Unconventional Meiosis in Plants: Holocentric chromosome recombination in beak-sedges and asymmetric inheritance in pentaploid dogroses

Doctoral thesis

for

the award of the doctoral degree

of the Faculty of Mathematics and Natural Sciences

of the University of Cologne

submitted by

Meng ZHANG

accepted in 2026

Table of Contents

N.b.: Pages in the attached paper are not counted in the pagination of the thesis.

Abstract

Meiosis is an essential cell division process for sexual reproduced eukaryotes, during which independent chromosome segregation and homologous recombination are usually indispensible, yet some species exhibit unconventional mechanisms to overcome challenges posed by holocentric chromosomes, achiasmy, or polyploidy. Recombination landscapes are typically shaped by centromere positioning and genomic features. However, the recombination patterns of holocentric species, that typically have multiple centromere units dipersed along the entire chromosomes, are understudied. Similarly, the mechanisms enabling stable reproduction in achiasmatic or polyploid species are poorly understood. Here, we investigate these exceptions through genomic, cytological, and evolutionary analyses in holocentric *Rhynchospora* and pentaploid dogroses.

Using single-cell sequencing of pollen grains, we first constructed a recombination landscape for *Rhynchospora breviuscula*, which presents a distal bias pattern despite uniformly dispersed (epi)genetic signals. Telomere-led chromosome pairing and synapsis were identified as key drivers of this pattern. Comparative analyses across multiple *Rhynchospora* species revealed that chromosome fusions and fissions drive the karyotype evolution in this genus, and further demonstrated that recombination landscapes are shaped by meiotic dynamics rather than genomic architecture. Strikingly, we discovered that *Rhynchospora tenuis* achieves efficient sexual reproduction without crossovers in both maternal and paternal meiosis, yielding clonal offspring. Finally, in pentaploid dogroses, we proposed a mechanism for its stabilised polyploid inheritance: 14 chromosomes form bivalents (transmitted biparentally), while 21 univalents are maternally retained. Genome assembly and centromere analysis revealed that bivalent-forming subgenomes tend to have ATHILA retrotransposons-based centromeres, whereas univalent subgenomes carry expanded tandem-repeat-based centromeres (CANR4) with higher CENH3 enrichment, suggesting centromere divergence drives their asymmetric segregation.

This work presents how meiosis progresses uncanonically in two different plant genus. These systems highlight nature's versatility in overcoming meiotic constraints, offering insights for genome evolution, sexual reproduction, crop engineering, and the preservation of genetic diversity in challenging genomic contexts.

1. General introduction

Meiosis is an essential cell division process for sexually reproducing eukaryotes, during which germ cells replicate their DNA once and divide twice through meiosis I and meiosis II, resulting in haploid gametes that contain one set of chromosomes for a diploid organism (**Figure 1a**). During meiosis I, homologous chromosomes pair, synapse, and exchange DNA fragments through genetic recombination, which shuffles alleles between homologs and generates novel genetic combinations in gametes (Zickler & Kleckner, 2015, 2023). This recombination process is fundamental for sexual reproduction, as it creates genetic diversity among offspring and thereby enhances a population's adaptation to changing environments (Barton, 2009; Parée & Teotónio, 2025).

Meiosis is a more complex cell division process than mitosis. The accurate pairing, synapsis, crossover (CO), segregation of homologs, etc., require precise orchestration of chromosome structure, centromere dynamics, and recombination machinery. These processes become even more complex in polyploid organisms, where multiple homologous or homeologous chromosomes must correctly identify pairing partners to prevent meiotic irregularities (Grandont et al., 2013; Lloyd & Bomblies, 2016). In holocentric species, which harbour distributed centromeric units along the entire chromosome, the distinct chromosome organisation leads to its unique meiosis process (Hofstatter et al., 2021; Melters et al., 2012). These different meiotic activities highlight the exquisite sensitivity of meiosis to genomic constitution.

In contrast to meiosis, some species reproduce via apomixis, a form of asexual reproduction that skips meiosis and fertilisation, producing clonal offspring genetically identical to the parent (Hojsgaard & Hörandl, 2019; Schmidt, 2025). The primary advantage of apomixis is the efficient propagation of well-adapted genotypes. However, this comes at the cost of genetic diversity, making populations vulnerable to pathogens and environmental changes (Hojsgaard & Hörandl, 2019). Nowadays, understanding the regulatory pathways of meiosis is also applied to emerging technologies such as apomixis engineering, which aims to fix heterosis by enabling clonal seed production to freeze the favourable genetic combination (Underwood & Mercier, 2022). Other popular practices that also aim to improve crop yield or phenotype, such as manipulating recombination rates to break undesirable genetic linkages (W. Wang et al., 2025) and stabilising meiosis in polyploids, take advantage of the power of genetic diversity while also addressing the challenge of inherent

difficulties of complex genomes (Y. Wang et al., 2024). Thus, insights into meiosis mechanisms are essential for both traditional breeding and modern genomic-based crop improvement strategies.

1.1 Genetic basis of meiosis regulation

Meiosis is a two-step specialised cell division that converts a diploid cell into four haploid gametes. Starting with a similar interphase as mitosis, meiosis also has one round of DNA replication, but then is followed by two successive divisions. In most species, meiosis I separates homologous chromosomes, and meiosis II separates sister chromatids (**Figure 1a**). In prophase I, homologous chromosomes pair, form bivalents via synapsis, and exchange genetic material through recombination. These interactions create a physical link called chiasmata, which ensures the correct homolog segregation. In metaphase I, the paired homologs align at the equator, with kinetochore microtubules attaching to the centromeres of each chromosome. In anaphase I, homologous chromosomes are pulled to opposite poles, reducing the chromosome number by half. Telophase I and cytokinesis then produce two nuclei containing one set of chromosomes in each. In the second division, meiosis II, chromosomes condense again, sister chromatids align at the equator, and then segregate to different cells. This results in four genetically distinct haploid nuclei. The accuracy of these steps is indispensable for fertility and genome stability, and is governed by a network of specialised meiotic proteins and regulatory pathways (**Figure 1b, c**).

Prophase I, in particular, is fundamentally important for genetic recombination and accurate segregation with a complex regulatory network. Prophase I starts with a substage called leptotene, during which sister chromatids are held together by cohesin complexes that contain the meiosis-specific kleisin REC8 (Eijpe et al., 2003; Sakuno et al., 2022; Watanabe & Nurse, 1999). Cohesins anchor chromatin loops to the linear meiotic chromosome axes, which are assembled by axial elements, HORMA-domain protein ASY1 and the coiled-coil proteins ASY3 and ASY4 (Armstrong et al., 2002; Chambon et al., 2018; Ferdous et al., 2012). These axes provide a scaffold for the initiation of programmed DNA double-strand breaks (DSBs) that are generated by the topoisomerase VI–like enzyme complex containing SPO11 with MTOPVIB (Keeney et al., 1997; Vrielynck et al., 2016); **Figure 1b**).

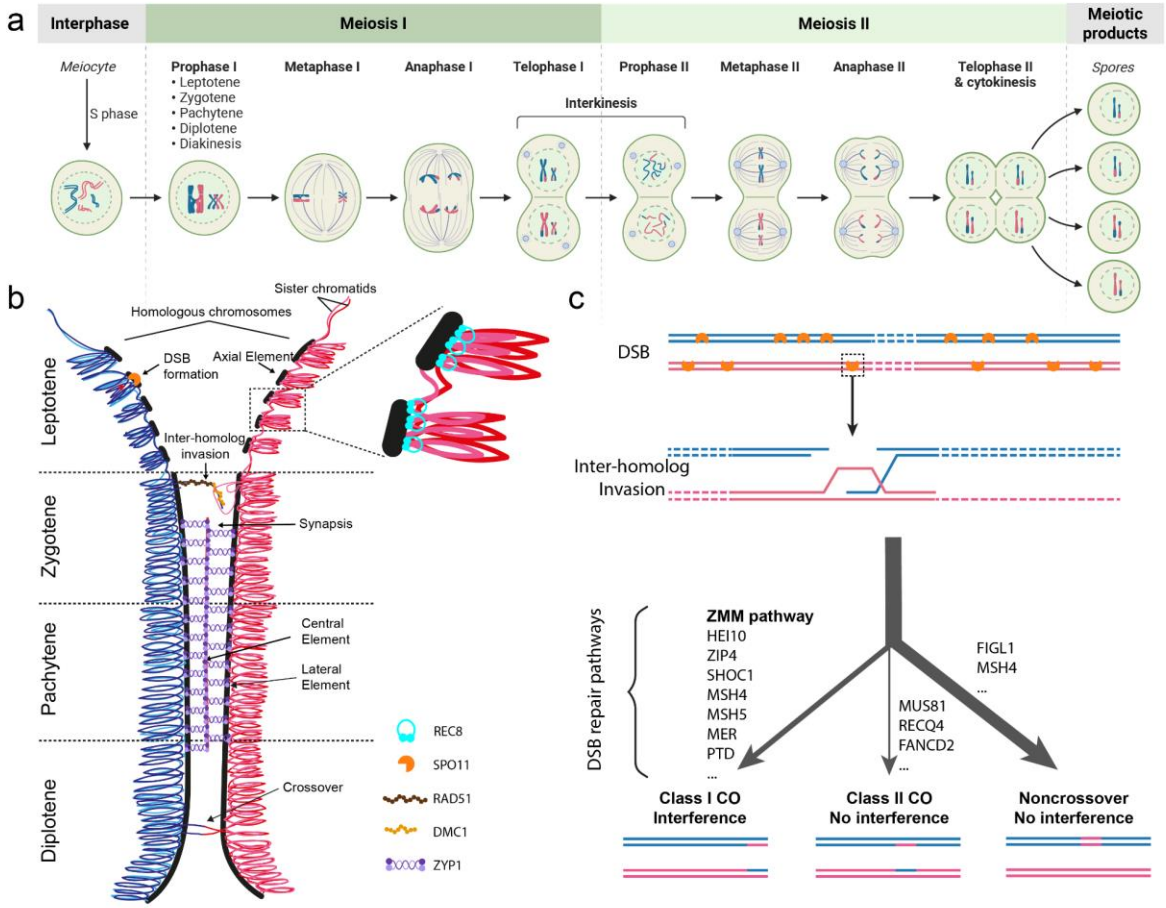


Figure 1. Meiosis process and mechanisms of recombination. (**a**) Complete meiosis process starting from DNA replication in interphase to four haploid spores. Two cell divisions, called meiosis I and meiosis II, follow once DNA replication, and every division can be further divided into four stages: prophase, metaphase, anaphase, and telophase. Blue and pink chromosomes in proximity are homologs. This diagram takes two pairs of homologous chromosomes as an example, shown in different lengths. The diagram was modified from a BioRender template “Plant Cell Meiosis” created by Nashed, M. (**b**) A zoom-in view of chromosome structure change during prophase I, as well as part of the proteins involved. Four substages, leptotene, zygotene, pachytene, and diplotene, are shown. Blue and pink threads still show homologous chromosomes. Each contains two sister chromatids, shaded in different hues of blue and pink. Sister chromatids are tethered into loops by cohesins, which link with axial elements. This diagram was modified from a template from (Wells, 2020)), with the license allowing for sharing and remixing (https://commons.wikimedia.org/wiki/File:Synaptonemal_Complex.svg). (**c**) Three major types of recombinations produced by resolving the inter-homolog invasions after DSB formation. The number of DSBs is usually much bigger than the detected crossover number because most DSBs are repaired as gene conversion via noncrossover. The class I crossover is produced by ZMM pathways, and its number is strongly controlled by interference, while the class II crossover is not affected by interference.

Concurrently, during the transition to the next substage, zygotene, telomeres and also centromeres in some species (Jin et al., 1998), attach to the inner nuclear envelope,

forming a “bouquet” configuration that helps to spatially organise homologous chromosomes and facilitate their pairing by rapid chromosome movement and accelerated homology search (Scherthan, 2007). This telomere-nuclear envelope attachment is mediated by protein complexes such as the LINC (linker of nucleoskeleton and cytoskeleton) complex and associated SUN- and KASH-domain proteins (Burke, 2018; Cain et al., 2018; Malone et al., 1999). In zygotene, the DNA ends are resected by nucleolytic factors after DSB formation, for example, EXO1 and the plant MRN complex (Mre11, Rad50 and Nbs1) (Lamarche et al., 2010; Lukaszewicz et al., 2015), producing 3′ single-stranded DNA that is bound by the strand exchange recombinases RAD51 and meiosis-specific DMC1. These recombinases guide the invaded single-strand into the homologous duplex of the partner chromosome in proximity rather than the sister chromatid, producing a displacement loop (D-loop)**,** thereby promoting inter-homolog interactions (Da Ines et al., 2013; Masson & West, 2001).

In Pachytene, homologs are fully synapsed through transverse filament protein ZYP1, which bridges the axial elements and assembles the synaptonemal complex (SC). In particular, the N-termini of two opposing ZYP1 molecules meet at the central region of SC, forming the central element, while the C-terminus links to chromosome axes (Higgins et al., 2005). At this stage, crossover sites are designated by the activity of ZMM proteins (e.g., HEI10, ZIP4, SHOC1, MSH4, MSH5, MER, PTD, etc.), with HEI10 foci marking recombination nodules that will mature into class I COs under interference control, which means one crossover formation prevents another crossover nearby (Börner et al., 2004; Chelysheva et al., 2012; Foss et al., 1993; Muller, 1916). Recent work supports an emergent “HEI10 coarsening” model in which the E3 ligase-related protein HEI10 initially forms many small foci along SC and then undergoes diffusion-mediated coarsening to concentrate at a subset of sites that become COs in the end. This dynamic explains interference and its sensitivity to SC structure and HEI10 dosage. Quantitative modelling and imaging studies across species have provided strong empirical support for HEI10-driven coarsening as a major mechanism of interference (Durand et al., 2022; C. Morgan et al., 2021; Zhang et al., 2021). In diplotene, the synaptonemal complex disassembles, leaving chiasmata that maintain homolog association. By diakinesis, chromosomes undergo further condensation and reorganisation, preparing for alignment and segregation at metaphase I.

However, early recombination intermediates can be processed along alternative routes besides the class I crossover via the ZMM pathway. Class II COs processed by structure-selective nucleases (e.g., MUS81) produce non-interfering COs (Hollingsworth & Brill, 2004; Santos et al., 2003). In addition, noncrossover (NCO) outcomes via synthesis-dependent strand annealing (SDSA), which restore the sequence without reciprocal exchange ((Allers & Lichten, 2001)**Figure 1c**).

Most core meiotic steps and proteins are widely conserved, while the relative contributions of these three ways (class I CO, class II CO, NCO) vary by species and sex. In *Arabidopsis thaliana*, class I COs constitute the majority (~80–90%), with class II providing roughly 10–20% and NCOs representing the remainder (Choi, 2017). Rice shows a similar dominance of class I COs, but the exact fraction varies. Disrupting anti-crossover factors such as RECQ4 and FANCM can dramatically increase total CO numbers and shift proportions (Li et al., 2021; Serra et al., 2018). But in most, if not all, species, each pair of homologous chromosomes or a bivalent must form at least one crossover during meiosis I to ensure proper chromosome segregation, which is called obligate CO. In many animals and plants, males and females display distinct recombination frequency and landscapes. This heterochiasmy reflects differences in chromatin organisation, SC length, and recombination dynamics. For instance, *Arabidopsis* wild type has longer SC in males than females during meiosis, correlated with higher CO frequency in males (Drouaud et al., 2007; Giraut et al., 2011). However, when SC is abolished by the zyp1 mutant, heterochiasmy is also abolished, highlighting the role of meiotic machinery in regulating crossover frequency and variation (Capilla-Pérez et al., 2021).

1.2 Variation of recombination frequency and distribution

Meiotic recombination is crucial because it not only ensures correct chromosome segregation but also generates new allelic combinations. Given these dual roles, studying the frequency and distribution of meiotic recombination is important both for understanding the mechanistic basis of meiosis and for assessing how variation in recombination shapes genome evolution and adaptation.

Recombination during meiosis occurs in two broad forms: 1) crossovers (COs), which exchange flanking chromosomal segments between homologs; and 2) non-crossovers (NCOs), which involve gene conversions without reciprocal exchange. The focus in

most evolutionary and population genetic studies is on COs, as they, on the one hand, break linkages between loci, thereby producing new genotypes and interrupting linkage disequilibrium (LD) (T. H. Morgan, 1911), and on the other hand, are easier to detect than gene conversions without high-resolution genomic markers. So most meiotic recombinations mentioned in the following are actually only COs.

1.2.1 Methods to detect COs or infer recombination rates

To investigate the recombination variation, the first step is to quantify meiotic recombination. The approach selection largely relies on the available material and the targeted resolution. There are two main approaches widely used, i.e. cytological method and genomic method (Peñalba & Wolf, 2020). Cytological methods rely on immunostaining of meiotic chromosome spreads using antibodies against proteins that mark crossover sites, such as mismatch repair protein MLH1, which can provide a direct visualisation of recombination foci along synapsed homologs (Anderson et al., 1999; Baker et al., 1996). This allows for the CO counting per cell and an intuitive visualisation of their spatial distribution along chromosomes. Despite the cell-level precision and its applicability to species without genomic resources, cytological assays are labour-intensive, have limited throughput, and lack molecular resolution to distinguish individual recombination breakpoints or detect NCOs, since only CO-associated markers are visible, and normally only class I COs. Before high-throughput sequencing, the genomic method usually used highly variable loci as markers to estimate the recombination rate indicated by loci segregation. This approach allows inferring recombinations from a large number of individuals, however, at a rather low resolution and with potential missing events (Buetow, 1991). Today, with high-throughput genomic sequencing, many genomic-based methods have been developed.

Gamete sequencing offers a powerful genomic alternative by sequencing individual gametes to directly infer recombination events from haplotype switches within single meioses. This approach provides base-pair resolution and is able to detect both COs and gene conversions if marker density is allowed. Its major advantage is the accuracy and completeness of recombination detection. Compared to other genomic methods, it requires only one individual (Campoy et al., 2020; Carioscia et al., 2022). However, this usually requires a large number of gametes to derive a recombination landscape,

which is usually not feasible for female gametes. Therefore, this approach is not suitable for a heterochiasmy study.

Lineage or family-based linkage mapping infers recombination from the segregation of markers in offspring derived from known parents, allowing estimation of CO numbers and construction of recombination maps (Johnston et al., 2017; Lander & Green, 1987). It captures both CO frequency and large-scale distribution and can be implemented with moderate genomic resources. However, its resolution also depends on marker density. With the integration of cross and backcross strategies in plants, both female and male meiotic COs can be indicated, so it is a popular practice to study differences between female and male recombinations (Kianian et al., 2018; Lloyd & Jenczewski, 2019).

Population-based approaches infer historical recombination rates indirectly from patterns of linkage disequilibrium among polymorphic sites in population genomic datasets. These methods are highly scalable, enabling inference across entire genomes and species (Myers et al., 2005; Stumpf & McVean, 2003; Zheng et al., 2019). However, they estimate long-term average recombination rates rather than per-meiosis events, which means the derived recombinations are likely confounded by demographic structure, selection, genetic drift, and various evolutionary activities (Chapman & Thompson, 2001; Charlesworth et al., 1997).

1.2.2 Variation in recombination rate and distribution

Recombination varies across sexes, individuals and species. At the sex level, many species show sex-specific recombination rates or distributions, which is usually called heterochiasmy. In humans, females exhibit higher genome-wide CO rates than males, though males have relatively elevated recombination near telomeres (Kong et al., 2002). For example, humans have an average of ~70 COs in females and ~50 in males (Barlow & Hultén, 1998; Cheng et al., 2009; Gruhn et al., 2013), while *Arabidopsis thaliana* has 6.65 COs in females and 11.15 COs in males on average (Giraut et al., 2011). More broadly, heterochiasmy is ubiquitous across taxa, though the direction and magnitude vary (Sardell & Kirkpatrick, 2020).

Within individuals, the recombination rate may also change due to both intrinsic developmental factors, such as age, and external stress, such as temperature and pathogen in plants. One well-studied field is in human females. The total number of

COs increases with maternal age, as well as a higher frequency of atypically positioned or closely spaced COs, leading to an elevated recombination rate (Martin et al., 2015). A similar trend was also found in male mice, where spermatocytes from juvenile male mice exhibit fewer COs compared to adult spermatocytes, associated with the risks of aneuploidy in young males (Zelazowski et al., 2017). Temperature stress is one of the best-studied factors in plants that can change recombination studies. For instance, in *Arabidopsis thaliana* and barley, exposure to heat stress increases recombination rates, especially in distal chromosomal regions, while extreme cold can suppress recombination and disrupt synapsis (De Storme & Geelen, 2020; Higgins et al., 2012; Modliszewski et al., 2018; Phillips et al., 2015). Similarly, pathogen infection and biotic stress can influence meiotic recombination in plants. Exposure to pathogens is associated with an increase in recombination frequency, potentially as a mechanism to enhance genetic variation and improve adaptability in the face of biotic challenges (Kerstes et al., 2012). Together, these modulations of recombination illustrate that recombination is not a static feature of the genome but a flexible and responsive process.

Between individuals within a species, substantial heritable variation in recombination rate exists. In a barley nested-association mapping population, sub-populations differed in average CO number and in landscape pattern (Dreissig et al., 2020). At the population level, different populations may have distinct recombination landscapes, influenced by allelic variation in recombination-modifier genes or by environment (Baudat et al., 2010; Kong et al., 2014; Stapley et al., 2017). At the taxonomic level, species differ markedly in baseline recombination rates (COs per meiosis) and in distributional patterns. For example, organisms with and without strict hotspots, or different genome architectures (Baudat et al., 2010; Mercier et al., 2015; Smukowski Heil et al., 2015).

1.2.3 Genetic features correlated with recombination rates

Variation in recombination rate across genomes and taxa is influenced by multiple correlated factors. Many act through chromosomal structure, sequence context, or epigenetic regulation. Chromosome size and architecture play a fundamental role. Smaller chromosomes, or chromosome arms, typically experience higher recombination per physical unit because at least one crossover for one homologous pair is required for proper segregation (Bascón-Cardozo et al., 2024; Farré et al., 2013;

Näsvall et al., 2023). Crossover events are frequently concentrated toward distal telomeric regions and are suppressed near centromeres and heterochromatin, reflecting the effect of structural and epigenetic environments on the crossover formation (Bascón-Cardozo et al., 2024; Dreissig et al., 2020; Kianian et al., 2018).

Epigenetic and chromatin features are also major determinants of recombination activity. Regions of open chromatin, characterised by low nucleosome occupancy and active histone marks such as H3K4me3, tend to coincide with hotspots of double-strand break formation and crossover occurrence (Borde et al., 2009; Castellani et al., 2024; Choi et al., 2013). In contrast, repressive chromatin features such as DNA methylation and histone H3K9me2 are typically associated with reduced recombination frequency (Kianian et al., 2018; Underwood et al., 2018). In addition, sequence compositions such as high GC content and low repetitive element density are also often linked to recombination hotspots (Beye et al., 2006; Chen et al., 2006). GC-biased gene conversion associated with recombination can further reinforce these patterns by driving the local enrichment of GC bases (Weber et al., 2014).

Genetic variation in recombination-modifier genes represents another important source of heterogeneity. For example, allelic differences in genes such as the cohesin subunit REC8 have been associated with differences in genome-wide recombination rate in plants such as barley (Dreissig et al., 2020; Sandor et al., 2012). In *Arabidopsis thaliana*, HEI10 overexpression, the zyp1 mutant, the recq4 mutant, and various combinations of these three variants show different levels of elevated chromosome-wide recombination rates and frequent heterochiasmy inversions (Jing et al., 2025). Structural genomic variation, such as inversions and segmental duplications, can locally suppress recombination by hindering proper homologous pairing (Rönspies et al., 2022; Sherizen et al., 2005), while juxtaposition of homozygous and heterozygous regions tends to displace crossovers to heterozygous regions (Ziolkowski et al., 2015).

1.2.4 Relations between recombination variation and adaptive evolution

Variation in recombination rate and pattern has important evolutionary implications. Recombination generally speeds up adaptation and increases the efficiency of natural selection. Conversely, it is also a heritable feature that responds to selection.

As recombination breaks genetic linkages, it can unlink the beneficial mutations from nearby deleterious mutations. Under the Hill-Robertson effect, this linkage usually

elongates the fixation time of beneficial alleles, thus slowing down the adaptation rate (Hill & Robertson, 1966; Whigham & Spencer, 2021). At the same time, recombination may combine beneficial mutations together and increase fitness and adaptation. The genetic variation derived from recombination provides a wider raw genetic substrate that natural selection can act upon. This is especially valuable in changing environments, as it increases the chances that some individuals will have a combination of alleles suited for survival (Bourguet et al., 2003; Uecker & Hermisson, 2016). However, recombination does not always have beneficial consequences because it can also break good combinations, but this situation should be against by selection. Therefore, recombination frequency cannot be too high. Even at the individual level, the chances of an error increase dramatically with an extremely high recombination rate, leading to genome instability and disease in humans and mice (Alves et al., 2017; Reliene et al., 2010). So recombination frequency through evolution must strike a balance. It needs enough recombination to purge bad mutations and create new combinations for adaptation, but not so much that it breaks up successful supergenes or causes catastrophic errors during meiosis.

Empirical evidence for population-level effects of recombination variation comes from diverse taxa. In *Drosophila melanogaster*, recombination rate differences between temperate and tropical populations correlate with adaptive divergence across climate-related loci. (Chan et al., 2012)) showed that genomic regions with higher recombination exhibit stronger signals of local adaptation, supporting the idea that recombination facilitates selective differentiation by reducing linkage constraints. Similar findings have also been reported in plants. For instance, in *Zea mays*, (Epstein et al., 2024)) demonstrated that recombination rate increases and landscape pattern changed after domestication, with the acquisition of recombinations in the originally non-recombinating regions in the wild type that linked to reduced efficacy of purifying selection.

In addition, the mechanisms controlling recombination may themselves be under selection. A well-studied example is PRDM9, which specifies recombination hotspots in most mammals. The PRDM9 gene exhibits rapid evolution in its DNA-binding zinc finger domain, with high allelic turnover between populations and species (Baudat et al., 2010; Smagulova et al., 2016). This rapid diversification is consistent with recurrent positive selection, likely driven by the “hotspot paradox”, i.e., recombination events alter their own recognition sites via gene conversion, favouring the emergence of new

PRDM9 alleles that target new sites (Myers et al., 2010). The resulting evolutionary arms race between PRDM9 and its binding motifs shows how selection avoids loss of recombination activity. Thus, understanding how recombination variation arises and is maintained is important both mechanistically and for broader evolutionary biology, including adaptation, speciation, and genome evolution.

In summary, recombination remains a crucial yet complex process. Its variation in rate and distribution across gametes, sexes, individuals, populations and taxa is shaped by molecular mechanisms (DSB distrubution, CO regulation, hotspot targeting), chromosome architecture (axis-loop structure, synapsis, interference), sequence and epigenetic context (GC content, chromatin marks), genetic variation (modifier genes), demographic and structural factors (population history, inversions), and even environmental and sex-specific influences. Methods to study recombination have advanced from cytology to high-resolution genomic maps and population-genetic inference, enabling richer insights into recombination landscapes. The concept of hotspots, such as the role of PRDM9 in mammals, highlights one paradigm of fine-scale variation, but also reveals that recombination systems are far from uniform across life. Ultimately, variation in recombination is a central driver of genome evolution, influencing linkage, diversity, adaptive potential and speciation. Further research is still expected to decipher how recombination is regulated, evolves, and contributes to adaptive evolution in diverse organisms.

1.3 Holocentric chromosome organisation and meiosis

Most meiosis and recombination variation mentioned in the two sections above is based on eukaryotes with one centromere on each chromosome. The centromere is essential for cell division as a site for attachment of spindle microtubules. Its fundamental role in both mitosis and meiosis is to ensure the faithful segregation of genetic material by orchestrating chromosome movement. However, the centromere cannot only reside at a single primary constriction site on chromosomes, but also diffuse to the entire chromosome in some species, which is usually called holocentromeres (**Figure 2**).

Holocentric chromosomes represent many unique features distinct from the conventional monocentric chromosome structure found in most eukaryotes. Monocentric chromosomes possess only one localised centromere that serves as the

primary site for kinetochore assembly. However, holocentric chromosomes exhibit multiple centromeric units distributed along their entire length, enabling multiple microtubule attachments along the entire chromosome. This feature has evolved multiple times independently in both animals and plants (Melters et al., 2012). Thus, monocentric chromosomes usually present a V shape during anaphase when spindle fibres are anchored to a localised centromere. While holocentric chromosomes show a parallel conformation, owing to the distributed tension along the entire length (**Figure 2**). These two observations, lack of a single primary constriction and parallel migration of chromatids during anaphase, are also commonly used to identify holocentricity cytologically. Genomic methods to confirm holocentrity usually rely on CENH3 ChIP-sequencing because the CENH3 (histone H3 variant) protein usually marks the positions of centromeres. However, holocentric lepidoptera species such as butterflies and moths lack CENH3. Kinetochores in the parasitic plant *Cuscuta europaea* attach to regions with and without CENH3 (Neumann et al., 2023). These examples suggest the centromere determination in holocentric species is highly plastic. Therefore, before confirming the fully functional centromere proteins, cytological observation is the only trustworthy approach.

The differences between monocentric and holocentric systems extend beyond simple kinetochore localisation because the centromere is essential for genome -organization, chromosome evolution, and meiotic behaviour.

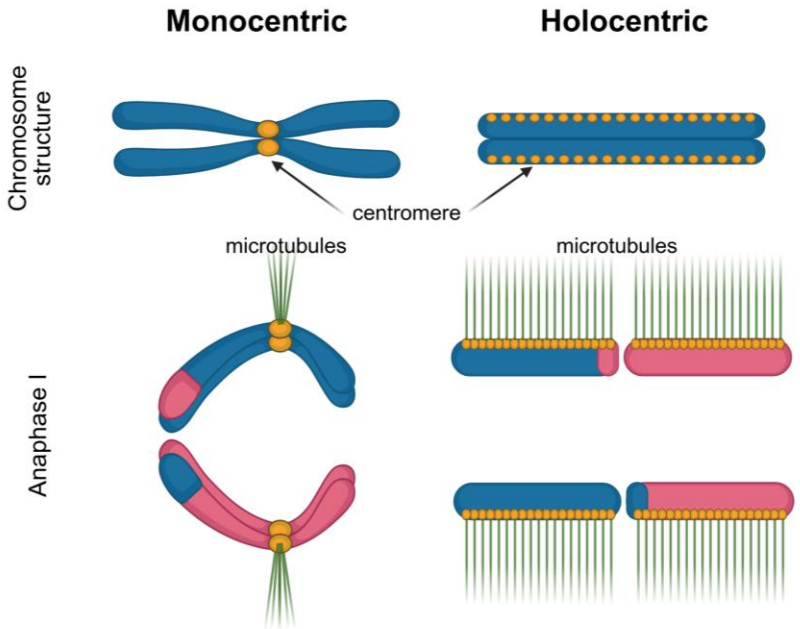


Figure 2. Comparison of monocentric and holocentric chromosomes. The monocentric chromosome has one centromere at one primary constriction. The holocentric chromosome can have multiple centromeres distributed along the entire length. In meiosis anaphase I, bivalents formed by homologous chromosomes segregate in monocentric species. But in holocentric species, sister chromatids segregate during anaphase I, a phenomenon called inverted meiosis.

In terms of genome organisation, holocentric chromosomes differ markedly from monocentric ones. A study of holocentric beak-sedges compared with a closely related monocentric species found that centromere function is distributed across hundreds of small domains rather than one large centromere region, and each centromere domain contains satellite DNA called *Tyba*. Furthermore, gene distribution across holocentric chromosomes appeared more uniform rather than being depleted at (peri)centromeric regions like monocentric species. Inter-chromosomal interactions were substantially reduced relative to monocentric relatives (Hofstatter et al., 2022). These findings suggest that holocentricity fundamentally alters genome architecture. Instead of a strong centromere-proximal domain of suppressed recombination and heterochromatin, the entire chromosome may participate more uniformly in gene expression and chromatin interaction.

Another important consequence due to the dispersed centromeres is that in holocentric chromosomes, nearly any chromosomal fragment retains centromere function. This means that fragments generated by fission or fusion events can segregate correctly rather than being lost, which is usually the case in monocentric systems. Thus, many studies argue that holocentricity confers increased tolerance to chromosomal rearrangements and potentially facilitates rapid karyotype evolution.

The meiotic behaviour of holocentric chromosomes poses distinct challenges compared to monocentric ones. Because kinetochores are distributed along the chromosome arms, the conventional mechanisms of sister-chromatid cohesion, mono- or bi-orientation of kinetochores, and stepwise loss of cohesion may require modifications. For example, in the holocentric plant *Luzula elegans*, sister chromatids behave as functional units during meiosis I, which remain unfused and segregate at anaphase I, rather than homologs segregating during meiosis I in monocentric species. Non-sister chromatids remain linked by heterochromatin threads enriched in satellite DNA until metaphase II, at which stage segregation of homologous non-sister chromatids occurs (Heckmann et al., 2014). This phenomenon is called “inverted meiosis”, which is one of the evolutionary adaptations to holocentric organisation.

However, not all holocentric species adopted this strategy. In holocentric nematodes like *Caenorhabditis elegans*, meiosis does not go through inverted meiosis, but, akin to monocentric species, segregates homologous chromosomes in meiosis I, then sister chromatids in meiosis II.

In summary, although holocentric species seem fewer compared to monocentric ones, they account for ~20% eukaryotes and have evolved convergently many times (Márquez-Corro et al., 2024), suggesting selective advantages under certain environmental or karyotypic pressures, for instance, fragmentation tolerance under high clastogenic stress. When the holocentric chromosome architecture imposes challenges for meiosis, the evolutionary solutions vary, indicating a remarkable plasticity of holocentric organisms.

1.4 Polyploid meiosis with challenges

The key challenge in polyploid meiosis lies in the ambiguity of partner recognition. In diploids, each chromosome has a unique homolog, whereas in polyploids, multiple nearly identical copies coexist, creating pairing competition. This can lead to multivalent associations, in which over two chromosomes pair simultaneously, or to erroneous pairing between homoeologous chromosomes originating from different ancestral genomes, which usually disrupt segregation at anaphase I, frequently resulting in unbalanced gametes and reduced fertility (Bomblies, 2023). Studies of natural *Arabidopsis arenosa* autotetraploids have shown how selection acts on genes controlling chromosome axis formation and synapsis to restore orderly bivalent pairing. These populations have evolved allelic variants of meiotic proteins such as ASY1, ASY3, REC8, ZYP1, and HEI10 (Gonzalo et al., 2025; Hollister et al., 2012; Wright et al., 2015), which tune the timing and organisation of synapsis, leading to more synchronous and complete homolog alignment, suggesting that meiotic stabilisation in polyploids is not achieved through a single genetic innovation but through the coordinated adjustment of multiple components controlling chromosome alignment and recombination.

Allopolyploids, formed by hybridisation of related species, face the additional difficulty of distinguishing homologous from homoeologous chromosomes during pairing. Suppressing recombination between subgenomes is crucial for fertility and genomic integrity. In wheat, this function is largely governed by the Pairing homoeologous 1

(Ph1) locus, which includes a diverged copy of the meiosis gene ZIP4. This gene enforces correct homolog recognition and effectively suppresses homoeologous crossover, allowing hexaploid wheat to maintain balanced segregation despite its three subgenomes (Rey et al., 2018). Similar mechanisms appear in *Brassica napus*, an allotetraploid derived from hybridisation between *B. rapa* and *B. oleracea*. In resynthesized *B. napus* lines, instability persists during early generations, with frequent homoeologous exchanges and fertility reduction, but over time, selection favours allelic combinations that improve meiotic regularity (Xiong et al., 2011). These observations suggest that stable polyploid meiosis can evolve through multiple molecular routes.

The reproductive biology of odd-ploidy species introduces further complications. Triploids or pentaploids, possessing an uneven number of chromosome sets, cannot normally produce balanced haploid gametes, as meiosis yields highly variable and often aneuploid chromosome complements. Nevertheless, odd-ploid lineages persist in nature through mechanisms that bypass or modify meiosis. Some reproduce asexually via apomixis or vegetative propagation, thereby avoiding the requirement for balanced reduction (Cornaro et al., 2025). Others produce unreduced (2n) gametes that, upon fusion, restore even ploidy levels in subsequent generations (Husband, 2004). In crops such as *Brassica*, triploid hybrids often show elevated recombination rates on certain subgenomes, indicating that while meiosis is error-prone, it can still proceed to completion, generating genetic diversity, albeit with lower fertility (Pelé et al., 2017). These pathways demonstrate that meiotic irregularity does not necessarily preclude evolutionary success but can act as a transient state leading toward new, stabilised polyploid lineages.

However, it was found that dogroses with a pentaploid genome (2*n*=5*x*=35), unlink most odd-polyploids, can reproduce sexually and produce viable seeds generation after generation. The question of how they manage to segregate chromosomes in an orderly, reproducible way, even though normal pairing rules shouldn’t allow it and how they do this through a highly unusual mechanism, the so-called Canina meiosis, has fascinated biologists for over a century. It was documented that dogroses have only seven bivalents (14 chromosomes) while the remaining 21 chromosomes remain as univalents during meiosis (Täckholm, 1922) thus the bivalent-forming chromosomes are preferentially transmitted through the pollen gametes as a haploid set, whereas the egg typically receives the bivalents plus the univalents, thus producing tetraploid

eggs (**Figure 3**). Univalent chromosomes exhibit a skewed distribution of rDNA families and expansion of particular pericentromeric satellite repeats (Kovarik et al., 2008; Lunerová et al., 2020). Dissection of the underlying mechanism of this balanced heterogamy will help to uncover more evolutionary solutions to complicated meiosis.

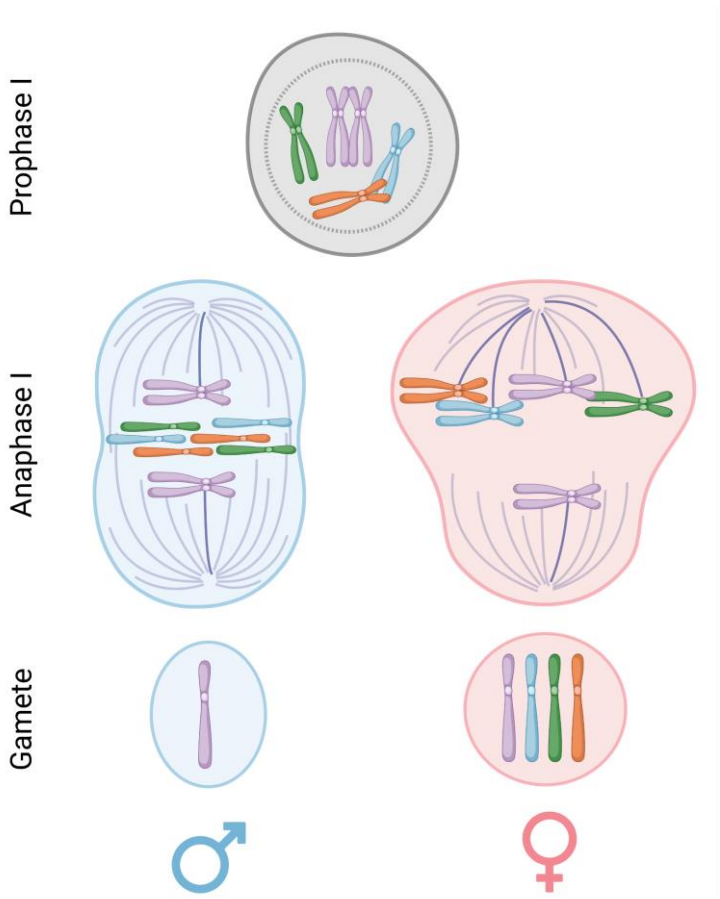


Figure 3. A diagram of the pentaploid dogrose meiotic prophase I, anaphase I, and gamete genotype. Here I use one chromosome represents one haplotypes, but dogroses normally have 7 chromosome in one haplotype (2*n*=5*x*=35). The bivalent-forming chromosomes are displayed in the same colour, which synapse during prophase I, while other chromosomes do not have synapsis partners. The left and right sides show male and female meiosis, respectively. Sister chromatids of univalent-forming chromosomes in prophase I already segregate and lag around the equator in anaphase I during male meiosis. In female meiosis, synapsed chromosomes (bivalents) segregate in anaphase I and all univalent chromosomes migrate in the same direction to one of the poles. In anaphase II, all unsegregated sister chromatids separate in both male and female meiosis, which is not shown here. The viable pollen gamete contains one chromosome set, and the egg contains four sets.

Across systems, the recurring themes are the refinement of synapsis dynamics, the modulation of crossover regulation, and the suppression of homoeologous pairing. These findings not only clarify how natural polyploids overcome the meiotic bottleneck but also inform breeding strategies aimed at generating fertile synthetic polyploids or

controlling recombination in polyploid crops. By integrating mechanistic insight with evolutionary perspective, future research may allow deliberate engineering of meiotic stability, which could profoundly expand the utility of polyploidy in agriculture.

1.5 Structure of this dissertation

This dissertation includes four projects, aiming to characterise the meiotic process in holocentric plants, *Rhynchospora*, and to solve the mechanism of meiosis in pentaploid dogroses.

The first study establishes the first recombination landscape of holocentric plants, and a complete pipeline for constructing the recombination landscape of *Rhynchoposra breviuscula* using scRNA sequences from pollen grains, from which we found that the recombination rate is distally biased despite even holocentromere dispersion. All detected (epi)genetic signals of *R. breviuscula* are uniformly distributed along its chromosomes. Therefore, we further studied meiotic dynamics in holocentric plants for the first time and revealed that the telomere-led pairing and synapsis shaped the recombination landscape of *R. breviuscula*.

The second study provides an insight into the impact of karyotype evolution on meiotic recombination distributions. We applied single-cell sequencing techniques described in the first study to other species in the genus *Rhynchospora* and compared their recombination landscapes and karyotypes. The massive chromosome rearrangements within the *Rhynchoposra* imply the holocentric plants' high tolerance to chromosome fusions and breaks. With the uniformly distributed (epi)genetic features across *Rhynchospora* species, we consolidate our previous hypothesis from the first study that chromatin dynamics during meiotic progression are more likely to shape recombination patterns than genomic sequences.

The third study investigates a unique species in the genus *Rhynchospora*, i.e. achiasmatic *Rhynchospora tenuis*. Neither our cytological observations nor genetic analysis showed any evidence of crossover occurring in *R. tenuis*, but it still retains efficient sexual reproduction. Furthermore, the selfed *R. tenuis* produced the same genotype as the parental individual, as clonal offspring. By applying the same genotyping method to 5 different accessions of *R. tenuis*, we revealed that achiasmata is a widespread phenomenon in this species despite variant segregation bias and distortion in male meiosis. This chapter mainly characterises this distinct species and

proposes a novel sexual reproduction mechanism bypassing crossovers and mimicking clonality.

The last project characterises a pentaploid dogrose *Rosa canina* (2*n*=5*x*=35), which achieves a stable sexual reproduction despite the uneven ploidy by segregating chromosomes asymmetrically during meiosis, i.e. 14 form bivalents (inherited biparentally), while 21 univalents are maternally transmitted. Genome analyses revealed this system relies on two homozygous bivalent-forming haplotypes from one subgenome and three divergent, univalent-forming subgenomes, suggesting hybrid origins from distinct *Rosa* clades. Surprisingly, bivalent-forming centromeres are enriched in ATHILA retrotransposons, while centromeres of univalents are dominated by tandem repeats CANR4, implying structural differences drive their meiotic fate. This centromere bimodality may enable exclusive univalent transmission during female meiosis. The study uncovers how dogroses bypass polyploid sterility through asymmetric meiosis and centromere diversity, offering insights into asymmetric inheritance and genome evolution.

2. Results

2.1 Meiotic recombination dynamics in plants with repeat-based holocentromeres shed light on the primary drivers of crossover patterning

Published in Nature Plants 10, pages 423–438 (2024)

Marco Castellani*, Meng Zhang*, Gokilavani Thangavel, Yennifer Mata Sucre, Thomas Lux, José A. Campoy, Magdalena Marek, Bruno Huettel, Hequan Sun, Klaus F. X. Mayer, Korbinian Schneeberger, and André Marques#
*These authors contributed equally
#Correspondence

Author Contributions:

A.M. conceived the research programm and coordinated the analyses. M.C. performed all cytogenetic analyses and microscopy. M.C. isolated the pollen nuclei and generated sequencing libraries with assistance from J.A.C. M.Z. performed all single-cell RNA sequencing and recombination-related analyses with assistance from H.S. G.T. performed the ChIP–seq analysis. Y.M.S. performed the immuno-FISH analysis. T.L. and K.F.X.M. performed the gene annotation and Ka/Ks ratio analysis. M.M. operated the FACS machine. B.H. performed all sequencing. K.S. supervised the single-cell analysis. M.C., M.Z. and A.M. wrote the first manuscript draft with input from all authors. All authors approved the final version of the manuscript.

nature plants

Article

https://doi.org/10.1038/s41477-024-01625-y

Meiotic recombination dynamics in plants with repeat-based holocentromeres shed light on the primary drivers of crossover patterning

Received: 24 August 2023

Accepted: 15 January 2024

Published online: 9 February 2024

Check for updates

Marco Castellani[1,10], Meng Zhang[1,10], Gokilavani Thangavel[1], Yennifer Mata-Sucre[1,2], Thomas Lux[3], José A. Campoy[1,8], Magdalena Marek[4], Bruno Huettel[4], Hequan Sun[1,5,9], Klaus F. X. Mayer[3,6], Korbinian Schneeberger[1,5,7] & André Marques[1,7]

Centromeres strongly affect (epi)genomic architecture and meiotic recombination dynamics, influencing the overall distribution and frequency of crossovers. Here we show how recombination is regulated and distributed in the holocentric plant *Rhynchospora breviuscula*, a species with diffused centromeres. Combining immunocytochemistry, chromatin analysis and high-throughput single-pollen sequencing, we discovered that crossover frequency is distally biased, in sharp contrast to the diffused distribution of hundreds of centromeric units and (epi)genomic features. Remarkably, we found that crossovers were abolished inside centromeric units but not in their proximity, indicating the absence of a canonical centromere effect. We further propose that telomere-led synapsis of homologues is the feature that best explains the observed recombination landscape. Our results hint at the primary influence of mechanistic features of meiotic pairing and synapsis rather than (epi)genomic features and centromere organization in determining the distally biased crossover distribution in *R. breviuscula*, whereas centromeres and (epi)genetic properties only affect crossover positioning locally.

During meiosis, homologous chromosomes (homologues) undergo meiotic recombination, exchanging genomic material between them. This exchange is initiated by physiologically induced DNA double-strand breaks (DSBs)[1,2]. The formation of meiotic DSBs is commonly resolved via crossovers (COs) or other recombination outcomes, called non-COs[3]. COs can be divided into two classes[4,5]. In plants, class I COs are the most prevalent and are sensitive to interference—they do not occur near each other along a chromosome—whereas class II

[1]Department of Chromosome Biology, Max Planck Institute for Plant Breeding Research, Cologne, Germany. [2]Laboratory of Plant Cytogenetics and Evolution, Department of Botany, Centre of Biosciences, Federal University of Pernambuco, Recife, Brazil. [3]Plant Genome and Systems Biology, German Research Centre for Environmental Health, Helmholtz Zentrum München, Neuherberg, Germany. [4]Max Planck Genome-Centre Cologne, Max Planck Institute for Plant Breeding Research, Cologne, Germany. [5]Faculty of Biology, Ludwig-Maximilians-Universität München, Planegg-Martinsried, Germany. [6]School of Life Sciences Weihenstephan, Technical University of Munich, Freising, Germany. [7]Cluster of Excellence on Plant Sciences (CEPLAS), Heinrich-Heine University, Düsseldorf, Germany. [8]Present address: Department of Pomology, Estación Experimental de Aula Dei (EEAD), Consejo Superior de Investigaciones Científicas, Zaragoza, Spain. [9]Present address: School of Automation Science and Engineering, Faculty of Electronic and Information Engineering, Xi'an Jiaotong University, Xi'an, China. [10]These authors contributed equally: Marco Castellani, Meng Zhang. e-mail: amarques@mpipz.mpg.de

COs are insensitive to interference and can accommodate around 10% of the total COs[4]. Although the repertoire of meiotic-specific proteins is largely conserved across eukaryotes[6,7], species-specific adaptations can occur[8].

The distribution of COs is typically associated with the distribution of genetic and epigenetic ((epi)genetic) features[9,10]. In most eukaryotes, the CO frequency correlates positively with gene/euchromatin density[11,12], but is lower in heterochromatic regions, including (peri)centromeres[13,14]. In monocentric species, centromeres are single-defined structural entities and are typically repeat-based[15]. Recombination is largely suppressed at and in the proximity of centromeres in these species, a phenomenon called the centromere effect[16–19]. This feature, combined with (epi)genetic elements, is believed to cause a distal bias in CO frequencies[9,12,20–23]. However, how these factors influence meiotic recombination patterning at broad and local scales is still elusive.

Monocentricity is not the only centromeric organization adopted by eukaryotes. For instance, holocentric species harbour multiple centromeric determinants over the entire length of their chromosomes[24,25]. Holocentricity has evolved independently multiple times in nematodes, insects and plants[26,27]. In the holocentric animal models *Caenorhabditis elegans* and silk moth (*Bombyx mori*), holocentromeres do not associate with a specific sequence and have variable dynamics[28,29]. By contrast, holocentric plants of the genus *Rhynchospora* (beaksedges) display specific repeats constitutively loaded with holocentromeres in both mitosis and meiosis[30,31]. Recently, we sequenced the genomes of three beaksedges (*R. breviuscula*, *R. pubera* and *R. tenuis*) and determined that each chromosome harbours hundreds of short arrays (~20 kb each) of the specific Tyba (meaning "abundance" in Tupi-Guarani, a language spoken by many Brazilian native tribes) tandem repeat that are distributed genome-wide and specifically associated with centromeric histone H3 protein (CENH3)[32]. This particular chromosome organization is associated with a remarkably uniform distribution of genes, repeats and (epi)genetic features, in contrast to the compartmentalized chromosome organization found in many monocentric eukaryotes[32]. Remarkably, each individual centromeric unit in *R. pubera* showed epigenetic regulation similar to that in other plant monocentromeres[23,32]. Moreover, meiosis has been only studied in holocentric plants regarding their intriguing 'inverted meiosis' and centromere organization[26,33–35]. Thus, *Rhynchospora* offers an excellent model to study the mechanisms of CO formation in the absence of the effect of the monocentromere, while sharing similar centromere chromatin (epi)genetic properties.

Here, we use *R. breviuscula* as a model to study meiotic recombination dynamics in the absence of both a localized centromere and a compartmentalized chromosome organization, features that potentially mask underlying factors affecting CO distribution in most genomes. Using a combination of immunocytochemistry, chromatin and DNA analysis, and CO calling from single-gamete sequencing, we offer an overview of meiotic recombination dynamics and distribution for a species with repeat-based holocentromeres. Remarkably, the megabase-scale CO distribution did not correlate with any (epi) genomic feature analysed. We show that the CO distal bias is achieved even in the absence of both a monocentromere and a correlation with (epi)genomic features. We found that COs are suppressed inside repeat-based centromeric units but not in their proximity, indicating the absence of a centromere effect. Moreover, our cytological data suggest that synapsis dynamics starting from chromosomal ends influences the broad-scale recombination landscape in *R. breviuscula*. We propose that centromere and (epi)genetic features play a role in CO positioning, but only at the fine scale.

Results

Canonical dynamics of early meiosis I in *R. breviuscula*

Chromosome spreads on male meiocytes of *R. breviuscula* revealed all the classical prophase I stages with the occurrence of five bivalents, indicating the formation of at least one CO per homologue pair (Extended Data Fig. 1a). Moreover, we confirmed the holocentric nature of *R. breviuscula* chromosomes by showing the localization of CENH3 in mitosis and meiosis. Similar to what has been reported in *R. pubera*[31], CENH3 appears as lines during mitotic metaphase, but undergoes restructuring into more irregular clusters during meiosis (Extended Data Fig. 1b–d).

We then investigated ASY1 (refs. 36,37) and ZYP1 (refs. 38–42) as indicators of conserved and functional axis formation and synapsis, respectively. The ASY1 signal was present along the entire length of unsynapsed chromosomes at leptotene (Extended Data Fig. 2a). During zygotene, paired ASY1 linear signals could be followed until they converged and lost intensity (Fig. 1a), whereas ZYP1 was gradually loaded onto synapsed chromosomes (Fig. 1a,b). In pachytene, complete synapsis was evidenced by the linear ZYP1 signal along the full length of chromosomes (Fig. 1c). We also tested whether the meiosis-specific alpha-kleisin REC8, a protein responsible for sister chromatid cohesion[43], is conserved in *R. breviuscula*. We detected a conserved linear REC8 signal at pachytene co-localizing with ZYP1, forming a linear signal along the entirely synapsed chromosome (Extended Data Fig. 2b). Thus, pairing, cohesion and synapsis are conserved in the holocentric plant *R. breviuscula*, resembling those in monocentric models.

We further studied the behaviour of HEI10, a RING-family E3 ligase that has been characterized in plants, fungi and animals[44–47]. HEI10 has been proposed to interact with both early and late recombination proteins, stabilizing recombination intermediates and advancing them into class I COs[44,46–50]. In *R. breviuscula*, when synapsis started in early zygotene, HEI10 was immediately loaded as a dotty linear signal co-localizing with the first ZYP1 signals (Fig. 1b). At early pachytene, HEI10 progressed to form an increasingly linear signal co-localizing with ZYP1 along the entirety of synapsed chromosomes (Fig. 1c). Throughout pachytene the HEI10 signal became discontinuous along chromosomes, and a few foci increased in intensity (Fig. 1d). We think that these are putative class I CO sites. In diplotene and diakinesis, only high-intensity foci remained (Fig. 1e and Supplementary Fig. 1). Thus, the dynamic behaviour of HEI10 is conserved in *R. breviuscula*, displaying the 'coarsening' dynamics recently proposed in *Arabidopsis*[51–53].

Another established marker for meiotic recombination is the mismatch repair protein MLH1 (MUTL-HOMOLOG 1), which has a meiosis-specific resolvase activity to process double Holliday junctions into final class I COs. MLH1 interacts with MSH4 and MSH5, specifically marking class I COs[54–56]. In *R. breviuscula*, the MLH1 signal appears exactly at diplotene, after synaptonemal complex (SC) disassembly, and persists as bright foci until the end of diakinesis (Fig. 1f and Supplementary Fig. 2). We observed a median number of six foci for both HEI10 ($n = 69$) and MHL1 ($n = 83$) (Fig. 1g), which is consistent with the formation of at least one CO per homologue pair. Foci counting of these two markers showed a small but significant difference (Fig. 1g). We also observed at least one homologue pair with more than one CO; that is, ring bivalent (Supplementary Fig. 3). Thus, in contrast to the holocentric *C. elegans* where strictly only one CO is allowed per bivalent[57], *R. breviuscula* can form more than one CO per bivalent. Furthermore, our results confirm a canonical and conserved early meiosis programme in *R. breviuscula*.

Pollen single-cell sequencing allows genome-wide CO detection

We set out to determine whether meiotic recombination in *R. breviuscula* is affected by the genome-wide distribution of holocentromeres. To identify COs from a single *R. breviuscula* individual, we took advantage of its 1% heterozygous genome[32] to generate a fully haplotype-phased chromosome-scale reference genome (Extended Data Fig. 3, Supplementary Tables 1 and 2 and Methods). Moreover, *R. breviuscula* is an outbred species with high levels of self-incompatibility, which hampers the standard detection of COs, typically involving the time-consuming

https://doi.org/10.1038/s41477-024-01625-y

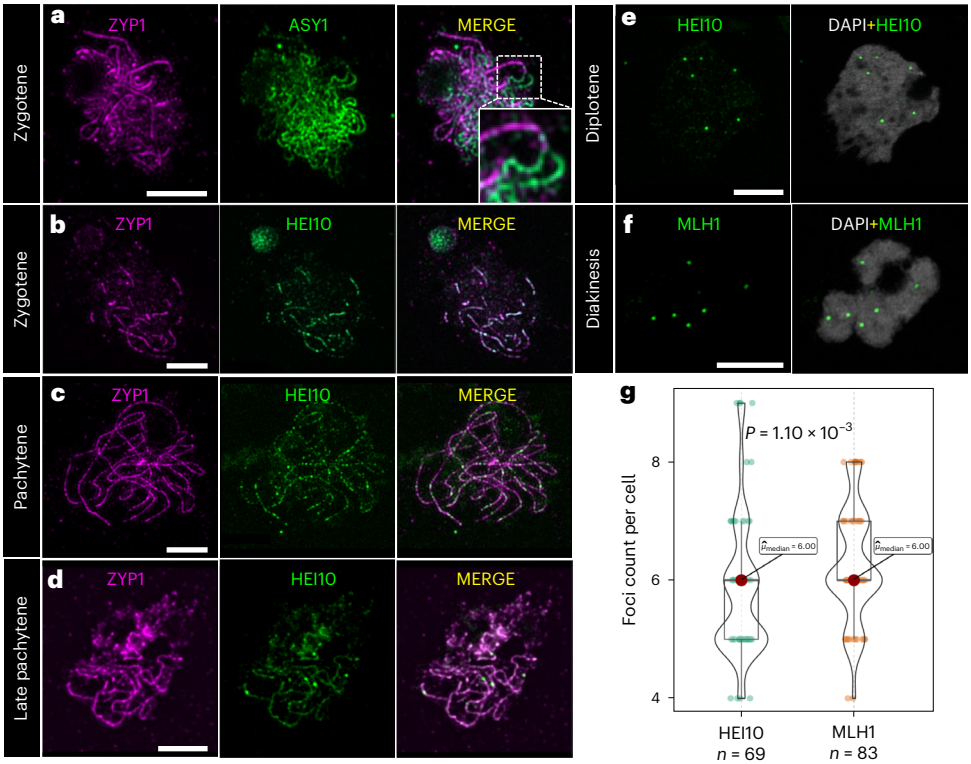


Fig. 1 | Immunolocalization of ZYP1, ASY1, HEI10 and MLH1 in prophase I. **a**, In zygotene, synapsis is visualized as the loading of ZYP1 while the ASY1 signal disappears. The insert shows a magnification of two unpaired chromosomes, represented by ASY1, coming together to synapse, with loss of the ASY1 signal and the loading of ZYP1. The behaviour of ASY1 + ZYP1 was consistent in all cells at zygotene ($n = 16$) in nine independent experiments. **b**, In early zygotene when synapsis starts, HEI10 is immediately loaded as many closely spaced foci, forming an irregular and patchy signal and co-localizing with ZYP1. **c**, In pachytene, HEI10 is visible as lines that co-localize with ZYP1. **d**, In late pachytene, the linear signal of HEI10 co-localizes with ZYP1 but becomes weaker, except for a few highly intense foci. The behaviour of HEI10 + ZYP1 was consistent in all cells from zygotene to late pachytene ($n = 61$) in ten independent experiments. **e**, During the diplotene and diakinesis stages, HEI10 appears as foci only on bivalents, with no linear signal. The behaviour of HEI10 from diplotene to diakinesis was consistent in all cells analysed ($n = 69$) in ten independent experiments. **f**, MLH1 appears in diplotene and diakinesis as foci on bivalents, representing chiasmata. The behaviour of MLH1 from diplotene to diakinesis was consistent in all cells analysed ($n = 83$) in eight independent experiments. **g**, Box plots of HEI10 and MLH1 foci count at late prophase I. Two-sided Mann–Whitney *U*-tests were performed. The box plots are comprised of minima, first quartile (Q1), median ($\hat{\mu}_{median}$; HEI10 = 6.00, MLH1 = 6.00), third quartile (Q3) and maxima following the definition of ggplot2 in R. Being produced in the same animal, co-localization of these two markers could not be performed. Maximum projections are shown for microscope images, with chromosomes counterstained with DAPI. Scale bar, 5 µm.

generation of segregating offspring. Thus, we obtained only 63 F_1 plants by manual selfing-pollination that were sequenced at threefold genome coverage. Although, we detected 378 CO events at a very high resolution (CO resolution: median 334 bp, mean ~2 kb), the limited number of detected COs did not allow comprehensive drawing of the recombination landscape (Supplementary Fig. 4).

Next, we aimed to provide a shortcut for constructing a robust broad-scale CO landscape for *R. breviuscula*. Gametes carry the outcome of meiotic recombination and can be obtained in large numbers in a relatively inexpensive manner from pollen grains. Thus, by adapting a strategy based on the gamete-binning method described by Campoy et al.[58], we identified genome-wide CO events by conducting 10x Genomics single-cell RNA sequencing (scRNA-seq) on extracted nuclei from pollen grains (male gametes) of *R. breviuscula* (Fig. 2a and Methods), with the caveat that pollen nuclei show a relatively low abundance of transcripts (Supplementary Fig. 5), limiting the resolution of CO detection.

To genotype the haploid gamete genomes, we obtained genome-wide haplotype-specific markers by aligning ~26 Gb Illumina whole-genome short reads to the phased haplotype 1 of *R. breviuscula*. We detected 820,601 haplotype-specific single nucleotide polymorphisms (SNPs; ~1 SNP per 449 bp) and used them as markers for genotyping (Fig. 2b and Supplementary Figs. 6 and 8a).

We pre-processed scRNA sequences by correcting barcodes, demultiplexing, and removing doublets and cells with a low number of reads (Supplementary Fig. 7 and Methods), resulting in a final set of 1,641 pollen nuclei with at least 400 markers (~1 marker per Mb). These markers covered almost the entire length of all five chromosomes (Supplementary Fig. 8b), ensuring genome-wide CO detection (CO resolution: median ~1.5 Mb, mean ~2.24 Mb). We detected 4,047 COs in the 1,641 pollen nuclei by inspecting genotype conversions, as indicated in Fig. 2c,d. Overall, we delineated a complete and detailed pipeline to detect COs in an economical way using high-throughput scRNA-seq of gametes from a single heterozygous individual (Fig. 2).

Deciphering the CO landscape in a holocentric plant

Counting the occurrence of COs in chromosome-wide genomic intervals across 1,641 pollen nuclei, we computed the CO rates along chromosomes and established a recombination map for *R. breviuscula* with known repeat-based holocentromeres (Fig. 3a and Supplementary Fig. 9). Unexpectedly, we observed an apparent distal bias in CO distribution, whereas centre regions maintain rather low CO rates, mostly lower than the mean CO rate genome-wide. Remarkably, chromosome 3 (chr3), chr4 and chr5 showed CO distal bias at both ends (Fig. 3a). By contrast, chr1 and chr2 showed a prominent increase in CO rate at only one chromosomal end (CO rate above the genome-wide mean),

https://doi.org/10.1038/s41477-024-01625-y

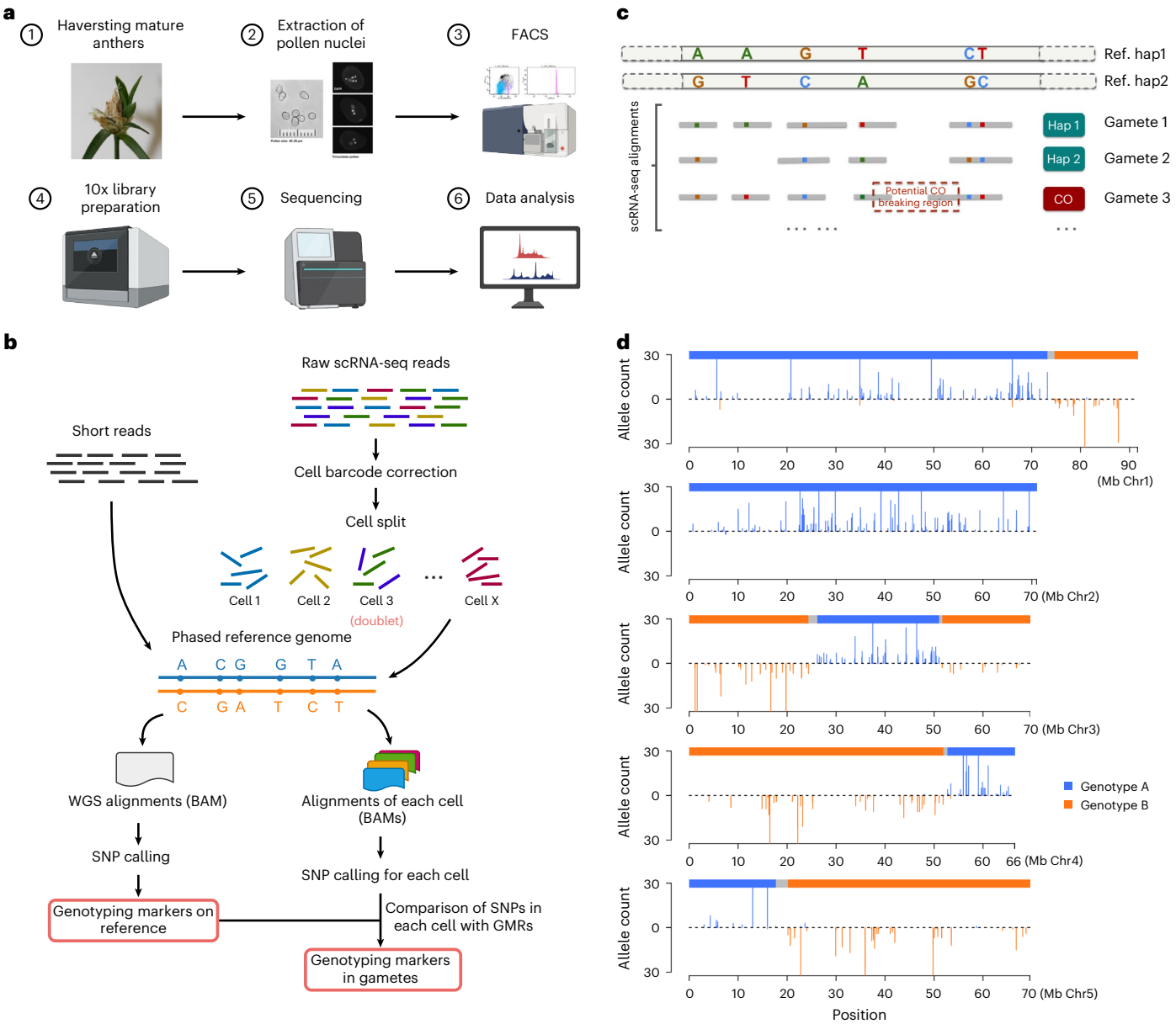


Fig. 2 | Overview of CO calling by adapting scRNA-seq to *R. breviuscula* gametes. a, Pollen sampling, library preparation and scRNA-seq pipeline. FACS, fluorescence activated cell sorting. **b**, Diagram of the strategy used for identifying genotyping markers on the reference genome by mapping short reads, and markers in gametes by mapping scRNA-seq reads, across a large number of gametes to the reference genome. GMR, genotyping marker on reference genome; WGS, whole-genome sequencing. **c**, Diagram of the identification of potential CO events after the alignment of scRNA reads from each gamete to the phased reference genome. **d**, An example of genotype definition by markers in a real pollen nucleus with the cell barcode AAGACTCTCATCCTAT.

whereas the other end showed comparatively low CO rates similar to centred regions. Interestingly, these low recombination ends were correlated with the localization of 35S ribosomal DNA loci (nucleolar organizing regions) (Fig. 3a,b). We thus revealed an uneven distribution of CO rates at the megabase-scale, despite the presence of holocentric chromosomes in *R. breviuscula*.

We further compared the number of genetically identified CO events derived from pollen nuclei with more than 2,000 markers (n = 81; 243 COs) and F_1 selfed offspring data. Although, the computed total linkage map length obtained from both pollen and F_1 offspring was 300 cM, slight differences in genetic length were observed among the chromosomes (Fig. 3b,c and Extended Data Fig. 4a,b). Differences in genetic length for the same chromosome between pollen and F_1 could be due to the occurrence of heterochiasmy[40], a phenomenon that causes strong differences in recombination rates between female and male meiosis. Remarkably, chr3, chr4 and chr5 had longer genetic lengths among all *R. breviuscula* chromosomes (Fig. 3b,c and Extended Data Fig. 4a,b). This is evident considering these three chromosomes have distal CO bias at both ends compared with only one in chr1 and chr2.

On average, we detected around three COs per haploid gamete, or 0.6 CO per chromatid (Fig. 3d,e), which is approximately half of the COs detected in our segregating offspring and HEI10/MLH1 foci (Figs. 1g and 3d), because gametes contain only one chromatid of each recombined chromosome. Furthermore, all single chromatids had, on average, one CO in half of these gametes (n = 81), whereas double COs appeared in approximately 5% of the 81 gametes considered (Fig. 3e). Chromosome 3 showed the highest frequency of double COs

https://doi.org/10.1038/s41477-024-01625-y

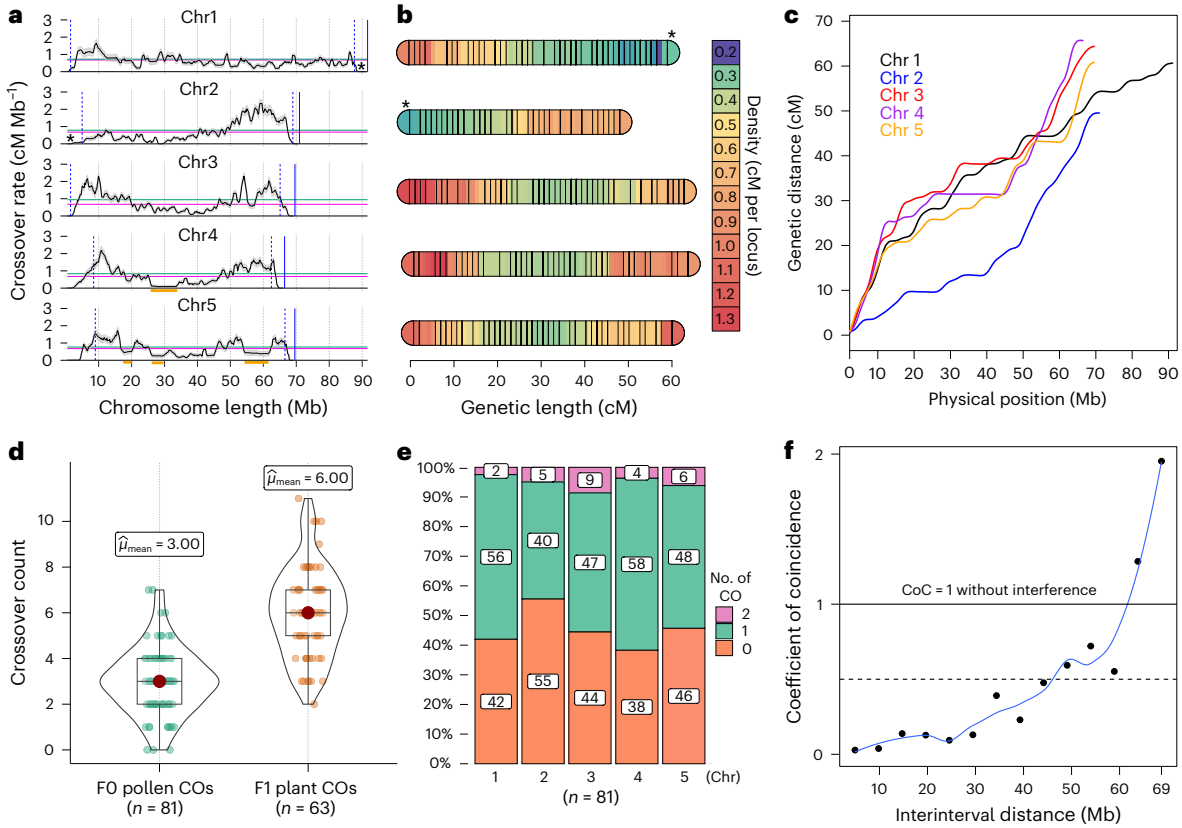


Fig. 3 | Gamete-sequencing derived meiotic recombination dynamics in *R. breviuscula*. **a**, Recombination landscape of the five chromosomes in *R. breviuscula* achieved by computing 4,047 COs from 1,641 pollen nuclei. The black line shows the CO rate; shadow ribbons indicate one standard deviation from mean CO rates; blue dashed vertical lines indicate the start and end of confident CO rate computation (Supplementary Fig. 9); blue solid vertical lines mark the chromosomal end; magenta horizontal lines indicate the genome-wide average CO rate; green horizontal lines are the chromosome-wide average CO rate; and orange bars indicate large (>2 Mb) homozygous regions with a reduced number of markers. **b**, Genetic linkage map computed from COs in 81 F_0 pollen nuclei having more than 2,000 markers. Genetic length density is indicated by the coloured scale. A set of 705 markers was selected using a 500 kb sliding window through all markers defined against the reference (Methods). The terminal locations of the *35S rDNA* locus on chr1 and chr2 are indicated by asterisks in **a** and **b**. **c**, Marey map calculated from the linkage map in **b**. Marey maps for each chromosome (colour lines) show genetic position as a function of physical position. **d**, CO number derived by counting CO events from the genetic analysis in pollen nuclei having more than 2,000 markers compared with the one extrapolated from the F_1 offspring. The box plots are comprised of minima, Q1, mean, Q3 and maxima following the definition of ggplot2 in R. N indicates biologically independent pollen nuclei and F_1 offspring individuals used for CO detection. **e**, Distribution of CO number for each single chromatid in gametes. Note the higher incidence of double COs on chr3. **f**, CoC curve in pollen nuclei (n = 1,641). Chromosomes were divided into 15 intervals, with random sampling at CO intervals, to calculate the mean CoC of each pair of intervals.

(9%; Fig. 3e and Extended Data Fig. 4c). Although we cannot discard the occurrence of class II COs, our results suggest that most COs formed in *R. breviuscula* are of class I.

We also tested whether CO interference occurred in *R. breviuscula*. A chi-squared goodness-of-fit test suggested that the observed COs did not reside randomly, and a further dispersion test revealed CO number distribution was underdispersed because of the effect of CO interference (Supplementary Fig. 10a). We also computed the coefficient of coincidence (CoC) for COs across the genome. The CoC measures the observed frequency of double COs over the expected frequency (Methods). The CoC curve of all chromosomes showed that the coefficients are below 1 for genomic intervals with distances of less than around 60 Mb (Fig. 3f and Supplementary Fig. 10b), meaning that the frequency of double COs is lower than expected. This result indicates the presence of strong CO interference in *R. breviuscula*.

CO landscape is independent of (epi)genomic features

The distally biased recombination landscape is ubiquitous across many eukaryotic species, and is typically explained by suppression of CO formation at pericentromeric regions and an association with large-scale epigenetic regulation[20,21]. Hence, we set out to correlate the recombination landscape of *R. breviuscula* with several (epi)genetic features. Surprisingly, a chromosome-wide comparison revealed no apparent correlation of CO rates with the even distribution of hundreds of repeat-based centromeric units and other genomic (genes, transposable elements (TEs), SNP densities or GC content) and epigenomic (such as H3K4me3, H3K27me3, H3K9me2 or DNA methylation) features (Fig. 4a and Extended Data Fig. 5a). Further, genome-wide comparison of the K_a/K_s ratio (a measurement of the relative rates of synonymous (K_s) and non-synonymous (K_a) substitutions at each gene) failed to identify fast-evolving genes correlated with higher recombination frequency regions (Extended Data Fig. 6). Quantification at the megabase scale of (epi)genomic features also revealed no strong correlation with CO distribution (Fig. 4b). These results indicate that, at the broad scale, meiotic recombination occurs independently of any (epi)genomic feature and chromosome-wide distribution of repeat-based holocentromeres.

Lack of centromere effect and the fine-scale CO regulation

Next, we tested whether COs are epigenetically affected at a fine scale and whether individual centromeric units have an effect on CO designation in *R. breviuscula*. For that we used only the set of 378 COs resolved

https://doi.org/10.1038/s41477-024-01625-y

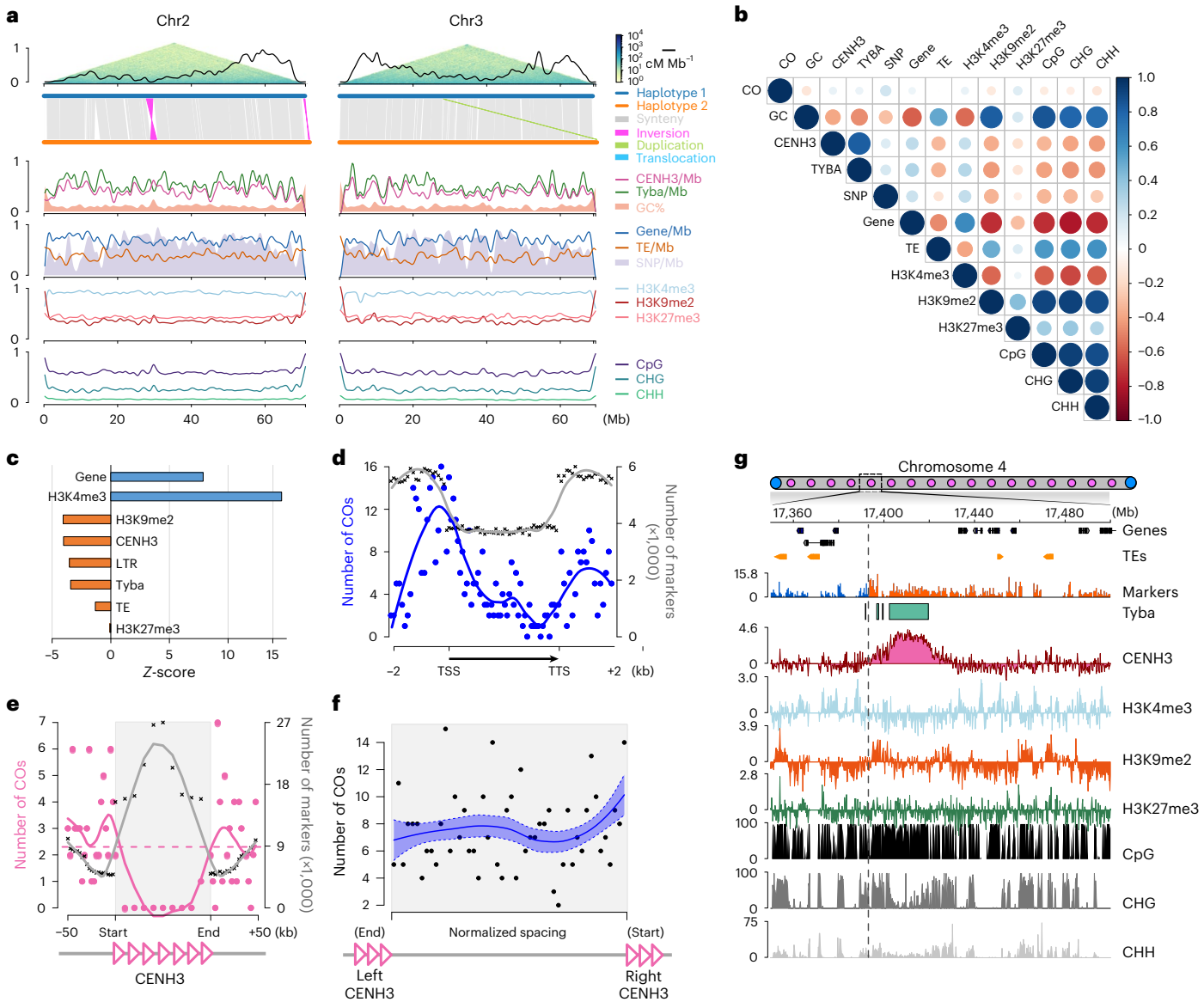


Fig. 4 | Broad- and fine-scale correlations between CO patterning and (epi) genomic features in *R. breviuscula*. **a**, Chromosome distribution of the CO rate coupled with different (epi)genetic features. (Upper) Recombination landscape (black line) created with COs detected in all single-pollen nuclei (n = 1,641), coupled with Omni-C data. Synteny analysis and detected structural variants between the two haplotypes. For the y axes, all features were scaled [0, 1], with 1 indicating a maximum of 2.34 for CO frequency (cM Mb^{-1}), 5 for Tyba array density, 6 for CENH3 domain density, 7,205 for SNP density, 88 for gene density and 227 for TE density. GC [33.3, 46.6], H3K4me3 [−1.494, 0.231], H3K9me2 [−1.20, 1.84] and H3K27me3 [−0.671, 0.491] are scaled to [0, 1] by their minima and maxima. Methylated CG (mCG), methylated CHG (mCHG) and CHH are original values (0–100%). All features were smoothed using a 1 Mbp sliding window and 250 kbp step size. **b**, Correlation matrix illustrating the correlation coefficients of 378 high-resolution COs detected in 63 selfed F_1 offspring with all available (epi)genomic features. Blue, positive correlations; red, negative correlations. Colour intensity and the size of the circle are proportional to the correlation coefficients. Pearson correlation coefficients for each pair of all features under a 1 Mb smoothing window and 250 kb step size. **c**, Z-score of the overlapped CO numbers with different (epi)genetic features. **d**, CO frequency (blue line) at genic regions. TSS, transcription start site; TTS, transcription termination site; grey line, marker density. **e**, Fine-scale CO frequency at CENH3 domains. The dashed horizontal line is the genome-wide mean CO count within a 2.5 kb interval. **f**, Random distribution of the relative distance of CO positions to the end of the left and to the start of the right CENH3 domain. The solid blue line was predicted by local polynomial regression fitting (loess function from R) using data from 378 COs from our F_1 offspring. The dashed blue band shows a range of one standard error above and below the fitted line. Pink-bordered triangles schematically represent CENH3 domains. **g**, Magnified view of one of the five COs placed closed to a region containing CENH3-positive chromatin and Tyba repeats. The CO resolution in this case is 200 bp. CO is indicated by the grey dashed line showing the haplotype switch (blue to orange) in the marker density track.

at high resolution (median 334 bp, mean ~2 kb) from our selfing offspring experiment.

The holocentromeres in *R. breviuscula* are repeat-based; that is, each centromeric unit is based on a specific array of the holocentromeric repeat Tyba associated with CENH3, with average sizes of ~20 kb and average spacings of ~400 kb, where each chromosome harbours hundreds of individual centromeric units (Extended Data Fig. 5b,c). Remarkably, we found the same epigenetic centromere identity in *R. breviuscula* (Extended Data Fig. 5d) as reported for *R. pubera*[32]. This organization makes it possible to identify centromeric units at the DNA level by annotating Tyba repeat arrays (Extended Data Fig. 5c). We then computed the observed versus expected by random

distribution fine-scale CO position across all available chromatin marks and genetic features. We found that COs are formed more frequently at H3K4me3 peaks and genes than expected by random distribution (Fig. 4c and Supplementary Fig. 11). Within genic regions, COs were preferentially formed at the promoter regions (Fig. 4d). Remarkably, COs were mostly suppressed inside the core of centromeric units and heterochromatic regions (Fig. 4c–e, Extended Data Fig. 5e and Supplementary Fig. 11). Detailed comparison of homologous Tyba arrays between the two haplotypes revealed a lack of extensive structural variations (Extended Data Fig. 7). Thus, CO suppression within centromeric units is likely caused by epigenetic features, where high levels of DNA methylation are typically found (Extended Data Fig. 5d). Furthermore, after computing the distances between the CO break intervals and the corresponding nearest CENH3 domains/Tyba arrays, the COs did not show a tendency to be positioned away from or close to centromeric units (Fig. 4e,f and Extended Data Fig. 5e,f), indicating that proximity to a centromeric unit does not affect CO formation. Moreover, we found five cases of a CO being placed in a region containing reduced CENH3-positive chromatin and Tyba repeats (Fig. 4g), confirming the absence of a centromere effect in its proximity. Our results point to the exciting finding that local CO formation in *R. breviuscula* is abolished at repeat-based centromeric units but enriched at genic promoter regions, supporting the role of chromatin features at a fine scale in contrast to the absence of correlation at a broad scale.

Telomere-led spatiotemporal dynamics of chromosome synapsis

We hypothesized that pairing and synapsis contribute to the distally biased recombination landscape observed in *R. breviuscula*. To investigate this, we performed immunolocalization on meiocytes with antibodies against ZYP1, ASY1 and HEI10, and fluorescence in situ hybridization (FISH) for telomeres. Signals detected for ZYP1, ASY1 and telomere probes indicated a tendency for telomeric signals to cluster together in one location, forming a typical bouquet configuration[59–61]. We found that ZYP1 loading likely started preferentially from the bouquet (Fig. 5 and Supplementary Fig. 12). The linear signal of ASY1 was still present and represented unpaired chromosomes on the opposite cell side to the bouquet (Fig. 5a and Supplementary Fig. 13). Next, we asked whether HEI10 loading also shows telomere-led dynamics. Indeed, we could determine that the first synapsed regions (ZYP1-stained) were also first loaded with HEI10 in the proximity of chromosome ends (Fig. 5b,c). These early ZYP1 signals at sub-telomeric regions showed positive and non-random co-localization with early loaded HEI10 signals, but this correlation decreased as HEI10 signals became less linear and more dot-like throughout pachytene (Fig. 5c,d and Supplementary Fig. 14). We consistently observed a few telomeres that did not participate in the meiotic bouquet ($n = 44$). These signals represent the terminal ends of chr1 and chr2 that harbour the *35S rDNA* loci and show low recombination rates; instead, these chromosome ends localized in the nucleolus (Fig. 3a,b and Extended Data Fig. 8). Remarkably, the nucleolus-positioned telomeres showed delayed ZYP1 loading compared with the telomeres involved in the bouquet (Fig. 5b–d and Supplementary Fig. 13). Indeed, there was an average of four telomeric signals, consistent with two unsynapsed chromosome ends, whereas at the bouquet there was an average of eight telomeric signals, consistent with eight synapsed chromosome ends ($n = 44$). Thus, our results are compatible with telomere-led HEI10 loading on early synapsed chromosome ends, which seems to better explain the broad-scale recombination landscape in *R. breviuscula* rather than centromeric or (epi)genetic effects.

Discussion

Deciphering the mechanisms controlling CO formation and distribution is key to understanding one of the main driving forces of genetic diversity in eukaryotes: meiotic recombination. Using *R. breviuscula*, a holocentric organism lacking both localized centromeres and compartmentalized chromosome organization, we show features that can potentially mask the factors underlying CO patterning and reveal important insights into CO control mechanisms. By showing that the CO distal bias observed in *R. breviuscula* is achieved even in the absence of compartmentalized chromosomal features, we propose that telomere-led (initiation at distal and sub-telomeric regions compared with more central regions) pairing and synapsis alone can impose bias in CO distribution. Indeed, the observed bouquet configuration has been shown to have a major role in synapsis and DNA DSB initiatiation[60–64]. Such telomere-led mechanisms have already been proposed to shift the concentration of COs towards the chromosome ends (Haenel et al.[22] and references therein). The clustering of telomeres during early prophase I has been proposed to be responsible for telomere-led recombination in *Arabidopsis thaliana asy1* mutants and wheat[37,64]. However, we cannot exclude that other factors, such as the density of DSBs along chromosomes, might contribute to the observed distribution of COs[65]. Although our observations are compatible with synapsis elongating from the telomeres, we cannot infer the directionality of HEI10 loading. Moreover, a recent study has shown that the bouquet configuration is required for the distal bias of DSB formation and repair, possibly playing an important role in the observed reduced recombination rates at interstitial positions[66]. Future experiments in *Rhynchospora* will be important to identify conserved and adapted mechanisms of spatiotemporal dynamics of meiotic DSB formation and HEI10 loading.

Considering the conservation of bouquet formation and synapsis progression in *R. breviuscula*, we propose that telomere-led synapsis and HEI10 loading are the driving force that shapes the observed distal bias of COs, independent of any centromere effect. Recently, a 'coarsening' model has been proposed to describe CO interference, CO assurance and CO patterning[52,53]. This model is based on the ability of HEI10 to aggregate and diffuse along synapsed chromosomes. Interestingly, in this model, enhanced loading of HEI10 at the chromosome ends leads to increased COs. Because the amount of loaded HEI10 accounts for the increased coarsening over time, early loading at the chromosome distal regions would accelerate the maturation of recombination intermediates compared with the interstitial regions. Thus, the coarsening model, proposed as a conserved mechanism among eukaryotes, fits our observations. It also explains the observed reduction in COs at telomeric ends in our genetic analyses. Although the presence of unidentified HEI10 interactors contributing to a bias in its loading at specific regions cannot be discarded. Moreover, we found low recombination frequencies at the 35S rDNA-harbouring ends of chr1 and chr2. We observed that these telomeres do not participate in bouquet formation and are consequently subject to late synapsis and HEI10 loading. In *A. thaliana* rDNA localizes in the nucleolus, where it is shielded from synapsis and meiotic recombination[67,68]. Similarly, we propose that in *R. breviuscula*, chromosome ends harbouring 35S rDNA are involved late in synapsis, resulting in late HEI10 loading, delayed coarsening and finally lower recombination frequencies (Fig. 6).

In the new era of accurate long-read genomics, haplotype-phased genomes are routinely available. By applying high-throughput scRNA-seq to individual pollen nuclei, we provide a powerful pipeline to investigate CO frequencies in the gametes of any heterozygous individual with an available phased genome. Using haplotype-specific markers, we detected and mapped CO events from thousands of gametes in a species with repeat-based holocentromeres. Although the found distal bias of CO is similar to that observed in numerous eukaryotes, including the holocentric *C. elegans*[21,22,69], the lack of correlation with the even distribution of (epi)genetic features (this study and Hofstatter et al.[32]) in *R. breviuscula* is remarkable. In *C. elegans* distinct chromosome domains ('centre' and 'arms') are characterized by differential gene density, repeats and histone modifications[21], and the unequal CO distribution corresponds to its chromosome domains—recombination rates are lower at the centre than at the arms[70]. Furthermore, a recent

https://doi.org/10.1038/s41477-024-01625-y

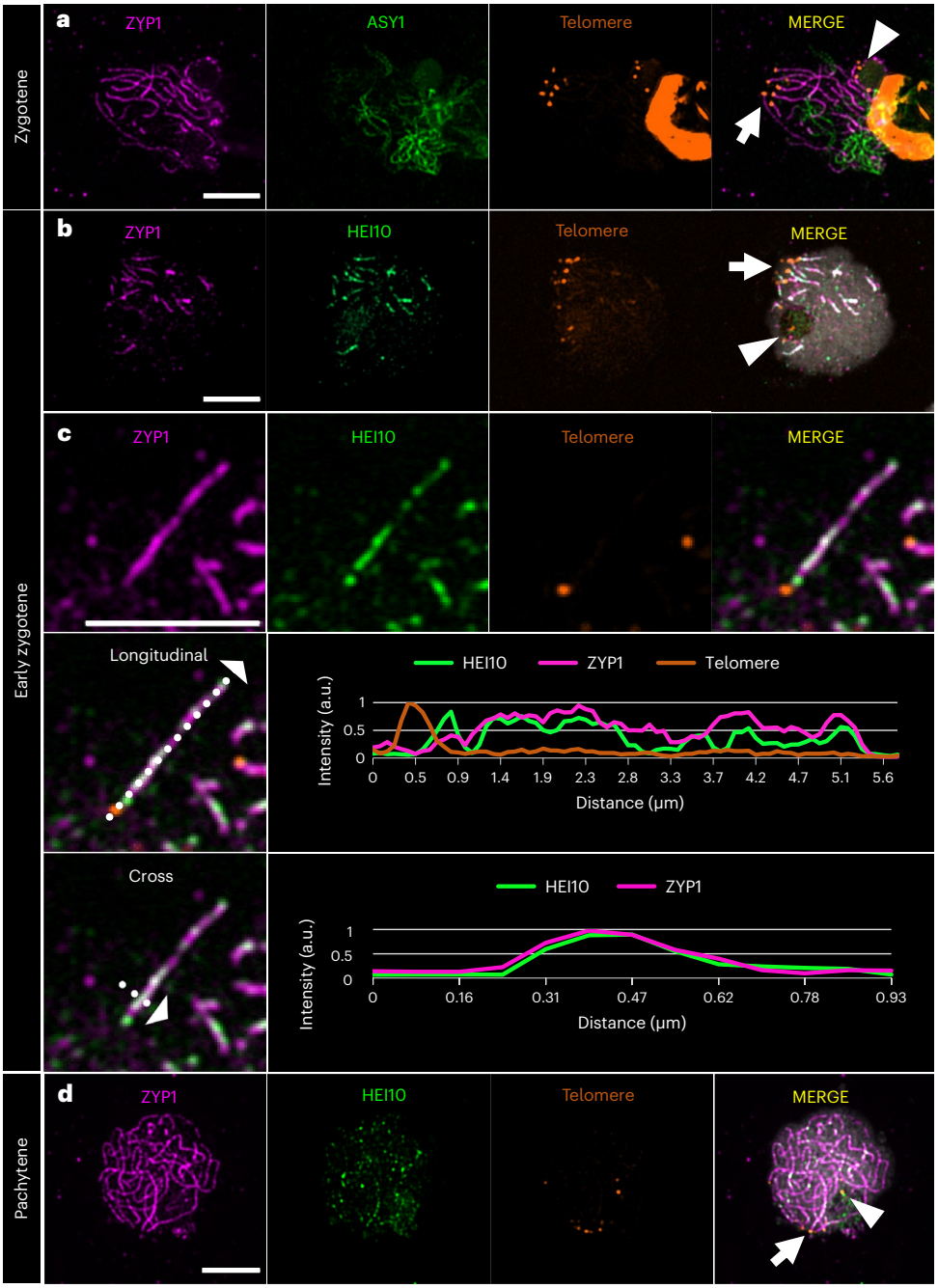


Fig. 5 | Telomere-led dynamics of synapsis formation and HEI10 loading. **a**, Telomeres cluster in a bouquet on one side of the cell, from where ZYP1 is elongating as the SC is being assembled. ASY1 represents unpaired chromosomes not yet reached by ZYP1. The behaviour of ASY1 + ZYP1 + telomeres was consistent in all immuno-FISH cells ($n = 8$) in three independent experiments. **b**, As ZYP1 lines elongate from the bouquet, HEI10 is quickly loaded onto synapsed chromosomes, whereas some telomeres localize to the nucleolus and lack the ZYP1 and HEI10 signals. **c**, Detail of early synapsis initiation and intensity profile of HEI10, ZYP1 and telomere: as soon as the SC (ZYP1) is assembled from telomeres, HEI10 is loaded and shows a high co-localization profile with ZYP1. a.u., arbitrary units. **d**, In late pachytene, ZYP1 occupies the whole chromosomal length, HEI10 signals become more dot-like and telomeres remain clustered in the bouquet, whereas few telomeres remain at the nucleolus. The behaviour of HEI10 + ZYP1 + telomeres was consistent in all immuno-FISH cells ($n = 44$) in three independent experiments. Telomeres at the bouquet (white arrow) and at the nucleolus (white arrowhead) are shown. Scale bar, 5 µm.

study showed that the megabase-scale CO landscape in *A. thaliana* is mostly explained by association with (epi)genetic marks, with COs forming inside open chromatin[10], and nucleotide polymorphisms only affecting COs locally[10,71]. By contrast, we could only find correlation of CO positioning with holocentromeres and (epi)genetic features at a fine scale, where COs preferentially formed within gene promoters rather than gene bodies, TEs and centromeres (Fig. 6b). This result appears to hold true for several eukaryotes and may be related to open chromatin states[9,10,72], suggesting that fine-scale CO regulation is associated with similar (epi)genetic factors independent of the chromosome organization. By contrast, the absence of a centromere effect in *R. breviuscula*, which seems to suppress CO formation only inside centromeric units but not in their vicinity (Fig. 6b), is probably due to the closed chromatin state of centromeric chromatin in *R. breviuscula*, as marked by high DNA methylation levels. Our findings suggest that the pericentromeric inhibition of COs observed in many monocentric eukaryotes[12,22] is

https://doi.org/10.1038/s41477-024-01625-y

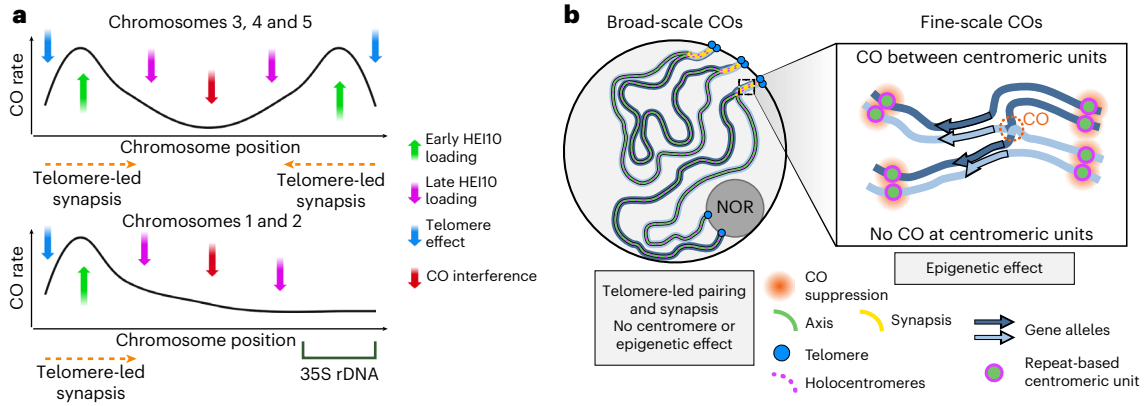


Fig. 6 | Model for CO regulation at a broad and fine scale in the holocentric plant *R. breviuscula*. **a**, Model for the role of telomere-led synapsis and HEI10 loading in shaping the broad-scale CO landscape in *R. breviuscula*. Whereas chr3, chr4 and chr5 have CO distal bias in both ends, chr1 and chr2 show bias at only one end, which is opposite to the localization of *35S rDNA* loci. **b**, Model for CO formation at a broad (left) and fine (right) scale. Telomere-led synapsis leads to the early loading of HEI10 at chromosomal ends that can potentially favour COs at distal regions, whereas 35S rDNA-harbouring chromosome ends do not show early synapsis and thus have less probability of making COs. At the local scale, COs are suppressed at core centromeric units, but not in their vicinity, where COs can be placed anywhere between two centromeric units. Remarkably, COs were preferentially placed at gene promoter regions.

likely a secondary effect of heterochromatin accumulation along large monocentric pericentromeres, and not a direct effect of centromeric chromatin. Understanding the molecular mechanisms of CO control in holocentric organisms will potentially unveil new strategies to address meiotic recombination within centromere proximal regions of monocentric chromosomes that rarely recombine.

Methods

Plant material

The same specimen of *Rhynchospora breviuscula* previously sequenced[32] was propagated clonally and cultivated in a greenhouse at the Max Planck Institute for Plant Breeding Research in Cologne, Germany. Flower buds were used for screening meiotic cells and pollen was collected for single-gamete sequencing.

Isolation of pollen nuclei, 10x Genomics scRNA-seq library preparation and sequencing

Protocols were adapted from Campoy et al.[58]. Briefly, to release pollen grains, anthers from fully developed flowers of *R. breviuscula* and *R. tenuis* (for multiplex purposes) were harvested and submerged in woody pollen buffer[73]. The nuclei were extracted using a modified bursting method. The solution containing the pollen grains was pre-filtered with a 100 µm strainer, and the pollen was crushed on a 30 µm strainer (Celltrics). The isolated nuclei were gathered in woody pollen buffer and stained with 4,6-diamidino-2-phenylindole (DAPI; 1 µg ml^{-1}) before being sorted using a BD FACSAria Fusion sorter with a 70 µm nozzle and 483 kPa sheath pressure. A total of 10,000 nuclei were sorted into 23 µl of sheath fluid solution and loaded into a 10x Chromium controller, according to the manufacturer's instructions. A library was created according to the chromium single-cell 3′ protocol. A CG000183 https://www.ncbi.nlm.nih.gov/nuccore/CG000183 Rev A kit from 10x Genomics was used for library preparation. The library was sequenced (100 Gb) on an Illumina NovaSeq instrument in 150 bp paired-end mode.

Whole-genome sequencing of F_1 recombinant offspring

To obtain a recombinant population of *R. breviuscula* plants, young inflorescences of the heterozygous reference *R. breviuscula* were bagged to force self-pollination. Because of its high self-incompatibility, only 63 seeds were obtained and germinated in soil to obtain F_1 plants. All 63 F_1 plants were sequenced to 3× coverage (~2 Gb) using an Illumina NextSeq2000 instrument in 150 bp paired-end mode.

Anther fixation and immunocytochemistry

Immunostaining was performed as described by Marques et al.[31] with some modifications. Anthers of *R. breviuscula* were harvested and fixed in ice-cold 4% (w/v) paraformaldehyde in phosphate-buffered saline (PBS; pH 7.5, 1.3 M NaCl, 70 mM Na_2HPO_4, 30 mM NaH_2PO_4) for 90 min. The anthers were separated according to size and dissected to release the meiocytes onto glass slides. The meiocytes were squashed with a coverslip that was later removed using liquid nitrogen. Meiocytes were then blocked with 1 h incubation in 3% (w/v) bovine serum albumin in PBS + 0.1% (v/v) Triton X-100 at 37 °C. The antibodies used were anti-AtASY1 raised in rabbits (inventory code PAK006)[36], anti-AtMLH1 raised in rabbits (PAK017)[74] and anti-*Rhynchospora*CENH3 raised in rabbits[30]. The anti-ZYP1 was raised in chickens against the peptide EGSLNPYADDPYAFD of the C-terminal end of AtZYP1a/b (gene ID: At1g22260/At1g22275) and affinity-purified (Eurogentec) (PAK048). Remarkably, this peptide region showed 100% similarity with the *Rhynchospora* ZYP1 C terminus (gene ID: RBREV_HAP1.r01.Chr2_h1G00222020.1). The anti-REC8 was a combination of two antibodies raised in rabbits against the *Rhynchospora*-specific REC8-peptides C-EEPYGEIQISKGPNM and C-YNPDDSVERMRDDPG (gene ID: RBREV_HAP1.r01.Chr4_h1G00395720.1) and affinity-purified (Eurogentec). The anti-HEI10 was a combination of two antibodies raised in rabbits against the *Rhynchospora*-specific HEI10-peptides C-NRPNQSRARTNMFQL and C-PVRQRNNKSMVSGGP (gene ID: RBREV_HAP1.r01.Chr4_h1G00387160.1) and affinity-purified (Eurogentec). Each primary antibody was diluted 1:200 in blocking solution. The slide-mounted samples were incubated with the primary antibodies overnight at 4 °C, after which they were washed three times for 10 min with PBS + 0.1% (v/v) Triton X-100. The slides were incubated with the secondary antibodies for 2 h at room temperature. The secondary antibodies were conjugated with Abberior STAR ORANGE (dilution 1:250; goat anti-rabbit immunoglobulin (Ig)G, catalogue no. STORANGE-1002-500UG, or goat anti-chicken IgY, catalogue no. STORANGE-1005-500UG) or Abberior STAR RED (dilution 1:250, goat anti-rabbit IgG, catalogue no. STRED-1002-500UG or goat anti-chicken IgY, catalogue no. STRED-1005-500UG) before being washed again three times for 10 min with PBS + 0.1% (v/v) Triton X-100 and allowed to dry. The samples were prepared with 10 µl of mounting solution (Vectashield + 0.2 µg of DAPI), covered with a coverslip and sealed with nail polish for storage. Images were taken with a Zeiss Axio Imager Z2 with Apotome system for optical sectioning or with a Leica Microsystems Thunder Imager dMi8 with computational clearing. The images

were deconvolved and processed with ZEN 3.2 or LAS X software. Co-localization analysis of ZYP1 and HEI10 signals was performed using the co-localization function of ZEN 3.2 software (Zeiss) and auto-thresholding was done using the Costes function[75,76].

Sequential immunostaining and fluorescence in situ hybridization

Sequential immunostaining and fluorescence in situ hybridization (immuno-FISH) was performed following Baez et al.[77]. The best slides obtained from immunostaining, as described above, were selected for FISH using a telomeric probe. Slides were washed with 1× PBS for 15 min, postfixed in 4% (w/v) paraformaldehyde in PBS for 10 min, dried with 70% (v/v) and 100% ethanol for 5 min each and probed with direct-labelled telomeric sequence (Cy3-[TTTAGGG]$_5$; MilliporeSigma). The hybridization mixture contained formamide (50% w/v), dextran sulfate (10%, w/v), 2× saline-sodium citrate (SSC) buffer and 50 ng μl^{-1} of telomeric probe. The slides were denatured at 75 °C for 5 min. Stringency washes were performed following Braz et al.[78] to give a final stringency of approximately 72%. The slides were counterstained with 10 μl of mounting solution (Vectashield + 0.2 μg of DAPI), and images were captured as described above.

Mitotic and meiotic chromosome spreads were performed as described by Ruban et al.[79], with some modifications. Briefly, tissue samples were fixed in a 3:1 ethanol/acetic acid (v/v) solution for 2 h with gentle shaking. The samples were washed with water twice for 5 min and treated with an enzyme mixture (0.7% w/v cellulase R10, 0.7% w/v cellulase R10, 1.0% w/v pectolyase and 1.0% w/v cytohelicase in citric buffer) for 30 min at 37 °C. The samples were dissected on slides under a binocular microscope in acetic acid (60%). The pTa71 probe (18S–5.8S–26S ribosomal RNA) from *Triticum aestivum*[80] was labelled with Alexa-448 by nick translation for 35S rDNA localization. FISH with telomeric and 35S rDNA probes was performed as described above.

Haplotype phasing and scaffolding

A phased chromosome-level genome of *R. breviuscula* was assembled using PacBio HiFi and Hi-C data available from Hofstatter et al.[32] under the NCBI BioProject no. PRJNA784789. First, a phased primary assembly was obtained by running hifiasm[81] using as inputs the 30 Gb of PacBio HiFi reads (~35× coverage per haplotype) in combination with Dovetail Omni-C reads, using the following command: hifiasm -o Rbrevi.phased.asm.hic --h1 hic.R1.fastq.gz --h2 hic.R2.fastq.gz hifi.reads.fastq.gz. The phased assemblies of each individual haplotype were further scaffolded to chromosome scale using Salsa2 (ref. 82), followed by successive rounds of manual curation and rescaffolding. The genome sizes of haplotypes 1 and 2 were 418,624,405 and 390,890,712 bp, respectively. Both haplotypes comprise five chromosomes with a length of ~370 Mb in total, as well as other unplaced sequences (Supplementary Table 1).

Hi-C map generation and haplotype comparison

The Hi-C heatmap (Extended Data Fig. 3c) was generated with juicer (v.1.6) by aligning Omni-C reads that were used for genome assembly to the phased *R. breviuscula* genome. The Hi-C triangles for each chromosome in Fig. 4a and Extended Data Fig. 5a were plotted using fancplot (v.0.9.1) with 500,000bp resolution and Knight-Ruiz (KR) normalization. Synteny blocks and structural rearrangements between two haplotypes (Extended Data Fig. 3d) were computed by SyRi (v.1.5.3) and plotsr[83,84] after aligning two haplotypes by minimap2 (v.2.20).

Definition of allelic SNPs as genotyping markers on the phased reference genome

To define genotyping markers for *R. breviuscula*, all available (NCBI BioProject no. PRJNA784789) raw Illumina HiSeq 3000 150 bp paired-end reads (25,899,503,075 bases, ~54× coverage) were first mapped to the five pseudochromosome scaffolds in haplotype 1 of the phased reference genome using bowtie2 (v.2.4.4)[85]. The alignment file was further sorted with SAMtools (v.1.9)[86]. The alignments of short reads to the reference genome were used for SNP calling by 'bcftools mpileup' and 'bcftools call' (v.1.9)[86] (with the –keep-alts, --variants-only and --multiallelic-caller flags enabled). A total of 1,404,927 SNPs excluding insertions/deletions were derived. To distinguish the two haplotypes using these SNPs, only allelic SNPs were selected as markers for genotyping; therefore, variant information was collected, including mapping quality, alternative base coverage and allele frequency resulting from SHOREmap conversion (v.3.6)[87], which converts SNP files (.vcf) into a read-friendly, tab-delimited text file. A final set of 820,601 alleles fulfilling certain thresholds (mapping quality >50; 5 ≤ alternative base coverage ≤ 30, 0.4 ≤ allele frequency ≤ 0.6) was selected as markers (Fig. 2b and Supplementary Fig. 6).

Preprocessing scRNA-seq data from pollen nuclei

Raw scRNA-seq data usually include barcode errors and contaminants such as doublets and ambient RNA. In the current study, cell barcodes (CBs) were first corrected in these data using 'bcctools correct' (v.0.0.1) based on the 10x v3 library complete barcode list with options '--alts 16 --spacer 12' because of the 16 bp CB and 12 bp unique molecular identifier (UMI). After correction, 952,535 viable CBs were detected. This step also truncated the CBs and UMIs from every pair of scRNA-seq reads. After counting the occurrence of CBs, the number of read pairs under each CB was determined. To ensure a sufficient number of reads for SNP calling, only CBs appearing more than 5,000 times were used for the subsequent analyses. Finally, each CB was seen as one viable cell, and reads corresponding to the CB were assigned to this cell (demultiplexing). A total of 8,001 viable cells were ultimately identified, with 365,771,748 (77.25% of all raw scRNA-seq) read pairs included. We also input the scRNA-seq data to the 10x standard analysis pipeline cellranger (v.7.1.0) to check the statistics. The clustering analysis and the gene number in each cell (Supplementary Fig. 5) were based on cellranger count results.

Alignments of single-pollen RNA sequences to genome and deduplication

To identify genotyping markers in the *R. breviuscula* gametes, scRNA reads of the pollen nuclei were first mapped to the haplotype 1 chromosomes (Fig. 2b) using hisat2 (v.2.1.0)[88]. Specifically, each cell-specific pair of reads was merged as one single-end FASTQ file, and hisat2 was run under single-end mode (-U) because the SNP calling approach does not detect SNPs on reads whose mated reads are not mapped. Before further analyses of the alignment results, UMIs were previously extracted from the read alongside the CBs; hence, a fast UMI deduplication tool, UMIcollapse[89], was employed to remove the PCR duplicates by collapsing reads with the same UMIs.

The sequencing library was prepared for mixed pollen nuclei of *R. breviuscula* and *R. tenuis* to enable multiple-potential analyses. The addition of gametes from *R. tenuis* was done for multiplexing purposes, and they are analysed in another study. To discriminate the single-cell data between the two species, we used a straightforward approach without gene expression profiling. For each cell: (1) the DNA sequences were mapped to both the *R. breviuscula* and *R. tenuis* chromosomal genomes; and (2) the alignment rates between the two species were compared to decide the cell identity (Supplementary Fig. 7). The alignment rates to *R. breviuscula* and *R. tenuis* were both bimodal distributions (Supplementary Fig. 7a,b); therefore, these cells can be grouped based solely on their mapping rates. It was estimated that 4,733 cells were from *R. breviuscula* and 2,709 cells were from *R. tenuis* (Supplementary Fig. 7c) based on the alignment fractions. The remaining 559 cells presented very similar alignment rates, which were potential doublets. Among the 4,733 *R. breviuscula* cells, those whose alignment rates were lower than 25% were discarded,

leaving 4,392 cells from *R. breviuscula* available for the next stage of the analysis.

SNP calling and selection of markers in gametes

SNP calling in all gametes adopted the same methods as the reference genome SNP calling; for example, via 'bcftools mpileup' and 'bcftools call' (v.1.9), with the difference that the '--variants-only' flag was not applied. After acquiring SNPs for every gamete, the SNP positions, allele counts of the reference and alternative bases were extracted through the 'bcftools query'. Comparing SNPs in every gamete with markers defined on the reference resulted in reliable genotyping markers in this gamete.

A total of 2,338 cells with fewer than 400 markers were first discarded to ensure accurate genotyping by sufficient markers. To remove doublets, the frequency of marker genotype switches across the remaining 2,054 cells was estimated. Cells with frequent switches—that is, a switching rate (genotype switching times/number of markers) greater than 0.07—were taken as doublets (Supplementary Fig. 7e). Ultimately, 402 doublets were identified, with the remaining 1,652 cells proving suitable for subsequent CO calling.

CO identification

Chromosome genotyping was performed by adapting the haplotype phasing method proposed by Campoy et al. The original approach was designed based on a scDNA-seq library that is commonly able to examine more SNPs than scRNA-seq data. Therefore, the smoothing function and parameters were adjusted accordingly to genotype genomic blocks with relatively sparse markers. Specifically, the markers were first smoothed using the allele frequencies of neighbouring markers (two ahead and two behind), and then smoothed by the genotypes of surrounding markers. After smoothing, the genomic segments harbouring the markers in the same genotype are merged to genotype blocks, and those containing at least five markers within 1 Mb length were qualified to be assigned a final genotype. The genomic regions that saw the conversion of the genotypes at the flanks were taken as CO break positions (Fig. 2c,d). Finally, the CO numbers in each cell were counted, and those with double COs were manually assessed and corrected.

Recombination landscape and CO interference from single-gamete sequencing

To gain an overview of the CO rates across the chromosomes of *R. breviuscula*, the CO positions in all viable cells (1,641 cells remaining after manual correction) were summarized, and the recombination landscape for each chromosome was plotted (Fig. 3a). The recombination rate (cM Mb^{-1}) was computed using a 1 Mb sliding window and 100 kb step size.

CO interference was analysed with MADpattern (v.1.1)[90], using 1,641 confident singleton pollen nuclei. Chromosome 1 was divided into 18 intervals and chr2–chr5 were divided into 15 intervals to compute the mean CoC of every pair of intervals.

F_1 offspring mapping and CO analysis

Sixty-three F_1 offspring were reproduced from selfed *R. breviuscula*. Each F_1 plant was sequenced with ~3× Illumina whole-genome sequencing data. To genotype F_1 offspring, whole-genome sequencing Illumina sequences of each plant were first mapped to the rhyBreHap1 reference genome using bowtie2 (v.2.4.4) paired-end mode, and then SNPs were called by 'bcftools mpileup' and 'bcftools call' (v1.9) (with --keep-alts, --variants-only and --multiallelic-caller flags enabled). Next, SNPs of each F_1 sample were input to TIGER[91] for genotyping and to generate potential CO positions. In addition, RTIGER[92] was used to identify the genotypes of chromosomal segments by utilizing the corrected markers that resulted from TIGER. Only COs that agreed using both tools were kept. The recombination landscape from F_1 COs (Supplementary Fig. 4) was plotted using the same strategy and sliding window, as illustrated for pollen nuclei.

Genetic linkage maps

To plot the genetic linkage maps (Fig. 3b and Extended Data Fig. 4a), 743 markers were extracted from the 820,601 reference markers by selecting the median marker within each 500 kb sliding window (step size was also 500 kb) from the first present marker to the last. The linkage map was then plotted using R package LinkageMapView. Marey maps (Fig. 3c and Extended Data Fig. 4b) were plotted using genetic length, computed based on pollen nuclei with more than 2,000 markers ($n = 81$) and F_1 offspring ($n = 63$), against physical length for each chromosome, respectively.

Chromatin immunoprecipitation

CENH3 chromatin immunoprecipitation sequencing (ChIP–seq) data were obtained from Hofstatter et al.[32]. Further chromatin immunoprecipitation (ChIP) experiments were performed for H3K4me3 (rabbit polyclonal to histone H3 tri-methyl K4; Abcam, catalogue no. ab8580, dilution 1:300), H3K9me2 (mouse monoclonal to histone H3 di-methyl K9; Abcam, catalogue no. ab1220, clone no. mAbcam 1220, dilution 1:200), H3K27me3 (mouse monoclonal to histone H3 tri-methyl K27; Abcam, catalogue no. ab6002, clone no. mAbcam 6002, dilution 1:200) and the IgG control (recombinant rabbit IgG, monoclonal; Abcam, catalogue no. ab172730, clone no. EPR25A, dilution 1:300) using the protocol described by Hofstatter et al.[32].

ChIP–seq and analysis

ChIP DNA was quality-controlled using the next-generation sequencing assay on a FEMTO pulse (Agilent Technologies). An Illumina-compatible library was prepared with the Ovation Ultralow V2 DNA-Seq library preparation kit (Tecan Genomics) and sequenced as single-end 150 bp reads on a NextSeq2000 (Illumina) instrument. An average of 20 million reads were obtained for each library.

Raw sequencing reads were trimmed using Cutadapt[93] to remove low-quality nucleotides (with a quality score <30) and the adaptors. Trimmed ChIPed 150 bp single-end reads were mapped to their respective reference genome using bowtie2 (ref. 85) with default parameters. All read duplicates were removed and only the single best-matching read was kept on the final alignment Binary Alignment Map (BAM) file. The BAM files were converted to BIGWIG coverage tracks using the bamCompare tool from deeptools[94]. Coverage was calculated as the number of reads per 50 bp bin and normalized as reads per kilobase per million mapped reads (RPKM). Magnified chromosome regions showing multiple tracks presented in Fig. 4g were plotted with pyGenomeTracks[95].

Tyba array and CENH3 domain annotation

Tyba repeats were annotated using a BLAST search with a consensus Tyba sequence, allowing a minimum of 70% similarity. Further annotation of the Tyba arrays was performed by removing spurious low-quality Tyba monomer annotations shorter than 500 bp. bedtools[96] was used to merge all adjacent Tyba monomers situated at a maximum distance of 25 kb into individual annotations to eliminate the gaps that arise because of fragmented Tyba arrays, and those smaller than 2 kb were discarded.

CENH3 peaks were called with MACS3 (ref. 97) using the broad peak calling mode:macs3 callpeak -t ChIP.bam -c Control.bam --broad -g 380000000 --broad-cutoff 0.1

The identified peaks were further merged using a stepwise progressive merging approach. CENH3 domains were generated by: (1) merging CENH3 peaks with a spacing distance less than 25 kb using bedtools to eliminate the gaps that arise because of fragmented Tyba arrays or the insertion of TEs; and (2) removing CENH3 domains less than 1 kb in size.

Transposable element annotation

TE protein domains and complete long terminal repeat retrotransposons were annotated in the reference haplotype genome using the REXdb database (Viridiplantae_version_3.0)[98] and the DANTE tool available from the RepeatExplorer2 Galaxy portal[99].

Enzymatic methyl sequencing and analysis

To investigate the methylome space in *R. breviuscula*, the relatively non-destructive NEBNext Enzymatic Methyl-seq Kit was employed to prepare an Illumina-compatible library, followed by paired-end sequencing (2 × 150 bp) on a NextSeq2000 (Illumina) instrument. For each library, 10 Gb of reads was generated.

Enzymatic methyl sequencing data were analysed using the Bismarck pipeline[100] following the standard pipeline described at https://rawgit.com/FelixKrueger/Bismark/master/Docs/Bismark_User_Guide.html. Individual methylation context files for CpG, CHG and CHH were converted into BIGWIG format and used as input tracks for the overall genome-wide DNA methylation visualization with pyGenomeTracks and R plots.

Quantitative correlation of COs and (epi)genetic features

The correlation matrix (Fig. 4b) was calculated for all pairwise features by Pearson correlation coefficient using sliding window: specifically, mean CO rates, mean GC contents, CENH3 peak density, Tyba array density, SNP density, TE density, H3K4me3 RPKM, H3K9me2 RPKM, H3K27me3 RPKM, mean CpG, mean CHG and mean CHH.

To inspect a possible centromere effect on CO positioning, the relative distance from the CO site was calculated to the closest left and right centromeric unit (the CENH3 domain or Tyba array) across the 378 COs in the F_1 offspring and normalized all distances to 0–1 such that all neighbouring centromeric units were displayed at the same scale (Fig. 4f, Extended Data Fig. 5f and Supplementary Fig. 11e,f). CO and marker positions over the transcript bodies, CENH3 domain or Tyba array were normalized by their distance to start sites and end sites and then counted by binning (Fig. 4d,e, Extended Data Fig. 5e and Supplementary Fig. 11).

To see the association of CO designations with a variety of (epi) genetic features at a local scale, we first counted the number of COs that overlap with CENH3, Tyba arrays, genes, TEs, long terminal repeats, H3K4me3 peaks, H3K9me2 peaks and H3K27me3 peaks using 'bedtools intersect' (v.2.29.0). Next, we assigned 378 pseudo-COs genome-wide at random. The number of COs on each chromosome was the same as that detected by F_1 individuals (for example, 72 COs on chr1, 69 on chr2, 76 on chr3, 84 on chr4 and 77 on chr5), whereas the CO break gap length was picked up from the 378 real F_1 CO gaps randomly. For each simulation round, the pseudo-COs were overlapped with (epi)genetic features, again using 'bedtools intersect'. Five thousand of these simulations were done, and the results were then plotted as the distribution of overlapped CO numbers for each feature (Supplementary Fig. 11). Finally, to evaluate the deviation of real overlapped COs with each feature from the expected overlapped CO number under the hypothesis of randomly distributed COs, *Z*-scores were calculated using the mean values and standard deviations of the simulated number of overlapped CO distribution (Fig. 4c).

Gene annotation

Structural gene annotation was done by combining de novo gene calling and homology-based approaches with *Rhynchospora* RNA-seq, IsoSeq and protein datasets already available[32].

Using evidence derived from expression data, RNA-seq data were first mapped using STAR[101] (v.2.7.8a) and subsequently assembled into transcripts by StringTie[102] (v.2.1.5, parameters -m 150 -t -f 0.3). Triticeae protein sequences from available public datasets (UniProt, https://www.uniprot.org, 5 October 2016) were aligned against the genome sequence using GenomeThreader[103] (v.1.7.1; arguments -startcodon -finalstopcodon -species rice -gcmincoverage 70 -prseedlength 7 -prhdist 4). IsoSeq datasets were aligned to the genome assembly using GMAP[104] (v.2018-07-04). All assembled transcripts from RNA-seq, IsoSeq and aligned protein sequences were combined using Cuffcompare[105] (v.2.2.1) and subsequently merged with StringTie (v.2.1.5, parameters --merge -m150) into a pool of candidate transcripts. TransDecoder (v.5.5.0; http://transdecoder.github.io) was used to identify potential open reading frames and to predict protein sequences within the candidate transcript set.

Ab initio annotation was initially done using Augustus[106] (v.3.3.3). GeneMark[107] (v.4.35) was additionally employed to further improve structural gene annotation. To avoid potential over-prediction, we generated guiding hints using the above-described RNA-seq, protein and IsoSeq datasets as described by Nachtweide and Stanke[106] A specific Augustus model for *Rhynchospora* was built by generating a set of gene models with full support from RNA-seq and IsoSeq. Augustus was trained and optimized using the steps detailed by Nachtweide and Stanke[106].

All structural gene annotations were joined using EVidenceModeller[108] (v.1.1.1), and weights were adjusted according to the input source: ab initio (Augustus: 5, GeneMark: 2), homology-based (10). In addition, two rounds of PASA[109] (v.2.4.1) were run to identify untranslated regions and isoforms using the above-described IsoSeq datasets.

We used DIAMOND[110] (v.2.0.5) to compare potential protein sequences with a trusted set of reference proteins (UniProt Magnoliophyta, reviewed/Swiss-Prot, downloaded on 3 August 2016; https://www.uniprot.org). This differentiated candidates into complete and valid genes, non-coding transcripts, pseudogenes and TEs. In addition, we used PTREP (release 19; https://trep-db.uzh.ch). Furthermore, functional annotation of all predicted protein sequences was done using the AHRD pipeline (https://github.com/groupschoof/AHRD).

K_a/K_s ratio calculation

We identified homologues between *Brachypodium distachyon* (v.3.0) (downloaded from ensemble plants plants.ensembl.org) and *Juncus effesus*[32] using the ortholog module from JCVI python library[111]. Subsequently, pairwise alignments were generated with ParaAT[112] (v.2) and the K_a/K_s ratio was calculated with KaKs_Calculator[113] (v.3) using the YN method[114]. Plots were generated using karyoploteR[115].

Structural comparison of Tyba arrays between two haplotypes

The Shapiro test was applied to Tyba array numbers, which suggested a normal distribution of array number in both haplotypes. A two-sided *F*-test was then performed and proved the equality of their variances. Thus, we applied a two-sided *t*-test, which suggested no significant difference of means of Tyba array number in two haplotypes ($P = 0.9881$). The sizes of Tyba arrays on both haplotypes also showed no difference in median based on Mann–Whitney *U*-test ($P = 0.83$). Next, we compared the relative positions of Tyba arrays within the colinearly syntenic blocks between haplotypes; that is, the position of the midpoint of each Tyba array was scaled to the syntenic block in which it resides. A two-sample Kolmogorov–Smirnov test was performed on this relative Tyba array position in two haplotypes, which implied that both followed the same distribution ($D = 0.021694$, $P = 0.9886$). We ended up with 823 pairs when further linking the Tyba arrays between two haplotypes that reside at the closest relative syntenic positions. Linear regression demonstrated that the difference in the relative positions of those Tyba array pairs almost fit to a horizontal line of $y = 0$ indicating that the relative positions of paired Tyba arrays in syntenic counterparts are mostly the same despite a few outliers (Extended Data Fig. 7).

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

All sequencing data used in this study have been deposited at NCBI under the BioProject ID PRJNA1059790 and are publicly available as of the date of publication. The reference genomes, sequencing data, annotations and all tracks presented in this work are made available for download at DRYAD: https://datadryad.org/stash/share/EvB3PRN-Vph5IiTkOM3jTZddgmS45cJhQYq2v3LI5InE. The REXdb database Viridiplantae v.3.0 [http://repeatexplorer.org/?page_id=918] is publicly available. All other data needed to evaluate the conclusions in the paper are provided in the paper and/or the supplemental information.

Code availability

The original code for the construction of recombination maps from single-cell RNA sequencing is available at https://github.com/Raina-M/detectCO_by_scRNAseq. Any additional information required to re-analyse the data reported in this paper is available from the corresponding author upon request.

References

1. Keeney, S., Giroux, C. N. & Kleckner, N. Meiosis-specific DNA double-strand breaks are catalyzed by Spo11, a member of a widely conserved protein family. *Cell* **88**, 375–384 (1997).
2. Keeney, S. Spo11 and the formation of DNA double-strand breaks in meiosis. *Genome Dyn. Stab.* **2**, 81–123 (2008).
3. Allers, T. & Lichten, M. Differential timing and control of noncrossover and crossover recombination during meiosis. *Cell* **106**, 47–57 (2001).
4. Mercier, R., Mézard, C., Jenczewski, E., Macaisne, N. & Grelon, M. The molecular biology of meiosis in plants. *Annu. Rev. Plant Biol.* **66**, 297–327 (2015).
5. Lambing, C., Franklin, F. C. H. & Wang, C.-J. R. Understanding and manipulating meiotic recombination in plants. *Plant Physiol.* **173**, 1530–1542 (2017).
6. Thangavel, G., Hofstatter, P. G., Mercier, R. & Marques, A. Tracing the evolution of the plant meiotic molecular machinery. *Plant Reprod.* **36**, 73–95 (2023).
7. Grishaeva, T. M. & Bogdanov, Y. F. Conservation of meiosis-specific nuclear proteins in eukaryotes: a comparative approach. *Nucleus* **61**, 175–182 (2018).
8. Loidl, J. Conservation and variability of meiosis across the eukaryotes. *Annu. Rev. Genet.* **50**, 293–316 (2016).
9. Zelkowski, M., Olson, M. A., Wang, M. & Pawlowski, W. Diversity and determinants of meiotic recombination landscapes. *Trends Genet.* **35**, 359–370 (2019).
10. Lian, Q. et al. The megabase-scale crossover landscape is largely independent of sequence divergence. *Nat. Commun.* **13**, 3828 (2022).
11. Mézard, C., Tagliaro Jahns, M. & Grelon, M. Where to cross? New insights into the location of meiotic crossovers. *Trends Genet.* **31**, 393–401 (2015).
12. Brazier, T. & Glémin, S. Diversity and determinants of recombination landscapes in flowering plants. *PLoS Genet.* **18**, e1010141 (2022).
13. Lambie, E. J. & Roeder, G. S. Repression of meiotic crossing over by a centromere (CEN3) in *Saccharomyces cerevisiae*. *Genetics* **114**, 769–789 (1986).
14. Topp, C. N. & Dawe, R. K. Reinterpreting pericentromeric heterochromatin. *Curr. Opin. Plant Biol.* **9**, 647–653 (2006).
15. Melters, D. P. et al. Comparative analysis of tandem repeats from hundreds of species reveals unique insights into centromere evolution. *Genome Biol.* **14**, R10 (2013).
16. Talbert, P. B. & Henikoff, S. Centromeres convert but don't cross. *PLoS Biol.* **8**, e1000326 (2010).
17. Choulet, F. et al. Structural and functional partitioning of bread wheat chromosome 3B. *Science* **345**, 1249721 (2014).
18. Mascher, M. et al. A chromosome conformation capture ordered sequence of the barley genome. *Nature* **544**, 427–433 (2017).
19. Kianian, P. M. A. et al. High-resolution crossover mapping reveals similarities and differences of male and female recombination in maize. *Nat. Commun.* **9**, 2370 (2018).
20. Yelina, N., Diaz, P., Lambing, C. & Henderson, I. R. Epigenetic control of meiotic recombination in plants. *Sci. China Life Sci.* **58**, 223–231 (2015).
21. Saito, T. T. & Colaiácovo, M. P. Regulation of crossover frequency and distribution during meiotic recombination. *Cold Spring Harb. Symp. Quant. Biol.* **82**, 223–234 (2017).
22. Haenel, Q., Laurentino, T. G., Roesti, M. & Berner, D. Meta-analysis of chromosome-scale crossover rate variation in eukaryotes and its significance to evolutionary genomics. *Mol. Ecol.* **27**, 2477–2497 (2018).
23. Naish, M. et al. The genetic and epigenetic landscape of the *Arabidopsis* centromeres. *Science* **374**, eabi7489 (2021).
24. Kursel, L. E. & Malik, H. S. Centromeres. *Curr. Biol.* **26**, R487–R490 (2016).
25. Schubert, V. et al. Super-resolution microscopy reveals diversity of plant centromere architecture. *Int. J. Mol. Sci.* **21**, 3488 (2020).
26. Melters, D. P., Paliulis, L. V., Korf, I. F. & Chan, S. W. Holocentric chromosomes: convergent evolution, meiotic adaptations, and genomic analysis. *Chromosome Res.* **20**, 579–593 (2012).
27. Marcial, E., Ignacio Márquez-Corro, J. & Andrew, L. H. The phylogenetic origins and evolutionary history of holocentric chromosomes. *Syst. Bot.* **41**, 580–585 (2016).
28. Steiner, F. A. & Henikoff, S. Holocentromeres are dispersed point centromeres localized at transcription factor hotspots. *eLife* **3**, e02025 (2014).
29. Senaratne, A. P. et al. Formation of the CenH3-deficient holocentromere in lepidoptera avoids active chromatin. *Curr. Biol.* **31**, 173–181 (2021).
30. Marques, A. et al. Holocentromeres in *Rhynchospora* are associated with genome-wide centromere-specific repeat arrays interspersed among euchromatin. *Proc. Natl Acad. Sci. USA* **112**, 13633–13638 (2015).
31. Marques, A., Schubert, V., Houben, A. & Pedrosa-Harand, A. Restructuring of holocentric centromeres during meiosis in the plant *Rhynchospora pubera*. *Genetics* **204**, 555–568 (2016).
32. Hofstatter, P. G. et al. Repeat-based holocentromeres influence genome architecture and karyotype evolution. *Cell* **185**, 3153–3168 (2022).
33. Cabral, G., Marques, A., Schubert, V., Pedrosa-Harand, A. & Schlögelhofer, P. Chiasmatic and achiasmatic inverted meiosis of plants with holocentric chromosomes. *Nat. Commun.* **5**, 5070 (2014).
34. Heckmann, S., Schubert, V. & Houben, A. Holocentric plant meiosis: first sisters, then homologues. *Cell Cycle* **13**, 3623–3624 (2014).
35. Hofstatter, P. G., Thangavel, G., Castellani, M. & Marques, A. Meiosis progression and recombination in holocentric plants: what is known? *Front. Plant Sci.* **12**, 658296 (2021).
36. Armstrong, S. J., Caryl, A. P., Jones, G. H. & Franklin, F. C. Asy1, a protein required for meiotic chromosome synapsis, localizes to axis-associated chromatin in *Arabidopsis* and *Brassica*. *J. Cell Sci.* **115**, 3645–3655 (2002).
37. Lambing, C., Kuo, P. C., Tock, A. J., Topp, S. D. & Henderson, I. R. ASY1 acts as a dosage-dependent antagonist of telomere-led recombination and mediates crossover interference in *Arabidopsis*. *Proc. Natl Acad. Sci. USA* **117**, 13647–13658 (2020).
38. Higgins, J. D., Sanchez-Moran, E., Armstrong, S. J., Jones, G. H. & Franklin, F. C. The *Arabidopsis* synaptonemal complex protein ZYP1 is required for chromosome synapsis and normal fidelity of crossing over. *Genes Dev.* **19**, 2488–2500 (2005).

39. Wang, M. et al. The central element protein ZEP1 of the synaptonemal complex regulates the number of crossovers during meiosis in rice. *Plant Cell* **22**, 417–430 (2010).
40. Capilla-Pérez, L. et al. The synaptonemal complex imposes crossover interference and heterochiasmy in *Arabidopsis*. *Proc. Natl Acad. Sci. USA* **118**, e2023613118 (2021).
41. France, M. G. et al. ZYP1 is required for obligate cross-over formation and cross-over interference in *Arabidopsis*. *Proc. Natl Acad. Sci. USA* **118**, e2021671118 (2021).
42. Barakate, A. et al. The synaptonemal complex protein ZYP1 is required for imposition of meiotic crossovers in barley. *Plant Cell* **26**, 729–740 (2014).
43. Lambing, C. et al. Interacting genomic landscapes of REC8-Cohesin, Chromatin, and Meiotic Recombination in *Arabidopsis* [CC-BY]. *Plant Cell* **32**, 1218–1239 (2020).
44. Chelysheva, L. et al. The *Arabidopsis* HEI10 is a new ZMM protein related to Zip3. *PLoS Genet.* **8**, e1002799 (2012).
45. Ward, J. O. et al. Mutation in mouse Hei10, an E3 ubiquitin ligase, disrupts meiotic crossing over. *PLoS Genet.* **3**, e139 (2007).
46. De Muyt, A. et al. E3 ligase Hei10: a multifaceted structure-based signaling molecule with roles within and beyond meiosis. *Genes Dev.* **28**, 1111–1123 (2014).
47. Zhang, L., Köhler, S., Rillo-Bohn, R. & Dernburg, A. F. A compartmentalized signaling network mediates crossover control in meiosis. *eLife* **7**, e30789 (2018).
48. Wang, K. et al. The role of rice HEI10 in the formation of meiotic crossovers. *PLoS Genet.* **8**, e1002809 (2012).
49. Nguyen, H., Labella, S., Silva, N., Jantsch, V. & Zetka, M. *C. elegans* ZHP-4 is required at multiple distinct steps in the formation of crossovers and their transition to segregation competent chiasmata. *PLoS Genet.* **14**, e1007776 (2018).
50. Qiao, H. et al. Antagonistic roles of ubiquitin ligase HEI10 and SUMO ligase RNF212 regulate meiotic recombination. *Nat. Genet.* **46**, 194–199 (2014).
51. Stauffer, W., Zhang, L. & Dernburg, A. Diffusion through a liquid crystalline compartment regulates meiotic recombination. in *Proc. SPIE 10888, Biophysics, Biology and Biophotonics IV: The Crossroads* 1088809 (2019); https://doi.org/10.1117/12.2513378
52. Morgan, C. et al. Diffusion-mediated HEI10 coarsening can explain meiotic crossover positioning in *Arabidopsis*. *Nat. Commun.* **12**, 4674 (2021).
53. Zhang, L., Stauffer, W., Zwicker, D. & Dernburg, A. F. Crossover patterning through kinase-regulated condensation and coarsening of recombination nodules. Preprint at *bioRxiv* https://doi.org/10.1101/2021.08.26.457865 (2021).
54. Lhuissier, F. G. P., Offenberg, H. H., Wittich, P. E., Vischer, N. O. E. & Heyting, C. The mismatch repair protein MLH1 marks a subset of strongly interfering crossovers in tomato. *Plant Cell* **19**, 862–876 (2007).
55. Anderson, L. K., Reeves, A., Webb, L. M. & Ashley, T. Distribution of crossing over on mouse synaptonemal complexes using immunofluorescent localization of MLH1 protein. *Genetics* **151**, 1569–1579 (1999).
56. Dai, J. et al. Molecular basis of the dual role of the Mlh1–Mlh3 endonuclease in MMR and in meiotic crossover formation. *Proc. Natl Acad. Sci. USA* **118**, e2022704118 (2021).
57. Hollis, J. A. et al. Excess crossovers impede faithful meiotic chromosome segregation in *C. elegans*. *PLoS Genet.* **16**, e1009001 (2020).
58. Campoy, J. A. et al. Gamete binning: chromosome-level and haplotype-resolved genome assembly enabled by high-throughput single-cell sequencing of gamete genomes. *Genome Biol.* **21**, 306 (2020).
59. Heckmann, S. et al. Alternative meiotic chromatid segregation in the holocentric plant *Luzula elegans*. *Nat. Commun.* **5**, 4979 (2014).
60. Niwa, O., Shimanuki, M. & Miki, F. Telomere-led bouquet formation facilitates homologous chromosome pairing and restricts ectopic interaction in fission yeast meiosis. *EMBO J.* **19**, 3831–3840 (2000).
61. Blokhina, Y. P., Nguyen, A. D., Draper, B. W. & Burgess, S. M. The telomere bouquet is a hub where meiotic double-strand breaks, synapsis, and stable homolog juxtaposition are coordinated in the zebrafish, *Danio rerio*. *PLoS Genet.* **15**, e1007730 (2019).
62. Rockmill, B. & Roeder, G. S. Telomere-mediated chromosome pairing during meiosis in budding yeast. *Genes Dev.* **12**, 2574–2586 (1998).
63. Zickler, D. & Kleckner, N. Recombination, pairing, and synapsis of homologs during meiosis. *Cold Spring Harb. Perspect. Biol.* **7**, a016626 (2015).
64. Osman, K. et al. Distal bias of meiotic crossovers in hexaploid bread wheat reflects spatio-temporal asymmetry of the meiotic program. *Front. Plant Sci.* **12**, 631323 (2021).
65. Choi, K. et al. Nucleosomes and DNA methylation shape meiotic DSB frequency in *Arabidopsis thaliana* transposons and gene regulatory regions. *Genome Res.* **28**, 532–546 (2018).
66. Valero-Regalón, F. J. et al. Divergent patterns of meiotic double strand breaks and synapsis initiation dynamics suggest an evolutionary shift in the meiosis program between American and Australian marsupials. *Front. Cell Dev. Biol.* **11**, 1147610 (2023).
67. Sims, J., Copenhaver, G. P. & Schlögelhofer, P. Meiotic DNA repair in the nucleolus employs a nonhomologous end-joining mechanism. *Plant Cell* **31**, 2259–2275 (2019).
68. Kuttig, V. et al. The cohesin subunit RAD21.2 functions as a recombination silencer of ribosomal DNA arrays. Preprint at *bioRxiv* https://doi.org/10.1101/2022.06.20.496767 (2022)
69. Rockman, M. V. & Kruglyak, L. Recombinational landscape and population genomics of *Caenorhabditis elegans*. *PLoS Genet.* **5**, e1000419 (2009).
70. Gerstein, M. B. et al. Integrative analysis of the *Caenorhabditis elegans* genome by the modENCODE Project. *Science* **330**, 1775–1787 (2010).
71. Szymanska-Lejman, M. et al. The effect of DNA polymorphisms and natural variation on crossover hotspot activity in *Arabidopsis* hybrids. *Nat. Commun.* **14**, 33 (2023).
72. Petes, T. D. Meiotic recombination hot spots and cold spots. *Nat. Rev. Genet.* **2**, 360–369 (2001).
73. Loureiro, J., Rodriguez, E., Dolezel, J. & Santos, C. Two new nuclear isolation buffers for plant DNA flow cytometry: a test with 37 species. *Ann. Bot.* **100**, 875–888 (2007).
74. Chelysheva, L. et al. An easy protocol for studying chromatin and recombination protein dynamics during *Arabidopsis thaliana* meiosis: immunodetection of cohesins, histones and MLH1. *Cytogenet. Genome Res.* **129**, 143–153 (2010).
75. Dunn, K. W., Kamocka, M. M. & McDonald, J. H. A practical guide to evaluating colocalization in biological microscopy. *Am. J. Physiol. Cell Physiol.* **300**, C723–C742 (2011).
76. Costes, S. V. et al. Automatic and quantitative measurement of protein–protein colocalization in live cells. *Biophys. J.* **86**, 3993–4003 (2004).
77. Baez, M. et al. Analysis of the small chromosomal *Prionium serratum* (Cyperid) demonstrates the importance of reliable methods to differentiate between mono- and holocentricity. *Chromosoma* **129**, 285–297 (2020).
78. Braz, G. T., Yu, F., do Vale Martins, L. & Jiang, J. in *In Situ Hybridization Protocols* (eds Nielsen, B. S. & Jones, J.) 71–83 (Springer, 2020); https://doi.org/10.1007/978-1-0716-0623-0_4
79. Ruban, A. et al. B chromosomes of *Aegilops speltoides* are enriched in organelle genome-derived sequences. *PLoS ONE* **9**, e90214 (2014).

80. Gerlach, W. L. & Bedbrook, J. R. Cloning and characterization of ribosomal RNA genes from wheat and barley. *Nucleic Acids Res.* **7**, 1869–1885 (1979).
81. Cheng, H., Concepcion, G. T., Feng, X., Zhang, H. & Li, H. Haplotype-resolved de novo assembly using phased assembly graphs with hifiasm. *Nat. Methods* **18**, 170–175 (2021).
82. Ghurye, J. et al. Integrating Hi-C links with assembly graphs for chromosome-scale assembly. *PLoS Comput. Biol.* **15**, e1007273 (2019).
83. Goel, M., Sun, H., Jiao, W.-B. & Schneeberger, K. SyRI: finding genomic rearrangements and local sequence differences from whole-genome assemblies. *Genome Biol.* **20**, 277 (2019).
84. Goel, M. & Schneeberger, K. plotsr: visualizing structural similarities and rearrangements between multiple genomes. *Bioinformatics* **38**, 2922–2926 (2022).
85. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359 (2012).
86. Danecek, P. et al. Twelve years of SAMtools and BCFtools. *GigaScience* **10**, giab008 (2021).
87. Schneeberger, K. et al. SHOREmap: simultaneous mapping and mutation identification by deep sequencing. *Nat. Methods* **6**, 550–551 (2009).
88. Kim, D., Paggi, J. M., Park, C., Bennett, C. & Salzberg, S. L. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nat. Biotechnol.* **37**, 907–915 (2019).
89. Liu, D. Algorithms for efficiently collapsing reads with unique molecular identifiers. *PeerJ* **7**, e8275 (2019).
90. Zhang, L., Liang, Z., Hutchinson, J. & Kleckner, N. Crossover patterning by the beam-film model: analysis and implications. *PLoS Genet.* **10**, e1004042 (2014).
91. Rowan, B. A., Patel, V., Weigel, D. & Schneeberger, K. Rapid and inexpensive whole-genome genotyping-by-sequencing for crossover localization and fine-scale genetic mapping. *G3 (Bethesda)* **5**, 385–398 (2015).
92. Campos-Martin, R., Schmickler, S., Goel, M., Schneeberger, K. & Tresch, A. Reliable genotyping of recombinant genomes using a robust hidden Markov model. *Plant Physiol.* **192**, 821–836 (2023).
93. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J.* **17**, 10 (2011).
94. Ramírez, F. et al. deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res.* **44**, W160–W165 (2016).
95. Lopez-Delisle, L. et al. pyGenomeTracks: reproducible plots for multivariate genomic datasets. *Bioinformatics* **37**, 422–423 (2021).
96. Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842 (2010).
97. Zhang, Y. et al. Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* **9**, R137 (2008).
98. Neumann, P., Novák, P., Hoštáková, N. & Macas, J. Systematic survey of plant LTR-retrotransposons elucidates phylogenetic relationships of their polyprotein domains and provides a reference for element classification. *Mobile DNA* **10**, 1 (2019).
99. Novák, P., Neumann, P. & Macas, J. Global analysis of repetitive DNA from unassembled sequence reads using RepeatExplorer2. *Nat. Protoc.* **15**, 3745–3776 (2020).
100. Krueger, F. & Andrews, S. R. Bismark: a flexible aligner and methylation caller for bisulfite-Seq applications. *Bioinformatics* **27**, 1571–1572 (2011).
101. Dobin, A. et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
102. Kovaka, S. et al. Transcriptome assembly from long-read RNA-seq alignments with StringTie2. *Genome Biol.* **20**, 278 (2019).
103. Gremme, G., Brendel, V., Sparks, M. E. & Kurtz, S. Engineering a software tool for gene structure prediction in higher organisms. *Inf. Softw. Technol.* **47**, 965–978 (2005).
104. Wu, T. D. & Watanabe, C. K. GMAP: a genomic mapping and alignment program for mRNA and EST sequences. *Bioinformatics* **21**, 1859–1875 (2005).
105. Ghosh, S. & Chan, C.-K. K. in *Plant Bioinformatics* Vol. 1374 (ed. Edwards, D.) 339–361 (Springer, 2016).
106. Nachtweide, S. & Stanke, M. in *Gene Prediction* Vol. 1962 (ed. Kollmar, M.) 139–160 (Springer, 2019).
107. Ter-Hovhannisyan, V., Lomsadze, A., Chernoff, Y. O. & Borodovsky, M. Gene prediction in novel fungal genomes using an ab initio algorithm with unsupervised training. *Genome Res.* **18**, 1979–1990 (2008).
108. Haas, B. J. et al. Automated eukaryotic gene structure annotation using EVidenceModeler and the Program to Assemble Spliced Alignments. *Genome Biol.* **9**, R7 (2008).
109. Haas, B. J. Improving the *Arabidopsis* genome annotation using maximal transcript alignment assemblies. *Nucleic Acids Res.* **31**, 5654–5666 (2003).
110. Buchfink, B., Reuter, K. & Drost, H.-G. Sensitive protein alignments at tree-of-life scale using DIAMOND. *Nat. Methods* **18**, 366–368 (2021).
111. Tang, H. et al. Synteny and collinearity in plant genomes. *Science* **320**, 486–488 (2008).
112. Zhang, Z. et al. ParaAT: a parallel tool for constructing multiple protein-coding DNA alignments. *Biochem. Biophys. Res. Commun.* **419**, 779–781 (2012).
113. Zhang, Z. KaKs_Calculator 3.0: calculating selective pressure on coding and non-coding sequences. *Genomics Proteomics Bioinformatics* **20**, 536–540 (2022).
114. Yang, Z. & Nielsen, R. Estimating synonymous and nonsynonymous substitution rates under realistic evolutionary models. *Mol. Biol. Evol.* **17**, 32–43 (2000).
115. Gel, B. & Serra, E. karyoploteR: an R/Bioconductor package to plot customizable genomes displaying arbitrary data. *Bioinformatics* **33**, 3088–3090 (2017).

Acknowledgements

We thank R. Mercier for the insightful comments and discussion. We also thank R. Mercier and N. Donnelly for reviewing the manuscript. This study was funded by the Max Planck Society (core funding to A.M.), by the Deutsche Forschungsgemeinschaft (DFG, grant no. MA 9363/2-1 to A.M. and grant no. INST 216/1190 in the context of TRR341 to K.S.) and by the European Union (European Research Council Starting Grant, HoloRECOMB, grant no. 101114879 to A.M.). M.Z. is financially supported by the DFG (grant no. MA 9363/2-1). Deutsche Forschungsgemeinschaft (DFG) founded this work under Germany's Excellence Strategy—EXC 2048/1–390686111 (K.S. and A.M.). We thank the PhD fellowship awarded to G.T. from the German Academic Exchange Service (DAAD), India. Y.M.S. received financial support from the PROBRAL (Coordination Foundation for the Improvement of Higher Education Personnel/DAAD) programme (grant no. 88881.144086/2017-01). J.A.C. is financially supported by the Marie Skłodowska-Curie Individual Fellowship PrunMut (grant no. 789673). We thank U. Pfordt and C. Philipp for the excellent technical assistance.

Author contributions

A.M. conceived the research programme and coordinated the analyses. M.C. performed all cytogenetic analyses and microscopy. M.C. isolated the pollen nuclei and generated sequencing libraries with assistance from J.A.C. M.Z. performed all single-cell RNA sequencing and recombination-related analyses with assistance from H.S. G.T. performed the ChIP–seq analysis. Y.M.S. performed the immuno-FISH analysis. T.L. and K.F.X.M. performed the gene annotation and Ka/Ks ratio analysis. M.M. operated the FACS machine. B.H. performed all sequencing. K.S. supervised the single-cell analysis. M.C., M.Z. and

A.M. wrote the first manuscript draft with input from all authors. All authors approved the final version of the manuscript.

Funding

Open access funding provided by Max Planck Society.

Competing interests

The authors declare no competing interests.

Additional information

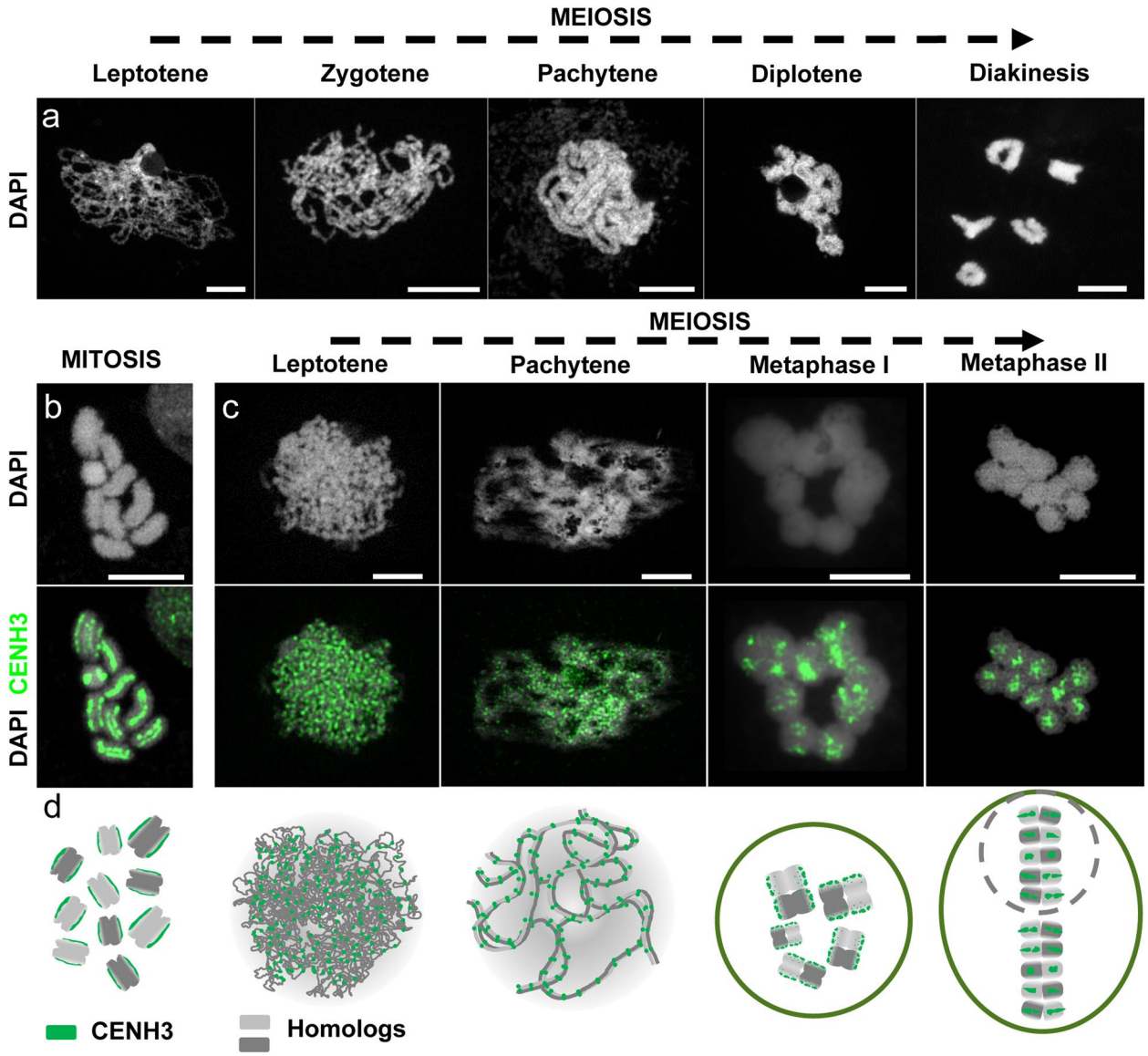
Extended data is available for this paper at https://doi.org/10.1038/s41477-024-01625-y.

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41477-024-01625-y.

Correspondence and requests for materials should be addressed to André Marques.

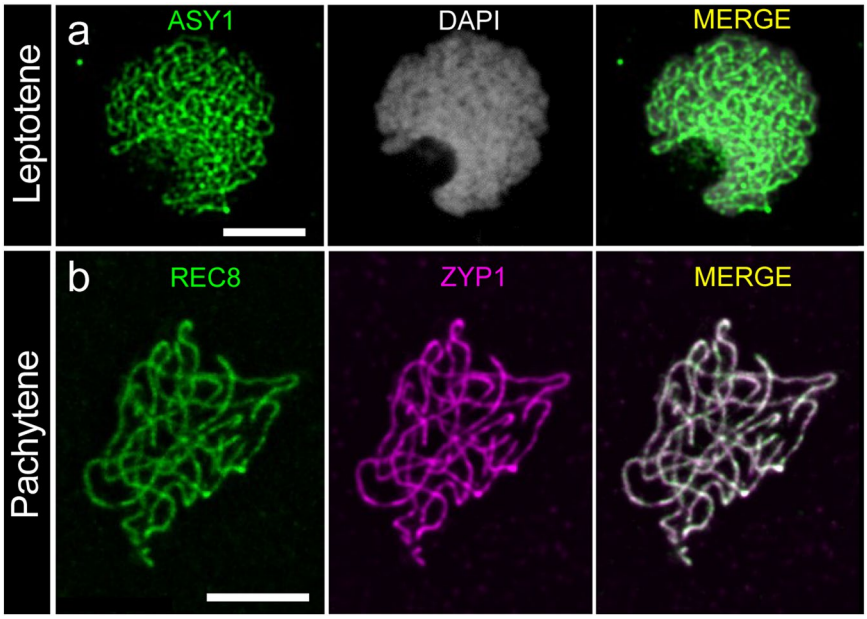
Peer review information *Nature Plants* thanks Ian Henderson and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

https://doi.org/10.1038/s41477-024-01625-y



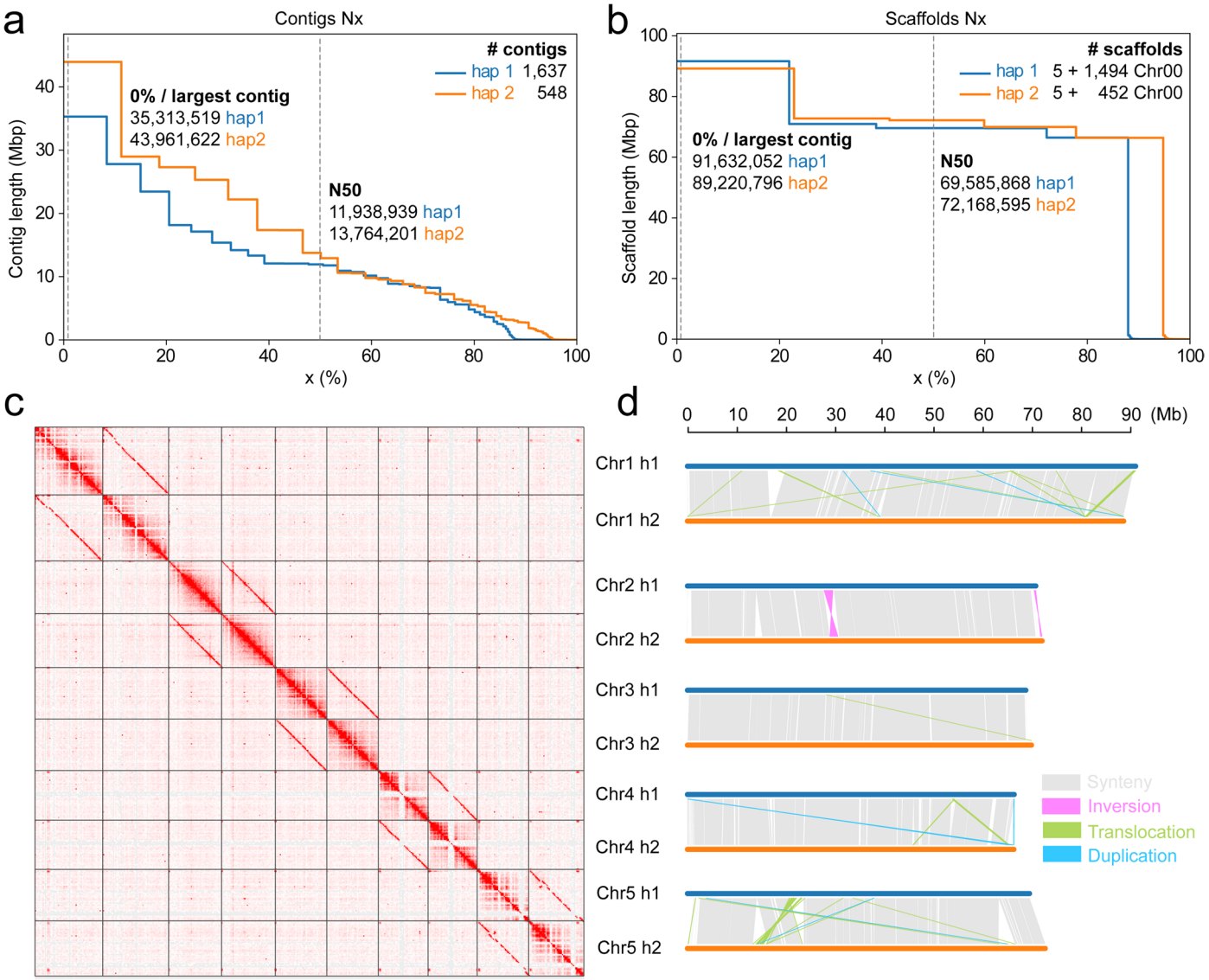
Extended Data Fig. 1 | Chromosome spreads and immunolocalisation in male *R. breviuscula* meiocytes. (a) Meiotic stages are displayed, including leptotene, zygotene, pachytene, diplotene and diakinesis. DAPI staining was consistent in all cells analysed in our immunocytochemistry experiments (see main text). **(b–c)** Immunolocalisation was performed against the centromeric protein CENH3, which appears as lines during mitosis **(b)** and as clusters during meiosis **(c)**. A maximum projection is shown, and the DNA was counterstained with DAPI. **(d)** Model of differential mitotic and meiotic holocentromere organization in the holocentric plant *R. breviuscula*. CENH3 behaviour was consistent in all cells (n = 31) in 4 independent experiments. Scale bars, 5 µm.

https://doi.org/10.1038/s41477-024-01625-y



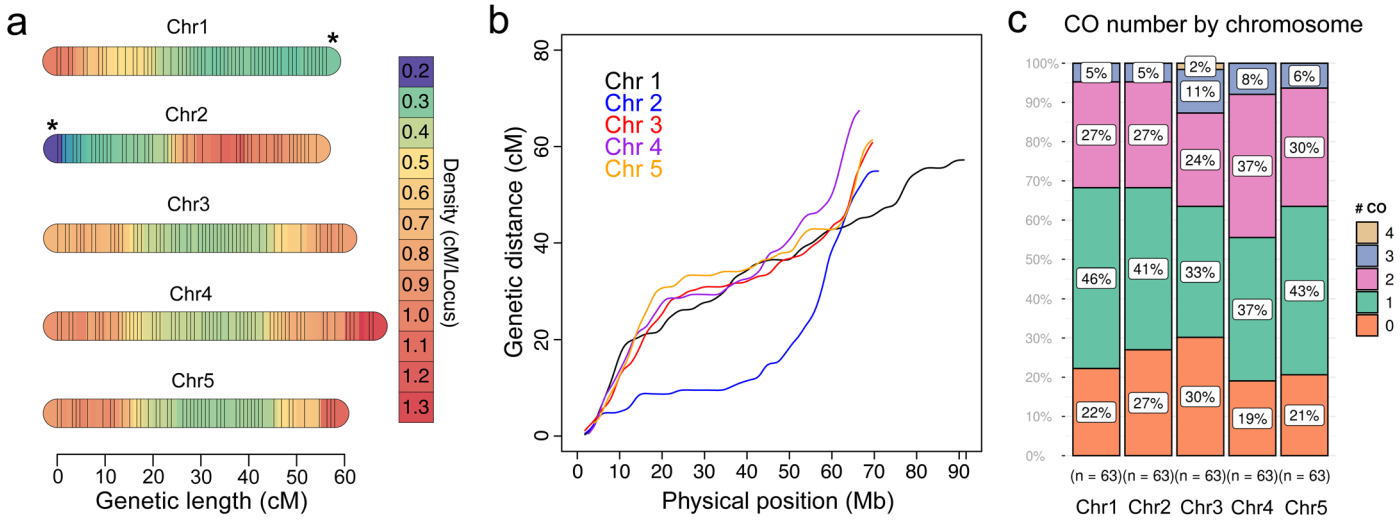
Extended Data Fig. 2 | Immunolocalisation of ASY1, REC8, and ZYP1 from leptotene to pachytene of meiotic prophase I. (a) ASY1 (green) appears as a linear signal on unpaired chromosomes. Behaviour of ASY1 was consistent in all meiocytes analysed (n = 10) among 6 independent experiments. **(b)** Full co-localisation of cohesin protein REC8 (green) and ZYP1 (magenta) at pachytene. Behaviour of REC8+ZYP1 was consistent in all meiocytes analysed (n = 12) among 3 independent experiments. A maximum projection is shown; chromosomes were counterstained with DAPI. Scale bars, 5 µm.

https://doi.org/10.1038/s41477-024-01625-y



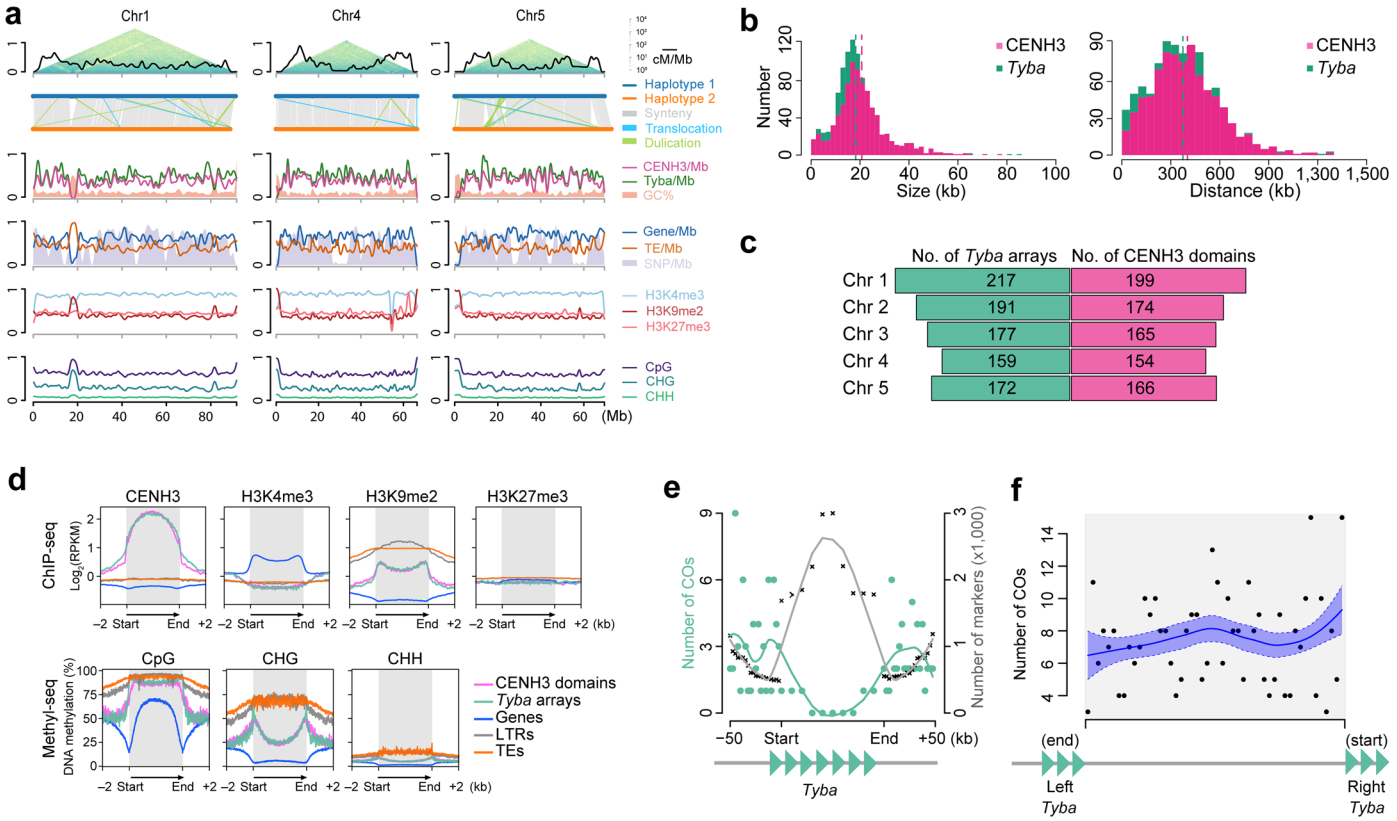
Extended Data Fig. 3 | Phasing and structural variation of the *R. breviuscula* heterozygous genome. **(a–b)** Assembly statistics of the phased contigs **(a)** and scaffolds **(b)** for haplotype 1 and haplotype 2. **(c)** Hi-C scaffolding of the five haplotype-phased pseudochromosomes. Homozygous regions between the haplotypes are seen as clear regions depleted of signals on the Hi-C map. **(d)** Synteny blocks and structural variants (>10 kb) identified between the two haploid assemblies. Note the overall high synteny between the two haplotypes. Synteny blocks were computed with SyRI and plotted with plotsr.

https://doi.org/10.1038/s41477-024-01625-y



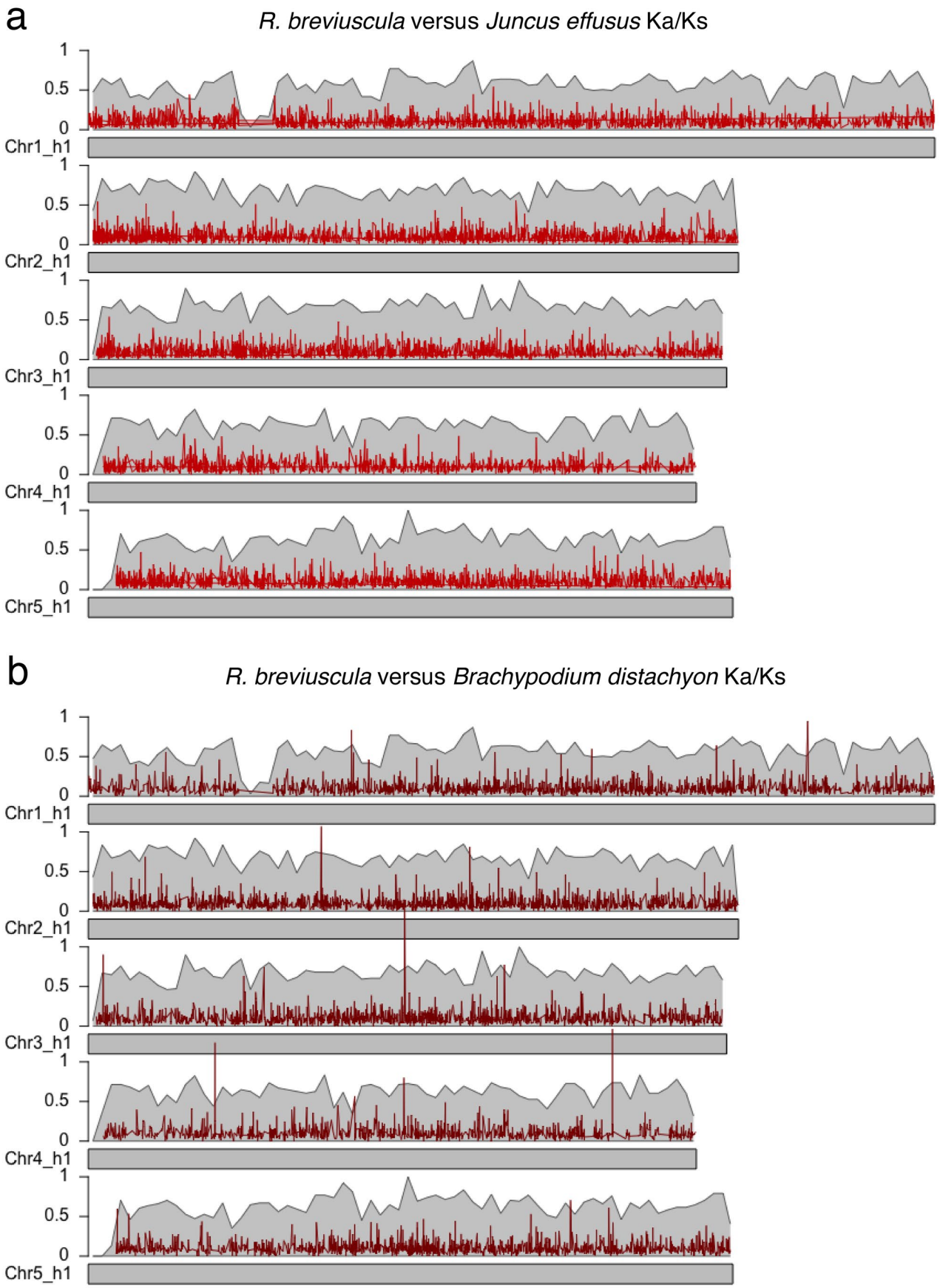
Extended Data Fig. 4 | Recombination frequencies in the F1 recombinant offspring of *R. breviuscula*. **(a)** Genetic linkage map using 378 COs detected in our 63 F1 individuals with density indicated by colouring. The 705 markers were selected using a 500-kb sliding window through all markers defined on the reference (**see Methods**). **(b)** Marey map calculated from the linkage map in **a**. Marey maps for each chromosome (colour lines) show genetic position as a function of physical position. **(c)** Distribution of CO numbers individual chromosomes in the 63 F_1 recombinant individuals.

https://doi.org/10.1038/s41477-024-01625-y



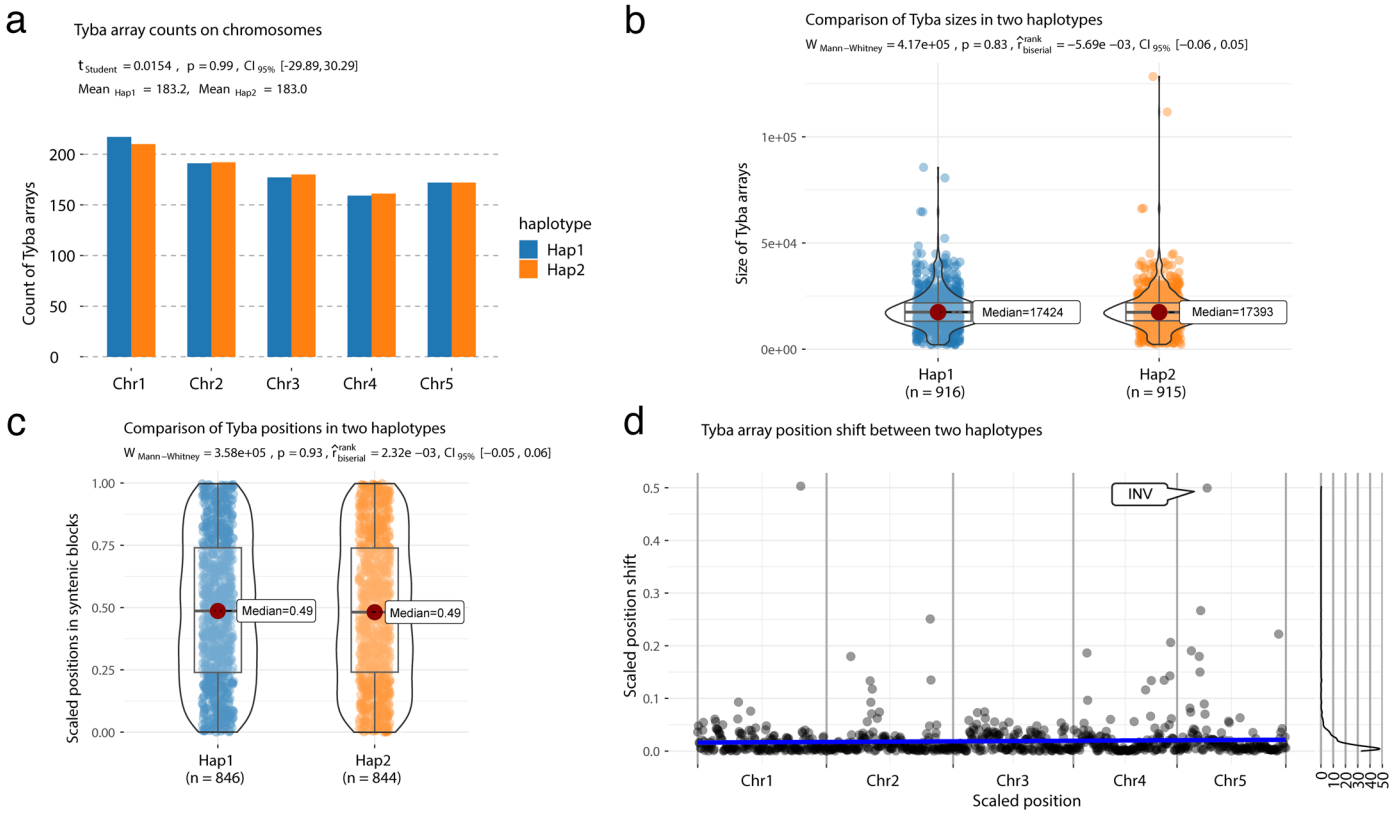
Extended Data Fig. 5 | Epigenetic regulation of repeat-based holocentromeres and fine-scale correlation of CO positions with *Tyba* repeats. (a) Chromosome distribution of the CO rate coupled with different (epi)genetic features. Top: recombination landscape (black line) created with COs detected in all single-pollen nuclei (n = 1,641), coupled with Omni-C chromosome conformation capture contacts. Synteny analysis and detected the structural variants between the two haplotypes. For the y-axes, all features were scaled [0,1], with 1 indicating a maximum of 2.34 for recombination frequency (cM/Mb), 5 for *Tyba* density, 6 for CENH3 density, 7205 for SNP density, 88 for gene density, and 227 for TE density. GC [33.3, 46.6], H3K4me3 [−1.494, 0.231], H3K9me2 [−1.20, 1.84], and H3K27me3 [−0.671, 0.491] are scaled to [0,1] by their minima and maxima. mCG, mCHG and CHH are original values (0 to 100%). All features were smoothed by 1Mbp sliding window and 250kbp step size. **(b)** Size (left) and spacing (right) length distribution of CENH3 domains and *Tyba* arrays. CENH3 domain median size is 19,156 bp and the mean size is 20,697 bp. The median of *Tyba* array size is 17,424 bp and the mean is 18,220 bp. CENH3 domain median spacing is 378,467 bp and the mean is 401,763 bp. The median of *Tyba* array spacing is 354,850 bp and the mean is 374,310 bp. **(c)** Number of *Tyba* arrays (left) and CENH3 domains (right) for each chromosome annotated in the reference haplotype genome. **(d)** Enrichment of CENH3, H3K4me3, H3K9me2, and DNA methylation in the CpG, CHG, and CHH contexts from the start and end of different types of sequences: CENH3 domains, *Tyba* repeats, genes, LTRs, and TEs. ChIP-seq signals are shown as log2 (normalised RPKM ChIP/input). Grey boxes highlight the modification enrichment over the body of each sequence type. **(e)** CO frequency within *Tyba* arrays. **(f)** Random distribution of the relative distance of CO positions to the end of the left and to the start of the right *Tyba* array. The median of CO resolution is 334 bp and the mean is about 2 kb. The solid blue line was predicted by Local Polynomial Regression Fitting (loess function from R) using data from 63 F_1 recombinant offspring and a total of 378 COs; whereas the dashed blue band presents the range of one standard error above and below the fitted line. Green-filled triangles schematically represent *Tyba* repeat arrays.

https://doi.org/10.1038/s41477-024-01625-y



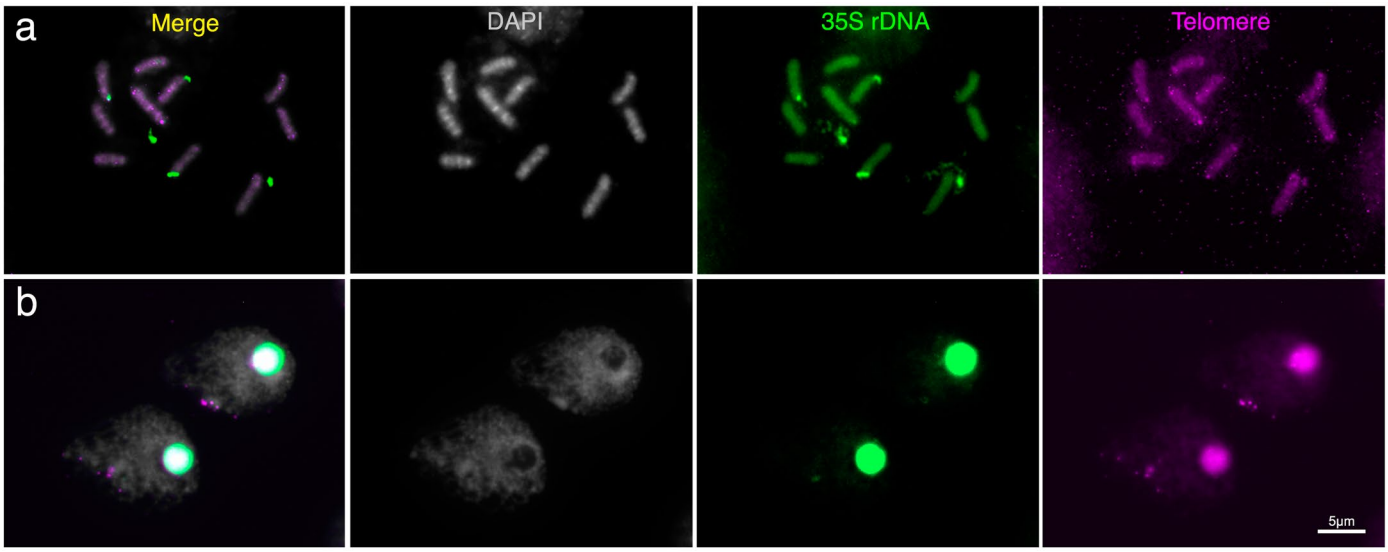
Extended Data Fig. 6 | Ka/Ks ratio estimation across the chromosomes of *R. breviuscula*. **(a)** Ka/Ks ratio comparison between *R. breviuscula* and *Juncus effusus* genomes. **(b)** Ka/Ks ratio comparison between *R. breviuscula* and *Brachypodium distachyon* genomes. X-axis: gene start, Y-axis: Ka/Ks values; grey background: gene density. High Ka/Ks values indicate fast-evolving regions.

https://doi.org/10.1038/s41477-024-01625-y



Extended Data Fig. 7 | Structural variation among *Tyba* arrays between haplotype 1 and 2 in the phased genome of *R. breviuscula*. **(a)** Number of *Tyba* arrays on each chromosome of two haplotypes. Since *Tyba* array numbers in both haplotypes are normally distributed (Shapiro-Wilk normality test p-values: 0.7571 and 0.945) and have no significant difference in variance (F-test p-values: 0.767), two-sided t-test was performed to compare the means of *Tyba* array number of two haplotypes. **(b)** Sizes of *Tyba* of arrays in two haplotypes. Mean sizes of *Tyba* arrays in two haplotypes are 17,424 bp and 17,393 bp. N shows the number of *Tyba* arrays in each haplotype. The box plots in figure b-c are comprised of minima, Q1, median, Q3, and maxima following the definition of ggplot2 in R. Two-sided Mann-Whitney U tests were performed to compare the *Tyba* sizes (panel **b**) and scaled positions in syntenic blocks (panel **c**) between two haplotypes. **(c-d)** Scaled position distribution in two haplotypes (panel **c**) and position shift between haplotypes (panel **d**) after pairing *Tyba* arrays by the closest relative positions in the corresponding synteny blocks. N in panel **c** shows the number of paired syntenic *Tyba* arrays in two haplotypes. Most outliers in panel **d** are from inversion syntenic blocks as the callout shown on Chr5, long indels, and highly divergent regions based on SyRI output (see **Methods**).

https://doi.org/10.1038/s41477-024-01625-y



Extended Data Fig. 8 | FISH with 35S rDNA and a telomeric probe in *R. breviuscula*. **(a)** Mitotic metaphase and **(b)** Prophase I. Telomeres of the *rDNA*-harbouring chromosomes 1 and 2 cluster in the nucleolus. Squares in **b** show telomeric sequences in chromosomes with 35 rDNA. FISH with 35S rDNA and telomeres showed consistent results (n = 25 cells) in two different experiments. Scale bar, 5 μm.

nature portfolio

Corresponding author(s): André Marques

Last updated by author(s): Jan 3, 2024

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed
☐	☒ The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒ A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒ The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐ A description of all covariates tested
☐	☒ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒ A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒ For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒ Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection

Sorting of pollen nuclei was performed by a BD FACSAria Fusion sorter (BD Biosciences). The Epifluorescence microscope Zeiss Axio Imager Z2 with Apotome system for optical sectioning and Leica Microsystems Thunder Imager dMi8 with Computational Clearing. were used for acquiring microscopic images.

Data analysis

Custom code:
CO detection pipeline by scRNA-seq (https://github.com/Raina-M/detectCO_by_scRNAseq)

Other available open source tools used in this study:
AHRD pipeline (v3.3.3; https://github.com/groupschoof/AHRD)
Augustus (v3.3.3)
Deeptools (v3.5.1)
Bcctools (v0.1.1)
Bcftools (v1.9)
Bedtools (v2.29.0)
Bismark (v0.23.0)
bowtie2 (v2.4.4)
BUSCO (v5.1.2)
Cuffcompare (v2.1.1)
Cutadapt (v4.7)
DIAMOND (v2.0.5)
EVidenceModeller (v1.1.1)

GeneMark (v4.35)
GenomeThreader (v1.7.1)
GMAP (2018-07-04)
HiFiasm (0.19.5-r592)
Hisat2 (v2.1.0)
KaKs_Calculator (v3)
karyoploteR (v3.18)
MACS3 (v3.0.0)
MADpattern (v1.1)
minimap2 (v2.26)
PASA (v2.4.1)
ParaAT (v2)
pyGenomeTracks (v3.8)
RepeatExplorer2 (v2.3.7)
RTIGER (v1.99.0)
SALSA2 (v2.3)
Samtools (v1.9)
SHOREmap (v3.6)
STAR (v2.7.8a)
StringTie (v2.1.5)
SyRI (v1.5)
TIGER (v1.0)
TransDecoder (v5.5.0)
UMIcollapse (v1.0.0)

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

All sequencing data used in this study have been deposited at NCBI under the BioProject ID PRJNA1059790 and are publicly available as of the date of publication. The reference genomes, sequencing data, annotations and all tracks presented in this work are made available for download at DRYAD: https://datadryad.org/stash/share/EvB3PRNVph5IiTkOM3jTZddgmS45cJhQYq2v3LI5InE. The REXdb database Viridiplantae v3.0 [http://repeatexplorer.org/?page_id=918] is publicly available. All other data needed to evaluate the conclusions in the paper are provided in the paper and/or the supplemental information.

Research involving human participants, their data, or biological material

Policy information about studies with human participants or human data. See also policy information about sex, gender (identity/presentation), and sexual orientation and race, ethnicity and racism.

Reporting on sex and gender	n/a
Reporting on race, ethnicity, or other socially relevant groupings	n/a
Population characteristics	n/a
Recruitment	n/a
Ethics oversight	n/a

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	Sample-size calculation was performed based on assessment of the literature in the field, our own experience from previous studies and requirement for corresponding protocols. For Immunocytochemistry analyis sample size was based on the number of cells obtained. For single-cell sequencing, sample size was determined based on the number of cells that can be typically obtained using 10X scRNAseq libraries. For F1 samples, the sample size was based on the maximum number of seeds obtained by selfing. The sample size used for all experiments provided sufficient resolving power.
Data exclusions	10X Genomics scRNA library was initially performed by combining pollen grains from R. breviuscula and R. tenuis, for multiplex purposes. However, all the pollen data from R. tenuis was excluded from the analysis, as this will be used in separate manuscript.
Replication	Replication was mainly used for our ChIP experiments, which were all done in two replicates. We confirm that all replicates showed similar results, confirming the reproducibility of our analysis. Furthermore, cytological analyses were performed several times in different days and with different fixations using the same set of antibodies. We confirm that our cytological data is very reproducible. Individual F1 plants WGS and scRNA sequencing approaches were not replicated, since each individual pollen nuclei or F1 plants already represent biological replicates of recombination events.
Randomization	A randomization is not relevant for this study because no genotype or treatment were compared with each other. However, the tissues for cytogenetic and ChIPseq experiments were randomly collected from different plant individuals grown under the same condition in a greenhouse. Only for scRNAseq, we used the pollen collected from a single R. breviuscula plant as this needs to match the same heterozygous polymorphisms of the reference mother plant in order to detect COs using our method.
Blinding	All the experiments were performed without prior knowledge of the final outcome, and therefore blinding was not applied.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☐	☒ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☒	☐ Clinical data
☒	☐ Dual use research of concern
☐	☒ Plants

Methods

n/a	Involved in the study
☐	☒ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used

Customized Rhynchospora-specific antibodies generated in this study:
1. anti-REC8 was a combination of two antibodies raised in rabbits against the Rhynchospora-specific REC8-peptides C-EEPYGEIQISKGPNM and C-YNPDDSVERMRDDPG (gene ID: RBREV_HAP1.r01.Chr4_h1G00395720.1) and affinity-purified (Eurogentec).
2. anti-HEI10 was a combination of two antibodies raised in rabbits against the Rhynchospora-specific HEI10-peptides C-NRPNQSRARTNMFQL and C-PVRQRNNKSMVSGGP (gene ID: RBREV_HAP1.r01.Chr4_h1G00387160.1) and affinity-purified (Eurogentec).

Previously designed antibodies:
3. anti-AtASY1 raised in rabbits (inventory code PAK006) (Armstrong et al. 2002)
4. anti-AtMLH1 raised in rabbits (PAK017) (Chelysheva et al. 2010)
5. anti-RhynchosporaCENH3 raised in rabbits (Marques et al. 2015).
6. anti-ZYP1 was raised in chickens against the peptide EGSLNPYADDPYAFD of the C-terminal end of AtZYP1a/b (gene ID: At1g22260/At1g22275) and affinity-purified (Eurogentec) (PAK048).
Each primary antibody above was diluted 1:200 in blocking solution.

Commercially available antibodies:
7. anti-H3K4me3 (rabbit polyclonal to Histone H3 tri-methyl K4; Abcam, UK, cat. no. ab8580, dilution 1:300)
8. anti-H3K9me2 (mouse monoclonal to Histone H3 di-methyl K9, Abcam, UK, cat. no. ab1220, clone no. mAbcam 1220, dilution 1:200)
9. anti-H3K27me3 (mouse monoclonal to Histone H3 tri-methyl K27, Abcam, UK, cat. no. ab6002, clone no. mAbcam 6002, dilution 1:200)
10. anti-IgG control (recombinant rabbit IgG, monoclonal Abcam, UK, cat. no. ab172730, clone no. EPR25A, dilution 1:300).

Commercially available secondary antibodies:
11. Abberior STAR ORANGE (dilution 1:250, Abberior, DE, goat anti-rabbit IgG, cat. no. STORANGE-1002-500UG or goat anti-chicken IgY, cat. no. STORANGE-1005-500UG)
12. Abberior STAR RED (dilution 1:250, Abberior, DE, goat anti-rabbit IgG, cat. no. STRED-1002-500UG or goat anti-chicken IgY, cat. no. STRED-1005-500UG)

Validation

Previously validated antibodies;
1. Anti-AtASY1 (inventory code PAK006) was validated in Armstrong et al. (2002).
2. Anti-AtMLH1 (PAK017) was validated in Chelysheva et al. (2010)
3. Anti-RhynchosporaCENH3 was validated in Marques et al. (2015, 2016)

Validation by commercial providers:
4. anti-H3K4me3 (cat. no. ab8580, https://www.abcam.com/products/primary-antibodies/histone-h3-tri-methyl-k4-antibody-chip-grade-ab8580.html)
5. anti-H3K9me2 (cat. no. ab1220, https://www.abcam.com/products/primary-antibodies/histone-h3-di-methyl-k9-antibody-mabcam-1220-chip-grade-ab1220.html)
6. anti-H3K27me3 (cat. no. ab6002, https://www.abcam.com/products/primary-antibodies/histone-h3-tri-methyl-k27-antibody-mabcam-6002-chip-grade-ab6002.html)
7. anti-IgG (cat. no. ab172730, https://www.abcam.com/products/primary-antibodies/rabbit-igg-monoclonal-epr25a-isotype-control-ab172730.html)

Newly validated antibodies:
4. Anti-ZYP1 antibody was generated by the company EUROGENTEC and validated by peptide ELISA tests. Further validation was demonstrated in A. thaliana by showing positive staining of ZYP1 in wildtype, but not in zyp1 mutants (data not shown). Remarkably, this peptide region showed 100% similarity with the Rhynchospora ZYP1 C-terminal (gene ID: RBREV_HAP1.r01.Chr2_h1G00222020.1).
Rhynchospora-specific REC8 and HEI10 antibodies were generated by the company EUROGENTEC and validated by peptide ELISA test. Further validation was obtained by the presence of specific signals on western blots (data not shown). Both ELISA and western blot information are available upon request. Furthermore, the observed indirect immuno-signals of anti-ZYP1, anti-HEI10 and anti-REC8 are compatible with data previously reported in the published literature for other species.

Plants

Seed stocks

Individual plants from R. breviuscula were previously obtained by Hofstatter et al. 2022 and kept under cultivation at growth chambers at the Max Planck Institute for Plant Breeding Research in Cologne, Germany. A F1 progeny was obtained by selfing the original heterozygous mother plant used in Hofstatter et al. 2022. Additionally, all cytology and pollen were collected from the same heterozygous mother plant genotype.

Novel plant genotypes

Describe the methods by which all novel plant genotypes were produced. This includes those generated by transgenic approaches, gene editing, chemical/radiation-based mutagenesis and hybridization. For transgenic lines, describe the transformation method, the number of independent lines analyzed and the generation upon which experiments were performed. For gene-edited lines, describe the editor used, the endogenous sequence targeted for editing, the targeting guide RNA sequence (if applicable) and how the editor was applied.

Authentication

Describe any authentication procedures for each seed stock used or novel genotype generated. Describe any experiments used to assess the effect of a mutation and, where applicable, how potential secondary effects (e.g. second site T-DNA insertions, mosiacism, off-target gene editing) were examined.

ChIP-seq

Data deposition

☒ Confirm that both raw and final processed data have been deposited in a public database such as GEO.

☒ Confirm that you have deposited or provided access to graph files (e.g. BED files) for the called peaks.

Data access links
May remain private before publication.

https://datadryad.org/stash/share/EvB3PRNVph5IiTkOM3jTZddgmS45cJhQYq2v3LI5InE

Files in database submission

Phased genome assemblies of two haplotypes of R. breviuscula

rhyBreHap1.fasta.gz

rhyBreHap2.fasta.gz

Annotation of phased genome assemblies

rhyBreHap1_functAnno.gff3.gz

rhyBreHap2_functAnno.gff3.gz

Single-cell RNA sequences from pollen nuclei by 10X Genomics

a4984_merged_R1.fastq.gz

a4984_merged_R1.fastq.gz

Please note that the scRNA-seq data were sequenced from mixed R. breviuscula and R. tenuis pollen nuclei. Therefore, you need to separate the pollen of two individuals before any further analysis. To do this, you can refer to the online methods of the above paper and our github docbumentation.

Whole-genome DNA short reads of 63 selfed F1 offspring by Illumina paired-end sequencing

F1_WGS.tar.gz

WGS DNA reads of all 63 F1 individuals are in this compressed folder. Sample IDs are 5445_A, 5445_B, 5445_C, 5445_D, 5445_E, 5621_A, 5621_B, 5621_C, 5621_E, 5621_F, 5621_G, 5621_H, 5621_I, 5621_J, 5621_K, 5621_L, 5621_M, 5621_N, 5621_O, 5621_P, 5621_Q, 5621_R, 5621_S, 5621_T, 5621_U, 5621_V, 5621_W, 5621_X, 5621_Y, 5844_A, 5844_B, 5844_C, 5844_D, 5844_E, 5844_F, 5844_G, 5844_H, 5844_I, 5844_J, 5844_K, 5844_L, 5844_M, 5844_N, 5844_O, 5844_P, 5844_Q, 5844_R, 5844_S, 5844_T, 5844_U, 5844_V, 5844_W, 5844_X, 5844_Y, 5844_Z, 5844_AA, 5844_AB, 5844_AC, 5844_AD, 5844_AE, 5844_AF, 5844_AG, 5844_AH. Each sample was sequenced under one single library, so you can merge the reads with the same sample names.

ChIPseq data

ChIPseq_raw_reads.tar.gz

5165_A CENH3 rep1 rabbit

5165_B CENH3 rep2 rabbit

5165_C H3K4me3 rep1 rabbit

5165_D H4K4me3 rep2 rabbit

5165_E H3K9me2 rep1 mouse

5165_F H3K9me2 rep2 mouse

5165_G Rabbit-IgG rep1 rabbit

5165_H Rabbit-IgG input rep2 rabbit

5165_I input chromatin rep1

5165_J input chromatin rep2

5165_M Mouse-IgG rep1 mouse

5165_N Mouse-IgG rep2 mouse

5165_O H3K27me3 rep1 rabbit

5165_P H3K27me3 rep2 rabbit

Each sample was sequenced under one single library, so you can merge the reads with the same sample names.

ChIP_peaks.zip

Methyl-seq data

Rhync_breviuscula_Methyl-seq.tar.gz

Each sample was sequenced under one single library, so you can merge the reads with the same sample names.

Complete CO detection pipeline by pollen scRNA-seq data can be found on this github page.

Genome browser session
(e.g. UCSC)

no longer applicable

Methodology

Replicates

ChIP experiments were performed in two technical replicates for each antibody used. All replicates agreed by showing similar enrichment results.

Sequencing depth

Each replicate was sequenced at an approx. 3.75 genome sequencing depth, i.e., 10 million of 150bp single-end reads. Alignment

Sequencing depth	rate >90% for all samples. CENH3 uniquely mapped reads >28% for both replicates. H3K4me3 uniquely mapped reads >45% for both replicates. H3K9me2 uniquely mapped reads >20% for both replicates. H3K27me3 uniquely mapped reads >37% for both replicates.
Antibodies	CENH3 ChIP-seq data were obtained from Hofstatter et al. (2022). Further ChIP experiments were performed for H3K4me3 (rabbit polyclonal to Histone H3 tri-methyl K4; Abcam ab8580), H3K9me2 (mouse monoclonal to Histone H3 di-methyl K9, Abcam ab1220), H3K27me3 (mouse monoclonal to Histone H3 tri-methyl K27, Abcam ab6002), and the IgG control (recombinant rabbit IgG, monoclonal Abcam ab172730) using the same protocol described by Hofstatter et al. (2022).
Peak calling parameters	MACS3 software was used for peak calling: macs3 callpeak -t ChIP.bam -c Control.bam --broad -g 380000000 --broad-cutoff 0.1
Data quality	Quality of peaks were checked by comparing the enrichment of immunoprecitated DNA in comparison to our controls, i.e., input chromatin and IgG, using deeptools. All peaks retained showed a fold enrichment above 5 and FDR less than 5%.
Software	bowtie2, MACS3, deeptools (bamCoverage, bamCompare, computeMatrix, plotHeatmap)

2.2 Holocentric karyotype plasticity alters recombination through changes in synapsis and chromatin configuration

Meng Zhang*, Stefan Steckenborn*, Marco Castellani*, Laia Marin Gual, Piotr Wlodzimierz, Laura A. Robledillo, Nafiseh Sargheini, Magdalena Marek, Eduardo Chacón, Wayt W. Thomas, Bruno Huettel, Leonardo P. Félix, Andrés Aria Gatica, Andrea Pedrosa-Harand, André L. L. Vanzela, Aurora Ruiz-Herrera, André Marques#

*These authors contributed equally
Correspondence

Author Contributions:

A.M. conceived the study. M.Z. performed genome assembly and bioinformatic analyses. S.S. performed all the single-cell pollen sequencing experiments. M.C. performed the cytological analysis of meiosis. All authors contributed to writing.

Holocentric karyotype plasticity alters recombination through changes in synapsis and chromatin configuration

Meng Zhang[1]*, Stefan Steckenborn[1]*, Marco Castellani[1]*, Laia Marin Gual[2], Piotr Wlodzimierz[3], Laura A. Robledillo[1], Nafiseh Sargheini[4], Magdalena Marek[4], Eduardo Chacón[5], Wayt W. Thomas[6], Bruno Huettel[4], Leonardo P. Félix[7], Andrés Aria Gatica[5], Andrea Pedrosa-Harand[8], André L. L. Vanzela[9], Aurora Ruiz-Herrera[2], André Marques[1,11]#

[1]Department of Chromosome Biology, Max Planck Institute for Plant Breeding Research, Cologne, Germany

[2]Department of Cellular Biology, Physiology and Immunology, Universitat Autònoma de Barcelona, Barcelona, Spain

[3]Institute of Biochemistry and Biophysics, Polish Academy of Sciences, Warsaw, Poland

[4]Max Planck Genome Centre Cologne, Max Planck Institute for Plant Breeding Research, Cologne, Germany

[5]Universidad de Costa Rica, San José, Costa Rica

[6]New York Botanical Garden, Bronx, New York, USA

[7]Federal University of Paraíba, João Pessoa, Brazil

[8]Department of Botany, Federal University of Pernambuco, Recife, Brazil

[9]State University of Londrina, Londrina, Brazil

[10]Cluster of Excellence on Plant Sciences (CEPLAS), Heinrich-Heine University, Düsseldorf, Germany

Correspondence: amarques@mpipz.mpg.de

Abstract

Chromosomal fissions, fusions, and whole-genome duplications can alter recombination landscapes, yet their impact is often masked by the strong suppression and positional constraints imposed by centromeres. Large chromosomal rearrangements in monocentric species frequently have deleterious meiotic consequences, limiting their persistence in populations[1–4]. Holocentric organisms, which distribute kinetochore activity along the entire chromosome, overcome this barrier and therefore offer a unique window onto the interplay between karyotype change and crossover control[2]. In this study, we assembled chromosome-scale genomes for 20 holocentric *Rhynchospora* species (including 54 haplotypes), representing all major clades of the genus, featuring satellite-based holocentromeres[5,6]. Breakpoint analysis shows that centromeric *Tyba* satellite arrays are recurrent hotspots for both chromosome fusions and fissions, driving the genus's extraordinary chromosome-number variation from $2n = 4$ to 36. Using single-gamete sequencing, we found crossover landscapes segregate into two classes: strongly distal-biased versus irregular distribution, which is determined not by sequence or epigenetic state but is correlated with divergent patterns of synapsis elongation. Chromosome architecture emerges as the dominant quantitative driver of recombination variations: crossover number scales with chromosome count, while crossover rate is negatively correlated to chromosome length and to chromatin-loop size inferred from Hi-C. We propose that shorter loops on smaller chromosomes extend the axial substrate accessible for double-strand-break formation, thus elevating recombination. Together, our results provide a structural link between large-scale structural chromosome evolution and meiotic recombination through coupled changes in chromosome number, size, loop geometry, and synapsis dynamics, offering a mechanistic framework for how karyotype diversity fuels genetic diversity, adaptation, and speciation.

Main

Chromosomal rearrangements, including end-to-end reciprocal chromosome translocations (hereafter called 'fusions', for simplicity), fissions, and whole-genome duplications, are major drivers of genome evolution, species divergence, and adaptation[2–4,7,8]. These structural changes can alter karyotypes and impact gene flow, recombination, and reproductive isolation[2,8]. While such changes are widespread, their consequences for chromosome behaviour and genetic diversity vary dramatically across taxa and chromosome architectures[2,8].

In monocentric species, where chromosome segregation relies on a single centromere, large-scale rearrangements often compromise chromosome stability, resulting in missegregation or even lethality[1,2]. In contrast, holocentric organisms, in which centromere activity is distributed along the entire chromosome, exhibit remarkable tolerance to chromosomal restructuring[2,9]. This architecture allows chromosomes fragmented by fission or joined by fusion while still maintaining proper segregation[9]. The sedge genus *Rhynchospora*, characterised by holocentromeres based on the *Tyba* satellite DNA arrays[5,6,10] and large chromosome number variation[11,12], provides a powerful model to investigate how karyotype evolution proceeds in the absence of centromere constraints.

Previous work from our lab has provided evidence that chromosomal fusions and, potentially, fissions underlie major karyotypic shifts in *Rhynchospora* and can be associated with *Tyba*-rich regions[6], suggesting that repetitive sequences may facilitate structural change. In parallel, studies of meiotic recombination in *R. breviuscula* revealed a striking departure from canonical monocentric patterns: crossover distribution is not associated with centromere distribution and shows no consistent association with the distributions of gene density, epigenetic marks, or transposable elements[13]. Instead, the spatial patterning of recombination appears to be shaped by chromosome pairing dynamics, particularly telomere-led synapsis initiation, rather than sequence or chromatin state.

Despite this, it remains unknown how the extreme variation in chromosome number and structure observed across *Rhynchospora* species ($2n = 4$ to $2n = 36$)[11,12] influences the distribution and frequency of meiotic crossovers (COs). Do chromosomal rearrangements

directly alter meiotic recombination? Are recombination landscapes conserved across the genus, or do they evolve in concert with karyotype change? And how do chromosome size and chromatin architecture modulate recombination rates in holocentric species?

Here, we combined high-quality haplotype-phased chromosome-level genome assemblies, single-cell pollen sequencing, synapsis cytology, and chromatin conformation data across 20 holocentric *Rhynchospora* species. By integrating structural and functional genomics, we show that chromosomal rearrangements, frequently driven by Tyba-mediated fusions and fissions, reconfigure recombination landscapes by altering chromosome size, loop architecture, and synapsis dynamics. Our results provide a mechanistic framework linking holocentric karyotype evolution to recombination plasticity, with broader implications for understanding how chromosomal structure shapes genetic diversity and speciation.

The *Rhynchospora* pangenome reveals extensive holocentric karyotype evolution

We assembled chromosome-scale genomes for 17 additional *Rhynchospora* species and integrated them with three previously published genomes[6], yielding 54 haplotypes from 20 species spanning all major clades. GENESPACE[14] synteny analysis recovered large conserved blocks across the pangenome, yet revealed pervasive structural reconfiguration among species. The predominant differences reflect recurrent interchromosomal fissions and fusions, accompanied by frequent intrachromosomal inversions, whole-genome duplications, and translocations occur at a lower frequency. Chromosome counts range from $2n = 4$ (*R. tenuis*) to $2n = 36$ (*R. rugosa*), underscoring the remarkable karyotype plasticity characteristic of holocentric chromosomes (**Fig. 1a, Extended Data Table 1**). The preservation of extensive collinearity despite widespread re-wiring of chromosomal structure indicates that speciation within *Rhynchospora* has proceeded primarily via rearrangement (i.e., chromosomal speciation) rather than rapid erosion of synteny.

The comparative framework enables precise reconstruction of rearrangement histories. For example, a fission of *R. alba* chromosome 1 accounts for the origin of *R. rugosa* chromosomes 12 and 16 (**Fig. 1b**), whereas a fusion between chromosomes 1 and 5 in *R. radicans* is found in chromosome 2 of *R. pubera* (**Fig. 1c**). In both cases, the inferred breakpoints localise within centromeric *Tyba* arrays, implicating repeat-rich

holocentromeric regions in mediating both breakage and joining. These exemplars reflect a common pattern across the genus[6].

To systematically test this association, we applied a breakpoint-detection pipeline and mapped 60 independent junctions across the pangenome. Comparative analyses with closely related species and outgroups to infer directionality showed that 53% of breakpoints overlap *Tyba* arrays or other LTR retrotransposon families. Both end-to-end fusions and reciprocal fissions recur within the same repeat classes, consistent with a model for ectopic recombination and/or defective double-strand-break repair in *Tyba*-rich chromatin. These findings identify holocentromeric *Tyba* arrays as hotspots for structural remodelling and provide a mechanistic basis for the dramatic chromosome-number shifts observed across *Rhynchospora*.

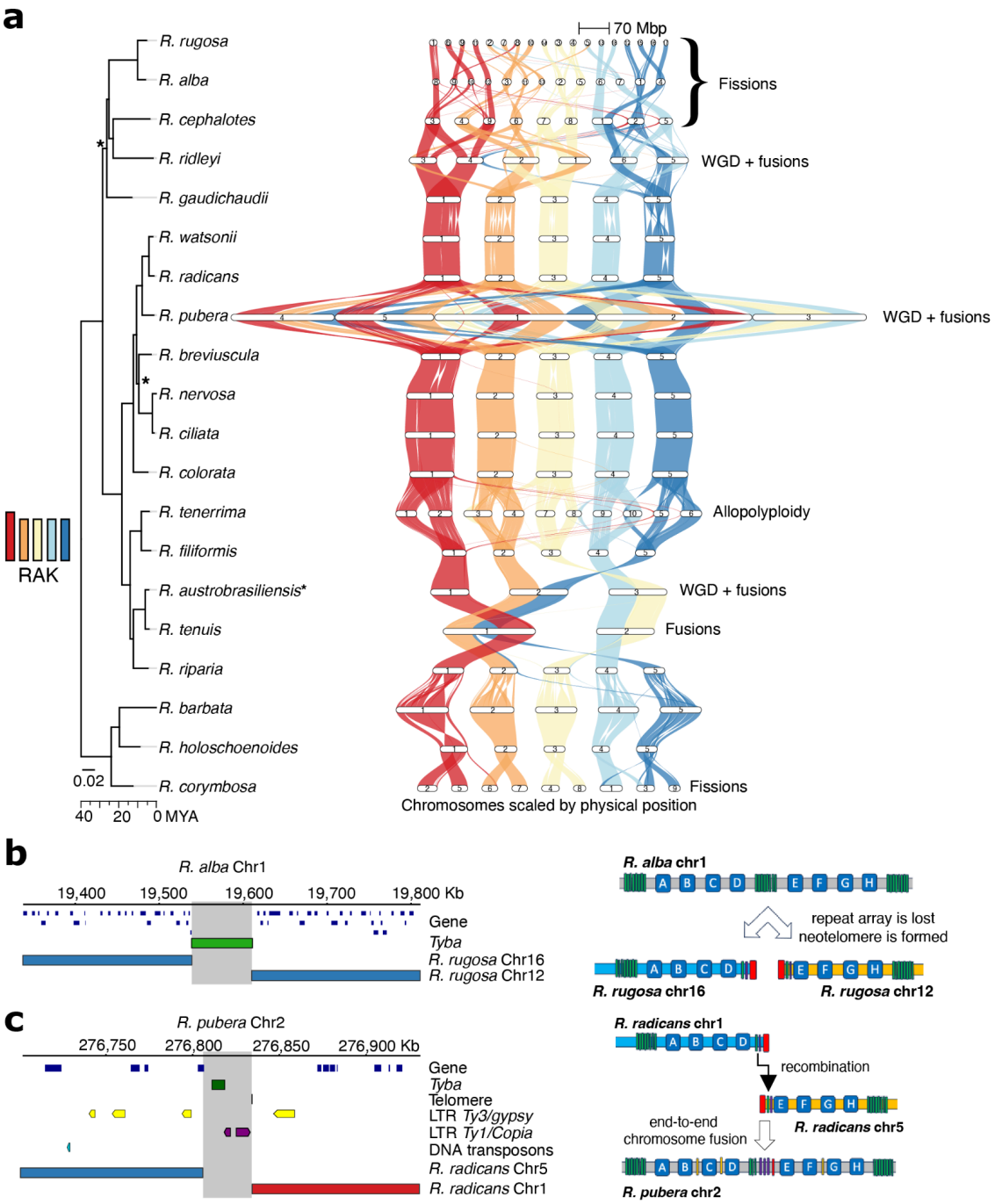


Fig. 1: Karyotype evolution in holocentric *Rhynchospora* revealed by a comparative pangenome. (**a**) Time-calibrated orthoMCL phylogeny of 20 *Rhynchospora* species with chromosome-level haploid assemblies, integrated with GENESPACE synteny. Conserved blocks are linked by colored ribbons (haploid comparison shown) based on the most parsimonious *Rhynchospora* ancestral karyotype (RAK). Extensive interspecific re-wiring is explained primarily by fissions, fusions and inversions; whole-genome duplications, allopolyploidy, and translocations occur at a lower frequency. Chromosome counts span $2n$ =

4 in *R. tenuis* to 2*n* = 36 in *R. rugosa*. Bootstrap values below 0.9 are indicated with asterisks. (**b**) Representative fission: *R. alba* chromosome 1 corresponds to *R. rugosa* chromosomes 12 and 16; breakpoints coincide with centromeric *Tyba* arrays. (**c**) Representative fusion: *R. radicans* chromosomes 1 and 5 found in the origin of *R. pubera* chromosome 2; the fusion junction maps within *Tyba* repeats. Across the pangenome, 60 rearrangement junctions were identified, 53% of which overlap *Tyba* arrays or other transposable elements, implicating holocentromeric repeats as hotspots for both fusion and fission.

Divergent recombination landscapes correlate with synapsis dynamics

Karyotype alterations are expected to influence CO frequency[15]. To assess how karyotype variation impacts meiotic recombination in holocentric plants, we generated genome-wide crossover (CO) landscapes for five additional *Rhynchospora* (*R. barbata*, *R. cephalotes*, *R. ciliata*, *R. colorata*, and *R. nervosa*) species using single-cell sequencing of pollen nuclei and integrated these with one from *R. pubera* F2 population (**Extended Data Fig. 1**) and our previously published *R. breviuscula* data[13].

Across the seven species, we detected 77,324 COs and found that recombination landscapes segregate into two classes (**Fig. 2a-b**): a strongly distal-biased pattern, with COs enriched at one or both chromosome ends (*R. cephalotes*, *R. breviuscula*, *R. pubera*), and an irregular pattern with COs distributed more evenly along chromosome arms (*R. barbata*, *R. ciliata*, *R. colorata*, *R. nervosa*). Notably, distal enrichment was observed irrespective of chromosome size and number, indicating that karyotype differences modulate CO number via assurance but do not dictate spatial patterning.

At the broad scale, CO location showed no correlation with canonical genomic features (**Fig. 2a-b**), the same as what has been shown in our previous study in *R. breviuscula*[13]. CO density did not track gene content, transposable elements, long terminal repeats, or *Tyba*/CENH3-marked holocentromeric arrays, nor with DNA methylation contexts (CpG, CHG, CHH) or chromatin marks (for example, H3K4me3, H3K9me2, H3K27me3). This contrasts with monocentric species, where centromere-proximal and heterochromatic regions typically suppress recombination[15], and indicates that CO patterning in *Rhynchospora* is largely decoupled from local sequence or epigenetic states.

Despite extensive conservation of synteny across species, syntenic intervals frequently display divergent recombination profiles (**Fig. 2c**), indicating that crossover patterning is not conserved with sequence collinearity. For example, the centre of chromosome 2 in *R. breviuscula* shows low recombination, whereas its syntenic counterpart resides at a chromosome end in *R. cephalotes* and exhibits a pronounced recombination peak, reinforcing that chromosome-scale dynamics rather than sequence collinearity shape recombination.

To investigate the mechanistic basis underlying the divergent recombination landscapes observed, we examined meiotic chromosome behaviour by high-resolution immunofluorescence of ZYP1, the synaptonemal complex transverse filament. Species with distal-biased CO landscapes (*R. breviuscula*, *R. cephalotes*, *R. pubera*) displayed clustered synapsis initiation consistent with a pronounced telomere bouquet, followed by synaptonemal complex elongation from chromosome ends. By contrast, species with irregular landscapes (*R. nervosa*, *R. ciliata*, *R. colorata*) exhibited a looser, more distributed onset of synapsis across chromosome arms (**Fig. 2d**). Among species lacking recombination maps due to low heterozygosity, *R. holoschoenoides* showed clustered synapsis, whereas *R. corymbosa* showed non-preferential early synapsis, suggesting corresponding divergence in recombination patterning. Together, these data establish a strong correlation between synapsis dynamics and CO spatial distributions, implicating early telomere-led pairing and subsequent synapsis progression as primary determinants of recombination landscapes in holocentric *Rhynchospora*, largely independent of underlying (epi)genomic features.

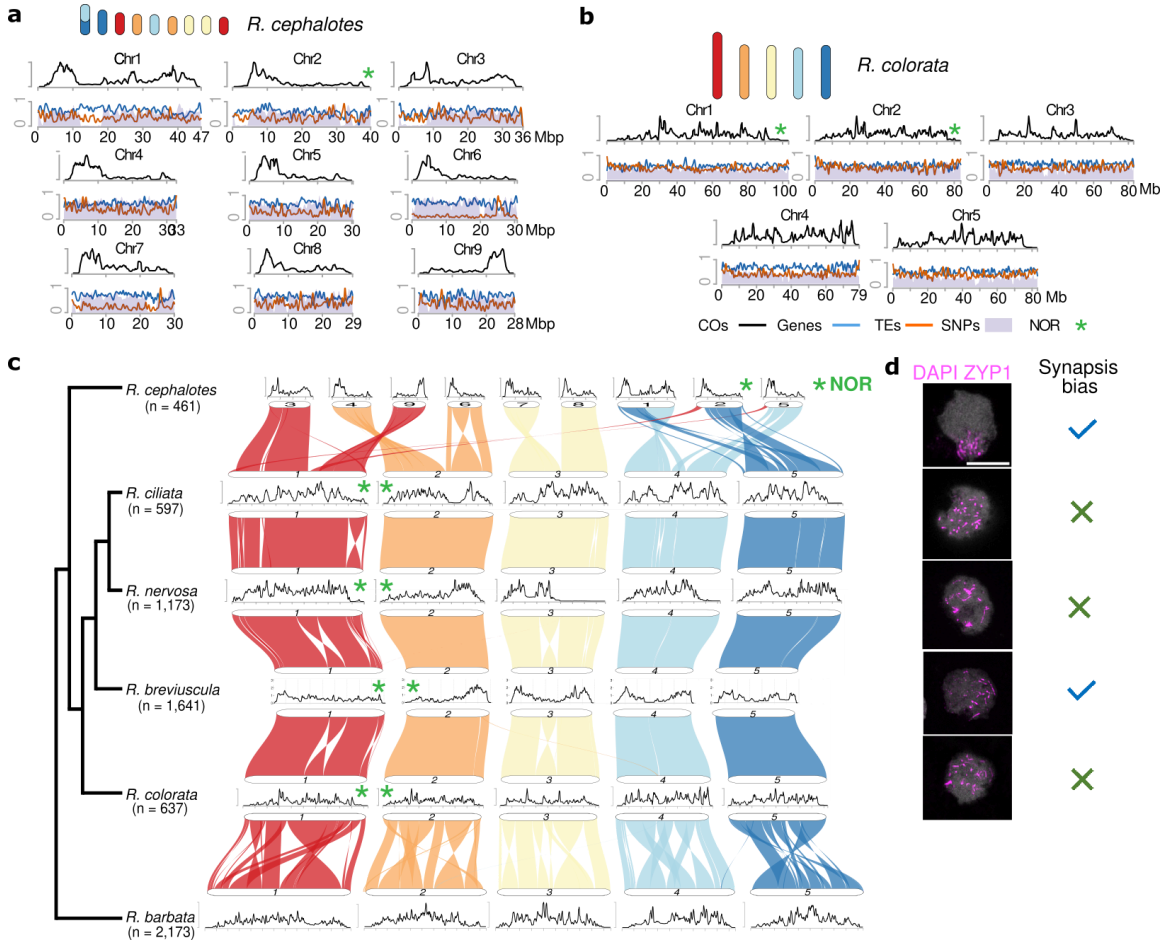


Fig 2: Divergent recombination landscapes correlate with synapsis dynamics in holocentric *Rhynchospora*. (**a**) Single-gamete sequencing across multiple species yields chromosome-scale recombination maps, revealing two CO landscape classes: distal-biased versus (**b**) irregular. (**c**) Despite high conservation of synteny across species, CO landscapes vary markedly, with syntenic regions exhibiting divergent recombination profiles. (**d**) Immunofluorescence of ZYP1 during early meiosis reveals differential synapsis dynamics: species with distal-biased COs show clustered, telomere-led synapsis initiation (bouquet), whereas species with irregular CO distributions show diffuse initiation across chromosomes. Across seven species, 77,324 COs were detected.

Chromosome size governs crossover frequency

We next asked how karyotype configuration influences the number and density of meiotic COs. At the species level, karyotypes with more chromosomes, often generated by fissions, exhibited higher total CO numbers, consistent with crossover assurance, which mandates at least one CO per chromosome to ensure proper segregation[16]. For example,

R. cephalotes (2*n* = 18), with an elevated chromosome count, shows increased genome-wide CO totals relative to species with fewer chromosomes, e.g., *R. breviuscula* (2*n* = 10). Across species, chromosome number was positively correlated with total CO number (**Fig. 3a**), indicating that changes in chromosome count directly scale recombination output. An exception is *R. pubera*, which displays fewer COs than expected given its chromosome count; its estimate is derived from cytological MLH1 foci (divided by two to match single-gamete maps) and likely reflects constraints imposed by its structurally complex karyotype composed of multiple ancestral syntenic blocks.

At the chromosome level, we quantified how physical length relates to recombination rate. Pooling all chromosomes from the seven *Rhynchospora* species with available CO maps, we observed a strong inverse relationship between chromosome length and CO density, measured as COs per megabase (linear regression on log10-transformed values, adjusted $R^2 = 0.9164$, $p < 2.2 \times 10^{-16}$; **Fig. 3b**). Thus, smaller chromosomes carry proportionally more crossovers per unit of DNA than larger chromosomes. This size dependence is consistent with structural constraints: assurance imposes at least one CO even on short chromosomes, elevating per-megabase rates, and shorter physical lengths likely increase the effective axial substrate per megabase available for double-strand break formation and CO designation. The relationship holds across diverse karyotypes, suggesting that chromosome size is a primary quantitative determinant of recombination rate in holocentric plants, largely overriding potential influences of sequence content or epigenetic landscape.

Marey maps further illustrate how recombination scales with chromosome size (**Fig. 3c**). In species with distal-biased landscapes, small chromosomes exhibit steep, near-linear cM–Mb relationships, with pronounced slopes at one or both ends. In contrast, large chromosomes exhibit extended central plateaus with steep terminal ramps, indicating concentration of COs near telomeres. Species with irregular landscapes display more variable, segmental slopes along chromosome arms, consistent with diffuse CO placement. Together, these trends show that karyotype evolution modulates recombination quantitatively through changes in chromosome number and size.

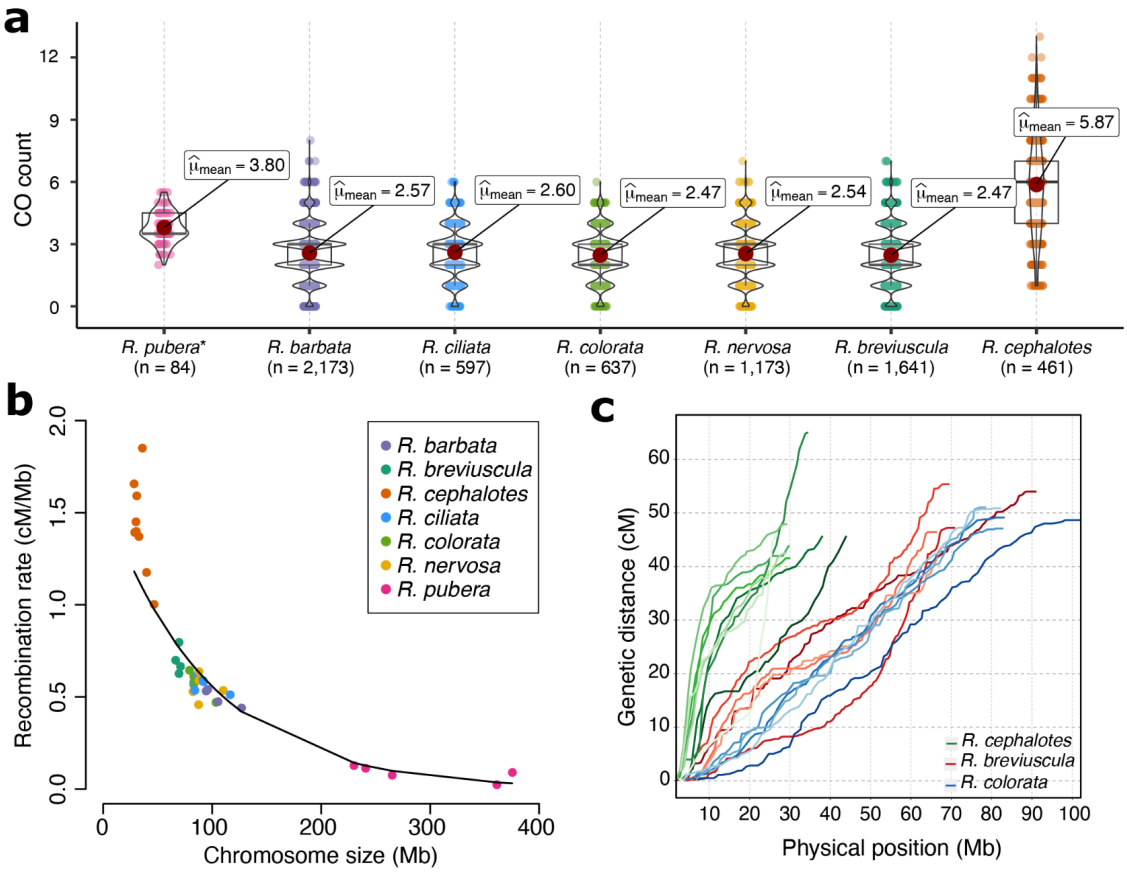


Fig. 3: Chromosome size and number govern crossover output in holocentric *Rhynchospora*. (**a**) CO count across seven different *Rhynchospora* species. Chromosome number directly influences total COs across the dataset, exemplified by *R. cephalotes* (higher chromosome count, elevated CO totals), consistent with crossover assurance. In contrast, *R. pubera* has a reduced CO number compared to what could be expected, due to the number of ancestral syntenic blocks found in its karyotype. *CO count in *R. pubera* was based on cytology count of MLH1 foci divided by two to better correlate with COs detected from pollen sequencing in the other six species. (**b**) Chromosome-level inverse relationship between physical length and recombination rate (centimorgan per Mb) across seven species, indicating that smaller chromosomes have higher CO density. (**c**) Marey map for three representative *Rhynchospora* species showing how the recombination rate differentially scales between small and large chromosomes.

Chromosome and loop size link axis length to crossover frequency

To uncover the mechanism underlying the inverse relationship between chromosome size and recombination rate, we quantified meiotic axis architecture and chromatin loop geometry across *Rhynchospora*. We first measured synaptonemal complex (SC) axes by

immunostaining ZYP1 on pachytene chromosomes and skeletonising the signal to extract axis length per bivalent and per chromosome (**Fig. 4a**). Across species and chromosome classes, shorter chromosomes exhibited greater axis length per megabase, whereas longer chromosomes showed reduced axis length per unit DNA. Axis metrics were predictive of recombination: absolute axis length correlated with the number of crossovers per chromosome, and axis length per megabase correlated positively with CO density (**Fig. 4b**). These relationships place the axial organisation of chromosomes as a primary determinant of quantitative CO output in holocentric meiosis.

We next estimated average chromatin loop size from somatic Hi-C contact decay profiles as a proxy for loop geometry (**Fig. 4b**). Chromosome size showed a clear positive correlation with loop length: larger chromosomes tended to be folded into fewer, longer loops, while smaller chromosomes were organised into shorter loops (**Fig. 4c**). Although Hi-C was generated from non-meiotic tissue, we reason that relative differences in loop architecture across chromosomes are likely conserved into meiosis, allowing loop size to serve as a meaningful proxy for axial organisation. Consistent with this view, loop length was inversely related to CO density, whereas axis length per megabase was positively related to CO density (**Fig. 4**), linking loop–axis geometry to crossover placement and frequency.

Together, our data support a structural model in which chromosome size sets loop geometry, which in turn defines the axial substrate accessible to the recombination machinery. This agrees with the proposed role of *Tyba* holocentromeric repeats on chromatin loop sizes as inferred from polymer chromatin modelling[10] and our estimated Hi-C contact profiles (**Fig. 4c**). Small chromosomes, with short loops and extended axis per megabase, present more opportunities for double-strand break formation and CO designation, elevating CO density. Large chromosomes, with long loops and shorter relative axes, restrict access and suppress COs. This loop–axis mechanism integrates with crossover assurance, scaling total CO number with chromosome count, to explain how karyotype evolution quantitatively and qualitatively remodels recombination landscapes in holocentric *Rhynchospora*.

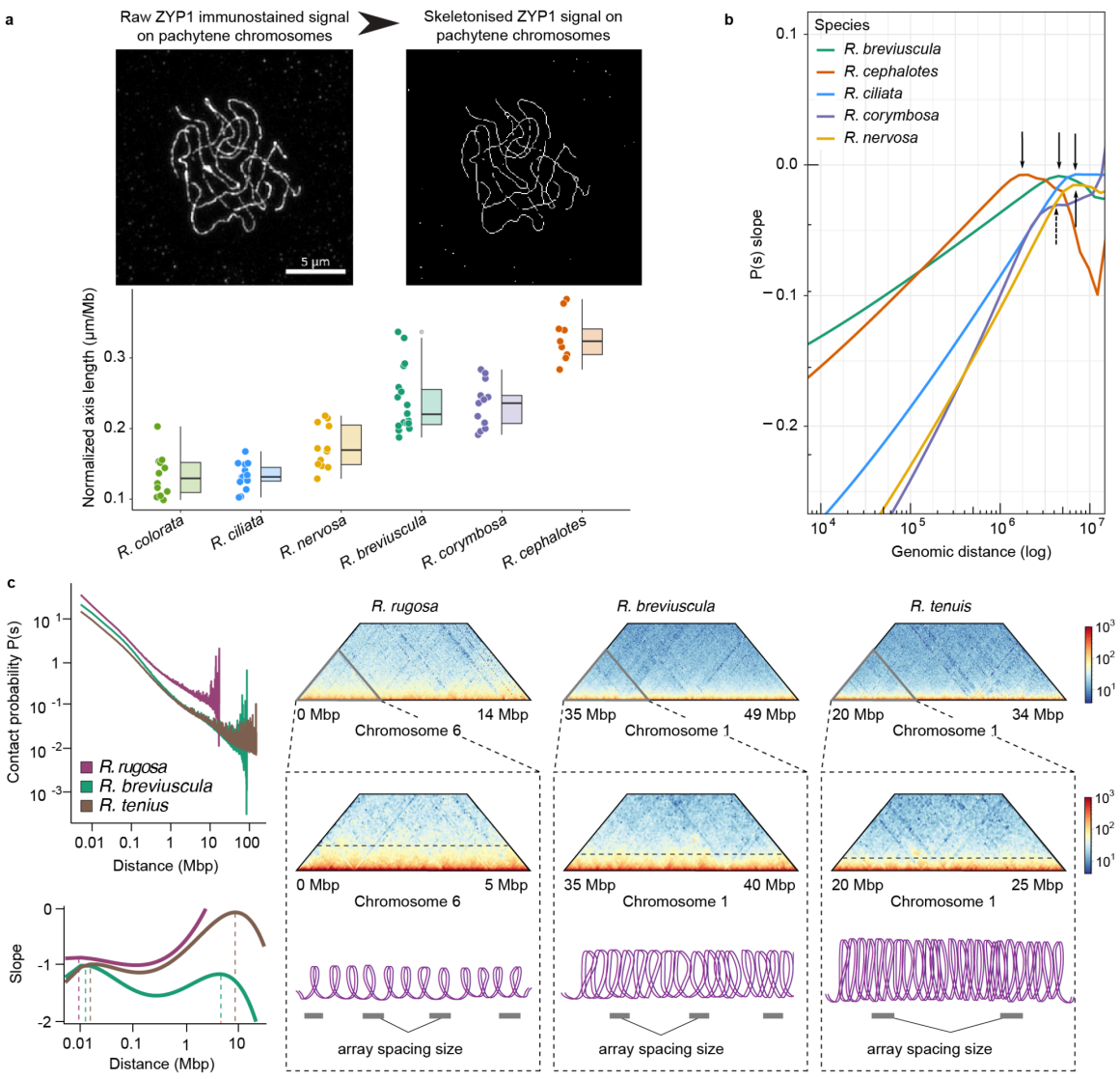


Fig 4: Chromatin loop geometry links chromosome size to axis length and crossover frequency in holocentric *Rhynchospora*. (**a**) Raw ZYP1 immunofluorescence on pachytene chromosomes used to measure synaptonemal complex axes (left). Skeletonised ZYP1 signal illustrating axis tracing and quantification per chromosome (right). Quantification of axis length per chromosome and per megabase across species (bottom). (**b**) Comparative analysis of the loop size measurement across five *Rhynchospora* species with different chromosome sizes. (**c**) Comparative analysis of chromatin conformation from somatic Hi-C data from three representative *Rhynchospora* species with different chromosome sizes to illustrate the effect of chromosome size on loop size architecture, hinting at a possible role of the spacing dynamics of holocentromeric *Tyba* repeats in overall axis length. (**e**) Positive relationship between axis length per megabase and crossover density (COs per Mb). (**f**) Negative relationship between average loop length and crossover density. (**g**) Positive correlation between chromosome size and average chromatin

loop length inferred from somatic Hi-C contact decay profiles, indicating that larger chromosomes are organised into fewer, longer loops and smaller chromosomes into shorter loops. Together, panels d–g support a model in which chromosome size determines loop geometry, which sets axis length per megabase and thereby modulates crossover frequency.

Discussion

Our comparative framework across 20 *Rhynchospora* species shows that holocentric chromosome architecture permits exceptional tolerance to fissions and fusions. Large syntenic blocks are preserved despite extensive re-wiring of chromosome structure, indicating that speciation proceeds primarily through rearrangement rather than loss of collinearity[8,17]. We identify *Tyba* satellite arrays and transposable elements as recurrent hotspots for both end-to-end fusion and reciprocal fission, positioning repeat-rich holocentromeric regions as fragile sites that both catalyse structural change and are buffered from lethality by distributed kinetochore activity. Our data thus also support *Tyba* repeats not merely as centromeric building blocks but as dynamic elements driving genome restructuring[10]. These combined features likely underpin the striking chromosome-number volatility of *Rhynchospora,* potentially linking to a similar pattern in other holocentric groups[18,19].

This structural plasticity has clear consequences for meiotic recombination. Single-gamete genotyping reveals two CO landscape classes across the genus, strongly distal-biased versus irregular distribution, and these are largely decoupled from gene density, transposable elements, DNA methylation, or histone marks[13]. Instead, cytology links CO patterning to synapsis dynamics: species with distal-biased landscapes show clustered synapsis initiation, whereas species with irregular landscapes show diffuse synapsis initiation (**Fig. 5a**). When synapsis initiates and persists at telomeres, then moves inward, distal chromosomal regions have more time in a recombination-competent state, meaning a higher probability to form crossovers. Even in species where recombination maps were not recoverable due to low heterozygosity (e.g., *R. corymbosa* and *R. holoschoenoides*), synapsis patterns (for example, clustered versus diffuse) align with the expected recombination biases. These observations reinforce a mechanistic role

for pairing and synapsis progression, not sequence or epigenetic features, in shaping CO distributions in holocentric *Rhynchospora* species, supporting our recent model[13].

Karyotype configuration also modulates the quantitative output of recombination. Across species, total CO number scales positively with chromosome number, as expected from crossover assurance ensuring at least one CO per chromosome[20]. Fissions thus elevate genome-wide CO counts, whereas fusions reduce them, without necessarily fixing the spatial pattern of COs (**Fig. 5b**). At the chromosome level, we observe a strong inverse relationship between chromosome length and CO density: smaller chromosomes carry more COs per megabase than larger chromosomes. This size dependence is consistent with structural constraints that elevate per-megabase CO rates on short chromosomes.

We explain the quantitative trends in recombination by linking chromosome size to loop–axis organisation. While absolute loop sizes likely differ between mitotic and meiotic contexts[21], relative differences are expected to persist[22], allowing loop length to serve as a meaningful proxy for recombination potential. In meiosis, chromosome axis length and chromatin loop size are inversely correlated[23–27]. We propose that shorter chromosomes, with smaller loops, have greater axis length per megabase and thus increased accessibility for double-strand break formation and subsequent higher CO rate. Larger chromosomes, by contrast, are constrained by longer loops and shorter relative axes, limiting DSB and CO opportunity[23,28]. These structural constraints, coupled with crossover interference[29,30], likely reinforce the observed chromosome size-dependent variation in CO frequency in *Rhynchospora* (**Fig. 5c**). In combination with synapsis mode, clustered versus dispersed initiation, these loop–axis principles account for both the quantitative scaling and spatial diversification of recombination across *Rhynchospora*.

Together, our findings support a mechanistic model in which *Tyba*-mediated fissions and fusions remodel the physical structure of chromosomes, altering loop architecture[10], axis length, and synapsis dynamics, and thereby rewire recombination landscapes. This structural-functional feedback loop demonstrates how chromosomal architecture can directly influence genetic diversity. By decoupling recombination from traditional sequence or epigenetic cues, holocentric genomes may access greater evolvability, particularly in lineages with frequent karyotypic change. Our study provides a unifying

framework linking chromosome structure to the evolutionary plasticity of recombination. These insights have broad implications for understanding how genome architecture shapes genetic diversity, adaptation, and speciation across eukaryotes.

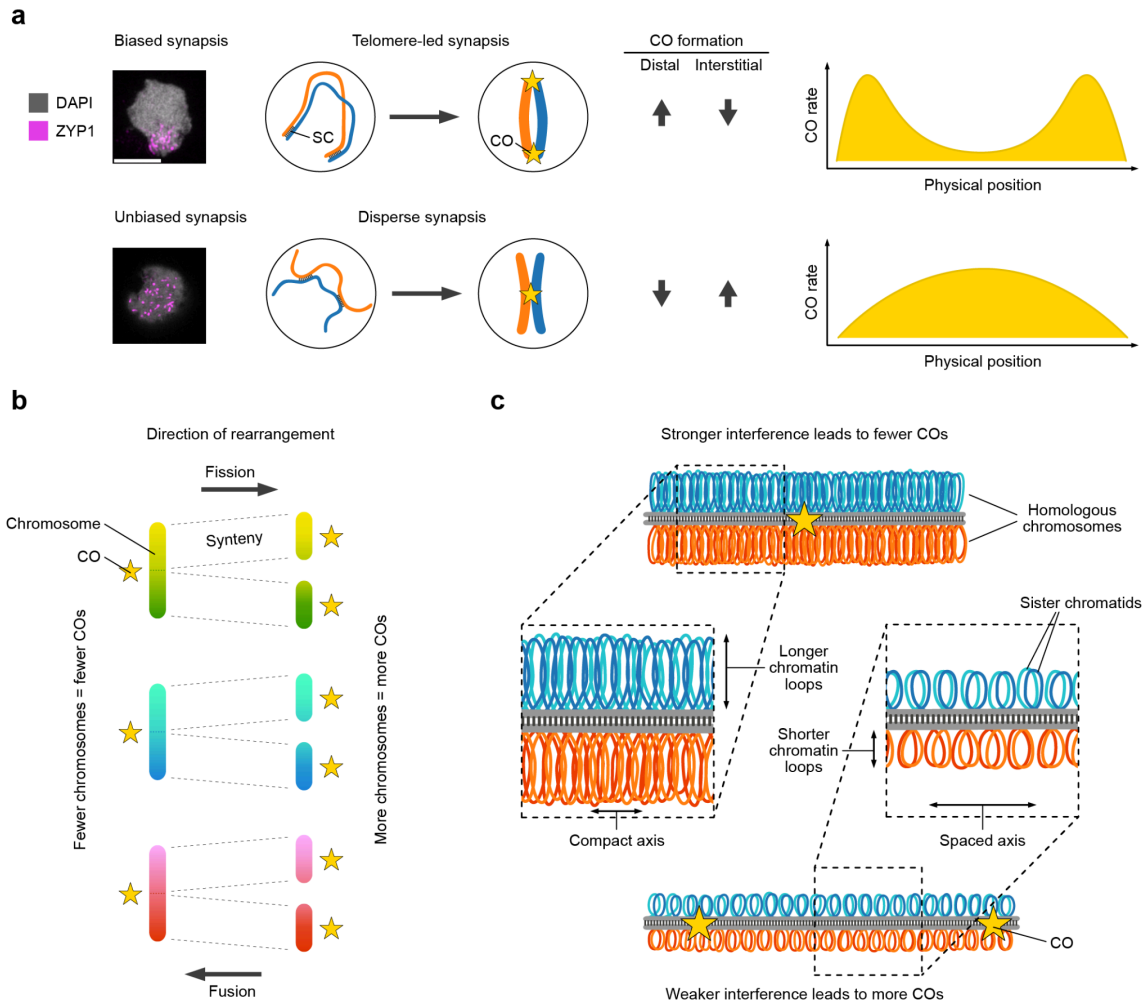


Fig. 5: A structural–mechanistic framework linking holocentric karyotype evolution to crossover patterning in *Rhynchospora*. (**a**) Synapsis modes determine CO spatial patterning: clustered, telomere-led synapsis (biased) generates distal CO enrichment and terminal ramps in CO rate along physical position; dispersed synapsis (unbiased) yields more interstitial or irregular CO placement. (**b**) Effects of chromosome number on total crossovers: fission increases chromosome number and elevates total COs via crossover assurance; fusion lowers chromosome number and reduces total COs (schematic chromosomes and SC, numbers illustrate expected differences in CO counts). (**c**) Loop–axis geometry sets CO capacity and interference: large chromosomes with longer loops and a compact axis show stronger interference and fewer COs; small chromosomes with shorter loops and an expanded axis show weaker interference and

more COs. Together, these panels synthesise how chromosome fissions and fusions, synapsis dynamics, and loop–axis organisation jointly control both the quantitative output and spatial distribution of crossovers in holocentric *Rhynchospora*.

Methods

Plant Material

Plants from naturally occurring populations growing in Brazil were collected and further cultivated under controlled greenhouse conditions (16h daylight, 26 °C, >70% humidity).

DNA isolation

High-molecular-weight DNA was isolated from 1.5 g of material with a NucleoBond HMW DNA kit (Macherey Nagel). Quality was assessed with a FEMTO-pulse device (Agilent), and quantity was measured with a Quantus fluorometer (Promega).

PacBio

A HiFi library was then prepared according to the "*Procedure & Checklist - Preparing HiFi SMRTbell® Libraries using SMRTbell Express Template Prep Kit 2.0*" manual, with an initial DNA fragmentation by g-Tubes (Covaris) and final library size binning into defined fractions by SageELF (Sage Science). Size distribution was again controlled by FEMTO-pulse (Agilent). Size-selected libraries were then sequenced on a Sequel II device with Binding kit 2.0 and Sequel II Sequencing Kit 2.0 for 30 h (*Pacific Biosciences*).

Arima HiC

Plant tissues were cross-linked with 1% formaldehyde for 30 minutes at room temperature, and the reaction was quenched with 125 mM glycine for 10 minutes. Subsequently, the tissues were ground using a TissueLyser at a frequency of 30 Hz for 3 minutes. Nuclei extraction was performed using the CelLytic PN Plant Nuclei Isolation/Extraction Kit (Sigma-Aldrich, Burlington, MA, USA) according to the

manufacturer's protocol. Hi-C libraries were prepared using the Arima High Coverage Hi-C Kit (Arima Genomics, A410110, Carlsbad, CA, USA) following the manufacturer's instructions, and were then sequenced paired-end (2 × 150 bp) on a NextSeq 2000 instrument (Illumina, San Diego, CA, USA).

Omni-C

For each species, a single chromatin-capture library was prepared from 0.5 g of fresh-weight material input. All treatments were administered according to the recommendations provided by the kit vendor for plants (Omni-C, Dovetail). As a final step, an Illumina-compatible library was prepared (Dovetail) and paired-end 2 x 150 bp deep-sequenced on a HiSeq 3000 (Illumina) device.

Genome assemblies

PacBio HiFi reads were generated for 20 *Rhynchospora* species and assembled with Hifiasm (v0.25.0) with default settings. Hi-C sequences were included when running HiFiasm when available. Scaffolding was performed by aligning Hi-C reads to contigs from HiFiasm using Juicer and 3D-DNA. Alternatively, phased scaffolding was performed using HapHiC.

Identification of break intervals in chromosome fusion and fission events

Whole genome alignments between different *Rhynchospora* species were conducted by minimap2 with '-cx asm10'. In both chromosome fusion and fission events, the larger chromosomes were taken as the target, and the smaller chromosomes were the query. Query chromosomes that account for no less than 5% of the target chromosome length were the main components of the target. Each query chromosome IDs were assigned to an integer. A sliding window with a 50kb window size and a 10kb step size was applied to each target chromosome to derive the query chromosome in this window by calculating the mean of the integer ID of the aligned query. The smoothed query integer ID was then input to a median filter for further denoising. The rough break intervals were defined at windows where consecutive non-zero first derivatives were found. To refine

the break intervals by the last aligned loci of the query sequence, a binary signal partition approach was adopted. The complete pipeline is available on GitHub: https://github.com/Raina-M/BreakID

Single-pollen nuclei library preparation

Pollen grains were released by mechanical disruption from mature flowers or anthers of the different *Rhynchospora* species and collected into Woody Pollen Buffer (scWPB: 200mM Tris-HCl; 4mM $MgCl_2$; 2mM EDTA; 86mM NaCl; 10mM $Na_2S_2O_5$; 250mM Sucrose; 0.5mM Spermine; 0.5mM Spermidine; 1% PVP-10). To remove large debris, the obtained solution was filtered through a 50µm CellTrics cell strainer. The pollen sample was then pelleted by centrifugation and stored at -70ºC until further use.

The frozen pollen samples were thawed on ice and collected into a 5µm CellTrics cell strainer. Pollen grains were macerated inside the cell strainer, releasing the pollen nuclei. The Pollen nuclei were recovered from the macerated pollen by adding scWPB with some additives (scWPC; 5mM DTT; 1%BSA; 0.2 U/µL Protector RNase Inhibitor). Pollen nuclei solution was stained with DAPI (1µg/µl) and used for Fluorescence-activated cell sorting (FACS) by employing BD FACSAriaIII Fusion Flow Cytometer with a 70µm nozzle. Nuclei were dispensed into a 96-well plate containing Collection Buffer (1xPBS, 1%BSA, and 0.2 U/µL Protector RNase Inhibitor). The quality and number of nuclei were evaluated with the LUNA-FX7 Automated Cell Counter. Nuclei solutions were used for library preparation using Chromium Next GEM Single Cell ATAC or Chromium Next GEM Single Cell 5′ Kits following the manufacturer's instructions. Libraries sequenced at BGI Genomics following Chromium 10x Kit recommendations.

Single-cell analysis

Crossovers were detected using the similar pipeline reported in Catellani and Zhang, et al. 2024[13]. An updated pipeline description was available on the GitHub page: https://github.com/Raina-M/detectCO_by_scRNAseq.

Markers for R. pubera F2 genotyping

The *R. pubera* with 10 chromosomes (Rp10) was crossed with *R. pubera* with 12 chromosomes (Rp12) and generated an F1 plant with 11 chromosomes (Rp11). Only the Rp10 genome was available[6]. Rp12 with a heterozygous genome and the F1 Rp11 both have whole-genome DNA sequencing short reads. We mapped the Rp12 short reads to the Rp10 genome (bwa mem, default parameters) to find the loci that differ from Rp10. There are three types of loci: 1) Only one alternate allele was found with zero reference allele (the same as Rp10) coverage, meaning two haplotypes of Rp12 at this locus have the same allele that differ from Rp10; 2) Only one alternate allele was found but the allele frequency of reference versus alternate is 1:1, meaning one haplotype of Rp12 is the same as Rp10, the other is different; 3) Two alternate alleles were found with zero reference allele coverage, meaning two haplotypes of Rp12 at this locus have the different alleles and both differ from Rp10. All loci in Rp12 with distinguishable alleles to Rp10 were included in these three categories. Then we mapped the short reads of F1/Rp11 to Rp10 and called SNPs with bcftools mpileup (--min-MQ 20) and bcftools call. The intersection of SNPs from F1/Rp11 with all alternate alleles found in Rp12 was defined as the genotyping marker for F1 crossover detection.

Crossover detection in R. pubera F1 offspring with whole-genome sequencing

F2 short reads were aligned to the Rp10 genome with bowtie2, and SNPs were called with bcftools mpileup (--min-MQ 1) and bcftools call. All SNPs were intersected with the genotyping markers defined previously on Rp11, and then the allele frequency of alternate alleles was converted to a [-1, 1] scale, with 1 as the reference and -1 as the alternate alleles. Each chromosome has a one-dimensional allele frequency array (AF array). The AF array was first smoothed with a median filter (median size 6000001). The smoothed signal was fed into a Gaussian low-pass filter with a Gaussian kernel size of four times sigma smaller than 0 and four times sigma greater than 0. Sigma denotes the standard deviation of the Gaussian distribution and is defined as 1000000*array size/chromosome size. After denoising by the Gaussian low-pass filter, the AF array gradient was calculated by the first derivative. Each group of non-zero gradients indicate a potential haplotype conversion location, and gradients greater than 2e-04 were defined as crossovers.

Hi-C contact analysis

The Hi-C heatmap was first generated by Juicer (v2.0). The output inter_30.hic was used to extract pairwise bins with certain distances using juicer_tools.jar with bin size 5kbp and normalisation method "VC_SQRT". Hi-C heatmaps were plotted with HiCExplorer's subfunction hicPlotMatrix.

Data availability

Genome assemblies and raw sequencing data have been deposited in the European Nucleotide Archive under BioProject PRJEB103769.

Acknowledgements

We thank … Funding was provided by the ERC (HoloRECOMB) and DFG …

This study was funded by the Max Planck Society (core funding to A.M.), the German Research Foundation (MA 9363/2-1 and MA 9363/3-1), and the European Union (European Research Council Starting Grant, HoloRECOMB, grant no. 101114879 to A.M.). The DFG also funded this work under Germany's Excellence Strategy—EXC 493 2048/1–390686111 (to K.S. and A.M.). M.Z. is financially supported by the DFG (grant no. MA 9363/2-1).

Author contributions

A.M. conceived the study. M.Z. performed genome assembly and bioinformatic analyses. S.S. performed all the single-cell pollen sequencing experiments. M.C. performed the cytological analysis of meiosis. All authors contributed to writing.

Competing interests

The authors declare no competing interests.

References

1. Barra, V. & Fachinetti, D. The dark side of centromeres: types, causes and consequences of structural abnormalities implicating centromeric DNA. *Nat. Commun.* **9**, (2018).
2. Steckenborn, S. & Marques, A. Centromere diversity and its evolutionary impacts on plant

karyotypes and plant reproduction. *New Phytol.* **245**, 1879–1886 (2025).

3. Van de Peer, Y., Mizrachi, E. & Marchal, K. The evolutionary significance of polyploidy. *Nat. Rev. Genet.* **18**, 411–424 (2017).
4. Schubert, I. & Lysak, M. A. Interpretation of karyotype evolution should consider chromosome structural constraints. *Trends Genet.* **27**, 207–216 (2011).
5. Marques, A. *et al.* Holocentromeres in Rhynchospora are associated with genome-wide centromere-specific repeat arrays interspersed among euchromatin. **112**, 13633–13638 (2015).
6. Hofstatter, P. G. *et al.* Repeat-based holocentromeres influence genome architecture and karyotype evolution. *Cell* **185**, 3153–3168 (2022).
7. Rieseberg, L. H. Chromosomal rearrangements and speciation. *Trends Ecol. Evol.* **16**, 351–358 (2001).
8. Lucek, K. *et al.* The Impact of Chromosomal Rearrangements in Speciation: From Micro- to Macroevolution. *Cold Spring Harb. Perspect. Biol.* **15**, a041447 (2023).
9. Jankowska, M. *et al.* Holokinetic centromeres and efficient telomere healing enable rapid karyotype evolution. *Chromosoma* **124**, 519–528 (2015).
10. Wlodzimierz, P. Pangenome analysis reveals modular satellite-driven holocentromere evolution in Rhynchospora. *submitted* (2025).
11. Mata-Sucre, Y. *et al.* Oligo-barcode illuminates holocentric karyotype evolution in Rhynchospora (Cyperaceae). *Front. Plant Sci.* **15**, 1330927 (2024).
12. Burchardt, P. *et al.* Holocentric Karyotype Evolution in Rhynchospora Is Marked by Intense Numerical, Structural, and Genome Size Changes. *Front. Plant Sci.* **11**, (2020).
13. Castellani, M. *et al.* Meiotic recombination dynamics in plants with repeat-based holocentromeres shed light on the primary drivers of crossover patterning. *Nat. Plants* **10**, 423–438 (2024).
14. Lovell, J. T. *et al.* GENESPACE tracks regions of interest and gene copy number variation

across multiple genomes. *eLife* **11**, e78526 (2022).

15. Brazier, T. & Glémin, S. Diversity and determinants of recombination landscapes in flowering plants. *PLOS Genet.* **18**, e1010141 (2022).
16. Mercier, R., Mézard, C., Jenczewski, E., Macaisne, N. & Grelon, M. The molecular biology of meiosis in plants. *Annu Rev Plant Biol* **66**, 297–327 (2015).
17. Lucek, K., Augustijnen, H. & Escudero, M. A holocentric twist to chromosomal speciation? *Trends Ecol. Evol.* **37**, 655–662 (2022).
18. Wright, C. J., Stevens, L., Mackintosh, A., Lawniczak, M. & Blaxter, M. Comparative genomics reveals the dynamics of chromosome evolution in Lepidoptera. *Nat. Ecol. Evol.* **8**, 777–790 (2024).
19. Márquez‐Corro, J. I. *et al.* Macroevolutionary insights into sedges (*Carex* : Cyperaceae): The effects of rapid chromosome number evolution on lineage diversification. *J. Syst. Evol.* **59**, 776–790 (2021).
20. Saito, T. T. & Colaiácovo, M. P. Regulation of Crossover Frequency and Distribution during Meiotic Recombination. *Cold Spring Harb Symp Quant Biol* **82**, 223–234 (2017).
21. Zuo, W. *et al.* Stage-resolved Hi-C analyses reveal meiotic chromosome organizational features influencing homolog alignment. *Nat. Commun.* **12**, 5827 (2021).
22. Oomen, M. E. *et al.* Mitotic chromosomes harbor cell type- and species-specific structural features within a universal loop array conformation. *Genome Res.* **35**, 1733–1744 (2025).
23. Kleckner, N., Storlazzi, A. & Zickler, D. Coordinate variation in meiotic pachytene SC length and total crossover/chiasma frequency under conditions of constant DNA length. *Trends Genet.* **19**, 623–628 (2003).
24. Zickler, D. & Kleckner, N. Recombination, Pairing, and Synapsis of Homologs during Meiosis. *Cold Spring Harb. Perspect. Biol.* **7**, a016626 (2015).
25. Kleckner, N. *et al.* A mechanical basis for chromosome function. *Proc. Natl. Acad. Sci.* **101**, 12592–12597 (2004).

26. Novak, I. *et al.* Cohesin Smc1beta determines meiotic chromatin axis loop organization. *J. Cell Biol.* **180**, 83–90 (2008).
27. Tease, C. & Hultén, M. A. Inter-sex variation in synaptonemal complex lengths largely determine the different recombination rates in male and female germ cells. *Cytogenet. Genome Res.* **107**, 208–215 (2004).
28. Kleckner, N. Chiasma formation: chromatin/axis interplay and the role(s) of the synaptonemal complex. *Chromosoma* **115**, 175–194 (2006).
29. Capilla-Pérez, L. *et al.* The synaptonemal complex imposes crossover interference and heterochiasmy in Arabidopsis. *Proc Natl Acad Sci U A* **118**, (2021).
30. Von Diezmann, L. & Rog, O. Let's get physical – mechanisms of crossover interference. *J. Cell Sci.* **134**, (2021).

Extended Data Table 1. 20 Rhynnchospora species karyotype and genome size. If only the haploid genome size is shown, the corresponding species usually has a highly homozygous genome, making the phasing impossible. Genome size here includes only pseudo-chromosomes.

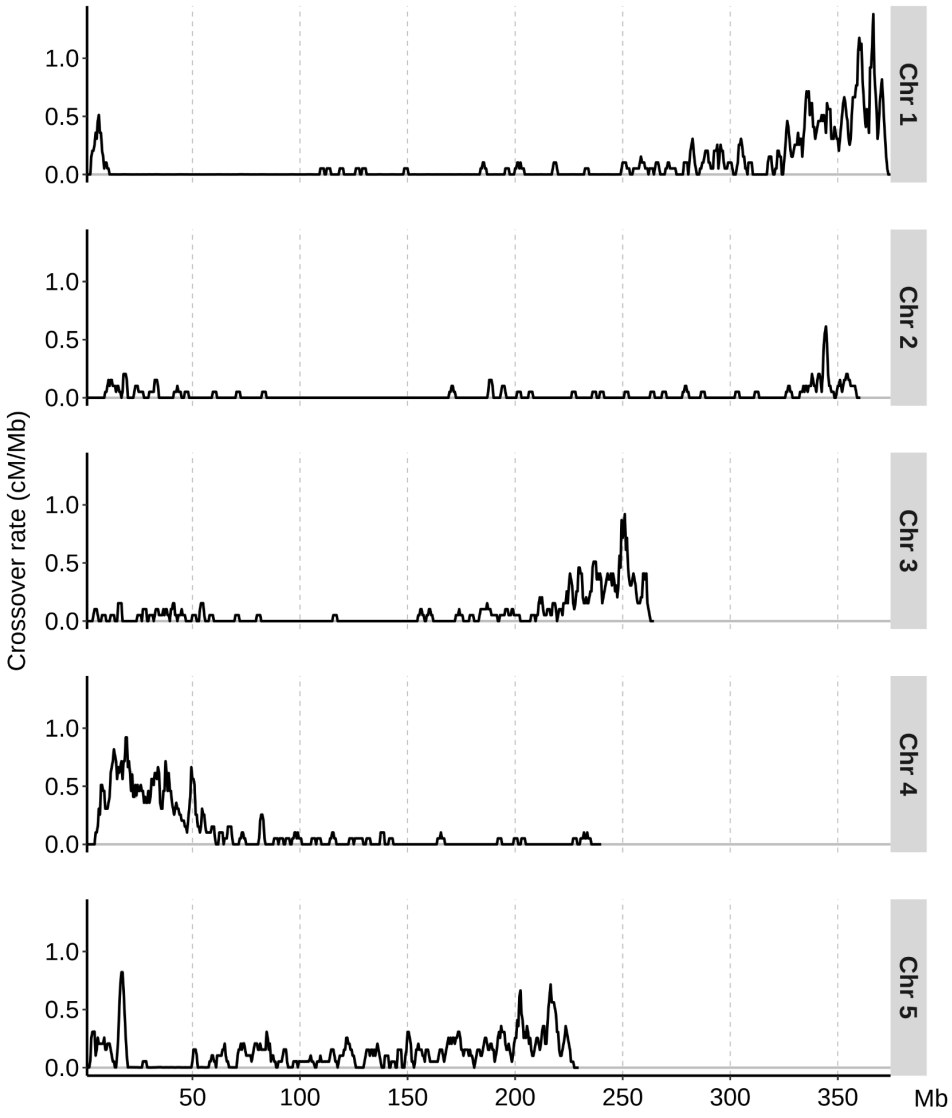
Species	Karyotype	Assembly type	Genome size (bp)
R. rugosa	2*n*=2*x*=36	Haploid	235,231,465
R. alba	2*n*=2*x*=26	Haploid	235,949,887
R. cephalotes	2*n*=2*x*=18	Haplotype 1	303,606,085
		Haplotype 2	303,631,510
R. ridleyi	2*n*=2*x*=12	Haplotype 1	423,345,332
		Haplotype 2	426,461,118
R. gaudichaudii	2*n*=3*x*=15	Haploid	337,173,723
R. watsonii	2*n*=2*x*=10	Haploid	357,361,827
R. radicans	2*n*=2*x*=10	Haploid	350,330,963
R. pubera	2*n*=2*x*=10	Haploid	1,694,878,336
R. breviuscula	2*n*=2*x*=10	Haplotype 1	368,174,147
		Haplotype 2	370,478,156
R. nervosa	2*n*=2*x*=10	Haplotype 1	450,937,689
		Haplotype 2	455,796,437
R. ciliata	2*n*=2*x*=10	Haplotype 1	462,117,430
		Haplotype 2	441,783,790
R. colorata	2*n*=2*x*=10	Haplotype 1	432,477,466
		Haplotype 2	430,351,217
R. tenerrima	2*n*=4*x*=20	Haplotype 1	232,714,659
		Haplotype 2	242,448,059
R. filiformis	2*n*=2*x*=10	Haplotype 1	239,492,786
		Haplotype 2	233,767,920

Species	Karyotype	Assembly type	Genome size
*R. austrobrasiliensis**	2*n*=6*x*=18	Pseudohaplotype 1	347,896,131
		Pseudohaplotype 2	402,874,894
		Pseudohaplotype 3	359,096,123
		Pseudohaplotype 4	352,179,589
		Pseudohaplotype 5	354,838,715
		Pseudohaplotype 6	348,573,314
R. tenuis#	2*n*=2*x*=4	Haplotype 1	348,158,620
		Haplotype 2	354,512,391
R. riparia	2*n*=2*x*=10	Haplotype 1	297,062,901
		Haplotype 2	286,276,137
R. barbata	2*n*=2*x*=10	Haplotype 1	514,099,728
		Haplotype 2	514,099,728
R. holoschoenoides	2*n*=2*x*=10	Haplotype 1	254,695,320
		Haplotype 2	255,930,137
R. corymbosa	2*n*=2*x*=18	Haplotype 1	348,126,674
		Haplotype 2	338,896,338

* *R. austrobrasiliensis* has an autohexaploid genome, so six haplotypes are barely distinguishable solely based on the genomic method. The pseudo-haplotype shown here is based on the *k*-mer distance clustering from the paper: *Sex without crossovers mimics clonal reproduction in the holocentric plant Rhynchospora tenuis*.
# *R. tenuis* has 9 accessions (18 haplotypes assembled); here, only the accession REC is shown. The other genome assemblies were shown in the above paper as well.

Extended Data Table 2: karyotype changes in 20 *Rhyncospora* species and break point sources. The paired original and derived species do not indicate the real evolutionary event. The one closest to the ancestral karyotype is regarded as the original karyotype.

Original	Derived	Fission	Fusion
R. alba	*R. rugosa*	11	4
R. cephalotes	*R. alba*	13	7
R. gaudichaudii	*R. cephalotes*	5	1
R. radicans	*R. pubera*	0	7
R. filiformis	*R. austrobrasiliensis*	0	2
R. austrobrasiliensis	*R. tenuis*	0	1
R. holoschoenoides	*R. corymbosa*	7	3



Extended Data Fig. 1. Recombination landscape of *R. pubera* through F2 whole genome sequencing. We crossed a homozygous *R. pubera* with 10 chromosomes and a heterozygous *R. pubera* with 12 chromosomes and generated an F1 *R. pubera* with 11 chromosomes. We got XX F2 after selfing this F1 individual. The recombination landscape was plotted with 489 plants and smoothed with a 2Mbp window size and a 500kbp step size.

2.3 Sex without crossovers mimics clonal reproduction in the holocentric plant *Rhynchospora tenuis*

Meng Zhang, Marco Castellani, Stefan Steckenborn, Maciej Majka, Georgios Tsipas, Thiago Nascimento, Ulla Neumann, Gokilavani Thangavel, Laura A. Robledillo, Thomas Lux, Lorraine Deberón, Ursula Pfordt, José A. Campoy, Athul Vijayan, Ton Timmers, Nafiseh Sargheini, Magdalena Marek, Hequan Sun, Paulo G. Hofstatter, Bruno Huettel, Steven Dreissig, Klaus F. X. Mayer, Duarte D. Figueiredo, André L. L. Vanzela, Korbinian Schneeberger, André Marques#

Correspondence

Author Contributions:

A.M. conceived the research and supervised the project. M.Z., P.G.H. and A.M. performed the genome assembly, scaffolding and genomics. M.Z., J.C., H.S., M.C., and S.S. designed and performed the single-cell experiments and processed the data. K.S. assisted the single-cell analysis. M.C. performed the immunostaining experiments. L.D. and M.C. performed the functional screening of meiotic genes. T.N. and M.Majka performed FISH experiments. M.Majka performed pollen-FISH genotyping experiments. G.T. and M.Z. performed the ChIP–seq analysis. M.Marek and S.S. performed the flow cytometry measurements and nuclei sorting. N.S., M.Marek and B.H. performed all nuclear sequencing libraries. M.Z., L.A.R., A.L.L.V. and A.M. performed the centromere and repeat characterisation. A.L.L.V. collected and provided the plant material. U.P. and M.Majka performed the plant crosses. S.S. performed the pollen viability assay. M.Z., S.D. and A.M. performed the dating analysis. M.Z., T.L. and K.F.X.M. performed the functional annotation. G.T., U.N., A.V., T.T. and D.F. performed and analysed the female gametophyte and embryo sac development microscopy data. M.Z., M.C., S.S. and A.M. wrote the first manuscript draft with input from all authors. All authors approved the final version of the manuscript.

Sex without crossovers mimics clonal reproduction in the holocentric plant *Rhynchospora tenuis*

Meng Zhang[1], Marco Castellani[1], Stefan Steckenborn[1], Maciej Majka[1,2], Georgios Tsipas[1], Thiago Nascimento[1,3], Ulla Neumann[4], Gokilavani Thangavel[1,5], Laura A. Robledillo[1], Thomas Lux[6], Lorraine Deberón[1,7], Ursula Pfordt[1], José A. Campoy[1,8], Athul Vijayan[4], Ton Timmers[4], Nafiseh Sargheini[9], Magdalena Marek[9], Hequan Sun[1,10], Paulo G. Hofstatter[1,11], Bruno Huettel[9], Steven Dreissig[12], Klaus F. X. Mayer[6,13], Duarte D. Figueiredo[14], André L. L. Vanzela[15], Korbinian Schneeberger[1,16,17], André Marques[1,17,#]

[1]Department of Chromosome Biology, Max Planck Institute for Plant Breeding Research, Carl-von-Linné-Weg 10, 50829 Cologne, Germany

[2]Institute of Plant Genetics, Polish Academy of Sciences, Strzeszyńska 34, 60-479 Poznań, Poland

[3]Department of Botany, Federal University of Pernambuco, Recife, PE, 50670-901 Brazil

[4]CeMic, Max Planck Institute for Plant Breeding Research, Carl-von-Linné-Weg 10, 50829 Cologne, Germany

[5]Molecular Cell Biology of Animals, Karlsruhe Institute of Technology, Fritz-Haber-Weg 4, 76131 Karlsruhe, Germany

[6]Helmholtz Center Munich, Ingolstädter Landstr. 1, 85764 Neuherberg, Germany

[7]Centre for Novel Agricultural Products (CNAP), Department of Biology, University of York, York YO10 5DD, UK

[8]Department of Pomology, EEAD-CSIC, Avda. Montañana 1.005, Zaragoza 50059, Spain

[9]Max Planck Genome Centre Cologne, Max Planck Institute for Plant Breeding Research, Cologne, Germany

[10]MOE Key Laboratory for Intelligent Networks & Network Security, Faculty of Electronic and Information Engineering, Xi'an Jiaotong University, Xi'an, China

[11]Universidade de São Paulo, Faculdade de Medicina de Ribeirão Preto, Brazil.

[12]Leibniz Institute of Plant Genetics and Crop Plant Research (IPK), 06466, Seeland, OT Gatersleben, Germany

[13]Department of Life Sciences, Technical University Munich, Freising, Germany

[14]Max Planck Institute of Molecular Plant Physiology, Potsdam Science Park, 14476, Potsdam, Germany
[15]Laboratory of Cytogenetics and Plant Diversity, Center for Biological Sciences, State University of Londrina, Londrina, 86051-980, Paraná, Brazil
[16]LMU Munich, Faculty of Biology, Martinsried, Germany
[17]Cluster of Excellence on Plant Sciences (CEPLAS), Heinrich-Heine University, Düsseldorf, Germany

#Correspondence: amarques@mpipz.mpg.de

Abstract

Meiotic recombination ensures accurate chromosome segregation and promotes genetic diversity by generating crossovers between homologous chromosomes[1]. While essential in most sexually reproducing organisms, recombination is variably regulated and can be absent in some lineages, a condition known as achiasmy[2]. However, obligate achiasmy in both sexes of a sexual species has not been previously documented. Here, we investigate the beak-sedge *Rhynchospora tenuis*, a holocentric plant with the lowest known chromosome number among flowering plants ($n = 2$) and inverted meiosis[3]. Using chromosome-scale genome assemblies from nine accessions, molecular cytogenetics, immunocytochemistry, high-throughput single-gamete sequencing and whole-genome sequencing of controlled crosses, we show that *R. tenuis* undergoes obligate, genome-wide achiasmy in both male and female meiosis. Despite normal early meiotic axis formation, synapsis fails, crossovers are not detected cytologically or genetically, and univalents persist at metaphase I. Extensive haplotype-specific accumulation of transposable elements (TEs) generates segregation distortion (e.g. meiotic drive), favouring the transmission of larger, TE-rich chromosomes. Remarkably, sexual reproduction is retained with fertilisation producing viable seeds only when translocation-compatible gametes meet, indicating strong post-meiotic selection that eliminates incompatible homozygous combinations. As a result, all surviving offspring are genetically identical to the maternal genotype, effectively restoring heterozygosity each generation and mimicking clonal reproduction. We propose that the combined effects of recombination loss, low chromosome number, holocentricity, inverted meiosis, and selective transmission of longer chromosomes enable faithful segregation and clonal-like inheritance despite sexual reproduction. These findings challenge the boundary between sex and clonality, revealing a unique evolutionary strategy linking genome architecture, recombination loss, and transmission bias.

Main

Although meiotic recombination is critical for most sexually reproducing species, its frequency varies widely, often differing between sexes (heterochiasmy)[4,5], and can be entirely suppressed in specific chromosomes or lineages, known as achiasmy[2]. In extreme cases, such as male *Drosophila melanogaster*, meiotic recombination and chiasmata are completely absent, requiring alternative mechanisms like chromatin threads or meiotic drive to ensure correct chromosome segregation[6–10]. The absence of recombination is expected to have profound consequences, including impaired chromosome segregation, reduced genetic diversity, and accelerated mutation accumulation through the so-called 'Meselson effect'[11–14]. However, obligate, genome-wide bisexual achiasmy of a sexually reproducing organism has never been described.

The beak-sedge *Rhynchospora tenuis* Link (n = 2) presents a compelling model to investigate these questions. It harbours only two pairs of holocentric chromosomes, the lowest known number in plants, and undergoes inverted meiosis, with sister chromatids separating at meiosis I, followed by the segregation of homologous chromosomes at meiosis II[3,15]. Its holocentric chromosomes, in which kinetochore activity is distributed along their entire length, are known to tolerate structural rearrangements such as chromosome end-to-end translocations (hereafter called 'fusions')[16], which may have facilitated the evolution of its extremely reduced chromosome number[17]. Previous cytological studies revealed that male meiosis in *R. tenuis* proceeds without homologous chromosome pairing and chiasmata formation, displaying four univalents at diakinesis, raising the possibility of achiasmy despite apparent meiotic DNA double-strand breaks (DSB) formation[3]. However, it remains unclear whether and how meiotic recombination is completely abolished, particularly during female meiosis, and whether genome transmission remains stable in the absence of crossovers (COs).

Here, we investigate the recombination landscape, meiotic progression, and inheritance patterns of *R. tenuis* using an integrative approach combining chromosome-scale pangenomics, molecular cytogenetics and immunocytochemistry, single-gamete sequencing of pollen nuclei, plus whole-genome sequencing of offspring from controlled crosses. We demonstrate that *R. tenuis* exhibits obligate bisexual achiasmy, with complete

absence of recombination, synapsis defects, and striking segregation distortion favouring larger chromosomes, enriched with TEs. Remarkably, this results in progeny that mirrors the mother plant genotype despite sexual reproduction. Our findings uncover a reproductive system in which meiotic drive and post-meiotic selection effectively replace the need for recombination, challenging established boundaries between sexual and clonal reproduction.

Genome assemblies reveal haplotype differentiation and reciprocal translocations

We assembled haplotype-phased chromosome-scale genomes for nine *R. tenuis* accessions (18 haplotypes) sampled across different locations in Brazil (see **Online Methods; Supplementary Table 1 and Supplementary Fig. 1**). Despite its predominantly autogamous mating system, *R. tenuis* exhibits extremely high levels of heterozygosity (>2% in all accessions), exceeding even that of the closely related outbreeding species *Rhynchospora breviuscula* of approximately 1% (**Supplementary Fig. 2**). Epigenome profiling further revealed a distinctive chromatin organisation of *R. tenuis* holocentromeres. Unlike *Rhynchospora pubera*[17] and *R. breviuscula*[18], which feature canonical centromeric domains enriched for CENH3 and CpG DNA methylation, *R. tenuis* holocentromeres display marked CpG hypomethylation and an unexpected enrichment of the active histone mark H3K4me3, coinciding with *Tyba* satellite arrays (**Extended Data Fig. 1**). These features indicate that *R. tenuis* has evolved an unconventional and atypically euchromatic centromeric environment, potentially linked to its unusual meiotic system and distinctive genome dynamics.

Macrosynteny with closely related species reveals a stepwise fusion trajectory underlying the remarkably reduced chromosome number of *R. tenuis*. Conserved collinear blocks reconstruct an ancestral *Rhynchospora* karyotype of $n = 5$, as seen in *R. breviuscula* and most *Rhynchospora* species with five chromosomes[17,19]. Two successive end-to-end translocations in a common ancestor appear to have produced an intermediate $n = 3$ karyotype by fusing ancestral chromosomes 2 + 5 and 3 + 4, an arrangement retained in the autohexaploid (6*x*) *R. austrobrasiliensis*, the closest known relative of *R. tenuis*, which shares identical fusion breakpoints (**Fig. 1a and Supplementary Fig. 3**). A subsequent fusion involving ancestral chromosome 1 and the 2 + 5 yielded the present *n*

= 2 karyotype, conserved across all sequenced *R. tenuis* accessions (**Fig. 1a and Extended Data Fig. 2**).

In addition to these chromosomal fusions, pangenome-scale synteny uncovered striking chromosome size asymmetry between opposing haplotypes (i.e., h1 and h2), varying from 8 Mbp in the REC (Recife) accession to up to 32 Mbp in JGV (Jaguariaíva) and PECP (Parque Estadual do Cerrado) accessions (**Fig. 1b, Extended Data Fig. 2**). Furthermore, population-specific telomeric reciprocal translocations were detected, ranging from 0.53 to 3 Mbp at one end of each chromosome in *R. tenuis*, with the exchanged segments mapping consistently to different chromosomal ends in different accessions, being classified into three translocation types (**Fig. 1a–b, Extended Data Fig. 2**). To further validate the reciprocal translocations, we designed translocation-specific oligo probes based on the translocation type 1, found in the REC accession (see **Methods**). Fluorescent *in situ* hybridisation (FISH) with translocation-specific oligo- and 45S rDNA (nucleolar organising regions, i.e., NORs) probes validated the reciprocal translocations in all accessions (**Fig. 1c, Extended Data Fig. 3**). The REC accession shows the signals for the translocation oligo-probe 1 (magenta) and oligo-probe 2 (cyan) at the chromosome ends of the non-homologous chromosomes (**Fig. 1b–c, Extended Data Fig. 2 and 3**). JGV accessions showed the same arrangement as REC for chromosome 1, but only a partial translocation was found at the chromosomal ends of chromosome 2 (**Fig. 1b–c, Extended Data Fig. 2 and 3**). As expected from the assemblies, the PECP accessions did not exhibit translocation for that specific chromosome ends. However, reciprocal translocations were found at the other chromosome ends, which in one of the combinations contain the NORs (**Fig. 1b–c, Extended Data Figs. 2 and 3**). No reciprocal translocations were found in the close relative *R. austrobrasiliensis* ($2n = 6x = 18$) (**Supplementary Fig. 4**). Our results highlight the role of population-specific structural rearrangements on the unique genome organisation of *R. tenuis*.

Since *R. tenuis* has been reported to be achiasmatic, neither recombination nor segregation of homologous chromosomes is expected, a condition consistent with the 'Meselson effect'. In such a system, homologous chromosomes accumulate mutations independently, leading to greater divergence between haplotypes than among individuals

or populations, effectively generating subgenome-like differentiation in an otherwise diploid genome[11,12]. *K*-mer similarity of all *R. tenuis* haplotypes shows a textbook signature of 'Meselson effect', where the same haplotypes from all accessions cluster together while two haplotypes from the same individual do not cluster (**Extended Data Fig. 4**), mimicking an obligately asexual species[12].

To test whether the *R. tenuis* genome evolves under this regime, we have used subphaser[20], a *k*-mer-based approach normally used to phase subgenomes in allopolyploid species (polyploid hybrids), on phased haplotypes to estimate divergence between haplotypes versus among-individual divergence. Remarkably, this analysis revealed pronounced subgenome-like differentiation between the two haplotypes across all accessions (**Fig. 1d–e; Extended Data Fig. 5**). This differentiation is mainly explained by haplotype-specific accumulation of TEs, which explains on average 84% of the striking chromosome size asymmetry found between haplotype 1 and 2, while other sequence types, e.g. *Tyba* repeats and genes, did not differ greatly between haplotypes (**Fig. 1d–e; Extended Data Fig. 6**). The extent and distribution of these structural variants, together with population-specific terminal reciprocal translocations, suggest long-term suppression, and effective absence, of meiotic recombination between haplotypes.

Next, we aimed to date the divergence between the two non-recombining haplotypes of *R. tenuis*. Considering that both haplotypes evolve independently without genetic exchange, we estimated their divergence time using synonymous substitution rates (*K*s) calibrated with a fixed molecular clock (6.13e-09 per site per year[21]), and contrasted these values with insertion age estimates of haplotype-specific LTR retrotransposons and SNP-based divergence estimates. The *K*s-based molecular clock suggests that the two haplotypes diverged approximately 1.7 million years ago (Mya; **Fig. 1f**), whereas LTR insertion times indicate a more recent split, between 400,000 and 600,000 years ago (**Fig. 1f**). Divergence times between haplotypes calculated from SNPs divergence were estimated to range from 230,000 - 310,000 years for chromosome 1 and 220,000 - 275,000 years for chromosome 2 (**Fig. 1f**). The higher divergence time inferred from *K*s likely reflects intrinsic differences between coding and non-coding evolutionary rates and the use of a fixed molecular clock, rather than artefactual inflation, whereas the more

recent SNP- and LTR-based estimates probably better capture the effective timescale of genome-wide divergence following recombination loss.

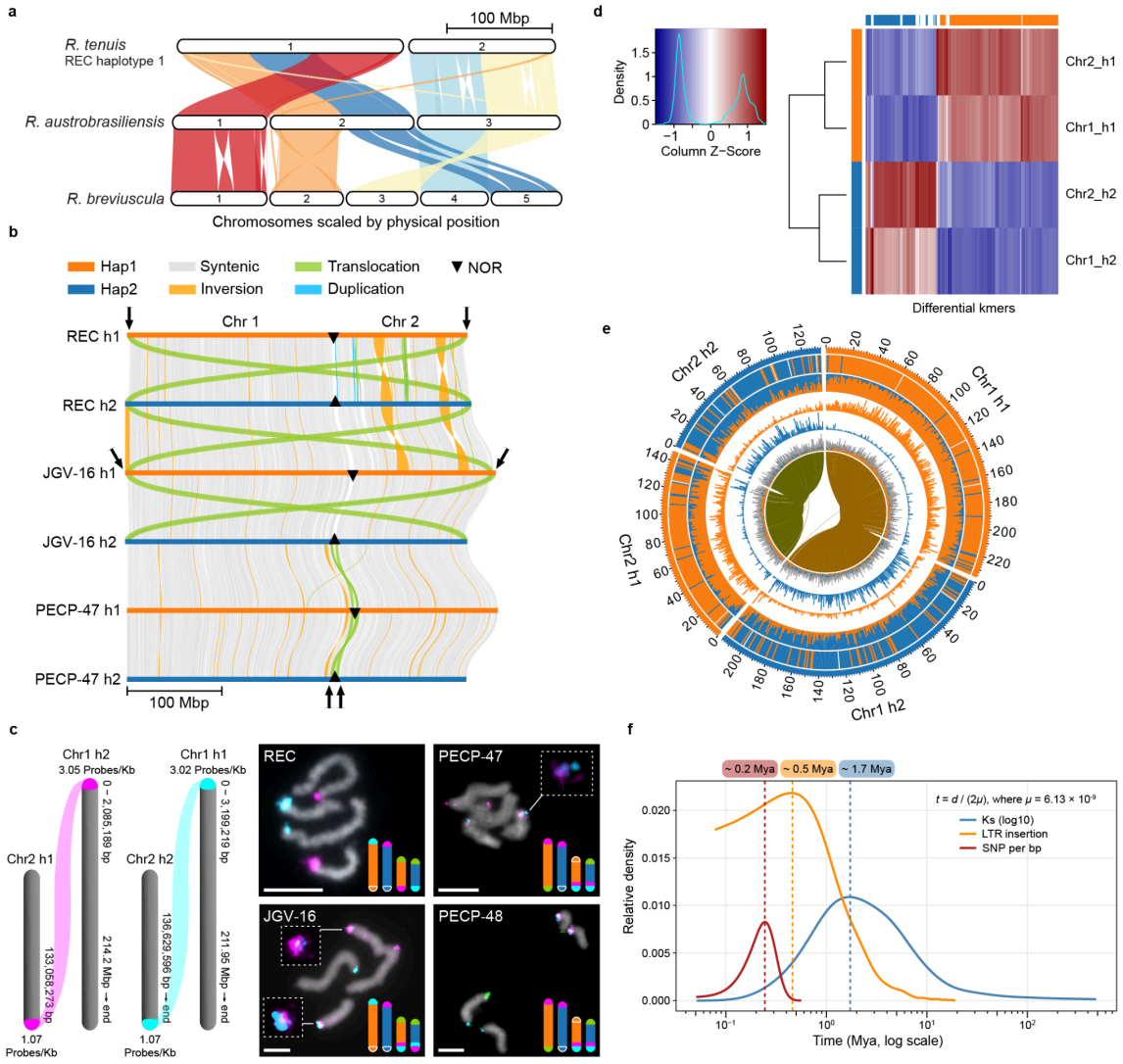


Fig. 1. Overview of *R. tenuis* pangenome organisation. (**a**) Macrosynteny across *Rhynchospora* delineates a stepwise reduction in *R. tenuis* chromosome number: conserved syntenic blocks support an inferred ancestral karyotype of $n = 5$, fusion events yielding $n = 3$ in *R. austrobrasiliensis*, and a final fusion producing $n = 2$ in all *R. tenuis* accessions. (**b**) Pangenome analysis also identifies ~1 Mb terminal reciprocal translocations at one end of each chromosome in *R. tenuis*. (**c**) Representation of the target regions used for haplotype-specific probes design and the final density of each synthesised probe. Oligo-FISH with translocation-specific probe sets (magenta, oligo-probe 1; cyan, oligo-probe 2) validates reciprocal exchanges in the REC and the JGV accessions, but not in PECP-47/PECP-48. 45S rDNA signals (green) in PECP-47/PECP-48 are found on different homologous chromosomes. (**d–e**) *k*-mer-based genome subphasing reports a striking haplotype differentiation found in all *R. tenuis* accessions, which is caused by a differential accumulation of LTR retroelements. (**d**) Unsupervised hierarchical clustering (the horizontal

colour bar at the top of the axis indicates to which subgenome the *k*-mer is specific; the vertical colour bar on the left of the axis indicates the subgenome to which the chromosome is assigned). The heatmap indicates the *Z*-scale relative abundance of *k*-mers. The larger the *Z* score is, the greater the relative abundance of a *k*-mer). (**e**) Chromosomal characteristics. From the outer to inner circles (1–9): (1) subgenome assignments based on the *k*-means algorithm; (2) significant enrichment of subgenome-specific *k*-mers − the same colour as the subgenome indicates significant enrichment for those subgenome-specific *k*-mers − white areas are not significantly enriched; (3) normalized proportion (relative) of subgenome-specific *k*-mers; (4–5) count (absolute) of each subgenome-specific *k*-mer set; (6) density of long terminal repeat retrotransposons (LTR-RTs) − if the colour is consistent with the subgenome, it indicates that LTR-RTs are significantly enriched in those subgenome-specific *k*-mers; grey indicates nonspecific LTR-RTs; and (7) homoeologous blocks. All statistics (2–7) are computed in sliding windows of 1 Mb. (**f**) Divergence time estimates between the two non-recombining haplotypes of *R. tenuis* based on the distribution density of pairwise *Ks* (log10) of orthologous coding sequences, haplotype-specific LTRs, and SNP density in the syntenic regions.

Absence of pairing and synapsis during *R. tenuis* achiasmatic inverted meiosis

Our previous study reported the absence of chiasmata and inverted meiosis in *R. tenuis*[3]. Here, we further tested whether achiasmy and inverted meiosis are consistent features across different populations of *R. tenuis* (REC, PECP and JGV). Indeed, all genotypes analysed always display four univalents at diakinesis and no bivalents (n = 10, n = 37; n = 51), followed by equational division of sister chromatids (n = 37, n = 28; n = 12), consistent with the concept of inverted meiosis; (**Fig 2a–d; Supplementary Fig. 5**). To deepen our understanding, we further tracked the localization of translocation-specific oligo- and 45S rDNA probes on meiotic chromosomes. During early meiotic prophase, translocation-specific oligo-probes and telomeric signals are never associated indicating absence of pairing and synapsis, while rDNA regions are associated at the nucleolus (n = 55; **Fig. 2a–b; Supplementary Fig. 6a**), followed by diakinesis with four univalents and no bivalents, indicating asynapsis and achiasmy (**Supplementary Fig. 6b**). At metaphase I, univalents aligned perpendicular to the equatorial plate and sister chromatids exhibited bi-orientation, leading to segregation of sisters at anaphase I (equational division) and formation of two diploid products at telophase I/interkinesis (**Fig. 2c; Supplementary Fig. 6c**). During meiosis II, the previously separated sisters realigned as homologous non-sisters at metaphase II and segregated at anaphase II, yielding four haploid products

(**Fig. 2d; Supplementary Fig. 6d**). Across all populations, the consistent absence of chiasmata/bivalents and the inversion of the canonical meiotic order (equational division in meiosis I and reductional division in meiosis II) demonstrate that achiasmy and inverted meiosis are conserved and consistent features of *R. tenuis* rather than accession-specific anomalies.

We extended our analysis of meiotic progression in *R. tenuis* by immunostaining male meiocytes with antibodies against the axis components ASY1[22] and REC8[23], the synaptonemal complex (SC) transverse filament ZYP1[24], and the CO marker HEI10[25]. Early prophase I appeared conserved: ASY1 (n = 41) and REC8 (n = 29) localised as linear signals along chromosomal axes in all cells analysed (**Fig. 2e–f and Extended Data Fig. 7a–b**). By contrast, ZYP1 failed to assemble as a continuous SC; instead, it formed a few persistent short elongated structures and polycomplexes while ASY1 was unloaded from the axes (n = 10; **Fig 2g and Extended Data Fig. 7c**). HEI10 was occasionally detected on ZYP1 polycomplexes but did not form discrete foci along chromosome axes (n = 12; **Extended Data Fig. 7d**). At diakinesis, meiocytes invariably displayed four univalents, and no HEI10 foci indicative of COs were observed (n = 21; **Fig 2h**). These features, fragmented SC signal, and absence of CO markers, culminated in univalents at metaphase I (**Fig. 2c and h**), consistent with achiasmy.

As a contrast, the closest relative, *R. austrobrasiliensis*[26], exhibited canonical chiasmatic meiosis. ASY1 localized to the axes and was progressively unloaded during synapsis (n = 12), the SC was assembled as continuous ZYP1 signals (n = 12; **Fig 2i and Extended Data Fig. 8a**), and HEI10 initially appeared as closely spaced foci on ZYP1-marked regions, later re-localizing as a few intense foci on bivalents, co-localizing with MLH1 (n = 59) and confirming normal CO maturation (**Fig. 2j and Extended Data Fig. 8b–d**).

To explore a genetic basis for recombination failure in *R. tenuis*, we surveyed meiotic gene content and transcription. Most core meiotic genes and their corresponding transcripts are present (**Supplementary Table 2**). However, we detected haplotype-specific copy-number variation at several meiotic loci, including additional copies for *ASY1* and *MSH5* specific to achiasmatic *R. tenuis* (**Supplementary Table 2**).

SHOC1, responsible for the maturation of mid-to-late recombination intermediates in plants and mammals[27,28], showed some ambiguities. Two independent RNAseq datasets did not contain traces of *SHOC1* expression. However, the gene displays a complete and apparently functional sequence, and PCR amplification of the functional domain from tissue-specific cDNA was successfully achieved. However, we cannot rule out that expression levels or post-transcriptional modifications might affect the activity of the SHOC1 protein.

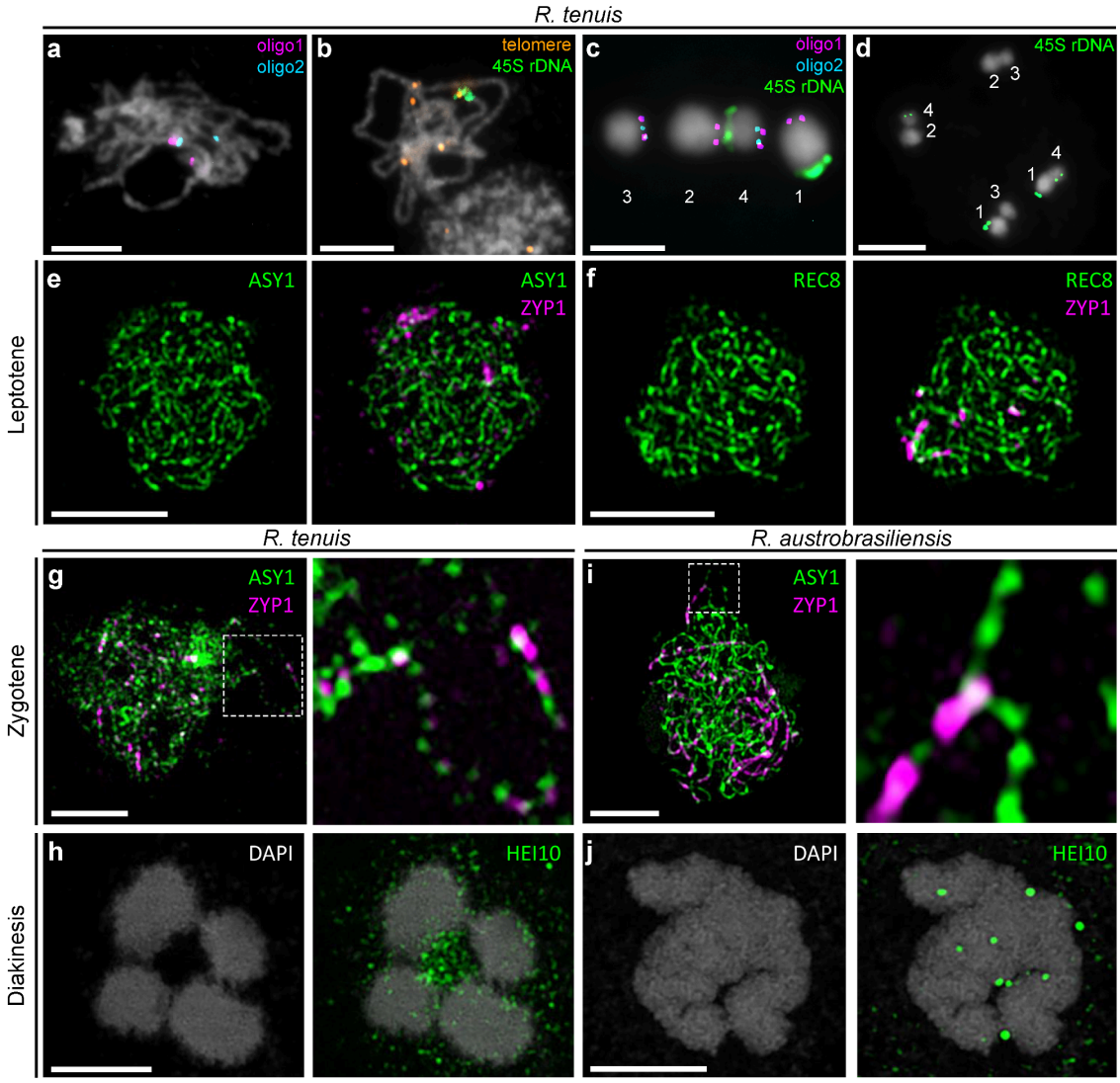


Fig. 2: FISH and Immunostaining confirms the failure of pairing, synapsis and crossover formation in *R. tenuis* (REC). FISH using (**a**) translocation-specific probe sets and (**b**) 45S rDNA and telomeric probes revealed neither pairing of translocation regions (**a**) nor bouquet formation (**b**) during early meiosis. FISH with translocation-specific probes (**c**) and 45S rDNA (**c–d**) in metaphase I (**c**) and telophase II (**d**). 1

= Chr1_h1, 2 = Chr1_h2, 3 = Chr2_h1 and 4 = Chr2_h2. Immunostaining with (**e**) ASY1 and (**f**) REC8 together with ZYP1 signal in Leptotene. Immunostaining with (**g; i**) ASY1 and ZYP1 depicts abnormal zygotene in *R. tenuis* (**g**) contrasted to zygotene in *R. austrobrasiliensis*, which shows normal axis and synapsis (**i**). (**h; j**) HEI10 immunostaining shows the lack of foci at diakinesis in *R. tenuis* (**h**), but the presence of foci in *R. austrobrasiliensis* (**j**). Scale bars correspond to 5 μm.

Genetic COs are absent in both male and female meiosis

To test genetically whether COs are absent in *R. tenuis*, we applied our single-gamete sequencing pipeline[18,19] to pollen nuclei across five geographically distinct accessions, leveraging chromosome-scale, haplotype-phased assemblies as references. Using high-throughput single-nucleus profiling (snRNA- and snATAC-seq), we obtained 10,997 pollen grains in total that yielded high-confidence, genome-wide genotypes (**Fig. 3a, Supplementary Table 3**). Haplotype blocks spanned entire chromosomes with no confident COs detected in any gamete, demonstrating that male achiasmy is pervasive across accessions. This means that, if crossovers occur at all, they must be extremely rare, consistent with a complete suppression of recombination during male meiosis.

Because our single-pollen sequencing interrogates only male meiosis, we assessed female recombination genetically by sequencing the offspring of controlled self-crossing heterozygous mother plants. In the highly heterozygous genome of *R. tenuis*, female COs would be visible as haplotype switchpoints in F1 individuals. Whole-genome sequencing of 49 and 62 F1 individuals coming from the controlled self-crossed seeds from heterozygous mothers of REC and PECP accessions, respectively, revealed no recombination breakpoints on either chromosome in any offspring (**Fig. 3b**). Remarkably, all F1 individuals carried the same two intact parental haplotypes as the mother, with no detectable gene conversion, demonstrating that female meiosis, like male meiosis, is achiasmatic. Furthermore, these results also indicate that all offspring show clonal-like formation as they mirror the maternal genotype.

We next examined transmission patterns among *R. tenuis* viable pollen grains, which are lower than 75% across all accessions analysed (**Supplementary Fig. 7**). In *R. tenuis*, as a conserved feature of the *Rhynchospora* genus, male microsporogenesis is asymmetric:

only one of the four meiotic products matures into a functional microspore (pollen), while the remaining products degenerate[29–32] (**Extended Data Fig. 9a**). Such asymmetry provides a cellular context for male meiotic drive, which is not expected under symmetric microsporogenesis. Consistent with this, allelic counts at haplotype-informative markers revealed strong, accession-specific segregation distortion among viable pollen (**Fig. 3c**). Under inverted meiosis, sister chromatids segregate at meiosis I, yielding diploid secondary spermatocytes; if meiosis II followed Mendelian segregation and all four products matured equivalently, the four genotype classes defined by the two chromosomes should be equally frequent. Instead, in 10,997 pollen nuclei with complete genotype calls across both chromosomes, we observed significant over-representation of specific chromosome combinations. Across the full dataset, distortion was particularly substantial in two closely related accessions (PECP-47, PECP-48), where the larger, TE-rich haplotype 1 was preferentially transmitted in mature pollen (>50%) over other combinations (**Fig. 3d**). Remarkably, pollen cells with both chromosomes of haplotype 2 were found at very low frequency (4%). Given that haplotype 1 exceeds haplotype 2 by ~32 Mb primarily due to LTR retrotransposon accumulation, these results point to a male meiotic drive system in which TE-associated features may directly or indirectly skew chromosomal inheritance during, or immediately following, meiosis II. Together with the complete absence of COs, these findings indicate that, in the asymmetric microsporogenesis of *R. tenuis*, transmission bias operates in the context of recombination loss to favour particular haplotype combinations, with chromosome size being the strongest factor.

A structural basis for this distortion emerges from the position of the nucleolar organiser region (NOR, 45S rDNA). In REC and JGV, the NOR is not interchanged between non-homologous chromosomes, whereas in PECP-47 and PECP-48, the NOR-bearing segments are reciprocally translocated between Chr1 and Chr2. Consequently, one genotype class, Chr1_h2 + Chr2_h1, lacks any NOR, predicting compromised ribosome biogenesis and gamete inviability. Intriguingly, single-nucleus genotypes retrieved pollen grains in PECP-47/PECP-48 showing the following frequencies: Chr1_h1 + Chr2_h1 (50%, bearing one NOR), Chr1_h1 + Chr2_h2 (28%, bearing two NORs), Chr1_h2 + Chr2_h1 (18%, NOR-null), and Chr1_h2 + Chr2_h2 (4%, bearing one NOR). However,

FISH on PECP-48 mature pollen grains (n = 33), using a combination of oligo-probes and 45S rDNA, recovered similar abundances as detected by our single-pollen sequencing, but only on genotype combinations that retain at least one NOR. However, we never found the NOR⁻null class (Chr1_h2 + Chr2_h1) genotype combination in mature pollen grains and in sperm cells from germinated pollen tubes (**Fig. 3d–e and Extended Data Fig. 9b**). Remarkably, we found in 90.5% (n = 21) of the germinated pollen tubes only one single NOR signal in both tube nucleus and sperm cells, which is consistent with chromosomes from the same haplotype being inherited together (**Extended Data Fig. 9b**). Further investigation in earlier pollen stages consistenly revealed segregation distortion of haploid products, with the NOR⁻null class genotype being found only at the degenerative cells (non-selected products) during pollen development (**Extended Data Fig. 9c–d**), indicating post⁻meiotic selection of haploid products during early pollen development. Together with a consistent bias favouring the larger, TE⁻rich haplotype 1 in PECP-47/PECP-48, these results reveal asymmetric male meiotic drive in *R. tenuis*, in which structural features, including NOR placement and TE⁻associated genome expansion, skew chromosomal transmission in the complete absence of recombination.

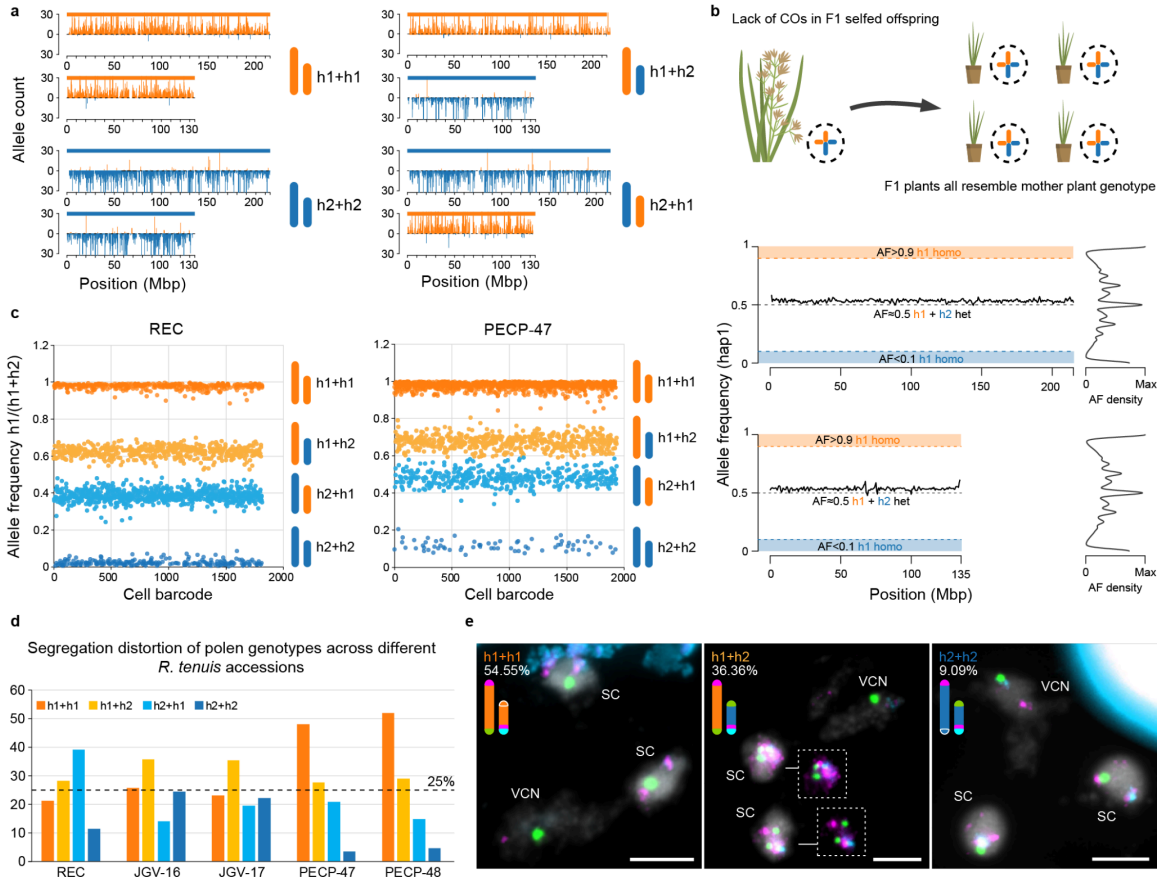


Fig. 3: Single-gamete sequencing reveals complete suppression of crossovers and male meiotic drive under asymmetric microsporogenesis in *R. tenuis*. (**a**) Four examples of pollen genotype combinations obtained from single-nucleus sequencing of pollen nuclei show chromosome-scale haplotype blocks with no recombinants. Failing to detect a single CO after analysing 10,997 pollen nuclei from five different accessions. (**b**) F1 offspring of controlled selfing-crossed mother plants show consistent chromosome-wide 1:1 allele ratios, indicating the absence of COs and mirroring of the mother genotype, indicating clonal-like seed formation. (**c**) Allele frequency, as defined by the ratio of h1 markers divided by the total markers (h1+h2) identified in each single pollen nucleus sequenced, reveals four main groups of pollen genotypes in *R. tenuis*. (**d**) Genotype class frequencies highlight segregation distortion among sequenced pollen genotypes across accessions, with the strongest bias in PECP-47 and PECP-48 favouring the larger, TE-rich haplotype 1. (**e**) Representative haplotype-specific barcoding across chromosomes illustrates intact, nonrecombinant transmission in individual mature pollen grains. In PECP-47/PECP-48, reciprocal translocations of the NOR eliminate NOR-null gametes post-meiotically, further skewing haplotype transmission. SC = sperm cell, VCN = vegetative cell nucleus. Scale bar = 5 µm. Oligo-probe 1 = magenta, oligo-probe 2 = cyan, 45S rDNA probe = green.

Post-meiotic selection yields clonal-like offspring despite sexual reproduction

To exclude that the obtained seeds from our F1 offspring are derived from apomixis, a form of asexual seed formation in which embryos develop without fertilisation or meiosis[33], we performed crossing experiments to confirm that seed formation requires fertilisation. Remarkably, viable seeds were only obtained when both parents carried the same reciprocal translocation type (e.g., PECP-47 × PECP-48), whereas all other combinations failed to set seed (**Fig. 4a**). Sequencing of offspring from these crosses confirmed that all plants were heterozygous, consistent with fertilisation but indicating that heterozygosity is obligatory (**Fig. 4b; Supplementary Dataset 1**).

To further examine female gametophyte and early seed development, we performed histological analyses of early and mature flowers. Female gametophytes showed normal development, with well-differentiated embryo sacs and initiation of seed development consistent with fertilisation (**Fig. 4c**). At early stages of megagametogenesis, multiple embryo sacs were frequently observed out of which only one continues to mature in 98% of the cases (n = 59, **Supplementary Fig. 8**). Pollen tube growth towards the ovule was consistently observed (**Fig. 4d and Extended Data Fig. 10a–b; Supplementary Fig. 9 and 10**). Although healthy embryos were consistently observed developing from the micropyle, indicating sexual origin (**Fig. 4e–f**), a high proportion of ovules seem to fail to develop into seeds (**Fig. 4h and Extended Data Fig. 10c–f**). Microscopic inspection of these undeveloped ovules revealed intact, likely unfertilised gametophytes, suggesting failure of fertilisation, potentially due to male gametophytic defects, whereas others likely collapsed shortly after fertilisation (**Fig. 4h and Extended Data Fig. 10d–f**). Furthermore, a high frequency (87%) of aborted seeds was observed (n = 167, 145 aborted: 22 fertile for PECP-48; **Extended Data Fig. 10d–f**), indicating strong post-meiotic and post-fertilisation selection.

Flow cytometry of full mature seeds confirmed diploid embryos and triploid endosperms (**Fig. 4i**), consistent with meiosis followed by double fertilisation. Single-nucleus ATAC-sequencing of flow-sorted embryo and endosperm nuclei verified biparental contributions: embryos were uniformly heterozygous, while endosperms showed the

expected 2:1 maternal:paternal dosage (**Fig. 4j**), enabling direct inference of female gamete haplotype composition. Strikingly, in both REC and PECP-48 accessions, a single maternal haplotype was always transmitted (haplotype 2 in REC and haplotype 1 in PECP-48; **Fig. 4k**). In both cases, the driving haplotype corresponded to the physically larger, TE-enriched chromosomes.

Importantly, none of the imaged ovules or seeds showed evidence of apomictic development. In all cases, embryos originated exclusively from the micropyle, never from sporophytic tissues, and only a single embryo was observed per seed. Flow cytometry of mature seeds confirmed the expected 2:3 embryo:endosperm ratio, consistent with double fertilisation. These results exclude diplospory, apospory, and adventitious embryony as mechanisms of seed formation, supporting sexual reproduction as the sole reproductive mode in *R. tenuis*.

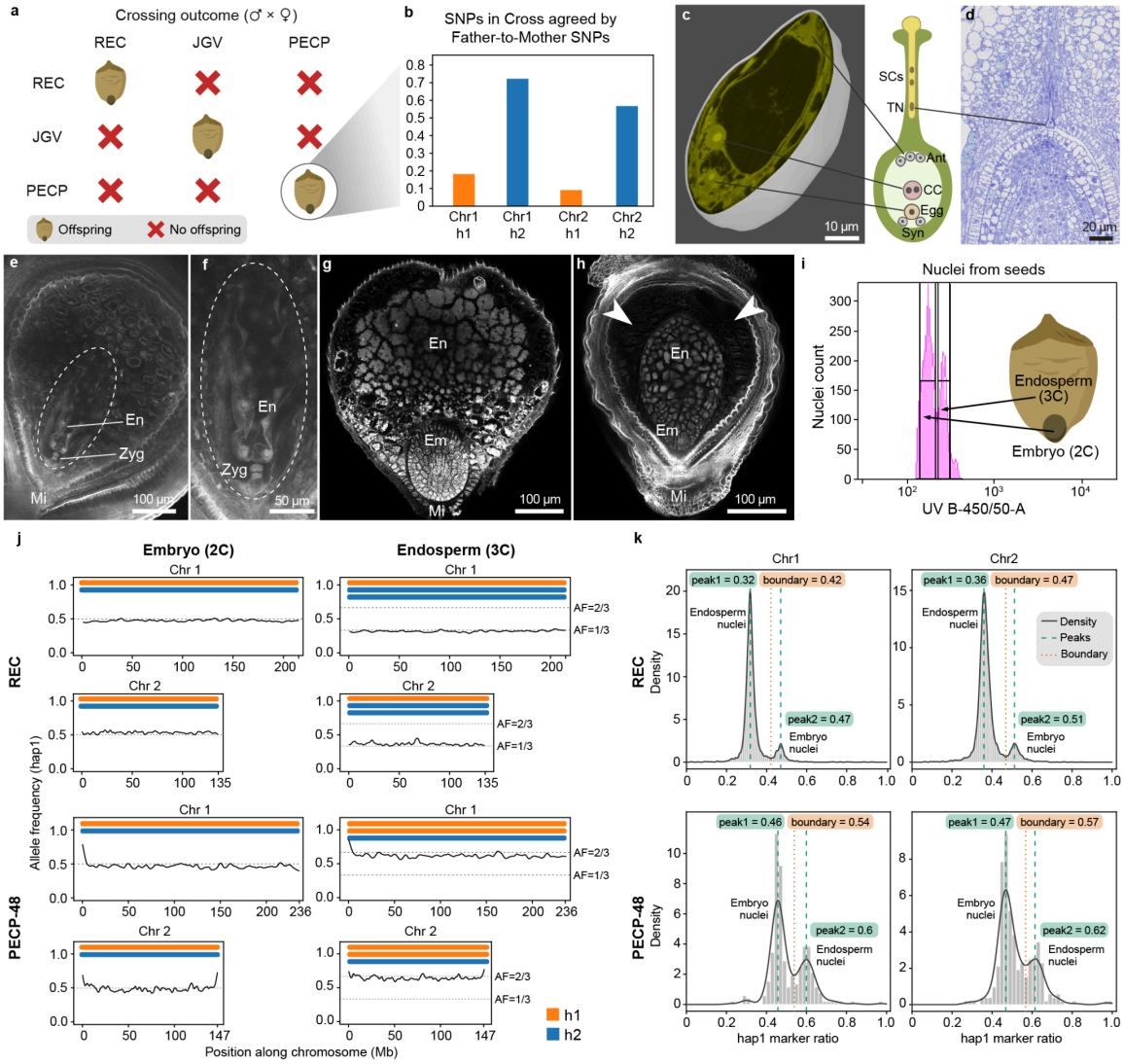


Fig. 4: Female achiasmy and post-meiotic selection produce clonal-like progeny despite sexual reproduction. (**a**) Crossing experiments between accessions with different reciprocal translocation genotypes show seed formation only when both parents share the same translocation type (example: PECP-48 × PECP-47). All successful crosses yield heterozygous offspring. (**b**) In the cross PECP-48 × PECP-47, alignment of progeny reads to the maternal genome (PECP-48) shows paternal-specific alleles across haplotype 2, indicating biparental fertilisation with haplotype 1 maternally inherited and haplotype 2 paternally inherited. (**c**) 3D reconstruction of serial histological sections reveals canonical female gametophyte development with a normal embryo sac containing egg cell (Egg), synergids (Syn), central cell (CC) with fused polar nuclei, and antipodals (Ant). (**d**) Pollen tube growth towards the ovule. (**e**) Very young stage after fertilisation showing the Zygote (Zyg) and coenocytic endosperm (En). (**f**) Enlargement of selected region shown in **e**. (**g**) Fully mature seed showing an embryo in a late stage of development surrounded by the endosperm. Embryos were consistently observed developing from the micropyle (Mi) consistent with sexual origin. (**h**) Late aborted seed showing underdeveloped and arrested embryo and

endosperm. Arrowheads point to collapsed area. (**i**) Flow cytometry profiles of seed nuclei reveal diploid embryos and triploid endosperms, consistent with canonical double fertilisation and the formation of seeds through sexual reproduction. (**j**) Single-nucleus sequencing of embryo and endosperm nuclei demonstrates biparental contributions: embryos are uniformly heterozygous, while endosperms exhibit the expected 2:1 maternal:paternal dosage. In REC and PECP-48, maternal transmission is biased toward the larger, TE-rich haplotype (haplotype 2 in REC, haplotype 1 in PECP-48). (**k**) Allelic ratio distributions confirm consistent dosage bias and post-meiotic selection that together restore heterozygosity and produce clonal-like offspring despite sexual reproduction.

Our findings indicate that seed production in *R. tenuis* is strictly sexual, but its success depends on translocation compatibility and the maintenance of heterozygosity, which appears to be essential for embryo viability. Together with the single-gamete sequencing data, these findings establish obligate, genome-wide achiasmy in both sexes. The apparent paradox of clonal-like offspring arising through sexual reproduction can be explained by extreme segregation distortion, reinforced by post-meiotic selection, which ensures that only heterozygous zygotes with specific haplotype combinations survive. As a result, viable progeny are genetically identical to the maternal genotype while retaining biparental genomic contributions (**Fig. 5**).

Discussion

Our study identifies a reproductive system without precedent: a sexually reproducing plant that has completely lost meiotic recombination while maintaining accurate chromosome transmission and normal fertility. In *R. tenuis*, we demonstrate obligate achiasmy in both sexes, characterised by the absence of crossovers and gene conversion, despite the normal initiation of meiotic processes. Thousands of sequenced gametes and multiple progenies reveal intact parental haplotypes and offspring genetically identical to their heterozygous mothers. This clonal-like inheritance arises through extreme segregation distortion, haplotype-biased transmission, and strong post-meiotic selection that restricts viable zygotes to specific heterozygous combinations.

These findings challenge the view that meiotic recombination is essential for proper chromosome segregation and for the generation of genetic diversity[34,35]. Although achiasmy occurs in limited contexts, such as *Drosophila* males[7], *Bombyx mori* females[10], or non-recombining sex chromosomes[2], *R. tenuis* appears to represent the first case of

complete, genome-wide achiasmy in both sexes of an otherwise sexually reproducing organism. Its unusual karyotype provides a permissive framework, as holocentric chromosomes stabilise segregation in the presence of rearrangements[17], inverted meiosis reduces reliance on chiasmata[3], and an extremely reduced chromosome number minimises the risk of missegregation. Together, these features create a cellular context in which recombination can be lost without catastrophic consequences.

The system both parallels and diverges from other non-Mendelian inheritance mechanisms. In *Ooceraea biroi* ants and *Mesorhabditis belari* nematodes, heterozygosity is retained through co-segregation of recombinant chromatids even when recombination is active[36,37]. A related plant analogue occurs in *Oenothera*, where balanced lethals and extensive reciprocal translocations enforce permanent translocation heterozygosity and restrict recombination to subterminal regions, thereby maintaining heterozygous haplotypes across generations[38–41]. By contrast, *R. tenuis* achieves the same functional outcome through a combination of bisexual achiasmy and gametic selection.

At first sight, clonal-like inheritance and recombination loss resemble the Meselson effect, which predicts that haplotypes in asexual lineages diverge through independent mutation accumulation[11,12]. Indeed, *R. tenuis* exhibits functional asexuality, and haplotypes are structurally divergent. However, unlike the parthenogenetic booklouse *Liposcelis bostrychophila*, where genomic data support Meselson-driven haplotype divergence[42], *R. tenuis* avoids this outcome. Meiosis and fertilisation are retained, homozygotes are eliminated, and meiotic drive biases transmission toward TE-rich haplotypes. These features limit independent haplotype evolution and reduce the risk of long-term genomic degeneration.

Recent work in the planarian *Schmidtea mediterranea* reveals that stepwise structural rearrangements can progressively suppress recombination before the evolution of fissiparous asexuality[43]. Both species show extensive structural variation that remodels the recombination landscape and produces divergent haplotypes without immediate genomic collapse. However, the evolutionary trajectories differ. In planarians, recombination suppression is followed by the loss of sex and reliance on somatic

selection to mitigate mutation load. In *R. tenuis*, recombination disappears mechanistically through obligate achiasmy, but meiosis and fertilisation persist. Clonal-like inheritance resulting from meiotic drive and post-meiotic selection rather than a shift to asexual reproduction. This highlights how similar ingredients, including recombination suppression and transmission bias, can generate different resolutions of the tension between sex and clonality.

The basis of recombination failure in *R. tenuis* remains unclear. Although *ASY1* and *MSH5* are duplicated, early meiotic processes appear to initiate normally before pairing and synapsis fail. Structural divergence between haplotypes, including reciprocal telomeric translocations and asymmetric TE accumulation, may disrupt homologous recognition. We observe consistent female drive of a single haplotype in independent accessions, in each case favouring the TE-rich, larger chromosomes. Because kinetochore activity extends along the entire chromosome in holocentric chromosomes during *Rhynchospora* meiosis[3,18,32], TE accumulation could enlarge the functional kinetochore and bias segregation in line with centromere and holokinetic drive models[44–46].

TE expansion on a single haplotype may reflect multiple processes. Germline-specific TE mobilisation during gametophyte development[47,48], could gradually enrich the haplotype that is preferentially transmitted through the ovule. Alternatively, haplotype asymmetry may derive from ancient hybridisation followed by recombination suppression. The opposite direction of female drive in different accessions (haplotype 1 in PECP and haplotype 2 in REC) argues that haplotype bias is dynamic rather than fixed, possibly reflecting local selection or stochastic changes in drive polarity.

Obligate bisexual achiasmy in *R. tenuis* expands the known limits of meiotic flexibility and blurs the distinction between sexual and clonal reproduction. Although recombination is widely regarded as essential for avoiding mutational meltdown and promoting adaptive potential[14,49–51], our results show that clonal-like inheritance can arise through modifications to meiotic mechanics, genome architecture, and post-meiotic selection without apomixis. Whether similar systems exist in other holocentric lineages with very low chromosome numbers remains unknown.

In conclusion, *R. tenuis* demonstrates that meiosis and fertilisation can persist in the complete absence of COs, creating a reproductive system that bridges sex and clonality. Holocentricity, inverted meiosis, and meiotic drive combine to stabilise obligate achiasmy. Our findings establish *R. tenuis* as a powerful model for studying how recombination can be lost, replaced, or reconfigured, providing a new framework for investigating the evolutionary consequences of extreme meiotic adaptations.

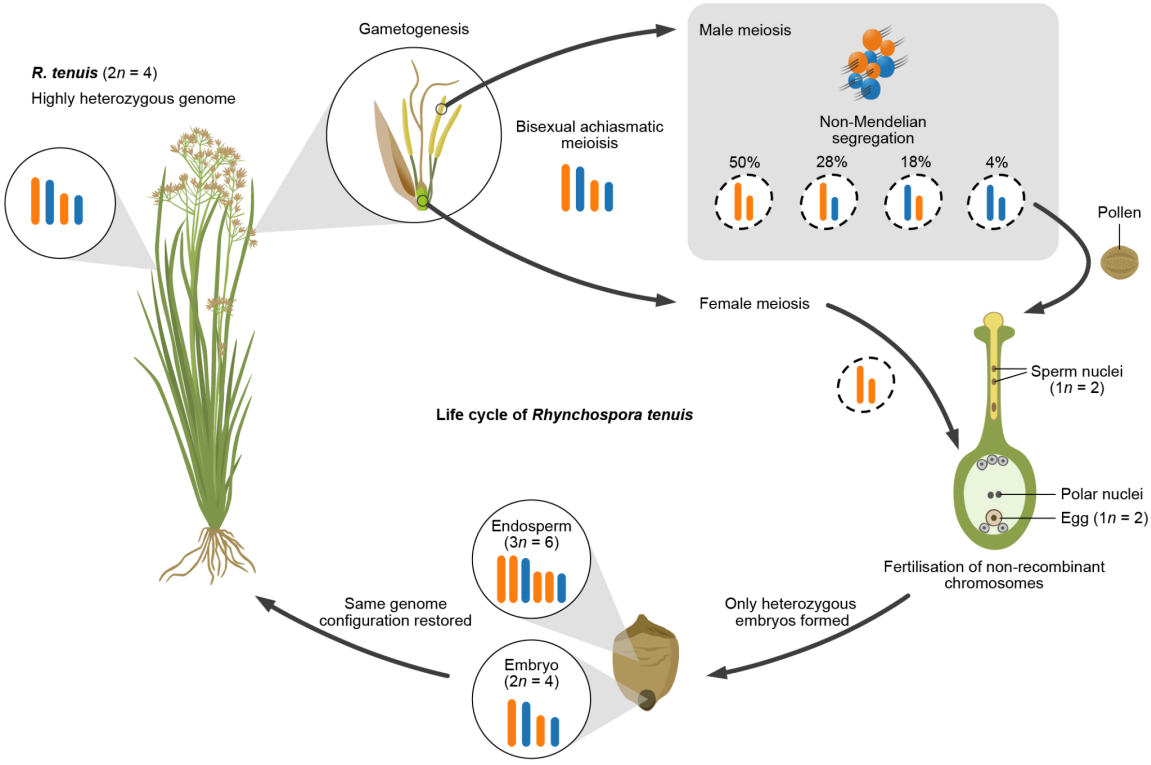


Fig. 5: Life cycle of *Rhynchospora tenuis* (i.e. PECP-48 genotype) under achiasmatic meiosis. Schematic representation of the *R. tenuis* life cycle ($2n = 4$). Both male and female meiosis occur without crossovers (achiasmatic inverted meiosis). Male gametogenesis is asymmetric, yielding a single functional sperm nucleus and non-Mendelian segregation. Fertilisation restores the diploid state with intact haplotypes, but only heterozygous seeds develop into viable embryos, ensuring faithful restoration of the highly heterozygous parental genome each generation.

Methods

Plant material

Plants of *R. tenuis* and *R. austrobrasiliensis* were collected from natural populations in Brazil under appropriate permits and cultivated under controlled greenhouse conditions. Nine geographically distinct accessions of *R. tenuis* were used for genome sequencing and cytological analyses. Individual flowers and pollen were staged for meiotic analyses. All experimental material was propagated clonally from field-collected individuals to ensure genetic identity across assays.

Sequencing

High-molecular-weight DNA was extracted from young leaf tissue using a modified CTAB protocol. PacBio HiFi libraries were prepared and sequenced on the Sequel IIe platform. Illumina short-read sequencing was used for genome polishing and transcriptome profiling. Hi-C libraries were constructed using Arima Hi-C kits and sequenced on an Illumina NovaSeq 6000 platform. RNA was extracted from young inflorescences for expression analyses.

DNA isolation

High-molecular-weight DNA from *R. tenuis* and *R. austrobrasiliensis* was isolated from 1.5 g of material with a NucleoBond HMW DNA kit (Macherey Nagel). Quality was assessed with a FEMTO-pulse device (Agilent), and quantity was measured with a Quantus fluorometer (Promega).

PacBio

HiFi libraries were prepared according to the "*Preparing whole genome and metagenome libraries using SMRTbell® prep kit 3.0*" manual, with an initial DNA fragmentation by Megaruptor-3 (Diagenode) and final size-selection by BluePippin (Sage Science). Size distribution was again controlled by FEMTO-pulse (Agilent). Size-selected libraries were then sequenced on a Revio device with Revio polymerase kit and Revio chemistry for 30 h (*Pacific Biosciences*).

Arima Hi-C

Plant tissues were cross-linked with 1% formaldehyde for 30 minutes at room temperature, and the reaction was quenched with 125 mM glycine for 10 minutes. Subsequently, the tissues were ground using a TissueLyser at a frequency of 30 Hz for 3 minutes. Nuclei extraction was performed using the CelLytic PN Plant Nuclei Isolation/Extraction Kit (Sigma-Aldrich, Burlington, MA, USA) according to the manufacturer's protocol. Hi-C libraries were prepared using the Arima High Coverage Hi-C Kit (Arima Genomics, A410110, Carlsbad, CA, USA) following the manufacturer's instructions, and were then sequenced paired-end (2 × 150 bp) on a NextSeq 2000 instrument (Illumina, San Diego, CA, USA).

Methyl-seq

To investigate the methylome space in *R. tenuis* (REC accession), the relatively non-destructive NEBNext Enzymatic Methyl-seq Kit was employed to prepare an Illumina-compatible library, followed by paired-end sequencing (2 × 150 bp) on a NextSeq2000 (Illumina) instrument. For each library, 10 Gb of reads was generated.

RNAseq

Total RNA was isolated from root, leaves, and flower buds (REC accession). Poly-A RNA was enriched from 1 µg total RNA using the NEBNext® Poly(A) mRNA Magnetic Isolation Module. RNAseq libraries were prepared as described in the NEBNext Ultra™ II Directional RNA Library Prep Kit for Illumina (New England Biolabs). A total of 11 cycles were applied to enrich library concentration. Sequencing was performed at BGI Genomics (Hong Kong) using a BGISEQ-500 system on the DNBseq platform in paired-end mode with a read length of 2 x 150 bp.

Illumina libraries (TPase and DNA FS)

Genomic DNA from 49 and 59 controlled self-crossed *R. tenuis* REC and PECP35-7 accessions, respectively, were deep-sequenced with an Illumina HiSeq 3000 or,

alternatively, with DNBseq short read sequencing (BGI Genomics, Hong Kong) in 150-bp paired-end mode.

Heterozygosity estimation

Heterozygosity levels of *R. tenuis*, *R. breviuscula,* and all *R. tenuis* accessions were estimated by GenomeScope2.0[52]. K-mer counting was performed using FASTK (v1.1) (https://github.com/thegenemyers/FASTK) with PacBio HiFi reads, which were used for genome assembly, as the input. Kmer size is 31 (Fastk -k31). The resulting binary output (.hist) from FASTK was then converted to a text histogram file by Histex (part of FASTK): Histex -G hifi_31mer.hist > hifi_31mer.histo. The histogram was then analysed with GenomeScope2 using k=31 and ploidy=2, except for *R. austrobrasiliensis*, where ploidy=6.

Genome assemblies

HiFi reads were assembled using Hifiasm v0.16.1 with default settings that combine both PAcBio HiFi reads and Hi-C reads to generate haplotype-resolved contigs. Assembly quality was assessed using BUSCO v5.2.2[53] and QUAST[54]. Duplicated haplotypes were retained for downstream analysis.

Scaffolding

Chromosome-scale scaffolding was performed using the Juicer and 3D-DNA pipelines[55,56]. Contigs from each haplotype generated from HiFiasm were taken as Hi-C alignment targets. Hi-C contact maps were visualised using Juicebox Assembly Tools to manually curate misjoins and validate the final pseudomolecules. Haplotypes were scaffolded independently and phased using *k*-mer-based alignment strategies.

Gene-based synteny analysis

Genes used for synteny analysis were annotated by a deep-learning-based tool Helixer (v0.3.4)[57,58] with the land plant mode. Gene synteny was analysed for *R. breviuscula* haplotype 1, *R. austrobrasiliensis* haplotype 1, and all haplotypes of nine *R. tenuis* accessions using GENESPACE (v1.3.1)[59] running in R (v4.2.0) along with the dependent tools OrthorFinder (v2.5.5)[60] and MCScanX (v1.0.0)[61]. The R script for running

GENESPACE (run_genespace.R) can be found in our project GitHub page: https://github.com/Raina-M/Rhynchospora_tenuis_project.

DNA sequence-based synteny analysis

Collinearity of all *R. tenuis* haplotypes was analysed by SyRI (v1.5.3)[62]. The alignment between haplotypes was conducted using minimap2 (v2.28)[63,64] with the following parameter settings: -ax asm5 --eqx. The plot was done by plotsr (v0.5.3) with a minimal 50kbp syntenic block size. NOR regions labelled on the haplotypes were searched by BLAST (v2.12.0) with the *Arabidopsis* rDNA sequences as the template.

Pangenome analysis

Pangenomes were generated using minigraph v0.17 with phased chromosome-scale assemblies from nine accessions. Variants were extracted using vg toolkit v1.40.0. Structural variants and sequence divergence were quantified across homologous haplotypes. Repeat annotation was performed with DANTE[65] and EDTA v2.0[66] using the plant TE library. Pairwise divergence and TE composition were visualised using custom R scripts.

ChIP-seq

ChIP was performed as described by Reimer and Turck[67] with minor modifications. Young leaves and flower buds were collected from greenhouse-grown plants and immediately frozen in liquid nitrogen before storage at −80 °C. Tissue samples were crosslinked with 4% formaldehyde under vacuum on ice for 1 hour, and the reaction was quenched by adding 1 M glycine. Nuclei were isolated using NIB buffer (50mM HEPES pH7.4, 5mM MgCl2, 25mM NaCl, 5% sucrose, 30% glycerol, 0.25% Triton X-100, 0.1 % β-mercaptoethanol and 0.1% protease inhibitor), and chromatin was extracted in TE-SDS buffer. Chromatin was sheared by sonication for 30 s ON/30 s OFF cycles for a total of 25 cycles. For immunoprecipitation, the sonicated chromatin was incubated overnight at 4 °C with 2 ng of the following antibodies: anti-CENH3[68], anti-H3K4me3 (Abcam, ab8580), anti-H3K9me2 (Abcam, ab1220), anti-H3K27me3 (Merck, 07-449). Control without antibody was performed simultaneously. The antibody–chromatin complexes were captured using rProtein A Sepharose® Fast Flow (Sigma,

GE17-1279-01) or Protein G Sepharose® 4 Fast Flow(Sigma, GE17-0618-01). Bound chromatin was eluted, and de-crosslinked by Proteinase K treatment. DNA was purified by ethanol/sodium acetate precipitation and resuspended in nuclease-free water. The purified DNA was sent for library preparation and sequencing.

Sequencing reads were aligned to the reference genome using Bowtie2[69] with the --very-sensitive-local option. ChIP and control alignments were compared using bamCompare[70] with the parameters --binSize 50 --normalizeUsing RPKM --operation log2. The resulting $\log_2$ ratio tracks were visualised using pyGenomeTracks[71]. Metaplots were generated using computeMatrix[70] (scale-regions mode) with the parameters --regionBodyLength 20000 --beforeRegionStartLength 10000 --afterRegionStartLength 10000 --binSize 50, and plotted using plotHeatmap.

Probes design and oligo-FISH

After phased genome assembly and synteny analysis, a haplotype-specific translocation in *R. tenuis* reference plant was identified at the chromosome ends. In order to validate this data, we designed oligo probes specific to each translocated region and hybridized in the reference plant, in different accessions, plus *R. austrobrasiliensis*. Probes were designed for the following regions to each chromosome/haplotype: for chromosome 1 (Chr1_h1: 0-3100219bp/ Chr1_h2: 0-2085189bp) in a density of 3 probes per Kilobase Pair (Kb) (**Fig. 1c**), and for chromosome 2 (Chr2_h1: 133059273bp - end/ Chr2_h2: 136628596bp - end) in a density of 1 probe per Kb. After designing and setting the regions corresponding to each haplotype, this data was submitted to Daicel Arbor Biosciences (Michigan, USA) for synthesis and labelling of the probes. Probes for haplotype one were labelled in Atto633 (far red), while probes for haplotype two were labelled in Alexafluor488 (green). For the slides preparation, young roots (mitosis) and flower buds (meiosis and microgametogenesis), from plants maintained in pots at greenhouse facilities of Max-Planck Institute for plant Breeding Research, were collected and fixed in Carnoy solution (methanol:acetic acid; 3:1, v/v) for 2-24h at room temperature then immediately used or kept at -20 ºC until the moment of use. After fixation, the material was digested in enzymatic solution containing 2% cellulase Onozuka (Serva), 2% pectolyase Y-23 (Duchefa Biochemie), 2% cyto-helicase (Sigma) and 10% pectinase (Sigma) in citrate-phosphate buffer (pH 4.5), washed in water and

macerated in 60% acetic acid, post-fixed with ice cold fresh Carnoy solution and air dried. For removal of cytoplasm, the slides were washed in 60% acetic acid for at least 30 minutes, then air dried. For the pollen tube germination experiment, 30 mature flower buds were collected into 3 ml of sdH_2O and vortexed to release pollen grains. From the bottom of the tube, 300 µl of the pollen suspension was collected and added to 3 ml of liquid medium for trinucleate pollen, following Tushabe and Rosbakh[72]. Pollen grains were incubated at 28 °C for 24 hours. After incubation, the medium containing germinating pollen grains was transferred into a conical tube and centrifuged at 5000 rpm for 3 minutes. After removing the supernatant, pollen nuclei were suspended in 60% AcAc, transferred onto a slide, post-fixed, and air-dried. For oligo-FISH, the procedures were similar to the ones published by Nascimento and Pedrosa-Harand[73], briefly the hybridization mix was composed of 50% formamide (Sigma), 2×SSC (Saline Sodium Citrate) solution (pH 7.0), 10% dextran sulphate, 350 ng of probe labelled in Alexa Fluor-488 and 200 ng of the probe labelled in Atto-633 (a total of 15 µL per slide). After denaturation and hybridisation, the slides were passed through stringency washes in 2× and 0.1× SSC at 42 ºC, which corresponds to a final stringency of ~76%. Slides were counterstained with 2 µg/mL DAPI in Vectashield H-100 antifade mounting medium (Vector Laboratories) and photographed on a Zeiss Microscope coupled with a Zeiss Axio Imager 2 camera system and software ZenBlue v3.2. Images were treated in brightness and contrast using Adobe Photoshop software.

Immunocytochemistry

Immunocytochemistry was performed as described in Castellani et al.[18], with some variations. Briefly, young flowers of *R. tenuis* and *R. austrobrasiliensis* were sampled and fixed in ice-cold 4% (w/v) paraformaldehyde (PFA) in phosphate buffered saline (PBS) solution (pH 7.5, 1.3 M NaCl, 70 mM Na2HPO4, 30 mM NaH2PO4) and 0.1% (v/v) Triton X-100 for 30 min in a vacuum. Flowers were dissected to select anthers of the appropriate size in order to have meiocytes in Prophase I. Anthers were opened from the tip in a drop of PBS with 0.1% (v/v) Triton X-100 and squeezed to release the meiocytes from the locules. The suspension of meiocytes was stirred to separate individual nuclei and remove excess debris. A coverslip was pressed onto the suspension to squash the

meiocytes and later removed with liquid nitrogen. Slides were mounted with Vectashield containing 0.2 µg/ml DAPI and checked for Prophase I stages of interest. Selected slides were incubated with blocking buffer (3% (w/v) bovine serum albumin (BSA) in PBS + 0.1% (v/v) Triton X-100) for 1 hour at 37 °C to block and permeabilize the cells. The following antibodies were used: anti-AtASY1 raised in rabbit (inventory code PAK006)[22], anti-AtMLH1 raised in rabbit (PAK017)[74]. The anti-RpZYP1 was raised in rat against the peptide KLTAERLVKDQASVKNDLEC (Gene ID: RP1G00482580/RP4G01479980/RP2G00810370/RP5G01738420) and affinity-purified (Lifeprotein). The anti-RpREC8 was a combination of two antibodies raised in rabbit against the peptides CEEPYGEIQISKGPNM and CYNPDDSVERMRDDPG (Gene ID: RP1G00316120/RP2G00915110/RP4G01319620/RP5G01638170) and affinity-purified (Eurogentec)[18]. The anti-RpHEI10 was a combination of four antibodies raised in rabbit and rat against the peptides CNRPNQSRARTNMFQL, CPVRQRNNKSMVSGGP, CIDIMDSRDMLRQGKREREEIW and CDTDSAVNMGPPSGDTSNRR (Gene ID: RP3G01271190/RP3G01008630/RP1G00269340/RP2G00699130) and affinity-purified (Eurogentec and Lifeprotein)[18]. Primary antibodies were diluted in the blocking buffer to a final volume ratio of 1:200. Slides were incubated with primary antibodies overnight at 4 °C. The following day, slides were washed three times for 5 min each with PBS + 0.1% (v/v) Triton X-100. Samples were incubated with secondary antibodies for 2h at room temperature or 1h at 37 °C. Secondary antibodies were conjugated with STAR ORANGE (Abberior, STORANGE-1007) or Alexa Fluor 488 (Thermofisher Goat anti-Rabbit IgG (H+L), Superclonal™ Recombinant Secondary Antibody, A27034) diluted 1:250 in blocking buffer. Slides were washed again three times for 5 min with PBS + 0.1% (v/v) Triton X-100 and allowed to dry. Samples were then prepared with 10 µl of mounting solution (Vectashield + 0.2 µg/ml DAPI). Specimens were covered with a coverslip and sealed with nail polish for storage. Images were taken with a Zeiss Axio Imager Z2 with Apotome system for optical sectioning. Images were deconvolved and processed with Zen 3.2 software and Adobe Photoshop.

Meiotic gene analysis

Meiotic genes with a role specifically in meiotic recombination were selected from the literature based on plants and other eukaryotes. Sequences were selected from

Arabidopsis or closer relatives in the case of poorly conserved genes (e.g. rice, maize, barley). Protein sequences were used as a query for tblastn with our in-house genomes, annotations and expression datasets. The number of hits on the same query sequence was used as a starting measure of copy number. Results were then manually curated for validation (**Supplementary Fig. 11**). Expression evidence for meiotic genes were based on the analysis of the inflorescence transcriptome obtained by RNA-seq data.

To further evaluate *SHOC1* transcriptional activity, total RNA from 100mg of flower buds was isolated using the Spectrum Plant Total RNA Kit (Sigma 000393619), followed by cDNA synthesis with the SuperScript IV Reverse Transcriptase (ThermoFisher 18090050). The obtained cDNA was used as a template for PCR with Phusion HF DNA Polymerase (NEB M0530), successfully amplifying the predicted coding sequence of *SHOC1*.

Pollen Nuclei Isolation and Sequencing

The protocol was adapted from Castellani, et al.[18]. For collecting the pollen, mature flowers of *Rhynchospora tenuis* were collected into 5ml tubes containing Woody Pollen Buffer (scWPB: 200mM Tris-HCl; 4mM $MgCl_2$; 2mM EDTA; 86mM NaCl; 10mM $Na_2S_2O_5$; 250mM Sucrose; 0,5mM Spermine; 0,5mM Spermidine; 1% PVP-10), and vortex for 30s at maximum speed. The obtained solution was filtered through a 50µm CellTrics cell strainer into a 5ml tube. Samples were centrifuged at 5000g for 5 minutes, the supernatant was removed, and the pollen pellet was frozen in liquid nitrogen. Pollen viability was assessed by Alexander staining (Morphisto 13441-00250). Pollen samples were stored at -70ºC until further use.

For extracting the pollen nuclei, frozen pollen samples were thawed on ice for 5 minutes and resuspended in scWPB. The solution was filtered through a 5µm CellTrics cell strainer, retaining the pollen. Pollen grains inside the cell strainer were crushed with a plastic rod. Pollen nuclei were filtered from the macerated pollen by adding scWPB with some additives to preserve RNA (scWPC; 5mM DTT; 1%BSA; 0.2 U/µL Protector RNase Inhibitor). Pollen nuclei solution was stained with DAPI (1µg/µl) and used for Fluorescence-activated cell sorting (FACS) to enrich the nuclei population. BD FACSAriaIII Fusion Flow Cytometer with a 70µm nozzle and 483 kPa sheath pressure

was employed for this purpose. Nuclei were dispensed into a 96-well plate containing Collection Buffer (1xPBS, 1%BSA, and 0.2 U/μL Protector RNase Inhibitor). The quality and number of nuclei were evaluated with the LUNA-FX7 Automated Cell Counter. Nuclei solutions were used for library preparation using Chromium Next GEM Single Cell ATAC or Chromium Next GEM Single Cell 5′ Kits following the manufacturer's instructions. Libraries sequenced at BGI Genomics following Chromium 10X Kit recommendations.

Single-cell analysis

Single-cell RNA data were analysed with a similar pipeline reported in the previous study in *R. breviuscula*[18]. A full description of the pipeline was updated in the GitHub page: https://github.com/Raina-M/detectCO_by_scRNAseq.

Subgenome-aware phasing of R. tenuis haplotypes

We used SubPhaser[20] (default parameters) to phase and partition the haplotypes of *R. tenuis* by assigning chromosomes to subgenomes based on differential repetitive *k*-mers. These were assumed to have expanded during the period of independent evolution after the loss of recombination. A subgenome is considered to be well phased when it displays distinct patterns of both differential *k*-mers and homoeologous chromosomes, confirming the presence of subgenome-specific features.

Estimating divergence times between haplotypes

Divergence times between haplotypes were estimated using the method outlined in Guo et al.[75]. In brief, all SNPs, derived from SyRI (v1.5.3)[62] analysis explained above, between two haplotypes were used to estimate divergence times according to the formula $g=d/2\mu$, where g is the number of generations, μ is the assumed mutation rate of 6.13×10^{-9} [76], and d is the number of SNPs per bp (*d=total number of SNPs between two haplotypes/chromosome size*). We assumed a generation time of one per year.

LTR insertion times were calculated by Subphaser as follows: LTR-TRs were de novo detected using LTRharvest (v.1.6.1)[77] and LTRfinder (v.1.07)[78]. To reduce false positives, TEsorter (v.1.3.0)[79] was used to reconstruct the classification of LTR-RTs and further refine this classification. The subgenome-specific *k*-mer sequences were mapped to the

LTR-RT sequences using a substring match procedure to identify the subgenome-specific LTR-RTs using Fisher's exact test. Two LTRs of each subgenome-specific LTR-RT were retrieved and the nucleotide divergence was estimated using the Jukes–Cantor 1969 model. The insertion time (T) was calculated using the equation $T = K/2r$, where $r = 6.13\times10^{-9}$ substitutions per year, and K represents the divergence of the LTRs from the LTR-RT.

Ks between two haplotypes was computed based on the CDS sequences annotated by Helixer by CoGe (https://ghibli.bti.cornell.edu/coge/). Genomes and gene annotations were uploaded to CoGe, and the option to calculate *Ks* was selected during the SynMap run. All *Ks* of coding sequences that overlapped with transposable elements annotation was removed by 'bedtools subtract -A'.

Phylogenetic tree construction

The phylogenetic tree was built based on single copy orthologs (SCOs) among *R. breviuscula*, *R. austrobrasiliensis*, and all haplotypes of *R. tenius*. SCOs were extracted form the output of OrthoFinder in GENESPACE. All SCO groups were aligned by MAFFT (v7.475)[80] with parameter '--auto –maxiterate 10'. All alignements were concatenated with a custom python script available in our project GitHub page (phylogeny_based_on_SCOs/concatenate_aln.py). The tree was constructed based on the catenated alignments with RAxML-NG (v1.2.2)[81], then visulised in FigTree (v1.4.4) (https://github.com/rambaut/figtree/).

F1 genotyping and analysis

F1 individuals were obtained from selfed heterozygous mother plants. Genomic DNA from seedlings was extracted and sequenced on an Illumina NovaSeq platform. Reads were mapped to haplotype 1 assembly, and SNP genotypes were called using bcftools (v1.9). Progeny haplotypes were reconstructed and compared to maternal alleles to assess inheritance patterns and identify crossover events. F1 genotyping method was the same as Castellani et al.[18].

Embryo and endosperm analysis

For extracting the nuclei from embryo and endosperm, frozen seeds were macerated with a plastic rod and resuspended in Seed Nuclei Isolation Buffer (100mM Tris, 5.2 mM MgCl2, 85.55 mM NaCl, 0.1% Triton, 0.2 U/µL Protector RNase Inhibitor, pH 7.5). The obtained solution was filtered through a 20µm CellTrics cell strainer into a fresh 2ml tube and stained with DAPI (1 µg/µl). BD FACSAriaIII Fusion Flow Cytometer with a 70µm nozzle was employed for FACS to enrich the 2c (embryo) and 3c (endosperm) nuclei population. The separate 2c and 3c nuclei populations were dispensed into a 96-well plate containing Collection Buffer (1xPBS, 1%BSA, and 0,2 U/µL Protector RNase Inhibitor). The quality and number of nuclei were evaluated with the LUNA-FX7 Automated Cell Counter. Nuclei solutions were used for library preparation using Chromium Next GEM Single Cell ATAC following the manufacturer's instructions. Libraries sequenced at BGI Genomics using the Chromium 10X Kit recommendations..

Female gametophyte sectioning and imaging

Developing ovaries and whole inflorescences were fixed according to Rocha et al.[29]. Dehydration and infiltration into Araldite 502/Embed 812 resin were performed in an EMS LYNX II automated tissue processor (EMS, Hatfield, PA, USA). Following resin polymerisation, serial semithin sections (1 µm) were produced from ovules at different stages of development, stained with 1% aqueous toluidine blue supplemented with 1% sodium tetraborate and imaged with a Zeiss Axio Imager M2. Images were stacked and aligned in FIJI using the StackReg plugin (BIG-EPFL). MorphoGrahX was used for 3D rendering of the aligned images and volumetric visualization of the embryo sac tissue[82]. 3D segmentation of the embryo sac was performed using ITK watershed (https://www.itk.org) in MorphoGraphX, taking PlantSeg[83] predictions from “generic confocal 3DUNET model” as input images for 3D cell segmentation. Immunodetection of ß-1,3-glucan to identify pollen tube cell walls was done as described in Freh et al.[84], with the exception that BSA was omitted from the buffer during all steps. Goat anti-mouse Alexa Fluor 647 (abcam ab150119, Cambridge, UK) was used as an alternative secondary antibody. Imaging was done on a Leica SP8.

For Feulgen staining of siliques, the pericarp was removed from the mature and fertile siliques; for the rest of the siliques, the pericarp was not removed, and the material was

fixed in ethanol:acetic acid (3:1) overnight. The samples were washed three times with distilled water for 15 min each wash and then incubated in 5 N HCl for 1 h, followed by another series of three washes with water, as before. The siliques were then incubated in Schiff's reagent for 3–4 h, after which they were washed three times with cold distilled water (4°C). Next, the siliques were incubated in ethanol series 10 min each, 10%, 30%, 50%, 70%, 95% and then washed several times with 99.5% ethanol until the ethanol came out colorless. The samples were incubated in 99.5% Ethanol:LR White series 15 minutes each solution 3:1, 2:1. The samples were then incubated for 1 hour in ethanol:LR White resin (1:1), followed by an overnight incubation in LR White resin. Afterwards, the seeds were mounted on microscope slides in LR White resin and polymerized for 24h at 60°C. The samples were imaged under Multiphoton microscopy, Leica Stellaris 8 and Leica SP8 Falcon Dive with emission around 580 nm and excitation between 580 nm - 700 nm.

Data availability

Genome assemblies and raw sequencing data have been deposited in the European Nucleotide Archive under accession PRJEB98016. The reference genomes, sequencing data, annotations and all tracks presented in this work are made available for download at EDMOND, the Open Research Data Repository of the Max Planck Society, https://doi.org/10.17617/3.IXRT5Y. All other data are available from the corresponding author upon request.

Contributions

A.M. conceived the research and supervised the project. M.Z., P.G.H. and A.M. performed the genome assembly, scaffolding and genomics. M.Z., J.C., H.S., M.C., and S.S. designed and performed the single-cell experiments and processed the data. K.S. assisted the single-cell analysis. M.C. performed the immunostaining experiments. L.D. and M.C. performed the functional screening of meiotic genes. T.N. and M.Majka performed FISH experiments. M.Majka performed pollen-FISH genotyping experiments. G.T. and M.Z. performed the ChIP–seq analysis. M.Marek and S.S. performed the flow

cytometry measurements and nuclei sorting. N.S., M.Marek and B.H. performed all nuclear sequencing libraries. M.Z., L.A.R., A.L.L.V. and A.M. performed the centromere and repeat characterisation. A.L.L.V. collected and provided the plant material. U.P. and M.Majka performed the plant crosses. S.S. performed the pollen viability assay. M.Z., S.D. and A.M. performed the dating analysis. M.Z., T.L. and K.F.X.M. performed the functional annotation. G.T., U.N., A.V., T.T. and D.F. performed and analysed the female gametophyte and embryo sac development microscopy data. M.Z., M.C., S.S. and A.M. wrote the first manuscript draft with input from all authors. All authors approved the final version of the manuscript.

Acknowledgements

This study was funded by the Max Planck Society (core funding to A.M.), the German Research Foundation (MA 9363/2-1 and MA 9363/3-1), and the European Union (European Research Council Starting Grant, HoloRECOMB, grant no. 101114879 to A.M.). The DFG also funded this work under Germany's Excellence Strategy—EXC 493 2048/1–390686111 (to K.S. and A.M.). M.Z. is financially supported by the DFG (grant no. MA 9363/2-1). M.Majka is supported by the Alexander von Humboldt Foundation (Humboldt Research Fellowship for Postdoctoral Researchers). G.T. is a recipient of the predoctoral fellowship HORIZON-MSCA-2021-COFUND-01, rePLANT GA 101081581, funded by the European Union and co-funded by the Max Planck Society. The views and opinions expressed are those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Executive Agency (REA). Neither the European Union nor the granting authority can be held responsible for them.

References

1. Ohkura, H. Meiosis: An Overview of Key Differences from Mitosis. *Cold Spring Harb. Perspect. Biol.* **7**, a015859 (2015).
2. Satomura, K., Osada, N. & Endo, T. Achiasmy and sex chromosome evolution. *Ecol. Genet. Genomics* **13**, 100046 (2019).
3. Cabral, G., Marques, A., Schubert, V., Pedrosa-Harand, A. & Schlögelhofer, P. Chiasmatic and achiasmatic inverted meiosis of plants with holocentric chromosomes. *Nat. Commun.* **5**, 5070 (2014).
4. Lenormand, T. & Dutheil, J. Recombination Difference between Sexes: A Role for Haploid Selection. *PLoS Biol.* **3**, e63 (2005).
5. Stapley, J., Feulner, P. G. D., Johnston, S. E., Santure, A. W. & Smadja, C. M. Variation in recombination frequency and distribution across eukaryotes: patterns and processes. *Philos. Trans. R. Soc. B Biol. Sci.* **372**, 20160455 (2017).
6. Morgan, T. H. Sex Limited Inheritance in *Drosophila*. *Science* **32**, 120–122 (1910).
7. Thomas, S. E. *et al.* Identification of Two Proteins Required for Conjunction and Regular Segregation of Achiasmate Homologs in Drosophila Male Meiosis. *Cell* **123**, 555–568 (2005).
8. Hawley, R. S. Meiosis: How Male Flies Do Meiosis. *Curr. Biol.* **12**, R660–R662 (2002).
9. Hughes, S. E. *et al.* Heterochromatic Threads Connect Oscillating Chromosomes during Prometaphase I in Drosophila Oocytes. *PLoS Genet.* **5**, e1000348 (2009).

10. Xiang, Y. *et al.* Multiple reorganizations of the lateral elements of the synaptonemal complex facilitate homolog segregation in Bombyx mori oocytes. *Curr. Biol.* **34**, 352-360.e4 (2024).
11. Ceplitis, A. Coalescence times and the Meselson effect in asexual eukaryotes. *Genet Res* **82**, 183–90 (2003).
12. Mark Welch, D. B. & Meselson, M. Evidence for the Evolution of Bdelloid Rotifers Without Sexual Reproduction or Genetic Exchange. *Science* **288**, 1211–1215 (2000).
13. Weir, W. *et al.* Population genomics reveals the origin and asexual evolution of human infective trypanosomes. *eLife* **5**, e11473 (2016).
14. Butlin, R. The costs and benefits of sex: new insights from old asexual lineages. *Nat. Rev. Genet.* **3**, 311–317 (2002).
15. Vanzela, A. L. L. Rhynchospora tenuis Link (Cyperaceae): a species with the lowest number of holocentric chromosomes (n=2). *1996* **88**, 219–228 (1996).
16. Schubert, I. & Lysak, M. A. Interpretation of karyotype evolution should consider chromosome structural constraints. *Trends Genet.* **27**, 207–216 (2011).
17. Hofstatter, P. G. *et al.* Repeat-based holocentromeres influence genome architecture and karyotype evolution. *Cell* **185**, 3153–3168 (2022).
18. Castellani, M. *et al.* Meiotic recombination dynamics in plants with repeat-based holocentromeres shed light on the primary drivers of crossover patterning. *Nat. Plants* **10**, 423–438 (2024).
19. Zhang, M. Karyotype shuffling drives crossover patterning through changes in synapsis dynamics and chromatin restructuring. *submitted* (2025).

20. Jia, K.-H. *et al.* SubPhaser: a robust allopolyploid subgenome phasing method based on subgenome-specific k-mers. *New Phytol.* **235**, 801–809 (2022).
21. Wang, L. *et al.* The architecture of intra-organism mutation rate variation in plants. *PLOS Biol.* **17**, e3000191 (2019).
22. Armstrong, S. J., Caryl, A. P., Jones, G. H. & Franklin, F. C. Asy1, a protein required for meiotic chromosome synapsis, localizes to axis-associated chromatin in Arabidopsis and Brassica. *J Cell Sci* **115**, 3645–55 (2002).
23. Lambing, C. *et al.* Interacting Genomic Landscapes of REC8-Cohesin, Chromatin, and Meiotic Recombination in Arabidopsis[CC-BY]. *Plant Cell* **32**, 1218–1239 (2020).
24. Higgins, J. D., Sanchez-Moran, E., Armstrong, S. J., Jones, G. H. & Franklin, F. C. The Arabidopsis synaptonemal complex protein ZYP1 is required for chromosome synapsis and normal fidelity of crossing over. *Genes Dev* **19**, 2488–500 (2005).
25. Chelysheva, L. *et al.* The Arabidopsis HEI10 Is a New ZMM Protein Related to Zip3. *PLOS Genet.* **8**, e1002799 (2012).
26. Michelan, V. S. *et al.* Morphological and genomic characterization of Rhynchospora tenuis complex (Cyperaceae) and its taxonomic implications. *Rodriguésia* **63**, 775–784 (2012).
27. Guiraldelli, M. F. *et al.* SHOC1 is a ERCC4-(HhH)2-like protein, integral to the formation of crossover recombination intermediates during mammalian meiosis. *PLOS Genet.* **14**, e1007381 (2018).
28. Macaisne, N. *et al.* SHOC1, an XPF Endonuclease-Related Protein, Is Essential for the Formation of Class I Meiotic Crossovers. *Curr. Biol.* **18**, 1432–1437 (2008).

29. Rocha, D. M. *et al.* Cryoimmobilized anther analysis reveals new ultrastructural insights into Rhynchospora (Cyperaceae) asymmetrical microsporogenesis. *Front. Plant Sci.* **15**, 1518369 (2025).
30. Rocha, D. M. *et al.* Developmental programmed cell death during asymmetric microsporogenesis in holocentric species ofRhynchospora(Cyperaceae). *J. Exp. Bot.* **67**, 5391–5401 (2016).
31. San Martin, J. A. B., De Jesus Andrade, C. G. T., Mastroberti, A. A., De Araújo Mariath, J. E. & Vanzela, A. L. L. Asymmetric cytokinesis guide the development of pseudomonads in *Rhynchospora pubera* (Cyperaceae). *Cell Biol. Int.* **37**, 203–212 (2013).
32. Marques, A., Schubert, V., Houben, A. & Pedrosa-Harand, A. Restructuring of Holocentric Centromeres During Meiosis in the Plant Rhynchospora pubera. *Genetics* **204**, 555–568 (2016).
33. Wang, Y. & Underwood, C. J. Apomixis. *Curr. Biol.* **33**, R293–R295 (2023).
34. Keeney, S. Spo11 and the Formation of DNA Double-Strand Breaks in Meiosis. *Genome Dyn Stab* **2**, 81–123 (2008).
35. Mercier, R., Mézard, C., Jenczewski, E., Macaisne, N. & Grelon, M. The molecular biology of meiosis in plants. *Annu Rev Plant Biol* **66**, 297–327 (2015).
36. Blanc, C. *et al.* Cosegregation of recombinant chromatids maintains genome-wide heterozygosity in an asexual nematode. *Sci. Adv.* **9**, (2023).
37. Lacy, K. D., Hart, T. & Kronauer, D. J. C. Co-inheritance of recombined chromatids maintains heterozygosity in a parthenogenetic ant. *Nat. Ecol. Evol.* **8**, 1522–1533 (2024).

38. Renner, O. Zur Kenntnis der nichtmendelnden Buntheit der Laubblätter). *Flora Oder Allg. Bot. Ztg.* **130**, 218-290.e1 (1936).

39. Harte, C. Propagation Systems and Evolution in Oenothera. in *Oenothera* vol. 20 221–222 (Springer Berlin Heidelberg, Berlin, Heidelberg, 1994).

40. Rauwolf, U., Golczyk, H., Meurer, J., Herrmann, R. G. & Greiner, S. Molecular Marker Systems for Oenothera Genetics. *Genetics* **180**, 1289–1306 (2008).

41. Hollister, J. D., Greiner, S., Johnson, M. T. J. & Wright, S. I. Hybridization and a loss of sex shape genome‐wide diversity and the origin of species in the evening primroses (*Oenothera* , Onagraceae). *New Phytol.* **224**, 1372–1380 (2019).

42. Feng, S., Pozzi, A., Opit, G., Stejskal, V. & Li, Z. Diploid genomes of the booklouse reveal evolutionary consequences of asexuality. Preprint at https://doi.org/10.1101/2025.09.02.673870 (2025).

43. Brand, J. N. *et al.* Stepwise emergence of recombination suppression precedes fissiparous asexuality in the planarian *Schmidtea mediterranea*. Preprint at https://doi.org/10.1101/2025.09.12.675808 (2025).

44. Bureš, P. & Zedek, F. HOLOKINETIC DRIVE: CENTROMERE DRIVE IN CHROMOSOMES WITHOUT CENTROMERES: BRIEF COMMUNICATION. *Evolution* n/a-n/a (2014) doi:10.1111/evo.12437.

45. Kursel, L. E. & Malik, H. S. The cellular mechanisms and consequences of centromere drive. *Curr. Opin. Cell Biol.* **52**, 58–65 (2018).

46. Henikoff, S., Ahmad, K. & Malik, H. S. The Centromere Paradox: Stable Inheritance with Rapidly Evolving DNA. *Science* **293**, 1098–1102 (2001).

47. Pachamuthu, K. & Borges, F. Epigenetic control of transposons during plant reproduction: From meiosis to hybrid seeds. *Curr. Opin. Plant Biol.* **75**, 102419 (2023).
48. Betancourt, A. J., Wei, K. H.-C., Huang, Y. & Lee, Y. C. G. Causes and Consequences of Varying Transposable Element Activity: An Evolutionary Perspective. *Annu. Rev. Genomics Hum. Genet.* **25**, 1–25 (2024).
49. Otto, S. P. & Lenormand, T. Resolving the paradox of sex and recombination. *Nat. Rev. Genet.* **3**, 252–261 (2002).
50. Veller, C., Kleckner, N. & Nowak, M. A. A rigorous measure of genome-wide genetic shuffling that takes into account crossover positions and Mendel's second law. *Proc. Natl. Acad. Sci.* **116**, 1659–1668 (2019).
51. Melamed-Bessudo, C., Shilo, S. & Levy, A. A. Meiotic recombination and genome evolution in plants. *Curr. Opin. Plant Biol.* **30**, 82–87 (2016).
52. Ranallo-Benavidez, T. R., Jaron, K. S. & Schatz, M. C. GenomeScope 2.0 and Smudgeplot for reference-free profiling of polyploid genomes. *Nat. Commun.* **11**, 1432 (2020).
53. Seppey, M., Manni, M. & Zdobnov, E. M. BUSCO: Assessing Genome Assembly and Annotation Completeness. in *Gene Prediction* (ed. Kollmar, M.) vol. 1962 227–245 (Springer New York, New York, NY, 2019).
54. Gurevich, A., Saveliev, V., Vyahhi, N. & Tesler, G. QUAST: quality assessment tool for genome assemblies. *Bioinformatics* **29**, 1072–1075 (2013).
55. Durand, N. C. *et al.* Juicebox Provides a Visualization System for Hi-C Contact Maps with Unlimited Zoom. *Cell Syst.* **3**, 99–101 (2016).

56. Dudchenko, O. *et al.* De novo assembly of the *Aedes aegypti* genome using Hi-C yields chromosome-length scaffolds. *Science* **356**, 92–95 (2017).

57. Holst, F. *et al.* Helixer– *de novo* Prediction of Primary Eukaryotic Gene Models Combining Deep Learning and a Hidden Markov Model. Preprint at https://doi.org/10.1101/2023.02.06.527280 (2023).

58. Stiehler, F. *et al.* Helixer: cross-species gene annotation of large eukaryotic genomes using deep learning. *Bioinformatics* **36**, 5291–5298 (2021).

59. Lovell, J. T. *et al.* GENESPACE tracks regions of interest and gene copy number variation across multiple genomes. *eLife* **11**, e78526 (2022).

60. Emms, D. M. & Kelly, S. OrthoFinder: phylogenetic orthology inference for comparative genomics. *Genome Biol.* **20**, 238 (2019).

61. Wang, Y. *et al.* MCScanX: a toolkit for detection and evolutionary analysis of gene synteny and collinearity. *Nucleic Acids Res.* **40**, e49–e49 (2012).

62. Goel, M., Sun, H., Jiao, W.-B. & Schneeberger, K. SyRI: finding genomic rearrangements and local sequence differences from whole-genome assemblies. *Genome Biol.* **20**, 277 (2019).

63. Li, H. New strategies to improve minimap2 alignment accuracy. *Bioinformatics* **37**, 4572–4574 (2021).

64. Li, H. Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics* **34**, 3094–3100 (2018).

65. Novák, P., Hoštáková, N., Neumann, P. & Macas, J. DANTE and DANTE_LTR: lineage-centric annotation pipelines for long terminal repeat retrotransposons in plant genomes. *NAR Genomics Bioinforma.* **6**, lqae113 (2024).

66. Ou, S. *et al.* Benchmarking transposable element annotation methods for creation of a streamlined, comprehensive pipeline. *Genome Biol.* **20**, (2019).
67. Genome-Wide Mapping of Protein-DNA Interaction by Chromatin Immunoprecipitation and DNA Microarray Hybridization (ChIP-chip). Part A: ChIP-chip Molecular Methods. in *Methods in Molecular Biology* 139–160 (Humana Press, Totowa, NJ, 2010). doi:10.1007/978-1-60761-646-7_12.
68. Marques, A. *et al.* Holocentromeres in Rhynchospora are associated with genome-wide centromere-specific repeat arrays interspersed among euchromatin. **112**, 13633–13638 (2015).
69. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359 (2012).
70. Ramírez, F. *et al.* deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res.* **44**, W160–W165 (2016).
71. Lopez-Delisle, L. *et al.* pyGenomeTracks: reproducible plots for multivariate genomic datasets. *Bioinformatics* **37**, 422–423 (2021).
72. Tushabe, D. & Rosbakh, S. A Compendium of in vitro Germination Media for Pollen Research. *Front. Plant Sci.* **12**, 709945 (2021).
73. Nascimento, T. & Pedrosa-Harand, A. High rates of structural rearrangements have shaped the chromosome evolution in dysploid Phaseolus beans. *Theor. Appl. Genet.* **136**, 215 (2023).
74. Chelysheva, L. *et al.* An Easy Protocol for Studying Chromatin and Recombination Protein Dynamics during Arabidopsisthaliana Meiosis: Immunodetection of Cohesins, Histones and MLH1. *Cytogenet. Genome Res.* **129**, 143–153 (2010).

75. Guo, Y. *et al.* A haplotype-based evolutionary history of barley domestication. *Nature* https://doi.org/10.1038/s41586-025-09533-7 (2025) doi:10.1038/s41586-025-09533-7.

76. Wang, L. *et al.* The architecture of intra-organism mutation rate variation in plants. *PLOS Biol.* **17**, e3000191 (2019).

77. Ellinghaus, D., Kurtz, S. & Willhoeft, U. LTRharvest, an efficient and flexible software for de novo detection of LTR retrotransposons. *BMC Bioinformatics* **9**, 18 (2008).

78. Xu, Z. & Wang, H. LTR_FINDER: an efficient tool for the prediction of full-length LTR retrotransposons. *Nucleic Acids Res.* **35**, W265-268 (2007).

79. Zhang, R.-G. *et al.* TEsorter: An accurate and fast method to classify LTR-retrotransposons in plant genomes. *Hortic. Res.* **9**, uhac017 (2022).

80. Kuraku, S., Zmasek, C. M., Nishimura, O. & Katoh, K. aLeaves facilitates on-demand exploration of metazoan gene family trees on MAFFT sequence alignment server with enhanced interactivity. *Nucleic Acids Res.* **41**, W22–W28 (2013).

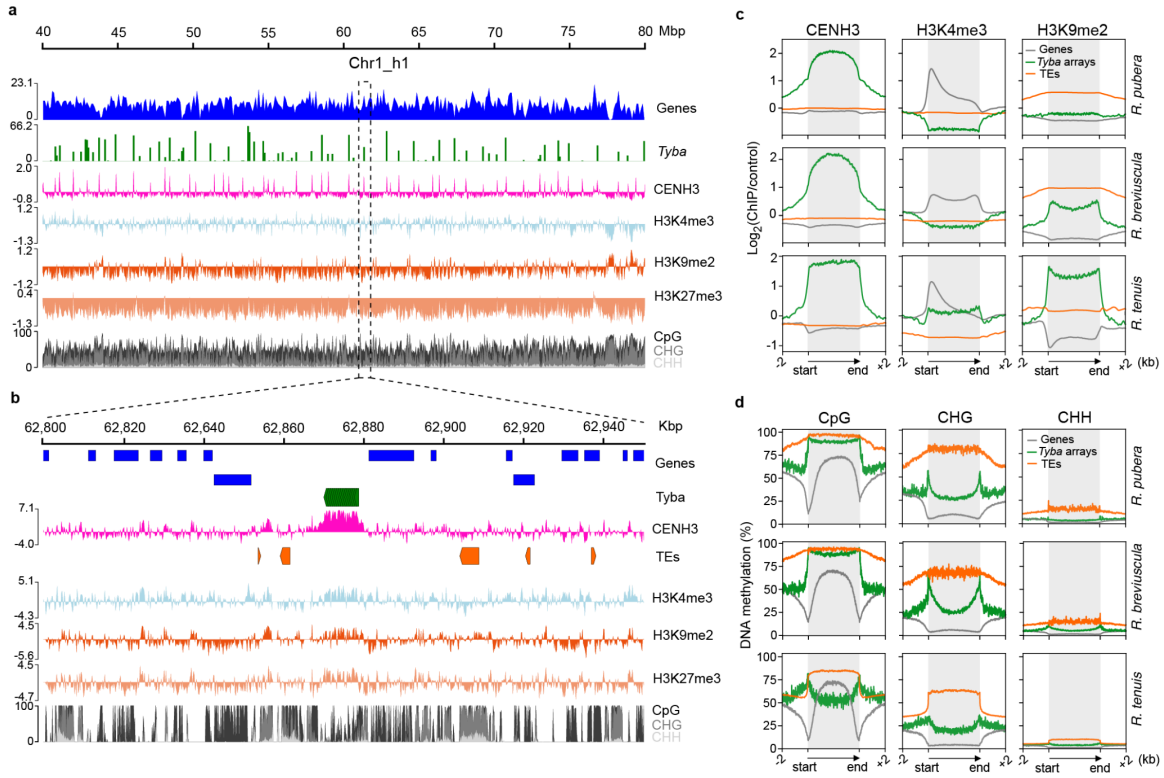
81. Stamatakis, A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30**, 1312–1313 (2014).

82. Barbier De Reuille, P. *et al.* MorphoGraphX: A platform for quantifying morphogenesis in 4D. *eLife* **4**, e05864 (2015).

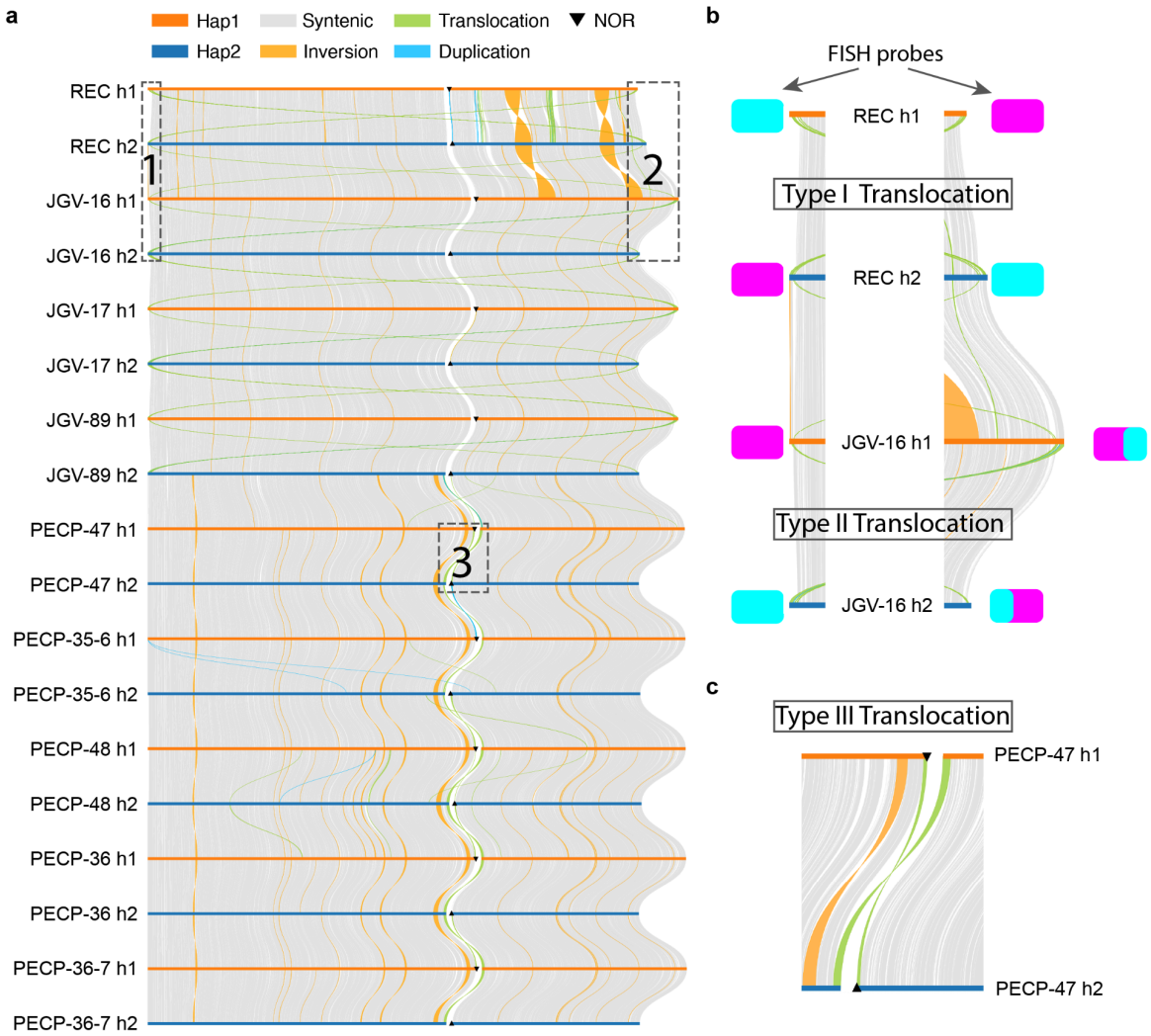
83. Wolny, A. *et al.* Accurate and versatile 3D segmentation of plant tissues at cellular resolution. *eLife* **9**, e57613 (2020).

84. Freh, M., Reinstädler, A., Neumann, K. D., Neumann, U. & Panstruga, R. The development of pleiotropic phenotypes in powdery mildew‐resistant barley and *Arabidopsis thaliana mlo* mutants is linked to nitrogen availability. *Plant Cell Environ.* **47**, 2360–2374 (2024).

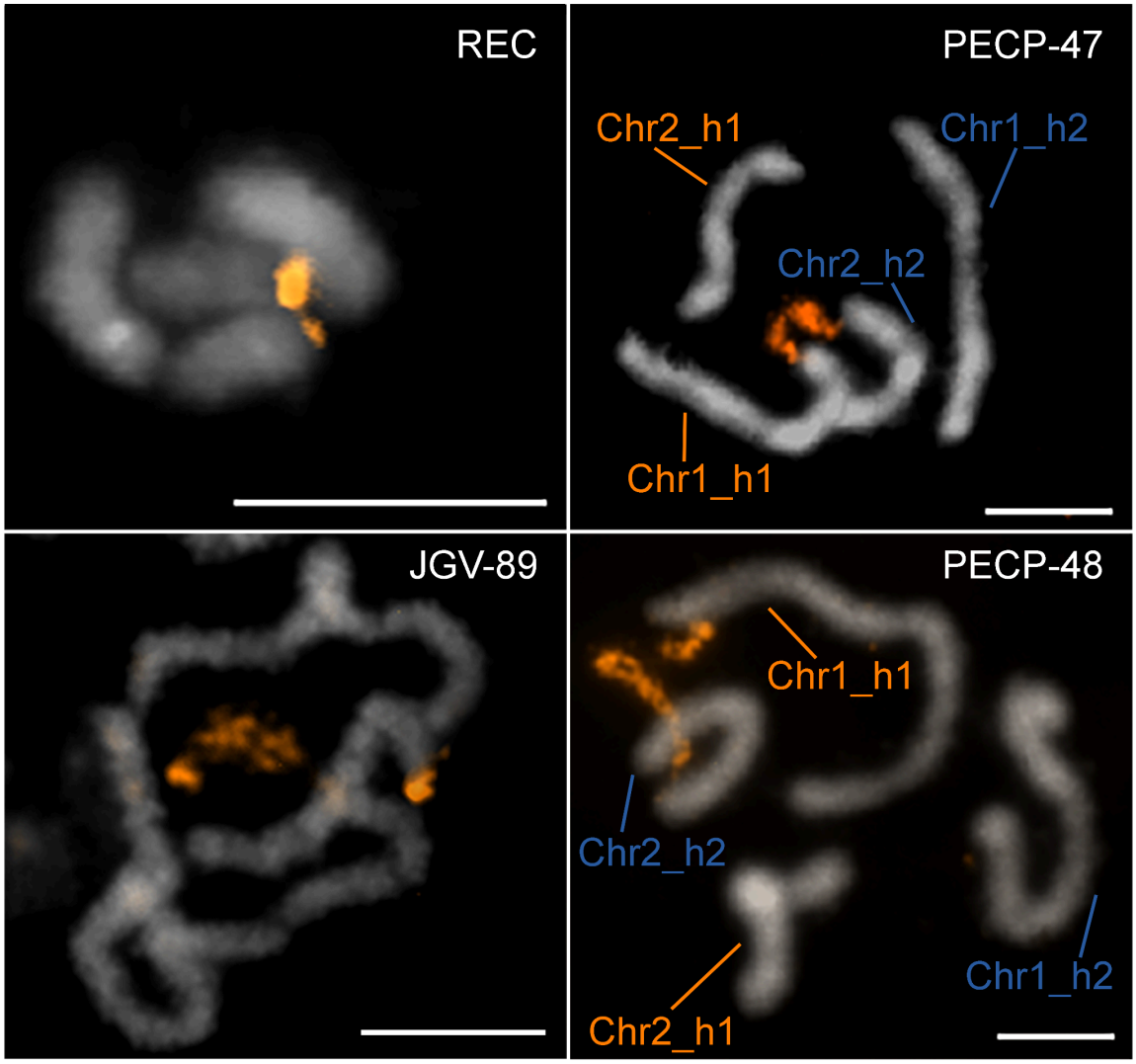
Extended Data Figures



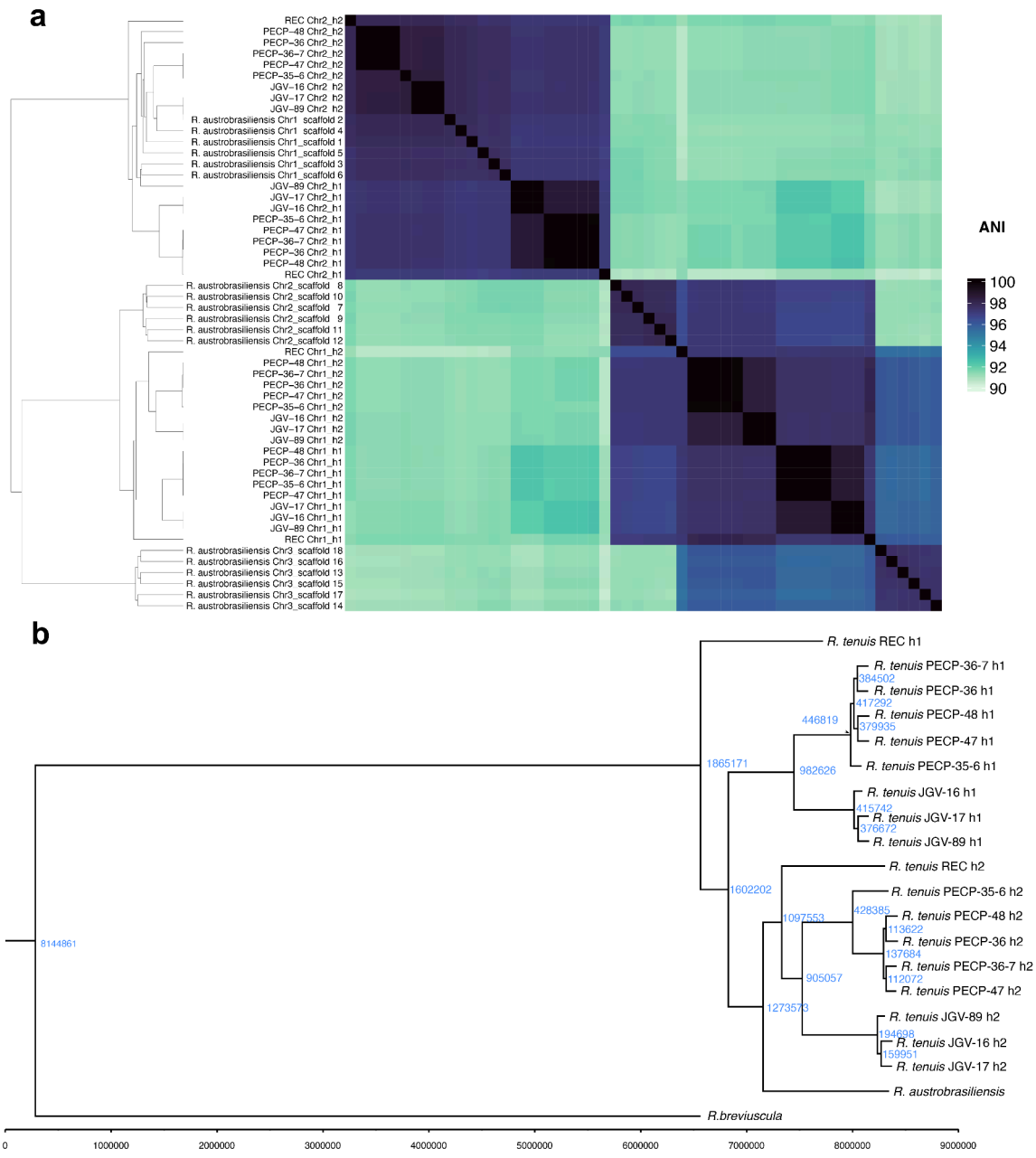
Extended Data Fig. 1: Divergent epigenetic landscape of holocentromeres in *Rhynchospora tenuis*. (**a**) Zoomed-in view of *R. tenuis* Chr1_h1 showing a 40-Mb region with multiple CENH3 domains that are closely correlated with *Tyba* repeat distribution. Gene and *Tyba* densities were calculated over 100-kb windows. (**b**) A detailed view of a representative ~150 kb holocentromeric region on *R. tenuis* Chr1_h1, illustrating the precise correlation between CENH3 binding, *Tyba* arrays and an ambiguous chromatin state, with low CpG DNA methylation and accumulation of both repressive and active histone marks. Values in **a** and **b** represent the log2 ratio of ChIP-seq signal to input control. (**c**) Enrichment of CENH3, H3K4me3, and H3K9me2 from the start and end of different types of sequences: genes (gray line), TEs (orange) and *Tyba* repeats (green) for *R. pubera*, *R. breviuscula* and *R. tenuis*. ChIP-seq signals are shown as $\log_2$ (normalized RPKM ChIP/input). (**d**) Enrichment of DNA methylation in the CpG, CHG, and CHH contexts for the same sequence types as shown in (**c**) for *R. pubera*, *R. breviuscula* and *R. tenuis*. Gray boxes in (**c–d**) highlight the modification enrichment over the body of each sequence type.



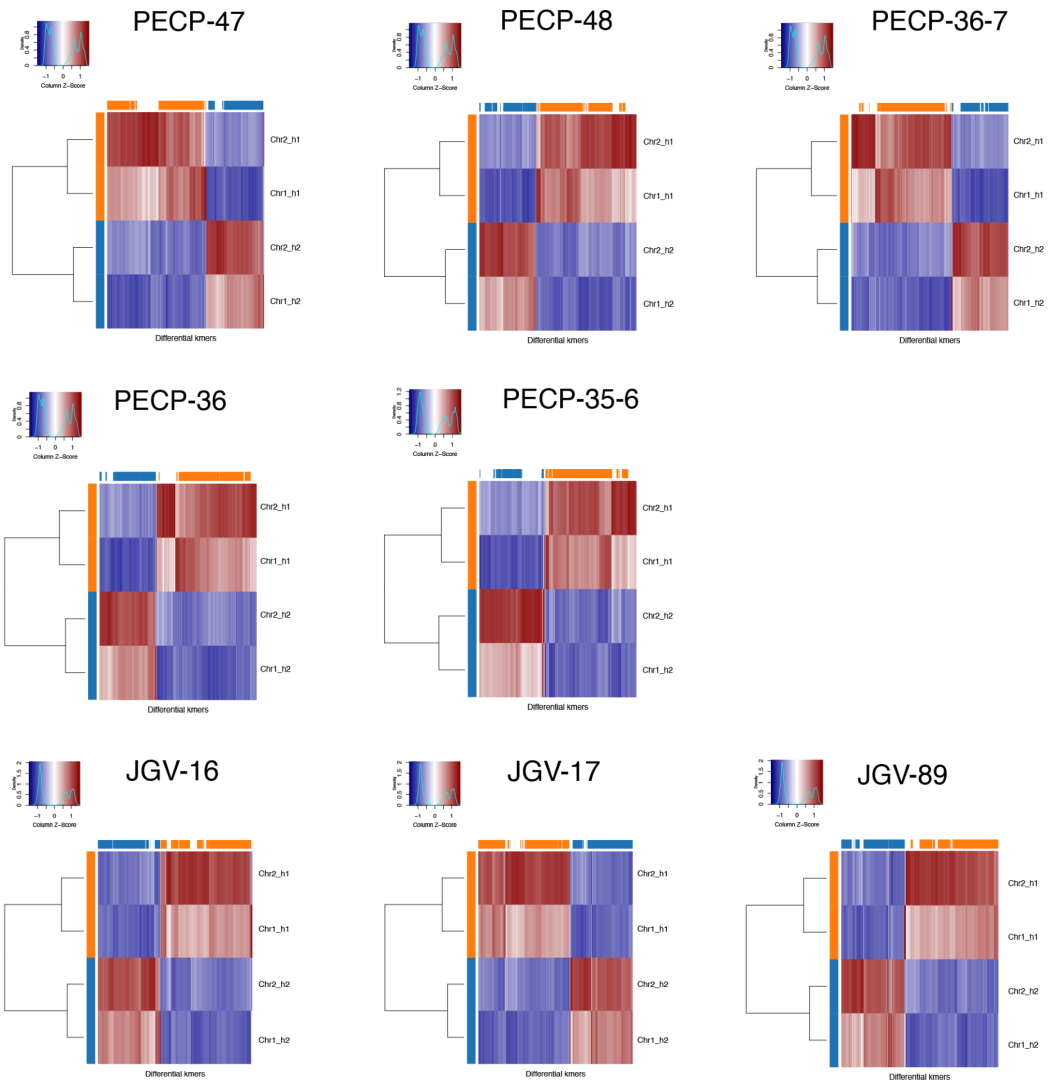
Extended Data Fig. 2: Colinearity across the two haplotypes of all *R. tenuis* accessions by SyRI. (**a**) Colinearity based on genomic sequences of all haplotypes of nice *R. tenuis* accessions (**b**) Zoom-in view of the regions 1 (left column) and 2 (right column) labeled in **a**. The translocation FISH probes designed for REC were also shown here, implying the Type I translocation. JGV-16 Chr 2 end (region 2) has a small translocation compared to REC Chr1_h2, so a mixed magenta and cyan probe signal was observed, indicating a different type of reciprocal translocation (Type II translocation). (**b**) Zoom-in view of the regions 3 labeled in **a**. (**c**) This displays the Type III translocation that found in PECP accessions, where the first reciprocal translocation is between Chr1_h1 end and Chr2_h2 start, and the second reciprocal translocation is between Chr1_h2 end and Chr2_h1 start.



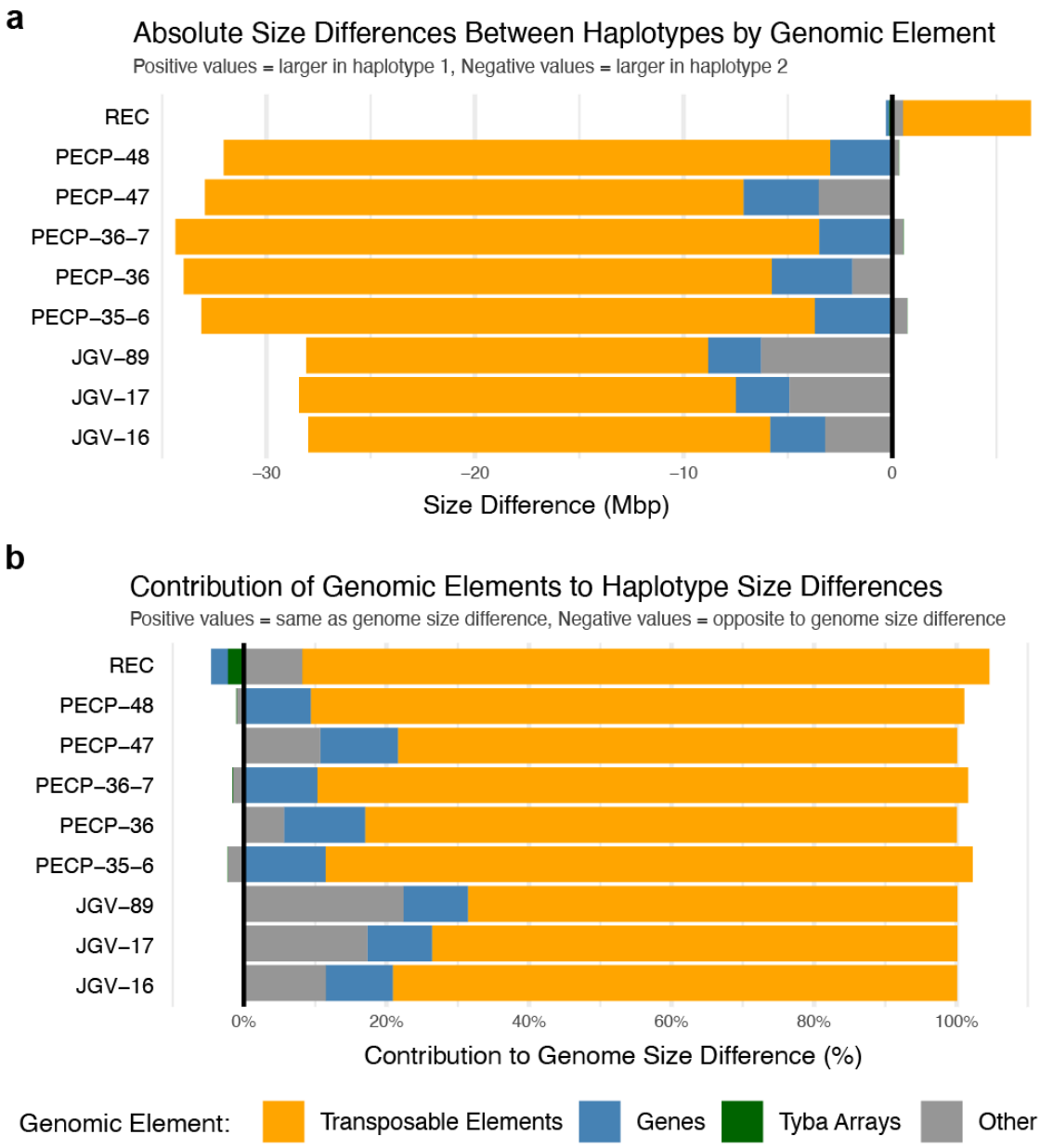
Extended Data Fig. 3: Distribution of 45S rDNA and identification of a reciprocal translocation in *R. tenuis* genotypes. FISH with a 45S rDNA probe (orange) on mitotic chromosomes of *R. tenuis* genotypes REC, JGV-89, PECP-47 and PECP-48. Genotypes PECP-47 and PECP-48 exhibit a reciprocal translocation involving 45S rDNA regions. The translocation occurs between chromosomes Chr1_h1 and Chr2_h2 and is absent in genotypes REC and JGV-89. Scale bars correspond to 5 µm**.**



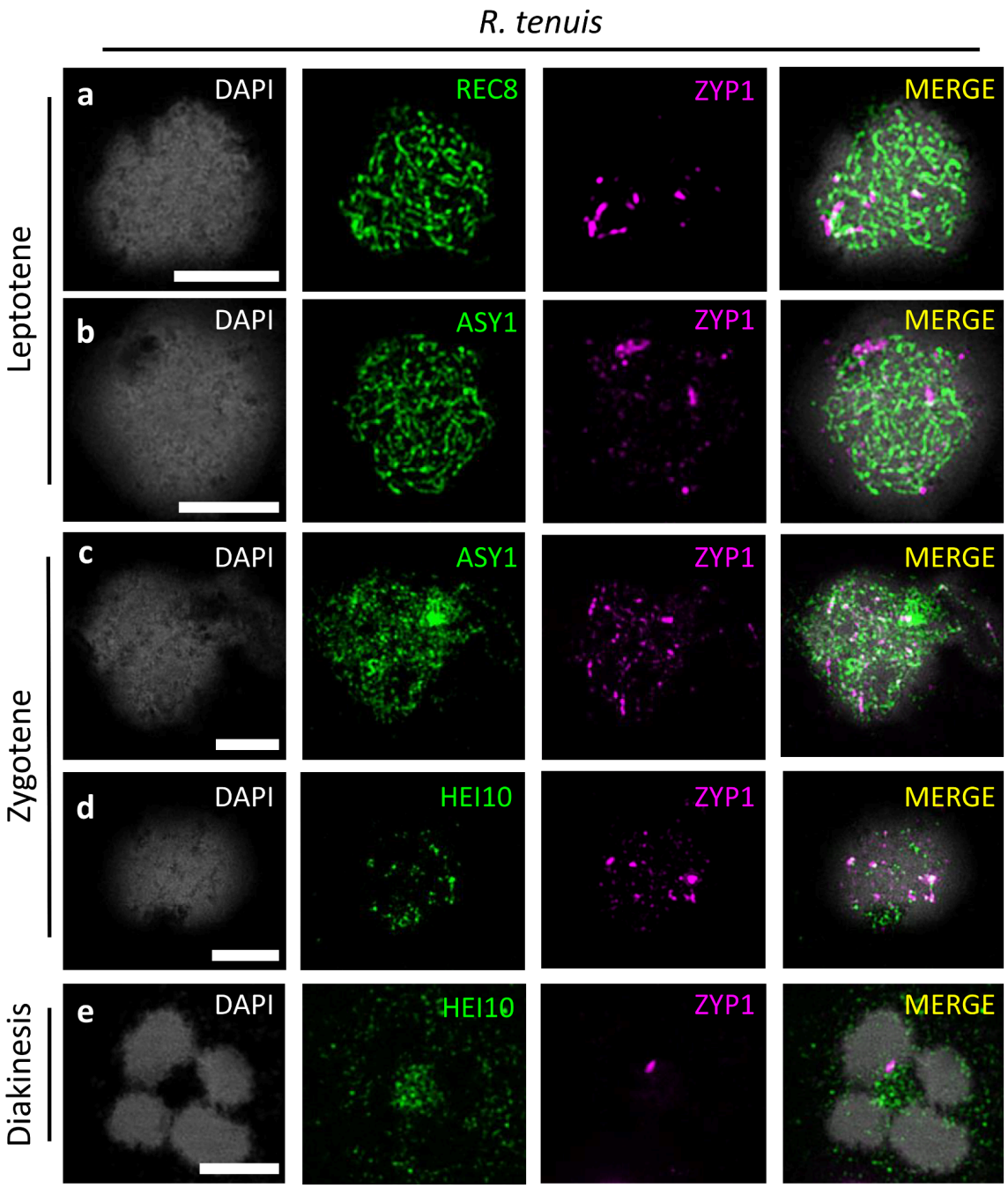
Extended Data Fig. 4: *K*-mer clustering and adjacency matrix of all 36 *R. tenuis* chromosomes from all accessions and 18 *R. austrobrasiliensis* pseudo-chromosomes. (**a**) *K*-mer size is 31 and the color bar on the right side indicates the similarity level measured by jaccard index. ANI: average nucleotide identity. (**b**) Phylogenetic tree rooted by *R. breviuscula* based on divergence of single-copy ortholog genes. All nodes were labeled with split time (blue number) of the corresponding branches. Time was calculated by *substitution per site* / (2 x *mutation rate per site per generation*). Mutation rate is 6.13e-9. Please note that this tree was constructed with only a pseudo-haplotype of *R. austrobrasiliensis* because the six haplotypes were not fully resolved. Here we used scaffold 1, 7, and 13.



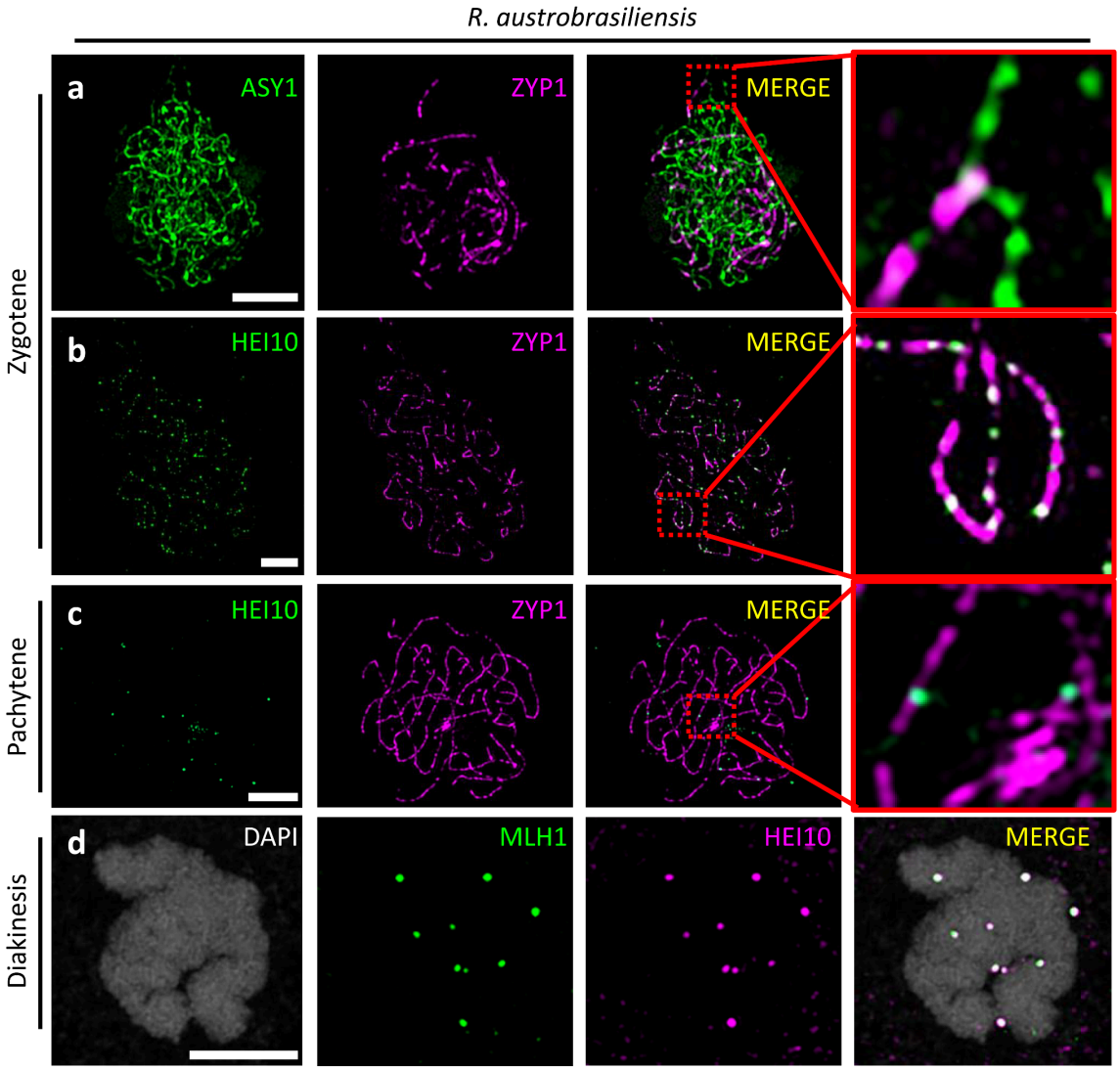
Extended Data Fig. 5: SubPhaser on phased haplotypes of different *R. tenuis* accessions. Unsupervised hierarchical clustering (the horizontal colour bar at the top of the axis indicates to which subgenome the *k*-mer is specific; the vertical colour bar on the left of the axis indicates the subgenome to which the chromosome is assigned). The heatmap indicates the *Z*-scale relative abundance of *k*-mers. The larger the *Z* score is, the greater the relative abundance of a *k*-mer).



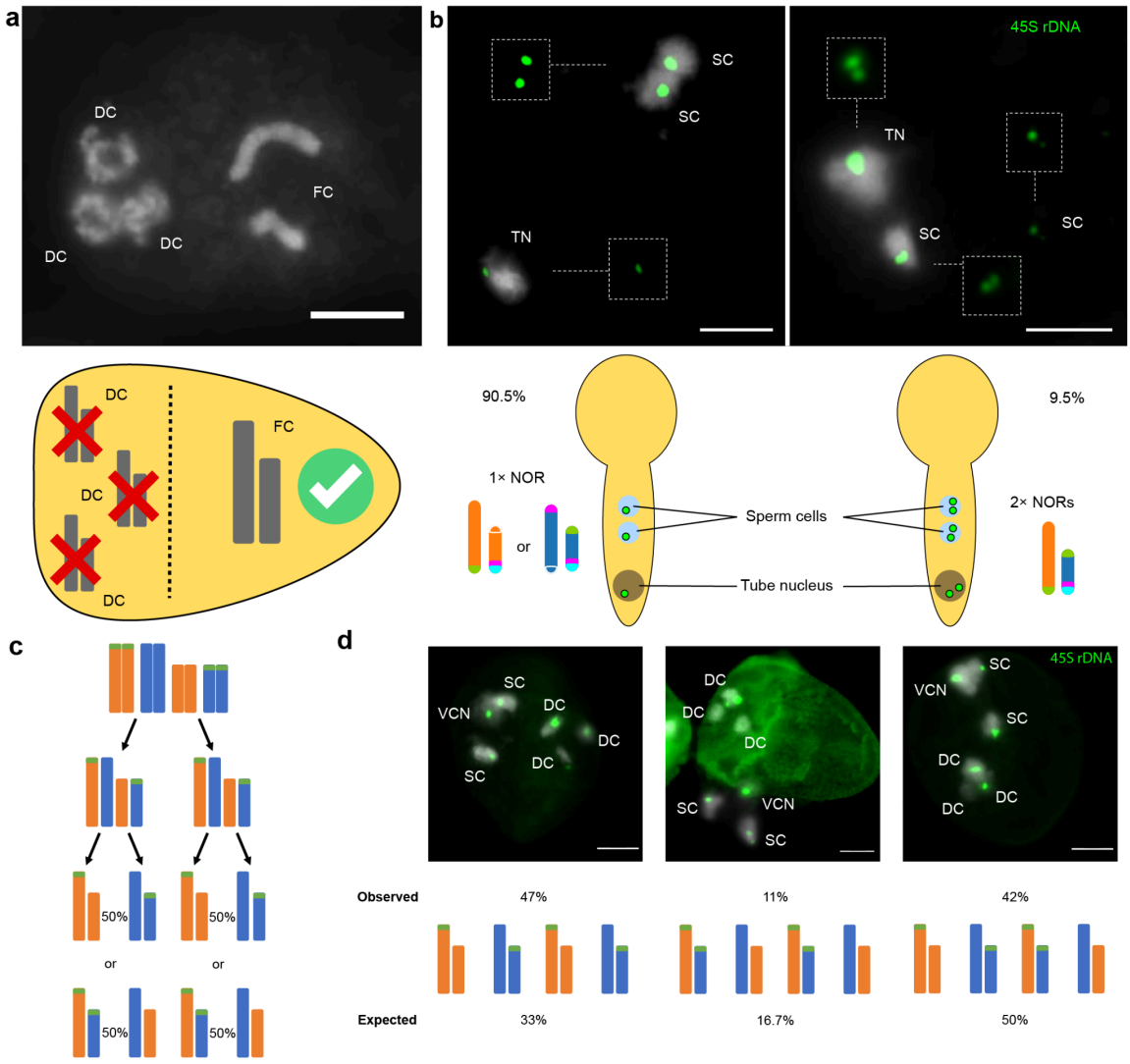
Extended Data Fig. 6: Genomic elements that contribute to the haplotype differences of all *R. tenuis* accessions. (**a**) Absolute size differences of different genomic elements between the two haplotypes of all *R. tenuis* accessions. A positive value means the total size of the element is larger in haplotype 1 and a negative value means a larger total size in haplotype 2. (**b**) Contribution of the genomic elements to the percentage of genome size difference of two haplotypes of *R. tenuis*. A positive value means this element difference is the same direction of genome size difference. For instance, haplotype 1 genome size of REC is smaller than haplotype 2. Transposable elements (TEs) in REC haplotype 1 is also shorter than haplotype 2, so the contribution of TEs has a positive value; while gene total length is longer in haplotype 1 so the contribution is negative.



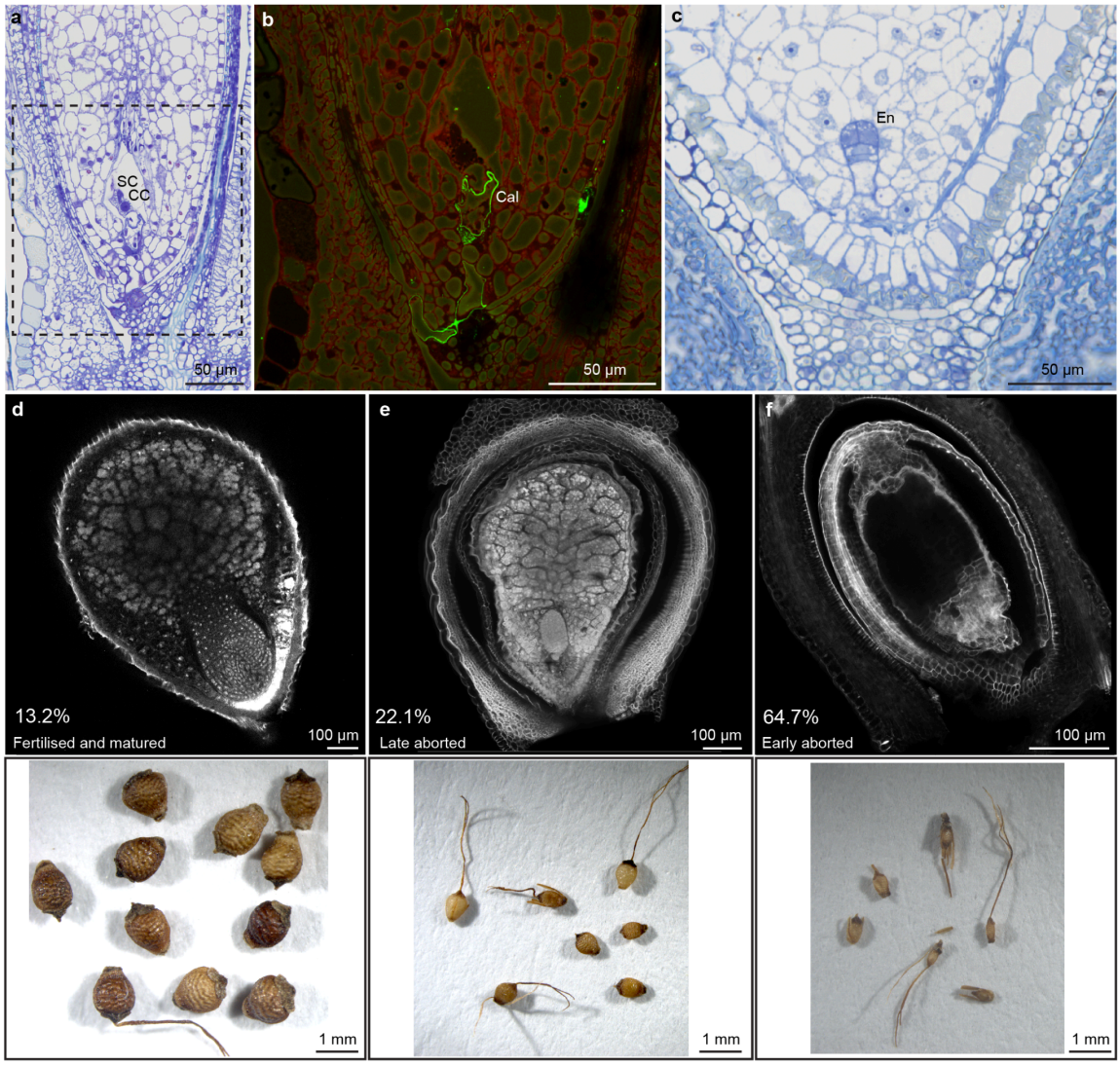
Extended Data Fig. 7: Additional channels and annotations of main figure, showing the meiotic atlas of *R. tenuis* and major recombination proteins. Early recombination proteins ASY1 and REC8 display regular behaviour, localising as linear signals (a-b). At zygotene, ASY1 is unloaded but ZYP1 shows only fragmented association with the axis. ZYP1 is able to attract unspecific signals of HEI10 (c, d). At diakinesis, HEI10 doesn't localise as foci and four univalents are observed. ZYP1 forms polycomplexes. Scale bars correspond to 5 µm.



Extended Data Fig. 8: Additional channels and annotations of the main figure of *R. austrobrasiliensis* prophase I. At zygotene, ZYP1 is loaded on paired chromosomes while ASY1 is unloaded (**a**). Simultaneously, HEI10 is loaded as multiple foci along the signal of ZYP1 (**b**). At pachytene, ZYP1 forms a continuous linear signal along the whole chromosome length and HEI10 signal condenses into few high intensity foci (**c**). At diakinesis, HEI10 forms foci that colocalize with MLH1 and mark class I crossovers. Scale bars correspond to 5 μm.



Extended Data Fig. 9. PECP-48 segregation distortion of NOR-bearing chromosomes during asymmetric male meiosis. (**a**) Male meiosis resulting in the formation of a single functional microspore (functional cell, FC) and three degenerating cells (DC), characteristic of monad-type microsporogenesis. (**b**) FISH detection of 45S rDNA loci (green) in sperm cells and tube nucleus of pollen tubes. Pollen grains exhibit either one (90.5%, left) or two (9.5%, right) 45S rDNA foci. The schematic representation below illustrates the relative frequencies of each class (90.5% and 9.5%) and possible chromosome haplotypes. (**c**) Predicted segregation outcomes and expected gamete genotype frequencies following achiasmatic inverted meiosis. (**d**) Expected versus observed (n = 100) frequencies of gamete genotypes based on FISH analysis of mature pollen grains hybridised with a 45S rDNA probe (green), illustrating segregation distortion favoring the larger, TE-rich haplotype 1. DC = degenerating cell; FC = functional cell; SC = sperm cell; TN = tube nucleus; VCN = vegetative cell nucleus. Scale bars correspond to 5 μm.



Extended Data Fig. 10: Embryology of *R. tenuis* (PECP-48). (**a**) Pollen tube entering the mature embryo sac to deliver the sperm cells. SC = sperm cell, CC = central cell. (**b**) The presence of the pollen tube in the ovule tissue is confirmed by the detection of callose (Cal) in green in the pollen tube cell wall (indirect immunofluorescent labeling of ß-1,3-glucan; see **Supplementary Figure 9**). (**c**) Abnormal development of seed with an embryo (En) but no visible endosperm formation. (**d**) Viable mature, (**e**) late aborted and (**f**) early aborted seeds shown by Feulgen staining (top) and respective seed morphology (bottom).

Supplementary Material

Supplementary Tables

Supplementary Table 1. Genome assembly statistics of all *R. tenuis* accessions and *R. austrobrasiliensis*

Supplementary Table 2: List of meiotic genes searched in *R. tenuis* genotypes, their location and comparison with *R. breviuscula* and *R. austrobrasiliensis.*

Supplementary Table 3. Cell number of snRNA- and snATAC-seq of different *R. tenuis* accessions.

Supplementary Figures

Supplementary Fig. 1: Hi-C heatmaps of diploid genomes of three *R. tenuis* accessions and *R. austrobrasiliensis*.

Supplementary Fig. 2: GenomeScope2 plots based on 31-mer of HiFi reads from all *R. tenuis* accessions as well as *R. austrobrasiliensis* and *R. breviuscula*.

Supplementary Fig. 3: Inferred tree by k-mer distance of autohexaploid *R. austrobrasiliensis* and the same fusion positions on both haplotypes of *R. tenuis*.

Supplementary Fig. 4: Probe distribution reveals absence of reciprocal translocations in *R. austrobrasiliensis*.

Supplementary Fig. 5: Meiotic chromosome behaviour of REC, PECP and JGV populations.

Supplementary Fig. 6: Distribution of 45S rDNA and REC translocation-specific probes (magenta - oligo-probe 1 and cyan - oligo-probe 2) during inverted meiosis in *R. tenuis* (PECP-48) visualised with FISH.

Supplementary Fig. 7: Pollen viability by Alexander Staining for each *R. tenuis* accession and *R. austrobrasiliensis* (percentage).

Supplementary Fig. 8: 3D rendering of an embryo sac from Toluidine blue-stained serial semithin physical sections stitched into a 3D z stack.

Supplementary Fig. 9: Pollen tube reaching the ovule.

Supplementary Fig. 10: Pollen tube reaching the mature embryo sac/female gametophyte to deliver the sperm cells.

Supplementary Fig. 11: Example of workflow for the genetic scan of meiotic genes.

Supplementary Table 1. Genome assembly statistics of all *R. tenuis* accessions and *R. austrobrasiliensis*.

Species	Accession	Hap	Contig total length (bp)	Chr. total length (bp)	Contig number	Largest contig (bp)	Contig N50 (bp)	# gap	GC%
R. tenuis	REC	h1	395,484,736	348,158,620	1,663	31,615,665	19,340,799	28	35.68
		h2	373,542,889	354,512,391	616	71,890,041	23,642,622	24	35.48
	JGV-16	h1	512,049,716	377,634,235	1,408	112,179,635	61,459,675	7	39.41
		h2	380,901,205	349,658,151	145	66,517,139	36,225,668	18	35.18
	JGV-17	h1	573,037,334	377,385,800	2,032	179,483,847	46,422,628	11	41.90
		h2	440,003,022	348,979,334	447	61,337,029	35,842,360	9	38.34
	JGV-89	h1	502,650,004	377,247,412	2,119	65,727,966	54,108,842	18	40.04
		h2	368,703,004	349,182,916	240	51,407,378	36,217,777	28	36.36
	PECP-47	h1	420,138,026	381,933,207	945	72,058,618	37,276,100	13	36.62
		h2	359,408,259	349,003,858	208	84,674,538	38,765,659	13	35.32
	PECP-48	h1	550,339,684	382,648,587	3,562	66,123,649	7,984,712	46	38.25
		h2	387,787,195	350,935,085	585	33,381,499	12,378,200	63	36.21
	PECP-35-6	h1	426,271,194	382,455,145	871	141,969,674	75,026,608	12	36.47
		h2	360,109,026	350,070,891	158	55,915,206	38,097,382	15	35.48
	PECP-36	h1	394,251,559	382,926,756	304	72,005,413	26,891,453	27	35.33
		h2	378,334,318	348,977,203	963	60,647,232	32,395,499	20	35.33
	PECP-36-7	h1	414,955,118	383,444,050	651	111,369,649	50,763,498	12	35.78
		h2	379,302,196	349,631,601	605	66,243,789	38,781,351	15	35.34
R. austrobrasiliensis			2,168,431,183	2,117,633,834	1,136	42,774,656	10,796,733	212	35.55

Supplementary Table 2: List of meiotic genes searched in *R. tenuis* genotypes, their location and comparison with *R. breviuscula* (Rb) and *R. austrobrasiliensis* (Ra). The table reports several meiotic genes involved in recombination and 1) whether they were found as intact copies; 2) how many copies were found; 3) the species and haplotype they were found in; 4) a comparison with *R. breviuscula* and *R. austrobrasiliensis*; 5) references.

Gene	REC hap1	REC hap2	JGV hap1	JGV hap2	PECP hap1	PECP hap2	Rb hap1	Rb hap2	Ra	RNA (REC)	Reference
SYN1/REC8	2	2	2	2	2	2	2	2	12	+	10.1371/journal.pgen.1010304
SYN3	1	1	1	1	1	1	1	1	6	+	10.1007/s00412-009-0220-x
Top3alpha	1	1	1	1	1	1	1	1	6	+	10.1371/journal.pgen.1000285
HEIP1	1	1	1	1	1	1	1	1	6	+	10.1073/pnas.2221746120
PMS1	1	1	1	1	1	1	1	1	6	+	10.1093/nar/gkaf187
PHS1	1	1	1	1	1	1	1	1	6	+	10.1155/2012/514398
PCH2	1	1	1	1	1	1	1	1	6	+	10.1093/nar/gkac1160
SCEP2	1	1	1	1	1	1	1	1	6	+	10.1038/s41477-023-01558-y
SPO22/ZIP4	1	1	1	1	1	1	1	1	6	+	10.1371/journal.pgen.0030083
SPO11/1	1	1	1	1	1	1	1	1	6	+	10.1105/tpc.107.054817
SPO11/2	1	1	1	1	1	1	1	1	6	+	10.1105/tpc.107.054817
SPO11/3	1	1	1	1	1	1	1	1	6	+	10.1111/pbi.13189

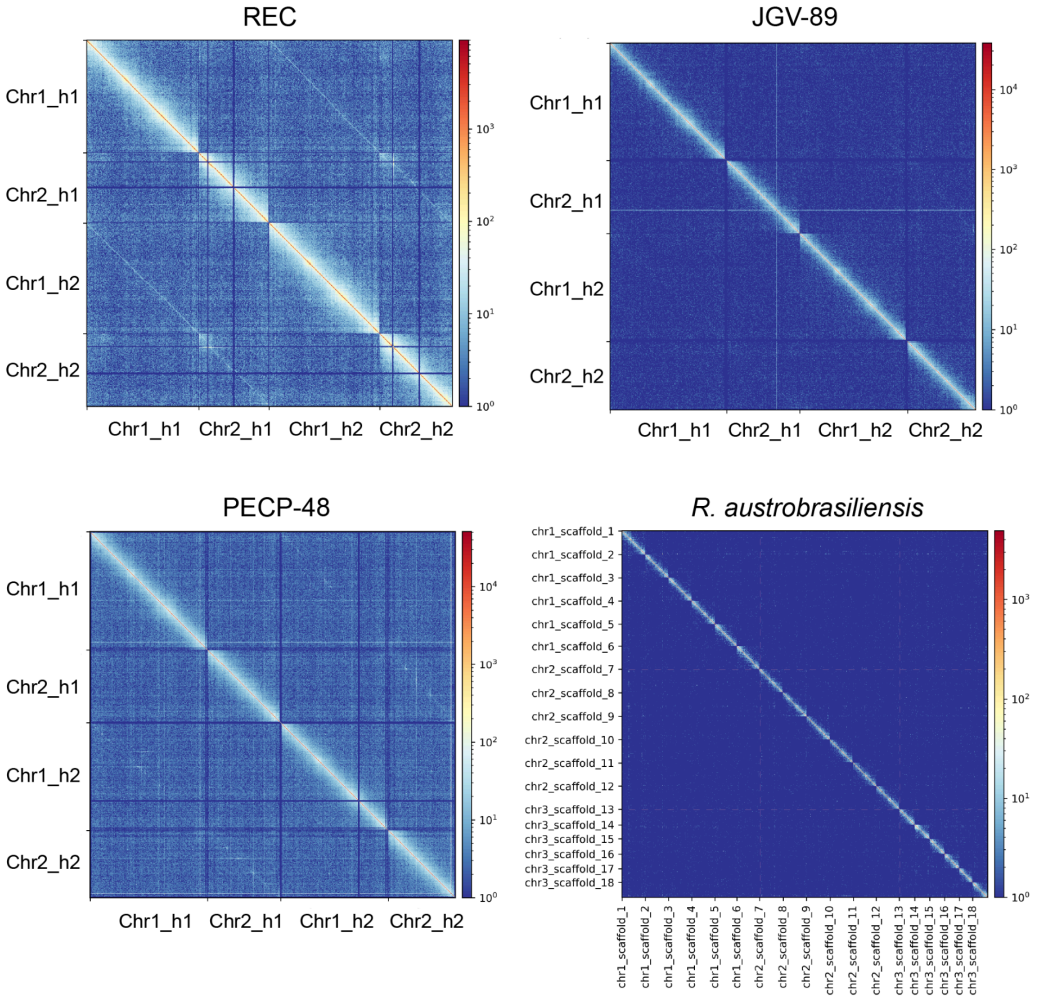
MLH3	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**6**	+	**10.1038/sj.emboj.7600992**
MRE11	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**6**	+	**10.1007/s00412-007-0147-z**
NBS1	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**6**	+	**10.1111/j.1365-313X.2007.03220.x**
PRD2	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**6**	+	**10.1371/journal.pgen.1000654**
PRD3	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**6**	+	**10.1371/journal.pgen.1010298**
FANCM	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**6**	+	**10.1093/nar/gkac1244**
MUS81	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**6**	+	**10.1073/pnas.2107543118**
RAD50	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**7**	+	**10.1073/pnas.0906273106**
RMI1	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**6**	+	**10.1093/nar/gkt730**
RPA1a	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**6**	+	**10.3389/fpls.2016.00033**
SEND1	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**7**	+	**10.1104/pp.114.237834**
DFO	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**6**	+	**10.1111/j.1365-313X.2012.05075.x**
ASY1	**2**	**2**	**2**	**2**	**2**	**2**	**1**	**1**	**6**	+	**10.1093/pnasnexus/pgac302**
ASY3	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**6**	+	**10.1104/pp.17.01725**
ASY4	**2**	**1**	**2**	**1**	**2**	**1**	**2**	**2**	**6**	+	**10.1104/pp.17.01725**
DMC1	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**6**	+	**10.3389/fpls.2020.00839**

HEI10	**2**	**2**	**2**	**2**	**2**	**2**	**3**	**3**	**12**	+	**10.7554/eLife.79408**
MER3	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**6**	+	**10.1371/journal.pone.0150482**
MLH1	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**6**	+	**10.1371/journal.pgen.1011197**
MSH4	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**5**	+	**10.1371/journal.pgen.1003922**
MSH5	**2**	**1**	**2**	**1**	**2**	**2**	**1**	**1**	**9**	+	**10.1093/mp/sss145**
SHOC1	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**6**	+	**10.1242/jcs.088229**
ZYP1	**1**	**3**	**1**	**1**	**1**	**1**	**1**	**1**	**6**	+	**10.1073/pnas.2021671118**
RAD51	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**6**	+	**10.26508/lsa.202402701**
MTOPVIb	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**6**	+	**10.1093/nar/gkae181**
PRD1	**2**	**2**	**2**	**2**	**2**	**2**	**2**	**2**	**12**	+	**10.1038/s41598-017-10270-9**
HAP2	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**6**	+	**https://doi.org/10.1038/s41467-023-42789-z**
NUF2	**2**	**2**	**2**	**2**	**2**	**2**	**1**	**1**	**12**	+	**10.1111/tpj.15347**
SCEP3	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**6**	+	**10.1038/s41477-025-02030-9**
COMET	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**6**	+	**10.1016/j.cub.2020.07.089**
SCEP1	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**6**	+	**10.1038/s41477-023-01558-y**
BRCA2	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**6**	+	**PubMed: 38213474**

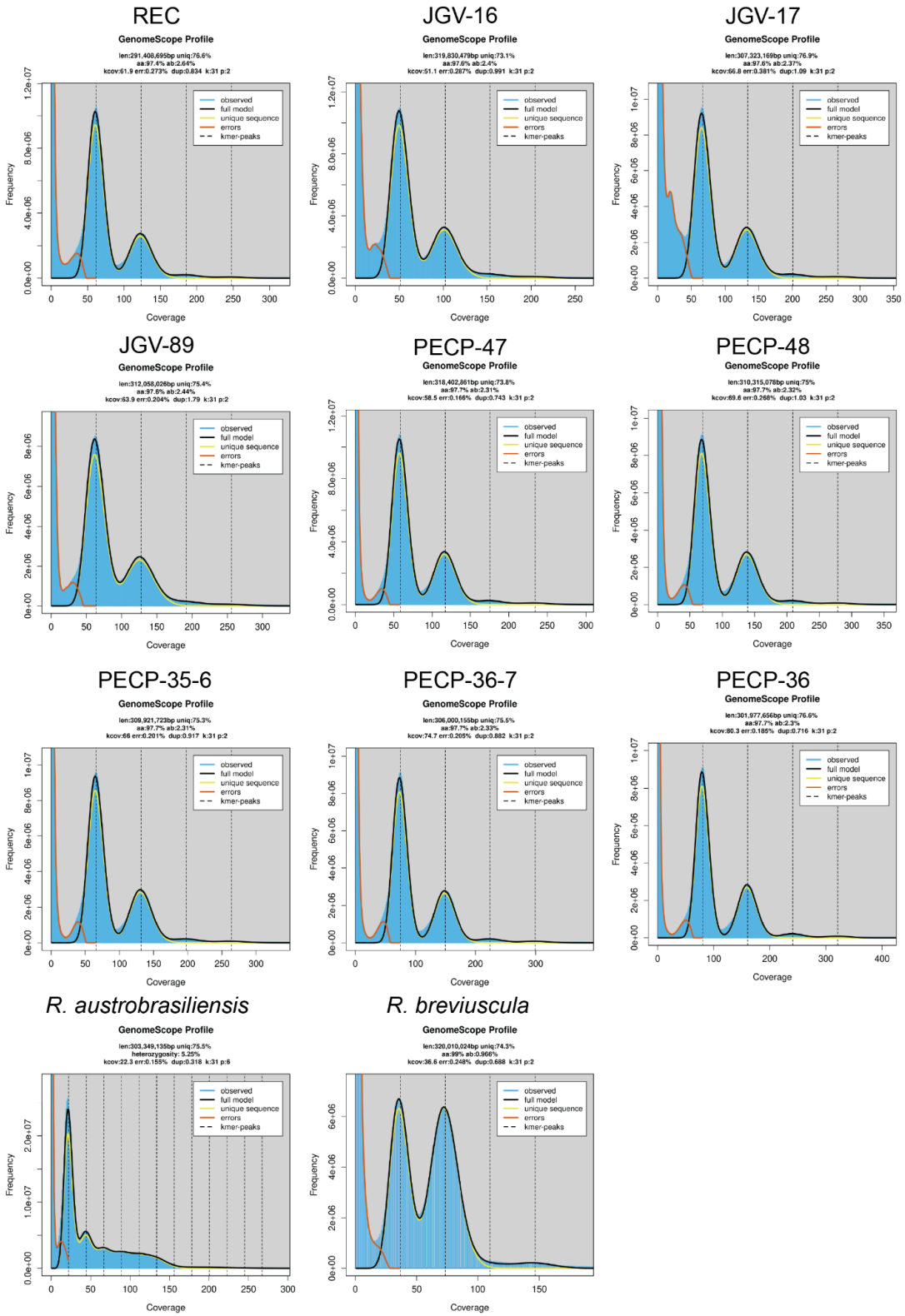
PTD	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**6**	**+**	**10.1016/j.jgg.2014.02.001**
RECQ4	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**6**	**+**	**10.1073/pnas.1713071115**
MCC1	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**6**	**+**	**10.1111/j.1365-313X.2010.04191.x**

Supplementary Table 3. Cell number of snRNA- and snATAC-seq of different *R. tenuis* accessions.

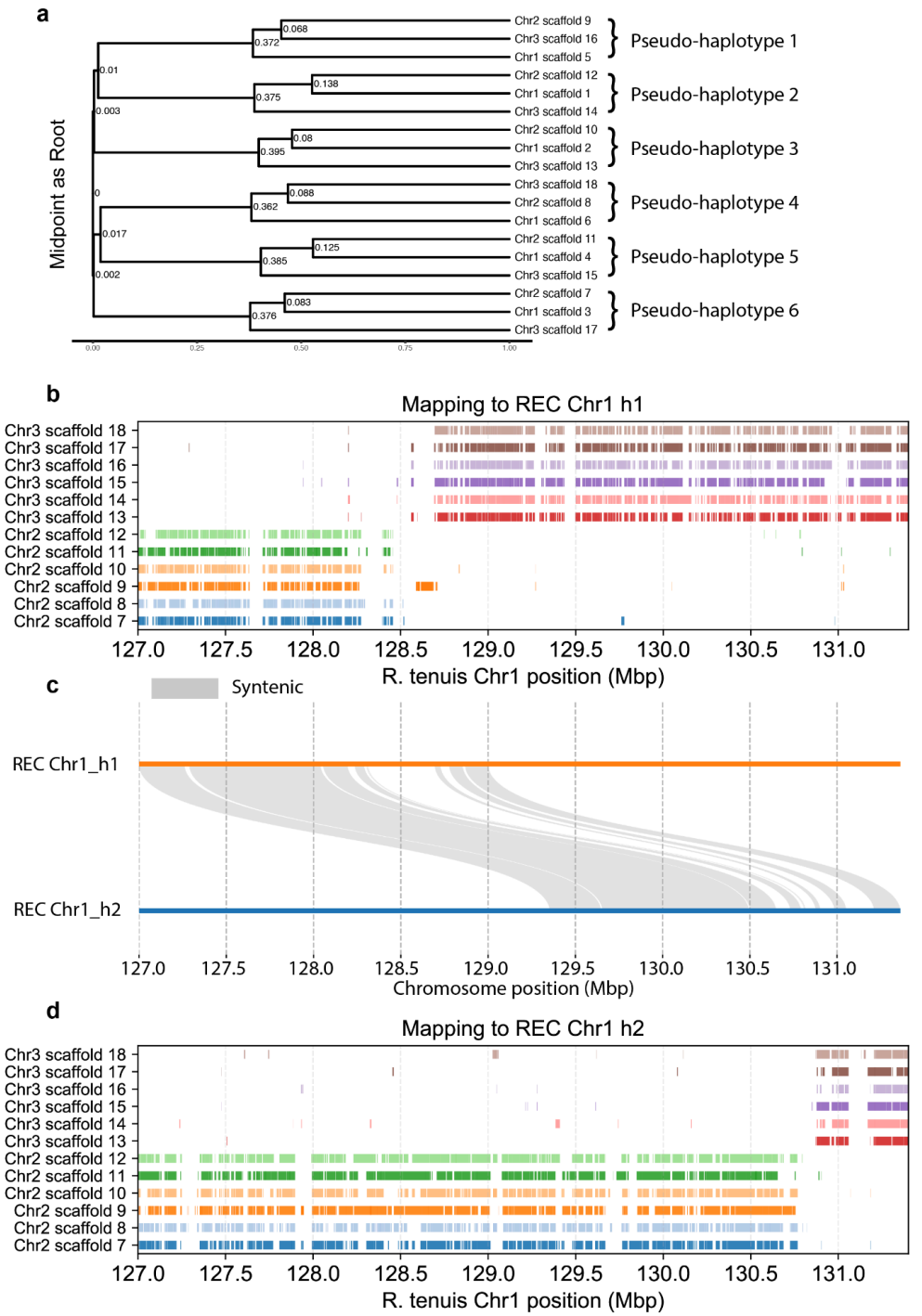
Source	**Accession**	**Library**	**snRNA-seq**	**snATAC-seq**
Pollen	REC	4984.A	1,816	x
		6067.C	x	3,296
		6067.D	x	2,131
	JGV-16	6451.C	976	x
	JGV-17	6451.B	844	x
	PECP-47	6385.B	1,029	x
	PECP-48	6451.A	905	x
Seed Embryo	REC	7248.B	x	327
	PECP-48	7470.B	x	185
Seed Endosperm	REC	7470.A	x	3,365
	PECP-48	7470.C	x	76



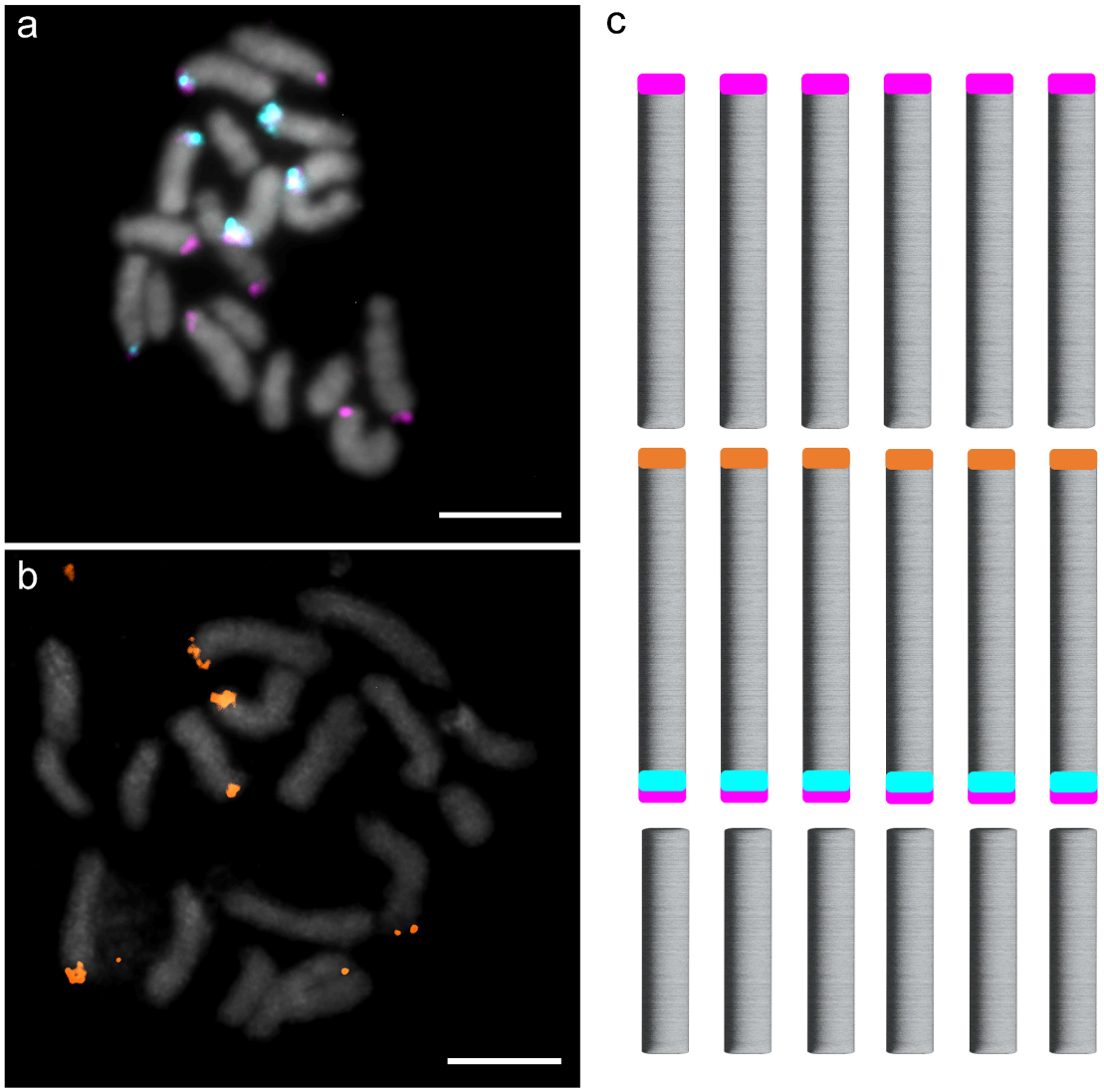
Supplementary Fig. 1: Hi-C heatmaps of diploid genomes of three *R. tenuis* accessions and *R. austrobrasiliensis*. The Hi-C maps of three *R. tenuis* accessions used 1Mb bin sizes. *R. austrobrasiliensis* used 2.5Mb resolution for plotting. The color bar on the right side of each Hi-C map indicates the Hi-C signal density.



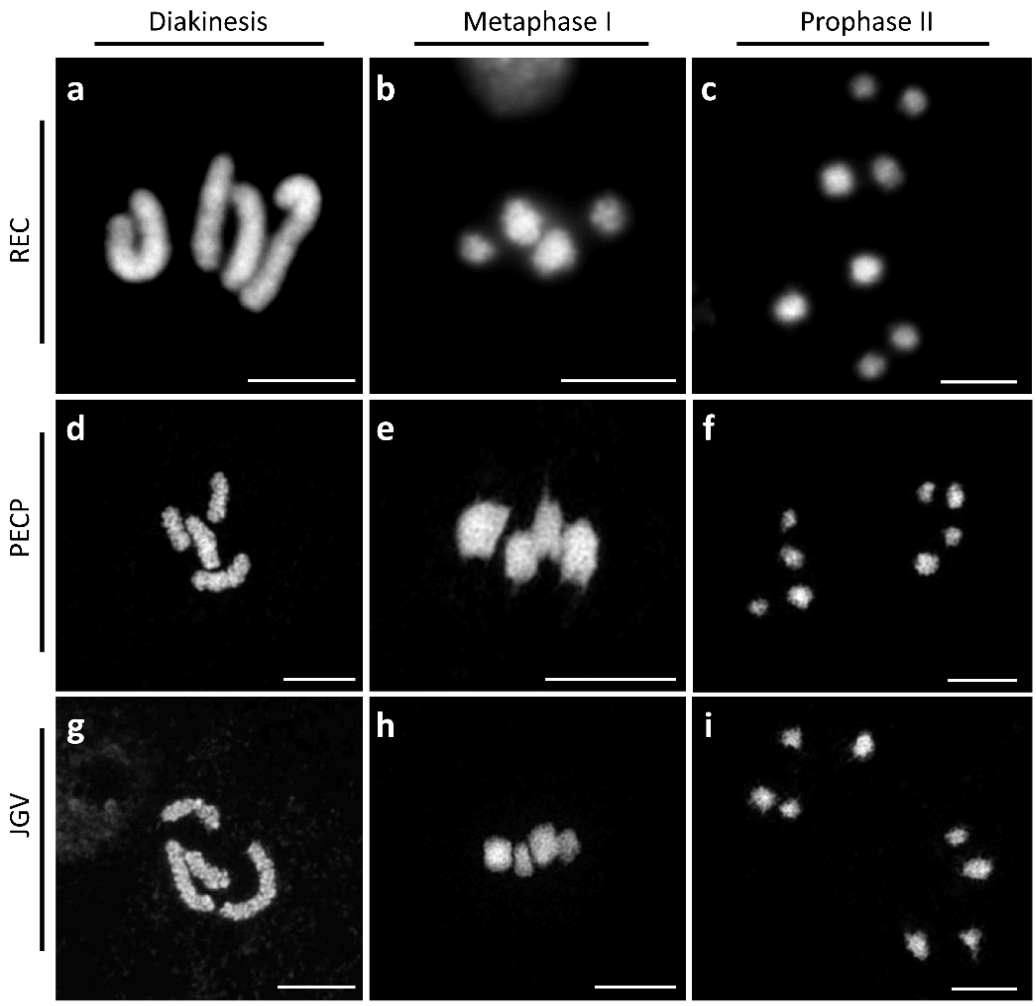
Supplementary Fig. 2: GenomeScope2 plots based on 31-mer of HiFi reads from all *R. tenuis* accessions as well as *R. austrobrasiliensis* and *R. breviuscula*.



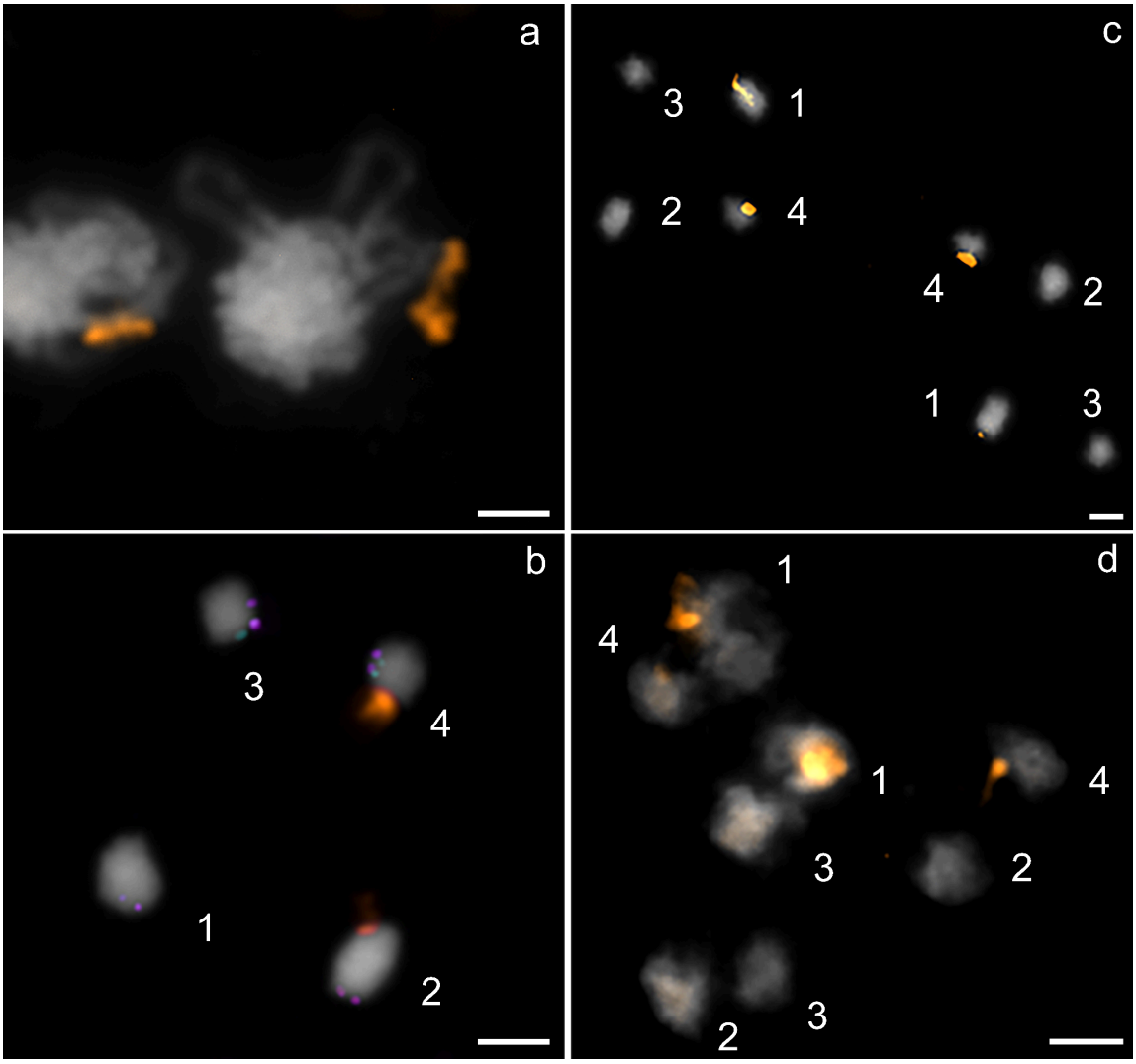
Supplementary Fig. 3: Inferred tree by *K*-mer distance of autohexaploid *R. austrobrasiliensis* and the same fusion positions on both haplotypes of *R. tenuis*. (**a**) Phylogenetic tree based on the k-mer distance matrix of all chromosome-level scaffolds of *R. austrobrasiliensis*. Labeled values are chronos (R function) scaled branch lengths assuming root edge has brach length 0 and sum of brach lengths from root to tip is 1. (**b**) Mapped fragments of *R. austrobrasiliensis* to *R. tenuis* (REC) haplotype 1 and (**d**) haplotype 2. (**c**) The SyRI collinearity of two haplotypes (REC) at the fusion positions.



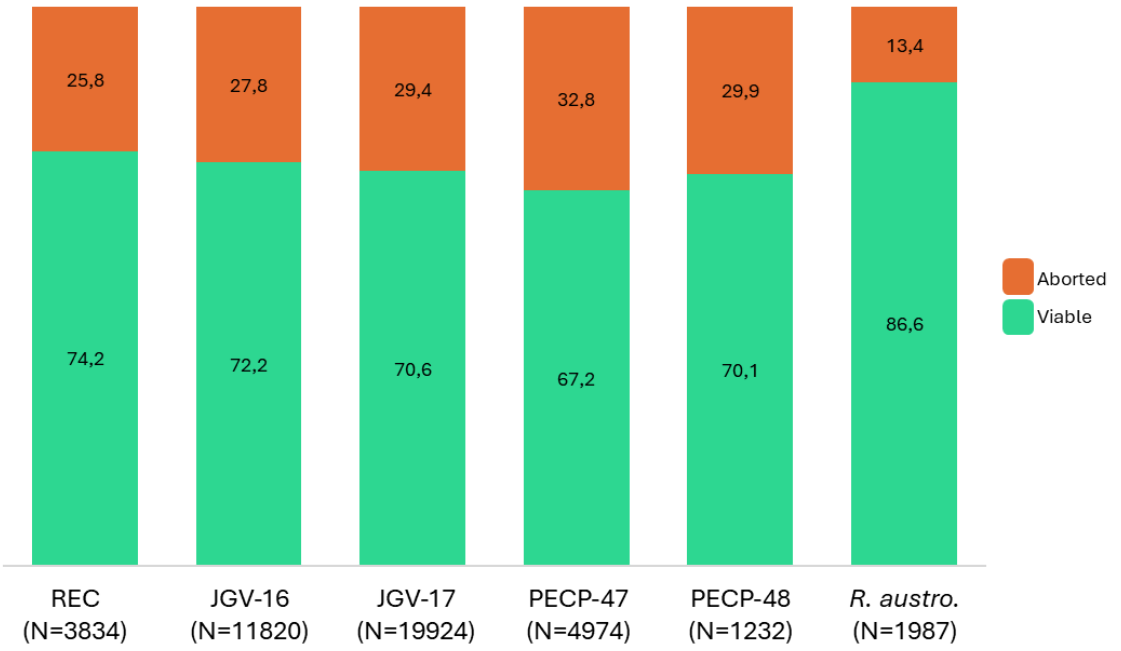
Supplementary Fig. 4: Probe distribution reveals absence of reciprocal translocations in *R. austrobrasiliensis*. FISH using translocation-specific probe sets (magenta, oligo-probe 1; cyan, oligo-probe 2) (**a**) and a 45S rDNA probe (orange) (**b**) on mitotic chromosomes of *R. austrobrasiliensis*. (**c**) Idiogram of *R. austrobrasiliensis* showing the distribution of probe signals from (**a**) and (**b**). The results indicate the absence of the reciprocal translocations found in *R. tenuis* genotypes. Scale bars correspond to 5 µm.



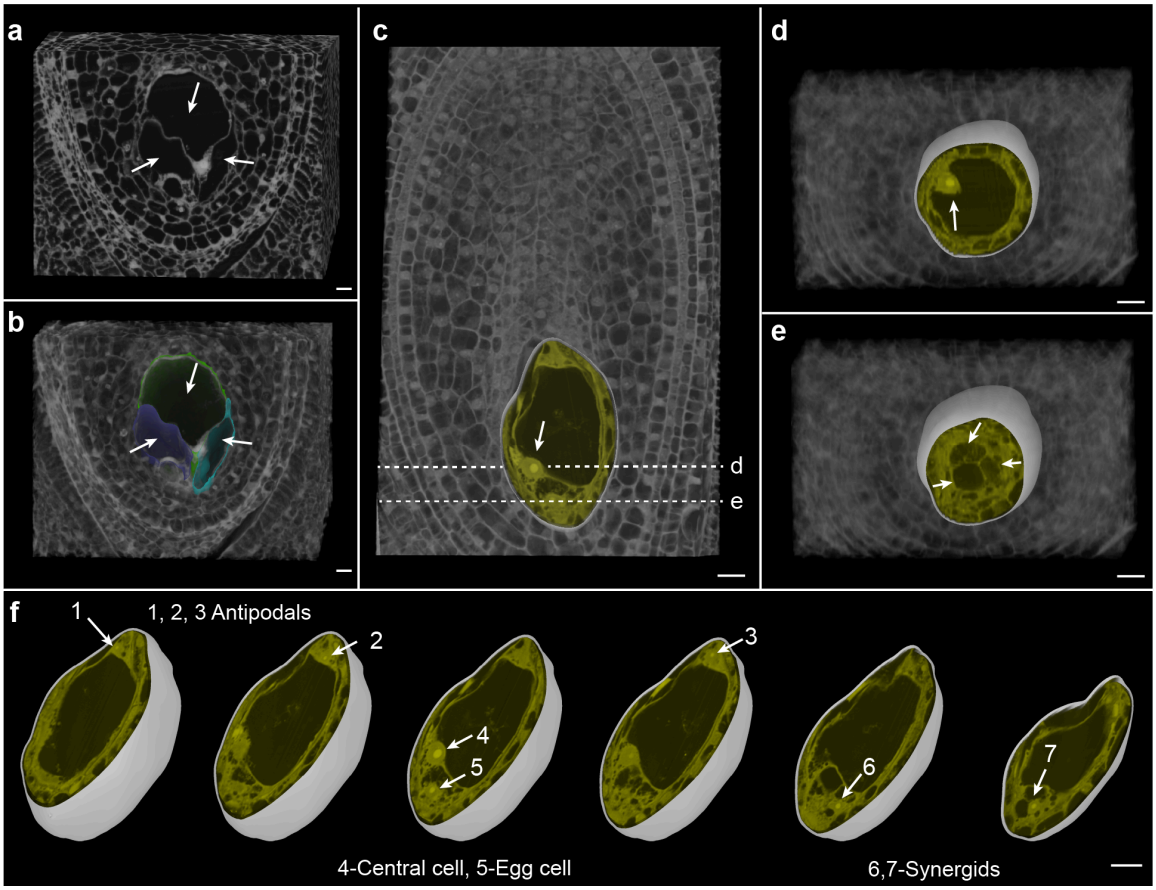
Supplementary Fig. 5: Meiotic chromosome behaviour of REC, PECP and JGV populations. At diakinesis, only 4 univalents are present and no bivalents are visible (**a**, n=10; **d**, n=21; **g**, n=40). Univalents align at metaphase I (**b**, n=15; **e**, n=16; **h**, n=11), followed by the segregation of sister chromatids that re-align at prophase II (**c**, n=22; **f**, n=26; **i**, n=12). Scale bars correspond to 5 µm.



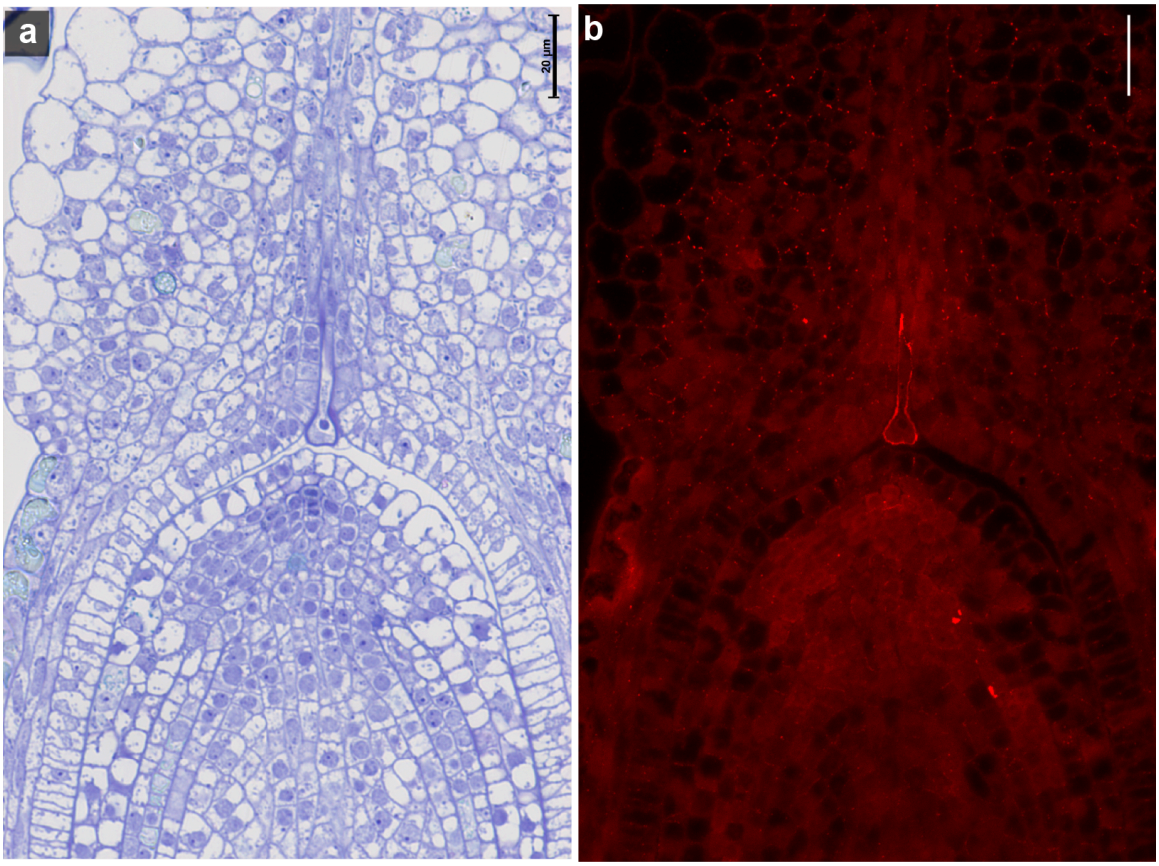
Supplementary Fig. 6: Distribution of 45S rDNA and REC translocation-specific probes (magenta - oligo-probe 1 and cyan - oligo-probe 2) during inverted meiosis in *R. tenuis* (PECP-48) visualised with FISH. (**a**) Zygotene, (**b**) Diakinesis, (**c**) Prophase II, (**d**) Anaphase II/Telophase II. 1 = Chr1_h1, 2 = Chr1_h2, 3 = Chr2_h1 and 4 = Chr2_h2. Note that the particular combination of 2 and 3 chromosomes in **d** does not harbour a NOR. Scale bars correspond to 2 μm.



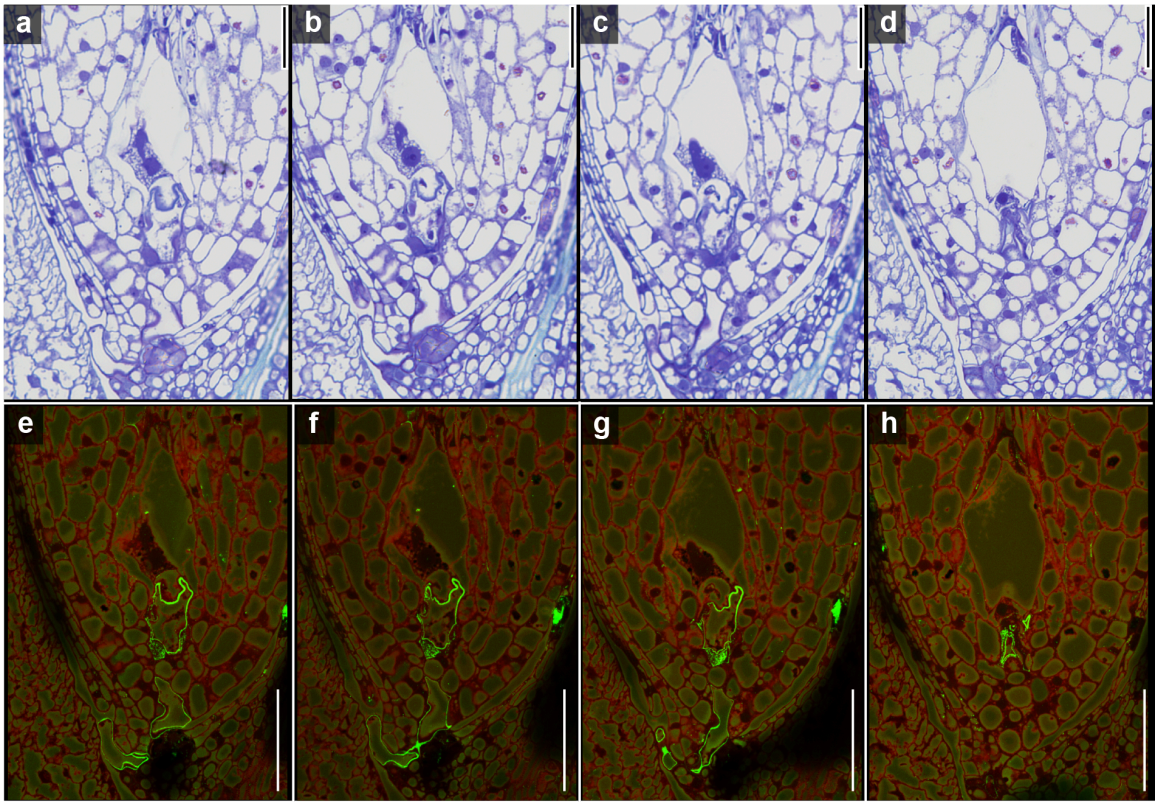
Supplementary Fig. 7: Pollen viability by Alexander Staining for each *R. tenuis* accession and *R. austrobrasiliensis* (percentage). Green indicates viable pollen, orange indicates aborted pollen. Total count (N) indicated in each accession label.



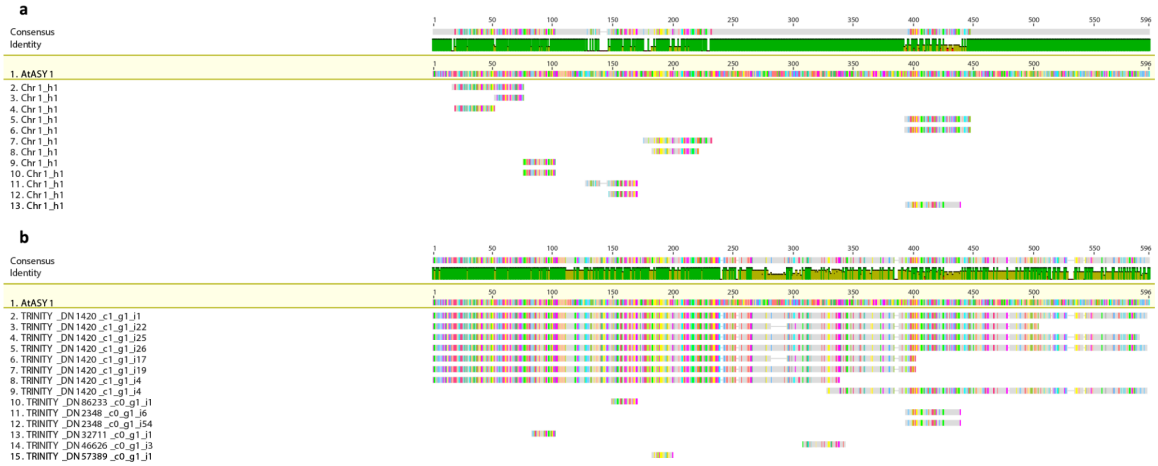
Supplementary Fig. 8: 3D rendering of embryo sacs from Toluidine blue-stained serial semithin physical sections stitched into a 3D z stack. (**a–b**) A case showing multiple embryo sacs in *R. tenuis*. Arrows point to the multiple embryo sacs (**b**) Embryo sac in (**a**) is visualized as a segmented object, with each one assigned a different color for distinction. (**c**) Longitudinal section view of a mature 7-nuclei embryo sac depicting the large central cell nucleus (arrow). Dotted lines represent the orientation of transverse sections shown in (**d–e**). (**d**) 3D cropped view of the sample along the transverse section showing the central cell nucleus (arrow). (**e**) 3D cropped view of the sample along the transverse section showing the egg apparatus containing the egg cell and the two adjacent synergids (arrow). (**f**) Series of cropped 3D views of the extracted embryo sac structure, showing the seven nuclei in different 2D planes through the mature embryo sac. (**c–f**) 3D segmentation of the embryo sac is shown as a white object, with the Toluidine blue-stained section overlaid in greyscale and the embryo sac region false-colored yellow. 3D rendering was performed using MorphoGraphX. Scale bar 10µm.



Supplementary Fig. 9: Pollen tube reaching the ovule. (**a**) Toluidine blue-stained semithin section depicting a pollen tube reaching the distal part of the ovule after growing down the style of the ovary. (**b**) Indirect immunofluorescent labeling of ß-1,3-glucan epitopes of the semithin section shown in (**a**) highlights the callose pollen tube cell wall in the surrounding ovary tissue. Note the punctate labeling of plasmodesmata decorating the cell walls in the beak-like cap tissue. Autofluorescence is caused by glutaraldehyde in the primary fixative and the previous Toluidine blue staining. Scale bar 20µm.



Supplementary Fig. 10: Pollen tube reaching the mature embryo sac/female gametophyte to deliver the sperm cells. (**a–d**) Consecutive toluidine blue-stained semithin sections depicting a pollen tube that has grown through the micropyle formed by the inner integument and reached the mature embryo sac. The globular nucleus of the central cell and - in direct contact - one of the two lens-shaped sperm nuclei can be seen. (**b–h**) Indirect immunofluorescent labeling of ß-1,3-glucan epitopes of the semithin sections shown in **a–g** highlights the callose pollen tube cell wall in the surrounding ovary tissue. Autofluorescence in both channels is caused by glutaraldehyde in the primary fixative and previous Toluidine blue staining. Scale bar 50µm.



Supplementary Fig. 11: Example of workflow for the genetic scan of meiotic genes. (**a**) ASY1 shows two overlapping hits for each query region, suggesting a duplication. (**b**) ASY1 can be found in our anther-specific transcriptome. Manual curation revealed a tandem duplication (156.939.594 -> 156.948.842 and 184.686.149 -> 184.695.397) on chromosome 1.

2.4 Bimodal centromeres in pentaploid dogroses shed light on their unique meiosis

Published in Nature (2025)

Herklotz, V.*, Zhang, M.*, Nascimento, T.*, Kalfusová, R., Lunerová, J., Fuchs, J., Harpke, D., Huettel, B., Pfordt, U., Wissemann, V., Kovařík, A#., Marques, A.#†, Ritz, C.M.#

* These authors contributed equally
# Correspondence
† Leading author and contact

Author Contributions:

A.M., A.K. and C.R. conceived the research. V.H. designed and performed the SCO experiments and processed the data. M.Z. and A.M. performed the genome assembly and genomics. D.H. performed the plastome assembly and annotation. B.H. performed all nuclear sequencing libraries. T.N., M.Z. and A.M. performed the ChIP-seq analysis. T.N. and J.L. performed the chromosome preparations, immunodetection and FISH. J.F. performed the flow cytometric measurements and the pollen nuclei sorting. A.K., J.L., R.K. and A.M. performed the centromere and repeat characterisation. U.P. performed the pollen viability assay. V.W. created and curated the synthetic hybrids. A.M., A.K. and C.R. supervised the project. V.H., M.Z., A.M. and C.R. wrote the original draft, and all authors discussed the results and contributed to writing the manuscript.

Article

Bimodal centromeres in pentaploid dogroses shed light on their unique meiosis

https://doi.org/10.1038/s41586-025-09171-z

Received: 9 July 2024

Accepted: 16 May 2025

Published online: 18 June 2025

Open access

Check for updates

V. Herklotz[1,10], M. Zhang[2,10], T. Nascimento[2,10], R. Kalfusová[3], J. Lunerová[3], J. Fuchs[4], D. Harpke[4], B. Huettel[5], U. Pfordt[2], V. Wissemann[6], A. Kovařík[3]✉, A. Marques[2,7]✉ & C. M. Ritz[1,8,9]✉

Sexual reproduction relies on meiotic chromosome pairing to form bivalents, a process that is complicated in polyploids owing to the presence of multiple subgenomes[1]. Uneven ploidy mostly results in sterility due to unbalanced chromosome pairing and segregation during meiosis. However, pentaploid dogroses (*Rosa* sect. *Caninae*; $2n = 5x = 35$) achieve stable sexual reproduction through a unique mechanism: 14 chromosomes form bivalents and are transmitted biparentally, while the remaining 21 chromosomes are maternally inherited as univalents[2,3]. Despite being studied for over a century, the role of centromeres in this process has remained unclear. Here we analyse haplotype-resolved chromosome-level genome assemblies for three pentaploid dogroses. Subgenome phasing revealed a bivalent-forming subgenome with two highly homozygous chromosome sets and three divergent subgenomes lacking homologous partners, therefore explaining their meiotic behaviour. Comparative analyses of chromosome synteny, phylogenetic relationships and centromere composition indicate that the subgenomes originated from two divergent clades of the genus *Rosa*. Pollen genome analysis shows that subgenomes from different evolutionary origins form bivalents, supporting multiple origins of dogroses and highlighting variation in subgenome contributions. We reveal that bivalent-forming centromeres are enriched with *ATHILA* retrotransposons, contrasting with larger tandem-repeat-based centromeres mainly found in univalents. This centromere structural bimodality possibly contributes to univalent drive during female meiosis. Our findings provide insights into the unique reproductive strategies of dogroses, advancing our understanding of genome evolution, centromere diversity and meiotic mechanisms in organisms with asymmetrical inheritance systems.

Whole-genome duplication or polyploidy is a frequent phenomenon across the phylogeny of land plants[4]. Meiosis is essential for sexual reproduction, ensuring the reduction in genomic content in gametes through chromosome pairing and exchanges between non-sister chromatids, that is, crossovers[5,6]. Polyploidy often results from meiotic failure, that is, the generation of unreduced gametes, which poses challenges to meiotic chromosome pairing and the maintenance of sexual reproduction[1,7]. Thus, polyploids often skip sexual reproduction by promoting vegetative propagation[8] or apomixis[9]. However, in many allopolyploids, in which distinct subgenomes come into contact through hybridization, recombination partners from homologous chromosomes (same parental subgenome) are preferred, while recombination between homoeologous chromosomes (different parental subgenomes) is suppressed[10,11].

The genus *Rosa*, which comprises approximately 150 species, is a typical example of evolution through frequent polyploidy and hybridization events[12], which is also reflected by the large variety of cultivated roses with a long breeding history that includes both processes. The genus comprises two major clades, the *Rosa* and allies clade and the *Synstylae* and allies clade with subg. *Hulthemia*, subg. *Hesperhodos* and sect. *Pimpinellifoliae* as the basalmost splits[12,13]. Available genomes from diploid roses of sect. *Synstylae*[14–16] and sect. *Rosa*[17] revealed high levels of synteny, enabling comparative studies in this taxonomically difficult genus. Studies on tetraploid cut roses (*Rosa hybrida*) undergoing regular meiosis have shown that most genomic markers were recombined freely from all four chromosome sets, but preferential recombination between chromosomes and even chromosome arms vary[18,19].

Within the *Synstylae* clade, allopolyploid dogroses (*Rosa* sect. *Caninae* (DC.) Ser.) exhibit a unique reproductive strategy known as *Canina* meiosis, in which the selective chromosome pairing results in a mixed mode of inheritance—combining biparental transmission of bivalents

[1]Senckenberg Museum for Natural History Görlitz, Senckenberg–Leibniz Institution for Biodiversity and Earth System Research, Görlitz, Germany. [2]Department of Chromosome Biology, Max Planck Institute for Plant Breeding Research, Cologne, Germany. [3]Department of Molecular Epigenetics, Institute of Biophysics, Czech Academy of Sciences, Brno, Czech Republic. [4]Leibniz Institute of Plant Genetics and Crop Plant Research (IPK) Gatersleben, Seeland, Germany. [5]Max Planck Genome Centre Cologne, Max Planck Institute for Plant Breeding Research, Cologne, Germany. [6]Institute of Botany, Systematic Botany Group, Justus-Liebig-University, Gießen, Germany. [7]Cluster of Excellence on Plant Sciences (CEPLAS), Heinrich-Heine University, Düsseldorf, Germany. [8]International Institute (IHI) Zittau, TUD Dresden University of Technology, Zittau, Germany. [9]German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig, Leipzig, Germany. [10]These authors contributed equally: V. Herklotz, M. Zhang, T. Nascimento. ✉e-mail: kovarik@ibp.cz; amarques@mpipz.mpg.de; christiane.ritz@senckenberg.de

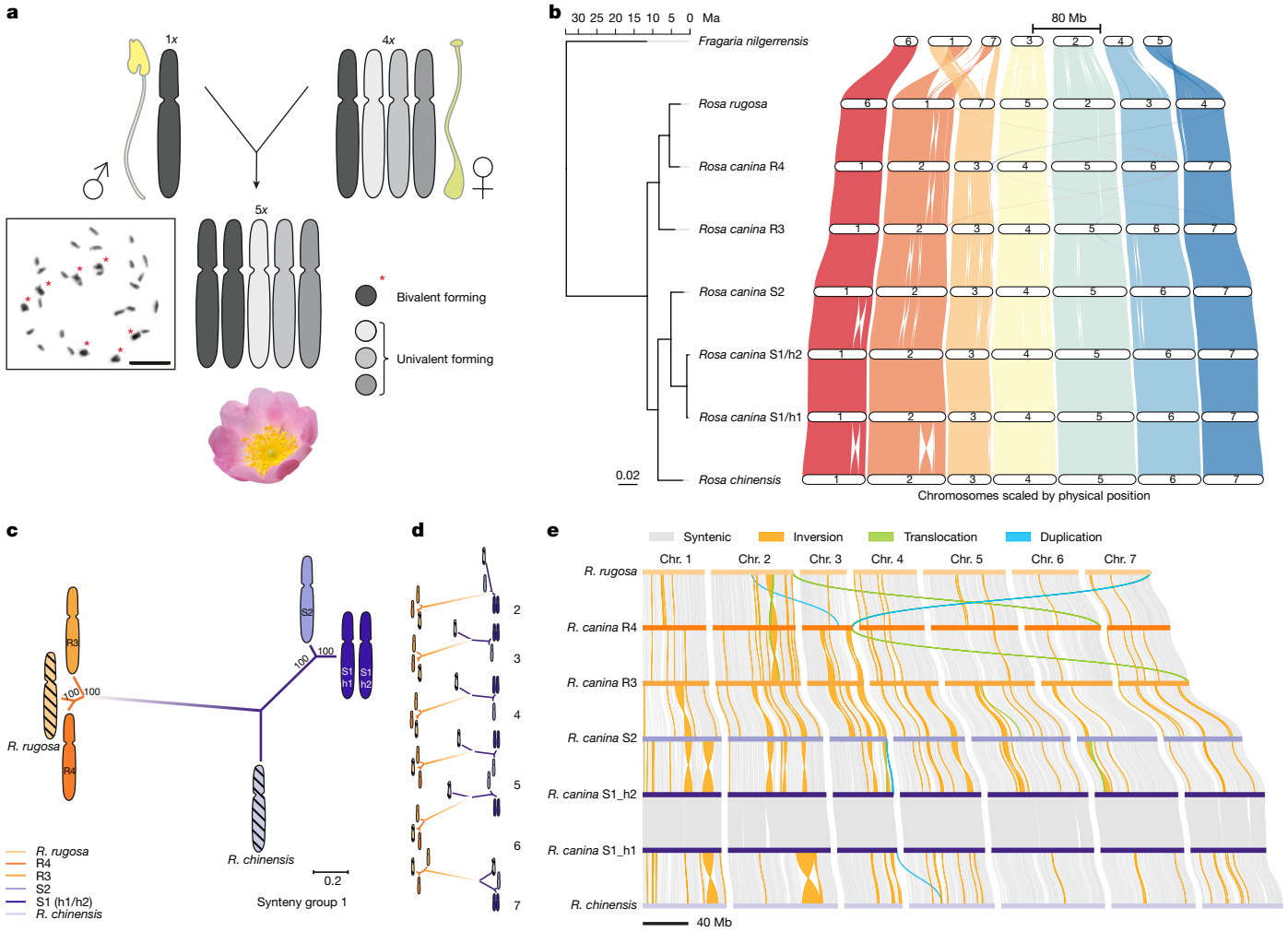


Fig. 1 | Synteny-based classification and phylogenetic relationships of *R. canina* subgenomes. a, The known sexual reproduction of the pentaploid *R. canina* ($2n = 5x = 35$). Different chromosome sets are represented by one chromosome each. During *Canina* meiosis[2,3,22,23], two chromosome sets form bivalents (dark grey and red asterisks) during meiosis and are transmitted through both pollen and egg cells[24–26]. The remaining three sets form univalents (lighter grey) and are transmitted through the egg cell only[24–26]. Diakinesis of male meiosis I ($n = 15$) of *R. canina* is shown on the left. Scale bar, 10 µm. **b**, GENESPACE synteny and phylogenetic relationships of the five chromosome sets of *R. canina* and their close diploid relatives *R. chinensis* (sect. *Synstylae*) and *R. rugosa* (sect. *Rosa*) with a dated phylogenetic tree constructed using 16,372 orthologous genes on the left. Each colour indicates synteny to each *R. chinensis* chromosome, which was used as reference to name the *R. canina* chromosomes. **c**,**d**, Unrooted maximum-likelihood phylogenies of the homoeologous *R. canina* chromosomes and chromosomes from the respective linkage groups of the diploid species *R. chinensis* (sect. *Synstylae*) and *R. rugosa* (sect. *Rosa*) based on alignments of whole-chromosome sequences. The phylogeny of synteny group 1 chromosomes is shown in **c**. Synteny groups 2–7 are shown in **d**. Filled chromosomes refer to subgenomes of *R. canina* belonging to the *Synstylae* clade (violet/light blue) and the *Rosa* clade (dark/light orange). Chromosomes from the diploid roses are indicated by hatching. **e**, Synteny and rearrangement analyses (SyRI) of the *R. canina* genome assembly. Pairwise comparisons of the synteny of all *R. canina* subgenomes (S1_h1/S1_h2, S2, R3 and R4) are juxtaposed against the corresponding chromosomes (chr.) of *R. chinensis* and *R. rugosa*. Only synteny blocks and rearrangement blocks greater than 50 kb in length are shown.

and uniparental transmission of univalents within the same nucleus[20,21]. First observed in the early twentieth century, this mechanism is most common in pentaploid dogroses ($2n = 5x = 35$), in which the male and female parents contribute 7 and 28 chromosomes, respectively, to the zygote[2,3,22,23]. During meiosis, 14 chromosomes form 7 bivalents, while the other 21 chromosomes remain as univalents. Bivalent-forming chromosomes from two highly homozygous sets are transmitted to both sperm and egg cells, whereas univalents are inherited exclusively through the egg cell and excluded from pollen grains, restoring pentaploidy in the offspring through the fusion of haploid male and tetraploid female gametes[24–27] (Fig. 1a). Despite extensive study, the precise cellular mechanisms underlying this asymmetric inheritance remain poorly understood.

Bivalent-forming and univalent chromosomes are thought to have originated from multiple ancient hybridization events[28–30]. Phylogenetic studies based on maternally inherited plastids suggest that dogroses are polyphyletic within the *Synstylae* clade, with subsects. *Rubigineae* and *Caninae* are separated by species exhibiting regular meiosis[13,31–33]. Cytogenetic evidence shows that bivalent-forming chromosomes in subsect. *Caninae* become univalents in subsect. *Rubigineae* and vice versa[34,35]. The rose-specific (peri)centromeric satellite repeat (*CANR4*) is notably enriched in dogrose univalents, possibly linking centromere expansion to their drive during female meiosis[36,37]. Although other uneven polyploid systems with hemisexual reproduction exist[38–40], the meiosis observed in dogroses is unique among eukaryotes. However, the lack of genomic studies has hindered understanding of how centromere properties contribute to this enigmatic reproductive behaviour.

Here we present a high-quality haplotype-phased chromosome-scale assembly of the pentaploid genome of *Rosa canina* (subsect. *Caninae*) and compare it with another *R. canina* individual and *R. agrestis*

(subsect. *Rubigineae*), both from the Darwin Tree of Life (DToL; https://www.darwintreeoflife.org/). Subgenome-aware analyses revealed that dogroses are composed of four subgenomes with one subgenome being present in two highly homozygous haplotypes and the other three in only one. Targeted sequencing of rose-specific single-copy orthologues (SCOs) from pollen DNA, together with the analysis of synthetic F_1 hybrids, confirmed that the two-copy subgenome is biparentally inherited and therefore forming bivalents, while 21 non-recombining univalents from three distinct subgenomes are exclusively inherited through the female germline. Our SCO-based phylogenetic analysis supports the multiple-origin nature of dogroses, as subgenomes from different evolutionary origins were found in pollen, that is, forming bivalents. We also identified a bimodal centromere architecture with small (retrotransposon *ATHILA*-based) and large (tandem repeat *CANR4*-based) centromeres. Notably, *CANR4*-based centromeres were prevalent in univalents, possibly contributing to their drive in asymmetric female meiosis. Our results therefore provide a valuable basis for studying the trade-offs between sexual and asexual reproduction within a single genome.

Unlocking dogrose pentaploid genomes

Although *R. canina* has been recognized as a pentaploid species ($2n = 5x = 35$) for decades, its genome has remained unresolved owing to its complex polyploid structure and hybrid origin. To address this, we assembled a de novo haplotype-resolved, chromosome-level genome using PacBio HiFi sequencing (23× coverage) and chromatin conformation capture (Hi-C) data (Supplementary Table 1). The total size of the assembled 35 pseudochromosomes is about 2.4 Gb, achieving 99.2% completeness in terms of gene content (Extended Data Fig. 1). This high-quality reference genome provides a critical resource for understanding *R. canina*'s genetic features, asymmetric meiosis and hybridization.

All-to-all chromosome alignments revealed seven syntenic groups in *R. canina*, each consisting of five chromosomes (Supplementary Fig. 1). In each group, two chromosome sets consistently exhibited 99–100% similarity, indicating two haplotypes of the same subgenome, while the remaining three chromosome sets showed lower similarities (95–98%; Supplementary Fig. 1 and Supplementary Data 1), which may derive from three different subgenomes. Phylogenetic analyses based on gene and chromosome data, using the diploid rose genomes of *Rosa chinensis* (sect. *Synstylae*)[16] and *Rosa rugosa* (sect. *Rosa*; https://www.darwintreeoflife.org/) as references, revealed that two subgenomes are closely related to *R. chinensis* and were therefore designated 'S', while the other two are more similar to *R. rugosa* (Fig. 1b–e) and were therefore designated 'R'. The two highly similar *Synstylae*-like haplotypes were named S1_h1 and S1_h2, while the more divergent *Synstylae*-like chromosome set was named S2. The *Rosa*-like subgenomes were named R3 and R4, respectively. Moreover, using subgenome-specific *k*-mers[41], we observed that chromosomes assigned to the same subgenome clustered together both in the *k*-mer heat map and the principal component analysis plot, confirming the correct assignment of the four primary subgenomes (S1, S2, R3 and R4; Extended Data Fig. 2a–c). These findings resolve the long-standing question regarding the identity of the five homoeologous chromosomes within each syntenic group.

To further validate the evolutionary relationships of the subgenomes in pentaploid *R. canina*, we conducted orthologous cluster analysis and genome-wide comparisons of synonymous substitution rates (K_s) between *R. canina*, *R. chinensis* and *R. rugosa*. These results were consistent with previous phasing assignments and confirmed the allopolyploid origin of *R. canina* (Supplementary Figs. 2–4). Structure-based pairwise chromosomal analysis across the subgenomes of *R. canina* and genomes of *R. chinensis* and *R. rugosa* revealed a strong conservation of synteny between the two haplotypes of the homozygous S1 subgenome (S1_h1/h2). By contrast, the S1 subgenome showed much lower synteny with the other three subgenomes (S2, R3, R4), which were characterized by large inversions, duplicated regions and translocations (Fig. 1e). Notably, the R3 and R4 subgenomes of *R. canina* exhibit greater synteny to the *R. rugosa* genome than to the other subgenomes of *R. canina* (Fig. 1e), supporting their origination from sect. *Rosa*[35]. Despite the distinct divergence and origins of four subgenomes, a comparison of *R. canina* chromosomes of all subgenomes against *R. rugosa* syntenic chromosomes revealed no evidence of differential fractionation (loss of one or the other copy of a duplicated gene; Supplementary Fig. 5). To detect differential evolutionary rates between subgenomes, we analysed the ratio of nonsynonymous versus synonymous substitution rates and revealed, besides a few outliers, strong purifying selection across orthologous genes in all subgenomes (Supplementary Fig. 6). Together, these results suggest an absence of large-scale subgenome dominance.

We next took advantage of the recent HiFi and Hi-C sequencing datasets from the DToL (https://www.darwintreeoflife.org/) for another *R. canina* accession (European Nucleotide Archive (ENA): PRJEB79802) and from *Rosa agrestis* (subsect. *Rubigineae*; ENA: PRJEB79880) to generate de novo pseudochromosomes (Supplementary Table 2). Comparative analysis revealed a high degree of synteny between our *R. canina* S27 genome and the DToL *R. canina*, both sharing the same subgenome composition (Extended Data Fig. 3a and Supplementary Fig. 7a,b). The *R. agrestis* genome, also pentaploid with 35 pseudochromosomes, displayed a different subgenome composition, with two highly similar haplotypes for the R4 subgenome (R4_h1/h2) and only one copy of the S1 subgenome (Extended Data Fig. 3b–e and Supplementary Fig. 7c,d). Comparative subgenome phasing revealed a gradient of differentiation between the subgenomes of *R. canina* and *R. agrestis*. The S1 subgenomes of both species were the least differentiated and clustered together (Supplementary Fig. 8a). This was followed by the R4 subgenomes, which exhibited a slightly higher degree of differentiation compared with S1 (Supplementary Fig. 8b). The most pronounced differentiation was seen in the S2 and R3 subgenomes, which were distinctly separated in both species (Supplementary Fig. 8c,d). This pattern suggests that the subgenomes S2 and R3 are accumulating more divergence over time.

Tracing the hybridizations of dogroses

To trace the hybridization history of dogroses, we identified subgenome-specific long terminal repeat retrotransposons (LTR-RTs) and estimated their insertion times in *R. canina* and *R. agrestis* to determine the timing of subgenome differentiation before hybridization. Notably, we observed a distinction in the median insertion times of LTR-RTs: the S1, S2 and R3 subgenomes in *R. canina* were estimated to have diverged around 0.7 million years ago (Ma), while the R4 subgenome was older, at approximately 1.2 Ma (median values, 95% confidence intervals; Extended Data Fig. 4a–c). In *R. agrestis*, the median insertion times for S1 and S2 (~0.7 Ma) and for R4 (~1.2 Ma) were the same as for *R. canina*, whereas a slightly older median insertion time was detected for the R3 subgenome (around 0.9 Ma; Extended Data Fig. 4c). These results suggest that the combination of S1, S2 and R3 subgenomes arose at different timepoints in *R. canina* and *R. agrestis*. This is further supported by the K_s-based divergence time estimation obtained from SCOs (Supplementary Fig. 6) and comparable findings from the high differentiation between their R3 subgenomes (Supplementary Fig. 8). Together, these findings suggest that modern dogroses originated through independent, stepwise hybridization events.

Unlocking dogrose reproduction mode

Only bivalent-forming chromosomes are able to segregate properly and produce viable haploid ($1x$) pollen in dogroses. We therefore used flow sorting to isolate pollen nuclei as a proxy to confirm which subgenomes are exclusively pollen-inherited and form bivalents in dogroses. We successfully collected around 200,000 generative nuclei from

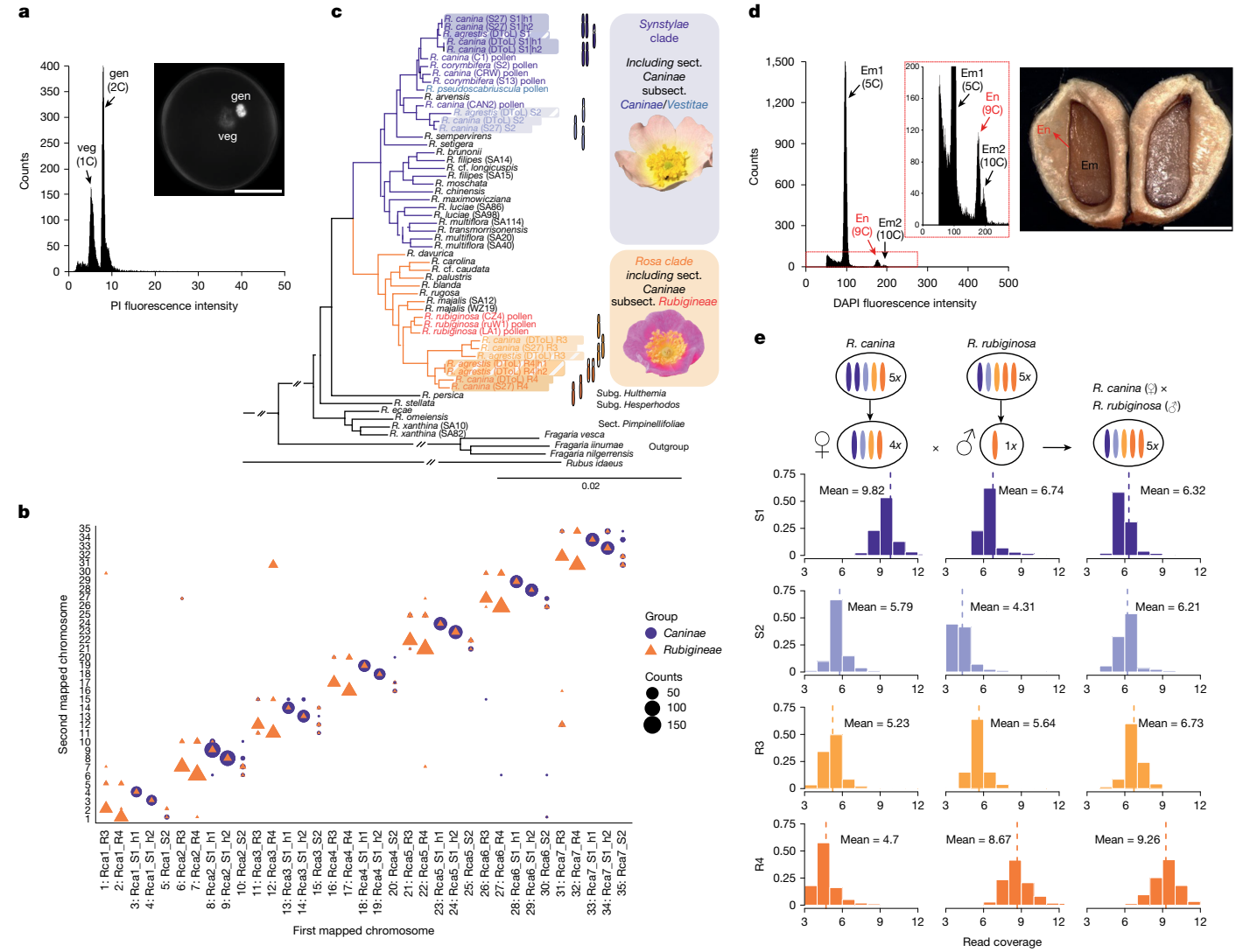


Fig. 2 | Experimental validation of the reproduction mode of pentaploid dogroses. a, Flow cytometry analysis of *R. canina* isolated pollen nuclei. Vegetative (veg) and generative (gen) nuclei differ in their DNA content in the binucleate pollen grains. Inset: an intact pollen grain after DAPI staining of both nucleus types. Scale bar, 10 µm. **b**, Genome-wide pollen SCO mapping of eight dogrose species (subsect. *Caninae*: three samples of *R. canina* and two samples of *R. corymbifera*; subsect. *Rubigineae*: three samples of *R. rubiginosa*) to the *R. canina* S27 genome. The bubble map represents chromosomal hits, which were selectively filtered to display loci with a single alternative hit. The size of the symbols corresponds to the mean counts of pollen SCOs mapped to each chromosomal pair, identifying seven pollen-inherited chromosomes from the S1 subgenome within the *R. canina* (subsect. *Caninae*) genome assembly. In *R. rubiginosa* (subsect. *Rubigineae*), pollen SCO mapped preferentially to the R3 and R4 subgenomes. **c**, Maximum-likelihood phylogeny of the genus *Rosa* based on SCO loci including those retrieved from chromosome-scale assemblies and all pollen samples. Nodes with less than 100% bootstrap are indicated by dashed lines. **d**, Flow cytometry analysis of nuclei isolated from nutlets of *R. canina*, showing an endosperm/embryo ratio of 1.8, corresponding to the expected 9C/9*x* endosperm and 5C/5*x* embryo ratio (in which C denotes unreplicated haploid DNA content; *x* is the basic chromosome number), confirming sexual reproduction and endosperm fertilization. Em1, embryo G0/G1; Em2, embryo G2; En, endosperm G0/G1. Scale bar, 1.5 mm. **e**, Assessment of the parental genomes contribution of a synthetic hybrid between *R. canina* (female donor) and *R. rubiginosa* (male donor). The *x* axis shows the coverage histogram of the short reads from this species mapped to each *R. canina* subgenome. The *y* axis shows the probability densities. The hybrid revealed doubled coverage for R4, indicating the presence of two sets of R4 copies, biparentally inherited, while only one set of maternally inherited S1 was detected, confirming sexual reproduction and the subgenome's inheritance through male and female meiosis.

pollen samples (Fig. 2a) of 9 pentaploid dogrose accessions (including 3 accessions of *R. canina*, 2 of *R. corymbifera*, both subsect. *Caninae*; and 3 of *R. rubiginosa*, subsect. *Rubigineae*) and analysed them using single-copy orthologous nuclear locus target enrichment (Methods). This enabled us to create sample-specific reference sequences for each SCO locus. We mapped a total of 5,794 SCO sequences to the *R. canina* S27 genome and identified 7 major chromosome pairs, with most hits in *R. canina* pollen located on the S1_h1/h2 chromosomes (Fig. 2b and Supplementary Data 2). By contrast, *R. rubiginosa* pollen showed hits primarily on the R3 and R4 chromosomes (Fig. 2b and Supplementary Data 2). By leveraging dogrose pollen SCO mappings on the *R. canina* genome, we unambiguously identified seven pairs of bivalent-forming chromosomes across several dogrose species. Moreover, mapping SCO loci to the *R. agrestis* genome revealed that the R4_h1/h2 subgenome forms bivalents in this species (Extended Data Fig. 5a). These results confirm that bivalent-forming chromosomes in subsect. *Caninae* form univalents in subsect. *Rubigineae* and vice versa[34,35].

Next, we aligned SCO loci obtained from the *R. canina* S27, *R. canina* DToL and *R. agrestis* DToL subgenomes, along with pollen DNA from section *Caninae* and different diploid rose samples and outgroups. This resulted in 58 sequences, totalling 642,158 positions derived from 1,904 concatenated SCO loci. Subgenome-wise as well as chromosome-wise

phylogenetic analysis delineated two large clades within the genus *Rosa*: the *Synstylae* clade and the *Rosa* clade, as well as the earlier splits of subg. *Hulthemia* (*Rosa persica*), subg. *Hesperhodos* (*Rosa stellata*) and sect. *Pimpinellifoliae* (Fig. 2c and Supplementary Data 3). The tree corroborates the allopolyploid origin of dogroses[13,28–30]. In *R. canina* (subsect. *Caninae*), the bivalent-forming subgenome (S1) and the univalent-forming subgenome S2 clustered in the *Synstylae* clade, while the univalent-forming subgenomes R3 and R4 were part of the *Rosa* clade, sister to the European species *R. majalis*. All of the subgenomes of *R. agrestis* grouped with the corresponding *R. canina* subgenome samples, supporting a common origin of individual subgenomes despite the high differentiation observed in S2 and R3 subgenomes (Fig. 2c and Supplementary Fig. 8). However, the pollen SCO data from all *R. rubiginosa* samples grouped as sister to both R3 and R4 subgenomes from *R. canina* and *R. agrestis*, respectively, implying multiple origins of the bivalent-forming chromosomes in subsect. *Rubigineae*. Notably, the pollen SCO data from *R. canina* 'CAN2' was sister to the univalent-forming subgenome S2, suggesting some intraspecific variation in the bivalent-forming subgenomes within *R. canina* (Fig. 2c and Supplementary Data 3), as indicated by variation in microsatellite alleles from bivalent-forming chromosomes in *R. canina*[26]. Supported by the respective clustering of bivalent-bearing pollen data from subsect. *Caninae* and subsect. *Rubigineae*, our data demonstrate an independent origin for the bivalent-forming subgenomes of dogroses and, consequently, the independent origins of asymmetric meiosis[34,35]. Furthermore, the finding that representatives of subsect. *Caninae* and subsect. *Rubigineae* (including the newly generated plastome assemblies for *R. canina* and *R. agrestis*) contain phylogenetically distant plastids from the *Synstylae* clade[13,31–33] (Extended Data Fig. 5b) supports the hypothesis that two progenitors from the *Synstylae* clade formed reciprocal hybrids, which subsequently incorporated *R* genomes through pollen donors.

While haploid (1*x*) pollen nuclei are clearly a product of *Canina* meiosis (Fig. 2a,b), tetraploid (4*x*) sexually derived egg cells were inferred by the respected embryo/endosperm ratio in seeds (Fig. 2d). In sexually reproducing diploids with double fertilization of the egg cell and the polar nuclei, the endosperm/embryo ratio is 1.5 (3*x* endosperm/2*x* embryo; Extended Data Fig. 5c). However, in sexually reproducing 5*x* dogroses, the ratio was found to be 1.8, indicating a 9*x* endosperm and a 5*x* embryo (Fig. 2d), similar to previous findings[42]. To further check the reproduction mode of dogroses, we have investigated the genome composition of two synthetic hybrids obtained from controlled crossing experiments[43]. In the first cross, the female gamete came from *R. canina* (subsect. *Caninae*) and the male donor was *R. rubiginosa* (subsect. *Rubigineae*). As anticipated from the result of the *Canina* meiosis, the subgenome contribution in the hybrid was consistent with the expected 4*x* egg cell containing one copy of each S1/S2/R3/R4 subgenome from *R. canina* and a 1*x* pollen nucleus with the R4 subgenome from *R. rubiginosa* (Fig. 2e and Extended Data Fig. 5d). In the second case, the female gamete came from *R. rubiginosa* and the male donor was *R. corymbifera*—a very close relative of *R. canina*[44]. Again, the hybrid showed the expected subgenome composition, consisting of a male haploid S1 subgenome and a female tetraploid S1/S2/R3/R4 subgenome (Extended Data Fig. 5e,f). These results are in agreement with S1 and R4 being bivalent-forming subgenomes and confirm the 1*x* male versus 4*x* female gamete composition. Our findings further suggest that different subgenomes are potentially interexchangeable in hybridization events; however, hybrids in extant populations originated mostly from unreduced eggs suggesting some subsection-specific differentiation subgenomes, which might impact bivalent formation[45].

The bimodal centromeres of *R. canina*

To gain further insights into the subgenome differentiation of *R. canina*, we aimed to characterize its global repeat composition, both genome-wide and specifically at centromeres. The *R. canina* genome exhibited a very high content of LTR *Ty1/Copia* elements, which made up 40% of the total repeat content, compared with 23% of *Ty3/Gypsy* elements. Among the *Ty1/Copia* elements, the *BIANCA* family accounted for more than 45% of all annotated full-length LTR-RTs while, among the *Ty3/Gypsy* elements, *RETAND* and *ATHILA* were the largest classes found, comprising 10% and 8% of all annotated full-length LTR-RTs, respectively. Tandem repeats, that is, satellite DNA, were mainly composed of the (peri)centromeric *CANR4* repeats[15,20] and rDNA sequences (Fig. 3a, Supplementary Fig. 9 and Supplementary Data 4). The repeat profile across the 35 chromosomes revealed prominent 2–3 Mb peaks of highly dense repeats probably corresponding to the centromeres (Fig. 3a).

To validate the DNA sequences associated with functional centromeres, we developed an *R. canina* centromeric histone H3 (CENH3)-specific antibody and performed chromatin immunoprecipitation followed by sequencing (ChIP–seq). Notably, our results revealed two main types of centromere composition—*Ty3/Gypsy ATHILA* and *CANR4* satellite-based centromeres (Fig. 3a,b). Analysis of the centromere-wide repeat structures revealed that *ATHILA*-based centromeres were most frequent in the chromosomes of the S1 and R4 subgenomes, while larger *CANR4*-based centromeres were found across all chromosomes in the S2 and R3 subgenomes (Fig. 3a–c and Supplementary Fig. 10). *CANR4* centromeric arrays were also found in 3 out of 7 syntenic groups (2, 4 and 5) of the S1_h1/h2 bivalent-forming chromosomes and in two R4 chromosomes, in which these arrays were frequently interrupted by *ATHILA* elements (Fig. 3a–c, Supplementary Fig. 10 and Supplementary Table 3). A similar centromeric sequence composition was observed in both *R. canina* and *R. agrestis* from DToL, despite considerable variation in sequence length (Supplementary Figs. 11 and 12 and Supplementary Data 5–8). Further structural sequence analysis of the diploid relatives *R. chinensis* and *R. rugosa* revealed that *CANR4* repeats are present in only four and three centromeric regions, respectively, while centrophilic *ATHILA* elements were found in all centromeres (Supplementary Data 9–11). Together, our results confirm the expansion and predominance of *CANR4*-based centromeres in exclusively maternally inherited univalent chromosomes in dogroses[20].

Moreover, we identified two centromeres of *R. canina* S27 from the R4 subgenome (Rca1_R4 and Rca4_R4) that lack *CANR4* repeats but exhibited high affinity for CENH3 in regions other than *ATHILA* elements. These two centromeres were characterized by the presence of several tandem-repeat sequences with very long monomers ranging from 1,425 to 2,596 bp. Detailed characterization of these tandem-repeat arrays has revealed that all of these sequences identified are probably derived from different centrophilic *ATHILA* elements, as they all share over 75% similarity with their LTR sequences and are therefore referred to *cenLTR1–4* (Extended Data Fig. 6 and Supplementary Fig. 13). Notably, the *cenLTR* arrays showed significantly higher CENH3 enrichment compared with neighbouring *ATHILA* elements, with the most pronounced enrichment observed in Rca1_R4, which contained a large array of *cenLTR1* (235 kb; Extended Data Fig. 6). Although *cenLTR* arrays found in Rca4_R4 were shorter and characterized by less CENH3 enrichment compared with *cenLTR1* in Rca1_R4, we found two different arrays of *cenLTR2* and *cenLTR3* with higher enrichment than neighbouring *ATHILAs*. Furthermore, the *cenLTR1* monomer sequence showed over 85% similarity to the LTR sequences of *ATHILAs* in Rca4_R4; however, it was not found in tandem arrays in this chromosome (Supplementary Fig. 13). Notably, these *cenLTR* arrays were not detected in either of the DToL genome assemblies of *R. canina* or *R. agrestis* (Supplementary Figs. 11 and 12), suggesting that the formation of these centromeric tandem repeat arrays is a very recent evolutionary event.

To investigate the epigenetic organization within centromeres, we analysed DNA methylation patterns across chromosome arms, scaling from the telomeres to the centromere midpoints. Methylation levels

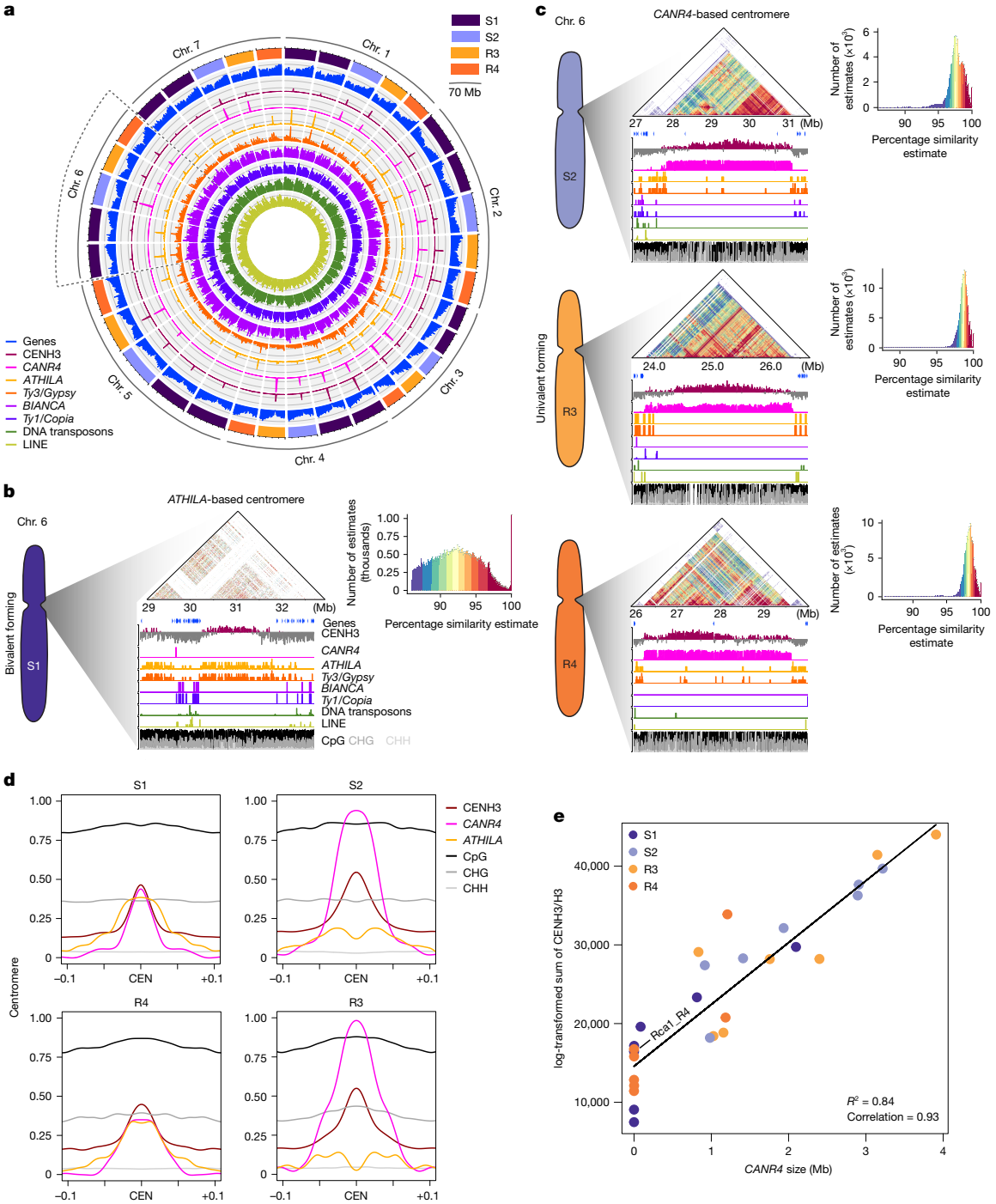


Fig. 3 | The bimodal sequence composition of *R. canina* centromeres. **a**, The global distribution of the main types of repeats identified across all chromosomes. Subgenomes (top right) and sequence tracks (bottom left) are assigned by coloured names. Window size, 100 kb. **b**,**c**, Centromere analysis of tandem-repeat structures in the chromosomes of synteny group 6 (additional detailed plots for the other chromosomes are provided in Supplementary Fig. 10). The sequence structure of bivalent (**b**) and univalent (**c**) centromeres was visualized using ModDotPlot (top triangles). The colour-intensity histograms (top right) show the number of alignments versus pairwise sequence similarity. The sequence tracks plotted below highlight the main classes of repeats identified and the respective association with CENH3 and DNA methylation. The window size in **b** and **c** is 10 kb. For the *y* axes in **a**–**c**, all features were scaled [0, 1]; the original values are provided in Supplementary Fig. 10. **d**, CENH3 ChIP–seq enrichment ($\log_2$[CENH3/H3]) compared with the densities of the centromeric elements *ATHILA* and *CANR4* in 50 kb windows, and DNA methylation in the CpG, CHG and CHH sequence contexts, for each subgenome. Only the centromere proximity regions are shown—10% of the centromere-to-telomere distance. Centromeres (CEN) were defined by the maximum CENH3 enrichment. All signal values (*y* axis) were scaled from 0 to 1 based on the global minimum to global maximum, except for DNA methylation, for which the original percentage values were retained. **e**, Linear regression of *CANR4* size and CENH3 abundance on the centromere across all chromosomes. Each dot coloured by its subgenome presents a centromere. The abundance of CENH3 was calculated by the sum of CENH3 ChIP–seq ($\log_2$[CENH3/H3]) signals on centromeres normalized to coverage. The R^2 of the linear regression model is 0.84. The Spearman's rank correlation is 0.93. Note the high CENH3 enrichment for the *cenLTR1*-based centromere in Rca1_R4, which lacks *CANR4* repeats. Source data are provided in Supplementary Data 15.

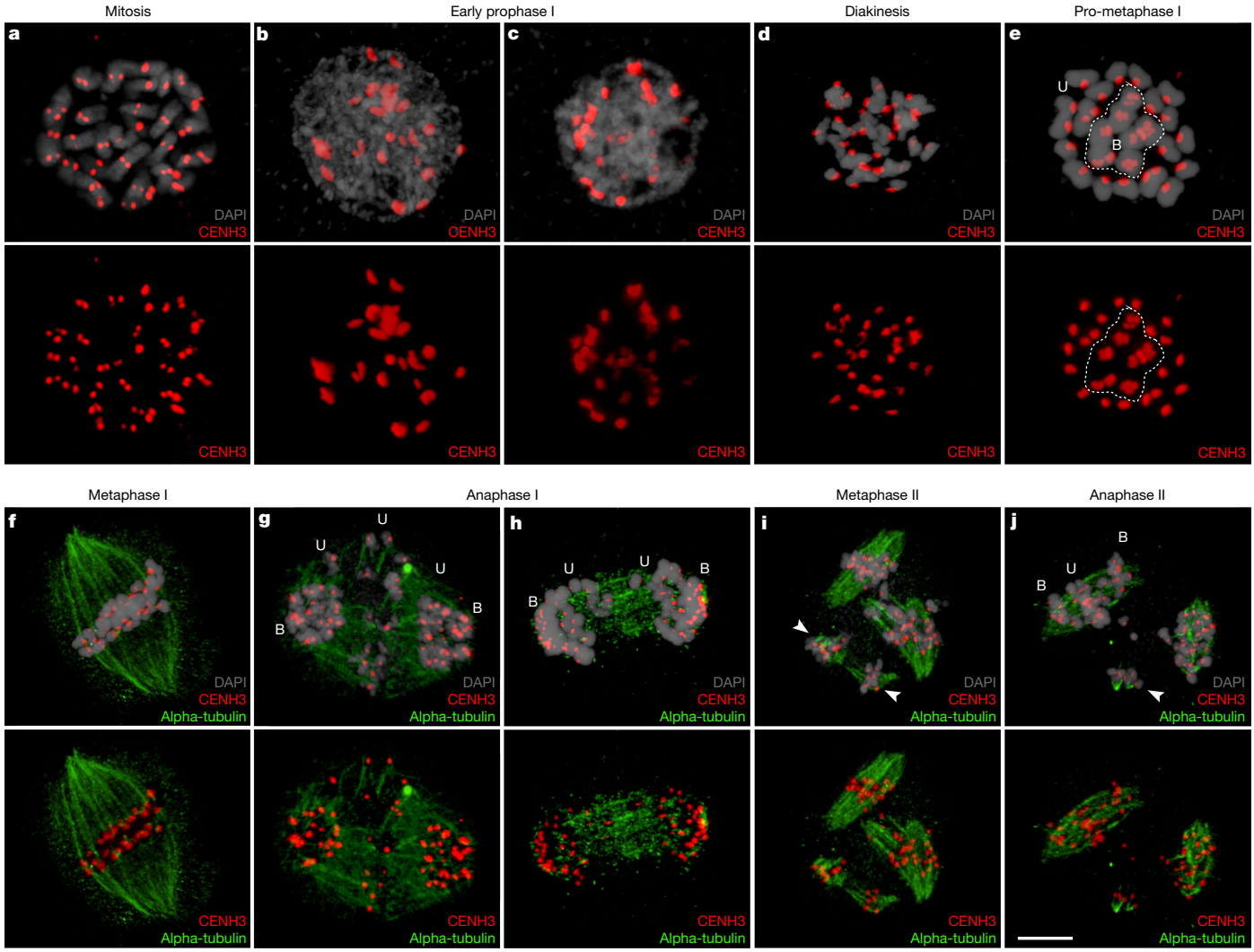


Fig. 4 | Immunodetection of CENH3 and alpha-tubulin in mitotic and meiotic cells of *R. canina*. **a**, Centromeres on all chromosomes were detected by CENH3 protein in mitotic metaphase. Note the size difference in CENH3 signals among different chromosomes. **b**,**c**, Centromeric organization during early prophase I. **d**, The orientation of centromeres during the diakinesis stage. **e**, Metaphase I with univalents (U) surrounding the bivalents (B); the dashed line highlights the typical clustering of bivalents in the middle. **f**, Metaphase I, with all chromosomes organized in the equatorial plate initiating the segregation of homologous pairs (in the case of bivalents) and separation of sister chromatids (in the case of univalents). **g**,**h**, Early (**g**) and late (**h**) anaphase I shows early separation of bivalents, while sister chromatids of univalents usually lagging behind. **i**,**j**, During metaphase II (**i**) and anaphase II (**j**), two main spindles are formed, while lagged chromosomes are still attached to additional abnormal spindles (arrowheads). At anaphase II (**j**), sister centromeres of the bivalent-forming chromosomes finally segregate to form haploid gametes, while single chromatids from univalents lag behind and are eliminated. Experiments were independently repeated at least ten times with similar results to track all meiotic stages represented here. Scale bar, 5 μm (**j**).

were generally elevated at centromeres across CpG, CHG and CHH contexts. However, DNA methylation was slightly reduced, particularly in *CANR4*-based centromeres (Fig. 3d, Extended Data Fig. 7 and Supplementary Fig. 14). These findings suggest distinct methylation patterns in *CANR4* centromeres compared with *ATHILA*-based centromeres. Indeed, *ATHILA* accumulation was less pronounced in the centromeres of univalents of S2 and R3 subgenomes, despite a few insertions being found within *CANR4* centromeric arrays (Fig. 3c and Supplementary Figs. 10 and 14). Notably, *ATHILA* insertions in *CANR4* arrays disrupted CENH3 binding, while *ATHILA*-based centromeres were smaller and showed a lower level of CENH3 association compared with *CANR4*-based centromeres (Fig. 3b–d, Extended Data Fig. 7 and Supplementary Figs. 10 and 14). Frequent insertion of *ATHILA* elements into *CANR4* arrays was also observed in *R. agrestis* and the diploid roses *R. chinensis* and *R. rugosa* (Supplementary Data 7–11). A similar disruption of centromere activity by centrophilic *ATHILA* has been recently found within *Arabidopsis* centromeres[46,47].

We further observed a positive correlation between the amount of *CANR4* repeats and CENH3 abundance along *R. canina* S27 centromeres (Fig. 3e). In fact, the total centromere length, as the measurement of the CENH3-binding regions per subgenome, confirmed that the S2 and R3 centromeres were larger compared with those of S1 and R4 (Fig. 3e, Supplementary Fig. 15 and Supplementary Data 5). Most *CANR4*-based centromeres showed increased levels of CENH3 accumulation compared with the *CANR4*-less ones. Notably, the centromere on Rca1_R4, which is mainly based on a *cenLTR1* array (Extended Data Fig. 6), showed one of the highest enrichments for CENH3 among *CANR4*-less centromeres (Fig. 3e). Thus, tandem repeats bearing different monomer composition seem to accumulate high CENH3 levels in the *R. canina* centromeres.

CENH3 immunostaining of *R. canina* chromosomes revealed differences in the size of individual centromeres (Fig. 4a and Supplementary Video 1; $n = 12$). The size difference was further confirmed by immunostaining analysis of the kinetochore component KNL1[48] (Extended

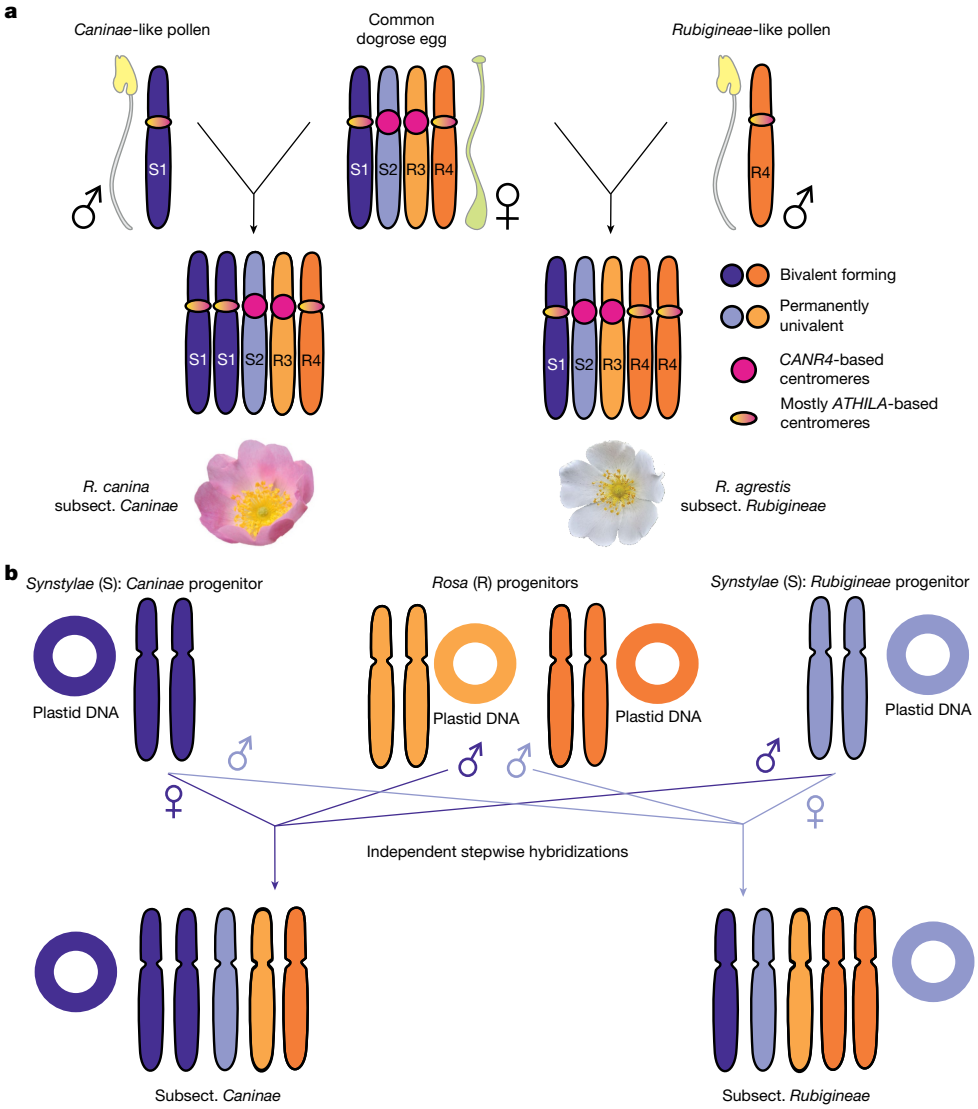


Fig. 5 | Model for the origin and evolution of dogrose centromeres and genomes. a, Reproduction mode and centromere evolution of the pentaploid dogroses. Dogroses of subsect. *Caninae* have two copies of the S1 subgenome, while dogroses of subsect. *Rubigineae* have two copies of the R4 subgenome. The S2 and R3 subgenomes are found as a single copy only. During meiosis in subsect. *Caninae*, the S1_h1/S1_h2 chromosomes form bivalents and are transmitted to both pollen and egg cells. The remaining univalent subgenomes (S2, R3 and R4) are transmitted by the egg cell only. By contrast, in subsect. *Rubigineae*, the R4_h1/R4_h2 chromosomes form bivalents and are transmitted to both pollen and egg cells, while the remaining univalent subgenomes S1, S2 and R3 are transmitted only by the egg cell. The analyses of centromeric sequence composition revealed a dominance in *ATHILA* LTRs in most of the bivalent-forming centromeres of S1 and R4 subgenomes. By contrast, larger *CANR4*-based centromeres were found in all chromosomes of the permanently univalent S2 and R3 subgenomes. **b**, The model for the origin of dogrose subsections. On the basis of the findings that representatives of subsect. *Caninae* and subsect. *Rubigineae* contain phylogenetic distant plastids from the *Synstylae* clade[13,31–33] (Extended Data Fig. 5b), we propose that two progenitors of the *Synstylae* clade formed reciprocal hybrids and additionally incorporated R genomes through pollen donors. Subgenomes are represented by one chromosome. *Synstylae* subgenomes S1 (violet) and S2 (light blue) and *Rosa* clade subgenomes R3 (light orange) and R4 (dark orange) are shown.

Data Fig. 8a,b; $n = 23$). Furthermore, in situ hybridization with probes for *CANR4* and *ATHILA* on mitotic ($n = 15$) and meiotic ($n = 18$) chromosomes of *R. canina* confirmed the predominance of large *CANR4* signals at the centromeres of univalents, while all bivalent and almost all univalents showed *ATHILA* centromeric signals (Extended Data Fig. 8c–k). In the diploids *R. chinensis* and *R. rugosa*, *CANR4* centromeric signals were observed in only three pairs of chromosomes while, again, *ATHILA* was found in all centromeric regions (Extended Data Fig. 8l,m). Thus, the observed difference in the size of CENH3 centromeric signals is probably associated with the accumulation of *CANR4* satellite repeats, supporting our ChIP–seq analysis (Fig. 3e). These results confirm the bimodal architecture of *R. canina* centromeres, which are preferentially *ATHILA*-based in bivalents and *CANR4*-based in univalents, with the caveat that our ChIP–seq experiment was performed using leaf tissue.

Next, we investigated the behaviour of centromeres and spindle dynamics during male meiosis of *R. canina*. Our immunostaining analysis using antibodies against CENH3 and alpha-tubulin clarified the asymmetric distribution of chromosomes during meiosis. In early stages of male meiosis, both small and large centromeres were visible (Fig. 4b–d and Supplementary Videos 2–4; $n = 31$). At onset of metaphase I, we observed seven bivalents organized at the centre of the cell surrounded by 21 univalents (Fig. 4e and Supplementary Video 5; $n = 31$), a configuration that was first proposed over a century ago[2,3,22,23]. During this stage, microtubules facilitated the separation of the homologous pairs through bipolar attachment, while univalents also showed bipolar

attachment (Fig. 4f and Supplementary Video 6; $n = 31$). In early and late anaphase I, homologous chromosomes migrated first, while single chromatids derived from univalents lagged behind (Fig. 4g,h and Supplementary Videos 7 and 8; $n = 31$). Notably, we observed two groups of univalents exhibiting different timing in sister-chromatid separation (Fig. 4g,h (arrows)). During metaphase and anaphase II, we frequently observed both normal and abnormal spindles, resulting from single chromatids lagging behind from anaphase I (Fig. 4i,j (arrowheads)). Homologous pairs derived from bivalents separated normally at end of anaphase II, forming haploid nuclei with seven chromosomes, while single chromatids from univalents lagged behind and were probably eliminated (Fig. 4i,j and Supplementary Videos 9 and 10; $n = 31$). Despite the apparent irregularities in male meiosis, viable pollen grains in *R. canina* (S27) were produced at a rate of approximately 20% (Supplementary Fig. 16a,b). Notably, while at the end of meiosis polyads are formed showing nuclei with varying number of centromere foci (Extended Data Fig. 9 and Supplementary Video 11), the mature binucleate pollen grains seem to contain only a haploid vegetative ($1n = 1x = 7$ chromosomes) and generative ($2n$) nucleus, as confirmed by their genome size and composition (Fig. 2a and Supplementary Fig. 16c). This suggests that, despite meiotic irregularities, a selective mechanism ensures the formation of haploid pollen. Our findings highlight how the atypical centromere behaviour and spindle dynamics during male meiosis in *R. canina* deviates from canonical meiotic processes but can still result in viable pollen production.

Discussion

By leveraging newly generated genome assemblies of dogroses, we shed light onto the long-standing century-old mystery of the unique *Canina* meiosis[2,3,22,23]. Through structural analysis of subgenomes and centromeres, combined with pollen-derived genomics and hybridization experiments, we demonstrate that the bivalent-forming subgenomes in dogroses evolved independently and exhibit distinct interaction patterns during meiosis (Fig. 5).

The bimodal architecture of centromeres in *R. canina* is particularly intriguing when considered alongside its asymmetric female meiosis, in which univalent chromosomes are obligatorily transmitted through the egg cell[3,49]. The prevalence of *CANR4* repeats in univalent centromeres could possibly link centromere expansion with their drive in female meiosis (Fig. 5a), a phenomenon in which larger centromeres are preferentially transmitted during meiosis[37,50,51]. The structural divergence and selective enrichment of *CANR4* repeats in the centromeres of univalents may underpin their larger size, ensuring the preferential transmission of univalents through the egg cell and, therefore, maintaining the pentaploid genome structure. This may represent a rare case of an obligate drive mechanism, functioning in a 'drive or die' manner to maximize the transmission of univalents. However, the occasional presence of *ATHILA*-based centromeres in some univalent chromosomes suggests that *CANR4* expansion alone does not fully explain univalent drive. Furthermore, the structural divergence of centromeres seems also to influence their behaviour in male meiosis, as large *CANR4*-based centromeres in univalents could possibly promote bipolar orientation and premature chromatid separation in male meiosis. Notably, in *Arabidopsis thaliana*, bipolar orientation of univalents happens only when sister chromatid cohesion is defective[52], but it appears to be more common in wheat univalents[53]. This observation contrasts with that of female meiosis, in which univalents seem to have monopolar orientation in dogroses[3] (Supplementary Fig. 17). Thus, a potential role for sexual dimorphism in sister chromatid cohesion regulation could be part of the adaptations enabling *Canina* meiosis.

It is possible that the absence of a homologous pair (and therefore a competing centromere) in the obligatory univalents (S2 and R3 subgenomes) may facilitate the expansion of *CANR4* repeats. By contrast, the S1 and R4 centromeres, which exist within a competitive pairing environment, experience counterbalancing forces that limit *CANR4* accumulation (Fig. 5a). This is further supported by the absence of solely *CANR4*-based centromeres in diploid roses *R. chinensis* and *R. rugosa*. Furthermore, we provide strong evidence for the emergence of few tandem repeats originating from LTR sequences, which outcompete neighbouring *ATHILA* elements for CENH3 binding in two R4 univalent chromosomes (Rca1_R4 and Rca4_R4). These findings highlight the higher affinity of tandem repeats for centromere function in dogroses, further emphasizing the role of centromere composition in shaping meiotic behaviour. However, while the observed correlation is intriguing, we acknowledge that future studies will be essential to confirm whether the expansion of *CANR4* in univalent centromeres is directly linked to their drive during female meiosis.

Despite the lack of recombination, univalent chromosomes retain functional protein-coding genes, as evidenced by high BUSCO completeness and the absence of differential selection pressures. This supports a relatively recent origin of modern dogroses and highlights the resilience of their polyploid genome. Phylogenetic analyses using pollen SCOs and maternally inherited plastid markers further corroborate the polyphyletic origin of dogroses[13,31–33] with multiple hybridization events contributing to their evolutionary history (Fig. 5b). The distinct subgenome ratios observed in *R. canina* (3:2 *Synstylae* to *Rosa*) and *R. agrestis* (2:3 *Synstylae* to *Rosa*) underscore the complexity of their hybrid origins[30,32]. Pollen SCO data align with cytogenetic studies, indicating that bivalent-forming subgenomes in the subsections *Caninae* and *Rubigineae* are phylogenetically distant[34,35]. Moreover, multiple origins for *R. canina* are suggested, as the S1 subgenome clusters with five *R. canina* individuals, while the S2 subgenome clustered with another individual. This model aligns with reports of *Canina*-like meiosis arising spontaneously in hybrids of diploid sexual *Synstylae* species[54], further highlighting the complex hybrid origin and evolutionary dynamics of dogroses.

Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-025-09171-z.

1. Bomblies, K. Learning to tango with four (or more): the molecular basis of adaptation to polyploid meiosis. *Plant Reprod.* **36**, 107–124 (2022).
2. Blackburn, K. B. & Harrison, J. H. W. The status of the British rose forms as determined by their cytological behaviour. *Ann. Bot.* **35**, 159–188 (1921).
3. Täckholm, G. Zytologische studien über die gattung *Rosa*. *Acta Horti Berg.* **7**, 97–381 (1922).
4. Van de Peer, Y., Maere, S. & Meyer, A. The evolutionary significance of ancient genome duplications. *Nat. Rev. Genet.* **10**, 725–732 (2009).
5. Mercier, R., Mezard, C., Jenczewski, E., Macaisne, N. & Grelon, M. The molecular biology of meiosis in plants. *Ann. Rev. Plant Biol.* **66**, 297–327 (2015).
6. Wang, Y. X. & Copenhaver, G. P. in *Annual Review of Plant Biology* Vol. 69 (ed. S. S. Merchant) 577–609 (2018).
7. Pelé, A., Rousseau-Gueutin, M. & Chevre, A. M. Speciation success of polyploid plants closely relates to the regulation of meiotic recombination. *Front. Plant Sci.* **9**, 907 (2018).
8. Herben, T., Suda, J. & Klimešová, J. Polyploid species rely on vegetative reproduction more than diploids: a re-examination of the old hypothesis. *Ann. Bot.* **120**, 341–349 (2017).
9. Wang, Y. Z. & Underwood, C. J. Apomixis. *Curr. Biol.* **33**, R293–R295 (2023).
10. Deb, S. K., Edger, P. P., Pires, J. C. & McKain, M. R. Patterns, mechanisms, and consequences of homoeologous exchange in allopolyploid angiosperms: a genomic and epigenomic perspective. *N. Phytol.* **238**, 2284–2304 (2023).
11. Moore, G. Meiosis in allopolyploids—the importance of 'Teflon' chromosomes. *Trends Genet.* **18**, 456–463 (2002).
12. Debray, K. et al. Unveiling the patterns of reticulated evolutionary processes with phylogenomics: hybridization and polyploidy in the genus *Rosa*. *Syst. Biol.* **71**, 547–569 (2022).
13. Fougere-Danezan, M., Joly, S., Bruneau, A., Gao, X. F. & Zhang, L. B. Phylogeny and biogeography of wild roses with specific attention to polyploids. *Ann. Bot.* **115**, 275–291 (2015).
14. Raymond, O. et al. The *Rosa* genome provides new insights into the domestication of modern roses. *Nat. Genet.* **50**, 772–777 (2018).
15. Saint-Oyant, L. H. et al. A high-quality genome sequence of *Rosa chinensis* to elucidate ornamental traits. *Nat. Plants* **4**, 473–484 (2018).

16. Zhang, X. et al. Haplotype-resolved genome assembly of the diploid *Rosa chinensis* provides insight into the mechanisms underlying key ornamental traits. *Mol. Hortic.* **4**, 14 (2024).
17. Chen, F. et al. A chromosome-level genome assembly of rugged rose (*Rosa rugosa*) provides insights into its evolution, ecology, and floral characteristics. *Hortic. Res.* **8**, 141 (2021).
18. Bourke, P. M. et al. Partial preferential chromosome pairing is genotype dependent in tetraploid rose. *Plant J.* **90**, 330–343 (2017).
19. Zhang, Z. et al. Haplotype-resolved genome assembly and resequencing provide insights into the origin and breeding of modern rose. *Nat. Plants* **10**, 1659–1671 (2024).
20. Lunerová, L. et al. Asymmetrical canina meiosis is accompanied by the expansion of a pericentromeric satellite in non-recombining univalent chromosomes. *Ann. Bot.* **125**, 1025–1038 (2020).
21. Lim, K. Y. et al. Evolutionary implications of permanent odd polyploidy in the stable sexual, pentaploid of *Rosa canina* L. *Heredity* **94**, 501–506 (2005).
22. Blackburn, K. B. Chromosomes and classification in the genus *Rosa*. *Am. Nat.* **59**, 200–204 (1925).
23. Täckholm, G. On the cytology of the genus *Rosa*. A preliminary note. *Sven. Bot. Tidskr.* **14**, 300–311 (1920).
24. Nybom, H., Esselink, G. D., Werlemark, G., Leus, L. & Vosman, B. Unique genomic configuration revealed by microsatellite DNA in polyploid dogroses, *Rosa* sect. *Caninae*. *J. Evol. Biol.* **19**, 635–648 (2006).
25. Nybom, H., Esselink, G. D., Werlemark, G. & Vosman, B. Microsatellite DNA marker inheritance indicates preferential pairing between two highly homologous genomes in polyploid and hemisexual dog-roses, *Rosa* L. sect. *Caninae* DC. *Heredity* **92**, 139–150 (2004).
26. Ritz, C. M. & Wissemann, V. Microsatellite analyses of artificial and spontaneous dogrose hybrids reveal the hybridogenic origin of *Rosa micrantha* by the contribution of unreduced gametes. *J. Hered.* **102**, 217–227 (2011).
27. Ritz, C. M., Köhnen, I., Groth, M., Theißen, G. & Wissemann, V. To be or not to be the odd one out—allele-specific transcription in pentaploid dogroses (*Rosa* L. sect. *Caninae* (DC.) Ser). *BMC Plant Biol.* **11**, 37 (2011).
28. Ritz, C. M., Schmuths, H. & Wissemann, V. Evolution by reticulation: European dogroses originated by multiple hybridization across the genus *Rosa*. *J. Hered.* **96**, 4–14 (2005).
29. Wissemann, V. Molecular evidence for allopolyploid origin in the *Rosa canina*-complex (Rosaceae, Rosoideae). *J. Appl. Bot.* **76**, 176–178 (2002).
30. Ballmer, D. *Dogrose Evolution and its Implications for Conservation*. MSc thesis, Univ. Zuerich (2018).
31. Bruneau, A., Starr, J. R. & Joly, S. Phylogenetic relationships in the genus *Rosa*: new evidence from chloroplast DNA sequences and an appraisal of current knowledge. *Syst. Bot.* **32**, 366–378 (2007).
32. Debray, K. et al. Unveiling the patterns of reticulated evolutionary processes with phylogenomics: hybridization and polyploidy in the genus *Rosa*. *Syst. Biol.* **71**, 547–569 (2021).
33. Wissemann, V. & Ritz, C. M. The genus *Rosa* (Rosoideae, Rosaceae) revisited: molecular analysis of nrITS-1 and *atp*B-*rbc*L intergenic spacer (IGS) versus conventional taxonomy. *Bot. J. Linn. Soc.* **147**, 275–290 (2005).
34. Herklotz, V. et al. The fate of ribosomal RNA genes in spontaneous polyploid dogrose hybrids (*Rosa* L. sect. *Caninae* (DC.) Ser.) exhibiting non-symmetrical meiosis. *Plant J.* **94**, 77–90 (2018).
35. Vozárová, R. et al. Ancient origin of two 5S rDNA families dominating in the genus *Rosa* and their behaviour in the canina-type meiosis. *Front. Plant Sci.* **12**, 643548 (2021).
36. Henikoff, S., Ahmad, K. & Malik, H. S. The centromere paradox: stable inheritance with rapidly evolving DNA. *Science* **293**, 1098–1102 (2001).
37. Henikoff, S. & Malik, H. S. Centromeres: selfish drivers. *Nature* **417**, 227 (2002).
38. Kolarčik, V., Zozomová-Lihová, J., Ducár, E. & Martonfi, P. Evolutionary significance of hybridization in *Onosma* (Boraginaceae): analyses of stabilized hemisexual odd polyploids and recent sterile hybrids. *Biol. J. Linn. Soc.* **112**, 89–107 (2014).
39. Dedukh, D. et al. A cyclical switch of gametogenic pathways in hybrids depends on the ploidy level. *Commun. Biol.* **7**, 424 (2024).
40. Dedukh, D. et al. Maintenance of pure hybridogenetic water frog populations: genotypic variability in progeny of diploid and triploid parents. *PLoS ONE* **17**, e0268574 (2022).
41. Jia, K. H. et al. SubPhaser: a robust allopolyploid subgenome phasing method based on subgenome-specific *k*-mers. *N. Phytol.* **235**, 801–809 (2022).
42. Kolarčik, V., Kocova, V. & Vaskova, D. Flow cytometric seed screen data are consistent with models of chromosome inheritance in asymmetrically compensating allopolyploids. *Cytometry A* **93**, 737–748 (2018).
43. Wissemann, V. & Hellwig, F. H. Reproduction and hybridisation in the genus *Rosa*, section *Caninae* (Ser) Rehd. *Bot. Acta* **110**, 251–256 (1997).
44. Reichel, K. et al. Untangling the hedge: genetic diversity in clonally and sexually transmitted genomes of European wild roses, *Rosa* L. *PLoS ONE* **18**, e0292634 (2023).
45. Herklotz, V. & Ritz, C. M. Multiple and asymmetrical origin of polyploid dog rose hybrids (*Rosa* L. sect. *Caninae* (DC.) Ser.) involving unreduced gametes. *Ann. Bot.* **120**, 209–220 (2017).
46. Naish, M. et al. The genetic and epigenetic landscape of the *Arabidopsis* centromeres. *Science* **374**, eabi7489 (2021).
47. Wlodzimierz, P. et al. Cycles of satellite and transposon evolution in *Arabidopsis* centromeres. *Nature* **618**, 557–565 (2023).
48. Oliveira, L. et al. KNL1 and NDC80 represent new universal markers for the detection of functional centromeres in plants. *Chromosome Res.* **32**, 3 (2024).
49. Blackhurst, H. T. Cytogenetic studies on *Rosa rubiginosa* and its hybrids. *Proc. Am. Soc. Hortic. Sci.* **52**, 510–516 (1948).
50. Malik, H. S. & Henikoff, S. Major evolutionary transitions in centromere complexity. *Cell* **138**, 1067–1082 (2009).
51. Fishman, L. & Saunders, A. Centromere-associated female meiotic drive entails male fitness costs in monkeyflowers. *Science* **322**, 1559–1562 (2008).
52. d'Erfurth, I. et al. Turning meiosis into mitosis. *PLoS Biol.* **7**, e1000124 (2009).
53. Lukaszewski, A. J. Behavior of centromeres in univalents and centric misdivision in wheat. *Cytogenet. Genome Res.* **129**, 97–109 (2010).
54. Wulff, H. D. Über das spontane Auftreten einer Canina-Meiosis bei der Mikrosporogenese der diploiden *Rosa ruga* Lindl. *Oesterreichische Bot. Z.* **101**, 539–557 (1954).

Methods

Plant material

For genome sequencing, we used the same individual of *R. canina* (S27), which has already been cytogenetically analysed[34] (voucher: GLM12396) from a natural stand (WGS84: 51.1732° N; 14.6271° E; Weißenberg, Saxony, Germany). A vegetative runner was dug on 28 March 2022 and planted in a pot. Clones of the collected plant specimen were cultivated in a greenhouse at the Max Planck Institute for Plant Breeding Research, Cologne, Germany.

Whole-genome sequencing

High molecular mass genomic DNA was isolated from leaves using the NucleoBond HMW DNA Kit (Macherey Nagel). A HiFi library was prepared according to the manual of the HiFi SMRTbell Libraries using SMRTbell Express Template Prep Kit 2.0 (Pacific Biosciences) with initial DNA fragmentation using Diagenode Megaruptor 3 and final library size binning into defined fractions by Blue Pippin with 0.75% agarose cassettes (Sage Science). The size distribution was again controlled by a Femto pulse system (Agilent). Size-selected libraries were then sequenced on the Sequel II device with a Binding Kit 2.0 and Sequel II Sequencing Kit 2.0 for 30 h using two SMRT cells (Pacific Biosciences). Moreover, a chromatin-capture library was prepared from 0.5 g of fresh-weight leaf material input. All treatments were performed according to the recommendations of the Dovetail Omni-C kit for plants (Dovetail Genomics). As a final step, an Illumina-compatible library was prepared (Dovetail) and paired-end 2 × 150 bp deep-sequenced on the HiSeq 3000 (Illumina) device. All libraries were sequenced at the Max Planck Genome Centre Cologne at the Max Planck Institute for Plant Breeding Research.

Genome assembly

A phased chromosome-level genome was assembled using the generated PacBio HiFi and Hi-C data. First, a phased primary assembly was obtained by running Hifiasm[55] using 50 Gb of PacBio HiFi reads in combination with Dovetail Omni-C reads with the following command: hifiasm -o out.phased.asm.hic --h1 hic.R1.fastq.gz --h2 hic.R2.fastq.gz hifi.reads.fastq.gz. In the default diploid mode, we generated two sets of phased contigs. Each set was further scaffolded to the chromosome scale using Salsa2[56], followed by successive rounds of manual curation and rescaffolding. We then identified 14 and 21 pseudochromosomes, respectively.

We used Benchmarking Universal Single-Copy Orthologues (BUSCO, v.5.4.0)[57] to evaluate the completeness of 35 chromosome-level scaffolds and for each of the four subgenomes. The lineage database used for running BUSCO was eudicots_odb10. The protein sequences were converted from the assembly using the GFF analysis toolkit AGAT: agat_sp_extract_sequences.pl -g annotation.gff -f genome.fasta -p --type cds -o protein.fasta. The *k*-mer-based tool Merqury (v.1.3)[58] was used to estimate both the completeness and base quality of the chromosome assembly. The quality value of the chromosome assembly was greater than 66.6, and the quality value of each chromosome was at least 62. Read *k*-mers were built from HiFi sequences by Meryl (v1.3) with a *k*-mer size of 31 bp.

Rca_S2 assembly correction. During the scaffolding step, we noted the absence of approximately 20 Mb (including the centromere) on chromosome Rca2_S1_h2. Further validation using fluorescence in situ hybridization (FISH) showed that this chromosome should indeed have a large array of *CANR4*, as found in its homologous chromosome Rca2_S1_h1 (Supplementary Fig. 1). We noted that this assembly error was probably generated by the presence of a small translocation found at the start of the missing region in Rca_S1_h2 compared with Rca_S1_h1. We further mapped the HiFi reads to the region present in Rca_S1_h1 and found robust evidence for the presence of five copies of this region in the genome. We then concluded that this region was incorrectly missing from the Hifiasm assembly due to its high degree of similarity between the S1_h1 and S1_h2 haplotypes. We therefore duplicated the Rca2_S1_h1 region (37265454–54065178 bp) and manually assigned it to the expected position on Rca_S1_h2 (from Rca2_S1_h2: 33518152 bp).

Assembly of DToL datasets. We also downloaded available data from the Darwin Tree of Life (DToL) project for another accession of *R. canina* (PRJEB79801) and for *R. agrestis* (PRJEB79880) and performed phased chromosome-level genome assemblies as described above. *R. agrestis* from DToL revealed two copies of the R4 subgenome; our previous studies suggested that some accessions of *R. agrestis* were of hybrid origin, which should then have two copies of S1 subgenome[45,59].

k-mer analysis for genome size and ploidy level estimation

k-mer analysis to estimate genome size was performed using jellyfish (v.2.3.0)[60] and Genomescope (v.2.0)[61]. The pentaploidy of *R. canina* was further confirmed and analysed using Smudgeplot (v.0.2.5)[61].

Chloroplast genome assembly and phylogeny

To clarify the maternal lineage of the allopolyploid *R. canina* (S27), we assembled the plastid genome of the sequenced individual. GetOrganelle (v.1.7.7.0)[62] was used to de novo assemble the first draft of the plastid genome using 2× 150 bp Illumina short-read data (Sequence Read Archive (SRA): ERS1370372). This toolkit implements Bowtie 2[63] to initially find reads mapped to a plant chloroplast database and SPAdes[64] for de novo assembly and iterative extension. During the assembly and iteration process, BLAST+[65] was used to identify off-target contigs, which were then removed or trimmed. The resulting plastid genome was then used as a reference for mapping the original reads back using Geneious Prime v.2023.2.1 (Biomatters), allowing only mapping of paired reads mapped nearby with a minimum overlap of 75 bp and a minimum overlap identity of 98%. The results were manually examined and corrected where necessary.

The initial annotation of the chloroplast genome was performed using GeSeq (v.2.03)[66]. The annotation included the chloroplast inverted repeats (IRs), *rps12* interspersed gene, protein-coding sequences, tRNAs and rRNAs using 55% identity as thresholds for annotation of proteins and 90% for DNA as well as RNAs. Furthermore, tRNAscan-SE (v.2.0.7)[67] and Chloë (v.0.1.0)[68] were used as additional annotators within GeSeq. The annotations were manually edited using Geneious Prime v.2023.2.1 (Biomatters). The presence of chloroplast genomes differing in the orientation of the single-copy units (large single-copy (LSC) region, small single-copy (SSC) region) was checked by selecting motifs from the border region of the IR and the single-copy units (LSC-trnH-GUG 5′-GGTTCAATTCCCGTCGTTC-3′ or LSC-rps19 5′-GTGACACGTTCACTGAAAAAA-3′ and IRb-rps19-rpl2-IGS 5′-AGACGAAGAAACAAATTCTAT-3′; SSC-ndhF 5′-TGTAATAATATAATAATTGAA-3′ or SSC-ycf15′-CGACCCTAAACGATGGAATCG-3′ and IRa-ycf15′-TTGAAAAACCCGTTGTAACTAT-3′), noting their relative orientation to each other on the same reads using SeqKit (v.2.6.1)[69].

The assembled *R. canina* chloroplast genome had a length of 156,650 bp and a classical quadripartite structure (Supplementary Fig. 18): a LSC of 85,634 bp (~56.57% of the plastid genome), a SSC of 18,878 bp (~12.05%) and two IR regions of 26,069 bp (~16.64% each). Different isomers were found to differ in the orientation of the SSC and LSC (flip-flop configuration).

We computed a chloroplast phylogeny using 37 samples, including sequences downloaded from GenBank and newly assembled data (Supplementary Data 17). The alignment was performed with MAFFT[70], and the phylogenetic tree was calculated using IQ-TREE[71] with the following settings: -m MFP --con-tree --burnin 250 -B 1000 -T 36 --wbtl.

Identification of the bivalent-forming subgenome and comparative analysis

The assembled chromosomes were subjected to pairwise comparisons presented as dot plots using the Synteny Mapping and Analysis (SyMAP)

tool[72]. Multiple alignments within the synteny groups comprising five *R. canina* chromosomes plus *R. chinensis* and *R. rugosa* assemblies were carried out in the CLC Genomics Workbench (CLC) using the 'whole genome alignment' plugin with the following parameters: minimum initial seed length of 250; minimum alignment block length of 250. The aligned chromosomes were subjected to pairwise comparisons. The similarity values were calculated as the block fraction of the two genomes that were aligned (that is, the alignment percentage) or as the percentage of exactly matching nucleotides within the aligned blocks (the average nucleotide identity).

Multifasta files containing assembled short-read sequences of pollen SCO loci from eight different dogrose individuals (three samples of *R. canina*, two samples of *R. corymbifera* (subsect. *Caninae*) and three samples of *R. rubiginosa* (subsect. *Rubigineae*)) that were not sequenced by long reads (Supplementary Data 17) were used as queries to map them with the software BWA[73] with the aln command to the *R. canina* chromosome assembly. From the sequence alignment map (.sam) file, those chromosome hits with only one alternative were filtered according to the 'XA:Z:' flag using a Python script written by GPT-4 (ChatGPT Plus, OpenAI). A bubble map displaying the mean counts of chromosome pairs within different subsections was drawn with ggplot2[74].

Synteny analysis

Chromosome synteny was analysed with the Synteny and Rearrangement Identifier (SyRI)[75]. For this purpose, chromosomes of subgenomes S1_h1, S1_h2, S3, R3 and R4 were aligned against each other within each linkage group (Rca1–Rca7) by minimap2[76,77] using the following command: minimap2 -ax asm5 --eqx -t 16 genome1.fa genome2.fa | samtools sort -@8 > aln.sorted.bam. Moreover, subgenome S1 was also aligned with *R. chinensis* (NCBI: GCA_041222415)[16] and *R. canina* subgenome R4 was aligned with *R. rugosa* (NCBI: GCA_958449725.1; https://www.darwintreeoflife.org/) to analyse its synteny. To keep all of the chromosomes arranged in the same order as *R. canina* and for better visualization, chromosomes 2, 5 and 7 of *R. chinensis* were inverted, and the chromosomes of *R. rugosa* were reordered to 6, 1 (inverted), 7, 5 (inverted), 2, 3 and 4, corresponding to chromosomes 1–7 in *R. canina*. SyRI was implemented for all of the aligned genome pairs using the following command: syri -c aln.sorted.bam -r genome1.fa -q genome2.fa -k -F B --nc 16. Visualization revealed only syntenic blocks over 50 kb, which was performed by plots: Python PLOTsr --sr rugosa_R4/syri.out --sr R4_R5/syri.out --sr R5_s3/syri.out –sr S3_S2/syri.out --sr S2_S1/syri.out --sr S1_chinensis/syri.out --genomes genomes.txt -o out_50k.pdf -S 0.7 -W 10 -H 9 -f 10 --itx -s 50000.

Syntenic orthologues among the primary annotations of diploid strawberry *Fragaria nilgerrensis*[78], *R. chinensis*[16], *R. rugosa* (GCA_958449725.1, DToL; https://www.darwintreeoflife.org/) and the five sets of chromosomes of *R. canina* were inferred using the GENESPACE (v.1.2.3)[79] pipeline with the default parameters. In brief, GENESPACE compares protein similarity scores into syntenic blocks using MCScanX[80] and Orthofinder (v.2.5.4)[81] to search for orthologues/paralogues within synteny constrained blocks. Syntenic blocks were used to query pairwise peptide differences among progenitor alleles, determine divergence among progenitor orthologues using *R. chinensis* syntenic anchors and search for specific orthogroups.

Self-synteny and fractionation bias

Synteny information was obtained using the SynMap tool on the CoGe platform[82,83]. Only genes within synteny blocks were considered, including not only gene pairs but also singleton genes in each genome that lost their counterpart in the other genome due to fractionation or other gene loss. The identification of syntelogues between species was performed using SynMap2 (https://genomevolution.org/wiki/index.php/SynMap2), which internally uses LAST for sequence alignments[84], and then fractionation bias was analysed with FractBias[85].

d*N*/d*S* analysis

Protein-coding sequences (CDSs) were extracted for each *R. canina* subgenome according to coordinates from the gene structural annotation file using GffRead (v.0.12.6)[86] and translated into amino acid sequences using the transeq command from EMBOSS[87]. Additional amino acid sequences and CDSs of *R. rugosa* (BioProject: PRJNA1061178; https://www.darwintreeoflife.org/), *R. chinensis* (NCBI: GCF_002994745.2)[14] and *Fragaria vesca* (NCBI: GCF_000184155.1)[88] were downloaded. The CDS and amino acid sequences were validated, for example, for correct start codons or methionine as the first amino acid in the proteins, using Python scripts. The confirmed proteomes were subsequently analysed using OrthoFinder[81] to identify common single-copy orthologues. According to the protein IDs, FASTA files for each orthologue gene containing five proteins of *R. canina* together with the three of outgroups were aligned with MAFFT (v.7.490)[70]. On the basis of the aligned proteins, corresponding CDSs were codon based aligned using PAL2NAL[89] and DNA alignments were transformed into PHYLIP format. The PAML pipeline[90] with yn00 was used in a looped pairwise mode over all PHYLIP files for each subgenome and outgroup to estimate the nonsynonymous (d*N*) and synonymous (d*S*) substitution rates, as well as their ratio (d*N*/d*S* = ω). The results based on the Yang–Nielsen[91] method were extracted from PAML output files and combined using a Python script and graphical visualized with ggplot2[74] in the R environment[92]. To covert the relative evolutionary time (Time$_t$) from yn00 into absolute divergence time (T_{Ma}) in millions of years ago (Ma), we used *F. vesca* as a fixed calibration point with its fossil record of 2.96 Ma (refs. 93,94). The relative divergence time Time$_t$ for each pairwise compared gene was multiplied with a scaling factor as follows:

$$T_{\mathrm{Ma}} = \mathrm{Time}_t \times \left(\frac{2.96}{\mathrm{Mean\ time}_t \text{ for } F.\ vesca \text{ versus } Rosa} \right)$$

All scripts were developed with the help of ChatGPT-4o (ChatGPT Plus, OpenAI).

Chromosome-level phylogenetic reconstruction

We first generated whole-chromosome multiple alignments in synteny groups 1–7 using the Whole Genome Alignment tool in ClC workbench (Qiagen). The algorithm identifies seeds, that is, short stretches of nucleotide sequences that are shared between multiple genomes but not present multiple times in the same genome. These seeds were then extended using a HOXD scoring matrix, and the HOXD substitution score was combined with an adjustment term based on *k*-mer frequency to avoid spurious matches to repetitive regions in the genome[95]. The program parameters were as follows: minimum initial seed length, 250; minimum alignment blocks, 250; and mismatches in seeds, allowed. The chromosome phylogenies were constructed from multiple alignments using RAxML (v.8.2.12)[96] with the GTRGAMMAI model. The diploid accessions were chromosome assemblies of *R. chinensis*[16] and *R. rugosa*. For *R. rugosa*, the original chromosomes were renamed to fit the *R. chinensis* synteny.

Subgenome-aware phasing of *R. canina*

We used SubPhaser[41] (default parameters) to phase and partition the subgenomes of the pentaploid *R. canina* and *R. agrestis* by assigning chromosomes to subgenomes based on differential repetitive *k*-mers. These were assumed to have expanded during the period of independent evolution after divergence from the nearest common ancestor and before the stepwise hybridization events (the divergence–hybridization period). A subgenome is considered to be well phased when it displays distinct patterns of both differential *k*-mers and homoeologous chromosomes, confirming the presence of subgenome-specific features, as expected. As the S1_h1 and S1_h2 chromosomes represent

haplotypes of the S1 genome, only the S1_h1 haplotype was used in the phasing analysis together with the sets of S2, R3 and R4 chromosomes.

LTR insertion times were calculated by Subphaser as follows: LTR-TRs were de novo detected using LTRharvest (v.1.6.1)[97] and LTRfinder (v.1.07)[98]. To reduce false positives, TEsorter (v.1.3.0)[99] was used to reconstruct the classification of LTR-RTs and further refine this classification. The subgenome-specific *k*-mer sequences were mapped to the LTR-RT sequences using a substring match procedure to identify the subgenome-specific LTR-RTs using the Fisher's exact test. Two LTRs of each subgenome-specific LTR-RT were retrieved and the nucleotide divergence was estimated using the Jukes–Cantor 1969 model. The insertion time (T) was calculated using the equation $T = K/2r$, where $r = 1.3 \times 10^{-8}$ substitutions per year (default)[100], and K represents the divergence of the LTRs from the LTR-RT.

Flow cytometric determination of the endosperm/embryo ratio

To isolate the nuclei from embryo and endosperm tissue, nutlets from fruits of the sequenced individual of *R. canina* S27 (voucher: GLM12396) were first cracked with pliers. The embryo and endosperm were then carefully transferred into a droplet of nuclei isolation buffer (CyStain PI Absolute P; Sysmex-Partec) in a Petri dish and chopped with a sharp razorblade. After adding additional nuclei isolation buffer to a final volume of 500 µl, the nuclei suspensions were filtered through 50 µm disposable filters (CellTrics, Sysmex-Partec), stained with 4′,6-diamidino-2-phenylindole (DAPI) to a final concentration of 1.5 µg ml^{-1} and stored on ice until use. The measurements were performed on a CyFlow Space flow cytometer (Sysmex-Partec) equipped with a high-power UV LED (365 nm).

SCOs

Plant material. To analyse the phylogenetic origin of the subgenomes of allopentaploid *R. canina*, we sampled 30 rose individuals of 24 diploid species across the genus *Rosa*. Thus, *R. stellata* (subgen. *Hesperhodos*), 10 species from sect. *Synstylae*, seven from sect. *Rosa* (*Cinnamomeae*, including sect. *Carolinae*), four from sect. *Pimpinellifoliae* and one individual of *R. chinensis* (sect. *Chinensis*) were sampled from the living collection of the Europa-Rosarium Sangerhausen. Moreover, one accession of *R. majalis* (sect. *Rosa*) was collected from the Botanical Garden Würzburg, and *R. persica* (subg. *Hulthemia*) from Botanical Garden Jena. Species were rechecked using their respective floras[101–103], and the material was compared with available herbarium specimens available online (JSTOR Global Plants, https://plants.jstor.org/; Moscow Digital Herbarium, https://plant.depo.msu.ru). Herbarium vouchers were deposited in the Herbarium Senckenbergianum Görlitz (GLM; Supplementary Data 17).

To determine bivalent-forming genomes, we sampled pollen from several individuals of *Rosa* sect. *Caninae* (subsect. *Caninae*: three 5*x R. canina*, two 5*x R. corymbifera*; subsect. *Vestitae*: one 5*x R. pseudoscabriuscula*; subsect. *Rubigineae*: three 5*x R. rubiginosa*; Supplementary Data 17). We collected anthers from 50 to 100 freshly opened flowers under dry weather conditions in early May 2021 in the field, stored them in open glass for 1 day to allow the anthers to open and subsequently transferred them to a 50 ml tube. Owing to electrostatic attraction, the pollen deposited on the walls of the tube. Anthers were carefully removed, and the pollen powder was collected at the bottom of the tube by gentle centrifugation. The pollen powder was then tapped out over clean paper and transferred to tubes with the help of a spatula. This procedure was repeated three times. Pollen grains were stored in a refrigerator until use.

Isolation and flow sorting of pollen nuclei. Nuclei of mature pollen grains were isolated by the filter bursting method[104] using the nuclear isolation buffer as described previously[105]. Pollen grains were burst on the surface of a 20 µm disposable CellTrics filter (Sysmex-Partec). The resulting nuclear suspension was stained with propidium iodide (50 µg ml^{-1}, PI) and run on a BD Influx cell sorter (BD Biosciences). Nuclear populations were identified in a dot plot showing the PI fluorescence signal (log scale) versus side scatter signal (SSC, log-scale). A sort gate was defined based on the corresponding fluorescence intensity (lin-scale) histogram. A total of 200,000 individual generative nuclei (volume, around 400 µl) were collected into a 1.5 ml reaction tube using the '1.0 Drop Pure' sorting mode of the BD FACS software (BD Biosciences). After adding 50 µl of 1× TE and 50 µl of NaN_3, the nuclei were sedimented by centrifugation (1,000*g* for 10 min at 4 °C). Next, 300 µl of the supernatant was removed, and the nuclei with the remaining liquid were stored at −20 °C. The gating strategy to isolate generative nuclei of *R. canina* is presented in Supplementary Fig. 19.

DNA extraction. DNA from diploid rose species was first extracted from 20 mg of silica-dried leaf tissue according to the ATMAB protocol[106] and subsequently purified using the Mag-Bind Total Pure NGS Kit (Omega Bio-Tek, Nocross) according to the manufacturer's manual. DNA from flow-sorted pollen nuclei was extracted using the Mag-Bind Plant DNA DS Kit with the modification that permanent but careful mixing was performed during binding and elution because the DNA quantities ranged from 37 ng to 236 ng. The DNA yield was quantified using the Qubit 4 Fluorometer (Thermo Fisher Scientific).

Target construction. To analyse nuclear single-copy regions in rose genomes, we used published SCO tags[107]. The SCO tags were originally developed to be amplifiable by PCR and covered coding as well as non-coding regions. We used the 29,000 sequences from additional file 3 from ref. 107, which consisted of SCO tags of 17 rose species and seven outgroup species of the Rosaceae family. These sequences were filtered for uniqueness so that duplicates were removed and searched with BLAST in the *R. chinensis* haploid line genome (v.1.0)[15]. Owing to the structural gene model annotation of the *R. chinensis* genome, we were able to identify 923 full-length nuclear genes with single-copy characteristics. The target-capturing baits were designed by the Agilent bioinformatics service (I. Kisakesen, Agilent Technologies) and covered exons + UTRs with flanking regions and small introns of the selected genes in the *R. chinensis* genome. Finally, the target consisted of 5,794 sequences of different lengths (the shortest at 179 bp and the longest at 6,544 bp) named according to *R. chinensis* gene prediction and had a total size of 2 Mb (Supplementary Data 12). All target sequences were covered by 2× tiling with a total of 85,670 specific baits.

Sequencing. For target enrichment, we used the SureSelect XT HS2 DNA system with precapture pooling (Agilent Technologies) and target design as described above. For diploid roses, 200 ng of input DNA was used, and for pollen DNA, 36–200 ng of input DNA was sheared with a Bioruptor Pico sonication device (Diagenode) to a recommended fragment size of 180–250 bp. The Illumina short-read libraries were amplified for 9 cycles after adapter ligation, pooled for precapture to 16 samples and then postcapture library pools were amplified again with 12 cycles of PCR amplification. The library pools were sequenced in 150 bp paired-end mode on the Illumina NovaSeq 6000 system by Novogene with approximately 1 GB of data output per sample.

To analyse the ploidy of the samples, in vitro flow cytometry was performed on silica-dried leaflets according to a protocol described previously[45] using *R. arvensis* ($2n = 2x = 14$) as an internal standard. The fluorescence intensity was measured using the CyFlow Ploidy Analyser (Sysmex Partec), and the data were analysed using Flowing Software v.2.5.1 (Turku Bioscience Centre). Each sample was measured three times with a minimum of 3,000 particles.

To estimate the ploidy of the samples in silico, we used K-Mer Counter (KMC) (v.3.1.1)[108,109] to generate a *k*-mer database from FASTQ sequence files containing short-read data covering SCOs. The setting was a *k*-mer size of 21, a minimum count for a *k*-mer to be included of 1 and an upper limit for *k*-mer counts of 5,000. To avoid noise, KMC database reduction

was performed using the transform operation with the L 30 and U 5000 settings. With smudgeplot[61] analysis and its hetkzer operation, the coverages of the identified *k*-mer pairs were written to a '_coverages.tsv' file. A custom R script with ggplot2[74] and data.table packages[92] was used to plot the distribution of frequencies of different SNP ratio classes. For each sample, the ploidy level was then estimated by visual inspection of the plots.

Target back-mapping, variant calling and creating a sample-specific reference. The raw SCO reads were trimmed using Trimmomatic (v.0.39)[110] with the following settings: 2:30:8 LEADING:13 TRAILING:13 SLIDINGWINDOW:4:19 MINLEN:36. We updated the script from a previous study[111] to run it with current package versions and used it for mapping, variant calling and sample-specific reference building. In brief, the trimmed short reads from the target enrichment sequencing were mapped against the SCO targets of the *R. chinensis* reference genome (5,794 sequences) using the BWA program[73]. Using SAMtools (v.1.16.1)[112], the reads were sorted and indexed, and duplicates were removed. Notably, approximately 98% of trimmed reads were successfully mapped to the target. Hits with exactly one alternative mapping position were subsequently filtered. After mapping, the Genome Analysis Toolkit (GATK) (v.4.1.9.0)[113] was used with the operation HaplotypeCaller[114] for variant calling, BaseRecalibrator and ApplyBQSR were used to realign around SNPs and indels, and FastaAlternateReferenceMaker was used to create a sample-specific consensus sequence as a reference for each SCO locus in each sample. The provided ploidy level for the HaplotypeCaller was 2 (diploid) for both the pollen and diploid roses with regular meiosis, and the --max-alternate-alleles flag was set to 6, so that although the pollen is monoploid, it would be possible to call potential variances.

Phylogenetic reconstructions based on SCO markers. The SCO target was used as a query for a local search with BLAST+[65] in our *R. canina* S27 genome assembly with a customized output table (-outfmt 6 qseqid sseqid pident length qstart qend sstart send evalue bitscore) and additional in the DToL *R. canina* (PRJEB79801) and *R. agrestis* (PRJEB79880; https://www.darwintreeoflife.org/) genomes also assembled by us (see below). Those SCO loci that had only five hits, one each on subgenomes S1_h1, S1_h2, S2, R3, R4 and R4_h1, R4_h2, R3, S2 and S1 for *R. agrestis*, respectively and within the same linkage group, were filtered and considered single copies. A main list of common single-copy loci for all three genomes was created to preserve the correct order and used to extract the filtered loci from the BLAST outputs with the grep command. The filtered BLAST output was then converted into a BED file containing the sequence coordinates using a bash script written with the help of GPT-4 (ChatGPT Plus, OpenAI). Using the BEDtools (v.2.30.0)[115] command getfasta, sequences for each SCO locus were extracted from the *R. canina* genome assembly and written into a multifasta file. To obtain sequences with the same strand orientation, two locus lists were also created: one of the loci with a positive strand orientation and one with a negative-strand orientation. Loci with negative-strand orientation were identified by calculating the end coordinates minus the start coordinates and filtering according to negative values. According to both lists, the sequences were extracted and stored in two separate multi-FASTA files. Sequences with negative-strand orientation were reversed and complemented with SeqKit[116] and combined with the positive strand-oriented SCO sequences in one fasta file. Finally, for each subgenome (S1_h1/S1_h2, S2, R3, R4 for *R. canina* and R4_h1/h2, R3, S2, S1 for *R. agrestis*), the extracted SCO sequences were concatenated in the same order according to the main locus list and written to subgenome-specific fasta sequences. The same procedure was used for the haploid genome assemblies of *Rubus ideaus* (GenBank: GCA_030142095.1)[117] and three strawberry species, *F. vesca* subsp. *vesca* (GCA_000184155.1)[88], *Fragaria iinumae* (GCA_009720345.1)[118] and *F. nilgerrensis* (GCA_010134655.1)[78] as outgroups. Moreover, the same single-copy loci considered in the genome assembly were extracted and concatenated in the same order with target enrichment samples from nine pollen samples and 30 leaf samples of 26 diploid rose species. The concatenated multilocus sequences were aligned using MAFFT (v.7.490)[70]. Finally, a maximum-likelihood phylogenetic tree was generated by applying IQTREE[71] with ModelFinder using the following settings: iqtree2 -s -m TEST --con-tree --burnin 250 -B 1000 -T 12 --wbtl. The tree figures were graphically finalized with MEGA X[119] and Inkscape v.0.92.3 (2405546, 2018-03-11) software.

Analyses of synthetic hybrids

Synthetic hybrid *R. canina* (seed parent) × *R. rubiginosa* (pollen parent) (sample ID, D62b_2; SRA: SRR15033882) was a cross between *R. canina* (sample ID, D3b_2; SRA: SRR15033883) and *R. rubiginosa* (sample ID, D145b_2; SRA: SRR15033877), and the second synthetic hybrid *R. rubiginosa* (seed parent) × *R. corymbifera* (pollen parent) (sample ID, D166b_2; SRA: SRR15033879) was a cross between *R. rubiginosa* (D145b_2; SRA: SRR15033877) and *R. corymbifera* (sample ID, D81b_2; SRA: SRR15033881). These hybrids were originally produced by Wissemann and Hellwig[43] and kept as a living plant in the Botanical Garden Gießen, Germany. Whole-genome short-read sequencing was performed for both hybrids and their parental plants. The mean coverage of the maternal plant (sample ID, D3b_2) is ~27×, and the paternal plant (sample ID, D145_b2) is ~27×. The hybrid's (sample ID, D62b_2) coverage is ~29×. The reciprocal hybrid (sample ID, D166_b2) has an average of ~30× coverage, whereas its paternal plant (sample ID, D81_b2) is ~19×.

The reads from these six samples were mapped to the S1 subgenome of our assembled *R. canina*, respectively, using bowtie2 (v.2.5.1)[63] with the default parameters. Filtering was applied for all alignments with the same setup 'samtools view -F 3340 --min-MQ 1'. The coverage of each sample was calculated by 'bedtools coverage' (v.2.30.0) with a 100 kb window size. SNPs were called with the filtered alignments by bcftools (v.1.9)[112]. Specifically, 'bcftools mpileup' ran first with the minimum mapping quality 1, then 'bcftools call' ran with flags '--keep-alts --variants-only --multiallelic-caller'. In the end, only the unique SNPs in each parent were selected to calculate the SNP contribution in the hybrids.

ModDotPlot analysis

Structural analysis of DNA sequences of whole chromosomes and centromere cuts were performed with ModDotPlot (v.0.9.0)[120] using the default parameters. ModDotPlot is a dot plot visualization tool designed for large sequences and whole genomes. The method outputs an identity heat map by rapidly approximating the average nucleotide identity between pairwise combinations of genomic intervals.

Gene and repeat sequence annotation

The predicted gene model structures in the nuclear genome were annotated by applying the full-length chromosome sequences to Helixer[121]. Moreover, complete LTR retrotransposons were annotated with the DANTE and DANTE-LTR tools implemented in RepeatExplorer2[122,123]. *R. canina* short-read data (SRA: ERR1662939) were subjected to clustering analysis using the RepeatExplorer2 pipeline, and the output library of repeats was subsequently used to annotate the genome with the implemented RepeatMasker[124]. Tandem-repeat annotation and genome abundance estimation were performed using TAREAN and TideCluster implemented in RepeatExplorer2[122].

RNA sequencing and analysis

Total mRNA was extracted from the leaf tissue of *R. canina* S27 using the Spectrum Plant Total RNA-Kit (Sigma-Aldrich). The RNA-sequencing library was prepared with poly(A) enrichment and then sent for sequencing on the NextSeq 2000 platform with 2 × 150 bp mode, resulting in 33,594,132 reads. For a more accurate mapping of RNA sequences, the annotated tandem repeats and transposable elements

were hard-masked from the genome. RNA alignment was done using hisat2 (v.2.1.0)[125] with the flag --no-mixed. The output was then filtered by only allowing for tag NM:0 and minimum mapping quality 2. To count the transcripts number for each gene, we converted the masked genome to protein sequences based on Helixer[121] structural annotation, and then functionally annotated the protein sequences by Mercator4 (v.7.0)[126] with both Prot-scriber and Swissprot databases, then htseq-count (v.2.0.1)[127] was applied to count the transcripts for all annotated proteins. The gene expression was analysed by DESeq2[128]. As the high homozygosity between the haplotypes of S1 subgenome, RNA reads were aligned to S1_h1, S2, S3, S4 genome, and the expression level of S1 was then halved (Supplementary Fig. 20 and Supplementary Table 4).

CENH3 ChIP–seq experiment and analysis

For detecting the functional centromeres of *R. canina* S27, we designed a specific polyclonal antibody against its CENH3 protein (ARVKHTAARKDRIKTARRQP-C, AB016310), synthetized by LifeTein with immunization in rabbits. The *CENH3* gene of *R. canina* S27 was identified using BLASTP with the parameter '-evalue 1e-5 -qcov_hsp_perc 50' and the *A. thaliana* CENH3 protein HTR12 (AT1G01370) was used as the reference. The ChIP experiment was performed as described previously[129] with a few modifications. Young leaves (around 2–5 g) of *R. canina* S27 were collected and cross-linked in 4% formaldehyde in 1× PBS on ice with vacuum infiltration applied for 30 min. The quenching was performed applying 1 M glycine in each sample followed by vacuum infiltration at room temperature for 15 min. The material was then macerated in liquid nitrogen and the chromatin was extracted. After extraction, the chromatin was sonicated for 30 min on a Bioruptor (Diagenode) until fragments of around 200–600 bp length were achieved (30 s on; 30 s off; in high mode). The sonicated chromatin was incubated over night at 4 °C with 1 µg of each polyclonal antibodies (anti-CENH3 specific for *R. canina* (LifeTein, AB016310) raised in rabbit and anti-histone H3 (Active Motif, 39064) raised in mouse). Samples with no addition of primary antibodies were also incubated as input control samples and at least two experimental replications were used for each ChIP combination. After incubation, protein beads (anti-rabbit: rProtein A Sepharose FastFlow 50% slurry; anti-mouse: rProtein G Sepharose FastFlow 50% slurry (GE Healthcare)) were washed and added to each complex protein–antibody and incubated for at least 2 h at 4 °C in slow rotation. The final recovered chromatin was eluted from the beads, followed by a de-cross-linking step and final DNA extraction. After quality control using the 4200 TapeStation System (Agilent Technologies), the samples were forwarded for 150 bp paired-end Illumina sequencing. For the analysis, the raw 150 bp paired-end reads were quality checked and then mapped to the *R. canina* haplotype phased reference genome using the default parameters in bowtie2[63]. The BAM file was converted to bigwig using the bamCompare tool from deeptools2[130], and then normalized to reads per kilobase of transcript per million reads mapped. Peak calling was then performed using the MACS3 pipeline[131] with the inclusion of the parameters --broad -g 1.9e+9. The plots showing the distribution of different genomic features per chromosome or specific region were constructed using pyGenomeTracks[132]. The ChIP–seq signals in metaplots to compare chromosome (Extended Data Fig. 7 and Supplementary Fig. 14) and subgenome CENH3 enrichment (Fig. 3e) were calculated by bamCompare with parameters '--ignoreDuplicates --scaleFactorsMethod readCount --operation log2' to normalize the CENH3/H3 by read coverage.

Functional centromere annotation

Functional centromere regions in the genome assembly of *R. canina* S27 were annotated based on the detection of CENH3 peaks with MACS3 (see above). The total centromere length was then calculated by the interval between the 5′ and 3′ CENH3 peaks. After alignment to the annotated functional centromeres in *R. canina* S27, comparable centromeric regions were extracted from DToL *R. canina* and *R. agrestis* (https://www.darwintreeoflife.org). The repeat abundance of *CANR4* satellite repeats and *Ty3/Gypsy ATHILA* retrotransposons in the predefined centromeric regions were determined in base pairs for each chromosome of the three investigated *Rosa* genomes (*R. canina* S27, *R. canina* DToL and *R. agrestis* DToL). To reduce data skewness the data were log-transformed. A Shapiro–Wilk normality test was used to check normal distribution of the data with R (v.4.3.3)[92] (29 February 2024). A bivariate Bayesian generalized linear mixed model was implemented using the MCMCglmm package[133]. The model included pairing type (bivalent B, univalent U and univalent in *R. canina* but bivalent in *R. agrestis* Ub) as a fixed effect, while subgenome, genome, and synteny group were random effects, with an unstructured covariance structure (us(trait):random_effect) to account for correlations between response variables. MCMC settings included 100,000 iterations, with a 50,000 burn-in and a thinning interval of 50 and the family parameter was set according to 'gaussian'. Data visualization was performed using the ggplot2[74], patchwork[134], tidyr[135] and dplyr[136] packages (Supplementary Fig. 15; source data are available in Supplementary Data 13 and 16). The correlation of *CANR4* size with CENH3 abundance (Fig. 3e) was calculated by Spearman's rank correlation as the Shapiro–Wilk normality test resulted in $P \ll 0.5$. Linear regression model was fitted using the lm function in R (v.4.4.0), with multiple R^2 value as 0.842 and adjusted R^2 value as 0.836.

cenLTR sequence characterization

cenLTR sequences were primarily annotated as tandem repeats using TAREAN and TideCluster implemented in RepeatExplorer2[122]. Further sequence similarity with LTR retrotransposons was performed using the transfer annotation tool of Geneious Prime v.2025.0.2 (https://www.geneious.com) with a minimum sequence similarity threshold of 75%. Using a Geneious Prime plugin for ClustalO[137], we performed alignments of consensus *cenLTR* sequences against the regions with the highest similarity found in the *R. canina* S27 genome, which all corresponded to different *ATHILA* elements on chromosomes Rca1_R4 and Rca4_R4. Consensus sequences of *cenLTR1–4* are available in Supplementary Dataset 14.

DNA methylation sequencing and analysis

To investigate the methylome of *R. canina*, we performed enzymatic methyl-sequencing (EM-seq). For this, we extracted genomic DNA from young leaves and the samples were then prepared for an Illumina-compatible library using the NEBNext Enzymatic Methyl-seq Kit and further sequenced on the HiSeq 3000 device with paired-end orientation. We ended up with 68,632,618 pairs of 150 bp reads. EM-seq data were first aligned to the S1_h1, S2, R3, R4 combined subgenomes with Bismark (v.0.23.0) with the flag '--local' and duplications were removed by deduplicate_bismark. CpG-, CHG- and CHH-context methylations were then extracted by bismark_methylation_extractor (v.0.23.0). The output was converted to bedgraph by bismark2bedGraph (v.0.23.0) with the flag '-CX' activated for CHG and CHH contexts to visualize the methylations chromosome-wide and on the centromeres.

Metaplots of CENH3 enrichment, DNA methylation, *ATHILA* and *CANR4* density

In the metaplots (Fig. 3d, Extended Data Fig. 7 and Supplementary Fig. 14), all signals were smoothed by the spline.smooth function with spar 0.3 in R (v.4.4.0). CENH3 enrichment was calculated by CENH3 ChIP–seq ($\log_2$[CENH3/H3]) signal normalized by coverage. CENH3 enrichment, DNA methylations, *ATHILA* density and *CANR4* were calculated in 50 kb adjacent windows and averaged by all chromosomes of the corresponding subgenome. All chromosome coordinates were scaled on the basis of their distance to centromere against the distance of centromere to telomere. Centromere position was defined

on the basis of where the maximum CENH3 enrichment was located. Mitochondrial sequences were masked when computing the CENH3 enrichment. All signal values (*y* axis of metaplots) were scaled from 0 to 1 based on the global minimum to global maximum except for DNA methylations, for which the original percentage values were retained. The p- and q-arm values were averaged and mirrored.

Immunodetection of CENH3 and microtubules

For immunodetecting the centromeres of *R. canina* S27, we used polyclonal antibodies against CENH3 protein (see above) and kinetochore protein KNL1 (C-EDHFFGPVSPSFIRPGRLSD, AB015677-3) described previously[48], also synthetized by LifeTein and raised in rabbits. To identify the microtubules, we used a commercial antibody against alpha-tubulin (Sigma-Aldrich, T6199) with immunization in mouse. For analysing the distribution of these markers in mitotic cells, root tips were fixed after a pretreatment in 0.2 mM 8-hydroxyquinoline for 4 h at 18 °C. For meiotic stages, the young anthers were directly fixed with no previous antimitotic pretreatment. The immunodetection experiment was performed according to a previously published protocol[138] with modifications to *R. canina* material. Young flower buds were collected on ice in buffer A (15 mM PIPES–NaOH, 80 mM KCl, 0.5 mM ethylene glycol tetraacetic acid, 80 mM sorbitol, 20 mM NaCl, 2 mM ethylenediaminetetraacetic acid (EDTA), 0.15 mM spermine, 0.5 mM spermidine and 1 mM dithiothreitol) and next incubated in 4% paraformaldehyde in buffer A for 1 h under vacuum infiltration on ice. After the fixation, the samples were washed three times with buffer A and then digested in enzymatic solution containing 1% cellulase-onozuka, 1% cellulase, 1% pectolyase Y23, 1% cytohelicase, 1% macerozyme and 10% pectinase in citrate buffer for 1 h in a humid chamber at 37 °C. To remove the excess of enzymatic solution, the material was gently washed with buffer A and left on ice until the preparation of the slides. A couple of anthers were placed and dissected in a drop of buffer A on a 18 × 18 mm high-precision coverslips, a few µl of polyacrylamide solution (25 µl 15% polyacrylamide (Sigma-Aldrich, A3574) in buffer A plus 1.25 µl of 20% sodium sulfite and 1.25 µl of 20% ammonium persulfate) were added to the dissected anthers, quickly mixed and a second coverslip was put above the first making a sandwich gently squeezing the anthers with a needle to liberate the meiocytes. The sandwiches were allowed to dry for up to 1 h until complete polymerization. After this, the coverslips were carefully separated and incubated in PBS with 1% Triton X-100 and 1 mM EDTA for at least 1 h, then more 2 h in blocking solution containing 3% BSA in PBS with 0.1% Tween-20. After this period, the primary antibodies were diluted in 1:500 (CENH3 and KNL1) and 1:200 (alpha-tubulin) ratios in blocking solution and applied on each sample, which were sequentially incubated at 4 °C for 48 h. After primary antibody incubation, primary antibodies were detected using secondary antibodies conjugated with specific fluorophores (Alexa Fluor 488 and Abberior StarRed and STARORange for STED microscopy, also diluted in blocking solution in a proportion of 1:250) and incubated in a dark humid chamber at room temperature for at least 2 h. The material was then washed four to five times for 20 min each in 1× PBS + 0.1% Triton X-100 and then mounted in SlowFade Gold medium containing DAPI. The slides were photographed using a super-resolution STED microscope (Abberior instrument facility line; https://abberior-instruments.com/) and posterior brightness and contrast adjustments were done in Photoshop.

Chromosome preparation and FISH

For mitotic chromosome preparations, root tips and young flower buds from *R. canina* S27 plants cultivated in the greenhouse were collected and then fixed in methanol:acetic acid solution (3:1 (v/v)) for 2–24 h at room temperature and then kept at −20 °C until use. After fixation, the root tips were pretreated with an enzymatic solution of 2% cellulase R10-onozuka (Duchefa Bioquemie)/20% pectinase (Sigma-Aldrich) in 0.1 M citric acid for 40 min at 37 °C in a humid chamber and then squashed in a drop of LB01 buffer (15 mM Tris, 2 mM Na_2EDTA, 80 mM KCl, 20 mM NaCl, 0.5 mM spermine, 15 mM β-mercaptoethanol, 0.1% Triton X-100 (pH 7.5)) and, after frozen in liquid nitrogen, the coverslips were removed.

For meiotic chromosome preparations, the anthers of *R. canina* C1 (GLM-P-0181117) were dissected from fixed flower buds around 0.5 cm in length. Anthers were washed with 1% (w/v) polyvinylpyrrolidone 40 (PVP-40; Sigma-Aldrich Chemie) and 0.5% (v/v) Triton X-100 for 15–20 min, followed by enzymatic digestion overnight in a humid chamber at 4 °C in 1% (w/v) cellulase Onozuka R-10 (Serva), 0.2% (w/v) pectolyase Y-23 (Sigma-Aldrich), 0.5% (w/v) hemicellulose (Sigma-Aldrich) and 0.5% (w/v) macerozyme R-10 (Duchefa Biochemie) dissolved in citric buffer (0.04 M citric acid and 0.06 M sodium citrate). The anthers were macerated on a slide, squashed in a drop of 70% acetic acid and fixed by freezing in liquid nitrogen.

For FISH experiments, a 22 bp oligo probe directly labelled with a Cy3 fluorophore at the 5′ terminus was designed, synthesized by Sigma-Aldrich and then used to detect the *CANR4* satellite repeat (Cy3-5′-ACCCTAGAAGCAAGAAGTTTGG-3′) or an insert of the plasmid carrying the *CANR4* dimer (GenBank MK069593) was used as a FISH probe[20]. For detection of the centromeric LTR *ATHILA* retrotransposon sequences, we designed a probe based on clustering analysis of Illumina reads (SRA: ERR1662939) using the RepeatExplorer2 pipeline. It was revealed that cluster 5 (CL5) contig contained *Ty3/Gypsy/ATHILA* sequences. The CL5 contig was used to design PCR primers amplifying a 180 bp product from *R. canina* genomic DNA. The primers were as follows: Rcan_centr_CL5_for: 5′- GCAAGCGCATAATTTAACC-3′ and Rcan-centr_CL5_rev: 5′-CAATCAAAAATATCCCCCC-3′. The PCR product was purified and cloned into the pDrive vector (Qiagen) and sequenced by the Sanger dideoxy method using the SP6 primer (Micosynth). Clone 11 was submitted to GenBank (PV030978). The inserts of plasmids were directly labelled in a nick translation reaction with Cy5 d-UTP or Cy3 d-UTP fluorochromes (Jena Bioscience) and used for FISH. To detect the 5S and 35S rDNA loci, the full-length 18S rRNA gene from tomato (GenBank: X51576.1) and the *Pta71* clone from *Triticum aestivum* were used to detect the 35S rDNA region, while a 5S rDNA unit (B variant) from *R. canina*[35] and the D2 clone from *Lotus japonicus* were used to detect the 5S rDNA locus. rDNA robes were directly labelled by Nick translation using Cy5 d-UTP (Jena Bioscience). The slides were prepared in accordance with the protocols described previously[35,139]. In brief, the slides were treated with pepsin solution (1 mg ml^{-1} diluted in 0.01 N HCl) for 30 min at 37 °C in a humid chamber, washed with 2× SSC (saline sodium citrate, pH 7.0) solution, post-fixed with 4% paraformaldehyde for 10 min at room temperature, washed again with 2× SSC and then dried in 70% and 100% ethanol. After air drying for at least 30 min, the slides were denatured with hybridization mix (50% formamide, 2× SSC, 10% dextran sulfate and ~50 ng of each probe (15 µl per slide)) for 5 min at 75 °C and then incubated for at least 18 h at 37 °C. After hybridization, stringency washes were performed with 2× and 0.1× SSC solutions at 42 °C, achieving around 76% stringency. The slides were then washed at room temperature with 2 × SSC solution and mounted with DAPI in the antifade mounting medium Vectashield (Vector Laboratories).

Alexander staining

Five mature and well open flowers were collected from the plant in the greenhouse. They were shaken above a microscope slide and their pollen was released on top of the slide. Then, 20 µl of Alexander staining solution (Morphisto, 13441.00250) was added and briefly mixed with the pollen by stirring with the pipette tip. A coverslip 24 × 40 mm was put on top of the mix. Pictures were taken with a Labscope microscope by Zeiss, using 10 × magnification. Five snapshots were counted with the help of the ZEN software.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

All raw sequencing data (HiFi, Hi-C, RNA, CENH3-ChIP, DNA methylation, SCO of pollens) and genome assembly of *R. canina* S27 isolate are available under NCBI BioProject: PRJNA1111045. The chloroplast genome of *R. canina* S27 isolate available under GeneBank accession number PV550499. Raw sequencing data of *R. canina* DToL and *R. agrestis* DToL are available from Darwin Tree of Life (DToL) data portal (https://portal.darwintreeoflife.org/). The corresponding NCBI BioProject accession numbers are PRJEB79802 and PRJEB79880, respectively. Genome assemblies, the sample-specific SCO reference sequences, variant calling format files, annotations and alignments presented in this work are also available for download at Dryad[140] (https://doi.org/10.5061/dryad.cc2fqz6fh). The REXdb database Viridiplantae v.3.0 (http://repeatexplorer.org/?page_id=918) is publicly available. All other data needed to evaluate the conclusions in the paper are provided in the Article and its Supplementary Information.

55. Cheng, H. Y., Concepcion, G. T., Feng, X. W., Zhang, H. W. & Li, H. Haplotype-resolved de novo assembly using phased assembly graphs with hifiasm. *Nat. Methods* **18**, 170–175 (2021).
56. Ghurye, J. et al. Integrating Hi-C links with assembly graphs for chromosome-scale assembly. *PLoS Comput. Biol.* **15**, e1007273 (2019).
57. Manni, M., Berkeley, M. R., Seppey, M., Simao, F. A. & Zdobnov, E. M. BUSCO update: novel and streamlined workflows along with broader and deeper phylogenetic coverage for scoring of eukaryotic, prokaryotic, and viral genomes. *Mol. Biol. Evol.* **38**, 4647–4654 (2021).
58. Rhie, A., Walenz, B. P., Koren, S. & Phillippy, A. M. Merqury: reference-free quality, completeness, and phasing assessment for genome assemblies. *Genome Biol.* **21**, 245 (2020).
59. Herklotz, V. & Ritz, C. M. Spontane Hybridisierung von Hundsrosen (*Rosa* L. sect. *Caninae* (DC). Ser.) an einem natürlichen Vorkommen in der Oberlausitz (Sachsen, Deutschland). *Peckiana* **9**, 119–131 (2014).
60. Marcais, G. & Kingsford, C. A fast, lock-free approach for efficient parallel counting of occurrences of *k*-mers. *Bioinformatics* **27**, 764–770 (2011).
61. Ranallo-Benavidez, T. R., Jaron, K. S. & Schatz, M. C. GenomeScope 2.0 and Smudgeplot for reference-free profiling of polyploid genomes. *Nat. Commun.* **11**, 1432 (2020).
62. Jin, J. J. et al. GetOrganelle: a fast and versatile toolkit for accurate de novo assembly of organelle genomes. *Genome Biol.* **21**, 241 (2020).
63. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359 (2012).
64. Bankevich, A. et al. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J. Comput. Biol.* **19**, 455–477 (2012).
65. Camacho, C. et al. BLAST+: architecture and applications. *BMC Bioinform.* **10**, 421 (2009).
66. Tillich, M. et al. GeSeq—versatile and accurate annotation of organelle genomes. *Nucleic Acids Res.* **45**, W6–W11 (2017).
67. Chan, P. P., Lin, B. Y., Mak, A. J. & Lowe, T. M. tRNAscan-SE 2.0: improved detection and functional classification of transfer RNA genes. *Nucleic Acids Res.* **49**, 9077–9096 (2021).
68. Small, I. & Castleden, I. Chloë v0.1.0 (GitHub); https://github.com/ian-small/chloe (2020).
69. Shen, W., Le, S., Li, Y. & Hu, F. SeqKit: a cross-platform and ultrafast toolkit for FASTA/Q file manipulation. *PLoS ONE* **11**, e0163962 (2016).
70. Katoh, K. & Standley, D. M. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* **30**, 772–780 (2013).
71. Nguyen, L. T., Schmidt, H. A., von Haeseler, A. & Minh, B. Q. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol. Biol. Evol.* **32**, 268–274 (2015).
72. Soderlund, C., Bomhoff, M. & Nelson, W. SyMAP v3.4: a turnkey synteny system with application to plant genomes. *Nucleic Acids Res.* **39**, e68 (2011).
73. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).
74. Wickham, H. *ggplot2: Elegant Graphics for Data Analysis* (Springer, 2016).
75. Goel, M., Sun, H. Q., Jiao, W. B. & Schneeberger, K. SyRI: finding genomic rearrangements and local sequence differences from whole-genome assemblies. *Genome Biol.* **20**, 277 (2019).
76. Li, H. Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics* **34**, 3094–3100 (2018).
77. Li, H. New strategies to improve minimap2 alignment accuracy. *Bioinformatics* **37**, 4572–4574 (2021).
78. Zhang, J. X. et al. The high-quality genome of diploid strawberry (*Fragaria nilgerrensis*) provides new insights into anthocyanin accumulation. *Plant Biotechnol. J.* **18**, 1908–1924 (2020).
79. Lovell, J. T. et al. GENESPACE tracks regions of interest and gene copy number variation across multiple genomes. *eLife* **11**, e78526 (2022).
80. Wang, Y. P. et al. MCScanX: a toolkit for detection and evolutionary analysis of gene synteny and collinearity. *Nucleic Acids Res.* **40**, e49 (2012).
81. Emms, D. M. & Kelly, S. OrthoFinder: phylogenetic orthology inference for comparative genomics. *Genome Biol.* **20**, 238 (2019).
82. Lyons, E. & Freeling, M. How to usefully compare homologous plant genes and chromosomes as DNA sequences. *Plant J.* **53**, 661–673 (2008).
83. Lyons, E. et al. Finding and comparing syntenic regions among *Arabidopsis* and the outgroups papaya, poplar, and grape: CoGe with Rosids. *Plant Physiol.* **148**, 1772–1781 (2008).
84. Kielbasa, S. M., Wan, R., Sato, K., Horton, P. & Frith, M. C. Adaptive seeds tame genomic sequence comparison. *Genome Res.* **21**, 487–493 (2011).
85. Joyce, B. L. et al. FractBias: a graphical tool for assessing fractionation bias following polyploidy. *Bioinformatics* **33**, 552–554 (2017).
86. Pertea, G. & Pertea, M. GFF utilities: GffRead and GffCompare. *F1000Research* **9**, 304 (2020).
87. Rice, P., Longden, I. & Bleasby, A. EMBOSS: the European Molecular Biology Open Software Suite. *Trends Genet.* **16**, 276–277 (2000).
88. Shulaev, V. et al. The genome of woodland strawberry (*Fragaria vesca*). *Nat. Genet.* **43**, 109–116 (2011).
89. Suyama, M., Torrents, D. & Bork, P. PAL2NAL: robust conversion of protein sequence alignments into the corresponding codon alignments. *Nucleic Acids Res.* **34**, W609–W612 (2006).
90. Yang, Z. H. PAML: a program package for phylogenetic analysis by maximum likelihood. *Comput. Appl. Biosci.* **13**, 555–556 (1997).
91. Yang, Z. H. & Nielsen, R. Estimating synonymous and nonsynonymous substitution rates under realistic evolutionary models. *Mol. Biol. Evol.* **17**, 32–43 (2000).
92. R Core Team. *R: A Language and Environment for Statistical Computing* (R Foundation for Statistical Computing, 2024).
93. Matthews, J. V., Westgate, J. A., Ovenden, L., Carter, L. D. & Fouch, T. Stratigraphy, fossils, and age of sediments at the upper pit of the Lost Chicken gold mine: new information on the late Pliocene environment of east central Alaska. *Quat. Res.* **60**, 9–18 (2003).
94. Matthews, J. V. & Ovenden, L. E. Late Tertiary plant macrofossils from localities in Arctic/Subarctic North America: a review of the data. *Arctic* **43**, 364–392 (1990).
95. Darling, A. E., Mau, B. & Perna, N. T. progressiveMauve: multiple genome alignment with gene gain, loss and rearrangement. *PLoS ONE* **5**, e11147 (2010).
96. Stamatakis, A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30**, 1312–1313 (2014).
97. Ellinghaus, D., Kurtz, S. & Willhoeft, U. LTRharvest, an efficient and flexible software for de novo detection of LTR retrotransposons. *BMC Bioinform.* **9**, 18 (2008).
98. Xu, Z. & Wang, H. LTR_FINDER: an efficient tool for the prediction of full-length LTR retrotransposons. *Nucleic Acids Res.* **35**, W265–W268 (2007).
99. Zhang, R. G. et al. TEsorter: an accurate and fast method to classify LTR-retrotransposons in plant genomes. *Hortic. Res.* **9**, uhac017 (2022).
100. Vitte, C., Panaud, O. & Quesneville, H. LTR retrotransposons in rice (*Oryza sativa*, L.): recent burst amplifications followed by rapid DNA loss. *BMC Genom.* **8**, 218 (2007).
101. Cuizhi, G. & Robertson, K. in *Flora of China* Vol. 9 (eds Wu, Z. Y. et al.) 339–381 (Science Press and Missouri Botanical Garden Press, 2003).
102. Henker, H. in *Gustav Hegi—Illustrierte Flora von Mitteleuropa* Vol. IV/2 C (eds H. J. Conert et al.) 1–108 (Parey, 2000).
103. Lewis, W., Ertter, B. & Bruneau, A. in *Flora of North America North of Mexico* Vol. 9, 75–119 (Oxford Univ. Press, 2015).
104. Kron, P. & Husband, B. C. Using flow cytometry to estimate pollen DNA content: improved methodology and applications. *Ann. Bot.* **110**, 1067–1078 (2012).
105. Galbraith, D. W. et al. Rapid flow cytometric analysis of the cell cycle in intact plant tissues. *Science* **220**, 1049–1051 (1983).
106. Doyle, J. J. & Doyle, J. L. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochem. Bull.* **19**, 11–15 (1987).
107. Debray, K. et al. Identification and assessment of variable single-copy orthologous (SCO) nuclear loci for low-level phylogenomics: a case study in the genus *Rosa* (Rosaceae). *BMC Evol. Biol.* **19**, 152 (2019).
108. Deorowicz, S., Kokot, M., Grabowski, S. & Debudaj-Grabysz, A. KMC 2: fast and resource-frugal *k*-mer counting. *Bioinformatics* **31**, 1569–1576 (2015).
109. Kokot, M., Dlugosz, M. & Deorowicz, S. KMC 3: counting and manipulating *k*-mer statistics. *Bioinformatics* **33**, 2759–2761 (2017).
110. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014).
111. Kamneva, O. K. et al. Evaluating allopolyploid origins in strawberries (*Fragaria*) using haplotypes generated from target capture sequencing. *BMC Evol. Biol.* **17**, 180 (2017).
112. Danecek, P. et al. Twelve years of SAMtools and BCFtools. *Gigascience* https://doi.org/10.1093/gigascience/giab008 (2021).
113. Van der Auwera, G. & O'Connor, B. *Genomics in the Cloud: Using Docker, GATK, and WDL in Terra*. Edn 1 (O'Reilly Media, 2020).
114. Poplin, R. et al. Scaling accurate genetic variant discovery to tens of thousands of samples. Preprint at *bioRxiv* https://doi.org/10.1101/201178 (2017).
115. Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842 (2010).
116. Shen, W., Le, S., Li, Y. & Hu, F. Q. SeqKit: a cross-platform and ultrafast toolkit for FASTA/Q file manipulation. *PLoS ONE* **11**, e0163962 (2016).
117. Price, R. J. et al. Chromosome-scale genome sequence assemblies of the 'Autumn Bliss' and 'Malling Jewel' cultivars of the highly heterozygous red raspberry *Rubus idaeus* L.) derived from long-read Oxford Nanopore sequence data. *PLoS ONE* **8**, e0285756. (2023).
118. Edger, P. P. et al. Reply to: Revisiting the origin of octoploid strawberry. *Nat. Genet.* **52**, 5–7 (2020).
119. Kumar, S., Stecher, G., Li, M., Knyaz, C. & Tamura, K. MEGA X: molecular evolutionary genetics analysis across computing platforms. *Mol. Biol. Evol.* **35**, 1547–1549 (2018).
120. Sweeten, A. P., Schatz, M. C. & Phillippy, A. M. ModDotPlot—rapid and interactive visualization of tandem repeats. *Bioinformatics* **40**, btae493 (2024).
121. Stiehler, F. et al. Helixer: cross-species gene annotation of large eukaryotic genomes using deep learning. *Bioinformatics* **36**, 5291–5298 (2020).

122. Novák, P., Neumann, P. & Macas, J. Global analysis of repetitive DNA from unassembled sequence reads using RepeatExplorer2. *Nat. Protoc.* **15**, 3745–3776 (2020).
123. Novák, P., Hoštáková, N., Neumann, P. & Macas, J. DANTE and DANTE_LTR: lineage-centric annotation pipelines for long terminal repeat retrotransposons in plant genomes. *NAR Genom. Bioinform.* **6**, lqae113 (2024).
124. RepeatMasker Open v.4.0 (2013–2015).
125. Kim, D., Paggi, J. M., Park, C., Bennett, C. & Salzberg, S. L. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nat. Biotechnol.* **37**, 907–915 (2019).
126. Bolger, M., Schwacke, R. & Usadel, B. MapMan visualization of RNA-Seq data using Mercator4 functional annotations. *Methods Mol. Biol.* **2354**, 195–212 (2021).
127. Putri, G. H., Anders, S., Pyl, P. T., Pimanda, J. E. & Zanini, F. Analysing high-throughput sequencing data in Python with HTSeq 2.0. *Bioinformatics* **38**, 2943–2945 (2022).
128. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).
129. Reimer, J. J. & Turck, F. Genome-wide mapping of protein-DNA interaction by chromatin immunoprecipitation and DNA microarray hybridization (ChIP-chip). Part A: ChIP-chip molecular methods. *Methods Mol. Biol.* **631**, 139–160 (2010).
130. Ramirez, F. et al. deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res.* **44**, W160–W165 (2016).
131. Zhang, Y. et al. Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* **9**, R137 (2008).
132. Lopez-Delisle, L. et al. pyGenomeTracks: reproducible plots for multivariate genomic datasets. *Bioinformatics* **37**, 422–423 (2021).
133. Hadfield, J. D. MCMC methods for multi-response generalized linear mixed models: the MCMCglmm R package. *J. Stat. Softw.* **33**, 1–22 (2010).
134. Pedersen, T. L. patchwork: the composer of plots. R package version 1.3.0; https://doi.org/10.32614/CRAN.package.patchwork (2024).
135. Wickham, H., Vaughan, D., Girlich, M. & Ushey, K. tidyr: tidy messy data. R package version 1.3.1; https://doi.org/10.32614/CRAN.package.tidyr (2024).
136. Wickham, H., François, R., Henry, L., Müller, K. & Vaughan, D. dplyr: a grammar of data manipulation. R package version 1.1.4 (2023).
137. Sievers, F. et al. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* **7**, 539 (2011).
138. Hurel, A. et al. A cytological approach to studying meiotic recombination and chromosome dynamics in *Arabidopsis thaliana* male meiocytes in three dimensions. *Plant J.* **95**, 385–396 (2018).
139. Nascimento, T. & Pedrosa-Harand, A. High rates of structural rearrangements have shaped the chromosome evolution in dysploid beans. *Theor. Appl. Genet.* https://doi.org/10.1007/s00122-023-04462-3 (2023).
140. Herklotz, V. et al. Three haplotype-resolved pentaploid Rosa assemblies with assembled and extracted single copy orthologue (SCO) sequences from *Rosa canina* genome, diploid *Rosa* species, and sect. *Caninae* pollen. *Dryad* https://doi.org/10.5061/dryad.cc2fqz6fh (2025).

Acknowledgements The authors thank A. Houben, H. Bruelheide, B. Usadel, M. Thines, M. Bagdevi and K. Wesche for discussions on the research; B. Schlitt, M. Schwager, S. H. Sourav, J. Wesenberg, A. Smolka and S. Stegmann for their technical assistance in the laboratory; colleagues F. Erdogan, I. Kisakesen and U. Abraham for their help with target enrichment; T. Schell, C. Sinai, C. Greve, L. Schardt, D. Baranski and A. Ben Hamadou for technical support; J. Bauernfeind for help with IT facilities; T. Hawel, G. Schulz, S. Arndt, G. Vogg, S. Ruge and Oberlausitz-Stiftung for providing access to their collections and plant material; M. Simon for providing rose photos; the staff at the Darwin Tree of Life Project at the Wellcome Sanger Institute for making the data of the additional accession of *R. canina* and *R. agrestis* available (https://www.darwintreeoflife.org/project-resources); and the staff at the Institute of Biophysics AS CR for internal support. This study was funded by the German Research Foundation (DFG Ri 2090/4-1, MA 9363/2-1 and MA 9363/3-1), the Czech Science Foundation (GAČR 22-16826S), the Max Planck Society (core funding to A.M.), and the European Union (European Research Council Starting Grant, HoloRECOMB, grant no. 101114879 to A.M.). The DFG also funded this work under Germany's Excellence Strategy—EXC 493 2048/1–390686111 (to A.M.). M.Z. is financially supported by the DFG (grant no. MA 9363/2-1).

Author contributions A.M., A.K. and C.M.R. conceived the research. V.H. designed and performed the SCO experiments and processed the data. M.Z. and A.M. performed the genome assembly and genomics. D.H. performed the plastome assembly and annotation. B.H. performed all nuclear sequencing libraries. T.N., M.Z. and A.M. performed the ChIP–seq analysis. T.N. and J.L. performed the chromosome preparations, immunodetection and FISH. J.F. performed the flow cytometry measurements and the pollen-nucleus sorting. A.K., J.L., R.K. and A.M. performed the centromere and repeat characterization. U.P. performed the pollen viability assay. V.W. created and curated the synthetic hybrids. A.M., A.K. and C.M.R. supervised the project. V.H., M.Z., A.M. and C.M.R. wrote the original draft, and all of the authors discussed the results and contributed to writing the manuscript.

Funding Open access funding provided by Max Planck Society.

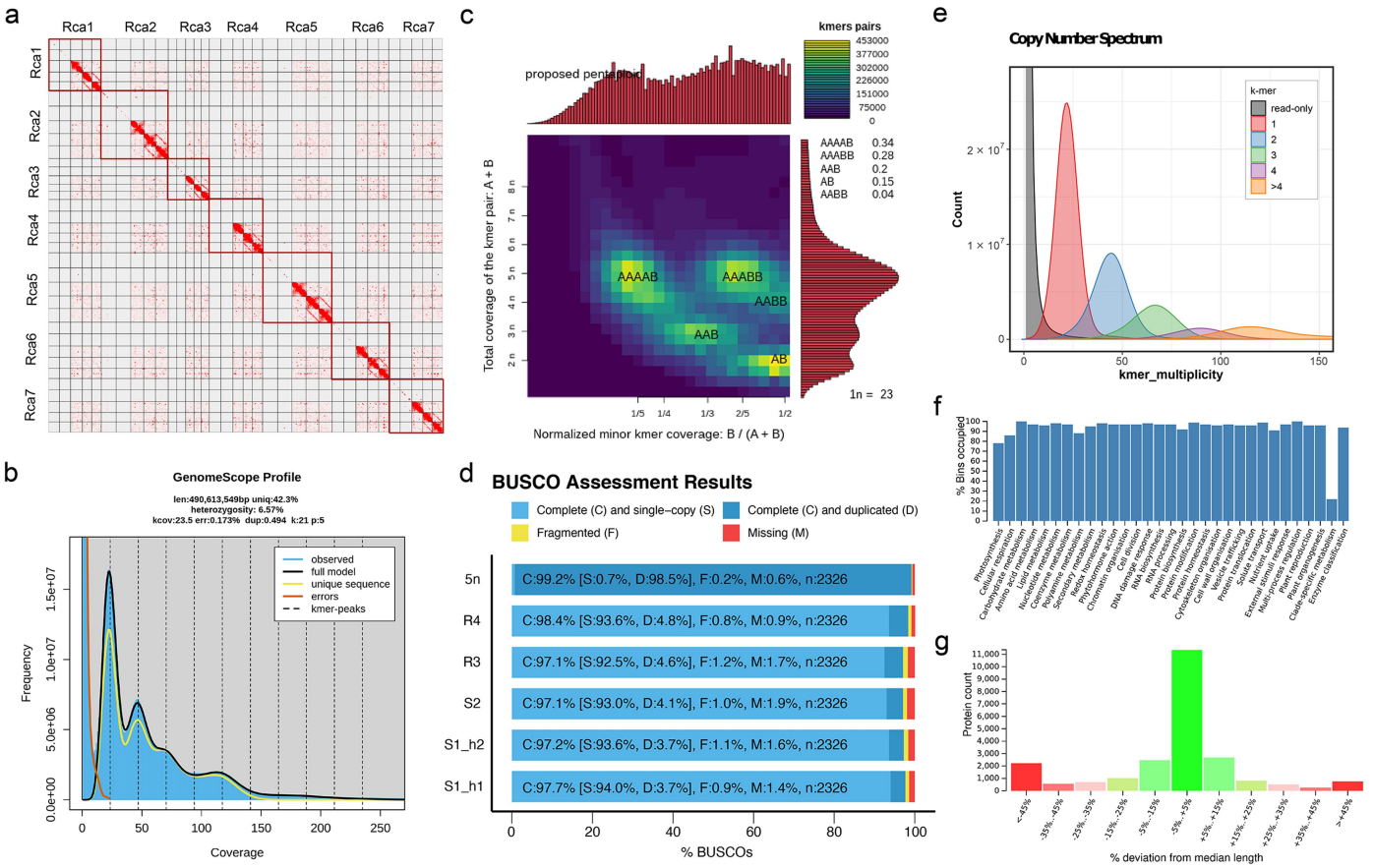
Competing interests The authors declare no competing interests.

Additional information

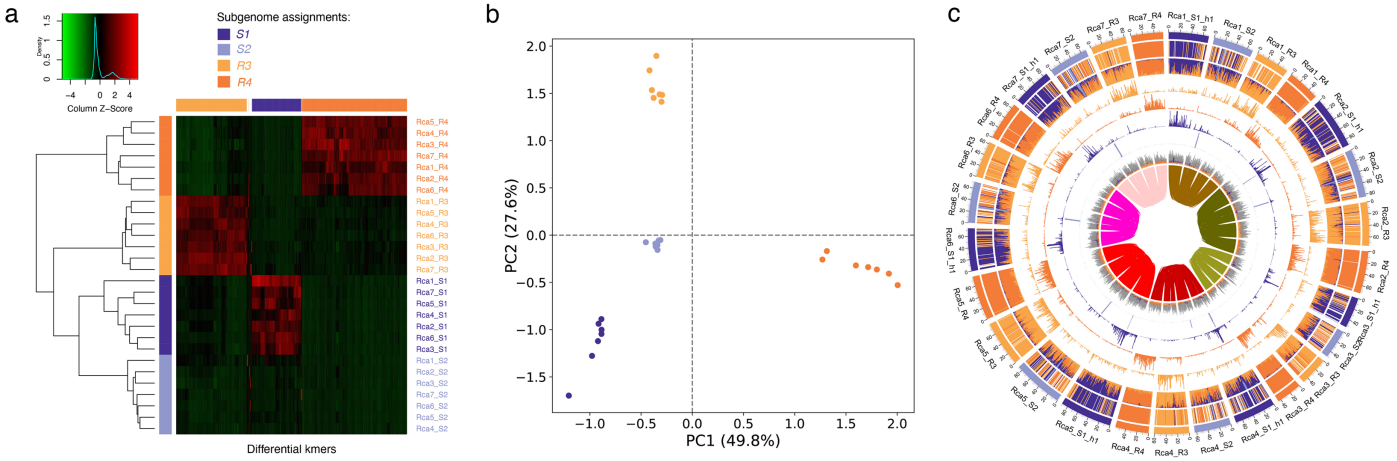
Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41586-025-09171-z.

Correspondence and requests for materials should be addressed to A. Kovařík, A. Marques or C. M. Ritz.

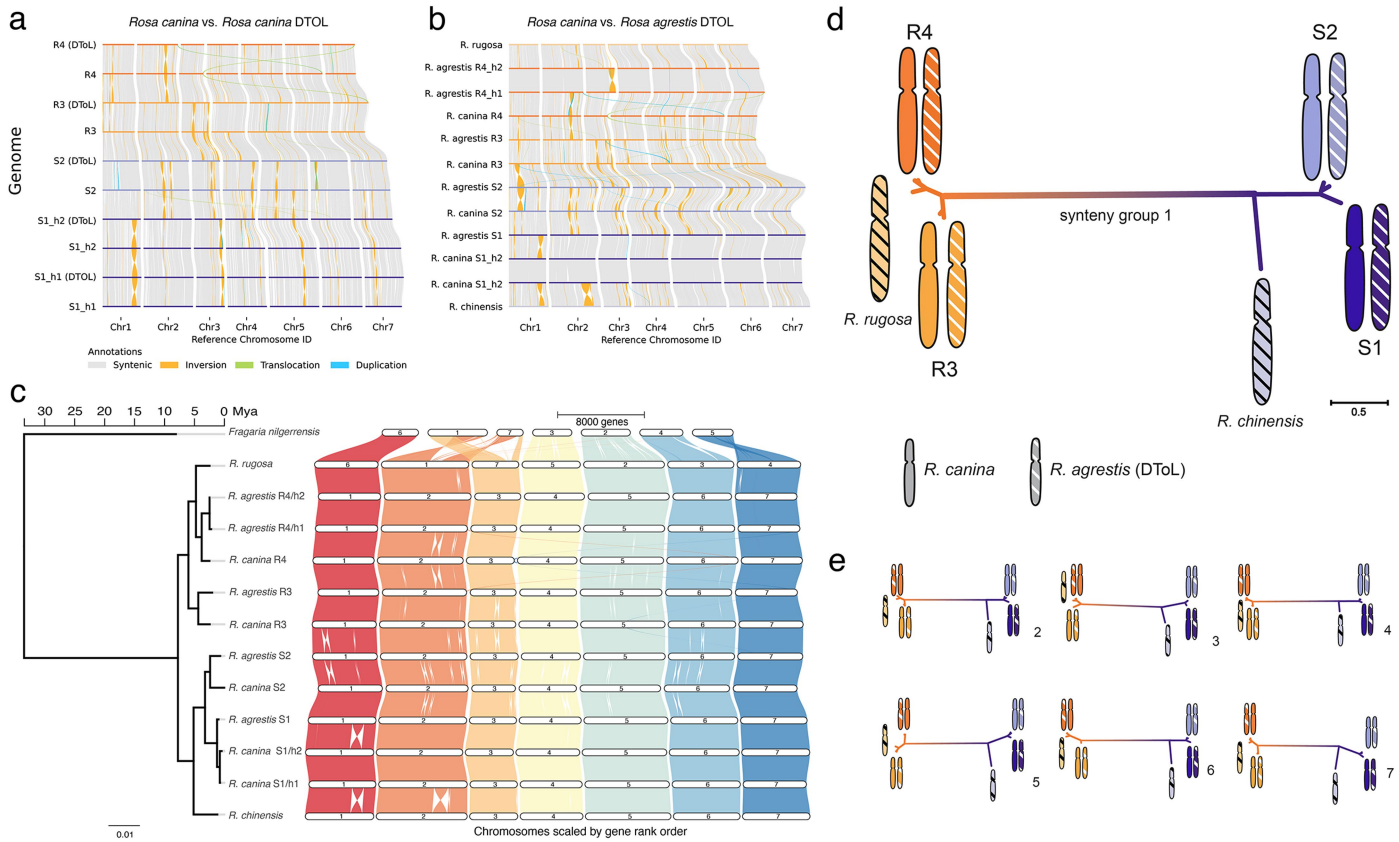
Peer review information *Nature* thanks Ian Henderson, Damon Lisch and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.



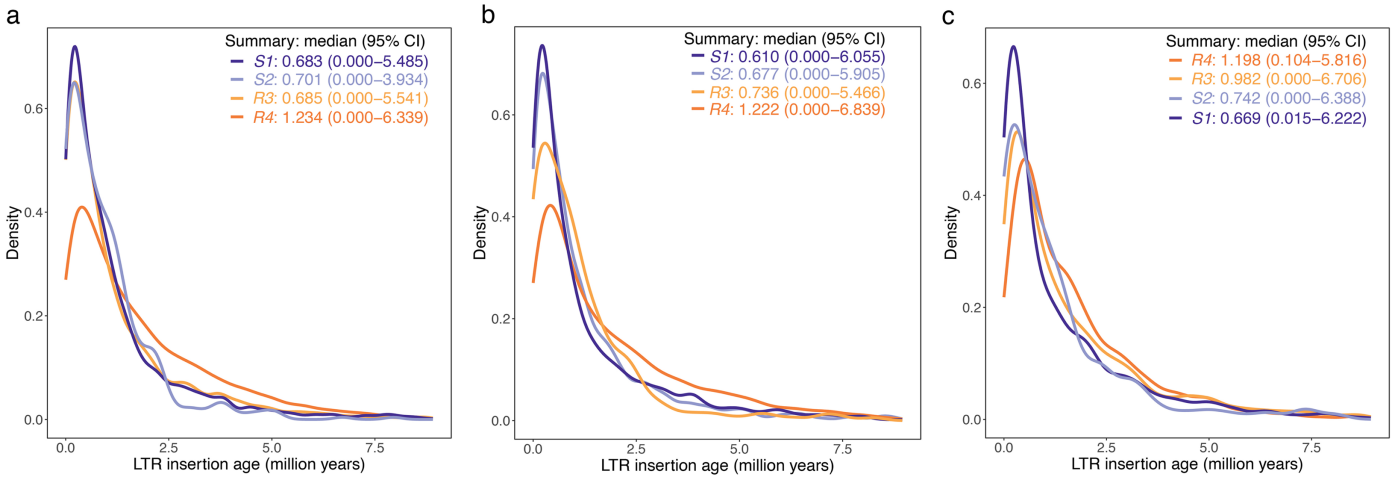
Extended Data Fig. 1 | Genome assembly assessment of *R. canina* (S27). (**a**) Hi-C map of 35 chromosome-level scaffolds. Each synteny group contains five chromosomes, which were ordered as S1_h1, S1_h2, S2, R3, and R4 in sequence. The vanished Hi-C signals of the first two chromosomes in each linkage group suggest the high homozygosity of S1_h1 and S1_h2 haplotypes. (**b**) GenomeScope2 *21*-mer distribution from HiFi sequences and genome size estimation confirmed the expected *R. canina* genome size, i.e., 5× 490 Mbp. 21-mer coverage is estimated as 23.5× coverage, consistent with read coverage (23×). The 4*x* peak at coverage 94 is less invisible, indicating two out of five haplotypes are highly similar. (**c**) Ploidy and genome structure inference based on *21*-mer Smudgeplot analysis. A and B represent the number of heterozygous *k*-mers identified. Note the high amount of pentaploid heterozygous *k*-mer combinations AAAAB and AABBB. (**d**) BUSCO (Universal Single-Copy Orthologues) genes assessment results based on the annotated protein sequences of the pentaploid genome and each individual subgenome set. (**e**) HiFi *k*-mer multiplicity frequency by Merqury. The number in the plot legend and different colours indicate the copy numbers found in the assembly. (**f**) and (**g**) are Mercator4 statistics of gene contents based on its functional annotation on *Rosa canina* protein sequences. (**f**) Percentage of genes occupying a certain gene category. (**g**) Protein length distribution based on the deviation to category-specific reference lengths. Each bar represents the number of proteins having a certain length difference to the reference length of the corresponding Mercator4 category. The overall small deviation of gene lengths in *R. canina* comparing to Mercator reference genes indicate most annotated genes in *R. canina* are complete. (**f**) and (**g**) show the results for the subgenome *S1* haplotype h1 only since the other three subgenomes show highly similar distribution.



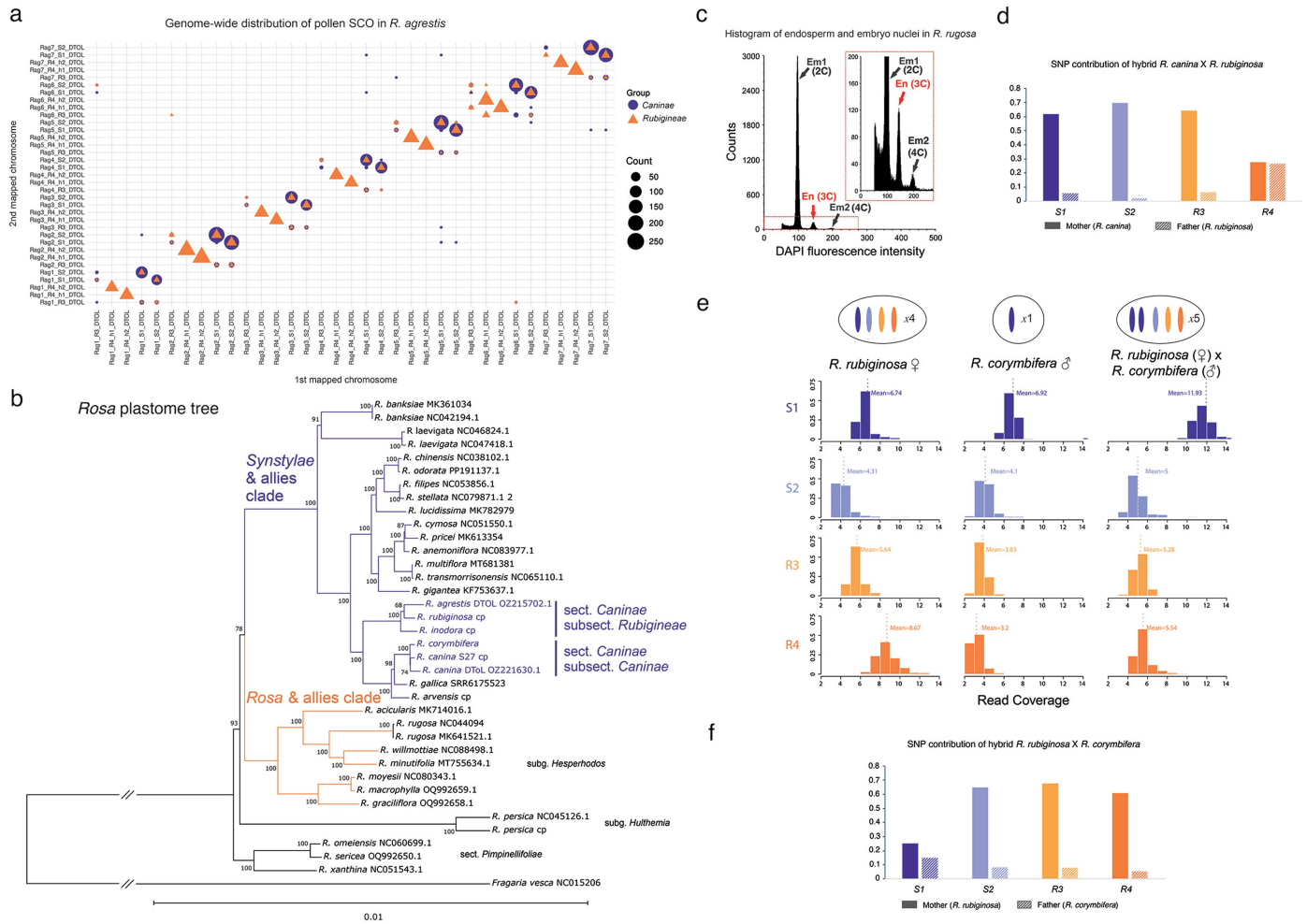
Extended Data Fig. 2 | *K*-mer-based subgenome phasing and characterization of the pentaploid *R. canina* genome. (**a**) Unsupervised hierarchical clustering (the horizontal colour bar at the top of the axis indicates to which subgenome the *k*-mer is specific; the vertical colour bar on the left of the axis indicates the subgenome to which the chromosome is assigned). The heatmap indicates the *Z*-scale relative abundance of *k*-mers. The larger the *Z* score is, the greater the relative abundance of a *k*-mer), and (**b**) Principal Component Analysis (PCA) of differential 15-mers confirmed that the genome was successfully phased into four subgenomes based on clearly distinct patterns of both differential *k*-mers and homoeologous chromosomes. (**c**) Chromosomal characteristics. From the outer to inner circles (1–9): (1) subgenome assignments based on the *k*-means algorithm; (2) significant enrichment of subgenome-specific *k*-mers – the same colour as the subgenome indicates significant enrichment for those subgenome-specific *k*-mers – white areas are not significantly enriched; (3) normalized proportion (relative) of subgenome-specific *k*-mers; (4–7) count (absolute) of each subgenome-specific *k*-mer set; (8) density of long terminal repeat retrotransposons (LTR-RTs) – if the colour is consistent with the subgenome, it indicates that LTR-RTs are significantly enriched in those subgenome-specific *k*-mers; grey indicates nonspecific LTR-RTs; and (9) homoeologous blocks. All statistics (2–7) are computed in sliding windows of 1 Mb. Notably, the fraction of subgenome differentiation between the S1 and S2 subgenomes was restricted to proximal regions, likely representing their different centromere compositions. Whereas R3 and R4 subgenomes are clearly distinct from each other in their specific *k*-mer spectrum.



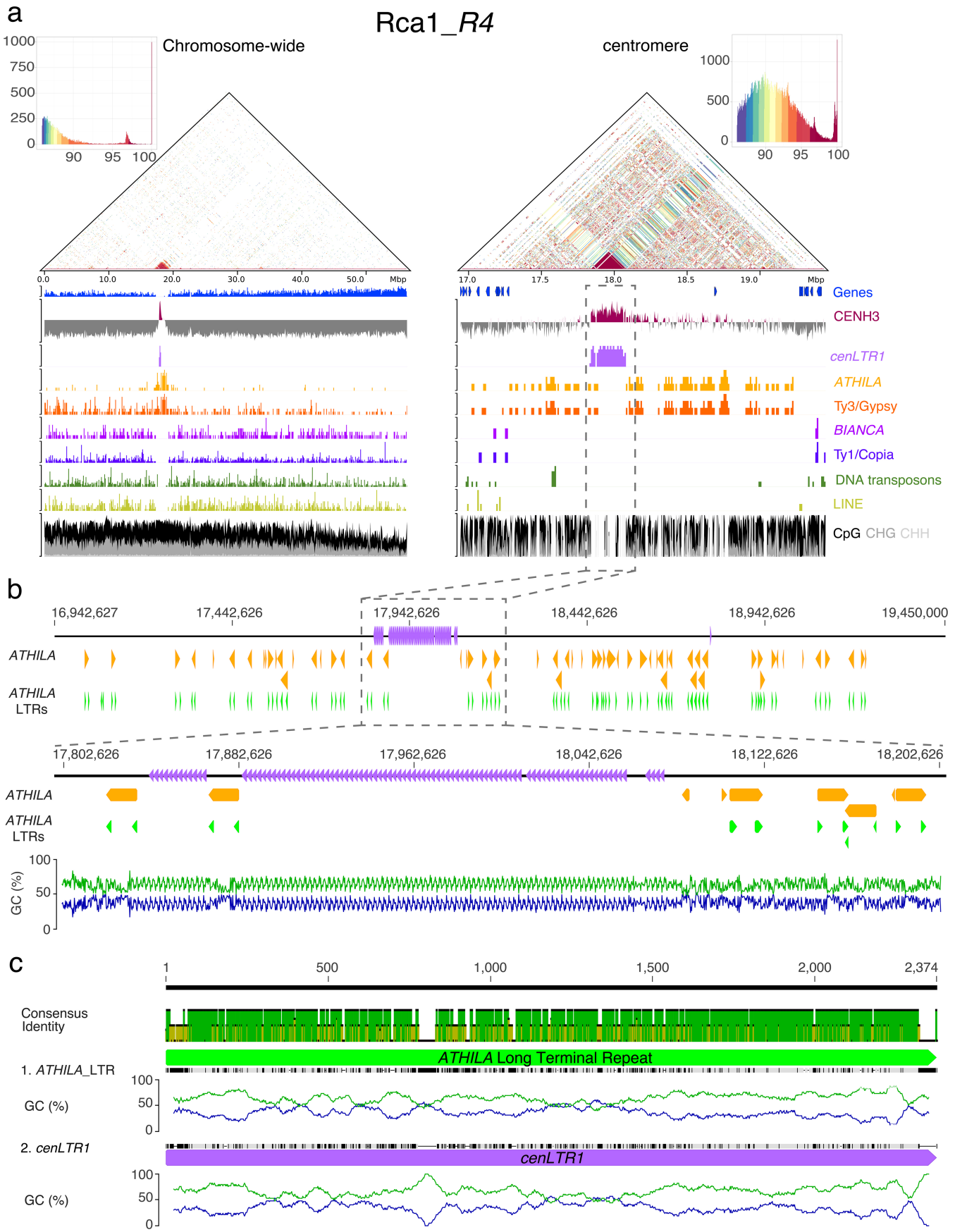
Extended Data Fig. 3 | Comparative synteny analysis of the assembled *R. canina* (S27) with the *R. canina* DToL and *R. agrestis* DToL. (**a**) Synteny and rearrangement analyses (SyRI) between our and the *R. canina* genome assembly from DToL. (b) Synteny and rearrangement analyses (SyRI) between our *R. canina* and the *R. agrestis* genome assembly from DToL. Pairwise comparisons of synteny of all *R. canina* subgenomes (S1_h1/S1_h2, S2, R3, and R4) and all *R. agrestis* subgenomes (S1, S2, R3, and R4_h1/R4_h2) juxtaposed against the corresponding chromosomes of *R. chinensis* (sect. *Synstylae*) and *R. rugosa* (sect. *Rosa*). Note that only synteny blocks and rearrangement blocks greater than 50 kb in length are shown here. (**c**) GENESPACE synteny and phylogenetic relationships of the five chromosome sets of our *R. canina* assembly, *R. agrestis* from DToL and their close relatives *R. chinensis* and *R. rugosa*. Chromosomes are normalized by number of genes. (**d**–**e**) Unrooted Maximum Likelihood phylogenies of the homoeologous *R. canina* and *R. agrestis* chromosomes and chromosomes from the respective syntenic groups of the diploid species *R. chinensis* and *R. rugosa* based on multiple alignments of whole chromosome sequences. The upper panel (**d**) exemplarily depicts the phylogeny of chromosomes of synteny group 1. Synteny groups 2 – 7 are presented in (**e**). Filled chromosomes refer to subgenomes of *R. canina* belonging to the *Synstylae* clade (violet/light blue) and *Rosa* clade (dark/light orange). Chromosomes from *R. agrestis* are marked with hatched white filling, while the diploid roses are marked with black hatched filling.



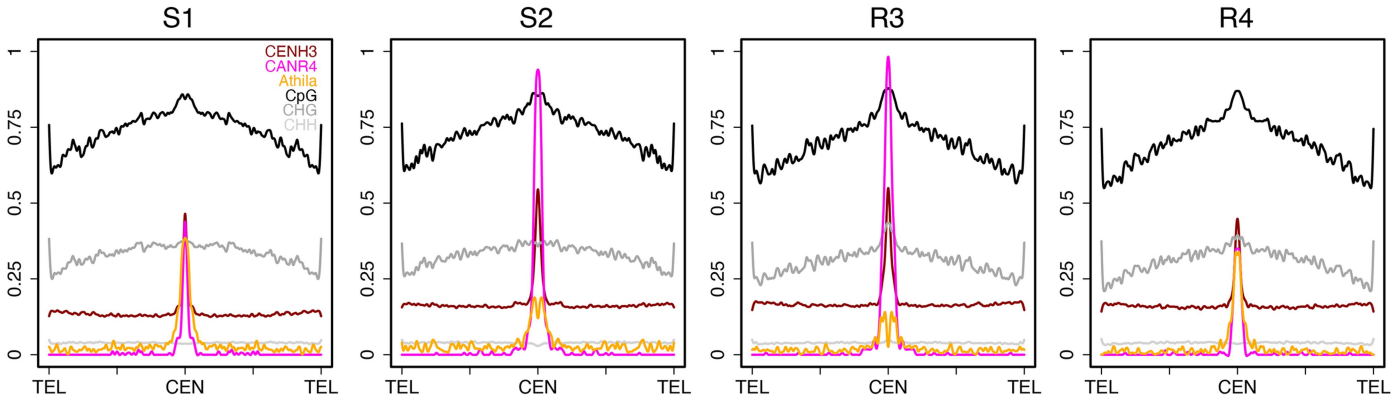
Extended Data Fig. 4 | Dating subgenome-specific LTR-RTs insertion times in dogrose genomes. The 95% confidence intervals (CIs) are marked in the upper right corner to predict the insertion time boundaries of LTR-RTs on the subgenome for *R. canina* (**a**), *R. canina* DToL (**b**) and *R. agrestis* DToL (**c**). The colours of the subgenomes are consistent throughout the Fig. panels.



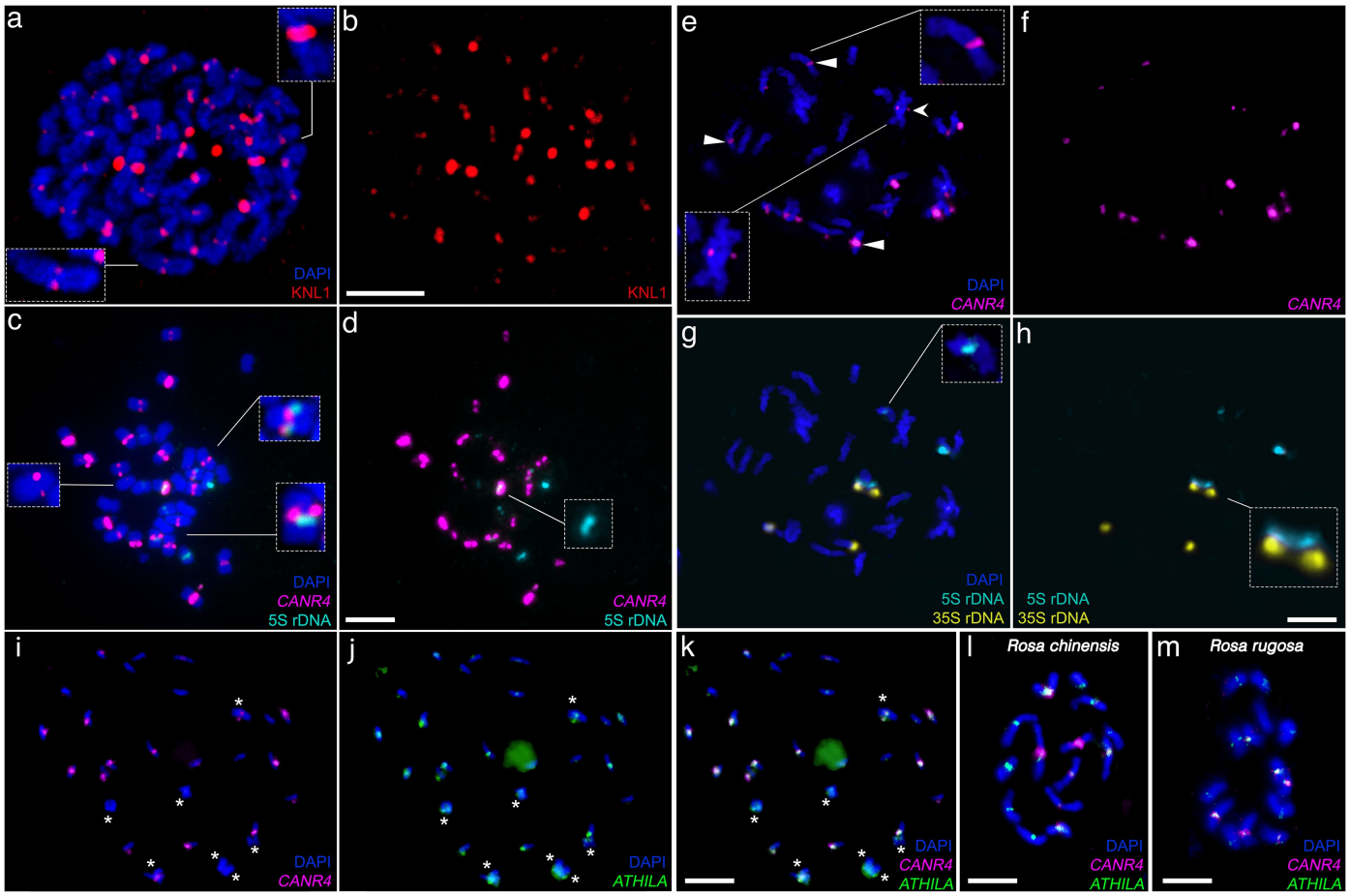
Extended Data Fig. 5 | Experimental validation of the reproduction mode of dogroses. (**a**) Genome-wide distribution of pollen SCOs from pollen of eigth dogrose species (subsect. *Caninae*: three samples of *R. canina*, two samples of *R. corymbifera*; subsect. *Rubigineae* three samples of *R. rubiginosa*) from two subsections in the *R. agrestis* genome assembly from DToL. The bubble map represents chromosomal hits from the SAM file output, which were selectively filtered to display loci with a single alternative hit. The size of the symbols corresponds to the mean count of pollen SCOs mapped to each chromosomal pair, identifying seven bivalent chromosome pairs within the *R. agrestis* (subsect. *Rubigineae*) genome assembly and seven different pairs in which *R. rubiginosa* (subsect. *Rubigineae*) pollen SCO mapped. (**b**) Plastome phylogeny of the genus *Rosa*. Maximum Likelihood phylogeny of the genus *Rosa* retrieved from plastome sequences obtained in this study and those available in Genbank (Supplementary Table 4). All nodes which were not supported by 100% bootstrap are indicated by dashed lines. (**c**) Representative histogram of flow cytometric measurements of nuclei isolated from nutlets of the sexual diploid *R. rugosa* to determine the endosperm/embryo ratios. The observed endosperm/embryo ratio 1.5, fits to the expectation of 3*x* endosperm and a 2*x* embryo originated by sexual reproduction including endosperm fertilization. Please note that the gain of the fluorescence was adjusted to position the embryo peak at channel 100. For a better visualization of the endosperm peak the scale of the Y-axis was manually adjusted to 200 counts (red rectangles). Em1: embryo G0/G1, Em2: embryo G2, En: endosperm G0/G1. (**d**) The parental SNP ratios detected in each subgenome of the hybrid *R. canina* × *R. rubiginosa* (Sample ID: D62b_2). SNPs were called by aligning short reads to our assembled *R. canina* (S27) *S1_h1*. The y-axis displays the proportion of SNPs in each subgenome. For instance, ~62% SNPs detected in the alignments to S1 subgenome are the same as the maternal *R. canina* (Sample ID: D3b_2), while only around 5% SNPs are from the paternal *R. rubiginosa* (Sample ID: D145_b2), indicating the S1 subgenome of the hybrid was supposed to be contributed by the maternal parent. Overall, the hybrid subgenome S1, S2, and R3 are all from maternal parent *R. canina*, while one R4 subgenome is from maternal *R. canina* and the other R4 is contributed by paternal *R. rubiginosa*. So, the subgenome composition of the hybrid should be S1, S2, R3, R4_h1/h2. (**e**) Assessment of the parental genomes contribution of a synthetic hybrid between *R. rubiginosa* (female donor) and *R. corymbifera* (male donor), confirming the sexual reproduction and subgenome's inheritance trough male and female meiosis. Note, that *R. corymbifera* belongs to subsect. *Caninae* and is a very close relative of *R. canina* and so it has two copies of S1 subgenome. (**f**) The parental SNP ratios detected in each subgenome of the hybrid *R. rubiginosa* × *R. corymbifera* (Sample ID: D166b_2). The plot interpretation is the same as (**d**).



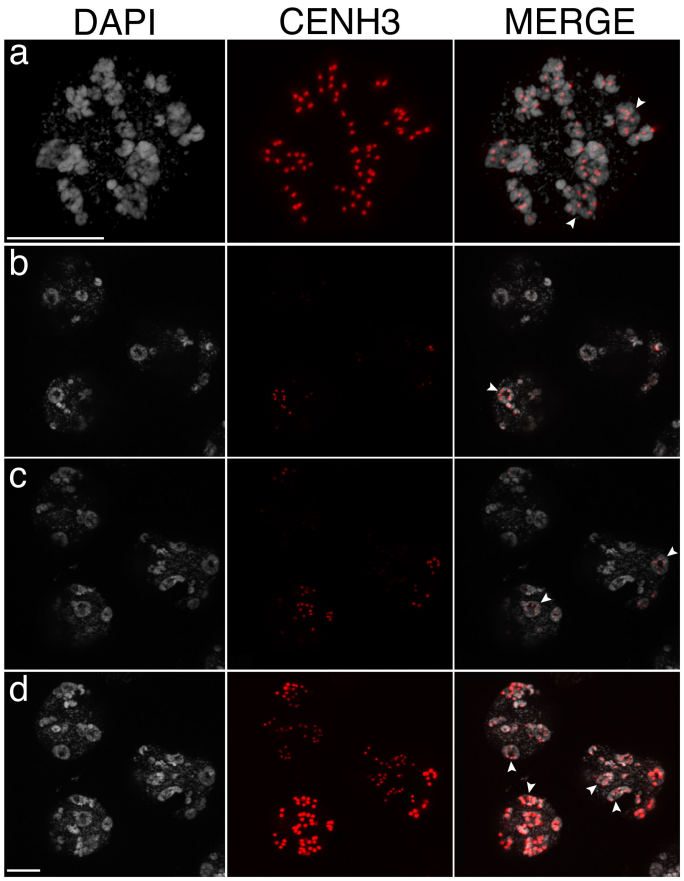
Extended Data Fig. 6 | Characterization of the centromeric LTR 1 (*cenLTR1*) found in the centromere of Rca1_*R4* of *R. canina* (S27) genome. (**a**) Structural visualization of the chromosome-wide (left) and centromere (right) close up. Please note the specificity of the CENH3 ChIP-seq signal only on the *cenLTR* sequence. (**b**) Close up visualization of the *cenLTR* array and insertion of an *ATHILA* inside the *cenLTR* array. (**c**) Alignment of the *cenLTR* consensus sequence with the LTR sequence from the inserted *ATHILA* element showing over 75% sequence similarity.



Extended Data Fig. 7 | Genetic and epigenetic centromere variation in *R. canina* (S27). Metaplot of CENH3 enrichment, DNA methylation, and centromeric elements—*ATHILA* and *CANR4* density—per subgenome. CENH3 enrichment marked in dark red was calculated by CENH3 ChIP-seq (log2 CENH3/H3) signal normalized by coverage. DNA methylations have three contexts with CpG marked in black, CHG maked in dark grey, and CHH marked in light grey. *ATHILA* (yolk yellow) and *CANR4* (magenta) were presented by their density. All signals were calculated in 50 kb adjacent windows. All chromosome coordinates were scaled based on their distance to centromere against the distance of centromere (CEN) to telomere (TEL). Centromere position (CEN) was defined by where maximum CENH3 enrichment was located. All signal values (y-axis) were scaled from 0 to 1 by global minimum to global maximum except for DNA methylations, which retained the original percentage values. The p- and q-arm values were averaged and mirrored. Please note the higher CENH3 association with *CANR4*-based centromeres compared with *ATHILA*-based centromeres.



Extended Data Fig. 8 | Detection of the kinetochore protein KNL1, *CANR4* satellite repeat, *ATHILA* and rDNA in the chromosomes of *R. canina*. (**a**–**b**) Immunodetection of the KNL1 protein (in red) at all centromeres of *R. canina* revealed a differential distribution, with the presence of large and small centromeres, as shown in more detail in the upper and lower insets (**a**). (**c**–**d**) Distribution of the centromeric satellite repeats *CANR4* (in magenta) and 5S rDNA (in cyan) in the *R. canina* miotic chromosomes, while the *CANR4* repeat is present at the centromeres of a set of chromosomes. The 5S rDNA is located at the pericentromeric regions of chr1, insets in **c** and **d** highlight the pericentromeric location of 5S loci (right) and a weak *CANR4* signal (left). (**e**–**h**) Distribution of the *CANR4* repeat in meiotic diakinesis, revealing its distribution mostly in univalents according (**e** and **f**), while the pattern of 5S and 35S rDNA (**g** and **h**) (cyan and yellow, respectively) hybridized to three sites, two of which were located near 35S sites on a bivalent (inset in **h**) corresponding to the Rca3_S1_h1/h2 chromosomes as found in our assembly annotation. Week signals are highlighted in the insets present in **g** and **h**. (**i**–**m**) FISH with *CANR4* and *ATHILA* on a diakinesis of *R. canina* (**i**–**k**) and mitotic cells of *R. chinensis* (**l**) and *R. rugosa* (**m**). Asterisks in **i**–**k** point to bivalent-forming chromosomes. The experiments for this set of data were repeated at least 10 times and independently presenting similar results. Scale bar = 5 µm.



Extended Data Fig. 9 | CENH3 immunostaining in polyads that are formed at the end of male meiosis in *R. canina* (S27). (**a**) A single polyad showing at least two nuclei with seven CENH3 signals (arrowheads), while several other nuclei show irregular number of centromeric foci. (**b**–**d**) Different Z-stacks of three polyads showing few nuclei with seven centromeric foci, while several other smaller nuclei are formed with irregular number of centromeres. Immunostaining experiments were repeated independently at least ten times outputting similar results. Scale bars = 10 µm.

nature portfolio

Corresponding author(s): André Marques

Last updated by author(s): May 5, 2025

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed
☐	☒ The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒ A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒ The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐ A description of all covariates tested
☐	☒ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒ A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☒	☐ For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐ Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	HiFi and Omni-C reads were obtained through own sequencing at the Max Planck Institute for Plant Breeding Research, Cologne, Germany. Images were analyzed using the ZEN software (Carl Zeiss GmbH) and the ZENBlack software (Carl Zeiss GmbH).
Data analysis	Available open source tools used in this study were: Bedtools (v2.29.0) bcftools 1.15.1 Bismark (v0.23.0) BLAST 2.13.0+ Bowtie2 (2.5.4) BUSCO (v5.1.2) BWA (0.7.17) CoGe (v7) Cutadapt (v4.7) DANTE_LTR (v0.3.5.2) Deeptools (v3.5.1) DESeq2 (1.46.0) Dotter (v0.13.1) EMBOSS (v2024.0419.155605) findGSE_v1.94.R GATK v4.1.9.0

GenomeScope2
Geneious (v2023.0.1)
GENESPACE (v1.3.1)
Helixer (0.3.4)
Hisat (2.2.1)
Hifiasm (0.19.8-r603)
htseq-count (v2.0.1)
IQ-TREE (2.4.0)
Jellyfish (v2.3.1)
Juicer (v1.6)
Kaks Calculator (v3)
MACS3 (3.0.1)
MAFFT (7.526)
minimap2 (v2.26)
ModDotPlot (v0.8.2)
PAML v4.10.6
plotsr (v0.5.3)
pyGenomeTracks (v3.8)
QUAST (v5.2.0)
RepeatExplorer2 (v2.3.7)
REXdb (v1.0)
SALSA2 (v2.3)
Samtools (v1.9)
StainedGlass (v0.6)
syri (v1.5.3)
VCFtools (0.1.16)
ZEN blue (3.1)

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

All raw sequencing data (HiFi, Hi-C, RNA, CENH3-ChIP, DNA methylation, SCO of pollens) and genome assembly of Rosa canina S27 isolate are available under NCBI BioProject: PRJNA1111045. The chloroplast genome of Rosa canina S27 isolate available under GeneBank accession number PV550499. Raw sequencing data of Rosa canina DToL and Rosa agrestis DToL are available from Darwin Tree of Life (DToL) data portal. The corresponding NCBI BioProject accession numbers are PRJEB79802 and PRJEB79880, respectively. Genome assemblies, the sample-specific SCO reference sequences, variant calling format files, annotations and alignments presented in this work are also made available for download at DRYAD: https://doi.org/10.5061/dryad.cc2fqz6fh. The REXdb database Viridiplantae v.3.0 [http://repeatexplorer.org/?page_id=918] is publicly available. All other data needed to evaluate the conclusions in the paper are provided in the paper and/or the supplemental information.

Research involving human participants, their data, or biological material

Policy information about studies with human participants or human data. See also policy information about sex, gender (identity/presentation), and sexual orientation and race, ethnicity and racism.

Reporting on sex and gender	n/a
Reporting on race, ethnicity, or other socially relevant groupings	n/a
Population characteristics	n/a
Recruitment	n/a
Ethics oversight	n/a

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	Sample-size calculation was performed based on assessment of the literature in the field, our own experience from previous studies and requirement for corresponding protocols. For Immunocytochemistry and in situ hybridisation analyes sample size was based on the number of cells obtained. The size of the sample used was performed according to the requirements for each protocol. For cytological analysis, different roots and anthers were collected and analysed to confirm the reproducibility of the results. For sequencing, sufficient coverage (>20x) was used to assemble and scaffold the R. canina genome. The sample size used for all experiments provided sufficient resolving power.
Data exclusions	No data was excluded form the analysis.
Replication	Cytogenetic analyses were performed on several cells, using the best superposition for the final figure. Experiments were independently repeated at least ten times with similar results, in order to track all meiotic stages.
Randomization	A randomization is not relevant for this study because no genotype or treatment were compared with each other. However, the tissues for cytogenetic and ChIPseq experiments were randomly collected from different plant individuals grown under the same condition in a greenhouse.
Blinding	The experiments were performed without knowing the final results.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☐	☒ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☒	☐ Clinical data
☒	☐ Dual use research of concern
☐	☒ Plants

Methods

n/a	Involved in the study
☐	☒ ChIP-seq
☐	☒ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used	Customized Rosa canina-specific antibodies generated in this study: rabbit anti-CENH3 (AB016310, LifeTein, generated in this study; dilution 1:500) Commercially available antibodies: mouse anti-alpha Tubulin (Sigma-Aldrich, St. Louis, MO; catalogue number T6199, Clone: MABI 0301, dilution 1:200) rabbit anti-Histone H3 (Active Motif, cat. No. 39064, ChIP only, 1µg used) Previously designed antibodies: rabbit anti-KNL1 (AB015677-3; GenScript, Piscataway, NJ, USA; Oliveira et al. 2024, dilution 1:500)
Validation	Newly validated antibodies: anti-Rosa canina CENH3 antibody was generated by the company LifeTein and validated by peptide ELISA tests and ChIP experiments. ELISA test information is available upon request. Furthermore, the observed indirect immuno-signals of anti-CENH3 on R. canina cells are compatible with centromere data previously reported in the published literature for other species. Previously validated antibodies: KNL1 was previously validated by Oliviera et al. (2024) by Immunostaining and Western blot. Validation by commercial providers: mouse anti-alpha Tubulin (validation; https://www.sigmaaldrich.com/DE/de/product/sigma/t6199?srsltid=AfmBOopZVI8rakd6EyMi2t1B9KcuqVPwcg6UFkhi3RDxHotJnOjCzZQW)

rabbit anti-Histone H3 (validation: https://www.activemotif.com/documents/tds/39763.pdf; https://www.activemotif.com/catalog/details/39763)

Plants

Seed stocks: Rosa canina plants were cultivated under controlled greenhouse conditions (16h daylight, 26 °C, >70% humidity).

Novel plant genotypes: n/a

Authentication: n/a

ChIP-seq

Data deposition

☒ Confirm that both raw and final processed data have been deposited in a public database such as GEO.

☒ Confirm that you have deposited or provided access to graph files (e.g. BED files) for the called peaks.

Data access links
May remain private before publication.

https://www.ncbi.nlm.nih.gov/sra/?term=SRR32424402

Files in database submission

6919_A_run867_AGGTTCCT_S30_L001_R1_001.fastq.gz
6919_A_run867_AGGTTCCT_S30_L001_R2_001.fastq.gz
6919_A_run867_AGGTTCCT_S30_L002_R1_001.fastq.gz
6919_A_run867_AGGTTCCT_S30_L002_R2_001.fastq.gz
6919_B_run867_GAACCTTC_S31_L001_R1_001.fastq.gz
6919_B_run867_GAACCTTC_S31_L001_R2_001.fastq.gz
6919_B_run867_GAACCTTC_S31_L002_R1_001.fastq.gz
6919_B_run867_GAACCTTC_S31_L002_R2_001.fastq.gz
6919_C_run867_AAGTCCTC_S32_L001_R1_001.fastq.gz
6919_C_run867_AAGTCCTC_S32_L001_R2_001.fastq.gz
6919_C_run867_AAGTCCTC_S32_L002_R1_001.fastq.gz
6919_C_run867_AAGTCCTC_S32_L002_R2_001.fastq.gz
6919_D_run867_CCACAACA_S33_L001_R1_001.fastq.gz
6919_D_run867_CCACAACA_S33_L001_R2_001.fastq.gz
6919_D_run867_CCACAACA_S33_L002_R1_001.fastq.gz
6919_D_run867_CCACAACA_S33_L002_R2_001.fastq.gz
6919_E_run867_ATAACGCC_S34_L001_R1_001.fastq.gz
6919_E_run867_ATAACGCC_S34_L001_R2_001.fastq.gz
6919_E_run867_ATAACGCC_S34_L002_R1_001.fastq.gz
6919_E_run867_ATAACGCC_S34_L002_R2_001.fastq.gz
6919_F_run867_CCGGAATA_S35_L001_R1_001.fastq.gz
6919_F_run867_CCGGAATA_S35_L001_R2_001.fastq.gz
6919_F_run867_CCGGAATA_S35_L002_R1_001.fastq.gz
6919_F_run867_CCGGAATA_S35_L002_R2_001.fastq.gz

Genome browser session
(e.g. UCSC)

no longer applicable

Methodology

Replicates: Two biological replicates of CENH3 sequenced as 6919_A and 6919_B; Two biological replicates of H3 sequenced as 6919_C and 6919_D; Two biological replicates of input control sequenced as 6919_E and 6919_F.

Sequencing depth:
6919_A: 21,902,101 pairs of reads (150bpx2); 30,512,245 reads uniquely mapped
6919_B: 21,890,693 pairs of reads (150bpx2); 31,961,379 reads uniquely mapped
6919_C: 21,798,217 pairs of reads (150bpx2); 34,691,504 reads uniquely mapped
6919_D: 21,798,366 pairs of reads (150bpx2); 34,835,465 reads uniquely mapped
6919_E: 21,304,165 pairs of reads (150bpx2); 33,647,990 reads uniquely mapped
6919_F: 21,682,527 pairs of reads (150bpx2); 34,367,915 reads uniquely mapped

Antibodies: The CENH3 gene of Rosa canina was identified using BLASTP with the parameter "-evalue 1e-5 -qcov_hsp_perc 50" and the A. thaliana CENH3 protein HTR12 (AT1G01370) was used as the reference. A specific polyclonal antibody against its CENH3 protein (ARVKHTAARKDRIKTARRQP-C / AB016310) was designed and synthetised by LifeTein with immunisation in rabbits.

Peak calling parameters: Genome indexing and read mapping were both done by bowtie2 (v2.5.4) with `--sensitive-local` flag activated for all experiments.

Peak calling parameters	After comparing the CENH3 domains with H3 or input samples woth bamCompare, peak calling was performed using the MACS3 pipeline with the inclusion of parameters --broad -g 1.9e+9.
Data quality	Mapped reads were not filtered by mapping quality considering the high similarity of four subgenomes. 251 out of 6105 broad peaks were 5 fold enriched when setting FDR value as 0.05.
Software	The raw 150bp pair-end ChIP -seq reads were checked by FastQC and then mapped to the R. canina haplotype phased reference genome using bowtie2 (as described in "Peak calling parameters"). The BAM file was converted to bigwig using the bamCompare tool from deeptools2, and then normalized to RPKM (reads per kilobase of transcript per million reads mapped). After this a peak calling, comparing the CENH3 domains with H3 or input samples, was performed using the MACS3 pipeline with the inclusion of parameters --broad -g 1.9e+9.

Flow Cytometry

Plots

Confirm that:

☒ The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).

☒ The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).

☒ All plots are contour plots with outliers or pseudocolor plots.

☒ A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation	Nuclei of mature pollen grains were isolated by applying the filter bursting method63 using the nuclei isolation buffer according to Galbraith et al. (1983)64. Pollen grains were burst on the surface of a 20 µm disposable CellTrics filter (Sysmex-Partec). The resulting nuclei suspension was stained with propidium iodide (50 µg/ml, PI) and run on a BD Influx cell sorter (BD Biosciences). After identifying the nuclear populations in a dotplot displaying the PI fluorescence signal (log-scale) versus side scatter signal (SSC, log-scale) a sort gate was defined in the corresponding fluorescence intensity (lin-scale) histogram. Per individual 200,000 generative nuclei (volume ca. 400 µL) were collected into a 1.5 ml reaction tube using the '1.0 Drop Pure' sort mode of the BD FACS Software (BD Biosciences). After adding 50 µL 1× TE and 50 µL NaN3 nuclei were sedimented by centrifugation (1000 × g for 10 min at 4°C). Afterwards 300 µL of the supernatant was removed and the nuclei with the remaining liquid stored at -20°C
Instrument	BD Influx cell sorter
Software	BD FACSDiva™ Software (v9.0)
Cell population abundance	Per individual 200,000 generative nuclei (volume ca. 400 µL) were collected into a 1.5 ml reaction tube using the '1.0 Drop Pure' sort mode of the BD FACS Software (BD Biosciences).
Gating strategy	After identifying the nuclear populations in a dotplot displaying the PI fluorescence signal (log-scale) versus side scatter signal (SSC, log-scale) a sort gate was defined in the corresponding fluorescence intensity (lin-scale) histogram.

☒ Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.

3. General discussion

Meiotic recombination is a key evolutionary process that ensures accurate chromosome segregation and generates genetic diversity, but its regulation remains remarkably diverse across species and chromosome architectures. In this thesis, I explored how chromosome structure and synapsis dynamics shape crossover patterning in the holocentric *Rhynchospora* genus. Unlike the classical monocentric model, where centromere localisation constrains chiasma formation, holocentric species offer an opportunity to dissect recombination control in the absence of a single primary centromere. The first project provided the first high-resolution characterisation of CO landscape and meiotic dynamics in *R. breviuscula*, revealing a striking distal-biased CO landscape despite diffused holocentromeres, and this biased landscape is associated with telomere-initiated synapsis. Building on this foundation, the second project expanded the analyses across multiple *Rhynchospora* species that cover all main clades and demonstrated that the distal bias of CO distribution occurs only when synapsis initiates at chromosome ends. In contrast, species with random synapsis initiation exhibit uniform CO distributions. This work also linked recombination rate to chromosome architecture, where we found that smaller chromosomes display higher recombination rates, more double COs, smaller chromatin loops, and proportionally longer chromosome axes. The third project uncovered an exceptional case within *Rhynchospora* in which both parents lack crossovers entirely, and only heterozygous progeny survive, highlighting an extreme of recombination suppression. Finally, the fourth project deviates from the holocentric plants, showing how the odd-ploidy dogroses circumvent the expected sterility of pentaploidy by deploying a dual mechanism of centromere structure and biased chromosome segregation. Together, these studies provide several cases of non-canonical meiosis in plants, from how holocentric chromosome architecture influences meiotic recombination to innovative meiotic drive to resolve odd-polyploid reproduction. Through the molecular and evolutionary consequences of altered meiotic control, we placed these findings into a broader genomic context and obtained an integrated view of how meiosis evolution and diversification evolved in different plants.

3.1 Synapsis initiation patterns as determinants of crossover landscape bias

One of the most consistent themes emerging from the comparison of different *Rhynhocpsora* species is the strong mechanistic link between the site of synapsis

initiation and the spatial organisation of meiotic crossovers. In *R. breviuscula*, detailed cytogenetic and genomic analyses revealed that synapsis begins predominantly at telomeric regions and proceeds inward, producing a characteristic distal bias in crossover distribution. This finding provided one of the first direct demonstrations in a holocentric plant that telomere-led synapsis can imprint the CO landscape. Extending this approach across multiple *Rhynchospora* species revealed that this pattern is not universal. Species where synapsis initiates randomly along chromosomes exhibit no consistent positional bias, showing instead an even or stochastic CO distribution. These results indicate that CO patterning in holocentric chromosomes is primarily dictated by synapsis initiation rather than by any other genomic features, considering the random distribution of all detected genetic and epigenetic signals so far.

This conclusion contrasts with classical monocentric models, where it is commonly believed that telomere clustering during the meiotic prophase I and centromere suppression jointly shape a distal CO enrichment. However, in *Rhynchospora*, the absence of a single primary centromere decouples these effects, exposing synapsis initiation as a major regulatory factor. We also found that CO does not form on the core centromere arrays in *Rhynchospora*, but this still results in a distally biased CO distribution pattern. Therefore, we speculate that centromere suppression may also act locally in monocentric species, probably with an extending or deepening effect because they usually have large pericentromere regions, but it's not the root cause of the distal bias. A more reliable determinant should be meiotic dynamics. Synapsis progression initiates at telomeres and moves inward, meaning distal regions spend more time in a recombination-competent state. A recent study in monocentric marsupials (Valero-Regalón et al., 2023) shows that, in the American marsupials (*Thylamys elegans* and *Dromiciops gliroides*), telomeres cluster in a strong bouquet configuration at leptotene/zygotene, and synapsis initiates exclusively from chromosome ends and then progresses inward from telomeres to interstitial regions. Conversely, the Australian species (*Macropus eugenii*) has a weak or incomplete bouquet, so the telomeres are less tightly clustered, and synapsis initiates both from telomeres and interstitial regions of chromosomes. At the same time, the RPA foci accumulate in the distal segments of bivalents in *T. elegans,* while more evenly distributed along the chromosomes in *M. eugenii*. Considering the role of RPA recruiting recombinase proteins RAD51 and DMC1 (Hinch et al., 2020; Osman et al., 2009), crossover distribution in *T. elegans* should be more likely to be telomere-biased than *M. eugenii*.

These findings further suggest that variation in synapsis initiation dynamics could represent an evolutionary mechanism generating diversity in recombination landscapes. Telomere clustering or bouquet formation may promote correct alignment and pairing at chromosome ends, but does not necessarily promote early synapsis from the telomeres, because we observed a telomere bouquet in early zygotene in some *Rhynchospora,* but the bouquet does not last, and the ZYP1 signal starts to load randomly.

To determine if the synapsis initiation in *Rhynchospora* or any other species, a comprehensive study of the influence of synapsis timing on crossover distribution is necessary. For example, one could check if synapsis initiation still starts from telomere proximity while disrupting telomere clustering in *R. breviuscula*. However, this may also disrupt the correct synapsis and fail to pass the meiosis checkpoints (Golubovskaya et al., 2002). Alternatively, one could also overexpress ZYP1 proteins, which may reduce the time difference of synapsis initiation along the chromosomes; then, reconstructing a new recombination landscape should give a hint at the temporal impact of synapsis initiation on the crossover distribution.

3.2 *Rhynchospora tenuis*: a mechatistic innovation but a potential evolutionary dead-end

Rhynchospora tenuis is, to our knowledge, the first case that both female and male meiosis lack crossovers in sexually reproducing eukaryotes. At the cytological and mechanistic level, the achiasmatic meiosis in *R. tenuis* is a remarkable innovation, not a simple loss of function. It demonstrates that accurate segregation of homologous chromosomes can occur without chiasmata, which has long been thought impossible in most eukaryotes. Paralleled with the inverted meiosis and reduced chromosome number, *R. tenuis* decrease its risk of missegregation dramatically even with the absence of chiasmata during meiosis.

From a population-genetic and evolutionary perspective, *R. tenuis* is potentially a short-term adaptation with long-term risk. Loss of recombination means loss of genetic shuffling, and all allelic combinations are inherited intact, drastically reducing the ability to generate new genotypes. Over time, this can increase mutational load (Muller's ratchet) (Felsenstein, 1974; Muller, 1964) and limit adaptability. Unless other mechanisms compensate, such as high heterozygosity, frequent outcrossing, polyploid buffering, etc (Fraser, 1972; Powell, 1997; Ronfort, 1999). The observed

“balanced lethal” or “heterozygote-only” viability in *R. tenuis* suggests a delicate equilibrium to ensure that only one genotype combination survives. This may ensure short-term stability by clonal preservation of a successful heterozygous genome, but may also limit evolvability.

Similar systems show that such states can persist for long evolutionary times, but they rarely diversify further. For example, *Oenothera* permanent translocation heterozygotes lead to genome-wide suppression of recombinations and heterozygous inheritance, analogous to *R. tenuis*, which is considered an evolutionary dead-end in diversification, despite persisting for millions of years (Holsinger & Ellstrand, 1984; Rauwolf et al., 2008). Hybridogenetic and parthenogenetic animals show similar semi-clonal inheritance, e.g. *Pelophylax esculentus*, a hybrid that does not contain mixed parental genomes (Holsbeek & Jooris, 2010), and *Poeciliopsis*, a sperm-dependent parthenogenetic fish (Schley et al., 2004). Both cases are effective for short-term niche persistence without recombinations, but rarely diversifying lineages. Holocentric Lepidoptera such as butterflies and moths display sex-specific CO suppression in females but compensate via male recombination, reinforcing that total recombination loss in both sexes is rare and evolutionarily costly (Rasmussen, 1977; Suomalainen et al., 2009).

In summary, we can think of *R. tenuis* as a frozen innovation. It rewires meiosis to bypass crossover-dependent segregation, demonstrating the flexibility of meiotic design. But it freezes genetic variation, making long-term diversification unlikely. It may represent a short-term adaptively optimal solution by maintaining the high heterozygosity, demonstrating the flexibility of the meiotic apparatus, i.e., how a lineage can retain the form of meiosis while radically altering its function. Our study has not reached an answer to why and how *R. tenuis* lost the ability of recombination. A more comprehensive transcriptome study will be followed to profile the detailed gene expression in *R. tenius* and find out the key responsible factors. Nevertheless, the fact that *R. tenuis* remains sexual rather than apomictic also suggests that developmental or ecological constraints still favour sexual cycles even without meiotic crossovers. But from an evolutionary standpoint, it may be walking a narrow ridge between stability and stagnation.

3.3 Centromere is more than a dock of the kinetochore complex

The centromere is a special chromosomal position that serves as the primary site for kinetochore assembly. Its fundamental role in both mitosis and meiosis is to ensure the faithful segregation of genetic material by orchestrating chromosome movement. During mitosis, centromeres ensure that sister chromatids are pulled apart to opposite poles of the cell. In canonical meiosis, the centromere's role becomes uniquely complex and critical for the two-stage division. During meiosis I, the centromeres of sister chromatids function as a single unit, co-orienting towards the same spindle pole to facilitate the segregation of homologous chromosomes. A key prerequisite for this is the protection of centromeric cohesion by the protein Shugoshin, which prevents the premature separation of sister chromatids in anaphase I (Kitajima et al., 2004). Subsequently, in meiosis II, centromeres must reprogram to act independently, like mitosis, to enable the separation of sister chromatids. This switch is essential for reducing ploidy and producing haploid gametes (Prosée et al., 2020).

The pentaploid dogroses ($2n = 5x = 35$) retain a stable sexual reproduction by deploying a dual mechanism of centromere structure and biased chromosome segregation. Only two of the chromosome sets form bivalents and are transmitted biparentally, while the remaining three sets continue as univalents and are transmitted exclusively through the maternal line. Our study finds a striking centromere bimodality. Bivalent-forming chromosomes carry smaller centromeres enriched in ATHILA retrotransposons, whereas univalent chromosomes bear larger tandem-repeat-rich centromeres, called CANR4 repeats, that recruit more CENH3 and apparently drive preferential transmission during asymmetric female meiosis. This biased transmission in female meiosis was also reported in other female meiosis, where the centromere is a central player in the phenomenon called meiotic drive, a process that does not follow Mendelian inheritance. Also known as centromere drive, this occurs when a stronger centromere variant on one chromosome can bias its segregation into the egg during the asymmetric female meiosis I. By attracting more microtubules or forming a larger kinetochore, a drive allele can exploit the meiotic machinery to be transmitted to the oocyte more frequently than the 50% expected by chance, even if it is detrimental to the organism (Akera et al., 2017; F. Finseth, 2023). For example, in monkeyflower (*Mimulus guttatus*), a centromeric region on haplotype *D* shows meiotic drive with increased transmission probability to eggs, which is associated with expanded arrays of centromeric repeats (Fishman & Saunders, 2008). Simultaneously, the driving

centromere (*D*) has led to adaptive evolution of a key kinetochore protein CenH3A, resulting in a recent selective sweep of a specific variant in natural populations, suggesting a mechanism for suppressing centromere drive, which can help restore fitness for organisms that are negatively affected by the drive's costs (F. R. Finseth et al., 2021). This genomic conflict is thought to be a powerful evolutionary force, driving the rapid evolution of both centromeric DNA and the kinetochore proteins that bind it to suppress this selfish activity (Henikoff et al., 2001; Malik, 2009). Thus, the centromere is not merely a passive attachment point but a dynamic genomic domain whose activity is pivotal for both the fidelity and the fairness of meiotic inheritance.

The dogrose story exemplifies how structural chromosome features such as centromeres and homolog availability can override classical homologous synapsis rules in specialised meiotic systems. While the other three projects in *Rhyncospora* show variation in synapsis, axis structure, and CO distribution across holocentric lineages, the dogrose system is driven by centromere-mediated non-recombinational segregation bias in a monocentric context and displays an extreme form of asymmetrical inheritance to maintain an odd chromosome set. This shows a rare polyploid lineage that can stabilise meiosis and maintain fertility through centromere diversity and transmission bias rather than through canonical recombination. Combined with the holocentric findings, it highlights that variations in chromosome architecture provide multiple evolutionary solutions for controlling meiotic outcomes.

3.4 Conserved core protein machinery and diversified chromosome-level behaviour

Core biochemical steps of meiosis, from DSB formation by SPO11, strand-exchange via recombinases, homolog search and resolution, cohesin function and synaptonemal complex formation, are deeply conserved because they represent constrained molecular solutions to the physical problems of homology search and DNA repair. By contrast, chromosome-level features such as centromere architecture, kinetochore enrichment, synapsis initiation pattern, axis variation, pairing strategies, etc., are relatively free to evolve because evolution can re-use the same molecular machinery while changing when and where it is deployed depending on complex chromatin contexts. The results from all projects show a wide diversity in how different species accomplish the same end, i.e., faithful segregation and gamete production without wholesale invention of new core proteins. Studies on holocentric

Rhynchospora show that when synapsis initiation is different, the CO distribution pattern varies, and that smaller chromosomes have a higher CO rate, smaller loops, and longer relative axis. These are precisely the kind of chromosome-level adjustments that use conserved recombination mechanisms but change meiotic outcomes. The existence of a complete meiosis protein repertoire in *R. tenuis* that appears to complete meiosis with no detectable COs but still produces viable heterozygous offspring shows that holocentric chromosome architecture permits alternative segregation strategies. In other words, rather than evolving a new meiosis pathway, homocentric species can alter the higher-level chromosomal behaviours to achieve reproduction. Work on other holocentrics has also repeatedly shown that kinetochore distribution and pairing modes can be remodelled in meiosis without changing core repair proteins. The dogrose system shows a different route in an odd-ploid monocentric context. Centromere bimodality and biased synapsis and transmission enforce a stable sexual system despite odd ploidy. The mechanism is centromere-level via sequence and CENH3 recruitment, not the invention of new core meiosis proteins, showing an example of how altering centromere architecture and pairing behaviour stabilises meiosis at the chromosome level.

In summary, the meiotic molecular factors perform enzymatic and structural tasks that are biophysically specific. Mutations that break those activities tend to be strongly deleterious, so purifying selection preserves their functions across deep phylogenies. By contrast, how those proteins are used, for example, whether DSBs are concentrated at telomeres, how wide the kinetochores can distribute, or how cohesin is released along an axis, can be modified by changes in regulatory elements, centromeric DNA repeats, spatial chromatin organisation, or expression of modulators, etc. These changes are often less pleiotropic and more easily tolerated by selection. Together, these projects show the same theme, that is, core proteins retained, higher-order chromosome behaviours evolve to meet particular constraints to improve adaptation.

4. References

Akera, T., Chmátal, L., Trimm, E., Yang, K., Aonbangkhen, C., Chenoweth, D. M., Janke, C., Schultz, R. M., & Lampson, M. A. (2017). Spindle asymmetry drives non-Mendelian chromosome segregation. *Science*, *358*(6363), 668–672. https://doi.org/10.1126/science.aan0092

Allers, T., & Lichten, M. (2001). Differential Timing and Control of Noncrossover and Crossover Recombination during Meiosis. *Cell*, *106*(1), 47–57. https://doi.org/10.1016/S0092-8674(01)00416-0

Alves, I., Houle, A. A., Hussin, J. G., & Awadalla, P. (2017). The impact of recombination on human mutation load and disease. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *372*(1736), 20160465. https://doi.org/10.1098/rstb.2016.0465

Anderson, L. K., Reeves, A., Webb, L. M., & Ashley, T. (1999). Distribution of Crossing Over on Mouse Synaptonemal Complexes Using Immunofluorescent Localization of MLH1 Protein. *Genetics*, *151*(4), 1569–1579. https://doi.org/10.1093/genetics/151.4.1569

Armstrong, S. J., Caryl, A. P., Jones, G. H., & Franklin, F. C. H. (2002). Asy1, a protein required for meiotic chromosome synapsis, localizes to axis-associated chromatin in *Arabidopsis* and *Brassica*. *Journal of Cell Science*, *115*(18), 3645–3655. https://doi.org/10.1242/jcs.00048

Baker, S. M., Plug, A. W., Prolla, T. A., Bronner, C. E., Harris, A. C., Yao, X., Christie, D.-M., Monell, C., Arnheim, N., Bradley, A., Ashley, T., & Liskay, R. M. (1996). Involvement of mouse Mlh1 in DNA mismatch repair and meiotic crossing over. *Nature Genetics*, *13*(3), 336–342. https://doi.org/10.1038/ng0796-336

Barlow, A., & Hultén, M. (1998). Crossing over analysis at pachytene in man. *European Journal of Human Genetics*, *6*(4), 350–358. https://doi.org/10.1038/sj.ejhg.5200200

Barton, N. H. (2009). Why Sex and Recombination? *Cold Spring Harbor Symposia on Quantitative Biology*, *74*(0), 187–195. https://doi.org/10.1101/sqb.2009.74.030

Bascón-Cardozo, K., Bours, A., Manthey, G., Durieux, G., Dutheil, J. Y., Pruisscher, P., Odenthal-Hesse, L., & Liedvogel, M. (2024). Fine-Scale Map Reveals Highly Variable Recombination Rates Associated with Genomic Features in the Eurasian Blackcap. *Genome Biology and Evolution*, *16*(1), evad233. https://doi.org/10.1093/gbe/evad233

Baudat, F., Buard, J., Grey, C., Fledel-Alon, A., Ober, C., Przeworski, M., Coop, G., & de Massy, B. (2010). PRDM9 is a major determinant of meiotic recombination hotspots in humans and mice. *Science (New York, N.Y.)*, *327*(5967), 836–840. https://doi.org/10.1126/science.1183439

Beye, M., Gattermeier, I., Hasselmann, M., Gempe, T., Schioett, M., Baines, J. F., Schlipalius, D., Mougel, F., Emore, C., Rueppell, O., Sirviö, A., Guzmán-Novoa, E., Hunt, G., Solignac, M., & Page, R. E. (2006). Exceptionally high levels of recombination across the honey bee genome. *Genome Research*, *16*(11), 1339–1344. https://doi.org/10.1101/gr.5680406

Bomblies, K. (2023). Learning to tango with four (or more): The molecular basis of adaptation to polyploid meiosis. *Plant Reproduction*, *36*(1), 107–124. https://doi.org/10.1007/s00497-022-00448-1

Borde, V., Robine, N., Lin, W., Bonfils, S., Géli, V., & Nicolas, A. (2009). Histone H3 lysine 4 trimethylation marks meiotic recombination initiation sites. *The EMBO Journal*, *28*(2), 99–111. https://doi.org/10.1038/emboj.2008.257

Börner, G. V., Kleckner, N., & Hunter, N. (2004). Crossover/Noncrossover Differentiation, Synaptonemal Complex Formation, and Regulatory Surveillance at the Leptotene/Zygotene Transition of Meiosis. *Cell*, *117*(1), 29–45. https://doi.org/10.1016/S0092-8674(04)00292-2

Bourguet, D., Gair, J., Mattice, M., & Whitlock, M. C. (2003). Genetic recombination and adaptation to fluctuating environments: Selection for geotaxis in Drosophila melanogaster. *Heredity*, *91*(1), 78–84. https://doi.org/10.1038/sj.hdy.6800283

Buetow, K. H. (1991). Influence of aberrant observations on high-resolution linkage analysis outcomes. *American Journal of Human Genetics*, *49*(5), 985–994.

Burke, B. (2018). LINC complexes as regulators of meiosis. *Current Opinion in Cell Biology*, *52*, 22–29. https://doi.org/10.1016/j.ceb.2018.01.005

Cain, N. E., Jahed, Z., Schoenhofen, A., Valdez, V. A., Elkin, B., Hao, H., Harris, N. J., Herrera, L. A., Woolums, B. M., Mofrad, M. R. K., Luxton, G. W. G., & Starr, D. A. (2018). Conserved SUN-KASH Interfaces Mediate LINC Complex-Dependent Nuclear Movement and Positioning. *Current Biology*, *28*(19), 3086-3097.e4. https://doi.org/10.1016/j.cub.2018.08.001

Campoy, J. A., Sun, H., Goel, M., Jiao, W.-B., Folz-Donahue, K., Wang, N., Rubio, M., Liu, C., Kukat, C., Ruiz, D., Huettel, B., & Schneeberger, K. (2020). Gamete binning: Chromosome-level and haplotype-resolved genome assembly enabled by high-throughput single-cell sequencing of gamete genomes. *Genome Biology*, *21*(1), 306. https://doi.org/10.1186/s13059-020-02235-5

Capilla-Pérez, L., Durand, S., Hurel, A., Lian, Q., Chambon, A., Taochy, C., Solier, V., Grelon, M., & Mercier, R. (2021). The synaptonemal complex imposes crossover interference and heterochiasmy in *Arabidopsis*. *Proceedings of the*

National Academy of Sciences, *118*(12), e2023613118. https://doi.org/10.1073/pnas.2023613118

Carioscia, S. A., Weaver, K. J., Bortvin, A. N., Pan, H., Ariad, D., Bell, A. D., & McCoy, R. C. (2022). A method for low-coverage single-gamete sequence analysis demonstrates adherence to Mendel's first law across a large sample of human sperm. *eLife*, *11*, e76383. https://doi.org/10.7554/eLife.76383

Castellani, M., Zhang, M., Thangavel, G., Mata-Sucre, Y., Lux, T., Campoy, J. A., Marek, M., Huettel, B., Sun, H., Mayer, K. F. X., Schneeberger, K., & Marques, A. (2024). Meiotic recombination dynamics in plants with repeat-based holocentromeres shed light on the primary drivers of crossover patterning. *Nature Plants*, *10*(3), 423–438. https://doi.org/10.1038/s41477-024-01625-y

Chambon, A., West, A., Vezon, D., Horlow, C., De Muyt, A., Chelysheva, L., Ronceret, A., Darbyshire, A., Osman, K., Heckmann, S., Franklin, F. C. H., & Grelon, M. (2018). Identification of ASYNAPTIC4, a Component of the Meiotic Chromosome Axis. *Plant Physiology*, *178*(1), 233–246. https://doi.org/10.1104/pp.17.01725

Chan, A. H., Jenkins, P. A., & Song, Y. S. (2012). Genome-Wide Fine-Scale Recombination Rate Variation in Drosophila melanogaster. *PLoS Genetics*, *8*(12), e1003090. https://doi.org/10.1371/journal.pgen.1003090

Chapman, N. H., & Thompson, E. A. (2001). Linkage disequilibrium mapping: The role of population history, size, and structure. *Advances in Genetics*, *42*, 413–437. https://doi.org/10.1016/s0065-2660(01)42034-7

Charlesworth, B., Nordborg, M., & Charlesworth, D. (1997). The effects of local selection, balanced polymorphism and background selection on equilibrium

patterns of genetic diversity in subdivided populations. *Genetical Research*, *70*(2), 155–174. https://doi.org/10.1017/s0016672397002954

Chelysheva, L., Vezon, D., Chambon, A., Gendrot, G., Pereira, L., Lemhemdi, A., Vrielynck, N., Le Guin, S., Novatchkova, M., & Grelon, M. (2012). The Arabidopsis HEI10 Is a New ZMM Protein Related to Zip3. *PLoS Genetics*, *8*(7), e1002799. https://doi.org/10.1371/journal.pgen.1002799

Chen, J.-F., Lu, F., Chen, S.-S., & Tao, S.-H. (2006). Significant positive correlation between the recombination rate and GC content in the human pseudoautosomal region. *Genome*, *49*(5), 413–419. https://doi.org/10.1139/g05-124

Cheng, E. Y., Hunt, P. A., Naluai-Cecchini, T. A., Fligner, C. L., Fujimoto, V. Y., Pasternack, T. L., Schwartz, J. M., Steinauer, J. E., Woodruff, T. J., Cherry, S. M., Hansen, T. A., Vallente, R. U., Broman, K. W., & Hassold, T. J. (2009). Meiotic Recombination in Human Oocytes. *PLoS Genetics*, *5*(9), e1000661. https://doi.org/10.1371/journal.pgen.1000661

Choi, K. (2017). Advances towards Controlling Meiotic Recombination for Plant Breeding. *Molecules and Cells*, *40*(11), 814–822. https://doi.org/10.14348/molcells.2017.0171

Choi, K., Zhao, X., Kelly, K. A., Venn, O., Higgins, J. D., Yelina, N. E., Hardcastle, T. J., Ziolkowski, P. A., Copenhaver, G. P., Franklin, F. C. H., McVean, G., & Henderson, I. R. (2013). Arabidopsis meiotic crossover hot spots overlap with H2A.Z nucleosomes at gene promoters. *Nature Genetics*, *45*(11), 1327–1336. https://doi.org/10.1038/ng.2766

Cornaro, L., Banfi, C., Cavalleri, A., Van Dijk, P. J., Radoeva, T., Cucinotta, M., & Colombo, L. (2025). Apomixis at high resolution: Unravelling diplospory in

Asteraceae. *Journal of Experimental Botany*, *76*(6), 1644–1657. https://doi.org/10.1093/jxb/erae477

Da Ines, O., Degroote, F., Goubely, C., Amiard, S., Gallego, M. E., & White, C. I. (2013). Meiotic Recombination in Arabidopsis Is Catalysed by DMC1, with RAD51 Playing a Supporting Role. *PLoS Genetics*, *9*(9), e1003787. https://doi.org/10.1371/journal.pgen.1003787

De Storme, N., & Geelen, D. (2020). High temperatures alter cross-over distribution and induce male meiotic restitution in Arabidopsis thaliana. *Communications Biology*, *3*(1), 187. https://doi.org/10.1038/s42003-020-0897-1

Dreissig, S., Maurer, A., Sharma, R., Milne, L., Flavell, A. J., Schmutzer, T., & Pillen, K. (2020). Natural variation in meiotic recombination rate shapes introgression patterns in intraspecific hybrids between wild and domesticated barley. *New Phytologist*, *228*(6), 1852–1863. https://doi.org/10.1111/nph.16810

Drouaud, J., Mercier, R., Chelysheva, L., Bérard, A., Falque, M., Martin, O., Zanni, V., Brunel, D., & Mézard, C. (2007). Sex-Specific Crossover Distributions and Variations in Interference Level along Arabidopsis thaliana Chromosome 4. *PLoS Genetics*, *3*(6), e106. https://doi.org/10.1371/journal.pgen.0030106

Durand, S., Lian, Q., Jing, J., Ernst, M., Grelon, M., Zwicker, D., & Mercier, R. (2022). Joint control of meiotic crossover patterning by the synaptonemal complex and HEI10 dosage. *Nature Communications*, *13*(1), 5999. https://doi.org/10.1038/s41467-022-33472-w

Eijpe, M., Offenberg, H., Jessberger, R., Revenkova, E., & Heyting, C. (2003). Meiotic cohesin REC8 marks the axial elements of rat synaptonemal complexes before cohesins SMC1β and SMC3. *The Journal of Cell Biology*, *160*(5), 657–670. https://doi.org/10.1083/jcb.200212080

Epstein, R., Wheeler, J., Hubisz, M., Sun, Q., Bukowski, R., Zhai, J., Lai, W.-Y., Buckler, E., & Pawlowski, W. P. (2024). *The maize recombination landscape evolved during domestication*. Evolutionary Biology. https://doi.org/10.1101/2024.11.04.621928

Farré, M., Micheletti, D., & Ruiz-Herrera, A. (2013). Recombination Rates and Genomic Shuffling in Human and Chimpanzee—A New Twist in the Chromosomal Speciation Theory. *Molecular Biology and Evolution*, *30*(4), 853–864. https://doi.org/10.1093/molbev/mss272

Felsenstein, J. (1974). The evolutionary advantage of recombination. *Genetics*, *78*(2), 737–756. https://doi.org/10.1093/genetics/78.2.737

Ferdous, M., Higgins, J. D., Osman, K., Lambing, C., Roitinger, E., Mechtler, K., Armstrong, S. J., Perry, R., Pradillo, M., Cuñado, N., & Franklin, F. C. H. (2012). Inter-Homolog Crossing-Over and Synapsis in Arabidopsis Meiosis Are Dependent on the Chromosome Axis Protein AtASY3. *PLoS Genetics*, *8*(2), e1002507. https://doi.org/10.1371/journal.pgen.1002507

Finseth, F. (2023). Female meiotic drive in plants: Mechanisms and dynamics. *Current Opinion in Genetics & Development*, *82*, 102101. https://doi.org/10.1016/j.gde.2023.102101

Finseth, F. R., Nelson, T. C., & Fishman, L. (2021). Selfish chromosomal drive shapes recent centromeric histone evolution in monkeyflowers. *PLOS Genetics*, *17*(4), e1009418. https://doi.org/10.1371/journal.pgen.1009418

Fishman, L., & Saunders, A. (2008). Centromere-Associated Female Meiotic Drive Entails Male Fitness Costs in Monkeyflowers. *Science*, *322*(5907), 1559–1562. https://doi.org/10.1126/science.1161406

Foss, E., Lande, R., Stahl, F. W., & Steinberg, C. M. (1993). Chiasma interference as a function of genetic distance. *Genetics*, *133*(3), 681–691. https://doi.org/10.1093/genetics/133.3.681

Fraser, A. S. (1972). An introduction to population genetic theory. By J. F. Crow and M. Kimura. Harper and Row, New York. 656 pp. 1970. *Teratology*, *5*(3), 386–387. https://doi.org/10.1002/tera.1420050318

Giraut, L., Falque, M., Drouaud, J., Pereira, L., Martin, O. C., & Mézard, C. (2011). Genome-Wide Crossover Distribution in Arabidopsis thaliana Meiosis Reveals Sex-Specific Patterns along Chromosomes. *PLoS Genetics*, *7*(11), e1002354. https://doi.org/10.1371/journal.pgen.1002354

Golubovskaya, I. N., Harper, L. C., Pawlowski, W. P., Schichnes, D., & Cande, W. Z. (2002). The pam1 gene is required for meiotic bouquet formation and efficient homologous synapsis in maize (Zea mays L.). *Genetics*, *162*(4), 1979–1993. https://doi.org/10.1093/genetics/162.4.1979

Gonzalo, A., Nayak, A., & Bomblies, K. (2025). Improved synapsis dynamics accompany meiotic stability in *Arabidopsis arenosa* autotetraploids. *Proceedings of the National Academy of Sciences*, *122*(19), e2420115122. https://doi.org/10.1073/pnas.2420115122

Grandont, L., Jenczewski, E., & Lloyd, A. (2013). Meiosis and Its Deviations in Polyploid Plants. *Cytogenetic and Genome Research*, *140*(2–4), 171–184. https://doi.org/10.1159/000351730

Gruhn, J. R., Rubio, C., Broman, K. W., Hunt, P. A., & Hassold, T. (2013). Cytological Studies of Human Meiosis: Sex-Specific Differences in Recombination Originate at, or Prior to, Establishment of Double-Strand Breaks. *PLoS ONE*, *8*(12), e85075. https://doi.org/10.1371/journal.pone.0085075

Heckmann, S., Jankowska, M., Schubert, V., Kumke, K., Ma, W., & Houben, A. (2014). Alternative meiotic chromatid segregation in the holocentric plant Luzula elegans. *Nature Communications*, *5*(1), 4979. https://doi.org/10.1038/ncomms5979

Henikoff, S., Ahmad, K., & Malik, H. S. (2001). The Centromere Paradox: Stable Inheritance with Rapidly Evolving DNA. *Science*, *293*(5532), 1098–1102. https://doi.org/10.1126/science.1062939

Higgins, J. D., Perry, R. M., Barakate, A., Ramsay, L., Waugh, R., Halpin, C., Armstrong, S. J., & Franklin, F. C. H. (2012). Spatiotemporal Asymmetry of the Meiotic Program Underlies the Predominantly Distal Distribution of Meiotic Crossovers in Barley. *The Plant Cell*, *24*(10), 4096–4109. https://doi.org/10.1105/tpc.112.102483

Higgins, J. D., Sanchez-Moran, E., Armstrong, S. J., Jones, G. H., & Franklin, F. Chris. H. (2005). The *Arabidopsis* synaptonemal complex protein ZYP1 is required for chromosome synapsis and normal fidelity of crossing over. *Genes & Development*, *19*(20), 2488–2500. https://doi.org/10.1101/gad.354705

Hill, W. G., & Robertson, A. (1966). The effect of linkage on limits to artificial selection. *Genetical Research*, *8*(3), 269–294.

Hinch, A. G., Becker, P. W., Li, T., Moralli, D., Zhang, G., Bycroft, C., Green, C., Keeney, S., Shi, Q., Davies, B., & Donnelly, P. (2020). The Configuration of RPA, RAD51, and DMC1 Binding in Meiosis Reveals the Nature of Critical Recombination Intermediates. *Molecular Cell*, *79*(4), 689-701.e10. https://doi.org/10.1016/j.molcel.2020.06.015

Hofstatter, P. G., Thangavel, G., Castellani, M., & Marques, A. (2021). Meiosis Progression and Recombination in Holocentric Plants: What Is Known?

Frontiers in Plant Science, *12*, 658296. https://doi.org/10.3389/fpls.2021.658296

Hofstatter, P. G., Thangavel, G., Lux, T., Neumann, P., Vondrak, T., Novak, P., Zhang, M., Costa, L., Castellani, M., Scott, A., Toegelová, H., Fuchs, J., Mata-Sucre, Y., Dias, Y., Vanzela, A. L. L., Huettel, B., Almeida, C. C. S., Šimková, H., Souza, G., … Marques, A. (2022). Repeat-based holocentromeres influence genome architecture and karyotype evolution. *Cell*, *185*(17), 3153-3168.e18. https://doi.org/10.1016/j.cell.2022.06.045

Hojsgaard, D., & Hörandl, E. (2019). The Rise of Apomixis in Natural Plant Populations. *Frontiers in Plant Science*, *10*, 358. https://doi.org/10.3389/fpls.2019.00358

Hollingsworth, N. M., & Brill, S. J. (2004). The Mus81 solution to resolution: Generating meiotic crossovers without Holliday junctions. *Genes & Development*, *18*(2), 117–125. https://doi.org/10.1101/gad.1165904

Hollister, J. D., Arnold, B. J., Svedin, E., Xue, K. S., Dilkes, B. P., & Bomblies, K. (2012). Genetic Adaptation Associated with Genome-Doubling in Autotetraploid Arabidopsis arenosa. *PLoS Genetics*, *8*(12), e1003093. https://doi.org/10.1371/journal.pgen.1003093

Holsbeek, G., & Jooris, R. (2010). Potential impact of genome exclusion by alien species in the hybridogenetic water frogs (Pelophylax esculentus complex). *Biological Invasions*, *12*(1), 1–13. https://doi.org/10.1007/s10530-009-9427-2

Holsinger, K. E., & Ellstrand, N. C. (1984). The Evolution and Ecology of Permanent Translocation Heterozygotes. *The American Naturalist*, *124*(1), 48–71. https://doi.org/10.1086/284251

Husband, B. C. (2004). The role of triploid hybrids in the evolutionary dynamics of mixed-ploidy populations: TRIPLOIDS IN MIXED-PLOIDY POPULATIONS.

Biological Journal of the Linnean Society, *82*(4), 537–546. https://doi.org/10.1111/j.1095-8312.2004.00339.x

Jin, Q., Trelles-Sticken, E., Scherthan, H., & Loidl, J. (1998). Yeast Nuclei Display Prominent Centromere Clustering That Is Reduced in Nondividing Cells and in Meiotic Prophase. *The Journal of Cell Biology*, *141*(1), 21–29. https://doi.org/10.1083/jcb.141.1.21

Jing, J., Lian, Q., Durand, S., & Mercier, R. (2025). Maximizing meiotic crossover rates reveals the map of Crossover Potential. *Nature Communications*, *16*(1), 5306. https://doi.org/10.1038/s41467-025-60663-y

Johnston, S. E., Huisman, J., Ellis, P. A., & Pemberton, J. M. (2017). A High-Density Linkage Map Reveals Sexual Dimorphism in Recombination Landscapes in Red Deer (*Cervus elaphus*). *G3 Genes|Genomes|Genetics*, *7*(8), 2859–2870. https://doi.org/10.1534/g3.117.044198

Keeney, S., Giroux, C. N., & Kleckner, N. (1997). Meiosis-Specific DNA Double-Strand Breaks Are Catalyzed by Spo11, a Member of a Widely Conserved Protein Family. *Cell*, *88*(3), 375–384. https://doi.org/10.1016/S0092-8674(00)81876-0

Kerstes, N. A., Bérénos, C., Schmid-Hempel, P., & Wegner, K. M. (2012). Antagonistic experimental coevolution with a parasite increases host recombination frequency. *BMC Evolutionary Biology*, *12*(1), 18. https://doi.org/10.1186/1471-2148-12-18

Kianian, P. M. A., Wang, M., Simons, K., Ghavami, F., He, Y., Dukowic-Schulze, S., Sundararajan, A., Sun, Q., Pillardy, J., Mudge, J., Chen, C., Kianian, S. F., & Pawlowski, W. P. (2018). High-resolution crossover mapping reveals similarities and differences of male and female recombination in maize.

Nature Communications, *9*(1), 2370. https://doi.org/10.1038/s41467-018-04562-5

Kitajima, T. S., Kawashima, S. A., & Watanabe, Y. (2004). The conserved kinetochore protein shugoshin protects centromeric cohesion during meiosis. *Nature*, *427*(6974), 510–517. https://doi.org/10.1038/nature02312

Kong, A., Gudbjartsson, D. F., Sainz, J., Jonsdottir, G. M., Gudjonsson, S. A., Richardsson, B., Sigurdardottir, S., Barnard, J., Hallbeck, B., Masson, G., Shlien, A., Palsson, S. T., Frigge, M. L., Thorgeirsson, T. E., Gulcher, J. R., & Stefansson, K. (2002). A high-resolution recombination map of the human genome. *Nature Genetics*, *31*(3), 241–247. https://doi.org/10.1038/ng917

Kong, A., Thorleifsson, G., Frigge, M. L., Masson, G., Gudbjartsson, D. F., Villemoes, R., Magnusdottir, E., Olafsdottir, S. B., Thorsteinsdottir, U., & Stefansson, K. (2014). Common and low-frequency variants associated with genome-wide recombination rate. *Nature Genetics*, *46*(1), 11–16. https://doi.org/10.1038/ng.2833

Kovarik, A., Werlemark, G., Leitch, A. R., Souckova-Skalicka, K., Lim, Y. K., Khaitová, L., Koukalova, B., & Nybom, H. (2008). The asymmetric meiosis in pentaploid dogroses (Rosa sect. Caninae) is associated with a skewed distribution of rRNA gene families in the gametes. *Heredity*, *101*(4), 359–367. https://doi.org/10.1038/hdy.2008.63

Lamarche, B. J., Orazio, N. I., & Weitzman, M. D. (2010). The MRN complex in double-strand break repair and telomere maintenance. *FEBS Letters*, *584*(17), 3682–3695. https://doi.org/10.1016/j.febslet.2010.07.029

Lander, E. S., & Green, P. (1987). Construction of multilocus genetic linkage maps in humans. *Proceedings of the National Academy of Sciences of the United States of America*, *84*(8), 2363–2367. https://doi.org/10.1073/pnas.84.8.2363

Li, X., Yu, M., Bolaños-Villegas, P., Zhang, J., Ni, D., Ma, H., & Wang, Y. (2021). Fanconi anemia ortholog FANCM regulates meiotic crossover distribution in plants. *Plant Physiology*, *186*(1), 344–360. https://doi.org/10.1093/plphys/kiab061

Lloyd, A., & Bomblies, K. (2016). Meiosis in autopolyploid and allopolyploid Arabidopsis. *Current Opinion in Plant Biology*, *30*, 116–122. https://doi.org/10.1016/j.pbi.2016.02.004

Lloyd, A., & Jenczewski, E. (2019). Modelling Sex-Specific Crossover Patterning in *Arabidopsis*. *Genetics*, *211*(3), 847–859. https://doi.org/10.1534/genetics.118.301838

Lukaszewicz, A., Shodhan, A., & Loidl, J. (2015). Exo1 and Mre11 execute meiotic DSB end resection in the protist Tetrahymena. *DNA Repair*, *35*, 137–143. https://doi.org/10.1016/j.dnarep.2015.08.005

Lunerová, J., Herklotz, V., Laudien, M., Vozárová, R., Groth, M., Kovařík, A., & Ritz, C. M. (2020). Asymmetrical canina meiosis is accompanied by the expansion of a pericentromeric satellite in non-recombining univalent chromosomes in the genus Rosa. *Annals of Botany*, *125*(7), 1025–1038. https://doi.org/10.1093/aob/mcaa028

Malik, H. S. (2009). The Centromere-Drive Hypothesis: A Simple Basis for Centromere Complexity. In D. Ugarkovic (Ed.), *Centromere* (Vol. 48, pp. 33–52). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-00182-6_2

Malone, C. J., Fixsen, W. D., Horvitz, H. R., & Han, M. (1999). UNC-84 localizes to the nuclear envelope and is required for nuclear migration and anchoring during *C. elegans* development. *Development*, *126*(14), 3171–3181. https://doi.org/10.1242/dev.126.14.3171

Márquez-Corro, J. I., Martín-Bravo, S., Blanco-Pastor, J. L., Luceño, M., & Escudero, M. (2024). The holocentric chromosome microevolution: From phylogeographic patterns to genomic associations with environmental gradients. *Molecular Ecology*, *33*(24), e17156. https://doi.org/10.1111/mec.17156

Martin, H. C., Christ, R., Hussin, J. G., O'Connell, J., Gordon, S., Mbarek, H., Hottenga, J.-J., McAloney, K., Willemsen, G., Gasparini, P., Pirastu, N., Montgomery, G. W., Navarro, P., Soranzo, N., Toniolo, D., Vitart, V., Wilson, J. F., Marchini, J., Boomsma, D. I., … Donnelly, P. (2015). Multicohort analysis of the maternal age effect on recombination. *Nature Communications*, *6*(1), 7846. https://doi.org/10.1038/ncomms8846

Masson, J.-Y., & West, S. C. (2001). The Rad51 and Dmc1 recombinases: A non-identical twin relationship. *Trends in Biochemical Sciences*, *26*(2), 131–136. https://doi.org/10.1016/S0968-0004(00)01742-4

Melters, D. P., Paliulis, L. V., Korf, I. F., & Chan, S. W. L. (2012). Holocentric chromosomes: Convergent evolution, meiotic adaptations, and genomic analysis. *Chromosome Research*, *20*(5), 579–593. https://doi.org/10.1007/s10577-012-9292-1

Mercier, R., Mézard, C., Jenczewski, E., Macaisne, N., & Grelon, M. (2015). The Molecular Biology of Meiosis in Plants. *Annual Review of Plant Biology*, *66*(1), 297–327. https://doi.org/10.1146/annurev-arplant-050213-035923

Modliszewski, J. L., Wang, H., Albright, A. R., Lewis, S. M., Bennett, A. R., Huang, J., Ma, H., Wang, Y., & Copenhaver, G. P. (2018). Elevated temperature increases meiotic crossover frequency via the interfering (Type I) pathway in Arabidopsis thaliana. *PLOS Genetics*, *14*(5), e1007384. https://doi.org/10.1371/journal.pgen.1007384

Morgan, C., Fozard, J. A., Hartley, M., Henderson, I. R., Bomblies, K., & Howard, M. (2021). Diffusion-mediated HEI10 coarsening can explain meiotic crossover positioning in Arabidopsis. *Nature Communications*, *12*(1), 4674. https://doi.org/10.1038/s41467-021-24827-w

Morgan, T. H. (1911). Random Segregation Versus Coupling in Mendelian Inheritance. *Science*, *34*(873), 384–384. https://doi.org/10.1126/science.34.873.384

Muller, H. J. (1916). The Mechanism of Crossing-Over. *The American Naturalist*, *50*(592), 193–221. https://doi.org/10.1086/279534

Muller, H. J. (1964). The relation of recombination to mutational advance. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis*, *1*(1), 2–9. https://doi.org/10.1016/0027-5107(64)90047-8

Myers, S., Bottolo, L., Freeman, C., McVean, G., & Donnelly, P. (2005). A fine-scale map of recombination rates and hotspots across the human genome. *Science (New York, N.Y.)*, *310*(5746), 321–324. https://doi.org/10.1126/science.1117196

Myers, S., Bowden, R., Tumian, A., Bontrop, R. E., Freeman, C., MacFie, T. S., McVean, G., & Donnelly, P. (2010). Drive against hotspot motifs in primates implicates the PRDM9 gene in meiotic recombination. *Science (New York, N.Y.)*, *327*(5967), 876–879. https://doi.org/10.1126/science.1182363

Näsvall, K., Boman, J., Höök, L., Vila, R., Wiklund, C., & Backström, N. (2023). Nascent evolution of recombination rate differences as a consequence of chromosomal rearrangements. *PLOS Genetics*, *19*(8), e1010717. https://doi.org/10.1371/journal.pgen.1010717

Neumann, P., Oliveira, L., Jang, T.-S., Novák, P., Koblížková, A., Schubert, V., Houben, A., & Macas, J. (2023). Disruption of the standard kinetochore in

holocentric Cuscuta species. *Proceedings of the National Academy of Sciences of the United States of America*, *120*(21), e2300877120. https://doi.org/10.1073/pnas.2300877120

Osman, K., Sanchez-Moran, E., Mann, S. C., Jones, G. H., & Franklin, F. C. H. (2009). Replication protein A (AtRPA1a) is required for class I crossover formation but is dispensable for meiotic DNA break repair. *The EMBO Journal*, *28*(4), 394–404. https://doi.org/10.1038/emboj.2008.295

Parée, T., & Teotónio, H. (2025). Experimental tests on the evolution of sex and recombination and their adaptive significance. *Journal of Evolutionary Biology*, *38*(7), 798–810. https://doi.org/10.1093/jeb/voaf028

Pelé, A., Falque, M., Trotoux, G., Eber, F., Nègre, S., Gilet, M., Huteau, V., Lodé, M., Jousseaume, T., Dechaumet, S., Morice, J., Poncet, C., Coriton, O., Martin, O. C., Rousseau-Gueutin, M., & Chèvre, A.-M. (2017). Amplifying recombination genome-wide and reshaping crossover landscapes in Brassicas. *PLoS Genetics*, *13*(5), e1006794. https://doi.org/10.1371/journal.pgen.1006794

Peñalba, J. V., & Wolf, J. B. W. (2020). From molecules to populations: Appreciating and estimating recombination rate variation. *Nature Reviews Genetics*, *21*(8), 476–492. https://doi.org/10.1038/s41576-020-0240-1

Phillips, D., Jenkins, G., Macaulay, M., Nibau, C., Wnetrzak, J., Fallding, D., Colas, I., Oakey, H., Waugh, R., & Ramsay, L. (2015). The effect of temperature on the male and female recombination landscape of barley. *New Phytologist*, *208*(2), 421–429. https://doi.org/10.1111/nph.13548

Powell, J. R. (1997). *Progress and Prospects in Evolutionary Biology: The Drosophila Model*. Oxford University PressNew York, NY. https://doi.org/10.1093/oso/9780195076912.001.0001

Prosée, R. F., Wenda, J. M., & Steiner, F. A. (2020). Adaptations for centromere function in meiosis. *Essays in Biochemistry*, *64*(2), 193–203. https://doi.org/10.1042/EBC20190076

Rasmussen, S. W. (1977). Meiosis in Bombyx mori females. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *277*(955), 343–350. https://doi.org/10.1098/rstb.1977.0022

Rauwolf, U., Golczyk, H., Meurer, J., Herrmann, R. G., & Greiner, S. (2008). Molecular marker systems for Oenothera genetics. *Genetics*, *180*(3), 1289–1306. https://doi.org/10.1534/genetics.108.091249

Reliene, R., Yamamoto, M. L., Rao, P. N., & Schiestl, R. H. (2010). Genomic instability in mice is greater in Fanconi anemia caused by deficiency of Fancd2 than Fancg. *Cancer Research*, *70*(23), 9703–9710. https://doi.org/10.1158/0008-5472.CAN-09-1022

Rey, M.-D., Martín, A. C., Smedley, M., Hayta, S., Harwood, W., Shaw, P., & Moore, G. (2018). Magnesium Increases Homoeologous Crossover Frequency During Meiosis in ZIP4 (Ph1 Gene) Mutant Wheat-Wild Relative Hybrids. *Frontiers in Plant Science*, *9*, 509. https://doi.org/10.3389/fpls.2018.00509

Ronfort, J. (1999). The mutation load under tetrasomic inheritance and its consequences for the evolution of the selfing rate in autotetraploid species. *Genetical Research*, *74*(1), 31–42. https://doi.org/10.1017/S0016672399003845

Rönspies, M., Schmidt, C., Schindele, P., Lieberman-Lazarovich, M., Houben, A., & Puchta, H. (2022). Massive crossover suppression by CRISPR–Cas-mediated plant chromosome engineering. *Nature Plants*, *8*(10), 1153–1159. https://doi.org/10.1038/s41477-022-01238-3

Sakuno, T., Tashiro, S., Tanizawa, H., Iwasaki, O., Ding, D.-Q., Haraguchi, T., Noma, K., & Hiraoka, Y. (2022). Rec8 Cohesin-mediated Axis-loop chromatin architecture is required for meiotic recombination. *Nucleic Acids Research*, *50*(7), 3799–3816. https://doi.org/10.1093/nar/gkac183

Sandor, C., Li, W., Coppieters, W., Druet, T., Charlier, C., & Georges, M. (2012). Genetic Variants in REC8, RNF212, and PRDM9 Influence Male Recombination in Cattle. *PLoS Genetics*, *8*(7), e1002854. https://doi.org/10.1371/journal.pgen.1002854

Santos, T. D. L., Hunter, N., Lee, C., Larkin, B., Loidl, J., & Hollingsworth, N. M. (2003). The Mus81/Mms4 Endonuclease Acts Independently of Double-Holliday Junction Resolution to Promote a Distinct Subset of Crossovers During Meiosis in Budding Yeast. *Genetics*, *164*(1), 81–94. https://doi.org/10.1093/genetics/164.1.81

Sardell, J. M., & Kirkpatrick, M. (2020). Sex Differences in the Recombination Landscape. *The American Naturalist*, *195*(2), 361–379. https://doi.org/10.1086/704943

Scherthan, H. (2007). Telomeres and meiosis in health and disease: Telomere attachment and clustering during meiosis. *Cellular and Molecular Life Sciences*, *64*(2), 117–124. https://doi.org/10.1007/s00018-006-6463-2

Schley, D., Doncaster, C. P., & Sluckin, T. (2004). Population models of sperm-dependent parthenogenesis. *Journal of Theoretical Biology*, *229*(4), 559–572. https://doi.org/10.1016/j.jtbi.2004.04.031

Schmidt, A. (2025). Developmental Plasticity of Apomixis Reproduction – Chance or Challenge for Evolution and Application? *Critical Reviews in Plant Sciences*, 1–33. https://doi.org/10.1080/07352689.2025.2480531

Serra, H., Lambing, C., Griffin, C. H., Topp, S. D., Nageswaran, D. C., Underwood, C. J., Ziolkowski, P. A., Séguéla-Arnaud, M., Fernandes, J. B., Mercier, R., & Henderson, I. R. (2018). Massive crossover elevation via combination of *HEI10* and *recq4a recq4b* during *Arabidopsis* meiosis. *Proceedings of the National Academy of Sciences*, *115*(10), 2437–2442. https://doi.org/10.1073/pnas.1713071115

Sherizen, D., Jang, J. K., Bhagat, R., Kato, N., & McKim, K. S. (2005). Meiotic recombination in Drosophila females depends on chromosome continuity between genetically defined boundaries. *Genetics*, *169*(2), 767–781. https://doi.org/10.1534/genetics.104.035824

Smagulova, F., Brick, K., Pu, Y., Camerini-Otero, R. D., & Petukhova, G. V. (2016). The evolutionary turnover of recombination hot spots contributes to speciation in mice. *Genes & Development*, *30*(3), 266–280. https://doi.org/10.1101/gad.270009.115

Smukowski Heil, C. S., Ellison, C., Dubin, M., & Noor, M. A. F. (2015). Recombining without Hotspots: A Comprehensive Evolutionary Portrait of Recombination in Two Closely Related Species of *Drosophila*. *Genome Biology and Evolution*, *7*(10), 2829–2842. https://doi.org/10.1093/gbe/evv182

Stapley, J., Feulner, P. G. D., Johnston, S. E., Santure, A. W., & Smadja, C. M. (2017). Variation in recombination frequency and distribution across eukaryotes: Patterns and processes. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *372*(1736), 20160455. https://doi.org/10.1098/rstb.2016.0455

Stumpf, M. P. H., & McVean, G. A. T. (2003). Estimating recombination rates from population-genetic data. *Nature Reviews. Genetics*, *4*(12), 959–968. https://doi.org/10.1038/nrg1227

Suomalainen, E., Cook, L. M., & Turner, J. R. G. (2009). Achiasmatic oogenesis in the Heliconiine butterflies. *Hereditas*, *74*(2), 302–304. https://doi.org/10.1111/j.1601-5223.1973.tb01134.x

Täckholm, G. (1922). *Zytologische Studien über die Gattung Rosa*. Almqvist & Wiksells boktr. https://doi.org/10.5962/bhl.title.15507

Uecker, H., & Hermisson, J. (2016). The Role of Recombination in Evolutionary Rescue. *Genetics*, *202*(2), 721–732. https://doi.org/10.1534/genetics.115.180299

Underwood, C. J., Choi, K., Lambing, C., Zhao, X., Serra, H., Borges, F., Simorowski, J., Ernst, E., Jacob, Y., Henderson, I. R., & Martienssen, R. A. (2018). Epigenetic activation of meiotic recombination near *Arabidopsis thaliana* centromeres via loss of H3K9me2 and non-CG DNA methylation. *Genome Research*, *28*(4), 519–531. https://doi.org/10.1101/gr.227116.117

Underwood, C. J., & Mercier, R. (2022). Engineering Apomixis: Clonal Seeds Approaching the Fields. *Annual Review of Plant Biology*, *73*(1), 201–225. https://doi.org/10.1146/annurev-arplant-102720-013958

Valero-Regalón, F. J., Solé, M., López-Jiménez, P., Valerio-de Arana, M., Martín-Ruiz, M., De La Fuente, R., Marín-Gual, L., Renfree, M. B., Shaw, G., Berríos, S., Fernández-Donoso, R., Waters, P. D., Ruiz-Herrera, A., Gómez, R., & Page, J. (2023). Divergent patterns of meiotic double strand breaks and synapsis initiation dynamics suggest an evolutionary shift in the meiosis program between American and Australian marsupials. *Frontiers in Cell and Developmental Biology*, *11*, 1147610. https://doi.org/10.3389/fcell.2023.1147610

Vrielynck, N., Chambon, A., Vezon, D., Pereira, L., Chelysheva, L., De Muyt, A., Mézard, C., Mayer, C., & Grelon, M. (2016). A DNA topoisomerase VI–like

complex initiates meiotic recombination. *Science*, *351*(6276), 939–943. https://doi.org/10.1126/science.aad5196

Wang, W., He, J., Wang, Y., Li, N., Sun, Q., Zhang, M., Li, Q., Xu, H., Shao, K., Zhang, S., Liu, S., Chen, L., Liu, X., Tian, Y., Zhao, Z., Liu, Y., & Wan, J. (2025). Breaking genetic linkage barriers to improve brown planthopper resistance and grain yield in japonica rice. *Theoretical and Applied Genetics*, *138*(8), 199. https://doi.org/10.1007/s00122-025-04978-w

Wang, Y., Fuentes, R. R., Van Rengs, W. M. J., Effgen, S., Zaidan, M. W. A. M., Franzen, R., Susanto, T., Fernandes, J. B., Mercier, R., & Underwood, C. J. (2024). Harnessing clonal gametes in hybrid crops to engineer polyploid genomes. *Nature Genetics*, *56*(6), 1075–1079. https://doi.org/10.1038/s41588-024-01750-6

Watanabe, Y., & Nurse, P. (1999). Cohesin Rec8 is required for reductional chromosome segregation at meiosis. *Nature*, *400*(6743), 461–464. https://doi.org/10.1038/22774

Weber, C. C., Boussau, B., Romiguier, J., Jarvis, E. D., & Ellegren, H. (2014). Evidence for GC-biased gene conversion as a driver of between-lineage differences in avian base composition. *Genome Biology*, *15*(12), 549. https://doi.org/10.1186/s13059-014-0549-1

Wells, D. (2020). *Delineation of meiotic gene expression in male mice* [PhD Thesis]. University of Oxford.

Whigham, P. A., & Spencer, H. G. (2021). Graph-structured populations and the Hill-Robertson effect. *Royal Society Open Science*, *8*(3), 201831. https://doi.org/10.1098/rsos.201831

Wright, K. M., Arnold, B., Xue, K., Šurinová, M., O'Connell, J., & Bomblies, K. (2015). Selection on meiosis genes in diploid and tetraploid Arabidopsis arenosa.

Molecular Biology and Evolution, *32*(4), 944–955. https://doi.org/10.1093/molbev/msu398

Xiong, Z., Gaeta, R. T., & Pires, J. C. (2011). Homoeologous shuffling and chromosome compensation maintain genome balance in resynthesized allopolyploid *Brassica napus*. *Proceedings of the National Academy of Sciences*, *108*(19), 7908–7913. https://doi.org/10.1073/pnas.1014138108

Zelazowski, M. J., Sandoval, M., Paniker, L., Hamilton, H. M., Han, J., Gribbell, M. A., Kang, R., & Cole, F. (2017). Age-Dependent Alterations in Meiotic Recombination Cause Chromosome Segregation Errors in Spermatocytes. *Cell*, *171*(3), 601-614.e13. https://doi.org/10.1016/j.cell.2017.08.042

Zhang, L., Stauffer, W., Zwicker, D., & Dernburg, A. F. (2021). *Crossover patterning through kinase-regulated condensation and coarsening of recombination nodules*. Cell Biology. https://doi.org/10.1101/2021.08.26.457865

Zheng, C., Boer, M. P., & van Eeuwijk, F. A. (2019). Construction of Genetic Linkage Maps in Multiparental Populations. *Genetics*, *212*(4), 1031–1044. https://doi.org/10.1534/genetics.119.302229

Zickler, D., & Kleckner, N. (2015). Recombination, Pairing, and Synapsis of Homologs during Meiosis. *Cold Spring Harbor Perspectives in Biology*, *7*(6), a016626. https://doi.org/10.1101/cshperspect.a016626

Zickler, D., & Kleckner, N. (2023). Meiosis: Dances Between Homologs. *Annual Review of Genetics*, *57*(1), 1–63. https://doi.org/10.1146/annurev-genet-061323-044915

Ziolkowski, P. A., Berchowitz, L. E., Lambing, C., Yelina, N. E., Zhao, X., Kelly, K. A., Choi, K., Ziolkowska, L., June, V., Sanchez-Moran, E., Franklin, C., Copenhaver, G. P., & Henderson, I. R. (2015). Juxtaposition of heterozygous and homozygous regions causes reciprocal crossover remodelling via

interference during Arabidopsis meiosis. *eLife*, *4*, e03708.

https://doi.org/10.7554/eLife.03708