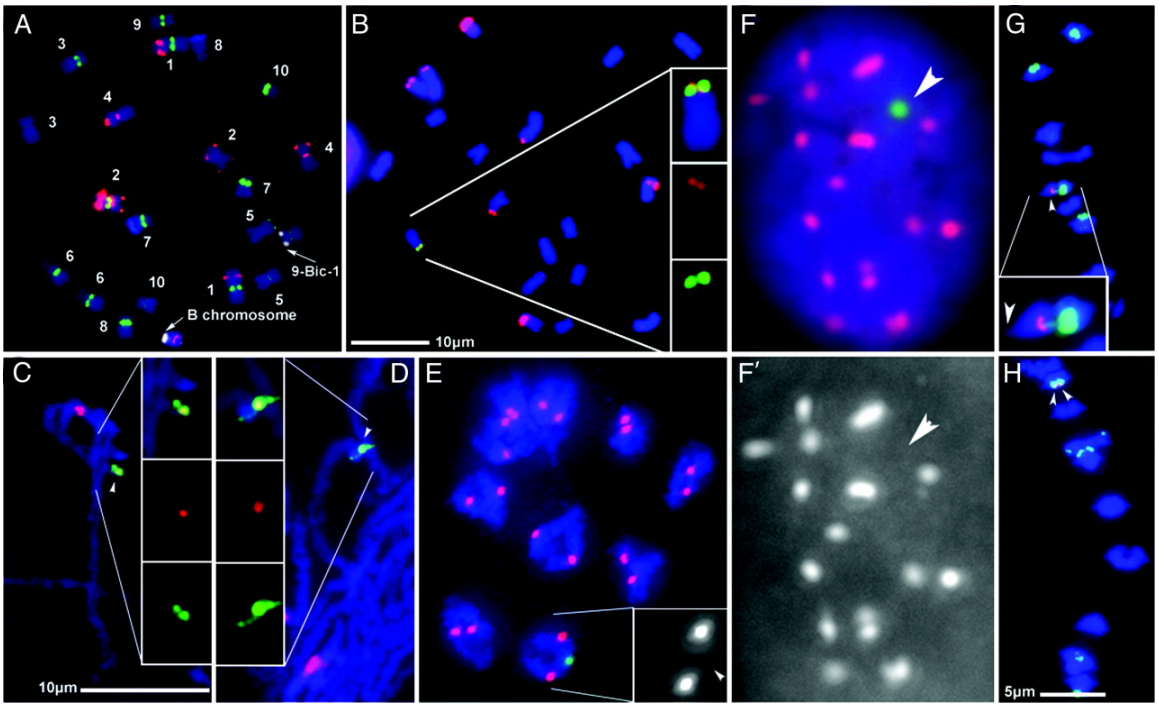
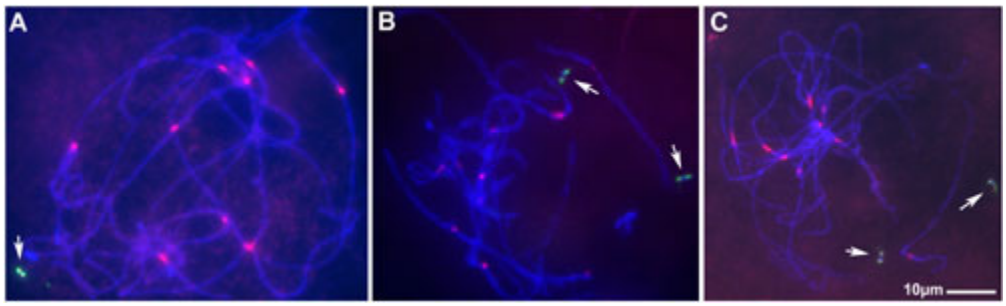


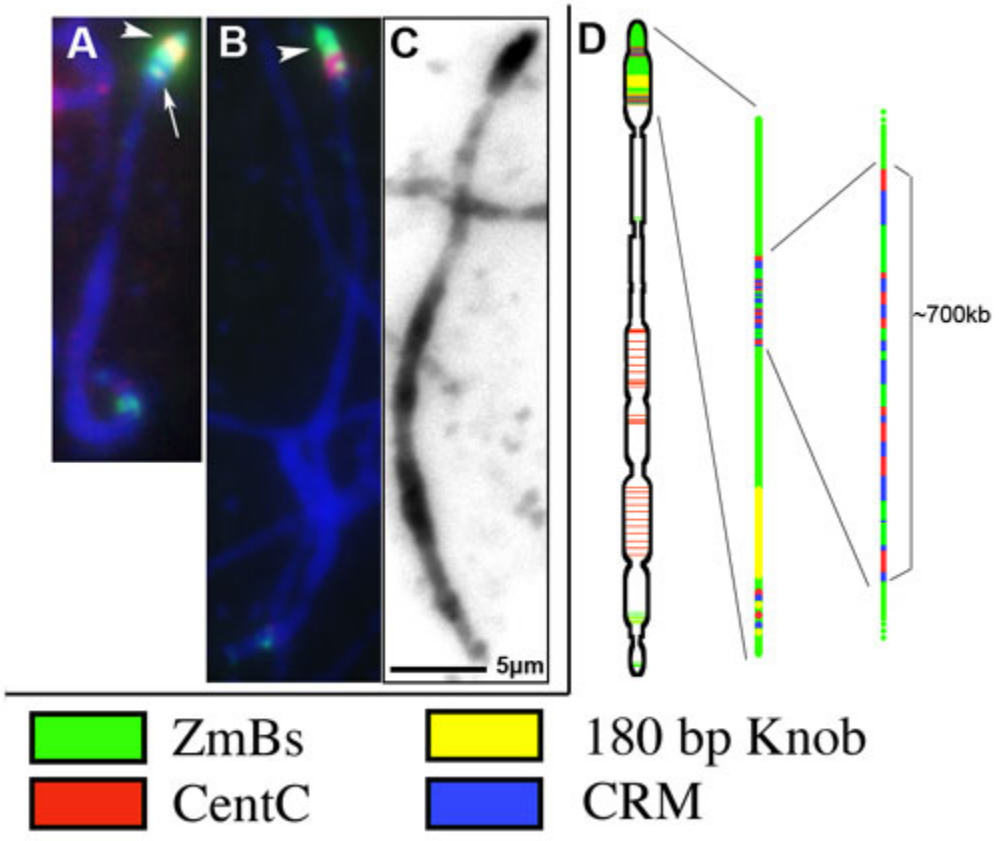
Figure 5. Cytological analysis of the translocation chromosome, 9-Bic-1.
For all figures, samples from a 9-Bic-1/+ heterozygote plant were examined. (A) Somatic chromosome spreads containing an intact B chromosome hybridized with ZmBs (white) and a cocktail of probes that allow the chromosomes to be identified. The probes include CentC (green), "TAG" microsatellite (red), 180bp knob repeat (blue), NOR (green), 5S rRNA (yellow), subtelomeric repeat 4-12-1 (green), and a 1.1 subtelomeric repeat (red). The chromosomes were counterstained with DAPI (blue). The arrows designate the B centromeric regions. (B) Knob signal (red) is observed adjacent to the ZmBs signal (green). (C) Pachytene chromosomes hybridized with ZmBs (green) and CRM (red). (D) Pachytene chromosomes hybridized with ZmBs (green) and CentC (red). (E) Diakinesis chromosomes with immunolabeled CENH3 (red) hybridized with ZmBs (green). The arrowhead in the inset indicates the location of the ZmBs signal in an enlarged image showing only the CENH3 signal. (F) A mitotic interphase nucleus with immunolabeled CENH3 (red) hybridized with ZmBs (green). The CENH3 signal alone is shown in (F') with an arrowhead indicating the location of ZmBs. (G) and (H) show anaphase chromosomes hybridized with ZmBs (green) and the 180bp knob repeat (red). Arrowheads indicate the ZmBs signal. The ZmBs repeat is not leading to the poles. The arrowhead in the inset indicates the portion of chromosome 9 that is leading. In (G) the presence of ZmBs signal on both separating chromosomes results from a crossover. The scale for (A) is the same for (B); (C), (D), and (E) are the same scale; (G) and (H) are the same scale.



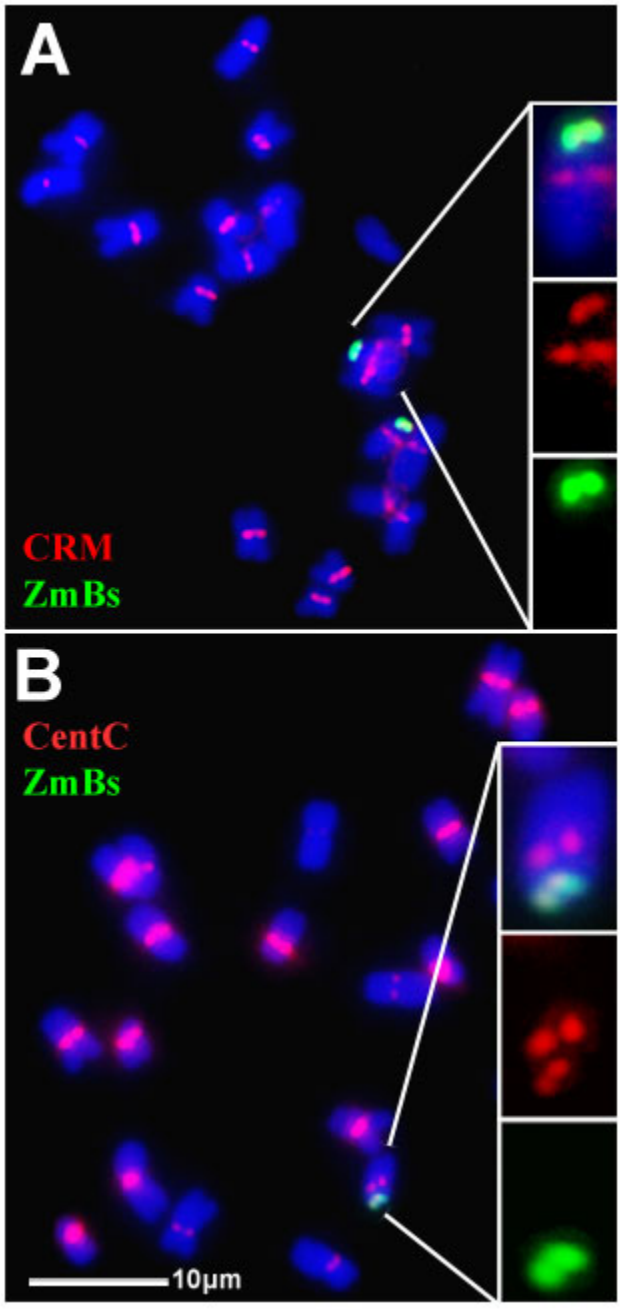
Supplementary Figure 1. Pachytene FISH of minichromosomes 2, 3, and 13. Pachytene spreads containing minichromosome 2 (*A*), 3 (*B*), and 13 (*C*) were hybridized with ZmBs (green) and CRM (red). Arrowheads indicate the minichromosomes.



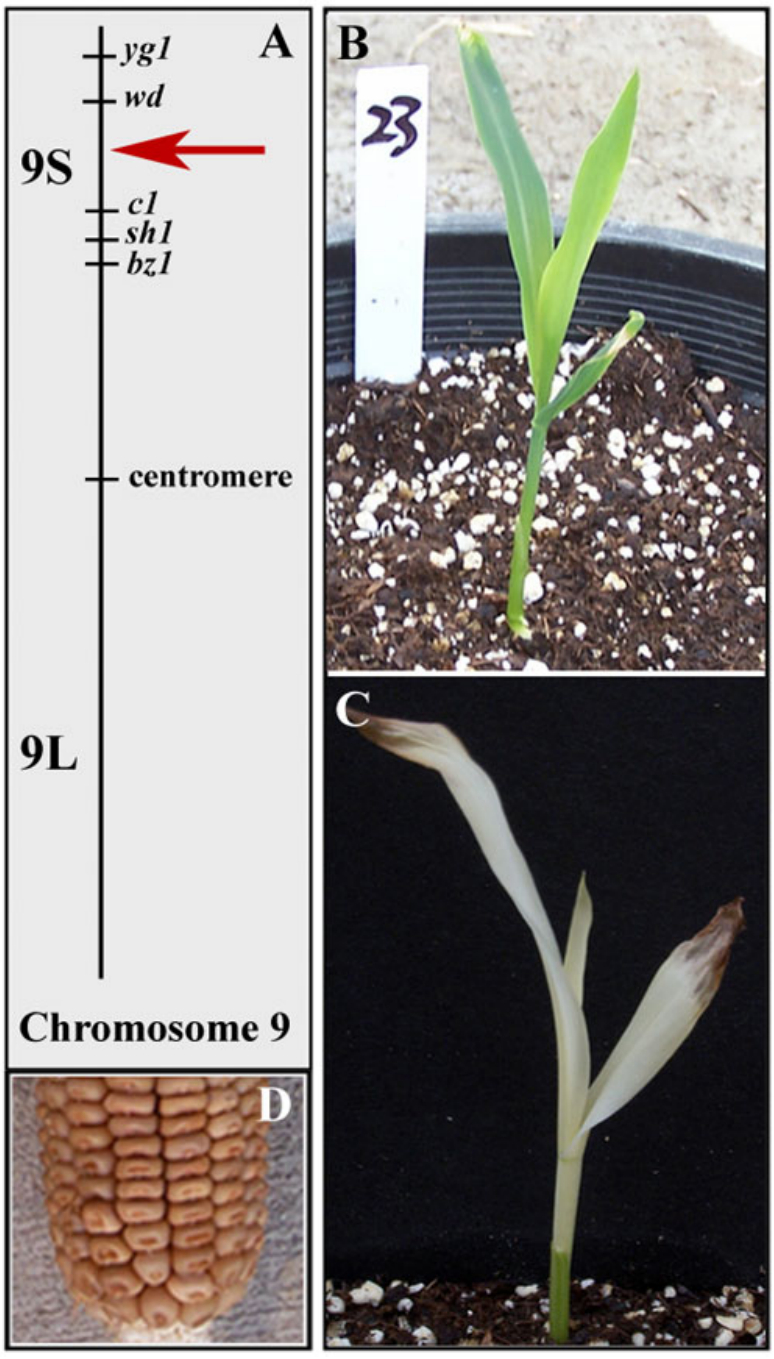
Supplementary Figure 2. Arrangement of centromeric elements in a normal B chromosome.
(*A*) A pachytene maize B chromosome hybridized with ZmBs (green) and CRM (red). The strongest CRM signal is located in an ≈700kb "core" region between two blocks of ZmBs signal (indicated with an arrowhead). The core contains the site of CenH3 localization. There is also a weaker CRM signal coincident with a third block of ZmBs (indicated by an arrow). (*B*) A pachytene maize B chromosome hybridized with ZmBs (green) and the 180-bp knob repeat (red). The decrease in ZmBs hybridization intensity at the core region is visible (indicated by an arrowhead). The knob signal is located between the second large ZmBs repeat region and the third smaller block of ZmBs. The knob element CentC and CRM are interspersed throughout the third block of ZmBs. (*C*) A pachytene maize chromosome labeled with DAPI. The color is inverted to more clearly illustrate the cytological features of the B chromosome. The scale bar applies to *A*–*C*. Note that the B chromosome in *A* is more condensed than those in *B* and *C*. (*D*) The B chromosome is represented as a cartoon, showing the locations of CentC (red), ZmBs (green), CRM (blue), and the 180-bp knob repeat (yellow). The region containing the centromere and the proximal heterochromatin is expanded to the right (not shown to scale). The ≈700-kb core domain is further expanded to the far right.



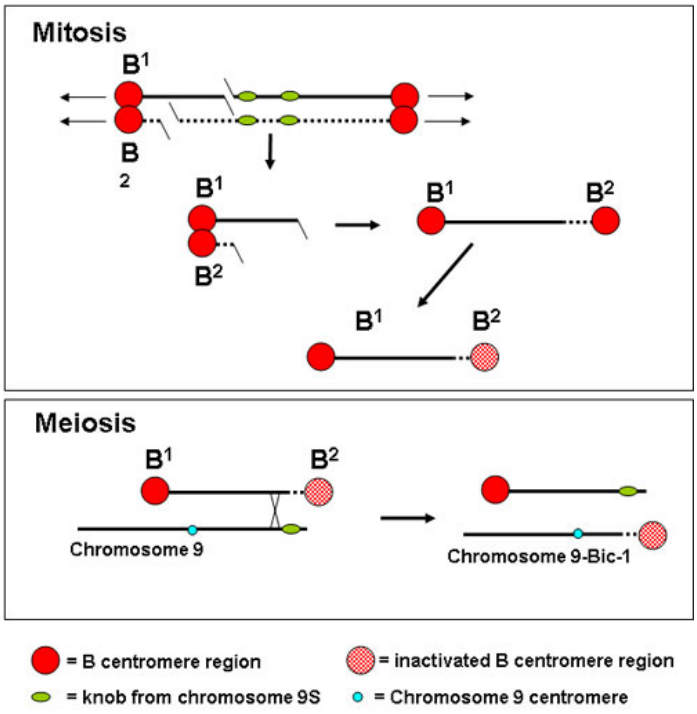
Supplementary Figure 3. Cytological analysis of 9-Bic-1.
Somatic chromosome spreads containing the 9-Bic-1 chromosome were hybridized with ZmBs (green) and CRM (red) (*A*) or CentC (red) (*B*).



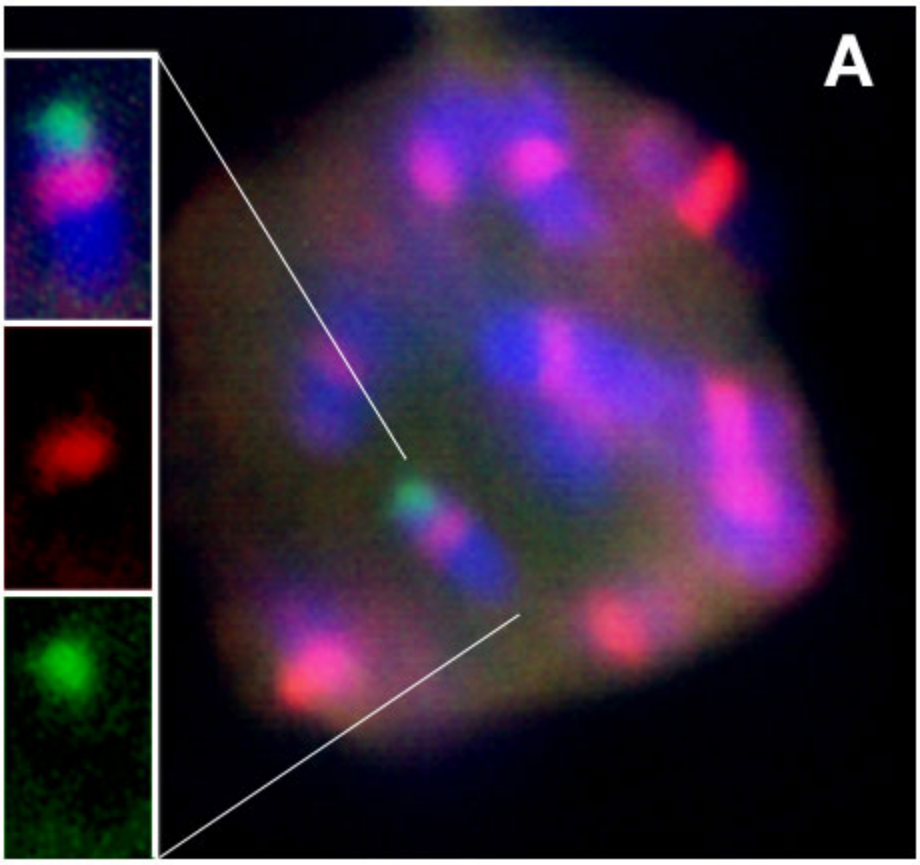
Supplementary Figure 4. The breakpoint of 9-Bic-1 is on 9S.
(*A*) The arrow indicates the approximate breakpoint on 9S. (*B*) Seedling of a 9-Bic-1/+ heterozygote. (*C*) Seedling of a 9-Bic-1/9-Bic-1 homozygote, showing the *wd* phenotype. (*D*) An ear of a self-pollinated 9-Bic-1/+ heterozygote.



Supplementary Figure 5. Formation of the 9-Bic-1 chromosome.
The presence of inactivated B centromeres in the minichromosomes that appear intact at the cytological level provides a possible explanation for the origin of the translocation chromosome 9-Bic-1, which contains an active centromere from chromosome 9 and an inactive B centromere. Because the original B9-Dp9 contained the dominant *Bz1* and was crossed to a *yg2*, *bz1* tester, the event leading to 9-Bic-1 must have placed the B centromere on the *bz1* tester chromosome 9. One explanation for how this could occur follows. Breaks occur at different locations in the two chromatid bridges that form in the chromosome type breakage–fusion–bridge (BFB) cycle, forming an asymmetric dicentric chromosome that is stabilized by inactivation of the B centromere region. During meiosis, the stabilized dicentric chromosome pairs with chromosome 9, and recombination places the inactivated B centromeric region onto the end of 9S, removing the *wd* and *yg2* genes and the distal knob of 9S.



Supplementary Figure 6. The H3 histone protein of the inactive B centromeric region of 9-Bic-1 is not phosphorylated at serine 10 in mitotic metaphase.
Mitotic chromosome spreads were labeled with antibodies against pSer10-H3 histone proteins (red) and hybridized with ZmBs (green).



6. RETROELEMENT GENOME PAINTING: CYTOLOGICAL VISUALIZATION OF RETROELEMENT EXPANSIONS IN THE GENERA *ZEA* AND *TRIPSACUM*

Centromeric elements in many species are rapidly evolving. I was interested in cytologically comparing the centromere regions of maize to a number of its closer relatives including other species in the same genus as well as the sister genus, *Tripsacum*. In particular, I was interested to find out if CentC was present and if the amount varied among centromeres to the same extent as observed in maize. It would be ideal if chromosomes from different species could be placed together on a single preparation allowing a direct comparison. Hybrids between maize and its relatives were available but it was not possible to distinguish the parent of origin for each chromosome in most cases. Therefore, I developed the technique described in this article.

Although the focus of this article is the technique of using abundant retroelements to differentially paint genomes, there are several observations that are relevant to centromere studies. Briefly, the amount of CentC does not vary to the same degree in either *Z. diploperennis* or *T. dactyloides* as it does in maize. The amount of CRM is greater in maize than either of the two relatives. Also, while retroelements' copy number varies greatly among the species, CentC is present in all, suggesting that it is maintained to a greater degree than other repetitive elements in the genome. In a naturally occurring allopolyploid containing chromosomes from *Zea* and *Tripsacum*, the amount of CRM is greater in the *Zea* chromosomes showing that homogenization for centromeres is not rapid. Finally, comparing GISH images using *Tripsacum* genomic DNA as FISH probe, reveals divergence between *Tripsacum* and *Zea* centromeres, either by the presence of a *Tripsacum*-specific element or variation in the sequence of shared centromeric elements.

LAMB, J. C., and J. A. BIRCHLER, 2006 Retroelement genome painting: cytological visualization of retroelement expansions in the genera *Zea* and *Tripsacum*. Genetics **173:** 1007-1021.

Retroelement Genome Painting: Cytological visualization of retroelement expansions in the genera *Zea* and *Tripsacum*

Jonathan C. Lamb and James A. Birchler[1]

Division of Biological Sciences, University of Missouri-Columbia

117 Tucker Hall, Columbia, MO 65211-7400, U S A

Sequence data from this article have been deposited with the EMBL/GenBank Data Libraries under accession nos. DQ223960-DQ223966.

To whom correspondence should be addressed.

E-mail: BirchlerJ@missouri.edu

117 Tucker Hall, Columbia, MO 65211-7400

Tel : (573) 882-4905 Fax : (573) 882-0123

Abstract

Divergence of abundant genomic elements among the *Zea* and *Tripsacum* genera was examined cytologically and a toolkit established for subsequent studies. The LTR regions from the CRM, Huck, Grande, Prem1, Prem2/Ji, Opie, Cinful-1, and Tekay retroelement families were used as FISH probes on mitotic chromosome spreads from a "tri-species" hybrid containing chromosomes from each of three species: *Zea mays* (2n=20), *Z. diploperennis* (2n=20), and *Tripsacum dactyloides* (2n=36). Except for Tekay, which painted both *Zea* and *Tripsacum* chromosomes with nearly equal intensity, the retroelement probes hybridized strongly to the *Zea* chromosomes allowing them to be distinguished from those of *Tripsacum*. Huck and Grande hybridized more intensely to maize than to *Z. diploperennis* chromosomes. *Tripsacum* genomic clones containing

retroelement sequences were isolated that specifically paint *Tripsacum* chromosomes. The retroelement paints proved effective for distinguishing different genomes in interspecific hybrids and visualizing alien chromatin from *T. dactyloides* introgressed into maize lines. Other FISH probes (180-bp knob, TR-1, 5S, NOR, Cent4, and CentC, rp1, rp3, and α-ZeinA) could be simultaneously visualized with the retroelement probes, emphasizing the value of the retroelement probes for cytogenetic studies of *Zea* and *Tripsacum*.

Introduction

The *Tripsacum* and *Zea* genera are sister taxa that share many morphological features as well as common elements of their genomes (DENNIS and PEACOCK 1984; GALINAT 1988; MANGELSDORF 1947; MEYERS *et al.* 2001). Their genomes are also similar in their gene density; both contain abundant LTR-type retroelements in contrast to a slightly more distantly related genus, Sorghum, which has a more compact, gene rich, genome (GAUT *et al.* 2000).

Except for one autotetraploid species, *Zea* species are diploids with a basal set of 10 chromosomes derived from an ancient polyploidization event. *Tripsacum* species represent a range of ploidies with total chromosome numbers in multiples of 18. The basal set of 18 chromosomes likely represents a recent polyploidization event (GAUT *et al.* 2000). The species in these genera constitute an excellent model of the types of changes that occur during the evolutionary divergence of species including chromosomal rearrangements, polyploidization events and divergence of repetitive elements such as retrotransposons (GAUT *et al.* 2000). A comparison of the retroelement distribution in

these related species would increase understanding of the mechanisms that shape plant evolution and the processes of retrotransposon expansion.

In addition to providing an evolutionary model of genome evolution, *Zea* and *Tripsacum* species have agronomic value. *T. dactyloides* is a forage crop species and can be hybridized with maize to produce partially fertile offspring (MANGELSDORF and REEVES 1931). Thus, *Tripsacum* species might serve as a source of germplasm for maize crop improvement and vice versa. Understanding how these genera are related would facilitate such reciprocal utilization of tools and traits. Cytological approaches could play a role in determining their relationship by visualizing the changes in genomic elements and their physical arrangement among the *Zea* and *Tripsacum* species and allow visualization of gene flow resulting from breeding programs.

Improvements in chromosome spreading procedures, development of a collection of DNA elements that can be used as Fluorescence In Situ Hybridization (FISH) probes to identify each chromosome (KATO *et al.* 2004) and detection of specific genetic loci on chromosomes (ANDERSON *et al.* 2004; KATO *et al.* 2005; KOUMBARIS and BASS 2003) have expanded the number of cytological tools available to maize researchers.

This report describes the application of karyotypic methods for the study of the evolution of abundant elements in the *Zea* and *Tripsacum* genomes, the extension of maize tools to *Tripsacum* and the development of additional techniques to create a toolkit for use in subsequent studies of these species. First, specific sequences from retrotransposons, including previously characterized maize retroelements and *Tripsacum* retroelement sequences isolated in this study, were used as FISH probes onto chromosomes from a "tri-species" hybrid containing chromosomes from *Z. mays*, *Z.*

diploperennis, and *T. dactyloides* to examine the distribution of retrotransposons among *Zea* and *Tripsacum*. The use of the "tri-species" hybrid allowed the relative intensities of hybridization of the retroelements to different genomes to be determined. Because many of the retroelement probes hybridized more intensely to one genome than another, they were used as genome paints and their effectiveness compared to Genomic In Situ Hybridization (GISH). Examination of GISH images revealed variation of centromeric elements and heterochromatic knob elements among *Zea* and *Tripsacum*.

Second, the retroelement genome paints were demonstrated to be a useful addition to a cytological tool kit for *Zea* and *Tripsacum*. Several FISH probes known to label maize chromosomes were used in combination with the retroelement FISH probes to simultaneously distinguish different genomes and detect the chromosomal location of specific DNA elements in interspecific hybrids. Additionally, retroelement genome paints were used to demonstrate introgression of *Tripsacum* chromatin into maize lines. Applications of these techniques to basic questions of genome evolution and organization as well as practical uses in breeding programs are discussed.

Materials and Methods

Plant materials

By crossing a maize (2n=20) X *Tripsacum dactyloides* (2n=36) hybrid that had been treated with colchicine to induce chromosome doubling with pollen from *Z. diploperennis* (2n=20), Dr. Walton Galinat produced an individual containing genomes of all three species (personal communication). This plant has been maintained vegetatively for many years, first by Dr. Galinat and after his retirement, by Dr. Mary Eubanks of

Duke University who provided a cutting of this "tri-species" hybrid for study. This plant provides an opportunity to examine chromosomes from three evolutionary branches at one time, facilitating direct comparisons of their genomes cytologically.

Maize X *Tripsacum* hybrids were produced using tetraploid *Tripsacum dactyloides* pollen for silks of a maize popcorn variety (Super Gold). Maize x *Z. diploperennis* hybrids (maize variety NC300 x *Z. diploperennis*, PI 441932) were provided as kernels by Jim Holland (North Carolina State University). *T. andersonii* was provided by Denise Costich (Cornell University).

The Maize Genetics Cooperation Stock Center (MGCSC) provided four lines derived from addition lines produced by Walton Galinat. Two did not have any detectable alien chromatin (Stock center field numbers 2000-9-1@ and 2000-8B-2@) and two did (described in this report, stock center field numbers 2000-7-5@ and 2000-W23-24/8A-2). The 2-T, T-2 translocation line produced by Marjorie Maguire was also obtained from the MGCSC (Stock record 220B).

Genomic DNA for library construction and Southern blotting was collected from leaf tissue. Maize genomic DNA was from the inbred line, W22. Genomic DNA from *T. dactyloides* was collected from variety "Pete". Leaf tissue from *T. laxum* and *Z. diploperennis* for genomic DNA isolation was provided by Mary Eubanks. *Z. luxurians* was grown from kernels provided by Mary Eubanks (PI# 306615). *S. bicolor* seed (variety Sudan) was grown for leaf tissue.

Genomic library production and screening

A Fosmid library made from *T. dactyloides* genomic DNA was constructed using a production kit from Epicentre according to the manufacturer's directions (Cat.

#CCFOS110). 96 colonies were hand picked and used to inoculate a 96-well square bottom plate. Plasmids were isolated using a 96-well plasmid isolation kit.

Small-insert libraries were made from nebulized genomic DNA from *T. dactyloides* and *Z. diploperennis*. The DNA was nebulized according to the manufacturer's directions (Invitrogen, Cat. #45-0072). Nebulized DNA was subjected to electrophoresis and DNA that was 1200 to 3000 bp in length was extracted. The extracted DNA was treated with Epicentre's End-It enzyme mixture (Cat. #ER0720) to produce blunt ends. The blunt end DNA was then incubated with Taq polymerase and dATP to add an "A" overhang and then cloned into Promega's pGemT TA-cloning vector (Cat. #A3600). 96 colonies were grown and plasmids isolated using a 96-well plasmid isolation kit.

Dot blots were prepared using the Bio-Dot apparatus (BioRad, Cat. #M1706545) according to the manufacturer's instructions. Maize Cot-100 DNA was prepared as described (ZWICK *et al.* 1997). Autoclaved genomic DNA (either maize, *T. dactyloides* or *Z. diploperennis*) and maize Cot-100 DNA was labeled with 32P dCTP using a random decamer priming reaction (Ambion's DECAprime II, Cat. #1456) and hybridized overnight to the dot blots in hybridization solution (Sigma, Cat. #H-7033) at 65º. Following a quick rinse in 0.2X SSC, 0.1X SDS, the blots were washed two times at 65º for 30 min each. Blots were then labeled with a refillable ink pen spiked with 32P dCTP and exposed to a phosphor-imaging plate for ~6 hours. Phosphor-imaging plates were analyzed using a FUJIFILM fluorescent image analyzer (FLA-2000) FUJI Image Reader V1.5E and processed using FUJI Image Gauge V3.45.

Tripsacum genomic clones TC#5 and TC#25 were labeled with P32 dCTP and probed onto a Southern blot prepared with genomic DNA from *S. bicolor, T. dactyloides, T. laxum, Z. mays, Z. diploperennis,* and *Z. luxurians* cleaved with *Hin*dIII. Hybridization was at 65º overnight followed by two washes thirty minutes each with 0.2 X SSC, 0.1% SDS at 65º and visualized by exposure to a phosphor-imaging plate for four to six hours.

Subclones from fosmids that hybridized in a *Tripsacum*-specific pattern were produced by digestion of fosmid DNA with *Hin*dIII or *Sac*I and ligated into a pBluescript vector that had been digested with the appropriate restriction enzyme. Plasmids were isolated from several colonies and used as templates for a PCR reaction using T3 and T7 primers. PCR products were used as templates for producing fluorescent probes, which were tested by FISH for specific hybridization to *Tripsacum*.

All plasmids from the small-insert library and fosmid subclones showing a *Tripsacum*-specific hybridization pattern were sequenced using M13F, M13R, T3, or T7 primers. Fosmids were end sequenced using the T7 primer. Sequences were compared against data in GenBank using the BlastN and/or BlastX algorithm (ALTSCHUL *et al.* 1990). The sequences of the elements that are enriched in *Tripsacum* chromatin have been deposited in GenBank (TF-B5-2, DQ223961; TF-B5-3, DQ223960; TF-B8-15, DQ223962; TC#5, DQ223963; TC#12, DQ223964; TC#24, DQ223965, TC#25, DQ223966).

Chromosome spread preparation and FISH

Chromosome preparations were prepared as described previously (KATO *et al.* 2004) with modifications described below. Root tips from maize x *Tripsacum* hybrids, maize x *Z. diploperennis* and from the "tri-species" hybrid were collected from adult

plants by choosing young rapidly growing root tips. Root tips with many cells in mitosis were obtained from plants that had recently been propagated (1-3 weeks) or from older plants whose roots were trimmed, watered, and allowed to re-grow for 3-6 days without watering. Root tips from other material were collected from seeds that had been germinated in moist vermiculite for ~3 days at 30º.

The maize retroelement probes Grande, Huck, Prem2/Ji, Tekay, Cinful, and Opie were cloned as described (MROCZEK and DAWE 2003) except a Promega pGemT TA-cloning kit was used. After sequencing to confirm the identity of the inserts, plasmids were used as templates for PCR reactions to produce DNA for fluorescent labeling. Prem1 was amplified from a clone provided by Dr. Jeffrey Bennetzen. DNA from the *Tripsacum* small-insert library was PCR-amplified from the respective plasmid using M13F and M13R (or T3 and T7) primers as were the 180bp knob repeat, 5S gene cluster, TR-1, NOR, Cent4, CentC, *rp1*, *rp3*, and the *α-zeinA* cluster inserts from plasmids described elsewhere (KATO *et al.* 2006; KATO *et al.* 2004). Fosmid DNA was prepared using a Qiagen miniprep kit after induction to high copy number using the induction solution provided in the fosmid production kit from Epicentre (Cat. #CCFOS110). Fluorescent labeling was as described previously (KATO *et al.* 2006).

FISH was performed as described previously (KATO *et al.* 2004) except that the slides were not fixed with formaldehyde after UV-crosslinking. For GISH, the following amounts of probe were included in a total volume of about 7 ul: FITC-labeled genomic DNA from *Tripsacum* - 50 ng, autoclaved unlabeled maize DNA - 2.5 µg (50X) or 5 µg (100X), Grande - 100 ng, Huck - 100 ng, Prem1 - 100 ng, and 180 bp knob repeat - 40

ng. For other FISH experiments, 40 to 200 ng of each probe was used per slide in a total volume of 5 µl per slide.

Images were taken either by an Optronics MagnaFIRE CCD camera mounted on a Zeiss Universal microscope using a 100× plan apo oil immersion lenses (for most images), or by a Sensys CCD camera (for the *rp1, rp3,* and *α-zeinA* pictures) using a 60× plan apo oil immersion lens on an Olympus IX70 inverted microscope controlled by Metamorph software. Images were processed as described previously (KATO *et al.* 2004; LAMB *et al.* 2005) with the following addition. For some images where light intensity was not uniform, the image was severely blurred using the "Gaussian blur" function of Photoshop 7.0 and the blurred image subtracted from the original to reduce the background.

Results

Application of abundant maize retroelement LTRs as FISH probes

To examine the relative abundance of different retroelement families in the three species that compose the "tri-species" hybrid, the LTR regions from seven families of abundant maize retroelements (Huck, Grande, Prem1, Prem2/Ji, Opie, Cinful-1, and Tekay) were amplified by PCR and used as FISH probes on mitotic chromosome spreads from the "tri-species" hybrid. For Tekay, all 37 chromosomes were strongly labeled, although Tekay hybridization appeared slightly less intense on the *Tripsacum* chromosomes (Figure 1). Huck, Grande, Prem1, Prem2/Ji, Opie, and Cinful-1 labeled 19 chromosomes and 18 chromosomes were unlabeled or relatively weakly labeled (Figure 1). Because the number of unlabeled chromosomes equals the haploid number for

Tripsacum (n=18) and many of the unlabeled chromosomes were very small, it was surmised that the retroelement probes paint *Zea* but not *Tripsacum* chromosomes and that one of the *Zea* chromosomes is missing in this plant. The missing chromosome was subsequently determined to be maize chromosome 2 (see below).

To confirm that the Huck, Grande, Prem1, Prem2/Ji, Opie, and Cinful-1 probes paint maize but not *Tripsacum* chromosomes, the retroelement probes were applied to chromosomes from a maize X *T. dactyloides* (2n=72) hybrid. Tekay hybridized to all 46 chromosomes (10 from maize, 36 from *T. dactyloides*) and the remaining probes labeled only 10 chromosomes. Opie and Huck hybridization is shown in Supplementary Figure 1.

In the "tri-species" hybrid, several of the retroelement probes labeled 9 chromosomes most intensely. To determine which of the *Zea* genomes is more strongly labeled and to confirm the pattern seen in the "tri-species" hybrid, these elements were also probed onto chromosomes from a *Z. mays* X *Z. diploperennis* F1 hybrid. Small heterochromatic blocks, called knobs, are present near the termini of chromosomes in *Z. diploperennis* whereas they tend to be larger and positioned interstitially in chromosomes of maize (KATO Y 1984; MCCLINTOCK *et al.* 1981). The locations of the knobs in the *Z. mays* x *Z. diploperennis* hybrid were determined by hybridization of the 180bp knob repeat together with simultaneous hybridization of the Huck and Opie retroelements, each probe labeled with a different fluorescent molecule. The chromosomes with large internal knobs hybridized intensely with the Huck and Opie probes (Figure 1A). Therefore, the 10 weakly labeled chromosomes were determined to be from *Z. diploperennis*.

T. andersonii is thought to be an allopolyploid that combines the genomes of a *Tripsacum* species and *Z. luxurians* (TALBERT *et al.* 1990). Consistent with this

hypothesis, chromosome spreads of *T. andersonii* contained 10 chromosomes that were strongly labeled by the maize retroelement probes except for Tekay which hybridized to all the chromosomes. The remaining chromosomes were labeled by the *Tripsacum* specific clones described below (Supplementary Figure 2).

Identification of *Tripsacum* clones (fosmids and small-inserts) that paint *Tripsacum* chromosomes

The distribution of maize retroelements in the "tri-species" hybrid indicates that many of the maize retroelement families have expanded after the divergence of the *Zea* and *Tripsacum* lineages. However, the *Tripsacum* genome size is comparable to maize (GAUT et al. 2000) suggesting that other retroelement families have expanded in the *Tripsacum* genome. This hypothesis would be supported by the discovery of retroelements that paint *Tripsacum* chromosomes but not *Zea* in the "tri-species" hybrid. Additionally, *Tripsacum* specific retroelement probes would complement the maize retroelements as useful tools to distinguish the chromatin of different species in interspecific hybrids.

To obtain probes that would label *Tripsacum* chromatin, small libraries of *Tripsacum* genomic DNA were screened for clones, which could be used as FISH probes to distinguish *Tripsacum* chromosomes in a similar fashion as the maize retroelements. A fosmid library was generated and DNA from 96 clones was spotted in triplicate onto nylon membranes. The membranes were then probed with radio-labeled maize DNA, Cot-100 maize DNA, or genomic DNA from *T. dactyloides*. The dot blots were then examined for clones that were labeled by the *Tripsacum* genomic DNA but not by maize DNA. This approach allowed many clones to be screened at one time.

Three fosmids, TFF-B5, TF-B8 and TF-F7, that were labeled on the dot blot more intensely by *Tripsacum* than by maize were used as red FISH probes simultaneously with the Huck probe labeled in green on a "tri-species" hybrid chromosome spread. Two of these fosmids, TF-B5 and TF-B8, hybridized intensely to *Tripsacum* but not to *Zea* chromosomes. TF-F7 hybridized to the 5S ribosomal gene clusters (Figure 2 and Supplementary Figure 3). A number of clones, TF-A4, TF-C12, TF-G6, and TF-H7, that had little or no signal on the dot blot when hybridized with either the maize or *Tripsacum* radio-labeled DNA were tested as FISH probes onto the "tri-species" hybrid. All of these probes gave weaker FISH signals than the fosmids that had strong dot blot signals and required longer exposure times to clearly visualize the fluorescent signals. Two fosmids, TF-C12 and TF-G6, hybridized more intensely to the *Tripsacum* chromosomes giving a signal that allowed the *Tripsacum* chromosomes to be clearly distinguished from the *Zea* ones. Fosmid clone TF-H7 specifically labeled the *Tripsacum* chromosomes but it hybridized in a punctate pattern instead of the more uniform hybridization seen for retroelement probes. Fosmid TF-A4 labeled the *Tripsacum* chromosomes slightly more intensely than the *Zea* ones. Several of the dot blot signals resulting from hybridization with maize genomic DNA were stronger than those with *Tripsacum* DNA. When fosmids of this category, TF-B6 and TF-E12, were used as FISH probes, they strongly labeled the *Zea* chromosomes (Figure 2 and Supplementary Figure 3).

To determine the sequences responsible for painting *Tripsacum* chromosomes, fosmids TF-B5 and TF-B8 (that had a *Tripsacum* specific pattern of hybridization) were subcloned and several resulting subclones tested for *Tripsacum* specific hybridization on "tri-species" hybrid chromosome spreads. Subclones TF-B5-2, TF-B5-3 and TF-B8-15

hybridized strongly to *Tripsacum* chromosomes but not those of maize or *Z. diploperennis*. The subclones were sequenced and compared to the GenBank database using BlastN (ALTSCHUL *et al.* 1990). TF-B5-3 was 1071 bp long and was homologous (>86% identity) to a retroelement sequence found in the *rp3* rust resistance gene cluster (gi:45934294). TF-B5-2 contained 2320 bp of non vector sequence including all of TF-B5-3. TF-B8-15 was 4093 bp in length and contained a 1075 bp region with 86% identity to TF-B5-3. This region also contains a 46 bp stretch that is homologous to a large number of rice genomic targets including some putative gag-pol genes (~89% identity). These 46 bps are contained in a 130 bp stretch that is 83% identical to a *Setaria italice* transposable element. In addition to the region shared between the subclones, TF-B8-15 contained a second region homologous to a maize retroelement (>86% identity for 1438bp to gi:46200524), and a region with homology to multiple rice genomic regions).

Additional *Tripsacum* specific clones were obtained from a small-insert library made from genomic DNA sheared to approximately 1.5 kb. Clones were dotted to duplicate nylon membranes and probed with P32 radio-labeled genomic DNA. Twenty-three clones that were strongly labeled on the dot blot (Supplementary Figure 4) were amplified via PCR to generate template DNA for fluorescent labeling. Many of the most intensely hybridizing spots produced multiple bands in a laddered pattern when subjected to gel electrophoresis suggesting they contained ribosomal DNA, knob sequences or other tandemly arranged elements (data not shown). These products were not used and the remaining PCR products were labeled and screened by hybridization to "tri-species" hybrid chromosome spreads allowing identification of four additional clones (TC#5, 12, 24, and 25) that hybridized strongly to *Tripsacum* chromosomes and weak to moderately

to *Zea* chromosomes. The remaining clones hybridized with slightly greater intensity to *Tripsacum* chromosomes than to *Zea* (TC#3, 13, 17), to specific chromosomal regions such as the NOR (TC#11, 19) or knobs (TC#4, 7, 14), or did not hybridize at all (Figure 3; Supplementary Figure 5).

Two of the *Tripsacum* specific clones (TC#5, 25) were labeled with 32P and hybridized to Southern blots of restriction enzyme digested genomic DNA from *Zea*, *Tripsacum* and Sorghum species. The probes hybridized in a smear along the genomic DNA and strongly labeled DNA from *T. dactyloides* and *T. laxum*, weakly labeled DNA from the *Zea* species and did not label DNA isolated from *S. bicolor* (Figure 3).

The clones that strongly labeled the *Tripsacum* chromosomes were sequenced and the results compared against the GenBank database using BlastN and/or BlastX. All of the clones were similar to many maize genomic sequences including sequences annotated as retrotransposons. TC#24 and TC#25 shared several blocks of highly similar sequence (92% identity for 165 bp, 81% for 215 bp, 79% for 126 bp, 96% for 30bp, 100% for 23 bp). The sequences identified in the current study that are highly abundant in *Tripsacum* appear to represent four different retroelements (1. TF-B5-3 and TF-B8-15; 2. TC#24 and TC#25; 3. TC#5; and 4. TC#12).

A small-insert library was also prepared from *Z. diploperennis* and dotted in the same manner as the *Tripsacum* library. The membranes were hybridized to maize genomic DNA, stripped, and hybridized to DNA from *Z. diploperennis*. Two clones were selected, one that hybridized more intensely to maize DNA (ZD#9) and one that hybridized equally (ZD#10). These clones were hybridized to the "tri-species" hybrid chromosome spreads. They labeled the chromosomes in a pattern consistent with the dot

blot hybridization; the *Tripsacum* chromosomes were only weakly labeled (Supplementary Figure 5). The clones were sequenced and both were found to be homologous to many maize sequences. ZD#9 was homologous to several sequences annotated as Huck elements (e.g. 94% identity to gi:67043718). ZD#10 was homologous to intergenic sequences from a variety of clones including centromeric BAC 15C5 (gi: 37514986) (NAGAKI *et al*. 2003). The regions of homology to BAC 15C5 included shadowspawn retroelement fragments (basepairs 43,649-43,409, 86% identity; 54,757-54,517, 86% identity; 61400-61082, 89% identity) and a region putatively identified as degenerate retroelement sequence (basepairs 5,587-5,887, 83% identity; 7,071-7,404, 87% identity). The latter region has previously been subcloned and used as a FISH probe on maize chromosomes giving a dispersed hybridization pattern similar to that of ZD#10 (NAGAKI *et al*. 2003).

In summary, many retroelements DNA sequences were assayed for their relative intensity of hybridization as FISH probes onto the "tri-species" hybrid. Sequences generally hybridized to the chromosomes in the following patterns: 1, strong on all three species; 2, strong and equal on *Z. mays* and *Z. diploperennis* but weak or not at all to *Tripsacum*; 3, strong on *Z. mays*, intermediate on *Z. diploperennis*, weak on *Tripsacum*; and 4, weak or intermediate on both *Zea* species and strong on *Tripsacum*. However, there were subtle differences in hybridization within each of these four general categories. This information is summarized in Table 1.

Retroelement labeling combined with GISH labeling

A common approach to labeling chromatin from one species in interspecific hybrids is to use Genomic In Situ Hybridization (GISH). In examining the distribution of retroelements among species, it became apparent that abundant retroelements could be used as chromosome paints to distinguish different genomes in interspecific hybrids. To compare the two approaches, GISH was performed by allowing labeled *Tripsacum* genomic DNA to hybridize to chromosomes from the "tri-species" hybrid in the presence of 50X or 100X excess unlabeled maize DNA. This approach was successful in labeling the *Tripsacum* chromatin, but the knobs of both *Tripsacum* and maize were also intensely labeled, preventing long exposure times that would increase the contrast between *Tripsacum* and maize chromatin.

Because knobs from all species were labeled using GISH, using this technique to detect interspecific rearrangements could be complicated. Retroelement painting was therefore attempted in conjunction with GISH to increase the ability to discriminate different genomes. GISH was performed by allowing a mixture of *Tripsacum* genomic DNA labeled in green, an excess of the maize retroelements Grande, Huck, and Prem1 labeled in red and the 180bp knob repeat labeled in red to hybridize to chromosome spreads in the presence of unlabeled maize DNA (Figure 4). This procedure painted the *Zea* chromosomes in one color, red, and the *Tripsacum* chromosomes in another, green, while the knobs are labeled with both colors.

GISH reveals divergence of repetitive elements among *Zea* and *Tripsacum*

GISH on the "tri-species" hybrid was undertaken to compare and evaluate the effectiveness of the retroelement painting approach to discern chromosomes from

different species. However, observation of the GISH images revealed divergence of a number of genomic components between the *Zea* and *Tripsacum* genera. While knobs of both *Tripsacum* and *Zea* were labeled with both colors using the combination of labeled *Tripsacum* genomic DNA, maize retroelements and the maize 180bp knob repeat, the *Zea* knobs appear reddish while the *Tripsacum* knobs appear greenish (Figure 4). This result is likely due to divergence among the knob sequences in the two genera (DENNIS and PEACOCK 1984). In addition to the knobs, the labeled *Tripsacum* DNA hybridized to the centromeres of both maize and *Tripsacum* chromosomes despite inclusion of unlabeled maize DNA as a competitor. The *Tripsacum* centromeres were clearly labeled in green while the maize centromeres were only slightly labeled; the latter often required digital enhancement to detect the signal (Figure 4).

To determine which centromeric components were responsible for the *Tripsacum* GISH signal, FISH with know centromeric components was performed. When "tri-species" hybrid chromosomes were probed with the maize tandemly repeated centromere element, CentC, all centromeres were labeled (Figure 4), although some of the maize centromeres had weaker signals than others, consistent with previous reports (ANANIEV et al. 1998; KATO et al. 2004). A full length clone of a centromere-specific retrotransposon, CRM, hybridized most intensely to maize, second most to *Z. diploperennis*, and least to the Trispacum centromeres (Figure 4). To determine whether divergence of the CRM element in *Tripsacum* explained the centromere-specific GISH signal, primers that amplify several regions of the CRM element (including the LTR, gag/pol genes, and reverse transcriptase) were used to obtain DNA for generating FISH probes from *Tripsacum* genomic DNA template using a low annealing temperature. The hybridization

pattern using those probes was similar to that observed for the complete labeled maize element (data not shown). This control indicates that the difference in hybridization intensity between the three species is likely due to copy number expansion of CRM in the maize lineage.

Because specific CRM variants are known to be present in different copy numbers among maize centromeres, chromosomes from a hybrid of maize X *Z. diploperennis* were probed with the LTRs of two different CRM variants, CRM1a and CRM2a, to determine if both variants were more abundant in maize than in *Z. diploperennis* (Supplementary Figure 6). CRM2a hybridized more intensely to maize chromosomes than those of *Z. diploperennis*. CRM1a also labeled the maize centromeres with greater intensity than those of *Z. diploperennis*, but the difference was less pronounced than that of CRM2a. CRM was also tested on a maize X *T. dactyloides* F1 hybrid and in *T. andersonii*. In both species, the intensity of *Tripsacum* centromere labeling was much less than the *Zea* centromeres (Supplementary Fig. 1 and data not shown).

Use of retroelement paints and GISH to identify introgressed *Tripsacum* chromatin in maize lines

Two of the *Tripsacum* specific retroelements, subclone TF-B5-3 and TC#25 (red) and two maize retroelements, Prem2 and Huck (green), were applied to material obtained from the Maize Genetics Cooperation Stock Center that was thought to contain *Tripsacum* chromatin in a maize background. This collection consisted of four lines descended from maize-*Tripsacum* addition lines produced by Dr. Walton Galinat (GALINAT and MANGELSDORF 1966; GALINAT et al. 1963) and one line identified by Dr.

Marjorie Maguire containing a translocation between maize chromosome 2 and a *Tripsacum* chromosome (MAGUIRE 1957; MAGUIRE 1960a; MAGUIRE 1960b). GISH was also performed on these lines using labeled *Tripsacum* genomic DNA (green), several maize retroelements (red) and the 180bp knob repeat (red). This experiment allowed the utility of the newly cloned *Tripsacum* elements in identifying introgressed *Tripsacum* chromatin to be tested and compared to a GISH approach (Figure 5).

In two of the Galinat lines, no *Tripsacum* chromatin was detected by either GISH or by chromosome painting using *Tripsacum* clones, indicating that the *Tripsacum* chromosome was lost. In the other two lines, 20 chromosomes were present instead of the expected 21 for an addition line. However, a portion of one chromosome was labeled by both GISH and by the *Tripsacum* retroelement probes. The remainder of the chromosome was labeled by either the knob sequence or by abundant maize retroelement probes (Figure 5). These lines were descended from addition lines that had extra chromosomes carrying dominant phenotypic markers that complemented mutations on maize chromosome 9 and 5. The Maguire material contained 21 chromosomes as previously reported (MAGUIRE 1957; MAGUIRE 1960a; MAGUIRE 1960b). The *Tripsacum* chromatin in the two reciprocal translocation chromosomes was easily detected.

When plants containing *Tripsacum* chromatin were selfed, individuals that were homozygous for the translocation chromosomes were recovered, and grown to maturity. The homozygous replacement line derived from the Maguire material was male-sterile as previously reported (MAGUIRE 1957; MAGUIRE 1960a; MAGUIRE 1960b), but the two homozygous replacement lines from the Galinat material were both male and female-fertile.

Additional probes

A collection of repetitive sequences has been applied to maize chromosome spreads as FISH probes allowing the somatic karyotype to be determined (KATO *et al.* 2004). Probes from this collection, and other probes that hybridize to single loci were applied to spreads containing both *Zea* and *Tripsacum* chromosomes to determine if maize probes can be universally applied to these closely related species and if they would be valuable in identifying specific *Tripsacum* chromosomes. The *Zea* chromosomes could be identified in these spreads because the probes were applied simultaneously with labeled retroelements to paint the maize and *Tripsacum* genomes differentially. The *Zea* chromosomes serve as a positive control because the probes have previously been applied to maize. Repetitive elements tested include a knob repeat (TR-1), a chromosome 4 specific probe (Cent4), the 5S ribosomal DNA, the 17S ribosomal RNA (the NOR), and a microsatellite repeat (TAG^n) (Figure 6). Probes that label specific loci included the *rp1*, *rp3*, and *α-zeinA* gene clusters (Figure 7). All probes tested gave signals on the *Z. diploperennis* and *T. dactyloides* chromosomes in the "tri-species" hybrid. No maize chromosomes were labeled by the 5S ribosomal RNA probed so the missing maize chromosome was identified as chromosome 2.

Discussion

Distribution and divergence of repetitive elements in *Zea* and *Tripsacum*

Retroelement expansion among Zea and Tripsacum specie

Examination of large genomic regions containing duplicated loci from maize and Sorghum suggests that the two progenitor genomes of maize and the progenitor sorghum

genome diverged about 11.9 million years ago (SWIGONOVA *et al.* 2004). Subsequently, the two diverged maize genomes were united through an polyploidization event about 4.8 million years ago followed by extensive rearrangements, gene loss and divergence, and other modifications leading to the current diploidized maize genome (SWIGONOVA *et al.* 2004). Comparing the amount of DNA between maize and sorghum shows that the maize genome is more than twice as large as that of sorghum and the difference is thought to result from expansion of LTR type retroelement families (GAUT *et al.* 2000; SANMIGUEL and BENNETZEN 1997).

When the *adh1-F* genomic region of maize was examined, it was found to contain retroelements inserted into other retroelements in a nested fashion that accounted for over 60% of the total sequence (SANMIGUEL *et al.* 1996). Because the process of transposition ensures that both LTRs are identical upon insertion into the genome, sequence divergence of the LTRs provides estimations of the timeframe of transposition. Using LTRs as a molecular clock, the approximate ages of several nested retroelements were determined in this region. Insertion dates for the various retroelements ranged from 5 Mya to near the present, including many elements, such as Ji and Huck, that show continual insertion over this period (SANMIGUEL *et al.* 1998).

Comparison of a number of large genomic regions revealed that the regions between genes are very different for different inbred lines due largely to the presence or absence of various retroelements (BRUNNER *et al.* 2005). The retroelements found to be present at genomic locations from only one inbred tend to be more recently inserted than shared elements (non-homologous elements: mean 0.91 Mya, median 0.61; shared retroelements, mean 1.34 Mya, median 1.16) while many older elements, common or not,

were truncated in some fashion (BRUNNER *et al.* 2005). Transposition is a continual process that can be documented to 5 million years ago in maize with several different families actively expanding throughout that entire time. Unless selection pressures act to maintain an element's integrity, mutations and sequence elimination deteriorate elements after they have inserted into a genome. After enough changes have accumulated, ancient insertions become unrecognizable. In rice, the half life for a retrotransposon has been estimated to be less than 6 million years (MA *et al.* 2004).

The four retroelement sequences identified in this study that are enriched in *Tripsacum* (1. TF-B5-3 and TF-B8-15; 2. TC#24 and TC#25; 3. TC#5; and 4. TC#12) have highly similar homologs in maize. Hybridization to Southern blots of several related species demonstrated that TC#5 and TC#25 are abundant in two different *Tripsacum* species, are present in the *Zea* species and are absent or highly diverged in Sorghum (Figure 3). The two sequences isolated from *Z. diploperennis* and several of the clones containing genomic fragments from *Tripsacum* hybridized with equal or greater intensity to maize than to the genome from which they were isolated. Thus, elements expanded in the maize lineage have closely related homologs in the *Tripsacum* genome and vice-versa. The reason why some retroelements have expanded in one lineage and remained dormant in others is unknown. However, because they are abundant and have a high rate of genomic turnover, retroelements make ideal candidates for probes that will hybridize to multiple locations in a genome but not to genomes of more distantly related species. Because individual families of retroelements can rapidly expand from a few copies in a genome to thousands, species that are only separated by a few million years will share many elements but are likely to have some families that have differentially expanded.

For closely related species, measuring retroelement abundance may contribute to understanding their taxonomy. The elements Prem1, Opie, Prem2/Ji but especially Huck and Grande hybridized with greater intensity to maize than to chromosomes from *Z. diploperennis,* while Cinful and Tekay hybridized with almost equal intensity. For many of the elements whose copy number has been estimated previously [Huck, Prem2/Ji, and Opie (MEYERS *et al.* 2001)], the hybridization pattern reflects the predicted copy number. In *Z. diploperennis*, many families of retroelements are less abundant than in maize (Meyers *et al*. 2001), suggesting that the *Z. diploperennis* lineage experienced a cessation of retroelement expansion subsequent to its divergence from the other *Zea* species. Since some retroelement families, including Tekay and TC#13, which both predate the split between *Zea* and *Tripsacum*, are present in nearly equal numbers among the *Zea*s, the presence of fewer copies of many retroelement families is probably not due to widespread retroelement elimination. Also, a study on the retroelement Grande found that the ratio of the LTR copy number to the internal sequence copy number was about 2 suggesting that most elements are intact instead of being found as partial remnants resulting from an elimination process. Additionally, all subfamilies of Grande were found in each of the *Zea* species, so expansion must have occurred before the split of the *Zea* species (GARCIA-MARTINEZ and MARTINEZ-IZQUIERDO 2003).

The element Huck exemplifies the differences in expansion of copy number that can occur among different lineages. This difference can be used to distinguish different genomes cytologically. Dot blot estimates of Huck in *T. dactyloides* (35,600–50,800 copies), *Z. diploperennis* (73,500–104,800), *Z. mays* (165,700–236,000) and *Z. luxurians* (246,900–351,800) show significantly different copy numbers among the species

(MEYERS *et al.* 2001). Huck had expanded in the progenitor genome but after the split between *Zea* and *Tripsacum*, it ceased rapid expansion in the *Tripsacum* lineage. In the *Zea* lineage, Huck continued to increase in copy number but in the line that led to *Z. diploperennis*, expansion stopped or slowed while the lineages leading to *Z. mays* and *Z. luxurians* experienced continued retrotransposition. The difference between *Z. luxurians* and *Z. mays* must have appeared rapidly as only 0.63-0.7 million years separate the two species (HILTON and GAUT 1998). Thus, the Huck element hybridizes with different intensities to chromosomes from the three species present in the "tri-species" hybrid individual (*Z. mays*, *Z. diploperennis*, *T. dactyloides*) allowing them to be readily classified (Figures 2 and 3).

In contrast to Huck, the Tekay element hybridized strongly to all chromosomes in the "tri-species" hybrid (Figure 1). Based on polymorphisms in the *glb*1 and *adh*1 genes, maize and *Tripsacum* lineages are estimated to have diverged 4.5-4.8 Mya (HILTON and GAUT 1998). Tekay was the oldest element identified in the *adh* region with an insertion timeframe of about 5 Mya (SANMIGUEL *et al.* 1998), which would predate the split between *Zea* and *Tripsacum*. Therefore, the abundance of the Tekay element in both *Zea* and *Tripsacum* may be explained by an amplification of the Tekay family in the progenitor line that gave rise to the two genera, although it is also possible that it has experienced transpositions in the separate lineages. Except for Tekay, the other elements tested (Huck, Prem1, Prem2/Ji, Opie, Grande, Cinful) hybridized only weakly to *Tripsacum* chromosomes, supporting the idea that Tekay has the oldest date of expansion of the major extant maize retroelement families.

Divergence of knobs and centromeric elements revealed by GISH

Previous work to characterize the knob sequences has revealed that there are many sequence differences between the knobs of maize and *Tripsacum* but that a strong sequence conservation exists for at least a portion of the 180-bp knob element (DENNIS and PEACOCK 1984). Because of the sensitivity of detection, even a small amount of probe can be visualized and because the knobs consist of thousands of tandemly arrayed elements, a small amount of unblocked probe will be concentrated at discrete spots on the chromosomes and visualized as *Tripsacum* labeling on maize chromosomes. Because *Tripsacum* chromatin introgression was suspected for some lines analyzed (discussed below), maize knob sequences were labeled in a different color than the *Tripsacum* genomic DNA and hybridized simultaneously. The rationale is that if hybridization of the labeled *Tripsacum* DNA is seen on a maize chromosome, the colocalization of that signal with the maize knob signal will allow the possibility of introgression of a *Tripsacum* segment to be ruled out.

This approach resulted in strong labeling of maize knobs by the labeled knob sequence (in red) and weak labeling by *Tripsacum* genomic DNA (in green). On *Tripsacum* chromosomes, the opposite was observed: strong labeling by *Tripsacum* genomic DNA and weak labeling by maize knob sequences. The difference in intensity of hybridization reflects the divergence between the knob sequences of the two genomes.

In addition to the maize knobs, discrete signals from the labeled *Tripsacum* genomic DNA were visualized at the centromeres of the *Tripsacum* and some *Zea* chromosomes of the "tri-species" hybrid. In an analogous fashion to the knobs, the intensity of labeling of the *Zea* centromeres was markedly less than those of *Tripsacum*.

Two types of elements are known to be abundant at the centromeres of grasses, a tandemly arrayed centromeric satellite repeat and a family of centromere-specific retrotransposons (ANANIEV et al. 1998; CHENG et al. 2002; ZHONG et al. 2002). When CRM was used as a probe, the maize chromosomes were labeled most intensely, even when CRM sequences amplified from *Tripsacum* genomic DNA were used as FISH probes. Therefore, the GISH signal at *Tripsacum* centromeres is not due to an increased abundance of CRM elements or to their divergence from maize homologs. When a maize version of CentC was used as a FISH probe, the centromeres of *Zea* and *Tripsacum* were labeled with equal intensity (although some maize centromeres had very little CentC signal). Therefore, copy number of CentC does not account for the GISH signal at *Tripsacum* centromeres. The presence of a *Tripsacum* specific subset of CentC elements or other *Tripsacum* specific centromere sequences is a potential explanation for *Tripsacum* centromere labeling in GISH preparations (Figure 4).

Sequencing of large stretches of DNA from centromeric regions of maize and rice has revealed the abundance of a family of centromere-specific retroelements at centromeres (NAGAKI *et al.* 2004a; NAGAKI *et al.* 2004b; ZHONG *et al.* 2002). Because many of the elements are intact and have nearly identical LTRs, it is assumed that these elements resulted from recent insertion events. Solo LTRs from these elements have also been identified and the pattern of distribution of these elements suggests that their frequent insertion and elimination (perhaps through unequal crossovers between the LTRs) plays a role in the evolution of sequences at centromeres. The insertion and elimination of CRM retrotransposons and flanking sequences could contribute to the widely varying amounts of centromeric satellite among centromeres of a given karyotype

and among homologs of different varieties (CHENG *et al.* 2002; KATO *et al.* 2004; LAMB *et al.* 2004). CRM1a signal intensity was fairly uniform among the chromosomes of maize and *Z. diploperennis* while CRM2a intensity was quite different (Supplementary Figure 6). The relatively weak signal intensity in centromeres from *Tripsacum* and *Z. diploperennis* using maize CRM elements could reflect the active expansion of different CRM versions in the different lineages of the three species. However, given the relative scarcity of CRM elements in *Z. diploperennis* and *Tripsacum* and the uniformity of the centromeric satellite copy number among the chromosomes (CentC is also fairly uniform on *Z. luxurians*, data not shown), it is possible that CRM retrotransposon activity is especially high at centromeres in maize.

Development of cytological tools for examination of *Zea* and *Tripsacum*

Use of abundant retroelements as FISH probes to identify genome

Although GISH has been traditionally used to distinguish genomes cytologically, there are some drawbacks to that approach. Although many sequences, especially retroelements, may have diverged between species, there may be tandem arrays of conserved elements found in both, such as the megabases of knob sequences present in tandem arrays in *Zea* and *Tripsacum*. When GISH was performed using labeled DNA from *T. dactyloides* on "tri-species" hybrid chromosome spreads, the *Tripsacum* chromatin was differentially hybridized but the knobs from both species were also intensely labeled. As a result, while entire chromosomes are readily distinguished, this approach could be limited when applied to material thought to contain introgressed alien chromatin. Large knobs are present in different locations in various maize lines and small

knobs can be identified on virtually every chromosome arm using FISH and long exposure times (ADAWY *et al.* 2004); J. LAMB, A. KATO, J. MEYERS and J. BIRCHLER, unpublished). Thus, there is a possibility of confusing small knobs in maize with introgressed material. In general, the presence of shared repetitive elements arranged in tandem repeats may also create "false" GISH signals when the procedure is used with any related species. The addition of abundant maize retroelements and the 180bp knob repeat labeled in a different color than the *Tripsacum* genomic DNA increased the power of the GISH technique by providing a second confirmation of the identity of the respective chromatin (Figure 4 and 7).

In contrast to GISH, using abundant retroelements to distinguish genomes allows for long exposure times and intense labeling of *Tripsacum* euchromatin. Use of previously characterized maize retroelements as FISH probes allowed the distinction of maize and *Tripsacum* genomes in the allopolyploid species, *T. andersonii* (Supplementary Figure 2), in a maize X *Tripsacum* intergeneric hybrid (Supplementary Figure 1) and in the "tri-species" hybrid. *Tripsacum* specific retroelements were not previously reported so it was necessary to isolate them from genomic libraries prior to testing them as FISH chromosome paints.

By screening either a fosmid or small-insert library, clones were identified that paint the *Tripsacum* genome. Many crop species have a similar (or lower) gene density than maize including barley, rye, wheat, and oat (BENNETZEN and KELLOGG 1997). Because of the ease in creating small-insert or fosmid libraries, the simplicity of using FISH to screen the clones, and the relatively high number of desirable clones, it would be reasonable to use this approach to create genome distinguishing "paints" for other

species. In some cases, such paints may allow genomes to be distinguished when GISH cannot. Using retroelement paints instead of GISH also simplifies the simultaneous detection of other FISH probes.

Maize-Tripsacum addition and replacement lines

Maize X *Tripsacum* F1 hybrids can be backcrossed with maize pollen yielding viable seeds that result from unreduced gametes (MANGELSDORF and REEVES 1932). In just a few backcross generations, most of the *Tripsacum* chromosomes are lost with a few exceptions (MAGUIRE 1957; MAGUIRE 1963). Having a collection of maize lines with each of the different *Tripsacum* chromosomes would be a valuable tool in anchoring and extending the *Tripsacum* genetic map. To create a number of maize-*Tripsacum* addition lines, maize lines with various recessive markers were hybridized to *Tripsacum* and used as the paternal parent in subsequent backcrosses. By selecting for individuals in which the recessive allele was complemented each generation, the presence of the corresponding *Tripsacum* chromosome would also be selected (GALINAT and MANGELSDORF 1966; GALINAT *et al.* 1963).

Five lines produced in this way were provided for analysis by the MGCSC. Three of the lines still retained detectable *Tripsacum* chromatin but as translocation chromosomes involving a maize chromosome rather than as a maize-*Tripsacum* addition line with an intact chromosome. One of the lines was described previously to contain a translocation between a maize chromosome and the added *Tripsacum* chromosome (MAGUIRE 1957) but the other two represent novel translocations that have arisen during maintenance at the stock center. Because the translocation chromosomes can be found as

homozygotes with a total chromosome count of 20, it appears that the *Tripsacum* chromatin has replaced the maize chromatin and is transmittable through both the male and female gametophytes. Although cross-overs between maize and *Tripsacum* chromosomes are probably rare events, they are the simplest explanation for the type of translocation observed. The recovery of three independent translocations that replace the maize chromatin with a *Tripsacum* functional equivalent raises the possibility of using *Tripsacum* as a source of germplasm for maize modification. The retroelement paints developed in this study will facilitate identification and tracking of introgressed *Tripsacum* chromatin in maize lines.

Extension of maize FISH probes to *T. dactyloides* and *Z. diploperennis*

In concert with the fluorescently labeled retroelements, additional probes can be used in a normal fashion because blocking genomic DNA is not added. Although the GISH procedure does not preclude the use of additional probes, repetitive DNA probes need to be added in larger amounts to obtain similar intensities (data not shown) and some signals may not be easy to visualize. Numerous probes (180bp knob, TR-1, NOR, 5S, Cent4, "TAG" microsattelite), which label maize chromosomes allowing each homolog to be distinguished in somatic preparations, and three probes (*rp1*, *rp3*, and *α-zeinA*), which label gene clusters and hybridize to single locations in maize, were applied to chromosome preparations from the "tri-species" hybrid. Each of these probes hybridized to the *Tripsacum* and *Z. diploperennis* chromosomes demonstrating the general applicability of probes developed in maize to other *Zea* species and to *Tripsacum*.

Advances in FISH techniques have allowed specific maize loci to be detected anchoring the genetic map to a physical location (ANDERSON *et al.* 2004; KATO *et al.* 2005; KOUMBARIS and BASS 2003; WANG *et al.* 2006). In the related Sorghum species, the use of BACs have allowed detailed cytogenetic maps to be developed (ISLAM-FARIDI *et al.* 2002; KIM *et al.* 2002; ZWICK *et al.* 1998) opening up the possibility of cytogenetic examination of chromosomal rearrangements among species in this genus. Many of these BACs can be used as FISH probes in maize or rice (KOUMBARIS and BASS 2003; ZWICK *et al.* 1998). As an alternative to using BACs as FISH probes to detect specific genetic loci, small and unique DNA elements, such as individual genes, can be used as probes. The current limit of consistent detection in maize is ~3 kb (WANG *et al.* 2006) and as sequencing efforts proceed on the maize genome, numerous probes against specific loci will be developed. The "tri-species" hybrid provides an excellent means of testing the functionality of additional maize FISH probes in *Tripsacum* and *Z. diploperennis* because it contains an internal positive control in the form of the maize chromosomes. As more probes become available, they could be used to improve the genetic map of *Tripsacum* and will facilitate cytogenetic comparisons of these species.

Acknowledgements

We thank Mayandi Sivaguru and Brady Hardiman at the Molecular Cytology Core, Life Sciences Center for their help in image acquisition. We thank W. Galinat for permission to cite articles in the Maize Genetics Cooperation Newsletter. This work was supported by National Science Foundation grants DBI 0421671 and DBI 0423898, and by an MU Life Science Fellowship to JL.

References

ADAWY, S. S., R. M. STUPAR and J. JIANG, 2004 Fluorescence in situ hybridization analysis reveals multiple loci of knob-associated DNA elements in one-knob and knobless maize lines. J Histochem Cytochem **52:** 1113-1116.

ALTSCHUL, S. F., W. GISH, W. MILLER, E. W. MYERS and D. J. LIPMAN, 1990 Basic local alignment search tool. J Mol Biol **215:** 403-410.

ANANIEV, E. V., R. L. PHILLIPS and H. W. RINES, 1998 Chromosome-specific molecular organization of maize (*Zea mays* L.) centromeric regions. Proc Natl Acad Sci U S A **95:** 13073-13078.

ANDERSON, L. K., N. SALAMEH, H. W. BASS, L. C. HARPER, W. Z. CANDE *et al.*, 2004 Integrating genetic linkage maps with pachytene chromosome structure in maize. Genetics **166:** 1923-1933.

BENNETZEN, J. L., and E. A. KELLOGG, 1997 Do plants have a one-way ticket to genomic obesity? Plant Cell **9:** 1509-1514.

BRUNNER, S., K. FENGLER, M. MORGANTE, S. TINGEY and A. RAFALSKI, 2005 Evolution of DNA sequence nonhomologies among maize inbreds. Plant Cell **17:** 343-360.

CHENG, Z., F. DONG, T. LANGDON, S. OUYANG, C. R. BUELL *et al.*, 2002 Functional rice centromeres are marked by a satellite repeat and a centromere-specific retrotransposon. Plant Cell **14:** 1691-1704.

DENNIS, E. S., and W. J. PEACOCK, 1984 Knob heterochromatin homology in maize and its relatives. J Mol Evol **20:** 341-350.

GALINAT, W. C., 1988 The origin of corn, pp. 1-31 in *Corn and corn improvement*, edited by G. F. SPRAGUE. Am. Soc. Agron., Madison, WI.

GALINAT, W. C., and P. C. MANGELSDORF, 1966 Genetic correspondence of *Tripsacum* chromosomes to their homeologs from corn. Maize Genetics Cooperation Newsletter **40:** 99-100.

GALINAT, W. C., P. C. MANGELSDORF and R. S. CHAGANTI, 1963 Segregation of genetically marked chromosomes in maize-*Tripsacum* hybrids. Maize Genetics Cooperation Newsletter **37:** 32-33.

GARCIA-MARTINEZ, J., and J. A. MARTINEZ-IZQUIERDO, 2003 Study on the evolution of the Grande retrotransposon in the *Zea* genus. Mol Biol Evol **20:** 831-841.

GAUT, B. S., M. LE THIERRY D'ENNEQUIN, A. S. PEEK and M. C. SAWKINS, 2000 Maize as a model for the evolution of plant nuclear genomes. Proc Natl Acad Sci U S A **97:** 7008-7015.

HILTON, H., and B. S. GAUT, 1998 Speciation and domestication in maize and its wild relatives: evidence from the globulin-1 gene. Genetics **150:** 863-872.

ISLAM-FARIDI, M. N., K. L. CHILDS, P. E. KLEIN, G. HODNETT, M. A. MENZ *et al.*, 2002 A molecular cytogenetic map of sorghum chromosome 1. Fluorescence in situ hybridization analysis with mapped bacterial artificial chromosomes. Genetics **161:** 345-353.

KATO, A., P. S. ALBERT, J. M. VEGA and J. A. BIRCHLER, 2006 Sensitive FISH signal detection in maize using directly labeled probes produced by high concentration DNA polymerase nick translation. Biotechnic and Histochemistry, in press.

KATO, A., J. C. LAMB and J. A. BIRCHLER, 2004 Chromosome painting using repetitive DNA sequences as probes for somatic chromosome identification in maize. Proc Natl Acad Sci U S A **101:** 13554-13559.

KATO Y, T. A., 1984 Chromosome morphology and the origin of maize and its races. Evol. Biol. **17:** 219-253.

KIM, J. S., K. L. CHILDS, M. N. ISLAM-FARIDI, M. A. MENZ, R. R. KLEIN *et al.*, 2002 Integrated karyotyping of sorghum by in situ hybridization of landed BACs. Genome **45:** 402-412.

KOUMBARIS, G. L., and H. W. BASS, 2003 A new single-locus cytogenetic mapping system for maize (*Zea mays* L.): overcoming FISH detection limits with marker-selected sorghum (*S. propinquum* L.) BAC clones. Plant J **35:** 647-659.

LAMB, J. C., A. KATO and J. A. BIRCHLER, 2005 Sequences associated with A chromosome centromeres are present throughout the maize B chromosome. Chromosoma **113:** 337-349.

LAMB, J. C., J. THEURI and J. A. BIRCHLER, 2004 What's in a centromere? Genome Biol **5:** 239.

MA, J., K. M. DEVOS and J. L. BENNETZEN, 2004 Analyses of LTR-retrotransposon structures reveal recent and rapid genomic DNA loss in rice. Genome Res **14:** 860-869.

MAGUIRE, M. P., 1957 Cytogenetic studies of *Zea* hyperploid for a chromosome derived from *Tripsacum*. Genetics **42:** 473-486.

MAGUIRE, M. P., 1960a A study of homology between a terminal portion of *Zea* chromosome 2 and a segment derived *Tripsacum*. Genetics **45:** 195-209.

MAGUIRE, M. P., 1960b A study of pachytene chromosome pairing in a corn-*Tripsacum* hybrid derivative. Genetics **45:** 651-664.

MAGUIRE, M. P., 1963 High transmission frequency of a *Tripsacum* chromosome in corn. Genetics **48:** 1185-1194.

MANGELSDORF, P. C., 1947 The origin and evolution of maize. Advances in Genetics **1:** 161-207.

MANGELSDORF, P. C., and R. G. REEVES, 1931 Hybridization of maize, *Tripsacum* and *Euchlaena*. Journal of Heredity **22:** 329-343.

MANGELSDORF, P. C., and R. G. REEVES, 1932 Genetic and cytological studies in hybrids of *Zea* and *Tripsacum*. Proc. 6th Intern. Congr. Genet. **2:** 134-135.

MCCLINTOCK, B., T. A. KATO Y and A. BLUMENSHEIN, 1981 *Chromosome constitution of races of maize*. Colegio do Postgraduados, Chapingo, Mexico.

MEYERS, B. C., S. V. TINGEY and M. MORGANTE, 2001 Abundance, distribution, and transcriptional activity of repetitive elements in the maize genome. Genome Res **11:** 1660-1676.

MROCZEK, R. J., and R. K. DAWE, 2003 Distribution of retroelements in centromeres and neocentromeres of maize. Genetics **165:** 809-819.

NAGAKI, K., Z. CHENG, S. OUYANG, P. B. TALBERT, M. KIM *et al.*, 2004a Sequencing of a rice centromere uncovers active genes. Nat Genet **36:** 138-145.

NAGAKI, K., P. NEUMANN, D. ZHANG, S. OUYANG, C. R. BUELL *et al.*, 2004b Structure, divergence, and distribution of the CRR centromeric retrotransposon family in rice. Mol Biol Evol **22:** 845-855.

SANMIGUEL, P., and J. L. BENNETZEN, 1997 Evidence that a recent increase in maize genome size was caused by the massive amplification of intergene retrotransposons. Annals of Botany **82:** 37-44.

SANMIGUEL, P., B. S. GAUT, A. TIKHONOV, Y. NAKAJIMA and J. L. BENNETZEN, 1998 The paleontology of intergene retrotransposons of maize. Nat Genet **20:** 43-45.

SANMIGUEL, P., A. TIKHONOV, Y. K. JIN, N. MOTCHOULSKAIA, D. ZAKHAROV *et al.*, 1996 Nested retrotransposons in the intergenic regions of the maize genome. Science **274:** 765-768.

SWIGONOVA, Z., J. LAI, J. MA, W. RAMAKRISHNA, V. LLACA *et al.*, 2004 Close split of sorghum and maize genome progenitors. Genome Res **14:** 1916-1923.

TALBERT, L. E., J. F. DOEBLEY, S. LARSON and V. L. CHANDLER, 1990 *Tripsacum andersonii* is a natural hybrid involving *Zea* and *Tripsacum*: molecular evidence. American Journal of Botany **77:** 722-726.

WANG, C. J., L. HARPER and W. Z. CANDE, 2006 High-resolution single-copy gene fluorescence in situ hybridization and its use in the construction of a cytogenetic map of maize chromosome 9. Plant Cell. in press.

ZHONG, C. X., J. B. MARSHALL, C. TOPP, R. MROCZEK, A. KATO *et al.*, 2002 Centromeric retroelements and satellites interact with maize kinetochore protein CENH3. Plant Cell **14:** 2825-2836.

ZWICK, M. S., R. E. HANSON, T. D. MCKNIGHT, M. N. ISLAM-FARIDI, D. M. STELLY *et al.*, 1997 A rapid procedure for the isolation of Cot-1 DNA from plants. Genome **40:** 138-142.

ZWICK, M. S., M. N. ISLAM-FARIDI, D. G. CZESCHIN, JR., R. A. WING, G. E. HART *et al.*, 1998 Physical mapping of the liguleless linkage group in Sorghum bicolor using rice RFLP-selected sorghum BACs. Genetics **148:** 1983-1992.

Table

Table 1. Probe intensity on "tri-species" hybrid chromosomes.

Probe name:	Hybridization pattern in "tri-species" hybrid
Cinful	**MAIZE, DIPLOPERENNIS**, *Tripsacum*
Prem1	**MAIZE, DIPLOPERENNIS**, *Tripsacum*
Tekay	**MAIZE, DIPLOPERENNIS**, ***TRIPSACUM***
Prem2/Ji	**MAIZE, DIPLOPERENNIS**, *Tripsacum*
Opie	**MAIZE, DIPLOPERENNIS**, *Tripsacum*
Huck	**MAIZE,** DIPLOPERENNIS, *Tripsacum*
Grande	**MAIZE,** DIPLOPERENNIS, *Tripsacum*
Fosmid A4	MAIZE, DIPLOPERENNIS, ***TRIPSACUM***
Fosmid B5	maize, diploperennis, ***TRIPSACUM***
subclone B5-2	maize, diploperennis, ***TRIPSACUM***
subclone B5-3	maize, diploperennis, ***TRIPSACUM***
Fosmid B6	**MAIZE**, DIPLOPERENNIS, *Tripsacum*
Fosmid B8	maize, diploperennis, ***TRIPSACUM***
subclone B8-15	maize, diploperennis, ***TRIPSACUM***
Fosmid C12	MAIZE, DIPLOPERENNIS, ***TRIPSACUM***
Fosmid E12	**MAIZE**, **DIPLOPERENNIS,** *Tripsacum*
Fosmid F7	Maize, diploperennis, TRIPSAUM Also, 4 distinct spots (3 *Tripsacum*, 1 diploperennis)
Fosmid G6	MAIZE, DIPLOPERENNIS, ***TRIPSACUM***
Fosmid H7	Speckles along *Tripsacum* chromosomes
TC#3	MAIZE, DIPLOPERENNIS, *TRIPSACUM*
TC#4	Maize, diploperennis, *TRIPSACUM* (knob signal also present)
TC#5	maize, diploperennis, ***TRIPSACUM***
TC#7	MAIZE, DIPLOPERENNIS, ***TRIPSACUM*** (knob signal also present)
ZD#9	**MAIZE**, DIPLOPERENNIS, *Tripsacum*
ZD#10	**MAIZE**, **DIPLOPERENNIS**, *Tripsacum*
TC#11	Three spots of hybridization
TC#12	MAIZE, DIPLOPERENNIS, ***TRIPSACUM***
TC#13	MAIZE, DIPLOPERENNES, ***TRIPSACUM***
TC#14	Knob sequences
TC#15	maize, diploperennis, *Tripsacum*
TC#17	maize, diploperennis, *TRIPSACUM*
TC#19	3 spots, one on each genome
TC#24	maize, diploperennis, ***TRIPSACUM***
TC#25	maize, diploperennis, ***TRIPSACUM***
CRM1a	**MAIZE**, DIPLOPERENNIS, *Tripsacum* (Only the centromeres are labeled)
CRM2a	**MAIZE**, DIPLOPERENNIS, *Tripsacum* (Only the centromeres are labeled)

Lower case non-bolded font indicates weak or no hybridization (i.e. maize). Uppercase and bold lettering indicates increasing degrees of hybridization intensity (i.e. MAIZE, **MAIZE,** **MAIZE**).

Figures

Figure 1. Hybridization of abundant maize retroelements to *Tripsacum* and *Zea*. (A) Chromosomes from a maize X *Z. diploperennis* hybrid labeled with Opie (red), Huck (green) and the 180bp knob repeat (white). The chromosomes of *Z. diploperennis* have small terminal knobs and are weakly labeled by the Opie and Huck elements. In the remaining images (B-F), chromosomes from a "tri-species" hybrid containing *T. dactyloides*, *Z. diploperennis* and *Z. mays* genomes are hybridized with pairs of abundant maize retroelements differently labeled in red and green. Chromosomes were counterstained with DAPI. The second row contains gray value images of only the red channel and the third row contains only the green channel with the probe in each case indicated.

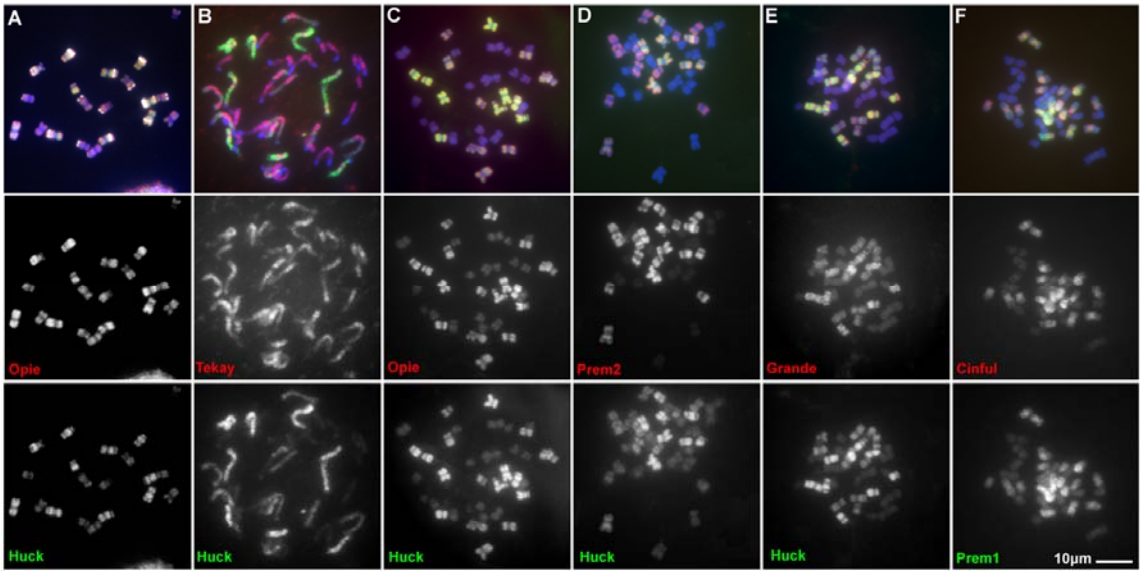


Figure 2. Selection of *Tripsacum* fosmids for chromosome painting.
Ninety-six fosmids were dotted on a nylon membrane and hybridized to maize Cot-100 DNA in (A) and to genomic DNA from *T. dactyloides* in (B). Circled dots represent fosmids that were tested as FISH probes. Fosmids circled in purple did not hybridize strongly to either maize or *Tripsacum* genomic DNA. Fosmids circled in blue hybridized to *Tripsacum* but not maize DNA and red circles indicate fosmids that hybridized to maize but not *Tripsacum* DNA. The "tri-species" hybrid chromosome spreads were hybridized with labeled fosmid DNA (red) and with Prem1 (C,D) or Huck (E,F). The fosmid used is indicated and the fosmid hybridization signals alone are presented in gray values below the merged images. TF-E12, although isolated from *Tripsacum*, hybridizes more strongly to *Zea* chromosomes. TF-B5, TF-C12 and TF-H7 hybridize more strongly to *Tripsacum*.

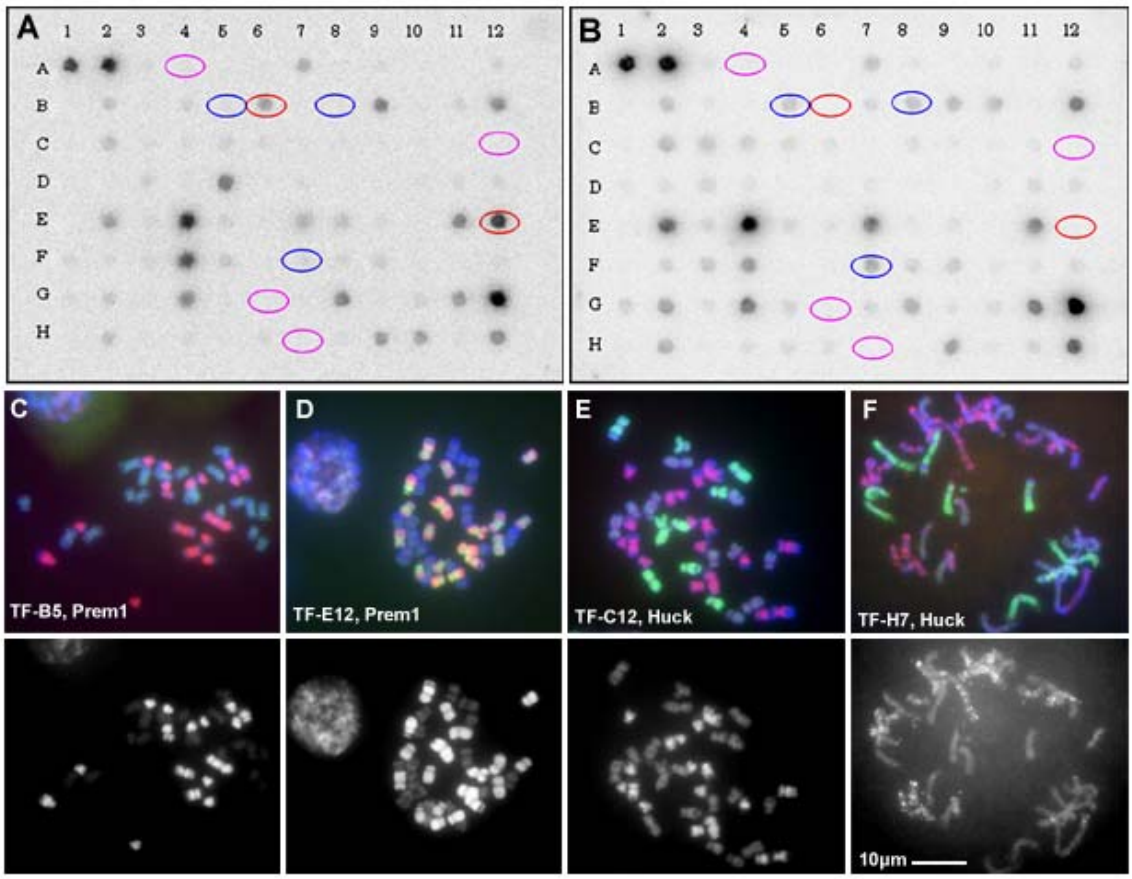


Figure 3. Selection of small DNA probes that distinguish *Zea* and *Tripsacum* genomes. The "tri-species" hybrid chromosomes were hybridized with the labeled clone TC#5 (A) and TC#25 (B) (red) and Huck (green). (A') and (B') show the gray value images for the TC probes. Genomic DNA from *Z. diploperennis* (1), *T. dactyloides* (2), *S. bicolor* (3), *T. laxum* (4), *Z. mays* (5), and *Z. luxurians* (6) was digested with *Hin*dIII, electrophoresed through an agarose gel, blotted to nylon membranes, and hybridized to 32P labeled TC#5 (C) or TC#25 (D). (C') and (D') show the ethidium bromide stained gels.

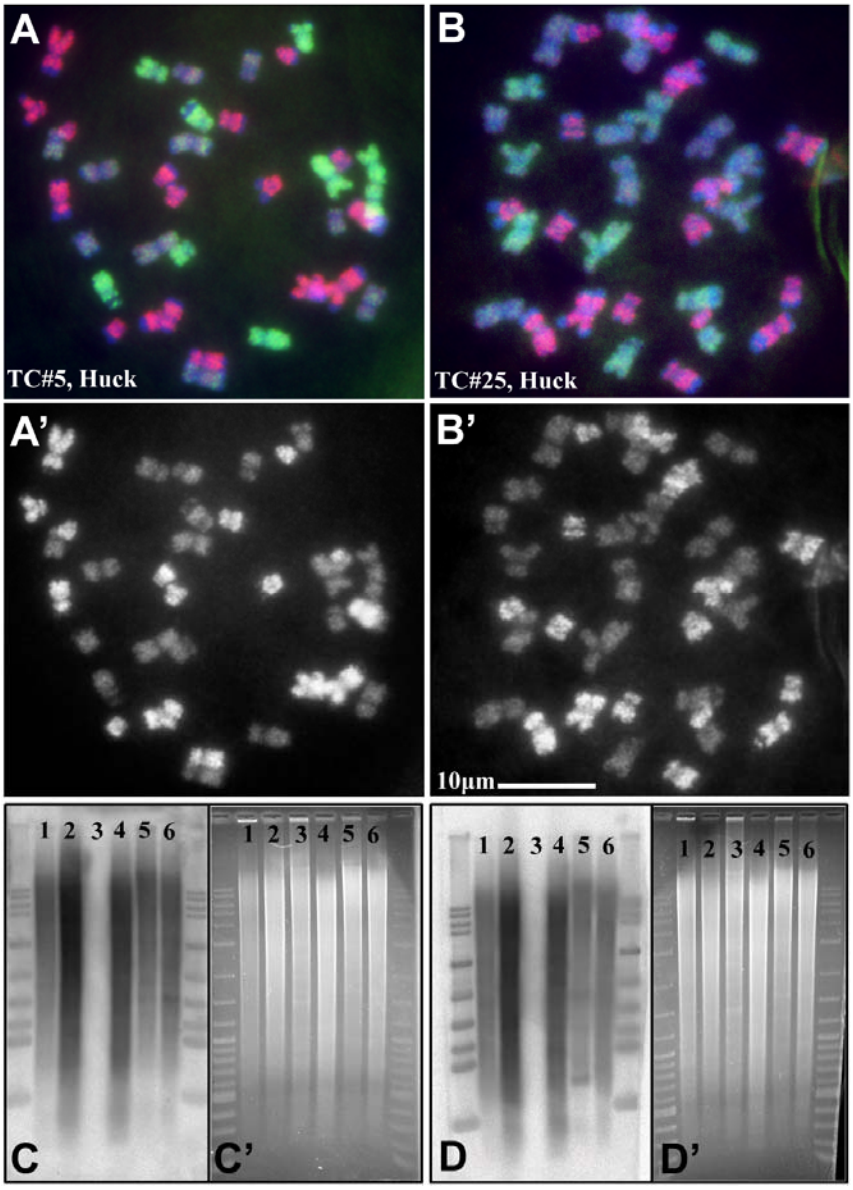


Figure 4. Divergence of centromeric elements between *Tripsacum* and *Zea*.
The "tri-species" hybrid chromosome spreads were hybridized with FITC-labeled genomic DNA from *T. dactyloides* (green), 50X unlabeled genomic maize DNA, Texas Red-labeled Huck, Grande, Prem1, and maize 180bp knob repeat (red) in (A), CentC (green) and Grande (red) in (B) and full length CRM (green) and Grande (red) in (C). (A') shows only the *Tripsacum* genomic DNA signal.

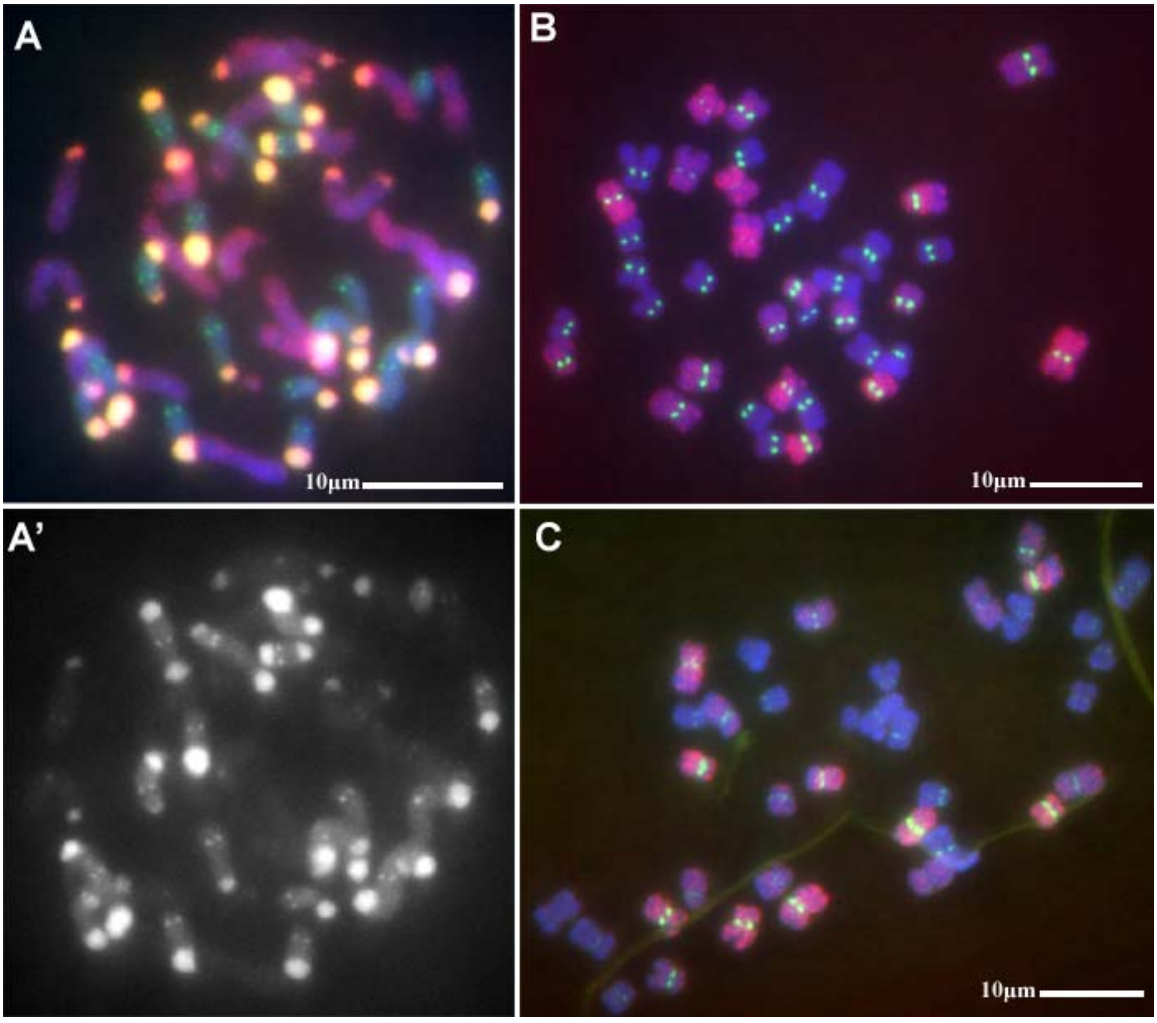


Figure 5. Use of abundant retroelements and GISH to paint *Tripsacum* and *Zea* chromosomes.
In the first column (A-D) chromosome preparations were hybridized with Oregon green-labeled *Tripsacum* genomic DNA (green), 50X unlabeled autoclaved maize genomic DNA and a cocktail of maize retroelements including Prem1, Huck, Opie, and Grande as well as the 180bp knob repeat (red). In the second column (A'-D'), chromosomes were hybridized to TC#5, TC#25, TF-B5-3 (red), Huck, Grande, and Prem1 (green). Chromosomes were obtained from a "tri-species" hybrid (A), lines thought to contain an additional *T. dactyloides* chromosome (B and C), and a reciprocal translocation involving maize chromosome 2 and a *Tripsacum* chromosome (D). Arrows in (B) and (C) indicate the sites of hybridization to *Tripsacum* probes. In (D), the red arrow indicates the T-2 translocation chromosome and the white arrow denotes the 2-T chromosome.

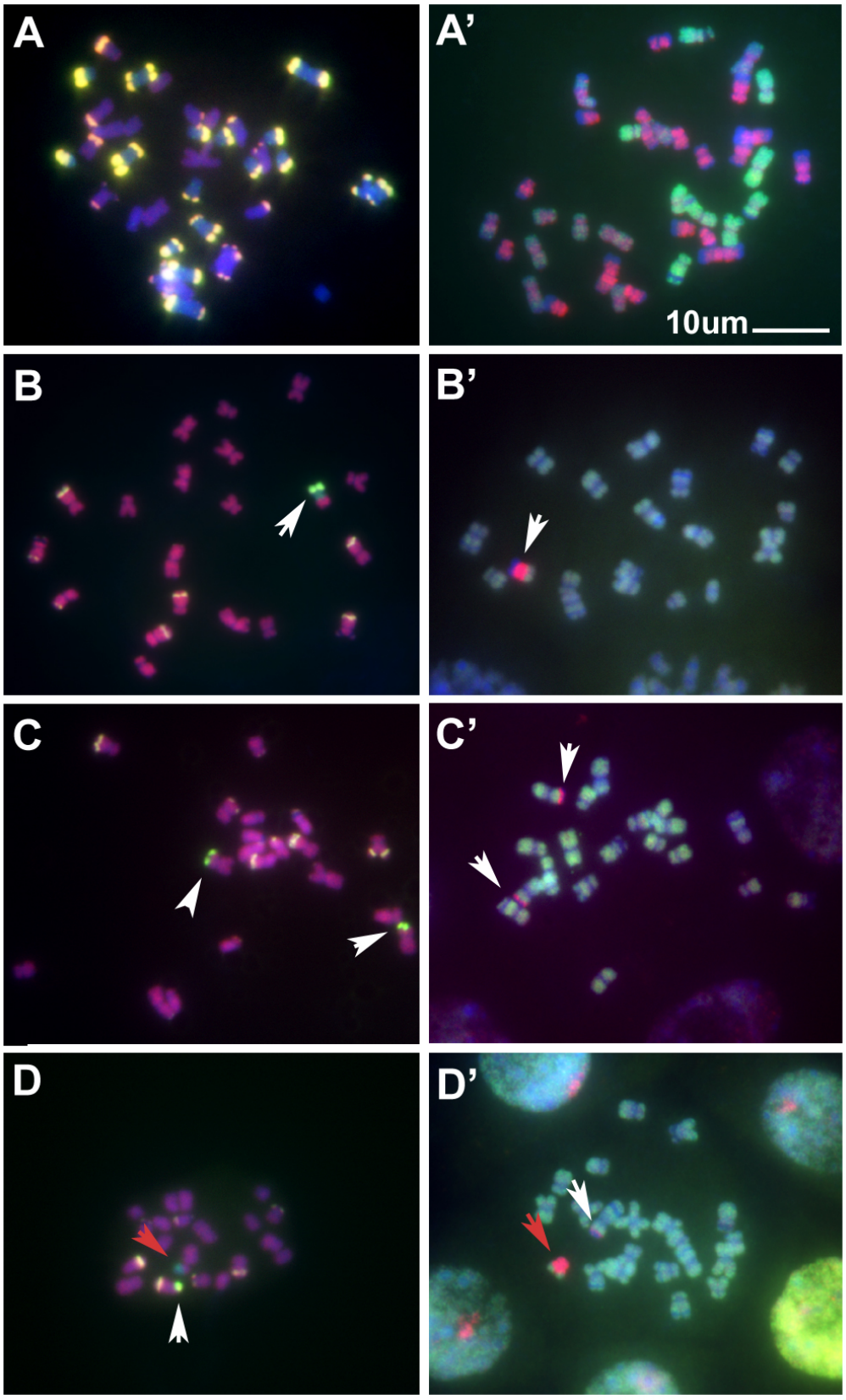

A
A'
10um
B
B'
C
C'
D
D'

Figure 6. Hybridization of maize repetitive sequences to "tri-species" hybrid chromosome spreads.
The "tri-species" hybrid chromosome spreads were hybridized simultaneously to a chromosome painting probe in green, Huck in (A) to (D) and TF-B5-2 in (E), and to various repetitive elements. Repetitive probes were (A) "TAG^{n}" oligonucleotide probe (red), (B) 5S ribosomal RNA (red), (C) TR-1 knob (red) and 180bp knob (blue), (D) Cent4 (red), and (E) 17S ribosomal RNA (red) sequences. The gray value images below (A-D) show only the red channel from the images above them. (E') shows the NOR signal only and (E") shows the chromosomes stained with DAPI. In (A-E), arrows indicate hybridization signals. In (E"), arrowheads denote the distal DAPI signals which are separated from the remainder of the chromosome because of a partially uncondensed NOR from maize and *Z. diploperennis* as illustrated in (E'). The site of hybridization of the NOR probe on the *Tripsacum* chromosome was smaller and did not have an uncondensed appearance.

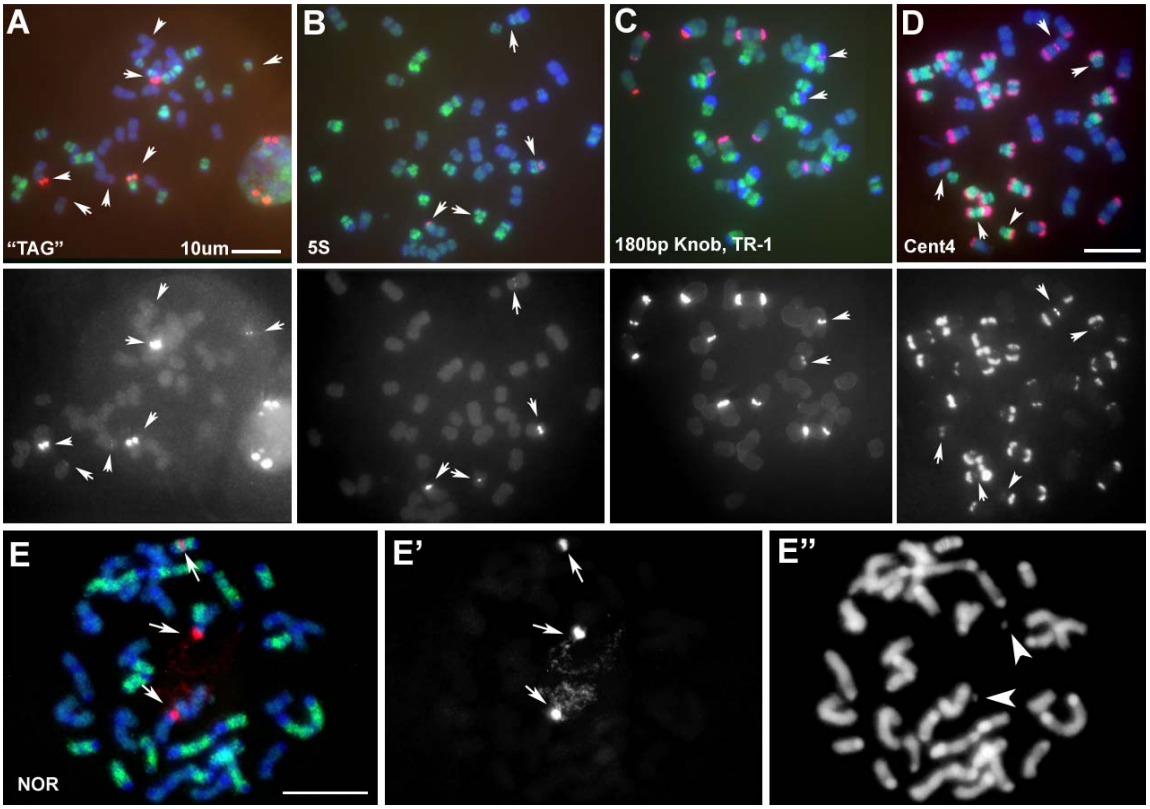
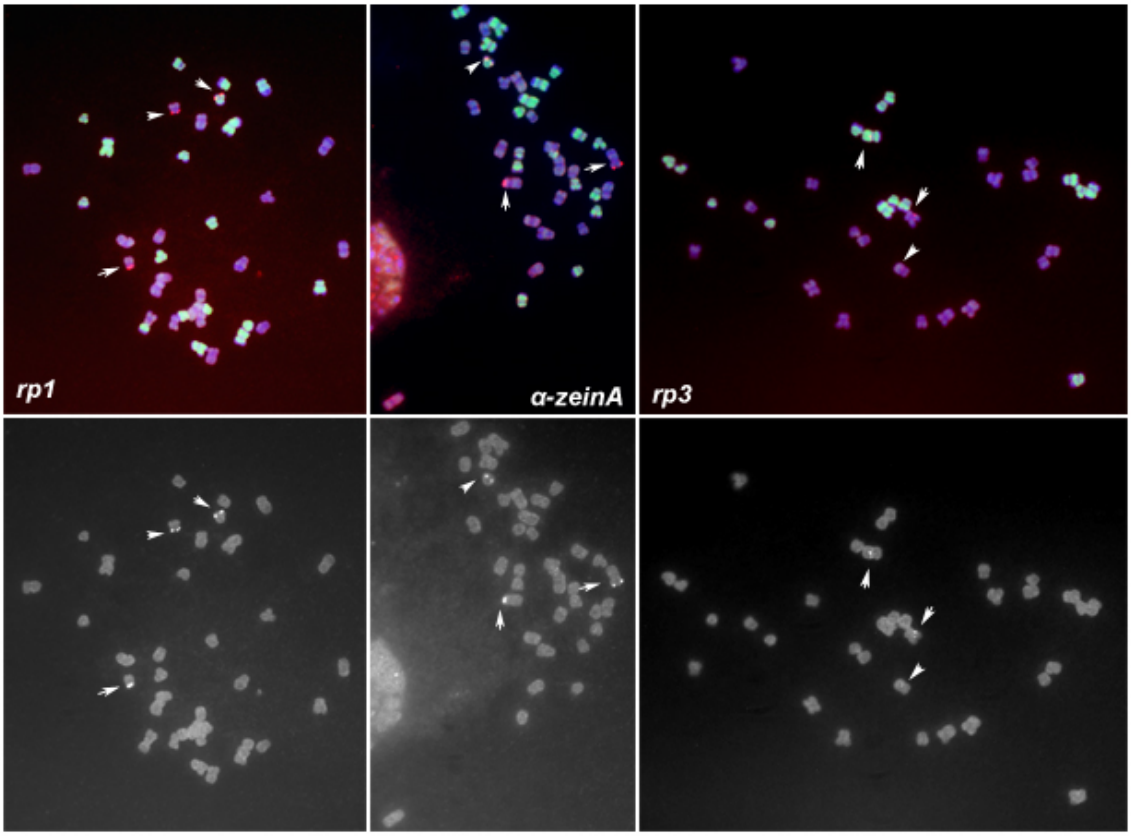
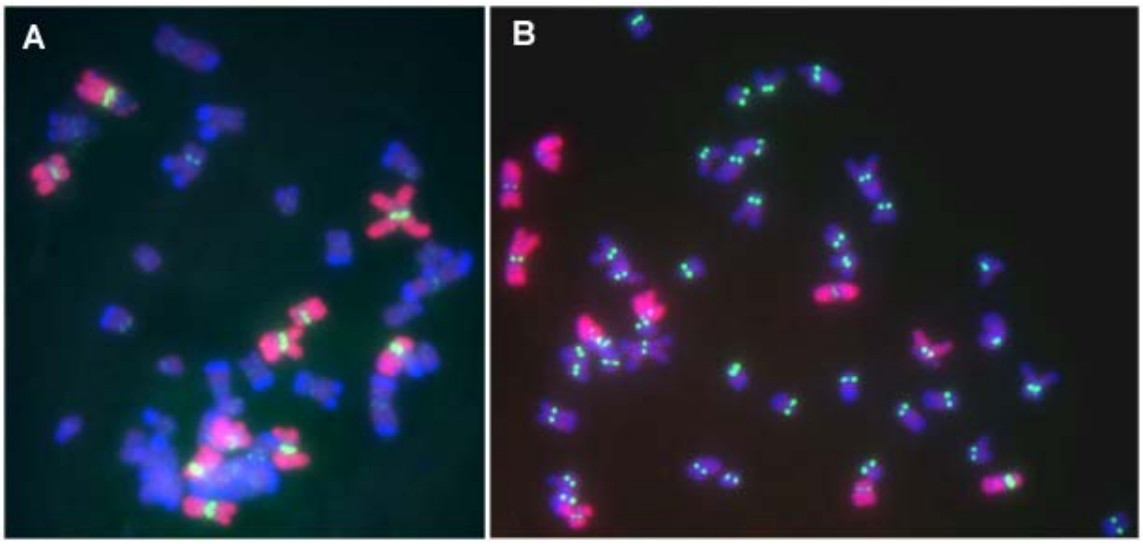


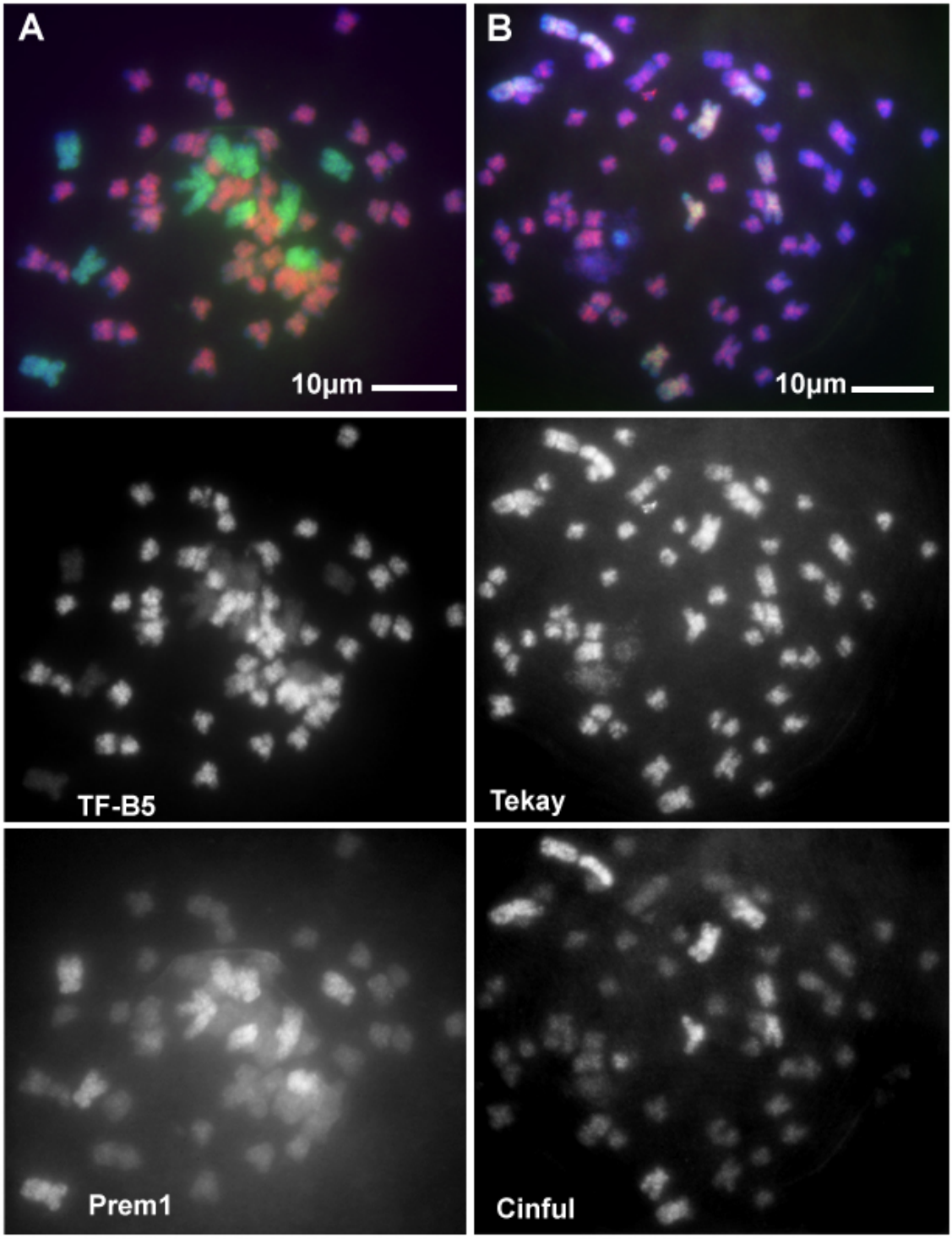
Figure 7. Hybridization of single loci probes.
The "tri-species" hybrid chromosome spreads were hybridized with probes that detect the maize *rp1, rp3,* or *α-zeinA* gene clusters (red) together with TC#5 (green). The gray value images show only the red channel and the arrows indicate hybridization sites.



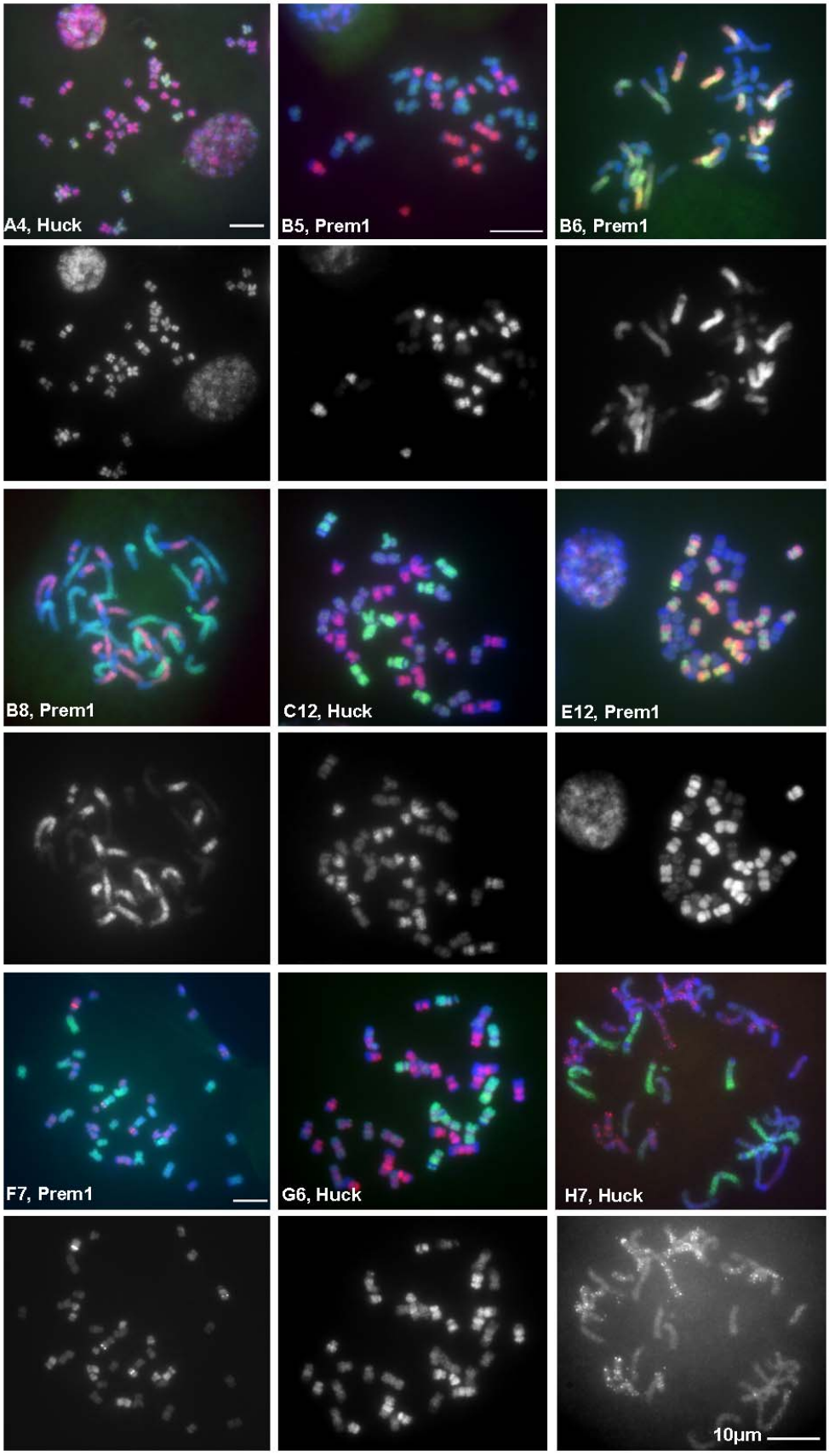
Supplementary Figure 1. Maize X *Tripsacum*
Z. mays X *T. dactyloides* F1 hybrid chromosome spreads were hybridized in (A) with Huck (red) and CRM (green) and in (B) with Opie (red) and CentC (green).



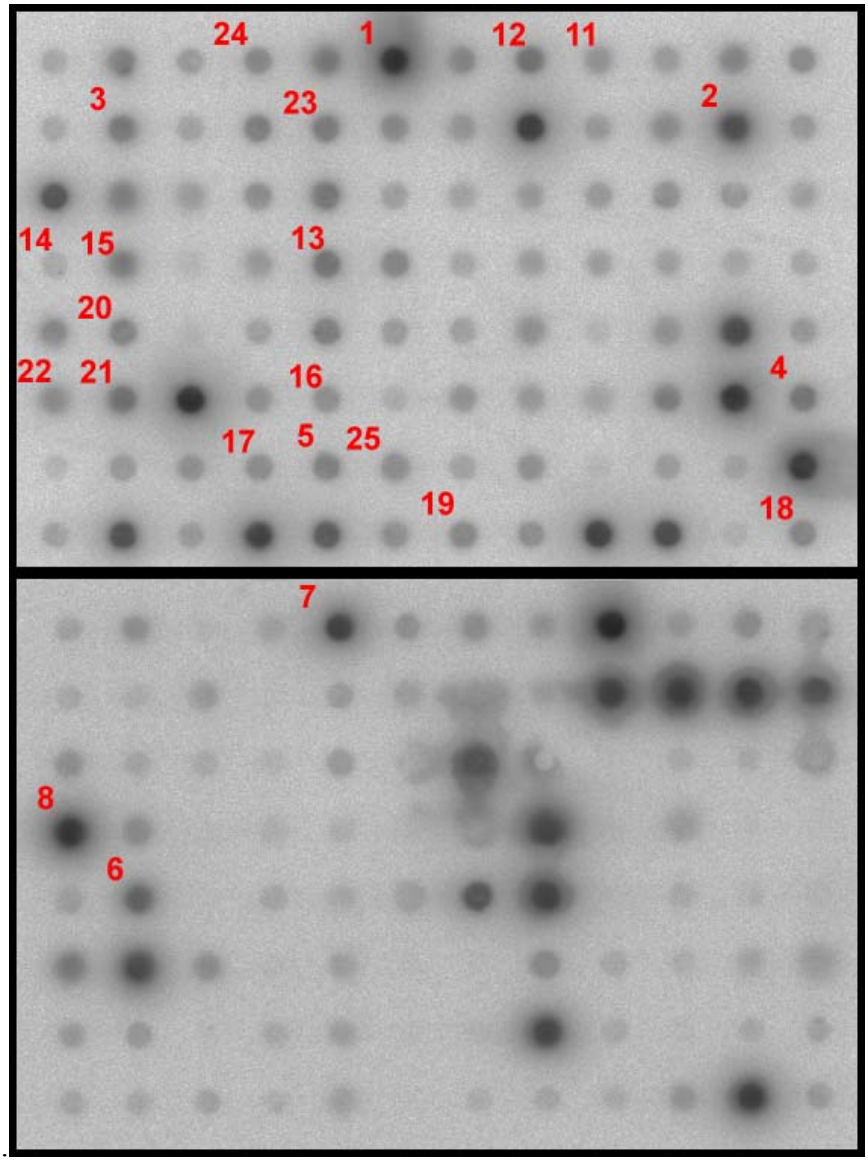
Supplementary Figure 2. Retroelement labeling of *T. andersonii*.
Chromosome spreads from *T. andersonii* were hybridized with TdB5 (red) and Prem1 in (A), and Tekay (red) and Cinful (green) in (B). The hybridization signals from individual probes are presented in gray values below the merged images.



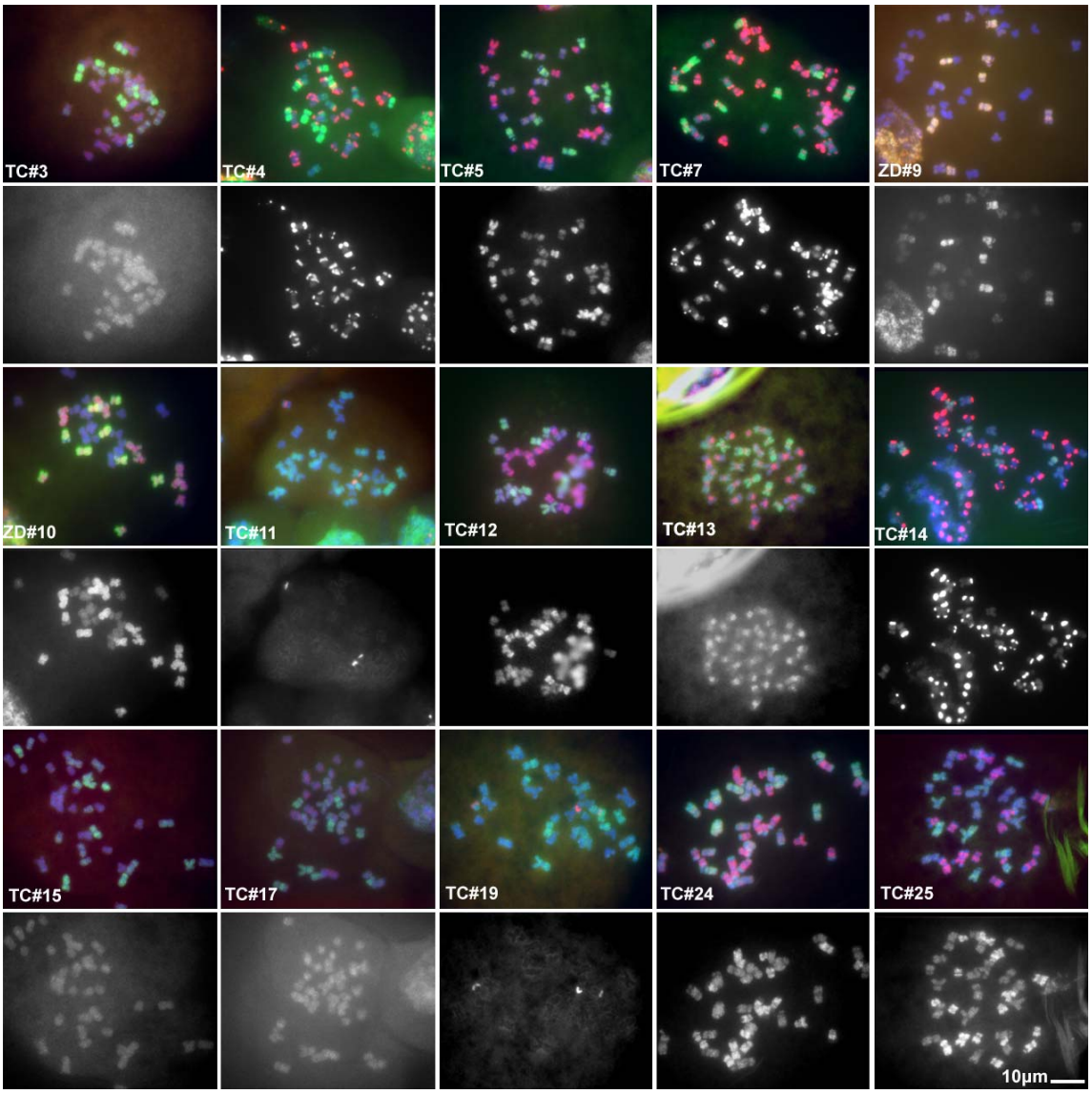
Supplementary Figure 3. FISH screening of *Tripsacum* fosmids.
Nine fosmids were fluorescently labeled (red) and hybridized to chromosomes from a "tri-species" hybrid. Huck or Prem1 (green) was also hybridized to allow the genome of origin for each chromosome to be determined. The name of the fosmid (TF-A4, B5, B6, B8, C12, E12, F7, G8, H7) is indicated on each image and the fosmid signal alone is presented in gray values below the merged image. The images of TF-B5, C12, E12 and H7 are the same as those of Figure 2. The bars represent 10 μM. All figures that do not have white bars are the same scale as TF-H7 (bottom right corner).


A4, Huck
B5, Prem1
B6, Prem1
B8, Prem1
C12, Huck
E12, Prem1
F7, Prem1
G6, Huck
H7, Huck
10µm

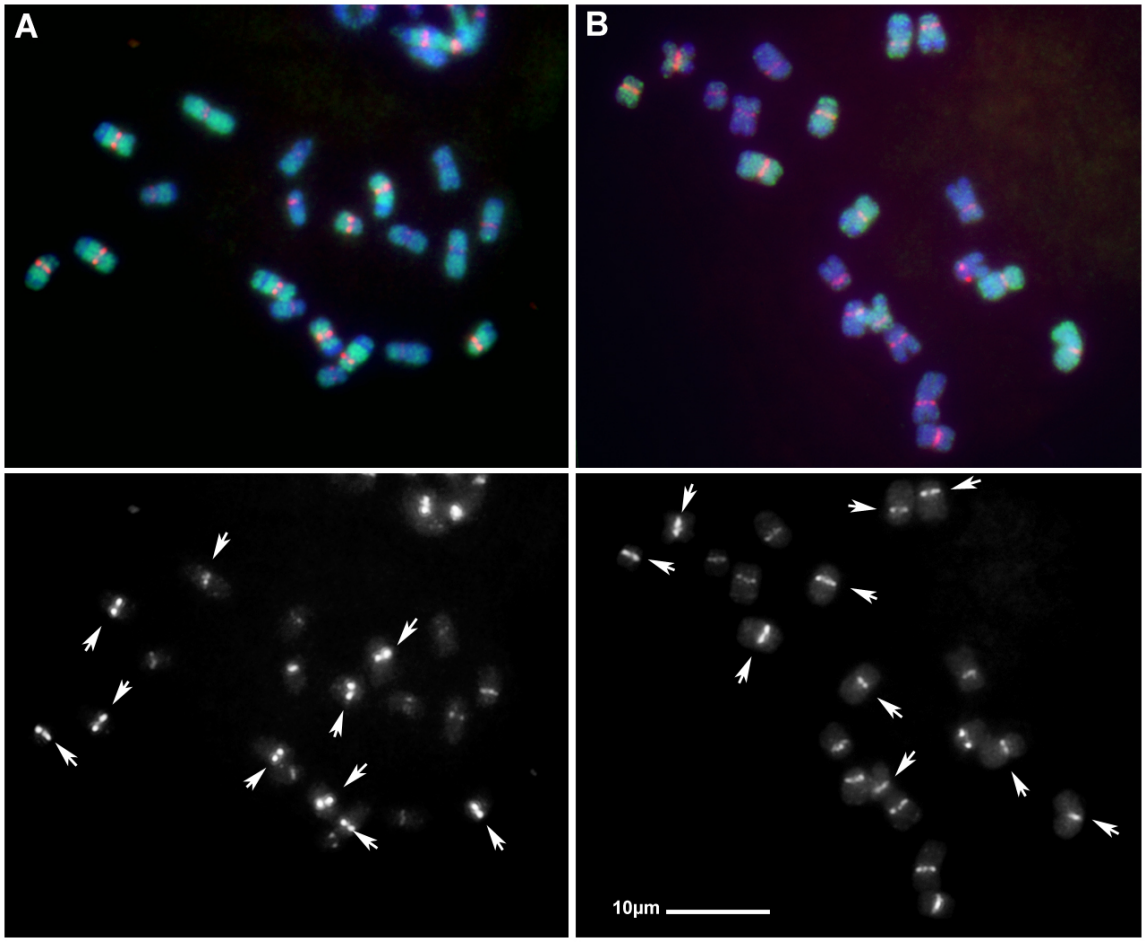
Supplementary Figure 4. Dot blots of small-insert library hybridized with 32P labeled *Tripsacum* genomic DNA.
Two 96-well mini-preps were performed on small-insert *Tripsacum* genomic clones and the plasmids dotted on two nylon membranes. The two membranes were hybridized with 32P labeled genomic DNA from *T. dactyloides*. Twenty-three clones were selected for further analysis (TC#1-8 and 11-25).



Supplementary Figure 5. FISH screening of *Tripsacum* genomic fragments.
Fifteen small genomic fragments were cloned, PCR-amplified and fluorescently labeled (red) for hybridization to chromosomes from a "tri-species" hybrid. Huck (green) was also hybridized to allow the genome of origin for each chromosome to be determined. The clone designation is indicated on each image and the clone signal alone is presented in gray values below the merged image. Clones from *T. dactyloides* are labeled "TD" and the two clones from *Z. diploperennis* are labeled "ZD." The images for TC#5 and 25 are the same as those in Figure 3.



Supplementary Figure 6. CRM hybridization on maize and *Z. diploperennis*. Chromosome spreads from a maize X *Z. diploperennis* hybrid plant were hybridized with the LTR from CRM2a (red) and Huck (green) in (A), and CRM1a (red) and Grande (green) in (B). The CRM signals alone are presented in gray values below the merged images and the arrows indicate the signals from maize chromosomes.



7. A HEMICENTRIC INVERSION IN MAIZE LINE KTF CREATED TWO SITES OF CENTROMERIC ELEMENTS AND MOVED THE KINETOCHORE FORMING REGION

This chapter has been formatted in preparation for submission to the journal Chromosoma and uses the corresponding referencing style.

A hemicentric inversion in the maize line KTF created two sites of centromeric elements and moved the kinetochore forming region.

Jonathan C. Lamb[1], Julie M. Meyer[2], James A. Birchler‡[1].

[1]Division of Biological Sciences, University of Missouri-Columbia, Columbia, MO 65211.

[2]Department of Agriculture, Truman State University, Kirksville, MO 63501.

‡Corresponding author:

James A. Birchler

117 Tucker Hall

Columbia, MO 65211

Phone: 573.882.4871

Fax: 573.882.0123

E-mail: BirchlerJ@Missouri.edu

Abstract

Here we report the identification of a maize line, knobless tama flint (KTF) containing a version of chromosome 8 with two spatially distinct regions of centromeric elements, one at the original genetic position and the other at a novel location on the long arm. The new site of centromeric elements functions as the kinetochore forming region resulting in a change of arm length ratios. Examination of FISH markers on chromosome 8 revealed an inversion between the two centromere sites relative to standard maize lines, indicating that dicentric chromosome 8 resulted from a hemicentric inversion with one breakpoint 20 cMc (20%) on the long arm and the other in the original cluster of

centromere repeats. This inversion moved the kinetochore forming region but left the remainder of the centromere. In a hybrid between a standard line (Mo17) and KTF, both chromosome 8 homologues were completely synapsed at pachytene despite the inversion. Although the homologous centromeres were not paired, they were always correctly oriented at anaphase and migrated to opposite poles. Additionally, recombination on 8L was severely repressed in the hybrid.

Introduction

Chromosomal rearrangements of various kinds frequently distinguish different species. Rearrangements can alter the position of centromeres, arm length ratios, and create new genetic linkages. Here, we identify a maize line, knobless Tama flint (KTF), that has a chromosome pair with two distinct sites of centromeric elements separated by a substantial distance equal to twenty percent of the long arm. We demonstrate that the region between the two centromere element sites is inverted relative to the standard maize karyotype, placing the functional kinetochore at a new location and changing the arm length ratio. This finding indicates that a novel type of rearrangement has occurred, an inversion with one breakpoint in the centromere, which we have termed a "hemicentric inversion."

Although dicentric chromosomes stabilized by inactivation of one of the centromeres have been frequently observed in mammalian systems (Murphy and Karpen 1998), only recently has an example of a stable chromosome containing two centromere regions, one of which is inactive, been reported in plants (Han et al. 2006). In that study, individuals containing a dispensable chromosome with an inverted duplication produced

progeny with small functionally monocentric chromosomes containing two centromere regions after multiple rounds of the breakage-fusion-bridge (BFB) cycle. The KTF line contains a second example of a chromosome with two centromere regions.

Individuals who are heterozygous for many chromosome rearrangements suffer from reduced fertility or sterility due to chromosome segregation during meiosis. In the case of inversions, recombination in the inverted region can lead to duplications, deletions or chromatin bridges resulting from placing two centromeres on the same chromosome. However, the KTF hemicentric inversion involves the pericentromeric region where little recombination occurs (Anderson et al. 2003), suggesting that this inversion may not cause problems during meiosis. To test this hypothesis, we observed the progression of the KTF chromosome 8 through Meiosis I in a hybrid between KTF and a standard maize line. Counter to expectations, the inverted region was completely synapsed from early through late pachynema. Although this arrangement physically separated the centromeres of the two homologous chromosomes, we did not observe any indications of aberrant chromosome transmission. However, recombination was severely repressed distal to the breakpoint, probably due to the proximity of the KTF centromere.

The mechanism that produced this rearrangement is unlikely to be different than other types of inversions, but it leaves a distinctive footprint: an inactive cluster of centromeric elements. We discuss the implications of these findings to centromere function and the possible contributions of this type of rearrangement to karyotype evolution.

Materials and methods

Plant lines

The Mo17 used in this study is from a stock maintained in our lab. The KTF line (Ames 21969) and the Tama line (PI 217411) were obtained from the Maize Genetics COOP Stock Center. The KTF line has been maintained first by the North Central Regional Plant Introduction Station and then later by the maize Genetics COOP since 1963. Because it was donated many years ago, it has not been possible to determine precisely from where it originated. However, the seeds have an unusual and distinctive flat shape that is identical to those of the Tama flint line, an open pollinated northern flint variety. Kernels of the Tama line are purple, red, and colorless while the KTF kernels are all colorless. Many early studies of maize cytological features focused on knobs. The Tama flint line would have been of special interest because it had few knobs and we believe that the KTF line was derived from the Tama line by selection for an individual or individuals that had no visible knobs.

Chromosome preparation, hybridization, image capture

Somatic chromosome spreads were made by treating root tips with nitrous oxide, fixing with acetic acid, digesting in an enzyme solution containing pectolyase and cellulase, suspending in 3:1 acetic acid:methanol and dropping the suspension on glass slides in a humid chamber as previously described (Yu et al. 2006). Meiotic cell spreads were prepared by digesting fixed and staged anthers in the same enzyme solution as the root tips as previously described (Kato et al. 2004; Lamb et al. 2005). After the slides

dried, they were subjected to UV crosslinking and used for hybridization within a few hours (Yu et al. 2006).

For hybridization, 5 ul of autoclaved salmon sperm DNA (120 ng/ul in 2X SSC, 1XTE) was applied to the slide and covered with a plastic coverslip. The probe mixture (~20 ng/ul of each probe in a total of 5ul) was placed in a 500 ul microcentrifuge tube. Both the slide and probe mixture were placed in a double boiler lined with wet paper towel and heated for 5 minutes. The slide was then placed on a piece of flat metal resting on ice and the probe placed in wet ice for 30 seconds. The coverslip was lifted with a scalpel, the probe applied and the coverslip replaced. Next, the slide was incubated at 55°C for 12 to 24 hours in a sealable plastic container lined with paper towel wetted with 2X SSC. Coverslips were removed by dipping in room temperature 2X SSC and then washed in 55C 2X SSC for 5-20 minutes (Kato et al. 2006; Lamb and Birchler 2006). After washing slides, Vectashield containing 4',6-diamino-2-phenilinodole (DAPI) (Vector Laboratories, cat. H-1200) was applied to preserve the preparations and to counter-stain the DNA.

Images were captured with an Olympus BX61 microscope using Applied Spectral Imaging (ASI) software and a cooled charge-coupled device (CCD) camera, Cool-1300QS. Some images were then processed using the "sharpen" feature of the ASI software. Background was reduced by overlaying and subtracting a severely blurred copy of the image using the "Gaussian blur" and "difference" function of Photoshop as previously described (Lamb and Birchler 2006; Yu et al. 2006).

FISH probe production

Probes were made from PCR products (below) by a nick translation reaction that incorporates fluorescently conjugated nucleotide analogs, TexasRed-dCTP (red), AlexaFluor488-dUTP (green), and Cy5-dUTP (far red) using a high concentration of DNA polymerase (Kato et al. 2006; Lamb et al. 2006b; Yu et al. 2006).

The BAC8L-AC157487 probe was described previously (Lamb et al. 2006a). It was designed by identifying three sequences free of repetitive elements from a BAC sequence (accession #: AC157487) located on chromosome 8L. The sequences were PCR amplified and pooled for use as a template for the nick translation reaction. CentC, CRM, and the 180bp knob repeat probes have been previously described (Kato et al. 2004; Lamb et al. 2005).

To identify candidate genes for use as FISH probes on 8L, the approximate genetic position of the new site of CentC was determined by converting the cytological distance to a genetic distance. First, the position relative to the Mo17 long arm of the new KTF centromere was measured for 14 well spread pachytene chromosomes using the computer application MicroMeasure version 3.3 (Reeves and Tear 2000) and averaged. Then the cytological position, 20cMc, was compared to genetic markers on the converted UMC98 genetic map produced by the Morgan2McClintock converter (Lawrence et al. 2005). 20 cMc corresponded to bin 8.02; therefore all gene sequences listed in the genome viewer at maizegdb (www.maizegdb.org/) in bins 8.01 to 8.04 were examined. Most of the available gene sequences were expressed sequence tags (ESTs) or cDNAs.

The minimum detection limit for maize genomic targets is around 3 kb, larger than most ESTs or cDNAs. Therefore, primers were designed from the EST or cDNA

sequences and used to amplify the genes from genomic templates in order to include the introns as well as the exons, thereby maximizing the target size. The *atp2* (cDNA accession #: X54233, ~2015bp, PCR product: ~3000 bp) and *cesa-1* (cDNA accession #: AF200525, 3736, PCR product: ~5000 bp) genes were PCR amplified from KYS genomic DNA using the following primers: *atp2*F_11, 5'-AACCCAGCGTCCCCCGAAGTCTAA; *atp2*R_1952, 5'-ATTCGCGTTACACAGGCCACAAGTTTC; cesa1F_308, GATTTGCGGTGACTCTGTGGGTGTTT; and cesa1R_3561, CGGCACGGAAGAAGCTAATGAACA. Many maize genes are located in clusters and make good FISH targets. Multiple gene sequences similar to the *tub2* (cDNA accession #: X52879, 1579 bp, PCR product: ~3500 bp), *rip1* (cDNA accession #: M83926, 1kb, PCR product: ~1000bp) and *actin* genes (e.g., U60513) are present in public databases. Therefore, they were suspected to be in clusters on 8L. The *tub2* and *rip1* genes were PCR amplified from KYS genomic DNA using the following primers: tub2F_31, CCCCCACATTCCCCAGAGTTAGC; tub2R_1459, AAAGCCAACCCCGCACCAACTG; rip1F_15, CAAAGAGAAGGGAATGGCCGAGATAAC; rip1R_966, GCAGCAGCAGATCATGATGTGTCGTT. For PCR, the AccuTaq system from Sigma was used following the manufacturer's suggestions. The PCR products were cloned using a TA-cloning kit after "A-tailing" according to the manufacturer's directions (pGemT-Easy, Promega) and the plasmids sequenced to confirm the identity of the insert. Clone pMAZ87, containing a portion of the *actin* gene (Moniz De Sa and Drouin 1996), was

provided by Dr. G. Drouin, University of Ottawa, and the insert amplified using M13F and M13R standard primers.

All of the new probes were labeled with TexasRed-dCTP (red) and tested on somatic chromosomes from the KYS line. Each probe was applied in combination with two other probes, CentC and "TAG," that allow all of the chromosomes to be identified in order to confirm the position of the signal on chromosome 8. Several of the probes produced multiple signals: *atp2*, 6L and 8L; *cesa-1*, 1S, 6L, 8L, 9L; *actin*, 8L, and several others. In the case of *cesa-1* (Holland et al. 2000), homologues with strong sequence similarity are present in the maize genome database. Many copies of the actin gene are present in the maize genome (Moniz De Sa and Drouin 1996) that map to various locations (information from maps at www.maizegdb.org), consistent with the FISH data. The presence of only a single signal from the *tub2* gene is somewhat surprising given that tubulin genes have been mapped to various locations (information from maps at www.maizegdb.org).

Results

Chromosome 8 from knobless Tama flint has two sites of centromeric elements

Mitotic chromosome spreads from many different maize lines were hybridized with a collection of fluorescently labeled repetitive elements that allows each chromosome to be identified (Kato et al. 2004). The repetitive element cocktail included CentC, the maize centromeric satellite. In the KTF line, two signals corresponding to CentC, one at the primary constriction and one on a chromosome arm, were observed on one homologous chromosome pair (Figure 1). The maize centromeric retrotransposon,

CRM, was also present at both of the CentC sites (Figure 1). The "dicentric" chromosome variant was homozygous in all individuals from the initial KTF sample we received from the Maize Genetics Cooperation stock center at the University of Illinois. We obtained additional samples from previous rounds of propagation and found that they were also homozygous for the chromosome 8 variant. However, in a sampling of 30 individuals from the presumed progenitor line, Tama flint, no cases of dicentric chromosomes were found.

Because it was unusual, it was not possible to identify the "dicentric" chromosome using the standard karyotyping cocktail. Therefore, FISH probes that label individual genetic locations were applied one at a time to the KTF line to identify the unusual chromosome. The BAC8L-AC157487 probe, designed from a BAC sequence located on chromosome 8L (accession #: AC157487), hybridized to the "dicentric" chromosome (Figure 2).

The centromere at the new cytological position is functional and the centromere at the original position is inactive

Although two sites containing centromere repeats are present on the KTF chromosome 8, only a single chromosome constriction is present, indicating that the chromosome is functionally monocentric (Figure 2). To compare the positions of the centromeres in KTF to those commonly found in maize lines, a hybrid was made with a standard maize line, Mo17. Mo17 was chosen because it has an easily discernable cytological landmark on chromosome 8L, a large heterochromatic block or knob (Kato et al. 2004). None of the other chromosomes in Mo17 or KTF have large knobs (Adawy et

al. 2004; Lamb et al. 2006c) although many have small clusters of the 180 bp satellite that composes maize knobs (Peacock et al. 1981). These small clusters are only detectable using sensitive FISH techniques (Adawy et al. 2004; Lamb et al. 2006c). In somatic metaphase, it is apparent that the arm length ratios are different between the Mo17 and KTF lines (Figure 2). In standard maize lines, the long arm of chromosome 8 is much longer than the short arm (Anderson et al. 2003; Longley 1939) but the primary constriction in KTF chromosome 8 was located at the more central of the two sites of centromeric elements. Indirect immuno-fluorescence using antibodies against the maize centromere H3 histone protein only produced signal at the CentC site corresponding to the primary constriction (Figure 3). The second, larger site of centromeric elements was located at a similar position relative to the end of the short arm as the regular chromosome 8 of the Mo17 line.

During meiotic prophase, the dicentric KTF chromosome synapses with the Mo17 chromosome. This association was consistent, and in over one hundred pachytene chromosome spreads, pairing was complete. No cases of pairing with other chromosomes were observed, nor were cases of incomplete synapsis in the region between the two CentC clusters at either early or late pachytene. The Mo17 centromere was positioned adjacent to the non centromeric KTF CentC site producing one to three distinct spots of hybridization (Figure 2, 3). The functional KTF centromere site was located ~20 centiMcClintocks (20.53, +/- 3.11, n=14) from the site of the paired Mo17 active/KTF inactive centromeres. A centiMcClintock (cMc) is a term of cytological measurement equivalent to one percent of the arm length (Lawrence et al. 2005). Indirect immuno-fluorescence using CenH3 antibodies at pachynema produced signal at both the Mo17

centromere and the KTF centromere 20 cMc away (Figure 3). Thus, the KTF line contains a cluster of centromeric elements at the original cytological position that is not active and a second cluster at a new cytological position on the long arm relative to a standard line that corresponds to the functional KTF centromere.

In both somatic and meiotic preparations, the CenH3 signal did not usually colocalize with the KTF CentC signal. Instead, the two signals were located adjacent to each other. Near the Mo17 centromere in the hybrid, CenH3 and CentC always colocalized at only one of the CentC foci. As seen at the location of the paired Mo17 active/KTF inactive centromeres, multiple CentC foci were observed at other centromeres during pachynema. Although CenH3/CentC colocalization was frequent, other centromeres were observed (in this hybrid as well as other lines) where CentC and CenH3 signals were adjacent.

The region between the two centromere repeat sites in KTF is inverted

An explanation for the presence of two centromeric regions is that an inversion occurred with one breakpoint in the long arm and the second breakpoint in the centromeric region. We have termed this type of rearrangement a "hemicentric" inversion because rather than including the centromere (pericentric) or being beyond the centromere (paracentric), it divides the centromere in two. To test this hypothesis, we developed several additional FISH probes that detect genes located on 8L and applied them to chromosome preparations from the Mo17 x KTF hybrid. A hemicentric rearrangement would invert the gene order along the chromosome between the two breakpoints but leave the gene order for the remainder of the chromosome unaltered. In

an inversion heterozygote where the chromosome is completely synapsed, the genes on different homologs would be separated from one another.

Consistent with the outcome expected for a hemicentric inversion heterozygote, probes located distal to the new centromere produced pairs of FISH signals, one corresponding to each homolog (Figure 4). The *atp2* probe was located between the two centromeres and usually appeared as two separated signals (Figure 4). However, in some images the *atp2* signals were close together and appeared paired. Because the signals were located nearly equidistant from both centromeres, the *atp2* genes may be inverted relative to each other but still be positioned adjacent to each other in the synapsed chromosome. The *cesa-1* probe provided strong evidence for an inversion. In root tips chromosomes, it produced a signal located away from the centromere on the long arm in the standard line. In KTF, the signal was located adjacent to the inactive centromere (Figure 4). FISH on pachytene chromosomes from the hybrid produced a signal on each paired homolog opposite the respective functional centromere (Figure 4). In the Tama flint line, the presumed progenitor of KTF, *cesa-1* hybridizes as it does in Mo17, away from the centromere indicating that the inversion is not present in that line (Figure 4).

The inverted region is composed of pericentromeric heterochromatin

As in most eukaryotes, the regions flanking maize centromeres are highly heterochromatic. In maize, certain families of repetitive elements are highly enriched in the proximal heterochromatin compared to the remainder of the genome (Lamb et al. 2006c). One clone, p6-9-11, which shows a particularly strong pericentromeric enrichment (Kato et al. 2004), was used as a FISH probe to determine the

hemicentromeric inversion breakpoint in the long arm relative to the proximal heterochromatin. In somatic chromosome spreads, an element that is enriched away from the centromere, Prem2, was included to provide additional contrast and more clearly define the boundary of the proximal heterochromatin (Lamb et al. 2006c). The KTF CentC signal was located at the edge of the 6-9-11 signal (Figure 5).

Centromere orientation and chromosome disjunction in meiosis I are normal in the hemizygous inversion heterozygote

The mechanism that ensures proper orientation of homologous centromeres leading to movement to opposite poles is not known. Because of the inversion, the functional centromeres are not properly aligned during pachynema in the hybrid. This situation allows the role of pairing in centromere orientation to be tested. The presence of the large knob in Mo17 and the extra centromere in KTF allowed chromosome 8 to be easily identified at anaphase I when the centromeres begin moving to their respective poles. In all 65 anaphase cells examined, the Mo17 and KTF functional centromeres moved in opposite directions (Figure 6). In early telophase I, it is possible to identify chromosome 8 when the chromosomes are sufficiently spread. In all such cases, the two homologues were oriented in opposite directions (Figure 6).

Male and female flowers were examined for gametophytic deficiencies that would result from missegregation. Pollen from the hybrid was examined with a handheld microscope in the field. Pollen was well shaped with no elevated levels of abortion. Ears grown in greenhouse conditions showed full, or nearly full, seed set.

Dramatically reduced recombination in the Mo17 x KTF hybrid

The recombination rate on 8L in the hybrid was measured by observing the behavior of the 8L knob at telophase I. Because the two lines are polymorphic for the presence of the large knob, a cross over proximal to the knob will separate the knob containing sister chromatids and link one of the knobs to the KTF centromere. At telophase, knobs will be observed moving toward both poles (Figure 6) and at the end of meiosis I, knob signal would be observed in both daughter cells. Split knob signal is seen in 26.4% of cells at telophase I (n=242). This percentage corresponds to a genetic map distance of ~13 cM. The knob is located ~75 cMc (75.4 +/- 2.7, n=14) from the Mo17 centromere. This cytological distance corresponds to a genetic position of 88-91 cM on the maize UMC98 genetic map. A conversion of cytological to genetic distance places the site of long arm inversion breakpoint at 37-41 cM. Subtracting the map position of the new centromere from the knob provides an estimate of the genetic distance between the two positions, 47-54 cM, indicating that the two positions should be transmitted as if they were unlinked. Using the map distances from a second genetic map available at the Morgan2McClintock converter (the classic maize genetic map) produces a slightly lower estimage of map distance, 35-42 cM. In either case, the actual genetic distance, 10 cM, and the predicted distance are quite different.

Although the homologues are fully synapsed, this synapsis is nonhomologous and does not form an inversion loop. Recombination is not expected to occur in the inverted region. Should recombination occur, it would be possible to visualize it cytologically because strand exchange would link two functional centromeres by a single chromatid. When the centromeres separate, a chromatin bridge would be produced at telophase I.

Among all the telophase figures examined, no cases of chromatin bridges involving chromosome 8 were observed (n=100). Thus, recombination is repressed in the inversion heterozygote, both within and beyond the inversion.

Discussion

Centromeric elements at sites distinct from the kinetochore

The presence of a hemicentric inversion creates an unusual situation: two sites of centromeric elements on a single chromosome. The role of centromeric elements in centromere formation is not fully understood and somewhat paradoxical (Henikoff et al. 2001). Centromeric sequences are neither sufficient nor necessary for centromere formation, but the DNA that underlies the centromeres of many eukaryotes is composed mainly of DNA elements that are not found elsewhere in the genome. Thus, while not strictly required, centromeric elements may be optimal for centromere stability (Lamb and Birchler 2003; Lamb et al. 2004).

In maize, the ~150 bp CentC satellite (Ananiev et al. 1998) is present at centromeres in tandem arrays intermixed with members of the centromere specific CRM family of retroelements (Nagaki et al. 2003; Zhong et al. 2002). As in humans (Spence et al. 2002), rice (Nagaki et al. 2004), Arabidopsis (Shibata and Murata 2004), and other model organisms, only a subset of the centromere elements (Zhong et al. 2002), often those near the edge of the repeat arrays (Jin et al. 2004), associate with CenH3 and form the kinetochore. Thus, most centromeric satellites do not underlie the kinetochore but compose the flanking region.

The factors that determine where CenH3 is located are unknown although it is clearly not dependent solely on the presence of centromeric satellites. We observed that in the KTF chromosome 8, CenH3 was located adjacent to, not at, the CentC site of the functional centromere. In other chromosomes as well, including KTF and other lines not included in this study, the CentC and CenH3 also do not colocalize (unpublished observation). The greater number of CentC copies at the Mo17 chromosome 8 centromere were organized into at least two distinct CentC clusters, one of which fully contained the CenH3 binding site. Thus, while CenH3 is not always associated with CentC, it is always located near it. The apparent repositioning of the KTF CenH3 binding domain out of the CentC array may have correctly positioned the centromeric elements for a centromeric role other than kinetochore formation; perhaps sister chromatid cohesion.

The presence of multiple CentC clusters is evident at many centromeres. This has been previously reported for chromosome 10 (Wang et al. 2006) and observed in this study. Perhaps very small hemicentric inversions are quite frequent as they could generate the multiple clusters. Clusters of centromeric elements are not found far away from the primary constriction except on the maize B chromosome (Lamb et al. 2005) and in the relatively new rearrangement in the KTF line. In contrast, other satellites such as the 180 bp knob repeat, the TR-1 knob repeat, and the "TAG" microsatellite are found at numerous locations on maize chromosomes (Kato et al. 2004; Lamb et al. 2006c). These observations suggest that tracts of centromeric elements are probably not stable at extra-centromeric positions. In mammalian cells, extra-centromeric clusters of centromere repeats can cause delayed sister chromatid separation and nondisjunction (Haaf et al.

1992; Warburton and Cooke 1997) and transgenic arrays of centromere repeats can acquire centromere function (Nakano et al. 2003). Although this was not observed in the case of KTF, the number of cells examined my not have been sufficient to detect such irregularities. Even a small increase in chromosomal instability would be strongly selected against in natural populations.

Synapsis and recombination in the Mo17 x KTF hybrid.

Complete synapsis of the inverted segment in KTF likely reflects the gene poor nature of the region involved. Similar synapsis of nonhomologous sequences resulting from inversions around centromeres has been previously reported in maize (McClintock 1933) and other species (Ashley et al. 1981). In maize, regions between genes are highly variable among lines and are composed of nested retrotransposons and other degenerated repeats (Fu and Dooner 2002; SanMiguel et al. 1996; SanMiguel et al. 1998). The region surrounding maize centromeres is gene poor and experiences very little recombination (Anderson et al. 2006). Furthermore, there are several families of retroelements which are very abundant in the gene poor portion of the chromosome around the centromere (Lamb et al. 2006c). Among Arabidopsis relatives, which have relatively gene-rich genomes, the intergenic regions in the proximal heterochromatin are highly variable for retroelements and other repetitive elements (Hall et al. 2006). Therefore, in the chromosome regions that experience limited recombination, the sequences are highly variable among lines. Even in hybrids that do not have chromosomal rearrangements, synapsis will place nonhomologous sequences adjacent to one another. Because the inversion is limited to

the proximal heterochromatin, synapsis in the inversion heterozygote probably proceeds much as it would in a hybrid without an inversion.

The lack of instability during meiosis in the inversion heterozygote is unusual. However, most genome instability in rearrangement heterozygotes is dependent on recombination and our results show that recombination is severely repressed along the hybrid chromosome, even extending beyond the breakpoints. Repression of recombination in and near inversions has previously been attributed to decreased chromosome association (synapsis). Concurrent with a decrease in the inverted region, recombination frequency distal to inversion breakpoints is increased in inversion heterozygotes (Brown et al. 1998; Luchessi 1976; Zetka and Rose 1992). In this case, complete synapsis was observed at early pachytene and recombination distal to the breakpoint was severely repressed. The difference is probably due to the involvement of the centromere in the inversion.

The "centromere effect," or repression of recombination in the vicinity of centromeres, is well documented for many species (Choo 1998). The lack of increased recombination in the distal part of chromosome 8 in the Mo17 x KTF hybrid may be due to repression by the KTF centromere extending into the long arm beyond the inversion breakpoint. This hypothesis will be testable by examining the recombination frequency in an inversion homozygote that is heterozygous for the knob or other genetic markers. Increased recombination in the short arm adjacent to the inactive centromere and decreased recombination near the new centromere site is predicted. Whatever the cause, in the hemicentric inversion heterozygote, recombination along most of chromosome 8 is eliminated.

Potential role for hemicentric inversions in karyotype evolution

Relocation of centromeres via other mechanisms, i.e. neocentromere formation, is well documented in many evolutionary lineages (Eder et al. 2003). Hemicentric inversions can also change the genetic position of centromeres. An inversion with a breakpoint in or very close to the centromere will alter the chromosome region that is subjected to recombination repression. This type of rearrangement could fix allele combinations for the region rendering them unable to recombine and increasing the genetic distance between other genes. However, for many rearrangements, the heterozygous condition results in decreased fertility. When rearrangements, such as the one described in this work, do not affect fertility, there is no reduced fitness for individuals that carry the alteration in either the heterozygous or homozygous condition. These types of rearrangements could create variation not only for particular allele combinations but for the frequency of recombination through alterations to genetic linkages.

An example of how alterations to recombination frequency can be selected for in populations is found in a classic study of natural populations of *Drosophila pseudoobscura* (Dobzhansky and Epling 1948). In some populations, an inversion for chromosome 3 was found to be highly prevalent and the number of heterozygous individuals exceeded the number of those homozygous for either chromosome variation. It was demonstrated that lack of recombination between the two chromosome types increased the likelihood of heterozygosity at multiple loci on chromosome 3, reducing the effects of inbreeding depression in these populations (Dobzhansky and Epling 1948).

For genes placed in the pericentromeric heterochromatin, lack of recombination prevents the shuffling of alleles and reduces potential variation. This could generate adaptive gene complexes; however, it could also prevent adaptations to new environmental pressures. For two linked pericentromeric genes, restoration of independent assortment would be accomplished by chromosome cleavage between the two loci that results in two separate chromosomes. However, in order to successfully form two chromosomes from one, a new centromere needs to be generated. In human cells, large clones containing centromeric elements that had been transgenically introduced to chromosomes were able to spontaneously form centromeres, splitting the chromosome into two pieces (Nakano et al. 2003). Also, in Drosophila, repeats placed next to centromeres were able to acquire centromere activity when the chromosome was broken using X-rays (Maggert and Karpen 2001). Therefore, a hemicentric inversion, especially a small one that results in centromere repeat cluster in proximity to the functional centromere, may predispose the chromosome to splitting by providing a hotspot for centromere formation.

Acknowledgements

We thank Marty Sachs (Maize Genetics COOP Stock Center) and Mark Milard (North Central Regional Plant Introduction Station) for their help in tracking the pedigree of the KTF line. This work was supported by grants from the National Science Foundation Plant Genome Initiative (DBI0421671 and DBI0423898).

References

Adawy SS, Stupar RM, Jiang J (2004) Fluorescence in situ hybridization analysis reveals multiple loci of knob-associated DNA elements in one-knob and knobless maize lines. J Histochem Cytochem 52: 1113-1116

Ananiev EV, Phillips RL, Rines HW (1998) Chromosome-specific molecular organization of maize (*Zea mays* L.) centromeric regions. Proc Natl Acad Sci U S A 95: 13073-13078

Anderson LK, Doyle GG, Brigham B, Carter J, Hooker KD, Lai A, Rice M, Stack SM (2003) High-resolution crossover maps for each bivalent of *Zea mays* using recombination nodules. Genetics 165: 849-865

Anderson LK, Lai A, Stack SM, Rizzon C, Gaut BS (2006) Uneven distribution of expressed sequence tag loci on maize pachytene chromosomes. Genome Res 16: 115-122

Ashley T, Moses MJ, Solari AJ (1981) Fine structure and behaviour of a pericentric inversion in the sand rat, *Psammomys obesus*. J Cell Sci 50: 105-119

Brown GM, Leversha M, Hulten M, Ferguson-Smith MA, Affara NA, Furlong RA (1998) Genetic analysis of meiotic recombination in humans by use of sperm typing: reduced recombination within a heterozygous paracentric inversion of chromosome 9q32-q34.3. Am J Hum Genet 62: 1484-1492

Choo KH (1998) Why is the centromere so cold? Genome Res 8: 81-82

Dobzhansky T, Epling C (1948) The suppression of crossing over in inversion heterozygotes of *Drosophila pseudoobscura*. Proc Natl Acad Sci U S A 34: 137-141

Eder V, Ventura M, Ianigro M, Teti M, Rocchi M, Archidiacono N (2003) Chromosome 6 phylogeny in primates and centromere repositioning. Mol Biol Evol 20: 1506-1512

Fu H, Dooner HK (2002) Intraspecific violation of genetic colinearity and its implications in maize. Proc Natl Acad Sci U S A 99: 9573-9578

Haaf T, Warburton PE, Willard HF (1992) Integration of human alpha-satellite DNA into simian chromosomes: centromere protein binding and disruption of normal chromosome segregation. Cell 70: 681-696

Hall AE, Kettler GC, Preuss D (2006) Dynamic evolution at pericentromeres. Genome Res 16: 355-364

Han F, Lamb JC, Birchler JA (2006) High frequency of centromere inactivation resulting in stable dicentric chromosomes of maize. Proc Natl Acad Sci U S A 103: 3238-3243

Henikoff S, Ahmad K, Malik HS (2001) The centromere paradox: stable inheritance with rapidly evolving DNA. Science 293: 1098-1102

Holland N, Holland D, Helentjaris T, Dhugga KS, Xoconostle-Cazares B, Delmer DP (2000) A comparative analysis of the plant cellulose synthase (CesA) gene family. Plant Physiol 123: 1313-1324

Jin W, Melo JR, Nagaki K, Talbert PB, Henikoff S, Dawe RK, Jiang J (2004) Maize centromeres: organization and functional adaptation in the genetic background of oat. Plant Cell 16: 571-581

Kato A, Albert PS, Vega JM, Birchler JA (2006) Sensitive FISH signal detection in maize using directly labeled probes produced by high concentration DNA polymerase nick translation. Biotechnic and Histochemistry 81: 71-78

Kato A, Lamb JC, Birchler JA (2004) Chromosome painting using repetitive DNA sequences as probes for somatic chromosome identification in maize. Proc Natl Acad Sci U S A 101: 13554-13559

Lamb JC, Birchler JA (2003) The role of DNA sequence in centromere formation. Genome Biol 4: 214

Lamb JC, Birchler JA (2006) Retroelement genome painting: cytological visualization of retroelement expansions in the genera *Zea* and *Tripsacum*. Genetics 173: 1007-1021

Lamb JC, Danilova T, Bauer MJ, Meyer JM, Holland JJ, Jensen MD, Birchler JA (2006a) Single loci detection and karyotyping using small target FISH on maize somatic chromosomes. **Submitted**:

Lamb JC, Kato A, Birchler JA (2005) Sequences associated with A chromosome centromeres are present throughout the maize B chromosome. Chromosoma 113: 337-349

Lamb JC, Kato A, Yu W, Han F, Albert PS, Birchler JA (2006b) Cytogenetics and chromosome analytical techniques. In: (ed) Floriculture, ornamental and plant biotechnology. Global Science Books, U.K., pp 244-248

Lamb JC, Meyer JM, Corcoran B, Kato A, Han F, Birchler JA (2006c) Distinct chromosomal distributions of highly repetitive sequences in maize. Chromosome Res In press:

Lamb JC, Theuri J, Birchler JA (2004) What's in a centromere? Genome Biol 5: 239

Lawrence CJ, Seigfried TE, Bass HW, Anderson LK (2005) Predicting chromosomal locations of genetically mapped loci in maize using the Morgan2McClintock translator. Genetics 172: 2007-2009

Longley AE (1939) Knob positions on corn chromosomes. J Agric Res 59: 475-490

Luchessi JC (1976) Interchromosomal effects. In: Ashburner, M and Novitski, E (ed) The genetics and biology of Drosophila. Academic Press, London, pp 315-329

Maggert KA, Karpen GH (2001) The activation of a neocentromere in Drosophila requires proximity to an endogenous centromere. Genetics 158: 1615-1628

McClintock B (1933) The association of non-homologous parts of chromosomes in the mid-prophase of meiosis in *Zea mays*. Z Zellforsch Mikrosk Anat 19: 191-237

Moniz de Sa M, Drouin G (1996) Phylogeny and substitution rates of angiosperm actin genes. Mol Biol Evol 13: 1198-1212

Murphy TD, Karpen GH (1998) Centromeres take flight: alpha satellite and the quest for the human centromere. Cell 93: 317-320

Nagaki K, Cheng Z, Ouyang S, Talbert PB, Kim M, Jones KM, Henikoff S, Buell CR, Jiang J (2004) Sequencing of a rice centromere uncovers active genes. Nat Genet 36: 138-145

Nagaki K, Song J, Stupar RM, Parokonny AS, Yuan Q, Ouyang S, Liu J, Hsiao J, Jones KM, Dawe RK, Buell CR, Jiang J (2003) Molecular and cytological analyses of large tracks of centromeric DNA reveal the structure and evolutionary dynamics of maize centromeres. Genetics 163: 759-770

Nakano M, Okamoto Y, Ohzeki J, Masumoto H (2003) Epigenetic assembly of centromeric chromatin at ectopic alpha-satellite sites on human chromosomes. J Cell Sci 116: 4021-4034

Peacock WJ, Dennis ES, Rhoades MM, Pryor AJ (1981) Highly repeated DNA sequence limited to knob heterochromatin in maize. Proc Natl Acad Sci U S A 78: 4490-4494

Reeves A, Tear J, 2000 MicroMeasure for Windows, version 3.3. (http://www.colostate.edu/Depts/Biology/MicroMeasure), pp.

SanMiguel P, Tikhonov A, Jin YK, Motchoulskaia N, Zakharov D, Melake-Berhan A, Springer PS, Edwards KJ, Lee M, Avramova Z, Bennetzen JL (1996) Nested retrotransposons in the intergenic regions of the maize genome. Science 274: 765-768

SanMiguel P, Gaut BS, Tikhonov A, Nakajima Y, Bennetzen JL (1998) The paleontology of intergene retrotransposons of maize. Nat Genet 20: 43-45

Shibata F, Murata M (2004) Differential localization of the centromere-specific proteins in the major centromeric satellite of Arabidopsis thaliana. J Cell Sci 117: 2963-70:

Spence JM, Critcher R, Ebersole TA, Valdivia MM, Earnshaw WC, Fukagawa T, Farr CJ (2002) Co-localization of centromere activity, proteins and topoisomerase II within a subdomain of the major human X alpha-satellite array. Embo J 21: 5269-5280

Wang CJ, Harper L, Cande WZ (2006) High-resolution single-copy gene fluorescence in situ hybridization and its use in the construction of a cytogenetic map of maize chromosome 9. Plant Cell 18: 529-544

Warburton PE, Cooke HJ (1997) Hamster chromosomes containing amplified human alpha-satellite DNA show delayed sister chromatid separation in the absence of de novo kinetochore formation. Chromosoma 106: 149-159

Yu W, Lamb JC, Han F, Birchler JA (2006) Cytological visualization of transposable elements and their transposition pattern in somatic cells of maize. Genetics in press:

Zetka MC, Rose AM (1992) The meiotic behavior of an inversion in *Caenorhabditis elegans*. Genetics 131: 321-332

Zhong CX, Marshall JB, Topp C, Mroczek R, Kato A, Nagaki K, Birchler JA, Jiang J, Dawe RK (2002) Centromeric retroelements and satellites interact with maize kinetochore protein CENH3. Plant Cell 14: 2825-2836

Figures

Figure 1: The knobless Tama flint line contains a chromosome with two sites of centromeric elements.

Chromosome spreads from the KTF line were hybridized with the two maize centromeric retrotransposon, CRM (red), and the centromeric satellite, CentC (green). One pair of chromosomes has two distinct sites of hybridization. The inset on the lower left shows the chromosome indicated by the arrow with the individual CRM and CentC channels displayed.

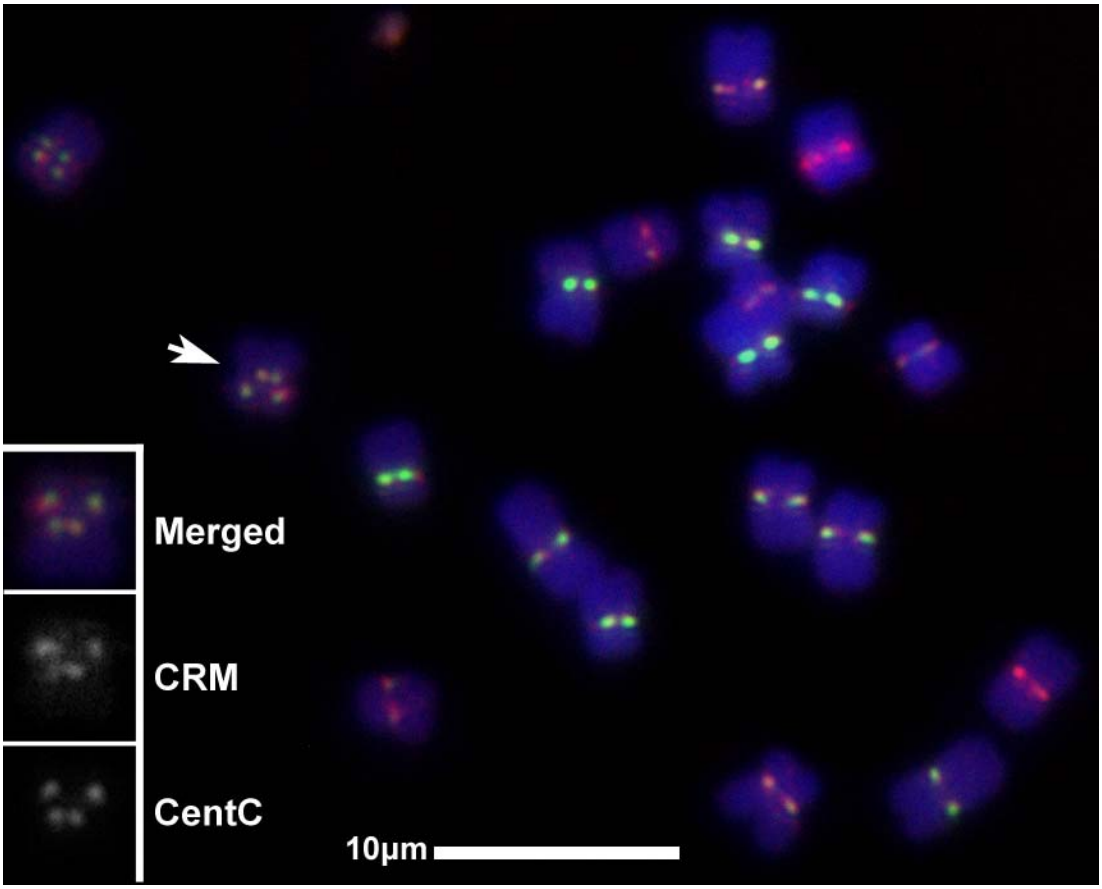


Figure 2. A new site of centromere activity on KTF chromosome 8.

A chromosome spread from a Mo17 x KTF hybrid was labeled with the BAC8L-AC157487 probe (red) that identifies chromosome 8 and CentC (green) (A). The two chromosomes 8 were cut out and placed next to each other (inset). Stars indicate the position of the primary constrictions where the kinetochore forms. At meiotic prophase, the homologous chromosomes from Mo17 and KTF fully synapse (B). Chromosome 8, labeled with the 180 bp knob repeat (red) and CentC (green) is displayed. The site of the active centromere from Mo17 (m) and KTF (k) are labeled. The inactive KTF centromere is positioned adjacent to the Mo17 centromere. In many cells, there are two or three distinct spots of hybridization of CentC near the Mo17 centromere. At several other centromeres, multiple CentC signals can also be observed.

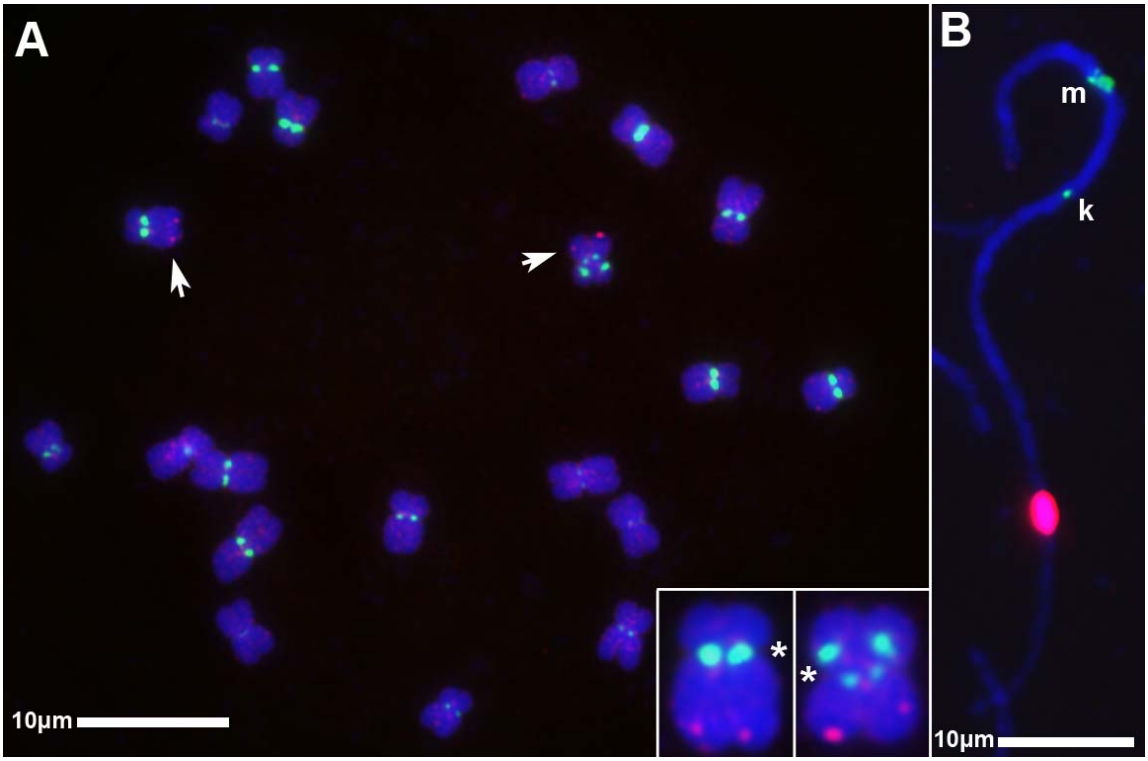


Figure 3. The new centromere site associates with CenH3.

Somatic KTF chromosomes (A) and meiotic preparations from the Mo17 x KTF hybrid (B) were labeled with antibodies against the centromere histone H3 variant, CenH3 (red) and CentC (green). In both (A) and (B), a single optical section is shown. No CenH3 labeling is observed near the inactive KTF centromere in somatic preparations. CenH3 is located at the primary constriction corresponding to the new position of CentC in KTF (arrows). However, CentC and CenH3 signals do colocalize at chromosome 8 or at some other chromosomes. In the paired chromosomes of the hybrid, CenH3 is found at both the KTF (k) and Mo17 (m) centromeres. At the site of the Mo17 centromere, CenH3 colocalizes with one site of CentC but not the other two.

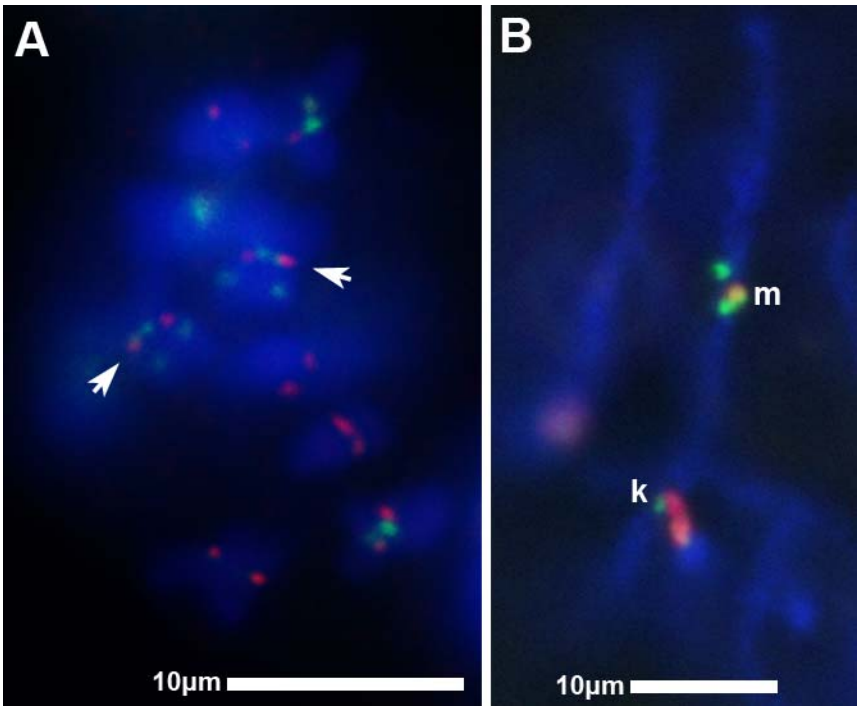


Figure 4. The region between the two KTF centromere element sites is inverted in relation to standard maize lines.

FISH probes were produced from genes located on 8L and applied to chromosome spreads from Mo17 x KTF hybrids. Somatic chromosome 8 pairs are shown below the synapsed meiotic chromosome pairs. (A) The *atp2* probe is located approximately halfway between the Mo17 (m) and KTF (k) centromeres. In most preparations, the two Mo17 and KTF signals are separated from each other. In some preparations (not shown) the *atp2* signals are close together as if paired. (B) The *cesa-1* probe produces two signals on the synapsed chromosomes, one close to each centromere. Somatic chromosomes show only one site per chromatid. The signal is located on the long arm away from the centromere in the Mo17 homologue. In KTF, the signal is adjacent to the inactive centromere. The *tub2* (C), *rip1* (D) and *actin* (not shown) probes are located distal to the new centromere site in KTF and produce a single pair of signals in the paired hybrid chromosomes.

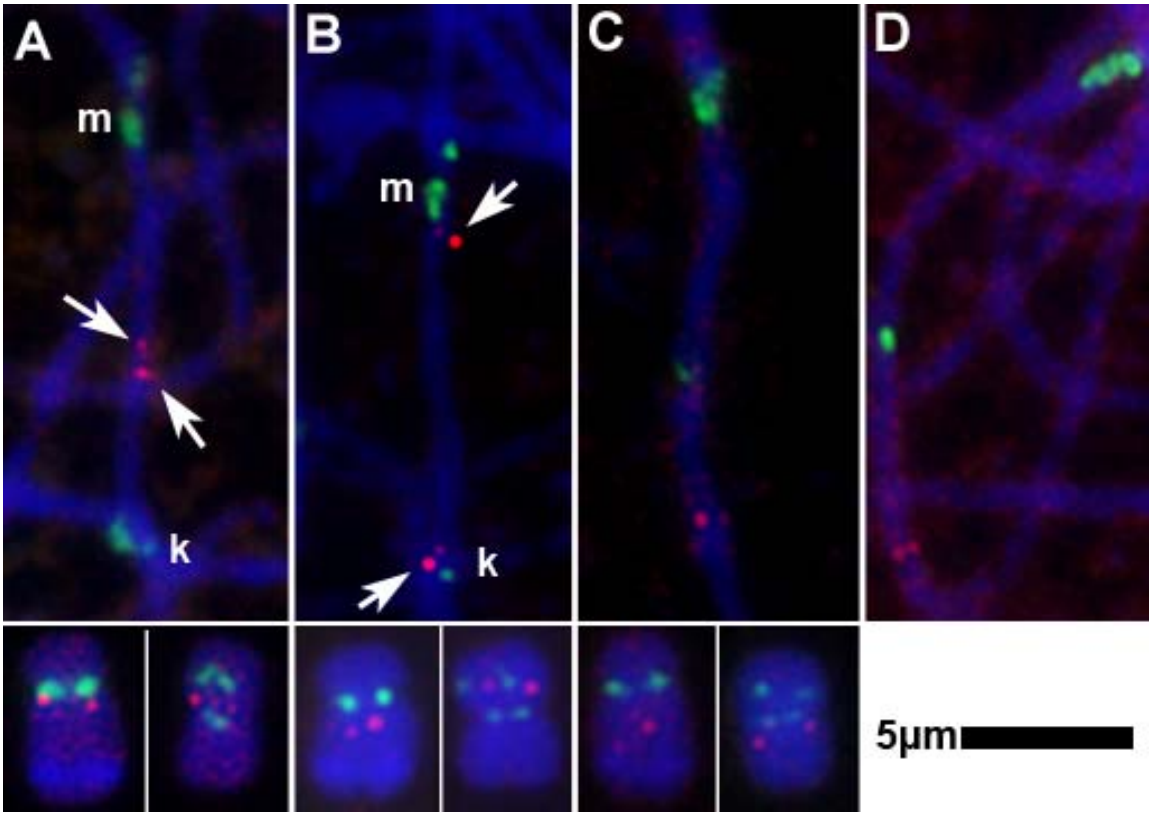


Figure 5. The inverted region consists of pericentromeric heterochromatin.

KTF somatic chromosomes are labeled with a pericentromere enriched element, 6-9-11 (red), a distally enriched retroelement, Prem2 (green), and CentC (white). The chromosome 8 indicated with the arrow is enlarged in the inset. In (1) the RGB image including the DAPI counter-stain is shown. In (2) the CentC signal has been minimized, the DAPI signal removed, and the Prem2 signal displayed as blue to aid the visualization of the position of CentC relative to 6-9-11 and Prem2. The two CentC signals on KTF chromosome 8 are contained in the region labeled with 6-9-11. (B) A synapsed meiotic prophase Mo17 x KTF chromosome pair labeled with CentC (green) and 6-9-11 (red) shows that the active KTF centromere is located at the edge of 6-9-11 signal.

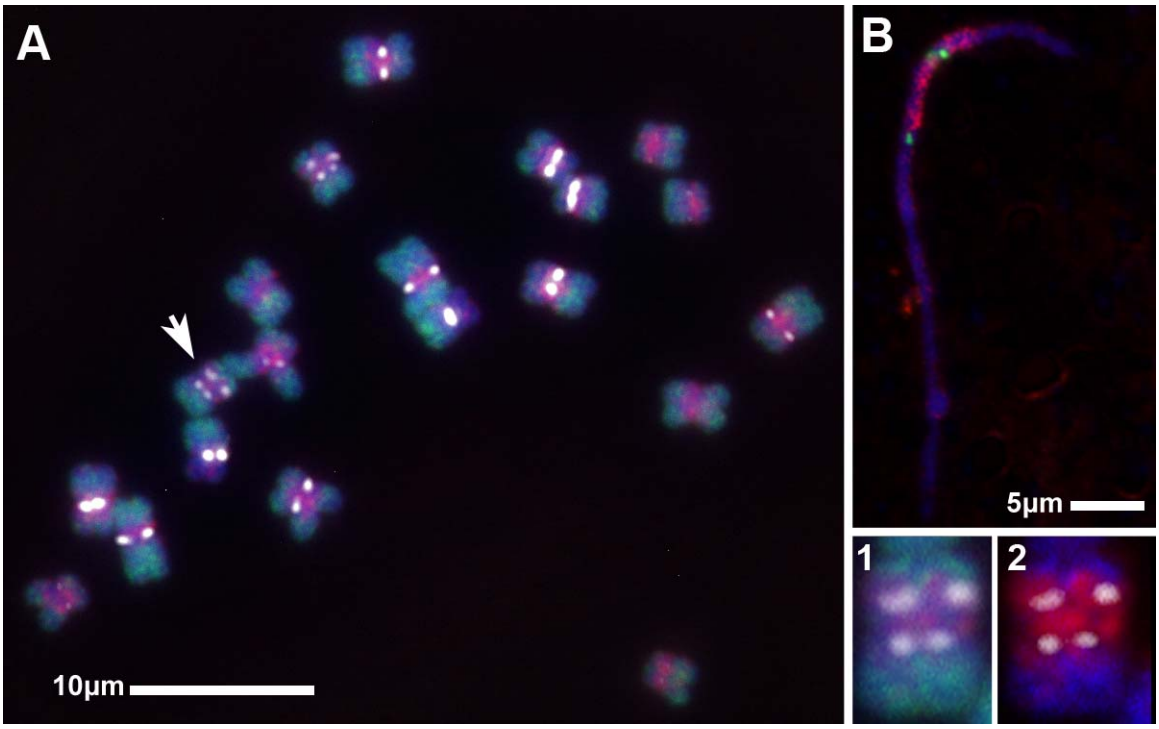
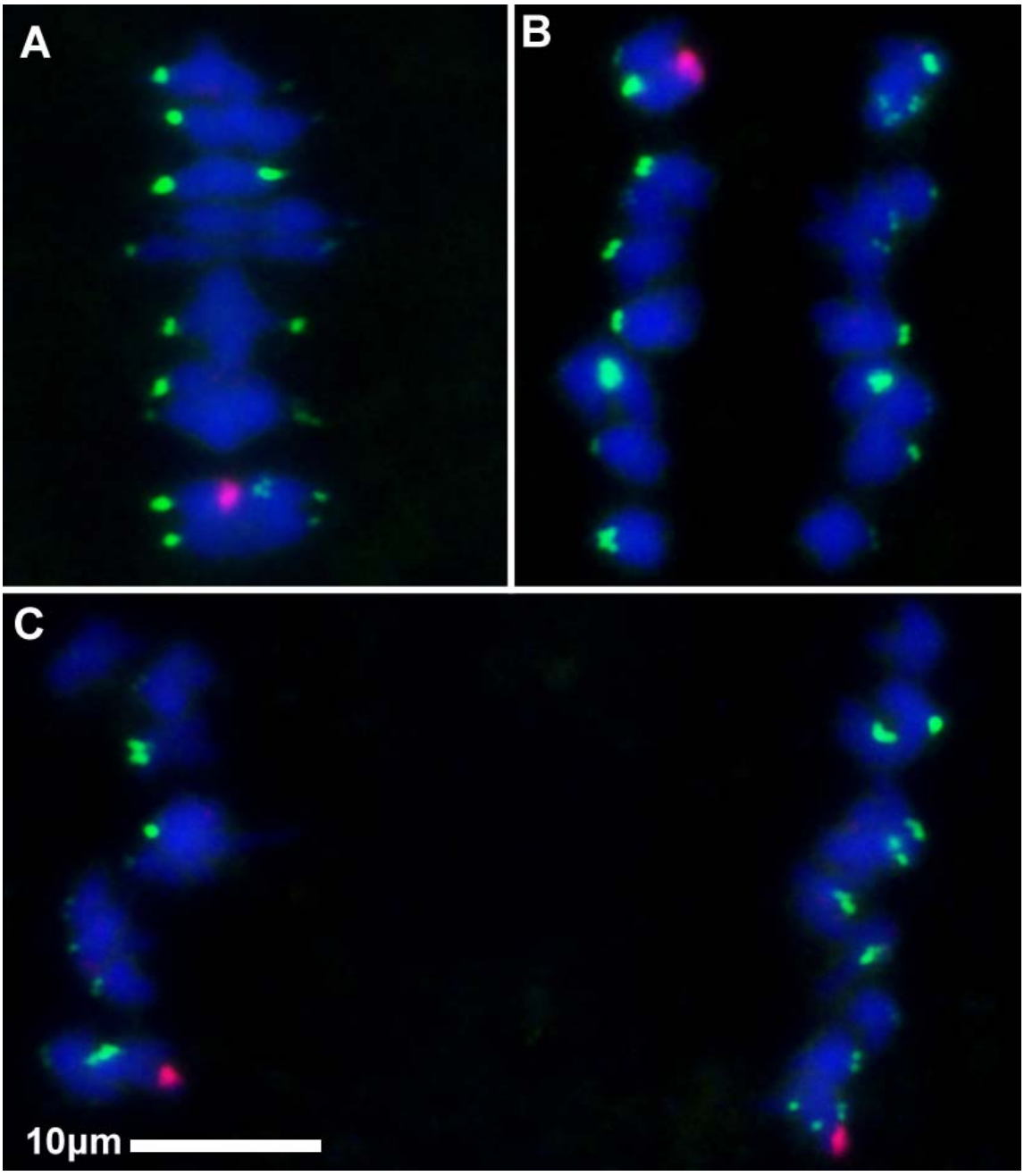


Figure 6. Proper centromere alignment and chromosome separation during meiosis I and cytological detection of recombination on 8L in the inversion heterozygote.

Meiotic cells were labeled with the 180 bp knob (red) and CentC (green). At anaphase I in the Mo17 x KTF hybrid, chromosome 8 centromeres move to opposite poles (A). Correct orientation for chromosome 8 was observed in 60 out of 60 cells examined. Proper disjunction of chromosome 8 is detected during telophase, (B) and (C). A crossover proximal to the knob on Mo17 is visualized cytologically by the presence of a bright knob signal in both groups of chromosomes in telophase I (C). Split knob signal is seen in 26.4% of cells (n=242).



8. CONCLUSIONS AND FUTURE DIRECTIONS

Although a discussion of the data is included in each chapter, I would like to summarize the major findings and discuss their significance in clarifying centromere determination and evolution. Then, I will briefly outline a number of specific questions that remain and describe ongoing research that will address these questions.

1. The size of CentC repeat clusters is extremely variable among different maize lines. However, the amount of CRM is much less variable. The lines chosen for the initial survey of repetitive elements were chosen because they are frequently used for scientific studies, not because they represent extremes in variation. However, among more diverse maize lines, and even *Z. mays parviglumis* lines that have been subsequently examined for other reasons, CentC variation is evident. The fluctuation of the copy number of CentC elements at centromeres would have a homogenizing effect on repeat diversity at individual centromeres in a similar fashion as a genetic bottleneck affects population diversity. When the overall size of the centromere contracts, repeat variants are lost and subsequent expansion must occur from the remaining variants. This process could result in the expansion of chromosome specific variants that would ultimately lead to substantially different sequences composing the centromeres of different chromosomes. Because classic recombination is absent from the centromeric region, even the centromeres of the same homologs could become dissimilar. Because the elements at different centromeres are very similar, there must be some force that acts to homogenize centromere repeats across the genome.

How does this homogenizing mechanism work? One possibility is that centromeric regions experience frequent breaks and are often repaired using the homologous end joining mechanism. At anaphase, the kinetochore region can be extremely extended by the tension applied by the microtubules. If this force produces breaks, then abundant copies of the centromeric elements would be available to serve as templates for repair. Assuming that repeats from other chromosomes could be used for repair, a constant shuttling of repeat sequences from one centromere to another would occur. One prediction of a break-repair homogenization mechanism is that inter-genome homogenization would occur predominantly at the sights of breakage.

To gain insight into the homogenization mechanisms, it would be valuable to compare variation among CentC elements located at the kinetochore and those that are located outside the kinetochore. Other labs that are participating on the centromere project have undertaken to sequence CentC repeats from different centromeres. Their approach will involve mapping BACs containing CentC copies to specific chromosomes and potentially to different functional domains within centromeres. To complement this approach, I have proposed isolating CentC elements from the B chromosome. On the B chromosome, CentC is found at the kinetochore forming region defined by the ~700kb core as well as at multiple locations away from the centromere on the long arm. A large insert library (fosmids) made from an oat +maize B addition line will be screened for clones that contain CentC. Those clones positive for CentC will be further screened for CRM and ZmBs allowing clones from the kinetochore domain to be distinguished from those found elsewhere. If the homogenization mechanism only affects the kinetochore, then CentC copies on the long arm will be more diverse than those at the B centromere.

2. In some close relatives of maize, this variation is absent and a consistently large amount of CentC is present. Although the mechanism that homogenizes centromeres is likely to be conserved across species and may in fact be a byproduct of centromere function, the maintenance of centromere repeat array size appears to be different for the different species. An interesting inverse correlation exists between the amount of CRM and the degree of CentC variation among these relatives. It would be interesting to compare the abundance of the centromeric retrotransposons to centromere satellite variation for additional species. For example, the satellite copy number variation in domesticated rice is also substantial as is the amount of CRR, the rice version of CRM (CHENG *et al.* 2002; NAGAKI *et al.* 2004). Comparisons to other rice or maize relatives would allow this potential correlation to be tested. Additional factors, especially geographic location could also be examined by comparisons of copy number variation among diverse taxa.

3. Variation exists between centromeres of *Zea* and *Tripsacum* although both contain CentC and CRM. To determine if the GISH pattern is a result of CentC variation or the presence of other unknown elements, sequencing of *Tripsacum* centromeric elements must be conducted. This effort is underway by members of the NSF plant centromere group. It will be interesting to compare the key centromere proteins that come into physical contact with the centromere DNA between these taxa as well. In any case, this result is consistent with other model organisms where centromeric elements have been shown to be changing more rapidly than predicted by a neutral model. It is not clear

that such models are applicable to long arrays of satellites, but few compelling explanations have been put forward by the scientific community to explain this observation.

One model of centromere evolution posits that competition between centromeres for inclusion in the female egg is the driving force behind centromere evolution (HENIKOFF and MALIK 2002). In the female, only one of the four products of meiosis is transmitted to the next generation. The cell divisions are regular, and linear in most species, and the cell that will become the egg is positionally determined. In plants, a centromere that is positioned at the base of the divided female meiocytes will transmit. This model assumes that the DNA sequence that underlies the centromere can influence its position through meiosis.

To test this model, I have crossed plants containing different variants of a chromosome that have centromeres which vary by size and amount of the kinetochore protein, CenH3. The centromeres' differences have resulted from misdivision and likely represent more extreme variation than would be normally found in natural populations. In the hybrid, the two chromosomes with different centromeres pair. After crossing with a tester line, the ratio of the variants among the progeny will be tested. Significant deviation from 1:1 segregation would indicate that one centromere type is not transmitting as well as the other.

4. Centromeric elements are not sufficient for centromere function.

In mammalian systems, the recovery of functionally monocentric chromosomes that result from Robertsonian translocations provides compelling evidence that the mere

presence of the DNA elements that typically underlie centromeres does not recruit kinetochore proteins or cause kinetochore activity (WARBURTON *et al.* 1997). Before the work described in this thesis, no similar observations had been made for any plant species. I have shown several cases where centromeric DNA was stably maintained on a chromosome without kinetochore activity; on the B chromosome long arm and in inactivated B centromere regions, and in the KTF hemicentric inversion chromosome. Furthermore, in the case of the dicentric minichromosomes, centromere inactivation was shown to be a frequent outcome of the BFB cycle.

5. Kinetochores can form over non-centromeric sequences in maize. In the KTF line and in the centromere misdivision deriviative, telo3-3(-), the site of CenH3 localization was adjacent to the CentC cluster. In many regular chromosomes CenH3 signal is not coincident (either partially together or completely separate) with CentC. Recently, other researchers have reported the formation of a functional centromere in plants without centromere repeats (NASUDA *et al.* 2005). These observations provide additional similarities to other model organisms and suggest that centromere identity is determined by common mechanism(s).

Early work with mammalian cells showed that transgenic arrays of human centromeric repeat would cause delayed sister chromatid separation but not kinetochore activity (WARBURTON and COOKE 1997). This finding raised the possibility that the primary function of centromeric repeats is in sister chromatid cohesion. A critical role for pericentromeric heterochromatin in proper cohesion and segregation has been demonstrated in several model organisms and the centromeric repeats compose most if

not all of the region associated with plant cohesins at metaphase (Shibata and Murata 2004; Zhang et al. 2005). The heterochromatic repeats have also been proposed to play a role in preventing merotely (attachment of microtubules from both poles to a single centromere), providing a platform for kinetochore assembly and acting as a "trigger" for kinetochore assembly (PIDOUX and ALLSHIRE 2005). However, since ChIP experiments have demonstrated that many of the centromeric repeat copies are associated with CenH3, most discussions of centromere repeats focus on their potential role in forming the kinetochore.

I propose that centromeric elements are optimized for centromere function. However, that optimization is not directly for association with CenH3 or other kinetochore proteins but instead for efficient pericentromeric heterochromatin formation. This model is consistent with the observed structure of pericentromeric heterochromatin in other model organisms that involves regular spacing of nucleosomes over long distances (SUN *et al.* 2001) and the requirement of pericentromeric heterochromatin in chromosome segregation. The repositioning that occurred in the KTF line positioned the CentC cluster adjacent to the CenH3 site to take advantage of structural properties of CentC arrays to form pericentromeric chromatin. Also, when the B chromosome centromere misdivisions occur in the kinetochore forming region, the location of the remaining kinetochore is unaltered in relation to the adjacent pericentromeric heterochromatin and therefore no repositioning occurs. In flies, a 290 kb repeat array was found to acquire kinetochore function when broken from the rest of the chromosome by X rays (MAGGERT and KARPEN 2001). However, centromere acquisition only occurred when the array was initially placed in close proximity to a functioning centromere. This

could indicate that the special chromatin structure immediately adjacent to the kinetochore reinforces kinetochore identity.

There are two predictions from this model that are testable. First, centromeres with very small centromere repeat clusters would frequently position the kinetochore, i.e. the CenH3 binding domain, adjacent to, not on top of, the repeat cluster. Because of the variation for CentC quantity among maize lines, it would be possible to examine many "small" centromeres and determine the relative position of CentC and CenH3. Second, if centromere repeats are optimized for efficient centromere formation and maintenance, *de novo* centromere formation would occur preferentially at site containing centromere repeat clusters. This second prediction is being tested in two different ways, one by other members of the NSF plant centromere group using transgenic centromere repeat cluster. Dr. Han and I are testing this prediction using the inactive B centromere found on the 9Bic-1 chromosome.

In a submitted manuscript, we demonstrate that the inactive B centromeres on sister chromatids are caused to strongly attach at the second pollen division when a factor(s) on the long arm of the intact B chromosomes is present. This attachment normally leads to failure of sister centromere separation causing nondisjunction of the B chromosome. However, on the 9Bic-1 chromosome, the active sister centromeres of chromosome 9 move to opposite poles leading to chromosome breaks in 9S when the inactive B centromere fails to separate. The breaks in 9S release acentric fragments which can be recovered if they acquire a centromere. As he documented the activation of the nondisjunction mechanism on the inactive B centromere, Fangpu screened many kernels looking for chromosomal abnormalities. In crosses where chromosome breaks were

induced, he observed a number of small fragments containing the ZmBs repeat. Using small target FISH, I have demonstrated that these fragments contain genes from 9S. We are currently attempting to document the site of the functional kinetochore and to recover additional examples of chromosome fragments. We expect that in the majority of cases, if not all, the kinetochore will be at the same position in relation to the underlying DNA elements as in the intact B chromosome.

References

Cheng Z, Dong F, Langdon T, Ouyang S, Buell CR, Gu M, Blattner FR, Jiang J (2002) Functional rice centromeres are marked by a satellite repeat and a centromere-specific retrotransposon. Plant Cell 14: 1691-1704

Henikoff S, Malik HS (2002) Centromeres: selfish drivers. Nature 417: 227

Maggert KA, Karpen GH (2001) The activation of a neocentromere in Drosophila requires proximity to an endogenous centromere. Genetics 158: 1615-1628

Nagaki K, Cheng Z, Ouyang S, Talbert PB, Kim M, Jones KM, Henikoff S, Buell CR, Jiang J (2004) Sequencing of a rice centromere uncovers active genes. Nat Genet 36: 138-145

Nasuda S, Hudakova S, Schubert I, Houben A, Endo TR (2005) Stable barley chromosomes without centromeric repeats. Proc Natl Acad Sci U S A 102: 9842-9847

Pidoux AL, Allshire RC (2005) The role of heterochromatin in centromere function. Philos Trans R Soc Lond B Biol Sci 360: 569-579

Shibata F, Murata M (2004) Differential localization of the centromere-specific proteins in the major centromeric satellite of Arabidopsis thaliana. J Cell Sci 117: 2963-70:

Sun FL, Cuaycong MH, Elgin SC (2001) Long-range nucleosome ordering is associated with gene silencing in Drosophila melanogaster pericentric heterochromatin. Mol Cell Biol 21: 2867-2879

Warburton PE, Cooke CA, Bourassa S, Vafa O, Sullivan BA, Stetten G, Gimelli G, Warburton D, Tyler-Smith C, Sullivan KF, Poirier GG, Earnshaw WC (1997) Immunolocalization of CENP-A suggests a distinct nucleosome structure at the inner kinetochore plate of active centromeres. Curr Biol 7: 901-904

Warburton PE, Cooke HJ (1997) Hamster chromosomes containing amplified human alpha-satellite DNA show delayed sister chromatid separation in the absence of de novo kinetochore formation. Chromosoma 106: 149-159

Zhang X, Li X, Marshall JB, Zhong CX, Dawe RK (2005) Phosphoserines on maize CENTROMERIC HISTONE H3 and histone H3 demarcate the centromere and pericentromere during chromosome segregation. Plant Cell 17: 572-583

APPENDIX

Complete List of The Author's Publications

In the course of my work on this dissertation, I have had the opportunity to publish several research articles. In addition to the research conducted specifically for inclusion in this dissertation, I have been able to conduct research in related fields. In some instances, the research was conducted to develop tools that could be used to address problems related to centromere studies. In other cases, I was invited to participate in research by other scientists because my expertise in an area or technique was necessary to their goals. In all cases, I was supported by my advisor, Dr. Birchler, who has been very encouraging towards all avenues of scientific inquiry. This extra work has resulted in several publications which are listed here. In some cases, the work is submitted for publication but has not yet passed throught the peer-review process.

LAMB, J. C., and J. A. BIRCHLER, 2003 The role of DNA sequence in centromere formation. Genome Biol **4:** 214.

KATO, A., J. C. LAMB and J. A. BIRCHLER, 2004 Chromosome painting using repetitive DNA sequences as probes for somatic chromosome identification in maize. Proc Natl Acad Sci U S A **101:** 13554-13559.

LAMB, J. C., J. THEURI and J. A. BIRCHLER, 2004 What's in a centromere? Genome Biol **5:** 239.

LAMB, J. C., A. KATO and J. A. BIRCHLER, 2005 Sequences associated with A chromosome centromeres are present throughout the maize B chromosome. Chromosoma **113:** 337-349.

JIN, W., J. C. LAMB, J. M. VEGA, R. K. DAWE, J. A. BIRCHLER *et al.*, 2005 Molecular and Functional Dissection of the Maize B Chromosome Centromere. Plant Cell **17:** 1412-1423.

KATO, A., J. M. VEGA, F. HAN, J. C. LAMB and J. A. BIRCHLER, 2005 Advances in plant chromosome identification and cytogenetic techniques. Curr Opin Plant Biol **8:** 148-154.

KATO, A., Y. Z. ZHENG, D. L. AUGER, T. PHELPS-DURR, M. J. BAUER *et al.*, 2005 Minichromosomes derived from the B chromosome of maize. Cytogenet Genome Res **109:** 156-165.

KATO, A., J. C. LAMB, F. HAN, J. M. VEGA and J. A. BIRCHLER, 2005 Molecular analysis of maize chromosomes. Maydica **50:** 311-320.

HAN, F., J. C. LAMB and J. A. BIRCHLER, 2006 High frequency of centromere inactivation resulting in stable dicentric chromosomes of maize. Proc Natl Acad Sci U S A **103:** 3238-3243.

HAN, F., J. C. LAMB, W. YU, Z. GAO and J. A. BIRCHLER, 2006 Centromere function and nondisjunction are independent components of the maize B chromosome accumulation mechanism, **Submitted**.

LAMB, J. C., A. KATO, W. YU, F. HAN, P. S. ALBERT *et al.*, 2006 Cytogenetics and chromosome analytical techniques, pp. 244-248 in *Floriculture, ornamental and plant biotechnology*. Global Science Books, U.K.

LAMB, J. C., and J. A. BIRCHLER, 2006 Retroelement genome painting: cytological visualization of retroelement expansions in the genera *Zea* and *Tripsacum*. Genetics **173:** 1007-1021.

LAMB, J. C., J. M. MEYER, B. CORCORAN, A. KATO, F. HAN *et al.*, 2006 Distinct chromosomal distributions of highly repetitive sequences in maize. Chromosome Res, **In press**.

LAMB, J. C., T. DANILOVA, M. J. BAUER, J. M. MEYER, J. J. HOLLAND *et al.*, 2006 Single loci detection and karyotyping using small target FISH on maize somatic chromosomes. Genetics**, In press**.

LAMB, J. C., F. HAN, D. L. AUGER and J. A. BIRCHLER, 2006 A trans-acting factor required for non-disjunction of the B chromosome is located distal to the TB-4Lb breakpoint on the B chromosome. Maize Genetics Cooperation Newsletter **80:** 51-54.

TAN, B., S. WU, L. LIU, A. SIMONNE, J. C. LAMB *et al.*, 2006 The dominant white endosperm factor White Cap encodes a large multiple copy array of the maize CCD1 carotenoid dioxygenase, **Submitted**.

VALDIVIA, E. R., J. SAMPEDRO, J. C. LAMB, S. CHOPRA and D. J. COSGROVE, 2006 Phylogenetic analysis and localization of maize genes encoding pollen group-1 allergens, **Submitted**.

YU, W., J. C. LAMB, F. HAN and J. A. BIRCHLER, 2006 Cytological visualization of transposable elements and their transposition pattern in somatic cells of maize. Genetics, **In press**.

YU, W., J. C. LAMB, F. HAN and J. A. BIRCHLER, 2006 Telomere-associated chromosomal truncation in maize. PNAS, **in press**.

LAMB, J. C., N. C. RIDDLE, Y. M. CHENG, J. THEURI and J. A. BIRCHLER, 2006 Localization and transcription of a retrotransposon-derived element on the maize B chromosome. **In preparation**.

LAMB, J. C., J. M. MEYER, and J. A. BIRCHLER, 2006 A hemicentric inversion in the maize line KTF created two sites of centromeric elements and moved the kinetochore forming region, **In preparation**.

VITA

Jonathan Carl Lamb was born on February 20th, 1977 in Apia, Western Samoa. He grew up in Granville, a small town in central Ohio. After graduating from high school, he enrolled at Brigham Young University in August, 1995. After studying for one year, he left for two years to serve as a missionary for the Church of Jesus Christ of Latter-day Saints. His service consisted of teaching English as a second language in various schools and universities in Mongolia as well as teaching the principles of the Mormon faith and organizing the church. He returned to BYU and continued his studies receiving a Bachelors of Science in Plant Genetics and Breeding in the spring of 2001. While at BYU, he met and married Elisabeth M. Maughan and they currently have two children, Archer M. Lamb and Ava Lamb, and have a third child on the way.

After completing his Ph.D. in Biological Sciences, he will continue research in the laboratory of Dr. Dorothy E. Shippen at Texas A&M University where he will study telomere maintenance in the model plant, *Arabidopsis thaliana*. Jonathan intends to pursue a career in academic science running an independent laboratory that studies basic problems in genetics and cell biology. Additionally, he hopes to start a research program that collaborates with plant breeders to assist in plant improvement in an understudied crop species.