


RESEARCH ARTICLE

# De Novo Centromere Formation and Centromeric Sequence Expansion in Wheat and its Wide Hybrids

Xiang Guo<sup>1,2</sup>, Handong Su<sup>1,2</sup>, Qinghua Shi<sup>1,2</sup>, Shulan Fu<sup>3</sup>, Jing Wang<sup>1,2</sup>, Xiangqi Zhang<sup>1</sup>, Zanmin Hu<sup>1</sup>, Fangpu Han<sup>1\*</sup>

**1** State Key Laboratory of Plant Cell and Chromosome Engineering, Institute of Genetics and Developmental Biology, Chinese Academy of Sciences, Beijing, China, **2** University of Chinese Academy of Sciences, Beijing, China, **3** State Key Laboratory of Plant Breeding and Genetics, Sichuan Agricultural University, Wenjiang, Chengdu, Sichuan, China

 These authors contributed equally to this work.

\* [fphan@genetics.ac.cn](mailto:fphan@genetics.ac.cn)



CrossMark  
click for updates

 OPEN ACCESS

**Citation:** Guo X, Su H, Shi Q, Fu S, Wang J, Zhang X, et al. (2016) *De Novo Centromere Formation and Centromeric Sequence Expansion in Wheat and its Wide Hybrids*. PLoS Genet 12(4): e1005997. doi:10.1371/journal.pgen.1005997

**Editor:** Gregory P. Copenhaver, The University of North Carolina at Chapel Hill, UNITED STATES

**Received:** February 2, 2016

**Accepted:** March 28, 2016

**Published:** April 25, 2016

**Copyright:** © 2016 Guo et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** All relevant data are within the paper and its Supporting Information files except for the anti-CENH3 ChIP-seq data, which were deposited in the Gene Expression Omnibus (GEO) database under number GSE63752.

**Funding:** This work was supported by NSFC (31130033 and 31320103912) and the 863 program (2011AA100101). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing Interests:** The authors declare that no competing interests exist.

## Abstract

Centromeres typically contain tandem repeat sequences, but centromere function does not necessarily depend on these sequences. We identified functional centromeres with significant quantitative changes in the centromeric retrotransposons of wheat (CRW) contents in wheat aneuploids (*Triticum aestivum*) and the offspring of wheat wide hybrids. The CRW signals were strongly reduced or essentially lost in some wheat ditelosomic lines and in the addition lines from the wide hybrids. The total loss of the CRW sequences but the presence of CENH3 in these lines suggests that the centromeres were formed *de novo*. In wheat and its wide hybrids, which carry large complex genomes or no sequenced genome, we performed CENH3-ChIP-dot-blot methods alone or in combination with CENH3-ChIP-seq and identified the ectopic genomic sequences present at the new centromeres. In addition, the transcription of the identified DNA sequences was remarkably increased at the new centromere, suggesting that the transcription of the corresponding sequences may be associated with *de novo* centromere formation. Stable alien chromosomes with two and three regions containing CRW sequences induced by centromere breakage were observed in the wheat-*Th. elongatum* hybrid derivatives, but only one was a functional centromere. In wheat-rye (*Secale cereale*) hybrids, the rye centromere-specific sequences spread along the chromosome arms and may have caused centromere expansion. Frequent and significant quantitative alterations in the centromere sequence via chromosomal rearrangement have been systematically described in wheat wide hybridizations, which may affect the retention or loss of the alien chromosomes in the hybrids. Thus, the centromere behavior in wide crosses likely has an important impact on the generation of biodiversity, which ultimately has implications for speciation.

## Author Summary

Chromosomal rearrangements during the formation of wheat aneuploids and their wide hybrids caused reduction, elimination or expansion of the centromeric retrotransposon sequences and the formation of multiple centromeres. Centromere function was not affected by centromeric sequence elimination, which was revealed by the *de novo* formation of centromeres on the rearranged chromosomes. Several retrotransposon-like elements near the former centromeres were embedded in the newly formed centromeres, and there were no obvious changes in six histone modifications between normal and new centromeres. The DNA sequences associated with the new centromeres are transcribed at a higher level after centromere formation. Chromosomes containing the neocentromeres can be stably transferred to the next generation. Chromosomes carrying two- or three-locus centromeres are unstable, which induces the formation of novel chromosomes through centromere breakage in wheat-*Th. elongatum* hybrid derivatives. The centromere-specific sequences on dicentric chromosomes are expanded to the chromosome arms in wheat-rye hybrids, and these sequences may function as a part of the active centromere to cause chromosome breakage in the next generation. Centromere variation and activity in wheat aneuploids and its wide hybrids may be associated with chromosome stability, rearrangements, and novel chromosome formations.

## Introduction

Centromeres, which are located at the primary constriction of the chromosome, are required for the accurate segregation of chromosomes and serve as the sites for kinetochore assembly during mitosis and meiosis. The main DNA components of the centromere are highly repetitive, such as the 171-bp  $\alpha$ -satellite repeat in humans and 150- to 180-bp simple tandem repeats in some flowering plants [1–5]. Long-terminal repeat (LTR) retrotransposons, also known as centromeric retrotransposons (CRs), are often intermingled with tandem repeats and are enriched in plant centromeric regions [6–11]. The highly conserved function of the centromere is correlated with its epigenetic features, including the histone H3 variant CENH3 in plants (CENP-A in mammals) [12–15], phosphorylation of histone H2A at Thr-133 [16] and H3 phosphorylation at Ser-10 [17,18]. Despite the conserved centromere function, centromeric repeat sequences apparently evolved rapidly in some species under specific circumstances. This phenomenon is known as the "centromere paradox" [13].

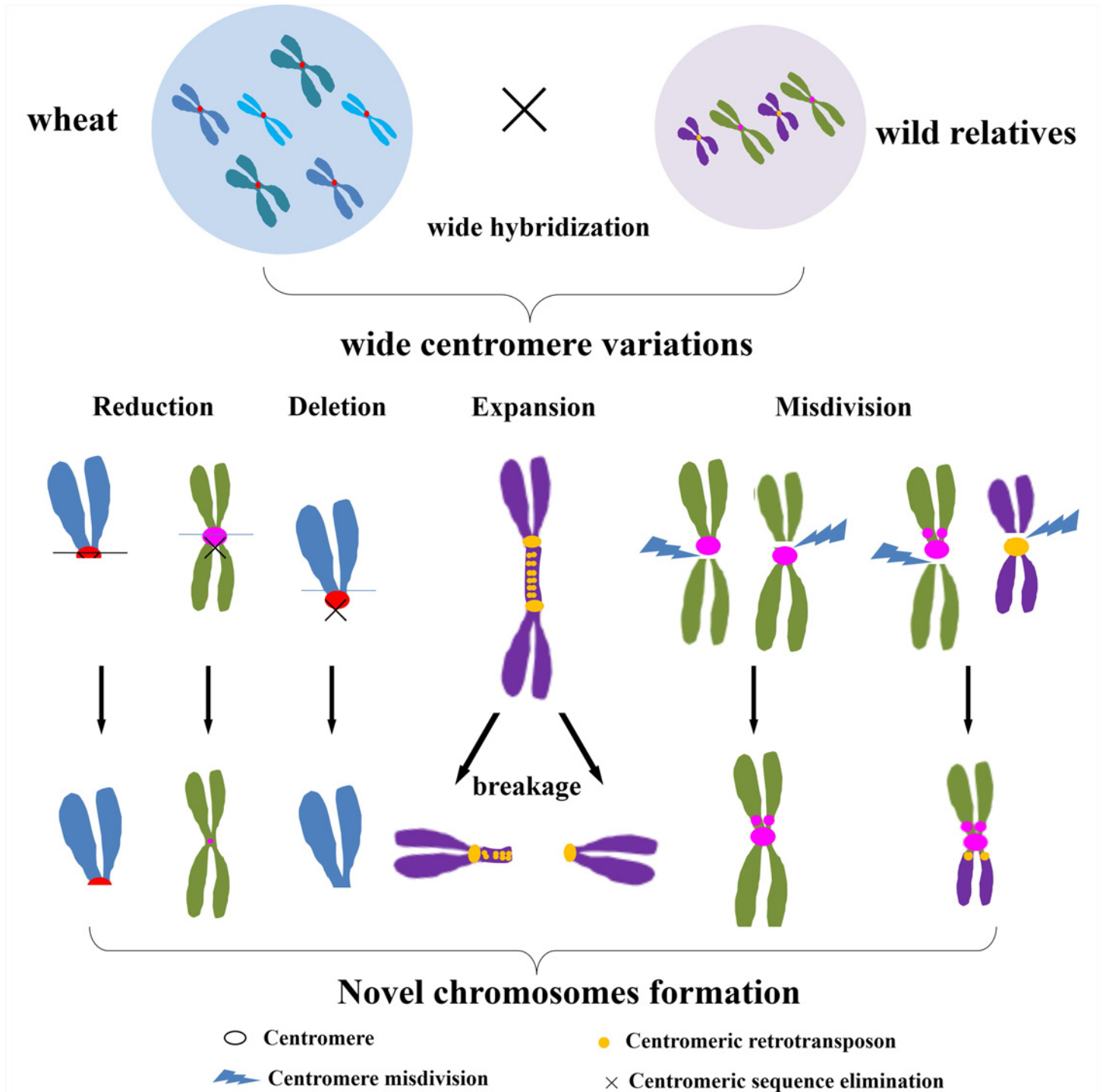
Centromeric sequences are highly variable between different species and different chromosomes and even between the same centromeres from different ecotypes or varieties [5,11,19,20]. Most of the centromeric tandem repeats in plants, such as CentO in rice (*Oryza sativa*), CentBd in *Brachypodium distachyon*, and CentC in maize (*Zea mays*), are likely to be species-specific [4,5,21]. Several wild *Oryza* species lack CentO and instead possess genome-specific satellite repeats [22]. Similarly, little homology was found between the centromeric sequences of the potato (*Solanum tuberosum*) and its wild relative *S. verrucosum* [23]. Moreover, centromeres showed diversity in the repeat-less and repeat-based sequences on different chromosomes of *S. verrucosum* [20]. Eukaryotic centromeres carrying novel satellites may have evolved from neocentromeres that experienced insertion and/or extensive amplification of satellite repeats [20,24]. Previous studies revealed that recent segmental duplication, abundant rearrangements, and reshuffling occurred in CEN4 and CEN8 of rice and that the changes in CEN8 seemed to appear after the divergence of the *O. sativa* subspecies *japonica* and *indica* from a common ancestor [24,25]. An analysis of centromere retention or loss indicated that

the major events during the evolution of maize from a supposed tetraploid ancestor (*Sorghum bicolor*) were chromosomal rearrangements, such as insertions and translocations, resulting in dysploidy and reduced chromosome numbers [26]. Despite the observation that substantial variations in centromeres occurred during evolution, the relationship between centromere variations and species evolution remains uncertain.

In most eukaryotes, the centromeric sequences alone are insufficient to maintain a functional centromere [27]. In humans and plants, many newly formed centromeres are devoid of typical centromeric sequences, and their formation was likely determined by epigenetic mechanisms [28–32]. Additionally, the centromere activity of dicentric chromosomes is independent of centromeric sequences. Many stable dicentric chromosomes in maize, including A-A and A-B centromeres (the centromere of the supernumerary B chromosome contains B-specific repeats), contain one active and one inactive centromere, as determined by examining the epigenetic modifications [18,33]. Furthermore, the inactive centromere recovered its activity by switching its epigenetic features under certain circumstances [34]. The essential structural and functional components for the core chromatin of centromeres include pericentromeric heterochromatin and active transcription of centromeric DNA [35–38]. The epigenetic components, rather than the DNA sequences, are essential for the establishment of centromere function. However, it remains a mystery why most functional centromeres contain highly repetitive sequences.

Allopolyploid wheat, either tetraploid or hexaploid, originates from interspecific hybridizations that trigger striking chromosomal rearrangements, genome reorganization, and chromatin remodeling in the parental genomes [39–43]. Wheat also has the capacity to hybridize with its wild relatives, which provides a broader gene source for wheat germplasm enhancement through addition, translocation, and substitution lines containing alien chromosomes [44–46]. In fact, wheat appears to prefer alien chromosomes or fragments from specific genomes [47]. However, the mechanisms regulating the stable transmission of these alien genomic sources in defined genetic backgrounds are still unclear. A previous study indicated that the size of the maize centromere was expanded in oat (*Avena sativa*)-maize addition lines, which may be a key factor for the survival of neocentromeric chromosomes in natural populations [48]. As such, an understanding of the adaptation of centromeres to "genome shock" and their evolutionary history in the wide wheat hybrid will require additional studies.

Due to their repetitive structures and low sequence conservation, it is difficult to compare centromeric sequences across different species. Complete centromeres on partial chromosomes have been sequenced in rice and maize [8,49–51]. In wheat, only partial centromere sequences have been released from published *bacterial artificial chromosome* (BAC) sequences [49,52–54]. Here, we observed that the content of classical centromeric retrotransposon sequences was reduced or apparently lost in both aneuploid wheats (4D, 1B, 5D chromosomes) and their wild relatives, such as *Th. intermedium* and *Th. elongatum*, when hybridized with wheat (Fig 1 and Table 1). With new developments in wheat genome sequencing, we first uncovered the detailed sequences in the new centromere of the 4DS chromosome in wheat aneuploids by ChIP (chromatin immunoprecipitation)-sequencing with wheat CENH3 antibodies. Additionally, for *Th. intermedium*, which does not have a sequenced reference genome, we developed a new ChIP-dot-blot strategy (see Methods) to identify the novel centromeric sequences in the wheat-*Th. intermedium* addition line TAI-14. We also detected the expansion of centromeric sequences and the formation of multiple centromeres in wheat and its wide hybrid offspring (Fig 1 and Table 1). Finally, we provide a detailed analysis of centromere variations and offer some new insights into centromere evolution in wheat and its wild relatives (Fig 1).



**Fig 1. Stable and novel chromosomes induced by universal centromere variations in wheat and wide hybrids.** Centromere alterations occurred in aneuploid wheat and in wheat hybrids. Reduction or even deletion of the centromeric sequences did not affect the transmission of the chromosomes or chromosome fragments to the next generation because of *de novo* centromere formation. Centromere expansion and breakage led to novel alien chromosomes in wheat and wide hybrids.

doi:10.1371/journal.pgen.1005997.g001

**Table 1. Centromere variation and *de novo* centromere formation in wheat aneuploids and derivatives of wheat wild hybrids.**

Materials	Types of centromere variation	Chromosome
Ditelosomic line 4DS	sequence deletion	4DS
Wheat- <i>Th. intermedium</i> addition lines	sequence deletion	alien chromosome
Wheat- <i>Th. elongatum</i> addition lines	sequence deletion	3E
Ditelosomic lines 1BS, 5DS, 5DL	sequence reduction	1BS,5DS,5DL
Wheat- <i>Th. elongatum</i> addition lines	sequence reduction	alien chromosome
N3AT3B× 8802	two-locus centromere	2E
N3DT3B	two functional centromeres	wheat and alien chromosome
N5AT5B× 8802	two functional centromeres	alien chromosome
N5BT5A× 8802	two-locus centromere	2E
N5DT5B× 8802	two-locus centromere	2E
	two functional centromeres	wheat and alien
	three functional centromeres	chromosome
	sequence reduction	alien chromosome
N6AT6B× 8802	two- and three-locus centromere	2E,5E
N6DT6A× 8802	two-locus centromere	alien chromosome
	two functional centromeres	wheat chromosome
	sequence reduction	alien chromosome
Wheat- <i>Th. poticum</i> addition lines	two functional centromeres	wheat chromosome
	three functional centromeres	
Wheat- <i>Th. intermedium</i> addition lines	two functional centromeres	wheat chromosome
Wheat- <i>Th. poticum</i> addition lines	two functional centromeres	wheat chromosome
Wheat and <i>Agropyron cristatum</i> addition lines	two functional centromeres	wheat chromosome
Wheat and <i>Hordeum vulgare</i> addition lines	two functional centromeres	wheat chromosome
Wheat-rye 2R addition lines	centromere expansion	2RS
Wheat-rye 6R addition lines	centromere expansion	6RS
Wheat and rye addition lines	two functional centromeres	wheat chromosome

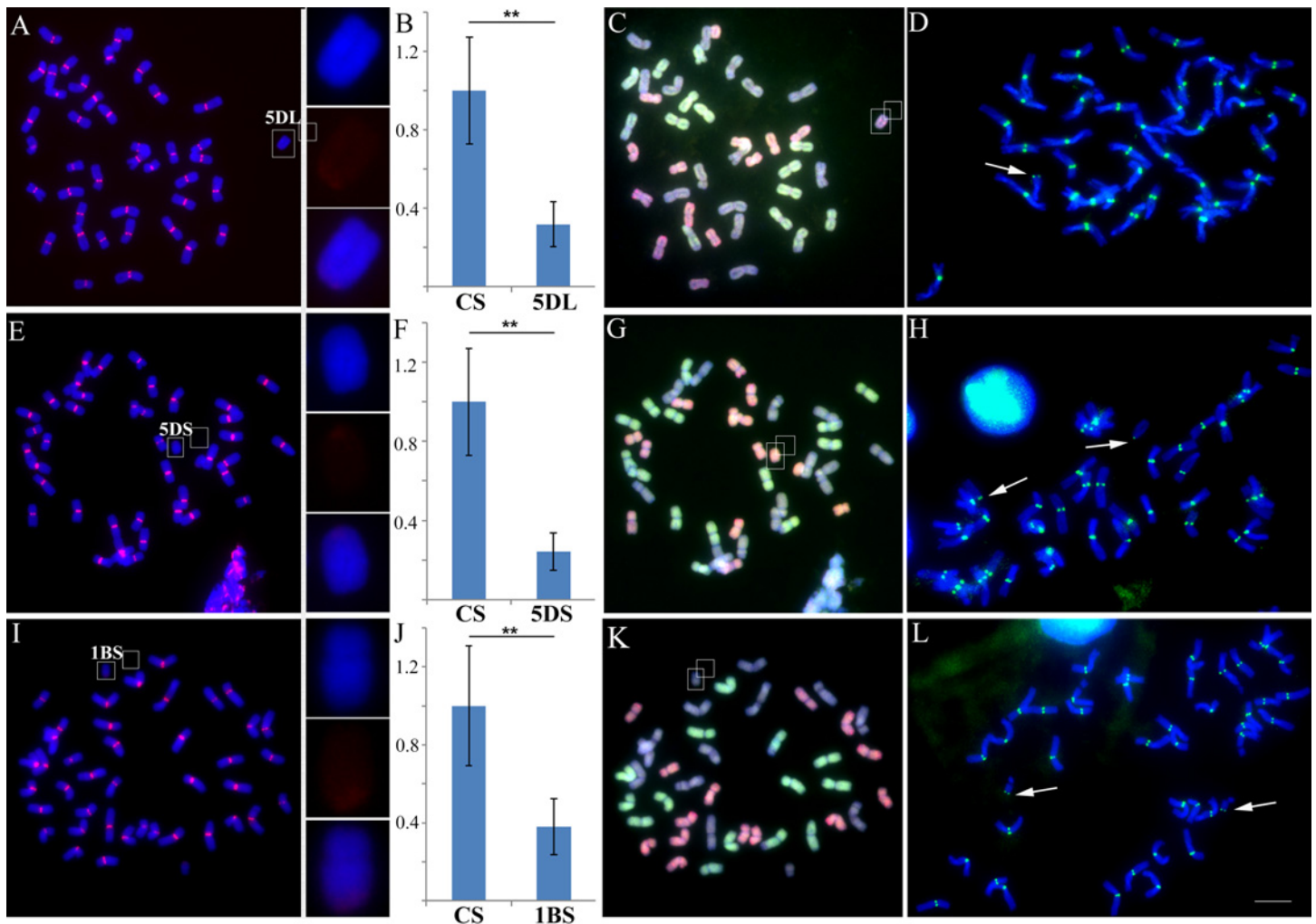
doi:10.1371/journal.pgen.1005997.t001

## Results

### Reduction and elimination of centromeric sequences in wheat aneuploids and derivatives of wheat wide hybrids

The loss of canonical centromere sequences can be induced by breakage, rearrangements and radiation at plant centromeres [31,32]. Here, we observed the elimination of centromeric sequences in both wheat aneuploids and their wide hybrids.

Compared with normal centromeres in the *T. aestivum* Chinese Spring background, weaker fluorescence *in situ* hybridization (FISH) signals from the CRW probes were detected in the ditelosomic lines 5DL, 5DS and 1BS [55] (Fig 2A–2C, 2E–2G and 2I–2K and S1 Fig). Thus, significant reductions of centromeric sequences can frequently occur in allopolyploid wheat. However, CENH3 immunostaining revealed that functional centromeres were present in these three lines (Fig 2D, 2H and 2L). Additionally, in the ditelosomic line 4DS [55], we were unable to detect any CRW signals with FISH in the centromere or the chromosome arms, which stands in stark contrast to the normal chromosome 4D (Figs 3A–3C and S1). However, the epigenetic marks of active centromeres, including CENH3 and H2A phosphorylation at Thr-133 and H3 phosphorylation at Ser-10, were correctly loaded on the short arm of the 4D chromosome, suggesting that a *de novo* centromere had formed that lacked the canonical centromeric sequences (Fig 3D–3F).

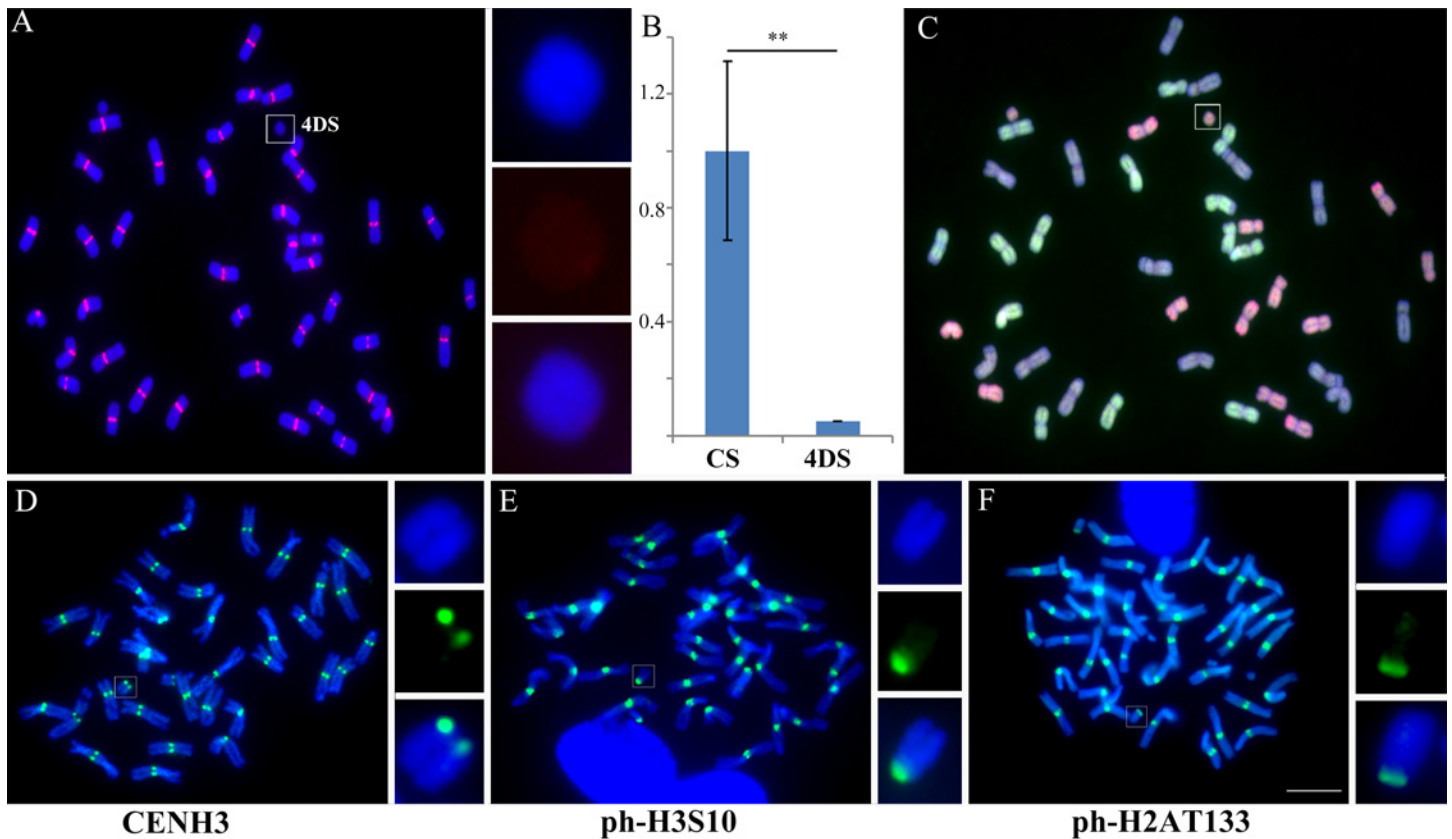


**Fig 2. Centromeric sequence reduction in the ditelosomic lines 5DL, 5DS and 1BS.** (A), (E) and (I). FISH analysis of 5DL, 5DS, and 1BS, respectively. The CRW sequences are labeled in red, and DAPI staining is labeled in blue. The insets show high-magnification images of chromosomes 5DL, 5DS and 1BS. (B), (F) and (J). Intensity of CRW fluorescence in 5DL, 5DS and 1BS, respectively, compared with the controls in Chinese Spring. The Y-axis represents the relative CRW fluorescence intensity. The error bars show the standard deviation (s.d.). The double asterisks denote significant differences at  $P < 0.001$  (two-tailed Student's t-test).  $P = 3.4E-16$  (B),  $P = 8.22E-18$  (F),  $P = 1.97E-13$  (J). (C), (G) and (K). Multi-color FISH analysis of 5DL, 5DS, and 1BS, respectively. The DNA for the wheat A genome is labeled in green, the DNA for the D genome is labeled in red, and the DNA for the B genome is used as a block. (D), (H) and (L). Immunostaining of 5DL, 5DS, and 1BS, respectively, with antibodies against CENH3 (green). The arrows indicate CENH3 on the 5DL, 5DS and 1BS chromosomes. Bar = 10  $\mu\text{m}$ .

doi:10.1371/journal.pgen.1005997.g002

The wheat-*Th. intermedium* addition line TAI-14 was generated from hybrids between *T. aestivum* Xinshuguang 1 and amphidiploids zhong2 ( $2n = 56$ ) [56]. The CRW sequences were heterogeneously distributed in the centromeric region of the 42 chromosomes of *Th. intermedium*, although some FISH-detected signals were very weak (S2A Fig). However, there were no detectable CRW signals on the *Th. intermedium*-derived chromosome in TAI-14 (Fig 4A). These chromosomes had functional centromeres, as revealed by the presence of CENH3 (Fig 4B).

Additionally, most copies of CRW in the two alien chromosomes from *Th. elongatum* were eliminated in the derivatives of the Chinese Spring nulli-tetrasomic lines N6AT6B ( $2n = 42$ )  $\times$  wheat-*Th. elongatum* amphidiploid 8802 ( $2n = 42$ , AABBE1E2) (Fig 5B). The genome of 8802 consists of 28 chromosomes from *T. durum* Kekeruite ( $2n = 28$ ) and 14 chromosomes from *Th. elongatum* AE31 ( $2n = 28$ ) [57]. No obvious CRW FISH signals were detected on the



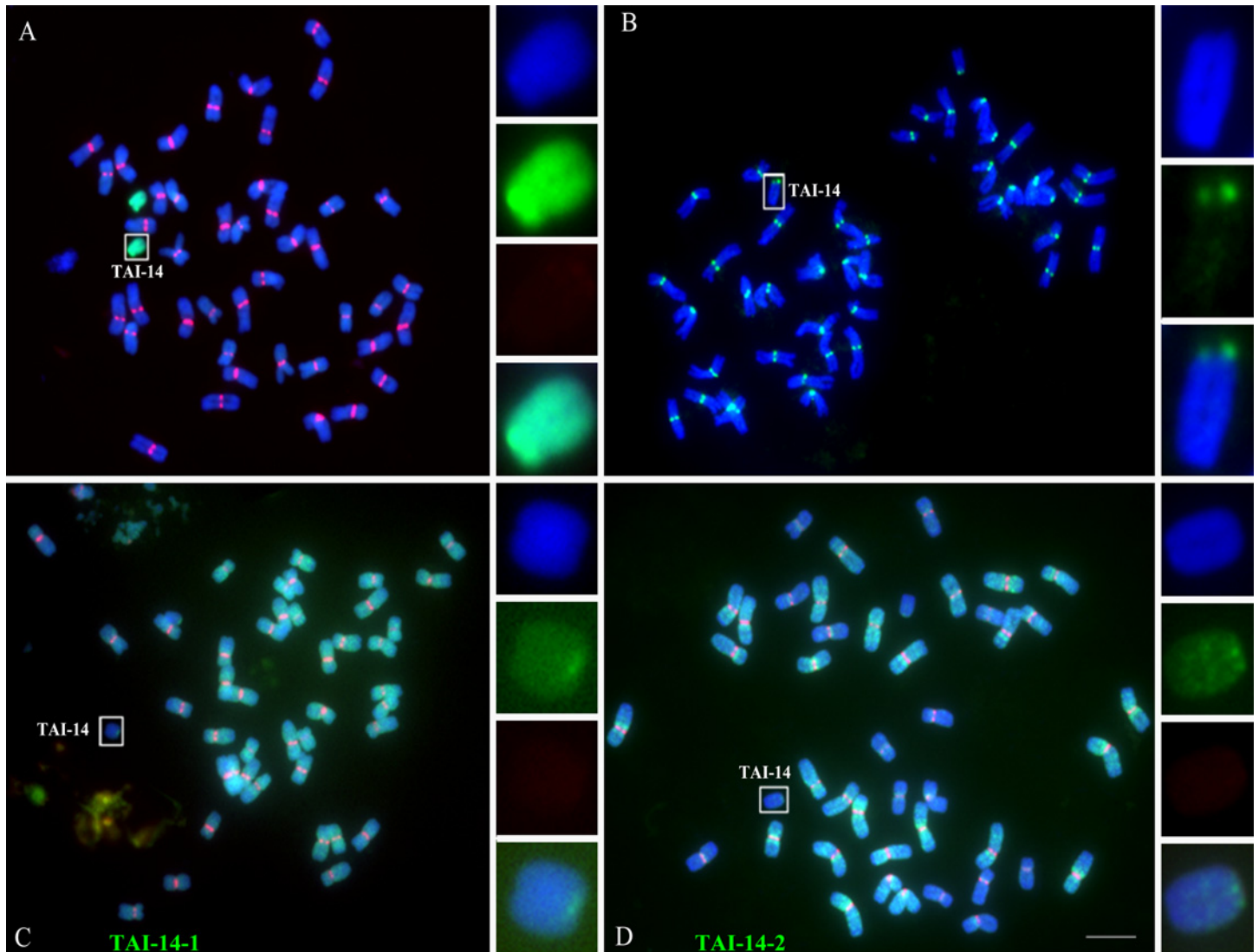
**Fig 3. Centromeric sequence deletion in the ditelosomic line 4DS.** (A). FISH on 4DS with CRW (red). DAPI staining is labeled in blue. (B). Intensity of CRW fluorescence in 4DS compared with the controls in Chinese Spring. The Y-axis represents the relative CRW fluorescence intensity. The error bars show the standard deviation (s.d.). The double asterisks denote significant differences at  $P < 0.001$  (two-tailed Student's t-test).  $P = 3.59E-15$ . (C). Multi-color FISH on 4DS. The DNA for the wheat A genome is labeled in green, the DNA for the D genome is labeled in red, and the DNA for the B genome is used as a block. (D)-(F). Immunostaining results of 4DS with antibodies against CENH3 (D), H3 phosphorylation at Ser-10 (E), and H2A phosphorylation at Thr-133 (F) are shown in green. The insets show high-magnification images of chromosome 4DS. Bar = 10  $\mu$ m.

doi:10.1371/journal.pgen.1005997.g003

chromosomes of *Th. elongatum* in the addition line derived from *T. durum* Kekeruite  $\times$  8802 (Fig 5A). However, the chromosomes that lack CRW sequences were stably transferred to the next generation, indicating that functional centromeres were formed on these chromosomes.

### Novel sequences are involved in *de novo* centromere formation on an alien chromosome

Without a reference genome sequence for *Th. intermedium*, it is difficult to characterize the sequences involved in the *de novo* formation of centromere in TAI-14. We designed a new strategy to isolate the neocentromere sequences based on CENH3-ChIP and dot-blot methods (see Methods). The CENH3-ChIP-enriched DNAs in the control Chinese Spring (abbreviated as CS) and TAI14 were further analyzed by dot-blotting. The signals from the dot-blots that were significantly different between CS and TAI14 (e.g., signals present in TAI14 and not in CS) were treated as potential elements involved in *de novo* centromere formation. Two sequences, TAI-14-1 and TAI-14-2, were identified as new centromeric sequences for the new centromere in TAI-14 (Fig 4C and 4D), and both sequences showed homology to known retrotransposons by alignment to a BAC genome sequence in wheat (S3 Fig). TAI-14-1 was widely dispersed on nearly all of the chromosomes, whereas TAI-14-2 was mainly detected in the pericentromeric regions of the wheat chromosomes (Fig 4C and 4D). We observed that TAI-14-2



**Fig 4. Deletion of centromeric sequences in the wheat-*Th. intermedium* addition line TAI-14.** (A). Genomic *in situ* hybridization (GISH) and FISH of somatic metaphase chromosomes in TAI-14. The genomic DNA of *Th. intermedium* is labeled in green, the CRW sequences are labeled in red, and DAPI staining is labeled in blue. (B). Immunostaining of TAI-14 with antibodies against CENH3 (green). (C) and (D). FISH analysis of the novel centromeric sequences TAI-14-1 (C) and TAI-14-2 (D), which are labeled in green in TAI-14. The CRW sequences are labeled in red. The insets show high-magnification images of the alien chromosomes in TAI-14. Bar = 10  $\mu$ m.

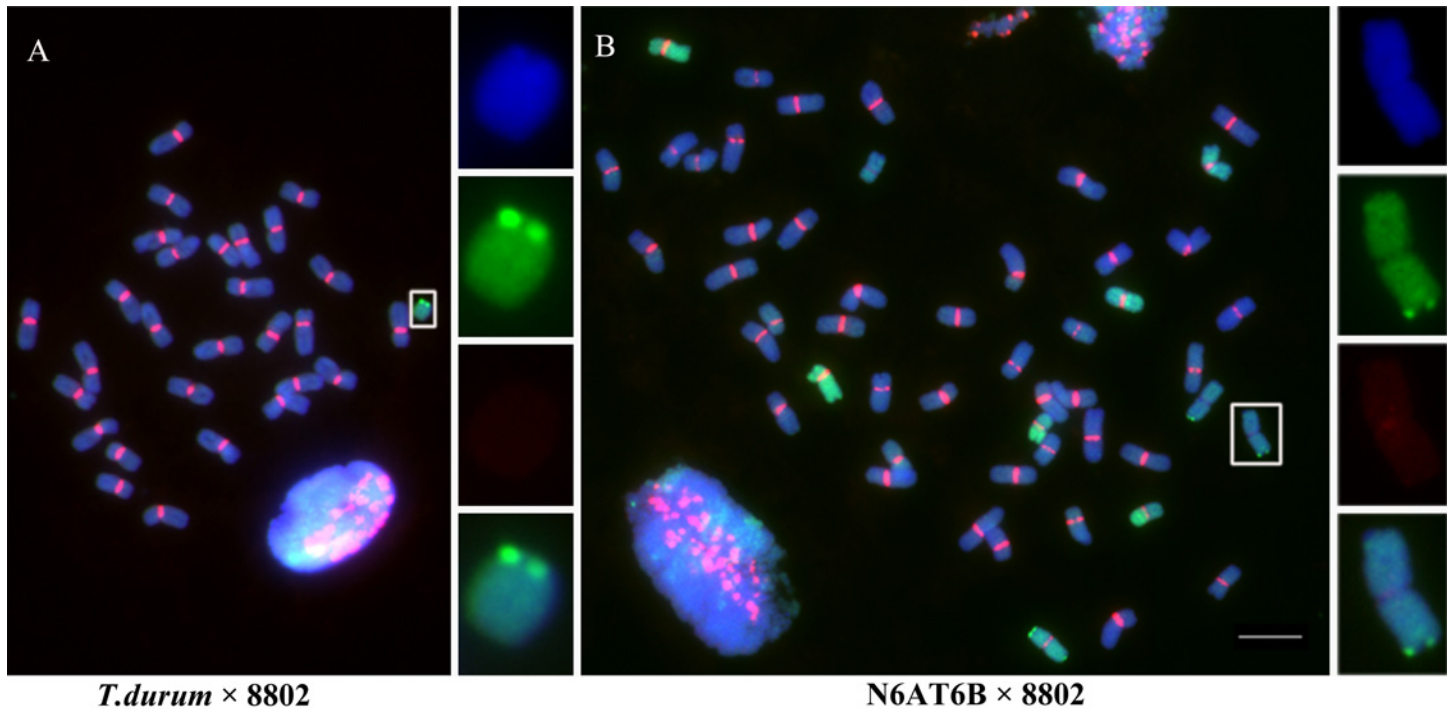
doi:10.1371/journal.pgen.1005997.g004

was located at both the centromeric and pericentromeric regions on different chromosomes in *Th. intermedium* (S1B Fig). Interestingly, some chromosomes of *Th. intermedium* that showed less CRW distribution were accompanied by more TAI-14-2 sequences occupancy in the centromeric region (S2D and S2E Fig). This result suggests that CRW and TAI-14-2 may be complementary centromeric sequences in some chromosomes of *Th. intermedium*.

### A 994-kb genomic sequence of the 4DS chromosome is involved in *de novo* centromere formation in 4DS

The total loss of CRW sequences and the presence of CENH3 in the 4DS ditelosomic line suggest that a new centromere formed on the 4DS chromosome (Fig 3). Because sequenced genomes published for *Ae. tauschii* (wheat D genome donor) and CS [58,59], we performed a





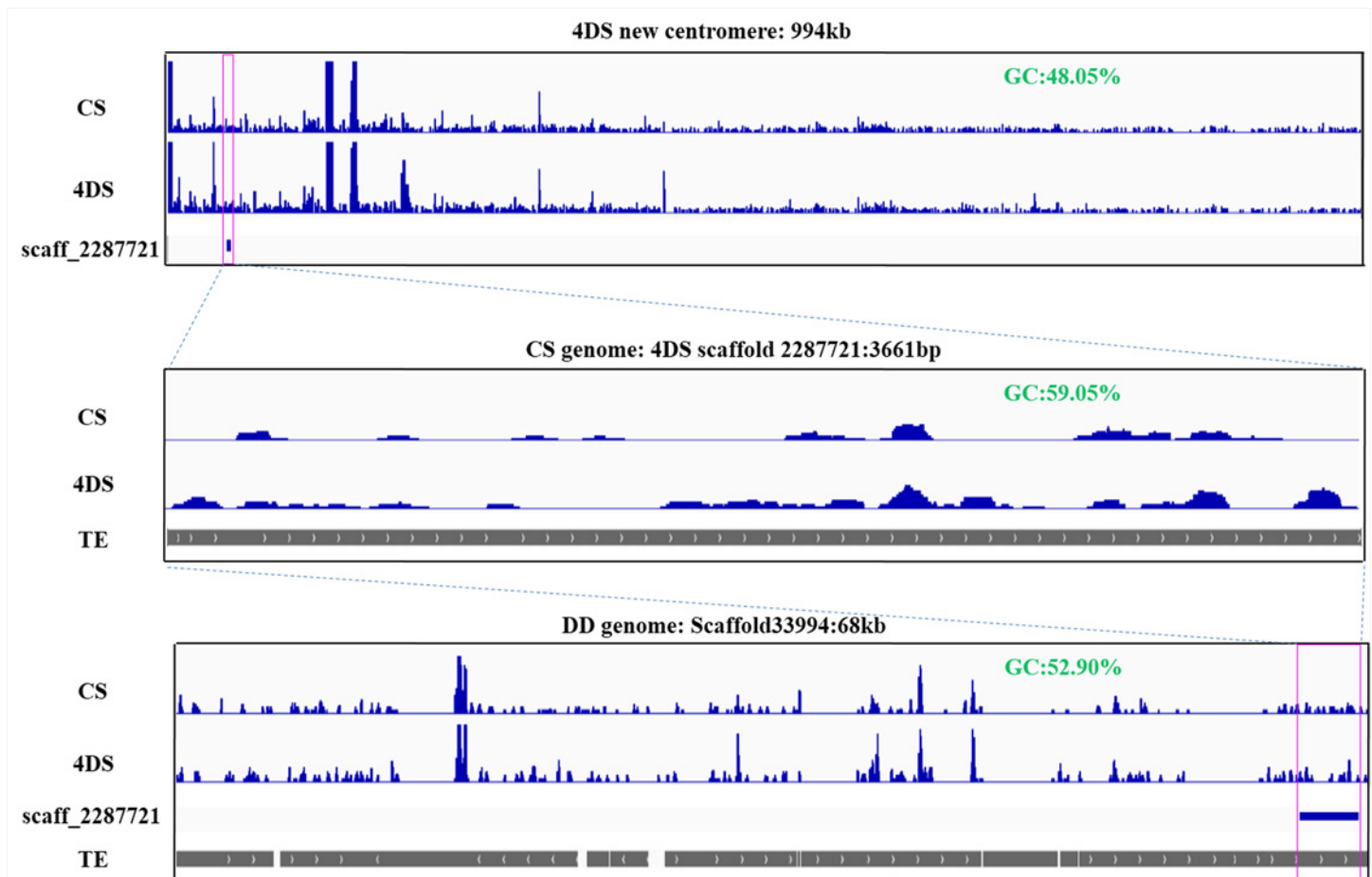
**Fig 5. Deletion of centromeric sequences in the wheat-*Th. elongatum* addition lines.** The genomic DNA of *Th. elongatum* is labeled in green, the CRW sequences are labeled in red, and DAPI staining is labeled in blue. (A) FISH analysis of a ditelosomic addition line from the hybrid of *T. durum* and 8802. There are no detectable FISH signals in the new telosomic chromosomes when CRW is used as a probe. (B) FISH analysis of an addition line from the hybrid of N6AT6B×8802 that shows weak CRW FISH signals. The insets show high-magnification images of the alien chromosomes with centromere changes from the addition lines. Bar = 10  $\mu$ m.

doi:10.1371/journal.pgen.1005997.g005

ChIP-seq analysis with CENH3 antibodies and identified potential sequences in the new centromere on the 4DS chromosome. Because the genomes were not completely assembled, we chose a mapping strategy that equally mapped all reads to multiple loci. The raw reads of the 4DS and CS (as control) samples were mapped to the wheat D and CS genomes, respectively, using BWA software (S1 Table) [60].

Using the sequence of the CS genome as a reference, we identified 107 scaffolds on the short arm of the 4D chromosome, with different CENH3 enrichments between 4DS and CS. The sizes of the 107 scaffolds ranged from 1,594 to 32,269 bp, and these scaffolds were combined into a 994-kb region containing only 11 genes that code for ribosomal and photosystem proteins (Fig 6, S2 Table). Furthermore, we selected one of the assembled scaffolds, IWGSC\_CSS\_4DS\_scaff\_2287721 (3665 bp), as a FISH probe and confirmed its localization in the 4DS *de novo* centromere (Fig 7A). We mapped this scaffold to the genome of *Ae. tauschii* using BLASTN [61] and identified a 68-kb fragment (Scaffold 33994) that contained most of the sequences of the scaffold that showed mapping differences between the 4DS and CS (Fig 6). Both of the sequences, the homologous 3,665-bp sequence from the CS genome and the 68-kb sequence from the D genome, contained many transposable elements and similar GC levels (48.05% and 52.90%, respectively, Fig 6). Due to the incomplete genome sequence, we tentatively suggest that the partial sequences of the 994-kb region in the wheat CS genome and the 68-kb region in wheat D genome may underlie *de novo* centromere formation in 4DS.

In addition, the same strategy that was used to isolate the neocentromere sequences on TAI-14 was employed with 4DS to better understand the sequences in the 994-kb region. A 769-bp fragment (named 4DS-1) near the original centromere was identified as a candidate



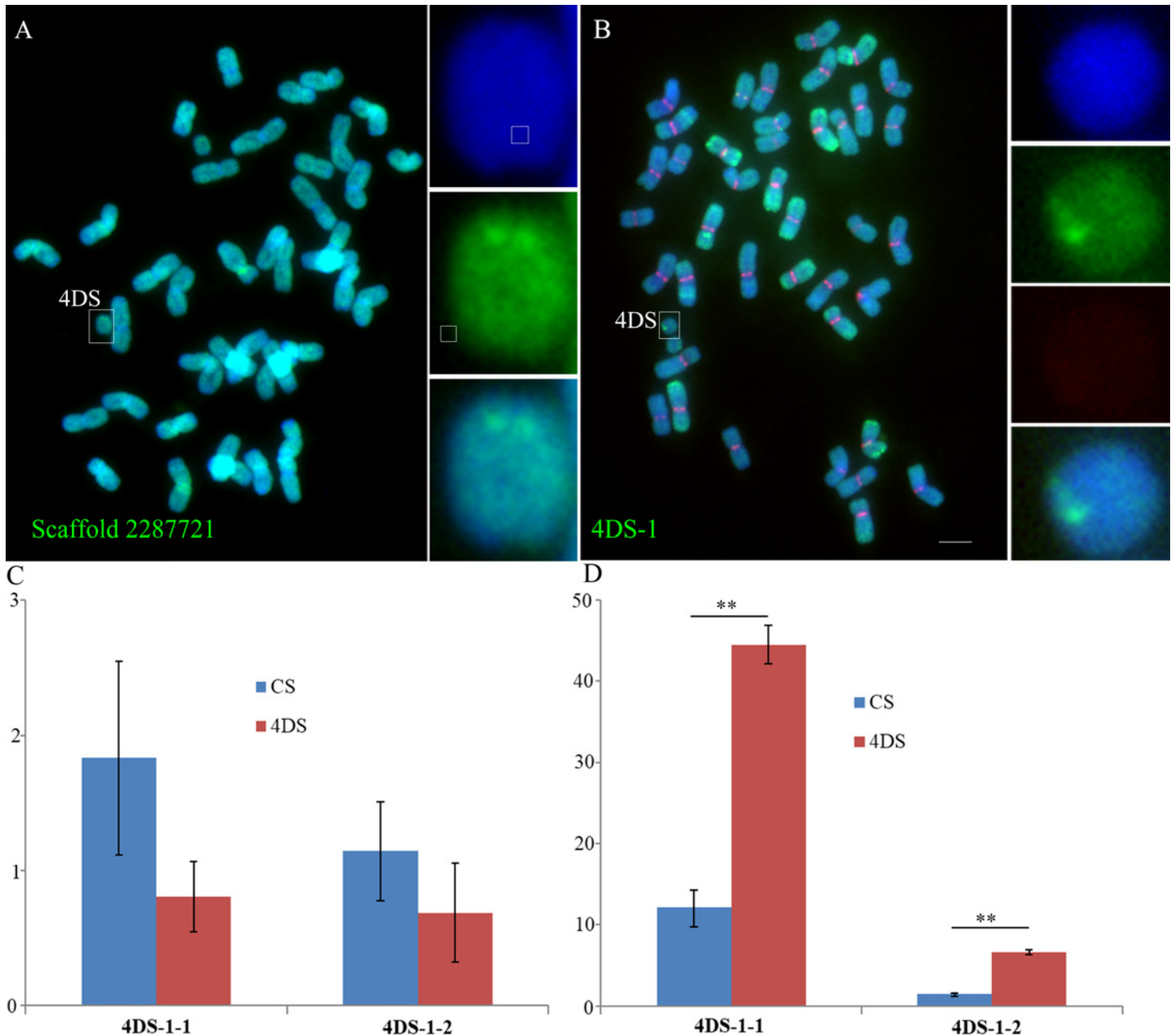
**Fig 6. ChIP-seq mapping results of the ditelosomic line 4DS with the CENH3 antibody.** The top box shows the ChIP-seq results of the control Chinese Spring (CS) and 4DS by mapping to the CS genome. A 994-kb region was detected as a difference between CS and 4DS and is involved in 4DS neocentromere formation. 4DS Scaffold 2287721 (3,661 bp, indicated by the magenta box) was selected as a probe for FISH confirmation. The scaffold was enriched with transposable elements, as shown in the middle box. In the wheat D genome, Scaffold 33994 (~68 kb, anchored to chromosome 4D) has high homology with Scaffold 2287721 and shows a mapping difference between CS and 4DS. The 994-kb region, Scaffold 2287721 and Scaffold 33994 had high GC contents.

doi:10.1371/journal.pgen.1005997.g006

sequence in the 4DS *de novo* centromere (Figs 7B and S4A). 4DS-1 was present at multiple locations on the chromosomes from the A and D genomes. FISH detection showed that it was localized to sites near the normal centromere on the 2A, 7A, 7D and 2D chromosomes (S4B and S4C Fig). However, FISH signals of 4DS-1 were not observed in the chromosomes from the B genome (Figs 7B and S4). 4DS-1 was mapped to two scaffolds (Scaffold 10770 and Scaffold 28550) from the 4DS chromosome and showed mapping differences between 4DS and CS, which confirmed that 4DS-1 was a part of the *de novo* centromeric sequence in the 4DS ditelosomic line (S5 Fig).

### Transcription and histone modification accompanied *de novo* centromere formation in 4DS

A previous study showed that the transcripts of centromeric sequences can function as essential components of centromere structure and activity [37]. Histone modifications, such as methylation and phosphorylation, are important regulators of centromere stability and activity [35,38]. For sequence 4DS-1 in the 4DS *de novo* centromere, we tested whether there were changes in



**Fig 7. FISH of a novel centromeric sequence in 4DS identified by ChIP-seq mapping and dot-blot hybridization.** (A). Scaffold 2287721 (labeled in green), which was obtained from ChIP-seq, contained sequences of the *de novo* centromere in 4DS. (B). 4DS-1 (labeled in green), which was obtained from the dot-blot, was detected as a novel centromere sequence in 4DS. The CRW sequences are labeled in red. The insets show high-magnification images of 4DS. Bar = 10  $\mu$ m. (C). RT-qPCR analysis of 4DS-1-1 and 4DS-1-2. (D). The RNA-ChIP-qPCR results for 4DS-1-1 and 4DS-1-2 suggest that transcription levels of both sequences at the 4DS centromeres were increased.

doi:10.1371/journal.pgen.1005997.g007

transcription and the transcripts that interacted with CENH3 via RT-qPCR and RNA-CENH3-ChIP. We selected two fragments (~300 bp) of the 4DS-1 sequence, termed 4DS-1-1 and 4DS-1-2. Compared with CS, the transcripts of 4DS-1-1 and 4DS-1-2 were slightly but not significantly decreased in 4DS (Fig 7C). However, the amounts of the 4DS-1-1 and 4DS-1-2 transcripts that were associated with CENH3 were remarkably increased in 4DS compared

with CS (Fig 7D). These results suggest that increased transcription of the corresponding sequences may accompany *de novo* centromere formation in 4DS. We also checked the possible changes in six histone modifications between the normal and *de novo* centromeres via immunoassay. No significant signals for the euchromatin marks H2AZ and H3K4me3 were detected in either centromere (S6A and S6B Fig), and enrichment of the euchromatin-related histone mark H3K4me2 was discernible for both centromeres (S6C Fig). Compared with non-centromeric chromosome ends, both centromere types revealed a reduction in the heterochromatin marks H3K27me2 and H3K27me3 (S6D and S6E Fig), whereas there were no obvious differences in H3K9me2 in the two centromere types (S6F Fig). In general, there were no significant differences in the accumulation of most euchromatic or heterochromatic histone markers between the normal and *de novo* centromeres.

### Multi-locus centromere formation in wheat wide hybrids

We crossed the hexaploid amphidiploid 8802 with the *T. aestivum* Chinese Spring nulli-tetrasomic lines to establish new substitution lines for the chromosomes from the E genome. All chromosomes from 8802 have only one centromere, as determined by the FISH signals of CRW (S7A Fig). Chromosomes with two regions containing CRW sequences were identified in the F1 hybrids of the nulli-tetrasomic (3, 5 and 6 homologous groups) lines × 8802 (S7C–S7F Fig). In the F5 generation of the hybrids between nulli-tetrasomic line N6AT6B ( $2n = 42$ ) × 8802, chromosomes containing two regions with centromeric sequences were inherited from the F1 generation, but only one region was functional (Fig 8A and 8E). In addition to the two-locus centromere, a three-locus centromere was discovered on an alien chromosome in the F5 generation (Fig 8B), and the middle centromere region was shown to act as the functional centromere (Fig 8F). Furthermore, two different three-locus centromeres were observed in the F6 generation (S8C and S8D Fig).

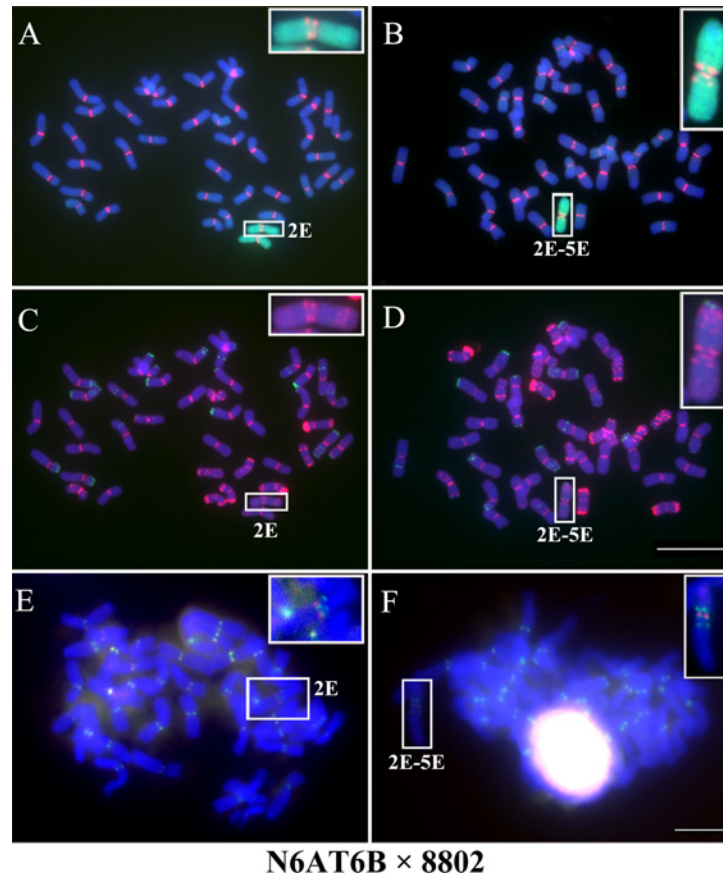
The repetitive sequences pAs1 and pSc119.2 were used to karyotype the chromosomes with two- or three-locus centromeres. The chromosomes with two centromeric regions originated from the 2E chromosome of 8802, but the three centromeric regions were produced by the combination of sequences from two different arms of chromosomes 2E and 5E, rather than direct inheritance from any chromosome (Figs 8B, 8D and S7B). Furthermore, two different three-locus centromeres were observed in the progeny. One progeny contained an isochromosome 2ES, and the other progeny contained a chromosome produced from the wheat 6BL chromosome and the *Th. elongatum* 2E chromosome (S8A and S8B Fig). However, neither was stably transferred to the next generation.

Multi-centric chromosomes are frequently formed in the hybrids of wheat and related species, such as *Th. elongatum*, *Th. poticum*, *Th. intermedium*, *Agropyron cristatum*, *Hordeum vulgare* and *S. cereale* (S9 Fig). However, unlike the two-locus centromere in the hybrids of N6AT6B × 8802, both centromeres in these dicentric chromosomes were active, which caused chromosomal loss in the next generation.

### Expansion of centromeric retroelements in wheat-rye addition lines

Heterochromatin alterations and chromosomal rearrangements associated with centromere changes have been reported in derivatives of wheat-rye hybrids [62]. The chromosomes containing altered centromeres were lost in the next generation. Here, we discovered another wheat-rye hybridization-promoted centromeric retrotransposon expansion in different rye addition lines. These changes can be stably transmitted to offspring.

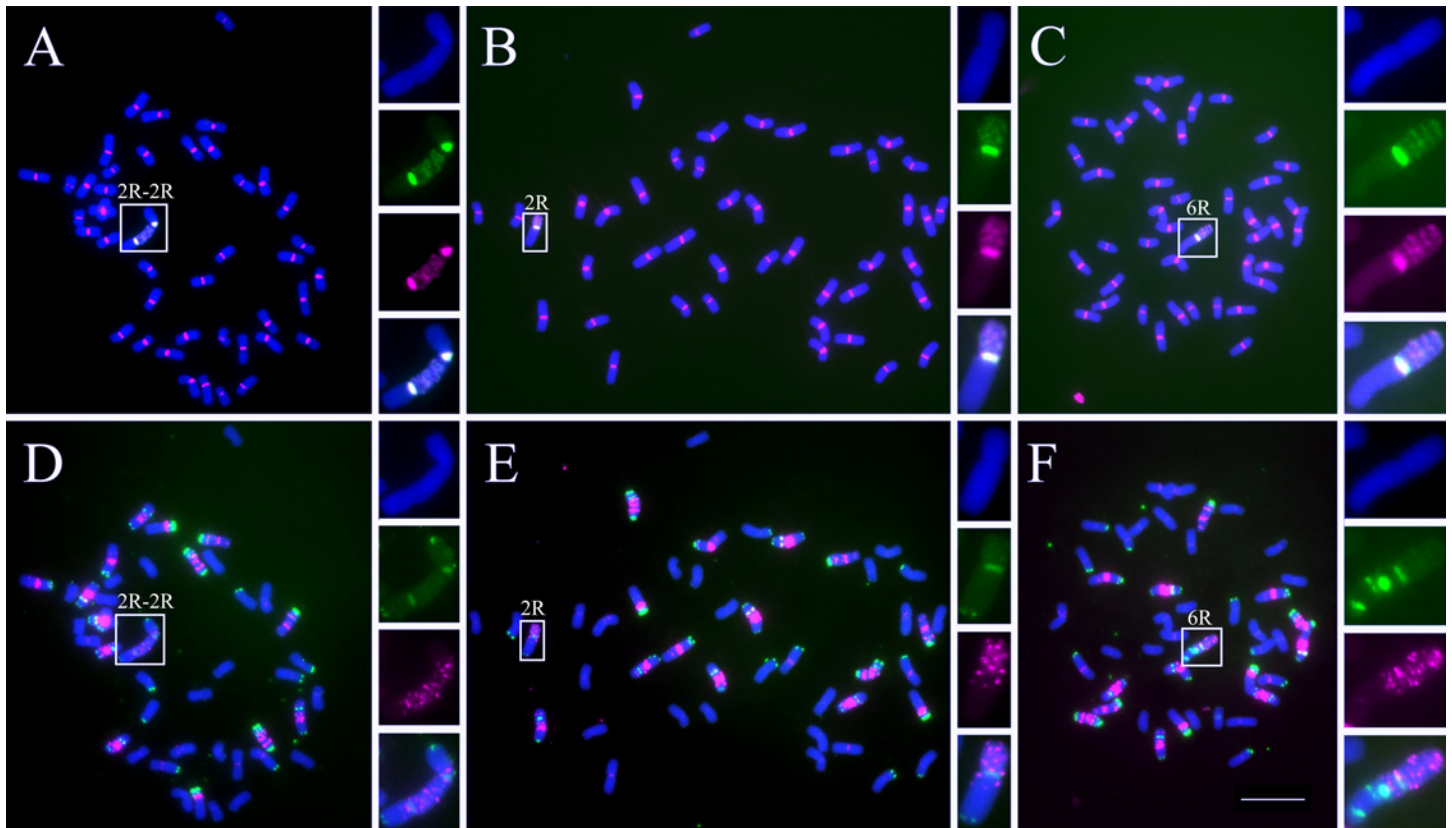
The rye addition lines were generally obtained from successive backcrossing between wheat and triticale. A novel octoploid triticale ( $2n = 56$ ) was generated by hybridization between *T.*



**Fig 8. Two- and three-locus centromeres in the F5 generation of hybrids between nulli-tetrasomic N6AT6B× 8802.** (A) and (B). FISH and GISH of chromosomes with two- and three-locus centromeres, respectively. The genomic DNAs of *Th. elongatum* are labeled in green, the CRW sequences are labeled in red. DAPI staining is labeled in blue. (C) and (D). Karyotype analysis of chromosomes with two- and three-locus centromeres using probes pAsI (green) and pSc119.2 (red). (E) and (F). Immunostaining-FISH analysis of chromosomes with two- and three-locus centromeres. The CRW sequences are labeled in green, and CENH3 is labeled in red. The insets show high-magnification images of the chromosomes with two- and three-locus centromeres. Bar = 10 μm.

doi:10.1371/journal.pgen.1005997.g008

*aestivum* Mianyang 11 (2n = 42) and *S. cereale* Kustro (2n = 14). In the wheat and octoploid triticale hybrids, a novel chromosome emerged after the joining of the two 2R chromosomes (2R-2R). Compared with the normal centromeres of chromosomes 2R and 2RL, the centromere in the 2R-2R chromosome was drastically expanded and was much larger than the 2RL arm (Figs 9A, 9D and S10). Further analysis showed that this large centromere consisted of two normal centromeres from chromosome 2R and a centromere-like region between them with dispersed pAWRC.1 (rye-specific centromeric retrotransposon) sequences [63] (Fig 9A). Dispersed centromeric retrotransposons may function as a part of an active centromere, as 2R-2R was broken into smaller fragments after self-pollination (Fig 9B and 9E). In the progeny, we detected a new 2R chromosome (smaller than the canonical 2R) that retained a region with dispersed pAWRC.1 sequences and was approximately half the size of 2R-2R (Figs 9B, 9E and S10B). A novel chromosome 6R contained pAWRC.1 sequences in a region near the functional centromere in the 6R addition line (Figs 9C, 9F and S10B). However, these regions did not have centromere activity in these progeny (S11 Fig).



**Fig 9. Centromere expansion in the wheat-rye addition lines.** (A)-(C). FISH indicates centromere expansion in the 2R-2R, novel 2R and 6R addition lines using the pAWRC.1 (green) and CRW (red) probes. (D)-(F). Karyotype analysis of centromere expansion in the 2R-2R (D), novel 2R (E) and 6R (F) addition lines using the AAC (red) and pSc119.2 (green) probes. The insets show high-magnification images of the chromosomes with expanded centromeres. Bar = 10 μm.

doi:10.1371/journal.pgen.1005997.g009

## Discussion

### *De novo* centromere formation adjacent to native centromeres depends on unique epigenetic modifications in wheat and its wide hybrid

Functional centromeres without classic centromeric sequences have been reported in humans, fungi, and plants [29–32,64,65]. Here, we found that most neocentromeres in wheat and its addition lines consist of genomic sequences that have loose resemblance to the sequences of normal centromeres. The sequences involved in the *de novo* centromere formation in the ditelosomic lines 4DS and TAI-14 were located at the chromosome arms adjacent to the native centromere. The 4DS-1 sequence was located very near the centromeres of chromosomes 4D, and sequence TAI14-2 was detected in the pericentromeric region before the new centromere had formed (Figs 7B and S4). The preferential localizations of the *de novo* centromeres in wheat and its addition lines were similar to some of the newly formed centromeres in chicken, *Candida albicans* and a wheat-barley addition line [66,67]. However, in contrast to wheat, human neocentromeres can originate at multiple positions on different chromosomes through the binding of essential kinetochore proteins [30]. In addition, neocentromeres also have been observed at multiple locations on chromosome arms in other plants [29,32,65]. This difference suggests that the *de novo* centromere formation was not dependent on location and occurred under appropriate epigenetic conditions, as indicated by previous studies [28,31,32].

We observed that CENH3, H2AT133ph and H3S10ph were deposited in all of the newly formed centromeres in wheat and its wide hybrids (Figs 2D, 2H, 2L, 3D–3F and 4B). With the exception of CENH3 or CENP-A, other histone modifications in the centromeric region are also critical elements for centromere stability and activity. CENP-A nucleosomes are interspersed with the H3K4me2 nucleosomes within the centromeric chromatin of humans and flies [68,69]. Enrichment of H3K9me2 and H3K9me3, but not H3K4me2, has been observed in maize centromeres [70,71]. Furthermore, H3K4me2, H4ac and H3K27me1 were enriched in the centromeric chromatin of rice [72]. We observed that H3K4me2 was slightly enriched in the normal and newly formed centromeres of wheat, similar to previous reports in humans, flies and rice. However, weak signals for H3K27me and H3K9me2 were observed in these two centromeres, which were different from the centromeres of maize and rice (S6 Fig). Based on our results, the histone modifications in the new centromeres are similar to the normal centromeres, highlighting that possibility that unique histone modifications in wheat centromeric regions may be required for *de novo* centromere formation and stability.

Noncoding RNAs from the centromeric sequences directly interact with the kinetochore and recruit CENPC to centromeres [73]. In maize, transcripts of centromeric retrotransposons and repeat sequences have been associated with the CENH3 protein, as detected by ChIP with anti-CENH3 antibodies [74]. Our results showed that the expression of neocentromeric sequence 4DS-1 was greatly elevated during the process of *de novo* centromere formation (Fig 7C and 7D). The transcripts of centromeric sequences may serve as structural and regulatory components of the centromere. In fact, its transcription process, rather than the transcripts themselves, may have facilitated CENP-A deposition and nucleosome assembly at the centromere [35,37]. *De novo* centromere formation in 4DS increased the expression of the 4DS-1 sequence, which may provide the RNAs that regulate CENH3 deposition and recruit epigenetic elements [35,37].

### Highly variable of new centromeric sequences between different chromosomes and species

For cereal centromeres, two common sequences, described as *cereal centromeric sequence* (CCS1) and Sau3A9, were first reported in *Brachypodium* and *Sorghum bicolor* [75,76]. In most cereals, these centromeric sequences represent parts of the *Ty3/gypsy* retroelement, e.g., ‘*cereba*’ of barley, implying that cereals maintain conserved retroelements in their centromeric regions [77]. However, the tandem repeats in the centromere are species-specific [77]. We observed that the pericentromeric sequence TAI-14-2 of wheat was located in the centromeric region of several chromosomes of the wild relative *Th. intermedium* (S2 Fig). The centromeric sequences may undergo insertion and amplification during species differentiation [20,25,51], providing insight into the mechanism of differential centromeric sequences change in wheat and its wild relatives.

Centromeres on different chromosomes of one species may evolve rapidly and are independent of each other [20]. The novel centromeric sequence 4DS-1 displayed a specific localization pattern in the different chromosomes of the wheat subgenomes, and its chromosomal location was very near the centromeres of chromosomes 7A, 7D, 2A and 2D (Figs 7B and S4). However, the sequence was not readily detectable on chromosomes 2B and 7B, and it was not found in other chromosomes, such as 4A and 4B. Similar to the situations in rice, maize, and potato, the centromeres on different chromosomes experienced differentiation by sequence loss and insertion [20,25,51].

Centromeric protein complexes, which are necessary for proper chromosome segregation, can be formed by the speciation factors HMR and LHR to mediate hybrid sterility and

incompatibility in *Drosophila* [78]. During hybridization, the intragenomic conflicts of different centromeres may cause incompatibilities in hybrids [79,80]. These results strongly suggest that centromere divergence has an important impact on the generation of biodiversity [78]. To adapt to highly variable centromere DNA sequences, the centromere proteins undergo rapid evolution to maintain a functional centromere [81]. We observed that the new centromeric DNA sequences are highly variable between different genomes and chromosomes, which may allow the centromeric proteins to mediate intragenomic incompatibility and genomic specificity in the nascent hybrids. Although detailed reports are not available, it is likely that centromere sequence diversity has an important impact on speciation. An understanding of the diversity of centromeric sequences and its link to speciation and genomic stability deserve further analysis.

### Most variations in the centromeric sequences occur during wide hybrid formation

Multiple centromeres, holocentromeres and neocentromeres formation implies that centromere positions may be alterable rather than permanently fixed [11,31,32,82–84]. Wide hybridization can trigger chromosomal rearrangements and genome reorganization, accompanied by centromere alterations [62,85]. Unstable di-centromeres and multiple centromeres have been associated with the formation of inter- and intra-chromosomal translocations in wheat-rye hybrids [62]. Our results demonstrated that multi-locus centromere formation, centromere expansion, and canonical centromeric sequence elimination may yield novel chromosomes in wheat and its wide hybrids (see summary Fig 1).

Chromosomes with two regions containing centromeric sequences were observed in the F1 hybrids of three null-tetra lines and 8802 (S7 Fig). During meiosis I of the F1 generation, several univalents of the E genome in amphidiploid 8802 may experience centromere breakage. Centromere misdivision, which depends on the orientation of the univalent, may occur across the either centromere or the pericentric chromatin, but chromosome fragments containing centromeric and pericentric regions may survive [86]. This observation suggests the possibility that chromosomal fragments containing centromeric and pericentromeric regions were rejoined to a novel chromosome and induced the formation of two- and three-locus centromeric regions. Centromere inactivation allows dicentric chromosomes with only one functional centromere to be stably transmitted to the next generation (Fig 8E), similar to stable dicentric A-A and A-B chromosomes in maize [18,33,34]. However, the chromosome containing a three-locus centromere still suffered from centromere breakage, which led to its structural alterations in the progeny. As a consequence of a translocation in the nulli-tetrasomic line N6AT6B, a chromosome with a three-locus centromere included the 2E chromosome from 8802 and the 6B chromosome from N6AT6B (S8D Fig). Similar to rye and maize [18,86,87], intrachromosomal recombination and centromere breakage likely promoted the formation of multi-locus centromeres and novel chromosomes in wheat and its wide hybrids.

Retrotransposons can be activated during wide hybridization [88]. Interspecific hybrids triggered the amplification of centromeric satellite repeats and retrotransposons [85]. In a wheat-rye hybrid, we observed two chromosomes that had likely lost their telomeres and fused into one chromosome, as previously suggested [89]. The wide hybridization affected the stability of the centromeric retrotransposons and the activation of rye-specific retrotransposons in the 2R-2R and 6R addition lines, causing centromeric sequence expansion (Fig 9A and 9C). These unstable centromeres may subsequently lead to chromosome breakage and different progeny with expanded centromeres, suggesting that centromere variants may trigger the formation of novel chromosomes. Total centromere size has been postulated to be positively correlated with genome size rather than chromosome size [90]. Centromere domains of several



maize chromosomes (average size 1.8 Mb) expanded to 3.6 Mb in the background of the oat genome [48]. However, the expanded centromere on chromosome 2R-2R may instead be a general response to genomic stress following the wheat-rye hybridization, rather than an adaptation to the wheat centromere size.

In summary, we observed that the elimination, rearrangement and expansion of centromeric sequences affect chromosome morphology and maintenance in wheat wide hybrids. *De novo* centromere formation promotes the accurate segregation of chromosomes that experienced centromere sequence elimination. The new centromeres have low sequence homology but high epigenetic similarity to normal centromeres. The highly variable centromeric sequences between genomes and chromosomes may facilitate genomic specificity and differentiation in hybrids. The centromeric sequences involved in *de novo* centromere formation are mainly retrotransposon-like sequences, and their RNAs are transcribed at high levels. Multiple centromeres and centromere sequence expansion strongly influence centromere activity and cause chromosome breakage and rearrangements. More importantly, centromere variations in the ditelosomic lines 4DS, 1BS, 5DL, 5DS and TAI-14 may specifically affect the size and DNA organization of normal chromosomes. Thus, these may be useful for future studies of chromosome sorting and sequencing.

## Methods

### Plant materials

The hexaploid amphiploid 8802 (AABBEE) originated from hybrids between *T. durum* and *Th. elongatum* [57]. The addition lines of wheat (Chinese Spring)-*Th. poticum*, *Th. intermedium* and *Agropyron cristatum* were produced by our laboratory. The addition lines of wheat (Mianyang 11) and *S. cereale* (Kustro) were produced by Dr. Shulan Fu, Sichuan Agriculture University, Chengdu, China. The nulli-tetrasomic lines and the ditelosomic lines 4DS, 1BS, 5DL and 5DS were kindly provided by Dr. Perry Gustafson, University of Missouri, Columbia, MO, USA.

### Cytological preparation, probe labeling and images processing

The root tips were prepared for the FISH experiments and probes as previously described [91]. The CRW and other genomic DNAs were labeled with Alexa Fluor-488-dUTP (green) or Alexa Fluor-594-5-dUTP (red) as needed. The root tip samples of the different wheat lines were treated with the same conditions, and with an equal amount of probe. The FISH images were acquired using an epifluorescence Olympus BX61 microscope (Olympus China Inc, Beijing, China) with the same exposure time and were processed with Adobe Photoshop CS 3.0. For analysis, the fluorescence was quantified using ImageJ [92]. At least 20 cells from three different plants were counted in each of the different wheat lines. Significant differences were calculated using Microsoft Excel and Student's t-test (two-tailed).

### Immunolocalization in somatic cells

The root tips were fixed with 4% formaldehyde in 1×PBS for 1 h, and the metaphase chromosomes were prepared as previously described [17]. The wheat-specific CENH3 antibodies were produced in our laboratory. The pH2AThr-133 and pH3Ser-10 antibodies were described previously [16,17].

### ChIP-seq, RNA-ChIP-qPCR and RT-qPCR

Chromatin immunoprecipitation (ChIP) was performed according to a previously described method [93]. Approximately 20 g of fresh leaf tissue from Chinese Spring and 4DS was

prepared for CENH3-ChIP (CENH3 antibody used as described above). ChIP-seq was conducted according to a previously described method [32]. Using an Illumina HiSeq2000 platform, the enriched DNA samples were sequenced to generate paired-end 100-bp sequence reads. RNA-ChIP was performed using a method that was similar to ChIP [93]. RNase activity was inhibited using Recombinant RNase inhibitor (RRI) in the RNA-ChIP process. The RNA was extracted from the sample using TRIzol reagent. The RNA was reverse-transcribed into cDNA using M-MLV reverse transcriptase (Promega) and random primers (New England Biolabs). The qPCR protocol was performed as described [94].

### ChIP-seq data analysis

Nearly 14,000 Mbp of raw ChIP-seq paired-end 100-bp reads were mapped to the wheat CS and wheat D genomes using BWA software [60]. All mappable reads were randomly assigned a locus from the possible options, and the duplicated reads were removed. The reads per million (RPM) values of every scaffold in the genome were calculated to show the normalized enrichment. We selected the scaffolds that had enriched reads in both 4DS and the control, with the RPM ratios between two samples  $\geq 20$ , or scaffolds that only had enriched reads in 4DS, with read counts  $\geq 20$ . IGV Tools was used to visualize the normalized read distribution of each scaffold [95]. The anti-CENH3 ChIP-seq data were deposited in the Gene Expression Omnibus (GEO) database under number GSE63752.

### Dot-blot hybridization of the ChIP-enriched DNAs

The ChIPed DNAs from CS and 4DS were amplified by Dop-PCR (primer sequence was 5'-CCGACTCGAGNNNNNNATGTG G-3') [96]. The DOP-PCR product was used as the DNA target, and the ChIPed DNAs were used as probes. The dot-blot protocol was performed as described [97]. The differences in the sequences between CS and 4DS that were determined by the dot-blot hybridization were selected as candidate centromeric sequences for FISH.

### Wheat genetic nomenclature

**Addition line.** An individual with addition chromosomes from other species.

**Ditelosomic line.** An individual with a pair of chromosome arms.

**Nulli-tetrasomic line.** An individual lacking one pair of chromosomes and containing a doubled pair of other chromosomes.

### Supporting Information

**S1 Fig. Centromeric sequences in the different chromosomes of Chinese Spring.** (A) FISH of the CS chromosomes with CRW (red). DAPI staining is labeled in blue. (B) Karyotype analysis of CS using the pAsI (green) and pSc119.2 (red) probes. DAPI staining is labeled in blue. Bar = 10  $\mu\text{m}$ . The 1B, 4D and 5D chromosomes are indicated by green, white, and red arrows, respectively.

(TIF)

**S2 Fig. Novel centromeric sequences, TAI-14-2, on chromosomes in *Th. intermedium*.** (A)-(C) FISH of TAI-14-2 on chromosomes in *Th. intermedium*. The CRW sequences are labeled in red, TAI-14-2 is labeled in green, and DAPI staining is labeled in blue. Bar = 10  $\mu\text{m}$ . (D)-(E) Line profile plots of the distributions of the intensity of the CRW and TAI-14-2 signals along the white line of selected chromosomes from (C).

(TIF)

**S3 Fig. Alignment of TAI-14-1 and TAI-14-2 DNAs to the BACs of wheat chromosome 3B.** (TIF)

**S4 Fig. Cytological analysis of 4DS-1 in Chinese Spring and the ditelosomic line 4DS.** (A) FISH of 4DS-1 on chromosomes in Chinese Spring. The CRW sequences are labeled in red, 4DS-1 is labeled in green, and DAPI staining is labeled in blue. The arrows indicate the 4D chromosomes. (B) Karyotype analysis of 4DS-1 in 4DS using the pAsI (green) and pSc119.2 (red) probes. DAPI staining is labeled in blue. The 2A, 7A, 2D, 4D and 7D chromosomes are indicated in the figure. (C) Multi-color GISH of line 4DS. The DNA for the wheat A genome is labeled in green, the DNA for the D genome is labeled in red, and the DNA for the B genome is used as a block. Bar = 10  $\mu$ m. (TIF)

**S5 Fig. Mapping results of the novel centromeric sequence 4DS-1 to the wheat D genome.** The novel centromeric sequence 4DS-1 shows mapping differences between CS and 4DS in Scaffolds10770 and 28550, as indicated by the magenta boxes. (TIF)

**S6 Fig. Histone modifications on the 4DS *de novo* centromere and a randomly selected centromere (control).** The immunostaining with antibodies against (A) H2AZ; (B) H3K4me2; (C) H3K4me3; (D) H3K27me2; (E) H3K27me3; and (F) H3K9me2 –labeled in green. CENH3 is labeled in red. Bar = 10  $\mu$ m. (TIF)

**S7 Fig. FISH on the 8802 and F1 hybrids between nulli-tetrasomic lines and 8802.** (A) Chromosomes of 8802. The genomic DNA of *Th. elongatum* is labeled in green, the CRW sequences are labeled in red, and DAPI staining is labeled in blue. (B) Karyotype analysis of 8802. pAsI is labeled in green, and pSc119.2 is labeled in red. The arrows indicate the 2E chromosome. (C)-(F) FISH analysis of the F1 hybrids of N3AT3B $\times$ 8802 (C), N5BT5A $\times$ 8802 (D), N5DT5B $\times$ 8802 (E) and N6AT6B $\times$ 8802 (F). The insets show high-magnification images of the chromosomes with two centromeric regions. The arrows indicate the chromosomes with changes in their centromeres. Bar = 10  $\mu$ m. (TIF)

**S8 Fig. Chromosomes with three-locus centromeres in the F5 generation of hybrids between nulli-tetrasomic N6AT6B  $\times$  8802.** (A) and (B). Karyotype analysis of two different progeny of the three-locus centromeres using the pAsI (green) and pSc119.2 (red) probes. (C) and (D). FISH and GISH on two different progeny with chromosomes bearing a three-locus centromere. The genomic DNA of *Th. elongatum* is labeled in green, the CRW sequences are labeled in red, and DAPI staining is labeled in blue. The insets show high-magnification images of the chromosomes with three-locus centromeres. Bar = 10  $\mu$ m. (TIF)

**S9 Fig. Dicentric and multi-centric chromosomes in hybrids of wheat and wild relatives.** (A) FISH and GISH of the dicentric and multi-centric chromosomes in hybrids of wheat and *Th. elongatum*. The DNA of *Th. elongatum* is labeled in red, the CRW sequences are labeled in green, and DAPI staining is labeled in blue. In (B) and (C), the CRW sequences are labeled in red. (B) FISH of the dicentric and multi-centric chromosomes in hybrids of wheat and *Th. poticum*. (C) FISH of the dicentric chromosomes in hybrids of wheat and *Th. intermedium*, *Agropyron cristatum*, *Hordeum vulgare* and *S. cereale*. Bar = 10  $\mu$ m. (TIF)

**S10 Fig. Cytological analysis of *S. cereale* Kustro and its 2R and 2RL addition lines.** (A) FISH of *S. cereale* Kustro with the CRW (red) and pAWRC.1 (green) probes. (B) Karyotype analysis of *S. cereale* Kustro using the AAC (red) and pSc119.2 (green) probes. (C) and (D) FISH analysis of the 2R and 2RL addition lines. The genomic DNA of rye is labeled in green, the CRW sequences are labeled in red, and DAPI staining is labeled in blue. Bar = 10  $\mu$ m. (TIF)

**S11 Fig. Immunodetection of CENH3 and GISH with rye DNA in the novel wheat-rye 2R and 6R addition lines.** (A) and (B). The novel 2R and 6R addition lines, respectively. The genomic DNA of rye is labeled in green, CENH3 is labeled in red, and DAPI staining is labeled in blue. The insets show high-magnification images of the chromosomes with expanded centromeres. Bar = 10  $\mu$ m. (TIF)

**S1 Table. Statistics of the ChIP-Seq mapping results.**  
(DOCX)

**S2 Table. Detailed scaffolds and genes located in 994-kb region of the new centromere in 4DS.**  
(XLS)

## Acknowledgments

We thank Dr. James A. Birchler, Dr. Ryan N. Douglas (University of Missouri) and Nathan D. Han (Washington University in St. Louis) for critically reading and editing the manuscript and for providing valuable suggestions. Dr. Ingo Schubert and Dr. Andreas Houben also read this manuscript and provided helpful suggestions and comments.

## Author Contributions

Conceived and designed the experiments: FH. Performed the experiments: XG HS QS SF JW FH. Analyzed the data: XG HS QS SF JW XZ ZH FH. Contributed reagents/materials/analysis tools: XZ ZH. Wrote the paper: XG HS FH.

## References

1. Manuelidis L, Wu JC. Homology between human and simian repeated DNA. *Nature*. 1978; 276: 92–94. PMID: [105293](#)
2. Martinez-Zapater JM, Estelle MA, Somerville CR. A highly repeated DNA sequence in *Arabidopsis thaliana*. *Mol Gen Genet*. 1986; 204: 417–423.
3. Willard HF, Wayne JS. Hierarchical order in chromosome-specific human alpha satellite DNA. *Trends Genet*. 1987; 3: 192–198.
4. Ananiev EV, Phillips RL, Rines HW. Chromosome-specific molecular organization of maize (*Zea mays* L.) centromeric regions. *Proc Natl Acad Sci U S A*. 1998; 95: 13073–13078. PMID: [9789043](#)
5. Cheng Z, Dong F, Langdon T, Ouyang S, Buell CR, Gu M, et al. Functional rice centromeres are marked by a satellite repeat and a centromere-specific retrotransposon. *Plant Cell*. 2002; 14: 1691–1704. PMID: [12172016](#)
6. Jiang J, Birchler JA, Parrott WA, Dawe RK. A molecular view of plant centromeres. *Trends Plant Sci*. 2003; 8: 570–575. PMID: [14659705](#)
7. Lamb JC, Theuri J, Birchler JA. What's in a centromere? *Genome Biol*. 2004; 5: 239. PMID: [15345039](#)
8. Wu J, Yamagata H, Hayashi-Tsugane M, Hijishita S, Fujisawa M, Shibata M, et al. Composition and structure of the centromeric region of rice chromosome 8. *Plant Cell*. 2004; 16: 967–976. PMID: [15037733](#)
9. Henikoff S, Dalal Y. Centromeric chromatin: what makes it unique? *Curr Opin Genet Dev*. 2005; 15: 177–184. PMID: [15797200](#)

10. Birchler JA, Han F. Maize centromeres: structure, function, epigenetics. *Annu Rev Genet.* 2009; 43: 287–303. doi: [10.1146/annurev-genet-102108-134834](https://doi.org/10.1146/annurev-genet-102108-134834) PMID: [19689211](https://pubmed.ncbi.nlm.nih.gov/19689211/)
11. Neumann P, Navrátilová A, Schroeder-Reiter E, Koblížková A, Steinbauerová V, Chocholová E, et al. Stretching the rules: monocentric chromosomes with multiple centromere domains. *PLoS Genet.* 2012; 8: e1002777. doi: [10.1371/journal.pgen.1002777](https://doi.org/10.1371/journal.pgen.1002777) PMID: [22737088](https://pubmed.ncbi.nlm.nih.gov/22737088/)
12. Earnshaw WC, Migeon BR. Three related centromere proteins are absent from the inactive centromere of a stable isodicentric chromosome. *Chromosoma.* 1985; 92: 290–296. PMID: [2994966](https://pubmed.ncbi.nlm.nih.gov/2994966/)
13. Henikoff S, Ahmad K, Malik HS. The centromere paradox: stable inheritance with rapidly evolving DNA. *Science.* 2001; 293: 1098–1102. PMID: [11498581](https://pubmed.ncbi.nlm.nih.gov/11498581/)
14. Talbert PB, Masuelli R, Tyagi AP, Comai L, Henikoff S. Centromeric localization and adaptive evolution of an *Arabidopsis* histone H3 variant. *Plant Cell.* 2002; 14: 1053–1066. PMID: [12034896](https://pubmed.ncbi.nlm.nih.gov/12034896/)
15. Birchler JA, Gao Z, Sharma A, Presting GG, Han F. Epigenetic aspects of centromere function in plants. *Curr Opin Plant Biol.* 2011; 14: 217–222. doi: [10.1016/j.pbi.2011.02.004](https://doi.org/10.1016/j.pbi.2011.02.004) PMID: [21411364](https://pubmed.ncbi.nlm.nih.gov/21411364/)
16. Dong Q, Han F. Phosphorylation of histone H2A is associated with centromere function and maintenance in meiosis. *Plant J.* 2012; 71: 800–809. doi: [10.1111/j.1365-313X.2012.05029.x](https://doi.org/10.1111/j.1365-313X.2012.05029.x) PMID: [22519817](https://pubmed.ncbi.nlm.nih.gov/22519817/)
17. Houben A, Wako T, Furushima-Shimogawara R, Presting G, Künzel G, Schubert I, et al. The cell cycle dependent phosphorylation of histone H3 is correlated with the condensation of plant mitotic chromosomes. *Plant J.* 1999; 18: 675–679. PMID: [10417719](https://pubmed.ncbi.nlm.nih.gov/10417719/)
18. Han F, Lamb JC, Birchler JA. High frequency of centromere inactivation resulting in stable dicentric chromosomes of maize. *Proc Natl Acad Sci U S A.* 2006; 103: 3238–3243. PMID: [16492777](https://pubmed.ncbi.nlm.nih.gov/16492777/)
19. Hall SE, Kettler G, Preuss D. Centromere satellites from *Arabidopsis* populations: maintenance of conserved and variable domains. *Genome Res.* 2003; 13: 195–205. PMID: [12566397](https://pubmed.ncbi.nlm.nih.gov/12566397/)
20. Gong Z, Wu Y, Koblizkova A, Torres GA, Wang K, Iovene M, et al. Repeatless and repeat-based centromeres in potato: implications for centromere evolution. *Plant Cell.* 2012; 24: 3559–3574. doi: [10.1105/tpc.112.100511](https://doi.org/10.1105/tpc.112.100511) PMID: [22968715](https://pubmed.ncbi.nlm.nih.gov/22968715/)
21. Wen R, Moore G, Shaw PJ. Centromeres cluster de novo at the beginning of meiosis in *Brachypodium distachyon*. *PLoS One.* 2012; 7: e44681. doi: [10.1371/journal.pone.0044681](https://doi.org/10.1371/journal.pone.0044681) PMID: [22970287](https://pubmed.ncbi.nlm.nih.gov/22970287/)
22. Lee H-R, Zhang W, Langdon T, Jin W, Yan H, Cheng Z, et al. Chromatin immunoprecipitation cloning reveals rapid evolutionary patterns of centromeric DNA in *Oryza* species. *Proc Natl Acad Sci U S A.* 2005; 102: 11793–11798. PMID: [16040802](https://pubmed.ncbi.nlm.nih.gov/16040802/)
23. Zhang H, Koblizkova A, Wang K, Gong Z, Oliveira L, Torres GA, et al. Boom-bust turnovers of megabase-sized centromeric DNA in *Solanum* species: rapid evolution of DNA sequences associated with centromeres. *Plant Cell.* 2014; 26: 1436–1447. PMID: [24728646](https://pubmed.ncbi.nlm.nih.gov/24728646/)
24. Ma J, Bennetzen JL. Recombination, rearrangement, reshuffling, and divergence in a centromeric region of rice. *Proc Natl Acad Sci U S A.* 2006; 103: 383–388. PMID: [16381819](https://pubmed.ncbi.nlm.nih.gov/16381819/)
25. Ma J, Jackson SA. Retrotransposon accumulation and satellite amplification mediated by segmental duplication facilitate centromere expansion in rice. *Genome Res.* 2006; 16: 251–259. PMID: [16354755](https://pubmed.ncbi.nlm.nih.gov/16354755/)
26. Wang H, Bennetzen JL. Centromere retention and loss during the descent of maize from a tetraploid ancestor. *Proc Natl Acad Sci U S A.* 2012; 109: 21004–21009. doi: [10.1073/pnas.1218668109](https://doi.org/10.1073/pnas.1218668109) PMID: [23197827](https://pubmed.ncbi.nlm.nih.gov/23197827/)
27. Burrack LS, Berman J. Neocentromeres and epigenetically inherited features of centromeres. *Chromosome Res.* 2012; 20: 607–619. doi: [10.1007/s10577-012-9296-x](https://doi.org/10.1007/s10577-012-9296-x) PMID: [22723125](https://pubmed.ncbi.nlm.nih.gov/22723125/)
28. Barry AE, Howman EV, Cancilla MR, Saffery R, Choo KH. Sequence analysis of an 80 kb human neocentromere. *Hum Mol Genet.* 1999; 8: 217–227. PMID: [9931329](https://pubmed.ncbi.nlm.nih.gov/9931329/)
29. Nasuda S, Hudakova S, Schubert I, Houben A, Endo TR. Stable barley chromosomes without centromeric repeats. *Proc Natl Acad Sci U S A.* 2005; 102: 9842–9847. PMID: [15998740](https://pubmed.ncbi.nlm.nih.gov/15998740/)
30. Marshall OJ, Chueh AC, Wong LH, Choo KH. Neocentromeres: new insights into centromere structure, disease development, and karyotype evolution. *Am J Hum Genet.* 2008; 82: 261–282. doi: [10.1016/j.ajhg.2007.11.009](https://doi.org/10.1016/j.ajhg.2007.11.009) PMID: [18252209](https://pubmed.ncbi.nlm.nih.gov/18252209/)
31. Fu S, Lv Z, Gao Z, Wu H, Pang J, Zhang B, et al. De novo centromere formation on a chromosome fragment in maize. *Proc Natl Acad Sci U S A.* 2013; 110: 6033–6036. doi: [10.1073/pnas.1303944110](https://doi.org/10.1073/pnas.1303944110) PMID: [23530217](https://pubmed.ncbi.nlm.nih.gov/23530217/)
32. Zhang B, Lv Z, Pang J, Liu Y, Guo X, Fu S, et al. Formation of a functional maize centromere after loss of centromeric sequences and gain of ectopic sequences. *Plant Cell.* 2013; 25: 1979–1989. doi: [10.1105/tpc.113.110015](https://doi.org/10.1105/tpc.113.110015) PMID: [23771890](https://pubmed.ncbi.nlm.nih.gov/23771890/)

33. Gao Z, Fu S, Dong Q, Han F, Birchler JA. Inactivation of a centromere during the formation of a translocation in maize. *Chromosome Res.* 2011; 19: 755–761. doi: [10.1007/s10577-011-9240-5](https://doi.org/10.1007/s10577-011-9240-5) PMID: [21947957](https://pubmed.ncbi.nlm.nih.gov/21947957/)
34. Han F, Gao Z, Birchler JA. Reactivation of an inactive centromere reveals epigenetic and structural components for centromere specification in maize. *Plant Cell.* 2009; 21: 1929–1939. doi: [10.1105/tpc.109.066662](https://doi.org/10.1105/tpc.109.066662) PMID: [19602622](https://pubmed.ncbi.nlm.nih.gov/19602622/)
35. Allshire RC, Karpen GH. Epigenetic regulation of centromeric chromatin: old dogs, new tricks? *Nat Rev Genet.* 2008; 9: 923–937. doi: [10.1038/nrg2466](https://doi.org/10.1038/nrg2466) PMID: [19002142](https://pubmed.ncbi.nlm.nih.gov/19002142/)
36. Chueh AC, Northrop EL, Brettingham-Moore KH, Choo KH, Wong LH. LINE retrotransposon RNA is an essential structural and functional epigenetic component of a core neocentromeric chromatin. *PLoS Genet.* 2009; 5: e1000354. doi: [10.1371/journal.pgen.1000354](https://doi.org/10.1371/journal.pgen.1000354) PMID: [19180186](https://pubmed.ncbi.nlm.nih.gov/19180186/)
37. Gent JI, Dawe RK. RNA as a structural and regulatory component of the centromere. *Annu Rev Genet.* 2012; 46: 443–453. doi: [10.1146/annurev-genet-110711-155419](https://doi.org/10.1146/annurev-genet-110711-155419) PMID: [22974300](https://pubmed.ncbi.nlm.nih.gov/22974300/)
38. Hall LE, Mitchell SE, O'Neill RJ. Pericentric and centromeric transcription: a perfect balance required. *Chromosome Res.* 2012; 20: 535–546. doi: [10.1007/s10577-012-9297-9](https://doi.org/10.1007/s10577-012-9297-9) PMID: [22760449](https://pubmed.ncbi.nlm.nih.gov/22760449/)
39. Shaked H, Kashkush K, Ozkan H, Feldman M, Levy AA. Sequence elimination and cytosine methylation are rapid and reproducible responses of the genome to wide hybridization and allopolyploidy in wheat. *Plant Cell.* 2001; 13: 1749–1759. PMID: [11487690](https://pubmed.ncbi.nlm.nih.gov/11487690/)
40. Kashkush K, Feldman M, Levy AA. Transcriptional activation of retrotransposons alters the expression of adjacent genes in wheat. *Nat Genet.* 2003; 33: 102–106. PMID: [12483211](https://pubmed.ncbi.nlm.nih.gov/12483211/)
41. Feldman M, Levy AA. Allopolyploidy—a shaping force in the evolution of wheat genomes. *Cytogenet Genome Res.* 2005; 109: 250–258. PMID: [15753584](https://pubmed.ncbi.nlm.nih.gov/15753584/)
42. Feldman M, Levy AA. Genome evolution due to allopolyploidization in wheat. *Genetics.* 2012; 192: 763–774. doi: [10.1534/genetics.112.146316](https://doi.org/10.1534/genetics.112.146316) PMID: [23135324](https://pubmed.ncbi.nlm.nih.gov/23135324/)
43. Li A, Liu D, Wu J, Zhao X, Hao M, Geng S, et al. mRNA and small RNA transcriptomes reveal insights into dynamic homoeolog regulation of allopolyploid heterosis in nascent hexaploid wheat. *Plant Cell.* 2014; 26: 1878–1900. PMID: [24838975](https://pubmed.ncbi.nlm.nih.gov/24838975/)
44. Khan IA. Molecular and agronomic characterization of wheat-*Agropyron intermedium* recombinant chromosomes. *Plant Breeding.* 2000; 119: 25–29.
45. Mullan DJ, Mirzaghaderi G, Walker E, Colmer TD, Francki MG. Development of wheat-*Lophopyrum elongatum* recombinant lines for enhanced sodium 'exclusion' during salinity stress. *Theor Appl Genet.* 2009; 119: 1313–1323. doi: [10.1007/s00122-009-1136-9](https://doi.org/10.1007/s00122-009-1136-9) PMID: [19727655](https://pubmed.ncbi.nlm.nih.gov/19727655/)
46. Mago R, Spielmeier W, Lawrence J, Lagudah S, Ellis G, Pryor A. Identification and mapping of molecular markers linked to rust resistance genes located on chromosome 1RS of rye using wheat-rye translocation lines. *Theor Appl Genet.* 2002; 104: 1317–1324. PMID: [12582587](https://pubmed.ncbi.nlm.nih.gov/12582587/)
47. Charpentier A, Feldman M, Cauderon Y. The effect of different doses of Ph1 on chromosome pairing in hybrids between tetraploid *Agropyron elongatum* and common wheat. *Genome.* 1988; 30: 974–977.
48. Wang K, Wu Y, Zhang W, Dawe RK, Jiang J. Maize centromeres expand and adopt a uniform size in the genetic background of oat. *Genome Res.* 2014; 24: 107–116. doi: [10.1101/gr.160887.113](https://doi.org/10.1101/gr.160887.113) PMID: [24100079](https://pubmed.ncbi.nlm.nih.gov/24100079/)
49. Zhang P, Li W, Fellers J, Friebe B, Gill BS. BAC-FISH in wheat identifies chromosome landmarks consisting of different types of transposable elements. *Chromosoma.* 2004; 112: 288–299. PMID: [14986017](https://pubmed.ncbi.nlm.nih.gov/14986017/)
50. Luce AC, Sharma A, Mollere OSB, Wolfgruber TK, Nagaki K, Jiang J, et al. Precise centromere mapping using a combination of repeat junction markers and chromatin immunoprecipitation–polymerase chain reaction. *Genetics.* 2006; 174: 1057–1061. PMID: [16951073](https://pubmed.ncbi.nlm.nih.gov/16951073/)
51. Wolfgruber TK, Sharma A, Schneider KL, Albert PS, Koo DH, Shi J, et al. Maize centromere structure and evolution: sequence analysis of centromeres 2 and 5 reveals dynamic loci shaped primarily by retrotransposons. *PLoS Genet.* 2009; 5: e1000743. doi: [10.1371/journal.pgen.1000743](https://doi.org/10.1371/journal.pgen.1000743) PMID: [19956743](https://pubmed.ncbi.nlm.nih.gov/19956743/)
52. Kishii M, Nagaki K, Tsujimoto H. A tandem repetitive sequence located in the centromeric region of common wheat (*Triticum aestivum*) chromosomes. *Chromosome Res.* 2001; 9: 417–428. PMID: [11448043](https://pubmed.ncbi.nlm.nih.gov/11448043/)
53. Cheng ZJ, Murata M. A centromeric tandem repeat family originating from a part of Ty3/gypsy-retroelement in wheat and its relatives. *Genetics.* 2003; 164: 665–672. PMID: [12807787](https://pubmed.ncbi.nlm.nih.gov/12807787/)
54. Li B, Choulet F, Heng Y, Hao W, Paux E, Liu Z, et al. Wheat centromeric retrotransposons: the new ones take a major role in centromeric structure. *Plant J.* 2013; 73: 952–965. doi: [10.1111/tpj.12086](https://doi.org/10.1111/tpj.12086) PMID: [23253213](https://pubmed.ncbi.nlm.nih.gov/23253213/)
55. Sears ER. The Aneuploids of common wheat. *Mo Agric Exp Stn Res Bull.*; 1954; 572.

56. He M, Xu Z, Zou M, Zhang H, Chen D, Piao Z, et al. The establishment of 2 sets of alien addition lines of wheat-wheatgrass. *Sci China B*. 1989; 32: 695–705.
57. Han F, Li J. Morphology and cytogenetics of intergeneric hybrids of crossing *Triticum durum* and *T. timopheevi* with tetraploid *Elytrigia elongata*. *J Genet Genomics*. 1993; 20: 44–49.
58. Brenchley R, Spannagl M, Pfeifer M, Barker GL, D'Amore R, Allen AM, et al. Analysis of the bread wheat genome using whole-genome shotgun sequencing. *Nature*. 2012; 491: 705–710. doi: [10.1038/nature11650](https://doi.org/10.1038/nature11650) PMID: [23192148](https://pubmed.ncbi.nlm.nih.gov/23192148/)
59. Jia J, Zhao S, Kong X, Li Y, Zhao G, He W, et al. *Aegilops tauschii* draft genome sequence reveals a gene repertoire for wheat adaptation. *Nature*. 2013; 496: 91–95. doi: [10.1038/nature12028](https://doi.org/10.1038/nature12028) PMID: [23535592](https://pubmed.ncbi.nlm.nih.gov/23535592/)
60. Li H, Durbin R. Fast and accurate short read alignment with burrows-wheeler transform. *Bioinformatics*. 2009; 25: 1754–1760. doi: [10.1093/bioinformatics/btp324](https://doi.org/10.1093/bioinformatics/btp324) PMID: [19451168](https://pubmed.ncbi.nlm.nih.gov/19451168/)
61. Benson DA, Karsch-Mizrachi I, Lipman DJ, Ostell J, Rapp BA, Wheeler DL. GenBank. *Nucleic Acids Res*. 2000; 28:15–18. PMID: [10592170](https://pubmed.ncbi.nlm.nih.gov/10592170/)
62. Fu S, Lv Z, Guo X, Zhang X, Han F. Alteration of terminal heterochromatin and chromosome rearrangements in derivatives of wheat-rye hybrids. *J Genet Genomics*. 2013; 40: 413–420. doi: [10.1016/j.jgg.2013.05.005](https://doi.org/10.1016/j.jgg.2013.05.005) PMID: [23969250](https://pubmed.ncbi.nlm.nih.gov/23969250/)
63. Houben A, Kumke K, Nagaki K, Hause G. CENH3 distribution and differential chromatin modifications during pollen development in rye (*Secale cereale* L.). *Chromosome Res*. 2011; 19: 471–480. doi: [10.1007/s10577-011-9207-6](https://doi.org/10.1007/s10577-011-9207-6) PMID: [21503764](https://pubmed.ncbi.nlm.nih.gov/21503764/)
64. Voullaire LE, Slater HR, Petrovic V, Choo KH. A functional marker centromere with no detectable alpha-satellite, satellite III, or CENP-B protein: activation of a latent centromere? *Am J Hum Genet*. 1993; 52: 1153–1163. PMID: [7684888](https://pubmed.ncbi.nlm.nih.gov/7684888/)
65. Topp CN, Okagaki RJ, Melo JR, Kynast RG, Phillips RL, Dawe RK. Identification of a maize neocentromere in an oat-maize addition line. *Cytogenet Genome Res*. 2009; 124: 228–238. doi: [10.1159/000218128](https://doi.org/10.1159/000218128) PMID: [19556776](https://pubmed.ncbi.nlm.nih.gov/19556776/)
66. Shang WH, Hori T, Martins NM, Toyoda A, Misu S, Monma N, et al. Chromosome engineering allows the efficient isolation of vertebrate neocentromeres. *Dev Cell*. 2013; 24: 635–648. doi: [10.1016/j.devcel.2013.02.009](https://doi.org/10.1016/j.devcel.2013.02.009) PMID: [23499358](https://pubmed.ncbi.nlm.nih.gov/23499358/)
67. Thakur J, Sanyal K. Efficient neocentromere formation is suppressed by gene conversion to maintain centromere function at native physical chromosomal loci in *Candida albicans*. *Genome Res*. 2013; 23: 638–652. doi: [10.1101/gr.141614.112](https://doi.org/10.1101/gr.141614.112) PMID: [23439889](https://pubmed.ncbi.nlm.nih.gov/23439889/)
68. Blower MD, Sullivan BA, Karpen GH. Conserved organization of centromeric chromatin in flies and humans. *Dev Cell*. 2002; 2: 319–330. PMID: [11879637](https://pubmed.ncbi.nlm.nih.gov/11879637/)
69. Sullivan BA, Karpen GH. Centromeric chromatin exhibits a histone modification pattern that is distinct from both euchromatin and heterochromatin. *Nat Struct Mol Biol*. 2004; 11: 1076–1083. PMID: [15475964](https://pubmed.ncbi.nlm.nih.gov/15475964/)
70. Shi J, Dawe RK. Partitioning of the maize epigenome by the number of methyl groups on histone H3 lysines 9 and 27. *Genetics*. 2006; 173: 1571–1583. PMID: [16624902](https://pubmed.ncbi.nlm.nih.gov/16624902/)
71. Jin W, Lamb JC, Zhang W, Kolano B, Birchler JA, Jiang J. Histone modifications associated with both A and B chromosomes of maize. *Chromosome Res*. 2008; 16: 1203–1214. doi: [10.1007/s10577-008-1269-8](https://doi.org/10.1007/s10577-008-1269-8) PMID: [18987983](https://pubmed.ncbi.nlm.nih.gov/18987983/)
72. Yan H, Jin W, Nagaki K, Tian S, Ouyang S, Buell CR, et al. Transcription and histone modifications in the recombination-free region spanning a rice centromere. *Plant Cell*. 2005; 17: 3227–3238. PMID: [16272428](https://pubmed.ncbi.nlm.nih.gov/16272428/)
73. Wong LH, Brettingham-Moore KH, Chan L, Quach JM, Anderson MA, Northrop EL, et al. Centromere RNA is a key component for the assembly of nucleoproteins at the nucleolus and centromere. *Genome Res*. 2007; 17: 1146–1160. PMID: [17623812](https://pubmed.ncbi.nlm.nih.gov/17623812/)
74. Topp CN, Zhong CX, Dawe RK. Centromere-encoded RNAs are integral components of the maize kinetochore. *Proc Natl Acad Sci U S A*. 2004; 101: 15986–15991. PMID: [15514020](https://pubmed.ncbi.nlm.nih.gov/15514020/)
75. Aragon-Alcaide L, Miller T, Schwarzacher T, Reader S, Moore G. A cereal centromeric sequence. *Chromosoma*. 1996; 105: 261–268. PMID: [8939818](https://pubmed.ncbi.nlm.nih.gov/8939818/)
76. Jiang J, Nasuda S, Dong F, Scherrer CW, Woo S-S, Wing RA, et al. A conserved repetitive DNA element located in the centromeres of cereal chromosomes. *Proc Natl Acad Sci U S A*. 1996; 93: 14210–14213. PMID: [8943086](https://pubmed.ncbi.nlm.nih.gov/8943086/)
77. Hudakova S, Michalek W, Presting GG, Hoopen RT, Santos KD, Jasencakova Z, et al. Sequence organization of barley centromeres. *Nucleic Acids Res*. 2001; 29: 5029–5035. PMID: [11812833](https://pubmed.ncbi.nlm.nih.gov/11812833/)

78. Thomae AW, Schade GO, Padeken J, Borath M, Vetter I, Kremmer E, et al. A pair of centromeric proteins mediates reproductive isolation in *Drosophila* species. *Dev Cell*. 2013; 27: 412–424. doi: [10.1016/j.devcel.2013.10.001](https://doi.org/10.1016/j.devcel.2013.10.001) PMID: [24239514](https://pubmed.ncbi.nlm.nih.gov/24239514/)
79. Fishman L, Saunders A. Centromere-associated female meiotic drive entails male fitness costs in monkeyflowers. *Science*. 2008; 322: 1559–1562. doi: [10.1126/science.1161406](https://doi.org/10.1126/science.1161406) PMID: [19056989](https://pubmed.ncbi.nlm.nih.gov/19056989/)
80. Malik HS, Henikoff S. Major evolutionary transitions in centromere complexity. *Cell*. 2009; 138: 1067–1082. doi: [10.1016/j.cell.2009.08.036](https://doi.org/10.1016/j.cell.2009.08.036) PMID: [19766562](https://pubmed.ncbi.nlm.nih.gov/19766562/)
81. Malik HS, Henikoff S. Adaptive evolution of Cid, a centromere-specific histone in *Drosophila*. *Genetics*. 2001; 157: 1293–1298. PMID: [11238413](https://pubmed.ncbi.nlm.nih.gov/11238413/)
82. Kynast RG, Friebe B, Gill BS. Fate of multicentric and ring chromosomes induced by a new gametocidal factor located on chromosome 4Mg of *Aegilops geniculata*. *Chromosome Res*. 2000; 8: 133–139. PMID: [10780702](https://pubmed.ncbi.nlm.nih.gov/10780702/)
83. Hiatt EN, Kentner EK, Dawe RK. Independently regulated neocentromere activity of two classes of tandem repeat arrays. *Plant Cell*. 2002; 14: 407–420. PMID: [11884683](https://pubmed.ncbi.nlm.nih.gov/11884683/)
84. Guerra M, Cabral G, Cuacos M, Gonzalez-Garcia M, Gonzalez-Sanchez M, Vega J, et al. Neocentrics and holokinetics (holocentrics): chromosomes out of the centromeric rules. *Cytogenet Genome Res*. 2010; 129: 82–96. doi: [10.1159/000314289](https://doi.org/10.1159/000314289) PMID: [20551611](https://pubmed.ncbi.nlm.nih.gov/20551611/)
85. Metcalfe CJ, Bulazel KV, Ferreri GC, Schroeder-Reiter E, Wanner G, Rens W, et al. Genomic instability within centromeres of interspecific marsupial hybrids. *Genetics*. 2007; 177: 2507–2517. PMID: [18073443](https://pubmed.ncbi.nlm.nih.gov/18073443/)
86. Lukaszewski AJ. Behavior of centromeres in univalents and centric misdivision in wheat. *Cytogenet Genome Res*. 2010; 129: 97–109. doi: [10.1159/000314108](https://doi.org/10.1159/000314108) PMID: [20551608](https://pubmed.ncbi.nlm.nih.gov/20551608/)
87. Kaszas E, Kato A, Birchler JA. Cytological and molecular analysis of centromere misdivision in maize. *Genome*. 2002; 45: 759–768. PMID: [12175080](https://pubmed.ncbi.nlm.nih.gov/12175080/)
88. Jiang B, Lou Q-F, Wang D, Wu Z-M, Zhang W-P, Chen J-F. Allopolyploidization induced the activation of Ty1-copia retrotransposons in *Cucumis hytivus*, a newly formed *Cucumis* allotetraploid. *Bot Stud*. 2011; 52: 145–152.
89. Schubert I, Lysak MA. Interpretation of karyotype evolution should consider chromosome structural constraints. *Trends Genet*. 2011; 27: 207–216. doi: [10.1016/j.tig.2011.03.004](https://doi.org/10.1016/j.tig.2011.03.004) PMID: [21592609](https://pubmed.ncbi.nlm.nih.gov/21592609/)
90. Zhang H, Dawe RK. Total centromere size and genome size are strongly correlated in ten grass species. *Chromosome Res*. 2012; 20: 403–412. doi: [10.1007/s10577-012-9284-1](https://doi.org/10.1007/s10577-012-9284-1) PMID: [22552915](https://pubmed.ncbi.nlm.nih.gov/22552915/)
91. Kato A, Lamb JC, Birchler JA. Chromosome painting using repetitive DNA sequences as probes for somatic chromosome identification in maize. *Proc Natl Acad Sci U S A*. 2004; 101: 13554–13559. PMID: [15342909](https://pubmed.ncbi.nlm.nih.gov/15342909/)
92. Schneider CA, Rasband WS, Eliceiri KW. NIH Image to ImageJ: 25 years of image analysis. *Nat Methods*. 2012; 9: 671–675. PMID: [22930834](https://pubmed.ncbi.nlm.nih.gov/22930834/)
93. Nagaki K, Talbert PB, Zhong CX, Dawe RK, Henikoff S, Jiang J. Chromatin immunoprecipitation reveals that the 180-bp satellite repeat is the key functional DNA element of *Arabidopsis thaliana* centromeres. *Genetics*. 2003; 163: 1221–1225. PMID: [12663558](https://pubmed.ncbi.nlm.nih.gov/12663558/)
94. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2<sup>-</sup>( $\Delta\Delta C_T$ ) Method. *Methods*. 2001; 25: 402–408. PMID: [11846609](https://pubmed.ncbi.nlm.nih.gov/11846609/)
95. Robinson JT, Thorvaldsdóttir H, Winckler W, Guttman M, Lander ES, Getz G, et al. Integrative Genomics Viewer. *Nat Biotechnol*. 2011; 29: 24–26. doi: [10.1038/nbt.1754](https://doi.org/10.1038/nbt.1754) PMID: [21221095](https://pubmed.ncbi.nlm.nih.gov/21221095/)
96. Deng C, Bai L, Fu S, Yin W, Zhang Y, Chen Y, et al. Microdissection and chromosome painting of the alien chromosome in an addition line of wheat—*Thinopyrum intermedium*. *PLoS One*. 2013; 8: e72564. doi: [10.1371/journal.pone.0072564](https://doi.org/10.1371/journal.pone.0072564) PMID: [23967319](https://pubmed.ncbi.nlm.nih.gov/23967319/)
97. Malinen E, Kassinen A, Rinttila T, Palva A. Comparison of real-time PCR with SYBR Green I or 5'-nuclease assays and dot-blot hybridization with rDNA-targeted oligonucleotide probes in quantification of selected faecal bacteria. *Microbiology*. 2003; 149: 269–277. PMID: [12576600](https://pubmed.ncbi.nlm.nih.gov/12576600/)