
. JOANA CARVALHO.

Evolution of centromere and meiosis in plants

Inaugural-Dissertation

Zur
Erlangung des Doktorgrades
der Mathematisch-Naturwissenschaftlichen Fakultät
der Universität zu Köln

vorgelegt von

Gokilavani Thangavel

aus

Sathyamangalam, India

Köln, August 2023

Die vorliegende Arbeit wurde am Max-Planck-Institut für Pflanzenzüchtungsforschung in Köln in der Abteilung Chromosomenbiologie (Direktor: Prof. Dr. Raphael Mercier) innerhalb der Arbeitsgruppe von Dr. André Marques angefertigt.

Gutachter und Prüfer

1. Prof. Dr. Raphael Mercier
2. Prof. Dr. Bart Thomma

Vorsitzender der Prüfungskommission
Prof. Dr. Alga Zuccaro

Beisitzender
Dr. André Marques

Zeitraum der letzten mündlichen Prüfung
25.07.2023

Table of contents

Acknowledgement

All my research work was performed in Max Planck Institute for Plant Breeding Research (MPIPZ), Cologne, Germany. My journey as a Ph.D. student was possible only because of the fellowship provided by German Academic Exchange Service (DAAD). I would like to thank my supervisor Dr. André Marques for giving me a chance to do my Ph.D. Despite of difficult situations, I thank him for supporting me till the end. I would like to convey my special thanks to other members of my thesis advisory committee (TAC) – Dr. Charles J. Underwood and Prof. Dr. Andreas Houben who have been there for me from the very beginning of my Ph.D. They have guided me, tracked my progress, given their valuable suggestions for improvement of the project and mentally supported me when there were hardships. I very much appreciate the time they have allotted to make it to my TAC meetings and evaluate my reports. I would like to thank my formal supervisor Dr. Raphael Mercier for his inputs for the project. He had created a very friendly working environment in our department, which very much helped me to adapt. Throughout my research period, there was great technical support from Christina, other technicians, green house team and IT team, for which I am thankful. I thank Neysan Donnelly and Stefan Strütt for their comments on the thesis.

A huge thanks from the bottom of my heart to the student coordinator, Stephan Wagner. Whenever there was an issue, Stephan was always there to look into the issue with empathy and provide the support needed. If it was not for him, I am not sure how I would have managed all the administrative works related to the university and survived the difficult and tensed situations of my Ph.D.

I personally would like to thank my family who trusted in me and didn't object me from travelling thousands of kilometres for pursuing my dreams. Doing a PhD is not just about research. The path was always tough with lots of ups and downs. It was always my friends especially Nami who was there to give me the mental support, I needed. Thanks to all my friends whom I didn't mention by name, members of the department, colleagues from the institute who had accompanied me in this tough yet beautiful journey.

Dedication

This thesis is dedicated to all the brave immigrants in science

Science knows no borders.

Neither do the immigrants in science.

Reminding us, the pursuit of knowledge transcends,

The challenges of geography and culture.

List of abbreviations

µL	microliter
APC/C	Anaphase Promoting Complex/Cyclosome
AT	Adenine – Thymine
BLAST	Basic Local Alignment Search Tool
BLOSUM	Blocks Substitution Matrix
bp	base pair
cenDNA	centromeric DNA sequence
ChIP-seq	Chromatin Immunoprecipitation sequencing
CPC	Chromosomal Passenger Complex
CRISPR-Cas9	clustered regularly interspaced short palindromic repeats and CRISPR-associated protein 9
DAPI	2 mg/ml 4,6-diamidino-2-phenylindole
DH5α	DH5-Alpha Cell
DNA	Deoxyribonucleic acid
EDTA	Ethylenediaminetetraacetic acid
FISH	fluorescence *in situ* hybridization
g	gram
gRNA	guide RNA
HMM	Hidden Markov Models
kDa	Kilodalton
M	Molar
MAFFT	Multiple Alignment using Fast Fourier Transform
mg	milligram
mL	millilitre
mM	millimolar
NGS	Next Generation Sequencing
OD	Optical Density
pH	potential of hydrogen
PMSF	phenylmethylsulfonyl fluoride
PP2A	Phosphatase 2 A
PSI-BLAST	Position-Specific Iterative Basic Local Alignment Search Tool

PVP	Polyvinylpyrrolidone
REC8	Recombination 8
RNA	Ribonucleic acid
SCC	Sister Chromatid Cohesion
SDS	Sodium dodecyl sulfate
SGO	Shugoshin
sgRNA	single guide RNA
SMC1	Structural Maintenance of Chromosome 1
SMC3	Structural Maintenance of Chromosome 3
ß-ME	beta-mercaptoethanol
TAIR	The Arabidopsis Information Resource
TEs	Transposable Elements
Tris-HCl	Tris hydrochloride
TUs	Transcriptional Units
U3	E3 ubiquitin ligase
UniProtKB	Universal Protein Resource Knowledgebase
v/v	volume by volume
w/v	weight in volume
WAPL	Wings apart-like protein homolog

List of publications

Thangavel, Gokilavani, Paulo G. Hofstatter, Raphaël Mercier, and André Marques. "Tracing the evolution of the plant meiotic molecular machinery." Plant Reproduction (2023): 1-23.

Hofstatter, Paulo G., **Gokilavani Thangavel**, Thomas Lux, Pavel Neumann, Tihana Vondrak, Petr Novak, Meng Zhang et al. "Repeat-based holocentromeres influence genome architecture and karyotype evolution." Cell 185, no. 17 (2022): 3153-3168.

Hofstatter, Paulo G., **Gokilavani Thangavel**, Marco Castellani, and André Marques. "Meiosis progression and recombination in holocentric plants: What is known?." Frontiers in Plant Science 12 (2021): 658296.

Abstract

The centromere is an essential region of the chromosome and a pre-requisite to ensure faithful segregation of the chromosomes during mitosis and meiosis. Holocentromeres are defined by the presence of several centromeric units along the entire chromosomes. Despite their existence in several animal and plant lineages, very little is known about holocentric chromosomes. The investigation of the holocentric worm *Cenorhabditis elegans* provided most of the current insights into nature and function of holocentromeres. Considering holocentricity having evolved several times independently, adaptations to its holocentric nature are also expected to be different.

The first chapter of the thesis aims to understand the epigenetic regulation of the repeat-based holocentromeres found in several *Rhynchospora* species belonging to the monocot family, Cyperaceae. Our study discovers that *Rhynchospora* chromosomes are composed of hundreds of small centromeric units defined by the presence of *Tyba* repeats associated with CENH3. The increase in H3K9me2 and CHG methylation at the border of each centromeric unit gives the cues about the exact centromere borders. We found this epigenetic regulation is strikingly similar to *Arabidopsis thaliana*, suggesting the existence of an evolutionarily conserved mechanism regulating repeat-based centromeres of both monocentric and holocentric organisms. We found evidence that holocentricity facilitates karyotype evolution by promoting chromosome fusions in this genus.

The second chapter of the thesis aims to understand how holocentricity is linked to the inverted meiosis found in several *Rhynchospora* species. I studied the cytological localisation of the key candidate genes involved in the protection of the sister centromere cohesion, Recombination 8 (REC8) and Shugoshin (SGO) during inverted meiosis. The signals for both REC8 and SGO begin to appear as early as interphase. They are observed on the chromosomes during leptotene, zygotene, and pachytene stages, and then start to disappear from diplotene. Only residual signals remain from metaphase I. Most importantly, colocalization studies showed no obvious correlation with the centromeric regions for REC8 and SGO. This indicates that the sister centromere cohesion persevered during the metaphase I of canonical meiosis is lost during the metaphase I of inverted meiosis. Moreover, the non-canonical localisation pattern of SGO to chromosomal regions other than centromeres, hints at a possible non-centromeric role in this species. Taken

together, our findings support the hypothesis of early centromeric cohesion loss, which allows sister-chromatids to segregate already at anaphase I promoting inverted meiosis.

Meiosis is ancestral to all the eukaryotes and, it is expected that most of the meiotic machinery is conserved across different eukaryotic lineages. However, studies elucidating the meiotic machinery in the whole kingdom Plantae are lacking. The third chapter of the thesis provides a detailed phylogenetic analysis of the proteins involved in the meiotic machinery. Most of the proteins were found to be conserved in the major lineages. Interestingly, ASY4, PHS1, PRD2, PRD3 were not detected in distant algal lineages. We could trace the origin of the proteins reported in plants only, DFO and HEIP1 to the ancestor of vascular plants and Streptophyta, respectively. We found that the SPO11 duplication is ancestral to all eukaryotes but lost in the green algal lineages, Chlorophyta and Charophyta.

This thesis presents three novel, first of its kind research studies on the unique characteristics of holocentric plants and their meiotic processes. Our findings offered novel insights into the regulation of repeat-based holocentromeres, adaptations to inverted meiosis, and the evolutionary conservation of plant meiotic machinery.

Zusammenfassung

Das Zentromer ist eine essenzielle Region des Chromosoms und wird zur Sicherstellung einer korrekten Segregation der Chromosomen während der Mitose und Meiose benötigt. Das Vorhandensein mehrerer zentromerischer Einheiten entlang des gesamten Chromosoms wird als Holozentromer oder Ganzchromosomen-Zentromer definiert. Trotz der Existenz in vielen Tier- und Pflanzenlinien des evolutionären Stammbaums, ist bisher wenig über holozentromerische Chromosomen bekannt. Studien am holozentrischen Wurm *Caenorhabditis elegans* lieferten bisher die einzigen wissenschaftlichen Kenntnisse über Natur und Funktion von Holozentromeren. Dass Holozentrizität im Laufe der Evolution mehrere Male unabhängig voneinander entstanden ist, lässt vermuten, dass jeweils spezifische Adaptionen an die holozentromerischen Beschaffenheiten vorhanden sind.

Im ersten Kapitel dieer Thesis werden die Ergebnisse der Untersuchung der epigenetischen Regulation der Repeat-basierten Holozentromere, die in einigen Spezies der *Rhynchospora* aus der Familie der Monokotylen entdeckt wurden, diskutiert. Die Ergebnisse zeigen, dass die Chromosomen der *Rhynchospora* aus hunderten kleiner zentromerischer Einheiten bestehen, die durch die Anwesenheit von *Tyba*-Repeats assoziiert mit CENH3 determiniert sind. Die erhöhte Methylierung von H3K9me2 und CHG am Rand jeder zentromerischen Einheit markiert das genaue lokale Vorkommen des Zentromers. Die Ergebnisse der Untersuchungen von *Arabidopsis thaliana* zeigen auf, dass es bemerkenswerte Ähnlichkeiten in der Beschaffenheit zwischen holo- und monozentrischen Strukturen gibt. Dies deutet darauf hin, dass die Regulation der Repeat-basierten Zentromere evolutionär konserviert und in vielen Organismen einheitlich sein könnte. Die Ergebnisse dieses Kapitels liefern Nachweise, dass Holozentrizität der Chromosomen in diesem Genus die Evolution des Karyotypen fördert.

Im zweiten Kapitel dieser Thesis werden die Zusammenhänge zwischen Holozentrizität und invertierter Meiose in verschiedenen *Rhynchospora* untersucht. Bisher gibt es noch keine Studien, in denen die miotische Maschinerie der Pflanzen untersucht wird. In diesem Kapitel wurde die zytologische Lokalisation der Schlüsselgene, REC8 und SGO, die essenziell zur Kohäsion der Schwesterchromadiden beitragen, im Kontext der invertierten Meiose untersucht. Genannte Gene sind während des Leptotäns, Zygotäns und des Pachytäns auf den Chromosomen zu finden und

beginnen dann während des Diplotäns wieder zu verschwinden. Lediglich Spuren des Signals verbleiben über die Metaphase I hinaus. Interessanterweise gibt es keine Hinweise auf eine örtliche Korrelation des Auftretens von der Zentromerregionen für REC8 und SGO. Dies deutet darauf hin, dass die Schwesterchromatidkohäsion bei der kanonischen Meiose konserviert, aber bei der invertierten Meiose in der Metaphase I verlorengegangen ist. Zudem deutet das Vorhandensein von SGO in anderen chromosomalen Regionen als den Zentromeren darauf hin, dass es möglicherweise in dieser Spezies eine andere oder weiterreichende Funktion hat. Diese Ergebnisse stützen die Hypothese des frühen zentromerischen Kohäsionsverlustes, unter der Schwesterchromatiden bereits in der Anaphase I segregieren, was wiederum zu einer Förderung der invertierten Meiose führt.

Die Meiose ist eine spezifische Zellteilung, um haploide Gameten zu erzeugen und deren Vielfalt durch homologe Rekombination zu fördern. Da Meiose in allen eukaryotischen Abstammungslinien evolutionär ursprünglich ist, kann davon ausgegangen werden, dass die meiotische Maschinerie in allen Eukaryoten konserviert ist. Im dritten Kapitel dieser Thesis wird die evolutionäre Geschichte der in die Meiose involvierten Proteine aus phylogenetischer Perspektive untersucht. Sequenzdaten unterschiedlicher Pflanzenlinien wurden verwendet. Die meisten Proteine waren in den größeren Pflanzenlinien konserviert. Bemerkenswerterweise wurden ASY4, PHS1, PRD2 und PRD3 in entfernt verwandten Algen nicht detektiert. Der Ursprung der Proteine, DFO und HEIP1, deren Existenz bisher nur in Pflanzen belegt wurde, konnte bis zu den Vorfahren der Gefäßpflanzen, den Streptophyta, zurückverfolgt werden. Die Untersuchungen zeigen, dass SPO11 allen Vorfahren der Eukaryoten gemein ist, jedoch in den Abstammungslinien der Chlorophyta und Charophyta verlorengegangen ist.

In dieser Thesis werden als erste ihrer Art drei Forschungsstudien der Charakteristiken holozentrischer Pflanzen und ihres meiotischen Prozesses vorgelegt. Sie bieten neue Einblicke in die Regulation Repeat-basierter Holozentromere, Adaptionen an invertierte Meiose und die evolutionäre Konservierung der pflanzlichen Meiose Maschinerie.

Introduction

Centromeres

Centromeres are essential regions of eukaryotic chromosomes, where kinetochores assemble and, microtubules attach, resulting in the segregation of chromosomes during both mitosis and meiosis. Centromeres take different forms and are diverse in terms of size, architecture, underlying DNA sequences and centromeric proteins. Centromere size varies from the 120bp-long point centromeres in the budding yeast, *Saccharomyces cerevisiae* (Clarke and Carbon, 1985) to 30 to 50 Kb long in *Schizosaccharomyces pombe* (Baum et al., 1994; Murakami et al., 1991) to several megabase pair-long regional centromeres in humans (Altemose et al., 2022). Some organisms have non-repetitive core centromeres flanked by repetitive sequences as in the case of *S. pombe* whereas other centromeres are completely made up of arrays of AT-rich satellite repeats as in the case of humans. In most plants and animals, centromeres are arranged in tandem arrays whose monomeric units are around 150bp in size, forming higher-order repetitive core centromeres (Kursel and Malik, 2016).

What defines a centromere?

Centromeres were initially considered to be genetically defined by the underlying DNA sequence (cenDNA) as in the case of *S. cerevisiae* (Clarke and Carbon, 1985). But later, many discoveries on regional centromeres led to this conclusion, to be challenged. In humans, when centromeres are lost due to chromosomal rearrangement, neocentromeres can arise *de novo* on non-repetitive regions which do not contain cenDNA (Harrington et al., 1997; Marshall et al., 2008). Centromeres are considered to be epigenetically defined and the presence of the CENH3 (CENP-A in mammals) protein, the histone H3 variant replacing the canonical histone H3 of the centromere nucleosomes usually marks functional centromeres (Kursel and Malik, 2016). CENH3 has two domains: a conserved histone fold domain which shares 62% similarity to histone H3 and a divergent amino-terminal tail domain. The histone fold domain targets CENH3 to centromeres. The amino-terminal domain is significantly divergent from histone H3 and CENH3 in other species (Goutte-Gattat et al., 2013; Sullivan et al., 1994). However, the later discovery of CENH3-independent holocentromeres in insects (Drinnenberg et al., 2014) and the presence of centromeric proteins without homology to CENH3 in

Trypanosoma brucei (Akiyoshi and Gull, 2014), has shown that centromere identity is even more diverse and complex than was previously thought. CENH3 localises to the both satellite repeat cenDNA-based centromeres and non-cenDNA-based neocentromeres in humans, meaning that CENH3 is still a reliable marker for centromeres in most organisms (Kursel and Malik, 2016). CENH3 is essential for the recruitment and formation of all the components of the kinetochore (Régnier et al., 2005).

Epigenetic regulation of the centromeres

DNA methylation and post-translational histone modifications such as acetylation, methylation, phosphorylation, and ubiquitination are the major epigenetic regulations which define the chromatin state (Jenuwein and Allis, 2001). As mentioned before, the key and most common epigenetic modification that regulates centromeric chromatin is the deposition of the specific histone variant CENH3, and its incorporation into the centromeric chromatin (McKinley and Cheeseman, 2016; Talbert et al., 2002; Yoda et al., 2000). CENH3 is necessary for the assembly of the kinetochores and the attachment of the spindle fibres. In the centromeric regions, CENH3-containing nucleosomes are interspersed with canonical H3-containing nucleosomes (Sullivan and Karpen, 2004). The permissive transcription euchromatin mark (open chromatin mark), H3K4me2 is enriched in the H3 subdomains of the centromeric regions in humans. Interestingly, the core centromeric chromatin is neither associated with the heterochromatin mark, H3K9me2 and H3K9me3 nor with the active transcription mark H3K4me3. On the contrary, the flanking pericentromeric region is enriched for the heterochromatin mark H3K9me2 (Sullivan and Karpen, 2004). *Schizosaccharomyces pombe* core centromeric chromatin is also associated with H3K4me and the flanking regions with H3K9me2 (Cam et al., 2005). In the case of *Arabidopsis thaliana*, H3K9me2 is generally enriched within the centromeres, while its levels are reduced in the core region compared to the flanking pericentromeric region. H3K4me3 is slightly increased in the core region compared to the pericentromeres (Naish et al., 2021). Similar to histone modifications, DNA methylation also regulates the centromeric chromatin state (Scelfo and Fachinetti, 2019). However, the highly repetitive nature of centromeres makes it difficult to study their DNA methylation patterns. Recently, the methylation landscape of *A. thaliana* was described, which included a complete assembly of the organism's centromeres. CG methylation was found to be dense in the centromeric region, whereas the non-CG methylation, i.e., CHH

and CHG, exhibited relatively reduced enrichment in the cores compared to the pericentromeres. This positively correlates with H3K9me2 pattern, which maintains DNA methylation in non-CG contexts (Naish et al., 2021).

Centromeres – the black boxes of chromosomes

Centromeres have remained as the black box of the chromosomes because of their complex nature. Regional centromeres are typically made up of a single unit of mega base pair-long satellite DNA arranged in head-to-tail tandem arrays making it complicated to study or to sequence them. The satellite arrays are considered to be the dark matter of the genome. Their nearly identical nature makes it difficult to assemble them, which results in huge gaps in genome assemblies (Talbert and Henikoff, 2022). The recently developed long and ultra-long sequencing technologies like Pacific Biosciences (PacBio) SMRT technology and Oxford Nanopore technology have enabled sequencing and assembly of complete regional centromeres. Complete centromere assemblies are now possible and the first releases for maize, *A. thaliana* and humans made using these technologies are now available (Logsdon et al., 2021; Miga et al., 2020; Naish et al., 2021; Nurk et al., 2022; Wolfgruber et al., 2016).

Holocentric organisms and adaptations

Centromeres were initially discovered as a single region to which the microtubules attach. They form the primary constriction on chromosomes, in monocentric organisms giving rise to the typical 'V'-shaped chromosomes during metaphase. Holocentric organisms, on the other hand, have centromeres all along their chromosomes, a diffuse formation of kinetochores, and microtubules attach from one end to the other end of the chromosomes. In the case of holocentric organisms, having diffuse centromeres that are not constrained to a single region appears to pose several challenges for chromosome segregation during meiosis. Different organisms have evolved different strategies to deal with holocentricity. Holocentric plants have evolved inverted meiosis or post-reductional meiosis to segregate their chromosomes during meiosis (Cabral et al., 2014; Heckmann et al., 2014a; Hofstatter et al., 2021; Malheiros et al., 1947; Pazy and Plitmann, 1991; Wahl, 1940).

Repeat based holocentromeres

Holocentromeres have been reported to have evolved independently at least 13 different times (Escudero et al., 2016; Melters et al., 2012). Thus, huge variation exists among the holocentromeres and different types of holocentromeres have been reported. Some holocentromeric units have been described to be associated with specific repeat DNA sequences while others do not (Gassmann et al., 2012; Marques et al., 2015). Similarly, holocentric insects have lost CENH3/CENPC (Senaratne et al., 2021), whereas *C. elegans* and plants have CENH3-dependent holocentromeres (Marques et al., 2016; Steiner and Henikoff, 2014). Unlike, most of the monocentric organisms, centromeric-specific DNA sequences are not commonly reported in the well-studied holocentric organisms like *C. elegans* (Gassmann et al., 2012; Steiner and Henikoff, 2014), insects (Senaratne et al., 2021) and plant species like *Luzula* (Heckmann et al., 2013). As an exception, holocentromeres of *Rhynchospora* species, were first reported to be associated with centromeric-specific repeats. In this case, hundreds of short tandem arrays of the repeat *Tyba* are found interspersed along the entire chromosome, showing specific association with CENH3 (Hofstatter et al., 2022; Marques et al., 2015).

Meiosis

Meiosis is a highly conserved, essential mechanism ancestral to all eukaryotes. It is a specialised cell division and differs from mitosis. Mitosis involves a single round of DNA replication followed by a single round of chromosome segregation. Whereas meiosis involves single round of DNA replication followed by two rounds to chromosome segregation, halving the ploidy level, to produce haploid gametes from diploid parental cells. The haploid gametes from both parents fuse together by fertilization to form the diploid zygote and give rise to the next generation. This forms the basis mode of reproduction for all sexually reproducing organisms. Successful meiosis is dependent on the action of several molecular pathways, including formation of DNA double-strand breaks, establishment of sister chromatid cohesion, pairing of homologous chromosomes and the formation of the synaptonemal complex, homologous recombination and the two-step segregation of the chromosomes. Hundreds of different proteins are involved in these different pathways. Since meiosis is ancestral to eukaryotes, these proteins should have a shared homology between the different eukaryotic lineages. However, lineage

specific proteins have been discovered and some proteins have been reported to be dispensable in certain lineages as well (Thangavel et al., 2023).

Canonical meiosis

The diploid microspore/megaspore mother cell has two homologous chromosomes, one coming from the maternal and the other from the paternal side. The mother cell undergoes DNA replication to produce sister chromatids resulting in two pairs of homologous chromosomes, held together by sister chromatid cohesion (SCC) (Uhlmann and Nasmyth, 1998). During meiosis I, the kinetochores assemble on the sister chromatids, act as a single unit, and mono-orient i.e. attach to the microtubules emanating from the same pole. As the result, the sister chromatids are kept together and segregate to the same pole, in a process called reductional division. During meiosis II, sister chromatid cohesion is entirely lost, releasing the sister chromatids that were kept together. As a result, kinetochores assemble, act as two different units, and bi-orient, i.e., attach to the microtubules coming from two different poles. The sister chromatids are segregated to two different poles at the end of meiosis II, which is known as the equational division (Hauf and Watanabe, 2004).

Inverted meiosis

Inverted meiosis is observed in holocentric plants and describes a form of meiosis in which the sequence of events is inverted with respect to canonical meiosis. As stated in the review (Hofstatter et al., 2021), inverted meiosis was first reported as early as 1940 in Carex (Wahl, 1940) and later discovered in other holocentric plants like *Cuscuta*, *Luzula* and *Rhynchospora* (Cabral et al., 2014; Heckmann et al., 2014a; Heckmann et al., 2014b; Malheiros et al., 1947; Pazy and Plitmann, 1991); the phenomenon has also been observed in holocentric insects (Battaglia and Boyes, 1955; Nokkala et al., 2002; Viera et al., 2009). In this type of meiosis, during metaphase I the sister- chromatids biorient, attaching to spindle fibres from opposite poles and resulting in segregation to opposite poles at the end of meiosis I (equational division during meiosis I). Thus, the daughter cells remain diploid, at the end of meiosis I. During meiosis II, some unknown mechanism helps the homologous non-sister chromatids to align and to segregate to the opposite poles (reductional division during meiosis II).

Not much is known about the candidates involved in the evolution of the inverted meiosis. As per the theory of inverted meiosis, sisters are segregated to two different poles during meiosis I itself. This hints at a scenario, in which the sister centromere cohesion preserved during the metaphase I of canonical meiosis is lost during inverted meiosis, enabling their segregation (equational division) to take place during meiosis I. The proteins involved in sister centromere cohesion protection, the meiotic specific cohesin Recombination 8 (REC8) and the cohesion protector Shugoshin (SGO) could possibly play role in the evolution of meiosis.

REC8

REC8 is the meiotic α-kleisin of the highly conserved sister chromatid cohesion complex along with SMC1, SMC3 and SCC3. Two step cohesion loss is essential for the proper segregation of chromosomes during canonical meiosis. In the first step, chromosome arm cohesion is lost and centromeric cohesion is retained in a WAPL dependent prophase pathway. SGO plays a vital role in regulating this process, by protecting the cohesins only around centromeres while exposing the cohesins along the chromosome arms, allowing their disassociation from the chromosomes. In the second step, upon entry into anaphase, when kinetochores exert tension on the sister centromeres, SGO is removed from the centromeres, resulting in centromeric cohesion loss. During meiosis, centromeric cohesion is maintained until anaphase II, which allows the sisters to be kept together until their segregation in meiosis II (Gregan et al., 2008; Kitajima et al., 2004; Liu et al., 2013; Marston, 2015; Watanabe and Kitajima, 2005). However, this is hypothesized to be different in the case of inverted meiosis. Understanding REC8 dynamics in a holocentric plant with inverted meiosis can help us elucidate more about centromeric cohesion, which is one of the main objectives of this study.

Shugoshin

The highly conserved Shugoshin/ Mei-S332 family of proteins was discovered for its primary role in centromeric cohesion protection. Other than its canonical cohesion protection roles, cohesion-independent functions which may be evolutionarily conserved but have been previously unrecognised, have also been reported. SGO ensures proper sister kinetochore bi-orientation and chromosome segregation by recruiting the

chromosome passenger complex (CPC). To ensure biorientation of the sister kinetochores at metaphase and the subsequent proper segregation of sister chromatids to opposite poles, kinetochores should attach to microtubules originating from a single pole. Initially, microtubules from both poles try to connect to the same kinetochore and finally proper attachment is achieved. This error correction between kinetochores and microtubules is regulated by the CPC (Foley and Kapoor, 2013; Tanaka, 2010). SGO is also proposed to play a role in centriolar cohesion during mitosis of humans (Wang et al., 2008). SGO controls checkpoint activity during meiotic prophase in *C. elegans* (Bohr et al., 2018). Recently, SGO was even found to be linked to cardiac pacing activity and a mutation of the protein causes Chronic Atrial and Intestinal Dysrhythmia Syndrome (Liu et al., 2021). Since inverted meiosis doesn't need the centromeric cohesion protection, studying SGO in a model like *Rhynchospora* can illuminate the possible conserved, non-centromeric roles of SGO.

Diversity of the kingdom Plantae

To study meiosis in plants, it is necessary to understand plant reproduction and the factors that influence genetic diversity (through meiotic recombination), insights that can be harnessed to improve crop production. Another important aspect of studying meiosis is an understanding of the evolutionary relationships, especially the evolution of sex determination and dioecy in plants. Generally, plants in a broader term is used to describe the photosynthetic organisms which can synthesize their own food. Archaeplastida or kingdom Plantae are largely comprised of Rhodophyta (photoautotrophic red algae), glaucophyta (unicellular freshwater algae) and Viridiplantae (green algae and land plants)(Palmer et al., 2004). They also include Rhodelphidia (non-photosynthetic algae sister to Rhodophyta) (Gawryluk et al., 2019) and the microscopic picozoans (Not et al., 2007). Viridiplantae are generally known as green plants and are comprised of aquatic green algae (chlorophytes and charophytes) and land plants (embryophytes). Land plants are hypothesized to have evolved within charophytes and the charophyte class Zygnematophyceae forms the sister group. Embryophytes can be further categorized into bryophytes (liverworts, hornworts, mosses), lycophytes, pteridophytes (ferns) and spermatophytes (gymnosperms and angiosperms) (Palmer et al., 2004; Puttick et al., 2018). Plants are diverse in nature and approximately 500,000 land plant species are estimated to exist (Corlett, 2016).

Advanced sequencing approaches allow comparative study of meiosis across plants

The rich diversity and the huge number of plant species makes it difficult to study not just meiosis but any particular feature in detail throughout the kingdom. The majority of studies on meiosis have been carried out in a few angiosperm models like *A. thaliana*, rice, maize, wheat, barley among others (Mercier and Grelon, 2008). However, these studies do not represent the breadth of diversity in the meiotic machinery across the whole kingdom. Advances in sequencing technology are resulting in the increasing availability of large amounts of sequencing data, primarily in the form of genomic data. This large amount of sequencing data enables scientists to better understand the relationships among different organisms with greater resolution. Even though more comprehensive meiosis studies are missing for the plant kingdom, one can make use of the available sequencing date from representatives of the different plant lineages, to identify key genes involved in the complex meiotic pathways.

Computational approaches for studying meiosis in plants

Homology searches are computational methods that look for similarities among different organisms based on simple DNA or protein sequences and complex structural and functional features. PSI-BLAST and HMMER are advanced homology search approaches used to identify homologs of the initial DNA/protein query sequence even in distantly related species. In PSI-BLAST (Position-Specific Iterative Basic Local Alignment Search Tool), an initial protein query is searched for in a database to identify matches based on the local alignment. Based on this initial protein-protein BLAST, PSI-BLAST creates a profile or position-specific scoring matrix (PSSM) by performing multiple sequence alignment of the hits above a certain threshold level. This profile is further used to find homology-based matches in the database, and the profile is updated after each subsequent iteration using the matches from that iteration. This process is continued until desired hits or no further new hits are obtained (Altschul et al., 1997; Bhagwat and Aravind, 2008). Similarly, HMMER (Hidden Markov Model-based Evolutionary Analysis of Sequences) make use of the probabilistic model, Hidden Markov Models (HMMs)-based profiles for homology searches against the desired database (Eddy, 1992). Since both PSI-BLAST and HMMER make use of the profile which stores the underlying conservation, they can find similarities in distant sequences of the initial query.

SPO11 duplication

SPO11 is the homolog of the archaeal topoisomerase VIA enzyme. It is meiosis-specific and forms a complex with topoisomerase VIB to create meiotic DNA double-strand breaks (Robert et al., 2016; Vrielynck et al., 2016). Several eukaryotes have more than one SPO11 homolog (Malik et al., 2007). The SPO11 scenario is even more complex in plants. The model plants *A. thaliana* encodes three (Hartung and Puchta, 2000, 2001) whereas rice encodes for five (An et al., 2011; Jain et al., 2006) SPO11 homologs. Interestingly in *A. thaliana*, both AtSPO11-1 and AtSPO11-2 are essential for meiotic recombination (Hartung and Puchta, 2000). Topoisomerase VI enzymes act as heterotetramers consisting of two A and two B subunits. The detected duplications hint at a scenario in which SPO11 homologs interact with each other to form heterotetramers (Robert et al., 2016).

Objectives of this study

The major aim of my research was to shed light on chromosome adaptations arising from the transition to holocentricity by studying the holocentric plants belonging to the genus *Rhynchospora* and to understand the evolution of meiotic genes in the kingdom Plantae. The first chapter of the thesis is aimed at understanding the genetic and epigenetic basis of repeat-based holocentromeres and how they influence genome architecture and karyotype evolution in the genus *Rhynchospora*. The second chapter of the thesis studies the role of cohesion during inverted meiosis in holocentric plants. We speculate that sister chromatid cohesion may be a critical point of difference between canonical and inverted meiosis and may have facilitated the evolution of inverted meiosis as an adaptation to holocentricity. To examine this hypothesis, this chapter studies the cohesion dynamics during inverted meiosis in the model plants, *R. pubera* and *R. breviuscula*. The third chapter of the thesis aims to understand the evolution and conservation of the meiotic machinery across the entire Plantae kingdom. Using phylogenetic analysis, I have traced the evolution of the different meiotic proteins which has allowed me to illuminate the evolution of the meiosis in plants.

Chapter 1

Repeat-based holocentromeres influence the genome architecture and karyotype evolution

Summary

Centromeres are epigenetically defined and influence the overall genome organisation, distribution of genes, eu- and hetero-chromatin domains, and occurrence of meiotic crossovers. The genus *Rhynchospora* is known to harbour repeat-based holocentromeres composed by a single tandem-repeat (*Tyba*). However, detailed characterisation of its holocentromeres have been missing due to the lack of a reference genome. This study sheds light on epigenetic regulation of holocentromeres by comparing three different holocentric beak sedges, i.e., *R. pubera, R. breviuscula, R. tenuis* and its closest monocentric relative *Juncus effusus*. Chromosome-scale reference genomes were developed for all these species. Chromatin Immunoprecipitation sequencing (ChIP-seq) experiments were performed for the centromeric protein-CENH3 in *R. pubera* and *R. breviuscula*. In addition to that, in case of *R. pubera*, ChIP-Seq was carried out for the well-defined euchromatin mark-H3K4me3, heterochromatin mark-H3K9me2 and methylation status was analysed in the context of CG, CHG, CHH. These epigenetic features are compared with the common genome features like presence of genes, repetitive elements, transposons, centromeric sequences and RNA transcription profiles to get an idea about the influence of holocentromeres on the genome architecture.

My contribution

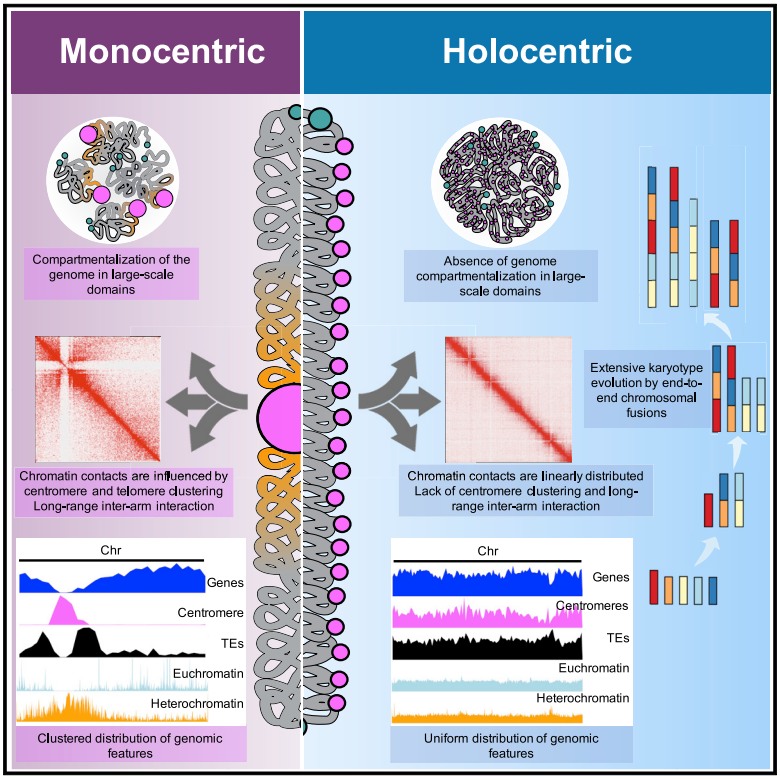
My contribution to this paper was deciphering the epigenetic regulation of the holocentromeres. I standardised the ChIP experiment for *Rhynchospora*, performed ChIP experiments for CENH3 in *R. pubera* and *R. breviuscula* and for H3K4me3 and H3K9me2 in *R. pubera*. The ChIP sequencing results were analysed by me. To understand the dynamics of these antibodies, I performed immunofluorescence in *Rhynchospora*. I contributed in writing and reviewing the manuscript.

Publication

Cell

Repeat-based holocentromeres influence genome architecture and karyotype evolution

Graphical abstract



Highlights

- Chromosome-scale genomes for holocentric plants with repeat-based holocentromeres
- Transition to holocentricity influenced the 3D (epi)genome architecture
- Regulation of repeat-based centromeres is conserved in mono- and holocentrics
- Chromosome fusions drive karyotype evolution and structural diploidization

Authors

Paulo G. Hofstatter, Gokilavani Thangavel, Thomas Lux, ..., Klaus F.X. Mayer, Andreas Houben, André Marques

Correspondence

amarques@mpipz.mpg.de

In brief

While most eukaryotes contain single regional centromeres, several plant and animal lineages assemble holocentromeres along the entire chromosome length. The assembly of chromosome-scale holocentric genomes with repeat-based holocentromeres from beak-sedges and their closest monocentric relative sheds light on important aspects of genome architecture and evolution influenced by centromere organization.

Hofstatter et al., 2022, Cell *185*, 3153–3168
August 18, 2022
https://doi.org/10.1016/j.cell.2022.06.045

Cell

CellPress
OPEN ACCESS

Article

Repeat-based holocentromeres influence genome architecture and karyotype evolution

Paulo G. Hofstatter,[1,12] Gokilavani Thangavel,[1,12] Thomas Lux,[2,12] Pavel Neumann,[3] Tihana Vondrak,[3,4] Petr Novak,[3] Meng Zhang,[1] Lucas Costa,[5] Marco Castellani,[1] Alison Scott,[1] Helena Toegelová,[6] Joerg Fuchs,[7] Yennifer Mata-Sucre,[5] Yhanndra Dias,[5] André L.L. Vanzela,[8] Bruno Huettel,[9] Cicero C.S. Almeida,[10] Hana Šimková,[6] Gustavo Souza,[5] Andrea Pedrosa-Harand,[5] Jiri Macas,[3] Klaus F.X. Mayer,[2,11] Andreas Houben,[7] and André Marques[1,13,*]

[1]Department of Chromosome Biology, Max Planck Institute for Plant Breeding Research, Cologne, NRW 50829, Germany
[2]Plant Genome and Systems Biology, German Research Center for Environmental Health, Helmholtz Zentrum München, Ingolstädter Landstraße 1, 85764 Neuherberg, Germany
[3]Biology Centre, Czech Academy of Sciences, Institute of Plant Molecular Biology, České Budějovice 37005, Czech Republic
[4]Faculty of Science, University of South Bohemia, České Budějovice 37005, Czech Republic
[5]Laboratory of Plant Cytogenetics and Evolution, Department of Botany, Centre of Biosciences, Federal University of Pernambuco, Recife, Pernambuco 50670-901, Brazil
[6]Institute of Experimental Botany of the Czech Academy of Sciences, Centre of Plant Structural and Functional Genomics, Olomouc 779 00, Czech Republic
[7]Breeding Research, Leibniz Institute of Plant Genetics and Crop Plant Research (IPK) Gatersleben, Seeland, Saxony-Anhalt 06466, Germany
[8]Laboratory of Cytogenetics and Plant Diversity, State University of Londrina, 86097-570 Paraná, Brazil
[9]Max Planck Genome-Centre Cologne, Max Planck Institute for Plant Breeding Research, Cologne, NRW 50829, Germany
[10]School of Agronomical Sciences, Campus Arapiraca, Federal University of Alagoas, Arapiraca 57309-005, Brazil
[11]School of Life Sciences Weihenstephan, Technical University of Munich, Alte Akademie 8, 85354 Freising, Germany
[12]These authors contributed equally
[13]Lead contact
*Correspondence: amarques@mpipz.mpg.de
https://doi.org/10.1016/j.cell.2022.06.045

SUMMARY

The centromere represents a single region in most eukaryotic chromosomes. However, several plant and animal lineages assemble holocentromeres along the entire chromosome length. Here, we compare genome organization and evolution as a function of centromere type by assembling chromosome-scale holocentric genomes with repeat-based holocentromeres from three beak-sedge (*Rhynchospora pubera*, *R. breviuscula*, and *R. tenuis*) and their closest monocentric relative, *Juncus effusus*. We demonstrate that transition to holocentricity affected 3D genome architecture by redefining genomic compartments, while distributing centromere function to thousands of repeat-based centromere units genome-wide. We uncover a complex genome organization in *R. pubera* that hides its unexpected octoploidy and describe a marked reduction in chromosome number for *R. tenuis*, which has only two chromosomes. We show that chromosome fusions, facilitated by repeat-based holocentromeres, promoted karyotype evolution and diploidization. Our study thus sheds light on several important aspects of genome architecture and evolution influenced by centromere organization.

INTRODUCTION

Most eukaryotes are monocentric, meaning that their centromeres are restricted to single regions on each chromosome. These centromeric regions can range from kilobases (kbs) to megabases (Mbs) in length and comprise often specific repeats (Gohard et al., 2014). Holocentromeres, by contrast, consist of multiple centromeric units distributed along the poleward surface of metaphase chromosomes, extending from one telomere to the other, and are thus typically visible as a line on each chromatid (Heckmann et al., 2013; Senaratne et al., 2021; Steiner and Henikoff, 2014). Holocentromeres are hypothesized to stabilize chromosomal fragments and fusions that favor karyotype rearrangements and speciation (Mandrioli and Manicardi, 2020), directly influencing chromosome evolution (Schubert and Lysak, 2011). This hypothesis is supported by the fact that holocentromeres have evolved independently several times in different plant and animal lineages (Escudero et al., 2016; Melters et al., 2012).

Aside from their function in cell division, centromeres have an evolutionarily conserved role in determining large-scale genome architecture and chromatin composition (Muller et al., 2019). Centromeres in monocentric chromosomes influence the distribution of genes, euchromatin- and heterochromatin-specific

post-translational histone modification domains, transposable elements (TEs), and meiotic crossovers (Fernandes et al., 2019; Fuchs et al., 2006; Muller et al., 2019; Naish et al., 2021). However, genome organization and chromatin composition of organisms with holocentric chromosomes is poorly understood, and it is likely that holocentric species differ markedly from the monocentric paradigm.

The beak-sedge *Rhynchospora pubera* (Cyperaceae, sedges) has repeat-based holocentromeres (Marques et al., 2015), as do other species from the same genus (Costa et al., 2021; Ribeiro et al., 2017). *R. pubera* holocentromeres are associated with a single tandem-repeat family (the centromeric 172-bp unit *Tyba* repeat) and the centromeric retrotransposon of *Rhynchospora* (*CRRh*), giving rise to thousands of small centromere units across the genome (Marques et al., 2015). The lack of a *Rhynchospora* reference genome has, however, hampered detailed studies about its intriguing centromere organization.

Here, we combined genomic and chromatin analyses to elucidate genomic adaptations related to different centromere organizations. We report the full characterization of a holocentric genome containing thousands of repeat-based centromere units. We show that this centromere organization influences the 3D genome architecture by redefining the extent of genomic compartments due to the lack of centromere clustering. Strikingly, despite substantial genome restructuring, the epigenetic regulation of centromere units in beak-sedges resembles that of monocentric centromeres, as in *Arabidopsis thaliana* (Naish et al., 2021). This observation suggests evolutionarily conserved epigenetic regulation of repeat-based centromeres in both monocentric and holocentric organisms. We further reveal that chromosome fusions facilitated by repeat-based holocentromeres reduce chromosome number and can act as an alternative to diploidization after genome doubling without the need for genome downsizing. Our work sheds light on the role of centromeres in overall genome organization and chromosome evolution.

RESULTS

Holocentricity affects spatial genome organization

To identify the genomic adaptations related to the transition to holocentricity, we constructed chromosome-scale reference genomes using PacBio HiFi sequencing and Dovetail Omni-C (DNase-based Hi-C) for three holocentric *Rhynchospora* species, *R. pubera* (n = 5; haploid nuclear genome size [1C] = 1.61 Gb), *R. breviuscula* (n = 5; 1C = 415 Mb), and *R. tenuis* (a plant with the fewest known chromosomes; n = 2; 1C = 394 Mb) (Castiglione and Cremonini, 2012; Vanzela et al., 1996), as well as their closest monocentric relative, the rush *J. effusus* (n = 21; 1C = 271 Mb) (Guerra et al., 2019; (Figures 1, 2, S1A, and S1B; Table S1; STAR Methods).

J. effusus showed a typical monocentric configuration of chromatin interaction within A (euchromatin) and B (heterochromatin) compartments, including some degree of a telomere-to-centromere axis (Figures 2A and 2B; see Hoencamp et al., 2021).

The concept of chromosome arms does not apply to holocentric species, as centromeres are ubiquitous. Consequently, we observed no large-scale compartmentalization or telomere-to-centromere axes, as evidenced by the chromatin configuration capture (Hi-C) contact matrices of our three *Rhynchospora* species (Figures 2C, 2D, S1A, and S1B). Further quantification of intrachromosomal (*cis*) and interchromosomal (*trans*) chromatin contacts revealed a significantly higher ratio ($p < 4.04e{-}05$) of *cis* versus *trans* interactions in all *Rhynchospora* species compared with the monocentric *J. effusus* (Figure S1C). Thus, holocentric beak-sedges are characterized by higher intrachromosomal spatial genome organization and lack of centromere clustering.

The distribution of genomic features differed markedly between holocentric *Rhynchospora* and monocentric *J. effusus* (Figures 2E and 2F). *Rhynchospora* had a uniform distribution of genes, transcriptional activity, *Tyba* centromeric repeats, TEs, and DNA methylation (Figures 2F, S1D, and S1E). By contrast, *J. effusus* genes were concentrated toward telomeric regions, while TEs and tandem repeats clustered toward centromeric regions (Figure 2E). Genome-wide gene distribution and transcriptional activity were positively correlated, while repeat distribution was positively correlated with overall DNA methylation levels (Figure 2E). Genome-wide CpG methylation (mCpG) was lower in *R. pubera* than in *J. effusus*, whereas CHG methylation was higher and CHH methylation was the same in both species (Figures S1F and S1G). Thus, transition to holocentricity likely affects 3D genome architecture by redefining the extents of genomic compartments and their relationships to each other.

Genetic and epigenetic composition of repeat-based holocentromeres

We analyzed the sequence organization and chromatin structure of the *Rhynchospora* repeat-based holocentromeres. The contiguity of our assemblies, coupled with the short array size of centromeric *Tyba* repeats, allowed us to resolve mostly complete *Tyba* arrays in the three *Rhynchospora* genomes. While total number and amount of *Tyba* arrays increased with chromosome size (Figures 3A and 3B), the density of arrays decreased (Figure 3C). Average array sizes of 20.3, 20.5, and 19.8 kb, and average spacing between two consecutive arrays of 368, 492, and 424 kb were found in *R. breviuscula*, *R. pubera*, and *R. tenuis*, respectively (Figures 3D and 3E). These results confirm a similar overall organization of centromeric *Tyba* repeats among the three *Rhynchospora* species. In common with monocentric centromeric repeats (Kasinathan and Henikoff, 2018), we also found a high frequency of dyad symmetries in the *Tyba* consensus sequences of all three *Rhynchospora* species (Figure 3F).

Chromatin immunoprecipitation followed by sequencing (ChIP-seq) confirmed the highest enrichment of centromeric histone H3 (CENH3) for the *Tyba* repeats and lower enrichment for *CRRh* throughout the entire *R. pubera* and *R. breviuscula* genomes (Figures 4A–4C and S1H; Table S2). We detected 2,753 and 995 CENH3-binding regions (hereafter CENH3 domains) evenly distributed across the five chromosomes of *R. pubera* and *R. breviuscula*, respectively. In both species, length, density, and spacing of CENH3 domains followed a similar pattern to the number of *Tyba* arrays detected (Figures 3A–3E). Considering that one CENH3 domain is equivalent to one centromere unit,

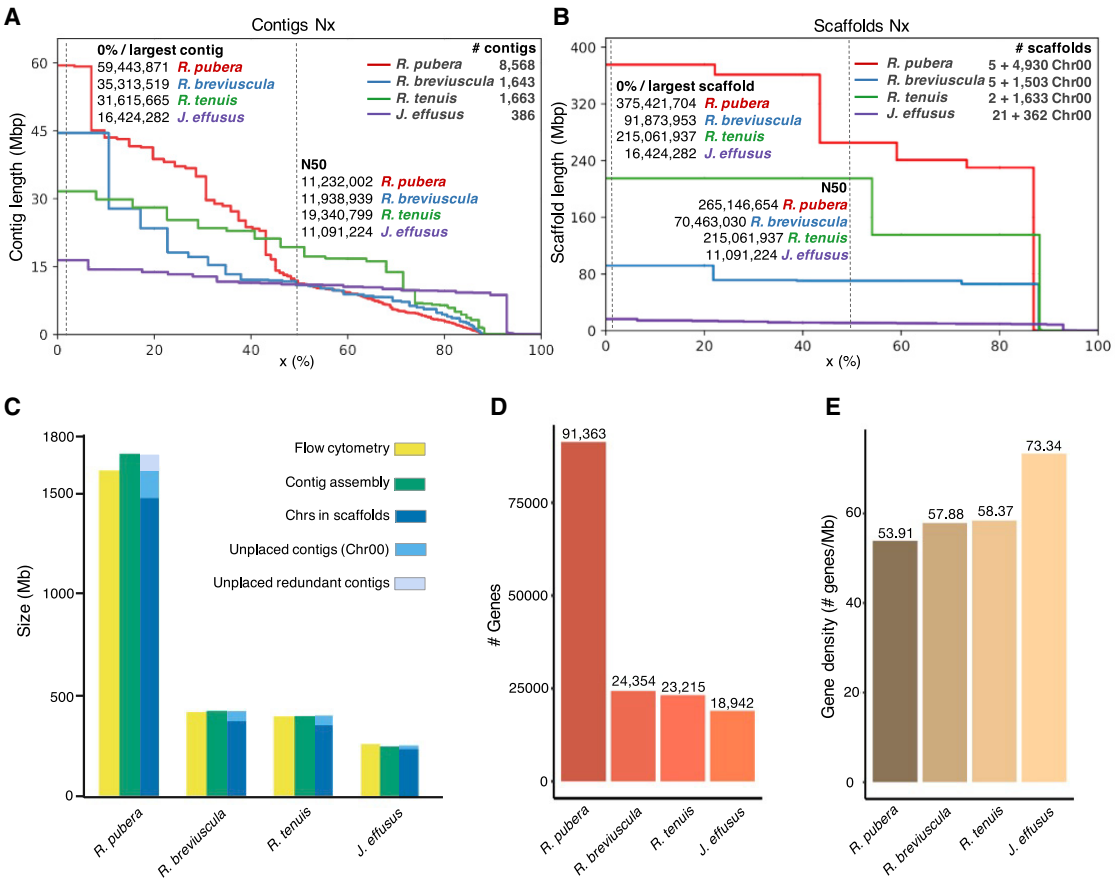


Figure 1. Summary of genome sizes, assemblies, scaffolding, and annotations

(A and B) Assembly (A) and final scaffolding (B) statistics.

(C) Comparison of estimated genome size and assembly and scaffolding sizes.

(D) Total number of high-confidence annotated genes.

(E) Gene density per Mb.

See also Table S1.

on average, each *R. pubera* chromosome carried 600 centromere units (1.88 domains/Mb), while the smaller chromosomes of *R. breviuscula* carried 200 centromere units each on average (2.69 domains/Mb) (Figures 3A–3C). Thus, genome/chromosome size may be negatively correlated with centromere unit density in beak-sedges. Genome-wide there was a significant association between CENH3 domains and *Tyba* repeats for both species ($p < 0.05$), confirming that *Tyba* repeats are the main CENH3-binding sites. Therefore, repeat-based holocentromeres are likely to be conserved and associated with *Tyba* repeats in beak-sedges.

In the monocentric *J. effusus*, the histone mark H3K4me3, which is euchromatin specific, showed dispersed labeling along chromosome arms, while H3K9me2 (heterochromatin specific) was concentrated at pericentromeric regions and co-localized with chromocenters in interphase (Figure S1I). By contrast, in the holocentric *R. pubera*, both euchromatin- and heterochromatin-specific histone marks were intermingled all along the chromosomes with a constant density even toward the subtelomeric and central chromosomal regions (Figures 4A–4C). Locally, H3K4me3 was mostly highly enriched at the promoter regions of protein-coding genes, whereas H3K9me2 was enriched on small heterochromatic islands, typically resembling TEs (Figure 4C). H3K4me3 was depleted at CENH3 domains, while H3K9me2 showed residual enrichment. We noticed a slight increase in H3K9me2 enrichment flanking CENH3 domains relative to the core region, mimicking the pericentromeric chromatin composition in monocentromeres (Figure 4C).

Irrespective of centromere type, gene bodies were highly enriched for mCpG in both *R. pubera* and *J. effusus*, with a sharp decrease at promoters and terminal regions. Methylation in the CHH and CHG contexts was much lower for the gene bodies than for intergenic regions (Figures 4D and 4E), as previously reported for other plants (Feng et al., 2020). Remarkably, despite the differences in chromosome organization, both the *Tyba* repeats in *R. pubera* and tandem repeats in centromeric regions of *J. effusus* chromosomes were highly enriched for mCpG at levels similar to those for TEs (Figures 4D and 4E). mCHG was sharply enriched flanking CENH3-binding regions in *R. pubera*, resembling the H3K9me2 pattern (Figure 4C). We obtained a

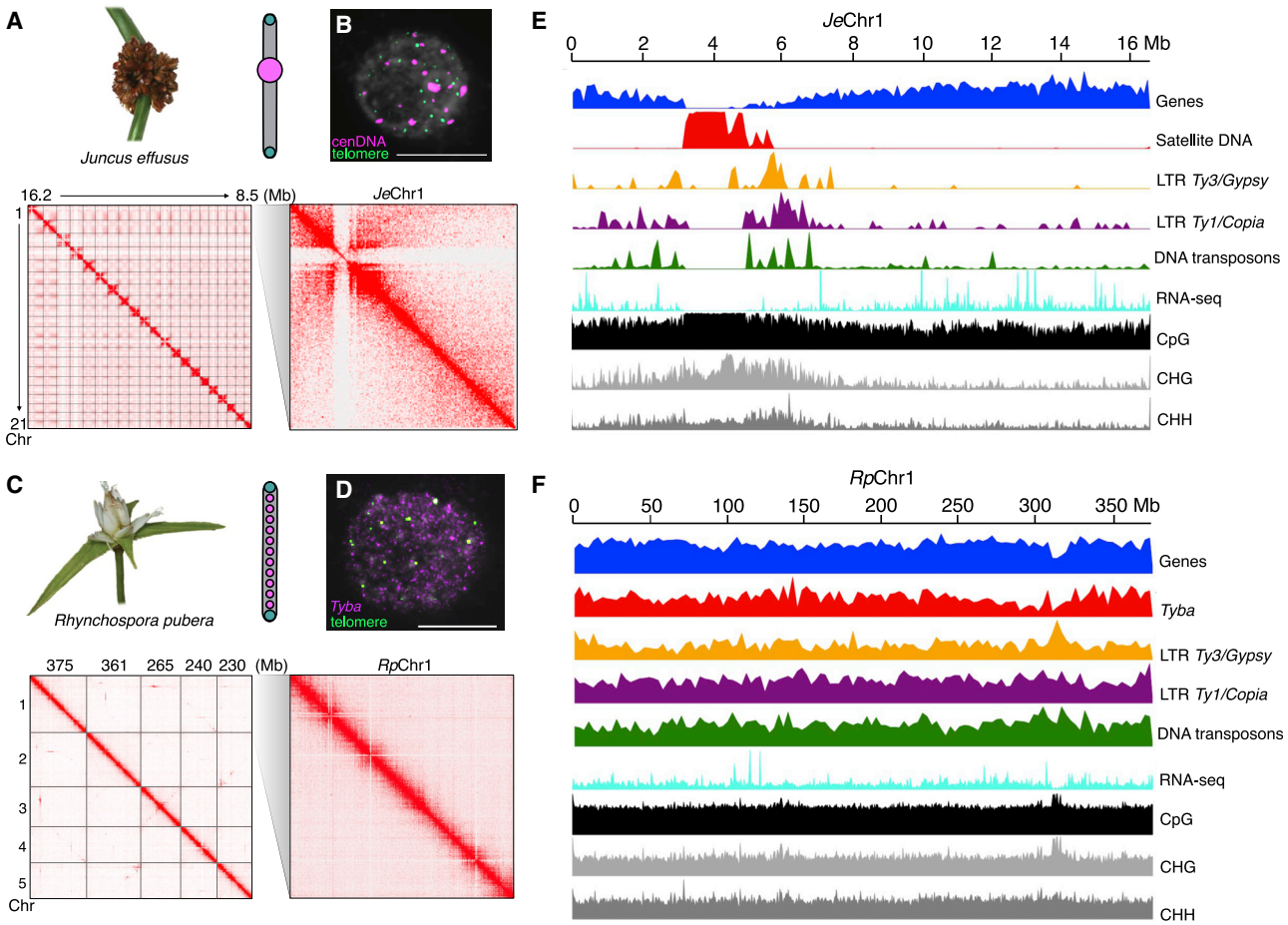


Figure 2. Spatial genome organization: monocentric versus holocentric chromosomes

(A) *J. effusus* (top left) genome contact map (bottom left) and chromosome 1 (*Je*Chr1) detailed view (bottom right). Centromere organization in monocentric chromosomes (top right).

(B) Interphase nucleus hybridized with DNA probes for the centromeric DNA (cenDNA) and telomeric sequence in *J. effusus*.

(C) *R. pubera* (top left) genome contact map (bottom left) and *Rp*Chr1 detailed view (bottom right). Centromere organization in holocentric chromosomes (top right).

(D) Interphase nucleus hybridized with DNA probes for the centromeric repeat *Tyba* and telomeric sequence in *Rhynchospora*.

(E) *Je*Chr1 and (F) *Rp*Chr1 detailed view showing the clustered (*Je*Chr1) and uniform (*Rp*Chr1) distribution of main genomic features, which are typical for monocentric chromosomes and holocentric chromosomes, respectively. Window sizes for sequence-type distribution density, 100 kb (*J. effusus*) and 3 Mb (*R. pubera*). Centromeres and telomeres in chromosome models are represented by magenta and green circles, respectively. Scale bars, 10 μm.

See also Figure S1.

similar pattern for mCHG at centromeric repeats in *J. effusus* (Figures 2D, 4D, and 4E). TEs showed the highest enrichment for mCHG and mCHH, while *Tyba* repeats displayed lower levels of mCHH, similar to genes (Figures 4D and 4E). Our results argue for the presence of a pericentromere-like chromatin state around the ends of centromere units in *Rhynchospora* that may mark the borders for CENH3 loading.

A typical centromere unit in *R. pubera* comprised a single *Tyba* array surrounded by genes and TEs (Figure 4F). We detected CENH3 domains all along the chromosomes, even in *Tyba* arrays located near telomeres like those at both ends of *R. pubera* chromosome 2 (*Rp*Chr2) (Figures 4F and 4G), confirming the telomere-to-telomere centromere activity of holocentric chromosomes. Notably, we observed an enrichment for H3K4me3 and actively transcribed genes close to centromere units, with an average distance of 6.3 kb (Figures 4H and 4I). We identified 313 genes that showed at least a 1-base-pair (bp) overlap with CENH3 domains. We even detected actively transcribed genes with typical H3K4me3 enrichment inside CENH3 domains (Figure 4H), a characteristic only rarely observed in monocentric organisms (Mizuno et al., 2011; Schotanus et al., 2021). Both CENH3 association and transcription were frequently reduced in genic regions inserted into centromere units, compared with genic regions residing outside the core centromere unit (Figure 4H), reflecting the precise regulation of chromatin composition of the *R. pubera* genome. *CRRh* was frequently inserted into *Tyba* arrays enriched for CENH3, but also H3K9me2 and some level of H3K4me3, suggesting a different epigenetic regulation

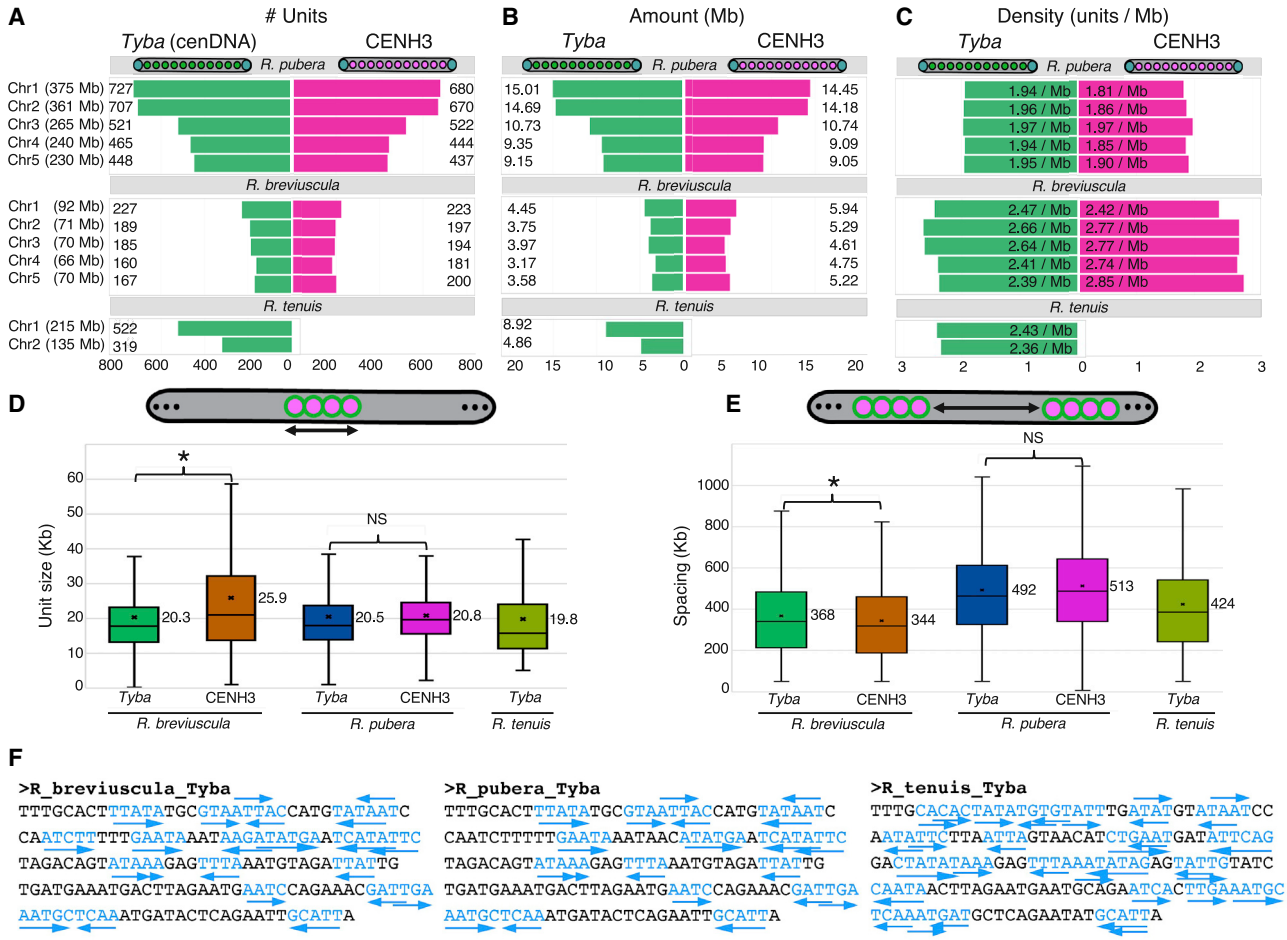


Figure 3. Features of *Tyba* centromeric DNA and CENH3 domains among *Rhynchospora* species

(A) Total number per chromosome of annotated *Tyba* arrays and CENH3 domains.

(B) Total amount of bases associated with *Tyba* arrays and CENH3 domains.

(C) Density of *Tyba* arrays and CENH3 domains per chromosome.

(D) Size distribution of *Tyba* arrays and CENH3 domains. NS = not significant.

(E) Spacing between two consecutive centromere arrays/domains among *Rhynchospora* species. Asterisks indicate Dunn's test, $p < 0.05$.

(F) Patterns of DNA dyad symmetry in the *Tyba* consensus sequences of the three *Rhynchospora* species.

of this retroelement compared with *Tyba* repeats (Figure 4I). Our results thus point to fine-scale epigenetic regulation of genomes with repeat-based holocentromeres.

Transposition partially explains genome-wide *Tyba* dispersal and expansion

Tyba repeats in *R. pubera* can be flanked by *TCR1* and *TCR2* repeats, suggesting that some *Tyba* arrays are part of larger repetitive elements (Marques et al., 2015). The consensus full-length *TCR1* element contained a *Tyba* array with a 5′ sequence of approximately 4.8-kb and a 136-bp 3′ sequence. The element possessed no open reading frame (ORF) and lacked terminal repeats, and its 5′ and 3′ ends harbored the ATC and CTAGT sequence motifs, respectively, suggesting that *TCR1* is a nonautonomous *Helitron* TE (Thomas and Pritham, 2015), from the same family as a fully autonomous *Helitron* element (*Helitron-27*) in the *R. pubera* genome. Despite sharing conserved terminal sequences, *TCR1* and *Helitron-27* exhibited no similarities in their internal regions. We identified three intact copies of the autonomous *Helitron-27* in the genome with high mutual similarity (Table S3), each encoding a full *Helitron* helicase (1,340 amino acids), indicating that *TCR1* and *Helitron-27* elements are still capable of transposition. We further identified an additional 322 full-length elements (Table S3) with both *TCR1* termini as well as *Tyba* and another 146 partial elements with the 3′-terminal sequence and containing *Tyba* within the upstream 500-bp region. We conclude that at least 468 *Tyba*-containing loci in the genome resulted from the transposition activity of *TCR1* elements. The full-length *TCR1* sequences were 6.9–49.6 kb (24.8 kb average), containing 1.2–31.3 kb from *Tyba* (15.7 kb average). In many *TCR1* elements, *Tyba* arrays were split into multiple segments due to insertions of other sequences, showing that multiple *Tyba* loci can originate from a single *TCR1* insertion (Figures 4J and 4K). Importantly, a comparison of *TCR1* and

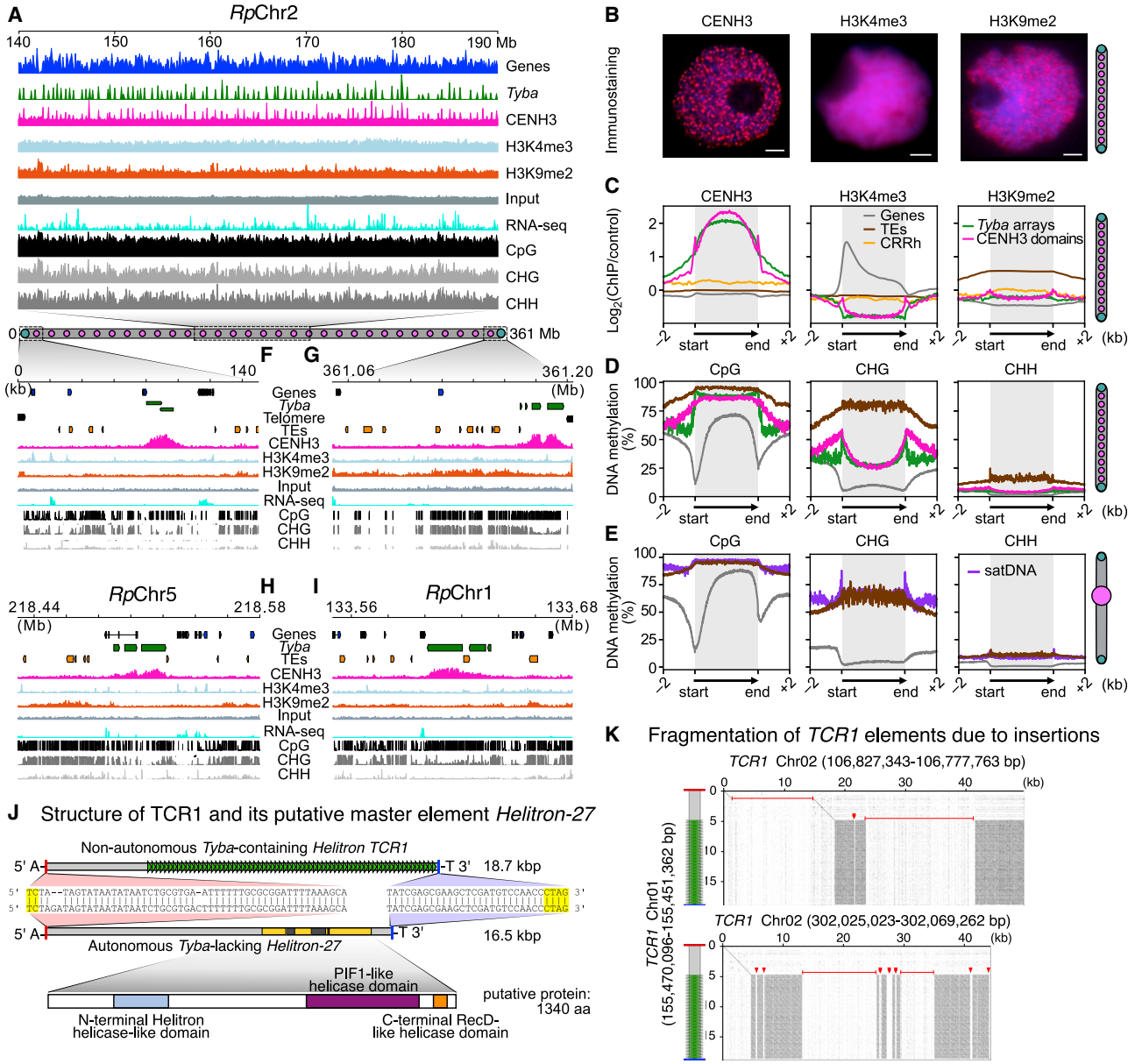


Figure 4. Genetic and epigenetic composition of repeat-based holocentromeres in *R. pubera*

(A) Zoomed-in view of *Rp*Chr2 showing a 50-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) Immunostaining of *R. pubera* interphase nuclei for CENH3, H3K4me3, and H3K9me2. Scale bars, 2 μm.

(C) Enrichment of CENH3, H3K4me3, and H3K9me2 from the start and end of different types of sequences: genes (gray line), TEs (brown), *CRRh* (yellow), *Tyba* repeats (green), and CENH3 domains (magenta). ChIP-seq signals are shown as $\log_2$ (normalized RPKM ChIP/input).

(D and E) Enrichment of DNA methylation in the CpG, CHG, and CHH contexts for the same sequence types as shown I(C) for *R. pubera* (D) and *J. effusus* (E), genes (gray line), satDNA (purple) and TEs (brown). Gray boxes in (C)–(E) highlight the modification enrichment over the body of each sequence type.

(F and G) Close-up view of the first (F) and last (G) centromere units of *Rp*Chr2, which are composed of a *Tyba* repeat array very close to the telomere and showing the typical CENH3 enrichment.

(H) A centromere unit where an active gene is intermingled with the *Tyba* repeat.

(I) A *Tyba* array showing an insertion of the centromeric retrotransposon *CRRh* and CENH3-binding activity.

(J) Structures of the typical nonautonomous *TCR1* element (Chr01:155470096–155451362) and its likely master element *Helitron-27* (Chr05:40901972–40918485). Similarities between *TCR1* and *Helitron-27* are mostly restricted to the terminal sequences. The 5′- and 3′-terminal sequences are in red and blue,

(legend continued on next page)

CENH3 domains revealed that the vast majority (98.7%) of full-length *TCR1* elements are embedded within or overlap with the centromere units (Table S2).

*Helitron*s with boundaries similar to *TCR1/Helitron-27* were present in *R. tenuis* and *R. breviuscula*; however, all but one of the full-length elements in these two genomes lacked *Tyba*. The sole exception was a single element from *R. breviuscula* (Chr1:69162288–69195619) with 5′ and 3′ boundary sequences characteristic of this *Helitron* family as well as a *Tyba* array; however, the remaining sections lacked any similarity to the *TCR1* of *R. pubera*. These results suggest that *Tyba* was amplified as a part of a *TCR1 Helitron* only in the genome of *R. pubera*.

The *TCR2* element was found to be a miniature inverted-repeat TE (MITE) and ranged from 672 to 1,235 bp, likely originated from the DNA transposon MuDR with shared similarity (up to 97%) in the terminal inverted repeats. All 158 full-length *TCR2* elements identified in the *R. pubera* genome were in *Tyba* arrays, but none were characterized by *Tyba* insertions. Thus, *TCR2* elements did not contribute to the dispersal of *Tyba* in the *R. pubera* genome.

R. pubera is a cryptic auto-octoploid with n = 5 chromosomes

The *R. pubera* genome is 4 times larger than that of its closely related species, despite sharing the same ancestral chromosome number (ACN) (x = 5) (Burchardt et al., 2020; Ribeiro et al., 2018) (Figure S2A). One explanation for this pronounced genome expansion would be a sudden and massive proliferation of repeat elements. However, we observed no accumulation of repeats when comparing repeat abundance profiles among closely related *Rhynchospora* species (Figure S2A). Thus, a different process must be responsible for the large genome size in *R. pubera*.

Completeness assessment of the *R. pubera* genome by calculating the benchmarking universal single-copy orthologs (BUSCOs) score revealed a surprisingly high level of gene duplications (96.0% duplicated BUSCOs) (Figure S2B). Annotation of the genome yielded far more high-confidence gene models (91,363) in *R. pubera* compared with the other species (Figures 1D and 1E; Table S1), confirming the high level of gene duplication (Figure S2C). Self-synteny analysis revealed that the *R. pubera* genome comprises two large syntenic blocks in four copies across the five chromosomes (Figure S3A). The larger syntenic block, named Block1, corresponded to the entire *Rp*Chr4 and *Rp*Chr5 and contributed to a large fraction of both *Rp*Chr1 and *Rp*Chr2. We identified the smaller block, named Block2, twice in an inverted arrangement in *Rp*Chr3, as well as in *Rp*Chr1 and *Rp*Chr2 (Figure S3A).

The distribution of synonymous substitutions per synonymous site (Ks) for coding sequences over the intragenomic syntenic blocks in *R. pubera* had a large peak indicative of recent and successive whole-genome duplication (WGD) events. An additional small peak was also observed, indicating an ancient WGD (Figure S3B). By filtering out the sequences showing the lowest Ks values, we determined that Block1 from *Rp*Chr1 shows higher sequence identity to *Rp*Chr4, which we renamed Block1A1 and Block1A2, respectively. Similarly, Block1 from *Rp*Chr2 showed higher sequence identity to *Rp*Chr5, which were thus named Block1B1 and Block1B2, respectively (Figures S3B and S3C). We confirmed the relationships of the four Block1 copies by comparative phylogenetic analysis (Figure S3D). A similar analysis of Block2 copies was inconclusive (Figures S3C and S3E). Using k-mer analysis, which provides information on genome size, ploidy, and genome structure through scrutiny of heterozygous k-mer pairs (Ranallo-Benavidez et al., 2020), we detected a higher incidence of homozygous and duplicated k-mers, favoring an autopolyploidy genome model for *R. pubera* (Figures S3F and S3G). Importantly, this analysis accurately determined the diploid heterozygous state of *R. breviuscula* and *R. tenuis* (Figure S4). Thus, *R. pubera* has an auto-octoploid genome shaped by two rounds of genome doubling explaining its large genome size. Post-polyploid genome shuffling events considerably reduced the chromosome number to n = 5.

Chromosome fusions explain karyotype evolution in beak-sedges

To explore the genome duplications seen in *R. pubera*, we compared its genome with its close relative *R. breviuscula*, which has the same chromosome number but a genome that is one-quarter the size (415 Mb) (Figure S2A). Assessment of the *R. breviuscula* genome revealed a high level of completeness, with a BUSCO score of 98.3%, and little gene duplication (2.1%) (Figures S2B and S2C), confirmed by the absence of self-synteny (Figure S1D). Gene annotation yielded 24,354 high-confidence gene models (Figures 1D and 1E; Table S1), 4 times fewer than in the *R. pubera* genome, as expected. Synteny analysis between both genomes illustrated how each *R. breviuscula* chromosome (*Rb*) is present in four copies in the *R. pubera* genome (Figure 5A). Remarkably, *Rp*Chr1 and *Rp*Chr2 contained all five putative *Rb*s in end-to-end configurations. *Rp*Chr3 contained *Rb*3 and *Rb*4 copied twice in an inverted order, comprising Block2, while *Rp*Chr4 and *Rp*Chr5 contained *Rb*1, R*b*2, and *Rb*5, comprising Block1A and Block1B, respectively (Figure 5A). Thus, *R. breviuscula* likely conserved the ancestral karyotype, while *R. pubera* restored the ACN (x = 5) of its clade due to descending dysploidy, which was mediated by a complex chain of chromosome fusions, e.g., end-to-end fusions (EEFs), with 15 EEF junctions detected. Remarkably, each chromosome pair had a unique combination of ancestral chromosomes. We conclude that descending dysploidy involving a unique combination of chromosomes may be a strategy to avoid meiotic pairing issues that could potentially arise from autopolyploidy, thereby acting as a rapid route to diploidization facilitated by holocentricity.

respectively. Yellow, conserved *Helitron* sequence motifs in the alignment of *TCR1* and *Helitron-27* terminal sequences. Light gray, noncoding regions. Green triangles, *Tyba* array in *TCR1*. Yellow and dark gray, putative exons and introns in the *Helitron-27* coding region, respectively.

(K) Dot-plot comparison of a typical *TCR1* element (vertical sequences) with two other elements (horizontal sequences) that have insertions of *TCR1*-unrelated sequences marked as red lines and triangles.

See also Figure S1 and Tables S2 and S3.

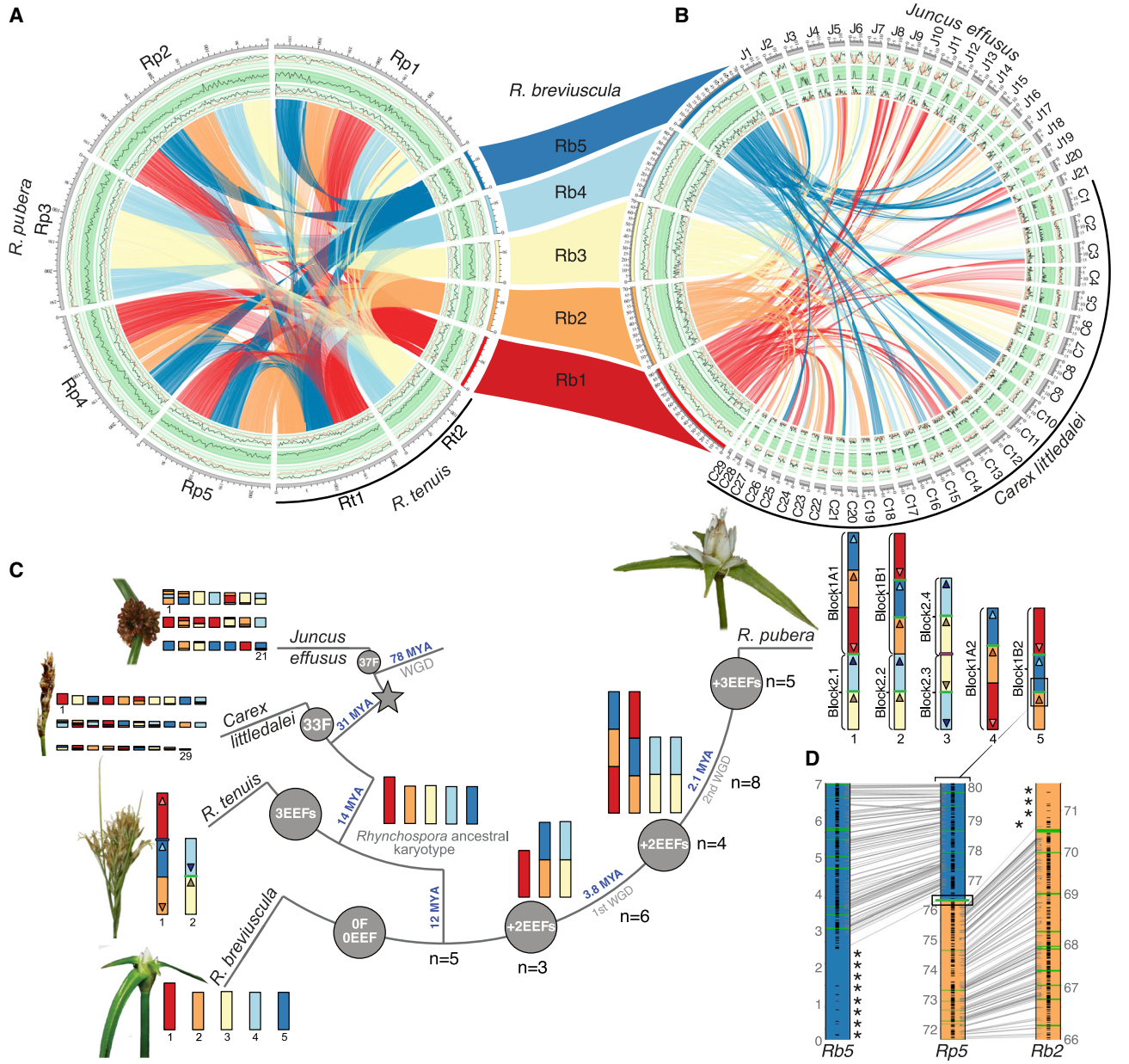


Figure 5. Genome organization and evolution of sedges and common rush

(A and B) Circos plots of *R. breviuscula* synteny to *R. pubera* and *R. tenuis* (A) and circos plots of *R. breviuscula* synteny to *J. effusus* and *C. littledalei* (B). Tracks from outside to inside: 1. genes (black line) and TEs (red line), 2. *Tyba*/tandem repeats (black line), and 3. LTR *Ty1/Copia* (black line) and *Ty3/Gypsy* (red line) retroelement distribution. Distribution of the main sequence classes was calculated in 3-Mb windows for *R. pubera*, *R. tenuis*, and *R. breviuscula* (A), in a 1-Mb window for *R. breviuscula*, and in 500-kb windows for *C. littledalei* and *J. effusus* (B).

(C) Karyotype evolution and synteny conservation in sedges and common rush. Transition to holocentricity is indicated by a star. Hypothetical ancestral karyotype for *Rhynchospora* based on the simplest karyotype of *R. breviuscula* illustrates frequent end-to-end fusions (EEFs) in beak-sedges. For reconstruction of karyotype evolution in *R. pubera* see also Figures S4 and S5. Arrow heads, orientation of the *R. breviuscula* chromosomes in the *R. pubera* and *R. tenuis* ideograms. For both *J. effusus* and *C. littledalei*, ideograms indicate the syntenic blocks to *R. breviuscula* chromosomes. Numbers of putative EEFs or fission (F) events necessary to transform the hypothetical *Rhynchospora* ancestral karyotype into the extant genomes are within the gray circles. Repeat sequences at the junctions between *Rb* blocks are indicated by colored bars (Tyba, green; rDNA, purple; telomeric DNA, blue) in *R. tenuis* and *R. pubera* ideograms.

(D) *R. pubera* Chr5 showing a *Tyba* array (black rectangle) at the junction between syntenic Rb2 and *Rb*5 blocks. Synteny from *Rp*Chr5 to Rb2 and *Rb*5 stops close to the last *Tyba* array, which is followed by a gene-poor, TE-enriched region, mainly LTR *Ty3/Gypsy* of the *Athila* clade (indicated by asterisks) that are frequently within *R. breviuscula* subtelomeric regions but absent in the fused chromosomes. Genes and *Tyba* arrays are annotated as black stripes and green lines, respectively.

See also Figures S2, S3, S4, S5, S6, and Table S4.

Because we detected several EEFs in multiple copies in *R. pubera*, we assessed whether they were derived from the same rearrangement or if they arose from multiple independent events. All duplicated EEFs in the *R. pubera* genome, e.g., Rb2/*Rb*5, *Rb*3/*Rb*4, *Rb*1/*Rb*2, and *Rb*1/*Rb*5 EEFs, share a fusion signature involving the same regions. This observation suggested that the Rb2/*Rb*5 and *Rb*3/*Rb*4 EEFs, which are present 4 times in the *R. pubera* genome emerged only once—before the first WGD event (Figures 5C and S5). The *Rb*1/*Rb*2 and *Rb*1/*Rb*5 EEFs, which were found twice, likely emerged after the first WGD event. Finally, we found the *Rb*1/*Rb*4, Rb2/*Rb*4, and *Rb*3/*Rb*3 EEFs only once, suggesting that they occurred after the second WGD (Figures 5C and S5).

The *Rb*3/*Rb*4 EEF, which forms Block2 in the *R. pubera* genome, was likely maintained as a duplicated fused chromosome after the first WGD, which might have allowed a longer period of tetrasomic inheritance. This hypothesis might explain the fact that the sequences of the four copies from Block2 cannot be distinguished from each other, in contrast to Block1.

We attempted to date the duplication events using a set of conserved genes shared among the four copies of Block1, which revealed the first WGD event as occurring around 3.8 million years ago (Mya) followed by a second WGD event around 2.1 Mya (Figure S3D). Based on this analysis, we deduced the origin and evolution of the *R. pubera* karyotype (Figure 5C). These results further support an autopolyploid origin for *R. pubera* and confirm a short interval between the two rounds of WGDs, indicating rapid chromosome number reduction in this species.

We carried out a number of analyses to determine the origin of the reduced karyotype in *R. tenuis* (n = 2). BUSCO analysis of its genome revealed high completeness (98.5% against the viridiplantae_odb10 dataset) and little duplication (3.7%) (Figures S2B and S2C). Gene annotation yielded 23,215 high-confidence gene models (Figures 1D and 1E). The absence of self-synteny in the *R. tenuis* genome ruled out large duplications (Figure S2E). Synteny comparison between *R. tenuis* and *R. breviuscula* genomes showed that again all *Rb*s were present in simple end-to-end configurations in the *R. tenuis* genome, explaining its karyotype by descending dysploidy from n = 5 to n = 2 (Figure 5A). Strikingly, we observed similar associations of syntenic *Rb* blocks as found in Block1 and 2 in both *R. pubera* and *R. tenuis*, where *Rt*Chr1 resembled Block1B and was composed of Rb2, *Rb*5, and *Rb*1, while *Rt*Chr2 resembled Block2, consisting of *Rb*3 and *Rb*4 (Figure 5A). However, the orientation of chromosome ends involved in the EEFs differed in the two instances, suggesting that the EEFs occurred independently (Figure 5C).

Despite their high chromosome number and centromere-type differences, *J. effusus* and the previously available genome for the sedge *Carex littledalei* (its homotypic synonym, *Kobresia littledalei*) (Can et al., 2020) showed a typical diploid gene content and no evidence of any recent WGD, outside of a shared ancient WGD between sedges and rushes (Figure S4). The *J. effusus* genome also revealed high completeness (100% viridiplantae_odb10 dataset) and little duplication (1.6%) (Figures S2B and S2C). Annotation of its genome yielded 18,942 high-confidence gene models (Figures 1D and 1E; Table S1). Synteny analysis further revealed that most *J. effusus* and *C. littledalei* chromosomes are present as highly collinear blocks across the five chromosomes of *R. breviuscula*, suggesting a high conservation of synteny although the group is ancient (78 Mya) (Figures 5B and 5C). Thus, neither the low nor the high chromosome numbers observed in many holocentric species necessarily reflect the absence or presence of recent polyploidy, respectively, and these numbers should be interpreted with caution in the absence of detailed genomic studies.

Tyba repeats are frequently present at the junctions of end-to-end fusions

TEs can influence chromosomal rearrangements (Lonnig and Saedler, 2002). To assess their possible role in the EEFs observed in *Rhynchospora* genomes, we looked for enrichment of specific repeats at the ends of *Rb*s and near the junctions of EEFs in *R. pubera* and *R. tenuis*. We detected a high density of TEs in almost all subtelomeric regions of *Rb*s. These repeat-rich regions varied from 500 kb to 3 Mb in size, were mainly enriched for an LTR *Ty3*/*Gypsy* element of the *Athila* clade, were poorly enriched for genes, and lacked *Tyba* repeats (Figures 5A, 5B, and S1D). Notably, the *R. breviuscula* subtelomeric repeat-rich regions were largely missing at the junctions of fused chromosomes in both *R. pubera* and *R. tenuis* (Figures 5D and S6). Remarkably, we detected *Tyba* repeats exactly at the EEF junctions in 10 out of the 15 EEFs of *R. pubera*, while we observed a small 45S rDNA remnant array (with only five 18S-5.8S-26S units) in one EEF (Figures 5C, 5D, and S6A–S6G). In *R. tenuis*, we also identified a *Tyba* repeat array in one out of three EEF junctions (*Rb*3/*Rb*4 junction on *Rt*Chr2), while an interstitial telomeric site (536 bp) was detected at the *Rb*5/*Rb*1 junction (Figures 5C and S6H).

Emergence and loss of CENH3 domains related to *Tyba*

We used the duplicated genome copies of *R. pubera* to study cases of paralogous CENH3 domains and *Tyba* arrays. Of 660 groups of paralogous regions, 66% of the CENH3 domains were present in all four copies (Table S5). We also identified 50 groups of paralogous regions in which the CENH3 domain was lost in one of the paralogs. Most cases (88%) were associated with *Tyba* loss (Figures 6A and S7A) or a reduced size of the *Tyba* region in loci devoid of CENH3 signal compared with their paralogous regions bound by CENH3. We observed the likely inverse event, i.e., the gain of a new CENH3 domain, in groups of paralogous regions where we only identified the CENH3 domain in only one of the four paralogous regions. In the newly acquired CENH3 domain, there was either a new *Tyba* insertion, likely due to a new insertion of *TCR1* (Figures 6B and S7B), a *Tyba* expansion, or the insertion of a new TE (which was most frequent). However, the ChIP/input ratios within these potentially new CENH3 domains containing a new TE insertion (1.5) were significantly lower ($p < 2.2e{-}16$) than the ChIP/input ratios in potentially new CENH3 domains associated with a new insertion of a *Tyba* element (4.1).

DISCUSSION

Here, we report high-quality and contiguous chromosome-scale reference genomes for three species with repeat-based holocentromeres, *R. pubera*, *R. breviuscula*, and *R. tenuis*, and their closest monocentric relative, *J. effusus*. These newly assembled genomes provide a valuable resource for comparative biology

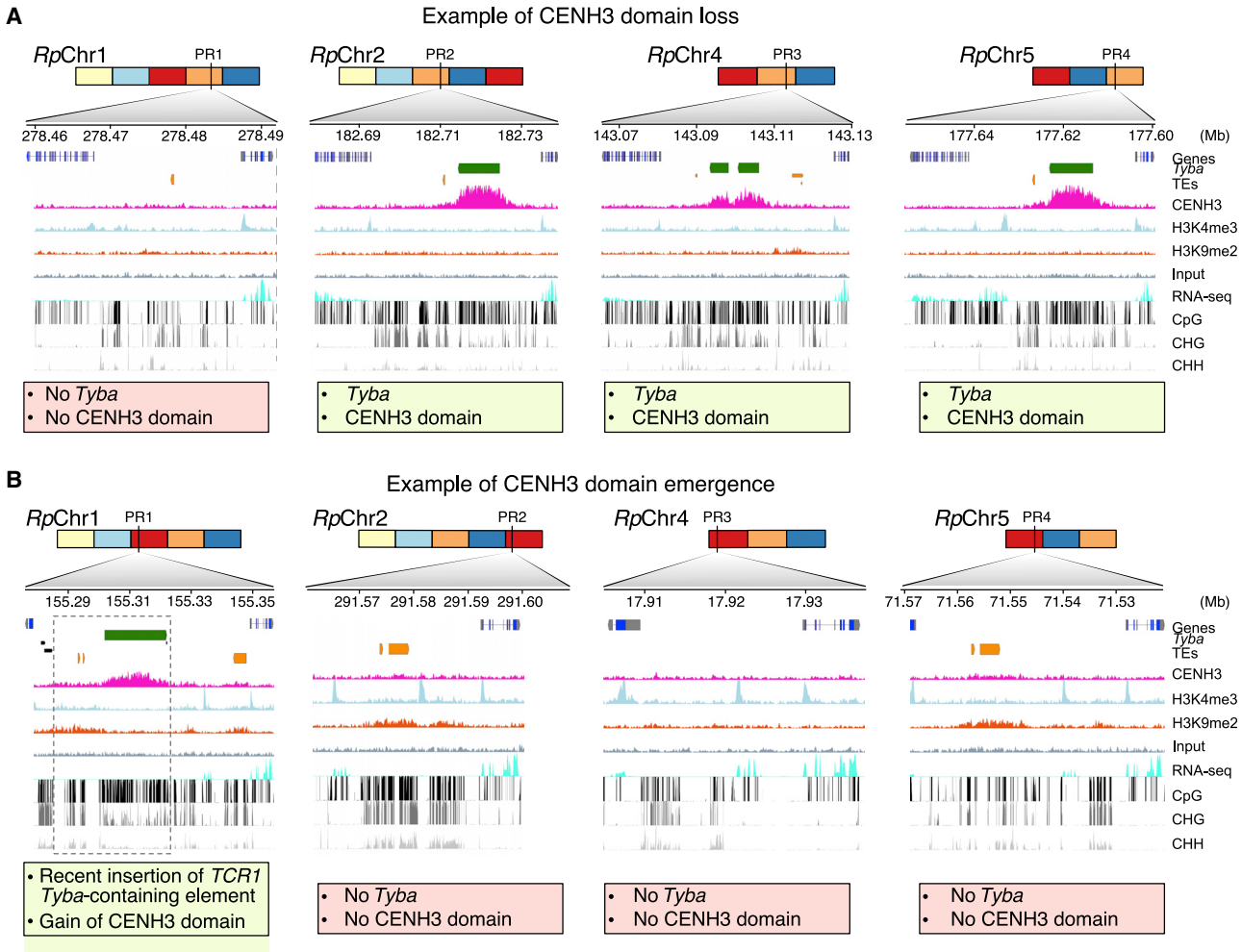


Figure 6. Emergence and loss of CENH3 domains in *R. pubera*

(A) CENH3 domain with *Tyba* array loss in one of the four paralogous regions, while the other three copies retain the *Tyba* array. Zoomed-in view of all four regions demonstrates the CENH3 domain loss only in the *Rp*Chr1 copy.

(B) CENH3 domain with *Tyba* array gain in one of four paralogous regions due to a transposition of a *Tyba*-containing *TCR1* in *Rp*Chr1, while the other three copies lack the *Tyba* array. The gained locus is indicated by the dashed box. Zoomed-in view of all four regions demonstrates the acquisition of a new CENH3 domain only in the *Rp*Chr1 copy. PR, paralogous region. Note that the four copies shared a similar chromatin composition.

See also Figures S5 and S7 and Table S5.

and studies related to genome adaptation to different centromere types.

Repeat-based holocentromeres influence genome organization and regulation

Repeat-based holocentromeres in beak-sedges comprise small islands (20–25 kb) of centromeric *Tyba* repeats, in which high mCpG, low H3K9me2, and depletion of H3K4me3 distinguish them from other holocentric genomes with and without repeat-based holocentromeres (Cortes-Silva et al., 2020; Despot-Slade et al., 2021; Nhim et al., 2022; Steiner and Henikoff, 2014). The association levels of H3K9me2 and mCHG at the core (low) and flanking (high) centromere units in *R. pubera* are strikingly similar to the recently reported *A. thaliana* centromeres (Naish et al., 2021). We also observed a similar pattern of mCHG methylation in monocentric *J. effusus*. Heterochromatinization of pericentromeres appears to be important for stabilizing the centromeric core, by preventing recombination between core repeats and stopping the spread of CENH3 into adjacent regions (Achrem et al., 2020; Wong et al., 2020). Thus, despite substantial genome restructuring, the epigenetic regulation of centromere units in beak-sedges resembles that in monocentric centromeres. This observation suggests an evolutionarily conserved epigenetic regulation of repeat-based centromeres in both mono- and holocentric organisms. We observed active genes close to and even within centromere units, which, although rare, is likely only possible with a plastic regulation of euchromatic and heterochromatic boundaries. We hypothesize that *R. pubera* achieves such a feat with a fine-scale epigenetic regulation of centromere units (Figures 7A and 7B).

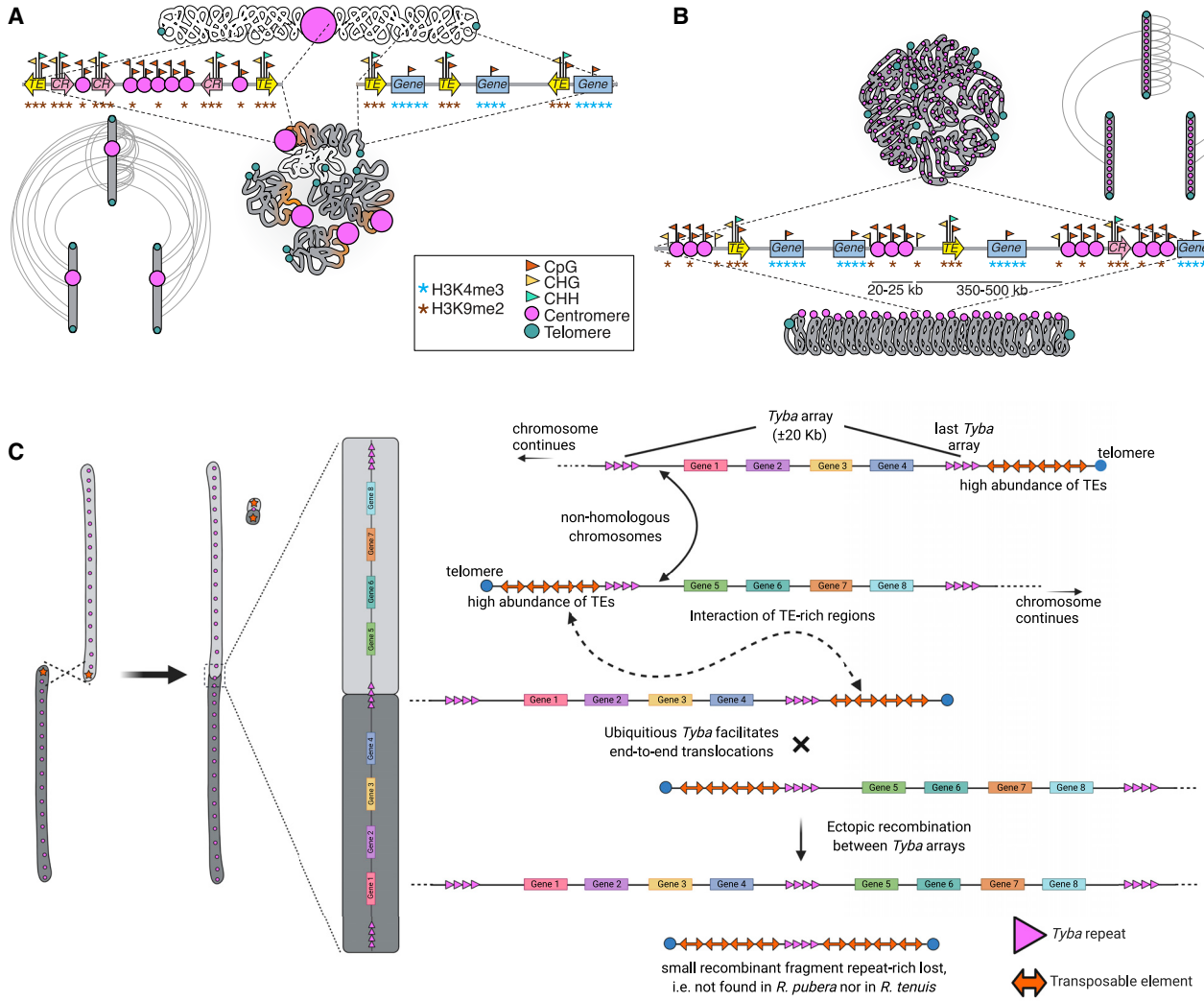


Figure 7. Genome organization in monocentric versus holocentric chromosomes and proposed model for end-to-end fusions

(A) Typically, in a monocentric chromosome, compartments of more compacted and silenced chromatin states extend along large megabase-long regions around centromeres and pericentromeres, while genes concentrate at subtelomeric regions. A telomere-to-centromere axis is frequently observed in genome contact maps in monocentric species due to the clustering of centromeres and telomeres, which increases the rate of interchromosomal chromatin contacts.

(B) *Rhynchospora* holocentric genome revealed uniform deposition of epigenetic marks at the macro scale and fine epigenetic regulation of repeat-based centromere units and silenced and active chromatin states at the micro scale. The regular spacing between centromere units (350–500 kb) appears to be the distance necessary to loop the chromatin back, aligning centromere units (20–25 kb) at the outer surface of the condensed chromosome. A telomere-to-centromere axis is absent in genome contact maps in holocentric species due to the lack of centromere clustering, affecting the spatial genome organization and decreases the rate of interchromosomal chromatin contacts. The model represents intra-/interchromosomal contacts among three different monocentric (A, bottom left) and holocentric (B, upper right) chromosomes.

(C) Possible mechanism for the involvement of centromeric *Tyba* repeats in end-to-end fusions (EEFs). Interaction of highly repetitive regions close to the telomere could facilitate ectopic recombinations of *Tyba* repeats.

Centromere units are regularly spaced (350–500 kb) in the *Rhynchospora* genomes, instead of randomly distributed. This specific spacing might point to a selection mechanism for establishing centromere units separated by an optimal spacing required to fold the chromatin during cell-cycle progression and for the recruitment of CENH3-positive nucleosomes to build the line-like holocentromere at metaphase (Figure 7B). *In silico* modeling based on polymer simulations of chromatin folding in holocentric chromosomes suggests that centromere units can act as anchors of loop extruders, facilitating the formation of line-like holocentromeres during chromosome condensation (Câmara et al., 2021).

A mechanism for the formation of repeat-based holocentromeres

The repeat-based holocentromeres of the *Rhynchospora* species analyzed here are almost exclusively composed of *Tyba* repeats. We cannot conclude from the available data whether the

accumulation of such repeats triggered the transition to holocentricity or whether CENH3 spreading preceded and/or also facilitated the subsequent expansion of holocentromeric repeats. However, we did demonstrate that a portion of *Tyba* arrays in *R. pubera* emerged in the genome as a result of the amplification of *TCR1*-type *Helitron*s and that most (98.7%) full-length *TCR1* elements possessing *Tyba* are associated with CENH3-bound chromatin. This either indicates that centromere units are at least partially determined genetically by the nucleotide sequence of *Tyba* or that *TCR1* transposition involves the transfer of epigenetic centromere marks, e.g., CENH3, which remain associated with the new copy of the element. The presence of tandem repeats within *Helitron*s is common in both plants and animals (Thomas and Pritham, 2015), but *TCR1* is unique because it possesses a centromeric satellite. The lack of *Tyba* repeats in *TCR1*-related elements in *R. breviuscula* and *R. tenuis* suggests that *Tyba* was captured by *TCR1* after the ancestors of *R. pubera* and the two other *Rhynchospora* species diverged.

It is conceivable, however, that the amplification and dispersal of *Tyba* occurred via mobilization by TEs earlier in the evolution of *Rhynchospora* and that the signatures of such events have long since been lost due to the accumulation of mutations, insertions and deletions, and DNA rearrangements. We also observed such changes in many *TCR1* loci identified in the genome of *R. pubera* that contained either truncated *TCR1* elements or full-length elements with nested insertions of other sequences (Figures 4K and 4L). The existence of a single *TCR1/Helitron27*-related element in *R. breviuscula* that possesses *Tyba* but lacks overall similarity to *TCR1* suggests that *Tyba* capture by *Helitron*s occurs recurrently in the evolution of *Rhynchospora* species and may result in waves of *Tyba* amplification via *Helitron* transposition.

The effect of holocentricity on karyotype evolution and diploidization

Our results are consistent with dysploidy as the main driver of karyotype evolution in holocentric organisms (Guerra, 2016; Mayrose and Lysak, 2021), where strong descending dysploidy restored the ACN ($x = 5$) in *R. pubera* and reduced the chromosome number in *R. tenuis*. In both cases, the same ancestral chromosomes were fused independently either without (*R. tenuis*) or following WGDs (*R. pubera*). Such tolerance of extensive chromosomal rearrangements seems to underlie rapid karyotype evolution, eventually leading to chromosomal speciation (Lucek et al., 2022; Lukhtanov et al., 2018).

Robertsonian translocations and chromosome fusions leading to descending dysploidy have been reported in some holocentric butterflies (Cicconardi et al., 2021; Hill et al., 2019). However, the incidence of EEFs as the sole mechanism of descending dysploidy in *Rhynchospora* is intriguing. Remarkably, meiotic pairing and segregation are not disturbed in the *R. pubera* genome (Marques et al., 2016), suggesting that selection has produced a balanced set of fewer chromosomes. Since *R. pubera* underwent two rounds of WGD, descending dysploidy by EEFs would be a way to effectively create chromosomes with different combinations of ancestral syntenic blocks, reducing the risk of meiotic multivalent pairing without the need of rapid genome downsizing. EEFs in genomes with monocentric chromosomes are normally associated with the formation of typically unstable dicentric chromosomes but may represent a way for chromosomal rearrangements when coupled with concurrent centromere elimination as part of structural diploidization after WGDs (Mandáková et al., 2010; Mandáková and Lysak, 2018; Murat et al., 2010). We argue that the prevalence of EEFs observed in *R. pubera* was facilitated by holocentricity, avoiding the deleterious effect of two centromeres after EEFs in monocentric species and likely promoting rapid structural diploidization.

In *Rhynchospora*, homologous non-sister chromatids are linked by terminal chromatin threads during inverted meiosis (Cabral et al., 2014). EEFs may occur with high(er) frequency in scenarios where chromatids of nonhomologous chromosomes are erroneously connected via repeat-based chromatin threads. However, this notion does not exclude the possibility of EEFs occurring during interphase or mitosis. It is tempting to speculate that the repeat-rich regions observed at chromosome ends in *R. breviuscula* are involved in the formation of chromatin threads, which may act as substrates for ectopic recombination. *Tyba* repeats near these repeat-rich regions may be preferentially used as the site for recombination and may thus facilitate the occurrence of EEFs (Figure 7C). Alternatively, the recruitment of *Tyba* repeats as DNA templates to seal double-stranded breaks involved in EEFs may explain their pronounced association with EEFs (Vu et al., 2017).

Limitations of the study

The three *Rhynchospora* species analyzed in this study are characterized by repeat-based holocentromeres associated with *Tyba* repeats. However, some *Rhynchospora* species lack *Tyba* repeats (Ribeiro et al., 2017); thus, it is not clear whether repeat-based holocentromeres evolved in all species of the genus. Extending our approach to other holocentric species lacking *Tyba*-like repeats will certainly reveal new insights into the evolution of repeat-based holocentromeres. In addition, the presence of holocentric chromosomes in multiple genera of sedges as well as in closely related rushes (e.g., *Luzula* species), but not in *Juncus*, suggests that the transition to holocentricity occurred a long time ago (>60 Mya), which makes temporal tracking challenging. Indeed, our analyses of orthogroups did not identify a clear pattern related to different centromere types.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Plant material
- METHOD DETAILS
  - Genome size measurement by flow cytometry
  - Library preparations and sequencing
  - Genome size estimation using k-mer frequency
  - Sequence-based ploidy assessment

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.cell.2022.06.045.

ACKNOWLEDGMENTS

We thank Martin Mascher, Raphael Mercier, Korbinian Schneeberger, Martin Lysak, and Ingo Schubert for their comments on the manuscript. We thank Saurabh Pophaly for assistance with NCBI sequence submission. We thank Neysan Donnelly for extensive manuscript editing. We thank Kristin Krause and Hequan Sun for support with ChIP experiments and k-mer frequency analysis, respectively. We acknowledge the excellent technical assistance of Zdeňka Dubská, Jitka Weiserová, and Eva Jahnová in preparing HMW DNA for optical mapping. This work was supported by the Max Planck Society, Germany, a PhD fellowship awarded to G.T. from the DAAD/India, and grants awarded to A.M. (DFG MA 9363/2-1), A.H., A.P.-H., A.M. (CAPES/DAAD PROBRAL 88881.144086/2017-01), and J.M. (GACR 20-24252S). We also thank the Elixir CZ Research Infrastructure Project (LM2018131), Czech Republic, for providing computing and data-storage facilities. This manuscript is dedicated to Prof. Marcelo Guerra for his long-standing work on beak-sedges cytogenetics, which paved the way for this work.

AUTHOR CONTRIBUTIONS

Conceptualization, A.M.; funding and resources, A.P.-H., J.M., K.F.X.M., A.H., and A.M.; data production, P.G.H., G.T., B.H., H.T., H.Š., and A.M.; formal analyses, investigation, and visualization, P.G.H., G.T., T.L., P. Neumann, T.V., P. Novak, M.Z., L.C., M.C., A.S., H.T., J.F., Y.M.-S., Y.D., A.L.L.V., C.C.S.A., G.S., and A.M.; writing – original draft, P.G.H., G.T., and A.M.; writing – review & editing: all co-authors.

DECLARATION OF INTERESTS

The authors declare no competing interests.

INCLUSION AND DIVERSITY

One or more of the authors of this paper self-identifies as an underrepresented ethnic minority in science. One or more of the authors of this paper self-identifies as a member of the LGBTQ+ community.

Received: January 3, 2022
Revised: May 24, 2022
Accepted: June 24, 2022
Published: August 4, 2022

REFERENCES

Achrem, M., Szućko, I., and Kalinka, A. (2020). The epigenetic regulation of centromeres and telomeres in plants and animals. Comp. Cytogenet. *14*, 265–311. https://doi.org/10.3897/CompCytogen.v14i2.51895.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. (1990). Basic local alignment search tool. J. Mol. Biol. *215*, 403–410. https://doi.org/10.1016/S0022-2836(05)80360-2.

Bremer, K. (2002). Gondwanan evolution of the grass alliance of families (Poales). Evolution *56*, 1374–1387. https://doi.org/10.1111/j.0014-3820.2002.tb01451.x.

Burchardt, P., Buddenhagen, C.E., Gaeta, M.L., Souza, M.D., Marques, A., and Vanzela, A.L.L. (2020). Holocentric karyotype evolution in Rhynchospora is marked by intense numerical, structural, and genome size changes. Front. Plant Sci. *11*, 536507. https://doi.org/10.3389/fpls.2020.536507.

Cabral, G., Marques, A., Schubert, V., et al. (2014). Chiasmatic and achiasmatic inverted meiosis of plants with holocentric chromosomes. Nat. Commun. *5*, 5070. https://doi.org/10.1038/ncomms6070.

Câmara, A.S., Schubert, V., Mascher, M., and Houben, A. (2021). A simple model explains the cell cycle-dependent assembly of centromeric nucleosomes in holocentric species. Nucleic Acids Res. *49*, 9053–9065. https://doi.org/10.1093/nar/gkab648.

Can, M., Wei, W., Zi, H., Bai, M., Liu, Y., Gao, D., Tu, D., Bao, Y., Wang, L., Chen, S., et al. (2020). Genome sequence of Kobresia littledalei, the first chromosome-level genome in the family Cyperaceae. Sci. Data *7*, 175. https://doi.org/10.1038/s41597-020-0518-3.

Capella-Gutiérrez, S., Silla-Martínez, J.M., and Gabaldón, T. (2009). trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. Bioinformatics *25*, 1972–1973. https://doi.org/10.1093/bioinformatics/btp348.

Castiglione, M.R., and Cremonini, R. (2012). A fascinating island: 2n=4. Plant Biosyst *146*, 711–726. https://doi.org/10.1080/11263504.2012.714806.

Cheng, H.Y., Concepcion, G.T., Feng, X.W., Zhang, H.W., and Li, H. (2021). Haplotype-resolved de novo assembly using phased assembly graphs with hifiasm. Nat. Methods *18*, 170–175. https://doi.org/10.1038/s41592-020-01056-5.

Cicconardi, F., Lewis, J.J., Martin, S.H., Reed, R.D., Danko, C.G., and Montgomery, S.H. (2021). Chromosome fusion affects genetic diversity and evolutionary turnover of functional loci but consistently depends on chromosome size. Mol. Biol. Evol. *38*, 4449–4462. https://doi.org/10.1093/molbev/msab185.

Conway, J.R., Lex, A., and Gehlenborg, N. (2017). UpSetR: an R package for the visualization of intersecting sets and their properties. Bioinformatics *33*, 2938–2940. https://doi.org/10.1093/bioinformatics/btx364.

Cortes-Silva, N., Ulmer, J., Kiuchi, T., Hsieh, E., Cornilleau, G., Ladid, I., Dingli, F., Loew, D., Katsuma, S., and Drinnenberg, I.A. (2020). CenH3-independent kinetochore assembly in Lepidoptera requires CCAN, including CENP-T. Curr. Biol. *30*, 561–572.e10. https://doi.org/10.1016/j.cub.2019.12.014.

Costa, L., Marques, A., Buddenhagen, C., Thomas, W.W., Huettel, B., Schubert, V., Dodsworth, S., Houben, A., Souza, G., and Pedrosa-Harand, A. (2021). Aiming off the target: recycling target capture sequencing reads for investigating repetitive DNA. Ann. Bot. *128*, 835–848. https://doi.org/10.1093/aob/mcab063.

Despot-Slade, E., Mravinac, B., Širca, S., Castagnone-Sereno, P., Plohl, M., and Meštrović, N. (2021). The Centromere histone Is Conserved and Associated with Tandem Repeats Sharing a Conserved 19-bp Box in the Holocentromere of meloidogyne Nematodes. Mol. Biol. Evol. *38*, 1943–1965. https://doi.org/10.1093/molbev/msaa336.

Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics *29*, 15–21. https://doi.org/10.1093/bioinformatics/bts635.

Dolezel, J., Bartos, J., Voglmayr, H., and Greilhuber, J. (2003). Nuclear DNA content and genome size of trout and human. Cytometry A *51*, 127–128. https://doi.org/10.1002/cyto.a.10013.

Drummond, A.J., and Rambaut, A. (2007). BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evol. Biol. *7*, 214. https://doi.org/10.1186/1471-2148-7-214.

Durand, N.C., Robinson, J.T., Shamim, M.S., Machol, I., Mesirov, J.P., Lander, E.S., and Aiden, E.L. (2016). Juicebox provides a visualization system for Hi-C contact maps with unlimited zoom. Cell Syst. *3*, 99–101. https://doi.org/10.1016/j.cels.2015.07.012.

Emms, D.M., and Kelly, S. (2019). OrthoFinder: phylogenetic orthology inference for comparative genomics. Genome Biol *20*, 238. https://doi.org/10.1186/s13059-019-1832-y.

Escudero, M., Márquez-Corro, J.I., and Hipp, A.L. (2016). The phylogenetic origins and evolutionary history of holocentric chromosomes. Syst. Bot. *41*, 580–585. https://doi.org/10.1600/036364416X692442.

Feng, S., Zhong, Z., Wang, M., and Jacobsen, S.E. (2020). Efficient and accurate determination of genome-wide DNA methylation patterns in Arabidopsis thaliana with enzymatic methyl sequencing. Epigenetics Chromatin *13*, 42. https://doi.org/10.1186/s13072-020-00361-9.

Fernandes, J.B., Wlodzimierz, P., and Henderson, I.R. (2019). Meiotic recombination within plant centromeres. Curr. Opin. Plant Biol. *48*, 26–35. https://doi.org/10.1016/j.pbi.2019.02.008.

Fuchs, J., Demidov, D., Houben, A., and Schubert, I. (2006). Chromosomal histone modification patterns–from conservation to diversity. Trends Plant Sci. *11*, 199–208. https://doi.org/10.1016/j.tplants.2006.02.008.

Gernhard, T. (2008). The conditioned reconstructed process. J. Theor. Biol. *253*, 769–778. https://doi.org/10.1016/j.jtbi.2008.04.005.

Ghosh, S., and Chan, C.K. (2016). Analysis of RNA-seq data using TopHat and cufflinks. Methods Mol. Biol. *1374*, 339–361. https://doi.org/10.1007/978-1-4939-3167-5_18.

Ghurye, J., Rhie, A., Walenz, B.P., Schmitt, A., Selvaraj, S., Pop, M., Phillippy, A.M., and Koren, S. (2019). Integrating Hi-C links with assembly graphs for chromosome-scale assembly. PLoS Comput. Biol. *15*, e1007273. https://doi.org/10.1371/journal.pcbi.1007273.

Gohard, F.H., Zhiteneva, A.A., and Earnshaw, W.C. (2014). Centromeres (eLS). Chichester. https://onlinelibrary.wiley.com/doi/epdf/10.1002/9780470015902.a0005785.pub2.

Grabherr, M.G., Haas, B.J., Yassour, M., Levin, J.Z., Thompson, D.A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q.D., et al. (2011). Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat. Biotechnol. *29*, 644–652. U130. https://doi.org/10.1038/nbt.1883.

Gremme, G., Brendel, V., Sparks, M.E., and Kurtz, S. (2005). Engineering a software tool for gene structure prediction in higher organisms. Inf. Software Technol. *47*, 965–978. https://doi.org/10.1016/j.infsof.2005.09.005.

Guerra, M. (2016). Agmatoploidy and symploidy: a critical review. Genet. Mol. Biol. *39*, 492–496. https://doi.org/10.1590/1678-4685-GMB-2016-0103.

Guerra, M., Ribeiro, T., and Felix, L.P. (2019). Monocentric chromosomes in Juncus (Juncaceae) and implications for the chromosome evolution of the family. Bot. J. Linn. Soc. *191*, 475–483. https://doi.org/10.1093/botlinnean/boz065.

Haas, B.J., Delcher, A.L., Mount, S.M., Wortman, J.R., Smith, R.K., Jr., Hannick, L.I., Maiti, R., Ronning, C.M., Rusch, D.B., Town, C.D., et al. (2003). Improving the Arabidopsis genome annotation using maximal transcript alignment assemblies. Nucleic Acids Res. *31*, 5654–5666. https://doi.org/10.1093/nar/gkg770.

Haas, B.J., Salzberg, S.L., Zhu, W., Pertea, M., Allen, J.E., Orvis, J., White, O., Buell, C.R., and Wortman, J.R. (2008). Automated eukaryotic gene structure annotation using EVidenceModeler and the Program to Assemble Spliced Alignments. Genome Biol. *9*, R7. https://doi.org/10.1186/gb-2008-9-1-r7.

Hao, Z., Lv, D., Ge, Y., Shi, J., Weijers, D., Yu, G., and Chen, J. (2020). RIdeogram: drawing SVG graphics to visualize and map genome-wide data on the idiograms. PeerJ Comput. Sci. *6*, e251. https://doi.org/10.7717/peerj-cs.251.

Haug-Baltzell, A., Stephens, S.A., Davey, S., Scheidegger, C.E., and Lyons, E. (2017). SynMap2 and SynMap3D: web-based whole-genome synteny browsers. Bioinformatics *33*, 2197–2198. https://doi.org/10.1093/bioinformatics/btx144.

Heckmann, S., Macas, J., Kumke, K., Fuchs, J., Schubert, V., Ma, L., Novák, P., Neumann, P., Taudien, S., Platzer, M., et al. (2013). The holocentric species *Luzula elegans* shows interplay between centromere and large-scale genome organization. Plant J. *73*, 555–565. https://doi.org/10.1111/tpj.12054.

Hill, J., Rastas, P., Hornett, E.A., Neethiraj, R., Clark, N., Morehouse, N., de la Paz Celorio-Mancera, M., Cols, J.C., Dircksen, H., Meslin, C., et al. (2019). Unprecedented reorganization of holocentric chromosomes provides insights into the enigma of lepidopteran chromosome evolution. Sci. Adv. *5*, eaau3648. https://doi.org/10.1126/sciadv.aau3648.

Hoencamp, C., Dudchenko, O., Elbatsh, A.M.O., Brahmachari, S., Raaijmakers, J.A., van Schaik, T., Sedeño Cacciatore, Á., Contessoto, V.G., van Heesbeen, R.G.H.P., van den Broek, B., et al. (2021). 3D genomics across the tree of life reveals condensin II as a determinant of architecture type. Science *372*, 984–989. https://doi.org/10.1126/science.abe2218.

Hoff, K.J., and Stanke, M. (2019). Predicting genes in single genomes with AUGUSTUS. Curr. Protoc. Bioinformatics *65*, e57. https://doi.org/10.1002/cpbi.57.

Kasinathan, S., and Henikoff, S. (2018). Non-B-form DNA is enriched at centromeres. Mol. Biol. Evol. *35*, 949–962. https://doi.org/10.1093/molbev/msy010.

Katoh, K., and Standley, D.M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol. Biol. Evol. *30*, 772–780. https://doi.org/10.1093/molbev/mst010.

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., Buxton, S., Cooper, A., Markowitz, S., Duran, C., et al. (2012). Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics *28*, 1647–1649. https://doi.org/10.1093/bioinformatics/bts199.

Kovaka, S., Zimin, A.V., Pertea, G.M., Razaghi, R., Salzberg, S.L., and Pertea, M. (2019). Transcriptome assembly from long-read RNA-seq alignments with StringTie2. Genome Biol. *20*, 278. https://doi.org/10.1186/s13059-019-1910-1.

Krueger, F., and Andrews, S.R. (2011). Bismark: a flexible aligner and methylation caller for bisulfite-Seq applications. Bioinformatics *27*, 1571–1572. https://doi.org/10.1093/bioinformatics/btr167.

Krumsiek, J., Arnold, R., and Rattei, T. (2007). Gepard: a rapid and sensitive tool for creating dotplots on genome scale. Bioinformatics *23*, 1026–1028. https://doi.org/10.1093/bioinformatics/btm039.

Krzywinski, M., Schein, J., Birol, I., Connors, J., Gascoyne, R., Horsman, D., Jones, S.J., and Marra, M.A. (2009). Circos: an information aesthetic for comparative genomics. Genome Res. *19*, 1639–1645. https://doi.org/10.1101/gr.092759.109.

Kurtz, S., Narechania, A., Stein, J.C., and Ware, D. (2008). A new method to compute K-mer frequencies and its application to annotate large repetitive plant genomes. BMC Genomics *9*, 517. https://doi.org/10.1186/1471-2164-9-517.

Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nat. Methods *9*, 357–359. https://doi.org/10.1038/nmeth.1923.

Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics *25*, 1754–1760. https://doi.org/10.1093/bioinformatics/btp324.

Lonnig, W.E., and Saedler, H. (2002). Chromosome rearrangements and transposable elements. Annu. Rev. Genet. *36*, 389–410. https://doi.org/10.1146/annurev.genet.36.040202.092802.

Lopez-Delisle, L., Rabbani, L., Wolff, J., Bhardwaj, V., Backofen, R., Grüning, B., Ramírez, F., and Manke, T. (2021). pyGenomeTracks: reproducible plots for multivariate genomic datasets. Bioinformatics *37*, 422–423. https://doi.org/10.1093/bioinformatics/btaa692.

Lucek, K., Augustijnen, H., and Escudero, M. (2022). A holocentric twist to chromosomal speciation? Trends Ecol. Evol. https://doi.org/10.1016/j.tree.2022.04.002.

Lukhtanov, V.A., Dincă, V., Friberg, M., Šíchová, J., Olofsson, M., Vila, R., Marec, F., and Wiklund, C. (2018). Versatility of multivalent orientation, inverted meiosis, and rescued fitness in holocentric chromosomal hybrids. Proc. Natl. Acad. Sci. USA *115*, E9610–E9619. https://doi.org/10.1073/pnas.1802610115.

Mandáková, T., Joly, S., Krzywinski, M., Mummenhoff, K., and Lysak, M.A. (2010). Fast diploidization in close mesopolyploid relatives of Arabidopsis. Plant Cell *22*, 2277–2290. https://doi.org/10.1105/tpc.110.074526.

Mandáková, T., and Lysak, M.A. (2018). Post-polyploid diploidization and diversification through dysploid changes. Curr. Opin. Plant Biol. *42*, 55–65. https://doi.org/10.1016/j.pbi.2018.03.001.

Mandrioli, M., and Manicardi, G.C. (2020). Holocentric chromosomes. PLOS Genet. *16*, e1008918. https://doi.org/10.1371/journal.pgen.1008918.

Marçais, G., and Kingsford, C. (2011). A fast, lock-free approach for efficient parallel counting of occurrences of k-mers. Bioinformatics *27*, 764–770. https://doi.org/10.1093/bioinformatics/btr011.

Marques, A., Ribeiro, T., Neumann, P., Macas, J., Novák, P., Schubert, V., Pellino, M., Fuchs, J., Ma, W., Kuhlmann, M., et al. (2015). Holocentromeres in *Rhynchospora* are associated with genome-wide centromere-specific repeat arrays interspersed among euchromatin. Proc. Natl. Acad. Sci. USA *112*, 13633–13638. https://doi.org/10.1073/pnas.1512255112.

Marques, A., Schubert, V., Houben, A., and Pedrosa-Harand, A. (2016). Restructuring of holocentric centromeres During meiosis in the plant Rhynchospora pubera. Genetics *204*, 555–568. https://doi.org/10.1534/genetics.116.191213.

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.journal *17*, 10–12. https://doi.org/10.14806/ej.17.1.200.

Mayrose, I., and Lysak, M.A. (2021). The evolution of chromosome numbers: mechanistic models and experimental approaches. Genome Biol. Evol. *13*. https://doi.org/10.1093/gbe/evaa220.

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

Minh, B.Q., Schmidt, H.A., Chernomor, O., Schrempf, D., Woodhams, M.D., von Haeseler, A., and Lanfear, R. (2020). IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. Mol. Biol. Evol. *37*, 1530–1534. https://doi.org/10.1093/molbev/msaa015.

Mizuno, H., Kawahara, Y., Wu, J., Katayose, Y., Kanamori, H., Ikawa, H., Itoh, T., Sasaki, T., and Matsumoto, T. (2011). Asymmetric distribution of gene expression in the centromeric region of rice chromosome 5. Front. Plant Sci. *2*, 16. https://doi.org/10.3389/fpls.2011.00016.

Muller, H., Gil, J., Jr., and Drinnenberg, I.A. (2019). The impact of centromeres on spatial genome architecture. Trends Genet. *35*, 565–578. https://doi.org/10.1016/j.tig.2019.05.003.

Murat, F., Xu, J.H., Tannier, E., Abrouk, M., Guilhot, N., Pont, C., Messing, J., and Salse, J. (2010). Ancestral grass karyotype reconstruction unravels new mechanisms of genome shuffling as a source of plant evolution. Genome Res. *20*, 1545–1557. https://doi.org/10.1101/gr.109744.110.

Nachtweide, S., and Stanke, M. (2019). Multi-genome annotation with AUGUSTUS. Methods Mol. Biol. *1962*, 139–160. https://doi.org/10.1007/978-1-4939-9173-0_8.

Naish, M., Alonge, M., Wlodzimierz, P., Tock, A.J., Abramson, B.W., Schmücker, A., Mandáková, T., Jamge, B., Lambing, C., Kuo, P., et al. (2021). The genetic and epigenetic landscape of the Arabidopsis centromeres. Science *374*, eabi7489. https://doi.org/10.1126/science.abi7489.

Neumann, P., Novák, P., Hoštáková, N., and Macas, J. (2019). Systematic survey of plant LTR-retrotransposons elucidates phylogenetic relationships of their polyprotein domains and provides a reference for element classification. Mobile DNA *10*, 1. https://doi.org/10.1186/s13100-018-0144-1.

Nhim, S., Gimenez, S., Nait-Saidi, R., Severac, D., Nam, K., d'Alençon, E., and Nègre, N. (2022). H3K9me2 genome-wide distribution in the holocentric insect Spodoptera frugiperda (Lepidoptera: Noctuidae). Genomics *114*, 384–397. https://doi.org/10.1016/j.ygeno.2021.12.014.

Novák, P., Neumann, P., and Macas, J. (2020). Global analysis of repetitive DNA from unassembled sequence reads using RepeatExplorer2. Nat. Protoc. *15*, 3745–3776. https://doi.org/10.1038/s41596-020-0400-y.

Nurk, S., Walenz, B.P., Rhie, A., Vollger, M.R., Logsdon, G.A., Grothe, R., Miga, K.H., Eichler, E.E., Phillippy, A.M., and Koren, S. (2020). HiCanu: accurate assembly of segmental duplications, satellites, and allelic variants from high-fidelity long reads. Genome Res *30*, 1291–1305. https://doi.org/10.1101/gr.263566.120.

Paten, B., Earl, D., Nguyen, N., Diekhans, M., Zerbino, D., and Haussler, D. (2011). Cactus: algorithms for genome multiple sequence alignment. Genome Res. *21*, 1512–1528. https://doi.org/10.1101/gr.123356.111.

Pedrosa, A., Sandal, N., Stougaard, J., Schweizer, D., and Bachmair, A. (2002). Chromosomal map of the model legume Lotus japonicus. Genetics *161*, 1661–1672. https://doi.org/10.1093/genetics/161.4.1661.

Quinlan, A.R., and Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics *26*, 841–842. https://doi.org/10.1093/bioinformatics/btq033.

Ramírez, F., Ryan, D.P., Grüning, B., Bhardwaj, V., Kilpert, F., Richter, A.S., Heyne, S., Dündar, F., and Manke, T. (2016). deepTools2: a next generation web server for deep-sequencing data analysis. Nucleic Acids Res. *44*, W160–W165. https://doi.org/10.1093/nar/gkw257.

Ranallo-Benavidez, T.R., Jaron, K.S., and Schatz, M.C. (2020). GenomeScope 2.0 and Smudgeplot for reference-free profiling of polyploid genomes. Nat. Commun. *11*, 1432. https://doi.org/10.1038/s41467-020-14998-3.

Reimer, J.J., and Turck, F. (2010). 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. https://doi.org/10.1007/978-1-60761-646-7_12.

Ribeiro, T., Buddenhagen, C.E., Thomas, W.W., Souza, G., and Pedrosa-Harand, A. (2018). Are holocentrics doomed to change? Limited chromosome number variation in Rhynchospora Vahl (Cyperaceae). Protoplasma *255*, 263–272. https://doi.org/10.1007/s00709-017-1154-4.

Ribeiro, T., Marques, A., Novák, P., Schubert, V., Vanzela, A.L., Macas, J., Houben, A., and Pedrosa-Harand, A. (2017). Centromeric and non-centromeric satellite DNA organisation differs in holocentric Rhynchospora species. Chromosoma *126*, 325–335. https://doi.org/10.1007/s00412-016-0616-3.

Rice, P., Longden, I., and Bleasby, A. (2000). EMBOSS: the European molecular biology open software suite. Trends Genet. *16*, 276–277. https://doi.org/10.1016/s0168-9525(00)02024-2.

Robinson, J.T., Thorvaldsdóttir, H., Winckler, W., Guttman, M., Lander, E.S., Getz, G., and Mesirov, J.P. (2011). Integrative genomics viewer. Nat. Biotechnol. *29*, 24–26. https://doi.org/10.1038/nbt.1754.

Schotanus, K., Yadav, V., and Heitman, J. (2021). Epigenetic dynamics of centromeres and neocentromeres in Cryptococcus deuterogattii. PLoS Genet. *17*, e1009743. https://doi.org/10.1371/journal.pgen.1009743.

Schubert, I., and Lysak, M.A. (2011). Interpretation of karyotype evolution should consider chromosome structural constraints. Trends Genet. *27*, 207–216. https://doi.org/10.1016/j.tig.2011.03.004.

Senaratne, A.P., Muller, H., Fryer, K.A., Kawamoto, M., Katsuma, S., and Drinnenberg, I.A. (2021). Formation of the CenH3-deficient holocentromere in

Lepidoptera avoids active chromatin. Curr. Biol. *31*, 173–181.e7. https://doi.org/10.1016/j.cub.2020.09.078.

Seppey, M., Manni, M., and Zdobnov, E.M. (2019). BUSCO: assessing genome assembly and annotation completeness. Methods Mol. Biol. *1962*, 227–245. https://doi.org/10.1007/978-1-4939-9173-0_14.

Šimková, H., Číhalíková, J., Vrána, J., Lysák, M.A., and Doležel, J. (2003). Preparation of HMW DNA from plant nuclei and chromosomes isolated from root tips. Biol. Plant. *46*, 369–373. https://doi.org/10.1023/A:1024322001786.

Smith, S., Collinson, M., Rudall, P., and Simpson, D. (2010). The Cretaceous and Palaeogene fossil record of Poales: review and current research. In Diversity, phylogeny, and evolution in monocotyledons. Proceedings of the fourth international conference on the comparative biology of the monocotyledons & the fifth international, O. Seberg, G. Petersen, A. Barfod, and J.I. Davis, eds. (Aarhus University Press), pp. 333–356.

Soderlund, C., Bomhoff, M., and Nelson, W.M. (2011). SyMAP v3.4: a turnkey synteny system with application to plant genomes. Nucleic Acids Res *39*, e68. https://doi.org/10.1093/nar/gkr123.

Soderlund, C., Nelson, W., Shoemaker, A., and Paterson, A. (2006). SyMAP: A system for discovering and viewing syntenic regions of FPC maps. Genome Res. *16*, 1159–1168. https://doi.org/10.1101/gr.5396706.

Steiner, F.A., and Henikoff, S. (2014). Holocentromeres are dispersed point centromeres localized at transcription factor hotspots. eLife *3*, e02025. https://doi.org/10.7554/eLife.02025.

Stovner, E.B., and Sætrom, P. (2019). epic2 efficiently finds diffuse domains in ChIP-seq data. Bioinformatics *35*, 4392–4393. https://doi.org/10.1093/bioinformatics/btz232.

Sun, H.Q., Ding, J., Piednoël, M., and Schneeberger, K. (2018). findGSE: estimating genome size variation within human and Arabidopsis using k-mer frequencies. Bioinformatics *34*, 550–557. https://doi.org/10.1093/bioinformatics/btx637.

Tang, H., Bowers, J.E., Wang, X., Ming, R., Alam, M., and Paterson, A.H. (2008). Synteny and collinearity in plant genomes. Science *320*, 486–488. https://doi.org/10.1126/science.1153917.

Ter-Hovhannisyan, V., Lomsadze, A., Chernoff, Y.O., and Borodovsky, M. (2008). Gene prediction in novel fungal genomes using an ab initio algorithm with unsupervised training. Genome Res. *18*, 1979–1990. https://doi.org/10.1101/gr.081612.108.

Thomas, J., and Pritham, E.J. (2015). Helitrons, the eukaryotic rolling-circle transposable elements. Microbiol. Spectr. *3*. https://doi.org/10.1128/microbiolspec.MDNA3-0049-2014.

Vanzela, A.L.L., Guerra, M., and Luceño, M. (1996). *Rhynchospora tenuis* Link (Cyperaceae), a species with the lowest number of holocentric chromosomes. Cytobios *88*, 219–228.

Vu, G.T.H., Cao, H.X., Fauser, F., Reiss, B., Puchta, H., and Schubert, I. (2017). Endogenous sequence patterns predispose the repair modes of CRISPR/Cas9-induced DNA double-stranded breaks in Arabidopsis thaliana. Plant J. *92*, 57–67. https://doi.org/10.1111/tpj.13634.

Wickham, H. (2016). ggplot2 : elegant graphics for data analysis. In Use R! Imprint (Springer International Publishing).

Wong, C.Y.Y., Lee, B.C.H., and Yuen, K.W.Y. (2020). Epigenetic regulation of centromere function. Cell. Mol. Life Sci. *77*, 2899–2917. https://doi.org/10.1007/s00018-020-03460-8.

Wu, T.D., and Watanabe, C.K. (2005). GMAP: a genomic mapping and alignment program for mRNA and EST sequences. Bioinformatics *21*, 1859–1875. https://doi.org/10.1093/bioinformatics/bti310.

Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., Li, W., et al. (2008). Model-based analysis of ChIP-Seq (MACS). Genome Biol. *9*, R137. https://doi.org/10.1186/gb-2008-9-9-r137.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
polyclonal rabbit anti-RpCENH3	Marques et al., 2015	anti-RpCENH3
Rabbit polyclonal to Histone H3 (tri methyl K4)	abcam	Cat# ab8580; RRID:AB_306649
Mouse monoclonal to Histone H3 (di methyl K9)	abcam	Cat# ab1220; RRID:AB_449854
Recombinant Rabbit IgG, monoclonal	abcam	Cat# ab172730; RRID:AB_2687931
Biological samples		
Rhynchospora pubera	Own greenhouse	N/A
Rhynchospora breviuscula	Own greenhouse	N/A
Rhynchospora tenuis	Own greenhouse	N/A
Juncus effusus var. spiralis	Own greenhouse (commercially acquired)	N/A
Critical commercial assays		
NucleoBond HMW DNA kit	Macherey Nagel	Cat# 740160.2
SMRTbell Express Template Prep Kit 2.0	Pacific Biosciences	Cat# 101-685-400
Dovetail® Omni-C® Kit	Dovetail	Cat# 21005
Ovation Ultralow V2 DNA-Seq library preparation kit	Tecan Genomics	Cat# 0344NB-08
Enzymatic Methyl-seq Kit	NEBNext®	Cat# E7120S
Poly(A) mRNA Magnetic Isolation Module	NEBNext®	Cat# E7490S
TeloPrime Version 2 kit	Lexogen	Cat# 013PF032V0200
Deposited data		
All sequence data	This study	PRJNA784789
Oligonucleotides		
(FAM)TTTAGGG(8)	Sigma-Aldrich	Telomeric FISH (FAM)oligo-labeled probe
(CY3)ATTGGATTATACATGGTAATTACGCATATAA AGTGCAAATAATGCAATTC	Sigma-Aldrich	*Tyba* repeat FISH oligo-labeled probe
(CY3)GCAAACCAAAATTTGTGTTCAATTTTAAAAT ATTTCTCCAC	Sigma-Aldrich	*Juncus effusus* cenDNA FISH oligo-labeled probe
Software and algorithms		
HiCanu 2.0	Nurk et al., 2020	https://github.com/marbl/canu
Hifiasm 0.16.1 (r375)	Cheng et al., 2021	https://github.com/chhylp123/hifiasm
BUSCO	Seppey et al., 2019	https://gitlab.com/ezlab/busco/
SALSA2	Ghurye et al., 2019	https://github.com/marbl/SALSA
Bowtie2	Langmead and Salzberg, 2012	https://github.com/BenLangmead/bowtie2
BWA	Li and Durbin, 2009	https://github.com/lh3/bwa
STAR (version 2.7.8a)	Dobin et al., 2013	https://github.com/alexdobin/STAR
DeepTools	Ramirez et al., 2016	https://github.com/deeptools/deepTools
PyGenomeTracks	Lopez-Delisle et al., 2021	https://github.com/deeptools/pyGenomeTracks
bedtools	Quinlan and Hall, 2010	https://github.com/arq5x/bedtools2
MCScan	Tang et al., 2008	https://github.com/tanghaibao/mcscan
SyMAP	Soderlund et al., 2011; Soderlund et al., 2006	https://github.com/csoderlund/SyMAP
Rideogram	Hao et al., 2020	https://github.com/TickingClock1992/RIdeogramsy
SynMap2	Haug-Baltzell et al., 2017	https://genomevolution.org/coge/SynMap.pl
FindGSE	Sun et al., 2018	https://github.com/schneebergerlab/findGSEsmudge

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Smudgeplot	Ranallo-Benavidez et al., 2020	https://github.com/KamilSJaron/smudgeplot
Juicer	Durand et al., 2016	https://github.com/aidenlab/juicer
Straw (strawC v0.0.9)	Durand et al., 2016	https://github.com/aidenlab/straw
Augustus (version 3.3.3)	Hoff and Stanke, 2019	https://github.com/Gaius-Augustus/Augustus
EVidenceModeller	Haas et al., 2008	https://github.com/EVidenceModeler
TRINITY (version 2.13.1)	Grabherr et al., 2011	https://github.com/trinityrnaseq/trinityrnaseq
BLAST+ (ncbi-blast-2.3.0+)	Altschul et al., 1990	ftp://ftp.ncbi.nlm.nih.gov/blast/executables/blast+/LATEST/
circus	Krzywinski et al., 2009	http://circos.ca/
ggplot2	Wickham, 2016	https://github.com/tidyverse/ggplot2
Geneious	Kearse et al., 2012	https://www.geneious.com/
RepeatExplorer2	Novák et al., 2020	https://repeatexplorer-elixir.cerit-sc.cz/galaxy/
jellyfish	Marçais and Kingsford, 2011	https://github.com/gmarcais/Jellyfish
Bismark	Krueger and Andrews, 2011	https://github.com/FelixKrueger/Bismark
MAFFT	Katoh and Standley, 2013	https://github.com/GSLBiotech/mafft
TrimAl	Capella-Gutiérrez et al., 2009	https://github.com/inab/trimal
IQ-tree2	Minh et al., 2020	http://www.iqtree.org/
BEAST v.1.10.4	Drummond and Rambaut, 2007	https://beast.community/
Other		
Assemblies, predicted transcripts and proteins, ChIP and DNA methylation tracks, repeat annotation tracks	This study	https://data.cyverse.org/dav-anon/iplant/home/dabitz66/marquesLabTrackHub/
Assemblies	This study	https://genomevolution.org/coge
Genome Browser	This study	https://genome-euro.ucsc.edu/cgi-bin/hgGateway?genome=rhyPub2m&hubUrl=https://data.cyverse.org/dav-anon/iplant/home/dabitz66/marquesLabTrackHub/hub.txt

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, André Marques (amarques@mpipz.mpg.de).

Materials availability

This study did not generate new unique reagents.

Data and code availability

All sequencing data used in this study have been deposited at NCBI under the Bioproject no. PRJNA784789 and are publicly available as of the date of publication. The reference genomes, annotations and all tracks presented in this work are made available at https://data.cyverse.org/dav-anon/iplant/home/dabitz66/marquesLabTrackHub/, the CoGe platform (https://genomevolution.org/coge) and the following UCSC Genome Browser hosted by CyVerse. All other data needed to evaluate the conclusions in the paper are provided in the paper and/or the supplemental information.

This paper does not report original code.

Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Plant material

Plants from naturally occurring populations of *R. pubera* and *R. tenuis* growing in Curado (Recife), Northeastern Brazil, and *R. breviuscula* growing in Londrina (Paraná state), Southern Brazil were collected in 2013 and further cultivated under controlled

greenhouse conditions (16h daylight, 26 °C, >70% humidity). As a monocentric outgroup an individual of the ornamental plant *Juncus effusus* var. *spiralis* was commercially obtained and cultivated under controlled greenhouse conditions (16h daylight, 20°C).

METHOD DETAILS

Genome size measurement by flow cytometry

The genome size of 1C=1.6 Gb for the *R. pubera* accession sequenced here has been previously measured (Marques et al., 2015). Thus, genome size estimations by flow cytometry were performed for the accessions of *R. breviuscula* and *R. tenuis* as well as for *Juncus effusus* var. *spiralis*. For that, roughly 0.5 cm^2 of young leaf tissue was chopped with a sharp razorblade in a Petri dish together with appropriate amounts of leaf tissue of the internal reference standard *Raphanus sativus* cv. Voran (2C=1.11 pg; Genebank Gatersleben, accession number: RA 34) using the 'CyStain PI Absolute P' nuclei extraction and staining kit (Sysmex-Partec). The resulting nuclei suspension was filtered through a 50-μm filter (CellTrics, Sysmex-Partec) and measured on a CyFlow Space flow cytometer (Sysmex-Partec). The absolute DNA content (pg/2C) was calculated based on the values of the G1 peak means and the corresponding genome size (Mb/1C) according to Dolezel et al. (2003).

Library preparations and sequencing

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).

Whole-genome shotgun sequencing (WGS)

Genomic DNA from *R. breviuscula* and *R. alba* were deep-sequenced with an Illumina HiSeq 3000 in 150-bp paired-end mode. Alternatively, DNBseq short read sequencing (BGI Genomics, Hong Kong) of genomic DNA was performed for *R. pubera*, *R. tenuis*, and *R. tenerrima*. Available WGS short reads from *R. cephalotes* (SRX9381225), *R. ciliata* (Ribeiro et al., 2017), *R exaltata* (SRX9381226), *R. globosa* (Ribeiro et al., 2017), and *C. littledalei* (SRX5833125, SRX5833124) were used.

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*). The numbers of SMRT cells for each species were as follows: *R. pubera* (3 cells), *R. breviuscula* (1 cell), *R. tenuis* (2 cells), and *J. effusus* (1 cell).

Omni-C

For each species, a single chromatin-capture library was prepared from 0.5 g fresh weight material input. All treatments were according to the recommendations of 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 for *R. breviuscula*, *R. tenuis*, and *J. effusus*. Alternatively, the *R. pubera* library was paired-end 2 x 150 bp deep-sequenced using DNBseq technology (BGI Genomics, Hong Kong).

ChIPseq

ChIP DNA was quality-controlled using the NGS-assay on a FEMTO-pulse (Agilent); then, an Illumina-compatible library was prepared with the Ovation Ultralow V2 DNA-Seq library preparation kit (Tecan Genomics) and single-end 1 x 150-bp sequenced on a HiSeq 3000 (Illumina) device. For each library, an average of 20 millions reads were obtained.

Enzymatic Methyl-seq

To investigate the methylome space in *R. pubera* and *J. effusus*, the relatively non-destructive NEBNext® Enzymatic Methyl-seq Kit was employed to prepare an Illumina-compatible library, followed by paired-end sequencing (2 x 150 bp) on a HiSeq 3000 (Illumina) device. For each library, 10 Gb of reads were generated.

RNAseq

Total RNA from root, leaves, and flower buds was isolated from *R. breviuscula*. For *R. tenuis*, total RNA was isolated from flower buds only. For *J. effusus*, RNAseq data from the NCBI (accession numbers SRX2268676, SRX2268675, and SRX1639021) were used to complement its genome annotation. For *R. pubera*, total RNA was extracted from six different tissues (i.e., roots, young leaves, old leaves, stem, early flower buds, and late flower buds). 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 done at BGI Genomics (Hong Kong) with a BGISEQ-500 system in the DNBseq platform in paired-end mode 2 x 150 bp.

IsoSeq

For the proper annotation of the complex *R. pubera* genome, total RNA was extracted from six different tissues (i.e., roots, young leaves, old leaves, stem, early flower buds and late flower buds) and quality-assessed by a Nanochip (Agilent Bioanalyser, Santa Clara, U.S.A.). Next, cDNA was synthesized according to the TeloPrime Version 2 kit (Lexogen, Vienna, Austria). We exchanged the Lexogen first-strand synthesis oligo-dT primer for the (5'-AAGCAGTGGTATCAACGCAGAGTACT(30)VN-3') primer to introduce a 3' anchor base. Then, the optimal number of cycles was determined by qPCR (Viia7, Applied Biosystems) with the 1x Evagreen

fluorochrome (Biotium, Fremont, U.S.A.), TeloPrime kit chemistry and 25% of the cDNA as input. The forward primer was FP from the TeloPrime kit, and the reverse primer was 5'-AAGCAGTGGTATCAACGCAGAGTAC-3'. The residual cDNA was mass-amplified with an extended Lexogen FP primer by adding 16mer barcodes as recommended by PacBio at the 5' end and a cycle number by which 80% of the maximal fluorescence signal was reached. The PCR products were bead-purified (Pronex beads, Promega) followed by PacBio library preparation with the SMRTbell Express Template Prep Kit 2.0 (Pacific Biosciences, Menlo Park, U.S.A.), and then quantity- (Quantus, Promega) and quality-assessed (Agilent Bioanalyser). Long-read sequencing was performed on a Sequel II sequencer with a Sequel II Binding kit 2.1, Sequel II Sequencing Kit 2.0 sequencing chemistry 2.0, and a single 8M SMRT cell (*Pacific Biosciences*, Menlo Park, U.S.A.). The movie time was 30 h after a 2-h immobilization step and 2-h pre-extension step to adjust for high-fidelity (HiFi) sequencing.

Genome size estimation using k-mer frequency

Genome sizes of the three *Rhynchospora* species and *J. effusus* were also confirmed by k-mer frequency analysis with the findGSE tool (Sun et al., 2018), after counting k-mers with Jellyfish (Marçais and Kingsford, 2011). High-coverage short reads were used as follows: *R. pubera* (60x), *R. breviuscula* (50x), and *R. tenuis* (130x). Since for *J. effusus* we did not have short-read data, we used our high-coverage HiFi PacBio reads (70x).

Sequence-based ploidy assessment

We used Smudgeplot (Ranallo-Benavidez et al., 2020) to visualize and estimate the ploidy and structure of the sequenced genomes. This tool can infer ploidy directly from the k-mers present in sequencing reads by analyzing heterozygous k-mer pairs.

Genome assembly

HiFi reads obtained by the sequencing process were subjected to assembly using the HiCanu function of Canu (Nurk et al., 2020), available at https://github.com/marbl/canu, for *R. pubera* with the following command line:

```
canu -assemble -p output.asm -d run1 genomeSize=1.6g maxThreads=40 useGrid=false -pacbio-hifi *.fastq.
```

Alternatively, Hifiasm (Cheng et al., 2021), available at https://github.com/chhylp123/hifiasm, was used for the assembly of *R. breviuscula*, *R. tenuis*, and *J. effusus* with the following command:

```
hifiasm -o output.asm -t 40 reads.fq.gz.
```

Preliminary assemblies were evaluated for contiguity and completeness with BUSCO (Seppey et al., 2019).

Optical map and hybrid scaffolding

We developed an optical mapping strategy to help resolve the complexity of the *R. pubera* genome. High-molecular-weight DNA was prepared from young leaves of *R. pubera*. A total of 3.15 million cell nuclei were purified by flow cytometry, pelleted by centrifugation (30 min at 300 *g*), and embedded in four agarose plugs of 20-μL volume. The nuclear DNA was purified in the plugs as described by Šimková et al. (2003) with an increased concentration of proteinase K (1 mg/mL of lysis buffer). The proteinase- and RNase-treated DNA was isolated from the agarose gel, and the resulting 525 ng DNA was directly labeled at DLE-1 recognition sites following the standard Bionano Prep Direct Label and Stain (DLS) Protocol (Bionano Genomics, San Diego, USA) and analyzed on the Saphyr platform of Bionano Genomics. A total of 1.27 Tbp of single-molecule data with N50 of 236 kb, corresponding to effective coverage of 96.8x of the *R. pubera* genome, was used in *de novo* assembly by Bionano Solve 3.6.1_11162020, using a standard configuration file "optArguments_nonhaplotype_noES_noCut_DLE1_saphyr.xml" (Table S6). A p-value threshold of 1e–11 was used to build the initial assembly, a p-value of 1e–12 was used for extension and refinement steps (five rounds), and a p-value of 1e–16 was used for final map merging. To improve the contiguity of the sequence assembly, an automatic hybrid scaffold pipeline integrated in Bionano Solve 3.6.1_11162020 was run with the *de novo* optical map assembly. The default *DLE-1 Hybrid Scaffold* configuration file was used with the "Resolve conflict" option for conflict resolution. The conflicts between sequences and the optical map were manually curated, and the pipeline was re-run using the modified *conflict_cut_statu.txt* file (Table S7). The results obtained from the optical mapping scaffolding of the genome assembly of *R. pubera* were used as input for Omni-C scaffolding.

Omni-C scaffolding

Dovetail Omni-C reads were first mapped using BWA (Li and Durbin, 2009) following the hic-pipeline available at https://github.com/esrice/hic-pipeline. Hi-C scaffolding was performed using SALSA2 (Ghurye et al., 2019), available at https://github.com/marbl/SALSA, with default parameters. After testing several minimum mapping quality values of bam alignments, final scaffolding was performed with MAPQ10. Several rounds of assembly correction guided by Hi-C contact maps and manual curation of scaffolds were performed to obtain the pseudomolecules.

Assembly and scaffolding strategy

The rather homozygous genome of *J. effusus* was estimated to be close to 1C=271 Mb (Figure 1; Figure S4C). Sequencing of *J. effusus* var. *spiralis* yielded 19 Gb of reads and an initial assembly of 258 Mb (79× coverage, N_{50} = 11 Mb, Figure 1), where 18 contigs corresponded to complete chromosomes. The assembly was further scaffolded to the expected 21 pseudomolecules (240 Mb), and unplaced contigs contained 18 Mb, corresponding to the complete haploid chromosome set of the species

(Figures 1 and 2A). The sequencing of *R. pubera*, which is an inbred species, yielded 66 Gb of PacBio HiFi reads, and the initial assembly spanned 1.7 Gb (38× coverage, N_{50} = 11.2 Mb). After removing redundant sequences likely due to some small residual heterozygosity, the assembly closely matched its estimated haploid genome size (Figures 1 and S3F). A first scaffolding using optical mapping was followed by a second scaffolding using chromosome conformation capture (Omni-C, Dovetail™) of the genome assembly, which yielded five very large pseudomolecules (1.47 Gb, N_{50} = 361 Mb), while unplaced contigs contained 141 Mb (Figure 1). The sequencing of *R. breviuscula* yielded 30 Gb of PacBio HiFi reads, resulting in an initial assembly that was 813 Mb in length. In contrast to *R. pubera*, *R. breviuscula* is outbred, which resulted in an assembly of its diploid genome size showing a high level of heterozygosity confirmed by k-mer analysis (Figures 1 and S4A). We pruned the resulting large contigs to the single largest representative haplotype (75× coverage, 421 Mb, N_{50} = 11 Mb; Figure 1) and then oriented and ordered it into five pseudomolecules using Omni-C scaffolding comprising 370 Mb (N_{50} = 71 Mb; Figure S1A). Unplaced contigs contained 50 Mb (Figure 1). The sequenced genome of *R. tenuis* yielded 45.9 Gb of PacBio HiFi reads resulting in an assembly of 770 Mb, which closely corresponds to its diploid genome size, showing a high level of heterozygosity (Figures 1 and S4B). We pruned the resulting large contigs to the single largest representative haplotype (120× coverage, 395 Mb, N_{50} = 19 Mb, Figure 1), which was oriented and ordered into two pseudomolecules of about 350 Mb (N_{50} = 215 Mb; Figure S1B). Unplaced contigs contained 47 Mb (Figure 1).

GENERATION OF HI-C MAPS

Final Hi-C maps of *R. pubera*, *R. breviuscula*, *R. tenuis*, and *J. effusus* were generated by Juicer (v1.6) (Durand et al., 2016) using the sequencing data from DNase *in situ* Hi-C (Omni-C) experiments. Specifically, technical replicates were aligned and deduplicated and then the results of each replicate were merged by MEGA from Juicer.

Quantitative analysis of Hi-C contacts

The python version of Straw (strawC v0.0.9) (Durand et al., 2016) was used to extract Hi-C counts from the illustrated Hi-C maps (Figure 2; Figure S1) in 1-Mb resolution and with the normalization approach of Vanilla Coverage (VC). To represent the intra- and interchromosomal interactions in an intuitive manner, the *cis* Hi-C contact of a chromosome was quantified as the sum of all Hi-C counts within the chromosome *per se*, while *trans* Hi-C contacts referred to the sum of Hi-C counts between the designated chromosome and all other chromosomes. The final intra- and interchromosomal contacts for each single chromosome were normalized through the percentages of Hi-C counts over the sum of all Hi-C signals in the corresponding Hi-C map. It is also noteworthy that the infinite extracted Hi-C counts through Straw were replaced by the mean of all other finite counts within the extracted chromosomal pair.

ChIP

ChIP experiments were performed following Reimer and Turck (2010), with adjustments for *R. pubera* and *R. breviuscula*. Unopened flower buds were harvested and frozen in liquid nitrogen until sufficient material was obtained. The samples were fixed in 4% formaldehyde for 30 min and the chromatin was sonicated for 25 min. Then, 7–85 μL of sonicated chromatin was incubated with 2 ng of respective antibody overnight. Immunoprecipitation was carried out for rabbit anti-*Rp*CENH3, for *R. pubera* and *R. breviuscula*, and for rabbit anti-H3K4me3 (abcam, ab8580), and mouse anti-H3K9me2 (abcam, ab1220). Recombinant rabbit IgG (abcam, ab172730) and no-antibody inputs were used as controls. Two experimental replications were also maintained for all the combinations. After overnight incubation of chromatin with antibody, protein beads (anti-mouse: Protein G Sepharose 4 Fast Flow, anti-rabbit: rProtein A Sepharose Fast Flow) were added to the chromatin-antibody mixture. The bound chromatin was finally eluted, de-crosslinked, precipitated, and sent for sequencing.

Synteny and self-synteny analyses

The synteny analysis shown in Figure 5 was performed using the MCscan pipeline implemented in the JVCI utility libraries (Tang et al., 2008). For this analysis, CDS sequences of the longest transcript were used. Circular plots were drawn with the circos package (Krzywinski et al., 2009).

Self-synteny analyses shown in Figures S1D, S1E, and S3A were performed with SyMAP v. 5.0.6 (Soderlund et al., 2006, 2011). Circular self-synteny plots were obtained with SyMAP or RIdeogram software (Hao et al., 2020) using the synteny calculation blocks obtained from SyMAP.

Whole-genome alignment (WGA)

A whole-genome alignment (WGA) between *R. pubera*, *R. tenuis*, *R. breviuscula*, *J. effusus*, and *C. littledalei* was generated using the Cactus pipeline (Version 1.0) (Paten et al., 2011). Prior to the alignment step, all nucleotide sequences were 20-kmer-softmasked to reduce complexity and facilitate construction of the WGA using the tallymer subtools from the genome tools package (Version 1.6.1) (Kurtz et al., 2008). The Cactus pipeline was run stepwise with the default settings described at https://github.com/ComparativeGenomicsToolkit/cactus#running-step-by-step.

Characterization of end-to-end fusions

For the characterization of the regions involved in EEFs observed in *R. pubera* and *R. tenuis*, we first compared the synteny alignment between their genomes with *R. breviuscula* used as a reference in SyMAP. This allowed us to pin the putative regions around the borders of the fusion events. To gain insights about the order of fusion events in the complex genome of *R. pubera* we extracted all duplicated fusion regions and aligned them against themselves in SyMAP (Figures 5C and S5). This detailed analysis further allowed us to reconstruct the order of the fusion events and tracing the karyotype history of *R. pubera* based on the shared EEF signatures found in the genome.

In order, to identify the underlying sequences at the fusion regions, we loaded annotation features for genes, TEs, *Tyba*, rDNA and telomeric repeats on SyMAP aligments. This allowed us to detect the sequence types in the putative translocated regions. In *R. pubera*, we counted 15 potential EEF regions, of which 11 regions had a *Tyba* array right in the middle between two ancestral syntenic chromosomes of *R. breviuscula*. Further inspection and characterization of such regions were done by checking the genome coordinates and annotation features with Geneious (Kearse et al., 2012), which revealed a remnant rDNA cluster involved in the EEF of two ancestral *Rb*3 in the *Rp*Chr3. The other three regions did not show any specific sequence (Figures 5C and S6A–S6G). Similar strategy was used to identify the underlying sequences within the three fusion events in *R. tenuis* chromosomes. In this species, we found a *Tyba* repeat underlying the fusion region between *Rb*3/*Rb*4 in *Rt*Chr2 and a telomeric repeat between *Rb*1/*Rb*5 in *Rt*Chr1, while the third fusion between had no specific sequence (Figures 5C and S6H). Common to most fusion events we found that the very ends of *R. breviuscula* chromosomes, which are enriched for TEs are mostly missing in the fused chromosomes (Figures 5D and S6).

Whole-genome duplication analysis

To identify ancient WGD events, we performed Synonymous (Ks) substitutions analysis on the fully annotated genomes with the SynMap2 (Haug-Baltzell et al., 2017) available at CoGe webportal (https://genomevolution.org/coge/SynMap.pl).

Gene annotation

Structural gene annotation was done combining *de novo* gene calling and homology-based approaches with RNAseq, IsoSeq, and protein datasets.

Using evidence derived from expression data, RNAseq data were first mapped using STAR (Dobin et al., 2013) (version 2.7.8a) and subsequently assembled into transcripts by StringTie (Kovaka et al., 2019) (version 2.1.5, parameters -m 150-t -f 0.3). *Triticeae* protein sequences from available public datasets (UniProt, https://www.uniprot.org, 05/10/2016) were aligned against the genome sequence using GenomeThreader (Gremme et al., 2005) (version 1.7.1; arguments -startcodon -finalstopcodon -species rice -gcmincoverage 70 -prseedlength 7 -prhdist 4). Isoseq datasets were aligned to the genome assembly using GMAP (Wu and Watanabe, 2005) (version 2018-07-04). All transcripts from RNAseq, IsoSeq, and aligned protein sequences were combined using Cuffcompare (Ghosh and Chan, 2016) (version 2.2.1) and subsequently merged with StringTie (version 2.1.5, parameters –merge -m150) into a pool of candidate transcripts. TransDecoder (version 5.5.0; http://transdecoder.github.io) was used to find potential open reading frames and to predict protein sequences within the candidate transcript set.

Ab initio annotation was initially done using Augustus (Hoff and Stanke, 2019) (version 3.3.3). GeneMark (Ter-Hovhannisyan et al., 2008) (version 4.35) was additionally employed to further improve structural gene annotation. To avoid potential over-prediction, we generated guiding hints using the above described RNAseq, protein, and IsoSeq datasets as described by Hoff and Stanke (2019). A specific Augustus model for *Rhynchospora* was built by generating a set of gene models with full support from RNAseq and IsoSeq. Augustus was trained and optimized using the steps detailed by Hoff and Stanke (2019).

To maximize uniformity across all annotated species, Augustus was also run in comparative annotation mode (Nachtweide and Stanke, 2019). The generated WGA served as sequence input together with the mapping of RNAseq data as described above.

All structural gene annotations were joined using EVidenceModeller (Haas et al., 2008) (version 1.1.1), and weights were adjusted according to the input source: *ab initio* (Augustus: 5, GeneMark: 2), homology-based (10), and comparative *ab initio* (7). Additionally, two rounds of PASA (Haas et al., 2003) (version 2.4.1) were run to identify untranslated regions and isoforms using transcripts generated by a genome-guided TRINITY (Grabherr et al., 2011) (version 2.13.1) assembly derived from *Rhynchospora* RNAseq data and the above described IsoSeq datasets.

We used BLASTP (Altschul et al., 1990) (ncbi-blast-2.3.0+, parameters -max_target_seqs 1 -evalue 1e–05) to compare potential protein sequences with a trusted set of reference proteins (Uniprot Magnoliophyta, reviewed/Swissprot, downloaded on 3 Aug 2016; https://www.uniprot.org). This differentiated candidates into complete and valid genes, non-coding transcripts, pseudogenes, and transposable elements. In addition, we used PTREP (Release 19; http://botserv2.uzh.ch/kelldata/trep-db/index.html), a database of hypothetical proteins containing deduced amino acid sequences in which internal frameshifts have been removed in many cases. This step is particularly useful for the identification of divergent transposable elements with no significant similarity at the DNA level. Best hits were selected for each predicted protein from each of the three databases. Only hits with an e-value below 10e–10 were considered. Furthermore, functional annotation of all predicted protein sequences was done using the AHRD pipeline (https://github.com/groupschoof/AHRD).

Proteins were further classified into two confidence classes: high and low. Hits with subject coverage (for protein references) or query coverage (transposon database) above 80% were considered significant and protein sequences were classified as high-confidence using the following criteria: protein sequence was complete and had a subject and query coverage above the threshold in the

UniMag database or no BLAST hit in UniMag but in UniPoa and not PTREP; a low-confidence protein sequence was incomplete and had a hit in the UniMag or UniPoa database but not in PTREP. Alternatively, it had no hit in UniMag, UniPoa, or PTREP, but the protein sequence was complete. In a second refinement step, low-confidence proteins with an AHRD-score of 3* were promoted to high-confidence.

BUSCO (Seppey et al., 2019) (version 5.1.2.) was used to evaluate the gene space completeness of the pseudomolecule assembly and structural gene annotation with the 'viridiplantae_odb10' database containing 425 single-copy genes.

Orthogroup analysis

Orthogroup assignments (Table S4) was performed with OrthoFinder (Emms and Kelly, 2019). For GO term enrichment, a GO annotation file (gaf; 2.1) was built using all GO terms assigned by the functional annotations of *R. pubera*, *R. breviuscula*, *R. tenuis*, and *J. effusus*. GO term enrichment was performed by feeding GO terms of the shared orthologos into Ontologiser (ontologiser.de). *P*-values were corrected using the Benjamini-Hochberg procedure. We used the UpSetR (Conway et al., 2017) package (http://gehlenborglab.org/research/projects/upsetr/) to analyze how many orthogroups are shared between the five species or are unique to a single species.

De novo repeat discovery and annotation

To identify the overall repetitiveness of genomes we performed *de novo* repeat discovery with RepeatExplorer2 (Novák et al., 2020) for nine species of *Rhynchospora*, *C. littledalei*, and *J. effusus*. We used a repeat library obtained from the RepeatExplorer2 analysis of Illumina paired-end reads. All clusters representing at least 0.005% of the genomes were manually checked, and the automated annotation was corrected if needed. Contigs from the annotated clusters were used to build a repeat library. To minimize potential conflicts due to the occasional presence of contaminating sequences in the clusters, only contigs with average read depths ≥ 5 were included and all regions in these contigs that had read depths < 5 were masked. Genome assemblies were then annotated using custom RepeatMasker (REF - Smit, AFA, Hubley, R & Green, P. RepeatMasker Open-4.0. 2013-2015 http://www.repeatmasker.org) search with options -xsmall -no_is -e ncbi -nolow. Output from RepeatMasker was parsed using custom scripts (https://github.com/kavonrtep/repeat_annotation_pipeline) to remove overlapping and conflicting annotations.

Transposable element protein domains (Neumann et al., 2019) found in the assembled genomes were annotated using the DANTE tool available from the RepeatExplorer2 Galaxy portal. To find master *Helitron* elements related to *TCR1*, we first searched the genome assembly for *Helitron* helicase-coding sequences using DANTE (https://repeatexplorer-elixir.cerit-sc.cz/galaxy/) exploiting the REXdb database (Neumann et al., 2019) (Viridiplantae_version_3.0) and then manually identified boundaries of full-length *Helitron* elements. We identified 111 putative autonomous *Helitron*s and compared their terminal sequences with *TCR1*. This revealed that *TCR1* is most similar to the *Helitron-27*, sharing 90% and 100% identity over 30-bp sequences at the 5' and 3' ends, respectively (Figures 4J and 4K), meeting the criteria for classification of *TCR1* and *Helitron-27* into the same family (Thomas and Pritham, 2015). To find *TCR1* insertions in the *R. pubera* genome, we performed iterative blastn searches using 30-bp sequences from their 5' and 3' termini and consensus sequences of *Tyba*.

To obtain the average number of *Tyba* arrays for each *Rhynchospora* genome, we first removed spurious low-quality *Tyba* monomer annotations with less than 500 bp and merged with bedtools (Quinlan and Hall, 2010) all adjacent *Tyba* monomers situated at a maximum distance of 50 kb into individual annotations to eliminate the gaps that arise because of fragmented *Tyba* arrays. Length and distance between *Tyba* arrays were then calculated using bedtools. Bar plots of the average distance and unit length used to compare the *Tyba* arrays among the three *Rhynchospora* species were made in RStudio using ggplot2 (Wickham, 2016) library available at https://github.com/tidyverse/ggplot2.

Detection of dyad symmetries in Tyba repeats

Dyad simmetries detection was performed as reported in Kasinathan and Henikoff (2018). We used EMBOSS palindrome (Rice et al., 2000) to detect perfect dyad symmetries in the *Tyba* consensus of the three *Rhynchospora* species with the following parameters:

-minpallen 4 -maxpallen 100 -gaplimit 20 -nummismatches 0 –overlap

ChIP-seq analysis

Raw sequencing reads were trimmed by Cutadapt (Martin, 2011) to remove low-quality nucleotides (with quality score less than 30) and adapters. Trimmed ChIPed 150-bp single-end reads were mapped to the respective reference genome with bowtie2 (Langmead and Salzberg, 2012) with default parameters, where all read duplicates were removed and only the single best matching read was kept on the final alignment BAM file. BAM files were converted into BIGWIG coverage tracks using the bamCompare tool from deeptools (Ramírez et al., 2016). The coverage was calculated as the number of reads per 50-bp bin and normalized by reads per kilobase per million mapped reads (RPKM). Plots of detailed chromosome regions showing multiple tracks presented in Figures 4 and 6 were done with pyGenomeTracks (Lopez-Delisle et al., 2021).

CENH3 domains were identified by comparing the ChIPed and input data using MACS3 (Zhang et al., 2008). The parameters for MACS3 included -B –broad –g 1470000000 –trackline. As an alternative method for detection of CENH3 domains, we compared input and ChIP using the epic2 program for detection of diffuse domains (Stovner and Sætrom, 2019). Parameters for epic2 included—bin-size 2000. Only CENH3 domains detected with both methods were kept for further analysis.

To determine the sizes and positions of centromere units, we merged with bedtools CENH3 peaks that were separated by less than 50 kb to eliminate the gaps that arise because of fragmented *Tyba* arrays or due to insertion of TEs. Small CENH3 domains of less than 1 kb were discarded. Length and distance between *Tyba* arrays and between CENH3 domains were then calculated using bedtools. Bar plots of the average distance and unit length used to compare CENH3 domains and *Tyba* arrays were made in RStudio using the ggplot library.

The obtained repeat annotation was used to evaluate the association of individual classes of repetitive sequences with the CENH3 domain in *R. pubera*. For each repeat type, we calculated the total abundance in the genome as a sum of repetitive element length and compared it with abundance of repetitive elements located within CENH3 domains. For each type of repetitive element, we calculated the observed/expected ratio using:

$$OE = \frac{\sum(R_{CENH3})}{\sum\left(\frac{L_{CENH3}}{L_G}\right)R_G}$$

where R_{CENH3} is length of repeat located within CENH3 domains, L_{CENH3} is the length of CENH3-binding regions, L_G is total genome size, and R_G is total length of repeat type in the genome.

Identification of paralogous CENH3 domains

To identify groups of paralogous CENH3 domains within the blocks of homologous regions of *R. pubera*, we identified the two nearest paralogous genes on both sides of each CENH3 domain. Subsequently, the groups of four genes surrounding CENH3 domains were used to identify corresponding regions on the other homologous blocks where we checked for the presence of the CENH3 domain. Resulting groups of four homologous regions were manually inspected using dotplot (Krumsiek et al., 2007) and the IGV browser (Robinson et al., 2011).

Methyl-seq analysis

To comparatively evaluate the DNA methylation context of a holocentric and monocentric genome, we applied enzymatic methyl-seq and used the Bismarck pipeline (Krueger and Andrews, 2011) to analyze the data using 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 to BIGWIG format and used as input track for overall genome-wide DNA methylation visualization with pyGenomeTracks.

Metaplots

Analysis of the enrichment of all ChIP treatment files was performed as follows: BAM files of each ChIP treatment were normalized to the ChIP Input BAM file by RPKM using bamCompare available from deeptools. The generated normalized BIGWIG files were used to calculate the level of enrichment associated with gene bodies, *Tyba* repeats, CENH3 domains, and TEs using computeMatrix scale-regions (parameters: –regionBodyLength 4000 –beforeRegionStartLength 2000 –afterRegionStartLength 2000) also available from deeptools. Finally, metaplots for all ChIPseq treatment files were plotted with plotHeatmap available from deeptools (Ramírez et al., 2016). Additionally, coverage BIGWIG files of transcriptional activity (RNAseq) and all DNA methylation contexts were also used to calculated their enrichment on gene bodies, *Tyba* repeats, CENH3 domains, and TEs with computeMatrix and plotting with plotHeatmap.

Dating WGD events

To date the two rounds of duplication of the genome of *R. pubera*, a large tree of concatenated single copy genes was produced. For this analysis, each of the four homologous regions of *R. pubera* were separated and treated as a tip in the subsequent phylogeny reconstructions. Only coding sequences were used. We used BUSCO (*Poales* dataset) (Seppey et al., 2019) to look for conserved single-copy genes that are shared by all selected datasets. We performed this analysis in three different ways: solely the large syntenic block (Block1) of *R. pubera*, solely the smaller syntenic block (Block2) of *R. pubera*, and the two blocks combined. For the analyses, we included the following nine datasets: *J. effusus*, *C. littledalei*, *R. tenuis*, *R. breviuscula*, and the four homologous blocks of *R. pubera*. BUSCO analyses were run for all datasets; all the resulting single-copy genes were selected for each dataset. The single-copy genes shared among all datasets were used for the analyses: 841 for the larger Block1, 400 for the smaller Block2, and 1,204 for the two blocks combined. All genes were then aligned with MAFFT (Katoh and Standley, 2013), trimmed with Trimal (Capella-Gutiérrez et al., 2009), and concatenated into a single large multi-fasta alignment, and used as input for a ML tree built with IQ-tree2 (Minh et al., 2020).

A molecular clock analysis was performed to explore genome evolution in *Rhynchospora* and related genera. Divergence times were estimated using BEAST v.1.10.4 (Drummond and Rambaut, 2007) through the CIPRES Science Gateway fixing the tree topology from the Bayesian inference of the *Rhynchospora* concatenated 1,204 BUSCO gene alignment. Uncorrelated relaxed lognormal clock (Drummond and Rambaut, 2007) and Birth-Death speciation model (Gernhard, 2008) were applied. Two independent runs of 100,000,000 generations were performed, sampling every 10,000 generations. After removing 25% of samples as burn-in, the independent runs were combined and a maximum clade credibility (MCC) tree was constructed using TreeAnnotator v.1.10.4

(Drummond and Rambaut, 2007). To verify the effective sampling of all parameters and assess convergence of independent chains, we examined their posterior distributions in TRACER. The MCMC sampling was considered sufficient at effective sampling sizes (ESSs) equal to or higher than 200. The phylogeny was dated using both fossils and secondary calibration from published dated phylogenies. We chose three calibration points: i) Juncaceae/Cyperaceae divergence at 72.0 Mya (Bremer, 2002); ii) a fossil for *Carex* at 37.8 MYA (Smith et al., 2010), and iii) *R. pubera*/*R. tenuis* divergence at 32.0 Mya (Unpublished data).

Fluorescence in situ hybridization (FISH)

Interphase nuclei were prepared using the air-drying method, after enzymatic digestion with 2% cellulase Onozuka and 20% pectinase Sigma (Ribeiro et al., 2017). Roots were fixed in Carnoy ethanol:acetic acid 3:1 (v/v) for 2 h and stored at –20 °C. The best slides were selected for FISH, performed as described by Pedrosa et al. (2002) and the slides were counterstained with 2 μg/mL DAPI in Vectashield (Vector) mounting buffer. *Juncus effusus* interphase nucleus was hybridized with directly labeled (FAM)TTTAGGG(8)-telomeric probe and a directly labeled (CY3) probe for its most abundant satellite repeat, while *R. breviuscula* nucleus was hybridized with the same telomeric probe and directly labeled *Tyba* (CY3) oligo-probe.

Immunostaining

Immunostaining was performed as described before by Marques et al. (2016) with some modifications. Roots or Anthers were fixed in ice-cold 4% paraformaldehyde in 1× PBS buffer pH 7.5 (1.3 M NaCl, 70 mM Na2HPO4, 30 mM NaH2PO4) for 1 hr and 30 min and squashed in a drop of the same buffer. Then, the slides were washed in 1× PBS and blocked with 3% BSA for 30 min at 37 °C. Rabbit anti-H3K4me3 (abcam, ab8580), mouse anti-H3K9me2 (abcam, ab1220), and previously generated *R. pubera* rabbit anti-CENH3 (Marques et al., 2015) were used for immunostaining. The slides were counterstained with 2 mg/ml 49,6-diamidino-2-phenylindole (DAPI) in Vectashield H-1000.

QUANTIFICATION AND STATISTICAL ANALYSIS

Comparison of Hi-C contacts

The chromosomal interactions between holo- and monocentric plant species were compared by the ratios of *cis* and *trans* Hi-C contacts, i.e., for each species, we quatified the ratios of *cis* and *trans* Hi-C counts for every chromosome and tested if they were significantly different across distinct species. For grouped comparison, we adopted the mutiple testing method of one-way ANOVA (Analysis of Variance), specifically the Kruskal-Wallis ranked test with Holm-Bonferroni correction, because the compared values and ratios of intra- and inter-chromosomal contacts were different in length among various species and were not supported by evidence such as normality. Pair-wise significance analysis was conducted using Dunn's *post hoc* test.

Tyba array and CENH3 domain size and spacing

The Dunn's test was used to compare pairwise distributions of values of interest between *Tyba* arrays and CENH3 domains size and spacing.

Supplemental figures

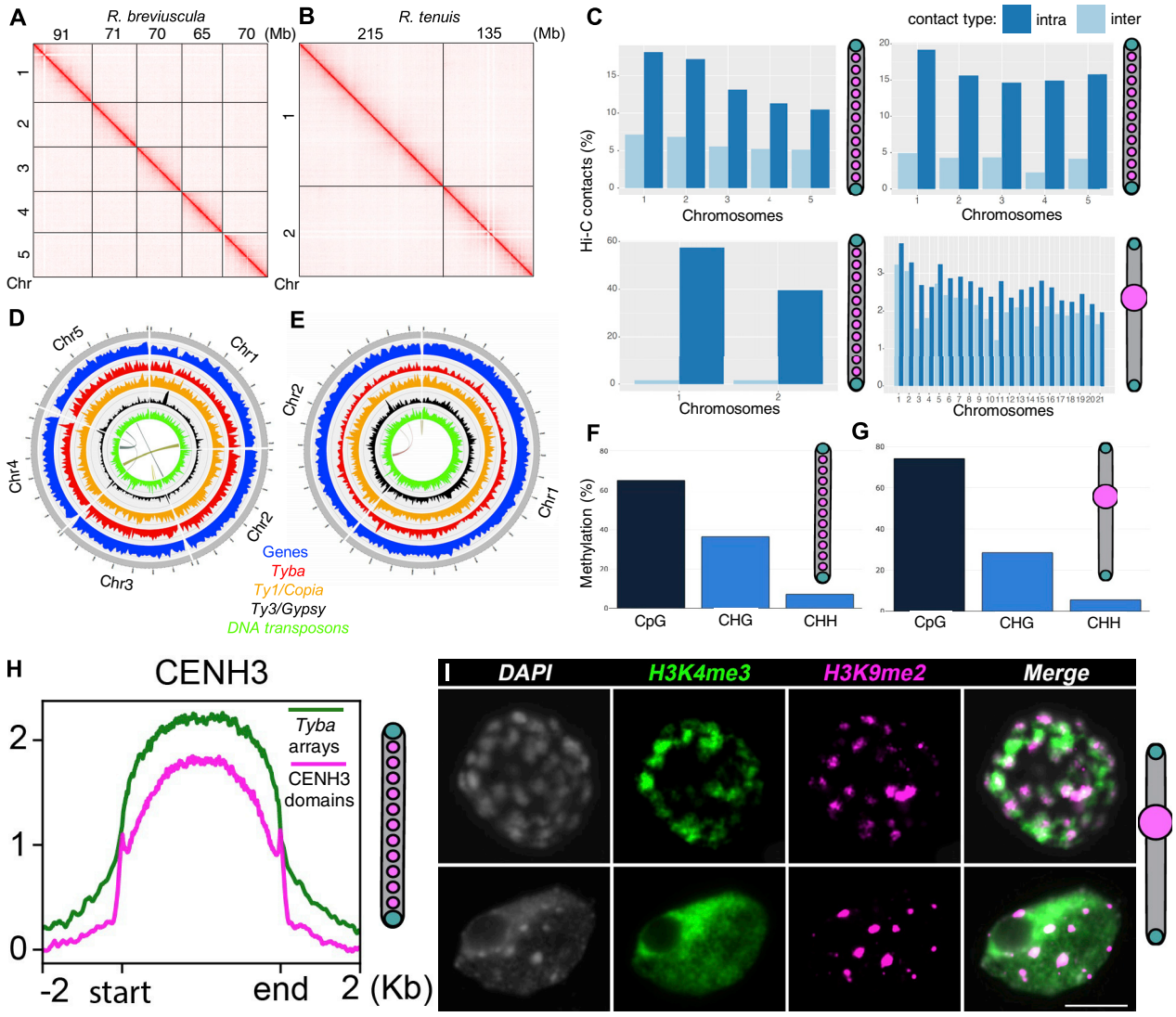


Figure S1. Characterization of the *Rhynchospora* and *J. effusus* genomes, related to Figures 2 and 4

(A and B) Contact maps for the five assembled pseudochromosomes of *R. breviuscula* (A) and the two assembled pseudochromosomes of *R. tenuis* (B). The intensity of pixels represents the normalized count of Hi-C links between 500-kb windows on a log scale.

(C) Hi-C contact counts (bin size, 1 Mb, normalization, VC) of intra- (*cis*) and interchromosomal (*trans*) chromatin contacts in the four species showing a significantly higher ratio ($p < 4.04e{-}05$) in holocentric compared with monocentric species, which implies relatively enriched *trans* interactions in the latter species.

(D and E) Distribution of the main classes of sequence types in *R. breviuscula* (D) and *R. tenuis* (E) with a 1-Mb sliding window. Note the high peaks of LTR *Ty3/Gypsy* density at most subtelomeric regions in *R. breviuscula* chromosomes. Self-synteny of *R. breviuscula* (D) and *R. tenuis* (E) genomes is shown in the inner circle.

(F and G) Summary of genome-wide DNA methylation contexts in *R. pubera* (F) and *J. effusus* (G).

(H) Metaplot showing the enrichment of CENH3 on *Tyba* repeat arrays (green) and CENH3 domains (magenta) in *R. breviuscula*.

(I) Immunostaining of metaphase chromosomes and an interphase nucleus of *J. effusus* for H3K4me3 and H3K9me2. Scale bars, 5 μm.

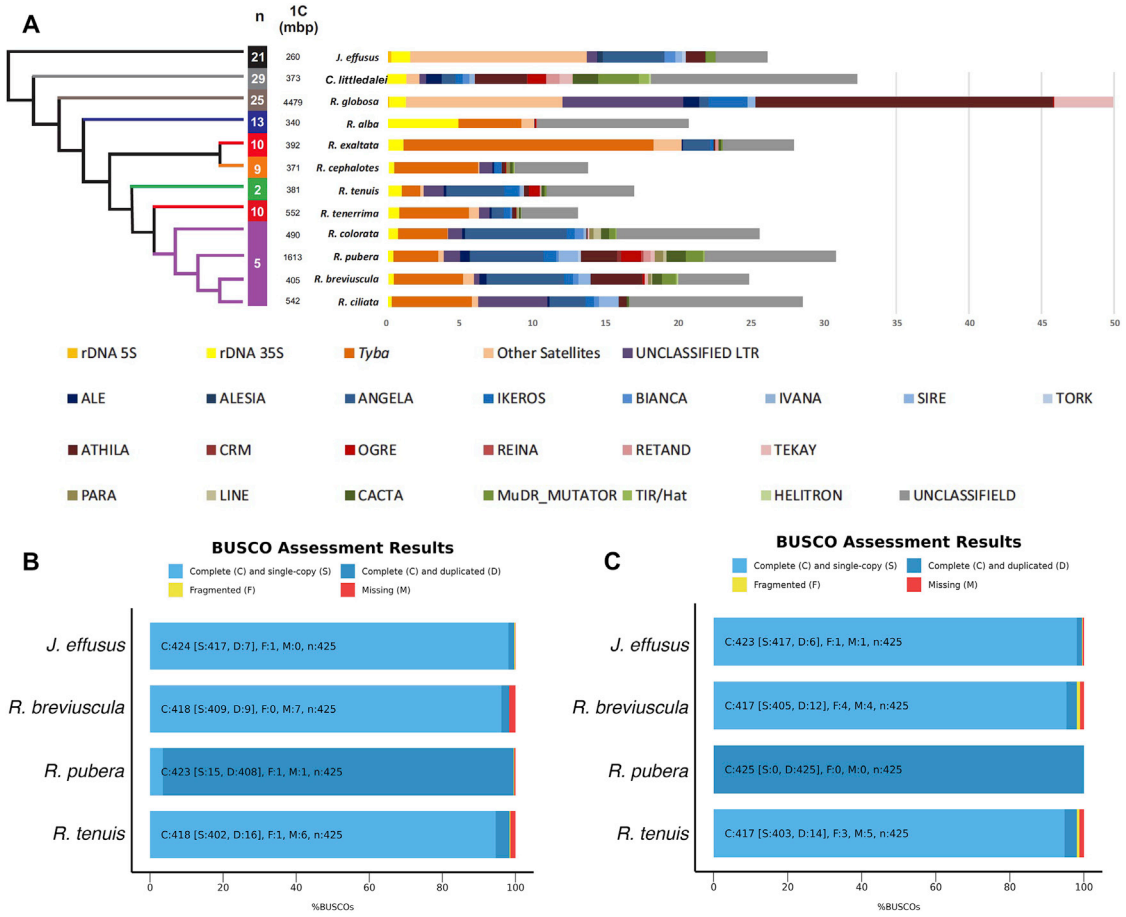


Figure S2. Composition and evolution of sedges and rush genomes, related to Figure 5

(A) Schematic phylogenetic tree and repeat composition of beak-sedge genomes and comparison with *C. littledalei* and *J. effusus*.

(B and C) BUSCO assessment for completeness of genic space with the viridiplantae_odb10 dataset, using the entire genome assembly (B) or the longest transcript (C).

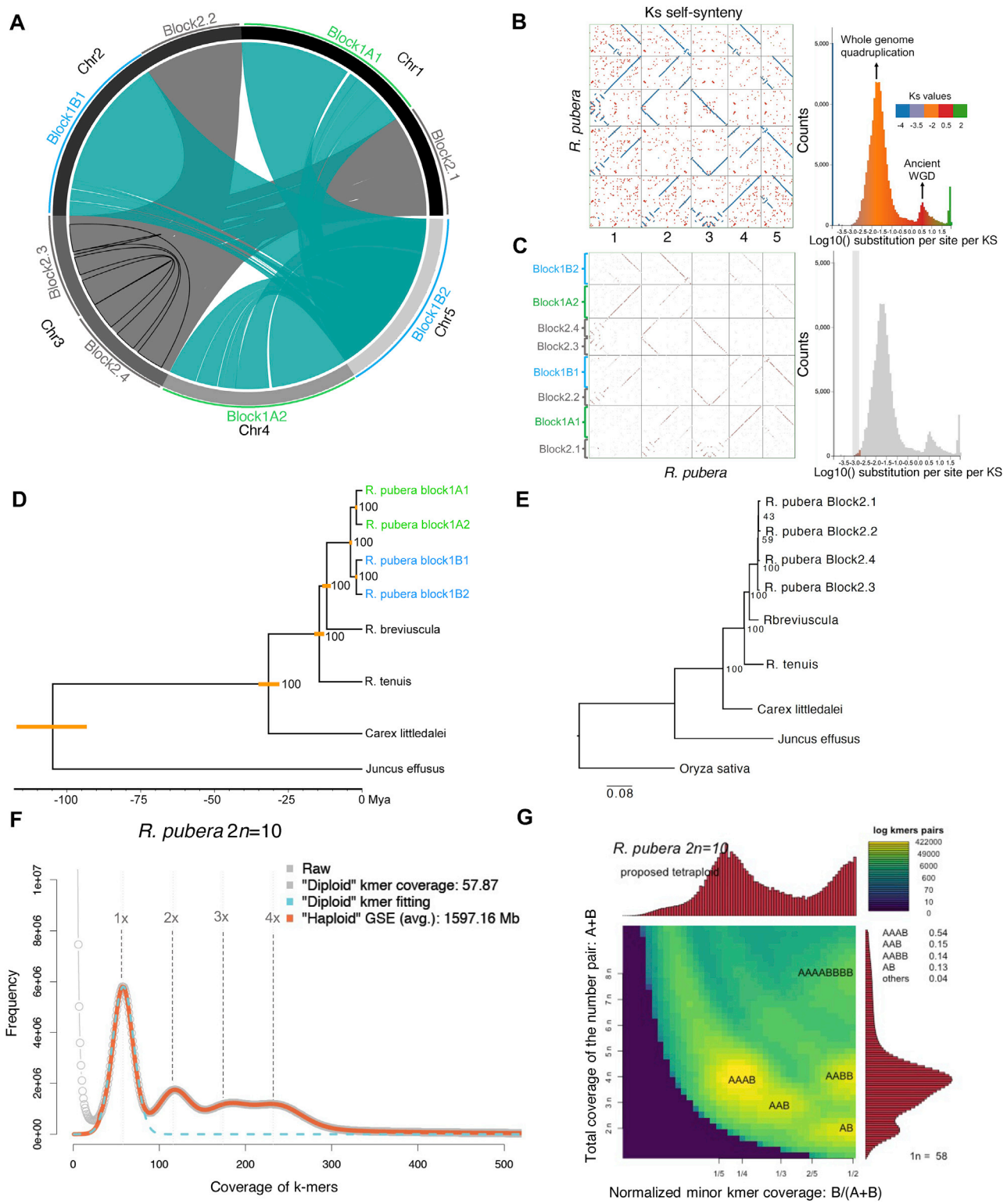


Figure S3. Identification, characterization, and dating of WGDs in *R. pubera*, related to Figure 5

(A) SyMap self-synteny plot of *R. pubera*. Block structure is indicated by outer arcs.

(B) SyMap self-synteny dot plot colored based on Ks values. Ks values on a log scale are shown to the right of the dot plot. Note the large peak that correlates with the large duplication events in *R. pubera* and a second small peak most likely representing an ancient WGD.

(legend continued on next page)

(C) Same plot as (B) but selecting only the sequences with the lowest number of synonymous substitutions, allowing the identification of intragenomic syntenic block relationships (Block1A and Block1B). We were unable to detect any relationships for Block2. The small colored block within the vertical gray bar represents the sequences with the lowest number of synonymous substitutions used in the dot-plot to the left. Ks values are indicated by the color scale in (B).
(D) Based on the assessment of the relationships among the syntenic blocks of *R. pubera*, we selected 1,204 BUSCO genes (Poales dataset) uniquely present in each block and also shared with *R. breviuscula*, *R. tenuis*, *C. littledalei*, and *J. effusus* to build a phylogenetic tree from a concatenated alignment, which was further used for dating the duplication events in *R. pubera*. We confirmed the Block1A and Block1B relationships with 100% bootstrap support and also determined that a first WGD occurred around 3.8 Mya, followed by a second event around 2.1 Mya. Note that the second WGD closely overlaps in both Block1A and Block1B branches. Yellow bars indicate the dating time interval.
(E) Phylogenetic analysis of Block2 genes did not resolve the relationships for this particular block and was not used for dating.
(F and G) K-mer based estimation of genome size and heterozygosity (F) and smudgeplot analysis of k-mer-based ploidy inference for *R. pubera* using 21-mers (G). GSE, genome size estimation. Smudgeplot infers ploidy directly from the k-mers present in sequencing reads by analyzing heterozygous k-mer pairs.

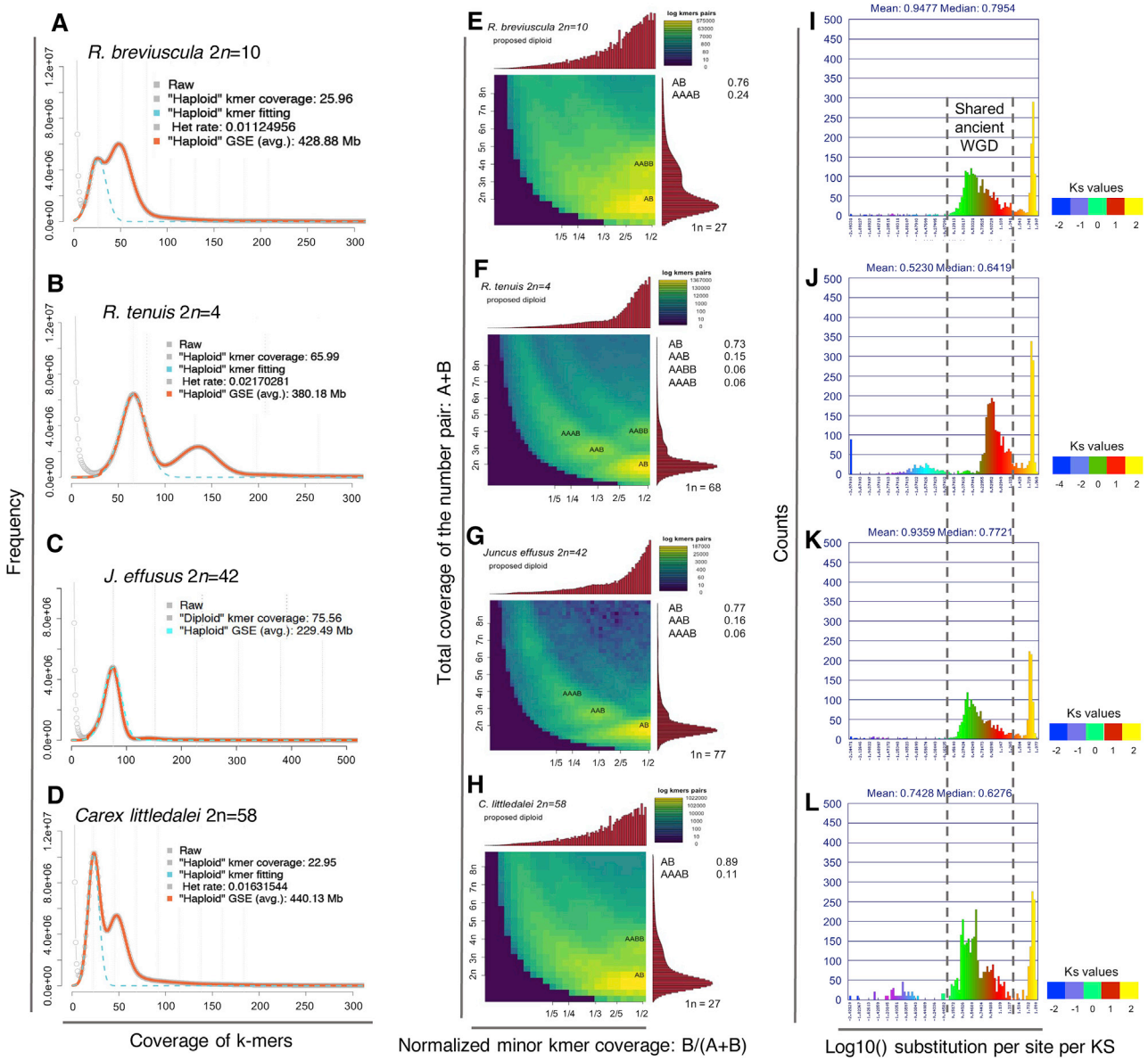


Figure S4. K-mer based genome size estimation and ploidy inference and WGD identification in sedges and rushes, related to Figure 5

(A–D) 21-mer based estimation of genome size and heterozygosity. GSE, genome size estimation.

(E–H) Ploidy and genome structure inference based on 21-mer Smudgeplot analysis.

(I–L) Ks values of coding sequences for each genome; a shared ancient WGD peak was observed for all species.

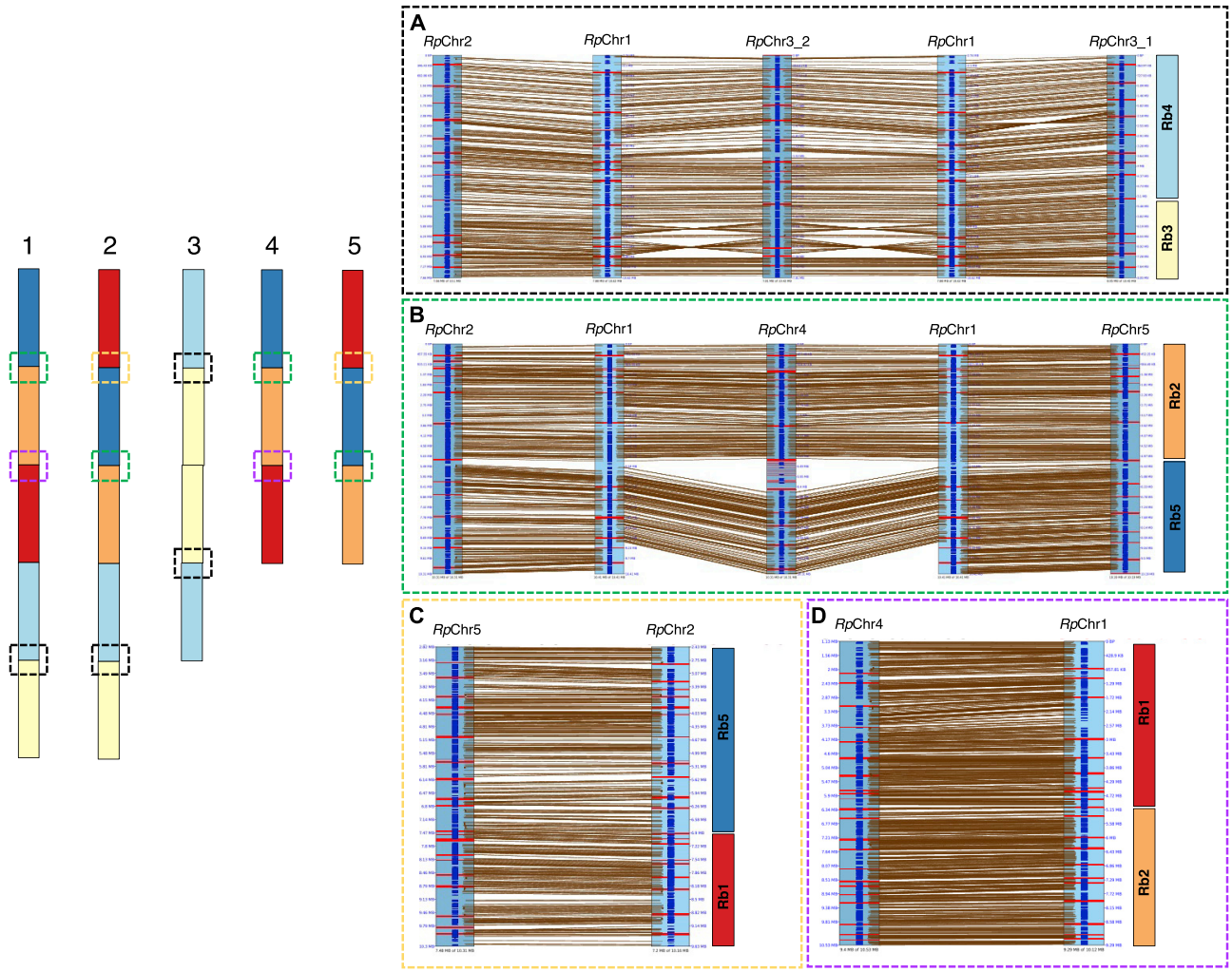


Figure S5. Comparative alignment of the duplicated end-to-end fusion (EEF) transition regions in the *R. pubera* genome, related to Figure 5

(Left) Ideogram model of *R. pubera* chromosomes, with the dashed boxes indicating the extracted and compared regions on the right.

(A) Alignment of the EEF of *Rb*3 and *Rb*4 found once on *Rp*Chr1 and *Rp*Chr2 and twice on *Rp*Chr3, showing the same fusion signature.

(B) Alignment of the EEF of Rb2 and *Rb*5, found on *Rp*Chr1, *Rp*Chr2, *Rp*Chr4, and *Rp*Chr5, also showing the same fusion signature.

(C) Alignment of the EEF of *Rb*1 and *Rb*5, found on *Rp*Chr2 and *Rp*Chr5 with the same fusion signature.

(D) Alignment of the EEF of *Rb*1 and Rb2, found on *Rp*Chr1 and *Rp*Chr4 with the same fusion signature. Colored boxes assign the synteny to *R. breviuscula* chromosomes. Red stripes on the synteny alignments depict *Tyba* repeats, while genes are annotated in dark blue.

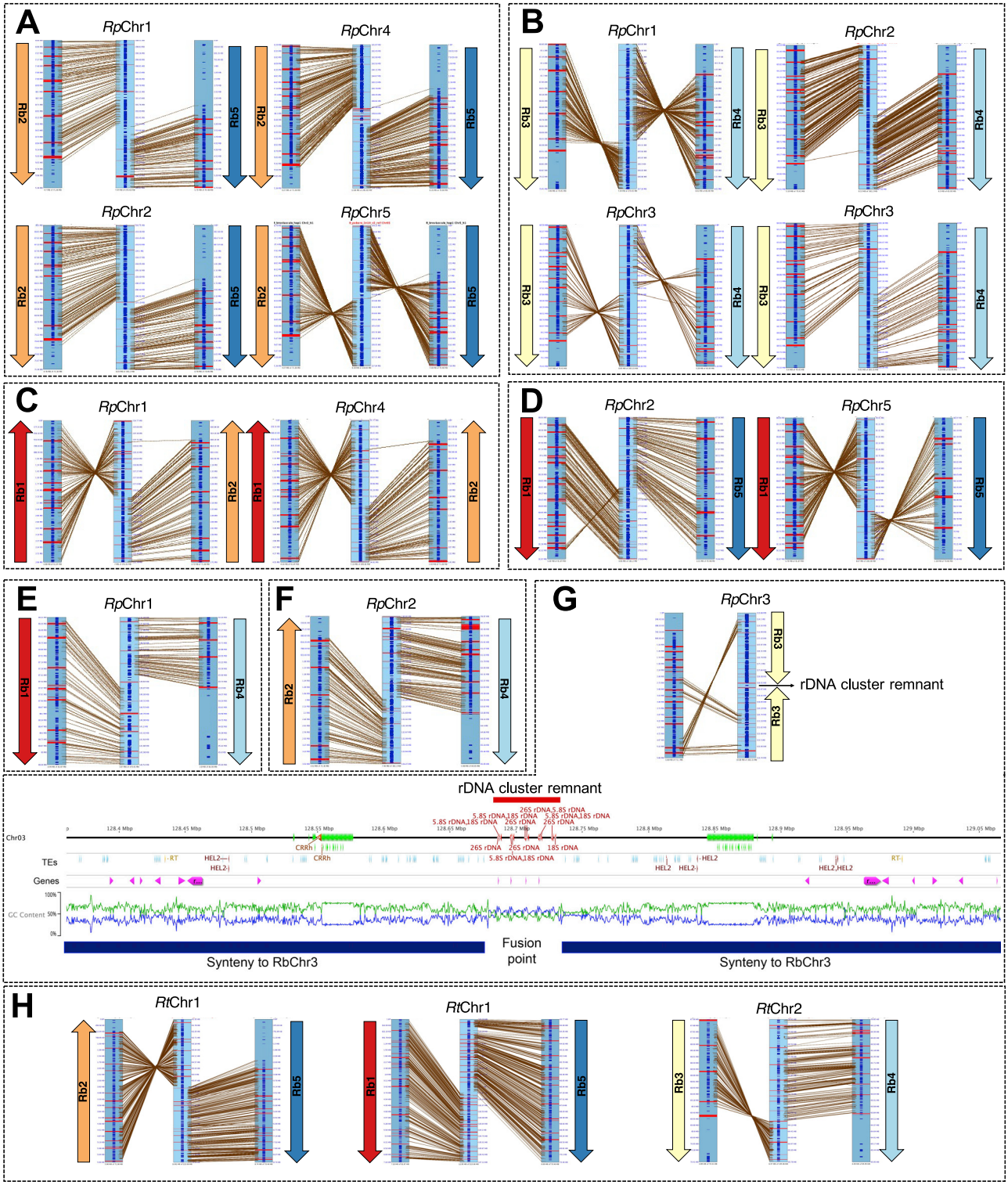


Figure S6. Identification of the sequences underlying the transitions between the syntenic regions to *R. breviuscula* chromosomes in the end-to-end fusions found in the *R. pubera* and *R. tenuis* genomes, related to Figure 5

(A) EEF of Rb2 and *Rb*5 found on *Rp*Chr1, *Rp*Chr2, *Rp*Chr4, and *Rp*Chr5. Similar fusion signatures are shared among the four chromosomes. In three of them, a *Tyba* repeat is found between them.

(legend continued on next page)

(B) EEF of *Rb*3 and *Rb*4 found on *Rp*Chr1 and *Rp*Chr2 and twice on *Rp*Chr3 with the same fusion signature. A *Tyba* repeat array is found between the transitions in all cases.
(C) EEF of *Rb*1 and Rb2 found on *Rp*Chr1 and *Rp*Chr4 with the same fusion signature, without a *Tyba* repeat in between.
(D) EEF of *Rb*1 and *Rb*5 found on *Rp*Chr2 and *Rp*Chr5 with the same fusion signature, with a *Tyba* repeat in between.
(E) EEF of *Rb*1 and *Rb*4 found only on *Rp*Chr1 with a *Tyba* repeat array in between.
(F) EEF of *Rb2* and *Rb*4 found only on *Rp*Chr2 with no *Tyba* repeat in between.
(G) EEF of *Rb*3 and *Rb*3 found only on *Rp*Chr3 and with a remnant of a rDNA cluster in the transition region (with detailed annotation shown to the right).
(H) Characterization of the three EEFs responsible for the chromosome reduction in *R. tenuis*. On *Rt*Chr1 we found an EEF involving Rb2 and *Rb*5, and a second event involving *Rb*5 and *Rb*1, while on *Rt*Chr2, we found a single EEF involving *Rb*3 and *Rb*4. Colored arrows indicate the *R. breviuscula* chromosomes and point to the telomeric region involved in the fusion event. Remarkably, although similar ancestral chromosome associations are found in *R. pubera* and *R. tenuis*, the chromosomal ends involved in the fusions are different. Red stripes on the synteny alignments depict *Tyba* repeats, while genes are annotated in dark blue.

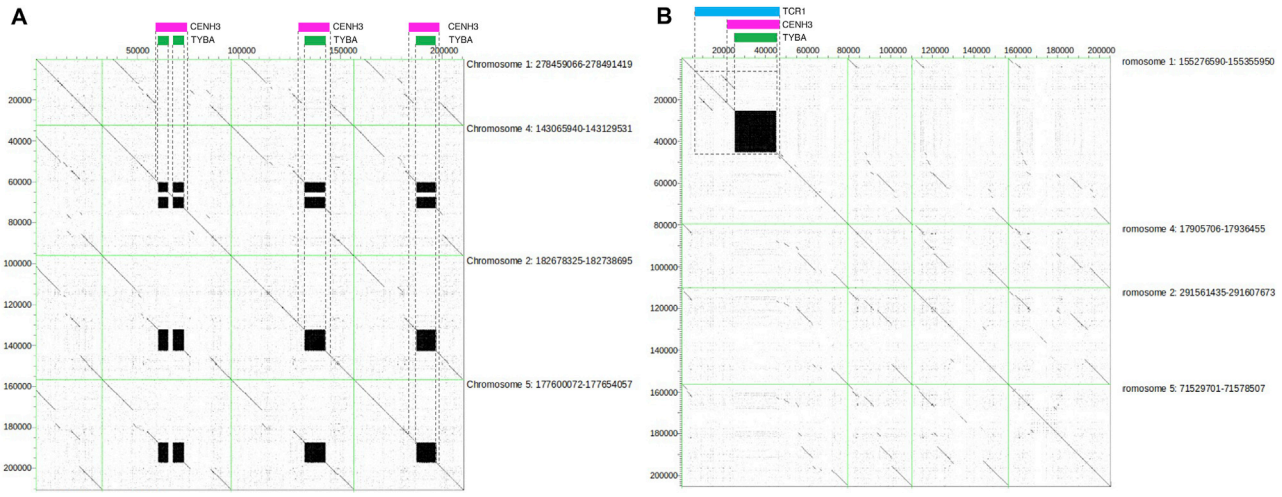


Figure S7. Characterization of emergence and loss of CENH3-binding regions in *R. pubera*, related to Figure 6

(A) Example of CENH3-binding region and *Tyba* array lost in one of four paralogous regions, while the other three copies retained the *Tyba* array and CENH3 binding. The conserved locus is indicated by the dashed box, along the x axis of the dot plot, with rectangles marking the area associated with CENH3 (magenta) and the *Tyba* array (green). The genome positions of the extracted regions are given to the right.

(B) Example of CENH3-binding region and *Tyba* array gain in one of four paralogous regions due to a transposition of *Tyba*-containing *TCR1* in *Rp*Chr1, while the other three copies lack the *Tyba* array. The gained locus is indicated by the dashed box, along the x axis of the dot plot, with rectangles marking the *TCR1* element (blue), the area associated with CENH3 (magenta), and the *Tyba* array (green). The genome positions of the extracted regions are given to the right.

Chapter 2

Cohesion dynamics during inverted meiosis

Gokilavani Thangavel[1] and André Marques[1]

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

Abstract

Canonical meiosis is characterized by a reductional segregation of homologous chromosomes during meiosis I followed by the equational segregation of sister chromatids during meiosis II. However, some plants harbour multiple centromere domains distributed along the chromosomes, i.e., holocentric chromosomes, which prompted them to evolve a different strategy. This is the case found in the holocentric plant genus *Rhynchospora*, in which equational segregation in meiosis I is followed by reductional segregation during meiosis II, i.e., inverted meiosis. Since the two-step cohesion loss is important during canonical meiosis, the meiotic cohesin REC8 and the cohesion protector SGO are likely to be related to the adaptation to inverted meiosis. Here I studied the dynamics of REC8 and SGO in *R. pubera* and *R. breviuscula.* The cytological pattern of REC8 expression showed a linear line-like pattern conserved in canonical meiosis during prophase I, while no colocalization with centromeres was observed in metaphase I. This observation suggests that, indeed, sister centromere cohesion is absent at metaphase I in these species, which seems to be a key factor to enable inverted meiosis. Moreover, Shugoshin is localised as multiple foci on chromosomes that do not colocalise with centromeres either, hinting at a possible non-centromeric role for SGO. However, further functional characterisation is needed to be certain. These findings, support the hypothesis that inverted meiosis occur due to the absence of sister centromere cohesion at the metaphase I.

Introduction

Centromeres appear as constricted regions on chromosomes and are the sites where kinetochores assemble and spindle microtubules attach to the chromatids; these structures also enable the movement of chromatids during cell division (Jiang et al., 2003). According to the distribution of centromeres, chromosomes can be classified as monocentric (a centromere at a single location) or holocentric (centromeres distributed along the chromosomes). Holocentricity has evolved independently among different animal and plant lineages (Melters et al., 2012). Holocentricity poses difficulties to segregate the chromosomes during meiosis. Different organisms have evolved different strategies to deal it. The three known strategies are (i) Chromosome remodelling as observed in the most common holocentric model *Caenorhabditis elegans* (Lui and Colaiácovo, 2013); (ii) Functional monocentricity as observed in different insect species of Heteroptera and *Parascaris* (Goday and Pimpinelli, 1989; Pe´rez et al., 1997); (iii) Inverted meiosis, which is observed among both plants (Cyperaceae and Juncaceae) and animals (*Planococcus citri*) (Cabral et al., 2014; Chandra, 1962; Heckmann et al., 2013). During meiosis I of canonical meiosis in a monocentric species, sister kinetochores attach to microtubules emanating from the same direction. As a result, the homologs segregate to opposite poles, whereas the sisters are kept together (reductional division). During meiosis II, the sister kinetochores attach to microtubules from opposite directions; thus, the sisters segregate to opposite poles at the end of meiosis II (equational division). However, during meiosis I of inverted meiosis, the sister chromatids bi-orient and segregate to two different poles (equational division). In meiosis II, some unknown mechanism is involved in aligning and segregating the homologues to the opposite poles (reductional division). Thus, in the case of inverted meiosis, equational segregation precedes reductional segregation. To enable this to happen, scientists have proposed three essential requirements: 1. Bipolar orientation of sister kinetochores and their attachment to microtubules from opposite spindle poles in meiosis I (amphitelic attachment); 2. Segregation of sister chromatids to opposite poles in anaphase I; 3. Mechanism to align and distribute homologous non-sister chromatids during the second meiotic division (Cabral et al., 2014; Heckmann et al., 2014a; Heckmann et al., 2014b).

Sister chromatid cohesion (SCC) is essential for the proper segregation of chromosomes during mitosis as well as meiosis in eukaryotic organisms (Orr-Weaver, 1999; van Heemst and Heyting, 2000). SCC not only holds the sister kinetochores in

position but also most importantly, prevents the microtubules from tearing the bivalents apart, thereby ensuring proper bi-orientation (Nasmyth and Haering, 2009). SCC is established by the highly conserved cohesin complex consisting of four subunits: SMC1, SMC3, α-kleisin subunit (SCC1/RAD21/MCD1) and SCC3 (SA1/SA2). Studies demonstrate that SCC1 binds to SMC1 and SMC3 to form a tripartite ring and this ring is proposed to encircle the sister chromatids and mediate the cohesion mechanism (Kurze et al., 2011; Nasmyth and Haering, 2009). The establishment and maintenance of cohesion is a highly regulated two-step process. First, 90% of the cohesion dissociates from the chromosome arms during prophase and pro-metaphase due to the action of a Wings apart-like protein homolog (WAPL)-dependent prophase pathway which re-associates with the chromosomes during telophase. Shortly before the onset of anaphase, the remaining cohesion at centromeres is lost through the cleavage of the phosphorylated α-kleisin subunit by the APC/C-dependent separase pathway, which prevents α-kleisin from re-associating with chromosomes. During meiosis, the separase pathway is needed to cleave both the α-kleisin of chromosome arms to resolve chiasmata and the α-kleisin around centromeres to release centromeric cohesion (Nasmyth and Haering, 2009).

Deeper insights into the mechanism of inverted meiosis depend on an understanding of the differences between chromosome structure during inverted meiosis and canonical meiosis. Due to the importance of cohesion genes in the two-step cohesion loss, in holocentric plants, these genes may play a major role in the mechanism of inverted meiosis. The meiotic specific cohesin REC8 and the cohesion protector, SGO are chosen as the candidates for this study. REC8, the meiotic α-kleisin was identified in *Schizosaccharomyces pombe* and several other organisms (Bai et al., 1999; Lin et al., 1992). REC8 mutants have defects in chromosome cohesion and condensation, resulting in chromosome fragmentation and formation of polyads. REC8 localises to the arms of the chromosomes (from interphase to anaphase I) and is also responsible for centromeric cohesion in the later stages of meiosis (Cai et al., 2003). Shugoshins (Mei-S332 in *Drosophila melanogaster*) along with phosphatase 2 A (PP2A) are essential for protecting cohesion at different stages by keeping α-kleisin dephosphorylated. SGO1 is needed to protect cohesins from the prophase I pathway and SGO2 protects centromeric cohesins from separase during meiosis I in humans (Llano et al., 2008; McGuinness et al., 2005; Salic et al., 2004). In this study, I explore the centromeric cohesion during meiosis I of holocentric plants to decipher the molecular mechanisms behind the evolution of inverted

meiosis. The results from REC8 localisation studies strongly suggests the absence of sister centromeric cohesion at metaphase I of inverted meiosis. SGO dynamics hints at a possible non-canonical, non-centromeric function in these species.

Materials and methods

Phylogeny analysis

Using protein-BLAST, one orthologue for both REC8 and SGO was found in *R. pubera*. Orthologs from related species were downloaded from the public databases UniProtKB or TAIR. To understand the conservation of these sequences, Geneious Prime software (https://www.geneious.com/) was used to create the multiple sequence alignment using the parameters MAFFT alignment and BLOSUM 62 matrix. For phylogeny analysis, the sequences were aligned by MAFFT (Katoh and Standley, 2013), trimmed by trimAl (Capella-Gutierrez et al., 2009) and the phylogenetic tree was constructed using IQ-TREE (Nguyen et al., 2015). The tree was manually interpreted using SeaView (https://doua.prabi.fr/software/seaview).

Western blot analysis

Newly generated antibodies for REC8 were tested by western blot analysis. Protein isolation was carried out from leaves (somatic tissue) and young flower buds (containing meiotic cells) following the protocol from (Nayar et al., 2013) with a few modifications. Total protein was extracted from around 1 g of young leaf tissues and 500 mg of unopened flower buds. The tissues (along with 1 g PVP) were ground using liquid nitrogen in lysis buffer containing 62.5 mM Tris-HCl buffer (pH 7.5), 2% (w/v) SDS, 30% (v/v) glycerol, 0.1 M beta-mercaptoethanol (ß-ME), 1.5 mM PMSF, 1 mM EDTA and 10 µL/mL Plant Protease Inhibitor from Sigma. The ground mixture was incubated on ice for 30 min. The homogenate was then centrifuged and the supernatant was transferred to a fresh tube. Protein was precipitated from this homogenate with chilled acetone containing 10 mM ß-ME. Proteins were pelleted and dissolved in Laemmli Buffer (0.5 M Tris-HCl, pH 6.8, 30% glycerol, 10% SDS, 0.5% bromophenol blue and 0.71 M ß-ME) and subsequently boiled for 5 min at 95 °C; supernatants were transferred to fresh tubes. Around 2.5 to 10 µL of protein extract were loaded onto 15% polyacrylamide SDS gels for electrophoresis at 130 V using a Mini-PROTEAN Tetra cell apparatus (Bio-Rad, USA). Blotting was performed in a semi-dry mini trans blot apparatus (Bio-Rad, USA) at 18 V for one hour. Rabbit anti-

RpREC8 (1:2,500 dilution) was used as the primary antibody. Rabbit anti-H3 (abcam-ab1791) was used as nuclear loading control. Anti-Rabbit (1:5,000 dilution) was used as secondary antibody. Immunodetection was carried out using the ECL Plus kit (Amersham, GE Healthcare Life Sciences) as per the manufacturer's protocol. Based on the intensity of anti-H3 band, anti-REC8 immunoblot was performed mimicking the loading volume.

Immunocytochemistry and fluorescence in situ hybridization (FISH)

Chromosome preparation and immunostaining experiments were performed in *R. breviuscula* and *R. pubera* as described in (Cabral et al., 2014). The primary antibodies used in this study were (1) anti-RpREC8 raised in rabbit against the peptides YNPDDSVERMRDDPG and EEPYGEIQISKGPNM by Eurogentec, (2) anti-RpSGO raised in rat against the peptides KLDDRKPIRRQSIAK and SYREQPVNVKMRRDO by Eurogentec, (3) anti-RpCENH3 raised in rabbit (Marques et al., 2015) and (4) anti-AtZYP1 raised in chicken against the peptide EGSLNPYADDPYAFD by Eurogentec. STAR ORANGE and STAR RED (Abberior) were the secondary antibodies used. To understand REC8 dynamics around centromeres, FISH against centromeric repeat (*Tyba*) probes was performed on the slides with the strongest immunodetection of REC8 as in (Braz et al., 2020). The slides were counterstained with 2 mg/ml 4,6-diamidino-2-phenylindole (DAPI) in Vectashield H-1000. Images were taken using a Zeiss Axio Imager Z2 with Apotome system for optical sectioning, deconvolved and processed with Zen 3.2.

Construction of CRISPR-Cas9 vectors

For further characterization of REC8 and SGO, CRISPR-Cas9 constructs were generated to mutate the candidate genes. The vector system currently used for barley in the lab of Dr. Ivan F. Acosta, Max Planck Institute for Plant Breeding Research, was adapted for our model species *R. pubera* (Kumar et al., 2018; Ordon et al., 2017). Two constructs were designed to target each gene. The *RpRec8* construct 1 has the sgRNAs, 5´-TGGCAGCAACGATGCGCTCG**AGG**-3´ and 5´-TATGTACCCTGCAGCGCCAA**TGG**-3´ targeting exon 2 and exon 3, respectively. This construct targets the RAD21/REC8 N-terminal domain of REC8. The *RpRec8* construct 2 has a single sgRNA, 5´-CCCTCAGTCAGTTAGCTCAG**GGG**-3´ targeting exon 19, coding for the C-terminal domain of REC8. *RpSgo* construct 1 has the sgRNAs 5´-AGGGTGCTCGTCAGCAGAAC**TGG**-3´ and 5´-AGGGTGCTCGTCAGCAGAAC**TGG**-3´ targeting exons 4 and 6, respectively, resulting in

targeting of the SGO N-terminal and coiled coil domains. *RpSgo* construct 2 has a single sgRNA, 5´-ACTGCGGCGAGCTGTGGAGA**AGG**-3´ targeting exon 13, encoding the C-terminal domain of SGO. CRISPR-Cas9 vectors were assembled in two steps by the Golden Gate system. The sgRNA Transcriptional Units (TUs) were first assembled in a shuttle vector system (pMGE625 and pMGE627 in case of 2 sgRNAs; pMGE624 in case of 1 sgRNA). These sgRNA TUs were then cloned into the recipient vector pMGE599, a plant binary vector with hygromycin as the plant selectable marker. Cas9 is under the control of the maize ubiquitin promoter and the gRNA transcriptional unit is driven by barley U3 promoter. All the bacterial transformations were carried out in *Escherichia coli* DH5α cells. The plant binary vector was finally cloned into *Agrobacterium tumefaciens* AGL1 strain.

Plant transformation

Immature seeds (green or green with brown spots) were used as the explant. The seeds were harvested and surface-sterilized with saturated trisodium phosphate + 0.1% Tween 20 for 30 min followed by 5% sodium hypochlorite + 0.1% Tween 20 for 1 hr. White and healthy embryos were dissected out of the immature seeds with the help of a binoculars under sterile conditions. Callus was induced at 28 °C, in the dark for two weeks. *Agrobacterium* culture (OD600 < 0.6) harbouring the plant binary vector pMGE599 with sgRNA(s) of our interest + acetosyringone, was used to infect the 2-week-old calli followed by co-cultivation for two days in the dark. After two rounds of selection under hygromycin 50mg/mL concentration, the white and healthy, proliferating calli were transferred to regeneration media. Once the shoots and roots regenerated, the young plants were transferred to soil (Figure 4c). The detailed media composition is available in appendix.

Genotyping

Putative transformants were tested for the plant selectable marker site hygromycin and Cas9 by PCR. hph_F (5'-CTCGGAGGGCGAAGAATCTC-3') and hph_R (5'-CTCCAGTCAATGACCGCTGT-3') were the hygromycin primers and JS1520/Cas9_FP (5'-CCTCTCTCTTGGCCTCACC-3') and JS1519/Cas9_RP (5'-CTTCTCGTTCTGGAGCTGG-3') were the Cas9 primers used. Taq polymerase was used for the PCR using an Eppendorf machine with the following cycling conditions for both targeted marker sites: 95 °C for 2

min as initial denaturation, 95 °C for 20 sec as denaturation, 64 °C for 30 sec as annealing, 72 °C for 2 min as extension and 72 °C for 5 min as final extension. The final PCR products were run in a 1% gel at 80 V and documented using a Biorad Gel Doc XR+ gel documentation unit. To confirm whether the plants were successfully edited at the targeted site, next generation sequencing (NGS) was performed. Primers were designed to amplify a 150-bp region around the sgRNA targeted site for 28 *Sgo* C-terminal transformant plants. The target region was amplified using Phusion polymerase and a second PCR was performed to add adapters to the amplicon ends for NGS. The pooled mixture was sent for Amplicon EZ sequencing at GENEWIZ, and the results were analysed for each individual plant.

Results and Discussion

REC8 signals are conserved as linear line pattern at pachytene, whereas centromere signals are lost during metaphase I

Based on the homology and phylogeny searches, we could detect one ortholog for REC8, homologous to *Arabidopsis* SYN1 in *Rhynchospora* (**Fig. 1a)**. Multiple sequence alignment among REC8 proteins from other species shows that the domains described for REC8 – the N-terminal REC8/RAD21-like domain belonging to the Pfam family PF04825 and the C-terminal REC8/RAD21-like domain belonging to the Pfam family PF04824 are also conserved in *Rhynchospora* (**Fig. 1b**). REC8, the meiotic α-kleisin, is reported to be specifically expressed only during meiosis (Lin et al., 1992). Western blot analysis from leaf (somatic tissues) and young flower buds (containing meiotic cells) could also reveal a signal of expected band size (around 67 kDa) only in the flower buds, which shows the meiotic specificity of REC8 in *Rhynchospora* (Figure 2a). REC8 localisation studies carried out in different organisms in the past have shown that it is present from premeiotic S phase until anaphase II. During prophase I, most of the signals are lost, reflecting a resolution of chiasmata in order to release the homologous chromosomes, whereas signals are present at the centromeres until anaphase II (Cai et al., 2003; Klein et al., 1999; Pasierbek et al., 2001; Watanabe and Nurse, 1999). Detailed immunocytological analysis of REC8 dynamics during meiotic phases revealed that REC8 signals start appearing as early as interphase in *Rhynchospora*. During prophase I, when the chromosomes pair and synapse, REC8 forms the backbone of the synaptonemal complex. As a result, when the chromosomes are fully synapsed at pachytene, REC8 forms a linear line-like pattern (Chelysheva et al., 2005) and co-localises with the synaptonemal complex protein ZYP1,

which was also observed in *Rhynchospora* (**Fig. 2 b–g, k–m**). After pachytene, most of the cohesion is lost to enable the resolution of chiasmata and the segregation of chromosomes during metaphase I. Accordingly, the REC8 signals start disappearing during the late prophase stages – diplotene and diakinesis – and some signals are observed in metaphase I (**Fig. 2 b–g**). In monocentric species, during canonical meiosis, cohesion is preserved around the centromeres in metaphase I to hold the sisters together until their separation in anaphase II (Cromer et al., 2013). During inverted meiosis of holocentric plants, on the other hand, the sisters are segregated during meiosis I itself. This suggests that centromeric cohesion is lost during metaphase I of inverted meiosis in holocentric plants. To test this hypothesis, co-localization studies were performed using centromeric repeat sequences (*Tyba*) and REC8 antibody during metaphase I. Even though some signals were observed during metaphase I, these signals did not correlate with centromeric regions (**Fig. 2 h–j**). Moreover, during monocentric meiosis, the sister centromeres and thus the sister kinetochores are fused and act as a single unit for microtubule attachment. However, earlier studies on *Rhynchospora* showed that while the centromere units from a single chromatid cluster together during metaphase I, the centromeres from sister chromatids are not fused together (Marques et al., 2015; Marques et al., 2016). They remain as separate units and attach to the microtubules from two different poles, thus enabling the sisters to separate to two different poles during meiosis I, which is also evident in our study (**Fig. 2 h–j**). These observations support the hypothesis that REC8 cohesion is lost around the centromeres during metaphase I, enabling the sister centromeres to act as independent units and thus allowing for their successful segregation during meiosis I. Given the fact that there are three other RAD21 proteins present in plants, we cannot rule out the possibility that an additional α-kleisin other than REC8 may be centromere-specific in these species. Regardless of this, our cytological observation that sister centromeres remain separate from metaphase I onwards provides strong evidence that sister centromere cohesion is absent during metaphase I of inverted meiosis.

Fig 1

Figure 1: Conservation of REC8 and SGO in *R. pubera*. A) Phylogenetic tree of RpREC8. *Rhynchospora* homolog is highlighted in red. B) Multiple sequence alignment of REC8 protein sequences from different species. The N terminal and C terminal Rad21/Rec8 domains belonging to the Pfam families PF04825 and PF04824 respectively are found to be conserved in RpREC8. C) Phylogenetic tree of RpSGO. Two SGOs are described for most of the species whereas only one SGO, is present in *R. pubera* (highlighted in red) and other related holocentric plants. D) Multiple sequence alignment of SGO protein sequences from different species. The N terminal coiled coil and C terminal basic regions are conserved in RpSGO as well. Zm=*Zea mays*, Sb=*Sorghum bicolor*, Os=*Oryza sativa*, At= *Arabidopsis thaliana*.

Fig 2

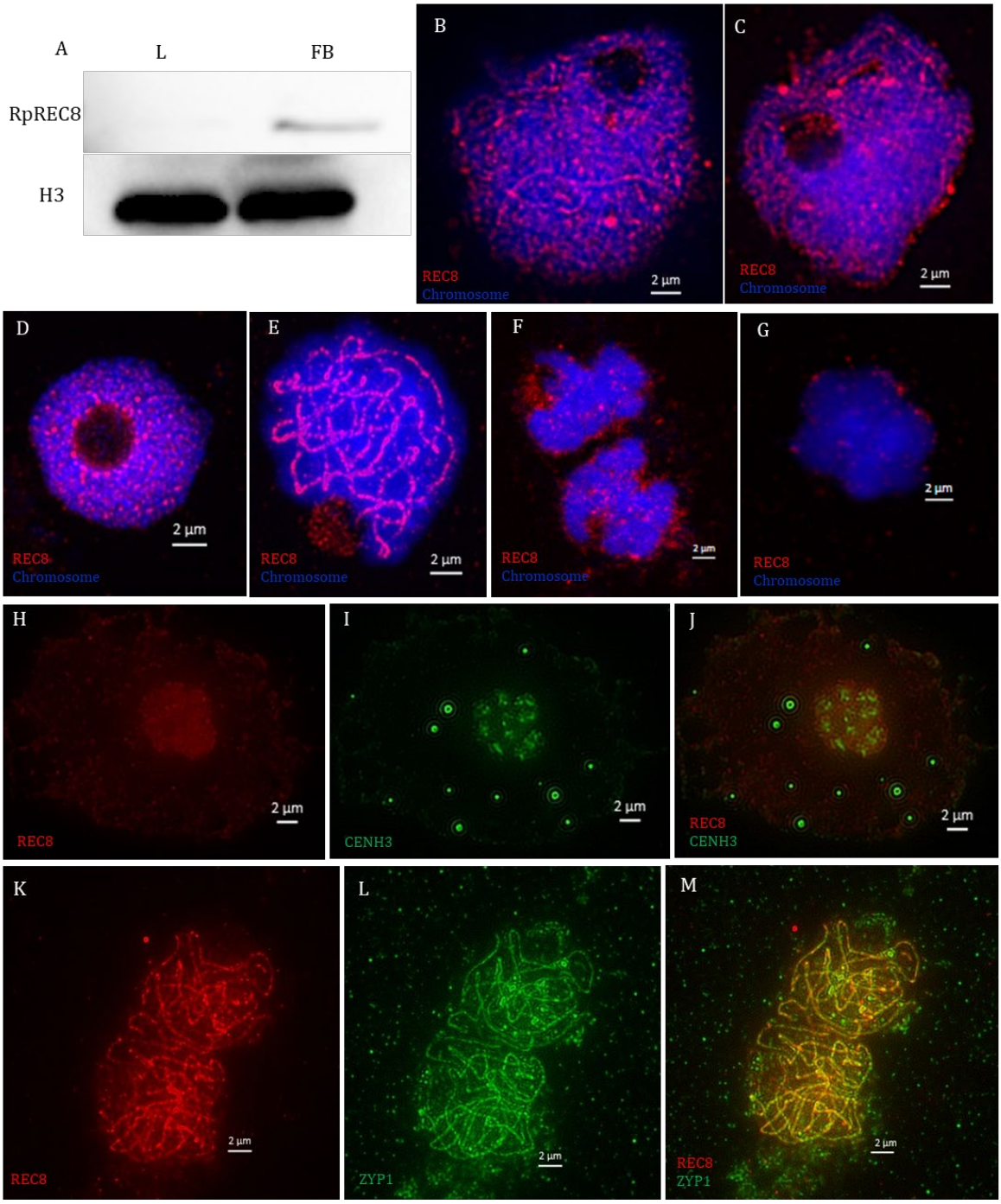


Figure 2: REC8 dynamics in *Rhynchospora* (**A**) Western blot for REC8 in *R. pubera*. Only flower bud tissue containing meiotic cells has the signal for REC8. H3 was used as the loading control. (**B, C**) Immunostaining for REC8 in *R. pubera* meiocytes shows the conserved REC8 signals. (**D–G**) Immunostaining for REC8 in *R. breviuscula* meiocytes. The signals start appearing in interphase (**D**), visualised as linear line signals in pachytene (**E**), start disappearing from diplotene chromosomes (**F**), and only a small fraction of signals is present in metaphase I (**G**). (**H–J**) Immuno-FISH of REC8 and Tyba in metaphase I of *R.*

breviuscula shows that REC8 signals are not preserved at the centromeric sites. (**K–M**) Co-localisation of REC8 and ZYP1 at pachytene stage of *R. breviuscula* shows that both the proteins have linear line signals and are co-localised.

Potential role of Shugoshins in *Rhynchospora*

Shugoshins were discovered as factors protecting centromeric cohesion, especially at metaphase I of canonical meiosis (Watanabe, 2005). However, inverted meiosis does not necessarily require centromeric cohesion protection and thus SGO as discussed earlier. This motivated us to characterise Shugoshin proteins in *Rhynchospora*. Based on homology and phylogeny analysis, we could detect one SGO in *Rhynchospora* and other related holocentric species (**Fig. 1c)**. Multiple sequence alignment showed that the protein has the conserved N-terminal coiled coil region and the C-terminal basic region (**Fig. 1d)**. The phylogenetic analysis was particularly interesting as it shows there is a trend towards duplication of Shugoshins. The tree shows a clear division between monocots and dicots. Inside the dicot group, an SGO duplication event is ancestral to all dicots. We observe the same happening in the monocots. However, some monocots subsequently lost one of the copies. Orchidaceae lost one of the paralogs, while Cyperaceae lost the other paralog independently (**Fig. 1d**). SGO duplication is also observed among non-plant species like humans and yeast (van Hooff et al., 2017). Why SGO has a strong tendency to undergo duplication, having done so several times independently, is currently unclear. Also, why some of the lineages have lost one of the paralogs also remains to be determined. Based on the analysis, however, it is clear that one SGO homolog is present in *Rhynchospora*. To further characterise the *Rhynchospora* SGO protein, SGO spatiotemporal dynamics were observed through different phases of inverted meiosis. Signals started appearing as early as interphase, and multiple foci were visualised all throughout the chromosomes during early prophase. During late prophase (diplotene and diakinesis), some of the signals started disappearing while signals were still present during metaphase I (**Fig. 3a–c**). During canonical meiosis, SGOs localise to the centromeric regions (Cromer et al., 2013). However, our colocalization studies with CENH3 show no obvious correlation between CENH3 and SGO (**Fig. 3d–f**). The potential role of SGO during inverted meiosis still remains poorly understood. Additional roles of SGO beyond regulatory activities at centromeres are also reported in *C. elegans*. Here, SGO1 regulates cohesins and plays a role in meiotic checkpoint activity (Bohr et al., 2018). Whether SGO has such non-centromere roles in *Rhynchospora* is yet to be explored.

Fig 3

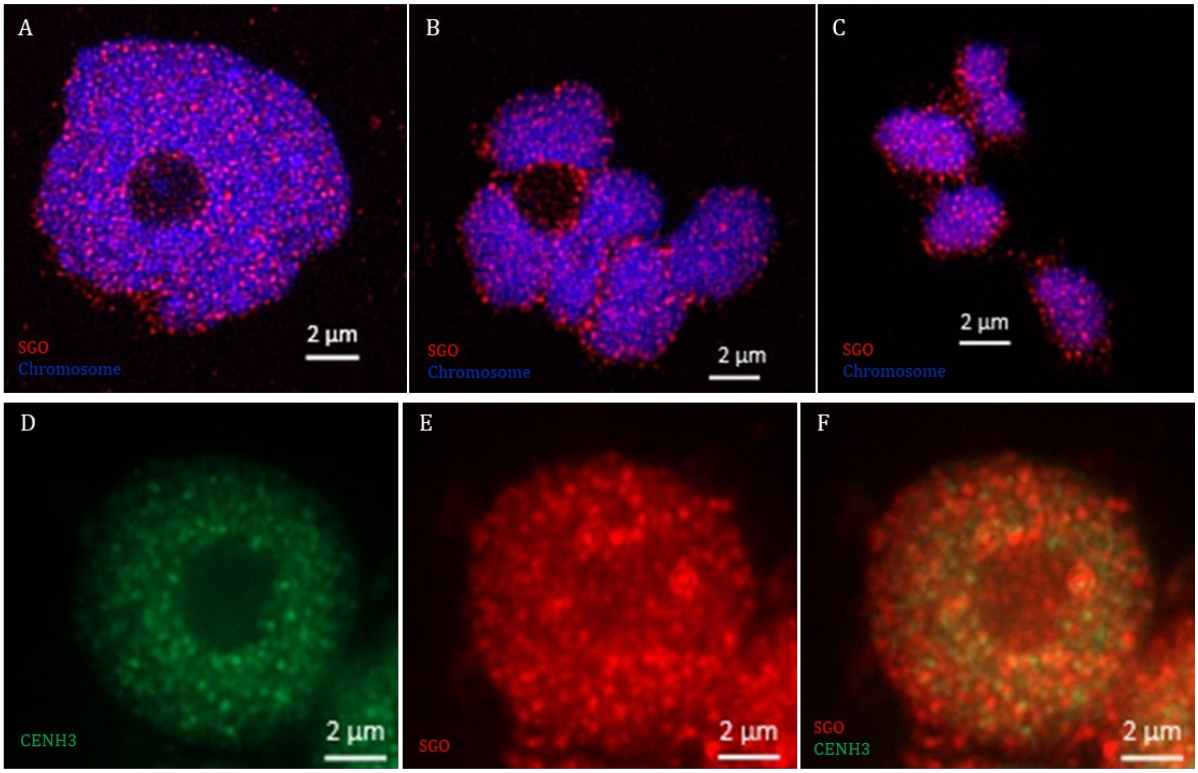


Figure 3: SGO dynamics in the meiocytes of *R. breviuscula*. (**A–C**) Immunostaining for SGO in *R. breviuscula*. Signals are observed in early prophase (**A**), diplotene (**B**) and metaphase I (**C**) cells. (**D–F**) Co-localisation of SGO and CENH3 in *R. breviuscula*. SGO signals do not show specific colocalization with CENH3 signals (**F**).

Functional characterisation of REC8 and SGO to elucidate its function

To functionally characterise REC8 and SGO, CRISPR–Cas9 vector constructs were constructed and genome editing was carried out in the plant *R. pubera*. N-terminal and C-terminal constructs targeting the conserved domains were designed for both the candidate genes (**Fig. 4a, b**). Eight events (34 plants), two events (6 plants) and seven events (28 plants) were generated for the *RpRec8* construct 2 (C-terminal), *RpSGO* construct 1 (N-terminal) and *RpSGO* construct 2 (C-terminal), respectively. We could not generate any plants for the *RpRec8* construct 1 (N-terminal), potentially because mutation at the N-terminal conserved domain is lethal for the plant. Most of the putative transformants were positive for the selectable markers hygromycin and Cas9, showing that the plants were indeed successfully transformed. Some of the plants were positive only for hygromycin and not for Cas9 which may represent a partial T-DNA integration event (**Fig. 5a**). Genotyping for the CRISPR-edited loci by conventional gene-specific PCR is challenging in *R. pubera* due to the multiple genome copies found in its genome

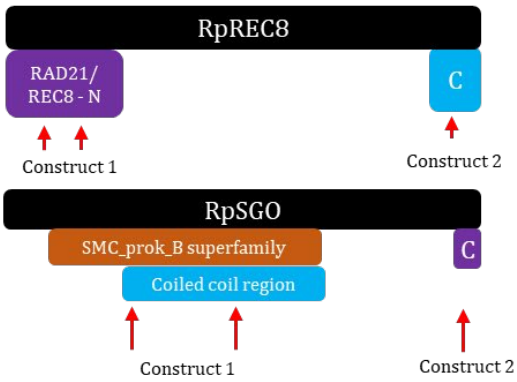
(Hofstatter et al., 2022). Therefore, we carried out next generation sequencing for the target region with a sufficient number of markers to enable mapping of all the alleles. Based on the results from analysing all *sgo* mutants, less than 15% of the sequences had the mutation at the desired site, which correlates to the fact that only one out of eight alleles were potentially mutated (**Fig. 5b–c**). The same observation was also seen in transformants from other on-going studies from our laboratory (personal communication with Dr. André Marques) confirming that the employed CRISPR–Cas9 system was not efficient in knocking out all the alleles. Meaningful characterisation of the candidates is not possible with only one allele mutated. Currently, the CRISPR–Cas9 system is being optimised to increase the efficiency of the genome editing. Unfortunately, we could not elucidate the functions of the candidate genes, but our first attempt to establish a transformation protocol in *Rhynchospora* was successful (**Fig. 4c**). We are currently working on improvising the CRISPR-Cas9 systems by using native promotors. These attempts will hopefully help us to improve the efficiency of the system and elucidate the function of the REC8 and SGO in the future.

Fig 4

A

Target gene	Construct name	No. of sgRNAs	Target regions	sgRNAs designed
Rec8	RpRec8 construct 1	2	Exon 2 Rec8 N terminal domain	5′-TGGCAGCAACGATGCGCTCGAGG-3′
			Exon 3 Rec8 N terminal domain	5′-TATGTACCCTGCAGCGCCAATGG-3′
	RpRec8 construct 2	1	Exon19 Rec8 C terminal domain	5′-CCCTCAGTCAGTTAGCTCAGGGG-3′
Sgo	RpSgo construct 1	2	Exon 4 SMC and coiled coil domain	5′-AGGGTGCTCGTCAGCAGAACTGG-3′
			Exon6 SMC and coiled coil domain	5′-TGAGTTGCACTGTACTACTCTGG-3′
	RpSgo construct 2	1	Exon13 SGO C domain	5′-ACTGCGGCGAGCTGTGGAGAAGG-3′

B



C

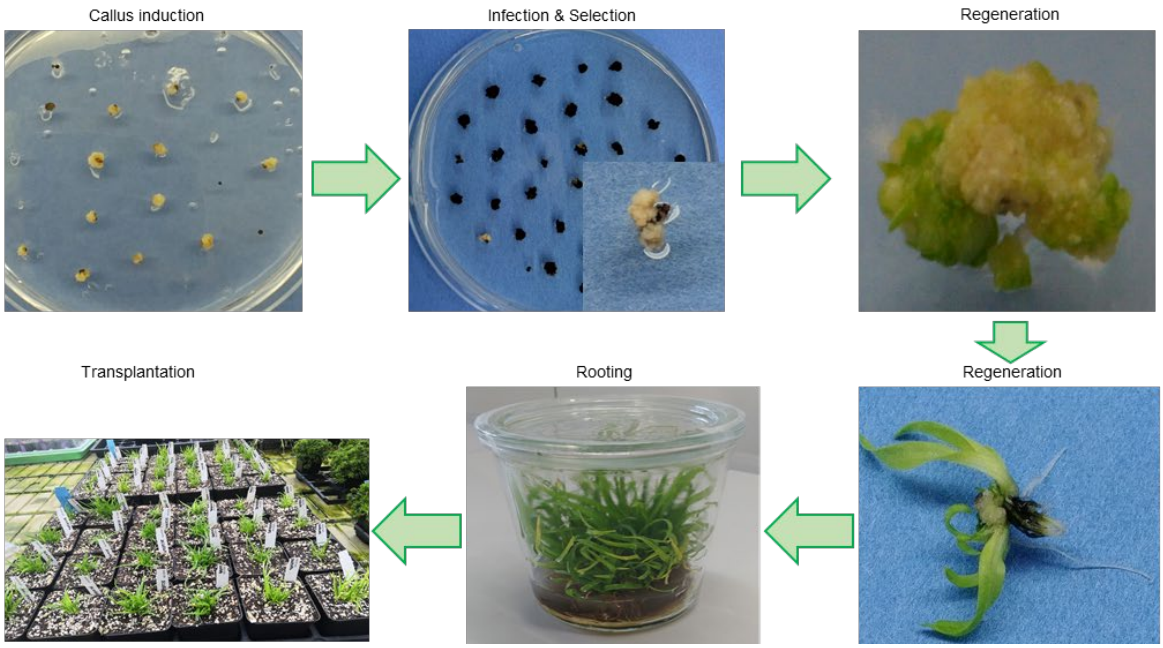


Figure 4: Editing of *Rec8* and *Sgo* using CRISPR/Cas9 constructs (**A**) List of sgRNAs designed to target *Rec8* and *Sgo*. PAM sequences are highlighted in red (**B**) *RpRec8* and *RpSgo* construct details. *RpRec8* construct 1 has 2 sgRNAs targeting the N terminal RAD21/REC8 domain and *RpRec8* construct 2 has a single sgRNA targeting the C terminal RAD21/REC8 domain. *RpSgo* construct 1 has 2 sgRNAs targeting the SMC and coiled coil domain, whereas the *RpSgo* construct 2 has a single sgRNA targeting the C terminal basic region. (**C**) Generation of putative transgenic plants for *Rec8* and *Sgo* constructs. The immature embryos were used for callus induction and co-cultivated with *Agrobacterium* culture containing the genome editing construct of interest. After two rounds of selection, only the healthy white calli resisting the hygromycin selection were forwarded for regeneration. Once they developed shoots and roots, the young plants were transferred to soil for further screening.

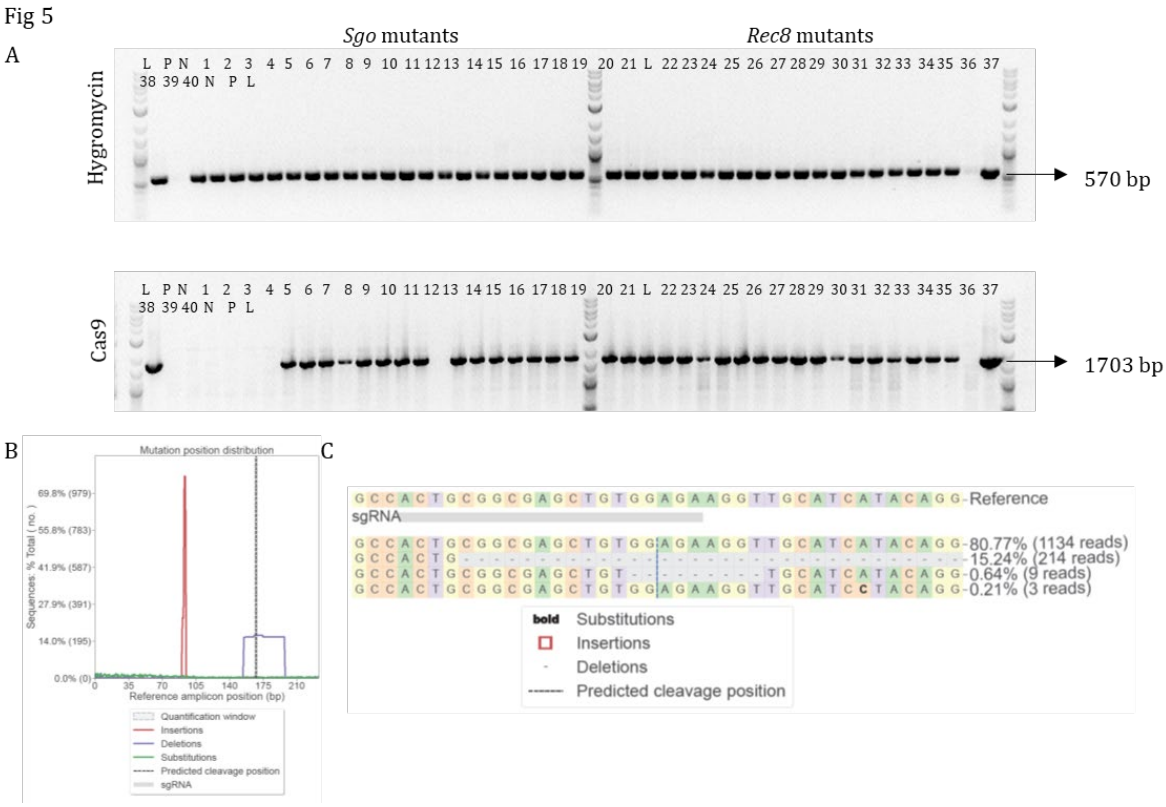


Figure 5: Genotyping results for *rec8* and *sgo* mutants. (**A**) Genotyping for the hygromycin and Cas9 markers. A representative image showing the results of PCR for hygromycin and Cas9 markers in the *rec8* and *sgo* mutants. Most of the plants have the desired amplicon size of 570 bp (hygromycin) and 1703 bp for (Cas9), confirming they are transformants. (**B, C**) NGS amplicon EZ sequencing results for SGO mutants. A representative image showing that 14% of the reads have deletions at the desired site.

Summary

Our results suggest that REC8 and SGO orthologs are present and functional in the holocentric plant *Rhynchospora*. Cytological evidence shows that REC8 has a conserved localisation pattern during the early stages of prophase I. However, during metaphase I, the phase that typically sees the major events of sister chromatid segregation during meiosis I of inverted meiosis, REC8 signals are not present around the centromeres. REC8 dynamics and the separate clustering of sister centromeres strongly supports the hypothesis that sister centromere cohesion is lost during metaphase I of inverted meiosis, which enables the sisters to segregate to opposite poles at the end of the meiosis I. The role of SGO in holocentric plants remains uncharacterised, but at least in *Rhynchospora* it may have non-centromeric roles. Increasing the efficiency of CRISPR–Cas9 genome

editing by using native promoters can contribute to a better understanding of its function in the future.

References

Bai, X., Peirson, B.N., Dong, F., Xue, C., and Makaroff, C.A. (1999). Isolation and Characterization of SYN1, a RAD21-like Gene Essential for Meiosis in Arabidopsis. The Plant Cell *11*, 417-430.

Bohr, T., Nelson, C.R., Giacopazzi, S., Lamelza, P., and Bhalla, N. (2018). Shugoshin Is Essential for Meiotic Prophase Checkpoints in C. elegans. Curr Biol *28*, 3199-3211 e3193.

Braz, G.T., Yu, F., do Vale Martins, L., and Jiang, J. (2020). Fluorescent In Situ Hybridization Using Oligonucleotide-Based Probes. In In Situ Hybridization Protocols, B.S. Nielsen, and J. Jones, eds. (New York, NY: Springer US), pp. 71-83.

Cabral, G., Marques, A., Schubert, V., Pedrosa-Harand, A., and Schlogelhofer, P. (2014). Chiasmatic and achiasmatic inverted meiosis of plants with holocentric chromosomes. Nat Commun *5*, 5070.

Cai, X., Dong, F., Edelmann, R.E., and Makaroff, C.A. (2003). The Arabidopsis SYN1 cohesin protein is required for sister chromatid arm cohesion and homologous chromosome pairing. J Cell Sci *116*, 2999-3007.

Capella-Gutierrez, S., Silla-Martinez, J.M., and Gabaldon, T. (2009). trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. Bioinformatics *25*, 1972-1973.

Chandra, H.S. (1962). Inverse meiosis in triploid females of the mealy bug, Planococcus citri. Genetics *47*, 1441-1454.

Chelysheva, L., Diallo, S., Vezon, D., Gendrot, G., Vrielynck, N., Belcram, K., Rocques, N., Marquez-Lema, A., Bhatt, A.M., Horlow, C., *et al.* (2005). AtREC8 and AtSCC3 are essential to the monopolar orientation of the kinetochores during meiosis. J Cell Sci *118*, 4621-4632.

Cromer, L., Jolivet, S., Horlow, C., Chelysheva, L., Heyman, J., De Jaeger, G., Koncz, C., De Veylder, L., and Mercier, R. (2013). Centromeric Cohesion Is Protected Twice at Meiosis, by SHUGOSHINs at Anaphase I and by PATRONUS at Interkinesis. Current Biology *23*, 2090-2099.

Goday, C., and Pimpinelli, S. (1989). Centromere organization in meiotic chromosomes of Parascaris univalens. Chromosoma *98*, 160-166.

Heckmann, S., Jankowska, M., Schubert, V., Kumke, K., Ma, W., and Houben, A. (2014a). Alternative meiotic chromatid segregation in the holocentric plant Luzula elegans. Nat Commun *5*, 4979.

Heckmann, S., Macas, J., Kumke, K., Fuchs, J., Schubert, V., Ma, L., Novak, P., Neumann, P., Taudien, S., Platzer, M.*, et al.* (2013). The holocentric species Luzula elegans shows interplay between centromere and large-scale genome organization. Plant J *73*, 555-565.

Heckmann, S., Schubert, V., and Houben, A. (2014b). Holocentric plant meiosis: first sisters, then homologues. Cell Cycle *13*, 3623-3624.

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

Katoh, K., and Standley, D.M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol *30*, 772-780.

Klein, F., Mahr, P., Galova, M., Buonomo, S.B., Michaelis, C., Nairz, K., and Nasmyth, K. (1999). A central role for cohesins in sister chromatid cohesion, formation of axial elements, and recombination during yeast meiosis. Cell *98*, 91-103.

Kumar, N., Galli, M., Ordon, J., Stuttmann, J., Kogel, K.H., and Imani, J. (2018). Further analysis of barley MORC1 using a highly efficient RNA-guided Cas9 gene-editing system. Plant Biotechnol J *16*, 1892-1903.

Kurze, A., Michie, K.A., Dixon, S.E., Mishra, A., Itoh, T., Khalid, S., Strmecki, L., Shirahige, K., Haering, C.H., Lowe, J.*, et al.* (2011). A positively charged channel within the Smc1/Smc3 hinge required for sister chromatid cohesion. EMBO J *30*, 364-378.

Lin, Y., Larson, K.L., Dorer, R., and Smith, G.R. (1992). Meiotically induced rec7 and rec8 genes of Schizosaccharomyces pombe. Genetics *132*, 75-85.

Llano, E., Gomez, R., Gutierrez-Caballero, C., Herran, Y., Sanchez-Martin, M., Vazquez-Quinones, L., Hernandez, T., de Alava, E., Cuadrado, A., Barbero, J.L.*, et al.* (2008). Shugoshin-2 is essential for the completion of meiosis but not for mitotic cell division in mice. Genes Dev *22*, 2400-2413.

Lui, D.Y., and Colaiácovo, M.P. (2013). Meiotic development in Caenorhabditis elegans. Adv Exp Med Biol *757*, 133-170.

Marques, A., Ribeiro, T., Neumann, P., Macas, J., Novak, P., Schubert, V., Pellino, M., Fuchs, J., Ma, W., Kuhlmann, M.*, et al.* (2015). Holocentromeres in Rhynchospora are associated with genome-wide centromere-specific repeat arrays interspersed among euchromatin. Proc Natl Acad Sci U S A *112*, 13633-13638.

Marques, A., Schubert, V., Houben, A., and Pedrosa-Harand, A. (2016). Restructuring of Holocentric Centromeres During Meiosis in the Plant Rhynchospora pubera. Genetics *204*, 555-568.

McGuinness, B.E., Hirota, T., Kudo, N.R., Peters, J.M., and Nasmyth, K. (2005). Shugoshin prevents dissociation of cohesin from centromeres during mitosis in vertebrate cells. PLoS Biol *3*, e86.

Melters, D.P., Paliulis, L.V., Korf, I.F., and Chan, S.W. (2012). Holocentric chromosomes: convergent evolution, meiotic adaptations, and genomic analysis. Chromosome Res *20*, 579-593.

Nasmyth, K., and Haering, C.H. (2009). Cohesin: its roles and mechanisms. Annu Rev Genet *43*, 525-558.

Nayar, S., Sharma, R., Tyagi, A.K., and Kapoor, S. (2013). Functional delineation of rice MADS29 reveals its role in embryo and endosperm development by affecting hormone homeostasis. J Exp Bot *64*, 4239-4253.

Nguyen, L.T., Schmidt, H.A., von Haeseler, A., and Minh, B.Q. (2015). IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol Biol Evol *32*, 268-274.

Ordon, J., Gantner, J., Kemna, J., Schwalgun, L., Reschke, M., Streubel, J., Boch, J., and Stuttmann, J. (2017). Generation of chromosomal deletions in dicotyledonous plants employing a user-friendly genome editing toolkit. Plant J *89*, 155-168.

Orr-Weaver, T.L. (1999). The Ties that Bind: Localization of the Sister-Chromatid Cohesin Complex on Yeast Chromosomes. Cell *99*, 1-4.

Pasierbek, P., Jantsch, M., Melcher, M., Schleiffer, A., Schweizer, D., and Loidl, J. (2001). A Caenorhabditis elegans cohesion protein with functions in meiotic chromosome pairing and disjunction. Genes Dev *15*, 1349-1360.

Pe´rez, R.n., Panzera, F., Page, J.s., Suja, J.A., and Rufas, J.S. (1997). Meiotic behaviour of holocentric chromosomes: orientation and segregation of autosomes in Triatoma infestans (Heteroptera). Chromosome Research *5*, 47-56.

Salic, A., Waters, J.C., and Mitchison, T.J. (2004). Vertebrate shugoshin links sister centromere cohesion and kinetochore microtubule stability in mitosis. Cell *118*, 567-578.

van Heemst, D., and Heyting, C. (2000). Sister chromatid cohesion and recombination in meiosis. Chromosoma *109*, 10-26.

van Hooff, J.J., Tromer, E., van Wijk, L.M., Snel, B., and Kops, G.J. (2017). Evolutionary dynamics of the kinetochore network in eukaryotes as revealed by comparative genomics. EMBO Rep *18*, 1559-1571.

Watanabe, Y. (2005). Shugoshin: guardian spirit at the centromere. Curr Opin Cell Biol *17*, 590-595.

Watanabe, Y., and Nurse, P. (1999). Cohesin Rec8 is required for reductional chromosome segregation at meiosis. Nature *400*, 461-464.

Chapter 3

Tracing the evolution of the plant meiotic molecular machinery

Summary

The kingdom Plantae is vast and studying meiosis throughout all lineages is essential to understand the whole plant meiotic machinery. Such studies are lacking till now due to the time required to functionally characterise the meiotic proteins in each lineage. This study is the first of its kind to study meiotic pathways in all major plant lineages by phylogeny of the meiotic proteins reported in the well-studied plant model, *A. thaliana*. Two parallel highly sensitive homology search approaches, PSI-BLAST and HMMER were employed to find the *Arabidopsis* orthologs in other plant lineages. Clustering by CLANS was performed when it was difficult to distinguish the right ortholog from other homologs. Multiple sequence alignment and phylogenetic tree were constructed for all the meiotic proteins. After careful analysis, the final tree was used to detect, whether a particular protein is present in a particular lineage. This shed light on conservation of meiotic pathways throughout kingdom Plantae and the evolutionary origin of few meiotic proteins reported only in plants.

My contribution:

My contribution to this publication is running the PSI-BLAST analysis, constructing the trees, developing the figures, writing the manuscript and corresponding the publication.

Publication

Plant Reproduction (2023) 36:73–95
https://doi.org/10.1007/s00497-022-00456-1

ORIGINAL ARTICLE

Tracing the evolution of the plant meiotic molecular machinery

Gokilavani Thangavel[1] · Paulo G. Hofstatter[2] · Raphaël Mercier[1] · André Marques[1]

Received: 22 July 2022 / Accepted: 12 December 2022 / Published online: 16 January 2023

Abstract
Meiosis is a highly conserved specialised cell division in sexual life cycles of eukaryotes, forming the base of gene reshuffling, biological diversity and evolution. Understanding meiotic machinery across different plant lineages is inevitable to understand the lineage-specific evolution of meiosis. Functional and cytogenetic studies of meiotic proteins from all plant lineage representatives are nearly impossible. So, we took advantage of the genomics revolution to search for core meiotic proteins in accumulating plant genomes by the highly sensitive homology search approaches, PSI-BLAST, HMMER and CLANS. We could find that most of the meiotic proteins are conserved in most of the lineages. Exceptionally, *Arabidopsis thaliana* ASY4, PHS1, PRD2, PRD3 orthologs were mostly not detected in some distant algal lineages suggesting their minimal conservation. Remarkably, an ancestral duplication of SPO11 to all eukaryotes could be confirmed. Loss of SPO11-1 in Chlorophyta and Charophyta is likely to have occurred, suggesting that SPO11-1 and SPO11-2 heterodimerisation may be a unique feature in land plants of Viridiplantae. The possible origin of the meiotic proteins described only in plants till now, DFO and HEIP1, could be traced and seems to occur in the ancestor of vascular plants and Streptophyta, respectively. Our comprehensive approach is an attempt to provide insights about meiotic core proteins and thus the conservation of meiotic pathways across plant kingdom. We hope that this will serve the meiotic community a basis for further characterisation of interesting candidates in future.

Keywords Meiotic proteins · Homology search · Phylogeny · Plant · Conservation · SPO11 duplication

Introduction

The mechanisms of meiosis, with a few notable exceptions, are highly conserved among sexually reproducing eukaryotes such as fungi, plants and animals (Gerton and Hawley 2005; Villeneuve and Hillers 2001). These processes include sister chromatid cohesion, homologous chromosome pairing, formation of the synaptonemal complex, double-stranded break (DSB) formation and processing, cross-over (CO) formation and resolution and two-step segregation of chromosomes, making meiosis special and different from mitosis. Therefore, typically, a common and shared set of specific meiotic genes can be found in all sexually reproducing organisms.

Formation of programmed double-stranded breaks (DSBs) during Prophase I is the upstream of many meiotic processes. First discovered in the budding yeast *Saccharomyces cerevisiae*, DSB initiation is catalysed by the highly conserved protein, SPO11 (Bergerat et al. 1997; de Massy et al. 1995; Keeney et al. 1997; Keeney and Kleckner 1995; Liu et al. 1995). In plants until now, many proteins have been isolated that function in DSB formation—PHS1/Rec114, PRD1/Mei1, PRD2/Mei4, PRD3/PAIR1/Mer2, DFO, PCH2 and MTOPVIB among which DFO have only been described in plants until now. DSBs are later loaded by the recombinases—RAD51 and DMC1. DMC1-mediated DNA repair using non-sister homologous chromatid appears to be the predominant pathway during *Arabidopsis thaliana* meiosis (Mercier et al. 2015). Chromosome axis mediates the formation of DSBs and its consecutive repair, resulting

Communicated by Frederic Berger.

✉ Gokilavani Thangavel
gthangavel@mpipz.mpg.de

✉ Raphaël Mercier
mercier@mpipz.mpg.de

✉ André Marques
amarques@mpipz.mpg.de

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

[2] University of Sao Paulo, São Paulo, Brazil

in the formation of inter-homolog COs. Cohesin complexes and axial element protein complexes form the components of chromosome axis formation. Cohesion complex is formed by the proteins—SMC1, SMC3, alpha-kleisin unit (SCC1/REC8) and SCC3 (Chelysheva et al. 2005; Onn et al. 2008). ASY1 and ASY2 are the HORMA domain containing axis proteins. ASY3 and ASY4 are the axis core proteins, essential for the recruitment of the HORMA domain proteins and the formation of axis (Caryl et al. 2000; Chambon et al. 2018; Ferdous et al. 2012; Sanchez-Moran et al. 2008, 2007; West et al. 2019). During the progression of prophase I, chromosome synapses and the axes of each homolog pair are connected to each other by coiled-coil transverse filaments (Dong and Roeder 2000; Liu et al. 1996; Meuwissen et al. 1992; Sym et al. 1993). ZYP1A and ZYP1B are identified as the proteins involved in the formation of synaptonemal complex (SC) in *A. thaliana* (Capilla-Perez et al. 2021; France et al. 2021; Higgins et al. 2005). There are two pathways for the formation of the COs—interference sensitive Class I and interference insensitive Class II pathways. Class I is the major one and depends on ZMM proteins (HEI10, HEIP1, MER3, MSH4, MSH5, PTD, ZIP2/SHOC1, ZIP4) and MLH1, MLH3 (Börner et al. 2004; Chelysheva et al. 2012; Dion et al. 2007; Franklin et al. 2006; Higgins et al. 2004, 2008b; Kuromori et al. 2008; Li et al. 2018; Lu et al. 2014; Macaisne et al. 2008; Mercier et al. 2005). Numerous DSBs are formed among which very few are processed to form COs. CO designation is still poorly understood (Berchowitz et al. 2007; Higgins et al. 2008a).

Understanding meiosis in plants can form a basis for advances in reproduction, fertility, genetics, breeding and thereby accelerate agricultural applications (Sanchez-Moran et al. 2008). Plants are also considered to be a good model system to study meiosis because in meiotic mutants, meiosis proceeds until the end of tetrad formation stage with meiotic defects like massive chromosome segregation defects but without confounding effects from the onset of meiotic arrest and apoptosis like in mammals (Higgins et al. 2004; Mercier and Grelon 2008). The kingdom Plantae or Archaeplastida in a broader sense includes freshwater unicellular algae (glaucophytes), photoautotrophic red algae (rhodophytes) and Viridiplantae which includes the paraphyletic group of green algae (chlorophytes and charophytes) and land plants. Land plants can be further classified into bryophytes (liverworts, hornworts, mosses), lycophytes, pteridophytes (ferns) and spermatophytes (gymnosperms and angiosperms) (Puttick et al. 2018). Plants are quite diverse and land plants alone are suggested to be approximately 500,000 species in comparison against 5400 mammalian species in total (Corlett 2016). Among plants, most studies investigating meiosis have been carried out in angiosperms, and the vast majority of studies characterising meiotic genes is done in the model plant *A. thaliana* and also in rice, maize, wheat, barley among others (Mercier and Grelon 2008). In total, around 100 genes involved in meiosis have been functionally studied in *A. thaliana* (Zhang et al. 2018). However, considering the diversity of plants, studying a few angiosperm models alone will not be sufficient to understand the evolution of meiosis in this kingdom. Functionally studying representative meiotic proteins from all plant lineages would be nearly impossible due to practical reasons. However, revolutionary advances in genomics means that sequence information is increasingly accumulating for many members of the Viridiplantae (green plants), and homology search can provide insights about the presence of meiotic machinery orthologs in a wide range of organisms.

To date, there is no comprehensive study that has aimed to search and detect core meiotic genes across all the main groups of the plant kingdom. Therefore, in this study, we searched for homologs of well-studied angiosperm meiotic genes among different plant lineages from algae to angiosperms. We bring to the attention of the readers that this paper discusses only Viridiplantae; however, rhodophytes and glaucophytes were included in our analysis as an outgroup. Our approach has allowed us to trace the conservation of the ancestral molecular machinery of plant meiosis and establish a correlation with the evolution of meiosis and the presence/absence of meiotic homologs across Viridiplantae. We found that proteins involved in DSB formation, chromosome axis formation and ZMM pathway are not detected in some early plant lineages, suggesting they are either missing or evolving rapidly during the diversification of the plant kingdom. Remarkably, our analysis confirms that land plants have two meiosis-expressed SPO11 paralogues, both essential for meiotic DSB formation and likely to act as a heterodimer, but only one homolog is retained in chlorophytes and charophytes. Our study shows how systematic analysis of the similarities and differences in meiotic regulation among plant species can provide insights into the fundamental elements of this critical process across evolution.

Materials and methods

Homology search using NCBI PSI-BLAST and phylogenetic tree construction

Twenty-seven genes with key meiotic function reported in *A. thaliana* were chosen for this study. Based on its function, the proteins were categorised into four pathways: chromosome axis/synaptonemal complex; double-strand break formation; strand invasion; and ZMM (Table 1). Protein sequences were downloaded from either UniProtKB or TAIR. TAIR has a list of plant homologs for all the proteins derived from the gene families of PANTHER 16.0 release which was used to create the initial multiple

Table 1 List of meiotic proteins used in this study

Protein id	Protein name in *Arabidopsis*	Alternative names in other species	Function in meiosis	Mutant phenotype	References	Sequence information
AT1G67370	ASY1	Hop1(yeast) HORMAD1 (mammals) PAIR2(rice)	Chromosome axis/SC	Failure in pairing, asynapsis or non-homologous synapsis, reduction in chiasma frequency	Caryl et al. 2000; Sanchez-Moran et al. 2008, 2007	> sp|F4HRV8|ASY1_ARATH Meiosis-specific protein ASY1 OS = Arabidopsis thaliana OX = 3702 GN = ASY1 PE = 1 SV = 1 MVMAQKLKEAEITEQDSLLLTRNLLRIAIFNISYIRGLFPEKYFNDKSVPALDMKIKKLM PMDAESRRLIDWMEKGVYDALQRKYLKTLMFSICETVDGPMIEEYSFSFSYSDSD- SQDVM MNINRTGNKKNGGIFNSTADITPNQMRSSACKMVRTLVQLMRTLDKMPDERTIVMKL- LYY DDVTPPDYEPPFFRGCTEDEAQYVWTKNPLRMEIGNVNSKHLVLTLKVKSVLDPCE- DEND DMQDDGKSIGPDSVHDDQPSDSDSEISQTQENQFIVAPVEKQDDDDGEVDEDDNTQD- PAE NEQQLARVKDWINSRHLDTLELTDILANFPDISIVLSEEIMDQLVTEGVLSKTGKDMYIK KRDKTPESEFTFVKEEADGQISPGKSVAPEDYLYMKALYHSLPMKYVTITKLHNMLD- GEA NQTAVRKLMDRMTQEGYVEASSNRRLGKRVIHSSLTEKKLNEVRKVLATDDMDVD- VTETI NKTNGPDAKVTADVSTCGGIHSIGSDFTRTKGRSGGMQQNGSVLSEQTISKAGNTPISNK AQPAASRESFAVHGGAVKEAETVNCSQASQDRRGRKTSMVREPILQYSKRQKSQAN
AT2G46980	ASY3	SYCP2(mammals) Red1(yeast) PAIR3(rice)	Chromosome axis/SC	Abnormal chromosome axis, disrupts SC formation, reduced meiotic COs, univalent formation and mis segregation of chromosomes	Ferdous et al. 2012	> sp|Q0WR66|ASY3_ARATH Meiosis-specific protein ASY3 OS = Arabidopsis thaliana OX = 3702 GN = ASY3 PE = 1 SV = 1 MSDYRSFGSNYHPSSQSRKISIGVMADSQPKRNLVPDKDDGDVIARVEKLKSATVTELQA NKKEKSDLAAKQRNSAQVTGHVTSPWRSPRSSHRKLGTLESVLCKQTSSLSGSK- GLNKGL NGAHQTPARESFQNCPISSPQHSLGELNGGRNDRVMDRSPERMEEPPSAVLQQKVASQRE KMDKPGKETNGTTDVLRSKLWEILGKASPANNEDVNSETPEVEKTNFKLSQDKGSND- DPL IKPRHNSDSIETDSESPENATRRPVTRSLLQRRVGAKGVQKKTKAGANLGRKCTEQVNSV FSFEEGLRGKIGTAVNSSVMPKKQRGRRKNTVVKCRKAHSRKKDEADWSRKEASK- SNTPP RSESTETGKRSSSSDKKGSSHDLHPQSKARKQKPDISTREGDFHPSPEAEAAALPEMSQG LSKNGDKHERPSNIFREKSVEPENEFQSPTFGYKAPISSPSPCCSPEASPLQPRNISPTL DETETPIFSFGTKKTSQGTTGQASDTEKRLPDFLEKKRDYSFRRESSPEPNEDLVLSDPS SDERDSDGSREDSPVLGHNISPEERETANWTNERSMLGPSSVKRNSNLKGIGRVVLSPPS PLSKGIDKTDSFQHCSEMDEDEDEGLGRAVALFAMALQNFERKLKSAAEKKSSEIIASVS EEIHLELENIKSHIITEAGKTSNLAKTKRKHAETRLQEQEEKMRMIHEKFKDDVSHHLED FKSTIEELEANQSELKGSIKKQRTSHQKLIAHFEGGIETKLDDATKRIDSVNKSARGKML QLKMIVAECLRDD

Table 1 (continued)

Protein id	Protein name in *Arabidopsis*	Alternative names in other species	Function in meiosis	Mutant phenotype	References	Sequence information
AT2G33793	ASY4	SYCP3(mammals)	Chromosome axis/SC	Defective chromosome axis, incomplete synapsis, reduction in formation of COs	Chambon et al. 2018	> tr|F4IFY5|F4IFY5_ARATH DNA ligase-like protein OS = Arabidopsis thaliana OX = 3702 GN = At2g33793 PE = 4 SV = 1 MSSTRRGTKRTRPEPPQSLKKPTPKAKLPDELDVDVSSDFKGIMSALQQFREKAHEDGRK KKEESISSVSTEVKSKIDELKSKLEKERQNFSKALSKSSKECENILKDEAAKFEELHKKF VKDKADHLQGLKDTISKFEEDKERLYMRYEQLRKKEKTMITEQEKFCTEKLAQLEES- LKK KKRGDKTFSILRKTLGSFLENEASDEEFPPDE
AT5G05490	REC8/SYN1/DIF1	Rec8(yeast) REC8(mammals)	Chromosome axis/SC	Defects in SCC, defects in chromosome condensation leading to the formation of chromosome fragmentation, chromosome mis segregation, formation of univalent and aneuploids, sterility	Bai et al. 1999; Bhatt et al. 1999; Cai et al. 2003; Chelysheva et al. 2005; Peirson et al. 1997	> sp|Q9S7T7|SCC11_ARATH Sister chromatid cohesion 1 protein 1 OS = Arabidopsis thaliana OX = 3702 GN = SYN1 PE = 2 SV = 2 MLRLESLIVTVWGPATLLARKAPLGQIWMAATLHAKINRKKLDKLDIIQICEEILNPSVP MALRLSGILMGGVVIVYERKVKLLFDDVNRFLVEINGAWRTKSVPDPTLLPKGK- THARKE AVTLPENEEADFGDFEQTRNVPKFGNYMDFQQTFISMRLDESHVNNNPEPEDLGQQF- HQA DAENITLFEYHGSFQTNNETYDRFERFDIEGDDETQMNSNPREGAEIPTTLIPSPPRHHD IPEGVNPTSPQRQEQQENRRDGFAEQMEEQNIPDKEEHDRPQPAKKRARKTATSAM- DYEQ TIIAGHVYQSWLQDTSDILCRGEKRKVRGTIRPDMESFKRANMPPTQLFEKDSSYPPQLY QLWSKNTQVLQTSSSESRHPDLRAEQSPGFVQERMHNHHQTDHHERSDTSSQNLDSPAEI LRTVRTGKGASVESMMAGSRASPETINRQAADINVTPFYSGDDVRSMPSTPSARGAASIN NIEISSKSRMPNRKRPNSSPRRGLEPVAEERPWEHREYEFEFSMLPEKRFTADKEILFET ASTQTQKPVCNQSDEMITDSIKSHLKTHFETPGAPQVESLNKLAVGMDRNAAAKLFFQSC VLATRGVIKVNQAEPYGDILIARGPNM
AT1G22260/AT1G22275	ZYP1 (*Arabidopsis* has ZYP1a and ZYP1b that are functionally redundant)	Zip1(yeast) SYCP1(mammals)	Chromosme axis/SC			

Table 1 (continued)

Protein id	Protein name in *Arabidopsis*	Alternative names in other species	Function in meiosis	Mutant phenotype	References	Sequence information
AT1G22260	ZYP1A			Reduction in fertility, Absence of SC, delayed prophase I, slight reduction in recombination, recombination between non-homologous chromosomes, absence of heterochiasmy	Capilla-Perez et al. 2021; France et al. 2021; Higgins et al. 2005	> sp|Q9LME2|SYCP1_ARATH Synaptonemal complex protein 1 OS = Arabidopsis thaliana OX = 3702 GN = ZYP1A PE = 2 SV = 1 MQKLGFPAMKSLDKPRSLSGSANMYSFSNRKPPDSVSSGSFSNLKLTAEKLVK-DQAAMRT DLELANCKLKKSMEHVYALEEKLQNAFNENAKLRVRKKEDEKLWRGLESKFSSTKTL-CDQ LTETLQHLASQVQDAEKDKGFFETKFSTSSEAIDSLNQQMRDMSLRLDAAKEEITSRDKE LEELKLEKQQKEMFYQTERCGTASLIEKKDAVITKLEASAAERKLNIENLNSQLEKVHLE LTTKEDEVKDLVSIQEKLEKEKTSVQLSADNCFEKLVSSEQEVKKLDELVQYLVAELTEL DKKNLTFKEKFDKLSGLYDTHIMLLQKDRDLALDRAQRSFDNLQGELFRVAAT-KEALESA GNELNEKIVELQNDKESLISQLSGLRCSTSQTIDKLESEAKGLVSKHADAESAISQLKEE METLLESVKTSEDKKQELSLKLSSLEMESKEKCEKLQADAQRQVEELETLQKESESHQLQ ADLLAKEVNQLQTVIEEKGHVILQCNENEKQLNQQIIKDKELLATAETKLAEAKKQY-DLM LESKQLELSRHLKELSQRNDQAINEIRRKYDVEKHEIINSEKDKVEKIIKDLSNKFDKEL SDCKEESKRQLLTIQEEHSSLILSLREEHESKELNLKAKYDQELRQSQIQAENELKERIT ALKSEHDAQLKAFKCQYEDDCKKLQEELDLQRKKEERQRALVQLQWKVMSDNP-PEEQEVN SNKNYSISKDSRLGGSKRSEHIRVRSDNDNVQDSPFVKAKETPVSKILKKAQNVNAGSVL SIPNPKHHSKVTHREYEVETNNGRVTKRRKTRNTTMFEEPQRRRTRATPKLTPQSIAKGT GMTSHARSANIGDLFSEGSLNPYADDPYAFD
AT1G22275	ZYP1B			Reduction in fertility, Absence of SC, delayed prophase I, slight reduction in recombination, recombination between non-homologous chromosomes, absence of heterochiasmy	Capilla-Perez et al. 2021; France et al. 2021; Higgins et al. 2005	> sp|P61430|SYCP2_ARATH Synaptonemal complex protein 2 OS = Arabidopsis thaliana OX = 3702 GN = ZYP1B PE = 2 SV = 1 MQKLGFPAMKSFDQLRSLPGSAKTYFFSTRPPQDSVSSGSFSNLKLTAEKLVKDQAAMRT DLELANCKLKKSMEHVYALEEKLQSAFNENAKLRVRQKEDEKLWRGLESKFSSTKTL-CDQ LTETLQHLASQVQDAEKDKGFFETKFNTSSEAINSLNQQMRDMSLRLDAAKEEITSRDKE LEELKLEKQHKEMFYQTERCGTASLIEKKDAVITELETTAAERKLKIEKLNSQLEKLHLE LTTKEDEVIHLVSIQEKLEKEKTNVQLSSDELFEKLVRSEQEVKKLDELVHYLIAELTEL DKKNLTFKEKFDKLSGLYDTHFMLLRKDRDLASDRAQRSFDQLQGELFRVAAEKEA-LESS GNELSEKIVELQNDKESLISQLSGVRCSASQTIDKLEFEAKGLVLKNAETESVISKLKEE IDTLLESVRTSEDKKKELSIKLSSLEIESKDKYEKLQADAQRQVGELETLQKESESHQLQ ADLLAKEVNQLQTIIEEKGHLILQCNENEKNINQQIIKDKELLATAETKLAEAKKQYDLM LESKQLELSRHLKELSQRNDQAINEIRRKYDVEKHEIINSEKDKVEKIIKELSTKYDKGL SDCKEESKRQLLTIQEEHSSRILNIREEHESKELNLKAKYDQELRQNQIQAENELKERIT ALKSEHDAQLKAFKCQYEDDCKKLQEELDLQRKKEERQRALVQLQWKVMSDNP-PEEQEVN SNKDYSHSSVKVKESRLGGNKRSEHITESPFVKAKVTSVSNILKEATNPKHHSKVTHREY EVETNNGRIPKRRKTRQTTMFQEPQRRSTRLTPKLMTPTIIAKETAMADHPHSANIGDLF SEGSLNPYADDPYAFD

Table 1 (continued)

Protein id	Protein name in *Arabidopsis*	Alternative names in other species	Function in meiosis	Mutant phenotype	References	Sequence information
AT1G07060	DFO	Reported to be plant specific so far	DSB	Reduced fertility, formation of polyads, defects in chromosome synapsis and segregation and recombination	Zhang et al. 2012	> sp|Q8RX33|DFO_ARATH Protein DOUBLE-STRAND BREAK FORMATION OS = Arabidopsis thaliana OX = 3702 GN = DFO PE = 1 SV = 1 MRHNIKFKSKGTLKIRNTAQISLWKKCSDSMIADQTYLFINRVQDRRFDEESLRILELSL VAMNVKSFLEVRSRLRDFMRSESVVIFGELTGESMVAKLSVLEFFARAFALLGDMESCLA MRYEALNLRQLKSPSCLWLGVSHSEWTKFAVQSMENGFPSIAGKASENALLSLKKDSLIE PKSEDNSDILDAAEKVRRLRDSAASLTSSHSGIFIYIVSSLKFAVCNRLLTTF
AT1G60460	MTOPVIB	TOPOVIBL (mouse)	DSB	Reduction in fertility, defects in synapsis and recombination	Tang et al. 2017; Vrielynck et al. 2016	> sp|Q5Q0E6|TO6BL_ARATH Type 2 DNA topoisomerase 6 subunit B-like OS = Arabidopsis thaliana OX = 3702 GN = MTOPVIB PE = 1 SV = 1 MENNAPVPKLLLQLISSAFQRCRLAEDLCRLSVLLDQSTERDPPITCISIADTGIGCNLE EFQNLRCPREFNGAKIWDGLLSVKTTCFTDDEVYYYHINLDEYIANKRLKRQP- SQAKNGA KFSGTEVSLSVFGSMDVLVAPIIGFFQKIIVLQILNVTLDLMVKQGTSPGNQTQYVFAVN ADKTPCFTASNLERLKSGLEDYVLRHANCLDTMCDYCFSDREHLKVGSGTVCQED- KHKRV GGTMEVVIVISDLLESTQHCSRSCNGKTEVLYFDNFLPSPVPHLALSALKKIDWKKYGLI LANVNDQDGHVFLEWDNFPSYVQIQIALHWYHNQYPTRQKNGPGISLLKK- GIKNALDNLK AKHEGFLLSSHSRKICSYVPDLARSIAGLIFSSTDLDFQGDCLSVLGFQTQEVERDTVEN YIQRKIVTVIGMNERKPQKDQEAAPFLFFDGESETSFFEDEEVEGEYYSSSLE
AT1G10710	PHS1	REC114 (mammals) Rec114 (Saccharomyces cerevisiae) Rec7 (Schizosaccharomyces pombe)	DSB	Defects in pairing—pairing of non-homologous chromosomes	Ronceret et al. 2009	> sp|Q45GQ7|POHS1_ARATH Protein POOR HOMOLOGOUS SYNAPSIS 1 OS = Arabidopsis thaliana OX = 3702 GN = PHS1 PE = 1 SV = 1 MAGSLTASNRRRNAEDSSEIYRWTIGFARFVHYPSSPSPHPVLKPLGKREQYHSPHGTWL SASSSTVSLHIVDELNRSDVILSVKLGQKVLEEHYISKLNFTWPQMSCVSGFPSRGSRAI FVTYMDSANQIQKFALRFSTCDAALEFVEALKEKIKGLKEASTQNQKNKTRCDVSFQSDY NPSDAIIPRATQKEPNMVRPLNSYVPEMLPRIVYEAQYQKSETRSEVSFQSDYNPSIEIF PRATEEEPNMVRFFDSSVPEVLPRPEYEAGQALYPSQSTLNQIPSLPPSFTTLLSGCFPD STLDAGQTTVKQNPDLKSQILKYMEDSSFQDMLQKVERIIDEIGGNWIT

Table 1 (continued)

Protein id	Protein name in *Arabidopsis*	Alternative names in other species	Function in meiosis	Mutant phenotype	References	Sequence information
AT4G14180	PRD1	MEI1(mammals)	DSB	Defects in synapsis, reduction of recombination rates, formation of achiasmatic univalents	De Muyt et al. 2007	>sp|O23277|PRD1_ARATH Protein PUTATIVE RECOMBINATION INITIATION DEFECT 1 OS=Arabidopsis thaliana OX=3702 GN=PRD1 PE=1 SV=3 MFFQHSQLQNSDHLLHESMADSNHQSLSPPCANGHRSTISLRDDQGGTFCLICFSNLVSD PRIPTVHVSYALHQLSIAISEPIFLRTLLSSHIHFLVSPLVHALSSIDDAPIAIQIMDMI SLLCSVEESSIGEDFVERISDQLSSGALGWSRRQLHMLHCFGVLMSCENININSHIRDKE ALVCQLVEGLQLPSEEIRGEILFALYKFSALQFTEQNVDGIEVLSLLCPKLLCLSLEALA KTQRDDVRLNCVALLTILAQQGLLANSHSNSASSMSLDEVDDDPMQTAENVAARP- CLNVL FAEAIKGPLLSTDSEVQIKTLDLIFHYISQESTPSKQIQVMVEENVADYIFEILRLSECK DQVVNSCLRVLDLFSLAEHSFRKRLVIGFPSVIRVLHYVGEVPCHPFQIQTLKLISSCIS DFPGIASSSQVQEIALVLKKMLERYYSQEMGLFPDAFAIICSVFVSLMKTPSFGETADVL TSLQESLRHSILASLSLPEKDSTQILHAVYLLNEVYVYCTASTSINKTICIELRHCVIDV CTSHLLPWFLSDVNEVNEEATLGIMETFHSILLQNSDIQAKEFAELLVSADWFSFSFGCL GNFCTDNMKQRIYLMLSSLVDILLEQKTGSHIRDALHCLPSDPQDLLFLLGQASSNNQEL ASCQSAALLIFHTSSIYNDRLADDKLVLASLEQYIILNKTSLICAISDSPALLNLVNLYG LCRSLQNERYQISYSLEAERIIFHLLNEYEWDLGSINIHLESLKWLFQQESISKSLIYQI QKISRNNLIGNEVHNVYGDGRQRSLTYWFAKLISEGDNYAATLLVNLLTQLAEKEEQEND VISILNLMNTIVSIFPTASNNLSMNGIGSVIHRLVSGFSNSSLGTSFRTLLLLVFNILTS VQPAVLMIDESWYAVSIKLLNFLSLRDTAIKQNHEDMVVIGILSLVLYHSSDGALVEASR NIVSNSYLVSAINTVVDVACSKGPALTQCQDETNIGEALAFTLLLYFFSLRSLQIVLAGA VDWQTFFGTSTSLETLPVVCIHCHNLCRLMHFGAPQIKLIASYCLLELLTGLSEQVDIKK EQLQCSSSYLKSMKAVLGGLVFCDDIRVATNSALCLSMILGWEDMEGRTEMLKTSSW- YRF IAEEMSVSLAMPCSASSTYVNHHKPAVYLTVAMLRLKNKPVWLRTVFDESCISSMIQNLN GINISREIVILFRELMQAELLNSQQVTKLDRAFQECRKQMHRNGTRDETVEEQVQRKIPS IHDHSEFCNYLVHLMVSNSFGHPSESETYTQKKKQILDEMEQFSELISTREGRVSPIQEE TRQMQTERIV
AT5G57880	PRD2/ MPS1	MEI4(mammals) Mei4(S. cerevisiae) Rec24(S. pombe)	DSB	Aberrant spindle formation and disordered chromosome segregation	Jiang et al. 2009	>sp|F4KDF5|MUPS1_ARATH Protein MULTIPOLAR SPINDLE 1 OS=Arabidopsis thaliana OX=3702 GN=MPS1 PE=2 SV=2 MSSSVAEANHTEKEESLRLAIAVSLLRSKFQNHQSSSSTSRCYVSSESDALRWKQKAKER KKEIIRLQEDLKDAESSFHRDLFPANASCKCYFFDNLGVFSGRRIGEASESRFNDVLRRR FLRLARRRSRRKLTRSSQRLQPSEPDYEEEAEHLRISIDFLLELSEADSNDSNFSNWSHQ AVDFIFASLKKLISMGRNLESVEESISFMITQLITRMCTPVKGNEVKQLETSVGFYVQHL IRKLGSEPFIGQRAIFAISQRISILAENLLFMDPFDESFPEMDECMFILIQLIEFLICDY LLPWANEAFDNVMFEEWIASVVHARKAVKALEERNGLYLLYMDRVTGELAKRVG- QITSFR EVEPAILDKILAYQEIE

Table 1 (continued)

Protein id	Protein name in *Arabidopsis*	Alternative names in other species	Function in meiosis	Mutant phenotype	References	Sequence information
AT1G01690	PRD3	Mer2(S. cerevisiae) Rec15(S. pombe)	DSB	Defects in DSB formation leading to defects in synapsis and chiasma formation	De Muyt et al. 2009	> sp|Q0WWX5|PRD3_ARATH Putative recombination initiation defects 3 OS = Arabidopsis thaliana OX = 3702 GN = PRD3 PE = 1 SV = 2 MKMNINKACDLKSISVFPPNLRRSAEPQASQQLRSQQSQQSFSQGPSSSQRGCGGFSQMT QSSIDELLINDQRFSSQERDLSLKKVSSCLPPINHKREDSQLVASRSSSGLSRRWSSASI GESKSQISEELEQRFGMMETSLSRFGMMLDSIQSDIMQANRGTKEVFLETERIQQKLTLQ DTSLQQLRKEQADSKASLDGGVKFILEEFSKDPNQEKLQKILQMLTTIPEQVETALQKIQ REICHTFTREIQVLASLRTPEPRVRVPTAPQVKAKENLPEQRGQAAKVLTSLKMPEPRVQ VPAAPQAKENFPEQRGPVAKSNSFCNTTLKTKQPQFPRNPNDASARAVKPYLSPKIQVGC WKTVKPEKSNFKKRATRKPVKSESTRTQFEQCSVVIDSDEEDIDGGFSCLINENTRGTNF EWDAEKETERILRTARRTKRKFGNPIIIN
AT3G13170	SPO11-1	SPO11(mammals) Spo11(S. cerevisiae) Rec12(S. pombe)	DSB	Defects in synapsis, bivalent formation, meiotic recombination reduction, defects in chromosome segregation, semi-sterile phenotype	Grelon et al. 2001	> sp|Q9M4A2|SPO11_ARATH Meiotic recombination protein SPO11-1 OS = Arabidopsis thaliana OX = 3702 GN = SPO11-1 PE = 1 SV = 1 MEGKFAISESTNLLQRIKDFTQSVVVDLAEGRSPKISINQFRNYCMNPEADCLCSSDKPK GQEIFTLKKEPQTYRIDMLLRVLLIVQQLLQENRHASKRDIYYMHPSAFKAQSIVDRAIG DICILFQCSRYNLNVVSVGNGLVMGWLKFREAGRKFDCLNSLNTAYPVPVLVEEVEDIVS LAEYILVVEKETVFQRLANDMFCKTNRCIVITGRGYPDVSTRRFLRLLMEKLHLPVHCLV DCDPYGFEILATYRFGSMQMAYDIESLRAPDMKWLGAFPSDSEVYSVPKQCLLPLTEEDK KRTEAMLLRCYLKREMPQWRLELETMLKRGVKFEIEALSVHSLSFLSEVYIPSKIRREVS SP
AT1G63990	SPO11-2	SPO11(mammals) Spo11(S. cerevisiae) Rec12(S. pombe)	DSB	Severe defects in synapsis, reduction in meiotic recombination, sterility	Stacey et al. 2006	> sp|Q9M4A1|SPO12_ARATH Meiotic recombination protein SPO11-2 OS = Arabidopsis thaliana OX = 3702 GN = SPO11-2 PE = 1 SV = 1 MEESSGLSSMKFFSDQHLSYADILLPHEARARIEVSVLNLLRILNSPDPAISDLSLINRK RSNSCINKGILTDVSYIFLSTSFTKSSLTNAKTAKAFVRVWKVMEICFQILLQEKRVTQR ELFYKLLCDSPDYFSSQIEVNRSVQDVVALLRCSRYSLGIMASSRGLVAGRLFLQEPGKE AVDCSACGSSGFAITGDLNLLDNTIMRTDARYIIIVEKHAIFHRLVEDRVFNHIPCVFIT AKGYPDIATRFFLHRMSTTFPDLPILVLVDWNPAGLAILCTFKFGSIGMGLEAYRYACNV KWIGLRGDDLNLIPEESLVPLKPKDSQIAKSLLSSKILQENYIEELSLMVQTGKRAEIEA LYCHGYNYLGKYIATKIVQGKYI
AT1G13330	HOP2/AHP2	Hop2(S. cerevisiae) Meu13(S. pombe)	Strand invasion	Sterility, defects in bivalent formation, chromosome fragmentation, chromatin bridges, unbalanced segregation	Schommer et al. 2003	> sp|Q9FX64|HOP2_ARATH Homologous-pairing protein 2 homolog OS = Arabidopsis thaliana OX = 3702 GN = HOP2 PE = 1 SV = 1 MAPKSDNTEAIVLNFVNEQNKPLNTQNAADALQKFNLKKTAVQKALDSLADAGKIT- FKEY GKQKIYIARQDQFEIPNSEELAQMKEDNAKLQEQLQEKKKTISDVESEIKSLQSNLTLEE IQEKDAKLRKEVKEMEEKLVKLREGITLVRPEDKKAVEDMYADKINQWRKRKRM- FRDIWD TVTENSPKDVKELKEELGIEYDEDVGLSFQAYADLIQHGKKRPRGQ

Table 1 (continued)

Protein id	Protein name in *Arabidopsis*	Alternative names in other species	Function in meiosis	Mutant phenotype	References	Sequence information
AT4G29170	MND1	Mnd1(yeast) MND1(mammals)	Strand invasion	Failure in pairing and synapsis, chromosome fragmentation, mis segregation, formation of inviable gametes, sterility	Kerzendorfer et al. 2006; Panoli et al. 2006	> sp|Q8GYD2|MND1_ARATH Meiotic nuclear division protein 1 homolog OS = Arabidopsis thaliana OX = 3702 GN = MND1 PE = 1 SV = 1 MSKKRGLSLEEKREKMLQIFYESQDFFLLKELEKMGPKKGVISQSVKDVIQSLVDDDLVA KDKIGISIYFWSLPSCAGNQLRSVRQKLESDLQGSNKRLAELVDQCEALKKGREESEERT EALTQLKDIEKKHKDLKNEMVQFADNDPATLEAKRNAIEVAHQSANRWTDNIFTL- RQWCS NNFPQAKEQLEHLYTEAGITEDFDYIELSSFPLSSSHEADTAKQLVQDEA
AT3G22880	DMC1	Dmc1(yeast) DMC1(mammals)	Strand invasion	Formation of univalents and reduced fertility	Couteau et al. 1999; Crismani et al. 2013	> sp|Q39009|DMC1_ARATH Meiotic recombination protein DMC1 homolog OS = Arabidopsis thaliana OX = 3702 GN = DMC1 PE = 1 SV = 2 MMASLKAEETSQMQLVEREENDEDEDLFEMIDKLIAQGINAGDVKKLQEAGIHTC- NGLMM HTKKNLTGIKGLSEAKVDKICEAAEKIVNFGYMTGSDALIKRKSVVKITTGCQALDDLLG GGIETSAITEAFGEFRSGKTQLAHTLCVTTQLPTNMKGGNGKVAYIDTEGTFRPDRIVPI AERFGMDPGAVLDNIIYARAYTYEHQYNLLLGLAAKMSEEPFRILIVDSIIALFRVDFTG RGELADRQQKLAQMLSRLIKIAEEFNVAVYMTNQVIADPGGGMFISDPKKPAGGHV- LAHA ATIRLLFRKGKGDTRVCKVYDAPNLAEAEASFQITQGGIADAKD
AT4G24710	PCH2	Pch2(yeast) PCH2/TRIP13 (mammals) CRC1(rice)	Strand invasion	Defects in CO maturation leading to the formation of univalents	Lambing et al. 2015	> sp|Q8H1F9|PCH2_ARATH Pachytene checkpoint protein 2 homolog OS = Arabidopsis thaliana OX = 3702 GN = PCH2 PE = 2 SV = 1 MVEDPIPLPNASMEVSYQNPIEAATIPVQIAVAEPVATPNPPPCLHENKFLVSVEVCLKP SSTARLEDVQRAVERMLENRSMSYADGLVLIPADDLFLVDNVQRICICDTEEWVKNNDVL LFWQVKPVVHTFQLIEEGPCEDLCADGQPASFNEWILPAKEFDGLWESLIYESGLKQRLL RYAASALLFTQKGVNPNLVSWNRIILLHGPPGTGKTSLCKALAQKLSIRCNSRYPHCQLI EVNAHSLFSKWFSESGKLVAKLFQKIQEMVEEDGNLVFVLIDEVESLAAARKAALSGSEP SDSIRVVNALLTQMDKLKSAPNVIILTTSNITTAIDVAFVDRADIKAYVGPPTLHVRYEI LRSCVEELISKGIISSFQGCDGLSIPSFSSLKEKLSESEVHDTNTVPWFCKQLIEAAKGC EGLSGRSLRKLPFLAHAALADPYSHDPSNFLCTMIETAKREKSEQPE
AT1G53490	HEI10	HEI10/RNF212 (mammals) Zip3/Hei10(yeast)	ZMM	Asymmetric tetrads or polyads, fertility defects leading to reduced seed number per silique	Chelysheva et al. 2012	> sp|F4HRI2|HEI10_ARATH E3 ubiquitin-protein ligase CCNB1IP1 homolog OS = Arabidopsis thaliana OX = 3702 GN = HEI10 PE = 2 SV = 1 MRCNACWRDLEGRAISTTCGHLLCTEDASKILSNDGACPICDQVLSKSLMKPVDINPNEE WINMAMAGISPQILMKSAYRSVMFYIAQRDLEMQYKMNRVVAQCRQKCEGMQAKF- SEKME QVHTAYQKMGKRCQMMEQEVENLTKDKQELQEKFSEKSRQKRKLDEMYDQLRSEYES- VKR TAIQPANNFYPRHQEPDFFSNPAVNMMENRETIRKDRSFFSPATPGPKDEIWPARQNSSN SGPFDISTDSPAIPSDLGNRRAGRGHPVYGGGGTANPQSTLRNLILSPIKRSQLSRSRPQ LFTL

Table 1 (continued)

Protein id	Protein name in *Arabidopsis*	Alternative names in other species	Function in meiosis	Mutant phenotype	References	Sequence information
AT2G30480	HEIP1	Reported to be plant specific so far	ZMM	Severe reduction of chaisma frequency, formation of univalent, sterility	Li et al. 2018	> tr|F4INT5|F4INT5_ARATH Uncharacterised protein OS = Arabidopsis thaliana OX = 3702 GN = At2g30480 PE = 4 SV = 1 MLQWMGGSRRKVAASHTSVKKRQKQYFEQRRQQQHQFTVGSESCSNDINNSNQHL-REHQS LDILNLLNLSTATPECKPSGPENGMQDLDADFYSLKDNMSGVGSSFNHIAEPTSSKRTLF SIPDNQTNDFKKANTDNQTNDFKKTNTTADLMDGTERKLSVFDLVGDDHTTTNLEEC-SPS EAHMAFSVEGLGKINTETPVNSPQPSDRTFVYRCSSPWKDTGQPDTSHVRGRLNDFE-NEV DTMIQSSKMFQDDSLYRSPIGIHAKDGGRKQKLQTFSDHLHKQYSDSRNYFCDVAD-FNNS RFSDDEWNAKPAFLDDGEDSFYWKAEQPCQKESLNPDFLKYCNDCTESRSSTEHHRK-KKR DYLETTWRSNIRDSPTRRSHLLKRNIDYPSFAKAATSDFDFDNVFDRPVWSSIVLEEDKD SHSLRSEESCSSSAVWTNETHNSQFETNTRQRKRETNKFSNLGDKKYINTDLFQESWEDW EVDDQHMKRQVRSGKQGRLSNSGKLKSTSQRKGGLDASYDWFGEGFTSAGINSEIT-SERN KPYPFLNPERGSSHWRSSRAPDSIPETWIPKFSVGGTGDDDDGDHDEEDYVNCLSANHKS KLAGDTCGFENDTLSENDNEQSREVNHPKNQGDETSSSIAKSLSDENDVVRCNPNK-EVME ARHQRNRESGEKTSRDPFQQMIMLERRTLQLVCFNKALLLDSLKT

Table 1 (continued)

Protein id	Protein name in *Arabidopsis*	Alternative names in other species	Function in meiosis	Mutant phenotype	References	Sequence information
AT3G27730	MER3, RCK, ROCK-N-ROLLERS	Mer3(yeast) HFM1(mammals)	ZMM	Low levels of fertility due to defects in synapsis and CO formation	Chen et al. 2005; Mercier et al. 2005	> sp|Q5D892|MER3_ARATH DExH-box ATP-dependent RNA helicase DExH17 OS = Arabidopsis thaliana OX = 3702 GN = MER3 PE = 2 SV = 1 MDTHTLKSVSDLPGNFRSAFSFRYFNSLQSECFPLCFHSDINMIISAPTGSGKTVLFELC ILRLFSKSISKEGSFLHAKGALKTVYISPSKALVQEKLRDWNQKFNSWGISCLELTGDNE TYSTKNIQDADIILTTPEKFDAVSRYRVTSGGLGFFSDIALVLIDEVHLLNDPRGAALEA IVSRLKILSSNHELRSSTLASVRLLAVSATIPNIEDLAEWLKVPTAGIKRFGEEMRPVKL TTKVFGYAAAKNDFLFEKRLQNYIYDILMQYSKGKSALVFCSTRK- GAQEAAQKLAQTAMT YGYSNPFIKSREQLERLREASPMCSDKQMQSYILQGVGYHNGGLCQKDRSLVEGL- FLNGD IQVICTTNTLAHGINLPAHTVVIKSTQHFNKEKGHYMEYDRSTLLQMSGRAGRPPFDDTG MVIIMTRRETVHLYENLLNGCEVVESQLLPCLIEHLTAEIVQLTISDITRAIEWMKCSYL YVRMKKNPENYAIKKGIPKDRVEKHLQELCLQKINELSQYQMIWTDTDGFVLK- PEEPGRL MTKYYLKFETMKYIINTPTSYSLDEALHIVCHAEEISWIQLRRNEKKTLNDVNADKEGRL RFHINDNKGKRKKRIQTREEKLFVLANDWLTGDPSVHDLSMTQDANSICSNGSRI- ARCMK EYFIYKKNYKGTLSSTLLAKSLYQKLWDDSPYLLKQLPGIGMVTAKALHSMGVRS- FEALA EADPRRIEIVTGRKYPFGNTIKESLSSLPPKVEIKVEEVDCQKQGISKLAVTLSRVSQPL QSTKRHYADLIVGSEEENLIHFHEKIRMEDFSSPYSVTVLLERPHQQTKVTVKADLIFEE YIGIDLHETLLLKKANNNKVNYKSENRMPQYYPPMASACIADDDNPVTSGPSNRKDK- KDD MPSFKLIDDDSEEEKEPYVTMEEDDCVIINEHTVFDHIREKAKCFPSLNPLNPTSSPASG KSILKRKSLVENNSPELDPLFQYDSVFDLPTNTKDIKQSAQQITSPGYASFAEKTETERP FSDETIFNYIRKRSKNSPALATSKIENPITISSQEGRNAEISPYRTYGLLVSPATKIPRI TSDAPSEILSFDISMVKRSDTSLEQTKGFCSTLAGKSNVSDSFLGFKSIFSFL

Table 1 (continued)

Protein id	Protein name in *Arabidopsis*	Alternative names in other species	Function in meiosis	Mutant phenotype	References	Sequence information
AT4G17380	MSH4	Msh4(yeast) MSH4(mammals)	ZMM	Delayed/incomplete synapsis but doesn't prevent meiosis, reduction in chiasma frequency, formation of univalents, nondisjunction of chromosomes leading to severe reduction in fertility	Higgins et al. 2004	> sp|F4JP48|MSH4_ARATH DNA mismatch repair protein MSH4 OS = Arabidopsis thaliana OX = 3702 GN = MSH4 PE = 2 SV = 1 MEDDGGERSSFVAGLIENRAKEVGMAAFDLRSASLHLSQYIETSSSYQNTKTLLRFYDPS VIIVPPNKLAADGMVGVSELVDRCYSTVRKVVFARGCFDDTKGAVLIQNLAAEEPLALGL DTYYKQHYLSLAAAAATIKWIEAEKGVIVTNHSLTVTFNGSFDHMNIDATSVENLELIDP FHNALLGTSNKKRSLFQMFKTTKTAGGTRLLRANLLQPLKDIETINTRLDCLDELMSNEQ LFFGLSQVLRKFPKETDRVLCHFCFKPKKVTEAVIGFENTRKSQNMISSIILLKTALDAL PILAKVLKDAKCFLLANVYKSVCENDRYASIRKKIGEVIDDDVLHARVPFVARTQQCFAL KAGIDGFLDIARRTFCDTSEAIHNLASKYREEFNLPNLKLPFNNRQGFFFRIPQKEVQGK LPNKFTQVVKHGKNIHCSSLELASLNVRNKSAAGECFIRTETCLEALMDAIREDISALTL LAEVLCLLDMIVNSFAHTISTKPVDRYSRPELTDSGPLAIDAGRHPILESIHNDFVSNSI FMSEATNMLVVMGPNMSGKSTYLQQVCLVVILAQIGCYVPARFATIRVVDRIFT- RMGTMD NLESNSSTFMTEMRETAFIMQNVTNRSLIVMDELGRATSSSDGLAMAWSCCEYLLSL- KAY TVFATHMDSLAELATIYPNVKVLHFYVDIRDNRLDFKFQLRDGTLHVPHYGLLLAE- VAGL PSTVIDTARIITKRITDKENKRIELNCGKHHEIHRIYRVAQRLICLKYSRQTEDSIRQAL QNLNESFTEERL
AT3G20475	MSH5	Msh5(yeast) MAH5(mammals)	ZMM	Reduction in chiasma frequency, formation of univalents, reduction in fertility	Higgins et al. 2008a, b; Lu et al. 2008	> sp|F4JEP5|MSH5_ARATH DNA mismatch repair protein MSH5 OS = Arabidopsis thaliana OX = 3702 GN = MSH5 PE = 2 SV = 1 MEEMEDTETEPQVYMACIQHGRRVGVSYYDCSVRQLHVLEFWEEDCSDFTLINM- VKYQAK PSIIYASTKSEESFVAALQQNDGTDETTMVKLVKSSTFSYEQAWHRLVYLRVTGMDDGLN IKERICYLSSMMDVGSEVQVRVSGGLLAILESERIVETLEQNESGSASIAIDSVMEVPLN KFLKLDAAAHEALQIFQTDKHPSHMGIGRAKEGFSVFGMMNKCATPMGR- RLLRSWFMRPI LDLEVLDRRLNAISFFISSVELMASLRETLKSVKDISHLLKKFNSPTSLCTSNDWTAFLK SISALLHVNKIFEVGVSESLREHMRRFNLDIIEKAGLCISTELDYVYELVIGVIDVTRSK ERGYQTLVKEGFCAELDELRQIYEELPEFLQEVSAMELEHFPHLHKEKLPPCIVYIQQIG YLMCIFGEKLDETALNRLTEFEFAFSDMDGETQRFFYHTSKTRELDNLLGDIYHKILDME RAIIRDLLSHTLLFSAHLLKAVNFVAELDCILSLACVAHQNNYVRPVLTVESLLDIRNGR HVLQEMAVDTFIPNDTEINDNGRIHIITGPNYSGKSIYVKQVALIVFLSHIGSFVPADAA TVGLTDRIFCAMGSKFMTAEQSTFMIDLHQVGMMLRQATSRSLCLLDEFGKGTLTEDGIG LLGGTISHFATCAEPPRVVVCTHLTELLNESCLPVSEKIKFYTMSVLRPDTESANMEEIV FLYRLIPGQTLLSYGLHCALLAGVPEEVVKRAAIVLDAFESNNNVDKLSLDKISSQDQAF KDAVDKFAELDISKGDIHAFFQDIFTS

Table 1 (continued)

Protein id	Protein name in *Arabidopsis*	Alternative names in other species	Function in meiosis	Mutant phenotype	References	Sequence information
AT1G12790	PTD	Spo16(yeast) SPO16(mammals)	ZMM	Reduction in the number of chiasmatas	Lu et al. 2014; Macaisne et al. 2011; Wijeratne et al. 2006	> sp|F4IDW9|PTD_ARATH Protein PARTING DANCERS OS = Arabidopsis thaliana OX = 3702 GN = PTD PE = 1 SV = 1 MATAGSSYSVSTDHQVSSPLVNLGNVAGVCIMSNAWKVEQEPSLINFISAFLSANSFRLN FVSIPPDLIFNCGGVSIAFVFVTKWDFSNVASIFSRVKRLKGQFAQLYVVATLSTKEQSD SFMRSYFQYEMEFGKPAFVQVTDAEMGFEKIVKIAHSRGVCKQQKVASKLKVERKRT- VQD TNIFIRFVTSIPNINKHDANTLYQAIGSIEAIAKASKEDILANTDLSSKKADTLTRFFQD PEFYLSPKFN
AT5G52290	SHOC1, ZYP2	Zip2(yeast) SHOC1/MZIP2 (mammals)	ZMM	Reduction in the number of COs	Macaisne et al. 2008	> sp|F4KG50|SHOC1_ARATH Protein SHORTAGE IN CHIASMATA 1 OS = Arabidopsis thaliana OX = 3702 GN = SHOC1 PE = 1 SV = 1 MRTRFLNIDYFSTPPSHVFETLGFLNLPAPDNFPAPIVYNGEEDRLRFGSIENVSIPIGN LPIEAALSKFLSDVVPDRVSVDYRVFEIDDSSLGVYYSDEKDDGDAIADKATPKIIELET PELDFEMENKLLCTSEDHLQCFSEVLEIKNDPVKYEGSDIILQNSKDIQEQIYSVDYIPS DYFTENNTSVAENECFRKIQPWFKDARFPLLEVDEVNLSELSSLSVLDKVFTVLETIEPQ DTNAGSSLIINSKELIGSKDYDLLDVLSTDCYLNKSGQSDVVPEDEFSEMDIVTILEISN AEEFQGKVAVPVTYEEFQILDVDISDVFDIFLCLQKAIEPEICYGMFSKEMNFKDFDELV VSSELAFTDDAFKSLPTPILHDYEMTRSLELIYEDVLSKIKPQSLSASNDIYLPWNLLEE RNHNHCDYPFEEIVTFNIDYNWEASEGDKWVYDFIFSEDAFCEPLVEKCTEPFYGISNLD EHAPVNTSHGLLENPFQKTGARDCAVDDNAKKATLLFKSMSAFDDLTFFMDPKKAVIEDN LESRVEAAKTTNHKCMSIDSKASCRSGGMHPNPKTEEMILHSVRPSENIQALVGEFVKSY LTLVKDESENLSEDKLKLLSISKGKLIDCIRKANVHKTQLADDKTFTFALLLAIKQMTWY MCFFGIHVAYIYLNKVCRSSNPMKIGLHTLYSAVETEHKSDETDITRSHPSLAVIQGILQ SEFARGNSKALLLAEKVFWSSLKRLLMSMGLSYNDLNSPSPSGNRPNVHEAIELGFLPIS DCLIISYEQISPSFPVENFSVIVEYGGPNASPRYSFPSKLDSFPSFHFIKVELDMPSACG QLCAGVTVPYSLKMIKGDEVETKTGWLEEVLNFVPLEKVCYAGSSETTNESEFISMPQES ERKRGIIEQGLSDQRSVIVVNTKTVDKEMIISRRSTYQKVLAMEKEGVQVVERDSDLPVD LMLSPAVCLLWYDSETVSKKSAATIGTSSSSLSWIGDIATNVLTSLSFSFSTCIMVFEGE PAFLAAVMDSSDELYAAAGSLGISLQMFCSSSANLTDEIILKCIKSSVKLSKLHVKMPES ESLAESFLTKFPSVNPLTAQVILSSSGSLLEFMKLPHKSKVERTQKYHVPEESVDLFSSV CRYGAREDSRSVMTDSSSSVSSGPDSDTHHVSVHSGSKKKQYIAEKDEIDMDDLVHFSPS IEFADTQLKSSGDFQLDDSWSSKDHEIFHFDPVTEFSDAPFKPSGISHPNDSWPSKDPER FDKKSGPGSSSKDTFWEKDQPDFSVEDSLPGIPELEDWSFPVKDKFMSQNRGCKFPVMRD FNLHDNRNSENFIADYKGEVIDRADKYLEEDFPPSPGYNRFARIVSDVNEEELPRKSKSS RKLSFFGSLQPNFPKAADIDSSSERYATEKDSKYDNNTSLRGYADNYPAKRQRTLLEEVL TRRSAVPTTELPFREEISHFGGSPLSNAIRSSNQVQSSPWTVDFLNRVRERSRARKQQQS LPSYASPPSLETPGNIKKANTKRKSPSILEFFKYKGGNKLQEEKRQKRSKNSSASPKNER FYSPLKSCTPIDKRAKQSLSYTANGTGQTKLVWK

Table 1 (continued)

Protein id	Protein name in *Arabidopsis*	Alternative names in other species	Function in meiosis	Mutant phenotype	References	Sequence information
AT5G48390	ZIP4	Zip4/Spo22(yeast) TEX11(mouse)	ZMM	Reduction in the formation of COs	Chelysheva et al. 2007; Kuromori et al. 2008	> sp|B0M1H3|ZIP4L_ARATH TPR repeat-containing protein ZIP4 OS = Arabidopsis thaliana OX = 3702 GN = ZIP4 PE = 2 SV = 1 MRIAEITTPDLRLHHRETDSHTHHPLLSEIELLIQQSEAISKDQPLPQSLPISLRQFLTR LSQLAPFPDNSFKLTIWKLSFRLWNACVDLANAASLQSSLTSAENIANLRHVAADMLFLA KDVTGVPSPTIKSSLFYYKTGLVYHSLKKFDLASDCFERATEIVSKIDIAKISDAGEKKL FLDLNLARSRTAWEISDRNLAVTLLNRAKNLLFGSPDHYKSLSNQFLAFGKSSLSRGDDD CSLNDALRLMNEALDLCEKGLGTAKTREDTTEFTAMRIKTLRFISAVHLQKGEFENVIKC VKVLRNGGNGSDGADQHASLPVLAMKAWLGLGRHSEAEKELRGMVGNNDIP- EAVWVSAVE AYFEVVGTAGAETAKGVFLGLLGRCHVSAKAALRVAHRVLGESRGGDNGSRIRANV- VAQL VSDERVVALFASEAVTKERKAIHSVLWNSASDHFRAKDYETSAEMFEKSMLYIPHDIENR VFRAKGFRVLCLCYLGLSQLDRALEYIEEAEKLEPNIACSFLKFKIYLQKKEHSCAIGQI DAMTSCLDFSPDYLSLSAHEAISCQALPVAVASLSKFLSFYISGKKMPTTEVVVFRTLVT ILTQDIGSETEALNFMLQAQSRASKLGTECFFGLGETGKREQNWFAATCWNLGSRCG- KEK KYELCGEFLRLASEFYGYIDTDESGEDKLMICRSIILSVTAMIALEKQTKSALTETQVKL AAELLVRAGKIMSSSLSDGKDCIMEPELIFMYTLLAYDIHGRLNNSAFQLLVVKTFAGSK SCHYNYLLQLGIFASQSPQSNPDVSTFALNECLSALIASASPEYPTIALIIRKLISIASV HKGDTDDEEAILKMYKQAYRIMVGLKEGEYPTEEGKWLAMTAWNRAALPVRLGQFE- TAKK WLSIGLEIADKVTGMDTYKACMQDYLAGFQTKVSSA
AT4G09140	MLH1	Mlh1(yeast) MLH1(mammals)	ZMM late	Reduction in the formation COs followed by reduction in fertility	Dion et al. 2007; Franklin et al. 2006	> sp|Q9ZRV4|MLH1_ARATH DNA mismatch repair protein MLH1 OS = Arabidopsis thaliana OX = 3702 GN = MLH1 PE = 2 SV = 1 MIDDSSLTAEMEEEESPATTIVPREPPKIQRLEESVVNRIAAGEVIQRPVSAVKELVENS LDADSSSISVVVKDGGLKLIQVSDDGHGIRREDLPILCERHTTSKLTKFEDLFSLSSMGF RGEALASMTYVAHVTVTTITKGQIHGYRVSYRDGVMEHEPKACAAVKGTQIMVENL- FYNM IARRKTLQNSADDYGKIVDLLSRMAIHYNNVSFSCRKHGAVKADVHSVVSPSRLDSIRSV YGVSVAKNLMKVEVSSCDSSGCTFDMEGFISNSNYVAKKTILVLFINDRLVECSALKRAI EIVYAATLPKASKPFVYMSINLPREHVDINIHPTKKEVSLLNQEIIIEMIQSEVEVKLRN ANDTRTFQEQKVEYIQSTLTSQKSDSPVSQKPSGQKTQKVPVNKMVRTDSSDPAGRLHAF LQPKPQSLPDKVSSLSVVRSSVRQRRNPKETADLSSVQELIAGVDSCCHPGMLETVRNCT YVGMADDVFALVQYNTHLYLANVVNLSKELMYQQTLRRFAHFNAIQLSDPAPLSE- LILLA LKEEDLDPGNDTKDDLKERIAEMNTELLKEKAEMLEEYFSVHIDSSANLSRLPVILDQYT PDMDRVPEFLLCLGNDVEWEDEKSCFQGVSAAIGNFYAMHPPLLPNPSGDGIQFYSKRGE SSQEKSDLEGNVDMEDNLDQDLLSDAENAWAQREWSIQHVLFPSMRLFLKPPAS- MASNGT FVKVASLEKLYKIFERC

Table 1 (continued)

Protein id	Protein name in *Arabidopsis*	Alternative names in other species	Function in meiosis	Mutant phenotype	References	Sequence information
AT4G35520	MLH3	Mlh3(yeast) MLH3(mammals)	ZMM late	Reduction in the number of COs, delayed Prophase I	Franklin et al. 2006; Jackson et al. 2006	> sp|F4JN26|MLH3_ARATH DNA mismatch repair protein MLH3 OS = Arabidopsis thaliana OX = 3702 GN = MLH3 PE = 2 SV = 2 MKTIKPLPEGVRHSMRSGIIMFDMARVVEELVFNSLDAGATKVSIFVGVVSCSVKVVDDG SGVSRDDLVLLGERYATSKFHDFTNVETASETFGFRGEALASISDISLLEVRTKAIGRPN GYRKVMKGSKCLHLGIDDDRKDSGTTVTVRDLFYSQPVRRKYMQSSPKKVLESIK- KCVFR IALVHSNVSFSVLDIESDEELFQTNPSSSAFSLLMRDAGTEAVNSLCKVNVTDGMLNVSG FECADDWKPTDGQQTGRRNRLQSNPGYILCIACPRRLYEFSFEPSKTHVEFKKWGPVLAF IERITLANWKKDRILELFDGGADILAKGDRQDLIDDKIRLQNGSLFSILHFLDADWPEAM EPAKKKLKRSNDHAPCSSLLFPSADFKQDGDYFSPRKDVWSPECEVELKIQNPKEQGTVA GFESRTDSLLQSRDIEMQTNEDFPQVTDLLETSLVADSKCRKQFLTRCQITTPVNINHDF MKDSDVLNFQFQGLKDELDVSNCIGKHLLRGCSSRVSLTFHEPKLSHVEGYESVVPMIPN EKQSSPRVLETREGGSYCDVYSDKTPDCSLGSSWQDTDWFTPQCSSDRGCVGIGEDFNIT PIDTAEFDSYDEKVGSKKYLSSVNVGSSVTGSFCLSSEWSPMYSTPSATKWESEYQKGCR ILEQSLRLGRMPDPEFCFSAANNIKFDHEVIPEMDCCETGTDSFTAIQNCTQLADKICKS SWGHADDVRIDQYSIRKEKFSYMDGTQNNAGKQRSKRSRSAPPFYREKKRFISLSCKSDT KPKNSDPSEPDDLECLTQPCNASQMHLKCSILDDVSYDHIQETEKRLSSASDLKASAGCR TVHSETQDEDVHEDFSSEEFLDPIKSTTKWRHNCAVSQVPKESHELHGQDGVFDISSGLL HLRSDESLVPESINRHSLEDAKVLQQVDKKYIPIVACGTVAIVDQHAADERIRLEELRTK VLAGKARTVTYLSADQELVLPEMGYQLLQSYSEQIRDWGWICNITVEGSTSFKKNMSIIQ RKPTPITLNAVPCILGVNLSDVDLLEFLQQLADTDGSSTIPPSVLRVLNSKACRGAIMFG DSLLPSECSLIIDGLKQTSLCFQCAHGRPTTVPLVDLKALHKQIAKLSGRQVWHGLQRRE ITLDRAKSRLDNAKS

alignment file using MAFFT (Katoh and Standley 2013). NCBI PSI-BLAST was performed against selected species (Supplementary table 1) representing all plant lineages using *A. thaliana* protein sequence as the query. Initial MAFFT alignment was used as a PSSM upload. E-value threshold of maximum 5e-05 and BLOSUM62 matrix was the parameters used for the analysis. PSI-BLAST was continued by increasing the iteration until desired hits were obtained or until no significant hits were able to be found by PSI-BLAST. FASTA sequence of all the hits was downloaded, aligned by MAFFT, trimmed by trimAl (Capella-Gutierrez et al. 2009), and phylogenetic tree was constructed using IQ-TREE (Nguyen et al. 2015). In cases, where the tree could not be resolved, clustering analysis was performed using CLANS (Frickey and Lupas 2004). Cluster containing the initial query was filtered out, and the phylogenetic tree was constructed as described above. The trees were interpreted manually one by one.

Similarity search with HMMER package and phylogenetic inference

HMMER is a more sensitive approach because it employs a whole profile of sequences as a query for similarity searches (Eddy 2011). This way, the program takes advantage of a diversity of amino acids for each position in order to find sequences with a lower level of conservation or more distantly related sequences. This is particularly important for comparisons of large assemblages of lineages of studies of large-scale evolution. In order to build a profile for HMMER searches, one needs to provide an initial trimmed multiple alignment of sequences, (we used MAFFT (Katoh and Standley 2013) and trimAl (Capella-Gutierrez et al. 2009) for alignment and trimming in this pipeline). This initial file is used as input for *hmmbuild* tool in order to generate the profile. The profile is then employed for searches against a database using *hmmsearch* tool. IDs obtained as an output of *hmmsearch* are selected up to an arbitrary threshold (normally e-6) which are used to recover the complete sequences from the database using another tool of the package, the *esl-sfetch* tool. Sequences obtained this way may be used for further analyses, especially phylogeny inference. For phylogenetic inferences, the sequences are aligned and trimmed using the same methods above and directed as input files for a powerful program for phylogeny inference, in this case, IQ-Tree (Minh et al. 2020). The phylogenies obtained this way are then analysed one by one for evolution patterns.

A comprehensive homology search was carried out by PSI-BLAST and HMMER throughout Archaeplastida. The results from both the analysis were compiled in the final figure. For a simplistic view, in some cases, only few representatives were mentioned for a lineage in the final figure and the rest were concatenated in the "Others" option

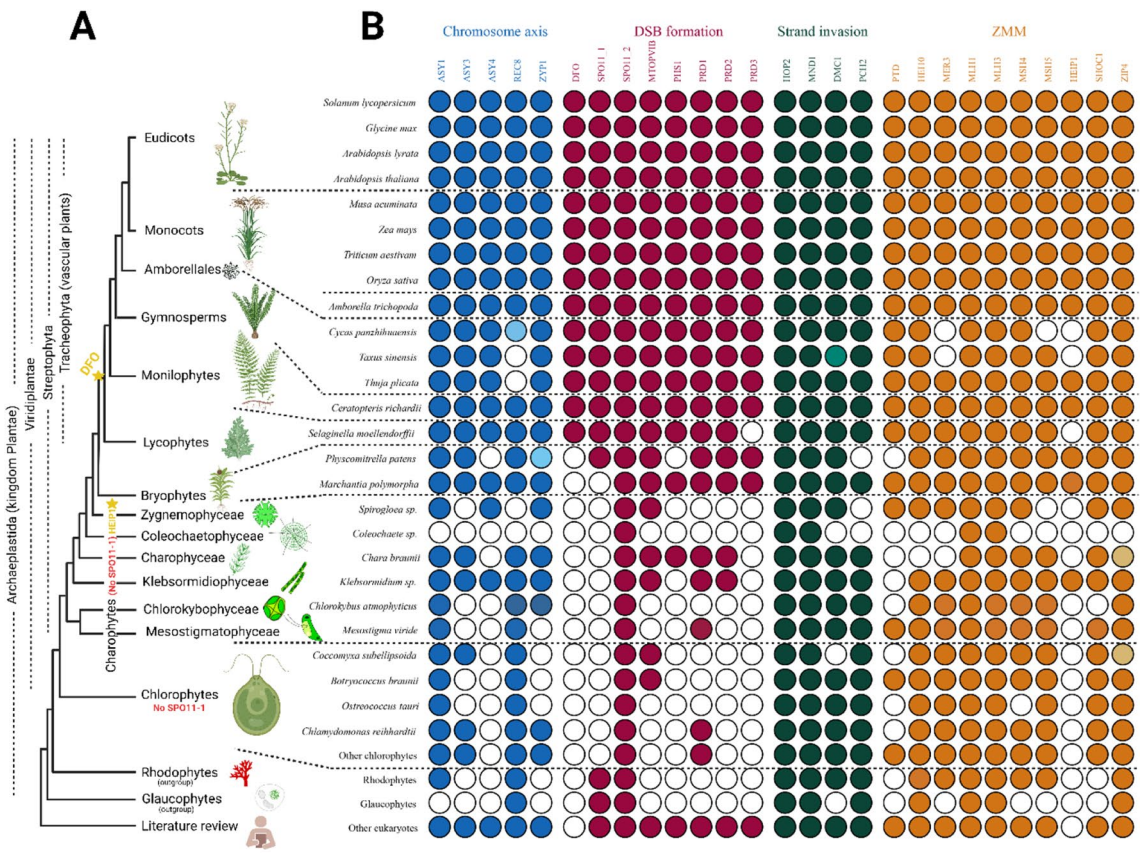
Fig. 1 Tracing the conservation of the meiotic machinery among plants. **A** Representative phylogenetic relationship illustration among the main plant lineages, showing the evolutionary events of important meiotic proteins. Loss of *SPO11-1* in Chlorophyta and Charophyta is indicated. Yellow star represents the possible emergence of the meiotic proteins described only in plants till now—HEIP1 and DFO. **B** Using protein homology searches, PSI-BLAST and HMMER, we inferred either presence (coloured circles) or absence (empty circles) of meiotic-specific proteins in all main Viridiplantae lineages. In case of chlorophytes and charophytes, only representative species are shown and the rest are represented as "Others" for chlorophytes. Members of Glaucophyta and Rhodophyta were included in the analysis and represented as outgroups in the figure. See the supplementary table 1 for the whole list of species used in the analysis. Additional information about non-plant homologs obtained based on literature review is added to the figure. Colour code represents the four meiotic pathways according to which the proteins are classified in our analysis. Fully coloured circles = ortholog is detected in our analysis, light coloured circles = a homolog was obtained as a hit but we are unsure whether it is the right ortholog, white coloured (empty) circle = ortholog was not detected. **C** Phylogenetic tree of *SPO11* showing its pattern of duplication across different lineages. Note that the meiotic-specific *SPO11-1* is missing in chlorophytes and charophytes

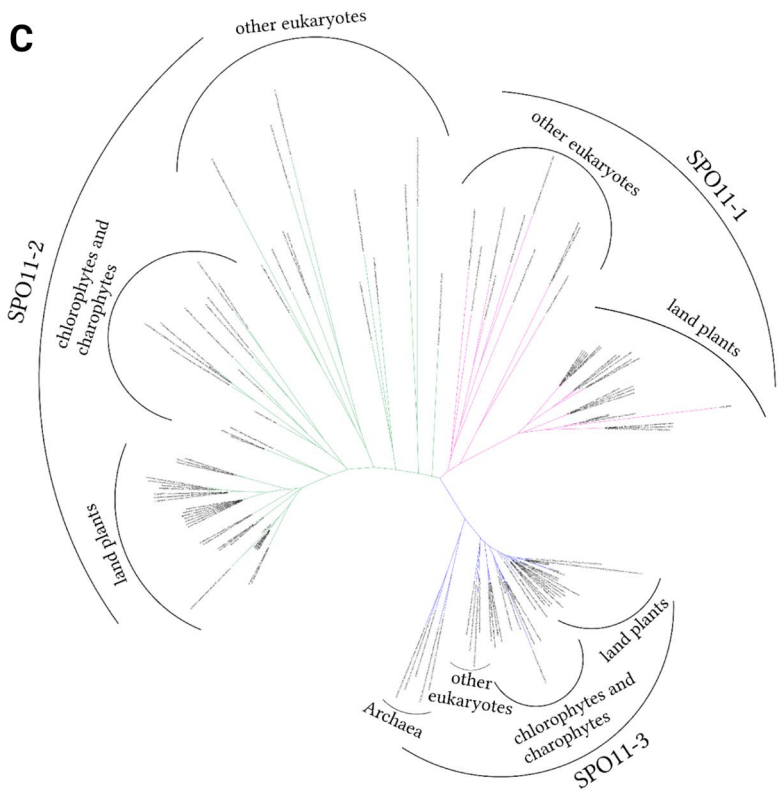
(Fig. 1A, B). For further details, we recommend the readers to look into the Supplementary Table 2 and the phylogenetic trees (https://data.cyverse.org/dav-anon/iplant/home/gokilavani/Tracing_the_evolution_of_the_plant_meiotic_molecular_machinery). Glaucophytes and rhodophytes were considered to provide a root for your analyses, and as mentioned above, this paper focusses only on discussing the meiotic machinery in Viridiplantae.

Results and discussion

Chromosome axis and synaptonemal complex elements are structurally highly conserved but markedly divergent at the sequence level

ASY1, ASY3, REC8 and ZYP1 were detected in all the species or at least in one representative species of all the major Viridiplantae lineages used for the analysis. Exceptionally, we detected ASY4 only in streptophytes, and not in chlorophytes (Fig. 1B). Supporting our analysis, ASY4 was also previously not identified outside land plants (Chambon et al. 2018). On the contrary, ASY3 which interacts with ASY4 (Chambon et al. 2018) was detected in chlorophytes as well. It is important to consider that ASY4 is reported to lack functional domains which constitutes the most conserved region of a protein sequence. Sequence divergence is a feature of the chromosome axis proteins. Axis elements and central elements of the SC exhibit poor similarity between species at the sequence level, but their structure and function are widely conserved (Chambon et al. 2018). The lower sequence conservation could explain why


A
B
Chromosome axis
DSB formation
Strand invasion
ZMM
ASY1
ASY3
ASY4
REC8
ZYP1
DFO
SPO11_1
SPO11_2
MTOPVIB
PHS1
PRD1
PRD2
PRD3
HOP2
MND1
DMC1
PCH2
PTD
HEI10
MER3
MLH1
MLH3
MSH4
MSH5
EIP1
SHOC1
ZIP4
Archaeplastida (kingdom Plantae)
Viridiplantae
Streptophyta
Tracheophyta (vascular plants)
Eudicots
Monocots
Amborellales
Gymnosperms
Monilophytes
Lycophytes
Bryophytes
Zygnemophyceae
Coleochaetophyceae
Charophyceae
Klebsormidiophyceae
Chlorokybophyceae
Mesostigmatophyceae
Charophytes
Chlorophytes
No SPO11-1
Rhodophytes
Glaucophytes
Literature review
Solanum lycopersicum
Glycine max
Arabidopsis lyrata
Arabidopsis thaliana
Musa acuminata
Zea mays
Triticum aestivum
Oryza sativa
Amborella trichopoda
Cycas panzhihuaensis
Taxus sinensis
Thuja plicata
Ceratopteris richardii
Selaginella moellendorffii
Physcomitrella patens
Marchantia polymorpha
Spirogloea sp.
Coleochaete sp.
Chara braunii
Klebsormidium sp.
Chlorokybus atmophyticus
Mesostigma viride
Coccomyxa subellipsoida
Botryococcus braunii
Ostreococcus tauri
Chlamydomonas reihhardtii
Other chlorophytes
Rhodophytes
Glaucophytes
Other eukaryotes


C
other eukaryotes
other eukaryotes
SPO11-1
land plants
SPO11-2
chlorophytes and charophytes
land plants
land plants
other eukaryotes
Archaea
chlorophytes and charophytes
SPO11-3

we could not detect *A. thaliana* homolog of ASY4 in distant algal species. For example, *A. thaliana* ASY3, mammalian SYCP2 and yeast Red1 ensures the same function but lacks sequence similarity, likewise *A. thaliana* ASY4 and mammalian SYCP3 (Chambon et al. 2018). Such possibilities cannot be ruled out in this case which is beyond the scope of algorithms used in our analysis.

The evolution of the meiotic DSB machinery in plants

Among the eight DSB formation proteins we analysed, *DFO* was not detected in Chlorophyta, Charophyta and Bryophyta, *PHS1* and *PRD2* in Chlorophyta and PRD3 and SPO11-1 in Chlorophyta and Charophyta. The rest of the candidates were detected in all Viridiplantae lineages. DFO is a plant-specific protein involved in the formation of DSBs. It has been not reported in other eukaryotic super-groups yet (Zhang et al. 2012). In our analysis, DFO homologs were detected only in the vascular plants and not in other plant lineages, suggesting that *DFO* evolved only in the common ancestor of vascular plants. The homologs of the other three missing candidates PHS1/Rec114, PRD2/Mei4 and PRD3/Mer2 were described to interact with each other and form the RMM complex in *Saccharomyces cerevisiae* (Maleki et al. 2007; Yadav and Claeys Bouuaert 2021). Recently, it has been described, PHS1, PRD2 and the plant-specific DFO forms the RMM-like complex also in *A. thaliana*. PRD3 does not interact with the RMM-like proteins and is proposed to have a different role, likely in coordinating DSB formation and repair mechanisms in *A. thaliana*. PHS1/Rec114 is characterised to have role in DSB formation in species studied so far including maize, except *A. thaliana* where it is proposed not necessary for DSB formation but in regulating meiotic recombination (Vrielynck et al. 2021). Therefore, it becomes evident, and RMM complex has divergent roles in some cases like PRD3 and PHS1. Notably, PHS1/Rec114, PRD2/Mei4, PRD3/Mer2 homologs are conserved across different phyla, but their conservation at the protein sequence level is very weak (Vrielynck et al. 2021). PRD2 and PRD3 have no functional domains reported, except for the presence of several alpha helixes and coiled-coil motifs (De Muyt et al. 2009; Jiang et al. 2009; Vrielynck et al. 2021). The divergence observed among RMM proteins and absence of conserved domains in PRD2, PRD3 explains why we could not detect RMM homologs and plant-specific DFO, part of *A. thaliana* RMM-like complex in distant relatives of our analysis, reconfirming the minimal conservation of RMM proteins.

SPO11 heterodimerisation has likely evolved in land plants

SPO11 is encoded by a single gene in most organisms (Malik et al. 2007); however, plants differ from yeasts and animals in having several *SPO11* homologs: two paralogs (*SPO11-1* and *SPO11-2*) are involved in meiosis of *A. thaliana* (Grelon et al. 2001; Hartung and Puchta 2001; Hartung et al. 2007; Stacey et al. 2006), where they seem to form a heterodimer that is required for meiotic DSB formation, whereas *SPO11-3* is involved in somatic DNA metabolism (Hartung et al. 2007; Sugimoto-Shirasu et al. 2002; Yin et al. 2002). However, the exact origin of *SPO11-1* and *SPO11-2* duplication and its relation to the heterodimerisation in plants outside *A. thaliana* remained unanswered. This caught our special attention and we further expanded our phylogenetic analysis by including more non-plant representatives from amoeba and archaea. This helped us in tracing the origin of SPO11 duplication in plants. *SPO11-3* (Fig. 1C), which is very similar to archaeal sequences, was detected in all the lineages analysed. Remarkably, among Viridiplantae lineages, our analysis could detect both *SPO11-1* and *SPO11-2* only in land plants, except for *Marchantia polymorpha*, whereas chlorophytes and charophytes have only *SPO11-2* and they seem to lack *SPO11-1* (Fig. 1C). Suggesting two scenarios: 1- heterodimerization of SPO11 evolved in land plants, 2- heterodimerization evolved earlier in eukaryotes but was later lost independently in several lineages and replaced by a homodimer. However, the duplication of *SPO11* is ancestral to eukaryotes, or happened very early in the evolution of eukaryotes as suggested by our phylogenetic analysis and is in agreement as reported earlier (Malik et al. 2007). Members of Amoebozoa, glaucophytes and red algae (grouped under other eukaryotes in Fig. 1C, B), share the same duplication with land plants and have both *SPO11-1* and *SPO11-2* paralogs (Fig. 1C). Thus, we propose that duplication of *SPO11* is ancestral to eukaryotes and most likely *SPO11-1* gene has been lost in both chlorophyte and charophyte lineages after the duplication event. Whether SPO11 activity function as a homodimer in these two lineages needs further investigation.

Strand invasion is the most conserved meiotic pathway

HOP2, MND1, DMC1, PCH2 are the proteins involved in strand invasion mechanism used for our analysis. It is noteworthy that it is the only group where all the proteins are found in all the lineages in our analysis except some specific cases (Fig. 1B). We observed DMC1 was not detected in glaucophytes analysed but the absence of a complete genome for these species makes it difficult to have a conclusion. DMC1 is the meiotic-specific homolog of bacterial RecA and is required for meiotic homologous recombination. MND1-HOP2 heterodimer promotes DMC1 activity at the DSB sites and promotes stable strand invasion and inter homologue bias (Kerzendorfer et al. 2006). However, some organisms lack DMC1, for example *Drosophila*

melanogaster, Caenorhabditis elegans, Sordaria macrospora, Neurospora crassa, which shows that DMC1 can be dispensable. These organisms also lack the accessary factors HOP2 and MND1. However, Viridiplantae and mammals were reported to have DMC1 (Brown and Bishop 2014; Neale and Keeney 2006). Our analysis also shows that all the major Viridiplantae lineages have DMC1 along with HOP2 and MND1 and it may be essential for meiotic homologous recombination in Viridiplantae. PCH2 has a role in chromosome remodelling during SC formation. The initial characterisation of all these proteins in *A. thaliana* revealed their conservation among eukaryotes and observed functional similarity with their non-plant orthologs (Couteau et al. 1999; Kerzendorfer et al. 2006; Lambing et al. 2015; Schommer et al. 2003). Our analysis also concludes the same that strand invasion proteins are the most conserved among the other meiotic proteins we analysed, even at the sequence level. We speculate that such high conservation is linked to their enzymatic function.

The ZMM pathway is highly conserved and detectable in all plant lineages

PTD, HEI10, MER3, MLH1, MLH3, MSH4, MSH5, SHOC1, ZIP4 are among the ten ZMM pathway proteins analysed, found to be highly conserved in all the major plant lineages. HEIP1 was not detected in chlorophytes. (Fig. 1B). HEIP1 was identified as an interacting partner of HEI10 and suggested to be a member of ZMM pathway as the mutants showed reduced chiasma frequency in rice. It contains a potential plant-specific domain (GCK domain) and not reported outside the plant kingdom till now (Li et al. 2018).This is confirmed in our analysis, and HEIP1 was not detected outside plants and also in the whole chlorophyte lineage. We could not detect HEIP1 in some cases other than chlorophytes as well but at least one species in all other major Viridiplantae lineages had its ortholog. Based on the pattern observed, we propose, HEIP1 is a member of ZMM pathway with possible emergence during the diversification of chlorophytes. PTD orthologs are distant relatives of ERCC1 proteins which are present in both plants and animals (Lu et al. 2014; Wijeratne et al. 2006). SHOC1, the interacting partner of PTD, is a member of XPF superfamily widely present among eukaryotes (Macaisne et al. 2011) and has also been detected in all plant lineages of our study. However, in our analysis, PTD was absent in most of the chlorophytes. PTD may be lost independently from these algae or the protein sequence may be too diverse to be detected by the algorithms given that PTD lacks the conserved motif for endonuclease activity (Wijeratne et al. 2006). Considering both ERCC1 and XPF are structure-specific endonucleases belonging to the XPF superfamily, this difference in the conservation of PTD and SHOC1 implies that individual proteins of the same complex can have different evolutionary trajectories. Another interesting observation is that MER3 was not detected in *Cycas panzhihuaensis* and *Taxus sinensis*. MER3 is highly conserved and *A. thaliana* orthologs were even detected in the most distant algal species used in our analysis. In this case, it may indicate a possible independent loss in the species mentioned above.

Final remarks

Our comprehensive analysis was able to characterise *SPO11* duplication in plant lineages. SPO11-1 is retained and possibly the heterodimerisation of SPO11-1, and SPO11-2 occurs only in land plants of Viridiplantae. We could also trace the possible origin of the meiotic genes, DFO and HEIP1, which is described only in plants till now. Although there is always a possibility that if the proteins are not detected, it does not necessarily mean they are absent. Notwithstanding the ever-growing volume of genome sequence information, some genomes remain incompletely annotated, which may result in the apparent absence of some proteins in the genome/proteome. Thus, although our results are based on more than one homology search approach, the non-detection of protein homologs in our analysis does not always imply their absence in a given species. Indeed, in a few instances, our failure to detect homologs seems suspicious, for example, the absence of MSH5 in *Cycas panzhihuaensis*, PCH2 in *Physcomitrella patens*, among others. These candidates are highly conserved and detected in all other species analysed. Here it becomes difficult to conclude, whether this is an independent loss scenario or it indicates an artefact. Such cases need more studies to give a concrete answer while other cases discussed had a clear pattern. ASY4, DFO, PHS1, PRD2, PRD3, HEIP1 are absent from all the species of a particular lineage. Here we can be more confident that they are putatively absent or have high sequence divergence to be identified by the algorithms. If meiosis is an ancestral characteristic of eukaryotes, then this raises the question of why some of the proteins in the highly conserved meiotic pathways are putatively absent/not recognised in certain lineages. Possible explanation would be either they are poorly conserved or evolved in some ancestor of the land plants but are absent in the others. If sequence divergence is the case, then it remains to be determined why, within the same pathway, some proteins are more divergent than others; moreover, such an explanation potentially hints at other, yet to identified, evolutionary pressures determining the evolution of these proteins. Most of the meiotic proteins which have enzymatic function or a described functional domain, for example ASY1, SPO11, HEI10, MLH1, MLH3 among others, are observed to be highly conserved in our analysis, whereas proteins like PRD2, PRD3 and ASY4, where

functional domains were reported to be absent and do not have an enzymatic function and were less conserved. What also remains to be elucidated is the relevance of lineage-specific loss/gain of certain proteins for meiotic adaptation. Functional validation of selected candidates will be necessary to answer the unanswered questions and to get a complete picture of the different meiotic strategies that have evolved across the massive plant kingdom but we hope our homology search is an attempt to provide first-hand information about the meiotic core proteins across the kingdom.

Limitations of the study

Arabidopsis thaliana protein sequence was used as the initial query in the analysis. We have considered using yeast homologs as the query. Considering, even though meiotic machinery is conserved, not all the proteins are conserved at sequence level between yeast and plants. In some cases, past studies have reported that the yeast and *Arabidopsis* homologs have functional conservation but divergent at the sequence level. The other way around, plant-specific protein like DFO is not reported in yeast. Considering the above points, we narrowed down our aim to look only for the proteins reported in the model plant *Arabidopsis thaliana* among other Viridiplantae lineages and not to look for all the reported meiotic proteins. However, the latter is very exciting but the sequence-based homology search algorithms used in this work will not suffice the needs. Involving structure-based algorithms and carefully looking for functional domains of each protein case by case can be considered but is not the scope of this manuscript.

The sensitivity of the algorithms decreased in the evolutionary distant lineages of *Arabidopsis thaliana* due to sequence divergence and one may think, this could bias our findings. To increase the chances of finding the orthologs, most of the algae which had omics data were included in our analysis. However, we would like to bring to your kind notice that the data sets available for algae were limited. In many cases, the data set available was either vegetative transcriptome or draft genome. This was particularly the case for *Coleochaete* and glaucophytes. Since we are dealing with meiotic-specific candidates, the transcriptome data from vegetative phase may not have their expression, and thus, no hits will be obtained. All the cases, where hits were not obtained, were carefully considered. Due to limitations of the analysis used, no hits do not necessarily mean the protein is absent. Only the cases, where hits were not obtained in the whole lineage was considered as a clear pattern unless specifically mentioned and interpreted further.

Author contribution statement GT and PGH performed the analysis. GT and AM wrote the first draft with subsequent input from PGH and RM. RM and AM conceived and coordinated the study.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s00497-022-00456-1.

Acknowledgements We thank Neysan Donnelly for helping with manuscript editing. We thank the DAAD/India for funding G.T. to do her Ph.D. This study was supported by the Max Planck Society.

Funding Open Access funding enabled and organized by Projekt DEAL. DAAD, Research Grants—Doctoral Programmes in Germany, Gokilavani Thangavel, 2019/20, Gokilavani Thangavel, Max-Planck-Gesellschaft.

Data availability All phylogenetic trees and alignments generated in this study can be freely accessed here: https://data.cyverse.org/davanon/iplant/home/gokilavani/Tracing_the_evolution_of_the_plant_meiotic_molecular_machinery.

Declarations

Conflict of interest The authors declare no competing interest.

References

Bai X, Peirson BN, Dong F, Xue C, Makaroff CA (1999) Isolation and characterization of SYN1, a RAD21-like gene essential for meiosis in arabidopsis. Plant Cell 11:417–430

Berchowitz LE, Francis KE, Bey AL, Copenhaver GP (2007) The role of AtMUS81 in interference-insensitive crossovers in A. thaliana. PLOS Genet 3:e132

Bergerat A, de Massy B, Gadelle D, Varoutas P-C, Nicolas A, Forterre P (1997) An atypical topoisomerase II from archaea with implications for meiotic recombination. Nature 386:414–417

Bhatt AM, Lister C, Page T, Fransz P, Findlay K, Jones GH, Dickinson HG, Dean C (1999) The DIF1 gene of Arabidopsis is required for meiotic chromosome segregation and belongs to the REC8/RAD21 cohesin gene family. Plant J 19:463–472

Börner GV, Kleckner N, Hunter N (2004) Crossover/noncrossover differentiation, synaptonemal complex formation, and regulatory surveillance at the Leptotene/Zygotene transition of meiosis. Cell 117:29–45

Brown MS, Bishop DK (2014) DNA strand exchange and RecA homologs in meiosis. Cold Spring Harb Perspect Biol 7:a016659

Cai X, Dong F, Edelmann RE, Makaroff CA (2003) The Arabidopsis SYN1 cohesin protein is required for sister chromatid arm cohesion and homologous chromosome pairing. J Cell Sci 116:2999–3007

Capella-Gutierrez S, Silla-Martinez JM, Gabaldon T (2009) trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. Bioinformatics 25:1972–1973

Capilla-Perez 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. Proc Natl Acad Sci USA. https://doi.org/10.1073/pnas.2023613118

Caryl AP, Armstrong SJ, Jones GH, Franklin FCH (2000) A homologue of the yeast HOP1 gene is inactivated in the Arabidopsis meiotic mutant asy1. Chromosoma 109:62–71

Chambon A, West A, Vezon D, Horlow C, De Muyt A, Chelysheva L, Ronceret A, Darbyshire A, Osman K, Heckmann S et al (2018) Identification of ASYNAPTIC4, a component of the Meiotic chromosome axis. Plant Physiol 178:233–246

Chelysheva L, Diallo S, Vezon D, Gendrot G, Vrielynck N, Belcram K, Rocques N, Marquez-Lema A, Bhatt AM, Horlow C et al (2005) AtREC8 and AtSCC3 are essential to the monopolar orientation of the kinetochores during meiosis. J Cell Sci 118:4621–4632

Chelysheva L, Gendrot G, Vezon D, Doutriaux MP, Mercier R, Grelon M (2007) Zip4/Spo22 is required for class I CO formation but not for synapsis completion in Arabidopsis thaliana. PLoS Genet 3:e83

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 Genet 8:e1002799

Chen C, Zhang W, Timofejeva L, Gerardin Y, Ma H (2005) The Arabidopsis ROCK-N-ROLLERS gene encodes a homolog of the yeast ATP-dependent DNA helicase MER3 and is required for normal meiotic crossover formation. Plant J 43:321–334

Corlett RT (2016) Plant diversity in a changing world: status, trends, and conservation needs. Plant Divers 38:10–16

Couteau F, Belzile F, Horlow C, Grandjean O, Vezon D, Doutriaux M-P (1999) Random chromosome segregation without meiotic arrest in both male and female meiocytes of a dmc1 mutant of Arabidopsis. Plant Cell 11:1623–1634

Crismani W, Portemer V, Froger N, Chelysheva L, Horlow C, Vrielynck N, Mercier R (2013) MCM8 is required for a pathway of meiotic double-strand break repair independent of DMC1 in Arabidopsis thaliana. PLoS Genet 9:e1003165

de Massy B, Rocco V, Nicolas A (1995) The nucleotide mapping of DNA double-strand breaks at the CYS3 initiation site of meiotic recombination in Saccharomyces cerevisiae. EMBO J 14:4589–4598

De Muyt A, Vezon D, Gendrot G, Gallois JL, Stevens R, Grelon M (2007) AtPRD1 is required for meiotic double strand break formation in Arabidopsis thaliana. EMBO J 26:4126–4137

De Muyt A, Pereira L, Vezon D, Chelysheva L, Gendrot G, Chambon A, Laine-Choinard S, Pelletier G, Mercier R, Nogue F et al (2009) A high throughput genetic screen identifies new early meiotic recombination functions in Arabidopsis thaliana. PLoS Genet 5:e1000654

Dion E, Li L, Jean M, Belzile F (2007) An Arabidopsis MLH1 mutant exhibits reproductive defects and reveals a dual role for this gene in mitotic recombination. Plant J 51:431–440

Dong H, Roeder GS (2000) Organization of the yeast zip1 protein within the central region of the synaptonemal complex. J Cell Biol 148:417–426

Eddy SR (2011) Accelerated profile HMM searches. PLoS Comput Biol 7:e1002195

Ferdous M, Higgins JD, Osman K, Lambing C, Roitinger E, Mechtler K, Armstrong SJ, Perry R, Pradillo M, Cunado N et al (2012) Inter-homolog crossing-over and synapsis in Arabidopsis meiosis are dependent on the chromosome axis protein AtASY3. PLoS Genet 8:e1002507

France MG, Enderle J, Rohrig S, Puchta H, Franklin FCH, Higgins JD (2021) ZYP1 is required for obligate cross-over formation and cross-over interference in Arabidopsis. Proc Natl Acad Sci USA. https://doi.org/10.1073/pnas.2021671118

Franklin FCH, Higgins JD, Sanchez-Moran E, Armstrong SJ, Osman KE, Jackson N, Jones GH (2006) Control of meiotic recombination in Arabidopsis: role of the MutL and MutS homologues. Biochem Soc Trans 34:542–544

Frickey T, Lupas A (2004) CLANS: a Java application for visualizing protein families based on pairwise similarity. Bioinformatics 20:3702–3704

Gerton JL, Hawley RS (2005) Homologous chromosome interactions in meiosis: diversity amidst conservation. Nat Rev Genet 6:477–487

Grelon M, Vezon D, Gendrot G, Pelletier G (2001) AtSPO11-1 is necessary for efficient meiotic recombination in plants. EMBO J 20:589–600

Hartung F, Puchta H (2001) Molecular characterization of homologues of both subunits A (SPO11) and B of the archaebacterial topoisomerase 6 in plants. Gene 271:81–86

Hartung F, Wurz-Wildersinn R, Fuchs J, Schubert I, Suer S, Puchta H (2007) The catalytically active tyrosine residues of both SPO11-1 and SPO11-2 are required for meiotic double-strand break induction in Arabidopsis. Plant Cell 19:3090–3099

Higgins JD, Armstrong SJ, Franklin FC, Jones GH (2004) The Arabidopsis MutS homolog AtMSH4 functions at an early step in recombination: evidence for two classes of recombination in Arabidopsis. Genes Dev 18:2557–2570

Higgins JD, Sanchez-Moran E, Armstrong SJ, Jones GH, Franklin FC (2005) The Arabidopsis synaptonemal complex protein ZYP1 is required for chromosome synapsis and normal fidelity of crossing over. Genes Dev 19:2488–2500

Higgins JD, Buckling EF, Franklin FCH, Jones GH (2008a) Expression and functional analysis of AtMUS81 in Arabidopsis meiosis reveals a role in the second pathway of crossing-over. Plant J 54:152–162

Higgins JD, Vignard J, Mercier R, Pugh AG, Franklin FC, Jones GH (2008b) AtMSH5 partners AtMSH4 in the class I meiotic crossover pathway in Arabidopsis thaliana, but is not required for synapsis. Plant J 55:28–39

Jackson N, Sanchez-Moran E, Buckling E, Armstrong SJ, Jones GH, Franklin FC (2006) Reduced meiotic crossovers and delayed prophase I progression in AtMLH3-deficient Arabidopsis. EMBO J 25:1315–1323

Jiang H, Wang FF, Wu YT, Zhou X, Huang XY, Zhu J, Gao JF, Dong RB, Cao KM, Yang ZN (2009) MULTIPOLAR SPINDLE 1 (MPS1), a novel coiled-coil protein of Arabidopsis thaliana, is required for meiotic spindle organization. Plant J 59:1001–1010

Katoh K, Standley DM (2013) MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol 30:772–780

Keeney S, Kleckner N (1995) Covalent protein-DNA complexes at the 5' strand termini of meiosis-specific double-strand breaks in yeast. Proc Natl Acad Sci USA 92:11274–11278

Keeney S, Giroux CN, Kleckner N (1997) Meiosis-specific DNA double-strand breaks are catalyzed by Spo11, a member of a widely conserved protein family. Cell 88:375–384

Kerzendorfer C, Vignard J, Pedrosa-Harand A, Siwiec T, Akimcheva S, Jolivet S, Sablowski R, Armstrong S, Schweizer D, Mercier R et al (2006) The Arabidopsis thaliana MND1 homologue plays a key role in meiotic homologous pairing, synapsis and recombination. J Cell Sci 119:2486–2496

Kuromori T, Azumi Y, Hayakawa S, Kamiya A, Imura Y, Wada T, Shinozaki K (2008) Homologous chromosome pairing is completed in crossover defective atzip4 mutant. Biochem Biophys Res Commun 370:98–103

Lambing C, Osman K, Nuntasoontorn K, West A, Higgins JD, Copenhaver GP, Yang J, Armstrong SJ, Mechtler K, Roitinger E et al (2015) Arabidopsis PCH2 mediates meiotic chromosome remodeling and maturation of crossovers. PLoS Genet 11:e1005372

Li Y, Qin B, Shen Y, Zhang F, Liu C, You H, Du G, Tang D, Cheng Z (2018) HEIP1 regulates crossover formation during meiosis in rice. Proc Natl Acad Sci USA 115:10810–10815

Liu J, Wu TC, Lichten M (1995) The location and structure of double-strand DNA breaks induced during yeast meiosis: evidence for a covalently linked DNA-protein intermediate. EMBO J 14:4599–4608

Liu J-G, Yuan L, Brundell E, Björkroth B, Daneholt B, Höög C (1996) Localization of the N-terminus of SCP1 to the central element of the synaptonemal complex and evidence for direct interactions between the N-termini of SCP1 molecules organized head-to-head. Exp Cell Res 226:11–19

Lu X, Liu X, An L, Zhang W, Sun J, Pei H, Meng H, Fan Y, Zhang C (2008) The Arabidopsis MutS homolog AtMSH5 is required for normal meiosis. Cell Res 18:589–599

Lu P, Wijeratne AJ, Wang Z, Copenhaver GP, Ma H (2014) Arabidopsis PTD is required for type I crossover formation and affects recombination frequency in two different chromosomal regions. J Genet Genomics 41:165–175

Macaisne N, Novatchkova M, Peirera L, Vezon D, Jolivet S, Froger N, Chelysheva L, Grelon M, Mercier R (2008) SHOC1, an XPF endonuclease-related protein, is essential for the formation of class I meiotic crossovers. Curr Biol 18:1432–1437

Macaisne N, Vignard J, Mercier R (2011) SHOC1 and PTD form an XPF-ERCC1-like complex that is required for formation of class I crossovers. J Cell Sci 124:2687–2691

Maleki S, Neale MJ, Arora C, Henderson KA, Keeney S (2007) Interactions between Mei4, Rec114, and other proteins required for meiotic DNA double-strand break formation in Saccharomyces cerevisiae. Chromosoma 116:471–486

Malik S-B, Ramesh MA, Hulstrand AM, Logsdon JM Jr (2007) Protist homologs of the Meiotic Spo11 gene and Topoisomerase VI reveal an evolutionary history of gene duplication and lineage-specific loss. Mol Biol Evol 24:2827–2841

Mercier R, Grelon M (2008) Meiosis in plants: ten years of gene discovery. Cytogenet Genome Res 120:281–290

Mercier R, Jolivet S, Vezon D, Huppe E, Chelysheva L, Giovanni M, Nogue F, Doutriaux MP, Horlow C, Grelon M et al (2005) Two meiotic crossover classes cohabit in Arabidopsis: one is dependent on MER3, whereas the other one is not. Curr Biol 15:692–701

Mercier R, Mezard C, Jenczewski E, Macaisne N, Grelon M (2015) The molecular biology of meiosis in plants. Annu Rev Plant Biol 66:297–327

Meuwissen RL, Offenberg HH, Dietrich AJ, Riesewijk A, van Iersel M, Heyting C (1992) A coiled-coil related protein specific for synapsed regions of meiotic prophase chromosomes. EMBO J 11:5091–5100

Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, von Haeseler A, Lanfear R (2020) IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. Mol Biol Evol 37:1530–1534

Neale MJ, Keeney S (2006) Clarifying the mechanics of DNA strand exchange in meiotic recombination. Nature 442:153–158

Nguyen LT, Schmidt HA, von Haeseler A, Minh BQ (2015) IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol Biol Evol 32:268–274

Onn I, Heidinger-Pauli JM, Guacci V, Ünal E, Koshland DE (2008) Sister chromatid cohesion: a simple concept with a complex reality. Annu Rev Cell Dev Biol 24:105–129

Panoli AP, Ravi M, Sebastian J, Nishal B, Reddy TV, Marimuthu MP, Subbiah V, Vijaybhaskar V, Siddiqi I (2006) AtMND1 is required for homologous pairing during meiosis in Arabidopsis. BMC Mol Biol 7:24

Peirson BN, Bowling SE, Makaroff CA (1997) A defect in synapsis causes male sterility in a T-DNA-tagged Arabidopsis thaliana mutant. Plant J 11:659–669

Puttick MN, Morris JL, Williams TA, Cox CJ, Edwards D, Kenrick P, Pressel S, Wellman CH, Schneider H, Pisani D et al (2018) The interrelationships of land plants and the nature of the ancestral embryophyte. Curr Biol 28:733-745.e732

Ronceret A, Doutriaux M-P, Golubovskaya Inna N, Pawlowski Wojciech P (2009) PHS1 regulates meiotic recombination and homologous chromosome pairing by controlling the transport of RAD50 to the nucleus. Proc Natl Acad Sci 106:20121–20126

Sanchez-Moran E, Santos JL, Jones GH, Franklin FC (2007) ASY1 mediates AtDMC1-dependent interhomolog recombination during meiosis in Arabidopsis. Genes Dev 21:2220–2233

Sanchez-Moran E, Osman K, Higgins JD, Pradillo M, Cunado N, Jones GH, Franklin FC (2008) ASY1 coordinates early events in the plant meiotic recombination pathway. Cytogenet Genome Res 120:302–312

Schommer C, Beven A, Lawrenson T, Shaw P, Sablowski R (2003) AHP2 is required for bivalent formation and for segregation of homologous chromosomes in Arabidopsis meiosis. Plant J 36:1–11

Stacey NJ, Kuromori T, Azumi Y, Roberts G, Breuer C, Wada T, Maxwell A, Roberts K, Sugimoto-Shirasu K (2006) Arabidopsis SPO11-2 functions with SPO11-1 in meiotic recombination. Plant J 48:206–216

Sugimoto-Shirasu K, Stacey NJ, Corsar J, Roberts K, McCann MC (2002) DNA topoisomerase VI is essential for endoreduplication in arabidopsis. Curr Biol 12:1782–1786

Sym M, Engebrecht J, Roeder GS (1993) ZIP1 is a synaptonemal complex protein required for meiotic chromosome synapsis. Cell 72:365–378

Tang Y, Yin Z, Zeng Y, Zhang Q, Chen L, He Y, Lu P, Ye D, Zhang X (2017) MTOPVIB interacts with AtPRD1 and plays important roles in formation of meiotic DNA double-strand breaks in Arabidopsis. Sci Rep 7:10007

Villeneuve AM, Hillers KJ (2001) Whence meiosis? Cell 106:647–650

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 351:939–943

Vrielynck N, Schneider K, Rodriguez M, Sims J, Chambon A, Hurel A, De Muyt A, Ronceret A, Krsicka O, Mézard C et al (2021) Conservation and divergence of meiotic DNA double strand break forming mechanisms in Arabidopsis thaliana. Nucleic Acids Res 49:9821–9835

West AM, Rosenberg SC, Ur SN, Lehmer MK, Ye Q, Hagemann G, Caballero I, Uson I, MacQueen AJ, Herzog F et al (2019) A conserved filamentous assembly underlies the structure of the meiotic chromosome axis. eLife. https://doi.org/10.7554/eLife.40372

Wijeratne AJ, Chen C, Zhang W, Timofejeva L, Ma H (2006) The Arabidopsis thaliana PARTING DANCERS gene encoding a novel protein is required for normal meiotic homologous recombination. Mol Biol Cell 17:1331–1343

Yadav VK, Claeys Bouuaert C (2021) Mechanism and control of meiotic DNA double-strand break formation in S. cerevisiae. Front Cell Dev Biol 9:642737–642737

Yin Y, Cheong H, Friedrichsen D, Zhao Y, Hu J, Mora-Garcia S, Chory J (2002) A crucial role for the putative Arabidopsis

topoisomerase VI in plant growth and development. Proc Natl Acad Sci 99:10191–10196

Zhang C, Song Y, Cheng ZH, Wang YX, Zhu J, Ma H, Xu L, Yang ZN (2012) The Arabidopsis thaliana DSB formation (AtDFO) gene is required for meiotic double-strand break formation. Plant J 72:271–281

Zhang L, Kong H, Ma H, Yang J (2018) Phylogenomic detection and functional prediction of genes potentially important for plant meiosis. Gene 643:83–97

General discussion

This thesis comprises three distinct studies: 1) investigations into the epigenetic regulation of holocentromeres, 2) evolutionary analysis of inverted meiosis, and 3) analysis of the conservation of plant meiotic machinery. Each of these three studies was the first of its kind and provided entirely novel insights. The first and third chapter were published successfully. The second chapter provided initial clues about the evolution of inverted meiosis. We are currently working on developing a highly efficient genome editing system for *Rhynchospora*, which could be used to further explore the evolution of inverted meiosis.

Holocentric chromosomes of *Rhynchospora*

I found that most genomic compartments, like the distribution of genes, centromeric repeats, and transposable elements (TEs) were uniform over the whole length of chromosomes in the holocentric species *Rhynchospora*. Transcriptional activity and methylation were also uniform. In contrast, in the monocentric plant *Juncus effusus*, repeats and TEs were concentrated towards the potential centromeric region. Genes were present along the chromosome arms. Transcriptional activity was correlated with the presence of genes along the chromosome arms; meanwhile, the overall methylation status was at its peak at the potential centromeric region. Immunofluorescence studies revealed that, in the case of the monocentric species *J. effusus*, H3K4me3(euchromatin mark) has a dispersed signal, whereas the signal for the heterochromatin mark H3K9me2 was concentrated in the chromocenters, which represents the centromeric and peri-centromeric domain. In *Rhynchospora*, the signals for both the eu- and hetero-chromatin marks were dispersed all along the nuclei confirming its holocentric nature. Based on the ChIP-seq results, we discovered that the centromeric regions are enriched for the centromeric protein CENH3. CENH3 domains are spaced at regular intervals (300–500kb), which may help with the folding of chromatin to achieve the line-like holocentromeric pattern observed during metaphase. CENH3 enrichment was found on the *Tyba* arrays closest to the telomeres, confirming the presence of active holocentromeres from telomere to telomere.

Epigenetic regulation of repeat-based centromeres is conserved

R. pubera was the first holocentric species whose holocentromeres were reported to be composed of the satellite repeat element *Tyba* (Marques et al., 2015). In our analysis, we found that holocentromeres in *R. pubera* are 20–25Kb average in size and that the species lacks larger pericentromeric domains, which are present in the case of monocentric organisms. High CpG methylation, depletion of H3K4me3 and low H3K9me3 are the epigenetic features that mark holocentromeres in *R. pubera*. Most interestingly, there is relatively higher enrichment of H3K9me2 and CHG methylation at the holocentromeric borders than the core region (Hofstatter et al., 2022). *A. thaliana* is a monocentric organism, which is very distantly related to *R. pubera*. Centromeres in A. thaliana are made up of 178bp alpha-satellite repeats (CEN180). The size of centromere 1 alone is predicted to be approximately 2.26Mb of core centromere flanked by a 2.12Mb-sized pericentromeric region (Haupt et al., 2001). Also, in this case, H3K9me2 and CHG methylation are enriched in the pericentromeric region when compared to the core (Naish et al., 2021). This large discrepancy in centromere size gives an idea of the profound genetic differences between the centromeres of *R. pubera* and *A. thaliana*. Defining centromeric borders is essential for the cell to properly mark its chromatin and thereby differentially regulate processes like transcription and recombination. This compelling evidence of H3K9me2 and CHG methylation enrichment at the centromeric borders suggests that the epigenetic regulation of repeat-based holo- and monocentromeres may be evolutionarily conserved.

Do holocentromeres facilitate karyotype evolution?

It has long been hypothesized that holocentromeres can facilitate chromosomal fusions, fissions and rearrangements, thereby facilitating karyotype evolution. Holocentric lineages have huge differences in their chromosomal number. For example, the genus *Rhynchospora* consists of species with chromosome numbers varying from 2n = 4 to 61 (Burchardt et al., 2020). Similarly, the genus *Carex* belonging to the same holocentric family Cyperaceae has a chromosome number ranging from 2n = 10–132 (Márquez-Corro et al., 2021). Thus, *Rhynchospora* is an excellent model in which we could test this hypothesis. By comparing the genome synteny of the three *Rhynchospora* species sequenced, i.e., *R. pubera* (n=5), *R. breviuscula* (n=5) and *R. tenuis* (n=2), we found that *R. pubera* is in fact a hidden octoploid that underwent a complex chain of chromosomal fusions. In this case, two rounds of whole genome duplication followed by end-to-end

chromosome fusions restored the ancestral chromosome number, x = 5. This was possible since the sister species *R. breviuscula* showed a diploid genome with the ancestral chromosome number x = 5. The high conservation of the synteny between these two species allowed us to detect that 10 out of 15 fusion points had the centromeric repeat element *Tyba* just at the junctions of the end-to-end fusions. Similarly, in the diploid species *R. tenuis* with chromosome number 2n = 4, we show that this very low chromosome number resulted from three end-to-end fusions (Hofstatter et al., 2022). This supports the hypothesis that holocentromeres could facilitate karyotype evolution.

Evolutionary implications of holocentricity

Holocentric organisms are present across several taxonomic groups including plants like *Luzula*, *Rhynchospora*, nematodes like the well-studied *Caenorhabditis elegans*, and several insect species. They have been reported to have evolved independently at least 13 times (Escudero et al., 2016; Melters et al., 2012), prompting the question of whether they are evolutionarily advantageous? Holocentric organisms are proposed to have advantages over monocentric organisms in cases of chromosomal breakage under stressful circumstances (Zedek and Bureš, 2018). In monocentrics, the spindle fibers attach to single centromeric regions. When chromosomes break, parts of chromosomes will not be segregated to the next generation which could be lethal. In holocentric organisms, however, the spindle fibers attach all along the chromosomes, resulting in faithful segregation even in the presence of chromosomal breaks.

When thinking about holocentromeres, one obvious question is whether numerous centromeres means that cells must produce an increased number of proteins involved in the regulation of centromeres, centromeric proteins and kinetochore proteins?". Numerous holocentromeres does not necessarily mean that such centromeres are larger than in monocentric organisms. However, limited studies are available on holocentric organisms and they remain uncharacterized. Despite the relative scarcity of research on holocentrics compared to monocentrics, there is a growing area of interest in understanding the genetic mechanisms and evolutionary implications of holocentric organisms.

Unanswered questions regarding holocentromere biology

Some *Rhynchospora* species lack the satellite repeat, *Tyba* (Costa et al., 2023; Ribeiro et al., 2017). It will be interesting to determine the genetic and epigenetic

composition of these holocentromeres: Expanding the focus to these *Rhynchospora* species can help us understand the following questions. 1) Are there satellite repeats other than *Tyba* that constitute the holocentromeres in the *Rhynchospora* genus? 2) Do these *Rhynchospora* holocentromeres lack any particular satellite repeats and thus have no genetic definition? 3) If there are *Rhynchsopora* species without *Tyba*, does that mean the holocentric transition in this genus is not associated with Tyba?

Sister centromeric cohesion protection might be dispensable during the inverted meiosis of *Rhynchospora*

In my investigations into the mechanism of inverted meiosis, I found that the meiotic cohesin REC8 has a conserved line-like pattern during early prophase I. Progressing towards metaphase I, most of the signals are lost and only residual signal was observed on the chromosomes. However, on comparing the REC8 signals with *Tyba* (cenDNA) signals, it is evident that sister centromere cohesion is lost. This is the major adaptation facilitating the evolution of inverted meiosis, and this marks the major difference between canonical and inverted meiosis. On the other hand, one ortholog for SGO was detected in *Rhynchospora*. However, based on its localization dynamics, this protein doesn't appear to protect sister centromere cohesion and may have other evolutionarily conserved roles.

Rhynchospora – an optimal system to characterize the non-canonical functions of SGO

Proteins are essential for biological processes and a single protein can have multiple roles and functions. Proteins are initially characterized for a particular function but as we continue to study, additional roles are discovered. This is not surprising because proteins can interact with a wide variety of molecules and can take part in many cellular processes. But often, due to insufficient scientific studies, a protein's primary function may be assumed to be the initially discovered role. SGO is one such perfect example. It is primarily known for its role in centromeric cohesion protection, and has been christened the "guardian spirit at the centromere", from which its nomenclature was derived (Watanabe, 2005). Later studies have reported non-centromeric roles for SGO as I discussed in the introduction. Since *Rhynchospora* has evolved to segregate its sister chromatids during meiosis I, it does not necessarily need SGO for centromeric cohesion protection during meiosis I, which was also evident from its cytological localization in this study. This makes *Rhynchospora* a perfect model to explore the potential non-centromeric

roles of SGO. This work is currently progressing and in future, hopefully will be able to get insights, with the development of an optimal functional characterization method.

Insights into the plant meiotic machinery

My comparative analysis of the meiotic machinery across different plant lineages revealed that most meiotic pathways are conserved in all the lineages. The strand invasion pathway was found to be the most conserved pathway among the tested species. Some proteins, like ASY4, PHS1, PRD2 and PRD3 belonging to the chromosome axis and DSB formation pathways were not detected in distant algal lineages, suggesting that they are only minimally conserved. DFO and HEIP1 proteins were not previously reported outside plants. Through phylogenetic analysis, I could trace the possible evolutionary origin of these two proteins. DFO evolved in the common ancestor of vascular plants and HEIP1 in the common ancestor of Streptophyta. SPO11-1 and SPO11-2 form a dimer in *A. thaliana*. To find the duplication history of SPO11, a phylogenetic tree was constructed by including non-plant lineages. Apparently, SPO11 duplication has occurred in the common ancestor of eukaryotes, and some lineages subsequently lost this duplication. Likewise, charophytes and chlorophytes, lost SPO11-1 secondarily.

How conserved are meiotic proteins?

Meiosis is a highly conserved process ancestral to all eukaryotes. By analyzing proteins from different meiotic pathways in this study, we could see proteins from some pathways are more conserved than the others. Some meiotic proteins were structurally conserved but divergent at the sequence level, making it difficult to detect them using the computational tools that we employed. Functional domains are the most conserved regions of the proteins. This correlation was also evident in this study. Proteins with conserved, characterized domains showed higher similarity than the ones without functional domains. Multiple meiotic proteins function as complexes in different pathways. It was thus interesting to note that proteins from the same complex had different degrees of conservation. Considering how conserved meiosis is, this study sheds light into conservation of meiotic proteins from the sequence point of view.

Future perspectives for studying meiotic proteins

In this study, I looked for *A. thaliana* meiotic proteins in different plant lineages. However, failure to detect a protein does not necessarily mean that the protein is absent

in a given lineage. The availability of high-quality genomic data, genomic data from meiotic tissues and improved computational tools can be a limiting factor for the analysis. The number of species with available genomic data for each plant lineage was also limited. With the advancement in sequencing technology, in future one may expect to have access to improved and high number of genomic data. Expanding such studies by including a greater number of species in the future can provide more insights about the evolution of plant meiotic machinery. As discussed earlier, proteins are more conserved structurally than at the sequence level. Employing algorithms which explore protein structures, like alpha fold, can help to detect distant homologs which has poor sequence similarity.

References

Akiyoshi, B., and Gull, K. (2014). Discovery of Unconventional Kinetochores in Kinetoplastids. Cell *156*, 1247-1258.

Altemose, N., Logsdon, G.A., Bzikadze, A.V., Sidhwani, P., Langley, S.A., Caldas, G.V., Hoyt, S.J., Uralsky, L., Ryabov, F.D., Shew, C.J., *et al.* (2022). Complete genomic and epigenetic maps of human centromeres. Science *376*, eabl4178.

Altschul, S.F., Madden, T.L., Schäffer, A.A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D.J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic acids research *25*, 3389-3402.

An, X.J., Deng, Z.Y., and Wang, T. (2011). OsSpo11-4, a rice homologue of the archaeal TopVIA protein, mediates double-strand DNA cleavage and interacts with OsTopVIB. PloS one *6*, e20327.

Bai, X., Peirson, B.N., Dong, F., Xue, C., and Makaroff, C.A. (1999). Isolation and Characterization of SYN1, a RAD21-like Gene Essential for Meiosis in Arabidopsis. The Plant Cell *11*, 417-430.

Battaglia, E., and Boyes, J. (1955). Post-Reductional Meiosis: Its Mechanism and Causes: (with 10 figures). Caryologia *8*, 87-134.

Baum, M., Ngan, V.K., and Clarke, L. (1994). The centromeric K-type repeat and the central core are together sufficient to establish a functional Schizosaccharomyces pombe centromere. Molecular Biology of the Cell *5*, 747-761.

Bhagwat, M., and Aravind, L. (2008). Psi-blast tutorial. Comparative genomics, 177-186.

Bohr, T., Nelson, C.R., Giacopazzi, S., Lamelza, P., and Bhalla, N. (2018). Shugoshin Is Essential for Meiotic Prophase Checkpoints in C. elegans. Curr Biol *28*, 3199-3211 e3193.

Braz, G.T., Yu, F., do Vale Martins, L., and Jiang, J. (2020). Fluorescent In Situ Hybridization Using Oligonucleotide-Based Probes. In In Situ Hybridization Protocols, B.S. Nielsen, and J. Jones, eds. (New York, NY: Springer US), pp. 71-83.

Burchardt, P., Buddenhagen, C.E., Gaeta, M.L., Souza, M.D., Marques, A., and Vanzela, A.L. (2020). Holocentric karyotype evolution in Rhynchospora is marked by intense numerical, structural, and genome size changes. Frontiers in Plant Science *11*, 536507.

Cabral, G., Marques, A., Schubert, V., Pedrosa-Harand, A., and Schlogelhofer, P. (2014). Chiasmatic and achiasmatic inverted meiosis of plants with holocentric chromosomes. Nat Commun *5*, 5070.

Cai, X., Dong, F., Edelmann, R.E., and Makaroff, C.A. (2003). The Arabidopsis SYN1 cohesin protein is required for sister chromatid arm cohesion and homologous chromosome pairing. J Cell Sci *116*, 2999-3007.

Cam, H.P., Sugiyama, T., Chen, E.S., Chen, X., FitzGerald, P.C., and Grewal, S.I. (2005). Comprehensive analysis of heterochromatin-and RNAi-mediated epigenetic control of the fission yeast genome. Nature genetics *37*, 809-819.

Capella-Gutierrez, S., Silla-Martinez, J.M., and Gabaldon, T. (2009). trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. Bioinformatics *25*, 1972-1973.

Chandra, H.S. (1962). Inverse meiosis in triploid females of the mealy bug, Planococcus citri. Genetics *47*, 1441-1454.

Chelysheva, L., Diallo, S., Vezon, D., Gendrot, G., Vrielynck, N., Belcram, K., Rocques, N., Marquez-Lema, A., Bhatt, A.M., Horlow, C.*, et al.* (2005). AtREC8 and AtSCC3 are essential to the monopolar orientation of the kinetochores during meiosis. J Cell Sci *118*, 4621-4632.

Clarke, L., and Carbon, J. (1985). THE STRUCTURE AND FUNCTION OF YEAST CENTROMERES. Annual Review of Genetics *19*, 29-55.

Corlett, R.T. (2016). Plant diversity in a changing world: status, trends, and conservation needs. Plant diversity *38*, 10-16.

Costa, L., Marques, A., Buddenhagen, C.E., Pedrosa-Harand, A., and Souza, G. (2023). Investigating the diversification of holocentromeric satellite DNA Tyba in Rhynchospora (Cyperaceae). Annals of Botany, mcad036.

Cromer, L., Jolivet, S., Horlow, C., Chelysheva, L., Heyman, J., De Jaeger, G., Koncz, C., De Veylder, L., and Mercier, R. (2013). Centromeric Cohesion Is Protected Twice at Meiosis, by SHUGOSHINs at Anaphase I and by PATRONUS at Interkinesis. Current Biology *23*, 2090-2099.

Drinnenberg, I.A., deYoung, D., Henikoff, S., and Malik, H.S. (2014). Recurrent loss of CenH3 is associated with independent transitions to holocentricity in insects. eLife *3*, e03676.

Eddy, S. (1992). HMMER user's guide. Department of Genetics, Washington University School of Medicine *2*, 13.

Escudero, M., Márquez-Corro, J.I., and Hipp, A.L. (2016). The phylogenetic origins and evolutionary history of holocentric chromosomes. Systematic Botany *41*, 580-585.

Foley, E.A., and Kapoor, T.M. (2013). Microtubule attachment and spindle assembly checkpoint signalling at the kinetochore. Nature reviews Molecular cell biology *14*, 25-37.
Gassmann, R., Rechtsteiner, A., Yuen, K.W., Muroyama, A., Egelhofer, T., Gaydos, L., Barron, F., Maddox, P., Essex, A., and Monen, J. (2012). An inverse relationship to germline transcription defines centromeric chromatin in C. elegans. Nature *484*, 534-537.
Gawryluk, R.M., Tikhonenkov, D.V., Hehenberger, E., Husnik, F., Mylnikov, A.P., and Keeling, P.J. (2019). Non-photosynthetic predators are sister to red algae. Nature *572*, 240-243.
Goday, C., and Pimpinelli, S. (1989). Centromere organization in meiotic chromosomes of Parascaris univalens. Chromosoma *98*, 160-166.
Goutte-Gattat, D., Shuaib, M., Ouararhni, K., Gautier, T., Skoufias, D.A., Hamiche, A., and Dimitrov, S. (2013). Phosphorylation of the CENP-A amino-terminus in mitotic centromeric chromatin is required for kinetochore function. Proc Natl Acad Sci U S A *110*, 8579-8584.
Gregan, J., Spirek, M., and Rumpf, C. (2008). Solving the shugoshin puzzle. Trends in Genetics *24*, 205-207.
Harrington, J.J., Bokkelen, G.V., Mays, R.W., Gustashaw, K., and Willard, H.F. (1997). Formation of de novo centromeres and construction of first-generation human artificial microchromosomes. Nature Genetics *15*, 345-355.
Hartung, F., and Puchta, H. (2000). Molecular characterisation of two paralogous SPO11 homologues in Arabidopsis thaliana. Nucleic acids research *28*, 1548-1554.
Hartung, F., and Puchta, H. (2001). Molecular characterization of homologues of both subunits A (SPO11) and B of the archaebacterial topoisomerase 6 in plants. Gene *271*, 81-86.
Hauf, S., and Watanabe, Y. (2004). Kinetochore orientation in mitosis and meiosis. Cell *119*, 317-327.
Haupt, W., Fischer, T.C., Winderl, S., Fransz, P., and Torres-Ruiz, R.A. (2001). The centromere1 (CEN1) region of Arabidopsis thaliana: architecture and functional impact of chromatin. The Plant Journal *27*, 285-296.
Heckmann, S., Jankowska, M., Schubert, V., Kumke, K., Ma, W., and Houben, A. (2014a). Alternative meiotic chromatid segregation in the holocentric plant Luzula elegans. Nat Commun *5*, 4979.

Heckmann, S., Macas, J., Kumke, K., Fuchs, J., Schubert, V., Ma, L., Novak, P., Neumann, P., Taudien, S., Platzer, M., *et al.* (2013). The holocentric species Luzula elegans shows interplay between centromere and large-scale genome organization. Plant J *73*, 555-565.

Heckmann, S., Schubert, V., and Houben, A. (2014b). Holocentric plant meiosis: first sisters, then homologues. Cell Cycle *13*, 3623-3624.

Hofstatter, P.G., Thangavel, G., Castellani, M., and Marques, A. (2021). Meiosis progression and recombination in holocentric plants: What is known? Frontiers in Plant Science *12*, 658296.

Hofstatter, P.G., Thangavel, G., Lux, T., Neumann, P., Vondrak, T., Novak, P., Zhang, M., Costa, L., Castellani, M., and Scott, A. (2022). Repeat-based holocentromeres influence genome architecture and karyotype evolution. Cell *185*, 3153-3168. e3118.

Jain, M., Tyagi, A.K., and Khurana, J.P. (2006). Overexpression of putative topoisomerase 6 genes from rice confers stress tolerance in transgenic Arabidopsis plants. The FEBS journal *273*, 5245-5260.

Jenuwein, T., and Allis, C.D. (2001). Translating the histone code. Science *293*, 1074-1080.

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

Katoh, K., and Standley, D.M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol *30*, 772-780.

Kitajima, T.S., Kawashima, S.A., and Watanabe, Y. (2004). The conserved kinetochore protein shugoshin protects centromeric cohesion during meiosis. Nature *427*, 510-517.

Klein, F., Mahr, P., Galova, M., Buonomo, S.B., Michaelis, C., Nairz, K., and Nasmyth, K. (1999). A central role for cohesins in sister chromatid cohesion, formation of axial elements, and recombination during yeast meiosis. Cell *98*, 91-103.

Kumar, N., Galli, M., Ordon, J., Stuttmann, J., Kogel, K.H., and Imani, J. (2018). Further analysis of barley MORC1 using a highly efficient RNA-guided Cas9 gene-editing system. Plant Biotechnol J *16*, 1892-1903.

Kursel, L.E., and Malik, H.S. (2016). Centromeres. Current Biology *26*, R487-R490.

Kurze, A., Michie, K.A., Dixon, S.E., Mishra, A., Itoh, T., Khalid, S., Strmecki, L., Shirahige, K., Haering, C.H., Lowe, J., *et al.* (2011). A positively charged channel within the Smc1/Smc3 hinge required for sister chromatid cohesion. EMBO J *30*, 364-378.

Lin, Y., Larson, K.L., Dorer, R., and Smith, G.R. (1992). Meiotically induced rec7 and rec8 genes of Schizosaccharomyces pombe. Genetics *132*, 75-85.

Liu, D., Song, A.T., Qi, X., van Vliet, P.P., Xiao, J., Xiong, F., Andelfinger, G., and Nattel, S. (2021). Cohesin-protein Shugoshin-1 controls cardiac automaticity via HCN4 pacemaker channel. Nature Communications *12*, 2551.

Liu, H., Rankin, S., and Yu, H. (2013). Phosphorylation-enabled binding of SGO1–PP2A to cohesin protects sororin and centromeric cohesion during mitosis. Nature cell biology *15*, 40-49.

Llano, E., Gomez, R., Gutierrez-Caballero, C., Herran, Y., Sanchez-Martin, M., Vazquez-Quinones, L., Hernandez, T., de Alava, E., Cuadrado, A., Barbero, J.L.*, et al.* (2008). Shugoshin-2 is essential for the completion of meiosis but not for mitotic cell division in mice. Genes Dev *22*, 2400-2413.

Logsdon, G.A., Vollger, M.R., Hsieh, P., Mao, Y., Liskovykh, M.A., Koren, S., Nurk, S., Mercuri, L., Dishuck, P.C., Rhie, A.*, et al.* (2021). The structure, function and evolution of a complete human chromosome 8. Nature *593*, 101-107.

Lui, D.Y., and Colaiácovo, M.P. (2013). Meiotic development in Caenorhabditis elegans. Adv Exp Med Biol *757*, 133-170.

Malheiros, N., Castro, D.d., and Camara, A. (1947). Chromosomas sem centromero localizado. O caso da Luzula purpurea Link. Agronomia lusitana *9*, 51-71.

Malik, S.-B., Ramesh, M.A., Hulstrand, A.M., and Logsdon Jr, J.M. (2007). Protist homologs of the meiotic Spo11 gene and topoisomerase VI reveal an evolutionary history of gene duplication and lineage-specific loss. Molecular biology and evolution *24*, 2827-2841.

Marques, A., Ribeiro, T., Neumann, P., Macas, J., Novak, P., Schubert, V., Pellino, M., Fuchs, J., Ma, W., Kuhlmann, M.*, et al.* (2015). Holocentromeres in Rhynchospora are associated with genome-wide centromere-specific repeat arrays interspersed among euchromatin. Proc Natl Acad Sci U S A *112*, 13633-13638.

Marques, A., Schubert, V., Houben, A., and Pedrosa-Harand, A. (2016). Restructuring of Holocentric Centromeres During Meiosis in the Plant Rhynchospora pubera. Genetics *204*, 555-568.

Márquez-Corro, J.I., Martín-Bravo, S., Jiménez-Mejías, P., Hipp, A.L., Spalink, D., Naczi, R.F., Roalson, E.H., Luceño, M., and Escudero, M. (2021). Macroevolutionary insights into sedges (Carex: Cyperaceae): The effects of rapid chromosome number evolution on lineage diversification. Journal of Systematics and Evolution *59*, 776-790.

Marshall, O.J., Chueh, A.C., Wong, L.H., and Choo, K.H.A. (2008). Neocentromeres: New Insights into Centromere Structure, Disease Development, and Karyotype Evolution. The American Journal of Human Genetics *82*, 261-282.

Marston, A.L. (2015). Shugoshins: tension-sensitive pericentromeric adaptors safeguarding chromosome segregation. Molecular and cellular biology *35*, 634-648.

McGuinness, B.E., Hirota, T., Kudo, N.R., Peters, J.M., and Nasmyth, K. (2005). Shugoshin prevents dissociation of cohesin from centromeres during mitosis in vertebrate cells. PLoS Biol *3*, e86.

McKinley, K.L., and Cheeseman, I.M. (2016). The molecular basis for centromere identity and function. Nature reviews Molecular cell biology *17*, 16-29.

Melters, D.P., Paliulis, L.V., Korf, I.F., and Chan, S.W. (2012). Holocentric chromosomes: convergent evolution, meiotic adaptations, and genomic analysis. Chromosome Res *20*, 579-593.

Mercier, R., and Grelon, M. (2008). Meiosis in plants: ten years of gene discovery. Cytogenetic and genome research *120*, 281-290.

Miga, K.H., Koren, S., Rhie, A., Vollger, M.R., Gershman, A., Bzikadze, A., Brooks, S., Howe, E., Porubsky, D., and Logsdon, G.A. (2020). Telomere-to-telomere assembly of a complete human X chromosome. Nature *585*, 79-84.

Murakami, S., Matsumoto, T., Niwa, O., and Yanagida, M. (1991). Structure of the fission yeast centromere cen3: Direct analysis of the reiterated inverted region. Chromosoma *101*, 214-221.

Naish, M., Alonge, M., Wlodzimierz, P., Tock, A.J., Abramson, B.W., Schmucker, A., Mandakova, T., Jamge, B., Lambing, C., Kuo, P.*, et al.* (2021). The genetic and epigenetic landscape of the Arabidopsis centromeres. Science *374*, eabi7489.

Nasmyth, K., and Haering, C.H. (2009). Cohesin: its roles and mechanisms. Annu Rev Genet *43*, 525-558.

Nayar, S., Sharma, R., Tyagi, A.K., and Kapoor, S. (2013). Functional delineation of rice MADS29 reveals its role in embryo and endosperm development by affecting hormone homeostasis. J Exp Bot *64*, 4239-4253.

Nguyen, L.T., Schmidt, H.A., von Haeseler, A., and Minh, B.Q. (2015). IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol Biol Evol *32*, 268-274.

Nokkala, S., Laukkanen, A., and Nokkala, C. (2002). Mitotic and meiotic chromosomes in Somatochlora metallica (Cordulidae, Odonata). The absence of localized centromeres and inverted meiosis. Hereditas *136*, 7-12.

Not, F., Valentin, K., Romari, K., Lovejoy, C., Massana, R., Töbe, K., Vaulot, D., and Medlin, L.K. (2007). Picobiliphytes: a marine picoplanktonic algal group with unknown affinities to other eukaryotes. science *315*, 253-255.

Nurk, S., Koren, S., Rhie, A., Rautiainen, M., Bzikadze, A.V., Mikheenko, A., Vollger, M.R., Altemose, N., Uralsky, L., Gershman, A.*, et al.* (2022). The complete sequence of a human genome. Science *376*, 44-53.

Ordon, J., Gantner, J., Kemna, J., Schwalgun, L., Reschke, M., Streubel, J., Boch, J., and Stuttmann, J. (2017). Generation of chromosomal deletions in dicotyledonous plants employing a user-friendly genome editing toolkit. Plant J *89*, 155-168.

Orr-Weaver, T.L. (1999). The Ties that Bind: Localization of the Sister-Chromatid Cohesin Complex on Yeast Chromosomes. Cell *99*, 1-4.

Palmer, J.D., Soltis, D.E., and Chase, M.W. (2004). The plant tree of life: an overview and some points of view. American journal of botany *91*, 1437-1445.

Pasierbek, P., Jantsch, M., Melcher, M., Schleiffer, A., Schweizer, D., and Loidl, J. (2001). A Caenorhabditis elegans cohesion protein with functions in meiotic chromosome pairing and disjunction. Genes Dev *15*, 1349-1360.

Pazy, B., and Plitmann, U. (1991). Unusual chromosome separation in meiosis of Cuscuta L. Genome *34*, 533-536.

Pe´rez, R.n., Panzera, F., Page, J.s., Suja, J.A., and Rufas, J.S. (1997). Meiotic behaviour of holocentric chromosomes: orientation and segregation of autosomes in Triatoma infestans (Heteroptera). Chromosome Research *5*, 47-56.

Puttick, M.N., Morris, J.L., Williams, T.A., Cox, C.J., Edwards, D., Kenrick, P., Pressel, S., Wellman, C.H., Schneider, H., and Pisani, D. (2018). The interrelationships of land plants and the nature of the ancestral embryophyte. Current Biology *28*, 733-745. e732.

Régnier, V., Vagnarelli, P., Fukagawa, T., Zerjal, T., Burns, E., Trouche, D., Earnshaw, W., and Brown, W. (2005). CENP-A Is Required for Accurate Chromosome Segregation and Sustained Kinetochore Association of BubR1. Molecular and Cellular Biology *25*, 3967-3981.

Ribeiro, T., Marques, A., Novák, P., Schubert, V., Vanzela, A.L., Macas, J., Houben, A., and Pedrosa-Harand, A. (2017). Centromeric and non-centromeric satellite DNA organisation differs in holocentric Rhynchospora species. Chromosoma *126*, 325-335.

Robert, T., Nore, A., Brun, C., Maffre, C., Crimi, B., Guichard, V., Bourbon, H.-M., and De Massy, B. (2016). The TopoVIB-Like protein family is required for meiotic DNA double-strand break formation. Science *351*, 943-949.

Salic, A., Waters, J.C., and Mitchison, T.J. (2004). Vertebrate shugoshin links sister centromere cohesion and kinetochore microtubule stability in mitosis. Cell *118*, 567-578.

Scelfo, A., and Fachinetti, D. (2019). Keeping the centromere under control: a promising role for DNA methylation. Cells *8*, 912.

Senaratne, A.P., Muller, H., Fryer, K.A., Kawamoto, M., Katsuma, S., and Drinnenberg, I.A. (2021). Formation of the CenH3-deficient holocentromere in Lepidoptera avoids active chromatin. Current Biology *31*, 173-181. e177.

Steiner, F.A., and Henikoff, S. (2014). Holocentromeres are dispersed point centromeres localized at transcription factor hotspots. Elife *3*, e02025.

Sullivan, B.A., and Karpen, G.H. (2004). Centromeric chromatin exhibits a histone modification pattern that is distinct from both euchromatin and heterochromatin. Nature structural & molecular biology *11*, 1076-1083.

Sullivan, K.F., Hechenberger, M., and Masri, K. (1994). Human CENP-A contains a histone H3 related histone fold domain that is required for targeting to the centromere. J Cell Biol *127*, 581-592.

Talbert, P.B., and Henikoff, S. (2022). The genetics and epigenetics of satellite centromeres. Genome Res *32*, 608-615.

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

Tanaka, T.U. (2010). Kinetochore–microtubule interactions: steps towards bi-orientation. The EMBO journal *29*, 4070-4082.

Thangavel, G., Hofstatter, P.G., Mercier, R., and Marques, A. (2023). Tracing the evolution of the plant meiotic molecular machinery. Plant Reproduction, 1-23.

Uhlmann, F., and Nasmyth, K. (1998). Cohesion between sister chromatids must be established during DNA replication. Curr Biol *8*, 1095-1101.

van Heemst, D., and Heyting, C. (2000). Sister chromatid cohesion and recombination in meiosis. Chromosoma *109*, 10-26.

van Hooff, J.J., Tromer, E., van Wijk, L.M., Snel, B., and Kops, G.J. (2017). Evolutionary dynamics of the kinetochore network in eukaryotes as revealed by comparative genomics. EMBO Rep *18*, 1559-1571.

Viera, A., Page, J., and Rufas, J. (2009). Inverted meiosis: the true bugs as a model to study. Meiosis *5*, 137-156.

Vrielynck, N., Chambon, A., Vezon, D., Pereira, L., Chelysheva, L., De Muyt, A., Mézard, C., Mayer, C., and Grelon, M. (2016). A DNA topoisomerase VI–like complex initiates meiotic recombination. Science *351*, 939-943.

Wahl, H.A. (1940). Chromosome numbers and meiosis in the genus Carex. American Journal of Botany, 458-470.

Wang, X., Yang, Y., Duan, Q., Jiang, N., Huang, Y., Darzynkiewicz, Z., and Dai, W. (2008). sSgo1, a major splice variant of Sgo1, functions in centriole cohesion where it is regulated by Plk1. Developmental cell *14*, 331-341.

Watanabe, Y. (2005). Shugoshin: guardian spirit at the centromere. Curr Opin Cell Biol *17*, 590-595.

Watanabe, Y., and Kitajima, T.S. (2005). Shugoshin protects cohesin complexes at centromeres. Philosophical Transactions of the Royal Society B: Biological Sciences *360*, 515-521.

Watanabe, Y., and Nurse, P. (1999). Cohesin Rec8 is required for reductional chromosome segregation at meiosis. Nature *400*, 461-464.

Wolfgruber, T.K., Nakashima, M.M., Schneider, K.L., Sharma, A., Xie, Z., Albert, P.S., Xu, R., Bilinski, P., Dawe, R.K., and Ross-Ibarra, J. (2016). High quality maize centromere 10 sequence reveals evidence of frequent recombination events. Frontiers in plant science *7*, 308.

Yoda, K., Ando, S., Morishita, S., Houmura, K., Hashimoto, K., Takeyasu, K., and Okazaki, T. (2000). Human centromere protein A (CENP-A) can replace histone H3 in nucleosome reconstitution in vitro. Proceedings of the National Academy of Sciences *97*, 7266-7271.

Zedek, F., and Bureš, P. (2018). Holocentric chromosomes: from tolerance to fragmentation to colonization of the land. Annals of botany *121*, 9-16.

Appendix

Media used for Plant Transformation

1. *Rhynchospora* Callus Induction Medium (RCI) *(to propagate calli)*

Part I: 2X Gelrite

Components	Stock concentration	Final concentration	Amount for 1 liter
Gelrite	-	4 g/l	4 g

Dissolve the components in 500 ml of water in a 1-liter bottle and sterilize by autoclaving.

Part II: 2X RCI

Components	Stock concentration	Final concentration	Amount for 1 liter
MS salts (M0221)	-	4.4 g/l	4.4 g
Sucrose	-	30 g/l	30 g
Casein hydrolysate	-	1 g/l	1 g

Adjust the pH to 5.8 – 5.6 with 1M KOH, make up the volume to ~490 ml and add the following components under sterile hood:

Components	Stock concentration	Final concentration	Amount for 1 liter
100X Vitamins RCI	100X	1X	10 ml
$CuSO_4.5H_2O$	1.25 mg/ml (5mM)	1.25 µg/ml (5 µM)	1 ml
Dicamba	2.5 mg/ml	5 µg/ml	2 ml

Filter sterilize the solution.

Part III: RCI

Warm the 2x gelrite and the 2x RCI at 50-60°C in a water bath, mix together and pour into petri dishes.

2. *Rhynchospora* Co-cultivation Medium (RAS-Co) *(To co-cultivate embryos with Agrobacterium culture and t-DNA transfer)*

Part I: 2X Gelrite

Components	Stock concentration	Final concentration	Amount for 500ml
Gelrite	-	4 g/l	2 g

Dissolve the components in 250 ml of water in a 500ml bottle and sterilize by autoclaving.

Part II: 2x Ras-Co

Components	Stock concentration	Final concentration	Amount for 500 ml
MS salts (M0221)	-	4.4 g/l	2.2 g
Sucrose	-	20 g/l	10 g
Glucose		10 g/l	5 g
Casein hydrolysate	-	1 g/l	0.5 g
100X Vitamins RCI	100X	1X	5 ml
L-Cysteine	8 g/l (1000X)	800 mg/l (1X)	0.5 ml
Dicamba	2.5 mg/ml	5 ug/ml	50 ul
Acetosyringone	1 M	250 µM	10 µl

Adjust the pH to 5.8 – 5.6 with 1M KOH, make up the volume to 250ml and filter sterilize.

Part III: Ras-Co

Warm the 2x gelrite and the 2x Ras-Co at 50-60°C in a water bath, mix together and pour into petri dishes. After the medium has solidified, place a sterile filter paper on top of it to avoid the diffusion of *Agrobacterium* into the medium thereby preventing it's growth.

3. *Rhynchospora* Callus Induction Medium + Hygromycin + Timentin (RCI + Hyg + Tim) *(To select the propagating calli)*

Part I: 2X Gelrite

Components	Stock concentration	Final concentration	Amount for 1 liter
Gelrite	-	4 g/l	4 g

Dissolve the components in 500 ml of water in a 1-liter bottle and sterilize by autoclaving.

Part II: 2X RCI

Components	Stock concentration	Final concentration	Amount for 1 liter
MS salts (M0221)	-	4.4 g/l	4.4 g
Sucrose	-	30 g/l	30 g
Casein hydrolysate	-	1 g/l	1 g

Adjust the pH to 5.8 – 5.6 with 1M KOH, make up the volume to ~490 ml and add the following components under sterile hood:

Components	Stock concentration	Final concentration	Amount for 1 liter
100X Vitamins RCI	100X	1X	10 ml
$CuSO_4.5H_2O$	1.25 mg/ml (5mM)	1.25 µg/ml (5 µM)	1 ml

Dicamba	2.5 mg/ml	5 µg/ml	2 ml
Hygromycin B	50mg/ml	50mg/l	1 ml
Timentin	300mg/ml	150mg/l	0.5 ml

Filter sterilize the solution.

Part III: RCI

Warm the 2x gelrite and the 2x RCI at 50-60°C in a water bath, mix together and pour into petri dishes.

4. *Rhynchospora* Regeneration Medium + BAP (K4N + BAP) (*To regenerate, selected calli - induce shoots and roots)*

Part I: 2X Gelrite

Components	Stock concentration	Final concentration	Amount for 1 liter
Gelrite	-	4 g/l	4 g

Dissolve the components in 500 ml of water in a 1-liter bottle and sterilize by autoclaving.

Part II: 2X K4N

Components	Stock concentration	Final concentration	Amount for 1 liter
Macronutrients	25X	1X	40 ml
Micronutrients	1000X	1X	1 ml
FeNaEDTA	1000X	27.5 mg/l	1 ml
KNO_3	-	3640 mg/l	3.64 grams
Sucrose	-	30 g/l	30 g
Glutamine	-	146 mg/l	146mg
Gamborg B5 Vitamins	112 mg/ml (1000X)	112 mg/l (1X)	1 ml

Adjust the pH to 5.8 – 5.6 with 1M KOH, make up the volume to ~490 ml and add the following components under sterile hood:

Component	Stock concentration	Final concentration	Amount for 1 liter
6-BAP	1 mg/ml	0.225 mg/l	225 µl
Hygromycin (Roche)	50 mg/ml	30 µg/ml	600 µl

Filter sterilize the solution.

Part III: RCI

Warm the 2x gelrite and the 2x K4N at 50-60°C in a water bath, mix together and pour into petri dishes.

4.1 *Rhynchospora* Alternative Regeneration Medium adapted from rice (RpReg)

(To regenerate selected calli - induce shoots and roots)

Part I: 2X Gelrite

Components	Stock concentration	Final concentration	Amount for 1l
Gelrite	-	4 g/l	4 g

Dissolve the components in 500 ml of water in a 1-liter bottle and sterilize by autoclaving.

Part II: 2x RpReg

Components	Stock concentration	Final concentration	Amount for 1l
MS salts (M0221)	-	4.4 g/l	4.4 g
Sucrose	-	30 g/l	30 g
Sorbitol		15 g/l	15 g

Adjust the pH to 5.8 – 5.6 with 1M KOH, make up the volume to 500ml and add the following components under sterile hood:

Components	Stock concentration	Final concentration	Amount for 1 liter
6-BAP	1 mg/ml	1 mg/l	1 ml
Hygromycin (Roche)	50 mg/ml	30 µg/ml	600 µl
NAA	1mg/ml	0.5mg/l	500 µl

Filter sterilize the solution.

Part III: RCI

Warm the 2x gelrite and the 2x K4N at 50-60°C in a water bath, mix together and pour into petri dishes.

5. *Rhynchospora* Rooting Medium (1/2 MS)

(To allow, regenerating calli to develop into full plants in a hormone-free medium, after the appearance of shoots and roots)

Components	Stock concentration	Final concentration	Amount for 1 liter
MS salts (M0221)	-	2.2 g/l	2.2 g
Gelrite	-	4 g/l	4 g
Sucrose	-	15g/l	15 g

Make up the volume to 1-liter, autoclave and pour into sterile glass pots.

6. *Rhynchospora* Infection Medium (RAS)

(To infect, embryos with Agrobacterium culture)

Part I: Concentrated medium without Cysteine and Acetosyringone

Components	Stock concentration	Final concentration	Amount for 500 ml
MS salts (M0221)	-	4.4 g/l	2.2 g
Sucrose	-	20 g/l	10 g
Glucose		10 g/l	5 g
Casein hydrolysate	-	1 g/l	0.5 g
100X Vitamins RCI	100X	1X	5 ml
L-Cysteine	8 g/l (1000X)	800 mg/l (1X)	0.5 ml

Mix everything, adjust the pH to 5.8 with 1M KOH, make up the volume to **500 ml**, filter sterilize and store at 4°C.

Part II: Final medium with Dicamba and Acetosyringone

Freshly prepared for each transformation experiment. For one transformation, take 25 ml of part I in a 50ml tube and add the following:

Components	Stock concentration	Final concentration	Amount for 25 ml
Dicamba	2.5 mg/ml	5 ug/ml	50 ul
Acetosyringone	1 M	250 µM	10 µl

Preparation of stocks

Stock solutions for RCI-related Media

100X Vitamin Stock for BCI

Components	Concentration in stock	Amount for 500 ml of stock	Final concentration in medium
Thiamine HCl	100 mg/l	0.05 g	1 mg/l
myo-inositol	35 g/l	17.5 g	350 mg /l
Proline	69 g/l	34.5 g	690 mg/l

Dissolve the components in ~400 ml of water, make up the volume to 500 ml, filter-sterilize and store at 4°C.

1.25 mg/ml (5 mM) $CuSO_4 \cdot 5H_2O$

Dissolve 125 mg of $CuSO_4 \cdot 5H_2O$ in 100 ml of water, filter sterilize and store at 4°C.

2.5 mg/ml Dicamba

Dissolve 30 mg of Dicamba in 10 ml of water, add few drops of 1M KOH (if necessary) to help with the dissolution. Make up the volume to 12 ml, filter-sterilize, divide in 1 ml aliquots and store at -20°C.

8 g/l L-cysteine (L-cysteine maximum solubility is 25 mg/ml)

Dissolve 400 mg of L-cysteine in 50 ml of water, filter-sterilize and store at 4°C.

100 mM Acetosyringone (3'5'dimethoxy-4'-hydroxy-acetophenone)

Dissolve 196 mg of acetosyringone in DMSO, make up the volume to 10 ml with DMSO and filter-sterilize. Make 0.5-ml aliquots and store at -20°C.

300 mg/ml Timentin

Add 20 ml water to the bottle containing 10 g of Timetin (Duchefa), dissolve and make up the volume to 33.3 ml. Filter-sterilize and store 1 ml aliquots at -20°C.

Gamborg B5 Vitamins

To prepare a 1000x stock solution dissolve 11.2g in 100ml of water. Filter sterilize and store aliquots at -20°C

Stock solutions for K4N Media - December 2020

25X Macronutrients

Components	Amount for 400 ml of stock (grams)	Final concentration in medium (mg/L)
NH_4NO_3	3.2	320
$CaCl_2 \cdot 2H_2O$	4.41	441
KH_2PO_4	3.4	340
$MgSO_4 \cdot 7H_2O$	2.46	246

Dissolve the components and make up the volume to 400 ml. Filter-sterilize and store at room temperature.

1000X Micronutrients

Components	Amount for 1 liter of stock (grams)	Final concentration in medium (mg/L)
$MnSO_4 \cdot H_2O$	9.6	9.6
H_3BO_3	3.1	3.1
$ZnSO_4 \cdot 7H_2O$	7.2	7.2
$CuSO_4 \cdot 5H_2O$	1.25	1.25
KI	0.17	0.17

$Na_2MoO_4 \cdot 2H_2O$	0.12	0.12
$CoCl_2 \cdot 6H_2O$	0.024	0.024

Dissolve the components and make up the volume to 1000 ml. Autoclave and store at room temperature.

1000X FeNaEDTA (27.5 g/l)

Dissolve 2.75 g in water and make up the volume to 100 ml. Filter-sterilize and store at room temperature in the dark.

1 mg/ml 6-BAP

Dissolve 12 mg in 9 ml water + ~300 µl of 1 M NaOH. Make up the volume to 12 ml, filter-sterilize and store aliquots at -20°C. Pre-made commercial solution is also fine.

1 mg/ml 1-NAA

Dissolve 50mg of 1-NAA powder in 50ml EtOH, filter sterilize and store at 4°C in darkness.

Review

Meiosis progression and recombination in holocentric plants: What is known?

frontiers in Plant Science

REVIEW
published: 22 April 2021
doi: 10.3389/fpls.2021.658296

Meiosis Progression and Recombination in Holocentric Plants: What Is Known?

*Paulo G. Hofstatter, Gokilavani Thangavel, Marco Castellani and André Marques**

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

Differently from the common monocentric organization of eukaryotic chromosomes, the so-called holocentric chromosomes present many centromeric regions along their length. This chromosomal organization can be found in animal and plant lineages, whose distribution suggests that it has evolved independently several times. Holocentric chromosomes present an advantage: even broken chromosome parts can be correctly segregated upon cell division. However, the evolution of holocentricity brought about consequences to nuclear processes and several adaptations are necessary to cope with this new organization. Centromeres of monocentric chromosomes are involved in a two-step cohesion release during meiosis. To deal with that holocentric lineages developed different adaptations, like the chromosome remodeling strategy in *Caenorhabditis elegans* or the inverted meiosis in plants. Furthermore, the frequency of recombination at or around centromeres is normally very low and the presence of centromeric regions throughout the entire length of the chromosomes could potentially pose a problem for recombination in holocentric organisms. However, meiotic recombination happens, with exceptions, in those lineages in spite of their holocentric organization suggesting that the role of centromere as recombination suppressor might be altered in these lineages. Most of the available information about adaptations to meiosis in holocentric organisms is derived from the animal model *C. elegans.* As holocentricity evolved independently in different lineages, adaptations observed in *C. elegans* probably do not apply to other lineages and very limited research is available for holocentric plants. Currently, we still lack a holocentric model for plants, but good candidates may be found among Cyperaceae, a large angiosperm family. Besides holocentricity, chiasmatic and achiasmatic inverted meiosis are found in the family. Here, we introduce the main concepts of meiotic constraints and adaptations with special focus in meiosis progression and recombination in holocentric plants. Finally, we present the main challenges and perspectives for future research in the field of chromosome biology and meiosis in holocentric plants.

Keywords: holocentric chromosome, meiotic recombination, cohesion, centromere, inverted meiosis

OPEN ACCESS

Edited by:
Mónica Pradillo, Complutense University of Madrid, Spain

Reviewed by:
Scott Hawley, Stowers Institute for Medical Research, United States
Marcial Escudero, Sevilla University, Spain

****Correspondence:***
André Marques
amarques@mpipz.mpg.de

Specialty section:
This article was submitted to Plant Cell Biology, a section of the journal Frontiers in Plant Science

Received: *25 January 2021*
Accepted: *22 March 2021*
Published: *22 April 2021*

Citation:
Hofstatter PG, Thangavel G, Castellani M and Marques A (2021) Meiosis Progression and Recombination in Holocentric Plants: What Is Known? Front. Plant Sci. 12:658296. doi: 10.3389/fpls.2021.658296

INTRODUCTION

Meiosis, Conserved Mechanisms and Adaptations

Meiosis is a type of cell division responsible for reducing the number of chromosomes in diploid cells by half to produce haploid cells. It is a central step responsible for shuffling genetic information through meiotic recombination and produce genetic variation in eukaryotic life-cycles (Zickler and Kleckner, 2015). This is possible due to two rounds of cell division after a single DNA replication event with the participation of a specific and specialized meiotic machinery (Schurko and Logsdon, 2008).

Preliminary evidence suggests that meiosis is an ancestral feature of eukaryotes, what can robustly explain the patterns of pervasive occurrence of sexual processes in all eukaryotic diversity (Speijer et al., 2015). Despite the extreme conservation of the main meiotic steps even in the most distantly related groups, several lineages have specific meiotic adaptations. In *Drosophila*, several components of the core eukaryotic machinery playing roles in meiosis have been lost or even replaced: the meiosis-specific DMC1 recombinase was replaced by a distant homolog of it, spin-D/RAD51C (Abdu et al., 2003). *Schizosaccharomyces pombe* has lost the main meiotic pathway to resolve crossovers (COs) and heavily relies on a secondary pathway for the resolution of COs (which lacks interference) (Cromie et al., 2006). As a result, CO numbers are significantly higher in *S. pombe* compared to other model organisms, such as *Arabidopsis thaliana* (this plant presents around 1.5 CO per bivalent and both crossover resolution pathways are present) (Mercier et al., 2015). *S. pombe* has also lost the synaptonemal complex (Lorenz et al., 2004) and, thus, performs meiosis with a highly reduced machinery when compared to other well-characterized models. However, meiotic specializations are not restricted to the molecular machinery underpinning the main steps of the process. Some organisms exhibit morphological specializations as a consequence of structural peculiarities of chromosomal organization. For instance, homologous chromosomes (homologs) from some species of the genus *Oenothera* do not synapse upon meiosis rendering them functionally asexual even though they perform meiotic divisions (Johnson et al., 2009). This is due to large scale rearrangements inside the chromosomes, what leads to a state of permanent translocation heterozygosity. Another challenge to the regular progression of meiosis is the evolution of holocentric chromosomes in several lineages. In the main holocentric model, the nematode *C. elegans*, meiosis progresses in such a way that only a single chiasma is formed for each chromosome pair (Martinez-Perez et al., 2008; Martinez-Perez and Colaiacovo, 2009). In the case of holocentric plants of the families *Cyperaceae* (sedges) and Juncaceae (rushes), inverted meiosis evolved to cope with the holocentric chromosome structure: sister-chromatids are separated in the first meiotic division, while homologs are separated only upon the second division (Cabral et al., 2014; Heckmann et al., 2014; Marques et al., 2016). In an even more extreme case, *Rhynchospora tenuis* (Cyperaceae) presents achiasmatic inverted meiosis, whose viability seems to be possible due to the very small number of holocentric chromosomes inside the nucleus (just two pairs) so that even random segregation would produce some viable offspring (Cabral et al., 2014).

Meiosis Progression and Recombination in Monocentric Plants

The sequence of events associated with canonical (monocentric) meiosis is well-established (**Figure 1A**). The homologs pair and synapse by the formation of the synaptonemal complex. After the introduction of double-strand breaks onto DNA, a process of DNA repair based on inter-homolog recombination ensues (Zickler and Kleckner, 2015). The sister chromatids are held together by cohesion along the chromosome arms and centromeres. By the end of prophase I, the homologs have recombined, are physically connected by chiasmata, and meiotic cohesin REC8 along chromosome arms is released (Xu et al., 2005). This segregation scheme necessitates a two-step loss of sister chromatid cohesion. Cohesin is removed distally to chiasmata to allow homologs to segregate during meiosis I while being partially maintained to enable sister-chromatids to partition correctly during meiosis II. In organisms that are monocentric, this sequential loss of cohesion is regulated by shugoshin which is specifically associated to centromeres (Kitajima et al., 2004). Shugoshin protects cohesin at the centromere until meiosis II by recruiting the conserved phosphatase, PP2A, to antagonize the phosphorylation and removal of the cohesin complex (Kitajima et al., 2006; Riedel et al., 2006). At metaphase I, the bivalents align to the metaphase plate with the sister kinetochores being poleward mono-oriented. At anaphase I, homologous centromeres are bi-oriented, the bivalents are detached, as chiasmata are resolved, and the homologs migrate to opposite poles. The sisters are held together until metaphase II by centromeric cohesion. The sister kinetochores now face opposite poles during metaphase II, centromeric cohesion is lost, the sister-chromatids are released and migrate to opposite poles as well. At the end of the meiosis, each nucleus has a haploid number of chromosomes (Mercier et al., 2015).

Meiotic recombination is essential to sexual reproduction and the generation of genetic diversity and, thus, has a profound effect on patterns of genetic variation and is an important tool for crop breeding (Taagen et al., 2020). Variation in recombination rates is of particular interest due to efforts to increase the rate of genetic gain in agricultural crops by breaking up large linkage blocks containing both beneficial and detrimental alleles. Meiotic recombination events (crossovers, i.e., COs) are unevenly distributed throughout eukaryotic genomes, some regions exhibiting higher recombination rates (hotspots), while other exhibiting lower rates (cold spots) (Petes, 2001; Fernandes et al., 2019). The causes of this observed uneven distribution are currently not well-understood.

In most eukaryotes there is at least one CO per chromosome per meiotic event, which is normally required for faithful segregation of chromosomes. Additionally, the average number of COs is relatively low, typically from 1 to 3 events per chromosome (Mercier et al., 2015). In monocentric chromosomes, the density of COs is extremely heterogeneous

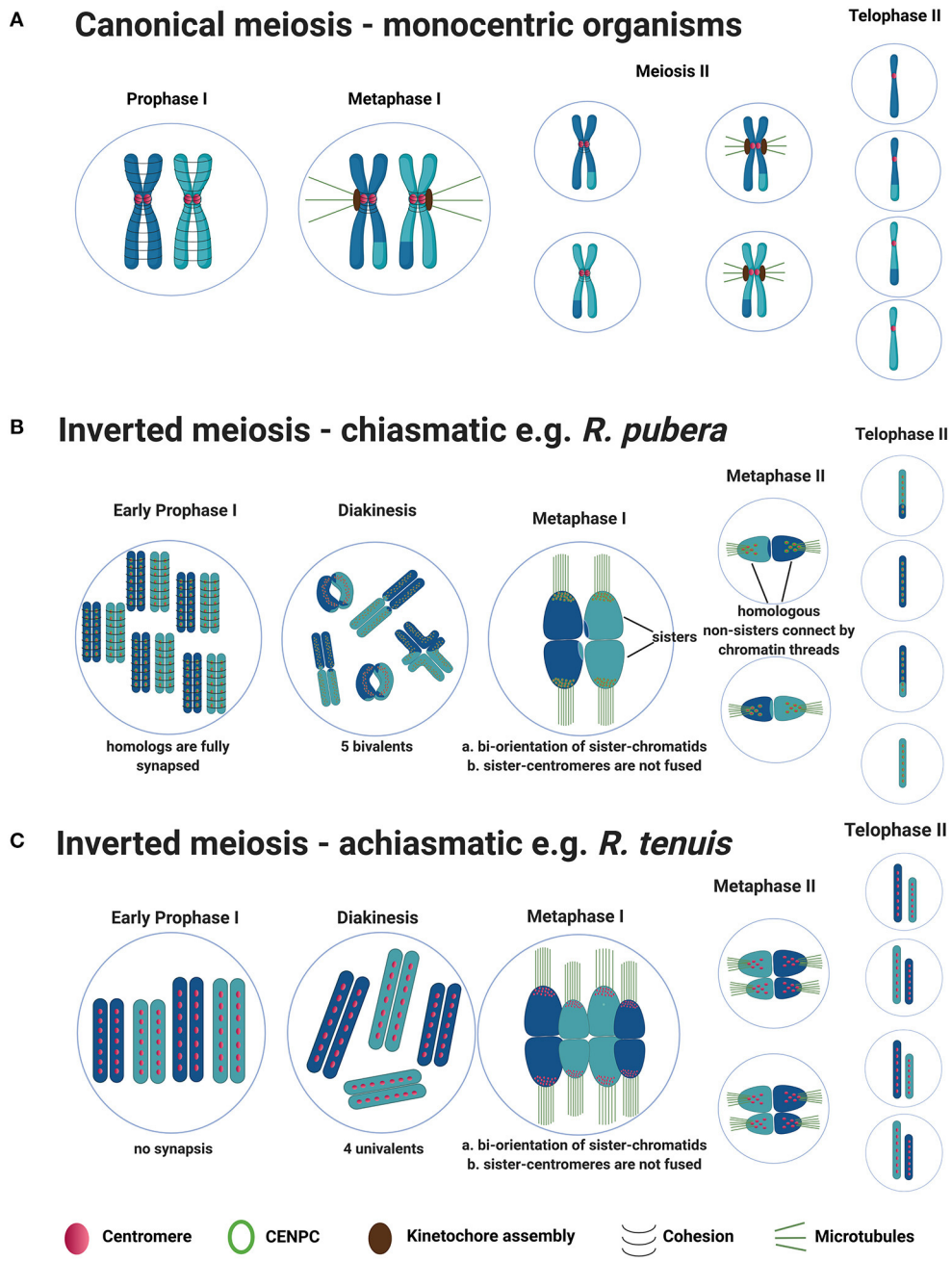


FIGURE 1 | General model for canonical meiosis in monocentric organisms vs. inverted meiosis (both chiasmatic and achiamatic) in holocentric plants. **(A)** Canonical meiosis: During meiosis I reciprocal genetic exchange between homologs (crossovers) occurs, sisters-chromatids mono-orient via fused sister-centromeres and segregate to the same poles. During meiosis II, sisters-chromatids bi-orient and segregate to the opposite poles, resulting in four haploid gametes at the end. **(B)** Schematic representation of chiasmatic inverted meiosis observed in *R. pubera* (from metaphase I only one bivalent is illustrated for better understanding). During meiosis I, COs take place but the difference is that, centromeres from sisters are not fused, sister chromatids bi-orient and segregate to the opposite poles already at anaphase I. During meiosis II homologous non-sisters align, bi-orient and segregate to the opposite poles, resulting in four haploid gametes similar to canonical meiosis. **(C)** Schematic representation of achiasmatic inverted meiosis observed in *R. tenuis*. The sequence of events during inverted meiosis observed in *R. tenuis* is similar to that of *R. pubera*, but meiosis in *R. tenuis* is reported to be achiasmatic i.e., crossover formation doesn't occur during prophase I. As a result, four univalents are observed during diakinesis instead of two bivalents.

at both large (chromosomal) and small scales (kb). Peri- and centromeric regions are largely depleted in COs (cold regions) (Petes, 2001; Fernandes et al., 2019). In some extreme cases, such as wheat, up to 80% of the genome hardly ever experience any COs (Choulet et al., 2014). These regions contain ∼30% of the genes which are thus out of reach for plant breeding.

To exchange DNA, the chromosomes must undergo double-strand breaks. This process of physiologically induced DNA fragmentation is conserved in the vast majority of eukaryotes and is carried out by the topoisomerase-like protein SPO11 (Keeney et al., 1997; Keeney, 2008). After SPO11 introduces double-strand breaks, the free 3′ ends left are targeted by the recombinases RAD51A and DMC1. These proteins help the 3′ ends to search for homologs as templates for repair. After the invasion of the single strand, a recombination intermediate structure is formed, the displacement loop (D-loop) (Brown and Bishop, 2014). DNA synthesis of both ends generate a new structure called double Holliday Junction (dHJ) (Wyatt and West, 2014). A CO is an outcome of the resolution of a dHJ, but other outcomes are possible (Allers and Lichten, 2001). In this case, the invading strand is ejected from the D-loop and anneal to the single-strand 3'end of the original double-strand break. Crossovers may be resolved in two main ways: the main pathway 1 (exhibiting interference) and a secondary pathway 2 (lacking interference). The pathway 1 is a meiosis-specific process with many associated proteins (the so-called ZMM proteins), namely MSH4, MSH5, MER3, HEI10, ZIP4, SHOC1, PTD (Mercier et al., 2015). This pathway is highly conserved among eukaryotes. The secondary pathway involves the protein MUS81. The existence of additional crossover pathways cannot be excluded (Mercier et al., 2015; Lambing et al., 2017).

Holocentric Chromosomes

Apart from the monocentric organization, another type of chromosomal organization, the holocentric (holokinetic) chromosomes, evolved independently in many lineages of unicellular eukaryotes, green plants, and metazoans (Melters et al., 2012; Escudero et al., 2016). Holocentric chromosomes have no distinct primary constriction visible while condensed, as they harbor multiple centromeric domains along their lengths (Heckmann et al., 2013; Steiner and Henikoff, 2014; Marques et al., 2015). Thus, spindle fibers attach along almost the entire poleward surface of the chromatids. As a result, sister-chromatids migrate to opposite poles parallel to each other during anaphase, while in the case of monocentric chromosomes microtubule spindles attach to a distinct kinetochore and the sister chromatids move together to opposite poles at anaphase with a clear attachment of microtubules onto the centromere.

Although organisms with holocentric chromosomes are considered relatively rare, clades possessing such chromosomal structure include more than 350,000 species (Kral et al., 2019). Between 1.5 and 2.0% of the flowering plants (∼5,500 species) are supposed to have holocentric chromosomes (Bures et al., 2013). Likely, due to the lack of chromosome studies, holocentricity should be even more common than reported.

A multiplication of centromeric sequences from one location to multiple sites along the chromosome arms has been proposed as a possible mechanism of holocentromere formation (Greilhuber, 1995). One common explanation for the evolution of holocentric chromosomes is their putative advantage over monocentric ones when it comes to chromosome breakages and consequent karyotypic variation (Zedek and Bures, 2018). The studies on artificial chromosomal rearrangements in various holocentric species showed that chromosome fragments retaining centromeric activity are stably transmitted during mitosis and meiosis (Heckmann et al., 2014; Jankowska et al., 2015).

Recent findings in holocentrics have brought back the discussion about the chromosome structure plasticity of holocentric lineages, including both CENH3-based and CENH3-less holocentromeres (Marques and Pedrosa-Harand, 2016; Drinnenberg and Akiyoshi, 2017). Such plasticity seems to be evolutionarily advantageous for it would increase the resistance of chromosomes against breaks and fusions. However, no difference in diversification rates between monocentrics and holocentrics seems to occur (Marquez-Corro et al., 2018).

Meiosis Progression in Holocentric Organisms

The best studied holocentric organism is the animal model *C. elegans*, and much of what we know about meiotic adaptations in organisms with this kind of chromosome structure derives from it (for additional information, see Wormbook (2005). However, due to the independent origin of holocentric organisms, adaptations in distantly related holocentric lineages are likely to be lineage-specific. In *C. elegans*, despite of its unique adaptations, meiosis progress resembles the process in monocentric organisms, in the way that homologs segregate at the end of meiosis I (Lui and Colaiacovo, 2013). During prophase I, chromosome remodeling processes occur, bivalents acquired a cruciform appearance with a long and a short arm and homologs are segregated to opposite poles l in a way similar to canonical meiosis. But the two-step loss of cohesion is accomplished through an alternate mechanism in a LAB-1 (a functional analog of shugoshin) dependent way (De Carvalho et al., 2008).

Meiotic Progression in Holocentric Plants Is Associated With Inverted Meiosis

Recently, several works have employed modern tools to better characterize the structure and function of holocentric centromeres (holocentromeres) during mitosis and meiosis in plants (Heckmann et al., 2013, 2014; Cabral et al., 2014; Marques et al., 2015, 2016; Oliveira et al., 2019; Neumann et al., 2020). However, the lack of genomic data and functional studies on holocentric plants hamper a better understanding of their cell-division-related adaptations. Upon mitosis, holocentricity does not affect sister chromosome segregation mechanisms, and a parallel migration of sister chromatids substitutes the typical V-shape migration of monocentric chromatids. In contrast, during meiosis several challenges appear because centromeres are not restricted to a single domain as in monocentrics, but rather dispersed across several domains genome-wide.

Thus, the stepwise cohesion release observed in monocentric chromosomes is not possible, since sister-holocentromeres are not associated in holocentric plants precluding their mono-orientation (Cabral et al., 2014; Heckmann et al., 2014; Marques et al., 2016). Additionally, the chromosome remodeling mechanism observed in *C. elegans* is unlike in holocentric plants, since they can have more than one CO per bivalent

and maintenance of holocentromeric activity during meiosis forces the bi-orientation of sister-holocentromeres. Therefore, holocentric plants have developed a different kind of meiosis called post-reductional or inverted meiosis to segregate their chromosomes. The phenomenon of inverted meiosis was first reported as early as 1940 in *Carex* (Wahl, 1940) and since then has been found in other holocentric plants of *Cuscuta*, *Luzula* and *Rhynchospora* (Malheiros et al., 1947; Pazy and Plitman, 1991; Cabral et al., 2014; Heckmann et al., 2014) but also in holocentric insects (Battaglia and Boyes, 1955; Nokkala et al., 2002; Viera et al., 2009). In this type of meiosis, the bivalents align themselves perpendicular to the equatorial plate during metaphase I with bi-orientation of sister-chromatids forcing them to separate to opposite poles during anaphase I (equational division during meiosis I) (**Figure 1B**). Thus, at the end of meiosis I, the daughter cells remain diploid. During meiosis II, thin chromatin threads are seen connecting the homologous non-sisters, which then separate to the opposite poles (reductional division during meiosis II). Although these chromatin threads are observed in both *Luzula* and *Rhynchospora*, it is not yet known what is the mechanism coordinating these connections (Cabral et al., 2014; Heckmann et al., 2014).

Furthermore, very little is known about the protein dynamics involved in the cohesion release and CO control during inverted meiosis in plants. Besides, in the plant genus *Rhynchospora* (beaksedge) both chiasmatic and achiasmatic inverted meiosis have been observed (Cabral et al., 2014). Apparently, meiotic recombination seems to occur in *R. pubera* (2n = 10), since chiasmata formation and the presence of meiosis-associated proteins (RAD51A, ASY1) have been observed, which represent the normal axis formation and occurrence and processing of DNA double strand breaks. In theory, inverted meiosis should be associated with a complete release of the meiotic cohesin REC8 between sister-chromatids already at end of meiosis I, allowing sisters to segregate at anaphase I (**Figure 1B**). However, sister-holocentromeres are not associated in holocentric plants, which could potentially interfere with the role of shugoshin. The behavior of cohesin or shugoshin in holocentric plants exhibiting inverted meiosis is unknown. Furthermore, the achiasmatic species *R. tenuis* ($2n = 4$) exhibits no chiasmata (**Figure 1C**). This species has the smallest reported number of chromosomes in the family and performs meiosis with the formation of four univalents, despite of RAD51 foci being observed, which suggests that DSBs are still occurring but being processed without crossovers (Cabral et al., 2014). Whether a defect in the meiotic machinery of this species is responsible for the achiasmy observed and whether the female meiosis is also achiasmatic is subject to current studies in our group. A similar phenomenon could be identified in a monocentric plant species, *Helianthemum squamatum*, which also exhibits a very small number of chromosomes when compared with close relatives (Aparicio et al., 2019).

The mechanisms behind the inverted meiosis have been further studied in *Luzula elegans* (Heckmann et al., 2014). Anti-CENH3 immunolabeling patterns appeared as linear lines during mitosis as well as meiosis. The authors propose that, a single linear functional centromere may be formed during meiosis and mitosis. Additionally, CENH3 signals from sisters-chromatids always remain separate. This may help, in the bi-polar orientation of the sisters. Each chromatid makes, end to end connection, by means of thin heterochromatin threads with its homologous partner which starts as early as pachytene. These connections are known to be established by satellite elements like LeSAT7, LeSAT11 and may represent chiasmata preserved at sub-telomeric regions (**Figure 2A**). A similar hypothesis was also proposed by Ris (1942) while researching on inverted meiosis in aphids. This connection may be involved in ensuring the correct segregation of homologous non-sister chromatids during the second meiotic division. In *Luzula elegans*, immunolocalization with anti-ASY1 and anti-ZYP1 signals were

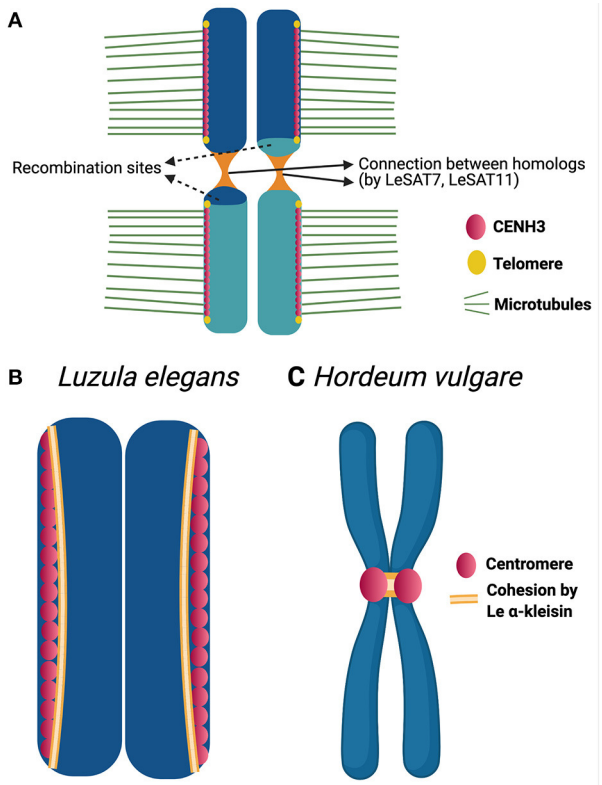


FIGURE 2 | Chromatin threads and cohesion in *Luzula* chromosomes. **(A)** Model highlighting the structural adaptations during inverted meiosis of *Luzula elegans* (see Heckmann et al., 2014) for further details). A rod bivalent with single crossover is illustrated in the model. CENH3 (centromeric protein) appears as a single linear line and centromeres of sisters are not fused. The sister chromatids bi-orient and attach to microtubules from opposite poles. Homologous non-sister chromatids associate with each other by end-to-end connections reported to be established by satellite elements, which maintain non-sister chromatids together up to meiosis II. **(B,C)** CENH3 and Le α-kleisin distribution during mitotic metaphase of *Luzula elegans* **(B)** and *Hordeum vulgare* **(C)** (see Ma et al., 2016) for further details). In the holocentric plant *Luzula*, CENH3(centromeric protein) appear as linear signals during mitotic metaphase. Le α-kleisin appears in the CENH3 regions and not between sister centromeric units. Whereas, in case of the monocentric plant *Hordeum vulgare*, the same Le α-kleisin is reported to present in the centromeric regions as well as establishing a connection between the sister centromeres.

observed as linear lines during early prophase I and telomere bouquet formation was also observed (Heckmann et al., 2014). Thus, early prophase events like DNA double strand break repair, pairing, synapsis and telomere bouquet formation appears the same as canonical meiosis.

Many questions remain a mystery with respect to inverted meiosis. What causes the sisters to separate during meiosis I? Is the cohesion mechanism, which plays a key role in holding the sisters together during meiosis I of canonical meiosis, evolved to enable inverted meiosis? How do the kinetochore proteins assemble and function during inverted meiosis? Monocentric organisms have mechanisms to prevent separase from degrading the cohesion in localized centromeric regions during anaphase I. This enables the sisters to be held together until anaphase II (Nasmyth, 2015). In holocentrics, which have diffuse centromeres, the mechanism of centromeric cohesion protection may be disabled. This may result in the loss of centromeric cohesion and allows the sisters to separate during meiosis I. Attempts to study cohesion mechanism during inverted meiosis were made in *Luzula elegans* (Ma et al., 2016). Signals of LeAlpha-kleisin-1 (cohesin ortholog of AtSYN4) appear during early prophase as reported for cohesin in monocentric meiosis, as demonstrated by immunolabeling. During both mitotic and meiotic metaphases I and II, these signals are observed in CENH3-positive regions but not between sister chromatids (**Figure 2B**). The authors also carried out the same experiment in the monocentric plant *Hordeum vulgare* (barley). In this experiment using the same antibody, the signals were observed in centromeric regions as well as in between the sister chromatids in mitotic metaphase (**Figure 2C**). Thus, the cohesion which connects the sister centromeres together in the monocentric species, barley, seems to not play the same role in the holocentric *Luzula*. This may be an early evidence that the function of a cohesin in monocentric may not be the same in a holocentric organism. It is speculated that LeAlpha-kleisin-1 may be involved in the centromere assembly but lost the function of establishing connection between sisters in *Luzula*. We cannot rule out the possibility of other cohesins involved in the connection between sisters. Thus, cohesins as potential candidates to be studied in future may give us more insights.

Anti-CENH3 immunolabeling patterns appeared as linear lines during mitosis in *R. pubera*. However, during meiosis centromeres form clusters (so-called cluster-holocentromeres) along the poleward side of the bivalents where the microtubules attach perpendicularly during meiosis I and the clusters are present in the middle of the chromatids during meiosis II (Marques et al., 2016). Additionally, CENPC, which represents the outer kinetochore protein, is also co-localized with CENH3 in meiosis which may refer to a conserved assembly of meiotic kinetochores on the holocentromeres (**Figure 1B**). This is the first report about kinetochore proteins in holocentric plants. But still, studies on kinetochore proteins like MIS12 (required for fusion of sister kinetochores), cohesion proteins like SMC1, SMC3, SCC3, REC8 (involved in centromeric cohesion during meiosis I) and shugoshin are necessary to provide more evidence to understand the observed phenomena during inverted meiosis.

The differences in the centromere organization during inverted meiosis of *Luzula* and *Rhynchopsora* show that the mechanisms differ in both cases and the regulation of inverted meiosis may be more complex. Regardless of the differences, in both cases the non-homologous chromatids appear to be connected by thin chromatin threads during meiosis II, as in case of *Luzula* specific tandem repeats were associated to such threads, but the nature of this connection is not yet identified in *Rhynchospora*. Heterochromatic threads seems to play an important role in the separation of achiasmate homologs during female meiosis in *Drosophila* (Hughes and Hawley, 2014). In this particular case the threads seem to be resolved by Topoisomerase II during meiosis I. However, chromatin threads in both *Luzula* and *Rhynchospora* are also observed in meiosis II, whether a similar mechanism occurs in the case of inverted meiosis in these holocentric plants is yet to be shown.

Meiotic Recombination in Holocentric Organisms

Being holocentric can have interesting implications for meiosis. In most eukaryotes and model plant species recombination is suppressed or highly reduced at centromeres (Copenhaver et al., 1998; Fernandes et al., 2019). Recombination at centromeres can disrupt their structural function, impair proper segregation and cause aneuploidy (Nambiar and Smith, 2016). Because of the meiotic recombination suppression at and near centromeres in monocentric organisms, it is of particular interest to understand how meiotic recombination works in organisms with holocentric chromosomes (**Figure 3A**). However, much of what we know about recombination in a holocentric organism comes from studies in *C. elegans*, wherein centromere proteins such as CENH3 and CENP-C are dispensable during meiosis (Monen et al., 2005) and likely do not affect meiotic recombination. In this case recombination rates broadly vary according to physical position in all six of its chromosomes. Each chromosome is comprised of three large domains: a low-recombining, gene-dense center, and two high-recombining arms (Barnes et al., 1995; Rockman and Kruglyak, 2009).

In Lepidoptera, the largest and most diverse holocentric lineage, meiotic recombination is restricted to male meiosis and frequent karyotype reorganization events are associated with wide variations in chromosome counts across species (Hill et al., 2019). Although high recombination densities were reported for some Lepidopteran insects (Wilfert et al., 2007), this does not seem to be linked to holocentricity.

Meiotic Recombination in Holocentric Plants

For the time being there are no detailed analysis about recombination frequencies in holocentric plants and all we know derive from basic cytological studies. Recently, the first linkage map for the presumed holocentric plant *Carex scoparia* (Escudero et al., 2018) has been reported, but without the physical map and holocentromere characterization the recombination landscape for a holocentric plant is still unknown. Understanding

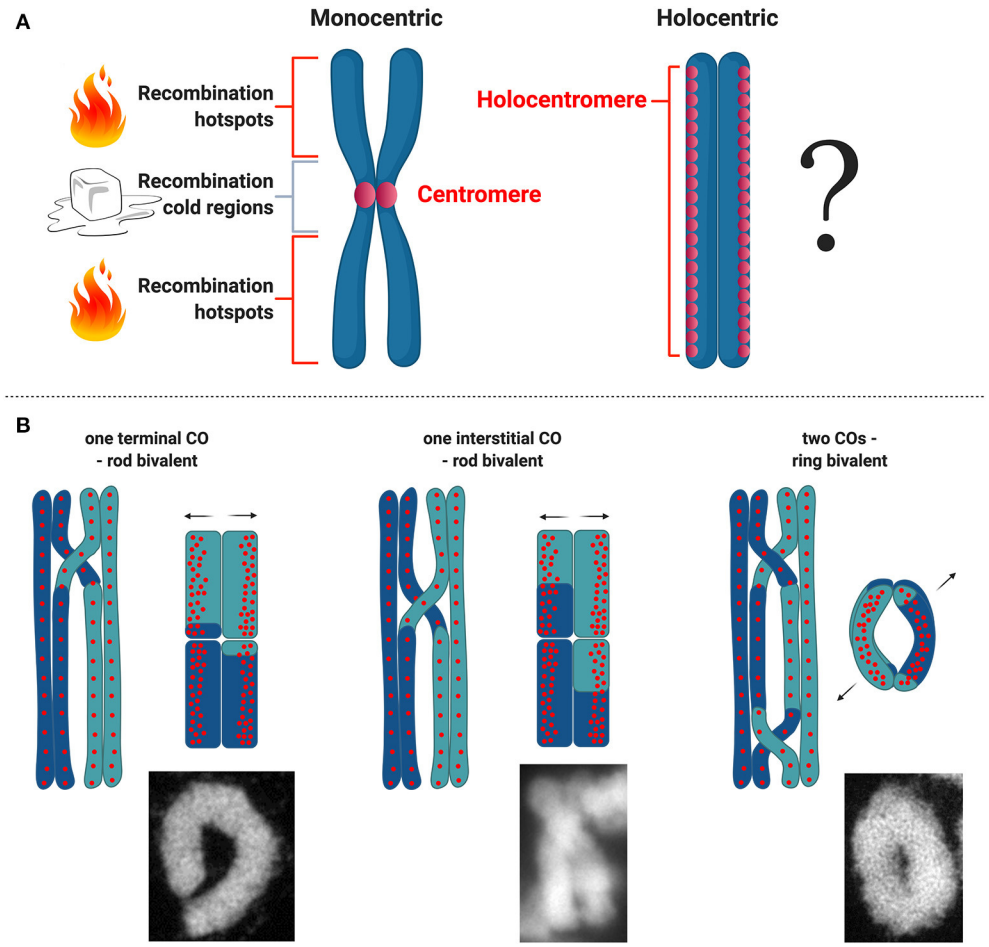


FIGURE 3 | Meiotic recombination in holocentric plants. **(A)** Scheme of what is known about distribution of hotspots for meiotic recombination with respect to centromere organization on monocentric and holocentric plant chromosomes. **(B)** Types of CO and bivalent formation and corresponding models with regard to centromeric units distribution in *R. pubera*. Bivalent microscopic images were made by *M. Castellani*.

how recombination is regulated in holocentric plants will potentially unveil new strategies to deal with this chromosome structure during meiosis. Specially in the case of holocentric plants where chromosomes maintain their holocentromere function during meiosis in contrast to *C. elegans* (Heckmann et al., 2014; Marques et al., 2016), which could potentially interfere with the designation of CO events. In the particular case of the plant *R. pubera*, holocentromeres of *R. pubera* extend linearly for the whole length of the chromosomes until their very ends (Cabral et al., 2014; Marques et al., 2015, 2016). Despite the observation that chiasmata frequently link homologs terminally, it seems that recombination in *R. pubera* also happens in internal regions (**Figure 3B**). Proximity of CO events to centromeric units cannot yet be quantified and recombination may happen in intervals where these units are not present. It is interesting though that centromeric units in *R. pubera* are associated with highly abundant repeats (Tyba repeats), which build short arrays of ∼15 kb long and are dispersed genome wide (Marques et al., 2015). In this sense the repeat-based holocentromeres of *R. pubera* seem to assemble in chromatin structures more similar to repeat-based monocentromes. It was estimated that each chromosome should have between 800 and 1,300 repeat-based centromere domains. Taking in account that RAD51 foci are found dispersed in early prophase I (Cabral et al., 2014) and that CENH3 does show similar signals (Marques et al., 2016), DSB sites could potentially occur very close or even within centromeric units.

Cytological observations in *R. pubera* show that at diakinesis five bivalents are present, and physically connected by chiasmata. In this species, ring-shaped bivalents are supposed to be connected by two chiasmata and rod-like bivalents to be connected by only one (Cabral et al., 2014). Observing the shapes of these bivalents, it seems that in *R. pubera* COs are happening mostly at the ends of the chromosomes, but, less frequent, internal COs are also observed. The occurrence of internal COs suggests that recombination events may take place in the vicinity of centromeric repeats (**Figure 3B**). Similar findings were observed in *Luzula* (Heckmann et al., 2014). Moreover,

this is an evidence that the final product of recombination, the crossover, is present at the end of prophase I and that CO interference is occurring as well as CO assurance. Considering the conservation level of the whole ZMM pathway, it seems that meiotic recombination in *R. pubera* is happening and that is not impaired by holocentromeres or inverted meiosis. These observations are quite interesting considering that the holocentromeres in *R. pubera* are repeat-based and distributed along the entire chromosomes in meiosis (Marques et al., 2016). It will be particularly interesting to study whether COs are somehow affected by such centromere distribution and where they are formed.

The molecular basis of recombination repression at centromeres is still not clear. Two possible ways are speculated to happen: either recombination is repressed at the DSB level by modulating the action or the binding of SPO11, or at the level of how DSBs are repaired and processed by the meiosis-specific DMC1 (Nambiar and Smith, 2016). Recent findings using budding and fission yeast has proposed a role for the kinetochore and cohesion as important regulators of DSBs formation within centromeres and surrounding regions (Vincenten et al., 2015; Kuhl and Vader, 2019). Considering the apparent proximity of recombination events and centromeres in *R. pubera*, it is still unclear whether these repression mechanisms exist and if so, how they are regulated. If we look at other well-studied model eukaryotes, the centromere effect appears to be highly conserved and also very efficient in avoiding COs in pericentromeric regions. In *Drosophila melanogaster* the DSB landscape appears to be flat along the chromosome arm, but downstream recombination is then affected by the centromere effect that eliminates pericentromeric recombination intermediates and models the recombination pattern (Hatkevich et al., 2017; Brady et al., 2018). In maize the centromeric effect seems to work with a different mechanism but with the same result. In centromeric regions of maize DSB can be detected, but COs are absent (He et al., 2017). In *Arabidopsis*, Spo11-oligos resulting from Chip-seq experiment are depleted at pericentromeric regions, where CO are also absent, indicating reduced levels of DSBs at these regions (Choi et al., 2018). In yeast, kinetochore complexes protect centromeric regions, reducing dramatically DSB and CO (Vincenten et al., 2015; Kuhl and Vader, 2019).

A similar question involves the presence of so-called hotspots and cold regions of recombination, regions on the chromosomes where recombination is more or less likely to take place. Multiple species, including plants, display hot and cold spots (e.g., centromeric regions) (Choi and Henderson, 2015). However, the presence of holocentromeres in *R. pubera* makes it difficult to predict the presence of hotspots or cold regions or to speculate about their location. Perhaps the situation is that there are no hotspots in *R. pubera* similar to *C. elegans*. A study in *C. elegans* has made a detailed analysis of recombination rate in a 2 Mb region, discovering that there are no clear hotspots, but recombination rates are constant, constrained only by the structural domain of the chromosome arm (Kaur and Rockman, 2014). This is a unique case similar only to S. *pombe*, which is not holocentric.

A different case is the one of the holocentric relative *Rhynchospora tenuis*. In this species chiasmata are not observed and at least male meiosis seems to be complete achiasmatic (**Figure 1C**). The further observation of RAD51 during early prophase I suggests, in principle, that DSBs are being formed (Cabral et al., 2014). The absence of recombination outcomes might be evidence of the disruption of the ZMM recombination pathway in one or more points. Mutations in the SC of *C. elegans* negatively affect recombination and crossover regulation (Colaiacovo et al., 2003). However, this behavior is not consistent among plant species. For instance, in barley it was reported that dramatic reduction of normal levels of ZYP1 by RNAi also drastically reduce CO formation (Barakate et al., 2014). However, in the case of *Arabidopsis* and rice a malfunctioning SC does not affect recombination and may even increase CO frequency and abolish CO interference (Wang et al., 2010; Capilla-Perez et al., 2021; France et al., 2021). In both holocentric *Rhynchospora* and *Luzula* it was shown that they apparently have conserved SC structures as immunostaining with SC proteins showed the conserved pattern for monocentric species (Cabral et al., 2014; Heckmann et al., 2014). Whether SC proteins are involved in CO regulation in holocentric plants is currently unknown and should be subject of future studies.

An interesting point in holocentric clades is that chromosome numbers tend to vary greatly within the group which could be a consequence of lack of centromere constrain. However, this may not be true for all holocentric clades (Ruckman et al., 2020). In the *Cyperaceae* family, which is the largest group of holocentric plants, chromosomes vary from $n = 2$ to $n = 108$ (Roalson, 2008). Although the lowest chromosome number in angiosperms is found in *Rhynchospora tenuis* ($n = 2$), we can also find extraordinarily very high chromosomes numbers in other genera within this family, for instance in *Carex* (Wieclaw et al., 2020) and *Cyperus* (Roalson, 2008). Since the number of chromosomes is proportional to recombination rates, high chromosome numbers would also impose higher recombination rates in holocentric plants, specially, in this case where the number of chiasmata tends to be typically low, with one or two CO per bivalent. However, a fitness balance must exist otherwise holocentric organisms would tend to have always high chromosome numbers, which is not the case. High chromosome numbers would potentially increase the complexity of the recombination process with likely more possibilities of mistakes in the segregation process.

Holokinetic Drive

Besides the occurrence of inverted meiosis, holocentric sedges (Cyperaceae) also exhibit another peculiar process: the formation of pseudomonads by the end of the microspore meiosis (Rocha et al., 2016, 2018). During this process, three microspores degenerate and only one proceeds with gametogenesis. Thus, only one pollen grain results from each meiotic event in these plants. This specific feature could relate the segregation process with the size of the chromosomes in a process called holokinetic drive, which was first introduced by Bures and Zedek (2014). According to this hypothesis, there would be a selection for chromosomal size upon meiosis. Either the smallest or the largest chromosomes would be favored depending on the case, and

formation of pseudomonads could accelerate this process. A negative correlation between chromosome number and total genome size observed in several holocentric groups seems to support this. For instance, this correlation has been recently reported for the genus *Rhynchospora* (Burchardt et al., 2020). Moreover, it has been recently proposed that centromere drive could occur in association with holokinetic drive in members of *Cyperaceae* and, thus, the meiotic asymmetry in both sexes of this family could increase the potential for selfish centromeres to gain an advantage in both male and female meiosis (Krátká et al., 2021). Alternatively, the selection of the survival cell could be related with the results of the recombination process, wherein the best combination of alleles resulting from the meiotic event would be selected.

PERSPECTIVES AND FUTURE AIMS

The mechanisms behind inverted meiosis in holocentric organisms are currently unknown. The occurrence of inverted meiosis demands modification in the conserved mechanisms of meiotic cohesion and chromosome segregation. New adaptations and differential regulation of meiotic cohesions such as REC8 and centromere cohesion guardians such as shugoshins are expected to have happened. Additionally, modification of the spindle attachment machinery also should be expected due to an alternative centromeric organization. Furthermore, the observed chiasmata formation between holocentric chromosomes demands adaptations of the mechanisms that prevent recombination at or around centromeres. The limited knowledge of holocentromeres and close relatives of *Cyperaceae* limits us to speculate about what to expect in terms of adaptations of the meiotic recombination machinery to holocentricity. Future studies aiming the molecular characterization of such mechanisms will be of interest for evolutionary and comparative biology studies.

AUTHOR CONTRIBUTIONS

PH drafted the section about the meiotic machinery in eukaryotes, incorporated the different contributions and reviewed the text. GT drafted the sections about meiosis progression and inverted meiosis in holocentrics. MC drafted the section about meiotic recombination. AM drafted the sections about holocentric plants, reviewed and supervised the production process of the manuscript. Figures were made by GT and AM. All authors contributed to the article and approved the submitted version.

FUNDING

PH, MC, and AM were funded by Max Planck Society. GT thanks to the DAAD/India Ph.D. fellowship.

ACKNOWLEDGMENTS

We thank both reviewers for the insightful comments and suggestions that helped the improvement of the manuscript. All figure illustrations were created with BioRender.com.

REFERENCES

Abdu, U., Gonzalez-Reyes, A., Ghabrial, A., and Schupbach, T. (2003). The *Drosophila* spn-D gene encodes a RAD51C-like protein that is required exclusively during meiosis. *Genetics* 165, 197–204.

Allers, T., and Lichten, M. (2001). Differential timing and control of noncrossover and crossover recombination during meiosis. *Cell* 106, 47–57. doi: 10.1016/S0092-8674(01)00416-0

Aparicio, A., Escudero, M., Valdes-Florido, A., Pachon, M., Rubio, E., Albaladejo, R. G., et al. (2019). Karyotype evolution in Helianthemum (Cistaceae): dysploidy, achiasmate meiosis and ecological specialization in *H. squamatum*, a true gypsophile. *Bot. J. Linn. Soc.*191, 484–501. doi: 10.1093/botlinnean/boz066

Barakate, A., Higgins, J. D., Vivera, S., Stephens, J., Perry, R. M., Ramsay, L., et al. (2014). The synaptonemal complex protein ZYP1 is required for imposition of meiotic crossovers in barley. *Plant Cell* 26, 729–740. doi: 10.1105/tpc.113.121269

Barnes, T. M., Kohara, Y., Coulson, A., and Hekimi, S. (1995). Meiotic recombination, noncoding DNA and genomic organization in *Caenorhabditis elegans*. *Genetics* 141, 159–179. doi: 10.1093/genetics/141.1.159

Battaglia, E., and Boyes, J. W. (1955). Post-reductional meiosis: its mechanismand causes. *Caryologia* 8, 87–134. doi: 10.1080/00087114.1955.10797554

Brady, M. M., Mcmahan, S., and Sekelsky, J. (2018). Loss of *Drosophila* Mei-41/ATR alters meiotic crossover patterning. *Genetics* 208, 579–588. doi: 10.1534/genetics.117.300634

Brown, M. S., and Bishop, D. K. (2014). DNA strand exchange and RecA homologs in meiosis. *Cold Spring Harb. Perspect. Biol.* 7:a016659. doi: 10.1101/cshperspect.a016659

Burchardt, P., Buddenhagen, C. E., Gaeta, M. L., Souza, M. D., Marques, A., and Vanzela, A. L. L. (2020). Holocentric karyotype evolution in rhynchospora is marked by intense numerical, structural, and genome size changes. *Front. Plant Sci.* 11:536507. doi: 10.3389/fpls.2020.536507

Bures, P., and Zedek, F. (2014). Holokinetic drive: centromere drive in chromosomes without centromeres. *Evolution* 68, 2412–2420. doi: 10.1111/evo.12437

Bures, P., Zedek, F., and Markova, M. (2013). "Holocentric chromosomes," in *Plant Genome Diversity,* eds J. Greilhuber, J. Dolezel, and J. F. Wendel (Vienna: Springer-Verlag Wien). doi: 10.1007/978-3-7091-1160-4_12

Cabral, G., Marques, A., Schubert, V., Pedrosa-Harand, A., and Schlogelhofer, P. (2014). Chiasmatic and achiasmatic inverted meiosis of plants with holocentric chromosomes. *Nat. Commun.* 5:5070. doi: 10.1038/ncomms6070

Capilla-Perez, L., Durand, S., Hurel, A., Lian, Q., Chambon, A., Taochy, C., Solier, V., Grelon, M., and Mercier, M. (2021). The synaptonemal complex imposes crossover interference and heterochiasmy in *Arabidopsis*. *Proc. Natl. Acad. Sci. U. S. A.* 118:e2023613118. doi: 10.1073/pnas.2023613118

Choi, K., and Henderson, I. R. (2015). Meiotic recombination hotspots - a comparative view. *Plant J.* 83, 52–61. doi: 10.1111/tpj.12870

Choi, K., Zhao, X. H., Tock, A. J., Lambing, C., Underwood, C. J., Hardcastle, T. J., et al. (2018). Nucleosomes and DNA methylation shape meiotic DSB frequency in *Arabidopsis thaliana* transposons and gene regulatory regions. *Genome Res.* 28, 532–546. doi: 10.1101/gr.225599.117

Choulet, F., Alberti, A., Theil, S., Glover, N., Barbe, V., Daron, J., et al. (2014). Structural and functional partitioning of bread wheat chromosome 3B. *Science* 345:1249721. doi: 10.1126/science.1249721

Colaiacovo, M. P., Macqueen, A. J., Martinez-Perez, E., Mcdonald, K., Adamo, A., La Volpe, A., et al. (2003). Synaptonemal complex assembly in *C. elegans* is dispensable for loading strand-exchange proteins but critical for proper completion of recombination. *Dev. Cell* 5, 463–474. doi: 10.1016/S1534-5807(03)00232-6

Copenhaver, G. P., Browne, W. E., and Preuss, D. (1998). Assaying genome-wide recombination and centromere functions with Arabidopsis tetrads. *Proc. Natl. Acad. Sci. U. S. A.* 95, 247–252. doi: 10.1073/pnas.95.1.247

Cromie, G. A., Hyppa, R. W., Taylor, A. F., Zakharyevich, K., Hunter, N., and Smith, G. R. (2006). Single holliday junctions are intermediates of meiotic recombination. *Cell* 127, 1167–1178. doi: 10.1016/j.cell.2006.09.050

De Carvalho, C. E., Zaaijer, S., Smolikov, S., Gu, Y., Schumacher, J. M., and Colaiacovo, M. P. (2008). LAB-1 antagonizes the aurora B kinase in *C. elegans*. *Genes Dev.* 22, 2869–2885. doi: 10.1101/gad.1691208

Drinnenberg, I. A., and Akiyoshi, B. (2017). Evolutionary lessons from species with unique kinetochores. *Prog. Mol. Subcell. Biol.* 56, 111–138. doi: 10.1007/978-3-319-58592-5_5

Escudero, M., Hahn, M., and Hipp, A. L. (2018). RAD-seq linkage mapping and patterns of segregation distortion in sedges: meiosis as a driver of karyotypic evolution in organisms with holocentric chromosomes. *J. Evol. Biol.* 31, 833–843. doi: 10.1111/jeb.13267

Escudero, M., Marquez-Corro, J. I., and Hipp, A. L. (2016). The phylogenetic origins and evolutionary history of holocentric chromosomes. *Syst. Bot.* 41, 580–585. doi: 10.1600/036364416X692442

Fernandes, J. B., Wlodzimierz, P., and Henderson, I. R. (2019). Meiotic recombination within plant centromeres. *Curr. Opin. Plant Biol.* 48, 26–35. doi: 10.1016/j.pbi.2019.02.008

France, M. G., Enderle, J., Röhrig, S., Puchta, H., Franklin, F. C. H., and Higgins, J. D. (2021). ZYP1 is required for obligate cross-over formation and cross-over interference in *Arabidopsis*. *PNAS*. 118:e2021671118. doi: 10.1073/pnas.2021671118

Greilhuber, J. (1995). "Chromosomes of the monocotyledons (general aspects)," in *Monocotyledons: Systematics and Evolution*, eds P. J. Rudall, P. J. Cribb, D. F. Cutler, and C. J. Humphries (Surrey: Kew Royal Botanic Gardens), 379–414.

Hatkevich, T., Kohl, K. P., Mcmahan, S., Hartmann, M. A., Williams, A. M., and Sekelsky, J. (2017). Bloom syndrome helicase promotes meiotic crossover patterning and homolog disjunction. *Curr. Biol.* 27, 96–102. doi: 10.1016/j.cub.2016.10.055

He, Y., Wang, M. H., Dukowic-Schulze, S., Zhou, A., Tiang, C. L., Shilo, S., et al. (2017). Genomic features shaping the landscape of meiotic double-strand-break hotspots in maize. *Proc. Natl. Acad. Sci. U. S. A.* 114, 12231–12236. doi: 10.1073/pnas.1713225114

Heckmann, S., Jankowska, M., Schubert, V., Kumke, K., Ma, W., and Houben, A. (2014). Alternative meiotic chromatid segregation in the holocentric plant *Luzula elegans*. *Nat. Commun.* 5:4979. doi: 10.1038/ncomms5979

Heckmann, S., Macas, J., Kumke, K., Fuchs, J., Schubert, V., Ma, L., et al. (2013). The holocentric species *Luzula elegans* shows interplay between centromere and large-scale genome organization. *Plant J.* 73, 555–565. doi: 10.1111/tpj.12054

Hill, J., Rastas, P., Hornett, E. A., Neethiraj, R., Clark, N., Morehouse, N., et al. (2019). Unprecedented reorganization of holocentric chromosomes provides insights into the enigma of lepidopteran chromosome evolution. *Sci. Adv.* 5:eaau3648. doi: 10.1126/sciadv.aau3648

Hughes, S. E., and Hawley, R. S. (2014). Topoisomerase II is required for the proper separation of heterochromatic regions during drosophila melanogaster female meiosis. *PLoS Genet.* 10:e1004650. doi: 10.1371/journal.pgen.1004650

Jankowska, M., Fuchs, J., Klocke, E., Fojtova, M., Polanska, P., Fajkus, J., et al. (2015). Holokinetic centromeres and efficient telomere healing enable rapid karyotype evolution. *Chromosoma* 124, 519–528. doi: 10.1007/s00412-015-0524-y

Johnson, M. T., Agrawal, A. A., Maron, J. L., and Salminen, J. P. (2009). Heritability, covariation and natural selection on 24 traits of common evening primrose (*Oenothera biennis*) from a field experiment. *J. Evol. Biol.* 22, 1295–1307. doi: 10.1111/j.1420-9101.2009.01747.x

Kaur, T., and Rockman, M. V. (2014). Crossover heterogeneity in the absence of hotspots in *Caenorhabditis elegans*. *Genetics* 196, 137–148. doi: 10.1534/genetics.113.158857

Keeney, S. (2008). Spo11 and the formation of DNA double-strand breaks in meiosis. *Genome Dyn. Stab.* 2, 81–123. doi: 10.1007/7050_2007_026

Keeney, S., Giroux, C. N., and Kleckner, N. (1997). Meiosis-specific DNA double-strand breaks are catalyzed by Spo11, a member of a widely conserved protein family. *Cell* 88, 375–384. doi: 10.1016/S0092-8674(00)81876-0

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

Kitajima, T. S., Sakuno, T., Ishiguro, K., Iemura, S., Natsume, T., Kawashima, S. A., et al. (2006). Shugoshin collaborates with protein phosphatase 2A to protect cohesin. *Nature* 441, 46–52. doi: 10.1038/nature04663

Kral, J., Forman, M., Korinkova, T., Lerma, A. C. R., Haddad, C. R., Musilova, J., et al. (2019). Insights into the karyotype and genome evolution of haplogyne spiders indicate a polyploid origin of lineage with holokinetic chromosomes. *Sci. Rep.* 9:3001. doi: 10.1038/s41598-019-39034-3

Krátká, M., Šmerda, J., Lojdová, K., Bureš, P., and Zedek, F. (2021). Holocentric chromosomes probably do not prevent centromere drive in *Cyperaceae*. *Front. Plant Sci.* 12:642661. doi: 10.3389/fpls.2021.642661

Kuhl, L. M., and Vader, G. (2019). Kinetochores, cohesin, and DNA breaks: controlling meiotic recombination within pericentromeres. *Yeast* 36, 121–127. doi: 10.1002/yea.3366

Lambing, C., Franklin, F. C., and Wang, C. R. (2017). Understanding and manipulating meiotic recombination in plants. *Plant Physiol.* 173, 1530–1542. doi: 10.1104/pp.16.01530

Lorenz, A., Wells, J. L., Pryce, D. W., Novatchkova, M., Eisenhaber, F., Mcfarlane, R. J., et al. (2004). *S. pombe* meiotic linear elements contain proteins related to synaptonemal complex components. *J. Cell. Sci.* 117, 3343–3351. doi: 10.1242/jcs.01203

Lui, D. Y., and Colaiacovo, M. P. (2013). Meiotic development in *Caenorhabditis elegans*. *Adv. Exp. Med. Biol.* 757, 133–170. doi: 10.1007/978-1-4614-4015-4_6

Ma, W., Schubert, V., Martis, M. M., Hause, G., Liu, Z., Shen, Y., et al. (2016). The distribution of alpha-kleisin during meiosis in the holocentromeric plant *Luzula elegans*. *Chromosome Res.* 24, 393–405. doi: 10.1007/s10577-016-9529-5

Malheiros, N., Castro, D., and Câmara, A. (1947). Cromosomas sem centrómero localizado: O caso de *Luzula purpurea* link. *Agron. Lusitana* 9, 51–74.

Marques, A., and Pedrosa-Harand, A. (2016). Holocentromere identity: from the typical mitotic linear structure to the great plasticity of meiotic holocentromeres. *Chromosoma* 125, 669–681. doi: 10.1007/s00412-016-0612-7

Marques, A., Ribeiro, T., Neumann, P., Macas, J., Novak, P., Schubert, V., et al. (2015). Holocentromeres in *Rhynchospora* are associated with genome-wide centromere-specific repeat arrays interspersed among euchromatin. *Proc. Natl. Acad. Sci. U.S.A.* 112, 13633–13638. doi: 10.1073/pnas.1512255112

Marques, A., Schubert, V., Houben, A., and Pedrosa-Harand, A. (2016). Restructuring of holocentric centromeres during meiosis in the plant rhynchospora pubera. *Genetics* 204, 555–568. doi: 10.1534/genetics.116.191213

Marquez-Corro, J. I., Escudero, M., and Luceno, M. (2018). Do holocentric chromosomes represent an evolutionary advantage? A study of paired analyses of diversification rates of lineages with holocentric chromosomes and their monocentric closest relatives. *Chromosome Res.* 26, 139–152. doi: 10.1007/s10577-017-9566-8

Martinez-Perez, E., and Colaiacovo, M. P. (2009). Distribution of meiotic recombination events: talking to your neighbors. *Curr. Opin. Genet. Dev.* 19, 105–112. doi: 10.1016/j.gde.2009.02.005

Martinez-Perez, E., Schvarzstein, M., Barroso, C., Lightfoot, J., Dernburg, A. F., and Villeneuve, A. M. (2008). Crossovers trigger a remodeling of meiotic chromosome axis composition that is linked to two-step loss of sister chromatid cohesion. *Genes Dev.* 22, 2886–2901. doi: 10.1101/gad.1694108

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

Mercier, R., Mezard, C., Jenczewski, E., Macaisne, N., and Grelon, M. (2015). The molecular biology of meiosis in plants. *Annu. Rev. Plant Biol.* 66, 297–327. doi: 10.1146/annurev-arplant-050213-035923

Monen, J., Maddox, P. S., Hyndman, F., Oegema, K., and Desai, A. (2005). Differential role of CENP-A in the segregation of holocentric *C. elegans* chromosomes during meiosis and mitosis. *Nat. Cell Biol.* 7, 1248–1255. doi: 10.1038/ncb1331

Nambiar, M., and Smith, G. R. (2016). Repression of harmful meiotic recombination in centromeric regions. *Semin. Cell Dev. Biol.* 54, 188–197. doi: 10.1016/j.semcdb.2016.01.042

Nasmyth, K. (2015). A meiotic mystery: how sister kinetochores avoid being pulled in opposite directions during the first division. *Bioessays* 37, 657–665. doi: 10.1002/bies.201500006

Neumann, P., Oliveira, L., Cizkova, J., Jang, T. S., Klemme, S., Novak, P., et al. (2020). Impact of parasitic lifestyle and different types of centromere organization on chromosome and genome evolution in the plant genus *Cuscuta*. *New Phytol.* 229, 2365–2377. doi: 10.1111/nph.17003

Nokkala, S., Laukkanen, A., and Nokkala, C. (2002). Mitotic and meiotic chromosomes in *Somatochlora metallica* (Cordulidae, Odonata). The absence of localized centromeres and inverted meiosis. *Hereditas* 136, 7–12. doi: 10.1034/j.1601-5223.2002.1360102.x

Oliveira, L., Neumann, P., Jang, T. S., Klemme, S., Schubert, V., Koblizkova, A., et al. (2019). Mitotic spindle attachment to the holocentric chromosomes of *Cuscuta europaea* does not correlate with the distribution of CENH3 chromatin. *Front. Plant Sc.* 10:1799. doi: 10.3389/fpls.2019.01799

Pazy, B., and Plitman, U. (1991). Unusual chromosome separation in meiosis of *Cuscuta L. Genome* 34, 533–536. doi: 10.1139/g91-082

Petes, T. D. (2001). Meiotic recombination hot spots and cold spots. *Nat. Rev. Genet.* 2, 360–369. doi: 10.1038/35072078

Riedel, C. G., Katis, V. L., Katou, Y., Mori, S., Itoh, T., Helmhart, W., et al. (2006). Protein phosphatase 2A protects centromeric sister chromatid cohesion during meiosis I. *Nature* 441, 53–61. doi: 10.1038/nature04664

Ris, H. (1942). A cytological and experimental analysis of the meiotic behavior of the univalent X chromosome in the bearberry aphid Tamalia (==Phyllaphis) coweni (CkLL). *J. Exp. Zool.* 90, 267–330. doi: 10.1002/jez.1400900207

Roalson, E. H. (2008). A synopsis of chromosome number variation in the *Cyperaceae*. *Bot. Rev.* 74, 209–393. doi: 10.1007/s12229-008-9011-y

Rocha, D. M., Marques, A., Andrade, C. G. T. J., Guyot, R., Chaluvadi, S. R., Pedrosa-Harand, A., et al. (2016). Developmental programmed cell death during asymmetric microsporogenesis in holocentric species of *Rhynchospora* (*Cyperaceae*). *J. Exp. Bot.* 67, 5391–5401. doi: 10.1093/jxb/erw300

Rocha, D. M., Vanzela, A. L. L., and Mariath, J. E. A. (2018). Comparative study of microgametogenesis in members of *Cyperaceae* and *Juncaceae*: a shift from permanent pollen tetrads to pseudomonads. *Bot. J. Linn. Soc.* 188, 59–73. doi: 10.1093/botlinnean/boy041

Rockman, M. V., and Kruglyak, L. (2009). Recombinational landscape and population genomics of *Caenorhabditis elegans*. *PLoS Genet.* 5:e1000419. doi: 10.1371/journal.pgen.1000419

Ruckman, S. N., Jonika, M. M., Casola, C., and Blackmon, H. (2020). Chromosome number evolves at equal rates in holocentric and monocentric clades. *PLoS Genet.* 16:e1009076. doi: 10.1371/journal.pgen.1009076

Schurko, A. M., and Logsdon, J. M. Jr. (2008). Using a meiosis detection toolkit to investigate ancient asexual "scandals" and the evolution of sex. *Bioessays* 30, 579–589. doi: 10.1002/bies.20764

Speijer, D., Lukes, J., and Elias, M. (2015). Sex is a ubiquitous, ancient, and inherent attribute of eukaryotic life. *Proc. Natl. Acad. Sci. U. S. A.* 112, 8827–8834. doi: 10.1073/pnas.1501725112

Steiner, F. A., and Henikoff, S. (2014). Holocentromeres are dispersed point centromeres localized at transcription factor hotspots. *eLife* 3:e02025. doi: 10.7554/eLife.02025.025

Taagen, E., Bogdanove, A. J., and Sorrells, M. E. (2020). Counting on crossovers: controlled recombination for plant breeding. *Trends Plant Sci.* 25, 455–465. doi: 10.1016/j.tplants.2019.12.017

Viera, A., Page, J., and Rufas, J. S. (2009). Inverted meiosis: the true bugs as a model to study. *Genome Dyn.* 5, 137–156. doi: 10.1159/000166639

Vincenten, N., Kuhl, L. M., Lam, I., Oke, A., Kerr, A. R., Hochwagen, A., et al. (2015). The kinetochore prevents centromere-proximal crossover recombination during meiosis. *eLife* 4:e10850. doi: 10.7554/eLife.10850.036

Wahl, H. A. (1940). Chromosome numbers and meiosis in the genus Carex. *Am. J. Bot.* 27, 458–470. doi: 10.1002/j.1537-2197.1940.tb14707.x

Wang, M., Wang, K. J., Tang, D., Wei, C. X., Li, M., Shen, Y., et al. (2010). The central element protein ZEP1 of the synaptonemal complex regulates the number of crossovers during meiosis in rice. *Plant Cell* 22, 417–430. doi: 10.1105/tpc.109.070789

Wieclaw, H., Kalinka, A., and Koopman, J. (2020). Chromosome numbers of Carex (*Cyperaceae*) and their taxonomic implications. *PLoS ONE* 15:e0228353. doi: 10.1371/journal.pone.0228353

Wilfert, L., Gadau, J., and Schmid-Hempel, P. (2007). Variation in genomic recombination rates among animal taxa and the case of social insects. *Heredity* 98, 189–197. doi: 10.1038/sj.hdy.6800950

Wormbook (2005). *WormBook, The Online Review of C. elegans Biology*. T.C.E.R. Community. Available online at: http://www.wormbook.org (accessed March 20, 2021).

Wyatt, H. D., and West, S. C. (2014). Holliday junction resolvases. *Cold Spring Harb. Perspect. Biol.* 6:a023192. doi: 10.1101/cshperspect.a023192

Xu, H., Beasley, M. D., Warren, W. D., Van Der Horst, G. T., and Mckay, M. J. (2005). Absence of mouse REC8 cohesin promotes synapsis of sister chromatids in meiosis. *Dev. Cell* 8, 949–961. doi: 10.1016/j.devcel.2005.03.018

Zedek, F., and Bures, P. (2018). Holocentric chromosomes: from tolerance to fragmentation to colonization of the land. *Ann. Bot.* 121, 9–16. doi: 10.1093/aob/mcx118

Zickler, D., and Kleckner, N. (2015). Recombination, Pairing, and synapsis of homologs during meiosis. *Cold Spring Harb. Perspect. Biol.* 7. doi: 10.1101/cshperspect.a016626

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Declaration

Erklärung zur Dissertation

gemäß der Promotionsordnung vom 12. März 2020

Diese Erklärung muss in der Dissertation enthalten sein.
(This version must be included in the doctoral thesis)

„Hiermit versichere ich an Eides statt, dass ich die vorliegende Dissertation selbstständig und ohne die Benutzung anderer als der angegebenen Hilfsmittel und Literatur angefertigt habe. Alle Stellen, die wörtlich oder sinngemäß aus veröffentlichten und nicht veröffentlichten Werken dem Wortlaut oder dem Sinn nach entnommen wurden, sind als solche kenntlich gemacht. Ich versichere an Eides statt, dass diese Dissertation noch keiner anderen Fakultät oder Universität zur Prüfung vorgelegen hat; dass sie - abgesehen von unten angegebenen Teilpublikationen und eingebundenen Artikeln und Manuskripten - noch nicht veröffentlicht worden ist sowie, dass ich eine Veröffentlichung der Dissertation vor Abschluss der Promotion nicht ohne Genehmigung des Promotionsausschusses vornehmen werde. Die Bestimmungen dieser Ordnung sind mir bekannt. Darüber hinaus erkläre ich hiermit, dass ich die Ordnung zur Sicherung guter wissenschaftlicher Praxis und zum Umgang mit wissenschaftlichem Fehlverhalten der Universität zu Köln gelesen und sie bei der Durchführung der Dissertation zugrundeliegenden Arbeiten und der schriftlich verfassten Dissertation beachtet habe und verpflichte mich hiermit, die dort genannten Vorgaben bei allen wissenschaftlichen Tätigkeiten zu beachten und umzusetzen. Ich versichere, dass die eingereichte elektronische Fassung der eingereichten Druckfassung vollständig entspricht."

Teilpublikationen:

Thangavel, Gokilavani, Paulo G. Hofstatter, Raphaël Mercier, and André Marques. "Tracing the evolution of the plant meiotic molecular machinery." Plant Reproduction (2023): 1-23.

Hofstatter, Paulo G., Gokilavani Thangavel, Thomas Lux, Pavel Neumann, Tihana Vondrak, Petr Novak, Meng Zhang et al. "Repeat-based holocentromeres influence genome architecture and karyotype evolution." Cell 185, no. 17 (2022): 3153-3168.

Hofstatter, Paulo G., Gokilavani Thangavel, Marco Castellani, and André Marques. "Meiosis progression and recombination in holocentric plants: What is known?." Frontiers in Plant Science 12 (2021): 658296.

13.04.2023, Gokilavani Thangavel T. Gokilavani

Datum, Name und Unterschrift

Declaration for the doctoral thesis (dissertation)

according to the doctoral regulations published 12th March 2020

Non-official English translation of the "Erklärung zur Dissertation"
(The German version must be included in the doctoral thesis)

"I hereby declare that I have completed the present dissertation independently and without the use of any aids or literature other than those referred to. All passages that have been taken, either literally or in sense, from published and unpublished works, are marked as such. I declare that this dissertation has not been submitted to any other faculty or university; that - apart from the partial publications and included articles and manuscripts listed below - it has not yet been published, and that I will not publish the dissertation before completing my doctorate without the permission of the PhD Committee. I am aware of the terms of the doctoral regulations. In addition, I hereby declare that I am aware of the "Regulations for Safeguarding Good Scientific Practice and Dealing with Scientific Misconduct" of the University of Cologne, and that I have observed them during the work on the thesis project and the written doctoral thesis. I hereby commit myself to observe and implement the guidelines mentioned there in all scientific activities. I assure that the submitted electronic version is identical to the submitted printed version".

Partial publications of the thesis:

Thangavel, Gokilavani, Paulo G. Hofstatter, Raphaël Mercier, and André Marques. "Tracing the evolution of the plant meiotic molecular machinery." Plant Reproduction (2023): 1-23.

Hofstatter, Paulo G., Gokilavani Thangavel, Thomas Lux, Pavel Neumann, Tihana Vondrak, Petr Novak, Meng Zhang et al. "Repeat-based holocentromeres influence genome architecture and karyotype evolution." Cell 185, no. 17 (2022): 3153-3168.

Hofstatter, Paulo G., Gokilavani Thangavel, Marco Castellani, and André Marques. "Meiosis progression and recombination in holocentric plants: What is known?." Frontiers in Plant Science 12 (2021): 658296.

13.04.2023, Gokilavani Thangavel T. Gokilavani

Date, name, and signature

Erklärung zum Gesuch um Zulassung zur Promotion

gemäß der Promotionsordnung vom 12. März 2020

1. Zugänglichkeit von Daten und Materialien

Die Dissertation beinhaltet die Gewinnung von Primärdaten oder die Analyse solcher Daten oder die Reproduzierbarkeit der in der Dissertation dargestellten Ergebnisse setzt die Verfügbarkeit von Datenanalysen, Versuchsprotokollen oder Probenmaterial voraus.

☐ Trifft nicht zu

☒ Trifft zu.
In der Dissertation ist dargelegt wie diese Daten und Materialien gesichert und zugänglich sind (entsprechend den Vorgaben des Fachgebiets beziehungsweise der Betreuerin oder des Betreuers).

2. Frühere Promotionsverfahren

Ich habe bereits einen Doktortitel erworben oder ehrenhalber verliehen bekommen.

Oder: Für mich ist an einer anderen Fakultät oder Universität ein Promotionsverfahren eröffnet worden, aber noch nicht abgeschlossen.

Oder: Ich bin in einem Promotionsverfahren gescheitert.

☒ Trifft nicht zu

☐ Zutreffend
Erläuterung:

3. Straftat

Ich bin nicht zu einer vorsätzlichen Straftat verurteilt worden, bei deren Vorbereitung oder Begehung der Status einer Doktorandin oder eines Doktoranden missbraucht wurde.

Ich versichere, alle Angaben wahrheitsgemäß gemacht zu haben.

13.04.2023	Gokilavani Thangavel	T. Gokilavani
Datum	Name	Unterschrift

Declaration on the application for admission to the doctoral examinations

according to the doctoral regulations published 12th March 2020

1. Accessibility of data and materials

The dissertation involves the acquisition of primary data or the analysis of such data or the reproducibility of the results presented in the dissertation requires the availability of data analyses, experimental protocols or sample material.

☐ not applicable

☒ applicable
I have described in the dissertation how these data and materials are secured and accessible (according to the specifications of the subject area or supervisor).

2. Previous doctoral examinations

I have already obtained a doctorate or been awarded an honorary doctoral degree.

Or: A doctoral examination has been opened at another faculty or university, but not yet completed.

Or: I have failed in a doctoral examination.

☒ not applicable

☐ applicable
explanatory note:

3. Criminal offense

I have not been convicted of a deliberate criminal offence in the preparation or commission of which the status as a doctoral candidate was abused.

I declare that all information provided is truthful.

13.04.2023	Gokilavani Thangavel	T. Gokilavani
Date	Name	Signature