Principal Investigator: R. Kelly Dawe, University of Georgia
Co-Principal Investigators: James A. Birchler, University of Missouri; Jiming Jiang, University of Wisconsin; Wayne A. Parrott, University of Georgia; Gernot Presting, University of Hawaii

Scientific Approaches and Objectives:
Centromeres have long been recognized as essential for cell division. It is now clear that centromeres are much more complex than ever envisioned; being composed of a mixture of rapidly evolving selfish DNA that binds to a conserved multi-protein kinetochore complex. The DNA-protein interactions are mediated by genetic and epigenetic mechanisms that appear to be unique to centromeres. Realizing that our ability to exploit plant centromeres is contingent upon understanding their complex biology, a comprehensive, integrated set of experiments will be carried out with the aim of answering fundamental questions in centromere research. The resulting data will not only provide insight into centromere biology and its role in evolution, it will also build the foundation for artificial chromosome construction and open the door for a new era in plant transformation.

In initial experiments, existing centromeric BAC contigs will be mapped onto chromosomes. Preliminary mapping will be followed by in-depth studies designed to develop a contiguous sequence representing maize centromere 4 or a suitable replacement. This information will be used to determine the abundance, distribution, and map position of known and novel centromere repeats. The distribution of three kinetochore proteins along mapped centromeric DNA will be determined, and the most effective centromere sequences identified using a novel in-vivo competition assay. Finally, new vectors will be constructed and experiments executed to identify the most efficient means of introducing and activating centromeric sequences in transgenic plants.

Broader Impacts:
To expand the training opportunities available through the project, two new programs will be initiated. One, centered at the University of Georgia (called GROUP), will combine two existing on-campus minority involvement programs and establish a plant biology-genomics focus. It is likely that GROUP will ultimately be integrated into all plant genetics research on campus.

In addition, hands-on training in plant cytogenetics will be offered through an Annual Cytogenetics Training Workshop (ACTW). The ACTW will provide training in fluorescent in situ hybridization, immunolocalization techniques, chromatin immunoprecipitation, and advanced microscopy methods. The workshop will be held at the University of Wisconsin and be led by Jiming Jiang and R. Kelly Dawe.

BACKGROUND

Centromeres are large repetitive DNA regions that organize kinetochores and mediate chromosome movement during cell division. In addition, centromeres are widely viewed as the key to producing the next generation of plant transformation vectors: artificial chromosomes (Brown et al., 2000; Somerville and Somerville, 1999). Artificial chromosomes offer the potential to introduce many genes at once without complications associated with multiple rounds of transformation. Although early success with yeast (Clarke and Carbon, 1980; Murray et al., 1986) suggested that plant artificial chromosomes (PACs) might not be far off, the complexity of plant centromeres has made progress towards PACs extraordinarily difficult (Henikoff, 2002). This is due to longstanding frustrations in cloning and sequencing centromeres, the fact that centromeres seem to evolve faster than any other region of the genome, and difficulties in developing the technologies required to construct an artificial centromere. Under our current award we have made significant progress towards overcoming these obstacles. Here we propose to continue our studies, focusing our efforts on three key questions, which collectively will provide the necessary understanding to manipulate centromeres:

1) What are the important sequences in plant centromeres?

Human centromeres are an estimated 3000-4000 kb in size (Schueler et al., 2001) and are primarily composed of simple ~171 bp repeats known as alpha-satellites. In *Drosophila*, the ~300 kb centromere is composed of small repeats of only 5, 7 or 10 bp units (Sun et al., 2003). In contrast, plant centromeres range from 500 kb to ~3 Mb in size and contain a variety of repeats and transposons (Jin et al., 2004; Kaszas and Birchler, 1996; Kumekawa et al., 2000; Nagaki et al., 2004). Little or no homology exists among the satellite repeats in these well-studied centromeres. However, several years ago two groups discovered a class of highly conserved Centromeric Retrotransposable (CR) elements within the *Poaceae* (Aragon-Alcaide et al., 1996; Jiang et al., 1996). The homology within the CR clade of retroelements is striking: CR probes from any of the major cereal grains (maize, rice, barley, sorghum, rye, wheat, or oat) will recognize the centromeres from all the other species (Jiang et al., 2003). CR elements tend to be in clusters intermingled with the centromere satellite arrays characteristic of each species (Cheng et al., 2002b; Hudakova et al., 2001; Jin et al., 2004).

Identifying abundant centromeric DNAs is one thing; determining which of the DNAs are functional is a more significant challenge. Sequences that lie within the primary constriction (by cytological criteria), sequences that map to the centromere but don't recombine with each other (by genetic criteria), and sequences that interact directly with key kinetochore proteins by Chromatin ImmunoPrecipitation (ChIP, a biochemical criterion) are all viewed as good candidates for functional centromeric DNA (Copenhaver, 2003; Houben and Schubert, 2003; Jiang et al., 2003). In recent years the biochemical definition has taken center stage (Jiang et al., 2003), although cytological and genetic data are often used to support ChIP data. The most widely used ChIP reagents are antisera to Centromeric Histone H3 (CENH3), a highly conserved histone H3 variant present only at centromeres (Meluh et al., 1998; Talbert et al., 2002; Vafa and Sullivan, 1997; Zhong et al., 2002). CENH3 antisera have been used to identify functional centromeric DNAs in a variety of species including maize (Zhong et al., 2002), *Arabidopsis* (Nagaki et al., 2003b), and rice (Nagaki et al., 2004). A surprising finding from rice (Nagaki et al., 2004), shared with similar data in humans (Saffery et al., 2003), is that CENH3 does not only interact with satellites and retroelements, but with single-copy DNA as well. Although DNA sequence is clearly one part of a centromere specification (e.g. Ohzeki et al., 2002) epigenetics is also a part of the process (Karpen and Allshire, 1997; Lamb and Birchler, 2003).

CENH3 is but one component of a large multi-protein inner kinetochore complex (Westermann et al., 2003). There are at least two other well-conserved inner kinetochore proteins that are critical for centromere function: Centromere Protein C (CENPC) (Dawe et al., 1999), and Mis12 (Goshima et al., 2003). New data suggesting that both CENH3 and Mis12 are required to recruit CENP-C (Goshima et al., 2003), and the fact that CENH3 does not always co-localize with other proteins of the inner kinetochore (Saffery et al., 2003) suggest that more than one inner kinetochore protein should be used to identify functional centromeric DNA.

2) How do centromeres evolve and what are the consequences for centromere structure?

A major question in centromere research is how the kinetochore proteins maintain their interactions with rapidly evolving centromeric tandem repeat arrays. Henikoff and colleagues have suggested that centromere evolution may be an outcome of meiotic drive (Malik and Henikoff, 2002).

Under this view, centromeric DNAs that are most efficient at binding kinetochore proteins (such as CENH3) are more likely to arrive at spindle poles first and be segregated to functional megaspores. The CENH3s in *Drosophila* and *Arabidopsis* show clear evidence of adaptive evolution, consistent with this hypothesis (Malik et al., 2002; Talbert et al., 2002). Satellites may evolve to bind more efficiently to CENH3 and enhance their chances of being segregated to progeny, while CENH3 evolves to modulate the interaction so that all chromosomes segregate with the same efficiency (Malik and Henikoff, 2002). The host benefits by maximizing interactions between the centromere and kinetochore, while satellite DNA benefits selfishly by increasing its copy number.

Another expectation of the centromere drive model is that newer repeats are more efficient at binding to the kinetochore proteins. Sequence data from the human X centromere suggests that the newest, most uniform satellite arrays lie within the centromere core, while the older more divergent copies lie in flanking pericentromeric regions (Kazakov et al., 2003; Schueler et al., 2001). By inference, it appears that centromeres 'grow outward', generating new (high affinity) repeats in the center and pushing older, less effective repeats to the flanks were they degenerate and become nonfunctional (Henikoff, 2002). Although both ‘centromere drive’ and the ‘centromere core’ ideas are evolutionary models, they have important practical implications. If we can mimic centromere drive in the field, we will have a strong functional assay for centromeric DNA; likewise, if we can identify the high affinity repeats of the centromere core, these would be the best sequences to use in artificial chromosomes.

3) How can we create artificial chromosomes?

Yeast artificial chromosomes were produced with apparent ease by combining a centromere, an origin of replication, and a telomere into a single molecule (Murray et al., 1986). A similar strategy of co-transfecting the constituent pieces of a chromosome – centromeric alpha-satellite and telomeric DNA - was also successfully employed in human cells, though the resulting artificial chromosomes often contained sequences derived from the host genome (Harrington et al., 1997). Other strategies have since been used with greater success, and it is now clear that alpha-satellite DNA alone is sufficient to organize an artificial centromere/kinetochore complex (Ikeno et al., 1998; Nakano et al., 2003; Ohzeki et al., 2002; Rudd et al., 2003). Nevertheless, many important questions remain about the efficiency of the process. The frequency of human artificial chromosome formation is generally low (Harrington et al., 1997; Nakano et al., 2003), strongly dependent on genetic background (Brown et al., 2000), and the resulting chromosomes are often poorly segregated (Rudd et al., 2003). In an effort to shed some light on these variables, Nakano et al (2003) focused on cases where artificial chromosomes did not form following transfection, but instead inserted into the genome as ectopic arrays. They showed that by selecting for the expression of a linked marker gene, or by treating cells with a drug that promotes active transcription (trichostatin A), they could increase the frequency of centromere-active ectopic satellite arrays by at least 3.6 fold and to as high as 11 fold (Nakano et al., 2003). The authors argued that introduced centromeric arrays tend to become heterochromatic and as a result unlikely to recruit key kinetochore proteins. However, when centromeric DNA is maintained in an open chromatin state, either by selecting for transcription nearby or by repressing the formation of heterochromatin, centromeric proteins are more likely to gain access and successfully organize a kinetochore.

The idea of a transcriptionally active centromere has gained considerable momentum in recent months. In human cancer lines, neocentromeres occasionally form over gene-containing regions (Choo, 2001). A recent analysis of one such neocentromere demonstrated that 15 of the 51 genes present produced apparently normal transcripts, proving that the unique structure of the centromere is not

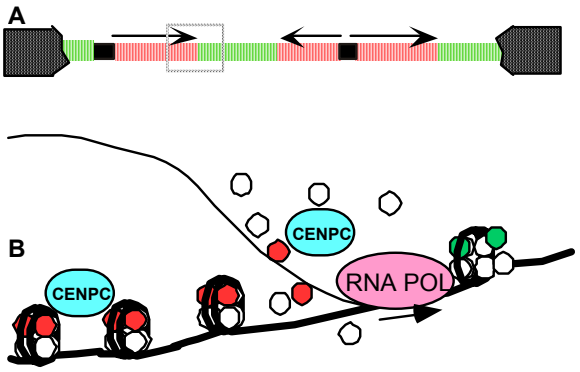


Figure 1. A model for CENH3 deposition (Jiang et al., 2003). **A**) Transcripts may be initiated from CR elements (black) and extend over flanking repeat arrays, replacing H3s (green) with CENH3s (red). **B**) RNA polymerase proceeding over repeat arrays destabilizes nucleosomes. Some RNA stays in contact with centromeres (Fig. 4), suggesting that the RNA may recruit CENH3 and/or associated proteins such as centromere protein (CENPC).

recalcitrant to transcription (Saffery et al., 2003). Subsequently, the laboratory of one of our Co-PIs (Jiang) showed that rice centromere 8 contains fourteen genes, four of which are actively expressed (Nagaki et al., 2004). Finally, as described below, we have shown that maize centromeric retroelements and satellites DNAs are actively transcribed. In our recent review of plant centromeres (Jiang et al., 2003) we emphasized a model that links centromere transcription with two other key observations: that CENH3 is deposited after DNA replication (Ahmad and Henikoff, 2001; Shelby et al., 2000) and that centromeres contain a mixture of histone H3 and CENH3 (Blower et al., 2002; Nagaki et al., 2004). Under this model transcription during G2 destabilizes nucleosomes and initiates the histone replacement process (Fig. 1). The fact that we are now able to transform long intact arrays of centromeric DNA into rice and maize (Phan et al, in preparation) should make it possible to test this model.

RESULTS OF FROM PRIOR NSF SUPPORT

Sequencing and analysis of centromeric DNA from A chromosomes

Two repetitive DNA elements, CentA and CentC, were isolated from maize centromeres before we began our project (Ananiev et al., 1998). However, it was not known whether these repeats were the only components of maize centromeres or how these elements were organized. One of the first goals of our project was to sequence several BACs derived from maize centromeres (GenBank AC114395, AC116034, and AC116033). The results revealed long arrays of the centromeric satellite CentC (Ananiev et al., 1998), two classes of what we termed CRM elements (CentA/CRM1 and CRM2) and a variety of other retroelements. Both classes of CR elements are located exclusively at centromeres and lie within or closely flanking tandem arrays of CentC. Sequence analysis suggests that CRM elements are relatively young (Nagaki et al., 2003), and that they target CentC arrays during transposition. We also characterized the abundance of six common retroelements in centromeres using 3D light microscopy. The data demonstrated that Opie, Huck, Cinful-1, Prem-2/Ji, Grande, and Tekay/Prem-1 are all relatively rare in centromeric satellite arrays, being only 30% as frequent in centromeres as in euchromatin (Mroczek and Dawe, 2003).

Large scale organization of individual maize centromeres

To reveal the large-scale organization of CentC and CRM we carried out fiber-FISH on oat-maize chromosome addition lines. In these strains, single maize chromosomes are propagated in oat, making it possible to analyze a single centromere at a time (Kynast et al., 2001). The centromeres from chromosomes 1, 2, 3, 4, 6, 7, and 9 were analyzed. The fiber-FISH signals from CentC and CRM were contiguous with few unambiguous gaps, confirming that these two sequences are indeed the dominant DNA components of maize centromeres. Fiber-FISH signals varied in length from 299±29 kb on centromere 3 to >2.8 Mb on centromere 7. Surprisingly, CRM is more abundant than CentC on three chromosomes (4, 6 and 9). We also showed that a previously identified centromeric sequence known as Cent4 (Page et al., 2001) does not co-localize with the CentC/CRM by fiber-FISH and does not co-immunoprecipitate with CENH3 (see below). Finally, we addressed the question of whether centromeres are damaged or rearranged during the process of introducing maize chromosomes into oat. The fiber-FISH patterns derived from the centromeres of three independently isolated chromosome 6 addition lines were nearly identical, indicating that substantial rearrangements had not occurred.

Organization of B chromosomes

One of the goals of the project was to examine the DNA sequence organization of small centromeres derived from the B chromosome centromere. Using the natural centromere breakage process known as misdivision (Darlington, 1937; Kaszas et al., 2002), a series of progressively smaller centromeres was derived (Kaszas and Birchler, 1996). We attempted to produce even smaller B centromere derivatives under this award (Phelps-Durr and Birchler, 2004). However, meiotic transmission drops to near zero when the B centromere is reduced to a size smaller than one or two hundred kilobases (Phelps-Durr and Birchler, 2004), indicating that we have reached the lower limit of the B centromere size.

Partial BAC libraries were produced from three small misdivision derivatives (Ring 4-8, Ring 4-9 and Telo 3-3). Several BACs were selected based on the presence of a B-specific centromere repeat (Alfenito and Birchler, 1993) and sequenced using a transposon-based method (Kimmel et al., 1997). Approximately 124 kb of submitted sequence indicates that B centromeres are similar to A centromeres, with the exception that that the B repeat is abundantly interspersed with CRM and CentC (Theuri et al., 2004). We also used fiber-FISH to analyze the structure of the normal B centromere and seven small-

centromere derivatives (Kaszas and Birchler, 1998). While the amount of B repeat in the seven small-centromere derivatives varied significantly (from 100 kb to >1 Mb), CentC and CRM were retained in all but one. The exception was Telo3-3, which has the lowest transmission among the seven lines analyzed (~20% of wild type; Kaszas and Birchler, 1998). It is likely that sequences other than CentC and CRM confer centromere activity in maize; but to what extent is not yet known.

Identification of maize CENH3 and its use to identify key centromeric DNAs

The kinetochore protein most tightly bound to DNA is CENH3, a conserved histone H3 homologue that replaces histone H3 in centromeric chromatin (Choo, 2001). Antisera to CENH3 have been used to identify centromeric DNA by ChIP (Meluh et al., 1998; Vafa and Sullivan, 1997). Therefore, one of our first efforts was to identify a full-length cDNA for maize *CenH3*. Polyclonal antibodies against maize CENH3 were generated using a synthetic peptide designed from the predicted protein sequence (Zhong et al., 2002). The resulting anti-CENH3 antibodies have an unusually high titer, specificity and broad utility; they work well for westerns, immunolocalization, and chromatin immunoprecipitation.

To identify the maize DNAs that interact with CENH3, we used a native ChIP method developed in the Choo laboratory (Lo et al., 2001). Nuclei were digested with micrococcal nuclease, incubated with antisera, and the immune complexes precipitated and separated into unbound (S) and bound (IP) fractions. Equal amounts of the S and IP fractions were blotted on membranes, hybridized with ^{32}P-labeled probes, and the extent of hybridization quantified using a phosphorimager. The results showed that 32-37% of CentC and CRM was immunoprecipitated by anti-CENH3 antisera, whereas negative controls such as the knob 180 bp repeat and Tekay retroelements were immunoprecipitated at less than 7%. These data provided the first biochemical evidence that centromeric satellite DNAs are key elements of the plant centromere, and support the view that the highly specialized CR elements cooperate with repeat arrays to form a functional centromere/kinetochore complex (Zhong et al., 2002).

The interaction between CENH3 and CRM/CentC was confirmed using a novel extended fiber immunolocalization technique (Jin et al., 2004). Nuclei were prepared from actively growing callus, centrifuged onto slides using a Cytospin 4 centrifuge, and manually stretched. Immunostaining signals derived from well-stretched centromeres were recorded and the same slides probed with CentC and CRM. Three types of co-localization were observed: (1) A nearly perfect overlap of CENH3 and CentC/CRM sequences; (2) CENH3 staining located in the middle of the fiber-FISH signals; and (3) CENH3 staining located to one end of the fiber-FISH signals. Examples of the first and last classes are shown in Figure 2. These results demonstrate that CentC and CRM underlie the majority of the CENH3-associated centromeric chromatin in maize. The fact that long tracts of CentC and CRM often do not interact with CENH3 (Fig. 2B) may be taken as evidence for the centromere core concept, or simply as further evidence of a stochastic/epigenetic component to centromere specification.

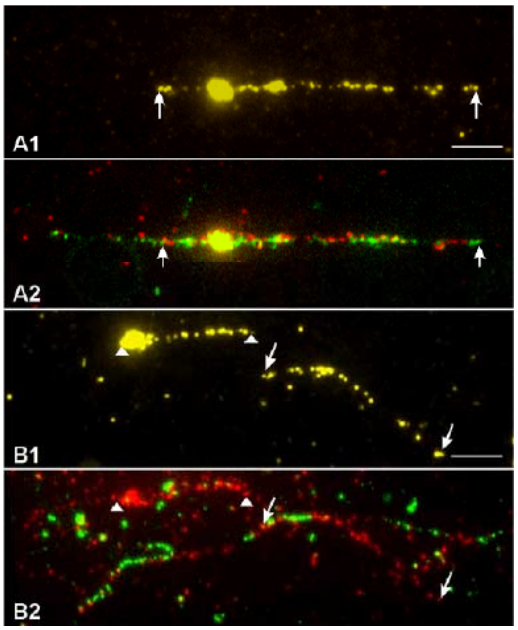


Figure 2. Association between CENH3 and CentC/CRM sequences on stretched maize centromeres. **A)** CENH3 (1) and CentC (green) and CRM (red) (2) on the same fiber, showing strong co-localization. **B** CENH3 (1) and CentC (green) and CRM (red) (2) on the same fibers. In this case CENH3 staining was located to one end of the fiber-FISH signals. Arrows and arrowheads point to the region associated with CENH3. Bar = 5 μm

Phosphorylation of CENH3 at serine 50 correlates with kinetochore maturation

Our CENH3 antisera strongly recognize maize kinetochores throughout the cell cycle except at prometaphase, metaphase, and anaphase, when the signal is weak or absent. Since the N-terminal peptide used to generate antibodies includes a serine (the 50th amino acid in the protein), and because staining at metaphase was restored after phosphatase treatment, we suspected the absence of staining on condensed chromosomes was caused by serine phosphorylation (similar to serine 10 phosphorylation of human CENH3; Zeitlin et al., 2001). We have now generated and extensively characterized an

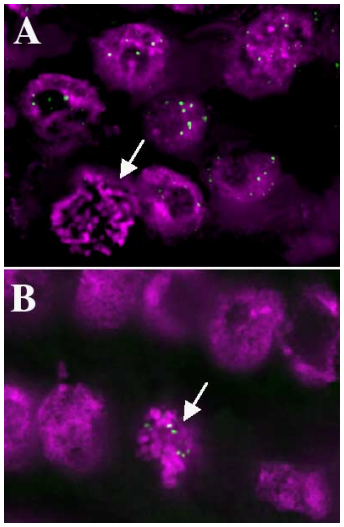


Figure 3. CENH3 and pCENH3 staining in root cells. A) anti-CENH3 (green) stains prophase cells, but not interphase cells. B) anti-pCENH3 stains prophase but not interphase cells. Arrows indicate prophase cells.

antibody to the serine 50-phosphorylated form of CENH3 (anti-pCENH3) (Zhang et al., in preparation). The anti-pCENH3 and anti-CENH3 antisera are both highly specific for centromeres, but they exhibit complementary staining patterns (Fig. 3). The two forms of the protein coexist only at early prometaphase and anaphase. At late prometaphase and metaphase all detectable CENH3 is phosphorylated. Anti-phosphoserine-50 CENH3 antisera provide a new reagent for ChIP analysis of condensed, active centromeres.

The maize centromere is transcribed

We became interested in chromatin-associated RNA when it was shown that RNaseA could change the structure of pericentric chromatin in mouse cells (Maison et al., 2002; Dawe, 2003). To test whether maize centromeres are transcribed and/or contain RNA, we used a modified ChIP protocol (Topp et al., submitted). Instead of blotting for DNA, ChIP samples were blotted and probed for RNA (like a northern blot). RNaseA was used to confirm the RNA hybridization signal. We detected strong in both the S and IP fraction when RNA blots were probed with CRM or CentC satellite repeats, but not when they were probed with the 180 bp knob repeat (Fig. 4). Riboprobes specific for the forward and reverse strands of CRM and CentC were also used to show that both strands of the repeats are present in roughly equal proportions. The simultaneous presence of forward and reverse CRM transcripts would be expected to activate the RNAi pathway and produce siRNAs (Hannon, 2002). However, we were not able to detect RNA in the size range expected for siRNAs (22-30 nt). Rather, the bulk of the RNA detected was between 40 and 250nt in length (Fig. 4B; longer RNAs of 900 nt could also be recovered using RTPCR for CRM). These and other data showing that CENH3 staining increases significantly during early G2 led to the model in Figure 1 whereby transcription facilitates CENH3 deposition. Under this model the primary function of transcription is to destabilize nucleosomes. However, the fact that some RNA remains in contact with the centromere (Fig. 4) suggests that RNA may also have structural or recruiting roles.

Long centromere arrays can be transformed into plants and remain intact

An important aim of our proposal was to take the first steps towards generating artificial chromosomes. We proposed to transform plants with centromeric BACs and determine whether these 'trans-centromeres' showed evidence of centromere activity. The feasibility of the first steps of this proposal had yet to be established: BAC-sized constructs had never been transformed into cereals and there were no data on the stability of repetitive DNAs during plant transformation. To answer these questions we took advantage of the easy transformation of rice and two well studied centromeric BACs, one from maize (BAC16H10, Nagaki et al., 2003) and one from rice (BAC 17p22, Cheng et al., 2002). Molecular data indicate substantial conservation between the CR elements and centromeric satellites of maize and rice (Cheng et al., 2002, Zhong et al., 2002, Nagaki et al., 2003),

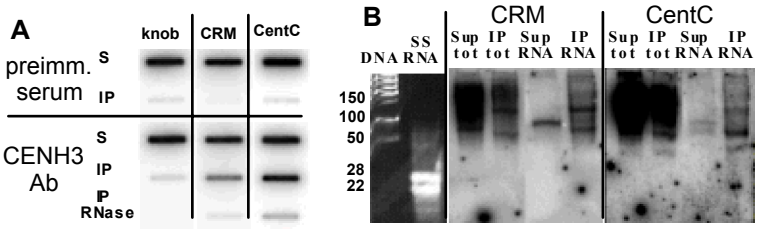


Figure 4. Centromere-encoded RNAs are co-immunoprecipitated with CENH3. **A**) Most of the RNA remains in the supernatant (S) when chromatin is immunoprecipitated with preimmune serum, but substantial quantities of centromeric RNA are co-immunoprecipitated (to the pellet, IP) with anti-CENH3 antibodies. RNaseA removes the bulk of the RNA. **B**) Centromeric RNA ranges in size from 40-250 nt in length. Supernatant (Sup) and IP fractions were run on 15% acrylamide gels untreated (lanes labeled 'tot') or after first treating the samples with DNaseI (RNA lanes). Blots were hybridized with the CRM GAG and CentC antisense riboprobes.

suggesting that it may be possible to produce artificial chromosomes that function in both species.

The procedure used throughout was to co-bombard BAC DNA with a second plasmid containing the selectable marker *bar* and the *gusA* gene. Two non-centromeric rice BACs (76A03 and 7K12) were used as controls. While transgenic calli were readily obtained, recovering regenerated plants with centromeric BAC inserts was much more difficult (overall frequencies were <1%). Seven plants containing the rice BAC17p22, 2 plants containing the maize centromeric BAC 16H10, and 5 plants containing control BACs are currently under study. Interestingly, all 5 plants transgenic for the control BACs were phenotypically normal, while all plants transgenic for centromeric BAC DNA were stunted and slow to mature.

Since BACs 17p22 and 16H10 contain the centromeric satellites CentO (rice) and CentC (maize), respectively (Cheng et al., 2002; Nagaki et al., 2003), these repeats can be used as probes for FISH. Six of the seven transgenic rice lines containing rice BAC 17p22 and two lines containing maize BAC 16H10 contain a single integration site (Fig. 5A,B), while the remaining 17p22 line contains four integration sites. Fiber-FISH analysis of one maize trans-centromere (Fig. 5C) suggests that the BAC molecule is largely intact.

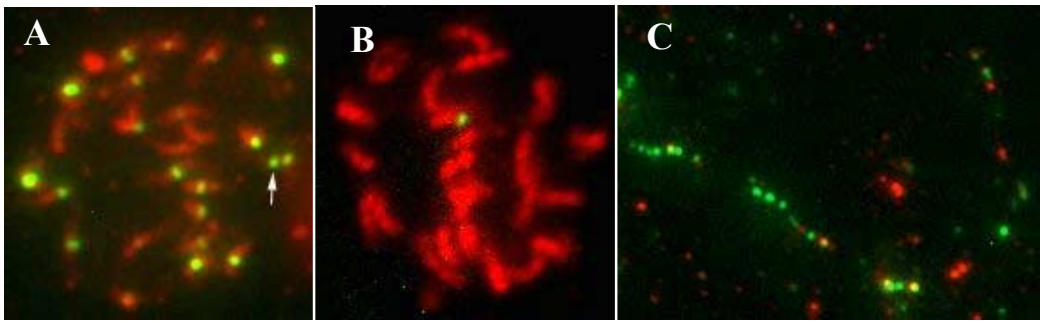


Figure 5. FISH localization of BAC-derived trans-centromeres. **A**) Rice centromeric BAC 17p22 inserted into rice, as detected using the rice CentO satellite (arrow). The signal from the trans-centromere is roughly the same intensity as the signal from the natural centromere adjacent to it. **B**) Maize centromeric BAC 16H10 inserted into rice, as detected using CentC. **C**) Fiber-FISH of the insert shown in B. The trans-centromere is roughly 120 kb long.

Publications arising from prior support (see **REFERENCES** for complete citations**)**

1 Page BT, Wanous MK, Birchler JA. 2001. Genetics 159:291-302.
2 Kaszas E, Kato A, Birchler, JA. 2002. Genome 45:759-76.
3 Zhong C, Marshall JB, Topp C, Mroczek R, Kato A, Nagaki K, Birchler JA, Jiang J, Dawe RK. 2002. Plant Cell 14:2825-2836.
4 Nagaki K, Song J, Stupar RM, Yuan Q, Ouyang S, Liu J., Dawe RK, Buell CR, Jiang J. 2003. Genetics 163:759-770
5 Nagaki K, Talbert PB, Zhong CX, Dawe RK, Henikoff S, Jiang J. 2003. Genetics 163:1221-1223
6 Lamb JC, Birchler, JA. 2003. Genome Biol. 4:214.
7 Mroczek RJ, Dawe RK. Genetics 165: 809–819.
8 Dawe RK. 2003. Plant Cell 15, 297-301
9 Jiang J, Birchler JA, Parrott WA, Dawe RK. 2003. Trends Plant Sci. 8:570-575
10 Jin W, Melo JR, Nagaki K, Talbert PB, Henikoff S, Dawe RK, Jiang J. 2004. Plant Cell, in press.
11 Topp CN, Zhong CX, Zhang X, Dawe RK. Submitted.
12 Phelps-Durr T, Birchler JA. 2004. Cytogenet. Genome Res. Submitted.
13 Theuri J, Phelps-Durr T, Mathews S, Birchler JA. 2004. Cytogenet. Genome Res. Submitted
14 Phan BH, Jin W, Topp CN, Zhong CX, Jiang J, Dawe RK, Parrott WA. In preparation
15 Zhang X, Li X, Marshall JB, Zhong CX, Dawe RK. In preparation.

RELEVANCE AND JUSTIFICATION TO THE PGRP

Roughly 259 million tons of maize were produced in the US during the year 2003, dwarfing production of the second-most commonly US-grown cereal grain (wheat at 12 million tons). These statistics and the PGRP have driven a large increase in funding for maize research. We now have impressive DNA sequence resources, a strong information database, several reverse genetics resources, and an extensive genetic map. Nevertheless, the centromere, as in many other species, remains the 'black hole' of the genome (Henikoff, 2002). Recent genomic sequencing efforts in maize excluded the repetitive DNAs of the centromere (Palmer et al., 2003; Whitelaw et al., 2003). Fortunately, the large maize chromosomes have attracted many cytologists over the years, and more is now known about maize kinetochores than those of any other plant (Houben and Schubert, 2003; Yu et al., 2000). This, combined with two unparalleled cytogenetic attributes, the well-studied maize B chromosomes (Birchler et al., 2004) and the more recent oat-maize addition lines (Kynast et al., 2001), have made maize the premier organism for plant centromere/kinetochore research.

The Plant Genome Research Program is designed to foster wide collaborations and develop virtual centers. Here we propose to continue the successful collaboration developed among Dawe, Jiang, Birchler and Parrott, and add the considerable expertise of Dr. Gernot Presting. Dr. Presting was one of the first plant biologists to begin molecular studies on centromeres (Presting et al., 1996), and was largely responsible for cloning the first full-length CR element (Hudakova et al., 2001; Presting et al., 1998). He spent the last three years training in bioinformatics and was heavily involved in the physical mapping of the rice genome (he is now an Assistant Professor in Hawaii). All of our current and proposed work is strongly collaborative. Each of our major papers have involved more than one of the cooperating labs (Jin et al., 2004; Nagaki et al., 2003; Zhong et al., 2002) and all reagents have been shared immediately, years in advance of publication. By bringing together the major centromere researchers in the U.S. into a single interactive group, we will continue to avoid duplication of effort, accelerate centromere research, increase collaboration among universities, expand our training opportunities, and bring down the overall cost of producing plant artificial centromeres in the public domain.

RESEARCH PLAN

Our long-term goals are to better understand the organization of native centromeres, unravel centromere evolution and its consequences for centromere structure and function, and use that information to delineate the steps necessary to create artificial chromosomes. For the current proposal, these broad goals are subdivided into smaller five-year aims, designed to take advantage of our combined expertise in bioinformatics, genetics, cytogenetics, and biochemistry.

GOAL1: MAPPING THE MAIZE KINETOCHORES

A major objective of this proposal is to complement the public effort to map and sequence the maize genome. While major projects are focusing on the gene-rich regions (http://w3.ag.uiuc.edu/maize-coop/Maize-Genome-Projects.html), we will carefully analyze centromeric regions. We will begin by generating as much contiguous centromere information as is possible from the existing data and all new public sequence/FPC (FingerPrint Contig) data produced over the next five years. These contigs will be low-pass sequenced to identify CentC variants and single-copy sequence for genetic mapping. In addition, we will generate a complete tiling path across chromosome 4 (or a suitable replacement) and sequence it at high redundancy. Markers obtained from high-pass sequence will be assayed for their interactions with four anti-kinetochore antisera, effectively mapping the inner kinetochore to the centromere.

AIM1A) SEQUENCING AND MAPPING OF CENTROMERIC FPC CONTIGS ONTO CHROMOSOMES

I) Identification of centromeric contigs using the public BAC library and database. BACs containing centromere sequences will be identified in two ways. First, the BAC end sequence generated by the Arizona Genomics Institute will be examined for the presence of known centromere repeats. A preliminary analysis of 89,000 BAC end sequences found 29 ends with homology to CentC and 571 ends with homology to CRM. Second, filters for the maize BAC library will be obtained and hybridized with centromeric repeats in order to identify all BAC clones with internal copies of the repeats. The distribution of centromere repeat-containing BACs in the Arizona fingerprint database will be examined, and contigs

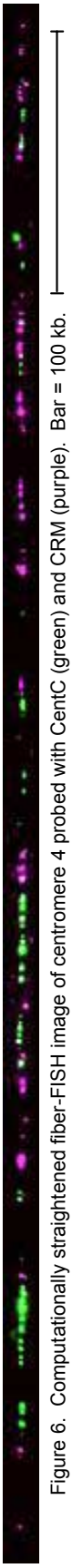


Figure 6. Computationally straightened fiber-FISH image of centromere 4 probed with CentC (green) and CRM (purple). Bar = 100 kb.

containing more than 1 BAC end sequence with repeat homology, or more than 1 BAC that hybridizes with a centromere repeat, will be identified.

II) Low-pass sequence and chromosome placement. Low pass (1X coverage) sequence will be obtained from roughly 10 Mb of centromeric contigs (~1Mb/chromosome) to reveal single or low-copy sequences that can be used as molecular markers. Sequencing will be performed at the Clemson University Genomics Institute (CUGI, see letter by Jeff Tomkins, director). These markers will be mapped by PCR onto the maize genetic map using 94 individuals from the B73 x Mo17 maize mapping population, thus assigning each centromeric FPC contig to a chromosome. Genomic DNA from the mapping population will be obtained from Georgia Davis at the University of Missouri. This step will bring together small, unplaced contigs and map them to chromosomes. It will identify single-copy markers, provide data on the frequency of CentC variants in each chromosome, and facilitate the identification of a full tiling path over a single centromere.

III) Cataloging the position and frequency of CentC variants. CentC is a key sequence in maize centromere specification, yet we do not know how it is organized, where the variants are located, or if any particular repeats are more likely to be associated with functional domains of the centromere (nor has this been studied in any other plant). We will carefully catalog the percentages of each CentC variants obtained from each of our low-pass BAC sequences. As described below, SNPs will be obtained from these data and assayed for their interaction with kinetochore proteins. The rationale for this aim is the popular meiotic drive/centromere core hypothesis for centromere evolution (Malik and Henikoff, 2002), which predicts that some repeats are more likely to interact with CENH3 than others, and that these high-affinity repeats will map to the center of centromeres.

AIM1B) SEQUENCING CENTROMERE 4, OR MORE SUITABLE CANDIDATE.

Many of the questions we wish to address can only be answered with a complete centromere contig. There are several reasons why we believe maize is a suitable system for such an undertaking, and why maize will be well served by the effort should we succeed. First, FPC mapping in maize is now in full swing, and any technical progress we make will directly benefit these efforts. Second, since the first native centromere to be fully sequenced was from rice (Nagaki et al., 2004), our data will make the concept of 'comparative centromere genomics' a reality. Third, the fact that the first sequenced centromere was from rice and not *Arabidopsis or* humans bodes well for the continued success of centromere sequencing in the cereals (a recently reported sequence in *Drosophila* was not derived from a native centromere; Sun et al., 2003). Not surprisingly, the key to success appears to be the length of the constituent satellite arrays. In rice centromere 8 the longest satellite array is less than 60 kb, and much of the rest of the centromere is composed of complex arrangements of CR elements and other retrotransposons rich in restriction fragment length polymorphisms. A perusal of maize centromeres reveals exactly the same sort of arrangement except that (with the exception of chromosomes 1 and 7) the average length of individual CentC arrays is even smaller. On maize centromere 4 (Figure 6) the longest continuous CentC array is only 50 kb, much smaller than the size of an average BAC clone.

I) Phase I FPC preparation: use of oat-maize addition lines to establish a centromere 4 tiling path. It is possible that an analysis of existing FPC data (including our contributions from AIM1A) will allow us to assemble a tiling path across one or more centromere(s). If so, this/these centromere(s) will be targeted for in-depth sequencing as described below. If we are unable to identify a complete tiling path over an entire centromere, we will take advantage of the oat-maize addition lines for chromosome 4 to fulfill this aim (Kynast et al., 2001; Okagaki et al., 2001). A BAC library will be prepared from oat-maize addition line 4, the ligation mixture transformed into *E. coli*, and plated at ca 10,000 colonies per plate. Filter lifts from these plates will be hybridized with maize centromere probes to identify colonies that contain centromeric BACs. At an estimated 533 kb in size (Jin et al. 2004; Fig. 6), centromere 4 will require the isolation of 107 BAC clones to achieve 20x coverage. We propose to use the high information content fingerprint technique (HICF) to fingerprint centromere 4 BACs. HICF not only is more automated than conventional fingerprinting, but also resolves more fragments per BAC clone and adds sequence information (Ding et al., 2001). We will conduct HICF using CUGI's ABI3730 sequencer and analyze the data using the GenoProfiler software. Given the high

efficiency of HICF and the fact that all BACs will have originated from the same 533 kb centromere region, it is possible that 10x or lower coverage will be sufficient to assemble a high quality fingerprint contig.

II) Phase II FPC preparation: editing by FISH confirmation. A powerful method for confirming contig assemblies is fiber-FISH (Nagaki et al., 2004). Centromeric BACs are inherently difficult to map due to the low single-copy DNA content (Nagaki et al., 2004), but the analysis will be simplified by taking advantage of the chromosome 4 oat-maize addition line. CentC is specific to maize and an excellent anchor sequence to distinguish maize centromeres from those in oat. Likewise, probes to the 5' untranslated region of CR elements are generally species specific. Southern analyses indicate that the UTR800 probe from maize (Zhong et al., 2002) does not recognize oat DNA. Thus, we should be able to use fiber-FISH to confirm the orientation and overlap of the BACs within the contig.

III) Sequencing and identification of single-copy markers. Once a minimum tile is identified and confirmed, it will be sequenced at 20X coverage in collaboration with CUGI. Based on other centromere sequence data, we expect to find single-copy markers within the resulting sequence. In rice centromere 8, 11 single-copy markers were identified within the CENH3-bound centromere (Jin et al., 2004) and in *Arabidopsis*, many single copy sequences were identified within the genetically mapped centromeres (though they have yet to be assayed for interactions with CENH3; Yamada et al., 2003). Although we cannot anticipate how many single-copy markers will be present in centromere 4, we will catalog and assay all low-copy markers that lie within the ~533 kb contig. Each marker will be verified as single or low-copy by Southern analysis, and primers will be identified that amplify the single sequence under study.

AIM1C: MAPPING THE INNER KINETOCHORE COMPLEX ONTO CENTROMERES

CENH3 interacts with a variety of other proteins that together make up the inner kinetochore. One of these is CENPC, a DNA-binding protein that interacts with CENH3 by co-immunoprecipitation in both humans and yeast (Ando et al., 2002; Westermann et al., 2003). Excellent antibodies to maize CENPC are already available in the Dawe lab (Dawe et al., 1999). In addition to CENH3 and CENPC, a third conserved inner kinetochore protein called Mis12 has recently been described (Goshima et al., 2003). Genetic data suggest that CENH3 and Mis12 share 'ground-level' status at the kinetochore, in that they are both required to recruit CENP-C (at least in humans; Goshima et al., 2003). These data, and the fact that CENH3 is not sufficient to organize a complete kinetochore (Ahmad and Henikoff, 2002; Hooser et al., 2001) strongly suggest that we should not rely entirely on CENH3 to identify centromeric DNA. Mis12 is conserved across fungi, plants and animals (pfam05859.1, Mis12). We identified a GSS contig containing the complete maize *Mis12* gene, amplified a full-length cDNA, and cloned it into the pet28a expression vector. By the end of the current granting period we should be able to add anti-Mis12 to our collection of kinetochore antisera, which now includes anti-CENH3, pCENH3 and CENPC.

The experiments described below will help us answer several important questions about the biology of the centromere/kinetochore interface: **1**) Is the inner kinetochore a large homogeneous protein complex, or do parts of the kinetochore interact with different DNAs? 2) Are CentC and CRM the key centromeric DNAs when assayed by the stringent definition of interacting with four different inner kinetochore antisera? **3**) Do some CentC variants have a higher affinity for CENH3 than others? **4**) Do the kinetochores in interphase occupy the same DNA sequences as those in active, metaphase chromosomes? **5**) What are the boundaries of the kinetochore on chromosome 4?

I) Stretched-fiber assays, interphase and metaphase. We have developed a technique to co-localize proteins and DNA sequences on stretched chromatin fibers (Jin et al., 2004; Fig. 2). We plan to use this technique extensively in future work to reveal whether CENH3 co-localizes with CENPC and Mis12, and to determine the association between CENH3/CENPC/Mis12 and CentC/CRM sequences. In addition we plan to extend the stretched chromatin technique from interphase (Jin et al., 2004) to metaphase, so that we can include anti-pCENH3 antisera in the analysis (pCENH3 is only present at metaphase, Fig. 3). Using a combination of mechanical and hypotonic treatments, Zinkowski et al (1991) were able to stretch human metaphase kinetochores up to 20 times their normal size. Since metaphase chromosomes tend to disperse (spread away from each other) in cytological preparations, it will be much easier to identify individual centromeres in metaphase than interphase (where they tend to tangle together). Reagents are now available that allow the rapid identification of maize chromosomes. In addition to the Cent4 probe for chromosome 4 (Page et al., 2001), the Birchler lab has developed a set of repetitive DNA probes that uniquely identify all 10 maize chromosomes (Kato et al, in prep).

II) Interphase-metaphase comparisons. An important question is whether the interphase kinetochores that others and we have used to define centromeres (e.g. Meluh and Koshland, 1997; Vafa and Sullivan, 1997; Lo et al., 2001; Alonso et al., 2003) accurately represent the active metaphase

kinetochores that generate chromosome movement. To this end we will treat roots with the microtubule-destabilizing drug oryzalin, which increases the number of pCENH3-stained metaphase cells by roughly 5X and the protein to levels detectable on western blots. By comparing the immunoprecipitation of single-copy markers (or CentC variants; below) obtained with non-p CENH3 antisera to the immunoprecipitation obtained with pCENH3 antisera, we will be able to provide a firm answer to this question. If we succeed, we will provide the first direct comparison between interphase and metaphase centromeres.

III) ChIP assays of single-copy markers on centromere 4. Using a sensitive PCR assay to analyze ChIP results (Nagaki et al., 2004) we will determine whether the single-copy markers identified on chromosome 4 interact with our four kinetochore antisera. CentC and CRM will be used as positive controls, and knob repeats as well as a collection of genes outside centromeric regions will serve as negative controls. These studies will allow us to map the complete kinetochore onto the continuous map of centromere 4. It is possible that very few of the single copy markers will show evidence of interaction with the kinetochore. This would be an interesting result, showing that for centromere 4 only the repeated DNAs interact with maize kinetochores (unlike in rice; Nagaki et al., 2004). It will not, however, address the general question of whether all four antisera identify the same sequences. Should this be the case, we will take alternative routes to answer this question. One avenue will be to use the single copy markers from other chromosomes identified in AIM1A. Assuming a subset of these markers is centromeric, the data should allow us to assess the concordance among different antisera.

IV) Analysis of CentC variants. An outcome of our survey sequencing will be a large collection of CentC sequences, organized with respect to chromosome and map position. A preliminary analysis of CentC (below) indicates that there are at least 7 single-nucleotide polymorphisms (SNPs) that are present in 10% of the repeats or more, and others are likely to be identified as we proceed with our sequencing. We will treat each CentC SNP as a potentially adaptive change that alters the interaction of the repeat with CENH3, CENPC, or Mis12. To assay for interactions with the kinetochore, we will make use of quantitative SNP detection assays (e.g. Rudi and Holck, 2003; Xiao and Kwok, 2003). In one protocol known as QUEX (Xiao and Kwok, 2003), a fluorescently labeled primer and an opposing non-labeled primer are first used to amplify all repeats. A second round of PCR then incorporates a chain-terminating quenching dye at the polymorphic site, which can be monitored using quantitative PCR equipment (available on the University of Georgia campus). QUEX was reported to accurately detect polymorphisms at frequencies as low as 5%. We will assay each SNP by comparing the frequency in precipitated DNA to its frequency in the input sample. For example if we find that a T at position 30 is present in 11% of the CentC monomers of the input sample, but 25% of the monomers precipitated by anti-CENH3, we will conclude that T30 is associated with domains that have high affinities for CENH3. By comparing affinity data to map positions, we will be able to test the idea that high-affinity repeats are more likely to lie in central domains of the centromere, as predicted by the 'centromere core' concept.

GOAL2: A COMPETITION ASSAY FOR CENTROMERE FUNCTION

To explain the rapid evolution of centromeres it has been proposed that homologous centromeres compete for access to female reproductive cells by a meiotic drive-like mechanism (Malik and Henikoff, 2002). The mechanism for preferential recovery is presumed to be analogous to the meiotic drive observed for maize Ab10 (Rhoades, 1952) where heterochromatic knobs are directed to the basal cell of the linear tetrad (the only cell that becomes an egg, shown in Figure 7). However, an equally plausible scenario for preferential recovery is that when an inferior centromere is placed in competition with a superior centromere, the inferior centromere is simply lost more often during chromosome alignment and segregation (Figure 7). Under both models centromere competition drives rapid changes in satellite arrays and the inner kinetochore proteins that interact with them.

If we place a 'superior' and an 'inferior' centromere together and test cross the plant, the superior centromere should be recovered at higher frequencies. This expectation was confirmed in 1992 by Carlson and Roseman (1992), who test crossed the heterozygote of a small B centromere derivative and its progenitor chromosome (Telo 2-2/TB-9Sb X wild type). The progenitor chromosome was preferentially recovered 27 to 10, a ratio that deviates significantly from the 1:1 Mendelian expectation. In the intervening years, the Birchler lab has accumulated a much larger collection of well-characterized B centromere derivatives (Kaszas and Birchler, 1996; Kaszas and Birchler, 1998) that have different sized centromeres (which can be recognized based on diagnostic FISH hybridization patterns) and correspondingly different meiotic transmission rates. It is important to emphasize that all features of these chromosomes are identical except for the size and content of the centromeres. The competitive assays

Table 1. Combinations of chromosomes to be tested in centromere competition assays.

	Deriv-ative [1]	Transm ission[2]	Size (kb) [3]
Pair 1	Telo 4-5	44 ± 6	2180
	Telo 4-4	10 ± 3	490
Pair 2	Telo 2-2	43 v 6	2150
	TB-9Sb*	ND	9000
Pair 3	Telo 2-2	43 ± 6	2150
	Telo 3-3	13 ± 2	280
Pair 4	Telo 2*	46 ± 10	3235
	Telo 3-5*	16 ±4	1665
Pair 5	Telo 2*	46 ± 10	3235
	Telo 2-2	43 ± 6	2150
Pair 6	Telo 4-11	42 ± 5	2360
	Telo 4-4	10 ±3	490
Pair 7	Telo 4-5	44 ± 6	2180
	TB-9Sb*	ND	9000
Pair 8	Telo 2*	46 ±10	3235
	TB-9Sb*	ND	9000

[1] Some derivatives (those designated with an asterisk) undergo nondisjunction at the second pollen mitosis. This can be factored into calculations.

[2] Transmission rates refer to the recovery of the univalent.

[3] Centromere size was estimated by the sum of *Pme*I fragments that contain the B repeat (a sequence specific to B centromeres; Kaszas and Birchler, 1998).

described below cannot be carried out using transgenic approaches (below), since trans-centromeres will map to different positions and as a result cannot be compared in a head-to-head fashion.

I) Determining whether the mechanism is meiotic drive or chromosome loss.

We plan to test cross at least eight B centromere heterozygotes (Table 1). Transmission frequencies for each centromere derivative will be determined by dot blot analysis of each progeny population using a common standard such as 5S rDNA, a technique we have used before to discriminate among derivatives (Kaszas and Birchler, 1996; Kaszas and Birchler, 1998). A minimum of 400 progeny from each cross will be analyzed, which is sufficient to detect a 5% segregation advantage. The rate of loss will also be determined for each B centromere homozygote. If the two centromeres exhibit the same transmission rates from a heterozygote, then the hypothesis that centromere competition is a factor in centromere evolution is refuted (unlikely given the prior results of Carlson and Roseman). Under the more likely scenario that a "superior" centromere exhibits preferential transmission, a key measure is whether the difference is more pronounced in females than in males. If the superior centromere shows a greater transmission through the female and no evidence of chromosome loss (below), the data will be consistent with the meiotic drive hypothesis (Malik and Henikoff, 2002).

The timing and frequency of chromosome loss during meiosis will be determined both genetically and cytologically. Genetically, we can estimate chromosome loss by measuring pollen and ovule abortion. Cytologically, FISH analysis at various stages of meiosis in the male will allow us to determine the behavior of the two B centromere derivatives in a heterozygote. Chromosome loss will be evident as laggards during anaphase I or II, or as micronuclei in tetrads. A minimum of three plants will be examined from each combination, and hundreds of meioses will be examined from

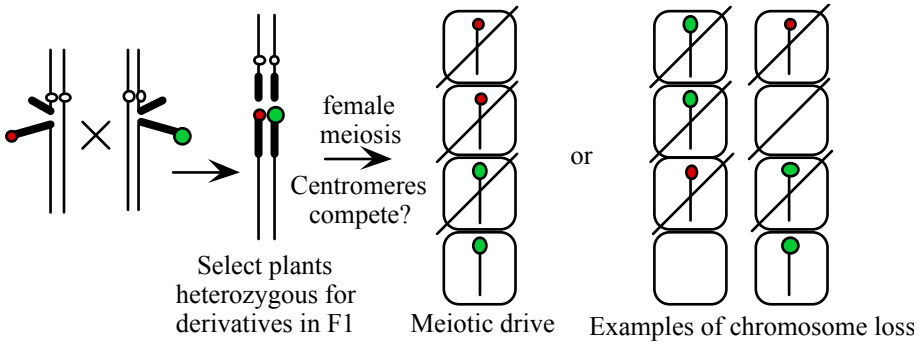


Figure 7. Genetic tests for centromere competition. The superior B centromere is shown in green and the inferior in red. Only the basal cell of the linear tetrad becomes an egg. If the superior centromere shows meiotic drive it will be recovered preferentially in the basal cells with no measurable gamete death. If chromosome loss occurs, the superior chromosome will be preferentially recovered due to loss of the inferior chromosome (causing gamete death). Only the half translocation carrying the B centromere is depicted in the products of meiosis.

each plant. New combinations of the most successful centromeres will be paired up in a tournament style, such that by the end of the granting period we will have identified the most successful centromere derivative of the eight under study.

II) Correlating physical features of the centromere/kinetochore complex with centromere success.

By scoring for the presence or absence of known centromere repeats we can confirm the correlations we have already made between sequence and function, and begin to quantify their effects. In addition, the enlarged sample size and the power fiber-FISH will allow us to determine if particular arrangements of repeats are associated with superior centromeres. For instance, we may be able to test genetically the widely accepted (though never tested) idea that long satellite arrays are required for accurate chromosome segregation. In addition, we will make use of the new protein reagents at our disposal. Antibodies for kinetochore proteins such as CENH3, pCENH3, CENPC, and Mis12 will allow us to determine whether transmission rate is correlated with the quantity of any particular kinetochore protein. These experiments should also give a clear idea of what repeats should be combined to create the smallest, most effective centromere.

GOAL3: DEVELOPING ARTIFICIAL CENTROMERES

We have already shown that we can introduce long centromeric arrays into rice cells and plants (Fig. 5). The goal of the next granting period will be to identify conditions necessary to 'activate' these introduced centromeric DNAs such that they become functional centromere/kinetochore complexes. Towards this aim, we were particularly impressed by the results of Nakano et al. (2003) who studied the activity of an ectopic alpha-satellite array on human chromosome 16. Initially, the trans-centromere did not show centromeric activity, as indicated by the absence (or reduced levels) of CENH3 (CENP-A) and CENP-C. However, long-term culture in selective medium (making essential the expression of the selectable marker gene) or short-term treatment with the histone deacetylase inhibitor Trichostatin A (TSA), promoted the assembly of CENP-A and CENPC at the trans-centromere and the release of minichromosomes containing the YAC integration site. TSA treatment increased both the acetylation of histone H3 and the transcriptional level of the marker gene. The simplest interpretation of these data is that inactivation by heterochromatinization is a major obstacle in establishing a functional centromeric satellite array. The studies described below are designed to rigorously test this hypothesis. Since somatic embryogenesis tends to screen out genetic abnormalities that would interfere with the regeneration process (Vasil, 1987; like ectopic centromere activity), we will begin our studies in transgenic cell lines rather than transgenic plants. In cell culture, transcriptional inducers and TSA can be applied with accuracy, and cells lines are easily bulked up for repeated experiments. Another strength of this approach is that our stretched chromatin fiber assay so far only works on cultured cells (Jin et al., 2004).

A significant problem when introducing cloned centromeric DNA back into the same species is the difficulty in discerning the introduced DNA from the centromeric DNA already present. ChIP assays and even standard FISH assays are confounded by this fact. One option for overcoming this problem is to deliberately introduce polymorphisms that differentiate the introduced from the native centromeres. However, since we do not yet understand the molecular interactions between satellites and kinetochores, we run the risk of changing key nucleotides that are required for centromere-kinetochore interactions. Fortunately, a close relative of maize, *Tripsacum dactyloides*, can be used to identify functionally neutral polymorphisms. Mangelsdorf and Reeves first demonstrated that the cross of maize and *Tripsacum* yields viable progeny (Mangelsdorf and Reeves, 1939). Few wide crosses have been studied as intensively (with over 200 publications listed on maizegdb.org). Maize centromeres function perfectly in *Tripsacum* and vice versa (Mangelsdorf and Reeves, 1939), and maize-*Tripsacum* addition lines have been maintained and analyzed for their phenotypic, cytological, and genetic consequences (DeWet, 1991). Southern analysis indicates that *Tripsacum* CentC homologs (CentT) can be readily

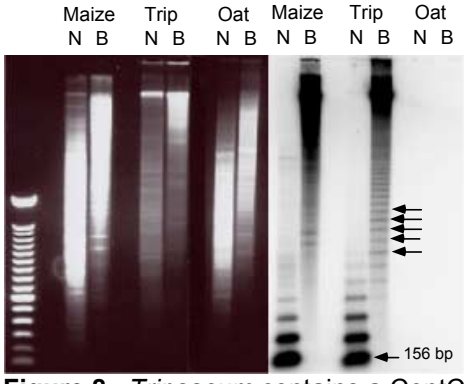


Figure 8. *Tripsacum* contains a CentC-like sequence. Various genomic DNAs were digested with NspI (N) and BamHI (B), blotted, and probed with CentC. *Tripsacum* contains strongly hybridizing bands similar to maize, as well as a BamHI polymorphism (series of arrows) not found in maize.

detected with the maize CentC probe (Figure 8). Since *Zea* diverged from *Tripsacum* roughly 4.5-4.8 million years ago (Hilton and Gaut, 1998), we expect significant polymorphism between maize CentC and *Tripsacum* CentT. The data in Figure 8 provide evidence for the predicted divergence: in *Tripsacum* CentT there is a *Bam*HI polymorphism that is not present in CentC (arrows).

AIM3A) Developing constructs and materials

A) Existing trans-centromeres in rice. As described above, we have already characterized several independently derived rice plants containing trans-centromeres. These include 6 plants containing the rice centromeric BAC 17p22 and 2 plants containing maize centromeric BAC 16H10. We have not yet determined the extent to which these trans-centromeres are inactive or active. Our first step will be to determine whether can detect anti-CENH3 staining at the integration sites (anti-rice CENH3 antisera are available; Nagaki et al., 2004). In addition, we will conduct ChIP-PCR experiments, which the Jiang lab used previously to identify singly-copy sequences in centromere 8 (Nagaki et al., 2004). The BAC vector and other unique sequences within the transformation constructs will be used for primer design. Ideally, we will find lines that are both active and inactive, providing us with both a positive control and material to use for improving the efficiency of the process.

B) An inducible system for activating transcription over artificial repeat arrays.

Most CentC and CentT repeats collapse into monomers when digested with *Nsp*I (Fig. 8). In preliminary experiments, we sequenced 72 CentC repeats by cutting the 156 bp band out of *Nsp*I genomic digests, cloning the sequences into plasmids, and identifying positive clones by colony hybridization (GenBank AY530216- AY530287). This method provides a better representation of CentC monomers than PCR, which selectively recovers clones with strong homology to the primers used (primers occupy a large fraction of a 156 bp repeat). Sequence analysis revealed a clear CentC consensus as well as 7 common SNPs that will be used under AIM1C-IV. We will reproduce these experiments in *Tripsacum* to identify a CentT consensus. With this, we will: 1) Design and test primers that amplify CentT but not CentC; 2) Design a synthetic CentC/T sequence that can be differentiated from CentC by PCR, and have it machine-synthesized; and 3) Prepare a synthetic array of the CentC/T repeat. We will follow the protocol of Ohzeki et al (2002), who prepared an array of human alpha-satellites by repeated cloning of larger arrays (from 1 to 2, 4, 8, 16, and finally 32 repeats in a series of 6 plasmids). 32 repeats were sufficient in humans, but we will increase it to 64 if possible. Stratagene SURE® cells, which lack all DNA repair mechanisms, will be used to stabilize the repeat array.

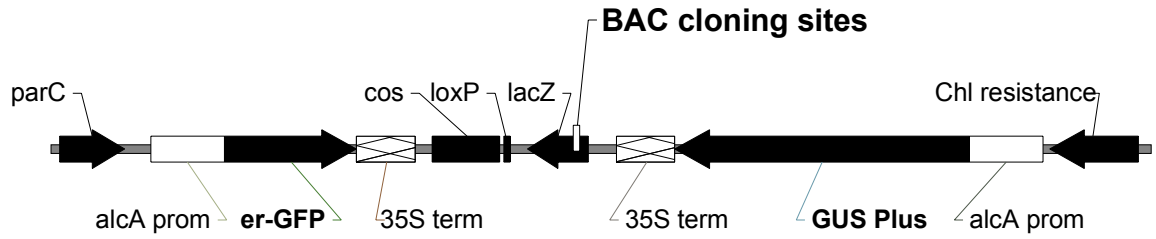


Figure 9. An inducible centromere construct. A CentC/T array or other centromeric sequences will be cloned into the region labeled 'BAC cloning sites'.

A special construct will be prepared to induce gene expression in and around centromeric sequences (Fig. 9). There are several inducible promoters available, but we have chosen the *alc* system from *Aspergillus nidulans* (Roslan et al., 2001), which is available from Syngenta for research purposes. It is highly induced by ethanol and acetaldehyde, without any apparent metabolic side effects (Junker et al., 2003). On one end of the CentC/T array we will insert the *alc-GUSPlus* gene (www.cambia.org) and at the other end we will insert an *alc-erGFP* gene. The effectiveness of the inducers will be determined using GUS and GFP assays. This 'inducible centromere' construct will be transformed by microprojectile bombardment into maize cell lines along with a separate plasmid containing a selectable marker. Having the selectable marker and alcohol response elements on separate plasmids will help avoid the confounding effects of constitutively expressed genes near centromeric DNA.

The modular nature of the construct, and the fact that rapid assays are possible in cultured cells, will allow us considerable flexibility. Instead of a synthetic CentC/T array, we can insert combinations of

CR elements and CentC/T arrays, or long native centromeric sequences. We can also remove the polyA addition sites from one or the other reporter genes to allow transcription to proceed over the CentC/T repeats.

AIM3B) Manipulation of the cell culture environment to allow centromere activation.

A) *TSA*. Nakano et al (2003) showed that TSA treatment increased the formation of active kinetochores over ectopically integrated alpha-satellite DNA. We propose to conduct similar experiments on the transgenic rice strains developed during our current funding period, and on cell lines carrying synthetic satellite arrays. In the existing rice plants, we are particularly interested in determining if both the ectopic CentC (BAC16H10) and CentO (BAC17P22) arrays show evidence of centromeric activity. Should we be able to activate maize BAC16H10 in rice plants, we will have taken the first step towards building a pan-cereal artificial centromere. In the event that we fail to see any affects when seedlings are treated, we will re-initiate callus and treat the resulting cell lines.

B) *Modulating transcription levels by varying the inducer.* Transcription from the *alc* promoter is induced by either ethanol or acetaldehyde in a strictly dose-dependent manner (Junker et al., 2003; Roslan et al., 2001). Various concentrations will be explored and monitored using GUS and GFP assays. The motivation for varying inducers is the prior data of Nakano et al (Nakano et al., 2003), which suggest that low levels of transcription were more effective than high levels in activating centromeres in human cell lines.

C) *Transcription combined with CENH3 overexpression.* Others and we have proposed that the incorporation of CENH3 at centromeres coincides with high-level expression of CENH3 (Shelby et al., 2000; Jiang et al., 2003). To test this model, we will build constructs where the *alc* promoter drives expression of the maize *CenH3* gene coincident with the induction of transcription over maize centromeric DNA. The combined burst of protein and transcription should maximize our chances of establishing active centromeres, which once established are likely to be stable.

AIM3C) Assays for CENH3 binding to trans-centromeres

A) *ChIP assays in cell lines*. The primary measure of centromere activity will be anti-CENH3-mediated ChIP (Zhong et al., 2002) followed by CentC/T-specific PCR (maize cell lines) or anchored PCR using BAC sequences (rice plants). ChIP is an inherently quantitative technique, especially when combined with quantitative or semi-quantitative PCR assays. In cases where positive results are obtained we will extend the assays to include CENPC and Mis12 antisera. These data will allow us to determine whether any of the treatments applied above facilitate the recruitment of kinetochore proteins to trans-centromeres.

B) *Extended fiber assays in cell lines*. Positive ChIP results will be followed with stretched chromatin assays. Depending on the degree of polymorphism between CentC/T and CentC, we may be able to use oligonucleotide probes to specifically identify the CentC/T arrays. However, we can also use flanking sequences (GUS, GFP, BAC vector) to identify trans-centromeres. By combining CENH3 localization with FISH, we should be able to confirm biochemical observations.

C) *Cytological assays*. A strength of our group is that all of us have expertise with standard cytological analyses (Parrott did cytogenetics research for his Ph.D. and teaches it at UGA). We intend to analyze callus (or meiotic cells in rice plants) to see if active trans-centromeres disrupt the segregation of chromosomes, as expected for effectively dicentric chromosomes. We will also study the progeny of transgenic plants. Progeny analysis will allow us to determine whether trans-centromeres are transmitted in a Mendelian fashion, and whether minichromosomes are released from their ectopic integration sites as observed in human cell lines (Nakano et al., 2003).

D) *ChIP assays in transgenic plants.* We view cell lines as a first step towards understanding the parameters required for maintaining artificial centromeres in growing plants. We have already transformed rice plants carrying native centromeric sequences. Once we have a better understanding of the procedures necessary to activate synthetic satellite arrays, we will proceed to regenerate maize plants containing these constructs. The Parrott and Birchler labs are adept at maize transformation. Since the *alc* inducers can be applied in vapor as well as aqueous form, we should be able to reproduce the cultured cell conditions in growth chambers containing regenerated plants. The resulting plants will be assayed for centromere activity by cytological criteria and by biochemical interactions with our collection of inner kinetochore antisera. We will also assay for the presence of the outer kinetochore proteins MAD2 and 3F3/2 antigen (Yu et al., 1999), both of which are found only on mature fully functional centromere/kinetochore complexes (Yu et al., 2000).

PLAN TO INTEGRATE RESEARCH AND EDUCATION

Minority outreach

All Co-PIs will take advantage of their respective University-sponsored minority recruitment programs. We have been very successfully in previous years, training 5 African Americans and 2 Americans of Hispanic descent. However, under this award we will put a strong emphasis on minority involvement at the University of Georgia. UGA is in the enviable position of having far more students from historically under represented groups than there are research opportunities for them. One of the current programs is the Summer Undergraduate Research Program (SURP), the mission of which is to address the national shortage in ethnic minorities in higher education. *SURP currently has funds for 40 interns per summer, yet receives between 140 and 160 applications per year*. In addition to research experience, students are provided with team building and skill-enhancing activities such as social receptions, writing workshops, graduate school information and GRE training. UGA also sponsors the Center for Undergraduate Research Opportunities (CURO-AP). Like SURP, the CURO Apprenticeship Program has far more applicants than it has funds to accommodate.

The plant genomics faculty at UGA (www.plantcenter.uga.edu) voted to establish a new undergraduate research initiative that will advance plant research on campus. The program, called the Georgia Research Opportunities for Undergraduates Program (GROUP), is being initiated with the support of the Office for Institutional Diversity, the Graduate School, Admissions Office, and the Center for Undergraduate Research Opportunities (see letter of support from Keith Parker, Associate Provost for Institutional Diversity). These offices will also provide active assistance with the recruitment process. GROUP will share elements of both SURP and CURO-AP, and may eventually grow to encompass them. A key component of this program is that students have a research project that is their own, and are therefore responsible for its success or failure. Students working over the summer receive a $4,500 stipend, which they use to cover room and board and all other activities, plus $1000 for the cost of team-building and training activities. Since this proposal is among the first to be submitted under GROUP, it will help support the inaugural class. We are requesting funds for 3 GROUP participants each year, one year-round student and two during summers only.

Annual cytogenetics training workshop

Chromosome- and cytology-based analyses were originally at the core of genetics, and plants like *Datura* and maize were among the first models because of their large chromosomes. However, it wasn't until FISH coupled with digital cameras/3D light microscopy arrived in the 1990s, that the small-genome barrier was broken and cytogenetics became fully integrated into the field of plant genetics (e.g. Fransz et al., 2000). Special techniques have been developed to fix meiotic samples for 3D studies of both DNA and proteins (Dawe et al., 1994; Yu et al., 1997), providing a significant advantage over mammalian systems in which meiotic samples are difficult to obtain and examine. Further genome sequencing will only increase the demand for cytological skills, as more and more researchers recognize the need to localize genes and gene products in the cell.

Many of the new cytological methods are highly technical. Scientists who intend to use these techniques need special training in well-established labs. Currently there is no annual training course or workshop available in the plant cytogenetics community. To address this need, we propose to organize an Annual Cytogenetics Training Workshop (ACTW). The specific arrangements and needs of the community were discussed during the plant cytogenetics workshop on Jan. 11, 2004 in San Diego (http://www.intl-pag.org/12/12-cyto.html). The ACTW will offer training in the following techniques: 1) FISH using mitotic and meiotic samples; 2) Fiber-FISH; 3) Immunoassays on mitotic and meiotic preparations; 4) 3D microscopy using the deconvolution system; 5) Protein mapping on stretched chromatin fibers; and 6) Chromatin immunoprecipitation. We will also invite guest lecturers to describe additional techniques that neither of the instructors are familiar with. The workshop will be a five-day event and host 10 people/year. Choices of techniques to be taught will depend on the stated interests of the participants.

Co-PI Jiang will hold the ACTW in his lab at the University of Wisconsin-Madison. Jiang will be the primary teacher for the workshop. PI Dawe will also participate in the teaching in the first, third, and fifth years of this project. Postdoctoral associates and graduate students working on this project in the Jiang lab will assist in technical training, and a dedicated coordinator will be hired for one month of each year to handle advertising, travel, and daily managerial duties. Ten Nikon Labophot biological microscopes and two Olympus fluorescent microscope workstations with CCD cameras/computers will be available for ACTW events.

REFERENCES

Ahmad, K., and S. Henikoff. 2001. Centromeres are specialized replication domains in heterochromatin. *J. Cell Biol.* 153:101-109.

Ahmad, K., and S. Henikoff. 2002. Histone H3 variants specify modes of chromatin assembly. *Proc. Nat. Acad. Sci. USA*. 99:6477-6484.

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

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

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. USA*. 95:13073-13078.

Ando, S., H. Yang, N. Nozaki, T. Okazaki, and K. Yoda. 2002. CENP-A, -B, and -C chromatin complex that contains the I-type alpha-satellite array constitutes the prekinetochore in HeLa cells. *Mol. Cell Biol.* 22:2229-2241.

Aragon-Alcaide, L., T. Miller, T. Schwarzacher, S. Reader, and G. Moore. 1996. A cereal centromeric sequence. *Chromosoma*. 105:261-268.

Birchler, J.A., D.L. Auger, and A. Kato. 2004. Cytogenetics of corn. *In* Corn: Origin, History, Technology and Production. W. Smith, J. Betran, and E. Runge, editors. John Wiley and Sons, New York.

Blower, M., B. Sullivan, and G. Karpen. 2002. Conserved organization of centromeric chromatin in flies and humans. *Dev. Cell*. 2:319-330.

Brown, R.A., P.J. Mee, and M.H. Shen. 2000. Artificial chromosomes: ideal vectors? *Trends Biotech.* 18:218-223.

Carlson, W.R., and R.R. Roseman. 1992. A new property of the maize B chromosome. *Genetics*. 131:211-223.

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

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

Clarke, L., and J. Carbon. 1980. Isolation of a yeast centromere and construction of functional small circular chromosomes. *Nature*. 187:504-509.

Copenhaver, G.P. 2003. Using Arabidopsis to understand centromere function: progress and prospects. *Chromosome Res.* 11:255-262.

Darlington, C.D. 1937. Misdivision and the genetics of the centromere. *J. of Genetics*. 37:342-364.

Dawe, R.K. 2003. RNA interference, transposons, and the centromere. *Plant Cell*. 15:297-301.

Dawe, R.K., L. Reed, H.-G. Yu, M.G. Muszynski, and E.N. Hiatt. 1999. A maize homolog of mammalian CENPC is a constitutive component of the inner kinetochore. *Plant Cell*. 11:1227-1238.

Dawe, R.K., J.W. Sedat, D.A. Agard, and W.Z. Cande. 1994. Meiotic chromosome pairing in maize is associated with a novel chromatin organization. *Cell*. 76:901-912.

DeWet, J.R. 1991. Cytogenetics of Zea and Tripsacum. *In* Chromosome engineering in plants : genetics, breeding, evolution. P.K. Gupta and T. Tsuchiya, editors. Elsevier, New York. 483-492.

Ding, Y., M.D. Johnson, W.Q. Chen, D. Wong, Y.J. Chen, S.C. Benson, J.Y. Lam, Y.M. Kim, and H. Shizuya. 2001. Five-color-based high-information-content fingerprinting of bacterial artificial chromosome clones using type IIS restriction endonucleases. *Genomics*. 74:142-154.

Fransz, P.F., S. Armstrong, J.H.d. Jong, L. Parnell, C.v. Drunen, C. Dean, P. Zabel, T. Bisseling, and G.H. Jones. 2000. Integrated cytogenetic map of chromosome arm 4S of *A. thaliana*: structural organization of heterochromatic knob and centromere region. *Cell*. 100:367-376.

Goshima, G., T. Kiyomitsu, K. Yoda, and M. Yanagida. 2003. Human centromere chromatin protein hMis12, essential for equal segregation, is independent of CENP-A loading pathway. *J. Cell Biol.* 160:25-39.

Hannon, G.J. 2002. RNA interference. *Nature*. 418:244-251.

Harrington, J.J., G.V. Bokkelen, R.W. Mays, K. Gustashaw, and H.F. Willard. 1997. Formation of de novo centromeres and construction of first-generation human artificial chromosomes. *Nature Genetics*. 15:345-355.

Henikoff, S. 2002. Near the edge of a chromosome's 'black hole'. Trends Genet. 18:165-167.

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.

Hooser, A.V., I. Ouspensik, H. Gregson, D. starr, T. Yen, M. Goldberg, K. Yokomori, W. Earnshaw, K. Sullivan, and B. Brinkley. 2001. Specification of kinetochore-forming chromatin by the histone H3 variant CENP-A. *J. Cell Sci.* 114:3529-3542.

Houben, A., and I. Schubert. 2003. DNA and proteins of plant centromeres. *Curr. Opin. Plant Biol.* 6: 967-973

Hudakova, S., W. Michalek, G.G. Presting, R. tenHoopen, K.d. Santos, Z. Jasencakova, and I. Schubert. 2001. Sequence organization of barley centromeres. Nucleic Acids Res. 29:5029-5035.

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

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

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

Jin, W., J.R. Melo, K. Nagaki, P.B. Talbert, S. Henikoff, R.K. Dawe, and J. Jiang. 2004. Maize centromeres: fine structure and functional adaptation in the genetic background of oat. *Plant Cell*. in press.

Junker, B.H., C.C. Chu, U. Sonnewald, L. Willmitzer, and A. Fernie. 2003. In plants the *alc* gene expression system responds more rapidly following induction with acetaldehyde than with ethanol. *FEBS Lett.* 535:136-140.

Karpen, G.H., and R.C. Allshire. 1997. The case of epigenetic effects on centromere identity and function. *Trends Genet.* 13:489-496.

Kaszas, E., and J.A. Birchler. 1996. Misdivision analysis of centromere structure in maize. *EMBO J.* 15:5246-5255.

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

Kaszas, E., A. Kato, and J. Birchler. 2002. Cytological and molecular analysis of centromere misdivision in maize. *Genome*. 45: 759-768

Kazakov, A.E., V.A. Shepelev, I.G. Tumeneva, A.A. Alexandrov, Y.B. Yurov, and I.A. Alexandrov. 2003. Interspersed repeats are found predominantly in the "old" alpha satellite families. *Genomics*. 82:619-627.

Kimmel, B.E., M.J. Palazzolo, C.H. Martin, J.D. Boeke, and S.E. Devine. 1997. Transposon-mediated DNA sequencing. *In* Genome Analysis: A laboratory Manual. Vol. 1. B. Birren, E.D. Green, S. Klapholz, R.M. Myers, and J. Roskams, editors. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY. 455-532.

Kumekawa, N., T. Hosouchi, H. Tsuruoka, and H. Kotani. 2000. The size and sequence organization of the centromeric region of *Arabidopsis thaliana* chromosome 5. *DNA Res.* 7:315-321.

Kynast, R.G., O. Riera-Lizarazu, M.I. Vales, R.J. Okagaki, S.B. Maquieira, G. Chen, E.V. Ananiev, W.E. Odland, C.D. Russell, A.O. Stec, S.M. Livingston, H.A. Zaia, H.W. Rines, and R.L. Phillips. 2001. A complete set of maize individual chromosome additions to the oat genome. *Plant Physiol.* 125:1216-1227.

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

Lo, A., J. Craig, R. Saffery, P. Kalitsis, D. Irvine, E. Earle, D. Magliano, and K. Choo. 2001. A 330 kb CENP-A binding domain and altered replication timing at a human neocentromere. *EMBO J.* 20:2087-2096.

Maison, C., D. Bailly, A.H. Peters, J.P. Quivy, D. Roche, A. Taddei, M. Lachner, T. Jenuwein, and G. Almouzni. 2002. Higher-order structure in pericentric heterochromatin involves a distinct pattern of histone modification and an RNA component. *Nat. Genet.* 30:329-334.

Malik, H., D. Vermaak, and S. Henikoff. 2002. Recurrent evolution of DNA-binding motifs in the *Drosophila* centromeric histone. *Proc. Nat. Acad. Sci. USA*. 99:1449-1454.

Malik, H.S., and S. Henikoff. 2002. Conflict begets complexity: the evolution of centromeres. *Curr. Opin. Genet. Dev.* 12:711-718.

Mangelsdorf, P.C., and R.G. Reeves. 1939. The origin of Indian corn and its relatives. *Texas Agric. Exp. Sta. Bull.* 574:1-315.
Meluh, P.B., and D. Koshland. 1997. Budding yeast centromere composition and assembly as revealed by in vivo cross-linking. *Genes and Development*. 11:3401-3412.
Meluh, P.B., P. Yang, L. Glowczewski, D. Koshland, and M.M. Smith. 1998. Cse4p is a component of the core centromere of Saccharomyces cerevisiae. *Cell*. 94:607-613.
Mroczek, R.J., and R.K. Dawe. 2003. Distribution of retroelements in centromeres and neocentromeres of maize. *Genetics*. 165:809-819.
Murray, A.W., N.P. Schultes, and J.W. Szostak. 1986. Chromosome length controls mitotic chromosome segregation in yeast. *Cell*. 45:529-536.
Nagaki, K., Z. Cheng, S. Ouyang, P.B. Talbert, M. Kim, K.M. Jones, S. Henikoff, C.R. Buell, and J. Jiang. 2004. Sequencing of a rice centromere reveals active genes. *Nature Genet.* Jan. 11
Nagaki, K., J. Song, R. Stupar, A.S. Parokonny, Q. Yuan, S. Ouyang, J. Liu, R.K. Dawe, C.R. Buell, and J. Jiang. 2003a. 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., P.B. Talbert, C.X. Zhong, R.K. Dawe, S. Henikoff, and J. Jiang. 2003b. Chromatin Immunoprecipitation reveals that the 180-bp satellite repeat Is the key functional DNA element of *Arabidopsis thaliana* centromeres. *Genetics*. 163:1221-1225.
Nakano, M., Y. Okamoto, J. Ohzeki, and H. Masumoto. 2003. Epigenetic assembly of centromeric chromatin at ectopic alpha-satellite sites on human chromosomes. *J. Cell Sci.* 116:4021-4034.
Ohzeki, J., M. NakanoT., Okada, and H. Masumoto. 2002. CENP-B box is required for de novo centromere chromatin assembly on human alphoid DNA. *J. Cell Biol.* 159:765-775.
Okagaki, R.J., R.G. Kynast, S.M. Livingston, C.D. Russell, H.W. Rines, and R.L. Phillips. 2001. Mapping maize sequences to chromosomes using oat-maize chromosome addition materials. *Plant Physiol.* 125:1228-1235.
Page, B., M. Wanous, and J. Birchler. 2001. Characterization of a maize chromosome 4 centromeric sequence: evidence for an evolutionary relationship with the B chromosome centromere. *Genetics*. 159:291-302.
Palmer, L.E., P.D. Rabinowicz, A.L. O'Shaughnessy, V.S. Balija, L.U. Nascimento, S. Dike, M.d.l. Bastide, R.A. Martienssen, and W.R. McCombie. 2003. Maize genome sequencing by methylation filtration. *Science*. 302:2115-2117.
Phelps-Durr, T., and J.A. Birchler. 2004. An asymptotic determination of minimum centromere size for the maize B chromosome. *Cytogenet. Genome Res.* submitted.
Presting, G., L. Malysheva, J. Fuchs, and I. Schubert. 1998. A TY3/GYPSY retrotransposon-like sequence localizes to the centromeric regions of cereal chromosomes. *Plant J.* 16:721-728.
Presting, G.G., A. Frary, K. Pillen, and S.D. Tanksley. 1996. Telomere-homologous sequences occur near the centromeres of many tomato chromosomes. *Mol Gen Genet*. 251:526-531.
Rhoades, M.M. 1952. Preferential segregation in maize. *In* Heterosis. J.W. Gowen, editor. Iowa State College Press, Ames, Iowa. 66-80.
Roslan, H.A., M.G. Salter, C.D. Wood, M.R.H. White, K.P. Croft, F. Robson, H. Coupland, J. Doonan, P. Laufs, A.B. Tomsett, and M.X. Caddick. 2001. Characterization of the ethanol-inducible alc gene-expression system in *Arabidopsis thaliana*. *Plant J.* 28:225-235.
Rudd, M.K., R.W. Mays, S. Schwartz, and H.F. Willard. 2003. Human artificial chromosomes with alpha satellite-based de novo centromeres show increased frequency of nondisjunction and anaphase. *Mol. Cell Biol.* 23:7689-7697.
Rudi, K., and A.L. Holck. 2003. Real-time closed tube single nucleotide polymorphism (SNP) quantification in pooled samples by quencher extension (QEXT). *Nucleic Acids Res.* 31:e117.
Saffery, R., H. Sumer, S. Hassan, L.H. Wong, J.M. Craig, K. Todokoro, M. Anderson, A. Stafford, and K.H. Choo. 2003. Transcription within a functional human centromere. *Mol. Cell*. 12:509-516.
Schueler, M.G., A.W. Higgins, M.K. Rudd, K. Gustashaw, and H.F. Willard. 2001. Genomic and genetic definition of a functional human centromere. *Science*. 294:109-114.
Shelby, R.D., K. Monier, and K.F. Sullivan. 2000. Chromatin assembly at kinetochores is uncoupled from DNA replication. *J. Cell Biol.* 151:1113-1118.
Somerville, C., and S. Somerville. 1999. Plant functional genomics. *Science*. 285:380-385.
Sun, X., H.D. Le, J.M. Wahlstrom, and G.H. Karpen. 2003. Sequence analysis of a functional Drosophila centromere. *Genome Res.* 13:182-194.

Talbert, P., R Masuelli, A. Tyagi, L. Comai, and S. Henikoff. 2002. Centromeric localization and adaptive evolution of an *Arabidopsis* histone H3 variant. *Plant Cell*. 14:1053-1066.

Theuri, J., T. Phelps-Durr, S. Mathews, and J.A. Birchler. 2004. A comparative study of retrotransposons in the centromeric regions of A and B chromosomes of maize. *Cytgenet. Genome Res.* submitted.

Topp, C.N., C.X. Zhong, X. Zhang, and R.K. Dawe. Maize centromeres are transcribed; implications for the mechanism of centromeric histone deposition. (submitted).

Vafa, O., and K.F. Sullivan. 1997. Chromatin containing CENP-A and a-satellite DNA is a major component of the inner kinetochore plate. *Curr. Biol.* 7:897-900.

Vasil, I.K. 1987. Developing cell and tissue culture systems for the improvement of cereal and grass crops. *Plant Physiol.* 128:193-218.

Westermann, S., I.M. Cheeseman, S. Anderson, J.R.Y. 3rd, D.G. Drubin, and G. Barnes. 2003. Architecture of the budding yeast kinetochore reveals a conserved molecular core. *J. Cell Biol.* 163:215-222.

Whitelaw, C.A., W.B. Barbazuk, G. Pertea, A.P. Chan, F. Cheung, Y. Lee, L. Zheng, S.v. Heeringen\, S. Karamycheva, J.L. Bennetzen, P. SanMiguel, N. Lakey, J. Bedell, Y. Yuan, M.A. Budiman, A. Resnick, S.V. Aken, T. Utterback, S. Riedmuller, M. Williams, T. Feldblyum, K. Schubert, R. Beachy, C.M. Fraser, and J. Quackenbush. 2003. Enrichment of gene-coding sequences in maize by genome filtration. *Science*. 302:2118-2120.

Xiao, M., and P.Y. Kwok. 2003. DNA analysis by fluorescence quenching detection. *Genome Res.* 13:932-939.

Yamada, K., J. Lim, J.M. Dale, H. Chen, P. Shinn, C.J. Palm, A.M. Southwick, H.C. Wu, C. Kim, M. Nguyen, P. Pham, R. Cheuk, G. Karlin-Newmann, S.X. Liu, B. Lam, H. Sakano, T. Wu, G. Yu, M. Miranda, H.L. Quach, M. Tripp, C.H. Chang, J.M. Lee, M. Toriumi, M.M. Chan, C.C. Tang, C.S. Onodera, J.M. Deng, K. Akiyama, Y. Ansari, T. Arakawa, J. Banh, F. Banno, L. Bowser, S. Brooks, P. Carninci, Q. Chao, N. Choy, A. Enju, A.D. Goldsmith, M. Gurjal, N.F. Hansen, Y. Hayashizaki, C. Johnson-Hopson, V.W. Hsuan, K. Lida, M. Karnes, S. Khan, E. Koesema, J. Ishida, P.X. Jiang, T. Jones, J. Kawai, A. Kamiya, C. Meyers, M. Nakajima, M. Narusaka, M. Seki, T. Sakurai, M. Satou, R. Tamse, M. Vaysberg, E.K. Wallender, C. Wong, Y. Yamamura, S. Yuan, K. Shinozaki, R.W. Davis, A. Theologis, and J.R. Ecker. 2003. Empirical analysis of transcriptional activity in the Arabidopsis genome. *Science*. 302:842-846.

Yu, H.-G., E.N. Hiatt, A. Chan, M. Sweeney, and R.K. Dawe. 1997. Neocentromere-mediated chromosome movement in maize. *J. Cell Biol.* 139:831-840.

Yu, H.-G., E.N. Hiatt, and R.K. Dawe. 2000. The plant kinetochore. *Trends Plant Sci.* 5:543-547.

Yu, H.-G., M.G. Muszynski, and R.K. Dawe. 1999. The maize homologue of the cell cycle checkpoint protein MAD2 reveals kinetochore substructure and contrasting mitotic and meiotic localization patterns. *J. Cell Biol.* 145:425-435.

Zeitlin, S.G., R.D. Shelby, and K.F. Sullivan. 2001. CENP-A is phosphorylated by Aurora B kinase and plays an unexpected role in completion of cytokinesis. *J. Cell Biol.* 155:1147-1157.

Zhang, X., X. Li, J.B. Marshall, C.X. Zhong, and R.K. Dawe. Phosphorylation of centromeric histone (CENH3) at serine 50 correlates with maturation of the maize inner kinetochore complex. *in preparation*.

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

Zinkowski, R.P., J. Meyne, and B.R. Brinkley. 1991. The centromere-kinetochore complex: a repeat subunit model. *J. Cell Biol*. 113:1091-1110.

BIOGRAPHICAL SKETCH

R. Kelly Dawe Associate Professor, Plant Biology/Genetics

a. Education:

B.S.	1983	Colorado State University	Landscape Horticulture
M.S.	1985	U.C. Riverside	Botany
Ph.D.	1989	U.C. Berkeley	Genetics

b. Appointments:

- Associate Professor, Departments of Botany and Genetics, Univ. Georgia 2001-
- Asst. Professor, Departments of Botany and Genetics, Univ. Georgia 1995-2000
- Postdoctoral fellow, 1/90-12/94, U.C. Berkeley

c. Publications - 5 most closely related:

Dawe, R. K., Reed, L. M., Yu, H.-G., Muszynski, M.G., and Hiatt, E.N. 1999. A maize homolog of mammalian CENP-C is a constitutive component of the inner kinetochore. *Plant Cell* 11, 1227-1238

Yu, H.-G., Muszynski, M. G., and Dawe, R. K. 1999. The maize homolog of the cell cycle checkpoint protein MAD2 reveals kinetochore substructure and contrasting mitotic and meiotic localization patterns. *J. Cell Biol*. 145, 425-435

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

Mroczek, R. J., and Dawe, R.K. 2003. Distribution of retroelements in centromeres and neocentromeres of maize. *Genetics*, 165: 809–819.

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

d. Publications - 5 other:

Dawe, R. K., Sedat, J. W., Agard, D. A., and Cande, W. Z. 1994. Chromosome synapsis in maize is associated with a novel chromatin organization. *Cell* 76, 901-912

Yu, H-G., Hiatt, E. N., Chan, A., Sweeney, M., and Dawe, R. K. 1997. Neocentromere-mediated chromosome movement in maize. *J. Cell Biol*. 139, 831-840

Yu, H.-G., Hiatt, E. N., and Dawe, R.K.. 2000. The plant kinetochore. *Trends Plant Sci*. 5, 543-547

Yu, H.-G. and Dawe, R. K. 2000. Functional redundancy in the maize meiotic kinetochore. *J. Cell Biol*. 151, 131-141

Hiatt, E.N., Kentner, E., and Dawe, R.K. 2002. Independently-regulated neocentromere activity of two classes of satellite sequences in maize. *Plant Cell*, 14, 407-420

e. Synergistic Activities

- Our DeltaVision light microscope workstation is broadly used by other labs, and the PI has participated each year in two courses designed to inform students about current methodologies in microscopy.
- The PI is actively involved in undergraduate education at UGA, having served as an advisor and teacher to hundreds of students. I am currently the undergraduate coordinator for the high-demand Genetics Major at UGA, where I work with students and the upper administration to enhance Genetics training on campus. The lab has hosted nine undergraduates in research internships, including three that qualify as underrepresented minorities.

f. Collaborators & Other Affiliations:

(i) Collaborators

- James A. Birchler, Univ. Missouri
- Edward Buckler, Cornell
- Robin Buell, TIGR
- Vicki Chandler, Univ. Arizona
- Karen Cone, Univ. Missouri
- John F. Doebley, Univ. Wisconsin
- Steven Henikoff, Fred Hutchinson Cancer Research Center
- Jiming Jiang, Univ. Wisconsin
- Heidi Kaeppler, Univ. Wisconsin
- Shawn Kaeppler, Univ. Wisconsin
- Russell Malmberg, UGA
- Richard Meagher, UGA
- Ronald Morris, Robert Wood Johnson Medical School, New Jersey
- Wayne Parrott, UGA
- Craig Pikaard, Washingtion Univ.
- Ingo Schubert, Gatersleben, Germany

(ii) Graduate and Postdoctoral Advisors

- Michael Freeling, doctoral, U.C. Berkeley
- W. Z. Cande, John W. Sedat, and David A. Agard, postdoctoral, U.C Berkeley and U.C. San Francisco

(iii) Thesis advisor and Postgraduate Scholar advisor:

- Hong-Guo Yu Ph. D., 2000
- Joshua Marshall M.S., 2001
- Evelyn Hiatt Ph. D, 2001
- Rebecca Mroczek Ph. D, 2003
- Carolyn Lawrence Ph. D, 2003
- Christopher Topp Ph.D. current
- Jing Wang Ph.D. current

- Xuexian li Ph.D. current
- Cathy Zhong Postdoctoral, 2000-2003

Biographical Sketch

James A. Birchler

EDUCATION:

B.S. 1972 Eastern Illinois University Botany/Zoology
Ph.D 1977 Indiana University Genetics/Biochemistry
Postdoctoral Investigator, 1977-81, Oak Ridge National Lab, Oak Ridge, TN

APPOINTMENTS:

Professor, 1996-present, Associate Professor, 1991-1996; University of Missouri
Visiting Scientist, University of Oxford, 1998
Associate Professor, 1987-91, Assistant Professor, 1985-87; Harvard University
Assistant Research Geneticist, 1983-85, Postdoctoral Research Affiliate, 1982-83
University of California-Berkeley

FIVE PUBLICATIONS RELATED TO THIS PROPOSAL:

J. A. Birchler, D. L. Auger and A. Kato, 2004. Cytogenetics of corn. IN: Corn: Origin, History, Technology and Production. Edited by C. Wayne Smith, Javier Betran and Ed Runge. John Wiley and Sons, New York.

J. A. Birchler, J. Vega and A. Kato, 2003. Molecular Analysis of Chromosome Landmarks. Encyclopedia of Plant and Crop Science. Marcel Dekker, New York, NY.

E. Kaszas, A. Kato and J. A. Birchler, 2002. Cytological and molecular analysis of centromere misdivision in maize. Genome 45: 759-76.

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

A. Kato, Y.-Z. Zheng, D. L. Auger, T. Phelps-Durr, M. J. Bauer, J. C. Lamb and J. A. Birchler, 2004. Minichromosomes derived from the B chromosome of maize. Cytogenetic and Genome Research, submitted.

FIVE OTHER PUBLICATIONS:

M. Pal Bhadra, U. Bhadra and J. A. Birchler, 2002. RNAi related mechanisms affect both transcriptional and post-transcriptional transgene silencing in Drosophila. Molecular Cell 9: 315-327.

B. T. Page, M. K. Wanous and J. A. Birchler, 2001. Characterization of a maize chromosome 4 centromeric sequence: evidence for an evolutionary relationship with the B chromosome centromere. Genetics 159: 291-302.

M. Pal Bhadra, U. Bhadra and J. A. Birchler, 1999. Cosuppression of non-homologous transgenes in Drosophila involves mutually related endogenous sequences. Cell 99: 35-46.

M. Guo and J. A. Birchler, 1994. Trans-acting dosage effects on the expression of model gene systems in maize aneuploids. Science 266: 1999-2002.

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

SYNERGISTIC ACTIVITIES:

1. Associate Editor: Genetics, 1994-present
2. Co-Editor, The Plant Cell, 2003-present
3. Genetics Coordinator, State of Missouri Science Olympiad (Secondary School), 2003-present
4. Named one of five "Teaching Legends", Mizzou Alumni Magazine, 2003
5. Life Sciences Foundation Postdoctoral Fellow Selection Committee, 1988-present
6. Co-Editor, Maize Genetics Cooperation Newsletter, 2000-present.
7. Maize Genetics Executive Committee, 2000-present.
8. Advisory Committee, NSF Maize Genome Sequencing Project, 2002-present.

Conflict of Interest List for James Birchler

Co-authors/Advisees: Mark Alfenito, Don Auger, Matthew Bauer, Elisaveta Benevolenskaya, Utpal Bhadra, Manika Bhadra, Chris Carson, Amy Csink, Jenny Cooper, Doug Davis, Chris Della Vedova, Anjali Dogra, Maxim Frolov, Mei Guo, John Hiebert, Sun Hui, Etienne Kaszas, Akio Kato, Harsh Kavi, Jon Lamb, Boris Leibovitch, Richard Linsk, Mai Nguyen, Brent Page, Tara Phelps-Durr, Nicole Riddle, Joy Sabl, Shoji Sonoda, James Theuri, Juan Vega, Michael Wanous, Laura Woody, Weichang Yu, Yinzhou Zheng

Collaborators: Jeffrey C. Chen, Edward H. Coe, Jr., Luca Comai, Kelly Dawe, Rebecca Doerge, Sarah Elgin, Jiming Jiang, Jan Larsson, Rob Martienssen, Kathleen Newton, Thomas C. Osborn, Wayne Parrot

Advisors: Ed Grell, Bruce Jacobsen, Ken Paigen, Drew Schwartz

Advisory Committee conflicts: Roger Beachy, Jeff Bennetzen, C. M. Fraser, Rob Martienssen, Jo Messing, Richard McCombie, Phillip SanMiguel, John Quakenbush, Karel Schubert, Cari Soderland, Rod Wing

Total number of graduate students advised: 14
Total number of postdoctoral scholars sponsored: 21

Biographical Sketch - Jiming Jiang

Date of Birth: July 6, 1963

Addresses: Department of Horticulture
University of Wisconsin-Madison
Madison, WI 53706
Phone: (608) 262-1878, Fax: (608) 262-4743
Email: jjiang1@wisc.edu

Education:

B.S.	Agronomy	1983	Zhejiang Agri. University
M.S.	Genetics	1986	Nanjing Agri. University
Ph.D.	Genetics	1993	Kansas State University

Academic Positions:

- 93-94, Postdoctoral Associate, Kansas State University
- 94-95, Postdoctoral Associate, Yale University
- 95-00, Assist. Prof.; 00-03, Assoc. Prof.; 03-present, Prof., University of Wisconsin

Publications - 5 most closely related:

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

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.** (2003) Molecular and cytological analyses of large tracks of centromeric DNA reveal the structure and evolutionary dynamics of maize centromeres. *Genetics* 163: 759-770.

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

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. *Nature Genet.* (in press).

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* (in press).

Publications - 5 other:

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

Dong, F., Miller, J.T., Jackson, S.A., Wang, G.-L., Ronald, P.C., and **Jiang, J.** (1998) Rice (*Oryza sativa*) centromeric regions consist of complex DNA. *Proc. Natl. Acad. Sci. USA* 95: 8135-8140.

Lilly, J.W., Havey, M.J., Jackson, S.A., and **Jiang, J.** (2001) Cytogenomic analyses reveal structural plasticity in the chloroplast genome of higher plants. *Plant Cell* 13: 245-254.

Cheng, Z.K., Buell, C.R., Wing, R.A., Gu, M., and **Jiang, J.** (2001) Toward a cytological characterization of the rice genome. *Genome Res*. 11: 2133-2141.

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

Collaborators in the last 4 years:

Barbara Baker	James Birchler	James Bradeen
Charles Brown	Robin Buell	Kelly Dawe
Molly Jahn	John Helgeson	Steve Henikoff
Georgiana May	Wayne Parrott	Rod Wing

Postdoctoral associates:

Name	Period	Current position
Joseph T. Miller	97-98	Scientist, University of Iowa
Zhukuan Cheng	99-02	Professor, Chinese Academy of Science, China
Junqi Song	97-03	Postdoctor, Yale University
Kiyotaka Nagaki	01-03	Asisstant Professor, Okayama Univ., Japan
Weiwei Jin	02-present	
Huihuang Yan	03-present	

Graduate students:

Name	Period	Degree	Current position
Scott A. Jackson	96-99	Ph.D.	Assistant Professor, Purdue University
Fenggao Dong	95-00	Ph.D.	Scientist, Monsanto Company
Karen Bresee	00-02	M.S.	Staff Scientisit, University of Minnesota
Ahmet L. Tek	97-03	Ph.D.	Postdoctor, University of Wisconsin
Robert M. Stupar	99-present	Ph.D.	
Lara M. Colton	03-present	Ph.D.	
Hye-Ran Lee	03-present	Ph.D.	

Doctoral advisor: Bikram S. Gill

Postdoctoral advisors: Bikram S. Gill
David C. Ward

BIOGRAPHICAL SKETCH - Parrott

i. Professional preparation:

University of Kentucky	Agronomy	B.S., 1981
Univ. of Wisconsin-Madison	Plant Breeding & Plant Genetics	M.A., 1983
Univ. of Wisconsin-Madison	Plant Breeding & Plant Genetics	Ph.D., 1985
Univ. of Wisconsin-Madison	Plant genetics	1985 - 1986
University of Kentucky	Biotechnology	1986 - 1987

ii. Appointments

- Professor, 1998-present, University of Georgia
- Associate. Professor., 1993-1998, University of Georgia
- Assistant Professor, 1988-1993, Univ of Georgia

iii. (A) Publications - 5 most closely related:

Thomson, J.M., and W.A. Parrott. pMECA: A size-based, blue/white selection multiple common and rare-cutter general cloning and transcription vector. Bio/Techniques. 24:922-928. 1998.

Dinkins, R., MS.S. Reddy, C.A. Meurer, B. Yan, J. Finer, F. Thibaud-Nissen, W. Parrott, and G.B. Collins. Increased sulfur amino acids in soybean plants expressing the maize 15 kDa zein protein. InVitro 37:742-747.

Thomson, J.M., P.R. LaFayette, M.A. Schmidt, and W.A. Parrott. 2002. Artificial gene-clusters engineered into plants using a vector system based on intron- and intein-encoded endonucleases. In Vitro-Plant 38:537-542.

Schmidt, M.A. and W.A. Parrott. 2001 Quantitative detection of transgenes in soybean (*Glycine max* (L.) Merrill) and peanut *Arachis hypogaea* L. by real-time PCR. Plant Cell Rep. 20:422-428.

Parrott, W.A. and T.E. Clemente. 2004. Transgenic soybean. In: J.E. Specht and H.R. Boerma (eds). Soybeans: Improvement, Production, and Uses, 3rd Ed. – Agronomy Monograph No. 16. ASA-CSA-SSSA, Madison, WI.

C. (ii) Publications - 5 other:

LaFayette, P.R. and W.A. Parrott. 2001 A non antibiotic marker for amplification of plant transformation vectors in *E. coli*. Plant Cell Rep. 20:338-342.

Magbanua, Z.V., H.D. Wilde, J.K. Roberts, K. Chowdhury, J. Abad, J.W. Moyer, H.Y. Wetzstein, and W.A. Parrott. 2000. Field resistance to tomato spotted wilt virus in transgenic peanut (*Arachis hypogaea* L.) expressing an antisense nucleocapsid gene sequence. Mol. Breed. 6:227-236.

Rosellini,D.,P. LaFayette, P. Barone F. Veronesi, and W.A. Parrott. 2004. A point mutation in the alfalfa chloroplast DNA conditions for kanamycin resistance. Plant Cell Rep. Accepted for publication

Schmidt, M.A., G.S. Martin, B.A. Artelt, and W.A. Parrott. 2004. Increased transgene expression level by breeding and selection in white clover. Submitted.

Walker, D.R., H.R. Boerma, J.N. All and W.A. Parrott. 2004. A soybean QTL that enhances the effectiveness of Bt. Submitted.

iv. Synergistic Activities

1. Dr. Parrott teaches a graduate course in plant cytogenetics, honors seminars in biotechnology, seminar technique for graduate students, and an undergraduate lecture course on tropical agroecology. He also supervises one or two undergraduate research projects in his laboratory every year. Three teaching-related matters of particular note:

- Creation and direction of a novel Applied Biotechnology major on campus. Wayne Parrott designed this major, and helped get it approved. It got started in the Fall semester of 2003, and he is now co-director for the major.
- Study abroad in tropical agroecology. This is another of Dr. Parrott's courses. It was the first study abroad course in the College of Agricultural and Environmental Sciences, and one of the first on campus. Dr. Parrott takes college students to a tropical location where they can get first hand experience in understanding agroecological concepts.
- Dr. Parrott is co-PI on an NSF teaching grant, "The Science Behind our Food" that is helping grade school, high school, and graduate students come up with hands-on exercises to help students learn science.

3. Dr. Parrott is active in the soybean community, the Society for In Vitro Biology and the Crop Science Society of America, where he has planned many symposia over the years. Of particular note:

- Dr. Parrott was elected to be chair for the C7 (Genomics, Molecular & Biotechnology) section of the Crop Science Society of America for 2002. As such, he helped plan symposia on the role of genomics in plant breeding, transgene flow, and the role of biotechnology in international development.
- He is also the chair for the upcoming world congress on In Vitro Biology to be held in San Francisco later this year.
- He helped establish and then got elected to be a charter member of the Soybean Genetics Executive Committee.
- He has been on the editorial board of *Plant Cell Reports* for 10 years, and has also been on the editorial boards of *Crop Science* and *Plant Cell, Tissue and Organ Culture*.

3. Finally, Dr. Parrott actively provides information and training on biotechnology topics. Of particular interest, he:
- Is on the expert list for the Council and Agricultural Science and Technology and for the International Food Information Council. These lists are made available to reporters looking for an expert to answer their questions.
- Is a member of the Agricultural Biotechnology Communications Group, which rapidly provides explanations and answers to issues related to GMOs as they appear in the popular press
- Does volunteer work for the USDA-FAS to help train Latin American journalists and government officials in the issues surrounding biotechnology.

v-a. Collaborators, past 4 years:

Jorge Abad, *North Carolina State University*
John All, *University of Georgia*
C. Arellano, *N.C. State University*
Matthew Bailey, *Pioneer*
Joanne Barton, *Pioneer*
C.M. Bianchi-Hall, *NC State University*
James Birchler, *University of Missouri*
Roger Boerma, *University of Georgia*
Joe Bouton, *Noble Foundation*
Thomas Carter, *North Carolina State University*
Kamal Chowhury, *International Paper*
Thomas Clemente, *University of Nebraska*
Glenn Collins, *University of Kentucky*
Miguel Dall'agnoll, *Unv. Rio Grande, Brazil*
Kelly Dawe, *University of Georgia*
Randy Dinkins, *University of Kentucky*
Julie Essig, *Kansas State University*
John Finer, *Ohio State University*
Holly Franz, *Ohio State University*
Richard Hussey, *University of Georgia*
Jiming Jiang, *University of Wisconsin*
Mark Jordan, *Agriculture Canada*
Heidi Kaeppler, *University of Wisconsin*
Gary Kochert, *University of Georgia*
Peter LaFayette, *University of Georgia*
Zenaida Magbanua, *Univ. of Mississippi*
Kay McAllister, *University of Kentucky*
Robert McPherson, *University of Georgia*
Curtis Meurer, *University of Kentucky*
Rouf Mian, *Noble Foundation*
James Moyer, *North Carolina State University*
Brian Rector, *USDA-ARS*
M.S.S. Reddy, *Noble Foundation*
Carl Redmond, *University of Kentucky*
James Roberts, *Monsanto*
Daniele Rosellini, *University of Perugia*
Thomas Rufty, *NC State University*
Randy Shoemaker, *USDA-ARS, Iowa State U.*
James Specht, *University of Nebraska*
Gary Stacey, *University of Missouri*
Françoise Thibaud-Nissen, *TIGR*
John Michael Thomson, *Univ. of North Carolina*
Harold Trick, *Kansas State University*
Anne Marie Thro, *USDA*
Donna Tucker, *University of Georgia*
Fabio Veronesi, *University of Perugia*
Lila Vodkin, *University of Illinois*
Hazel Wetzstein, *University of Georgia*
Jack Widholm, *University of Illinois*
Dayton Wilde, *Arborgen*
Bo Yan, *University of Kentucky*

b. Graduate and Post Doctoral Advisors:

- Glenn Collins, University of Kentucky
- Elizabeth Heij, CSIRO

Thesis Advisor and Post-Graduate-Scholar Sponsor, past 5 years

Piero Barone, *University of Illinois*
Patrick Kane, *USDA*
Elizabeth Little, *UGA*
Zenaida Magbanua, *University of Mississippi*
James Narvel, *Monsanto*
Monica Schmidt, *UGA*
Mary K. Sledge, *Noble Foundation*
David Walker, *UGA*
Shuquan Zhu, *UGA*

BIOGRAPHICAL SKETCH

Gernot Presting | Assistant Professor, Bioinformatics | 8/30/64

a. Education:

B.S.	1984	University of Delaware	Entomology and Plant Pathology
Ph.D.	1991	U.W. Madison	Plant Pathology

b. Appointments:

- Asst. Professor of Bioinformatics, Departments of Molecular Biosciences and Bioengineering, University of Hawaii 2004-
- Senior Scientist for data mining, Torrey Mesa Research Institute, 2000-2003.
- Research Associate Professor, Clemson University Genomics Institute, Clemson University 1998-2000.

c. Publications - 5 most closely related:

Chen M[a], Presting G[a], Barbazuk WB[a], Goicoechea JL, Blackmon B, Fang G, Kim H, Frisch D, Yu Y, Sun S, Higingbottom S, Phimphilai J, Phimphilai D, Thurmond S, Gaudette B, Li P, Liu J, Hatfield J, Main D, Farrar K, Henderson C, Barnett L, Costa R, Williams B, Walser S, Atkins M, Hall C, Budiman MA, Tomkins JP, Luo M, Bancroft I, Salse J, Regad F, Mohapatra T, Singh NK, Tyagi AK, Soderlund C, Dean RA, Wing RA. 2002. An integrated physical and genetic map of the rice genome. Plant Cell 14:1-10.

Cheng Z, Presting G, Buell CR, Wing RA, Jiang J. 2001. High-resolution pachytene chromosome mapping of bacterial artificial chromosomes anchored by genetic markers reveals the centromere location and the distribution of genetic recombination along chromosome 10 of rice.

Presting G, Malysheva L, Fuchs J and Schubert I. 1998. A Ty3/GYPSY retrotransposon-like sequence localizes to the centromeric region of cereal chromosomes. Plant Journal 16:721-728.

Presting G, Frary A, Pillen KP and Tanksley SD. 1996. Telomere-homologous sequences occur near the centromeres of many tomato chromosomes. Mol. Gen. Genet. 251:526-531.

Frary A, Presting G and Tanksley SD. 1995. Molecular mapping of the centromeres of tomato chromosomes 7 and 9. Mol. Gen. Genet. 250:295-304.

d. Publications - 5 other:

Presting G. 2003. Mapping multiple co-sequenced T-DNA integration sites within the Arabidopsis genome. BioInformatics 19(5):579-586.

Goff SA, Ricke D, Lan T, Presting G, et al. 2002. A Draft Sequence of the Rice Genome (Oryza sativa L. ssp. japonica). Science Apr 5 2002: 92-100.

Wechter WP, Begum D, Presting G, Kim JJ, Wing RA, Kluepfel DA. 2002. Physical mapping, BAC-end sequence analysis, and marker tagging of the soilborne nematicidal bacterium, Pseudomonas synxantha BG33R. OMICS – A Journal of Integrative Biology 6:11-21.

Hudakova S, Michalek W, Presting G, ten Hoopen R, dos Santos K, Jasencakova Z, Schubert I. 2001. Sequence organization of barley centromeres. Nucleic Acids Research 29:5029-5035.

Presting G, Budiman MA, Wood T, Yu Y, Kim HR, Goicoechea JL, Fang E, Blackman B, Jiang J, Woo SS, Dean RA, Frisch D, Wing RA. A framework for sequencing the rice genome. Novartis Found Symp. 2001;236:13-24; discussion 24-7.

e. Synergistic Activities

- Gernot Presting joined the University of Hawaii (UH) faculty in January of 2004 and teaches a course on bioinformatics during the Spring 2003 semester. He is also actively recruiting minority students into his laboratory at UH.

f. Collaborators & Other Affiliations:

(i) Collaborators

- Ray Ming
- Paul Moore
- Dulal Borthakur
- Jeff Tomkins

(ii) Graduate and Postdoctoral Advisors

- John Helgeson, doctoral, U.W. Madison
- Chuck Brown, Karen Kindle and Steve Tanksley, Ingo Schubert, postdoctoral, Prosser, Cornell, IPK Gatersleben

(iii) Thesis advisor and Postgraduate Scholar advisor:

-

FACILITIES, EQUIPMENT & OTHER RESOURCES

FACILITIES: Identify the facilities to be used at each performance site listed and, as appropriate, indicate their capacities, pertinent capabilities, relative proximity, and extent of availability to the project. Use "Other" to describe the facilities at any other performance sites listed and at sites for field studies. USE additional pages as necessary.

Laboratory: **The Dawe lab is composed of a main laboratory of 1000 sq. ft. and a seed storage room of 300 sq. ft. The laboratory includes two separated microscope rooms and an isolated area for use with radioactive chemicals. The laboratory is equipped with micro, low speed and superspeed**

Clinical:

Animal:

Computer: **As a member of the University of Hawaii faculty, Dr. Presting has free access to the Maui High Performance Computing Center (http://www.mhpcc.edu/). Central computer facilities are also available at Georgia and Missouri.**

Office: **Birchler has office space for the PI as well as associates is immediately adjacent to the laboratory. Dawe has cubicles for graduate students and postdocs in a separate room, and the PI has a 270 sq ft. office across from the laboratory. Parrott has 8 cubicles for technicians and graduate**

Other: **Core facilities available to all faculty at Georgia include the Molecular Cytology Facility (transmission EM and cryo and paraffin sectioning), the Molecular Genetics Instrumentation Facility (oligonucleotide primer synthesis, DNA sequencing), a central computing resource, and the Center for Applied Genetic Technologies core facility, which is down the hall**

MAJOR EQUIPMENT: List the most important items available for this project and, as appropriate identifying the location and pertinent capabilities of each.

The Jiang laboratory is equipped with necessary facilities for molecular cytogenetic research, including two Olympus fluorescent microscope workstations installed with digital CCD cameras and computers. One workstation is set up for 3D deconvolution microscopic analysis. The microscopes are equipped with high-quality 60X and 100X ApoPlan objectives specifically designed for fluorescence analysis. Equipment related to FISH analysis also includes four G4 Power Macintosh computers.

OTHER RESOURCES: Provide any information describing the other resources available for the project. Identify support services such as consultant, secretarial, machine shop, and electronics shop, and the extent to which they will be available for the project. Include an explanation of any consortium/contractual arrangements with other organizations.

Parrott and Dawe have full access to their respective departmental secretarial pools. All the associated universities have outstanding support infrastructures that include electronics and machine maintenance and repair.

LABORATORY FACILITIES (continued):

centrifuges, a variety of different gel rigs for DNA, RNA and protein electrophoresis, power supplies, a thermal cycler, a cold room, refrigerators, freezers and ultracold (-80) freezer, water and air incubators, a hybridization oven, a UV crosslinker, an incubator/shaker for bacterial cultures, spectrophotometers, a cryostat, three Zeiss dissecting microscopes.

Parrott has a BL-2 laboratory (2400 ft2) that is fully equipped for plant tissue culture and molecular genetics. Equipment includes 3 growth rooms for maintenance of plant cultures (8' x 8'); 3 growth chambers for cell cultures and acclimatization of transgenic plantlets, autoclaves; 6 six-ft laminar flow hoods; 2 Class IIA containment cabinets; Biolistic HE gene gun and a particle inflow gene gun; 2 incubators; 2 water baths; vacuum oven; 1 tabletop and 1 refrigerated; Type I water purifier; 2 fume hoods; 4 dissecting microscopes (3 for use inside a laminar flow hood, the other has photographic capabilities); a microscope (Olympus BH2) equipped with brightfield, darkfield, and Nomarski optics is also available in the laboratory, as is an Olympus inverted microscope. These microscopes are equipped with full microphotographic capabilities, including a video camera, color thermal printer, and an image grabber for computer use; microscope (Olympus BH2) for epifluorescence; hot plate stirrers; pH meters; analytical balances; 2 dishwashers; multiple refrigerators, 2 ultralow freezer; DNA transilluminator and camera; Stratagene Eagle Eye II Still Video System with image analysis software; fluorometer; 3 thermocyclers for PCR; 3 electrophoresis power supplies and equipment; automatic X-ray developer; 3 Eppendorf microfuges; 3 shaking incubators; 5 double-tiered and 2 single-tiered New Brunswick orbital shakers.

The Birchler lab at the University of Missouri is 1600 sq. ft. with bench space for 16 people. Included in the lab are a small darkroom, microscope room and two chemical fume hoods. Immediately adjacent is an equipment room, which houses a high speed centrifuge, Beckman Optima ultracentrifuge, appropriate rotors, Beckman liquid scintillation counter, microtome, speed-vac concentrator, gel drier, vacuum drier, vacuum oven, -80 freezer (2), bacterial incubator shaker, chromatography cabinet and other incubators. the lab is equipped with many small items such as electrophoresis power supplies and rigs (for agarose and acrylamide), water baths, heat blocks, balances, pH meter, homogenizers, hybridization ovens, need puller, sonicator, laser densitometer, pulsed field electrophoresis system (2), isoelectric focusing rig, spectrophotometer, Zeiss Universal compound microscope plus fluorescence and CCM camera, etc.

The Jiang lab at the University of Wisconsin-Madison has about 1,000 sq. ft. of laboratory space, including a 250 sq. ft. fluorescence microscopy lab. The lab space is enough to host ten associates. The Jiang lab is equipped with facilities for routine DNA work, including Perkin Elmer 9600 and 2400 PCR machines, a speed vacuum, a Bio-Rad CHEF-DR II apparatus,

(A-1) Sharing Of Results And Management Of Intellectual Property

Sharing of Results

We will produce sequence data, mapping data, and antibodies under this award. All low-pass sequence from Aim1A will be released to GenBank as soon as proper verification and annotation is complete. FPC and associated mapping data will be released through the Maize Mapping Project (http://www.maizemap.org/) as quickly as it is acquired and verified (see attached letter of support from Cari Soderlund). New antisera will be announced on our web site and released upon request. Complete sequence information from centromere 4 or a suitable replacement (Aim1B), and kinetochore mapping data (Aim1C) will be released within six months of completion or as soon as the work is accepted for publication.

Management of Intellectual Property:

Any use of inventions for academic research purposes will be unrestricted. Participating universities may file patent applications for work completed through this project according to their normal procedures. However, all investigators agree to disclose their intent to file a patent to all other investigators in the program before submitting the application. In deciding inventorship, ownership, and rights to any invention, fair consideration will be given to any and all relevant contributions by other participants in the program. The University of Georgia Technology Commercialization Office will coordinate with other institutions prior to the start of funding to ensure full agreement to these terms by all parties.

(A-2) Management Plan

Within-group communication

All PIs and senior personnel will communicate on a monthly basis through conference calls. During these calls we will discuss results, problems and solutions, and ways to maximize coordination among laboratories. A within-group listserv will be established and notes from the conference calls will be forwarded to all project participants (including those not involved in the conference call).

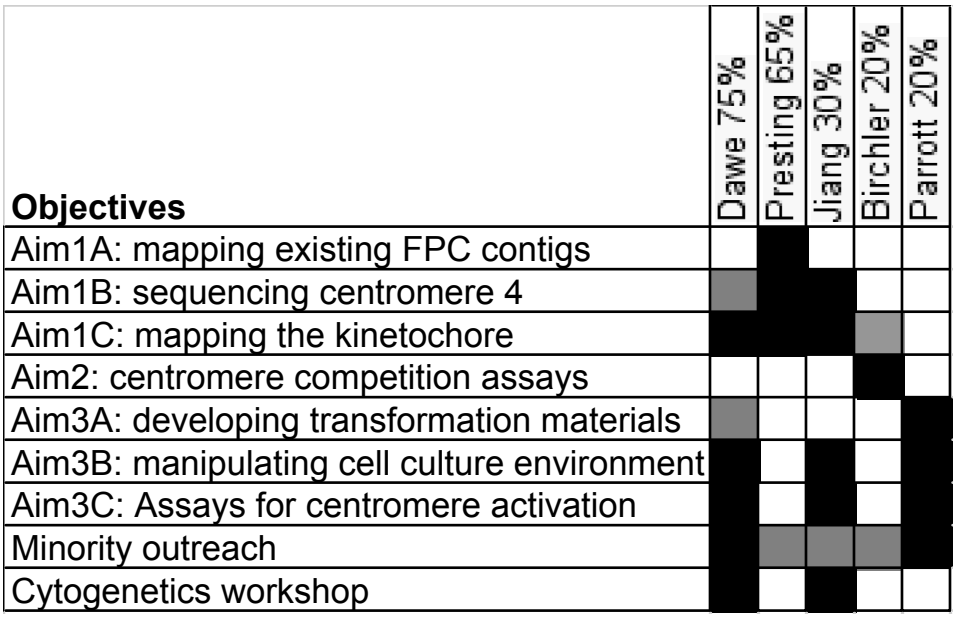
In addition, all PIs and senior personnel will attend annual meetings. Meeting locations will rotate from university to university, so that all non-senior personnel can be involved in at least one annual meeting. The meeting rotation will be: year 1-UGA, year 2-MO, year 3-UW, year 4-UH, and year 5-UGA. Meetings will occur in May or June, following the submission of annual progress reports to NSF. During annual meetings we will reinforce and expand on progress from the previous year and establish clear goals for the coming year.

Communication with the public

To communicate our results as quickly as possible to the wider community, we will establish and maintain a Plant Centromeres web site administered from the University of Hawaii. A graduate student with a strong background in computer science and an interest in bioinformatics will be hired for this work. Links to all deliverables will be available here – lists of BAC end sequences with centromere repeat homology, identification of BAC contigs containing centromere DNA, map data for centromeric contigs, antibodies, and progress on a minimum sequencing tile of centromere 4. In addition, the web site will include the project description, a list of the PIs with their respective areas of expertise and subprojects, links to related sites, a list of publications and centromere-related sequence information. Gernot Presting will be the bioinformatics contact person.

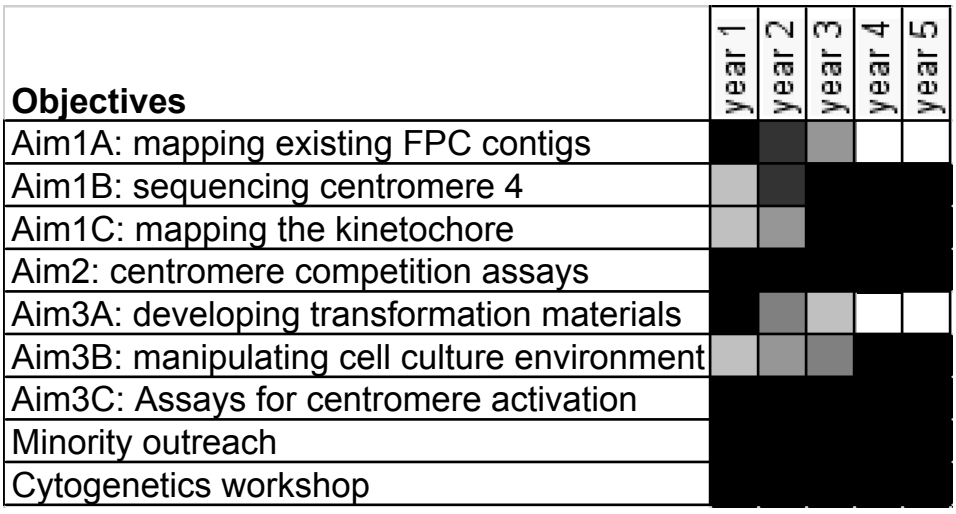
Role of Project Participants

The following table indicates the percent commitment of each PI to this project (e.g. Dawe will contribute 75% of his research appointment to this project), as well as the extent to which each PI will contribute to the major aims of the proposal. Major contributions are indicated in black, and minor contributions are indicated in shades of gray.



Project Timeline

The following table illustrates in broad terms the year each aim will be carried out. Periods of major effort are indicated in black, while periods of minor effort are indicated with shades of gray. A detailed listing of yearly goals and benchmarks is shown below. Minority outreach and the cytogenetics workshop are will be carried out on an annual basis.



Yearly goals and benchmarks

Year 1: Aim1A: Identify BACs with end sequence homology to centromere repeats, hybridize BAC filters with centromere repeats, identify FPC contigs with multiple centromeric BACs.
Aim1B: Order appropriate oat-maize addition line and multiply seed.
Aim1C: Characterize anti-MIS12 antisera, develop stretched chromatin assay for metaphase chromosomes, identify CentC polymorphisms, develop QEXT
Aim2: Score control crosses, develop markers for identifying each centromere derivative
Aim3A: Prepare centromere induction constructs
Aim3B: Characterize rice trans-centromeres with respect to their interaction with CENH3
Aim3C: Test TSA on transgenic rice seedlings

Year 2: Aim1A: Order BAC clones and single-pass sequence all centromeric contigs, identify low/single copy probes, map probes to maize mapping population.
Aim1B: Generate BAC library from oat-maize addition line.
Aim1C: Identify single copy markers, work out conditions for single-copy ChIP, quantify CentC SNPs following ChIP with anti-CENH3, pCENH3, CENPC, and MIS12 Ab.
Aim2: Begin scoring segregation data. Score chromosome loss. Characterize molecular structure of B centromeres.
Aim3A: Finish constructs.
Aim3B: Begin transformation with centromere induction constructs, start induction and TSA experiments, prepare CENH3 expression construct for maize studies.
Aim3C: Analyze first-round experiments by anti-CENH3 mediated ChIP.

Year 3: Aim1A: Identify maize centromere with best coverage in BAC library, order appropriate oat-maize addition line and multiply seed.
Aim1B: Identify BACs containing centromere repeats by hybridization, high information fingerprint analysis of BACs to generate centromeric contig.
Aim1C: Initiate kinetochore mapping by ChIP and extended fiber assays, using all four antisera
Aim2: Score segregation data, score chromosome loss. Characterize molecular structure of B derivatives.
Aim3A: -
Aim3B: Centromere induction, TSA, and CENH3 overexpression studies in maize and rice.
Aim3C: Analyze centromere activation by ChIP and extended fiber assays.

Year 4: Aim1A: -
Aim1B: Continue high information fingerprint analysis of BACs to generate centromeric contig. Develop methods for confirmation of contig assembly by low-copy FISH.
Aim1C: Continue kinetochore mapping by ChIP and extended fiber assays, using all four antisera
Aim2: Score segregation data. Score chromosome loss and assay meiotic drive. Correlate molecular features with meiotic behavior.
Aim3A: -
Aim3B: Centromere induction, TSA, and CENH3 overexpression studies in maize cells and rice plants. Start transforming maize plants with centromere induction constructs.
Aim3C: Analyze centromere activation by ChIP & fiber assays, extend experiments anti-CENPC and anti-MIS12 antisera.

Year 5: Aim1A: -
Aim1B: Sequence centromeric contig at 10x coverage, confirm contig using FISH.
Aim1C: Fully map centromere 4 by ChIP and extended fiber assays, using all four antisera
Aim2: Finish segregation data, chromosome loss, and assays for meiotic drive. Correlate molecular features with meiotic behavior.
Aim3A: -
Aim3B: Induce centromeres in regenerated plants.
Aim3C: Analyze centromere activation in regenerated maize plants by ChIP & fiber assays, using anti-CENH3, anti-pCENH3, CENPC, MIS12, MAD2, 3F3/2 Ab.