

CHARACTERIZATION OF TWO MEIOTIC DRIVE KINESINS AND LOCI  
CONTROLLING PERENNIALITY IN *ZEA*

by

KYLE WILLIAM SWENTOWSKY

(Under the Direction of R. Kelly Dawe)

ABSTRACT

The genomes of maize and its wild relatives in the genus *Zea* are regarded widely for their genotypic and phenotypic diversity. The two major overarching biological themes in this dissertation relate to phenomena that evolved within teosinte genomes long before the domestication of maize. Maize exhibits a preferential segregation system that is genetically transmitted through a variant of the chromosome 10 haplotype referred to as Abnormal chromosome 10 (Ab10). Ab10 elicits meiotic drive by manipulating the process of female meiosis. This phenomenon occurs through a mechanism whereby genomic regions composed of heterochromatic tandem repeats known as knobs become motile during meiosis. Knobs can be composed of two types of tandem repeats. The active motile forms of knobs are called neocentromeres and genetic control over the two types of neocentromeres is controlled by distinct loci which we have characterized in this work. Both loci encode kinesin-14 motor proteins that interact with knob repeats and pull them towards the minus-end poles during anaphase I and II of meiosis. The major class of knob repeats, knob180, is controlled by the kinesin-14 motor, KINDR. KINDR undergoes an indirect interaction with knob180 knobs and this interaction may be facilitated by a newly-identified kinesin-10-like protein that is encoded

near *Kindr* on the Ab10 distal tip. Motility for the second class of knob repeats, TR-1, is controlled by another kinesin-14 protein called TRKIN. KINDR and TRKIN are not directly related and TRKIN is highly divergent from its closest maize homolog. Future work on the Ab10 kinesin-14s may help elucidate this mechanism of meiotic drive and how the two neocentromere systems interplay during this process. The other focus of this dissertation pertains to the perennial growth habit seen in the teosinte species *Zea diploperennis*. Annual maize (*Z. mays* ssp. *mays*) lines were crossed with *Z. diploperennis* and perennial regrowth was assessed in the F2 population. Using QTL-seq and PCR genotyping of individual plants, regrowth was significantly associated with two loci located on chromosomes 2 and 8. Back-crossed and RIL populations may be used in the future to more precisely map the locations of these loci.

INDEX WORDS: meiotic drive, abnormal chromosome 10, kinesin motor protein, cell division, perenniality, QTL mapping, maize

CHARACTERIZATION OF TWO MEIOTIC DRIVE KINESINS AND LOCI  
CONTROLLING PERENNIALITY IN *ZEA*

by

KYLE WILLIAM SWENTOWSKY

BS, University of California, Santa Barbara, 2014

MA, University of California, Santa Barbara, 2016

A Dissertation Submitted to the Graduate Faculty of The University of Georgia in Partial  
Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2021

© 2021

Kyle William Swentowsky

All Rights Reserved

CHARACTERIZATION OF TWO MEIOTIC DRIVE KINESINS AND LOCI  
CONTROLLING PERENNIALITY IN *ZEA*

by

KYLE WILLIAM SWENTOWSKY

Major Professor:	R. Kelly Dawe
Committee:	Katrien Devos
	Wolfgang Lukowitz
	Wayne Parrott
	Esther van der Knaap

Electronic Version Approved:

Ron Walcott  
Dean of the Graduate School  
The University of Georgia  
May 2021

## ACKNOWLEDGMENTS

I would like to extend my deepest thanks to my advisor, Kelly Dawe, whose outstanding mentoring and support was essential for my graduate education and growth as a scientist. I would also like to thank my committee members Katrien Devos, Wolfgang Lukowitz, Wayne Parrott, and Esther van der Knaap for providing their expertise to enhance the quality of my research.

I would like to acknowledge the help of several members of the Dawe Lab for the important roles they played in my doctorate training and research. In particular, the broad expertise provided by Jonathan Gent has been instrumental in my work. Jianing Liu has also been a tremendous aid in helping me to navigate bioinformatics. I would also like to thank Harrison Bell whose help in the lab and field allowed me to accomplish much during my doctorate training.

This work was supported by a Palfrey Grant and a Research, Teaching, and Outreach Award (RTOA) from the Plant Biology department at the University of Georgia, an Interdisciplinary and Innovative Research Grant (IIRG) from the Graduate School at the University of Georgia, and the National Science Foundation.

## TABLE OF CONTENTS

	Page
ACKNOWLEDGEMENTS .....	iv
LIST OF TABLES .....	vii
LIST OF FIGURES .....	viii
CHAPTER	
1 INTRODUCTION AND LITERATURE REVIEW .....	1
KNOBS, THE AB10 HAPLOTYPE, AND MEIOTIC DRIVE ...	1
DESCRIPTIONS OF AB10 AND K10L2 HAPLOTYPES .....	3
AB10 NEOCENTROMERE MOTORS .....	5
PERENNIALITY IN GRASSES .....	7
REFERENCES .....	13
2 MOLECULAR CHARACTERIZATION OF THE KNOB180 NEOCENTROMERE MOTOR.....	17
INTRODUCTION .....	17
RESULTS .....	21
DISCUSSION .....	32
MATERIALS AND METHODS .....	36
REFERENCES .....	47
3 DISTINCT KINESIN MOTORS DRIVE TWO TYPES OF MAIZE NEOCENTROMERES .....	71

ABSTRACT .....	72
INTRODUCTION .....	73
RESULTS .....	76
DISCUSSION .....	84
MATERIALS AND METHODS .....	88
ACKNOWLEDGEMENTS .....	97
REFERENCES .....	97
4 QTL MAP OF EARLY- AND LATE-STAGE PERENNIAL REGROWTH IN <i>ZEA</i>	
<i>DIPLOPERENNIS</i> .....	121
ABSTRACT .....	122
INTRODUCTION .....	123
MATERIALS AND METHODS .....	125
RESULTS .....	129
DISCUSSION .....	133
ACKNOWLEDGEMENTS .....	138
REFERENCES .....	138
5 CONCLUSIONS AND DIRECTIONS FOR FUTURE STUDY ....	155
REFERENCES .....	158

## LIST OF TABLES

	Page
Table 2.1: Annotation-based differentially expressed <i>smd13</i> genes.....	51
Table 2.2: Proteins identified from KINDR co-IP/Mass Spectrometry.....	52
Table 3.1: SRA project and run codes for Illumina reads used in this study.....	101
Table 4.1: Number of regrown (RG) and non-regrown (NRG) plants in P39/'Gigi' F2 greenhouse-grown plants planted in 2018 and scored after one and two cycles of regrowth in 2019.....	142
Table 4.2: Genotype distribution of markers in greenhouse-regrown P39/'Gigi' F2 plants.....	143
Table 4.3: Numbers of regrown (RG) and non-regrown (NRG) plants in F2 populations scored during different time points following flowering and senescence in 2019.....	144
Table 4.4: Genotype distribution of markers in field-grown P39/'Gigi' F2 plants either showing regrowth (RG) or no regrowth (NRG).....	145
Table 4.5: List of Illumina sequenced samples for this study.....	146
Table 4.6: Number of filtered, homozygous SNPs per chromosome present in <i>Z. diploperennis</i> 'Gigi' from Illumina reads aligned to maize P39 reference.....	147

## LIST OF FIGURES

	Page
Figure 2.1: KINDR is expressed in meiotic tissues and localizes to knob180 knobs.....	58
Figure 2.2: KINDR does not interact with knob180 DNA and characterization of <i>smd13</i> .....	59
Figure 2.3: DNA-seq coverage along Ab10 in Ab10 and <i>smd</i> genotypes and repeat analysis of Ab10 distal tip .....	61
Figure 2.4: Ab10 SNPs in <i>smd13</i> .....	63
Figure 2.5: Differentially-expressed Ab10 annotated genes .....	64
Figure 2.6: RNA-seq analysis of Ab10 using Trinity transcripts .....	65
Figure 2.7: <i>Kindr</i> and <i>Kinesin-10-like</i> gene copy locations .....	67
Figure 2.8: Identification of <i>Kinesin-10-like</i> .....	69
Figure 2.9: Protein gel analysis of Ab10 anti-KINDR IP experiment .....	70
Figure 3.1: Ab10 types and the Rhoades model for meiotic drive .....	102
Figure 3.2: <i>Trkin</i> and its divergence from other kinesin-14s.....	104
Figure 3.3: Conservation and functionality of the TRKIN motor.....	106
Figure 3.4: TRKIN localization and presence/absence in different Ab10 types.....	108
Figure 3.5: Immunolocalization of TRKIN and KINDR in male meiosis.....	110

Figure 3.6: TR-1 sequence diversity and distribution.....	112
Figure 3.7: Possible TR-1 neocentromere functions.....	114
Figure 3.8: Sizes and positions of exons and transposable elements in <i>Vks1</i> , <i>Dv1</i> , <i>KindrE9</i> , pseudo- <i>Trkin1</i> , and <i>Trkin</i> .....	115
Figure 3.9: Kinesin-14 protein alignments.....	116
Figure 3.10: Histogram showing microtubule-gliding velocity distribution of GFP- TRKIN.....	117
Figure 3.11: Expression of <i>Trkin</i> as assayed by quantitative RT-PCR using cDNA from Ab10 reproductive and vegetative tissues.....	118
Figure 3.12: Clustering dendrogram of anther mRNA-seq data that align to the Ab10 haplotype from 135Mb to the end of the chromosome.....	119
Figure 3.13: Structure of <i>Trkin</i> and the pseudo- <i>Trkin</i> loci.....	120
Figure 4.1: Growth conditions for scoring regrowth in F2 populations...	148
Figure 4.2. QTL-seq for regrowth reveals two major loci in a P39/'Gigi' F2 population.....	149
Figure 4.3: Association between regrowth and markers on chromosomes 2, 7, and 8 in the 2019 P39/'Gigi' F2 population.....	150
Figure 4.4: Association between regrowth and markers on chromosomes 2, 7, and 8 in the 2019 Hp301/'Gigi' F2 population.....	151
Figure 4.5. QTL-seq for tiller number supports numerous tillering QTL in the 2019 P39/'Gigi' F2 field population.....	152
Figure 4.6. Distribution of filtered homozygous 'Gigi' SNPs mapped to P39v1 reference assembly.....	153

Figure 4.7. Distribution of tiller number phenotype in field-grown

P39/'Gigi' F2 plants..... 154

## CHAPTER 1

### INTRODUCTION AND LITERATURE REVIEW

Maize is one of humanity's most shining accomplishments as it represents man's extraordinary ability to mold nature to our will and unravel biological mysteries long hidden to us. Much of the agricultural and scientific progress we have made in this major crop and model organism can be attributed to the vast diversity present in the genomes of different maize lines. This diversity is also present in teosinte, the other members of the *Zea* genus that have existed since long before maize was domesticated. My dissertation work tells two stories of fascinating biological phenomena that arose within the genomes of teosinte.

#### **KNOBS, THE AB10 HAPLOTYPE, AND MEIOTIC DRIVE**

The first of these stories begins near the beginning of the study of genetics itself. Barbara McClintock was among the first to recognize the potential of using maize as a cytological model after describing heterochromatic features on maize chromosomes she called knobs (McClintock 1929). The early part of her career would be spent using knobs to demonstrate cytological phenomena that are fundamental to eukaryotic genetics. One such example is the relationship between cytological crossing over and genetic recombination (Creighton and McClintock 1931).

In addition to their usefulness as a biological tool, it would later become evident that knobs evolved to serve a biological function.

A colleague of McClintock's named Marcus Rhoades became interested in a particularly large knob on the long arm of maize's shortest chromosome in a haplotype known as Abnormal Chromosome 10 or Ab10. This haplotype is strikingly different than the more common maize chromosome 10 haplotype referred to as Normal Chromosome 10 or N10. Rhoades undertook an effort to genetically map the distance between the Ab10 knob and a known marker on this chromosome knob known as *Colored1 (RI)* that dominantly confers purple kernel color. Ab10 was linked to the recessive allele (*ri*) which produces yellow kernels and found that when the heterozygous *ri-Ab10 / RI-N10* was used as a female in a cross with *ri-N10* homozygous pollen for the purposes of mapping, the resulting ear had about 70% yellow kernels (Rhoades 1942). After Rhoades observed a crossover that placed Ab10 on a chromosome with *RI* (i.e. *RI-Ab10*) he repeated this test cross and observed ears containing around 70% purple kernels. Because the number of progeny containing Ab10 and its linked kernel color allele deviated significantly from 50%, Rhoades concluded that Ab10 is able to violate Mendel's first law of segregation and would spend the greater portion of his life discovering how this system operates.

After observing meiotic anaphase of Ab10 plants, Rhoades found that knobs, which are present on the chromosome arms, moved towards the spindle poles ahead of the centromeres that normally arrive at the poles before any part of the chromosome (Rhoades and Vilkomerson 1942). He called these knobs that had been activated in meiotic anaphase "neocentromeres" and since he believed neocentromeres during meiosis were the cause of the difference from observed Mendelian expectations, this process became known as "meiotic drive".

Rhoades devised a model to explain how neocentromeres lead to meiotic drive that is now referred to as the “Rhoades Model” ((Rhoades 1952); Figure 3.1). First, during prophase I, recombination occurs between one chromatid with a knob and another without a knob. This forms a pair of “heteromorphic dyads” that each contain one knobbed and one non-knobbed chromatid. Second, neocentromeres can be seen at anaphase I and II. At anaphase II, the sister chromatids separate and the chromatid containing the knob preferentially moves towards the bottom of the tetrad due to spindle polarity. Finally, as part of normal female meiosis, the top three cells in the tetrad degenerate, leaving only the bottom cell behind. Since the knobbed chromatid more frequently enters this cell, the knob and all alleles linked to it are preferentially transmitted.

## **DESCRIPTIONS OF AB10 AND K10L2 HAPLOTYPES**

Like many features of the maize genome, knobs show remarkable diversity. Knobs appear cytologically as large, compact heterochromatin but we now know much about knobs at the sequence level. Knobs are composed of two satellite tandem repeats: the knob180 repeat (Peacock et al. 1981) and the TR-1 repeat (Ananiev et al. 1998). Individual knobs appear to be composed of one type or both types together and, to little surprise, maize karyotypes display tremendous diversity for knob position and composition (Albert et al. 2010). The 180bp repeat appears to be much more prevalent in maize and teosinte genomes and is essential for Ab10 meiotic drive (Dawe et al. 2018). Ab10 itself contains a large distal 180bp knob and three smaller TR1 knobs at the proximal end of the haplotype (Liu et al. 2020) and both repeats can form active neocentromeres (Hiatt et al. 2002). Ab10 also contains euchromatin in two regions. One euchromatic domain known as the shared region lies between the TR-1 and knob180 knobs

and is shared but structurally varied with N10. Another euchromatic region is Ab10-specific and is located beyond the knob180 knob called the distal tip (Liu et al. 2020).

An additional chromosome 10 haplotype called K10L2 was discovered that contains a large TR-1-only knob (Kanizay et al. 2013). Unlike Ab10, K10L2 does not appear to have large euchromatic regions outside of those normally observed on this chromosome and the TR-1 knob exists on a haplotype that resembles N10. Plants containing K10L2 have TR-1 but not knob180 neocentromeres. Very moderate but statistically significant levels of drive (~51%) were associated with K10L2. Interestingly, when heterozygous Ab10/K10L2 plants are assayed for meiotic drive, the percentage of progeny containing Ab10 is significantly reduced compared to Ab10/N10 plants. In conjunction with the Rhoades model that connects neocentromeres to meiotic drive, it is thought that Ab10 and K10L2 neocentromeres compete with each other for access to the basal megaspore cell. This scenario highlights the intragenomic conflict between the two types of knob repeats. This raises an interesting possibility that the TR-1 neocentromere system evolved to suppress meiotic drive and the harmful long-term evolutionary consequences associated with it (Rhoades 1942; Higgins et al. 2018).

Meiotic drive observed with Ab10 does not only happen on this chromosome. When Ab10 was present, meiotic drive was observed for knob-linked markers on chromosomes 3 (Rhoades and Dempsey 1966), 9 (Kikudome 1959; Rhoades and Dempsey 1966), 6 (Kanizay et al. 2013; Dawe et al. 2018), and 4 (Dawe et al. 2018) so it was concluded that Ab10 contains one or more *trans*-acting factors that enable meiotic drive. Ab10 variants have been recovered that are viable but do not confer meiotic drive, such as terminal deficiencies (Emmerling 1959; Rhoades and Dempsey 1986; Hiatt et al. 2002; Hiatt and Dawe 2003) and mutants without

structural changes (Dawe and Cande 1996; Dawe et al. 2018). These Ab10 variants have been used to assign four cytological functions associated with meiotic drive to different regions of the haplotype. It was shown that neocentromere activity of knob180 and TR-1 knobs is controlled independently and that these functions are located on the distal tip and in the region between the TR-1 knobs, respectively (Hiatt et al. 2002; Hiatt and Dawe 2003). Another locus encoded in the shared region euchromatin enhances the recombination rate between centromeres and knobs. This increases the strength of preferential transmission by improving the rate of heteromorphic dyad formation (Hiatt and Dawe 2003). The final known meiotic drive-related activity is characterized by heritable loss of meiotic drive with unaffected neocentromeres. This activity was mapped to the distal tip and is thus known as “distal tip function” (Hiatt and Dawe 2003). A thorough understanding of the maize meiotic drive mechanism will require the discovery and characterization of the genes required for the various Ab10 cytological activities and the recapitulation of this system using the discovered components. Such a goal could affect the synthetic biology field, for example, by creating an artificial gene drive system to achieve biotechnological goals.

## **AB10 NEOCENTROMERE MOTORS**

(Yu et al. 1997) discovered that neocentromeres interact with the meiotic spindle laterally, and thus hypothesized that Ab10 neocentromere movement was mediated by a minus end-directed motor. Kinesin-14s make up the major class of microtubule minus end-directed motors (Gicking et al. 2018) so the search for neocentromere motors has centered around this class of proteins. The knob180 neocentromere motor was recently discovered to be a kinesin-14 protein called KINDR that is encoded on the Ab10 distal tip (Dawe et al. 2018). *Kindr* is

composed of nine tandemly-repeated gene copies and the loss of *Kindr* expression in several *suppressor of meiotic drive (smd)* mutants explains the lack of knob180 neocentromeres and meiotic drive.

**Chapter 2** of this dissertation describes the further characterization of KINDR as well as steps taken to identify the hypothetical “distal tip function” gene. KINDR is a functional, minus end-directed kinesin-14 that is expressed in Ab10 meiotic tissues and co-localizes specifically with knob180 knobs. An Ab10 mutant called *smd13* with an unknown lesion shows unaffected KINDR expression but the protein does not localize to knob180 knobs. We hypothesize that this mutant lacks a gene required for proper KINDR localization and fulfills the distal tip function. *smd13* lacks structural variants but is not complemented by *smd8*, a 2-Mb deletion within the distal tip in a region directly adjacent to the *Kindr* complex. This suggests that the *smd13* mutation lies within this *smd8*-deleted region. This area is composed of six tandem repeats each containing a copy of an undescribed *Kinesin-10-like* gene, but further study will be required to determine if this novel gene fulfills the distal tip function.

In **Chapter 3**, the identification of a TR-1 neocentromere motor named TRKIN is described. This protein is encoded at a locus between the first two TR-1 knobs on Ab10. TRKIN is a functional minus end-directed kinesin-14 that localizes specifically to TR-1 knobs. The sequence of TRKIN is highly divergent from its closest kinesin-14 homologs including KINDR, suggesting TRKIN is either extremely ancient or evolved under substantial selective pressures. Natural Ab10 variants display diversity for full-length *Trkin* expression and this correlates with TR-1 neocentromere activity. We hypothesize that in addition to their known roles in Ab10 drive

suppression, TR-1 neocentromeres may enhance meiotic drive or facilitate chromosome integrity when neocentromeres are present.

## **PERENNIALITY IN GRASSES**

For millennia, maize has been cultivated as an annual crop throughout a regular, yearly growth cycle. Maize kernels are sown, the plants grow and reach maturity, the female inflorescences produce new kernels, and finally the plants senesce and die. This annual life cycle strategy has been highly successful for all cereal crops, but we now face novel agricultural challenges unique to the 21st century. It has been hypothesized that we can produce cereals more sustainably and exploit new growing environments by converting our annual crop species into perennials.

Perennial plants undergo a similar life cycle to that of annuals except they can also regrow following senescence. Most perennials are capable of living for multiple years and many use the perennial life cycle to survive seasonal stresses such as drought or freezing. During these expanded life cycles, perennial plants develop vast rooting systems that are typically deeper than those of their annual counterparts. These more developed roots allow perennial species to access more nutrients and water, which may both be limiting when soil conditions become more unpredictable during the 21st century. Additionally, since perennials store nutrients below-ground and reuse them in the next cycle of growth, they may require less fertilizer compared to annual crops. Because perennial agriculture may harbor significant benefits for sustainable agriculture, converting our cereal crops into perennial species has begun to attract attention in recent decades (Zhang et al. 2011; Food and Agriculture Organization of the United Nations 2018; Murray and Jessup 2014; L. Fernando et al. 2018; Crews and Cattani 2018).

Virtually all cereal crops cultivated today have a closely-related perennial relative and in some cases, breeding can take place between these species. This observation has led to two major avenues for the production of perennial crops: selecting for beneficial traits in a wild perennial species or breeding alleles from the perennial relative into the cultivated species. Utilizing the former approach, an alternative to bread wheat, trademarked as Kernza® was recently bred from the intermediate wheatgrass species *Thinopyrum intermedium*. Kernza® has shown remarkable potential as a perennial crop and researchers have observed some of the hypothesized benefits of perennials in this crop. Notably, Kernza® has high water use efficiency, likely due to its substantial rooting depth (de Oliveira et al. 2018; Oliveira et al. 2020). Still, *de novo* domestication involves slow progress since improvement must occur using standing variation within the wild species.

An alternative approach to *de novo* domestication of perennial crops is to introgress alleles from perennial species into annual crops. Progress has been made using this approach in three of our major cereal crops: rice, sorghum, and maize. By crossing the annual and perennial species, mapping populations have been generated that segregate perennial-related traits that can be mapped and subsequently bred into annuals. Rhizomes are winter-hardy below-ground storage organs derived from stems and have been the most well-studied perennial trait. Since they frequently occur in perennial but not annual species, it is generally thought that rhizomes are crucial for a successful perennial system. When spring regrowth following a winter killing-frost was mapped in Johnsongrass (*Sorghum halepense*), spring regrowth was associated with rhizome QTL (Paterson et al. 1995). This highlights the value of mapping rhizomes as a trait: a functional perennial crop in a temperate environment must be capable of withstanding winter colds and rhizomes simply give the best opportunity for plants to do so. Mapping studies in all

three species have shown that rhizome development is controlled by many different loci and rhizome QTL in rice and sorghum map to conserved locations (Paterson et al. 1995; Hu et al. 2003; Westerbergh and Doebley 2004; Kong et al. 2015).

In the first published study to map perennial traits in the close maize relative *Zea diploperennis*, (Westerbergh and Doebley 2004) found 38 QTL for various perennial traits, namely, rhizomes, elongated underground stems, number of tillers, slender tillers, compactness of tillers, thin roots, number of side branches, and withered stems. Slender tillers and thin roots are observed in *Z. diploperennis* but not in the annual parent used in this mapping study, *Z. mays* ssp. *parviglumis* so it was thought these traits may be associated with perenniality. These QTL each explained a relatively small proportion of the variation but it was noted that most of these QTL mapped to overlapping positions on chromosomes 1, 2, and 6. In particular, only two QTL were found for rhizome development that each explained less than 10% of the variation observed. The authors hypothesized that either low heritability or high epistatic effects of their loci could explain why they did not explain a large portion of the variation.

There is utility in learning about other factors that contribute to the perennial life cycle so that we may fully understand how annual and perennial life cycles are genetically controlled. In contrast to the attention given to loci that control rhizome development, little effort has been directed towards mapping QTL that directly affect regrowth. Rather than focusing on a morphological trait, regrowth in the context of perenniality reflects the plant's ability to produce new vegetative tissue following flowering. Perennial grasses will exhibit regrowth by producing new tillers following organ senescence, whereas annual grasses do not regrow.

Through the process of regrowth, dormant axillary buds become activated and grow out as vegetative organs. Axillary bud dormancy and activation has been well-studied, and it is likely that perennial regrowth involves modulation of these pathways. The gene *Teosinte branched1* (*Tb1*) encodes a transcription factor that activates *Grassy tillers1* (*Gt1*) that in turn suppresses axillary buds from developing into tillers (Doebley et al. 1995; Whipple et al. 2011; Dong et al. 2019). Numerous inputs can suppress either of these factors to modulate tiller growth. Perennial regrowth likely functions by similarly acting upon this signaling module to suppress then activate tillers in response to temporal or seasonal cues. Two known physiological signals that affect *Tb1* and *Gt1* expression could play a potential role in regulating tillers in the context of perennial regrowth: plant age and carbohydrate mobilization.

One of the major transitions in a plant's life cycle is the juvenile-to-adult transition (JAT) which reflects morphological changes and the plant gaining competency to flower. JAT is typically considered terminal, where the reversion back to the juvenile form only occurs in the seedling produced by the adult plant. However during perennial regrowth of some species, juvenile tissues form from a previously-transitioned adult plant. It is unclear how this reversion to the juvenile form occurs, but the answer likely involves a well-conserved microRNA module that controls plant age and the JAT. Juvenile plants express high levels of miR156, a microRNA that inhibits expression of the *SQUAMOSA PROMOTER BINDING PROTEIN-LIKE* (*SPL*) genes (Wu et al. 2009). *SPLs* in turn activate another microRNA, miR172 that is highly-expressed in adult-transitioned tissues. miR156 and miR172 expression display an antagonistic relationship where, plant tissues will transition from a juvenile state with high miR156 and low miR172 levels, and the adult state containing low miR156 and high miR172. *SPL* transcription factors and miR172 target modules that affect flowering competency and epidermal patterning,

two traits classically associated with plant JAT. High miR156 levels are also associated with increased branching (Chuck et al. 2007; Jiao et al. 2010) and in rice, this was shown to be at least in part due to the activation of *OsTB1* by the SPL14-homolog, IPA (Lu et al. 2013). This link between plant age and branching pathways, along with the reversion of plant age seen in perennials, leads us to speculate that this intersection might lay host to a developmental mechanism that controls axillary bud growth in the context of perennial regrowth.

Another signal that might affect perennial regrowth by modulating the *Tb1/Gt1* pathway is carbohydrate mobilization. Carbohydrates produced through photosynthesis are transported to sink tissues in their mobile form of sucrose. During flowering, an annual plant dedicates most of its carbohydrates to the developing inflorescence. In addition to inflorescences, perennials also direct sucrose towards storage tissues and these resources will be re-invested into new developing axillary branches (Komor 2000; Purdy et al. 2015). Trehalose 6-phosphate (T6P) is derived from sucrose and acts as a signal of available sucrose (Figueroa and Lunn 2016). Dormant axillary buds become de-repressed when sucrose is made available to them, and this process is mediated by T6P (Fichtner et al. 2017; Dong et al. 2019). Genes involved in T6P metabolism are direct targets of TB1 in axillary buds (Dong et al. 2019), providing a direct link between the *Tb1/Gt1* module and sucrose/T6P signaling. Because it relies on remobilization of sugars into newly-growing axillary buds, perennial regrowth may also involve elements of sucrose or T6P signaling pathways.

We hypothesize that perennial regrowth involves modulation of the conserved branching pathway by processes involved in the age-related pathway and/or sugar signaling, but very little is currently known about the genetic control of regrowth. One challenge in scoring regrowth in a

perennial species is this trait manifests long after seed sowing which may account for the lack of attention in this field. Maize, for example, typically sets seed 3-4 months after planting, and robust perennial regrowth must be scored after this period when annuals would normally senesce and die. This scenario presents technical challenges involved in growing and keeping alive many plants for prolonged growth periods in population sizes suitable for quality genetic mapping studies. Regrowth was scored in Johnsongrass but this was preceded by a winter killing-frost. In this study, regrowth was associated with rhizome QTL which is consistent with the role of rhizomes in over-wintering (Paterson et al. 1995).

We now know that, at least in some cases, rhizome development and perennial regrowth can be genetically decoupled. In a study of the perennial wheatgrass species *Thinopyrum elongatum*, perennial regrowth was looked at independently of rhizome growth (Lammer et al. 2004). The authors used chromosome addition lines into the annual Chinese spring wheat background to determine that just the addition of chromosome 4E from *T. elongatum* was sufficient for perennial regrowth. Rhizomes were not observed in *T. elongatum* and it was noted that when these plants were screened during a harsh winter, regrowth was not observed possibly due to freezing conditions. Two main points can be concluded from this study. First, perennial regrowth in grasses can occur entirely without rhizomes, although rhizomes can help the plant over-winter. Second, since the entire perennial trait could be conferred to an annual species with only the addition of a single chromosome, perennial regrowth may not be a complex multigenic trait like rhizome development.

Regrowth was recently mapped in the teosinte species *Z. diploperennis*. In F2 mapping populations derived from *Z. diploperennis* crossed to B73, two major loci conferring perenniality

were discovered on chromosomes 2 and 7, and were named *regrowth1* (*reg1*) and *reg2*, respectively (Ma et al. 2019). Both QTL were fully dominant and complementary to each other, such that either QTL was sufficient to confer perennial regrowth. Although rhizomes were previously found in *Z. diploperennis* mapping populations (Westerbergh and Doebley 2004), rhizomes were infrequently observed in the mapping population where *reg1* and *reg2* were segregating (Ma et al. 2019). This study therefore supports the hypothesis that in grasses, perennial regrowth may be under simple genetic control and can be decoupled from rhizome growth.

**Chapter 4** highlights our work in mapping perennial regrowth in *Z. diploperennis*. We used two new mapping populations derived from crosses between *Z. diploperennis* and two maize NAM lines: the sweetcorn P39 and the popcorn Hp301. We used QTL-seq to map two perennial loci in the P39-derived population: one of which was located near *reg1* on chromosome 2 and a novel locus which we named *reg3* on chromosome 8. When we genotyped F<sub>2</sub> individuals using markers on chromosomes 2, 7, and 8, *reg1* and *reg3* were significantly associated with regrowth in the P39-derived population but only *reg1* was associated in the Hp301-population. When QTL-seq was performed for tiller number in the P39-derived population, tiller number QTL did not appear to be associated with either *reg1* or *reg3*, indicating that genetic control of tiller number is distinct from that of perennial regrowth. Our results will contribute to future efforts to clone genes that regulate the perennial life cycle.

## REFERENCES

Albert PS, Gao Z, Danilova TV, Birchler JA. 2010. Diversity of chromosomal karyotypes in maize and its relatives. *Cytogenet Genome Res* **129**: 6–16.

- Ananiev EV, Phillips RL, Rines HW. 1998. A knob-associated tandem repeat in maize capable of forming fold-back DNA segments: are chromosome knobs megatransposons? *Proc Natl Acad Sci U S A* **95**: 10785–10790.
- Chuck G, Cigan AM, Saeteurn K, Hake S. 2007. The heterochronic maize mutant Corngrass1 results from overexpression of a tandem microRNA. *Nat Genet* **39**: 544–549.
- Creighton HB, McClintock B. 1931. A Correlation of Cytological and Genetical Crossing-Over in *Zea Mays*. *Proc Natl Acad Sci U S A* **17**: 492–497.
- Crews TE, Cattani DJ. 2018. Strategies, Advances, and Challenges in Breeding Perennial Grain Crops. *Sustain Sci Pract Policy* **10**: 1–7.
- Dawe RK, Cande WZ. 1996. Induction of centromeric activity in maize by suppressor of meiotic drive 1. *Proc Natl Acad Sci U S A* **93**: 8512–8517.
- Dawe RK, Lowry EG, Gent JI, Stitzer MC, Swentowsky KW, Higgins DM, Ross-Ibarra J, Wallace JG, Kanizay LB, Alabady M, et al. 2018. A Kinesin-14 Motor Activates Neocentromeres to Promote Meiotic Drive in Maize. *Cell* **173**: 839–850.e18.
- de Oliveira G, Brunzell NA, Sutherlin CE, Crews TE, DeHaan LR. 2018. Energy, water and carbon exchange over a perennial *Kernza* wheatgrass crop. *Agric For Meteorol* **249**: 120–137.
- Doebley J, Stec A, Gustus C. 1995. *teosinte branched1* and the origin of maize: evidence for epistasis and the evolution of dominance. *Genetics* **141**: 333–346.
- Dong Z, Xiao Y, Govindarajulu R, Feil R, Siddoway ML, Nielsen T, Lunn JE, Hawkins J, Whipple C, Chuck G. 2019. The regulatory landscape of a core maize domestication module controlling bud dormancy and growth repression. *Nat Commun* **10**: 1–15.
- Emmerling MH. 1959. Preferential Segregation of Structurally Modified Chromosomes in Maize. *Genetics* **44**: 625–645.
- Fichtner F, Barbier FF, Feil R, Watanabe M, Annunziata MG, Chabikwa TG, Höfgen R, Stitt M, Beveridge CA, Lunn JE. 2017. Trehalose 6-phosphate is involved in triggering axillary bud outgrowth in garden pea (*Pisum sativum* L.). *Plant J* **92**: 611–623.
- Figueroa CM, Lunn JE. 2016. A Tale of Two Sugars: Trehalose 6-Phosphate and Sucrose. *Plant Physiol* **172**: 7–27.
- Food and Agriculture Organization of the United Nations. 2018. *Perennial Crops for Food Security: Proceedings of the FAO Expert Workshop 28-30 August, 2013, Rome, Italy*. Food & Agriculture Org.
- Gicking AM, Swentowsky KW, Dawe RK, Qiu W. 2018. Functional diversification of the kinesin-14 family in land plants. *FEBS Lett* **592**: 1918–1928.
- Hiatt EN, Dawe RK. 2003. Four loci on abnormal chromosome 10 contribute to meiotic drive in maize. *Genetics* **164**: 699–709.
- Hiatt EN, Kentner EK, Dawe RK. 2002. Independently regulated neocentromere activity of two classes of tandem repeat arrays. *Plant Cell* **14**: 407–420.
- Higgins DM, Lowry EG, Kanizay LB, Becraft PW, Hall DW, Dawe RK. 2018. Fitness Costs and Variation in Transmission Distortion Associated with the Abnormal Chromosome 10 Meiotic Drive System in Maize. *Genetics* **208**: 297–305.
- Hu FY, Tao DY, Sacks E, Fu BY, Xu P, Li J, Yang Y, McNally K, Khush GS, Paterson AH, et al. 2003. Convergent evolution of perenniality in rice and sorghum. *Proc Natl Acad Sci U S A* **100**: 4050–4054.

- Jiao Y, Wang Y, Xue D, Wang J, Yan M, Liu G, Dong G, Zeng D, Lu Z, Zhu X, et al. 2010. Regulation of OsSPL14 by OsmiR156 defines ideal plant architecture in rice. *Nat Genet* **42**: 541–544.
- Kanizay LB, Albert PS, Birchler JA, Dawe RK. 2013. Intragenomic conflict between the two major knob repeats of maize. *Genetics* **194**: 81–89.
- Kikudome GY. 1959. Studies on the Phenomenon of Preferential Segregation in Maize. *Genetics* **44**: 815–831.
- Komor E. 2000. The physiology of sucrose storage in sugarcane. In *Developments in Crop Science* (eds. A.K. Gupta and N. Kaur), Vol. 26 of, pp. 35–53, Elsevier.
- Kong W, Kim C, Goff VH, Zhang D, Paterson AH. 2015. Genetic analysis of rhizomatousness and its relationship with vegetative branching of recombinant inbred lines of Sorghum bicolor × S. propinquum. *Am J Bot* **102**: 718–724.
- Lammer D, Cai X, Arterburn M, Chatelain J, Murray T, Jones S. 2004. A single chromosome addition from *Thinopyrum elongatum* confers a polycarpic, perennial habit to annual wheat. *J Exp Bot* **55**: 1715–1720.
- L. Fernando A, Rettenmaier N, Soldatos P, Panoutsou C. 2018. 8 - Sustainability of Perennial Crops Production for Bioenergy and Bioproducts. In *Perennial Grasses for Bioenergy and Bioproducts* (ed. E. Alexopoulou), pp. 245–283, Academic Press.
- Liu J, Seetharam AS, Chougule K, Ou S, Swentowsky KW, Gent JI, Llaca V, Woodhouse MR, Manchanda N, Presting GG, et al. 2020. Gapless assembly of maize chromosomes using long-read technologies. *Genome Biol* **21**: 121.
- Lu Z, Yu H, Xiong G, Wang J, Jiao Y, Liu G, Jing Y, Meng X, Hu X, Qian Q, et al. 2013. Genome-wide binding analysis of the transcription activator ideal plant architecture1 reveals a complex network regulating rice plant architecture. *Plant Cell* **25**: 3743–3759.
- Ma A, Qiu Y, Raihan T, Paudel B, Dahal S, Zhuang Y, Galla A, Auger D, Yen Y. 2019. The Genetics and Genome-Wide Screening of Regrowth Loci, a Key Component of Perennialism in *Zea diploperennis*. *G3* **9**: 1393–1403.
- McClintock B. 1929. CHROMOSOME MORPHOLOGY IN ZEA MAYS. *Science* **69**: 629.
- Murray SC, Jessup RW. 2014. Breeding and genetics of perennial maize: Progress, opportunities and challenges. In *Perennial Crops for Food Security, Proceedings of the FAO Expert Workshop*, pp. 103–111.
- Oliveira G de, de Oliveira G, Brunsell NA, Crews TE, DeHaan LR, Vico G. 2020. Carbon and water relations in perennial Kernza (*Thinopyrum intermedium*): An overview. *Plant Science* **295**: 110279. <http://dx.doi.org/10.1016/j.plantsci.2019.110279>.
- Paterson AH, Schertz KF, Lin YR, Liu SC, Chang YL. 1995. The weediness of wild plants: molecular analysis of genes influencing dispersal and persistence of johnsongrass, *Sorghum halepense* (L.) Pers. *Proc Natl Acad Sci U S A* **92**: 6127–6131.
- Peacock WJ, Dennis ES, Rhoades MM, Pryor AJ. 1981. Highly repeated DNA sequence limited to knob heterochromatin in maize. *Proc Natl Acad Sci U S A* **78**: 4490–4494.
- Purdy SJ, Cunniff J, Maddison AL, Jones LE, Barraclough T, Castle M, Davey CL, Jones CM, Shield I, Gallagher J, et al. 2015. Seasonal Carbohydrate Dynamics and Climatic Regulation of Senescence in the Perennial Grass, *Miscanthus*. *Bioenergy Res* **8**: 28–41.
- Rhoades M, Dempsey E. 1986. Evidence that the K10 knob is not responsible for preferential segregation and neocentromere activity. *Maize Genet Coop News Lett* **60**: 26–27.
- Rhoades MM. 1942. Preferential Segregation in Maize. *Genetics* **27**: 395–407.

- Rhoades MM, Dempsey E. 1966. The Effect of Abnormal Chromosome 10 on Preferential Segregation and Crossing over in Maize. *Genetics* **53**: 989–1020.
- Rhoades MM, Others. 1952. Preferential segregation in maize. *Preferential segregation in maize*. <https://www.cabdirect.org/cabdirect/abstract/19531603828>.
- Rhoades MM, Vilkomerson H. 1942. On the Anaphase Movement of Chromosomes. *Proc Natl Acad Sci U S A* **28**: 433–436.
- Westerbergh A, Doebley J. 2004. Quantitative trait loci controlling phenotypes related to the perennial versus annual habit in wild relatives of maize. *Theor Appl Genet* **109**: 1544–1553.
- Whipple CJ, Kebrom TH, Weber AL, Yang F, Hall D, Meeley R, Schmidt R, Doebley J, Brutnell TP, Jackson DP. 2011. grassy tillers1 promotes apical dominance in maize and responds to shade signals in the grasses. *Proc Natl Acad Sci U S A* **108**: E506–12.
- Wu G, Park MY, Conway SR, Wang J-W, Weigel D, Poethig RS. 2009. The sequential action of miR156 and miR172 regulates developmental timing in Arabidopsis. *Cell* **138**: 750–759.
- Yu HG, Hiatt EN, Chan A, Sweeney M, Dawe RK. 1997. Neocentromere-mediated chromosome movement in maize. *J Cell Biol* **139**: 831–840.
- Zhang Y, Li Y, Jiang L, Tian C, Li J, Xiao Z. 2011. Potential of Perennial Crop on Environmental Sustainability of Agriculture. *Procedia Environmental Sciences* **10**: 1141–1147.

## CHAPTER 2

# MOLECULAR CHARACTERIZATION OF THE KNOB180 NEOCENTROMERE MOTOR

### INTRODUCTION

Maize chromosomes have been long known to contain features called knobs that have served as cytological markers. When viewed under the microscope, knobs appear as heterochromatin on characteristic chromosomal locations. We now know that knobs are composed at the molecular level of two types of tandemly-repeated sequences: knob180 (Peacock et al. 1981) and TR-1 (Ananiev et al. 1998). A large haplotype on the end of the long arm of chromosome 10 called Abnormal chromosome 10 (Ab10) has three small TR-1 knobs and a large knob180 knob (Hiatt et al. 2002; Liu et al. 2020). This differs from the standard version of this chromosome, Normal chromosome 10 (N10) that has no knobs or distinctive cytological features. At the base of both haplotypes is a gene that controls kernel color called *R1*, where dominant *R1* causes purple kernels over the recessive *r1*. When a heterozygote (i.e. *R1/r1*) containing N10 is test-crossed as a female, each copy of *R1* segregates evenly such that the resulting ears will have 50% purple and 50% yellow kernels. Rhoades noted however that if Ab10 is present, the *R1* allele linked to it is typically transmitted at frequencies of 70-80% (Rhoades 1942). To help explain this fascinating phenomenon during the golden age of cytogenetics, Rhoades began to analyze the cytology of Ab10.

When observing Ab10 in meiosis, Rhoades discovered an astonishing property of this haplotype. In plants where Ab10 was present, Rhoades observed anaphase chromosomes were led not by centromeres but by their knobs so he called these structures neocentromeres (Rhoades and Vilkomerson 1942). He then devised a model to connect the cytological nature of neocentromeres to the genetic observation of segregation distortion (Rhoades 1952). Meiotic drive occurs in female meiosis when Ab10 is heterozygous (i.e. Ab10/N10). First, a crossover occurs between the centromere and Ab10, resulting in a pair of Ab10/N10 heteromorphic dyads. Meiosis proceeds as normal with the exception that neocentromeres pull Ab10 towards the poles in anaphase I and II. Because anaphase II separates sister chromatids, a cell division with a heteromorphic dyad will result in one megaspore with Ab10 and another with N10. Cells would normally receive either sister chromatid randomly with equal chance, but neocentromere activity dictates that the basal cell in the linear tetrad will more frequently receive Ab10. Through the natural process of female meiosis in plants, only the bottom of the four megaspore cells develops into the megagametophyte and transmits the knobbed chromosome (Figure 3.1).

Meiotic drive was initially described for Ab10, but this form of segregation distortion can be initiated by knobs in other locations of the genome (Kikudome 1959; Rhoades and Dempsey 1966; Kanizay et al. 2013). This is consistent with the observation that when Ab10 is present, all knobs -- not just the knobs on Ab10 -- become neocentromeres. With our current view of molecular biology, this can be understood as a gene product that is encoded on Ab10 and acts throughout the cell when expressed. Kinesin-14s are a family of eukaryotic motor proteins that have expanded in plants (Gicking et al. 2018), so it was hypothesized that the gene responsible for neocentromeres encoded a kinesin (Yu et al. 1997). Finding such a *trans*-acting factor that is

present on Ab10 and is sufficient to convert knobs to neocentromeres has been an area of focus for research on maize meiotic drive.

It is not possible to map neocentromere activity using traditional recombination-based mapping approaches because recombination does not occur between the Ab10 and N10 haplotypes (Mroczek et al. 2006). Several Ab10 terminal deletions have been generated with breakpoints at different locations along the haplotype. Mapping with terminal deficiencies has shown that distinct factors control the neocentromere activity of the two knob repeats independently (Hiatt et al. 2002). When successive terminal deficiencies were scored for knob180 and TR-1 neocentromere activity, the positions of both factors were revealed (Hiatt and Dawe 2003a). The TR-1 neocentromere factor was mapped to a locus in the proximal region (Hiatt et al. 2002; Hiatt and Dawe 2003a). We have recently characterized the gene in this region that encodes the TR-1 neocentromere motor, TRKIN (Chapter 3, (Swentowsky et al. 2020)).

Knob180 neocentromeres are controlled by a locus distal to the large knob180 knob, in a euchromatic region known as the distal tip. A series of Ab10 mutants that did not show meiotic drive were identified through a *Mu* transposable element screen and were thus termed *suppressor of meiotic drive (smd)* mutants (Dawe and Cande 1996; Hiatt and Dawe 2003a; Dawe et al. 2018). A family of nine tandemly-repeated kinesin-14 genes was discovered on the distal tip of Ab10 and the genes were subsequently named *Kindr (Kinesin driver)* to reflect their demonstrated role in neocentromere activity and meiotic drive. Most *smd* mutants can be attributed to the loss of *Kindr* expression in either Ab10 deletions (*smd3* and *smd8*) or epimutants with hypermethylated DNA (*smd1* and *smd12*) (Dawe et al. 2018). Characterized *smd* mutants show a loss of knob180 neocentromeres but retain TR-1 neocentromere activity, further

supporting the idea that the two classes of knobs are controlled by different motors (Dawe et al. 2018). Although we have discovered the motors that power the two types of neocentromeres, other components that are involved in these complexes remain a mystery.

Here we present molecular characterization of the KINDR protein and further work in elucidating the molecular nature of the knob180 neocentromere motor. We show that KINDR is expressed in Ab10 meiotic tissues and localizes specifically to knob180 knobs during meiosis. Since the publication of these results (Dawe et al. 2018), additional work was directed towards understanding how KINDR interacts with DNA. Data described here demonstrate that purified KINDR does not physically associate with knob180 *in vitro*, suggesting that the interaction between KINDR and knob repeats is indirect. The gene encoded by the *smd13* locus is a likely candidate for an adapter protein that bridges KINDR to DNA. In support of this view, genetic data show that KINDR fails to localize to knobs in *smd13* mutants. *smd8* is another Ab10 mutant that is caused by a ~2Mb deletion on the Ab10 distal tip. We performed a complementation test by assaying KINDR localization in a *smd13/smd8* mutant background. In this genotype we also observed improper KINDR localization, suggesting the *smd13* mutation lies within the *smd8* deletion. Using *de novo* transcriptome assembly, we identified a *Kinesin-10-like* gene that is differentially expressed in *smd13*. *Kinesin-10-like* is encoded by genes at a six tandem copy locus immediately distal to the *Kindr* complex. This region is included within the *smd8* deletion, but no putative mutations or epi-mutations to explain the loss of *kinesin-10-like* expression in *smd13* were observed. We also performed immunoprecipitation followed by mass spectrometry (IP-MS) using a rabbit anti-KINDR antibody in Ab10 ear tissue but did not observe any likely KINDR-binding proteins. This work highlights candidates that may be used for future study of

proteins that associate with KINDR and help form the interface between the knobs and microtubules.

## RESULTS

KINDR is expressed in Ab10 meiotic tissues and localizes to knob180 knobs during meiosis

RNA-seq and quantitative PCR results showed that *Kindr* transcripts are expressed in male and female meiotic tissues (Dawe et al. 2018). To study KINDR at the protein level, we generated an antibody against an epitope specific to the N-terminal end of the predicted KINDR protein (Dawe et al. 2018). We found this antibody to have highly specific affinity to KINDR when tested on a western blot. A signal around the predicted molecular weight of KINDR (69 kDa) was detected in Ab10 ears and anthers but was absent in lanes where N10 tissue was loaded (Figure 2.1A). Although *Kindr* transcripts were previously found in Ab10 leaves by quantitative-RT-PCR, we did not detect KINDR protein in Ab10 leaves assayed by western blot. Such a discrepancy could be due to different ages that leaves were collected or reflect a true difference between transcript and protein abundance in leaves. Nevertheless, the presence of KINDR protein within Ab10 anthers and ears that contain meiotic cells supports the hypothesis that KINDR facilitates meiotic drive.

We next performed immunofluorescence microscopy using the KINDR antibody to probe Ab10 meiocytes and observe this protein's localization pattern. In anaphase II cells, KINDR was localized to puncta at neocentromeres near the spindle poles (Figure 2.1B). This localization

pattern is consistent with a motor that binds knobs and facilitates neocentromere movement. To determine if KINDR localizes to a specific type of knob, we combined immunofluorescence with FISH using probes against the knob180 and TR-1 knob repeats. The data show specific association between KINDR and knob180 knobs, whereas TR-1 knobs are completely devoid of KINDR signal (Figure 2.1C). This is consistent with findings that *kindr* mutants are associated with loss of knob180 neocentromeres but still retain TR-1 neocentromere activity. Together with the finding that KINDR is a functional minus end-directed motor (Dawe et al. 2018), these data strongly indicate that KINDR is the knob180 neocentromere motor that facilitates Ab10 meiotic drive.

KINDR likely does not physically associate with knob180 knobs

The physical nature of the knob-KINDR interface and what components are required for neocentromere activity is still unclear. The predicted KINDR protein does not contain a known DNA-binding domain, but a physical association between KINDR and knobs has not previously been tested. To test if KINDR can directly bind knob180 *in vitro*, a 6xHis-tagged version of KINDR was expressed in *E. coli* cells and purified. A biotinylated knob180 probe was used to test for a physical interaction between 6xHis-KINDR and knob180 using an electrophoretic mobility shift assay (EMSA). As a positive control for this assay, a 6xHis-tagged version of the kinetochore protein CENPC that was previously shown to bind DNA non-specifically was tested (Du et al. 2010). A gel shift of the knob180 probe was observed when 6xHis-CENPC was used (Figure 2.2A). When 6xHis-KINDR was tested using this assay, no gel shift was seen, suggesting that KINDR does not directly interface with knob180 to facilitate neocentromere activity (Figure 2.2A). This assay is not definitive, however, as EMSAs can lead to false

negatives. Since 6xHis-KINDR was purified from *E. coli*, it is possible that this protein does not reflect the *in vivo* form of KINDR found in maize and that these differences are responsible for DNA-binding. A false negative result could also have been obtained due to improper binding conditions, although many different buffer mixtures were tested which yielded the same negative result. In conclusion, EMSA results suggest that KINDR does not directly bind knob180 DNA, although it is possible this result is a false negative. Additional protein-DNA interaction assays such as ChIP or yeast one-hybrid should be used to conclude whether or not KINDR directly associates with knob180 knobs.

*smd13* lacks proper KINDR localization and is allelic to the *smd8* deletion

Ab10 *suppressor of meiotic drive* (*smd*) mutants may help dissect the nature of the KINDR-knob180 interaction. Whereas the *smd1*, *smd3*, *smd8*, and *smd12* show decreased *Kindr* expression, wild-type *Kindr* levels were detected in *smd13* (Higgins 2017). When *smd13* anthers and ears were tested for KINDR protein levels on a western blot, KINDR levels were similar to the Ab10 control (Figure 2.2B). This finding indicates that the lack of meiotic drive observed in *smd13* is not due to a loss of KINDR expression. We assayed the localization of KINDR during meiosis in an *smd13* background, since we had previously shown that KINDR localizes to knobs during meiosis (Figure 2.1). Although KINDR expression was found in meiotic tissues, localization of KINDR at knobs was not observed in *smd13* (Figure 2.2C). This suggests *smd13* lacks an unknown gene product required for proper KINDR localization. When *smd13* was complemented with the wild-type form of Ab10 (i.e. *smd13*/Ab10 heterozygote), proper KINDR localization on knobs was restored, although localization was only assayed in one plant (Figure 2.2D). This result demonstrates that *smd13* carries a recessive mutation that disrupts KINDR

localization. We next attempted to characterize the nature of the *smd13* mutation to find the gene required to localize KINDR properly.

To characterize the nature of the *smd* mutants, we first searched for large deletions in these mutant genotypes. We resequenced wild-type Ab10 as well as *smd3*, *smd8*, and *smd13* mutants and mapped the resulting Illumina reads to the Ab10 assembly. *smd3* and *smd8* display large regions with missing coverage compared to wild-type Ab10 originating within the *Kindr* complex and extending towards the distal tip (Figure 2.3A). This loss of coverage indicates that *smd3* and *smd8* both contain deletions along the distal tip. Since the *Kindr* complex is composed of nine tandemly-repeated units, reads do not uniquely map in this region and it is not possible to determine the exact proximal breakpoint in *smd3* and *smd8* but in both cases, the deletion appears to begin in the *Kindr* complex at about 190 Mb on chromosome 10. Low coverage is observed along the rest of the Ab10 chromosome in *smd3*, so we determined this line carries a terminal deletion of the Ab10 haplotype originating in the *Kindr* complex at about 190 Mb. The loss of low coverage in *smd8* is again observed at about 192.3 Mb, so we can conclude this line contains an internal Ab10 deletion from ~190-192.3 Mb. Interestingly, the *smd8* region includes six tandemly-repeated units, much like the nine-copy *Kindr* complex immediately proximal to it (Figure 2.3B). Since *smd13* and *smd8* fail to complement (Figure 2.2C), the *smd13* lesion may lie within these repeats.

In contrast to large deletions found in *smd3* and *smd8*, reads obtained from *smd13* resequencing do not display any obvious coverage differences relative to the resequenced Ab10 control line (Figure 2.3A). This rules out the possibility that *smd13* is caused by a large deletion of the Ab10 chromosome. Furthermore, *smd1* and *smd12* do not contain visible Ab10 deletions

(data not shown) but are *kindr* epi-mutants and show lower *Kindr* expression due to increased DNA methylation in the *Kindr* promoters and gene bodies (Dawe et al. 2018). Since the *smd13* phenotype is not caused by a large deletion, we attempted to map this mutation using additional techniques.

### *smd13* SNP Calling

Since we failed to observe an obvious deletion in the *smd13* genotype, we next attempted to identify SNPs or short indels that might explain the lack of meiotic drive in this mutant. Ab10 and *smd13* genotypes were resequenced using Illumina short reads to ~20X coverage. Reads were mapped to the Ab10 assembly and only those mapping to the Ab10 haplotype (chr10:142500000-195026473) were retained for analysis. SNP calling was performed using the GATK4 best practices (DePristo et al. 2011) and SNPs were filtered to retain SNPs that were called homozygous for the alternate allele in *smd13*, homozygous for the reference allele in Ab10, *smd13* genotype quality above 19, and *smd13* depth between 3-30. In total, 159 SNPs were retained including 26 within genes and 37 SNPs located on the Ab10 distal tip (Figure 2.4A). There are eleven SNPs within the *smd8* deleted region, and ten of these lie within exon 5 and the intron immediately downstream of it within *KindrF6*. However, the raw alignment shows the reads mapped to this region contain numerous SNPs and mapped poorly (Figure 2.4B), so it is likely that these resulted from mismapped reads originating from another copy of *Kindr* or pseudo-*Kindr*. The single remaining SNP within the *smd8* deletion lies 56 kb from the nearest

annotated gene whose predicted encoded protein does not align to any known proteins. In summary, analysis of *smd13* SNPs did not reveal any obvious gene candidates to facilitate a KINDR-knob180 interaction.

RNA-seq using gene annotations does not reveal a likely *smd13* candidate gene

To attempt to identify the causal deficient gene in *smd13*, we performed RNA-seq using young ear tissue of Ab10, N10, Df(L) (a terminal deletion of the Ab10 distal tip), *smd13*, and *smd8* all back-crossed into a B73 background. Reads were mapped to the Ab10 assembly and expression was quantified using predicted transcripts from the Ab10 annotation (Liu et al. 2020). The majority of transcripts along Ab10 have low expression in the B73 genotype because the entire ~45 Mb region is absent in Ab10 (Figure 2.5). Df(L) is missing the entire Ab10 distal tip which spans approximately 183-195 Mb on chromosome 10 and has a similar significant decrease in expression of transcripts in this region. Along the Ab10 haplotype, nine genes are differentially expressed in *smd13*, with an adjusted p-value below 0.95, and four of these are similarly differentially expressed in *smd8* (Table 2.1). None of these genes, however, lie within the predicted *smd8* deletion of ~190-192Mb or contained *smd13* SNPs. Of the nine genes, eight are predicted to encode proteins that have close matches to previously annotated proteins in maize. Additionally, none of these predicted proteins contain known domains involved in DNA-binding or protein-protein interactions.

Trinity assembly reveals a likely candidate for the gene deficient in *smd13*

Gene annotations within Ab10 are less reliable than those elsewhere in the genome, since predicted transcripts are shorter and have greater overlap with transposable elements (Liu et al. 2020). Additionally, genes within tandemly-repeated blocks are difficult to accurately annotate. For these reasons, we pursued a *de novo* transcriptome approach to identify putative *smd13* genes. RNA-seq reads generated from three replicates each of Ab10 ears and anthers were used to assemble 496,228 raw transcripts using Trinity *de novo* transcript assembly (Grabherr et al. 2011). We took advantage of the Ab10 variant Df(L), a terminal deficiency lacking the entire Ab10 distal tip (Hiatt and Dawe 2003b), to retain transcripts that are likely to be encoded on the Ab10 distal tip. Using DESeq2 (Love et al. 2014), we found that 1,572 transcripts were differentially expressed between Ab10 and Df(L) ears, and 549 of these had  $-2.5$   $\log_2$ FoldChange expression (Figure 2.6A). These 549 transcripts represent those that are expressed in Ab10 but are missing from Df(L) plants. We then performed a BLAST alignment between each of these transcripts and the transcripts annotated within the Ab10 assembly. We found a significant enrichment of 124/549 (22.6%) of Df(L)-depleted genes that are located within the Ab10 distal tip. Therefore, this set of 549 Df(L)-depleted genes contains many of the distal tip-encoded genes, likely including those that are missing from the Ab10 assembly-based gene annotation (Figure 2.6B).

We next looked for differential expression of the Df(L)-depleted transcripts between Ab10 and *smd13* using RNA-seq Illumina reads generated from ear tissue. 66 transcripts were differentially expressed, including one transcript that can be mapped to the Ab10 distal tip (Figure 2.6C). This transcript is predicted to encode a cytochrome P450, the largest enzyme family found in plants. Cytochrome P450s are almost exclusively involved in biosynthesis of secondary metabolites (Xu et al. 2015), so it is unlikely that this transcript is involved in the

KINDR-knob180 interaction. Another transcript, TRINITY\_DN45531\_c2\_g1\_i24, matches an annotated transcript found on one of the unincorporated scaffolds that were not incorporated into assembled chromosomes. A BLAST search using this Trinity transcript as a query finds a 100% identical match to the BAC containing *KindrB10* used in the original identification of *Kindr* on Ab10 (GenBank: KX759203.1). This is the most proximal copy of *Kindr*, so the BAC likely contains additional sequence proximal to the *Kindr* complex, into the *smd8*-deleted region which itself contains a tandemly-repeated unit. Only one copy of TRINITY\_DN45531\_c2\_g1\_i24 is found within the sequenced BAC but when a BLAST search is performed against the Ab10 assembly querying this predicted transcript, six tandem copies are found in the region proximal to the *Kindr* complex (Figure 2.7).

The full predicted TRINITY\_DN45531\_c2\_g1\_i24 is 3,923 bp long, but aligning Ab10 RNA-seq reads to this sequence reveals that most reads pile up between about 2.4-3.4 kb (Figure 2.8A) indicating the biologically-relevant form of this transcript is likely to be represented by this 1kb segment. This transcript was expressed at about 50% the level of Ab10 in *smd13* ears (Figure 2.8B), although this difference was not statistically significant ( $p=0.0918$ , unpaired T-test). Expression was absent in *smd8*, Df(L), and B73 ears, which was expected since the chromosomal segment this transcript is produced from is missing in these three genotypes. Two predicted ORFs can be found within this 1 kb transcript (Figure 2.8A). The 5' ORF is 300 bp and its predicted protein produces no significant alignment with known proteins. The 3' ORF is 306 bp and is predicted to encode the C-terminal end of a kinesin-10 protein. The two ORFs overlap each other by 4-bp. Recent examples from cotton (Wang et al. 2019) and green algae (Gallaher et al. 2021) have demonstrated that multiple proteins can be translated from a single ORF on a polycistronic transcript. Therefore, it is conceivable the 3' ORF that encodes a predicted kinesin-

10 ortholog may be translated. In animals, kinesin-10s act as chromokinesins that provide the force to push chromosome arms towards the spindle equator and are important for proper chromosome segregation in cell division (Funabiki and Murray 2000; Antonio et al. 2000). Since chromokinesins can directly bind DNA and serve as an interface between chromatin and the microtubule spindle (Almeida and Maiato 2018), this predicted protein fits the hypothesized *smd13* function and was further analyzed.

*Kinesin-10-like* is predicted to encode the C-terminal end of a kinesin-10 protein, so we predicted this protein's putative domains (Figure 2.8C). Kinesin-10s contain an N-terminal kinesin motor domain, which is missing from *Kinesin-10-like*. The N-terminal portion of *Kinesin-10-like* contains a short ~14-residue coiled-coil domain. Coiled-coils are frequently involved in protein-protein interactions and in kinesins often facilitate either homo- or hetero-oligomerization. The most notable feature of kinesin-10s is a helix-hairpin-helix (HhH) motif found near the C-terminal end of the protein (Almeida and Maiato 2018). HhH domains bind DNA and are composed of two alpha helices separated by a GhG hairpin, where h is one of the nonpolar residues I, L, or V (Shao and Grishin 2000). Two short alpha-helices separated by a KGIG motif are found near the center of *Kinesin-10-like*, so this motif is likely a functional HhH DNA-binding domain (Figure 2.8C).

We were unable to determine an underlying genetic cause for the loss of *Kinesin-10-like* expression in *smd13* ears. No SNPs are present among all *Kinesin-10-like* copies and no obvious changes in read coverage which would indicate a deletion are observed in *smd13* (Figure 2.4). Several *kindr* mutants are epimutants that display both hypermethylation and an increase in siRNAs that map to *Kindr* (Dawe et al. 2018). Neither of these epigenetic signatures were

observed at *Kinesin-10-like* in *smd13* (Jonathan Gent, personal communication) so the genetic cause for its expression decrease is still unknown.

We have formulated a testable hypothesis that the Kinesin-10-like may associate with KINDR through its N-terminal coiled-coil and binds knobs using its HhH domain. Kinesin-10-like does not contain a motor domain, but hetero-oligomerization between Kinesin-10-like and KINDR may facilitate minus-end motility for this protein. Future work into characterizing this *kinesin-10-like* gene could help determine if it plays a role in Ab10 meiotic drive and the association of KINDR with knobs.

#### IP-Mass Spectrometry

We also took a biochemical approach to find the KINDR-knob180 linking factor. Our rabbit anti-KINDR antibody is both specific and has high affinity to the KINDR protein, as assayed by both western blot and immunofluorescence that represent unfolded and folded forms of the protein, respectively. We carried out immunoprecipitation using rabbit anti-KINDR followed by LC-MS/MS to identify any proteins that co-eluted with KINDR. We observed high levels of KINDR in developing ears (Figure 2.2B). Ear tissue has also been used for successful IP-MS previously (Jia et al. 2020), so we chose to perform IP-MS from native protein isolated from ~3-4 cm developing Ab10 ears. To test for an efficient immunoprecipitation, equal fractions of input, flow-thru, and elution were tested on a western blot using the rabbit anti-KINDR antibody. KINDR was detected in input at its standard 69 kDa region, as well as an unknown signal that is frequently detected when KINDR is present at a much higher molecular weight (Figure 2.9). The flow-thru lacks any KINDR staining, indicating rabbit anti-KINDR efficiently bound KINDR protein and removed it from the solution. The elution lane shows

KINDR staining of similar intensity to the input lane, which shows that KINDR was efficiently eluted from beads. This level of KINDR is not detectable on the accompanying colloidal Coomassie gel, which can detect as little as approximately 10 ng of protein.

IP was performed using rabbit anti-KINDR on three replicates of Ab10 ears and one replicate of Ab10 root tips. The resulting western blot indicates a successful KINDR immunoprecipitation was consistent across two additional biological replicates from ears and one from root tips (data not shown). Following elution, samples were briefly run on an SDS-PAGE gel, then submitted for mass spectrometry analysis using a tryptic digest. Tryptic fragments were searched against the Ab10 protein annotation using the Mascot search engine (Matrix Science). For each protein, Mascot scores are reported that scale logarithmically, and roughly 95% confidence corresponds to a score of around 90 (Perkins et al. 1999; Koenig et al. 2008).

In all samples, the top hit was the KINDR protein KINDR B9, which confirms that this immunoprecipitation procedure efficiently eluted our bait protein (Table 2.2). Unfortunately, no other proteins displayed high confidence across multiple replicates to consider them for additional study. Proteins identified from all immunoprecipitated samples, particularly from root tips, were enriched for proteins involved in protein translation (4/20; 1/9; 4/30; 29/54 proteins from three ear and one root tip samples, respectively). This includes a homolog of elongation factor 1-alpha, the only other protein to be identified in all four samples (Table 2.2). Protein translation is unlikely to be directly involved in the KINDR-knob180 interaction and therefore likely represents KINDR units being translated during the time of immunoprecipitation. Interestingly, a homolog of actin was identified in two Ab10 ear and one Ab10 root tip replicates. This result was unexpected since tubulin, the major cytoskeletal component of the

meiotic spindle that was previously shown to physically associate with KINDR (Dawe et al. 2018), was not identified in our mass spectrometry data. Whether KINDR truly associates with actin and what role this may play in meiotic drive will require additional study to determine. Other notable proteins identified from this experiment include homologs of histone-lysine N-methyltransferase EZ2, cyclin-T1-3, histone H4, histone H2A variant 3, histone H1, Cinfl1 polyprotein, and Alba DNA/RNA-binding protein, but these proteins were predicted with low Matrix score and were inconsistently identified across multiple replicates.

## **DISCUSSION**

Maize knobs are capable of undergoing meiotic drive when the chromosome 10 haplotype, Ab10, is present. Meiotic drive in maize is facilitated by genes on Ab10 that transform knobs into neocentromeres that arrive at spindle poles before centromeres during meiosis. The knob180 motor protein is encoded by a kinesin-14 gene called *Kindr* that is necessary for Ab10 meiotic drive and knob180 neocentromeres. We showed that KINDR is expressed in Ab10 meiotic tissues and co-localizes to knob180 knobs during meiotic anaphase. Another Ab10 kinesin-14, TRKIN, localizes specifically to TR-1 neocentromeres, but how these two proteins achieve such specific localization is currently unknown.

We know of two classes of proteins that facilitate interactions between microtubules and chromatin: the kinetochore component NDC80 and chromokinesins. Kinetochores are sophisticated multi-protein complexes that link centromeres to microtubule K-fibers. In plants, the interaction between the kinetochore and chromatin is facilitated by the centromeric histone variant CENH3 along with CENPC and KNL2 (Wang and Dawe 2018). CENPC recruits the

MIS12 complex which in turn recruits NDC80, the protein that directly interfaces with microtubules (Cheeseman et al. 2006). Chromokinesins perform additional roles during cell division such as moving chromosome arms towards the metaphase plate (Almeida and Maiato 2018). Two classes of chromokinesins, Kif4a and Kid, possess DNA-binding domains and associate directly with chromatin (Afshar et al. 1995; Lee et al. 2001) but another chromokinesin, Kif15, binds chromatin through interaction with its binding partner KBP (Brouwers et al. 2017).

How KINDR localizes to knob180 knobs and the nature of its association with chromatin is still a mystery. KINDR itself does not contain a known DNA-binding motif and did not show a direct interaction with knob180 DNA by EMSA. While this EMSA assay was not definitive, it is likely that the KINDR-knob180 interaction is facilitated by additional proteins. Our genetic data demonstrate that *smd13* is likely to encode a factor that provides this linking function. Neither *smd13* nor *smd8* mutants show proper KINDR localization, and we observed the heterozygous *smd13/smd8* mutant also had mislocalized KINDR. Since these mutants were incapable of complementing each other, we have genetically proven that *smd13* and *smd8* share a common mutation that disrupts proper KINDR localization. There is a 2 Mb-deletion immediately proximal to the *Kindr* complex present in *smd8* so we can infer the mutation affecting the *smd13* phenotype is present in this location as well. Like the *Kindr* complex that contains nine tandem copies of the *Kindr* gene, the *smd8*-deleted region is composed of six tandemly-repeated units. The quality of gene annotations in this region are poor but using a Trinity-derived *de novo* transcriptome approach, we found a *kinesin-10-like* gene encoded in tandem here.

*Kinesin-10-like* is missing in *smd8* and its expression is significantly reduced in *smd13* compared to Ab10, but we could not determine a genetic cause for this expression difference. We did not detect any SNPs, deletions, or epimutations on Ab10 in the *smd13* genotype that could easily explain the loss of KINDR localization or meiotic drive in this genotype. Still, there are several possibilities why we found no mutation. The *smd8*-deleted region is composed of six tandem copies, and repetitive DNA has long been known to pose significant analytical challenges. Sequence assembly errors frequently occur in tandemly-repeated DNA and an error during this early step may introduce artifacts that are compounded during subsequent analyses (Tørresen et al. 2019). The Ab10 assembly was generated using PacBio and Oxford Nanopore long reads which are superior in handling tandemly-repetitive regions, although these sequencing methods contain a substantially higher rate of errors. To attempt to correct these known issues, long reads were polished using Illumina reads which in turn could introduce assembly errors (Liu et al. 2020). We attempted to search for *smd13* mutations using SNP calling and DNA methylation analyses which both rely on mapping of short reads to this genomic region. Short read coverage is sparse and patchy across this area (Figures 2.3A, 2.7) so the challenging nature of this genomic region raises the concern that we have poor ability to detect mutations here. Therefore, our inability to detect *smd13* mutations to explain either the decrease of *kinesin-10-like* expression or loss of meiotic drive may reflect a false negative due to an inaccurate sequence assembly.

It is also possible that the *smd13* phenotype is indeed unrelated to the decrease in *kinesin-10-like* expression and another gene is responsible for linking KINDR to knobs. We determined that *smd13* and *smd8* do not genetically complement one another which shows they each lack a common gene required for knob180 neocentromeres. The 2 Mb distal tip deletion is the most

obvious defect in *smd8*, but it is possible this genotype contains additional mutations that are shared with *smd13*. We reported four genes that are differentially expressed in both *smd13* and *smd8* based on RNA-seq results from maize ears. These genes could potentially represent the mis-expressed gene that gives rise to both *smd8* and *smd13* phenotypes, but none of these four genes represented obvious candidates for involvement in meiotic drive. While we have high read coverage in the *smd13* genotype, which was sufficient for SNP calling, we only have low coverage reads for *smd8* that are not suitable for accurately calling SNPs. Future efforts may attempt to sequence *smd8* at higher coverage to determine if the two genotypes share an additional mutation.

We attempted to identify KINDR binding partners with a biochemical approach using immunoprecipitation followed by mass spectrometry. While we consistently and confidently identified KINDR itself from immunoprecipitated samples, we failed to detect a putative KINDR binding partner that displayed high confidence across multiple replicates (Table 2.2). While an association between KINDR and these proteins has not conclusively been demonstrated, several of the proteins found in the immunoprecipitated samples may hint at the molecular activity of KINDR. A homolog of actin was found to co-immunoprecipitate with KINDR across three of four biological samples. While microtubules are the major cytoskeletal component involved in meiotic chromosome segregation, actin microfilaments are also found in the meiotic spindle (Staiger and Cande 1991) and it has recently been shown that actin can play roles in meiosis in mammals as well (Uraji et al. 2018). The Arp2/3 complex is involved in the RanGTP-dependent spindle assembly mechanism (Yi et al. 2011) which we proposed may be involved in generating spindle asymmetry during Ab10 meiotic drive (Swentowsky et al. 2020). It was also shown that

actin is required for formation of K-fibers, the thick microtubule threads that attach to kinetochores (Mogessie and Schuh 2017).

Although the statistical support for each of these proteins is weak, three histones were identified as proteins that co-immunoprecipitated with KINDR: histone H4, histone H2A variant 3, histone H1. Histones are the major protein components of nucleosomes and post-translational histone modifications are well-known to affect chromatin compaction such as lysine methylation of histone H3 on knobs (Shi and Dawe 2006). If confirmed, this result could suggest a close interaction between KINDR and chromatin.

There are several biological and technical possibilities why a probable KINDR binding partner was not discovered using IP-MS. KINDR does not contain a nuclear localization signal (NLS) and its localization on knobs is only observed in meiosis after nuclear envelope breakdown in late prophase I (Swentowsky et al. 2020). It is possible that a KINDR binding partner is localized to knobs only during cell division. The ear and root tip tissues assayed contain very few cells undergoing cell division so if this is the case, only a small proportion of KINDR bound to its binding partner would be present and the eluted amount of this protein would be well below the detectable limit of mass spectrometry.

## **MATERIALS AND METHODS**

### Western Blot

Sibling plants either homozygous for Ab10 or homozygous for N10 were grown in parallel. Seedling second leaves, anthers containing meiotic cells, and 5 cm long young ears were

flash frozen in liquid nitrogen. Tissue was ground to a fine powder and protein was extracted in 1mM EDTA, 50mM Tris HCl, 10% glycerol, 150mM NaCl, 0.1% Triton X-100, 1mM DTT, and cOmplete Mini Protease Inhibitor Cocktail (Roche cat no. 11836170001). Protein concentrations were determined using a Bradford assay and 20 mg of total protein was loaded in each well of a 12% mini-PROTEAN TGX Precast Gel (Bio-Rad Cat# 4561043). Protein was transferred to a nitrocellulose membrane (Bio-Rad Cat# 1620168). All washes were performed using Tris-buffered saline with 0.01% tween-20 (TBST). The membrane was blocked for two hours using TBST containing 5% powdered milk. Anti-KINDR antibody was added at 1:500 and incubated overnight at 4C. After washing, the membrane was incubated with blocking buffer containing Rabbit IgG HRP Linked Whole Ab (Sigma cat no. GENA934-1ML) at 1:5000 and incubated at room temperature for two hours. After additional washes, SuperSignal West Dura Extended Duration Substrate (Thermo Fisher cat no. 34075) was applied and the results visualized on a FluorChem E FE0528 machine.

## Immunofluorescence and FISH

Immunolocalization was performed using a protocol modified from (Higgins et al. 2016). Meiotic anthers of Ab10 homozygous plants were fixed in fixative solution (4% paraformaldehyde, 1% Triton X-100 diluted in PHEMS (30mM PIPES, 12.5mM HEPES, 5mM EGTA, 1mM MgCl<sub>2</sub>, 175mM D-Sorbitol, pH6.8)) for one hour, then washed three times in PBS. Meiocytes were extruded from anthers and immobilized onto polylysine-coated coverslips by centrifuging at 100xG for one minute. The coverslips were incubated in permeabilization solution (PBS containing 1% Triton X-100 and 1mM EDTA) for one hour, then washed three

times in PBS. The coverslips were incubated in a blocking solution of 10% goat serum diluted in PBS for two hours and washed three times with PBS. Antibodies were diluted in an antibody dilution buffer (3% BSA diluted in PBS). The primary antibodies were mAb mouse  $\alpha$ -tubulin (Asai et al. 1982) diluted 1:200 and the pAb rabbit anti-KINDR (Dawe et al. 2018) diluted 1:100. The primary antibody solution was pipetted onto coverslips and left overnight at 37C, then coverslips were washed three times with PBS. Following an additional two hour blocking step, a secondary antibody solution (containing Rhodamine-conjugated AffiniPure Donkey Anti-Rabbit IgG H+L (Jackson cat no. 711-025-152) and Fluorescein-conjugated AffiniPure Goat Anti-Mouse IgG H+L (Jackson cat no. 115-095-146), both diluted 1:200) was pipetted onto coverslips and incubated at room temperature for two and a half hours. After three final washes with PBS, the coverslips were mounted on microscope slides using ProLong Gold with DAPI (Thermo Fisher cat no. P36931) and sealed with clear nail polish before imaging. Data showing KINDR localization specifically at knobs was collected from at least eight meiotic metaphase/anaphase cells from each of four different homozygous Ab10 plants.

Combined FISH-immunolocalization was carried out using a modified version of the suspended coverslip method described in (Yu et al. 1997). Fixed Ab10 meiocytes were incubated on coverslips for one hour in permeabilization solution. Coverslips were then suspended on a microscope slide using small broken bits of coverslips. A solution containing oligo probes (FITC-labeled 180bp repeat oligos and Rhodamine-labeled TR1 repeat oligos, see (Kanizay et al. 2013)) was pipetted under the coverslip and the edges were sealed using nail polish. The slides were heated at 95C for five minutes then incubated at room temperature for two hours in the dark. The coverslips were removed and washed, and immunolocalization performed as described above using the primary rabbit anti-KINDR antibody and secondary Cy5 AffiniPure Donkey

Anti-Rabbit IgG H+L (Jackson cat no. 711-175-152). Coverslips were mounted on slides using ProLong Gold with DAPI (Thermo Fisher cat no. P36931). Co-localization of KINDR with knobs containing 180 bp repeats (and not TR1 repeats) was observed in 10 metaphase/anaphase cells from a single plant showing strong neocentromere activity. Cells were imaged on a Zeiss Axio Imager.M1 fluorescence microscope with a 63 3 Plan-APO Chromat oil objective, and data analyzed using Slidebook software (Intelligent Imaging Innovations, Denver, CO, USA).

### Protein Expression and Purification

6xHis-tagged versions of KINDR and CENPC were generated by cloning these coding sequences into the pET28a vector. The KindrE9 coding sequence was synthesized and cloned into the pET28a vector by Genscript. Full-length CENPC was previously inserted into pET28a (Du et al. 2010). Both plasmids were transformed into One Shot™ BL21(DE3) Chemically Competent *E. coli* (Invitrogen, cat no. C600003). To express 6xHis-tagged versions of KindrE9 and CENPC, 10mL LB+Kanamycin (50 µg/ml) starter cultures of each transformed *E. coli* strain were incubated overnight at 37C. These were used to inoculate 500mL of LB+Kanamycin which were incubated for several hours at 37C until culture reached OD600 of 0.6. 1mM IPTG was added to this culture and it was incubated for 20 hours at 30C.

Cells were harvested by centrifuging at 4000 x g for 20 minutes at 4C and pellet was resuspended in 10mL lysis buffer (50mM NaH<sub>2</sub>PO<sub>4</sub>, 300mM NaCl, 10mM imidazole, pH 8.0). Ground powder of cComplete Mini Protease Inhibitor Cocktail (Roche cat no. 11836170001) was added periodically in all subsequent purification steps. Lysozyme (100K U/mg) in 50% glycerol was supplemented to 20ug/mL and lysate was incubated on ice for 30 minutes. Lysate was

sonicated using a probe sonicator at full-strength over ice with six 10 second pulses, each followed by a 10 second rest period. Sonicated lysate was centrifuged at 10,000 x g for 20 minutes at 4°C to collect cleared lysate. 2.5mL Ni-NTA Agarose slurry (Qiagen cat no. 30210) was added to cleared lysate and this mixture was incubated for 30 minutes at 4°C with gentle rotation to facilitate protein binding to Ni-NTA beads. The slurry/bead mixture was poured into a 5mL polypropylene column (Qiagen cat no. 34964) and the flow-thru was collected. Beads were washed twice with 5mL wash buffer (50mM NaH<sub>2</sub>PO<sub>4</sub>, 300mM NaCl, 20mM imidazole, pH 8.0). To elute, 5mL elution buffer (50mM NaH<sub>2</sub>PO<sub>4</sub>, 300mM NaCl, 250mM imidazole, pH 8.0) was applied to beads and elution was captured in tubes three drops at a time. One drop was tested in Bradford dye between tubes to determine which fraction contained the majority of protein, and ~500µl of elution was collected once a high protein concentration was found. Protein concentration of fractions with highest protein amounts were measured with the Qubit™ Protein Assay Kit (Invitrogen cat no. Q33212). Protein aliquots containing 1µg of protein each were flash frozen in liquid nitrogen then stored at -80°C. Protein purity and size were checked using Western blots and Coomassie-stained protein gels.

#### Probe Generation and EMSA

Biotinylated knob180 probes were generated by PCR using biotinylated primers to amplify knob DNA. A 50µl PCR reaction was performed with the Phusion High-Fidelity PCR Master Mix (Thermo Scientific cat no. F531S) and B73 genomic DNA diluted 1:250 with biotinylated primers at 1µM each. Manufacturer-recommended cycling conditions were used with an annealing temperature of 52°C, an extension time of 30 seconds, and 30 total cycles.

PCR products were run on a 1% agarose TAE gel and two bands at approximately 180 and 360bp were observed. The 180bp band was excised and gel purified using the Monarch® DNA Gel Extraction Kit (NEB cat no. T1020S) and DNA concentration was measured using the Qubit™ dsDNA HS Assay Kit (Invitrogen cat no. Q32851).

To test for protein-DNA interactions, we performed Electrophoretic Mobility Shift Assays (EMSAs) using the Gelshift™ Chemiluminescent EMSA kit (Active Motif cat no. 37341). Many conditions for binding reactions were tested by varying concentrations of multiple components but the following binding reaction conditions, which were similar to those previously published for an EMSA involved 6xHis-CENPC (Du et al. 2010), were ultimately used: 0.5mM DTT, 0.5mM EDTA, 1mM MgCl<sub>2</sub>, 150mM KCl. Biotinylated knob180 probes at 2fmol/μl were added to the binding reaction along with 1μg of protein or the equivalent volume of elution buffer for free probe controls. Binding reactions were incubated at room temperature for 20 minutes and samples were loaded on 5% Mini-PROTEAN TBE Precast Gel (Bio-Rad cat no. 456-5013) and samples run for about half an hour. DNA probe was transferred to a positively-charged nylon membrane (Roche cat no. 11209299001) in a 1X TBE solution for one hour, and subsequent blocking and chemiluminescent detection steps were performed according to the Active Motif kit instructions.

#### Resequencing Ab10 lines, Genomic Alignment, and SNP Calling

Genomic DNA was extracted from Ab10 and Ab10-*smd* variants with the Genomic Plant DNA Mini Kit (IBI Scientific cat no. IB47231) and Illumina libraries were prepared using the KAPA HyperPrep Kit (KAPA cat no. KK8502). Libraries were sequenced on an Illumina

NextSeq500 sequencer. DNA-seq reads from Ab10, *smd8*, *smd3*, and *smd13* were obtained previously and accessed from the SRA Database (PRJNA339461). For SNP calling, higher coverage Ab10 and *smd13* reads were used. Sequences of new reads obtained were also deposited into the SRA database under BioProject PRJNA339461.

Illumina reads were mapped to the Ab10 genomic assembly (Liu et al. 2020) using BWA MEM v0.7.15 (Li 2013). Only mapped reads on the Ab10 haplotype (chr10:122341516-195026473) were retained using -F 4 and -L options in SAMtools view v1.3.1 (Li et al. 2009). Reads were sorted with SAMtools sort and duplicate reads were marked using Picard (<http://broadinstitute.github.io/picard/>). Reads with a quality score of at least 20 were retained for further analysis. Read counts were obtained using the count function in 5000bp windows in IGVtools in Integrated Genome Viewer v2.6.3 and IGV was used to visualize read coverage (Robinson et al. 2011).

For *smd13* SNP calling, Illumina reads were compared between Ab10 and *smd13* genotypes. Reads from an Ab10 plant inbred into a B73 background were previously generated at high-coverage (SRA Accession: ERR3773575) for the Ab10 assembly project (Liu et al. 2020) and 223,521,050 paired-end 150 reads were randomly subsampled for SNP-calling analysis. The *smd13* library described above was sequenced again at higher coverage and 182,056,626 paired-end 150 reads were obtained.

The GATK Best Practices Pipeline for Germline short variant discovery was applied to call SNPs from Illumina data (DePristo et al. 2011; Poplin et al. 2017). Briefly, reads were mapped to the Ab10 reference using BWA MEM v0.7.15 (Li 2013), sorted using SAMtools sort v1.3.1 (Li et al. 2009), and duplicate reads were marked using Picard

(<http://broadinstitute.github.io/picard/>). SNPs on the Ab10 haplotype were called using HaplotypeCaller (GATK/4.0.11.0) using default settings and the coordinates chr10:140000000-195026473 followed by GenotypeGVCFs. SelectVariants was used to filter SNPs on all of the following criteria: homozygous REF in Ab10; homozygous ALT in *smd13*; *smd13* depth between 3-30; and *smd13* GQ of at least 20. This left 159 *smd13* SNPs that were retained for analysis.

### Analysis of Ab10 Repetitive Regions

The Ab10 distal tip was extracted and aligned to itself using minimap2 v2.13 (Li 2018) with the -cx asm5 option. Alignments of at least 1000bp were retained and an alignment dot-plot was generated using ggplot2 (Wickham 2016).

### RNA-seq, Differential Expression Analysis, and *de novo* Transcript Assembly

To analyze differential transcript expression in Ab10 and related lines, plants that had been back-crossed into B73 at least 6 times while selecting for purple kernels were used. We analyzed Ab10, Df(L), *smd13*, *smd8*, and B73 genotypes. Ear tissue was collected from plants when ears were ~2-3cm long and flash frozen in liquid nitrogen and stored at -80°C. Meiotic anthers from Ab10 in a mixed genetic background were independently collected and stored at -80°C. Three biological replicates per genotype were processed for RNA-seq analysis. Tissue was ground in liquid nitrogen in a mortar and pestle, and total RNA was extracted using the RNeasy Plant Mini Kit (Qiagen, cat no. 74904). RNA was flash frozen in liquid nitrogen and stored at -

80°C. Samples were shipped on dry ice and libraries were prepared by Novogene with the NEBNext Ultra II Direction RNA Library Prep Kit for Illumina (NEB cat no. E7760L). Libraries were sequenced on a Novaseq6000. Illumina read sequences were deposited in the SRA database under BioProject PRJNA339461.

For differential analysis, RNA-seq reads were trimmed using sickle v1.33 (Joshi and Fass 2011) and trimmed reads were aligned to the Ab10 reference (Liu et al. 2020) using STAR v2.7.1 (Dobin et al. 2013). Read counts were obtained using the htseq-count function from htseq v0.9.1 (Anders et al. 2015). Differential expression between Ab10 and each other genotype was measured in RStudio using DEseq v1.30.1 (Love et al. 2014). Genomic positions of transcripts were identified using the coordinates specified in the annotation file. Plots from differential expression data were generated using ggplot2 v3.3.3 (Wickham 2016).

A reference-free Ab10 *de novo* transcriptome was generated using Trinity v2.6.6 (Grabherr et al. 2011) with default settings. Three replicates each of stranded mRNA-seq reads from Ab10 ears and anthers were used to generate a Trinity transcriptome. The following Trinity accessory scripts were used to process transcripts and obtain differential expression:

abundance\_estimates\_to\_matrix.pl, count\_matrix\_features\_given\_MIN\_TPM\_threshold.pl, filter\_low\_expr\_transcripts.pl, run\_DE\_analysis\_from\_samples\_file.pl. The DESeq2 method was used for differential expression analysis. Trinity transcripts were aligned to Ab10-annotated transcripts using BLAST+ v2.7.1 (Camacho et al. 2009) to find genomic positions of Trinity transcripts.

After an individual transcript from the Trinity assembly, TRINITY\_DN45531\_c2\_g1\_i24, was found to encode a KINESIN-10-LIKE protein, we

attempted to identify the biologically-relevant portion of this sequence that was represented by the majority of RNA-seq reads. RNA-seq reads were mapped to the sequence of this transcript using BWA MEM v1.3.1 (Li 2013) to determine the region where reads tended to pile up. Reads were counted from the pileup region between 2318-3437bp and expression was calculated in RPKM (reads / (geneLength/1000 \* totalReads/1000000) ).

### Prediction of Kinesin-10-like motifs

Protein BLAST analysis of the Kinesin-10-like protein sequence predicted a putative helix-hairpin-helix domain spanning amino acids 41-81 (Marchler-Bauer et al. 2017). For independent prediction of these alpha helices, the Kinesin-10-like protein sequence was queried using the JPred4 server (Drozdetskiy et al. 2015) and two helices spanning residues 46-52 and 57-69 were predicted. Coiled-coil domains were predicted using COILS (Lupas et al. 1991). A NLS was searched for using NLS Mapper (Kosugi et al. 2009) but no NLS was predicted.

### Immunoprecipitation and LC-MS/MS Analysis

Rabbit anti-KINDR-conjugated beads were prepared prior to immunoprecipitation using the Crosslink Magnetic IP/Co-IP Kit (Pierce, cat no. 88805). 10µg of rabbit anti-KINDR antibody was conjugated to 50µl of Protein A/G Magnetic Beads and subsequently cross-linked with disuccinimidyl suberate (DSS) according to manufacturer's protocol. Immunoprecipitation from maize ears was performed by first grinding flash-frozen Ab10 ears in liquid nitrogen in a porcelain pestle and mortar. Ground tissue was suspended in 500ul of Plant IP Lysis Buffer

(50mM Tris, 150mM NaCl, 5mM EDTA, 5mM EGTA, 10% glycerol, 0.2% Triton X-100, 1mM PMSF, 1X Plant Protease Inhibitor Cocktail; Sigma cat no. P9599), gently mixed, and incubated on ice for ten minutes to extract total protein. Lysate was centrifuged at 15,000 x G for ten minutes at 4°C, the supernatant was extracted, and this centrifugation step was repeated to remove all solids from lysate. Magnetic beads conjugated to the anti-KINDR antibody were resuspended in cleared lysate and incubated for one hour at room temperature with gentle rotation. Beads were collected using a magnetic stand and we performed four washes with IP Lysis/Wash Buffer (pH 7.4, 25mM Tris, 150mM NaCl, 1mM EDTA, 1% NP40, 5% glycerol), changing tubes during the last wash. One final wash was performed using 500µl water and two subsequent elutions were performed each using 20µl of Elution Buffer (from Crosslink Magnetic IP/Co-IP Kit) for five minutes at room temperature, each. Following elution, 5µl of neutralization buffer (from Crosslink Magnetic IP/Co-IP Kit) was added and this solution was mixed with 50µl 2X Laemmli buffer. Equal 1% fractions of input, flow-through, and IP were run on a polyacrylamide gel and a Western blot probed with rabbit anti-KINDR was performed for analysis.

Protein elution mixture was briefly run on a 12% mini-PROTEAN TGX Precast Gel (Bio-Rad Cat# 4561043) and gel slices were excised with a razor blade. Protein gel slices were stained with QC Colloidal Coomassie Stain (Bio-Rad cat no. 1610803) according to manufacturer's recommendations. Gel slices were sent to the University of Georgia Proteomics and Mass Spectrometry facility for protein identification. Briefly, a tryptic digest was performed on these gel slices and peptides were extracted and dried before identification with a 60-minute LC-MS/MS run. Tryptic fragment sequences were searched against the Ab10 protein annotation database (Liu et al. 2020) using the Mascot Server software (Matrix Science).

## REFERENCES

- Afshar K, Barton NR, Hawley RS, Goldstein LS. 1995. DNA binding and meiotic chromosomal localization of the *Drosophila* nod kinesin-like protein. *Cell* **81**: 129–138.
- Almeida AC, Maiato H. 2018. Chromokinesins. *Curr Biol* **28**: R1131–R1135.
- Ananiev EV, Phillips RL, Rines HW. 1998. A knob-associated tandem repeat in maize capable of forming fold-back DNA segments: are chromosome knobs megatransposons? *Proc Natl Acad Sci U S A* **95**: 10785–10790.
- Anders S, Pyl PT, Huber W. 2015. HTSeq--a Python framework to work with high-throughput sequencing data. *Bioinformatics* **31**: 166–169.
- Antonio C, Ferby I, Wilhelm H, Jones M, Karsenti E, Nebreda AR, Vernos I. 2000. Xkid, a chromokinesin required for chromosome alignment on the metaphase plate. *Cell* **102**: 425–435.
- Asai DJ, Thompson WC, Wilson L, Brokaw CJ. 1982. Two different monoclonal antibodies to alpha-tubulin inhibit the bending of reactivated sea urchin spermatozoa. *Cell Motil* **2**: 599–614.
- Brouwers N, Mallol Martinez N, Vernos I. 2017. Role of Kif15 and its novel mitotic partner KBP in K-fiber dynamics and chromosome alignment. *PLoS One* **12**: e0174819.
- Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL. 2009. BLAST+: architecture and applications. *BMC Bioinformatics* **10**: 421.
- Cheeseman IM, Chappie JS, Wilson-Kubalek EM, Desai A. 2006. The Conserved KMN Network Constitutes the Core Microtubule-Binding Site of the Kinetochore. *Cell* **127**: 983–997.
- Dawe RK, Cande WZ. 1996. Induction of centromeric activity in maize by suppressor of meiotic drive 1. *Proc Natl Acad Sci U S A* **93**: 8512–8517.
- Dawe RK, Lowry EG, Gent JI, Stitzer MC, Swentowsky KW, Higgins DM, Ross-Ibarra J, Wallace JG, Kanizay LB, Alabady M, et al. 2018. A Kinesin-14 Motor Activates Neocentromeres to Promote Meiotic Drive in Maize. *Cell* **173**: 839–850.e18.
- DePristo MA, Banks E, Poplin R, Garimella KV, Maguire JR, Hartl C, Philippakis AA, del Angel G, Rivas MA, Hanna M, et al. 2011. A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat Genet* **43**: 491–498.
- Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR. 2013. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**: 15–21.
- Drozdetskiy A, Cole C, Procter J, Barton GJ. 2015. JPred4: a protein secondary structure prediction server. *Nucleic Acids Res* **43**: W389–94.
- Du Y, Topp CN, Dawe RK. 2010. DNA binding of centromere protein C (CENPC) is stabilized by single-stranded RNA. *PLoS Genet* **6**: e1000835.
- Funabiki H, Murray AW. 2000. The *Xenopus* chromokinesin Xkid is essential for metaphase chromosome alignment and must be degraded to allow anaphase chromosome movement. *Cell* **102**: 411–424.
- Gallaher SD, Craig RJ, Ganesan I, Purvine SO, McCorkle SR, Grimwood J, Strenkert D, Davidi L, Roth MS, Jeffers TL, et al. 2021. Widespread polycistronic gene expression in green algae. *Proc Natl Acad Sci U S A* **118**. <https://www.pnas.org/content/118/7/e2017714118> (Accessed March 31, 2021).

- Gicking AM, Swentowsky KW, Dawe RK, Qiu W. 2018. Functional diversification of the kinesin-14 family in land plants. *FEBS Lett* **592**: 1918–1928.
- Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng Q, et al. 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat Biotechnol* **29**: 644–652.
- Hiatt EN, Dawe RK. 2003a. Four loci on abnormal chromosome 10 contribute to meiotic drive in maize. *Genetics* **164**: 699–709.
- Hiatt EN, Dawe RK. 2003b. The meiotic drive system on maize abnormal chromosome 10 contains few essential genes. *Genetica* **117**: 67–76.
- Hiatt EN, Kentner EK, Dawe RK. 2002. Independently regulated neocentromere activity of two classes of tandem repeat arrays. *Plant Cell* **14**: 407–420.
- Higgins DM. 2017. Characterization of Kinesin Motor Proteins and Evolution of a Meiotic Drive Element in *Zea mays*. ed. R.K. Dawe., Doctor of Philosophy, University of Georgia.
- Higgins DM, Nannas NJ, Dawe RK. 2016. The Maize Divergent spindle-1 (dv1) Gene Encodes a Kinesin-14A Motor Protein Required for Meiotic Spindle Pole Organization. *Front Plant Sci* **7**: 1277.
- Jia H, Li M, Li W, Liu L, Jian Y, Yang Z, Shen X, Ning Q, Du Y, Zhao R, et al. 2020. A serine/threonine protein kinase encoding gene KERNEL NUMBER PER ROW6 regulates maize grain yield. *Nat Commun* **11**: 988.
- Joshi NA, Fass JN. 2011. *Sickle: A sliding-window, adaptive, quality-based trimming tool for FastQ files*. <https://github.com/najoshi/sickle>.
- Kanizay LB, Albert PS, Birchler JA, Dawe RK. 2013. Intragenomic conflict between the two major knob repeats of maize. *Genetics* **194**: 81–89.
- Kikudome GY. 1959. Studies on the Phenomenon of Preferential Segregation in Maize. *Genetics* **44**: 815–831.
- Koenig T, Menze BH, Kirchner M, Monigatti F, Parker KC, Patterson T, Steen JJ, Hamprecht FA, Steen H. 2008. Robust prediction of the MASCOT score for an improved quality assessment in mass spectrometric proteomics. *J Proteome Res* **7**: 3708–3717.
- Kosugi S, Hasebe M, Tomita M, Yanagawa H. 2009. Systematic identification of cell cycle-dependent yeast nucleocytoplasmic shuttling proteins by prediction of composite motifs. *Proc Natl Acad Sci U S A* **106**: 10171–10176.
- Lee YM, Lee S, Lee E, Shin H, Hahn H, Choi W, Kim W. 2001. Human kinesin superfamily member 4 is dominantly localized in the nuclear matrix and is associated with chromosomes during mitosis. *Biochem J* **360**: 549–556.
- Li H. 2013. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv [q-bioGN]*. <http://arxiv.org/abs/1303.3997>.
- Li H. 2018. Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics* **34**: 3094–3100.
- Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, 1000 Genome Project Data Processing Subgroup. 2009. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**: 2078–2079.
- Liu J, Seetharam AS, Chougule K, Ou S, Swentowsky KW, Gent JI, Llaca V, Woodhouse MR, Manchanda N, Presting GG, et al. 2020. Gapless assembly of maize chromosomes using long-read technologies. *Genome Biol* **21**: 121.
- Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**: 550.

- Lupas A, Van Dyke M, Stock J. 1991. Predicting coiled coils from protein sequences. *Science* **252**: 1162–1164.
- Marchler-Bauer A, Bo Y, Han L, He J, Lanczycki CJ, Lu S, Chitsaz F, Derbyshire MK, Geer RC, Gonzales NR, et al. 2017. CDD/SPARCLE: functional classification of proteins via subfamily domain architectures. *Nucleic Acids Res* **45**: D200–D203.
- Mogessie B, Schuh M. 2017. Actin protects mammalian eggs against chromosome segregation errors. *Science* **357**. <http://dx.doi.org/10.1126/science.aal1647>.
- Mroczek RJ, Melo JR, Luce AC, Hiatt EN, Dawe RK. 2006. The maize Ab10 meiotic drive system maps to supernumerary sequences in a large complex haplotype. *Genetics* **174**: 145–154.
- Peacock WJ, Dennis ES, Rhoades MM, Pryor AJ. 1981. Highly repeated DNA sequence limited to knob heterochromatin in maize. *Proc Natl Acad Sci U S A* **78**: 4490–4494.
- Perkins DN, Pappin DJ, Creasy DM, Cottrell JS. 1999. Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis* **20**: 3551–3567.
- Poplin R, Ruano-Rubio V, DePristo MA, Fennell TJ, Carneiro MO, Van der Auwera GA, Kling DE, Gauthier LD, Levy-Moonshine A, Roazen D, et al. 2017. Scaling accurate genetic variant discovery to tens of thousands of samples. *Cold Spring Harbor Laboratory* 201178. <https://www.biorxiv.org/content/10.1101/201178v2.abstract> (Accessed January 18, 2021).
- Rhoades MM. 1942. Preferential Segregation in Maize. *Genetics* **27**: 395–407.
- Rhoades MM, Dempsey E. 1966. The Effect of Abnormal Chromosome 10 on Preferential Segregation and Crossing over in Maize. *Genetics* **53**: 989–1020.
- Rhoades MM, Others. 1952. Preferential segregation in maize. *Preferential segregation in maize*. <https://www.cabdirect.org/cabdirect/abstract/19531603828>.
- Rhoades MM, Vilkomerson H. 1942. On the Anaphase Movement of Chromosomes. *Proc Natl Acad Sci U S A* **28**: 433–436.
- Robinson JT, Thorvaldsdóttir H, Winckler W, Guttman M, Lander ES, Getz G, Mesirov JP. 2011. Integrative genomics viewer. *Nat Biotechnol* **29**: 24–26.
- Shao X, Grishin NV. 2000. Common fold in helix-hairpin-helix proteins. *Nucleic Acids Res* **28**: 2643–2650.
- Shi J, Dawe RK. 2006. Partitioning of the maize epigenome by the number of methyl groups on histone H3 lysines 9 and 27. *Genetics* **173**: 1571–1583.
- Staiger CJ, Cande WZ. 1991. Microfilament Distribution in Maize Meiotic Mutants Correlates with Microtubule Organization. *Plant Cell* **3**: 637–644.
- Swentowsky KW, Gent JI, Lowry EG, Schubert V, Ran X, Tseng K-F, Harkess AE, Qiu W, Dawe RK. 2020. Distinct kinesin motors drive two types of maize neocentromeres. *Genes Dev* **34**: 1239–1251.
- Tørresen OK, Star B, Mier P, Andrade-Navarro MA, Bateman A, Jarnot P, Gruca A, Grynberg M, Kajava AV, Promponas VJ, et al. 2019. Tandem repeats lead to sequence assembly errors and impose multi-level challenges for genome and protein databases. *Nucleic Acids Res* **47**: 10994–11006.
- Uraji J, Scheffler K, Schuh M. 2018. Functions of actin in mouse oocytes at a glance. *J Cell Sci* **131**. <http://dx.doi.org/10.1242/jcs.218099>.
- Wang N, Dawe RK. 2018. Centromere Size and Its Relationship to Haploid Formation in Plants. *Mol Plant* **11**: 398–406.

- Wang K, Wang D, Zheng X, Qin A, Zhou J, Guo B, Chen Y, Wen X, Ye W, Zhou Y, Zhu Y. 2019. Multi-strategic RNA-seq analysis reveals a high-resolution transcriptional landscape in cotton. *Nat Commun* **10**: 4714.
- Wickham H. 2016. ggplot2: Elegant Graphics for Data Analysis. <https://ggplot2.tidyverse.org>.
- Xu J, Wang X-Y, Guo W-Z. 2015. The cytochrome P450 superfamily: Key players in plant development and defense. *J Integr Agric* **14**: 1673–1686.
- Yi K, Unruh JR, Deng M, Slaughter BD, Rubinstein B, Li R. 2011. Dynamic maintenance of asymmetric meiotic spindle position through Arp2/3-complex-driven cytoplasmic streaming in mouse oocytes. *Nat Cell Biol* **13**: 1252–1258.
- Yu HG, Hiatt EN, Chan A, Sweeney M, Dawe RK. 1997. Neocentromere-mediated chromosome movement in maize. *J Cell Biol* **139**: 831–840.

**Table 2.1. Annotation-based differentially expressed *smd13* genes.**

Name	<i>smd</i> <b>8</b> DE	Position (Mb)	L2FC	padj	BLAST hits	%ID	Predicted Domains
Zm00043a049636_T001	N	159.8698	-1.419	9.73E-02	Thioesterase superfamily protein	100	thioesterase
Zm00043a049920_T001	N	163.8076	-0.777	7.99E-01	uncharacterized protein	100%	none
Zm00043a049987_T001	Y	164.6951	-2.062	1.30E-01	Tyrosine-sulfated glycopeptide receptor 1	87.62	LRR-RLK
Zm00043a050066_T001	Y	165.6760	2.4061	4.62E-03	hypothetical protein	98.83	none
Zm00043a050167_T001	N	167.2489	1.0994	1.28E-01	hypothetical protein	95.24	DUF3774
Zm00043a050240_T002	Y	168.6378	-3.394	1.40E-01	CLP protease regulatory subunit; mRNA-decapping enzyme subunit 2	99.06	DCP2; Nudix-Hydrolase
Zm00043a050491_T001	Y	179.9972	-2.317	4.64E-04	none		none
Zm00043a050667_T001	N	185.7112	-2.924	2.25E-01	C2 and GRAM domain-containing protein	97.36	PH
Zm00043a051039_T001	N	194.1887	-0.816	6.90E-01	Sterile alpha motif (SAM) domain-containing protein	97.67	none

**Table 2.2. Proteins identified from KINDR co-IP/Mass Spectrometry.**

Protein Name	BLAST hit	Mascot Score				n
		Ear 1	Ear 2	Ear 3	Root Tip 1	
Zm00043a050837_P001	KINDR B9	333	303	543	354	4
Zm00043a015101_P001	elongation factor 1-alpha	47	38	63	82	4
Zm00043a000006_P001	none	45	NA	34	51	3
Zm00043a006404_P001	putative actin family protein	NA	80	66	46	3
Zm00043a033133_P001	Auxin-responsive protein SAUR50	NA	41	34	36	3
Zm00043a035792_P001	60S ribosomal protein L9-like	48	NA	61	77	3
Zm00043a053384_P001	histone-lysine N-methyltransferase EZ2	35	NA	40	32	3
Zm00043a000231_P001	cyclin-T1-3	38	NA	39	NA	2
Zm00043a000251_P001	histone H4	NA	NA	35	34	2
Zm00043a003663_P001	Cinfull1 polyprotein	34	39	NA	NA	2
Zm00043a006610_P001	uncharacterized protein	29	NA	30	NA	2
Zm00043a018465_P001	40S ribosomal protein S14	31	NA	43	NA	2

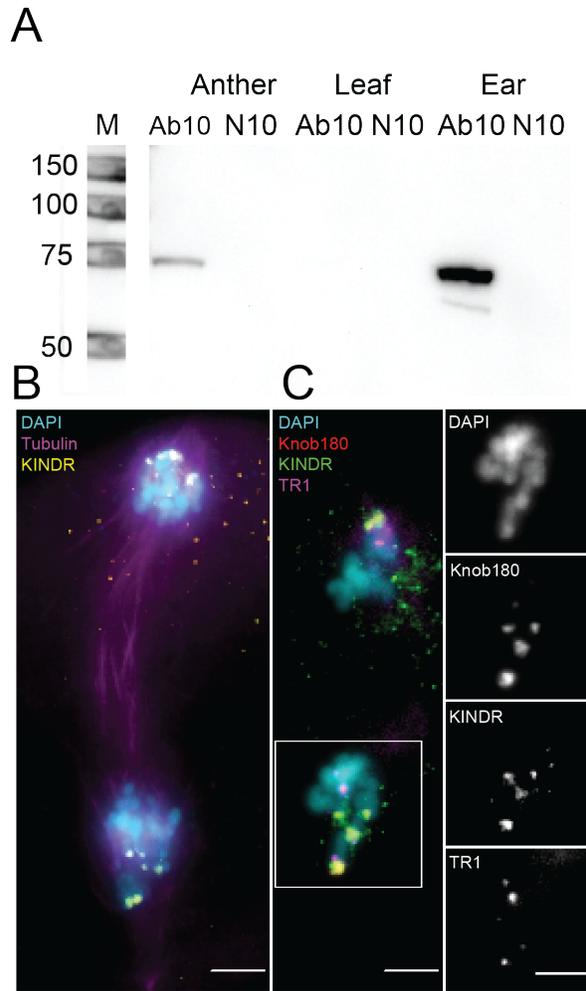
Zm00043a020659_P001	26S protease regulatory subunit 8 homolog B	40	NA	39	NA	2
Zm00043a023707_P001	Alba DNA/RNA-binding protein	28	NA	99	NA	2
Zm00043a031619_P001	THO complex subunit 4C	NA	44	35	NA	2
Zm00043a032879_P001	Replication factor C subunit 3	40	NA	49	NA	2
Zm00043a038321_P001	uncharacterized protein	NA	NA	39	32	2
Zm00043a041596_P001	salt stress-induced protein	90	NA	71	NA	2
Zm00043a043660_P001	Protein NUCLEAR FUSION DEFECTIVE 4	30	NA	43	NA	2
Zm00043a045251_P001	60S ribosomal protein L28-like	46	NA	NA	138	2
Zm00043a000508_P001	60S ribosomal protein L23-like	NA	NA	NA	31	1
Zm00043a000531_P001	60S ribosomal protein L21-like	NA	NA	NA	46	1
Zm00043a000743_P001	histone H2A variant 3	NA	32	NA	NA	1
Zm00043a000844_P001	homeobox-leucine zipper protein HAT7	NA	NA	31	NA	1
Zm00043a000974_P001	Aspartyl protease 25	NA	NA	NA	102	1
Zm00043a001773_P001	heat shock cognate 70 kDa protein 2	NA	NA	NA	53	1
Zm00043a002360_P004	40S ribosomal protein S21	NA	NA	NA	37	1
Zm00043a002473_P001	phospholipase A1-II 7	NA	NA	32	NA	1

Zm00043a002721_P001	FCS-Like Zinc finger 13	NA	NA	NA	34	1
Zm00043a002826_P001	40S ribosomal protein S19	NA	NA	NA	61	1
Zm00043a003062_P001	probable lysophospholipase BODYGUARD 1	35	NA	NA	NA	1
Zm00043a003283_P001	expansin-B4 precursor	NA	NA	NA	40	1
Zm00043a004255_P001	60S ribosomal protein L8	NA	NA	NA	35	1
Zm00043a004785_P001	5- methyltetrahydropteroyltriglutamate-- homocysteine methyltransferase 2	NA	NA	35	NA	1
Zm00043a005255_P001	60S ribosomal protein L32-1	NA	NA	NA	54	1
Zm00043a005298_P001	60S ribosomal protein L17	NA	NA	NA	81	1
Zm00043a005832_P001	60S ribosomal protein L26-1	NA	NA	NA	65	1
Zm00043a006354_P001	heat shock 70 kDa protein	NA	NA	NA	59	1
Zm00043a007207_P001	60S ribosomal protein L10-3	NA	NA	NA	36	1
Zm00043a007735_P001	Tetratricopeptide repeat (TPR)-like superfamily protein	NA	NA	39	NA	1
Zm00043a007829_P001	pinin/SDK/memA/protein conserved region containing protein	NA	NA	NA	34	1
Zm00043a008240_P001	exoglucanase precursor	NA	NA	NA	119	1
Zm00043a008599_P001	40S ribosomal protein S18	NA	NA	NA	54	1

Zm00043a008667_P001	histone H1	NA	NA	NA	121	1
Zm00043a008814_P001	40S ribosomal protein S26	NA	NA	NA	46	1
Zm00043a009020_P001	40S ribosomal protein S4-like	NA	NA	NA	34	1
Zm00043a009896_P001	pentatricopeptide repeat-containing protein	NA	NA	30	NA	1
Zm00043a010030_P001	60S ribosomal protein L12	NA	NA	NA	48	1
Zm00043a010919_P001	60S ribosomal protein L23A	NA	NA	NA	91	1
Zm00043a012038_P001	60S ribosomal protein L37a	NA	NA	NA	105	1
Zm00043a012298_P001	dirigent	NA	NA	79	NA	1
Zm00043a012299_P001	uncharacterized protein	NA	NA	NA	44	1
Zm00043a014733_P001	60S ribosomal protein L14	NA	NA	NA	45	1
Zm00043a015345_P001	60S ribosomal protein L27a-3	NA	NA	NA	46	1
Zm00043a016302_P001	60S ribosomal protein L4-like	NA	NA	NA	34	1
Zm00043a019930_P001	putative UPF0481 protein	31	NA	NA	NA	1
Zm00043a021262_P001	60S ribosomal protein L6	NA	NA	NA	149	1
Zm00043a021323_P001	myosin-9	NA	29	NA	NA	1
Zm00043a022570_P001	none	NA	NA	37	NA	1

Zm00043a025071_P001	ATP synthase subunit alpha, mitochondrial	NA	NA	NA	86	1
Zm00043a025479_P001	4-hydroxy-7-methoxy-3-oxo-3,4-dihydro-2H-1,4-benzoxazin-2-yl glucoside beta-D-glucosidase 1, chloroplastic	NA	NA	NA	39	1
Zm00043a026652_P001	beclin 1 protein	NA	NA	31	NA	1
Zm00043a026657_P001	60S ribosomal protein L13	NA	NA	NA	69	1
Zm00043a026806_P001	fructose-bisphosphate aldolase, cytoplasmic isozyme	NA	NA	NA	94	1
Zm00043a028482_P001	ATP synthase subunit beta, mitochondrial-like	NA	NA	NA	34	1
Zm00043a029227_P001	uncharacterized protein	NA	NA	NA	48	1
Zm00043a030647_P001	putative oxysterol binding domain family protein	NA	NA	35	NA	1
Zm00043a030792_P001	40S ribosomal protein S13	NA	NA	NA	67	1
Zm00043a031114_P001	protein-L-isoaspartate O-methyltransferase	NA	NA	NA	135	1
Zm00043a032245_P001	60S ribosomal protein L6-like	NA	NA	NA	124	1
Zm00043a036895_P001	pentatricopeptide repeat-containing protein	NA	NA	NA	34	1
Zm00043a037233_P001	60S ribosomal protein L28-like	NA	NA	NA	134	1

Zm00043a037782_P001	P-loop containing nucleoside triphosphate hydrolase superfamily protein	35	NA	NA	NA	1
Zm00043a041610_P001	E3 UFM1-protein ligase 1 homolog	NA	45	NA	NA	1
Zm00043a041919_P001	60S ribosomal protein L39	NA	NA	33	NA	1
Zm00043a042631_P001	40S ribosomal protein S14-like	NA	NA	NA	191	1
Zm00043a043946_P001	mitochondrial 2-oxoglutarate/malate carrier protein	NA	NA	NA	31	1
Zm00043a046510_P001	4-hydroxy-7-methoxy-3-oxo-3,4-dihydro-2H-1,4-benzoxazin-2-yl glucoside beta-D-glucosidase 1, chloroplastic	NA	NA	NA	29	1
Zm00043a050813_P001	KINDR C4	59	NA	NA	NA	1
Zm00043a052844_P001	phosphoenolpyruvate carboxylase 3	NA	NA	47	NA	1
Zm00043a053710_P001	glucose-6-phosphate isomerase	NA	NA	32	NA	1
Zm00043a055653_P001	60S ribosomal protein L22-2	NA	NA	NA	106	1
Zm00043a055704_P001	Ricin B-like lectin R40G2	NA	NA	NA	137	1
Zm00043a057712_P001	RING/FYVE/PHD-type zinc finger family protein	30	NA	NA	NA	1

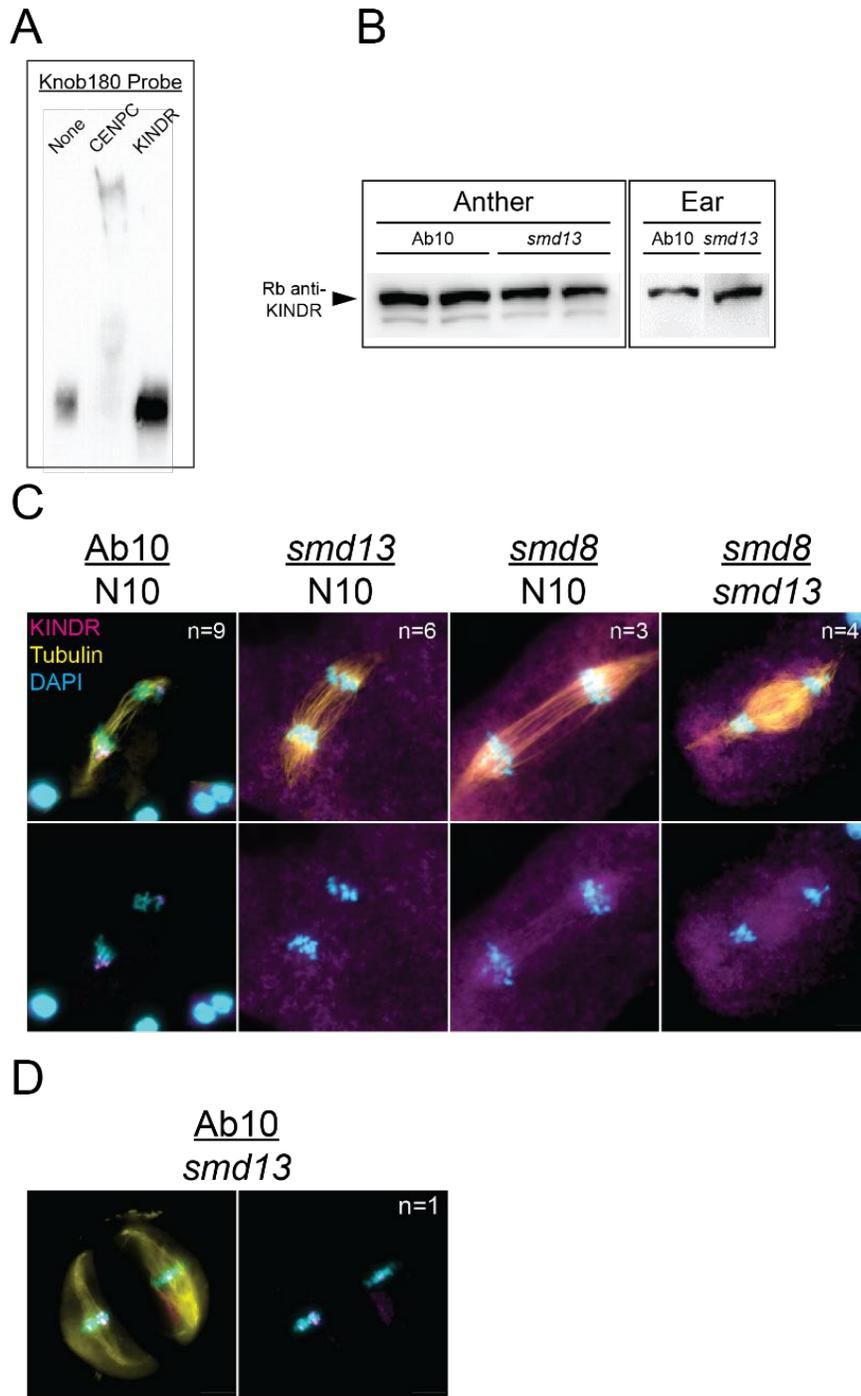


**Figure 2.1. KINDR is expressed in meiotic tissues and localizes to knob180 knobs.**

(A) Protein blot analysis of KINDR expression in three tissues. A strong band is observed at ~74 kDa (predicted size is 69 kDa). A white bar has been added to block out the right half of the marker lanes, which were highlighted by hand using a ChemiPen.

(B) Localization of KINDR to neocentromeres at meiosis I. The upper spindle pole is pointed toward the viewer; in this projection image the pole appears rounded.

(C) Co-localization of KINDR and knob repeats at meiosis II. KINDR colocalizes with 180 bp knob repeats but not with TR1 repeats. The panels to the right show each channel separately.



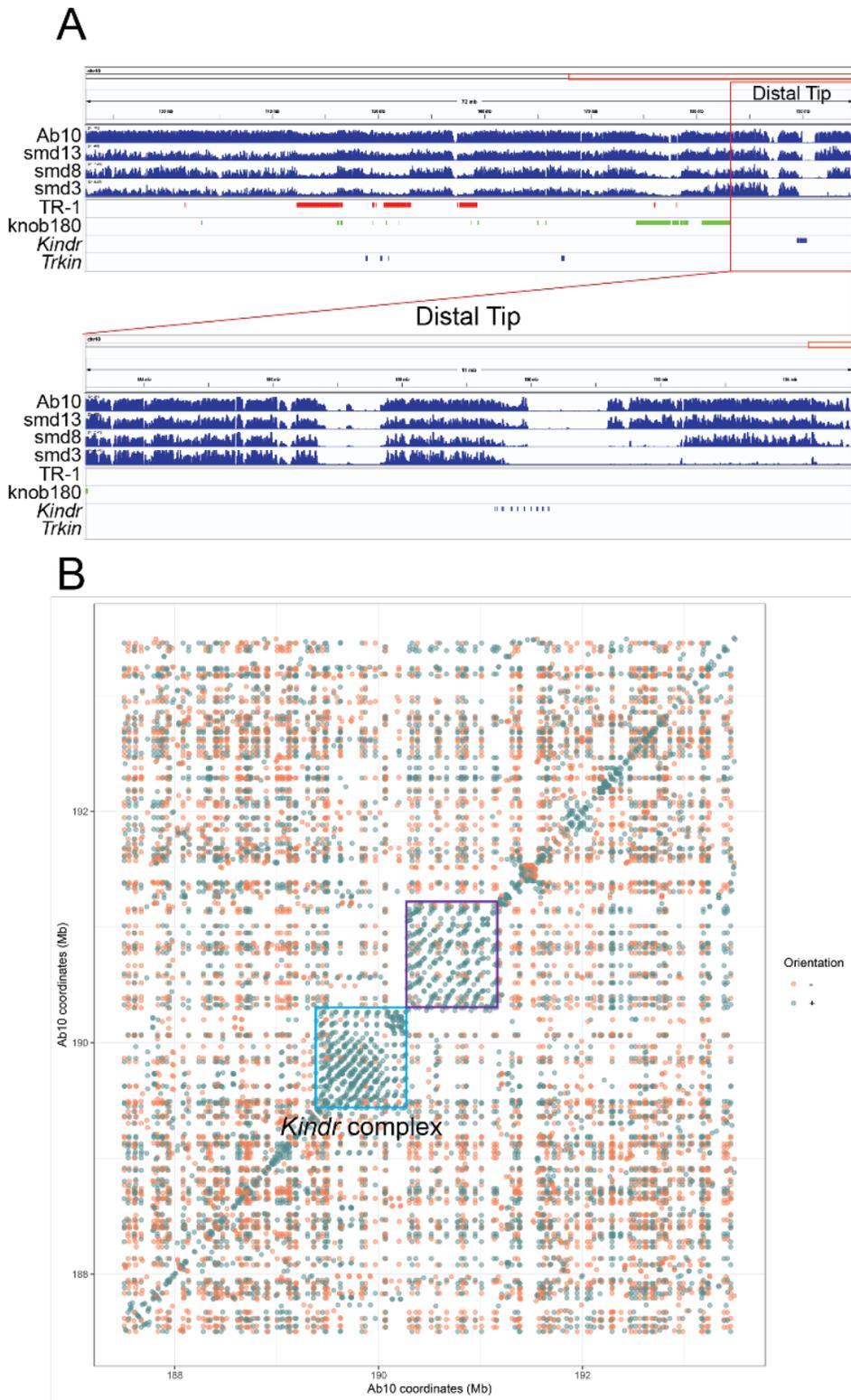
**Figure 2.2. KINDR does not interact with knob180 DNA and characterization of *smd13*.**

(A) EMSA using biotinylated-knob180 DNA as a probe. Probe was incubated with no protein (left), 6xHis-CENPC (middle), or 6xHis-KINDR (right).

(B) Protein blot analysis of KINDR expression in Ab10 and smd13 anthers and ears. Two replicates each are shown for anthers, and one replicate for ears.

(C) Immunolocalization of KINDR and tubulin in male meiocytes of Ab10 and related genotypes. Genotypes are indicated at top and at least ten cells per biological replicate were observed. The number of biological replicates scored are indicated in the top right each panel and representative images are shown. Top panel set shows composite images of KINDR, tubulin, and DAPI channels; bottom panel set shows composite images of KINDR and DAPI channels.

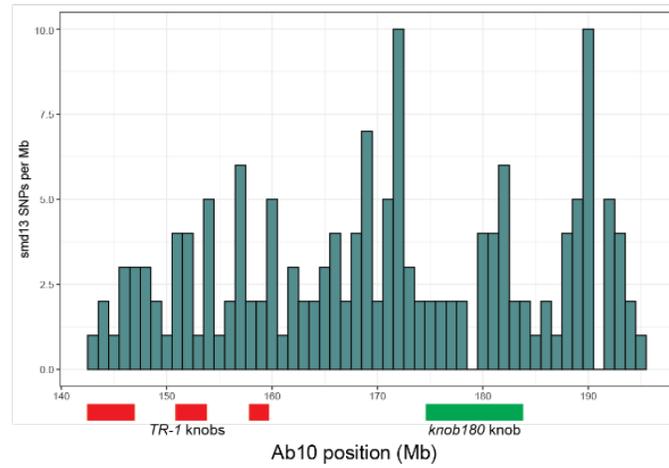
(D) Immunolocalization of KINDR and tubulin in male meiocytes in an Ab10/smd13 heterozygous genotype. Ten cells were observed with this localization pattern in one biological replicate, and a representative image is shown. Left panel shows composite image of KINDR, tubulin, and DAPI channels; right panel set shows composite images of KINDR and DAPI channels.



**Figure 2.3. DNA-seq coverage along Ab10 in Ab10 and *smd* genotypes and repeat analysis of Ab10 distal tip.**

(A) Coverage of Illumina DNA-seq reads generated from Ab10, *smd13*, *smd8*, and *smd3* genotypes aligned to the Ab10 haplotype. Entire haplotype is shown (top panel), and the Ab10 distal tip is expanded in the bottom panel.

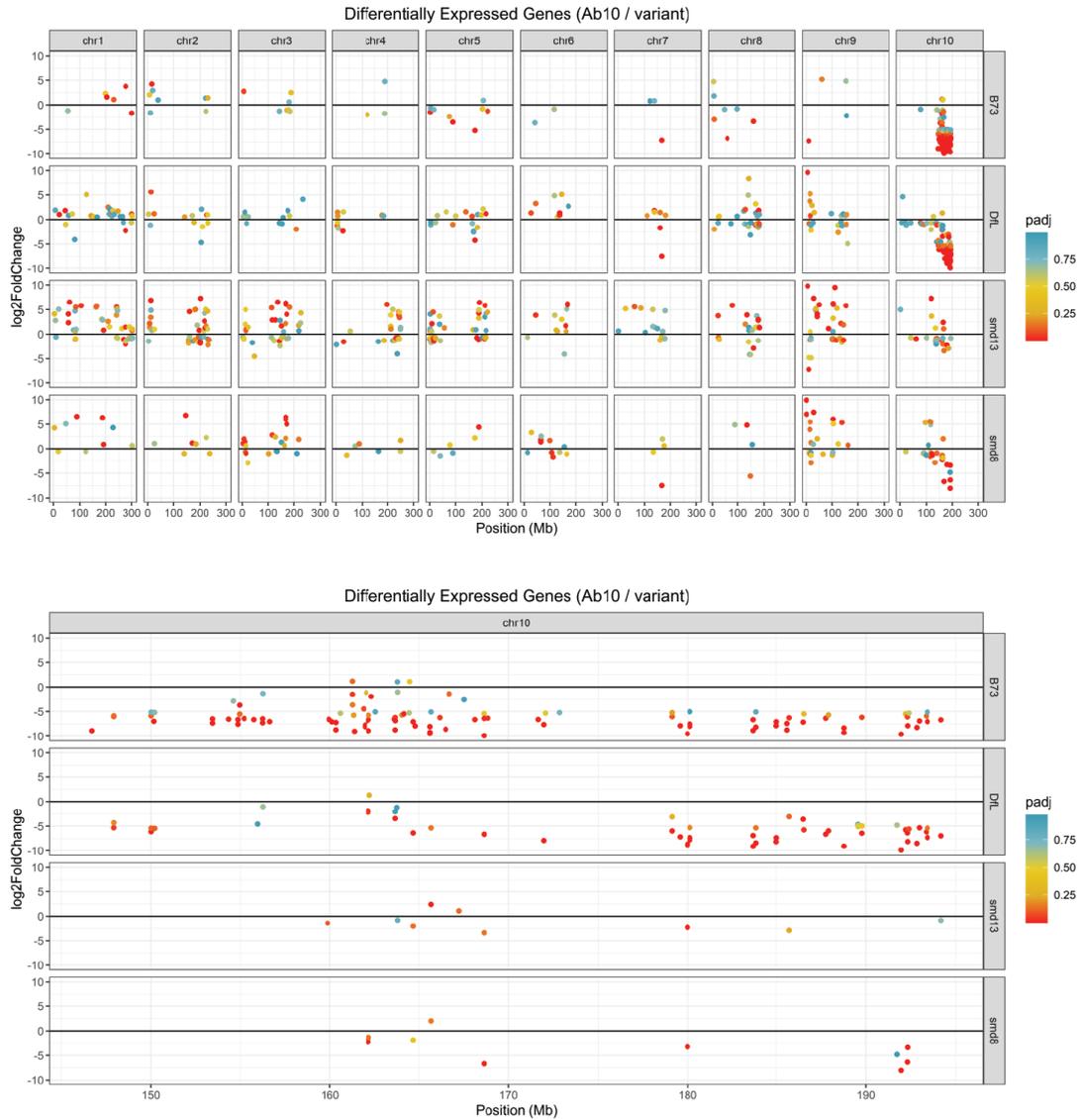
(B) Dot-plot of Ab10 distal tip showing aligned sequences when Ab10 was aligned to itself. Same-strand (blue) and opposite-strand (orange) alignments are displayed as dots. The *Kindr* complex (blue square) and a newly-described (purple) tandemly-repeated block are shown near the center of the distal tip.

**A****B**

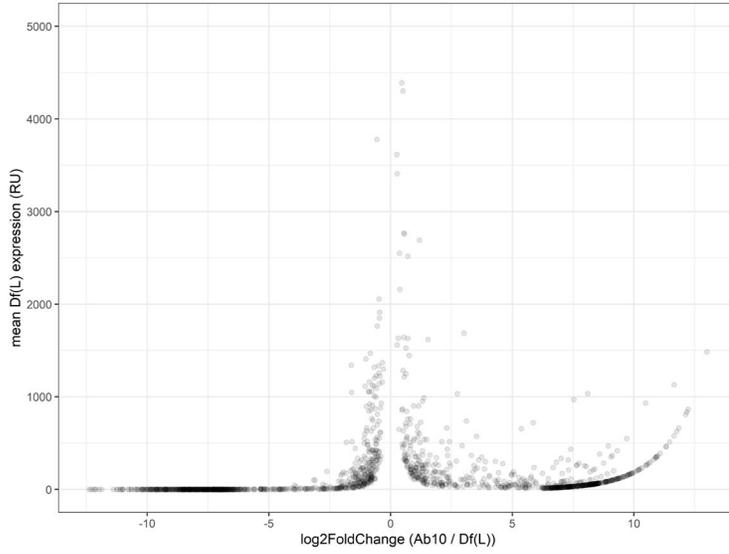
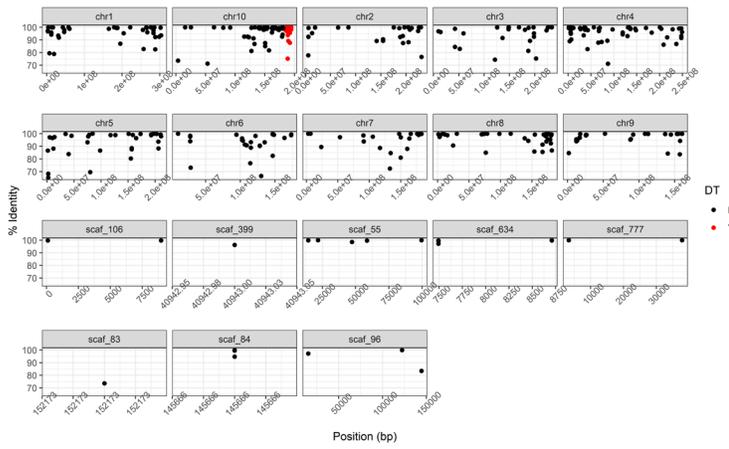
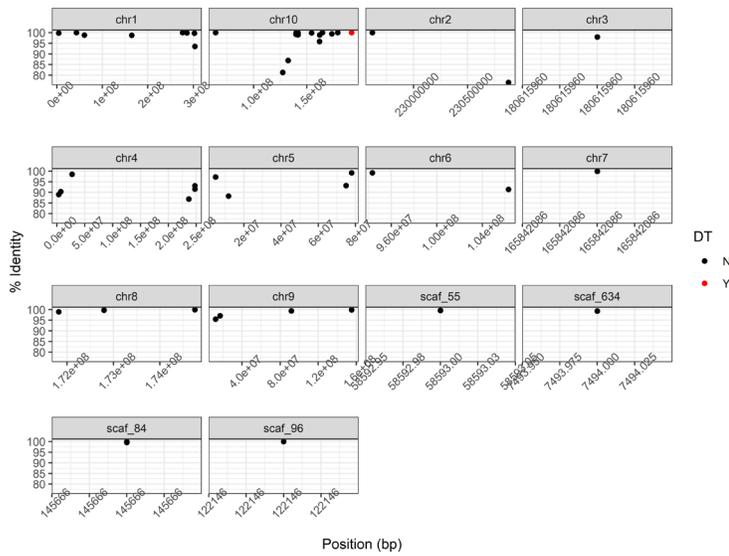
**Figure 2.4. Ab10 SNPs in *smd13*.**

(A) SNP density of *smd13* SNPs on the Ab10 haplotype.

(B) *smd13* SNPs (top) and aligned reads (bottom) near *KindrF6*. Colored bars indicate SNPs relative to the Ab10 reference.



**Figure 2.5. Differentially-expressed Ab10 annotated genes.** Fold-change of each gene relative to Ab10 is shown on the y-axis in four genotypes: B73, Df(L), smd13, and smd8. Genes are plotted on each chromosome, and genomic position is displayed on the x-axis. Adjusted p-value is shown by the color of the dot. Gene expression is shown for all genomic transcripts (top) and the Ab10 haplotype (bottom) is expanded for detail.

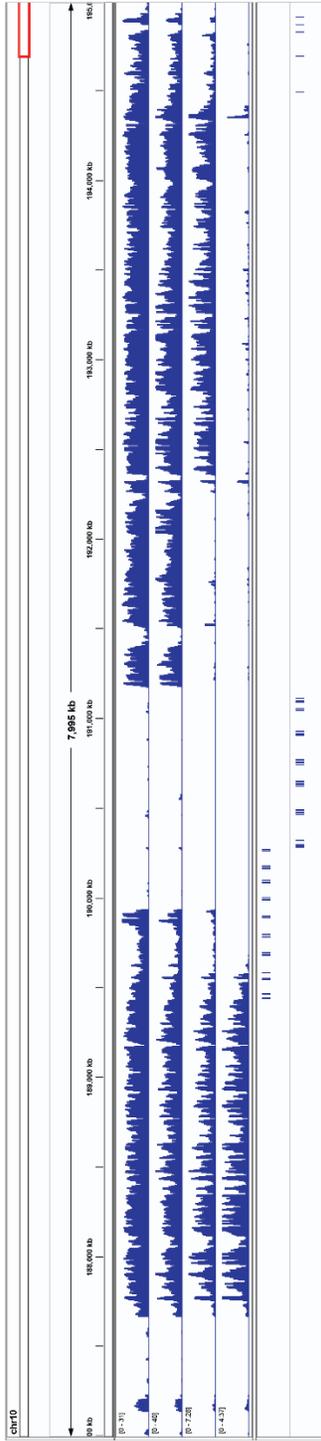
**A****B****C**

**Figure 2.6. RNA-seq analysis of Ab10 using Trinity transcripts and *Kindr* and *Kinesin-10-like* gene copy locations.**

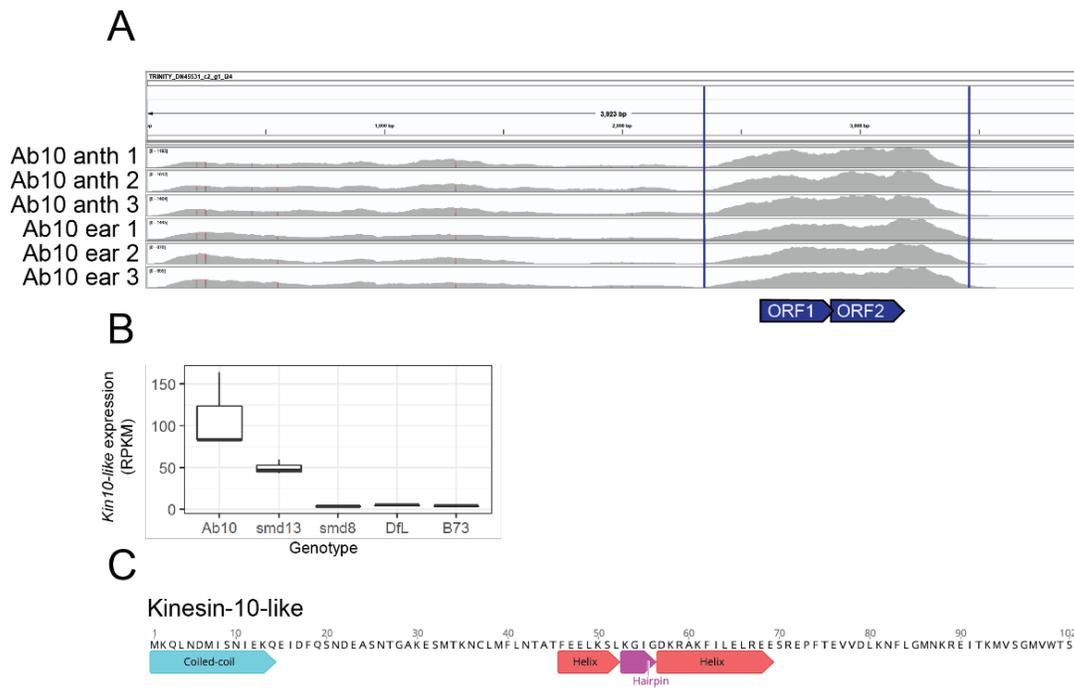
(A) Dot-plot comparing log fold change expression between Ab10 and Df(L) and mean expression of transcript in Df(L).

(B) Plot comparing genomic position with BLAST percent identity among transcripts differentially expressed in Df(L).

(C) Plot comparing genomic position with BLAST percent identity among transcripts differentially expressed in both Df(L) and *smd13*.



**Figure 2.7. *Kindr* and *Kinesin-10-like* gene copy locations.** Coverage of Illumina DNA-seq reads generated from Ab10, smd13, smd8, and smd3 genotypes aligned to the Ab10 haplotype. Bottom two tracks show alignment from *Kindr* and *Kinesin-10-like* genes. The gap in alignment is a result of errors in the assembly in this region (uncorrected Nanopore data were used here).

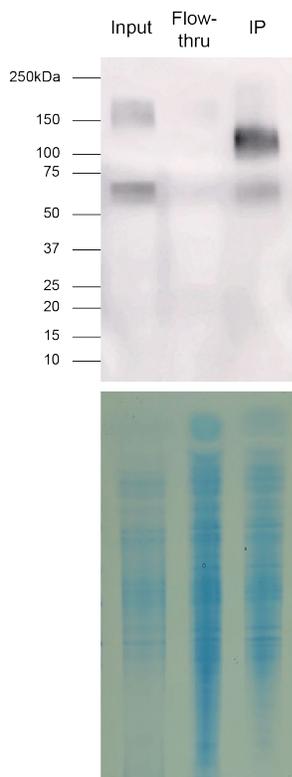


**Figure 2.8. Identification of *Kinesin-10-like*.**

(A) RNA-seq read coverage along TRINITY\_DN45531\_c2\_g1\_i24 in three replicates each of Ab10 anthers and ears.

(B) Boxplot of expression measured in RPKM detected in ears of Ab10, *smd13*, *smd8*, Df(L), and B73 genotypes with RNA-seq.

(C) Kinesin-10-like predicted protein sequence with putative domains shown.



**Figure 2.9. Protein gel analysis of Ab10 anti-KINDR IP experiment.** 1% of total input, flow-thru, and IP was run in each lane. Western blot (top) was probed using rabbit anti-KINDR antibody and colloidal coomassie (bottom) shows total protein loaded on gel.

## CHAPTER 3

### **Distinct kinesin motors drive two types of maize neocentromeres<sup>1</sup>**

---

<sup>1</sup> Swentowsky, Kyle W; Gent, Jonathan I; Lowry, Elizabeth G; Schubert, Veit; Ran, Xia; Tseng, Kuo-Fu; Harkess, Alex E; Qiu, Weihong; and Dawe, R Kelly. Manuscript published in *Genes and Development*, 2020. Reprinted here with permission of publisher.

## ABSTRACT

A maize chromosome variant called Abnormal chromosome 10 (Ab10) converts knobs on chromosome arms into neocentromeres, causing their preferential segregation to egg cells in a process known as meiotic drive. We previously demonstrated that the gene *Kinesin driver* (*Kindr*) on Ab10 encodes a kinesin-14 required to mobilize neocentromeres made up of the major tandem repeat knob180. Here we describe a second kinesin-14 gene, *TR-1 kinesin* (*Trkin*), that is required to mobilize neocentromeres made up of the minor tandem repeat TR-1. *Trkin* lies in a 4-Mb region of Ab10 that is not syntenic with any other region of the maize genome and shows extraordinary sequence divergence from *Kindr* and other kinesins in plants. Despite its unusual structure, *Trkin* encodes a functional minus end-directed kinesin that specifically colocalizes with TR-1 in meiosis, forming long drawn out neocentromeres. TRKIN contains a nuclear localization signal and localizes to knobs earlier in prophase than KINDR. The fact that TR-1 repeats often co-occur with knob180 repeats suggests that the current role of the TRKIN/TR-1 system is to facilitate the meiotic drive of the KINDR/knob180 system.

## INTRODUCTION

Meiotic drive describes any process whereby a selfish gene, complex of genes, or chromosomal feature reproducibly segregates to progeny at frequencies higher than Mendelian expectations (Fishman and McIntosh 2019). Most are male meiotic drive systems that use genetic elements to alter sperm or pollen viability and do not alter the mechanics of meiosis. In contrast are several female meiotic drive systems that directly impact the segregation of chromosomes (Dawe et al. 2018; Chmátal et al. 2014, 2015; Akera et al. 2017; Wu et al. 2018; Iwata-Otsubo et al. 2017; Fishman and Saunders 2008). Asymmetric segregation in mice is explained by differences in centromeric tandem repeat abundance that causes a preferential orientation of one chromosome toward the egg pole in meiosis I (Wu et al. 2018; Iwata-Otsubo et al. 2017). In maize, meiotic drive is caused by a chromosome variant, Abnormal chromosome 10 (Ab10), that converts tandem repeat arrays called knobs into spindle attachments called neocentromeres (Rhoades 1942; Longley 1945). According to Rhoades (Figure 3.1A), recombination between centromeres and knobs creates the opportunity for neocentromeres to move preferentially to the basal daughter cell that will ultimately become the female gametophyte that gives rise to the egg cell (Rhoades and Others 1952). Knobs can be tens of megabases in length, are strikingly polymorphic between lines, and exist in an extreme form of tightly condensed heterochromatin (McClintock 1929; Albert et al. 2010; Dawe et al. 2018; Liu et al. 2020). The ability of knobs to directly engage meiotic spindles as neocentromeres and undergo meiotic drive explains their prevalence in all maize lines (Buckler et al. 1999).

Recent results revealed that neocentromeres are activated by a kinesin-14 motor on Ab10 called *Kinesin-driver* (*Kindr*) (Dawe et al. 2018). Kinesins are a superfamily of microtubule-based motor proteins that display a wide range of functions including vesicle and organelle

movement and small molecule transport (Marx et al. 2009; Endow et al. 2010). The kinesin-14s are the only class that moves cargoes towards the minus ends of microtubules that accumulate at spindle poles (Lawrence et al. 2004; Gicking et al. 2018). The closest homologs of *Kindr* are maize *Variable kernel size1 (Vks1)* (Huang et al. 2019) and *Divergent spindle1 (Dv1)* (Clark 1940; Higgins et al. 2016), both of which function in spindle organization. KINDR differs by binding specifically to knobs and powering them towards spindle poles at both meiosis I and II to mediate their preferential segregation (Dawe et al. 2018).

Knobs can be found at 34 distinct positions and are typically located in midway positions along the chromosome arms in gene-dense areas (Kato and Others 1976; Buckler et al. 1999; Albert et al. 2010; Ghaffari et al. 2013). They are polymorphic in their presence/absence, size, and sequence composition, but most maize lines contain between 3-12 visible knobs (Kato and Others 1976; Albert et al. 2010). At the molecular level knobs are composed of two types of tandem repeats: the more abundant 180-bp tandem repeat (knob180) (Peacock et al. 1981), and the less abundant 359-bp tandem repeat (TR-1) (Ananiev et al. 1998). The KINDR protein localizes specifically to knob180 repeats (Dawe et al. 2018). A subset of the known Ab10 types, including the reference form Ab10-I-MMR, also promote neocentromere activity at TR-1 repeat arrays (Figure 3.1B, 2A). TR-1 neocentromeres remain active in *kindr* mutants, indicating that there are two separate genes that independently control neocentromere activity of knob180 and TR-1 neocentromeres (Hiatt et al. 2002). A factor that controls TR-1 neocentromere activity was previously mapped to a region of Ab10 that contains three TR-1 knobs (Figure 3.1B) (Hiatt et al. 2002).

Despite the prevalence of TR-1 repeats in *Zea* and its sister genus, *Tripsacum*, their role in meiotic drive has remained an open question. One hypothesis, that TR-1 knobs are required

for meiotic drive, is supported by data showing that a TR-1 knob on the chromosome variant K10L2 displays weak (51-52%) meiotic drive and that a TR-1-rich knob on chromosome 6 is preferentially segregated when crossed in the presence of Ab10 (Kanizay et al. 2013a). A second hypothesis is that TR-1 currently has no role in meiotic drive; this view is supported by the fact that TR-1 neocentromeres are not sufficient to cause meiotic drive of Ab10 when the KINDR/knob180 system is absent (Dawe et al. 2018). Yet a third view is that the TR-1 neocentromere system functions primarily as an antagonist of Ab10-meiotic drive. Support for the latter view comes from the fact that the TR-1-rich knob on K10L2 reduces the meiotic drive of Ab10 when paired in opposition (Kanizay et al. 2013a).

We recently completed a high quality genome assembly of Ab10 including the distal ~53 Mb region containing the meiotic drive haplotype (Figure 3.2A and (Liu et al. 2020)). The assembly gives molecular details to the many known features of Ab10 that were described using classical approaches, including its three TR-1-rich knobs (Hiatt et al. 2002), large knob composed primarily of knob180 repeats (Longley 1938; Peacock et al. 1981), two inversions containing genes shared with N10 (Mroczek et al. 2006), and the nine-gene *Kindr* complex that is required for knob180 neocentromere activity and meiotic drive (Dawe et al. 2018). Here we demonstrate that the new genome assembly includes another divergent kinesin-14 gene that we call *TR-1 kinesin (Trkin)*. TRKIN is a functional kinesin-14 protein that localizes specifically to TR-1 neocentromeres during meiosis. Analysis of neocentromeres in Ab10 types differing in the presence or absence of *Trkin* support the view that it is necessary for TR-1 neocentromere formation. Surprisingly, *Trkin* is not a homolog of *Kindr* and shows extreme sequence divergence suggestive of an ancient origin. We argue that in modern Ab10 types the TRKIN/TR-

1 system functions primarily to ensure the efficiency and fidelity of meiotic drive caused by the KINDR/knob180 system.

## RESULTS

*Trkin* encodes a divergent kinesin-14 protein

Whereas *Kindr* is encoded by nine genes distal to the large knob180-rich knob (Dawe et al. 2018; Liu et al. 2020), cytological analysis of Ab10 deletion derivatives indicate that the TR-1 neocentromere-activating factor is encoded near or among the small TR-1 knobs (Hiatt et al. 2002). In line with expectations, the complete Ab10 assembly revealed an expressed kinesin-14 gene between two of the TR-1 knobs in a ~4-Mb region that is not shared with N10 (Figure 3.2A). We named the gene *TR-1 Kinesin*, or *Trkin* for short. *Trkin* is a large gene extending over 130 kb (Figure 3.2A). It has 19 exons, 14 of which are organized in an alternating pattern of short introns followed by long introns. The large introns are composed primarily of transposable elements, including 27 kb of L1 line elements (Figure 3.8). To confirm the accuracy of the *Trkin* annotation, we sequenced 23 cDNA clones. Of these, 22 agreed with the original annotation. The single exception was an alternative splicing event that affected the last exon and a small amount of terminal sequence.

The predicted TRKIN protein contains three known motifs: a short ~54-amino acid coiled-coil domain, a kinesin-14 motor domain, and a nuclear localization sequence (NLS) (Figure 3.2B). Coiled-coil domains in kinesins are usually involved in protein dimerization, but the TRKIN coiled-coil region is unusually small. For comparison with other maize kinesin-14s, the coiled-coil region in KINDR is 128aa, in DV1 308aa, and in VKS1 303aa. The 126 amino

acids that comprise the N-terminal end of TRKIN are uncharacterized, but would normally be the location of a cargo-binding domain (Marx et al. 2009). The NLS lies in a 68-amino-acid C-terminal extension beyond the motor domain. Neither a C-terminal extension nor an NLS are present in other characterized kinesin-14 homologs (Cross and McAinsh 2014). The 1668-nt *Trkin* coding sequence does not produce significant DNA alignments outside of its conserved motor domain to annotated maize transcripts or complete maize genomes (except for pseudo-*Trkin*, see below).

To explore the relationship of *Trkin* to other kinesin-14s, we constructed a maximum likelihood phylogeny using cDNA sequences from the conserved motor domains. Kinesin-14s from angiosperms formed a cluster with two main lineages (Figure 3.2C). One of these contains the maize kinesin-14s *Kindr* and *VksI*, which were estimated to have diverged from each other roughly 12 mya (Dawe et al. 2018). Another lineage contains *Trkin* and its closest maize homolog *Dv1*. It is readily apparent that *Trkin* is not related by recent ancestry to *Kindr*. A simple interpretation based on its position within the tree is that the *Trkin* motor diverged before a split between rice (*Oryza sativa*) and the panicoid grasses (represented here by maize and *S. bicolor*) (Fig 2C). *Trkin* either evolved early in the grass lineage, or has rapidly evolved in a short time to become a kinesin-14 that differs significantly from any other kinesins in the grass clade.

TRKIN is a functional kinesin

The unusually short coiled-coil domain and low amino acid identity between TRKIN and other kinesins raises the question of whether it is a functional kinesin. To determine if key motifs

within the kinesin motor domain are conserved, we compared protein alignments of TRKIN with KINDR, DV1, VKS1, two homologs from rice (OsKIN14H and OsKIN14N), a homolog from *Arabidopsis* (AtATK3), one from the moss *Physcomitrella patens* (PpKIN14N), and another from the green algae *Chlamydomonas reinhardtii* (CHLREDRAFT120667). Despite the fact that the motor domain shows only 43.6% amino acid identity to its closest maize homolog DV1 (Higgins et al. 2016), the exon positions are conserved (Figure 3.2B) and all residues of the ATP binding site are present (Figure 3.9) (Gulick et al. 1998). There is weak similarity to the neck region at the N-terminal end of kinesin-14 motor domains that influence minus end-directed motility (Figure 3.9) (Yamagishi et al. 2016). TRKIN also lacks 16 residues (relative to other plant kinesin-14s) that would comprise the loop-8 motif (Figure 3.3A). The loop-8 motif resides near the microtubule-binding interface (Kozielski et al. 1997; Woehlke et al. 1997) and mutations affecting this region in the yeast kinesin-5, Cin8, have been shown to alter its microtubule-binding activity (Gerson-Gurwitz et al. 2011; Bell et al. 2017).

To test if the dramatic sequence divergence of TRKIN affects its ability to function as a minus-end directed motor, we purified a 6xHis/GFP-tagged version of TRKIN and performed *in vitro* microtubule-gliding experiments using total internal reflection fluorescence microscopy (Figure 3.3B). The results revealed that, like KINDR, TRKIN is a functional minus-end directed motor, moving microtubules along a microscope slide with brightly-labeled plus ends leading (Figure 3.3C). However, TRKIN generates considerably slower (38 nm/s; Figure 3.10) rates of gliding than KINDR (154 nm/s), suggesting that the mechanism of TRKIN movement is different from KINDR. *In vivo*, TRKIN appears to be at least as effective as KINDR, as TR-1 neocentromeres are long and streaking, often stretching across entire half spindles (Figure 3.4A and (Hiatt et al. 2002)).

TRKIN colocalizes specifically with TR-1 neocentromeres

Quantitative RT-PCR of cDNA from anther and ear tissue containing meiotic cells revealed *Trkin* expression (Figure 3.11). We also observed expression in leaf tissue, similar to what was observed for *Kindr* (Dawe et al. 2018). To determine the localization of TRKIN in meiosis, we generated antibodies against a peptide corresponding to 19 residues near the N-terminal end of TRKIN (Figure 3.4C). We combined TRKIN immunofluorescence with knob fluorescence *in situ* hybridization (FISH) on male meiocytes carrying Ab10-I-MMR. Meiotic drive does not occur in male meiosis because all four products of meiosis survive; however, neocentromeres are evident in microsporocytes and are traditionally scored in these cells. The data revealed a clear and precise overlap between TRKIN staining and TR-1 neocentromeres and an absence of staining over knob180 neocentromeres (Figure 3.4A). As a control we also carried out immunofluorescence on another Ab10 type, Ab10-II-MMR, which lacks TR-1 neocentromeres and appears to lack the N-terminal region that the antibody was generated against (Figure 3.4C). As predicted, there was no above-background TRKIN signal in Ab10-II-MMR meiocytes (Figure 3.4B).

Chromosome 10 variants that express full-length *Trkin* show TR-1 neocentromeres

Ab10 chromosomes have been categorized into groups based on their cytological characteristics: Ab10-I, with three small TR-1-rich knobs and a large knob180-rich knob; Ab10-II, with one TR-1-rich knob and two knob180-rich knobs; Ab10-III, which is similar to Ab10-I but with a large mixed knob (with both types of repeats); and K10L2, a chromosome 10 variant

that has a single large TR-1-rich knob (Figure 3.1B). Each of the Ab10 types differ in their capacity to activate TR-1 neocentromeres (Kanizay et al. 2013b), suggesting that they also differ in *Trkin* expression.

Ab10-I-MMR and Ab10-II-MMR are reference forms from Marcus M. Rhoades that trace to Mexico. Additional variants are from Mexico (Ab10-I-Jal, Ab10-II-Sal, Ab10-II-Tel, Ab10-III-Oax), from Guatemala (Ab10-III-Gua) and from Colombia (Ab10-III-Hui, Ab10-III-Caq) (Kanizay et al. 2013b; Higgins et al. 2018). We examined RNA-seq data from meiotic anthers taken from each of these haplotypes and wild-type siblings (Higgins et al. 2018), focusing on the genes expressed from the ~53 Mb Ab10 haplotype (Liu et al. 2020). Hierarchical clustering of the gene expression profiles (Figure 3.12) revealed that the K10L2 and N10 samples form a distinct cluster, consistent with the previous conclusion that K10L2 is an N10-like chromosome (Kanizay et al. 2013a). All of the Ab10 samples clustered into a single group, with the exception of one Ab10-III-Gua replicate that is probably from an N10 sibling that was mis-genotyped. While the Ab10-I and Ab10-III samples generally clustered together, Ab10-I-Jal clustered with the Ab10-II samples and is likely a chimeric form generated by recombination between Ab10 haplotypes (Kanizay et al. 2013a).

Among the diverse Ab10 types, we observed three categories of *Trkin* expression (Figure 3.4C). One group includes Ab10-I-MMR, Ab10-II-Tel, K10L2, and all Ab10-III accessions which encode a full-length *Trkin* transcript. The second includes Ab10-I-Pue, Ab10-II-MMR, and Ab10-II-Sal, which have an apparent truncation of the first five coding exons. The third is Ab10-I-Jal, which lacks *Trkin* expression altogether. The *Trkin* expression data correlate with the presence or absence of TR-1 neocentromeres as reported in the prior literature. Both Ab10-I-MMR and K10L2 express a full length *Trkin* transcript and show TR-1 neocentromeres (Hiatt et

al. 2002; Kanizay et al. 2013a) while Ab10-II-MMR does not express a full length *Trkin* transcript and does not show TR-1 neocentromeres (Mroczek et al. 2006).

We also scored neocentromeres in Ab10-II-Tel and Ab10-I-Pue by analyzing meiotic anaphase samples. The data reveal that while Ab10-II-Tel is structurally and phylogenetically different from Ab10-I-MMR and K10L2, all three express full length *Trkin* transcripts and display TR-1 neocentromeres (Figure 3.4D). Ab10-I-Pue is structurally and phylogenetically different from Ab10-II-MMR, yet neither express full length *Trkin* transcripts or show active TR-1 neocentromeres (Figure 3.4D). The correlation between *Trkin* expression and TR-1 neocentromere activity in five diverse lines suggests that *Trkin* is required for TR-1 neocentromere activity, although we cannot rule out other theoretical possibilities without a *Trkin* knockout mutation in an isogenic background.

Read alignment revealed that Ab10-I-MMR has two pseudogenes homologous to *Trkin*. One, pseudo-*Trkin1*, is structurally similar to *Trkin* but has numerous SNPs and indels, including an early frameshift in the 2nd exon that would render the transcript non-coding. It is part of a ~1 Mb inverted duplication encompassing *Trkin* (Figure 3.13). The other, pseudo-*Trkin2*, is located in a region of Ab10 that is shared with normal chromosome 10 (Figure 3.1B,2A), though non-recombining due to two large inversions. The maize inbred B73 carries a similar pseudo-*Trkin* in a position syntenic with Ab10 pseudo-*Trkin2* (Figure 3.13). Transcription of pseudo-*Trkin* could explain the apparent heterozygosity and non-uniform coverage over *Trkin* mRNA in some accessions (Figure 3.4C).

TRKIN localizes to knobs earlier in meiosis than KINDR

The availability of specific antibodies to TRKIN and KINDR made it possible to compare their localization patterns during each stage of male meiosis in Ab10-I-MMR plants. TRKIN localized to a few small puncta on chromosomes throughout meiosis I, meiotic interphase, and meiosis II (Figure 3.5). It was first visible before nuclear envelope breakdown in the pachytene substage of prophase I, suggesting that the predicted NLS in TRKIN is functional. In contrast, KINDR staining was not visible on knobs until after nuclear envelope breakdown. Both TRKIN and KINDR stained brightly throughout metaphase I, anaphase I and telophase I (Figure 3.5), and were detectable in meiotic interphase where they localized inside the nucleus at the site of the latent spindle poles. Both proteins were undetectable after the completion of telophase II. These data are consistent with the hypothesis that both TRKIN and KINDR serve as molecular motors for neocentromere motility, and suggest that TRKIN may be active in the earliest stages of spindle morphogenesis.

TR-1 sequences are conserved in *Zea* and *Tripsacum* and often occur in mixed arrays

One way to interpret the function of *Trkin* is to measure its impact on the sequence and distribution of TR-1 repeats throughout the genome. Long repeat arrays are inherently unstable (Garrido-Ramos 2017), and in the absence of selection, should accumulate sequence variants and ultimately be lost (Charlesworth et al. 1994). Both TR-1 and knob180 are present throughout *Zea* and within its close relative, *Tripsacum dactyloides* (Albert et al. 2010). To explore the variation among TR-1 sequences we produced TR-1 consensus sequences for individual plants from *Zea mays*, *Z. luxurians*, and *Tripsacum dactyloides* using short reads from previously published

studies (Novák et al. 2010; Gent et al. 2017). The TR-1 consensus sequences were similar among species with the consensus in *Z. mays* being 95% identical to *T. dactyloides* (Figure 3.6A). For comparison, the knob180 consensus sequence in *Z. mays* has 99% identity with *T. dactyloides*, and the centromere repeat *CentC* consensus has 97% identity with *T. dactyloides* (Gent et al. 2017). Mapping short reads to the consensus sequences revealed that the majority of TR-1 sequences were more than 90% identical to the consensus sequence in each case (Figure 3.6B), consistent with a repeat involved in meiotic drive with selection on the primary sequence. Although TR-1 is abundant in *T. dactyloides*, we were unable to find *Trkin* in publically available Illumina reads from this species, although it may be present in other accessions that were not assayed (data not shown).

Ab10 and many other knobs appear to contain both TR-1 and knob180 at the level of FISH (Albert et al. 2010; Kanizay et al. 2013a). To get a more detailed picture of knob mixing, we examined knobs and knob fragments that were assembled in the recently completed genomes of B73 and 25 other inbred lines known as the NAM founders (<https://nam-genomes.org/>). The data revealed that most knobs contain a significant amount of both repeat types. Only considering arrays of at least 100 kb in length, mixed knobs were the most common type, with a median of nine mixed arrays per line, compared with three containing knob180 alone and two with TR-1 alone (Figure 3.6C). The number of mixed arrays is likely higher, because most large knob180 knobs were not completely assembled and may contain TR-1 repeats that were not detected. These data support the prior assertion that the TRKIN/TR-1 system may be most effective in mixed knobs (Kanizay et al. 2013a).

## DISCUSSION

The plant kinesin-14 superfamily normally functions in spindle and phragmoplast morphogenesis, with lesser roles in nuclear migration and plastid distribution (Gicking et al. 2018). In keeping with these known roles, two native maize kinesin-14s function to focus meiotic spindle poles (*Dvl*) and ensure accurate spindle formation in endosperm (*Vks1*) (Higgins et al. 2016; Huang et al. 2019). *Kindr* shares a recent ancestor with *Vks1*, yet its function is to bind knob180 repeats and facilitate their segregation as neocentromeres to egg cells (Dawe et al. 2018). Here we describe *Trkin* as a fourth member of this clade and show that its function is to mobilize TR-1 repeats. The likely existence of *Trkin* among the three TR-1 rich knobs on Ab10 was predicted by early mapping studies of the TR-1 neocentromere phenotype (Hiatt et al. 2002). That *Trkin* might be a kinesin was a natural expectation after discovering that *Kindr* is a Kinesin-14 motor (Dawe et al. 2018). However, we were surprised to discover that *Trkin* is not related to *Kindr* by descent. The encoded protein shows a highly unusual structure that distinguishes it from all other known kinesin-14 proteins (Figure 3.2B, 3A).

Like KINDR, TRKIN is a functional kinesin-14 motor that interacts with a unique class of knob repeats at meiosis. The binding of kinesin to chromatin is remarkably sequence-specific for both the KINDR/knob180 and TRKIN/TR-1 systems, yet we do not understand how this selectivity is achieved. KINDR and TRKIN may either bind directly to DNA or bind as a complex with other proteins that are sequence specific. The interaction occurs primarily at metaphase and anaphase when neocentromeres are visibly active (Figure 3.5). However, unlike KINDR, TRKIN has a nuclear localization signal and can be seen associating with knobs in late prophase I (Figure 3.5). TR-1 neocentromeres are visible immediately upon nuclear envelope breakdown, whereas knob180 neocentromeres are slower to move (Hiatt et al. 2002). Another

notable difference is that TR-1 neocentromeres appear to uncoil as they move from the metaphase plate to the spindle pole, while knob180 neocentromeres retain their compact structure (Hiatt et al. 2002). The difference in appearance may reflect the unique properties of TRKIN or the genomic constitution of TR-1 arrays, which have far more retrotransposon insertions than knob180 arrays (Liu et al. 2020).

It remains unclear whether the TRKIN/TR-1 neocentromere system can confer meiotic drive on its own. When *Kindr* is absent, the TRKIN/TR-1 system is not sufficient to cause meiotic drive. However, the K10L2 variant, which does not have *Kindr* or knob180 repeats, shows a low level of meiotic drive (~51-52%). The transmission of K10L2 was statistically different from 50% in three independent experiments where over 10,000 kernels were counted (Kanizay et al. 2013a). The fact that K10L2 can compete with Ab10 when paired in opposition also supports the view that the TRKIN/TR-1 system is capable of supporting a low level of meiotic drive (Kanizay et al. 2013a). It is possible that the extreme elasticity of TR-1 neocentromeres reduces their efficiency as neocentromeres. If the elasticity is caused by transposon insertions that reduced the density of TR-1 repeats (Liu et al. 2020), it is reasonable to suggest that the TRKIN/TR-1 system was once a more powerful meiotic drive system but is now losing its effectiveness.

The current prevalence of mixed knobs suggests there is a benefit to having both work together. A likely scenario is that the early prophase movement and/or inherent elasticity of TR-1 neocentromeres enhances the accuracy of the primary KINDR/knob180 system, while the KINDR/knob180 system assures a high level of meiotic drive for the TR-1 system. One way that the TRKIN/TR-1 system may facilitate meiotic drive is by helping to resolve directional conflicts. Classic literature on dicentric chromosomes (McClintock 1939, 1941; Novitski 1952)

suggests that linked centromeres on dicentric chromosomes are as likely to move in opposite directions as they are to move in the same direction on the bipolar spindle. Linked knobs and centromeres should behave similarly. This problem is particularly acute in meiosis I where neocentromeres are physically separated from centromeres by chiasmata, which are held in place by sister chromatid cohesion (Figure 3.7). Sister chromatids do not separate until anaphase, at which point the centromeres are already in motion (Dawe 1998; Buonomo et al. 2000). Both chromatids may move to the same pole while their respective centromeres migrate in opposite directions, potentially causing chromosome lagging and breakage. However, since bridges and breakage are not regularly observed in Ab10 lines (Yu et al. 1997), the neocentromere traveling in the direction opposite to the centromere must switch directions. For the same reasons, directional conflicts are likely in meiosis II when both sister chromatids contain knobs. It is possible that TRKIN has a higher microtubule dissociation rate than KINDR and can change direction more effectively.

Some mouse strains demonstrate a process known as centromere drive where larger centromeres preferentially engage with the egg pole side of the spindle, which differs by having more tyrosinated tubulin (Akeru et al. 2017, 2019). This raises the question of whether there is any asymmetry during maize female meiosis. Indeed, there is visible asymmetry after both meiosis I and meiosis II, where the basal cells are larger (Evans and Grossniklaus 2009). There are also differences in callose deposition (a cell wall component) and the number of plasmodesmata (connections between cells) (Russell 1979). An asymmetrically distributed molecule, most likely one related to the spindle, could in theory help direct KINDR and TRKIN toward the basal cell in meiosis I or II. However, as originally postulated, the Rhoades model for meiotic drive in maize requires neither spindle asymmetry nor other molecular asymmetry

(Rhoades 1952). According to this view, meiosis I simply sets the stage by pulling all knobs towards the polar periphery of telophase nuclei (Figure 3.1A), where they appear (at least in male meiosis) to stay throughout interphase (Dawe and Cande 1996). The peripheral location has been postulated to facilitate the movement of knobbed chromatids to the basal cell in meiosis II (Yu et al. 1997; Dawe and Hiatt 2004) as a natural outcome of Ran-regulated spindle self-organization (Zhang and Dawe 2011). As the spindle forms around chromatin, basally-oriented neocentromeres could apply tension to the linked kinetochores such that knobbed chromatids are more likely to develop a connection to the spindle facing the basal cell than their corresponding knobless sister chromatids (Figure 3.7). The fact that TRKIN contains a nuclear localization signal whereas KINDR does not suggests that it may be specially adapted to this key early movement event.

In summary, our results support the conclusion that KINDR functions as the primary driver (Dawe et al. 2018) while TRKIN serves to improve the efficiency of drive and/or limit the negative consequences of Ab10 on the organism. Our data showing that the majority of knobs contain both knob180 and TR-1 repeats argues against the proposal that TR-1 is under selection primarily an antagonist of the KINDR/knob180 system, and supports the view that the two repeats cooperate to promote meiotic drive (Kanizay et al. 2013a). Further progress in understanding the role of TRKIN will require making a *Trkin* knockout on an otherwise intact Ab10 chromosome. A *Trkin* knockout may reduce meiotic drive, but based on our results showing that multiple Ab10 chromosomes naturally lack *Trkin*, we believe the effects will be minimal. The more important role of *Trkin* may be to reduce the frequency of anaphase errors that cause chromosome loss and reduced fertility. Mathematical modeling suggests that high pollen and ovule viability when Ab10 is heterozygous is critical for the drive system to spread

into new populations and maintain a stable equilibrium (Hall and Dawe 2017). Careful analyses of meiotic anaphase in Ab10 lines with and without *Trkin*, as well measurements of pollen viability and seed set, will be needed to fully test this hypothesis.

## MATERIALS AND METHODS

### Genome references and annotation

The *Trkin* gene was annotated as a part of the Ab10 assembly project (Liu et al. 2020). Sequences from genomic regions containing pseudo-*Trkin* arrays were extracted from the maize Ab10 assembly (Liu et al. 2020) and B73v5 assembly (<https://www.maizegdb.org/genome/assembly/Zm-B73-REFERENCE-NAM-5.0>). Knob composition data were obtained from publicly available maize NAM genome assemblies (<https://jbrowse.maizegdb.org/?data=all>). *Trkin* cDNA sequence is available on NCBI GenBank, accession MT459824.

### mRNA sequencing and analysis

All mRNA samples were derived from whole anthers at the meiotic stage (2-3mm) and Illumina sequenced with paired-end 75-nt reads, as described previously (Higgins et al. 2018). In addition to the seven accessions from the previous study (NCBI BioProject PRJNA339442), ten additional chromosome 10 variant genotypes were used for this study (NCBI BioProject PRJNA285341). These include six N10, one K10L2, three Ab10-I, three Ab10-II, and four

Ab10-III accessions that are listed along with their SRR identifiers in Table 3.1. For each Ab10 type we collected data from a minimum of three different plants. For the analysis of *Trkin* transcript levels among various genotypes, the following protocol was used. Replicates were pooled and mapped to the *Trkin* cDNA sequence using BWA MEM v0.7.15 (Li 2013) with default parameters. Mapped reads were filtered to allow up to one mismatch in samples known to contain a SNP (Ab10-I-Pue, Ab10-II-MMR, Ab10-II-Tel, Ab10-II-Sal, and K10L2) but no mismatches were allowed for all other samples. Reads were sorted using SAMtools sort v1.3.1 (Li et al. 2009) and exported for visualization in IGV v2.6.3 (Robinson et al. 2011).

For expression clustering analysis, reads were initially trimmed using Sickle v1.33 (Joshi and Fass 2011) and mapped to the Ab10 reference using default parameters in the STAR v2.7.1 package (Dobin et al. 2013). Gene expression in TPM was calculated using the package TPMCalculator v0.2 (Vera Alvarez et al. 2019). TPM values were imported into RStudio for analysis and only transcripts present in the Ab10 haplotype region (after coordinate 135 Mb) were retained for analysis. The heatmap.2 function in gplots (Warnes et al. 2005) was used for hierarchical clustering and heatmap visualization.

#### Analysis of *Trkin* orthologs

*Trkin* homologs from other species were identified by BLAST (NCBI) using the *Trkin* cDNA sequence as a query. The gene with the best match was chosen for comparison. Gene sequences and their predicted proteins were aligned in Geneious Prime 2019.2.3 (<http://www.geneious.com/>) using the default MUSCLE alignment options.

For phylogenetic gene tree construction, kinesin-14 motor domain coding sequences were aligned in Geneious Prime 2019.2.3 using the default MUSCLE alignment options. The alignment was uploaded into IQ-TREE (Minh et al. 2020) and a phylogeny generated using maximum likelihood gene tree construction using the GTR + gamma + i models (GTR+G+I+F4 options) to account for rate heterogeneity.

#### Analysis of pseudo-Trkin homologs

Positions of pseudo-*Trkin* exon orthologs in Ab10 and B73v5 were determined using NCBI Blast v2.2.26. Large genomic sequences were aligned in MiniMap2 (Li 2018) using the -cx asm5 option. Alignments with a quality score of 60 or higher were retained for analysis. Genomic alignments were visualized along with pseudo-*Trkin* exon alignments using the `geom_point` function in `ggplot2` (Wickham 2016).

#### TR-1 consensus sequences

Consensus sequences for TR-1 were derived from 150-nt single-end Illumina reads from micrococcal nuclease digested chromatin from a previous study (Gent et al. 2017). SRA run accession numbers are as follows: *Zea mays* (B73), SRR5466391; *Zea luxurians* (PI 422162), SRR5466389; *Tripsacum dactyloides* (PI 421612), SRR5466393. Consensus sequences were produced with RepeatExplorer (Version 2018.11.09) (Novák et al. 2010) using parameters “--max\_memory 500000000 -tax VIRIDIPLANTAE3.0 -opt ILLUMINA”. Sample-specific consensus sequences were aligned to each other with Geneious Prime 2019.2.3

(<http://www.geneious.com/>). The percent identity of individual reads with consensus sequences was determined using blastall with parameters as follows: “-p blastn -e 1e-5 -W 7 -G 2 -E 1 -r 1 -q -1”. Only reads producing alignments of at least 125 bp in length to consensus sequence dimers were included.

### Analysis of knob composition in NAM lines

Tandem repeat array data for the 26 maize NAM founder genome assemblies were annotated as part of the NAM sequencing project (<https://nam-genomes.org/>). Here “arrays” are defined as large areas of assemblies composed primarily of short tandem repeats but may also be interspersed with other sequences. These data were filtered to only include arrays with over 100kb of tandem repeat sequences. Only arrays within the final chromosome assemblies (pseudomolecules) were considered. Knob fragments that were not assigned to a chromosome were not included in the analysis. Knobs were called “knob180” or “TR-1” if more than 95% of the knob repeats in the array were composed of one type of knob repeat. Knobs were called “mixed” if they contained both repeats and one type made up less than 95% of the total.

### Plant Materials

The Immunolocalization of TRKIN in Figure 3.4A was performed on a heterozygous K10L2/Ab10-Df(L) line that is homozygous for *Trkin* but lacks *Kindr* (Dawe et al. 2018). All other lines used for immunolocalization or FISH (Ab10-I-MMR, Ab10-II-MMR, Ab10-I-Pue and Ab10-II-Tel) were previously described (Higgins et al. 2018) and heterozygous with N10.

## TRKIN Antibody preparation

Antibodies were prepared by Pacific Immunology, Ramona, CA. The peptide TRKIN(6-24): RGEEPKVVAHREDIKAQFK-Cys was injected into rabbits and antibodies affinity purified using the same peptide.

## Quantitative RT-PCR

cDNA samples from Ab10 and N10 (B73 inbred) ears, anthers, leaf tips, and leaf bases originally described in Dawe et al. (2018) were assayed for the presence of *Trkin* using the SYBR Green PCR Master Mix (Applied Biosystems Cat# 4367659) and assayed in a C1000 Touch Thermal Cycler with CFX96 Real Time System (BioRad Cat# 1855195). *Trkin* primers were designed to capture the splice junction between exons 13-14: Trkin qPCR F1 (5'-GGAGGCCAAGGCCATAAAT-3') and Trkin qPCR R1 (5'-TGAGACAGAGTCGATCCTCTAAA-3'). Control primers were from the LUG gene (Manoli et al. 2012): LUGF (5'-TCCAGTGCTACAGGGAAGGT-3') and LUGR (5'-GTTAGTTCTTGAGCCCACGC-3'). PCR was performed using an initial denaturing step at 95°C for 10 min followed by 40 cycles of denaturation (95°C, 15 s) and annealing/extension (60°C, 1 min). Expression was measured from three different plants per tissue and error bars represent biological variation.

## Immunolocalization and fluorescence in situ hybridization (FISH)

Immunolocalization was performed as described in (Dawe et al. 2018). Briefly, tassels were removed from plants and meiotic anthers fixed in 4% paraformaldehyde in PHEMS buffer and 1% Triton X-100 for one hour. Meiocytes were extruded from anthers and immobilized onto polylysine-coated coverslips by centrifugation at 100xG for one minute. Coverslips were incubated in a permeabilization solution for one hour and then in a blocking solution containing 10% goat serum diluted in PBS for two hours. Antibodies were diluted in an antibody dilution buffer (3% BSA diluted in PBS) and incubated overnight at 37°C. The primary antibodies were mAb mouse anti-tubulin (Asai et al. 1982) diluted 1:200, pAb rabbit anti-KINDR (Dawe et al. 2018) diluted 1:100 or 1:250, and the pAb rabbit anti-TRKIN described above diluted 1:100. An additional blocking step was performed as above then coverslips were incubated in a secondary antibody solution containing Rhodamine-conjugated AffiniPure Donkey Anti-Rabbit IgG H+L (Jackson 711-025-152) and Fluorescein-conjugated AffiniPure Goat Anti-Mouse IgG H+L (Jackson 115-095-146), both diluted 1:200 for two and a half hours at room temperature in the dark. The coverslips were mounted on microscope slides using ProLong Gold with DAPI (Thermo Fisher Cat# P36931) before imaging on a Zeiss Axio Imager.M1 fluorescence microscope with a 63 X Plan-APO Chromat oil objective, and data analyzed using Slidebook software (Intelligent Imaging Innovations, Denver, CO, USA).

Combined FISH-immunolocalization for Figure 3.4A was performed as described in (Dawe et al. 2018). Briefly, coverslips containing fixed meiocytes were suspended over a slide using broken bits of coverslips and a solution containing oligo probes (FITC-labeled 180bp repeat oligos and Rhodamine-labeled TR-1 repeat oligos, see (Kanizay et al. 2013b)) was pipetted under the coverslip and the edges were sealed using nail polish. The slides were heated

at 95°C for five minutes then incubated at room temperature for two hours in the dark. The coverslips were removed and washed, and immunolocalization performed as described above using the primary rabbit anti-TRKIN antibody and secondary Cy5 AffiniPure Donkey Anti-Rabbit IgG H+L (Jackson Cat# 711-175-152). Coverslips were mounted on slides using ProLong Gold with DAPI (Thermo Fisher Cat# P36931) and imaged using Structured Illumination Microscopy.

FISH for Figure 3.4D was performed as described in (Dawe and Cande 1996). Briefly, coverslips containing fixed meiocytes were incubated with the following solutions for 10 mins each: 1X SSC (4.38g/liter sodium citrate/8.75g/liter NaCl), 20% formamide; 2X SSC/30% formamide; and 2X SSC/50% formamide. Broken bits of coverslips were placed at four corners of a slide, and the coverslips with meiocytes were placed upside down over the broken pieces. 75µl of a solution containing 2X SSC/50% formamide and 1 µg/ml of fluorescently-labeled oligos (FITC-labeled 180bp repeat oligos, Rhodamine-labeled TR-1 repeat oligos, and Cy5-labeled CentC repeat oligos, see (Kanizay et al. 2013b)) was pipetted under the coverslip and the edges sealed with rubber cement. The slide was placed on a 95°C heating block for five minutes and the probes were allowed to anneal at room temperature overnight. The rubber cement was removed, and the coverslip was incubated with the following solutions for 10 mins each: 2X SSC/20% formamide/0.01% Tween-20; 1X SSC/10% formamide/0.01% Tween-20; 1X SSC/1X PBS; 1X PBS; 1X PBS/0.1µg/ml DAPI. The coverslips were mounted on microscope slides and imaged on a Zeiss Axio Imager.M1 fluorescence microscope as described above.

## Structured Illumination Microscopy

Super-resolution microscopy was used to analyze the spatial arrangement of TRKIN, TR-1 and knob180 in Figure 3.4A. Spatial structured illumination microscopy (3D-SIM) was applied using a Plan-Apochromat 63×/1.4 oil objective of an Elyra PS.1 microscope system and the software ZENblack (Carl Zeiss GmbH). Image stacks were captured separately for each fluorochrome using 642, 561, 488, and 405 nm laser lines for excitation and appropriate emission filters. The 3D-SIM stacks with a step size of 110 nm were acquired consecutively for each fluorophore starting with the highest wavelength dye to minimize bleaching. Maximum intensity projections were calculated based on the image stacks using the ZENblack software (Weisshart et al. 2016). The same image stacks were also used to produce 3D movies by the Imaris 8.0 (Bitplane) software.

## TRKIN protein purification and microtubule-gliding experiments

An *E. coli* codon-optimized version of the *Trkin* CDS was synthesized by Genscript and cloned into the expression vector pET-17b (Novagen) modified to include N-terminal 6xHis- and GFP-tags for affinity protein purification. Subsequent steps were performed as described in (Dawe et al. 2018). Briefly, expression was induced in BL21(DE3) Rosetta cells (Novagen, Cat #70954-3) with 0.1mM IPTG for 12-14 hours at 20°C. Cells were centrifuged and re-suspended in 50 mM sodium phosphate buffer (pH 7.2) containing 250 mM NaCl, 1 mM MgCl<sub>2</sub>, 0.5 mM ATP, 10 mM β-mercaptoethanol, 20 mM imidazole, and a protease inhibitor cocktail, then lysed by sonication. The lysate was centrifuged and soluble protein in the supernatant was purified by

Talon resin (Clontech) and eluted in 50 mM sodium phosphate buffer described above, but with 250 mM imidazole. Protein was flash-frozen and stored at -80°C.

*In vitro* gliding assays were performed as described in (Dawe et al. 2018). Briefly, polarity-labeled microtubules were first prepared as described in (Hyman 1991). Time-lapse microtubule gliding experiments were performed at room temperature using a Zeiss Axio Observer Z1 objective-type TIRF microscope equipped with a 100X 1.46 NA oil-immersion objective and a back-thinned electron multiplier CCD camera (Photometrics). Flow chambers were made by attaching a coverslip to a glass slide by double-sided tape as described in (Popchok et al. 2017) and the chamber was perfused with an anti-His antibody diluted in BRB12 buffer and incubated for two minutes at room temperature. Unbound antibody was washed away and one chamber volume of 6xHis-TRKIN diluted in BRB50 supplemented with 20µM taxol and 1.3 mg/ml casein was added to the chamber. Following two minutes of incubation, unbound protein was washed away using BRB50 supplemented with 20 µM taxol and 1.3 mg/ml casein. Polarity-labeled TMR (tetramethylrhodamine) microtubules diluted in BRB50 supplemented with 20 µM taxol and 1.3 mg/ml casein were then added to the chamber and unbound microtubules were removed with 2 chamber volumes of this buffer after a two minute incubation. The flow chamber was then perfused with one chamber volume of BRB50 motility buffer containing 1 mM ATP and 25 µM taxol, 1.3 mg/ml casein, and an oxygen scavenger system. Time-lapse images were taken at 1 frame per 5s for 5 minutes. The measurements of 100 microtubules were collected from 14 different gliding experiments. Although the KINDR gliding assays published in (Dawe et al. 2018) and the TRKIN assays reported here were not performed in parallel, they were carried out under similar conditions with the same microscopy system.

## ACKNOWLEDGEMENTS

This study was supported by NSF grants 1412063 and 1925546 to RKD, 1616462 to WQ, and an NSF Doctoral Dissertation Improvement Grant to EGL. We relied on computational resources and technical expertise from the Georgia Advanced Computing Resource Center, a partnership between the University of Georgia's Office of the Vice President for Research and Office of the Vice President for Information Technology.

## REFERENCES

- Akera T, Chmátal L, Trimm E, Yang K, Aonbangkhen C, Chenoweth DM, Janke C, Schultz RM, Lampson MA. 2017. Spindle asymmetry drives non-Mendelian chromosome segregation. *Science* 358: 668–672.
- Akera T, Trimm E, Lampson MA. 2019. Molecular Strategies of Meiotic Cheating by Selfish Centromeres. *Cell* 178: 1132–1144.e10.
- Albert PS, Gao Z, Danilova TV, Birchler JA. 2010. Diversity of chromosomal karyotypes in maize and its relatives. *Cytogenet Genome Res* 129: 6–16.
- Ananiev EV, Phillips RL, Rines HW. 1998. A knob-associated tandem repeat in maize capable of forming fold-back DNA segments: are chromosome knobs megatransposons? *Proc Natl Acad Sci U S A* 95: 10785–10790.
- Asai DJ, Thompson WC, Wilson L, Brokaw CJ. 1982. Two different monoclonal antibodies to alpha-tubulin inhibit the bending of reactivated sea urchin spermatozoa. *Cell Motil* 2: 599–614.
- Bell KM, Cha HK, Sindelar CV, Cochran JC. 2017. The yeast kinesin-5 Cin8 interacts with the microtubule in a noncanonical manner. *J Biol Chem* 292: 14680–14694.
- Buckler ES 4th, Phelps-Durr TL, Buckler CS, Dawe RK, Doebley JF, Holtsford TP. 1999. Meiotic drive of chromosomal knobs reshaped the maize genome. *Genetics* 153: 415–426.
- Buonomo SB, Clyne RK, Fuchs J, Loidl J, Uhlmann F, Nasmyth K. 2000. Disjunction of homologous chromosomes in meiosis I depends on proteolytic cleavage of the meiotic cohesin Rec8 by separin. *Cell* 103: 387–398.
- Charlesworth B, Sniegowski P, Stephan W. 1994. The evolutionary dynamics of repetitive DNA in eukaryotes. *Nature* 371: 215–220.
- Chmátal L, Gabriel SI, Mitsainas GP, Martínez-Vargas J, Ventura J, Searle JB, Schultz RM, Lampson MA. 2014. Centromere strength provides the cell biological basis for meiotic drive and karyotype evolution in mice. *Curr Biol* 24: 2295–2300.
- Chmátal L, Yang K, Schultz RM, Lampson MA. 2015. Spatial Regulation of Kinetochore Microtubule Attachments by Destabilization at Spindle Poles in Meiosis I. *Curr Biol* 25: 1835–1841.

- Clark FJ. 1940. Cytogenetic Studies of Divergent Meiotic Spindle Formation in *Zea Mays*. *Am J Bot* 27: 547–559.
- Cross RA, McAinsh A. 2014. Prime movers: the mechanochemistry of mitotic kinesins. *Nat Rev Mol Cell Biol* 15: 257–271.
- Dawe RK. 1998. MEIOTIC CHROMOSOME ORGANIZATION AND SEGREGATION IN PLANTS. *Annu Rev Plant Physiol Plant Mol Biol* 49: 371–395.
- Dawe RK, Cande WZ. 1996. Induction of centromeric activity in maize by suppressor of meiotic drive 1. *Proc Natl Acad Sci U S A* 93: 8512–8517.
- Dawe RK, Hiatt EN. 2004. Plant neocentromeres: fast, focused, and driven. *Chromosome Res* 12: 655–669.
- Dawe RK, Lowry EG, Gent JI, Stitzer MC, Swentowsky KW, Higgins DM, Ross-Ibarra J, Wallace JG, Kanizay LB, Alabady M, et al. 2018. A Kinesin-14 Motor Activates Neocentromeres to Promote Meiotic Drive in Maize. *Cell* 173: 839–850.e18.
- Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR. 2013. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* 29: 15–21.
- Endow SA, Kull FJ, Liu H. 2010. Kinesins at a glance. *J Cell Sci* 123: 3420–3424.
- Evans MMS, Grossniklaus U. 2009. The Maize Megagametophyte. In *Handbook of Maize: Its Biology* (eds. J.L. Bennetzen and S.C. Hake), pp. 79–104, Springer New York, New York, NY.
- Fishman L, McIntosh M. 2019. Standard Deviations: The Biological Bases of Transmission Ratio Distortion. *Annu Rev Genet* 53: 347–372.
- Fishman L, Saunders A. 2008. Centromere-associated female meiotic drive entails male fitness costs in monkeyflowers. *Science* 322: 1559–1562.
- Garrido-Ramos MA. 2017. Satellite DNA: An Evolving Topic. *Genes* 8. <http://dx.doi.org/10.3390/genes8090230>.
- Gent JI, Wang N, Dawe RK. 2017. Stable centromere positioning in diverse sequence contexts of complex and satellite centromeres of maize and wild relatives. *Genome Biol* 18: 121.
- Gerson-Gurwitz A, Thiede C, Movshovich N, Fridman V, Podolskaya M, Danieli T, Lakämper S, Klopfenstein DR, Schmidt CF, Gheber L. 2011. Directionality of individual kinesin-5 Cin8 motors is modulated by loop 8, ionic strength and microtubule geometry. *EMBO J* 30: 4942–4954.
- Ghaffari R, Cannon EKS, Kanizay LB, Lawrence CJ, Dawe RK. 2013. Maize chromosomal knobs are located in gene-dense areas and suppress local recombination. *Chromosoma* 122: 67–75.
- Gicking AM, Swentowsky KW, Dawe RK, Qiu W. 2018. Functional diversification of the kinesin-14 family in land plants. *FEBS Lett* 592: 1918–1928.
- Gulick AM, Song H, Endow SA, Rayment I. 1998. X-ray crystal structure of the yeast Kar3 motor domain complexed with Mg-ADP to 2.3 Å resolution. *Biochemistry* 37: 1769–1776.
- Hall DW, Dawe RK. 2017. Modeling the evolution of female meiotic drive in maize. G3: Genes, Genomes. *Genetics DOI* 10: g3.
- Hiatt EN, Kentner EK, Dawe RK. 2002. Independently regulated neocentromere activity of two classes of tandem repeat arrays. *Plant Cell* 14: 407–420.
- Higgins DM, Lowry EG, Kanizay LB, Becraft PW, Hall DW, Dawe RK. 2018. Fitness Costs and Variation in Transmission Distortion Associated with the Abnormal Chromosome 10 Meiotic Drive System in Maize. *Genetics* 208: 297–305.
- Higgins DM, Nannas NJ, Dawe RK. 2016. The Maize Divergent spindle-1 (dv1) Gene Encodes a

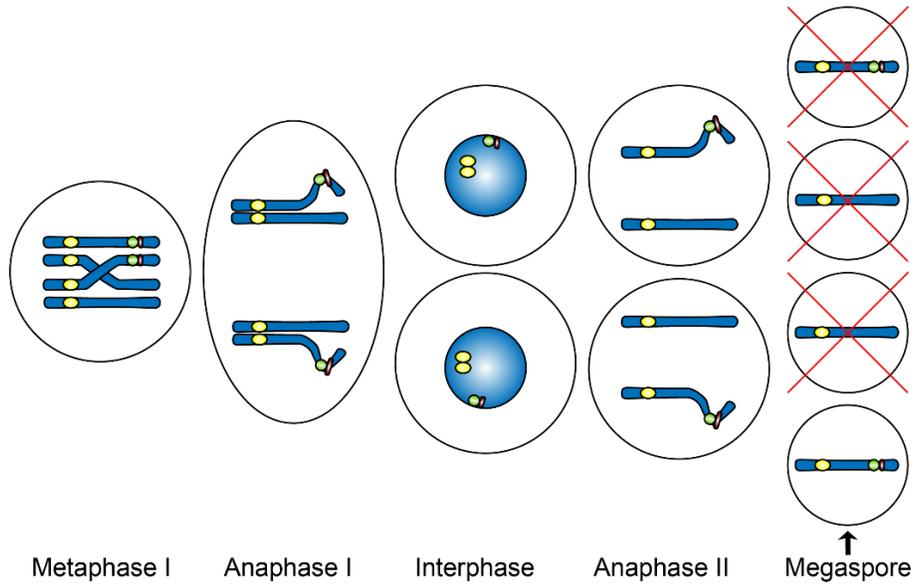
- Kinesin-14A Motor Protein Required for Meiotic Spindle Pole Organization. *Front Plant Sci* 7: 1277.
- Huang Y, Wang H, Huang X, Wang Q, Wang J, An D, Li J, Wang W, Wu Y. 2019. Maize VKS1 Regulates Mitosis and Cytokinesis During Early Endosperm Development. *Plant Cell* 31: 1238–1256.
- Hyman AA. 1991. Preparation of marked microtubules for the assay of the polarity of microtubule-based motors by fluorescence. *J Cell Sci Suppl* 14: 125–127.
- Iwata-Otsubo A, Dawicki-McKenna JM, Akera T, Falk SJ, Chmátal L, Yang K, Sullivan BA, Schultz RM, Lampson MA, Black BE. 2017. Expanded Satellite Repeats Amplify a Discrete CENP-A Nucleosome Assembly Site on Chromosomes that Drive in Female Meiosis. *Curr Biol* 27: 2365–2373.e8.
- Joshi NA, Fass JN. 2011. *Sickle: A sliding-window, adaptive, quality-based trimming tool for FastQ files*. <https://github.com/najoshi/sickle>.
- Kanizay LB, Albert PS, Birchler JA, Dawe RK. 2013a. Intragenomic conflict between the two major knob repeats of maize. *Genetics* 194: 81–89.
- Kanizay LB, Pyhäjärvi T, Lowry EG, Hufford MB, Peterson DG, Ross-Ibarra J, Dawe RK. 2013b. Diversity and abundance of the abnormal chromosome 10 meiotic drive complex in *Zea mays*. *Heredity* 110: 570–577.
- Kato YTA, Others. 1976. Cytological studies of maize (*Zea mays* L.) and teosinte (*Zea mexicana* Schrader Kuntze) in relation to their origin and evolution. *Mass Agric Exp Stn Bull* 635: 1–185.
- Kozielski F, Sack S, Marx A, Thormählen M, Schönbrunn E, Biou V, Thompson A, Mandelkow EM, Mandelkow E. 1997. The crystal structure of dimeric kinesin and implications for microtubule-dependent motility. *Cell* 91: 985–994.
- Lawrence CJ, Dawe RK, Christie KR, Cleveland DW, Dawson SC, Endow SA, Goldstein LSB, Goodson HV, Hirokawa N, Howard J, et al. 2004. A standardized kinesin nomenclature. *J Cell Biol* 167: 19–22.
- Li H. 2013. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv [q-bioGN]*. <http://arxiv.org/abs/1303.3997>.
- Li H. 2018. Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics* 34: 3094–3100.
- Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, 1000 Genome Project Data Processing Subgroup. 2009. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 25: 2078–2079.
- Liu J, Seetharam AS, Chougule K, Ou S, Swentowsky KW, Gent JI, Llaca V, Woodhouse MR, Manchanda N, Presting GG, et al. 2020. Gapless assembly of maize chromosomes using long-read technologies. *Genome Biol* 21: 121.
- Longley AE. 1945. Abnormal Segregation during Megasporogenesis in Maize. *Genetics* 30: 100–113.
- Longley AE. 1938. Chromosomes of Maize from North American Indians. *J Agric Res* 56: 177–195.
- Marx A, Hoenger A, Mandelkow E. 2009. Structures of kinesin motor proteins. *Cell Motil Cytoskeleton* 66: 958–966.
- McClintock B. 1929. CHROMOSOME MORPHOLOGY IN ZEA MAYS. *Science* 69: 629.
- McClintock B. 1939. The Behavior in Successive Nuclear Divisions of a Chromosome Broken at Meiosis. *Proc Natl Acad Sci U S A* 25: 405–416.

- McClintock B. 1941. The Stability of Broken Ends of Chromosomes in *Zea Mays*. *Genetics* 26: 234–282.
- Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, von Haeseler A, Lanfear R. 2020. IQ-TREE 2: New models and efficient methods for phylogenetic inference in the genomic era. *Mol Biol Evol*. <http://dx.doi.org/10.1093/molbev/msaa015>.
- Mroczek RJ, Melo JR, Luce AC, Hiatt EN, Dawe RK. 2006. The maize Ab10 meiotic drive system maps to supernumerary sequences in a large complex haplotype. *Genetics* 174: 145–154.
- Novák P, Neumann P, Macas J. 2010. Graph-based clustering and characterization of repetitive sequences in next-generation sequencing data. *BMC Bioinformatics* 11: 378.
- Novitski E. 1952. The Genetic Consequences of Anaphase Bridge Formation in *Drosophila*. *Genetics* 37: 270–287.
- Peacock WJ, Dennis ES, Rhoades MM, Pryor AJ. 1981. Highly repeated DNA sequence limited to knob heterochromatin in maize. *Proc Natl Acad Sci U S A* 78: 4490–4494.
- Popchok AR, Tseng K-F, Wang P, Karplus PA, Xiang X, Qiu W. 2017. The mitotic kinesin-14 KlpA contains a context-dependent directionality switch. *Nat Commun* 8: 13999.
- Rhoades MM. 1942. Preferential Segregation in Maize. *Genetics* 27: 395–407.
- Rhoades MM, Others. 1952. Preferential segregation in maize. *Preferential segregation in maize*. <https://www.cabdirect.org/cabdirect/abstract/19531603828>.
- Robinson JT, Thorvaldsdóttir H, Winckler W, Guttman M, Lander ES, Getz G, Mesirov JP. 2011. Integrative genomics viewer. *Nat Biotechnol* 29: 24–26.
- Russell SD. 1979. Fine structure of megagametophyte development in *Zea mays*. *Can J Bot* 57: 1093–1110.
- Vera Alvarez R, Pongor LS, Mariño-Ramírez L, Landsman D. 2019. TPMCalculator: one-step software to quantify mRNA abundance of genomic features. *Bioinformatics* 35: 1960–1962.
- Warnes G, Bolker B, Bonebakker L, Gentleman R, Huber W, Liaw A, Lumley T, Mächler M, Magnusson A, Möller S. 2005. *gplots: Various R programming tools for plotting data*.
- Weisshart K, Fuchs J, Schubert V. 2016. Structured illumination microscopy (SIM) and photoactivated localization microscopy (PALM) to analyze the abundance and distribution of RNA polymerase II molecules on flow-sorted Arabidopsis nuclei. *Bio-protocol* 6: e1725.
- Wickham H. 2016. ggplot2: Elegant Graphics for Data Analysis. <https://ggplot2.tidyverse.org>.
- Woehlke G, Ruby AK, Hart CL, Ly B, Hom-Booher N, Vale RD. 1997. Microtubule interaction site of the kinesin motor. *Cell* 90: 207–216.
- Wu T, Lane SIR, Morgan SL, Jones KT. 2018. Spindle tubulin and MTOC asymmetries may explain meiotic drive in oocytes. *Nat Commun* 9: 2952.
- Yamagishi M, Shigematsu H, Yokoyama T, Kikkawa M, Sugawa M, Aoki M, Shirouzu M, Yajima J, Nitta R. 2016. Structural Basis of Backwards Motion in Kinesin-1-Kinesin-14 Chimera: Implication for Kinesin-14 Motility. *Structure* 24: 1322–1334.
- Yu HG, Hiatt EN, Chan A, Sweeney M, Dawe RK. 1997. Neocentromere-mediated chromosome movement in maize. *J Cell Biol* 139: 831–840.
- Zhang H, Dawe RK. 2011. Mechanisms of plant spindle formation. *Chromosome Res* 19: 335–344.

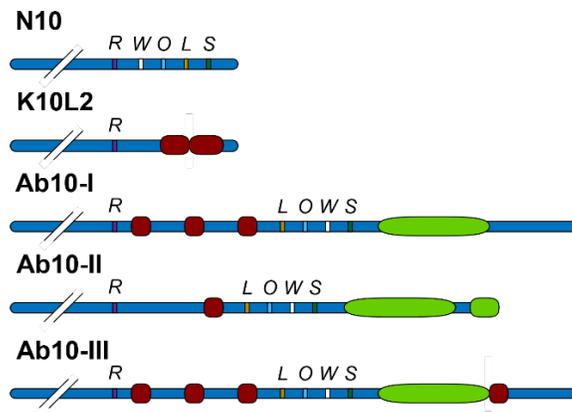
**Table 3.1. SRA project and run codes for Illumina reads used in this study. All data can be found at <https://www.ncbi.nlm.nih.gov/sra>**

Sample	SRA Project	Rep 1	Rep 2	Rep 3	Rep 4	Rep 5
Ab10I-MMR	PRJNA285341	SRR2045277	SRR2045279	SRR2045282		
Ab10II-MMR	PRJNA285341	SRR2045283	SRR2045444	SRR2045452		
K10L2	PRJNA285341	SRR2045453	SRR2046269	SRR2046270	SRR2046271	SRR2046272
Ab10-II-Tel	PRJNA285341	SRR2046273	SRR2046276	SRR2046278		
N10 control (for Ab10I-MMR)	PRJNA285341	SRR2046282	SRR2046286	SRR2046287		
N10 control (for Ab10II-MMR)	PRJNA285341	SRR2046290	SRR2046293	SRR2046294		
N10 control (for Ab10-II-Tel)	PRJNA285341	SRR2046418	SRR2046419	SRR2046421		
N10 control (for K10L2)	PRJNA285341	SRR2046429	SRR2046448	SRR2046445		
Ab10I-Pue	PRJNA285341	SRR2046758	SRR2046763	SRR2046764		
N10 control (for Ab10I-Pue)	PRJNA285341	SRR2046765	SRR2046767	SRR2046766		
Ab10-I-Jal	PRJNA339442	SRR4041256	SRR4042223	SRR4042225		
Ab10-II-Sal	PRJNA339442	SRR4044254	SRR4044255	SRR4044256		
Ab10-III-Caq	PRJNA339442	SRR4044269	SRR4044270	SRR4044271		
Ab10-III-Gua	PRJNA339442	SRR4044272	SRR4044273	SRR4044284		
Ab10-III-Hui	PRJNA339442	SRR4044382	SRR4044384	SRR4044388		
Ab10-III-Oax	PRJNA339442	SRR4044579	SRR4044581	SRR4044586		
N10-Rnj	PRJNA339442	SRR4042715	SRR4043650	SRR4043652		

**A**



**B**



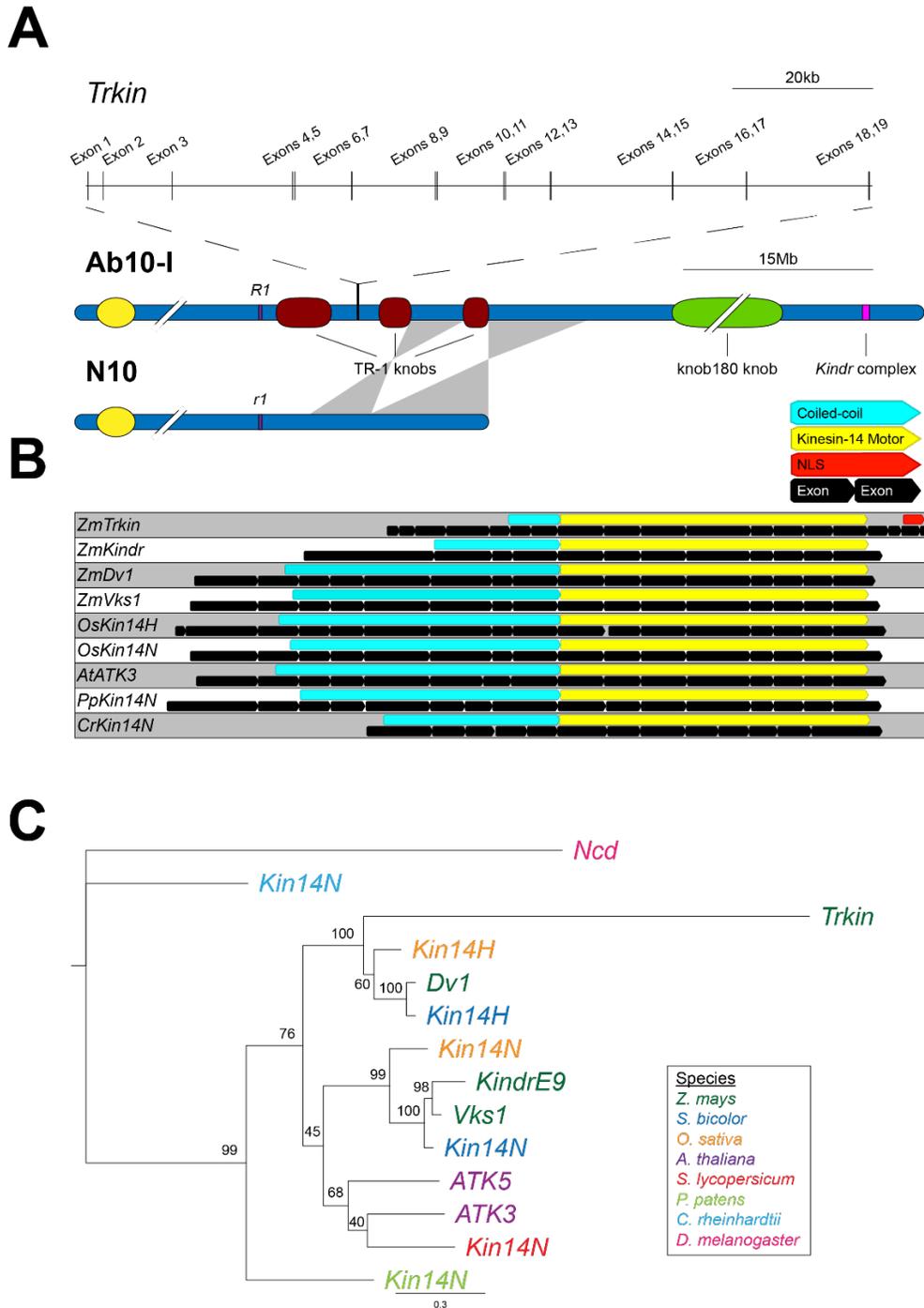
**Figure 3.1. Ab10 types and the Rhoades model for meiotic drive.**

(A) Schematic of known maize chromosome 10 haplotypes where green and red indicate knob180 and TR-1 knobs, respectively. *R*, *W*, *O*, *L*, and *S* denote relative positions of genetic

markers, signifying known inversions that exist between N10 and Ab10.

(B) The Rhoades model for neocentromere-mediated meiotic drive (Rhoades and Others 1952).

Centromeres are shown in yellow, knob180 in green and TR-1 in red. Recombination must occur between centromeres and knobs to create heteromorphic dyads, where one chromatid contains a knob and the other does not. The extreme poleward orientation of knobs is maintained through meiotic interphase, and in meiosis II, knobs again move towards the upper and lower cells of the linear tetrad. Since the basal cell becomes the megaspore (progenitor of the egg), knobs are preferentially transmitted.



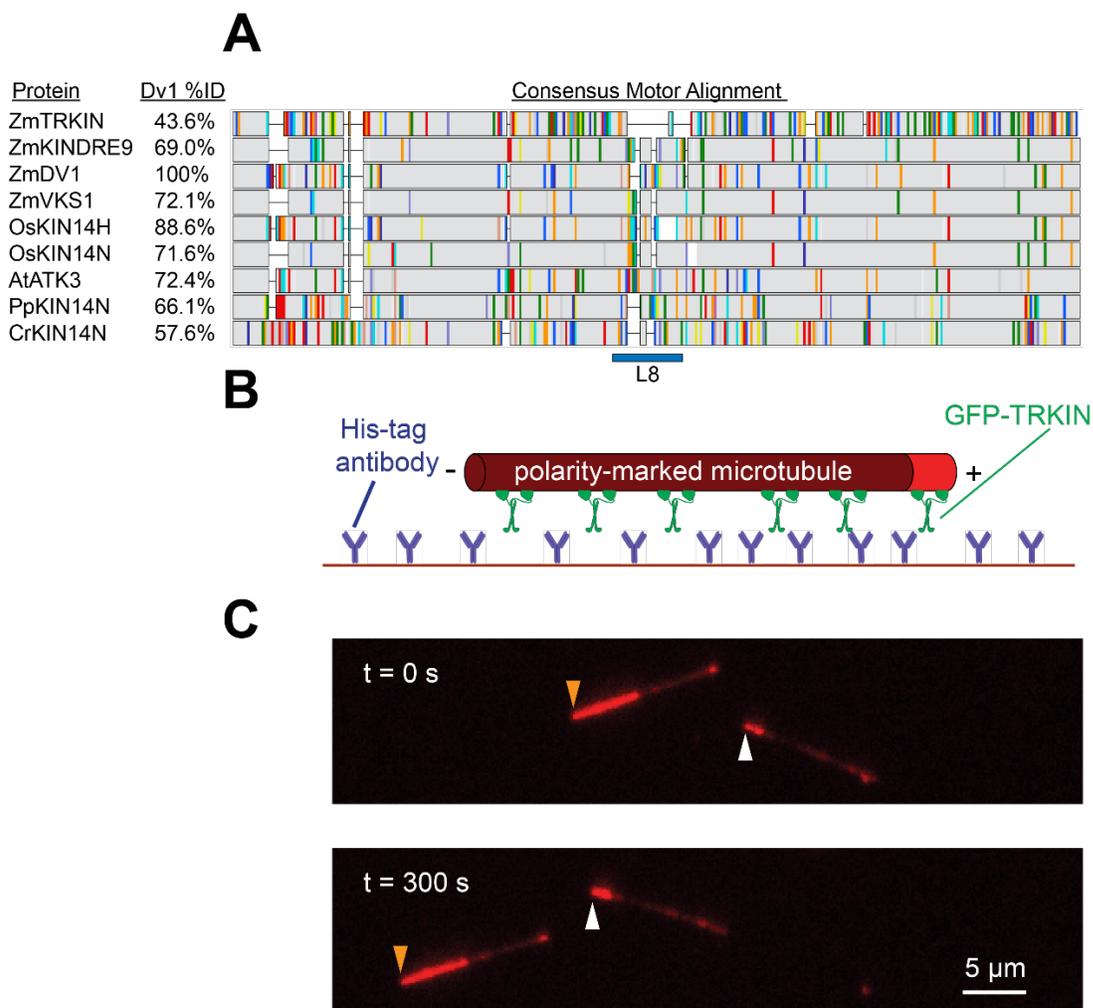
**Figure 3.2. *Trkin* and its divergence from other kinesin-14s.**

(A) Comparison of Ab10-I-MMR and normal chromosome 10 from B73 (v5 assembly). Regions

of shared, inverted synteny between N10 and Ab10 are highlighted in grey. The ~130 kb *Trkin* gene is expanded above Ab10. The *Kindr* complex, highlighted in purple, contains nine *Kindr* genes in tandem.

(B) Alignment of the *Trkin* gene and selected orthologs. The black boxes show exon boundaries. Predicted protein domains are also highlighted, including coiled-coil (cyan), kinesin-14 motor (yellow), and nuclear localization signal (red).

(C) Phylogeny of *Trkin* and related kinesin-14s motor domains. Bootstrap values are shown at nodes.



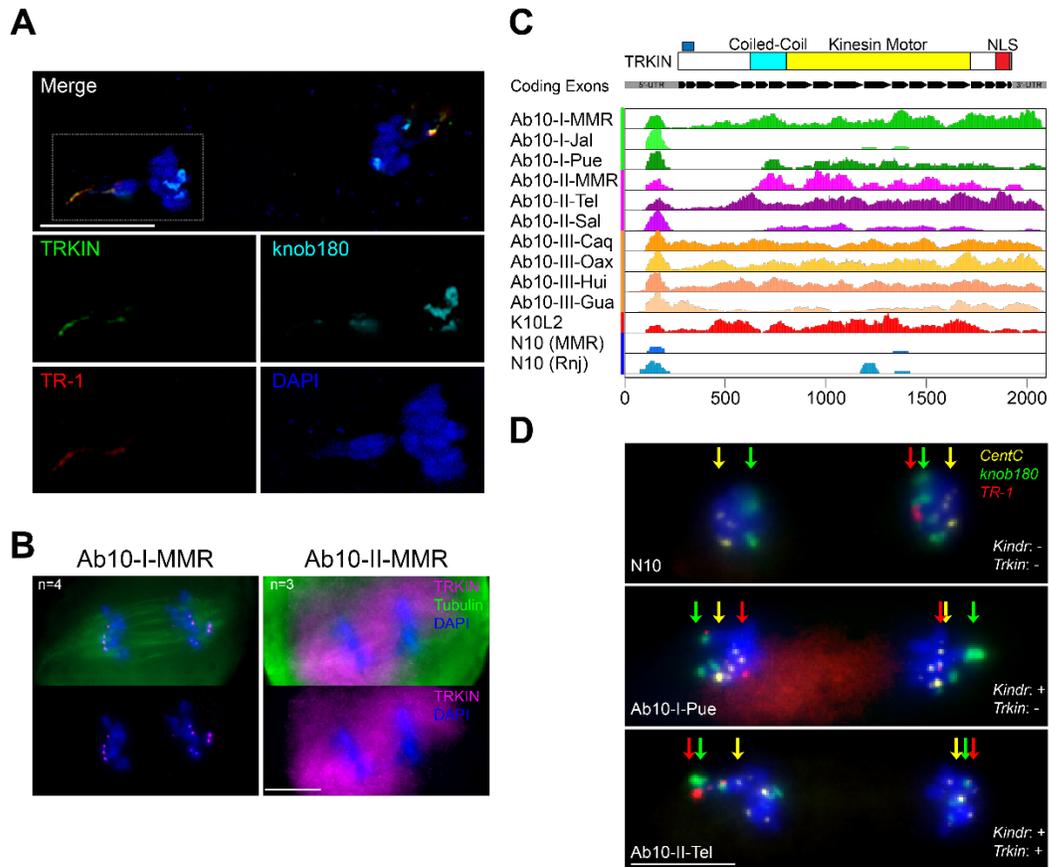
**Figure 3.3. Conservation and functionality of the TRKIN motor.**

(A) Alignment of the TRKIN motor domain and selected orthologs. Percent identity of each protein to ZmDV1 is indicated on the left. Sequence matches (in grey boxes), deletions (in black lines), and substitutions (in RasMol colors) are shown relative to the consensus of all proteins in the alignment. The position of the loop-8 motif is highlighted in blue.

(B) Schematic of the microtubule-gliding assay. His-tag antibodies (blue) adhere to the microscope slide surface and immobilize GFP-TRKIN proteins. Motor activity moves the

polarity-marked microtubules along the slide.

(C) Representative images showing that surface-immobilized GFP-TRKIN causes polarity-marked microtubules to move with the bright plus ends leading; yellow and white arrowheads indicate the plus ends of two different microtubules.



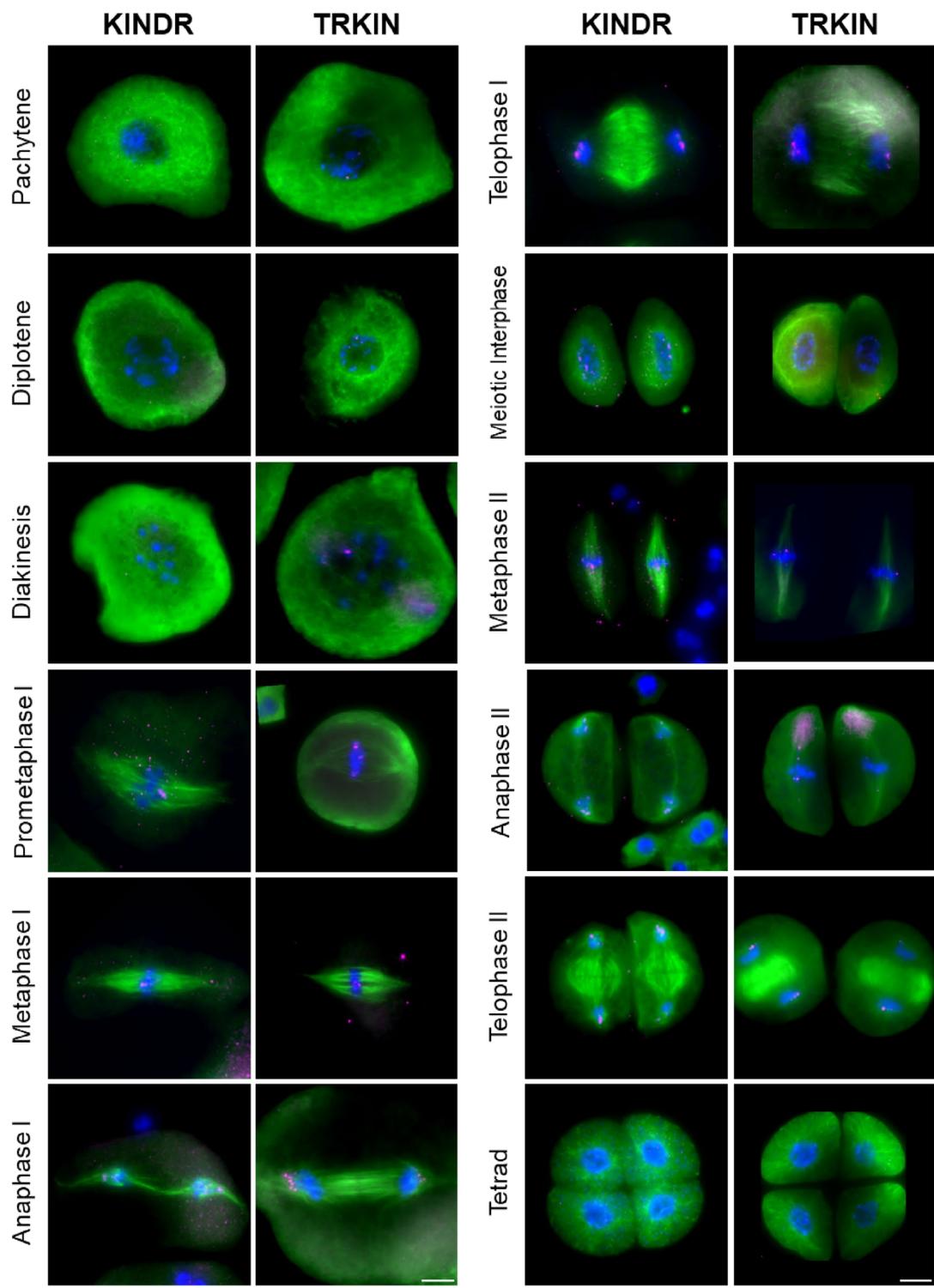
**Figure 3.4. TRKIN localization and presence/absence in different Ab10 types.**

(A) Visualization of TRKIN during anaphase II by structured illumination microscopy. TRKIN protein immunolocalization (green) and TR-1 (red) and knob180 (cyan) sequences labeled by FISH in a line homozygous for *Trkin* but lacking *Kindr*. The image is representative of results obtained from eight different plants.

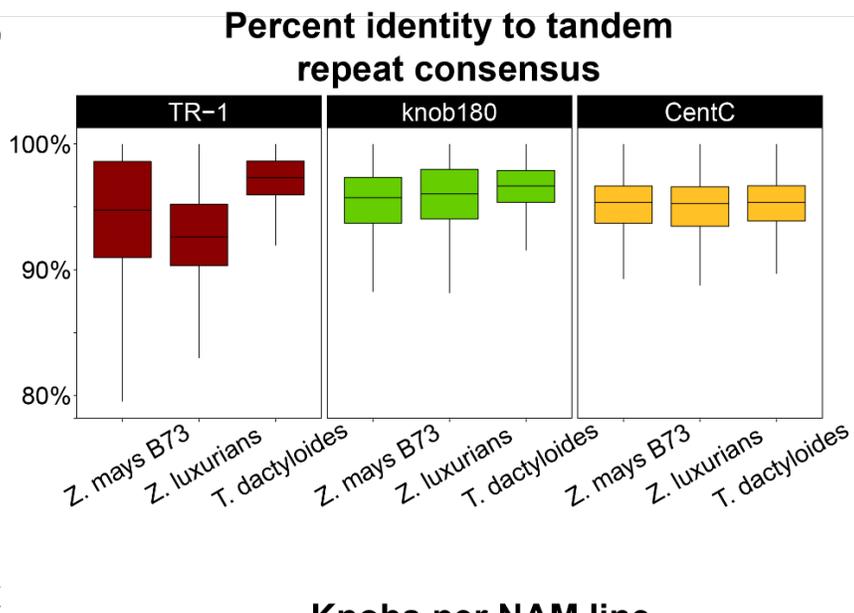
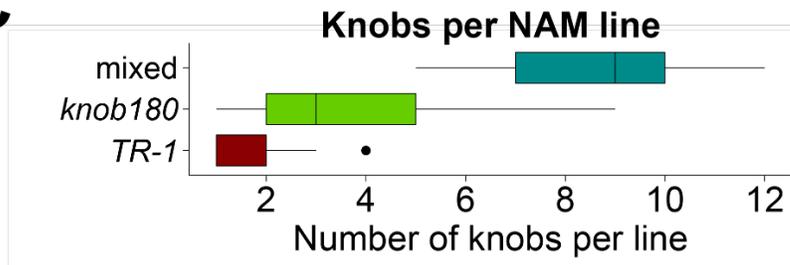
(B) Localization of TRKIN (magenta) and tubulin (green) in Ab10-I-MMR and Ab10-II-MMR male meiocytes during anaphase II. Note the absence of specific staining in the Ab10-II-MMR line. The images shown are representative of results obtained from four (Ab10-I) and three (Ab10-II) plants.

(C) Distribution of mRNA-seq reads from various libraries aligned to the *Trkin* transcript. The predicted TRKIN protein and coding exons are shown, along with the location of the peptide used to prepare antibodies (blue) for reference. Track colors reference cytologically distinguishable haplotypes: Ab10-I (green), Ab10-II (purple), Ab10-III (orange), K10L2 (red), and N10 (blue). Vertical black lines indicate mismatched bases. The Y axes are scaled relative to the read count in the various samples and should not be interpreted as absolute values.

(D) Correlation between *Trkin* and TR-1 neocentromeres in different genotypes. Merged FISH images show CentC (yellow), knob180 (green), and TR-1 (red) during anaphase II. Knob positions are denoted by colored arrows. TR-1 neocentromeres generally move faster than knob180 neocentromeres, and neocentromeres of both types arrive at the poles before centromeres. The images are representative of results obtained from analyzing three plants for each genotype. All scale bars are 10um.



**Figure 3.5. Immunolocalization of TRKIN and KINDR in male meiosis.** TRKIN and KINDR are shown in magenta and tubulin is shown in green. The stages of meiosis are indicated at left. The images shown are representative of data collected from eight plants. The TRKIN staining at spindle poles in prometaphase I is background staining and was not observed in all images. Scale bars are 10um.

**A****B****C**

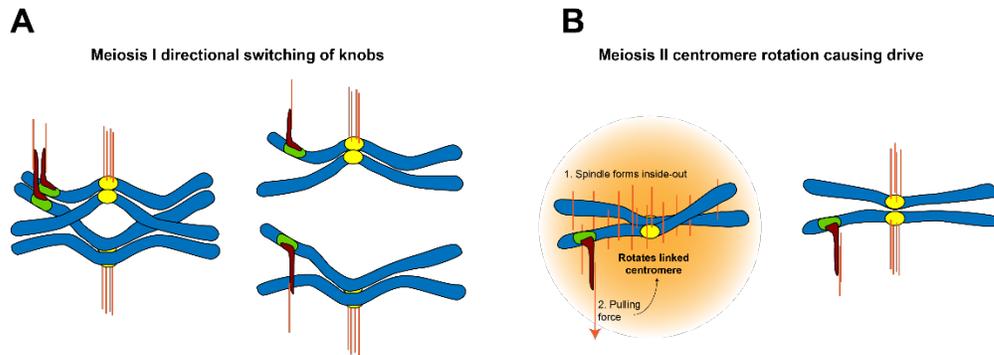
**Figure 3.6. TR-1 sequence diversity and distribution.**

(A) Multiple sequence alignment of TR-1 consensus sequences made from *T. dactyloides*, *Z. luxurians*, and *Z. mays* (B73). Vertical colored bars indicate SNPs and horizontal lines indicate deletions relative to the consensus made by the three sequences.

(B) Sequence variation within repeat families. Illumina reads were mapped to the consensus

sequences and the distribution of percent identities plotted as boxplots.

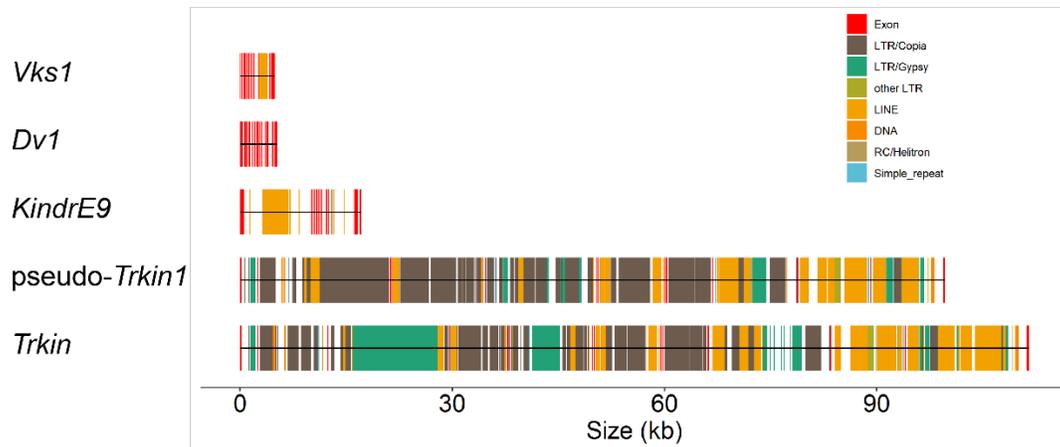
(C) The number of mixed (knob180 and TR-1), knob180-only, and TR-1-only repeat arrays in B73 and 25 other sequenced maize genomes plotted as boxplots. For boxplots in B and C, the bounds of boxes indicate first and third quartiles and whiskers show the distribution of 1.5 interquartile range.



**Figure 3.7. Possible TR-1 neocentromere functions.**

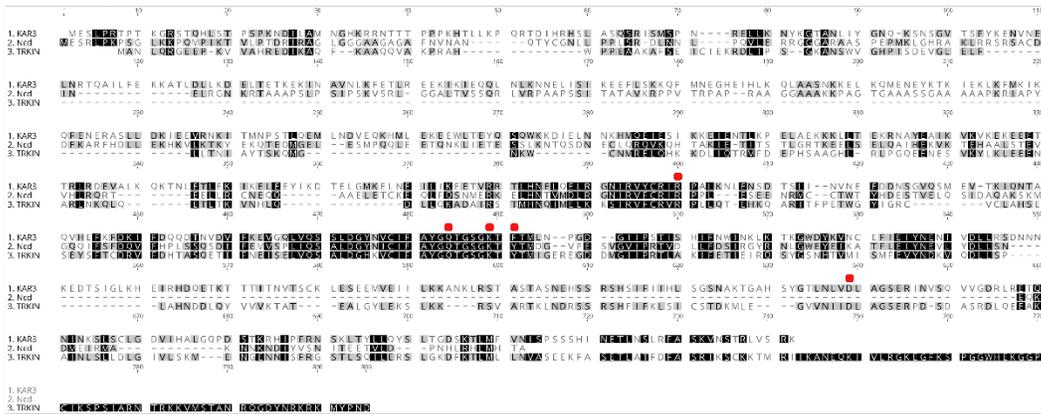
(A) Meiosis I directional switching. During meiosis I (left), sister chromatids remain attached along their entire lengths by cohesin (Dawe 1998; Buonomo et al. 2000). Because they are physically attached, the knobs on sister chromatids are likely to migrate towards one pole, creating a knob-centromere conflict on one pair of homologs. The unusual properties of TRKIN may help resolve these conflicts by facilitating a directional switch.

(B) Meiosis II centromere rotation. Neocentromeres in meiosis I leave knobs near the nuclear periphery during interphase, already prepositioned to move towards the basal cell of the meiotic tetrad (see Figure 3.1A). The meiotic spindle forms by a self-organization mechanism (Zhang and Dawe 2011). Microtubules are stabilized around chromatin where RanGTP (orange) concentration is highest, and the spindle grows out to form poles. We proposed that neocentromeres move with the forming spindle and rotate linked centromeres in the same direction before stable kinetochore-microtubule contacts are made (Yu et al. 1997; Dawe and Hiatt 2004). The fact that TR-1 neocentromeres move early and fast may facilitate this event.

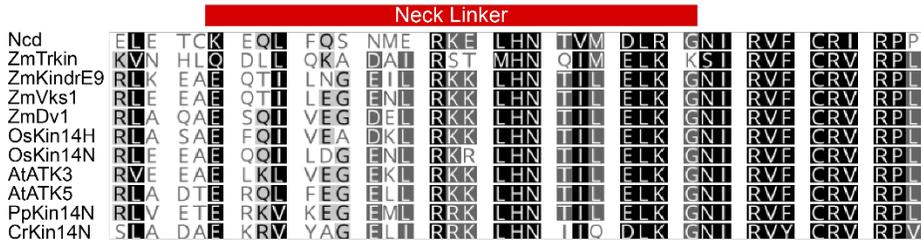


**Figure 3.8. Sizes and positions of exons and transposable elements in *Vks1*, *Dv1*, *KindrE9*, *pseudo-Trkin1* and *Trkin*.**

**A**



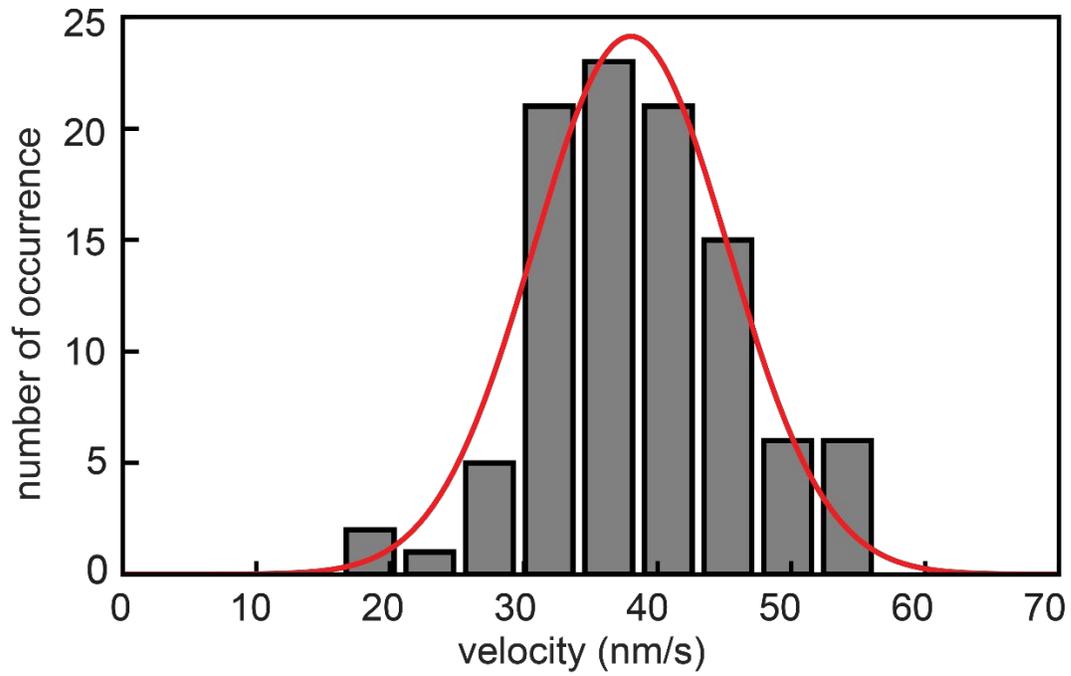
**B**



**Figure 3.9. Kinesin-14 protein alignments.**

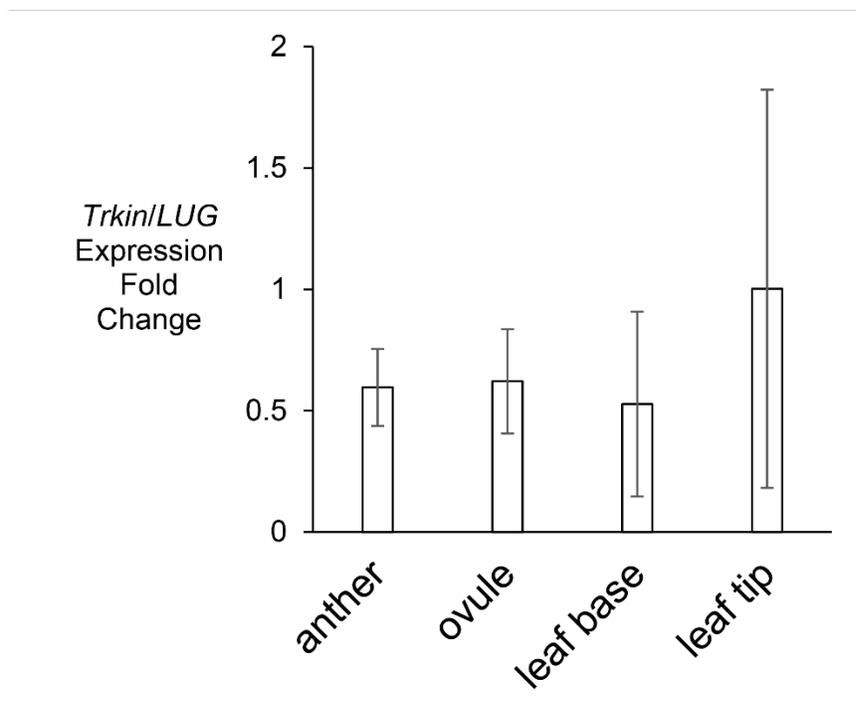
(A) Alignment showing conservation of the five ATP-binding residues (highlighted red) between yeast KAR3, *Drosophila* Ncd, and TRKIN.

(B) Alignment of kinesin-14 proteins in the predicted neck-linker region. Residues are generally conserved between plant kinesin-14s and *Drosophila* Ncd but are divergent in TRKIN. Darker colors indicate a higher degree of amino acid similarity.

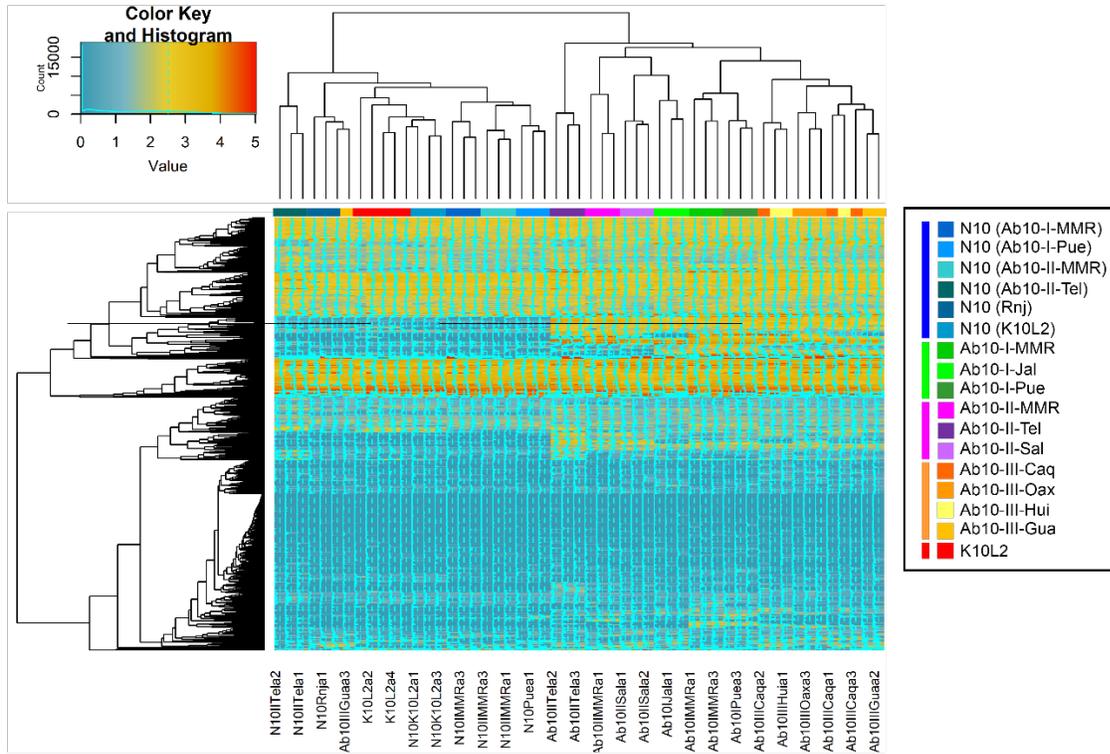


$$V_{\text{TRKIN}} = 38.0 \pm 0.7 \text{ nm/s (mean } \pm \text{ s.d.; } n = 100)$$

**Figure 3.10. Histogram showing the microtubule-gliding velocity distribution of GFP-TRKIN. Red line indicates a Gaussian fit to the velocity histogram.**



**Figure 3.11. Expression of *Trkin* as assayed by quantitative RT-PCR using cDNA from Ab10 reproductive and vegetative tissues.** Primers against LUG were used as a normalization control. Values are expression fold change relative to Ab10 leaf tip and error bars indicate standard deviation from three biological replicates. B73 was used as a negative control and values were zero in all cases.



**Figure 3.12. Clustering dendrogram of anther mRNA-seq data that align to the Ab10 haplotype from 135 Mb to the end of the chromosome. Colors indicate cytologically distinguishable haplotypes: Ab10-I (green), Ab10-II (purple), Ab10-III (orange), K10L2 (red), N10 (blue).**



## CHAPTER 4

### QTL MAP OF EARLY- AND LATE-STAGE PERENNIAL REGROWTH IN *ZEA DIPLOPERENNIS*<sup>2</sup>

---

<sup>2</sup> Swentowsky, Kyle W; Bell, Harrison S; Wills, David M; and Dawe, R Kelly. Manuscript to be submitted to *Theoretical and Applied Genetics*.

## ABSTRACT

Numerous climate change threats will necessitate a shift towards more sustainable agricultural practices during the 21st century. Conversion of annual crops to perennials that are capable of regrowing over multiple yearly growth cycles could help to facilitate this transition. Perennials can capture greater amounts of carbon and access more water and soil nutrients compared to annuals. In principle it should be possible to identify genes that confer perenniality from wild relatives and transfer them into existing breeding lines to create novel perennial crops. Two major loci controlling perennial regrowth in the maize relative *Zea diploperennis* were previously mapped to chromosome 2 (*reg1*) and chromosome 7 (*reg2*). Here we extend this work by mapping perennial regrowth in segregating populations involving *Z. diploperennis* and the maize inbreds P39 and HP301 using QTL-seq and linkage mapping approaches. The results confirmed the existence of a major perennial regrowth QTL on chromosome 2 (*reg1*). Although we did not observe the *reg2* QTL in these populations, we discovered a third QTL on chromosome 8 which we named *regrowth3* (*reg3*). The *reg3* locus exerts its strongest effect late in the regrowth cycle. Neither *reg1* nor *reg3* overlap with tiller number QTL scored in the same population, suggesting these loci specifically affect perennial regrowth and not tillering. Our data, along with prior work, indicate that perennial regrowth in maize is conferred by a few major QTL.

## INTRODUCTION

Plant life cycles fall into one of two categories. Annual species undergo their complete life cycle that involves germination, vegetative growth, reproduction, and senescence, in one year. Perennials exhibit these same growth stages, but do not fully senesce and are capable of new vegetative growth following senescence. Some perennial species are seasonally dormant while others stay green year round. Evolutionary transitions between annual and perennial growth modes have occurred numerous times during land plant evolution. Perenniality is usually the ancestral trait and the annual growth mode is derived (Friedman and Rubin 2015). For instance, a transition from perennial to annual growth may have allowed plants to survive seasonal stresses such as drought (Sherrard and Maherali 2006; Friedman and Rubin 2015). Taxa that allocate relatively more resources to above-ground organs are more likely to have evolved annuality (Lindberg et al. 2020).

Although our major cereal crops today are annuals, it has been hypothesized that growing perennial grains could result in sustainable agricultural benefits (Zhang et al. 2011; L. Fernando et al. 2018). Compared to annuals, perennials grow deeper, more established root systems that allow them to tap into water or nutrients further below the soil surface. Perennials can also emerge earlier in the growing season which allows them to capture more light (Dohleman and Long 2009). The increased photoperiod could theoretically lead to higher productivity at a given latitude while mitigating effects of climate change through more effective capture and sequestration of carbon dioxide.

New perennial cereal crops must be developed to recognize the sustainable benefits of perennial agriculture (Crews and Cattani 2018). One way to achieve this goal would be to select for domestication traits in populations of an existing perennial species. A new cereal

trademarked as Kernza® has been developed from the intermediate wheatgrass species *Thinopyrum intermedium* using this approach. Some of the hypothetical benefits of perennial agriculture have been achieved with Kernza® (de Oliveira et al. 2018; de Oliveira et al. 2020), but its yield is currently much lower than that of wheat (Culman et al. 2013). Perennial grains generally have lower yield potential than congeneric annuals (Vico et al. 2016), and will likely require intensive long-term selection to improve yields.

An alternative approach is to breed perenniality into existing annual species (Murray and Jessup 2014). It may be possible to leverage the already high yields of modern cultivars while bringing in alleles that provide some degree of perenniality. Several groups have initiated these efforts by focusing on the formation of rhizomes as a proxy for perennial regrowth. Rhizomes are below-ground organs derived from stems that can store resources used for regrowth. Trait mapping studies in *Oryza* (Hu et al. 2003; Fan et al. 2020), *Sorghum* (Paterson et al. 1995; Kong et al. 2015), and *Zea* (Westerbergh and Doebley 2004) revealed QTL associated with the presence of rhizomes in F<sub>2</sub> progeny. These are important first steps but it remains unclear whether lines bred for rhizome formation will demonstrate perennial regrowth. A recent study took a different approach of mapping genes that promote regrowth itself by crossing domesticated maize (*Zea mays* L.) to its close perennial ancestor, *Zea diploperennis* (Iltis et al. 1979; Ma et al. 2019). Regrowth is generally scored as the presence of new branches at the base of the plant after the plant has flowered and senesced. The authors found two *Zea* regrowth QTL termed *regrowth1* (*reg1*) and *regrowth2* (*reg2*) and noted that these are distinct from previously found rhizome QTL in *Z. diploperennis*. The authors also noted that rhizomes are almost never found in their regrown plants and that regrowth can occur from dormant tiller buds.

Here we report the results from mapping regrowth in two F<sub>2</sub> populations between *Z. diploperennis* and maize inbred lines with high tillering. We found two genomic regions associated with regrowth using a modified bulk segregant analysis approach (Michelmore et al. 1991; Takagi et al. 2013) in combination with targeted PCR-based genotyping. One QTL maps to the same location as *reg1* (Ma et al. 2019), while a novel QTL maps to chromosome 8 and shows a strong effect in the later stages of regrowth.

## **MATERIALS AND METHODS**

### **Plant Materials**

*Zea diploperennis* ‘Gigi’ is the clone of a plant grown from seed obtained from Germplasm Resources Information Network (GRIN), Ames, Iowa (PI 462368). F<sub>2</sub> populations were generated by applying pollen from ‘Gigi’ to ears of a sweet corn line P39 (PI 587133) and a popcorn line Hp301 (PI 587131). Approximately five F<sub>1</sub> progeny from each cross were sown in isolation and allowed to intermate freely to generate F<sub>2</sub> seeds.

### **Plant Growth Conditions and Phenotyping**

For greenhouse experiments, seeds were sown in square four-inch pots using a pine bark soil mixture supplemented with Osmocote slow-release fertilizer. Approximately one-third of seeds germinated. Pots with germinated seedlings were placed close together on top of a coconut mat. They were watered daily and roots were allowed to grow into the coconut mat. The first experiment (for QTL-seq) used 496 P39/‘Gigi’ F<sub>2</sub>s and began in September 2018. A second experiment (for QTL mapping using PCR primers) involved 196 P39/‘Gigi’ F<sub>2</sub>s and 419 Hp301/‘Gigi’ F<sub>2</sub>s and began in September 2019. In both cases, plants flowered in November and

December, then began to senesce. Fully-senesced (brown) stems were cut back to the soil line. Positive regrowth was scored by the presence of green leaves on young tillers. Regrowth was scored twice in the QTL-seq experiment during April and August 2019. In the subsequent QTL mapping experiment, regrowth was scored on four dates in 2020: January 23, February 19, March 24, and May 6.

Field-grown plants were initially sown in pine bark soil in September 2019 and 227 P39/'Gigi' F<sub>2</sub>s were transplanted into a field site in Athens, GA approximately two weeks after sowing. Plants were treated with Osmocote slow-release fertilizer and kept watered. Tiller number and regrowth were scored on December 1, 2019 before a freeze of 32°F on December 2 killed all plants.

#### DNA Extraction and Illumina Library Preparation/Sequencing

Regrowth (RG) and non-regrowth (NRG) bulks were prepared by pooling leaves of equal sizes from 90 and 30 plants per bulk, respectively. Tiller number bulks were created by pooling leaves of equal sizes from 26 plants per bulk using plants with 0-2 tillers (low tiller) and 9-15 tillers (high tiller). Genomic DNA from 'Gigi' was extracted from a single leaf using the Genomic DNA Mini Kit (Plant) (IBI, cat no. IB47231), and libraries for Illumina sequencing were generated using the KAPA HyperPrep Kit for NGS (Roche, cat no. KK8502). The NRG bulk library was sequenced twice on two separate flow cells and the resulting reads were combined for analysis. All reads used in this study were deposited in the NCBI Sequence Read Archive under BioProject PRJNA700589.

## SNP Calling and QTL-Seq Analysis

The GATK Best Practices Pipeline for germline short variant discovery was applied to call SNPs and determine SNP-indices from Illumina data (DePristo et al. 2011; Poplin et al. 2017). Briefly, reads were mapped to the P39 reference (Hufford et al. 2021) using BWA MEM v0.7.15 (Li 2013), sorted using SAMtools sort v1.3.1 (Li et al. 2009), and duplicate reads were marked using Picard (<http://broadinstitute.github.io/picard/>). HaplotypeCaller (GATK/4.0.11.0) using default settings followed by GenotypeGVCFs were used to call SNPs from 'Gigi', RG, and NRG. SelectVariants was used to filter SNPs based on all of the following criteria: homozygous in 'Gigi' with depth (DP) of at least three reads; GenotypeQuality (GQ) of at least 99 in RG and NRG. The filtered VCF file was exported using the VariantsToTable command for analysis using RStudio v1.2.1335. SNPs for high tillered (HT) and low tillered (LT) bulks were called independently with the same workflow and only homozygous 'Gigi' SNPs were retained for SNP-index analysis.

The allele frequency of each SNP in each bulk was calculated using RStudio by dividing the number of reads supporting a SNP by the number of reads covering the position of the SNP (i.e. AD/DP). The genome was divided into 1-Mb bins and the mean SNP-index was calculated for each bin. The SNP-index and  $\Delta$ SNP-index (RG - NRG or HT - LT) were plotted according to their respective genomic locations using ggplot2 (Wickham 2016).

## PCR Genotyping

Genomic DNA was extracted from flash-frozen and ground leaf material using a modified CTAB method described by (Saghai-Marooof et al. 1984). DNA pellets were suspended in 100  $\mu$ l TE.

PCR genotyping for individual F<sub>2</sub> plants was performed using GoTaq Green Master Mix (Promega, cat no. M7123) diluted to 1X with 1 µl of genomic DNA and a final primer concentration of 1 µM. Manufacturer recommendations for PCR cycling conditions were used with an annealing temperature of 56°C, an extension time of 30 seconds per cycle, and 30 cycles total. We used a draft 'Gigi' genome assembly (generously provided by Matthew Hufford, Iowa State University) to design primers that capture indels or SNPs on chromosomes 2, 7, and 8 that differentiate 'Gigi' from P39 or Hp301. A total of 22 and 20 markers were designed to capture *Z. diploperennis* polymorphisms against P39 and Hp301, respectively (Supplementary File 1). The majority of markers captured indels, however two were CAPS markers that required digestion with the restriction enzyme EcoRV-HF (NEB, cat no. R3195T). 10 µl restriction digests were performed using 5 µl of PCR products and 1 µl of EcoRV-HF for three hours at 37°C. A full list of primer sequences and genomic positions are provided in Supplementary File 1.

For the 2018 greenhouse trial, 96 regrown P39/'Gigi' F<sub>2</sub>s of the 496 total were genotyped for three markers. Data was obtained for DCP23/DCP24 on chromosome 2 (92 plants), DPP7/DPP8 on chromosome 7 (94 plants), and DCP84/DCP85 on chromosome 8 (96 plants). For the 2019 greenhouse trial, we genotyped all 196 of the P39/'Gigi' F<sub>2</sub> plants, however we only genotyped 40 regrown and 152 non-regrown individuals from the population of 419 Hp301/'Gigi' F<sub>2</sub> plants. For the 2019 field trial, we genotyped 35 regrown and 71 non-regrown plants from the population of 227 P39/'Gigi' F<sub>2</sub> individuals.

## QTL Analysis

QTL analysis using PCR genotyping data was carried out using the R package R/qtl (Broman et al. 2003). Genotypic and phenotypic data for the P39/'Gigi' F<sub>2</sub> population

(Supplementary File 3) or Hp301/'Gigi' F<sub>2</sub> population from the 2019 greenhouse trials (Supplementary File 4) were imported using the `read.cross()` function with the option `map.function="kosambi"`. Individuals with data for fewer than 11/22 (P39/'Gigi') or 8/20 (Hp301/'Gigi') total markers were removed. Segregation distortion was analyzed using the `geno.table()` function, however, no markers showing significant segregation distortion over the chromosome 2, 7, and 8 intervals were found. The markers were placed in the order they occur in the P39v1 genome assembly (Hufford et al. 2021) and genetic map distances were estimated with the `est.rf()` function. Individuals showing excessively large numbers of crossovers (P39/'Gigi' > 7 and Hp301/'Gigi' > 6) were removed from the analysis. With these individuals removed, the genetic map was calculated again using the `est.rf()` function. Linkage between genotypic data and phenotypes (regrowth scored during January, February, March, and May) was calculated using the `scanone()` function with options `method="em"` and `model="binary"`. LOD score data for each marker were organized into a new data frame that incorporated each marker's physical position and plotted using `ggplot2` (Wickham 2016).

## RESULTS

### Greenhouse studies of regrowth

In the first experiment, planted in September of 2018, 496 plants flowered in November-December and the above-ground parts of the plants senesced and turned brown. A total of 203 (41%) plants grew fresh stems that flowered in early 2019. Another round of senescence occurred in spring of 2019, followed by regrowth in 104 (21%) of the original 496 plants (Table 1). After the second flowering, bulks were prepared from frozen leaf tissue of 90 regrown (RG)

and 30 non-regrown (NRG) individuals and sequenced using Illumina short reads to ~15X coverage (Supplementary Table 1). Reads were aligned to the P39 assembly (Hufford et al. 2021) along with sequenced genomic DNA from the *Z. diploperennis* parent 'Gigi', and 34,235,573 raw SNPs were called using the GATK4 Best Practices pipeline (DePristo et al. 2011). Strict filtering based on coverage, quality scores, and homozygosity in 'Gigi' resulted in 2,180,254 SNPs (Supplementary Table 2, Supplementary Figure 1). The SNP-index, defined by the number of reads containing a particular SNP divided by the total number of reads spanning the SNP site, was calculated for all SNPs in the two bulks.

Two genomic regions displayed a higher SNP-index in the RG compared to the NRG bulk (Figure 2). A large difference in SNP-index is observed broadly across chromosome 2 with a peak on the short arm which corresponds to the dominant QTL previously identified as *regrowth1* (*reg1*) (Ma et al. 2019). The second is located on the long arm of chromosome 8 with the largest difference near position 150 Mb on the P39 physical map and will be referred to as *regrowth3* (*reg3*). The *regrowth2* locus on the short arm of chromosome 7 (Ma et al. 2019) did not show significant differences in SNP-index between the two sequenced bulks.

The bulk sequencing results were confirmed on individual plants by designing codominant PCR markers spanning polymorphisms near each of the SNP-index peaks. The markers were scored in a sample of 96 regrown plants from the same 2018 population. The *Z. diploperennis* markers for *reg1* and *reg3* were represented at frequencies significantly higher than Mendelian expectations. An excess of heterozygotes was observed at both loci indicating dominant inheritance patterns (Table 2). There was also a low number of RG individuals that were homozygous for the P39 allele suggesting that *reg1* and *reg3* are not fully penetrant or that other loci can compensate for their function. Only 1 of 92 RG individuals positively genotyped at

these loci were homozygous for the P39 allele of both *reg1* and *reg3*. The marker near *reg2* on chromosome 7 segregated at a 1:2:1 Mendelian ratio, consistent with the SNP-index data indicating that *reg2* does not contribute to regrowth in the P39/'Gigi' population (Table 2). These results support the existence of two *Z. diploperennis* alleles that co-segregate with RG plants in our population: the previously described *reg1* and the novel *reg3*.

We planted a second population in September 2019, but instead of a single late observation, plants were scored for green tissue at four time points: once in January, February, March, and May. This time course revealed that early regrowth does not necessarily translate to continued regrowth. A number of plants that showed evidence of regrowth early in the season ultimately senesced and died. In January, we scored 42.3% of the plants as regrown, whereas in May, only 25% of the plants were still flourishing (Table 3). PCR genotyping revealed that markers linked to *reg2* on chromosome 7 showed no significant association with regrowth at any time point (Figure 3). However, we observed significant associations between regrowth and markers on chromosomes 2 (*reg1*) and 8 (*reg3*). The highest linkage was between regrowth scored in January and a marker on chromosome 2 (at 44.6 Mb) with a LOD of 4.10. Association between regrowth and the chromosome 2 marker dropped during subsequent sampling points, and there was no significant linkage detected at the latest time point in May. In marked contrast, a marker on chromosome 8 at 139.6 Mb became more significantly associated with regrowth at later sampling points, such that significant LOD scores were only observed in March and May (with a LOD of 3.37). These data suggest that *reg1* is important for the initiation of regrowth while *reg3* has a more significant impact later in the growth cycle.

The second wave planting also included plants from the Hp301/'Gigi' F<sub>2</sub> mapping population. The observed regrowth rate was considerably lower for this group, ranging from 6.4

to 10.3% (Table 3). A total of 192 Hp301/'Gigi' F<sub>2</sub> individuals were genotyped, including 40 plants that showed regrowth during at least one time point. Significant associations were not observed between regrowth and any markers on chromosomes 7 or 8, suggesting that neither *reg2* nor *reg3* segregate in the Hp301/'Gigi' F<sub>2</sub> population (Figure 4). In contrast, several chromosome 2 markers were significantly associated with regrowth, with the highest LOD score of 3.95 between the marker at 44.6 Mb and regrowth scored in May. We observed the highest trait-marker association in the Hp301/'Gigi' population when regrowth was scored in May, second highest in January, third highest in February, and non-significant association in March (Figure 4; Table 3).

#### Field studies of regrowth and tillering

An outdoor field planting of the P39/'Gigi' population was carried out in the fall of 2019, both to confirm the mapping data and to test whether the plants could survive a winter freeze and dormant season. A total of 228 P39/'Gigi' F<sub>2</sub> individuals were planted in September, and flowered and began to senesce by mid to late-November before a freeze in early December. At the time of the freeze, 38/228 (16.7%) of the plants had regrown. The 35 regrown and 71 non-regrown plants were genotyped for a single marker on chromosomes 2, 7, and 8 (Table 4). The only significant deviation from Mendelian expectations was observed on chromosome 2 (*reg1*), where regrown plants were significantly enriched for the 'Gigi' allele. This is consistent with our greenhouse studies, since *reg3* was only observed later in the growth cycle. The field site was monitored for the next five months, but no plants survived to regrow the next spring.

Because regrowth in *Z. diploperennis* occurs through the re-activation of tiller buds, a potential explanation for our results is that *reg1* and *reg3* simply promote strong tillering. To

address this possibility, we performed a QTL-seq experiment utilizing high and low tiller number bulks from the field plot (Figure 5, Supplementary Figure 2). We obtained paired-end 150 reads at ~20X coverage for each bulk. After SNP calling, we found the strongest enrichment of 'Gigi' alleles in the high tiller bulk were observed on chromosomes 2 and 3, where multiple large-effect QTL have previously been mapped for tiller number (Westerbergh and Doebley 2004; Chen et al. 2019, 2020). The region on chromosome 2 corresponding to *reg1* (44.6Mb) showed a relatively low difference in SNP-index between high and low tiller bulks (compare Fig. 2A to Fig. 5). There is also a tiller number QTL on chromosome 8, however it is located at ~164 Mb, a significant distance from the *reg3* peak at 139.6 Mb. These data suggest that *reg1* and *reg3* are distinct from the major loci that control tiller number.

A QTL corresponding to the known tiller number gene *tin1* on the short arm of chromosome 7 was not detected in our QTL-seq analysis. Prior data show that the P39 inbred used in our study contains the high-tiller *tin1* allele whereas the B73 inbred used by (Ma et al. 2019) has the low-tiller *tin1* allele (Zhang et al. 2019). The *tin1* gene is located within the *reg2* interval so these data raise the possibility that the *reg2* locus (Ma et al. 2019), which we did not detect in our studies, is the tillering gene *tin1*.

## DISCUSSION

Previous genetic studies of perenniality in grasses have focused on morphological traits related to perenniality, most notably tiller and rhizome growth, and not the key phenotype of regrowth after senescence. Studies in rice (Hu et al. 2003; Fan et al. 2020), sorghum (Paterson et al. 1995; Kong et al. 2015), and maize (Westerbergh and Doebley 2004) have highlighted that rhizome development is a complex, multigenic trait that is influenced by the environment. While

rhizomes may be necessary for over-wintering in some grasses (Washburn et al. 2013), rhizomes alone are unlikely to promote regrowth. Many perennial grasses regrow year after year without forming obvious rhizomes (Derner and Briske, 2001). Further, it is possible that the genetics of regrowth is simpler than the genetics of rhizome formation. In support of this notion is *Thinopyrum elongatum*, which regrows from tillers alone (Lammer et al. 2004). When wheat lines containing individual *T. elongatum* chromosomes were screened for regrowth, just the addition of chromosome 4E was enough to confer perennial regrowth, implying that regrowth is under relatively simple, dominant genetic control (Lammer et al. 2004).

The first two perennial regrowth QTL in *Z. diploperennis* were named *regrowth1* and *regrowth2* and these loci have peaks near 33.0 Mb on chromosome 2 and 4.2 Mb on chromosome 7 using B73v4 coordinates (Ma et al. 2019). They were described as dominant, fully penetrant, and complementary to one another. However, the QTL corresponding to the *reg2* locus was not observed in either of our mapping populations (Figures 2-4, Tables 3, 5). One explanation is that there was no significant polymorphism for this trait in our crosses involving P39 and Hp301. A QTL for tiller number called *tin1* was recently mapped to a location near *reg2* on chromosome 7 (Zhang et al. 2019). The high-tiller allele of this QTL is present in most sweet corn and popcorn lines but absent from other lines such as the inbred B73. The *tin1* polymorphism is a SNP that affects a splice-site to increase transcript instability in a C2H2-zinc-finger transcription factor (Zhang et al. 2019). TIN1 controls tillering by directly repressing a previously identified tillering locus in maize, *grassy tillers1* (Whipple et al. 2011; Wills et al. 2013).

In contrast, our data support the existence of a QTL for regrowth, independent from tillering, that corresponds to *reg1*, located on chromosome 2 at 43.8 Mb (in B73v4 coordinates).

This locus is roughly 10 Mb away from the reported location of *reg1* in the B73 background (Ma et al. 2019). QTL positions are inherently approximations that are affected by population size, recombination rate, crossing partner, errors in genotyping and phenotyping, and effect size in a given environment. Other types of mapping populations such as RILs or recombinant substitution lines will be necessary to fine-map the position of *reg1*.

We also observed a QTL on chromosome 8 that we refer to as *regrowth3* (Figures 2-3, Table 2). Interestingly, *reg1* and *reg3* had their greatest impact at different stages of the growth cycle: *reg1* showed the highest LOD for QTL identified when regrowth was scored in the earlier months of January and February while LOD scores declined below the threshold of significance when scoring was done in May. In contrast, *reg3* showed its lowest LOD for January scorings and reached its highest point in May scorings. The fact that *reg3* exerted its strongest effects when regrowth was scored about six months after sowing could explain why it was not identified previously, since (Ma et al. 2019) scored regrowth soon after senescence. Similarly, we only scored early regrowth in our field study, and detected *reg1* but not *reg3*. A plausible scenario is that *reg1* promotes the initiation of regrowth while *reg3* contributes to the continuation of regrowth.

We observed a lower percentage of regrown plants than what was reported in prior work by Ma et al. (2019). These authors generated F<sub>2</sub> populations using the maize inbred B73 that does not tiller and the heirloom landrace Rhee Flint that develops numerous tillers. They reported 60% (B73-Zd) and 57% (Rhee Flint-Zd) regrowth following flowering and senescence, while we observed regrowth frequencies of 43% and 42% (P39/'Gigi' - greenhouse), 8% (Hp301/'Gigi' - greenhouse), and 17% (P39/'Gigi' - field) when scored at a similar developmental stage. Genotype-by-environment (GxE) interactions are well-known to play significant roles in

determining plant phenotypes (Zeng et al. 1999; Yadav et al. 2003; El-Soda et al. 2014; Buescher et al. 2014; Frey et al. 2016; Li et al. 2018; AlKhalifah et al. 2018; McFarland et al. 2020).

Unless the environmental contribution is excessively large, the same QTL are typically identified when grown in different environments (Buckler et al. 2009; Peng et al. 2011; Liu et al. 2014).

Greenhouse and field study sites differ dramatically in a plant's access to water, nutrients, microbes, growth space, and light and these factors likely contribute to the differences in regrowth rate we observed. The maize parents used in the cross also had a major impact on the regrowth phenotype, as evidenced by the differences in regrowth observed in the P39 and Hp301 backgrounds (Table 3). Finally, *Z. diploperennis* is an outcrossing species and it is possible that the *Z. diploperennis* alleles in our crosses were not identical to those in the Ma et al (2019) study although the same accession was used.

Our results show that *regrowth* QTL are distinct from those that regulate tiller number (Figures 4.2, 4.5), which mirrors the conclusion reached by (Ma et al. 2019) that there was no significant association between tiller number and regrowth. Nevertheless, some degree of tillering is a prerequisite for perenniality (Galinat 1981; Westerbergh and Doebley 2004). We hypothesize that *tin1* may correspond to *regrowth2*, and be an important factor in the tillering pathway that allows perennial regrowth in *Z. diploperennis*. Branching architecture is also regulated by numerous other developmental and physiological cues, including plant age and carbon status. For instance, the microRNA miR156 is known to regulate juvenility in angiosperms (Wu et al. 2009) and to promote branching through a module affecting *SPL15*, *Tb1*, and *Gt1* (Chuck et al. 2007; Liu et al. 2017). The perennial growth habit requires coordinated recurring switches between juvenile and adult forms, which are likely reflected in their miR156 levels. In *Arabis alpina*, a perennial closely related to Arabidopsis, miR156 levels remain high in

some axillary meristem buds as plants age, which may affect the branching fate of these buds (Park et al. 2017). Sugar availability also positively affects tiller bud growth (Mason et al. 2014; Fichtner et al. 2017; Wang et al. 2020). In perennial grasses, mobilization of remaining sugars to the below-ground organs following flowering has been well documented (Komor 2000; Purdy et al. 2015). Trehalose 6-phosphate (T6P) is derived from sucrose and serves as a signal of available sucrose levels (Figuroa and Lunn 2016). Induction of axillary bud growth by sucrose is mediated by T6P (Fichtner et al. 2017; Dong et al. 2019) suggesting that the coordination of sucrose transport may play an important role in activating and/or maintaining regrowth following senescence.

What types of genes might underlie *regrowth* QTL? Proper spatiotemporal control of tiller regrowth for a perennial life strategy must utilize genes that respond to seasonal and physiological cues. Tillers form from basal axillary buds, and we understand much about how these buds are suppressed. A key regulator of tiller bud suppression is the TCP transcription factor encoded by *Teosinte branched-1* (*Tb1*) that dominantly initiates axillary bud dormancy (Doebley et al. 1995; Dong et al. 2019). *Tb1* coordinates cues from light quality (Kebrom et al. 2006), nutrient (Mason et al. 2014; Dong et al. 2019), age (Liu et al. 2017), and phytohormonal (Wang et al. 2019; Dong et al. 2019) pathways, suggesting many points where *Tb1* or downstream genes in this pathway may be modulated in the context of perennial regrowth. Although *Tb1* (on chromosome 1) was not detected as a QTL in our study, *Tin1* promotes tiller bud growth by repressing *grassy tillers-1*, a downstream component of the *Tb1* pathway (Zhang et al. 2019). Future studies will focus on identifying *reg1* and *reg3*, and additional *regrowth* QTL to better understand the developmental processes of axillary bud dormancy and activation.

## ACKNOWLEDGEMENTS

We would like to thank Mike Boyd and Kevin Turner from the UGA greenhouse staff for their support on this project. In addition, we thank Katrien Devos and Esther van der Knaap for providing thoughtful comments on the manuscript, and Arun Seetharam and Matthew Hufford for sharing their preliminary *Z. diploperennis* assembly. We relied on computational resources and technical expertise from the Georgia Advanced Computing Resource Center, a partnership between the University of Georgia's Office of the Vice President for Research and Office of the Vice President for Information Technology.

## REFERENCES

- AlKhalifah N, Campbell DA, Falcon CM, et al (2018) Maize Genomes to Fields: 2014 and 2015 field season genotype, phenotype, environment, and inbred ear image datasets. *BMC Res Notes* 11:452
- Broman KW, Wu H, Sen S, Churchill GA (2003) R/qtl: QTL mapping in experimental crosses. *Bioinformatics* 19:889–890
- Buckler ES, Holland JB, Bradbury PJ, et al (2009) The Genetic Architecture of Maize Flowering Time. *Science* 325:714–718
- Buescher EM, Moon J, Runkel A, et al (2014) Natural variation at sympathy for the ligule controls penetrance of the semidominant Liguleless narrow-R mutation in *Zea mays*. *G3* 4:2297–2306
- Chen Q, Samayoa LF, Yang CJ, et al (2020) The genetic architecture of the maize progenitor, teosinte, and how it was altered during maize domestication. *PLoS Genet* 16:e1008791
- Chen Q, Yang CJ, York AM, et al (2019) TeoNAM: A Nested Association Mapping Population for Domestication and Agronomic Trait Analysis in Maize. *Genetics* 213:1065–1078
- Chuck G, Cigan AM, Saeteurn K, Hake S (2007) The heterochronic maize mutant *Corngrass1* results from overexpression of a tandem microRNA. *Nat Genet* 39:544–549
- Crews TE, Cattani DJ (2018) Strategies, Advances, and Challenges in Breeding Perennial Grain Crops. *Sustain Sci Pract Policy* 10:1–7
- Culman SW, Snapp SS, Ollenburger M, et al (2013) Soil and water quality rapidly responds to the perennial grain *kernza* Wheatgrass. *Agron J* 105:735–744
- de Oliveira G, Brunsell NA, Sutherlin CE, et al (2018) Energy, water and carbon exchange over a perennial *Kernza* wheatgrass crop. *Agric For Meteorol* 249:120–137
- de Oliveira G, de Oliveira G, Brunsell NA, et al (2020) Carbon and water relations in perennial *Kernza* (*Thinopyrum intermedium*): An overview. *Plant Science* 295:110279
- DePristo MA, Banks E, Poplin R, et al (2011) A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat Genet* 43:491–498

- Derner JD, Briske DD (2001) Below-ground carbon and nitrogen accumulation in perennial grasses: A comparison of caespitose and rhizomatous growth forms. *Plant and Soil* 237:117–127
- Doebley J, Stec A, Gustus C (1995) teosinte branched1 and the origin of maize: evidence for epistasis and the evolution of dominance. *Genetics* 141:333–346
- Dohleman FG, Long SP (2009) More productive than maize in the Midwest: How does *Miscanthus* do it? *Plant Physiol* 150:2104–2115
- Dong Z, Xiao Y, Govindarajulu R, et al (2019) The regulatory landscape of a core maize domestication module controlling bud dormancy and growth repression. *Nat Commun* 10:1–15
- El-Soda M, Malosetti M, Zwaan BJ, et al (2014) Genotype × environment interaction QTL mapping in plants: lessons from *Arabidopsis*. *Trends in Plant Science* 19:390–398
- Fan Z, Wang K, Rao J, et al (2020) Interactions Among Multiple Quantitative Trait Loci Underlie Rhizome Development of Perennial Rice. *Front Plant Sci* 11:591157
- Fichtner F, Barbier FF, Feil R, et al (2017) Trehalose 6-phosphate is involved in triggering axillary bud outgrowth in garden pea (*Pisum sativum* L.). *Plant J* 92:611–623
- Figueroa CM, Lunn JE (2016) A Tale of Two Sugars: Trehalose 6-Phosphate and Sucrose. *Plant Physiol* 172:7–27
- Frey FP, Presterl T, Lecoq P, et al (2016) First steps to understand heat tolerance of temperate maize at adult stage: identification of QTL across multiple environments with connected segregating populations. *Theor Appl Genet* 129:945–961
- Friedman J, Rubin MJ (2015) All in good time: understanding annual and perennial strategies in plants. *Am J Bot* 102:497–499
- Galinat WC (1981) The inheritance and linkage of perennialism derived from *diploperennis*. *MNL* 55:107
- Hufford MB, Seetharam AS, Woodhouse MR (2021) De novo assembly, annotation, and comparative analysis of 26 diverse maize genomes. *bioRxiv*
- Hu FY, Tao DY, Sacks E, et al (2003) Convergent evolution of perenniality in rice and sorghum. *Proc Natl Acad Sci U S A* 100:4050–4054
- Iltis HH, Doebley JF, M RG, Pazy B (1979) *Zea diploperennis* (Gramineae): A New Teosinte from Mexico. *Science* 203:186–188
- Kebrom TH, Brutnell TP (2015) Tillering in the sugary1 sweet corn is maintained by overriding the teosinte branched1 repressive signal. *Plant Signal Behav* 10:e1078954
- Kebrom TH, Burson BL, Finlayson SA (2006) Phytochrome B represses Teosinte Branched1 expression and induces sorghum axillary bud outgrowth in response to light signals. *Plant Physiol* 140:1109–1117
- Komor E (2000) The physiology of sucrose storage in sugarcane. In: Gupta AK, Kaur N (eds) *Developments in Crop Science*. Elsevier, pp 35–53
- Kong W, Kim C, Goff VH, et al (2015) Genetic analysis of rhizomatousness and its relationship with vegetative branching of recombinant inbred lines of *Sorghum bicolor* × *S. propinquum*. *Am J Bot* 102:718–724
- Lammer D, Cai X, Arterburn M, et al (2004) A single chromosome addition from *Thinopyrum elongatum* confers a polycarpic, perennial habit to annual wheat. *J Exp Bot* 55:1715–1720
- L. Fernando A, Rettenmaier N, Soldatos P, Panoutsou C (2018) 8 - Sustainability of Perennial Crops Production for Bioenergy and Bioproducts. In: Alexopoulou E (ed) *Perennial Grasses for Bioenergy and Bioproducts*. Academic Press, pp 245–283

- Li H (2013) Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. arXiv [q-bio.GN]
- Li H, Handsaker B, Wysoker A, et al (2009) The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 25:2078–2079
- Lindberg CL, Hanslin HM, Schubert M, et al (2020) Increased above-ground resource allocation is a likely precursor for independent evolutionary origins of annuality in the Pooideae grass subfamily. *New Phytol* 228:318–329
- Li P, Zhang Y, Yin S, et al (2018) QTL-By-Environment Interaction in the Response of Maize Root and Shoot Traits to Different Water Regimes. *Front Plant Sci* 9:229
- Liu J, Cheng X, Liu P, Sun J (2017) miR156-Targeted SBP-Box Transcription Factors Interact with DWARF53 to Regulate TEOSINTE BRANCHED1 and BARREN STALK1 Expression in Bread Wheat. *Plant Physiol* 174:1931–1948
- Liu Y, Wang L, Sun C, et al (2014) Genetic analysis and major QTL detection for maize kernel size and weight in multi-environments. *Theor Appl Genet* 127:1019–1037
- Ma A, Qiu Y, Raihan T, et al (2019) The Genetics and Genome-Wide Screening of Regrowth Loci, a Key Component of Perennialism in *Zea diploperennis*. *G3* 9:1393–1403
- Mason MG, Ross JJ, Babst BA, et al (2014) Sugar demand, not auxin, is the initial regulator of apical dominance. *Proc Natl Acad Sci U S A* 111:6092–6097
- McFarland BA, Alkhalifah N, Bohn M, et al (2020) Maize genomes to fields (G2F): 2014–2017 field seasons: genotype, phenotype, climatic, soil, and inbred ear image datasets. *BMC Res Notes* 13:71
- Michelmore RW, Paran I, Kesseli RV (1991) Identification of markers linked to disease-resistance genes by bulked segregant analysis: a rapid method to detect markers in specific genomic regions by using segregating populations. *Proceedings of the National Academy of Sciences* 88:9828–9832
- Murray SC, Jessup RW (2014) Breeding and genetics of perennial maize: Progress, opportunities and challenges. In: *Perennial Crops for Food Security, Proceedings of the FAO Expert Workshop*. pp 103–111
- Park J-Y, Kim H, Lee I (2017) Comparative analysis of molecular and physiological traits between perennial *Arabis alpina* Pajares and annual *Arabidopsis thaliana* Sy-0. *Sci Rep* 7:13348
- Paterson AH, Schertz KF, Lin YR, et al (1995) The weediness of wild plants: molecular analysis of genes influencing dispersal and persistence of johnsongrass, *Sorghum halepense* (L.) Pers. *Proc Natl Acad Sci U S A* 92:6127–6131
- Peng B, Li Y, Wang Y, et al (2011) QTL analysis for yield components and kernel-related traits in maize across multi-environments. *Theor Appl Genet* 122:1305–1320
- Poplin R, Ruano-Rubio V, DePristo MA, et al (2017) Scaling accurate genetic variant discovery to tens of thousands of samples. *Cold Spring Harbor Laboratory* 201178
- Purdy SJ, Cunniff J, Maddison AL, et al (2015) Seasonal Carbohydrate Dynamics and Climatic Regulation of Senescence in the Perennial Grass, *Miscanthus*. *Bioenergy Res* 8:28–41
- Saghai-Marooif MA, Soliman KM, Jorgensen RA, Allard RW (1984) Ribosomal DNA spacer-length polymorphisms in barley: mendelian inheritance, chromosomal location, and population dynamics. *Proc Natl Acad Sci U S A* 81:8014–8018
- Sherrard ME, Maherali H (2006) The adaptive significance of drought escape in *Avena barbata*, an annual grass. *Evolution* 60:2478–2489
- Takagi H, Abe A, Yoshida K, et al (2013) QTL-seq: rapid mapping of quantitative trait loci in

- rice by whole genome resequencing of DNA from two bulked populations. *Plant J* 74:174–183
- Vico G, Manzoni S, Nkurunziza L, et al (2016) Trade-offs between seed output and life span - a quantitative comparison of traits between annual and perennial congeneric species. *New Phytol* 209:104–114
- Wang F, Han T, Song Q, et al (2020) The Rice Circadian Clock Regulates Tiller Growth and Panicle Development Through Strigolactone Signaling and Sugar Sensing. *Plant Cell* 32:3124–3138
- Wang M, Le Moigne M-A, Bertheloot J, et al (2019) BRANCHED1: A Key Hub of Shoot Branching. *Front Plant Sci* 10:76
- Washburn JD, Whitmire DK, Murray SC, et al (2013) Estimation of Rhizome Composition and Overwintering Ability in Perennial Sorghum spp. Using Near-Infrared Spectroscopy (NIRS). *Bioenergy Res* 6:822–829
- Westerbergh A, Doebley J (2004) Quantitative trait loci controlling phenotypes related to the perennial versus annual habit in wild relatives of maize. *Theor Appl Genet* 109:1544–1553
- Whipple CJ, Kebrom TH, Weber AL, et al (2011) Grassy Tillers1 Promotes Apical Dominance in Maize and Responds To Shade Signals in the Grasses. *Proc Natl Acad Sci U S A* 108:E506–12
- Wickham H (2016) ggplot2: Elegant Graphics for Data Analysis
- Wills DM, Whipple CJ, Takuno S, et al (2013) From Many, One: Genetic Control of Prolificacy during Maize Domestication. *PLoS Genet* 9:e1003604
- Wu G, Park MY, Conway SR, et al (2009) The sequential action of miR156 and miR172 regulates developmental timing in Arabidopsis. *Cell* 138:750–759
- Yadav RS, Bidinger FR, Hash CT, et al (2003) Mapping and characterisation of QTL x E interactions for traits determining grain and stover yield in pearl millet. *Theor Appl Genet* 106:512–520
- Zeng ZB, Kao CH, Basten CJ (1999) Estimating the genetic architecture of quantitative traits. *Genet Res* 74:279–289
- Zhang X, Lin Z, Wang J, et al (2019) The tin1 gene retains the function of promoting tillering in maize. *Nat Commun* 10:5608
- Zhang Y, Li Y, Jiang L, et al (2011) Potential of Perennial Crop on Environmental Sustainability of Agriculture. *Procedia Environmental Sciences* 10:1141–1147

**Table 4.1. Number of regrown (RG) and non-regrown (NRG) plants in P39/'Gigi' F<sub>2</sub> greenhouse-grown plants planted in 2018 and scored after one and two cycles of regrowth in 2019. All plants that did not regrow during the first cycle also did not regrow in the second cycle.**

	<b>First regrowth, April 2019</b>	<b>Second regrowth, August 2019</b>
<b>RG</b>	203	104
<b>NRG</b>	293	392
<b>Percent RG</b>	40.9%	21.0%

**Table 4.2. Genotype distribution of markers in greenhouse-regrown P39/'Gigi' F<sub>2</sub> plants.**

Numbers of plants with specified genotype (rows) at each genomic marker (columns) are displayed. Bottom rows depict chi-square and p-values to analyze significant deviation from Mendelian (1:2:1) expectations. p-values below 0.05 are bold.

<b>Genotype</b>	<b>chr2 (44.6Mb)</b>	<b>chr7 (4.5Mb)</b>	<b>chr8 (139.5Mb)</b>
Gigi/Gigi	30	18	43
Gigi/P39	55	44	41
P39/P39	7	32	12
$\chi^2$	15.022	4.254	22.063
p =	<b>0.0005</b>	0.1192	<b>&lt;0.0001</b>

**Table 4.3. Numbers of regrown (RG) and non-regrown (NRG) plants in F<sub>2</sub> populations scored during different time points following flowering and senescence in 2019.**

	P39 / 'Gigi' F <sub>2</sub> s			Hp301 / 'Gigi' F <sub>2</sub> s		
Date Scored	RG	NRG	Percent RG	RG	NRG	Percent RG
Jan	83	113	42.3%	35	384	8.4%
Feb	66	130	33.7%	43	376	10.3%
Mar	61	135	31.1%	33	386	7.9%
May	50	146	25.5%	27	392	6.4%

**Table 4.4: Genotype distribution of markers in field-grown P39/'Gigi' F<sub>2</sub> plants either showing regrowth (RG) or no regrowth (NRG). Bottom rows depict chi-square and p-values to analyze significant deviation from Mendelian (1:2:1) expectations. p-values below 0.05 are bolded.**

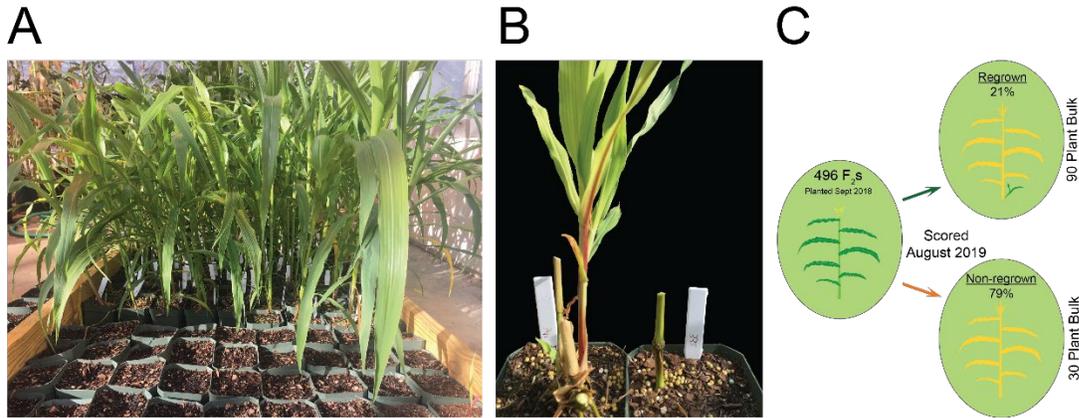
	<b>chr2 (44.6Mb)</b>		<b>chr7 (4.5Mb)</b>		<b>chr8 (139.5Mb)</b>	
<b>Genotype</b>	<b>RG</b>	<b>NRG</b>	<b>RG</b>	<b>NRG</b>	<b>RG</b>	<b>NRG</b>
'Gigi'/'Gigi'	14	15	6	22	7	20
'Gigi'/P39	18	43	19	35	15	29
P39/P39	3	13	10	14	13	22
$\chi^2$	6.837	3.717	1.346	1.778	2.458	2.140
p =	<b>0.0328</b>	0.1559	0.5101	0.4111	0.2927	0.3431

**Table 4.5:** List of Illumina sequenced samples for this study. SE150=Single End 150 sequencing; PE150=Paired End 150 sequencing.

<b>Sample</b>	<b>Sequencing Type</b>	<b>Reads/Read Pairs</b>	<b>Sequencer</b>	<b>SRA Accession</b>
<i>Z. diploperennis</i> 'Gigi'	PE150	449,924,498	NextSeq 500	SAMN17833852
Regrowth Bulk	SE150	246,314,039	HiSeq 4000	SAMN17833853
Non-regrowth Bulk	PE150	21,841,531	NextSeq 500	SAMN17833854
Non-regrowth Bulk	PE150	374,940,596	HiSeq 4000	SAMN17833854
High-tiller Bulk	PE150	325,222,732	HiSeq 4000	SAMN17833855
Low-tiller Bulk	PE150	330,008,372	HiSeq 4000	SAMN17833856

**Table 4.6:** Number of filtered, homozygous SNPs per chromosome present in *Z. diploperennis* 'Gigi' from Illumina reads aligned to maize P39 reference.

<b>Chromosome</b>	<b>Number of SNPs</b>
chr1	294,198
chr2	256,571
chr3	267,237
chr4	338,163
chr5	222,403
chr6	172,588
chr7	124,049
chr8	165,662
chr9	205,810
chr10	133,573
Total	2,180,254

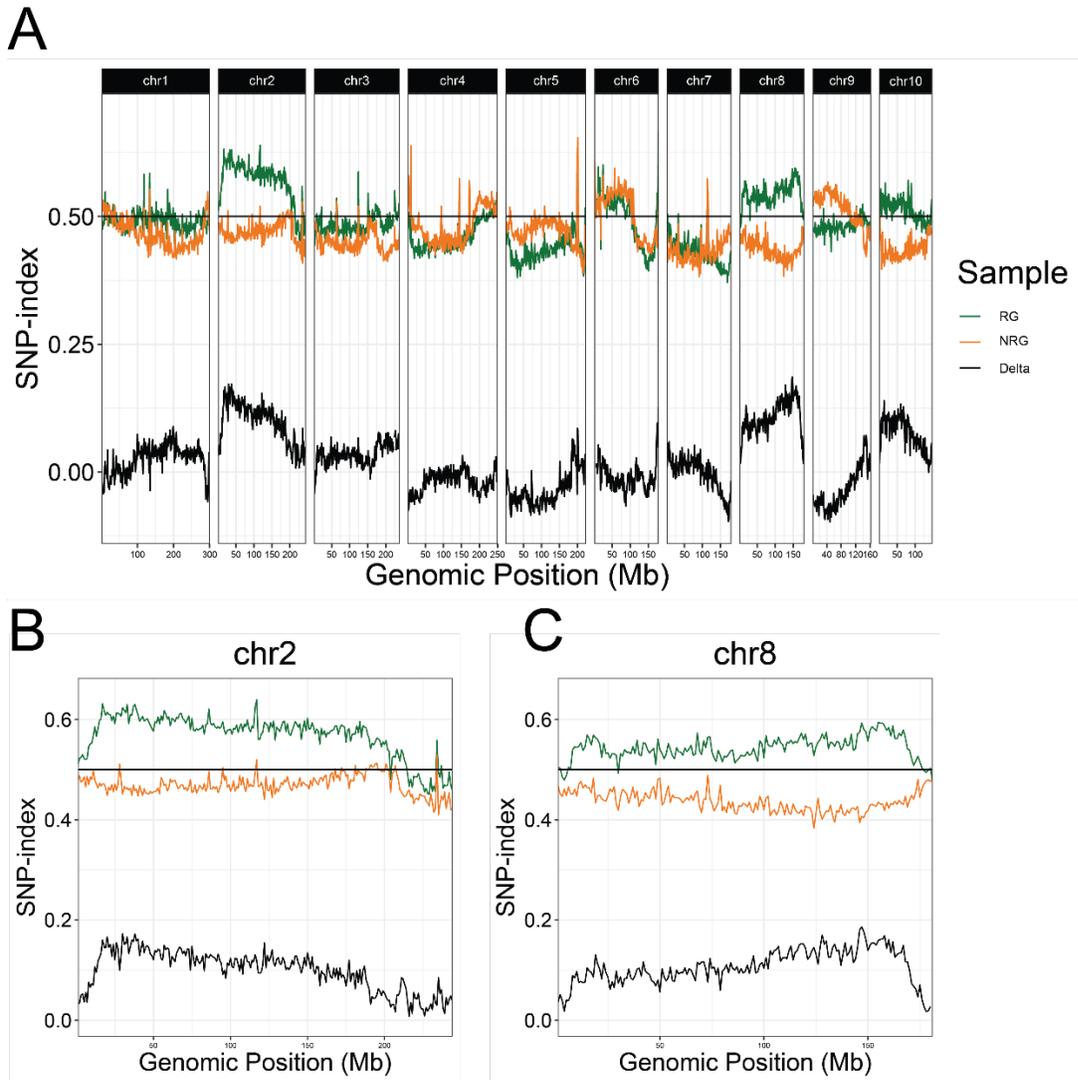


**Figure 4.1. Growth conditions for scoring regrowth in F<sub>2</sub> populations.**

(A) Seedlings were grown in square four-inch pots over a moist coconut mat.

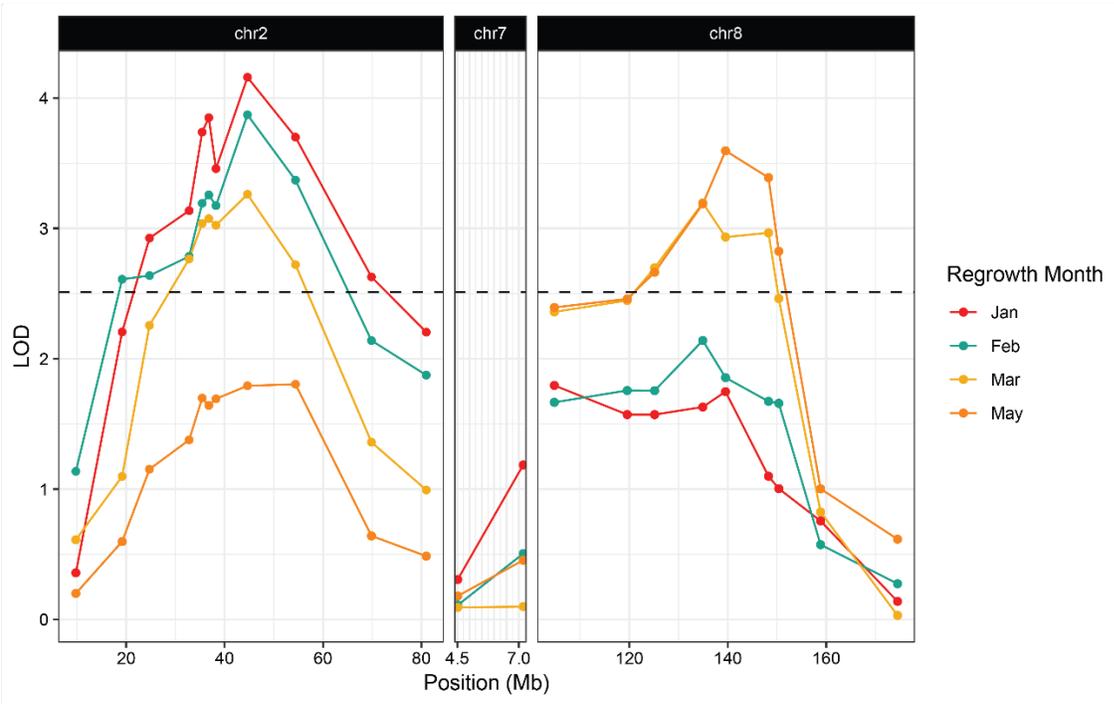
(B) Following flowering and senescence, regrown (left) and non-regrown (right) plants were scored.

(C) Schematic for QTL-seq experiment. 496 P39/'Gigi' F<sub>2</sub>s were planted in September 2018 and scored in August 2019 for regrowth (21%; top) and non-regrowth (79%; bottom). QTL-seq was carried out on a 90-plant regrown bulk and a 30-plant non-regrown bulk.

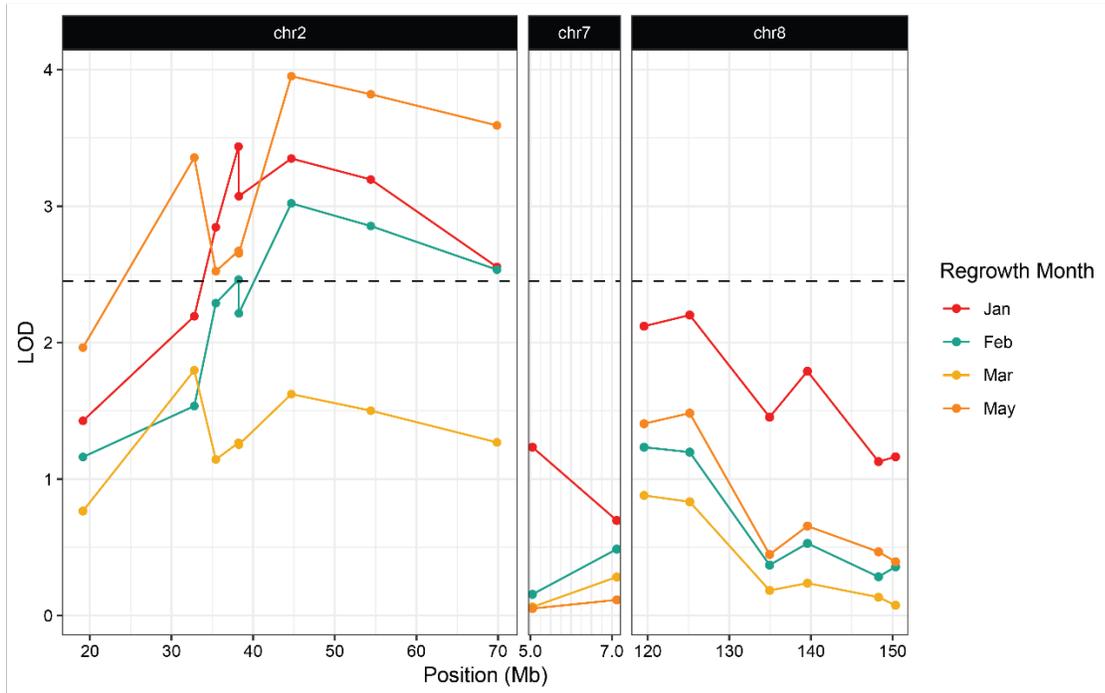


**Figure 4.2. QTL-seq for regrowth reveals two major loci in a P39/'Gigi' F<sub>2</sub> population.**

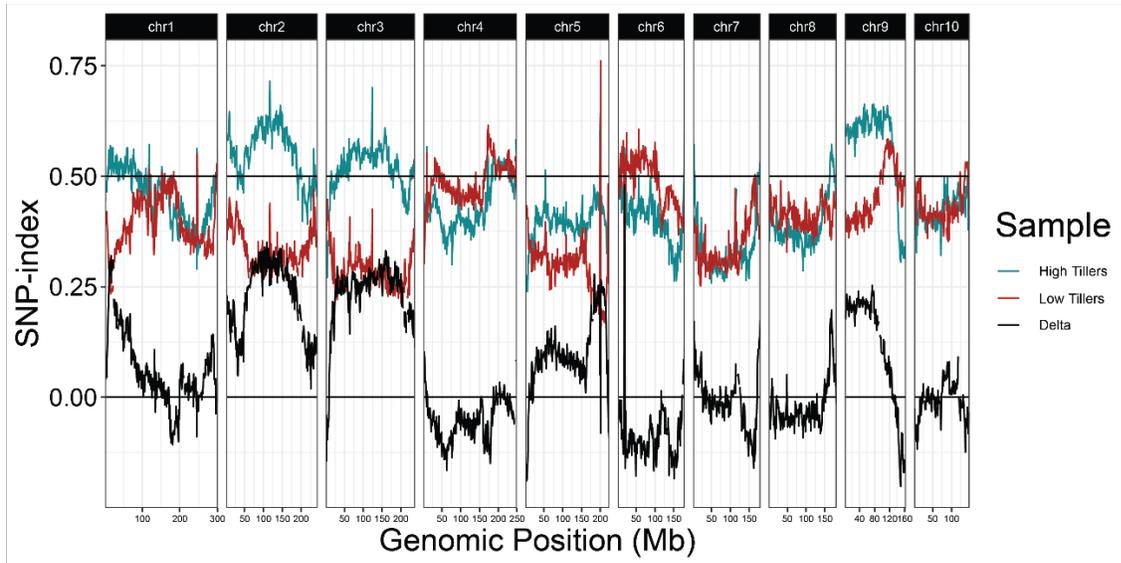
(A) Average SNP-index per 1 Mb window is plotted across all ten chromosomes. Chromosomes 2 (B) and 8 (C) are shown in greater detail. For each plot, lines display SNP-index of regrown (RG) bulk in green, non-regrown (NRG) bulk in orange, and the difference between the two bulks (RG - NRG) in black.



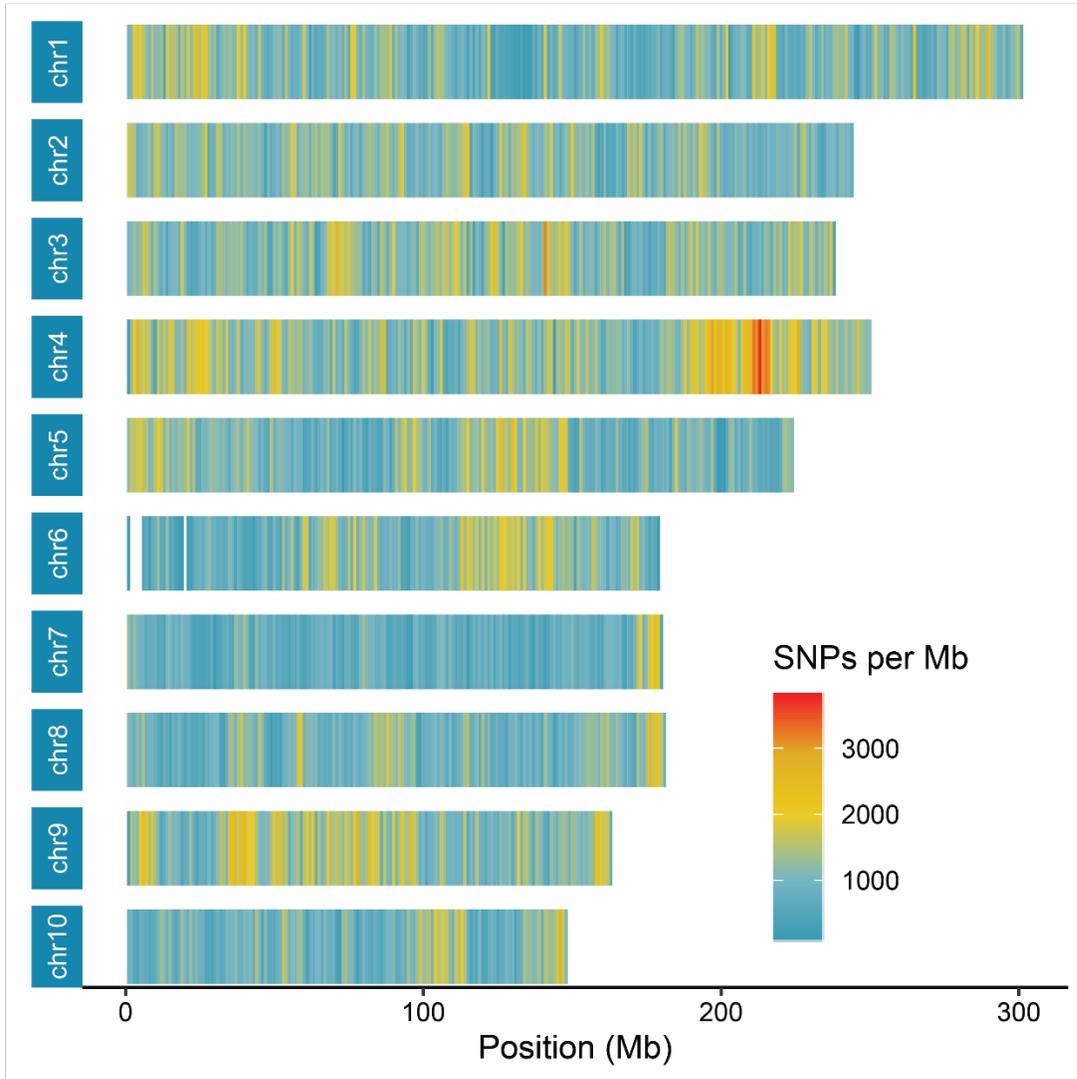
**Figure 4.3. Association between regrowth and markers on chromosomes 2, 7, and 8 in the 2019 P39/'Gigi' F<sub>2</sub> population.** QTL were considered significant if the LOD score exceeded the 95% threshold determined by 1000 random permutations of the data for each trait. The threshold was calculated independently for each time point. February had the highest threshold of LOD=2.51 and was used here as the significance cutoff (dashed line).



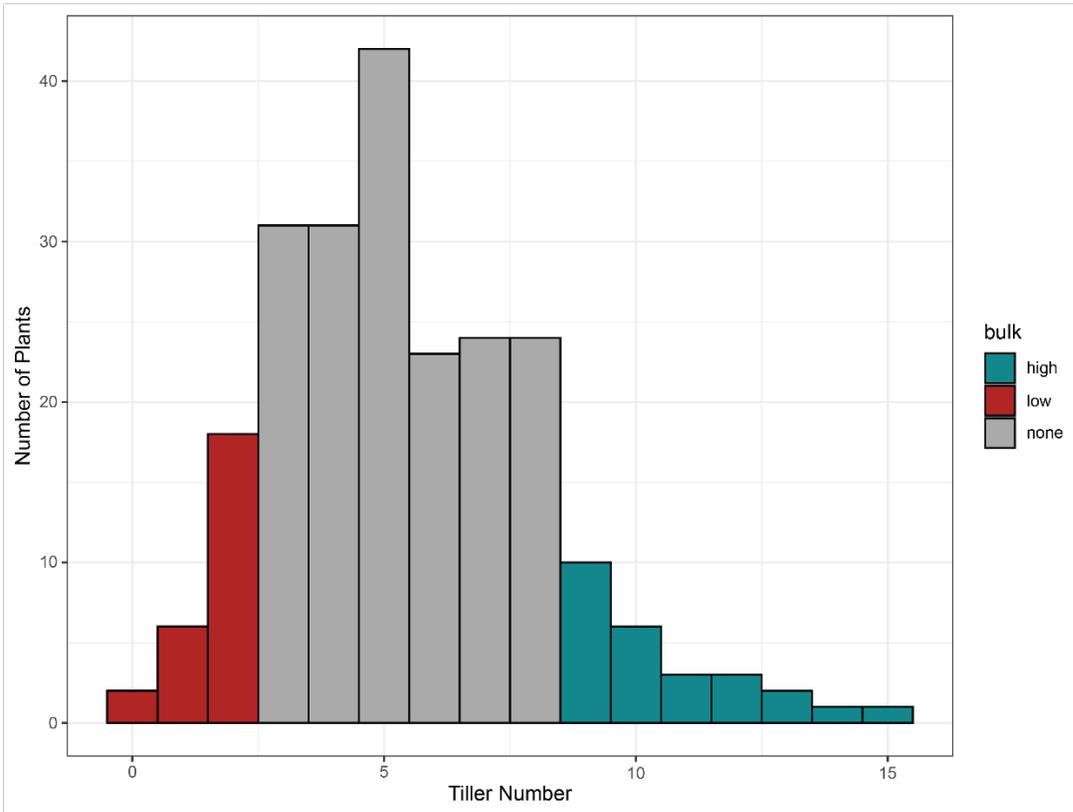
**Figure 4.4. Association between regrowth and markers on chromosomes 2, 7, and 8 in the 2019 Hp301/'Gigi' F<sub>2</sub> population.** QTL were considered significant if the LOD score exceeded the 95% threshold determined by 1000 random permutations of the data for each trait. The threshold was calculated independently for each time point. May had the highest threshold of LOD=2.45 and was used here as the significance cutoff (dashed line).



**Figure 4.5. QTL-seq for tiller number supports numerous tillering QTL in the 2019 P39/'Gigi' F<sub>2</sub> field population.** Average SNP-index per 1 Mb window is plotted across all ten chromosomes. Lines display SNP-index in high (blue) and low (red) tiller bulks and black is the difference between the two bulks (HT-LT).



**Figure 4.6. Distribution of filtered homozygous 'Gigi' SNPs mapped to P39v1 reference assembly.**



**Figure 4.7. Distribution of tiller number phenotype in field-grown P39/'Gigi' F<sub>2</sub> plants.** Red and blue shaded regions indicate plants that were included in low and high tiller QTL-seq bulks, respectively.

## CHAPTER 5

### CONCLUSIONS AND DIRECTIONS FOR FUTURE STUDY

Since Rhoades' initial discovery that the Ab10 haplotype confers meiotic drive in maize (Rhoades 1942), we have come to a nearly-complete understanding of many aspects of this meiotic drive system. Knobs become motile neocentromeres in meiosis (Rhoades 1952) and we have characterized the molecular motors that provide motility for the two classes of knobs in maize. Some questions surrounding Ab10 meiotic drive do still persist and future work on this biological question may help provide answers and lay foundations to aid in genome engineering.

The major class of maize knobs are composed of the knob180 unit and we discovered a gene complex on the distal tip of Ab10 that encodes a kinesin-14 protein capable of providing the motile force necessary for knob180 neocentromeres. This encoded protein, KINDR, localizes specifically to knob180 knobs during meiotic anaphase (Figure 2.1) and is necessary for Ab10 meiotic drive (Dawe et al. 2018). It remains unclear how the physical association between KINDR and knob180 knobs occurs. Biochemical and genetic evidence suggests an intermediate molecule may assist in the KINDR-knob180 interface and the gene required for this function is missing in the Ab10 mutant genotype *smd13* (Figures 2.2, 2.4). Given that the gene product required for proper KINDR localization is missing from *smd13*, we attempted to identify the causal mutation using a variety of bioinformatic techniques. However, we did not ultimately identify a causative mutation to explain the *smd13* mutant phenotype. We identified a candidate

gene, a *Kinesin-10-like*, that is found at a locus on the Ab10 distal tip near the *Kindr* complex (Figure 2.6) and has lower expression in *smd13* compared to Ab10 (Figure 2.7).

We have formulated a hypothesis that the *Kinesin-10-like* gene product physically associates with both KINDR and knob180 knobs, thus physically providing a connection between the kinesin motor protein and the DNA element that comprises the major class of maize knobs. Future work in validating this hypothesis should focus on its biochemical aspects.

Namely, the following criteria must be met in order to conclude that the KINESIN-10-LIKE protein connects KINDR to knob180 knobs: 1) KINESIN-10-LIKE must physically interact with KINDR; 2) KINESIN-10-LIKE must physically interact with knob180 DNA, likely in a sequence-specific manner; and 3) KINESIN-10-LIKE must localize to knob180 knobs. An independent mutation of *Kinesin-10-like* showing that a loss of its expression results in a decrease in meiotic drive will also provide genetic evidence that this gene is involved in this process. A more complete understanding of the Ab10 meiotic drive system is not only fascinating from an evolutionary biology perspective but could have synthetic biology applications as well. For example, learning how a broadly-successful chromosome motility system functions may allow us to develop ways to tether and move chromosomes. These could be used to engineer an artificial drive system or a synthetic centromere capable of sustaining a newly-engineered chromosome.

TR-1 repeats comprise the minor class of knobs in the maize genome, and we have demonstrated the *Trkin* gene encodes the molecular motor required for TR-1 neocentromeres (Figures 3.3, 3.4). TR-1 neocentromeres play an important role in meiotic drive suppression, where the large TR-1 knobs of the K10L2 haplotype can compete with Ab10, thus suppressing Ab10 meiotic drive (Kanizay et al. 2013). However, it is still unclear if TR-1 neocentromeres

themselves are necessary or sufficient to confer meiotic drive. K10L2 itself appears to drive very weakly, but a definitive test of whether or not TR-1 neocentromeres can confer meiotic drive may come from a comparison of closely-related genotypes that differ in their ability to provide the TR-1 neocentromere force. Future efforts to characterize the genetic effects of TR-1 neocentromeres may focus on loss-of-function *trkin* mutants generated using CRISPR/Cas9 mutation.

In our *Z. diploperennis* mapping studies, we found two major QTL, *regrowth1* and *regrowth3*, that confer perennial regrowth (Figures 4.2, 4.3, 4.4). The positions of these two loci have only been mapped to rough genomic coordinates, so an obvious future direction is to fine-map each locus and determine the genetic variants underlying the regrowth phenotype. Given the complex nature of the regrowth trait, this will require specialized mapping populations such as backcrosses and NILs which are currently being generated. To achieve the goal of breeding perennial maize lines, a more thorough understanding of the physiological and developmental processes that dictate resource homeostasis and axillary meristem regulation will be crucial. Isolation of NILs segregating *regrowth* loci will also enable further characterization of the associated phenotypes. For example if NILs segregating *reg1* and *reg3* can be isolated, we can determine if these loci have direct effect on traits related to perenniality such as tiller number, prolificacy, or juvenile-to-adult transition. Maize genes involved in tillering regulation have been well-characterized due to analysis of extensive maize mutant collections. Another approach to identifying genes involved in perennial regrowth is to cross maize mutants into *Z. diploperennis* and look for defects in the regrowth phenotype when mutations in tillering genes are present. Developmental genetic analysis of perennial regrowth in *Z. diploperennis* should provide a more detailed understanding of this phenotype and hopefully aid in the eventual production of

perennial maize.

## REFERENCES

- Dawe, R. Kelly, Elizabeth G. Lowry, Jonathan I. Gent, Michelle C. Stitzer, Kyle W. Swentowsky, David M. Higgins, Jeffrey Ross-Ibarra, et al. 2018. “A Kinesin-14 Motor Activates Neocentromeres to Promote Meiotic Drive in Maize.” *Cell* 173 (4): 839–50.e18.
- Kanizay, Lisa B., Patrice S. Albert, James A. Birchler, and R. Kelly Dawe. 2013. “Intragenomic Conflict between the Two Major Knob Repeats of Maize.” *Genetics* 194 (1): 81–89.
- Rhoades, M. M. 1942. “Preferential Segregation in Maize.” *Genetics* 27 (4): 395–407.
- Rhoades, M. M., and Others. 1952. “Preferential Segregation in Maize.” *Preferential Segregation in Maize*. <https://www.cabdirect.org/cabdirect/abstract/19531603828>.