the plant journal

SEB SOCIETY FOR EXPERIMENTAL BIOLOGY

The Plant Journal (2024) doi: 10.1111/tpj.16712

How diverse a monocentric chromosome can be? Repeatome and centromeric organization of *Juncus effusus* (Juncaceae)

Yhanndra Dias[1,2,†], Yennifer Mata-Sucre[1,3,†], Gokilavani Thangavel[3], Lucas Costa[1], Mariana Báez[1,4], Andreas Houben[2], André Marques[3,*] and Andrea Pedrosa-Harand[1,*]

[1]*Laboratório de Citogenética e Evolução Vegetal, Departamento de Botânica, Centro de Biociências, Universidade Federal de Pernambuco, Recife, Pernambuco 50670-901, Brazil,*

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

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

[4]*Plant Breeding Department, University of Bonn, Bonn, Germany*

Received 20 September 2023; revised 19 February 2024; accepted 28 February 2024.
*For correspondence (e-mail amarques@mpipz.mpg.de; andrea.harand@ufpe.br).
†These authors contributed equally.

SUMMARY

Juncus **is the largest genus of Juncaceae and was considered holocentric for a long time. Recent findings, however, indicated that 11 species from different clades of the genus have monocentric chromosomes. Thus, the** ***Juncus*** **centromere organization and evolution need to be reassessed. We aimed to investigate the major repetitive DNA sequences of two accessions of** ***Juncus effusus*** **and its centromeric structure by employing whole-genome analyses, fluorescent** ***in situ*** **hybridization, CENH3 immunodetection, and chromatin immunoprecipitation sequencing. We showed that the repetitive fraction of the small** ***J. effusus*** **genome (~270 Mbp/1C) is mainly composed of Class I and Class II transposable elements (TEs) and satellite DNAs. Three identified satellite DNA families were mainly (peri)centromeric, with two being associated with the centromeric protein CENH3, but not strictly centromeric. Two types of centromere organization were discerned in** ***J. effusus***: **type 1 was characterized by a single CENH3 domain enriched with JefSAT1-155 or JefSAT2-180, whereas type 2 showed multiple CENH3 domains interrupted by other satellites, TEs or genes. Furthermore, while type 1 centromeres showed a higher degree of satellite identity along the array, type 2 centromeres had less homogenized arrays along the multiple CENH3 domains per chromosome. Although the analyses confirmed the monocentric organization of** ***J. effusus*** **chromosomes, our data indicate a more dynamic arrangement of** ***J. effusus*** **centromeres than observed for other plant species, suggesting it may constitute a transient state between mono- and holocentricity.**

Keywords: centromere organization, centromeric sequences, ChIP-sequencing, chromosome evolution, repetitive elements, rushes, satellite DNAs.

INTRODUCTION

Centromeres are essential for genome stability through the mediation of chromosome segregation in mitosis and meiosis (Cuacos et al., 2015; McKinley & Cheeseman, 2016; Schubert et al., 2020). Chromosomes that have a single size-restricted centromere, visible as a primary constriction, are called monocentric, while holocentric chromosomes have centromeric activity distributed along a substantial portion of the chromosome axis (Melters et al., 2012). Holocentric chromosomes also differ in mitotic and meiotic organization and behavior (Marques et al., 2016; Schubert et al., 2020). Most studied organisms have monocentric chromosomes, but holocentricity originated independently at least 14 times throughout the evolution of eukaryotes, nine of them in animals and at least five times in plants (Escudero et al., 2016; Melters et al., 2012; Senaratne et al., 2022). However, identifying the centromere type can be challenging, particularly in groups with small chromosomes, and different approaches were applied to demonstrate monocentricity in species previously considered to be holocentric (Báez et al., 2020; Guerra et al., 2019; Král et al., 2008). In addition, when a chromosome type is defined for a single species only, it may be incorrectly extrapolated to related species (Heckmann et al., 2013; Melters et al., 2012; Šmarda et al., 2014).

The mechanism behind the transition from mono- to holocentricity is unknown, but monocentricity is usually considered the ancestral state (Melters et al., 2012; Nagaki et al., 2005; Neumann et al., 2012; Schubert et al., 2020). Nevertheless, some authors proposed the opposite direction of evolution (Escudero et al., 2016; Moore et al., 1997). The differences in genome organization and composition between these two types of centromeres show how complex and impactful the transition from mono- to holocentricity can be. In insects, the transition to holocentricity was associated with the loss of the centromere-specific histone H3 variant (CENH3), an otherwise highly conserved centromeric protein (Drinnenberg et al., 2014). In the genus *Cuscuta*, the centromere-type transition had a significant impact on the genomes of these species, such as changes in epigenetic marks, in the composition of repetitive sequences and kinetochore proteins (Neumann et al., 2021, 2023). Another type of complex centromeric organization based on repetitive DNA was found in animals and a few species of legumes, which present metapolycentric centromeres. This centromere variant is characterized by an extended primary constriction containing multiple CENH3 domains and satellites, denoting a possible intermediate state between mono- and holocentric chromosomes (Gržan et al., 2020; Huang et al., 2016; Macas et al., 2023).

Repetitive DNAs are classified as tandem (mainly satellite DNAs) or dispersed (mainly transposable elements [TEs]) repeats (Neumann et al., 2011, 2019). Satellite DNAs (satDNAs) are usually arranged in long and, frequently, head-to-tail arrays. They are a substantial part of many genomes and have been widely reported as a core component of plant and animal centromeres (Aldrup-MacDonald & Sullivan, 2014; Dong et al., 1998; Sullivan & Sullivan, 2020; Wang et al., 2022). In plants, centromeric satDNAs were identified in rice (Cheng et al., 2002), beans (Iwata et al., 2013), *Vicia faba* (Ávila Robledillo et al., 2018), *Prionium serratum* (Báez et al., 2020), pea (Macas et al., 2023) and many other monocentric species. In holocentrics, centromeric satellite DNAs were identified in the genus *Rhynchospora* from Cyperaceae (Marques et al., 2015; Ribeiro et al., 2017) and in *Chionographis japonica* from Melanthiaceae (Kuo et al., 2023), while in other species, such as *Luzula elegans* (Juncaceae) and holocentric *Cuscuta* species (Convolvulaceae), satDNAs were found in non-centromeric terminal chromosome regions (Heckmann et al., 2013; Neumann et al., 2021).

Transposable elements are, in general, the most abundant repeat type in plant genomes and are largely responsible for genome size variations in many species (Ambrožová et al., 2011; Macas et al., 2015; Sader et al., 2021; Schnable et al., 2009). They can exhibit a dispersed distribution in larger genomes and often a pericentromeric distribution in species with small genomes (Báez et al., 2019, 2020; De Souza et al., 2018; Ibiapino et al., 2022; Ribeiro et al., 2020; Sader et al., 2021). In *Cuscuta*, the variation in genome size seems to be mainly driven by the differential accumulation of LTR retrotransposons and satDNAs, especially in the large genomes of some monocentric species (Ibiapino et al., 2022; Neumann et al., 2021). In contrast, small genomes tend to have a smaller fraction of TEs (Báez et al., 2020; Ribeiro et al., 2017; Sader et al., 2021). Class I LTR-retroelements are the most abundant TEs, but the prevalence and organization of Ty1/copia and Ty3/gypsy varies in a given species or even among related plant species (Báez et al., 2020; Marques et al., 2015; Ribeiro et al., 2020; Sader et al., 2021; Van-Lume et al., 2019). Ty1/copia elements have been described as having a dispersed distribution across the genome (De Souza et al., 2018; Heckmann et al., 2013), whereas the Ty3/gypsy tend to show restricted accumulation in certain genomic regions, such as found in some lineages from the Chromovirus clade which are found at (peri)centromeres of mono- and holocentric chromosomes (Marques et al., 2015; Neumann et al., 2021; Van-Lume et al., 2019).

Only few plant genera harbor both mono- and holocentric species among their representatives, such as *Drosera* and *Cuscuta* (Junichi et al., 2011; Neumann et al., 2021; Pazy & Plitmann, 1994; Shirakawa et al., 2011). Until recently, the cyperid clade, containing Juncaceae, Cyperaceae and Thurniaceae families, was considered as exclusively possessing holocentric chromosomes (Greilhuber, 1995). However, recent findings have challenged this view and showed that *P. serratum* (Thurniaceae; Báez et al., 2020) and some species of *Juncus* (*J. effusus*, *J. marginatus*, *J. microcephalus* and *J. tenuis*; Juncaceae) are, in fact, monocentric (Guerra et al., 2019). More recently, CENH3 immunostaining confirmed monocentricity for six other *Juncus* species, suggesting a possible synapomorphy for the genus (Mata-Sucre et al., 2023). Thus, most likely, at least two independent transition events to holocentricity occurred within the cyperid clade, one in Cyperaceae and one in Juncaceae (Figure S1). *Juncus* is the largest genus in Juncaceae, with about 332 species, followed by *Luzula*, the holocentric sister genus, with about 124 species (POWO, 2023; Roalson, 2005). Both show a wide range of chromosome numbers, with *Juncus* varying from $2n = 18$ to $2n = 170$, being $2n = 40$ the modal count (Záveská Drábková, 2013). *J. effusus* (soft rush; $2n = 42$) is a species with an almost cosmopolitan distribution, having biotechnological importance for phytoremediation of contaminated water and medicinal properties (Liao et al., 2011; Vymazal, 2014). Two recently performed genome assemblies of *J. effusus* and *J. inflexus* confirmed their monocentricity (Hofstatter et al., 2022; Planta et al., 2022). However, the repeat composition and centromeric organization were not investigated in detail.

Here, we performed a detailed analysis of the main repetitive DNA elements and elucidated the sequence and

chromatin composition of centromeres in *J. effusus*. Our results allowed us to uncover a complex centromere landscape, revealing two different types of centromere organization based on the extension and continuity of CENH3 domains, as well as the diversity and organization of centromeric repetitive DNA sequences. We discuss our results in light of a dynamic centromere organization possibly in a transient state between mono- and holocentricity.

RESULTS

The repetitive DNA composition of *J. effusus* accessions

The nuclear DNA content of *J. effusus* subsp. *effusus* was estimated to be 1C = 270.42 Mbp (n = 21; Table S1), similar to the genome size of 1C = 271 Mbp for *J. effusus* var. *spiralis* (n = 21; Hofstatter et al., 2022). A total of 250 and 120 clusters with abundance above 0.01% were generated by RepeatExplorer2 after a comparative analysis of unassembled short reads, indicating that the repetitive DNA sequences of *J. effusus* subsp. *effusus* and *J. effusus* var. *spiralis* comprised 23.84 and 24.97% of the genomes, respectively (Table 1; Figure S2). The genome characterization by low-coverage sequencing of both genotypes was similar in terms of proportion and number of repeat families. Nevertheless, the proportion of 35S ribosomal DNA (rDNA) repeats was around three times higher in *J. effusus* subsp. *effusus*. Considering that a whole-genome characterization of *J. effusus* var. *spiralis* is available in Hofstatter et al. (2022) and a more comprehensive approach was used for its centromere characterization below, we based the following repeat characterization on the genome of *J. effusus* subsp. *effusus*.

Few repeats remained unclassified (1.39%), and among those annotated, the majority were LTR-retrotransposons (7.52%), with seven lineages of Ty1/copia and two of Ty3/gypsy (Table 1). The Ty1/copia superfamily was dominant, making up about 5.59% of the genome, with Angela being the most abundant lineage (3.57%). Ty3/gypsy contributed to 1.93% of this genome, where 1.91% was from the non-Chromovirus Athila lineage (Table 1). The CRM lineage from the Chromovirus clade, usually associated with centromeres, was the second most abundant Ty3/gypsy lineage, but represented only 0.03% of *J. effusus* genome. Class II elements and DNA transposons made up a significant fraction of the genome (1.30%; Table 1), with TIR/MuDR_MUTATOR being the most abundant.

Satellite DNA was a highly abundant repetitive DNA type, making up 9.28% of the genome (Table 1). Ten clusters corresponded to this type of sequence, but only the most abundant one represented a high-confidence repeat (CL 1, Table S2). Two clusters, despite being classified as low-confidence putative satellites, were not in a tandem organization in the genome assembly, so they were not considered satDNA (Figure S3). CL10 (5235 bp) showed similarity to LTR elements of the Ty1/copia-Bianca lineage (7.45%), with a dispersed distribution in the genome and with detection of protein domains within it (Figure S3a). CL18 (1595 bp) showed similarity with plastid sequences (60%) and it was therefore considered contamination (Figure S3b). Based on sequence similarity, the satDNA clusters were grouped into six families, five of them distributed into two superfamilies, SF1 and SF2 (Table 2; Figure S4a).

The most abundant satellite in the genome (JefSAT1, 6.25%) showed monomers of 155 (CL1) and 154 (CL30) bp, corresponding to two JefSAT1 subfamilies with 91% identity along its length (Table 2; Figure S4a). Two other satDNA families (JefSAT3 and JefSAT4, both 364 bp long) were grouped into superfamily SF1 (1.48% abundance), with 75% identity between them (Table 2; Figure S4a). The remaining three satDNA families (JefSAT2, 5 and 6), with monomers varying between 122 and 180 bp, were grouped into SF2 (1.53%), with identities between 51 and 80% (Table 2; Figure S4a). Other tandem repeats, such as 5S and 35S rDNA, represented together 4.35% of the genome (Table 1).

Chromosome mapping indicates a diverse distribution of repeats

To investigate the chromosomal distribution of *J. effusus* repeats, the three most abundant satellite families (JefSAT1, 2 and 3), belonging to different superfamilies, as well as the most abundant Ty1/copia (Angela) and Ty3/gypsy (Athila) lineages, were selected for *in silico* and fluorescence *in situ* hybridization (FISH) mapping (Table 2; Table S3; Figure S4b). Despite the small chromosome size, which impaired the correct assignment of repeats to a particular chromosomal domain, dot-like hybridization patterns of satDNAs covering the primary chromosome constrictions of several chromosomes suggest a (peri)centromeric location (Figure 1). In some cases, satellite signals covered the pericentromeric extending to the adjacent proximal regions of chromosomes, as observed for JefSAT1-155 and JefSAT2-180 (Figure 1a,b). JefSAT1-155 labeled large blocks of decreasing intensity in three chromosome pairs, and weaker signals on 18 chromosome pairs, which represented most of the chromosome complement (Figure 1a). Low stringency hybridization showed JefSAT2-180 signals in 15 chromosome pairs (probably representing the distribution of all satellite families of SF1), with strong signals in one pair (Figure 1b). JefSAT3-365 showed weaker signals in most chromosomes (Figure 1c).

The Ty1/copia/Angela probe, in contrast to the pattern observed for the satellites, showed scattered signals along the chromosomes, with small blocks in distal regions on some chromosomes (Figure 1d). The Ty3/gyspy/Athila retroelement also showed some scattered signal, but was

Table 1 Composition and relative abundance of repetitive DNA in the *Juncus effusus* genomes characterized by low-coverage sequencing and Repeatexplorer pipeline

Repetitive sequence	Superfamily	Lineage	*Juncus effusus* subsp. *effusus* (%)	*Juncus effusus* var. *spiralis* (%)
LTR retrotransposons			7.52	8.49
	Ty1/copia		5.59	6.60
		Angela	3.57	5.08
		Bianca	0.63	0.82
		SIRE	0.51	0.26
		Alesia	0.36	-
		Tork	0.3	0.40
		Ale	0.18	0.03
		Ivana	0.04	0.01
	Ty3/gypsy		1.93	1.89
		Non-chromovirus/OTA/Athila	1.91	1.87
		Chromovirus/CRM	0.03	0.02
DNA transposon			1.3	1.00
35S rDNA			4.29	1.22
5S rDNA			0.06	0.27
Satellite DNA			9.28	10.75
Unclassified			1.39	3.24
Total			23.84	24.97

Table 2 Satellite DNAs identified in the *Juncus effusus* subsp. *effusus* genome

Satellite family	Superfamily	Cluster	Monomere size (bp)	Abundance (%)	Confidence TAREAN
JefSAT1	-	1, 30	154, 155	6.25	High, Low
JefSAT2	SF-2	6, 21	122, 180	1.53	Low
JefSAT3	SF-1	5	364	1.48	Low
JefSAT4	SF-1	44	364	0.09	Low
JefSAT5	SF-2	66	174	0.04	Low
JefSAT6	SF-2	78	179	0.03	Low
Total				9.28	

strongly enriched in one pair of chromosomes (Figure 1e). These patterns were analyzed *in silico* to obtain a better resolution of the elements using full annotation and thus associating the sequences to centromeres and/or pericentromeres in the genome (Supporting Information, Data S1; Figure 2). We confirmed the high abundance of JeSAT1-155 satDNA in the genome of *J. effusus* with a distribution in (peri)centromeric regions. JeSAT1-155 covered an extensive region along chromosome 3, as observed by FISH analyses (Figure 1a). On the other hand, at least three chromosome pairs (e.g. 3, 16, 18; Figure 2) showed a high enrichment of the Ty3/Athila retroelement, distinct from what was observed by FISH (Figure 1e).

Centromere organization in *J. effusus*

To confirm monocentricity and to identify the position of active centromeres, a *Juncus* CENH3-specific antibody (anti-*Je*CENH3) was generated. CENH3 immunostaining showed monocentromere-typical dot-like signals in interphase nuclei, compatible with the number of (pro) metaphase chromosomes in both accessions (Figure 3; Figure S5). Subsequent FISH revealed a partial overlap between JefSAT1-155 and CENH3 signals (Figure 3a,b). Strong JefSAT1-155 signals expanded beyond the CENH3-containing chromatin, and some centromeres did not reveal JefSAT1-155 signals (Figure 3b). The observed interaction of CENH3-labeled regions with alpha-tubulin fibers at metaphase demonstrated the centromere specificity of the CENH3 antibody (Figure 3c,d).

Centromere sequence composition differs between chromosomes

To further investigate the centromere regions, we used the assembled and available genome of *J. effusus* var. *spiralis* and carried out chromatin immunoprecipitation sequencing (ChIP-seq) using antibodies against *Je*CENH3, the heterochromatin-specific histone mark H3K9me2 and euchromatin-specific histone modification H3K4me3. The *Juncus* genome had one or more ChIP enrichment peaks for CENH3 per chromosome, with different intensities

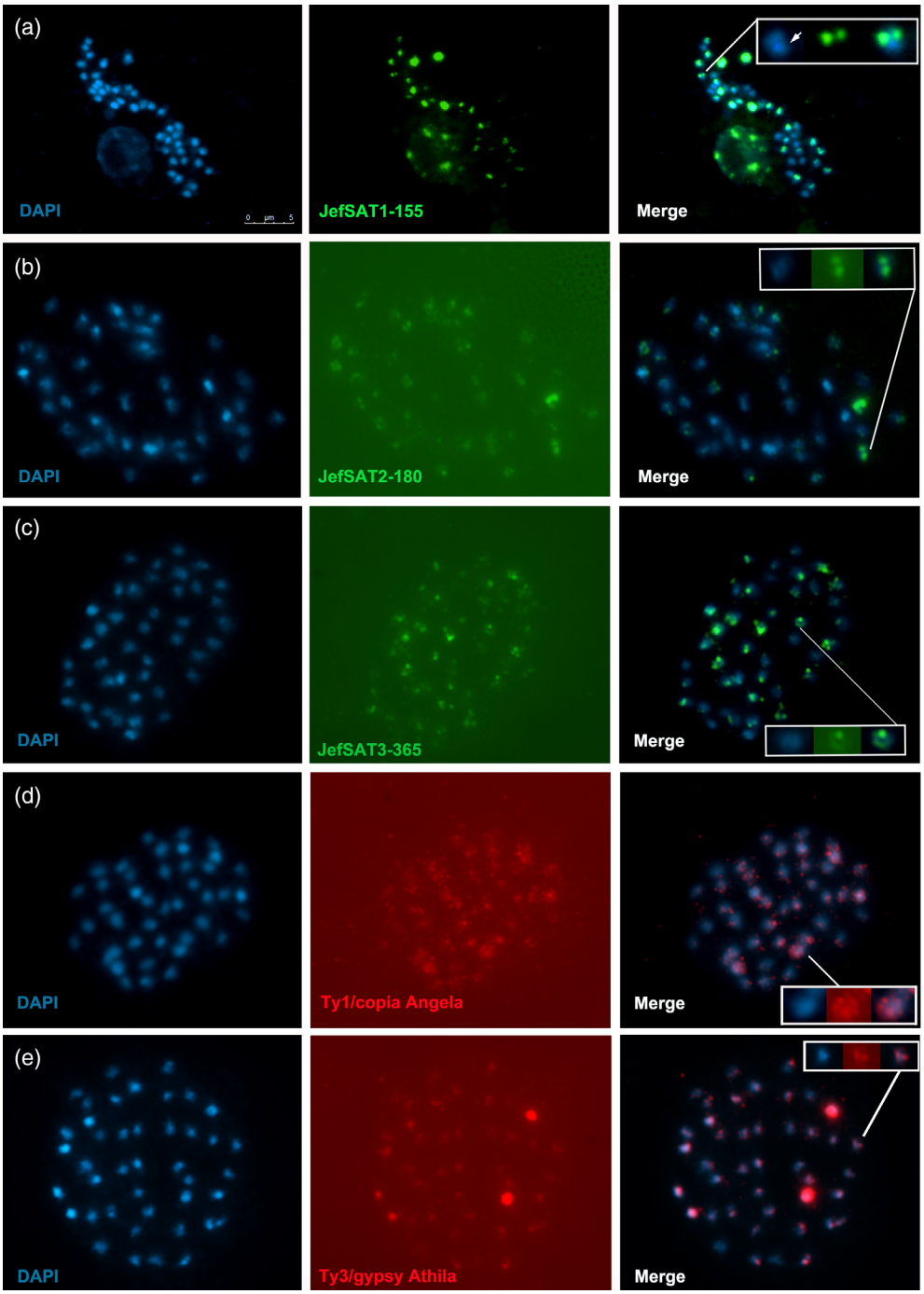

Figure 1. Mitotic metaphases of *Juncus effusus* after FISH showing the (peri)centromeric distribution of three satellites (green) and scattered, non-uniform patterns of two retroelements (red).
(a) JefSAT1-155. (b) JefSAT2-180. (c) JefSAT3-365. (d) Ty1/copia Angela. (e) Ty3/gypsy Athila. Arrow indicates a primary constriction, while insets show dot-like signals (or scattered), in (d) in amplified chromosomes. Bar size = 5 μm.

(Figure 4). The location and size of centromeric regions on each chromosome were estimated based on the positions of the ChIP-seq peaks, for chromosomes with more than one peak, the centromeric region was determined through the distance of the first and last CENH3 enrichment peaks. Thus, centromere size ranged from 0.37 Mbp on Chr. 8 (which is 11.33 Mbp in size) to ~1 Mbp on chromosome 16 (9.7 Mbp in size, Figure 4; Figure S6).

ChIP-seq of histone marks confirmed the enrichment of H3K9me2 at pericentromeric regions, while H3K4me3

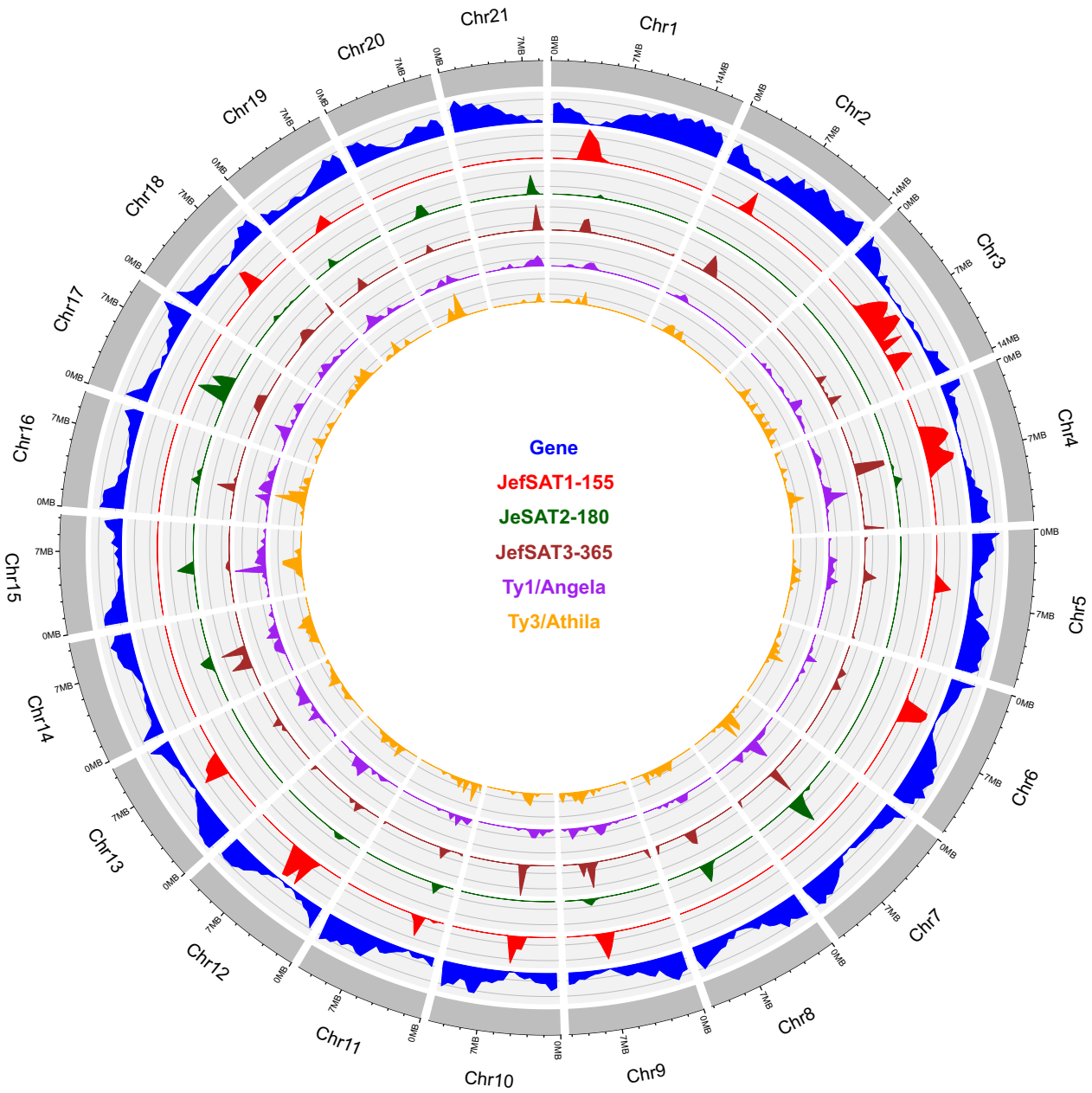

Figure 2. *In silico* distribution of the major DNA repeats in *Juncus effusus* pseudomolecules (gray bars in the outer circle).
JefSAT1 (red), JefSAT2 (dark green) and JefSAT3 (dark red) are interspersed in different proportions at different centromere regions (CENH3 in green). The graph also shows the distribution of genes (blue), Ty1/copia/Angela (purple) and Ty3/gypsy/Athila (orange). The height of the peaks represents the relative frequency of a sequence in a given position.

presented a dispersed distribution along chromosomal arms, except on chromosome 3 in which it was restricted to the terminal parts of the chromosome (Figure 4a; Figure S6). Detailed analysis of available DNA methylation data (Hofstatter et al., 2022) revealed an increase in the concentration of CpG methylation toward the centromere core, while CHG methylation typically showed a decrease at the centromere core, especially along JefSAT1 arrays (Figure 4; Figure S6).

To characterize centromere organization and centromeric repeat profile, we performed a detailed comparison of the repeat distribution along the 21 centromeres of the assembled chromosomes using DANTE-LTR (https://github.com/kavonrtep/dante_ltr) and E.D.T.A. pipelines (Ou et al., 2019, see Supporting Information, Data S1). CENH3 regions were highly enriched with sequences of JefSAT1-155 and JefSAT2-180, and low-enriched by JefSAT3-365 in different combinations, with TEs in and around regions of

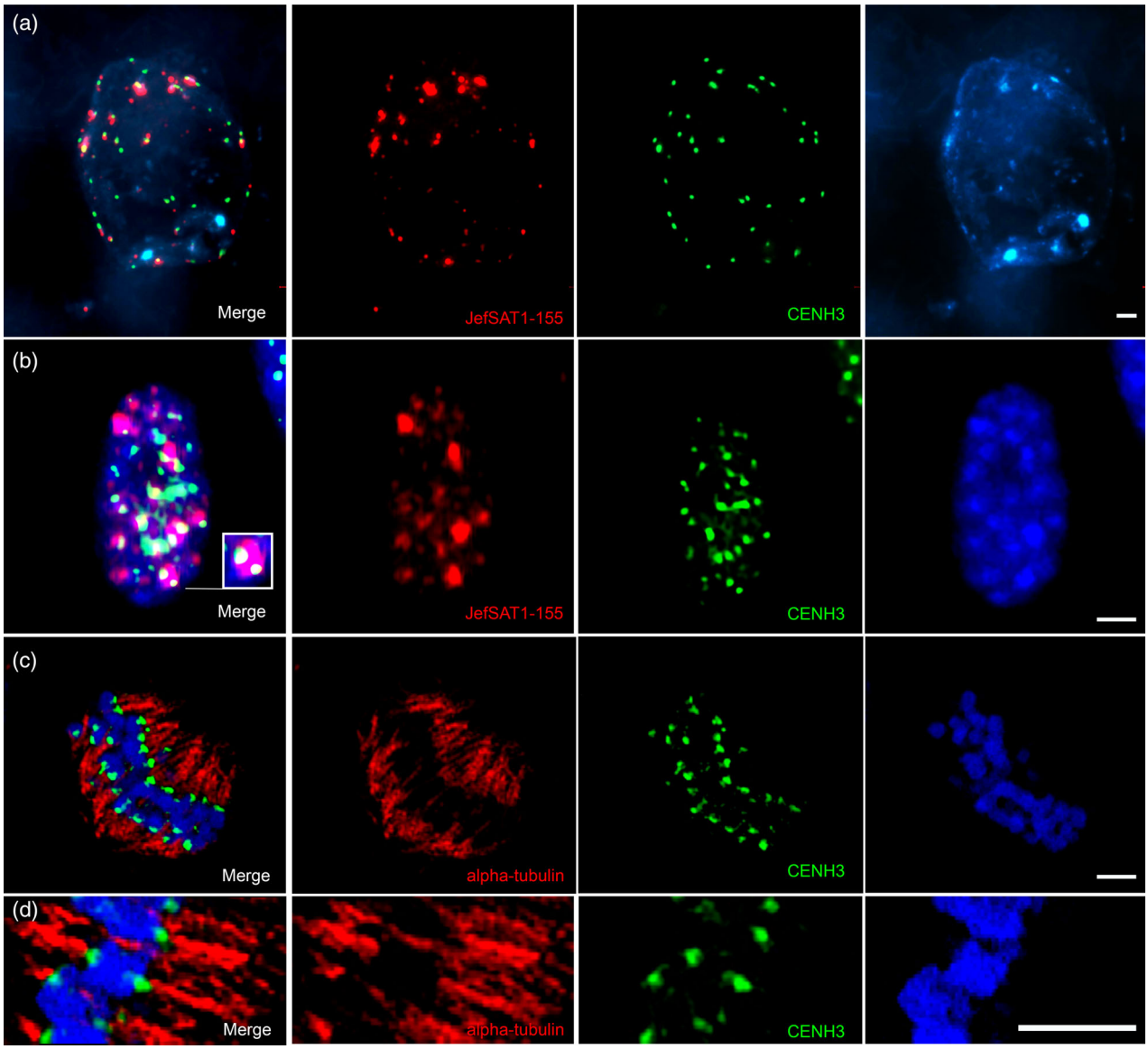

Figure 3. Immunostaining using the CENH3 (green) antibody developed for *Juncus effusus* in interphase nuclei (a) and prometaphase (b) of the same species, followed by FISH with JefSAT1-155 (red).
A partial colocalization of the centromeric protein with the satellite is observed in some centromeres. Inset in (b) shows the JefSAT1-155 signal labeling the whole chromocenter, while CENH3 signal is more restricted.
(c, d) CENH3 (green) interaction with alpha-tubulin (red) in metaphase chromosomes. (d) Bar = 5 μm.

low gene density (Figure 4a; Figures S6 and S7). These satDNAs were enriched within or adjacent to CENH3 domains, although more than one CENH3-binding hotspot was observed, for example, on chromosomes 14 and 21 (Figure 4a). Chromosome 3 presents a distinct pericentromeric configuration, with a large ~9 Mbp low-gene density region mainly populated by long arrays of JefSAT1-155, but a restricted CENH3 enrichment peak (Figure 4a). Centromeres of chromosomes Chr2 and Chr14 present two or three CENH3 peaks that are enriched with JefSAT1-155 or JefSAT2-180, and interrupted by a JefSAT3-365 enrichment between the two CENH3 peaks, respectively (Figure 4b; Figure S6). LTRs belonging to the most abundant lineages Ty3/gypsy-Athila and Ty1/copia-Angela were the most predominant in (peri)centromeric regions (Figure 4b; Figures S6, S8 and S9). Also, several lineages of class II elements were present, especially the non-TIR-Helitron and TIR-Mutator elements. In addition, different class II lineages were also found in a scattered pattern along chromosome arms (Figure 4a; Figures S6, S8 and S9).

Based on the centromere sequence composition and organization among chromosomes the centromeres were classified into two different types (Figure 4b). In type 1 centromeres, which occurs in 14 of the 21 *J. effusus*

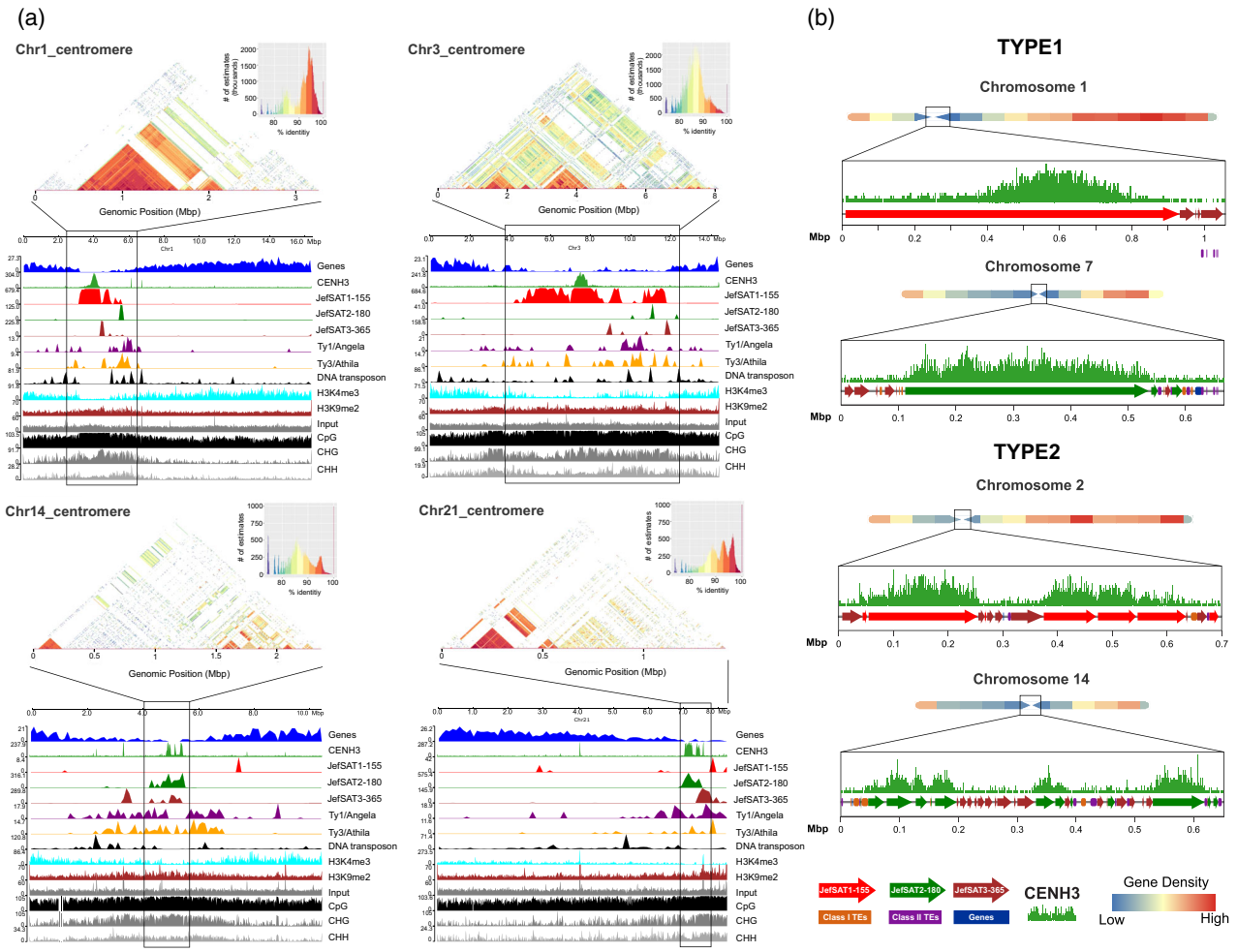

Figure 4. Centromere annotation of different *Juncus effusus* chromosomes.
(a) Pairwise sequence identity of the centromeres of chromosomes 1, 3, 14 and 21 of *Juncus effusus* (top). Each centromere has a mirror organization formed by evolutionarily distinct layers. The size of each centromeric region is defined by the centromeric repeat arrays JefSAT1 (red) or JefSAT2 (dark green) together with CENH3 as $\log_2$ enrichment (RPKM ChIP/input normalized in light green). Sequence identity maps are plotted with ModDotPlot in the same color scale (top right). (Bottom) Distribution of CENH3-interacting sequences, high copy number satellite repeats and retrotransposons along chromosomes. Some satellite sets colocalize with CENH3-enriched, gene-poor, dimethylated sites in H3K9. Ty3/Athila and Ty1/Angela show a sparse distribution in the (peri)centromeric regions.
(b) Two different types of centromere configuration based on the distribution of CENH3, satDNAs and retrotransposons. In type 1 centromeres, one satDNA (either JefSAT1-155 or JefSAT2-180) is present in a long array corresponding to a continuous CENH3 enrichment peak. Type 2 centromeres contain either JefSAT1-155 or JefSAT2-180 on CENH3-rich peak regions, but these regions are interrupted by small low-CENH3 regions populated by JefSAT3-365 arrays, transposable elements and/or genes.

chromosomes, one specific satDNA (either JefSAT1-155 or JefSAT2-180) is present in a long array corresponding to a continuous CENH3 enrichment peak (Figure 4b; Figure S8). While type 2 centromeres, present in 7 of the 21 chromosomes, also contain either JefSAT1-155 or JefSAT2-180 on CENH3-rich peak regions, but these regions are interrupted by low-CENH3 regions (>50 kb long) populated by JefSAT3-365 arrays or TEs (mainly Ty1/Angela, Ty3/Athila and/or non-TIR/Helitrons), and eventually genes (Figure 4b; Figure S9). In addition, the spaces between (type 2) and the flanking regions of the CENH3 core (pericentromeres of both types) are populated by a diversity of TEs in which class II elements stand out (Figures S8 and S9). Remarkably, in contrast to type 1 centromeres, mostly composed of JefSAT1, CHG methylation was frequently not reduced at the core of centromere type 2 peaks, especially along JefSAT2 arrays (Figure 4; Figures S6 and S10), suggesting a higher heterochromatinization of the latter ones.

To further investigate the organization and evolution of these repeat-based centromeres, we generated pairwise sequence identity heatmaps of the satDNA sequences along the centromere and pericentromere, which revealed a relationship between the degree of satellite identity and type 1 and 2 centromeres (Figure 4a; Figure S10). We observed that there are homogenized inner cores of

JefSAT1-155 and JefSAT2-180, where the sequences are highly similar to each other >90% but diverge from the rest of the centromere, where the CENH3 enrichment peaks reside (Figure 4a; Figure S10). However, once these regions start to diverge from each other (<90% similarity) transitions to multiple CENH3 peaks are observed, showing a tendency for the formation of type 2 centromeres (Figure 4a). This confirms that JefSAT1-155 and JefSAT2-180 are canonical centromeric satDNAs on different chromosomes, but not exclusively centromeric, while JefSAT3-365 and TEs, although present in the centromeric region, interrupting the centromeric domain, are not tightly associated with CENH3-rich regions (Figure 4a; Figures S6, S8 and S9). Although present, neither Athila nor Angela elements were enriched in pericentromeric regions, but rather extended over proximal chromosome regions, as observed by FISH (Figures 1 and 4a; Figure S6).

DISCUSSION

J. effusus centromeres differ in size, CENH3 density and repeat arrangement

Satellite DNAs are major components of *J. effusus* centromeres, as in most monocentric species (Hartley & O'Neill, 2019; Melters et al., 2013; Ruiz-Ruano et al., 2018; Šatović-Vukšić & Plohl, 2023). In *J. effusus*, satDNA represents around 10% of the genome and three satDNAs are found within the centromeric region, with two satellites (JefSAT1-155 and JefSAT2-180) being associated, although not exclusively, with CENH3-rich regions. JefSAT3-365 is found mainly in regions not enriched for CENH3, frequently interrupting the centromere core.

Variation in centromere composition has also been observed in some plants, with different centromeric satellite repeats composing these regions or even being absent (repeat-free centromeres; Gong et al., 2012; Liu et al., 2023; Macas et al., 2023). In *Phaseolus vulgaris*, CentPv1 and CentPv2 satellites are present in two distinct chromosomal subsets that evolved independently within the common bean centromeres (Iwata et al., 2013). Similarly, potato centromeres showed chromosome-specific satellite DNAs for six chromosome pairs, while five centromeres did not contain any satDNA (Gong et al., 2012). On the other hand, the monocentric *P. serratum* showed two different subfamilies of centromeric repeats interspersed in different proportions in all centromeres (Báez et al., 2020). A complex centromeric composition was also observed in *Beta* species, where chromosome-specific sets of repeated DNA elements, including satellite DNAs, were present at *Beta* centromeres, with a satDNA present in only one centromere, but exhibiting signals in other chromosomal regions (Gindullis et al., 2001).

Similar to other repeat-based centromeres (Despot-Slade et al., 2021; Hofstatter et al., 2022; Naish et al., 2021), centromeric and pericentromeric regions of *J. effusus* were distinguishable based on CENH3, satellite DNA, TEs and histone methylation distribution in most chromosomes. Nevertheless, the TEs identified in the centromeric regions are predominantly LTR elements belonging to the Ty3/gypsy-Athila and Ty1/copia-Angela lineages and few lineages of Class II, which correspond to the most abundant lineages also present in pericentromeres and along the arms. Although Ty3/gypsy elements tend to extend less toward chromosomal ends than Ty1/copia, they are still abundant along at least half of the chromosomal length of *J. effusus*. In contrast, the enrichment of Athila in *Arabidopsis* pericentromeres is more predominant; and it was shown to be an important driver of centromere evolution (Wlodzimierz et al., 2023).

Satellite DNAs are distributed in numerous clusters along each chromosome, also interrupted by TEs, extending along several Mbp toward chromosomal arms in some cases. Moreover, CENH3 enrichment profiles were non-uniform across the genome. Notably, in type 2 centromeres, CENH3-rich regions appear to be interrupted by CENH3-poor regions associated with either JefSAT3-365 or TEs (especially non-TIR/Helitron) and eventually genes. On the other hand, particularly in type 1 centromeres, the CENH3 core might extend considerably less than the centromeric repeat array, maintaining itself in the most preserved region of the array, as observed in humans (Logsdon et al., 2021) and *Arabidopsis* (Naish et al., 2021). Furthermore, we observed that type 2 centromeres lacking array homogenization showed increased levels of CHG methylation compared to more homogenized type 1 centromeres, which resembles the case as reported for *Arabidopsis* centromeres (Naish et al., 2021; Wlodzimierz et al., 2023). In soybean, widespread centromere repositioning was observed, and satellites seemed important in constraining centromere positions (Liu et al., 2023). These observations suggest that the association of CENH3 with centromeric satellites is influenced by their sequence homogenization, with divergence causing a loss of binding of CENH3 or vice versa, giving rise to an organization in multiple conserved domains along the centromere.

Our results demonstrate a diversity of the centromeric organization of *J. effusus*, in contrast to the typical monocentric organization, with a single satellite repeat containing centromeric core enriched in CENH3, as observed in *Arabidopsis thaliana* (Wang et al., 2022; Wlodzimierz et al., 2023). Centromeres differing in size, CENH3 density and arrangement of centromeric repeats have been described also for maize, where CENH3 domains separated by several megabases are strongly influenced by centromeric and non-centromeric repeats (Wolfgruber et al., 2009). In maize, chromosomes and centromeres are larger than in *Juncus*, but the distance between CENH3 domains does not expand beyond 1–2% of chromosome

length (Wolfgruber et al., 2009), whereas in pea, with metapolycentric chromosomes, centromeres comprise, in average, 46% of chromosome length (Macas et al., 2023). In *Juncus*, the discontinuous expansion of CENH3-rich domains (as seen in type 2) in addition to the presence of chromosome 3 with an extensive heterochromatic pericentromeric region rich in JefSAT1-155 (~7 Mbp), low density of genes, H3K4me3 hypomethylation, and decrease in CHG methylation along JefSAT1-155 arrays, may suggest an early stage of centromere expansion, similar to the observed in maize and pea (Macas et al., 2023; Wolfgruber et al., 2009).

Given the phylogenetic proximity to the holocentric sister genus *Luzula* and the diverse centromeric organization in multiple repeat-based CENH3-rich domains (type 2) found in *J. effusus*, we speculate that this could potentially represent an early intermediate state between mono- and holocentric chromosomes. When more telomere-to-telomere assembled genomes become available, particularly in *Juncus* and other cyperids, it will be possible to assess whether the centromeric organization of *J. effusus* may be common among other monocentrics, and provide further insights about the mechanisms of transition between centromere types.

Genomic composition of *J. effusus* is mainly associated with its small genome size

Chromosome organization in *J. effusus* is determined by its compact genome, but also influenced by its monocentric nature (Hofstatter et al., 2022). The amount of repetitive DNA in this species (~49% for a 1C = 268.95 Mbp) is similar to previously observed (Planta et al., 2022), and larger than monocentric species with similar genome sizes, such as *P. serratum* (26.9% for a 1C = 335 Mbp; Báez et al., 2020), from the same phylogenetic clade (cyperids; Figure S1). On the other hand, holocentric species with large genomes from this clade possess variable repetitive proportions: *L. elegans* (61% for a 1C = 3.81 Gbp genome; Heckmann et al., 2013) and *Rhynchospora pubera* (41.16% for a 1C = 1.61 Gbp; Marques et al., 2015).

In *J. effusus* genome, ~25% of all repeats are TEs, with the majority being LTR-Ty1/copia and non-TIR-Helitron transposons, distributed especially in type 2 centromeric regions. The Ty1/copia-Angela lineage comprised more than half of all annotated LTR retrotransposons of the *J. effusus* genome, a feature shared with other *Juncus* species (Mata-Sucre et al., 2023) and *L. elegans*, from *Juncus* sister genus *Luzula*. In this holocentric species, rich in repetitive elements, Angela comprised more than half of all repeats (33% of the genome; Heckmann et al., 2013). In other species of cyperids, such as *R. pubera*, Angela is significantly represented in the genome and its centromeric satellite *Tyba* was at least partially amplified as part of a Helitron element (Hofstatter et al., 2022; Marques et al., 2015). Recently, it has been proposed that environmental stress may have contributed to the evolution of holocentric chromosomes in eukaryotes (Márquez-Corro et al., 2018; Zedek et al., 2021). Thus, the significant stress-mediated activity of Ty1-copia-Angela and Helitron sequences may be an important driving force in the dynamics of centromeric organization, thus impacting the evolution of monocentric and holocentric cyperids' genomes.

In terms of chromosomal distribution, Angela exhibits a scattered distribution in the small chromosomes of *J. effusus* and is less enriched in the distal, gene-rich regions. It differs from the uniformly dispersed pattern in the large chromosomes of *L. elegans* (Heckmann et al., 2013). Outside cyperids, in the large monocentric genome of *Hordeum murinum* (1C = ~8 Gbp), a set of Ty1/copia probes, including different lineages from the Angela clade, also showed a uniform distribution on chromosomes, except in the centromeric region and nucleolus-organizing regions (Ourari et al., 2020), indicating that genome size also has a major influence on genome organization.

Generally, monocentric species with small genomes show a tendency for repeats to be accumulated at (peri) centromeric and, eventually, subtelomeric, chromosome regions (Báez et al., 2020; Naish et al., 2021; Ribeiro et al., 2020; Sader et al., 2021). *Juncus effusus* satellite DNAs were essentially (peri)centromeric, but only a small fraction of the centromeric-like CRM clade was detected. This situation was similar to *L. elegans*, where the absence of CRM and a small number of Ty3/gypsy elements (~1%) were observed (Heckmann et al., 2013). The CRM clade typically colonizes centromeric regions in monocentric plants, as well as in holocentric species (Jin et al., 2004; Marques et al., 2015; Neumann et al., 2021), although a centromeric distribution is not always observed, as seen in *Eleocharis* (De Souza et al., 2018). The Ty3/gypsy Athila, which is highly enriched in pericentromeric regions in the small *Arabidopsis* chromosomes (Naish et al., 2021; Wlodzimierz et al., 2023), is also scattered in *Juncus* genome, but enriched, together with Angela, in proximal chromosomal regions. Nevertheless, neither Athila nor Angela are strictly (peri)centromeric, as observed for the JefSAT1-155 and JefSAT2-180, suggesting a weak interplay between chromosome domains and the primary DNA sequences. Furthermore, while both *Juncus* and *Luzula* are satDNA-rich genomes, *Juncus* has repeat-based centromeres and *L. elegans* does not (Heckmann et al., 2013). Overall, the genomic composition of *J. effusus* is similar to that of *L. elegans*, with differences in abundance mainly associated with differences in genome size, while the distribution of sequences is influenced by the genome size and its centromeric organization.

EXPERIMENTAL PROCEDURES

Plant material

Samples of *J. effusus* var. *spiralis* (a commercially available genotype with curled foliage) and of a wild, straight *J. effusus* subsp. *effusus* (collected in Biritiba Mirim, SP, Brazil) were grown in the Experimental Garden of the Department of Botany at the Federal University of Pernambuco, Leibniz Institute of Plant Genetics and Crop Plant Research (IPK) and at the greenhouse of the Max Planck Institute (MPIPZ; 16 h daylight, 20°C, >70% humidity). A sample of *J. effusus* subsp. *effusus* (voucher LPF 16950) was deposited in the EAN herbarium (Prof. Jayme Coelho de Moraes, Federal University of Paraíba, Areia, Paraíba, Brazil).

Estimation of genome size

The genome size of *J. effusus* subsp. *effusus* was estimated by flow cytometry. Sample preparation was performed according to Loureiro et al. (2007). Young leaves of *J. effusus* were chopped simultaneously with leaves of *Raphanus sativus* L. cv. Saxa (2C = 1.11 pg, Doležel et al., 1992) in a Petri dish (kept on ice) containing 2 ml of Woody Plant Buffer. The sample was filtered through a 30-μm disposable mesh filter (CellTrics; SYSMEX, Norderstedt, Germany) and 50 μg ml^{-1} propidium iodide (Sigma-Aldrich, St. Louis, MO, USA) was added to the final mixture. Seven replicates were made and samples were measured in a CyFlow Space flow cytometer (SYSMEX) equipped with a green laser (532 nm). Histograms of relative fluorescence were obtained using the software Flomax v.2.3.0. (SYSMEX). Mean fluorescence and coefficient of variation were assessed at half of the fluorescence peak. The absolute DNA content (pg/2C) was calculated by multiplying the ratio of the G1 peaks by the genome size of the internal standard. The genome size of *J. effusus* var. *spiralis* was obtained from Hofstatter et al. (2022).

DNA extraction, genome sequencing and repeat characterization

Young leaves from *J. effusus* subsp. *effusus* were used for genomic DNA extraction, using the DNeasy Plant Mini Kit (QIAGEN, Hilden, Germany). The DNA was quantified on 1% agarose gel and Nanodrop and sequenced using the HiSeq 2500 platform (Illumina, San Diego, CA, USA) at low coverage by BGI Americas Corporation (Cambridge, MA, USA), generating 150-bp paired-end reads (Genbank SRR26090479). Reads from *Juncus effusus* var. *spiralis* were obtained from the NCBI database (SRR17245910).

The characterization of repetitive DNA was performed by a graphic-based clustering using the RepeatExplorer2 pipeline (Novák et al., 2020). A total of 213 648 reads from *J. effusus* subsp. *effusus* and 319 938 reads from *J. effusus* var. *spiralis* were analyzed, corresponding to a coverage of ca. 0.12× and 0.17×, respectively. Reads were filtered by quality with default settings (95% of bases equal to or above the quality cut-off value of 10) and interlaced. Clustering analysis was performed for each species and comparatively for both accessions with default settings of 90% similarity over a 55% minimum sequence overlap. The *Find RT Domains* tool and additional database searches (BLASTx) were used to identify protein domains for repeat annotation. Graph layouts of individual clusters were examined interactively using the SeqGrapheR tool (Novák et al., 2013). All clusters with at least 0.01% of the genome were annotated and manually checked to identify the most abundant families in the genome. The proportion of the main types of repeats was calculated from the amount of reads in the individual annotated clusters to the total reads, after excluding those of putative contamination (similarity to mitochondrial or plastid DNA).

Genome-wide protein domain-based annotation of TEs in *J. effusus* var. *spiralis* was performed using the REXdb Viridiplantae v3.052 database, the DANTE (https://github.com/kavonrtep/dante) and DANTE-LTR (https://github.com/kavonrtep/dante_ltr) tools implemented on the RepeatExplorer server. In addition, the Extensive De-novo TE Annotator (E.D.T.A., https://github.com/oushujun/EDTA, Ou et al., 2019) was used to annotate non-LTR elements after filtering false positives in de novo TE predictions (see Supporting Information, Data S1).

For the identification of satDNAs, the TAREAN tool (Tandem Repeat Analyzer), also implemented in RepeatExplorer2, was used (Novák et al., 2017). This tool classifies satDNAs based on the shapes of the pie charts, connected component index *C* and pairwise completeness index *P*. Clusters with high cutoff values ($P > 0.8$ and $C > 0.8$) are identified as high-confidence satellites. Clusters with less stringent values of $P > 0.4$ and $C > 0.7$ are referred to as low-confidence satellites (Novák et al., 2017). In addition, TideCluster and TideHunter available on the RepeatExplorer platform were used to annotate and masking tandem repeats in the genome assembly. To assess the degree of similarity between the satellites, an alignment using the software Geneious version 9.1 (https://www.geneious.com; Kearse et al., 2012) was performed using default parameters. We classify families and superfamilies of satDNA based on the percentage of identity between the satellites. Sequences with similarities between 50 and 80% were considered to be different satDNAs which belong to the same superfamily, while repeats with similarities between 80 and 95% were considered subfamilies of the same family. Clusters with similarities greater than 95% were considered variants of a single family without subfamilies (adapted from Ruiz-Ruano et al., 2016). Moreover, a dot-plot using the consensus sequences of all identified repeats was performed in DOTTER to grasp the all-to-all similarity (Sonnhammer & Durbin, 1995).

Probe generation and FISH

For chromosome analyses, the satellites JefSAT1-155, JefSAT2-180 and JefSAT3-364, representing the most abundant superfamilies identified, as well as two LTR – retrotransposons Angela and Athila, corresponding to the most abundant elements of the Ty1/copia and Ty3/gypsy superfamilies, respectively, were selected (see "Results" section). Primers were designed with Primer 3 (Untergasser et al., 2012) implemented in Geneious version 9.1 (Biomatters, Inc., Boston, MA, USA). In the case of satDNAs, the consensus sequences were used as targets. When satDNAs were represented by more than one cluster, the most abundant one was selected for this purpose. For LTR-retrotransposons, primers were designed from the integrase domain detected by blast at NCBI (https://www.ncbi.nlm.nih.gov; Marchler-Bauer et al., 2017). SatDNA-specific primers were designed facing outwards, while, for the retrotransposons, primers faced inwards (Table S3).

PCR amplification was performed in 50 μl reactions containing 10 ng of genomic DNA from *J. effusus* subsp. *effusus*, 0.1 mM dNTP, 2 mM $MgCl_2$, 1× PCR buffer, 0.4 μM of each primer (Table S3), 0.4× TBT and a Taq DNA polymerase. The PCR program involved 30 cycles of amplification (1 min at 95°C, 1 min at the annealing temperature and 1 min at 72°C; Table S3). All PCR products were checked by sequencing and labeled by nick translation using Alexa-488 (Jena Bioscience GmbH, Jena, Germany) or Cy3-dUTP (Thermo Scientific, Waltham, MA, USA; Table S3) with DNase I (Thermo Scientific) and DNA polymerase I (Invitrogen,

Boston, MA, USA) or a nick translation kit following manufacturer's instructions (Jena Bioscience).

Mitotic chromosomes from *J. effusus* subsp. *effusus* were prepared using the air-drying method after 2 h of enzymatic digestion with 2% cellulase Onozuka and 20% pectinase (Sigma, St. Louis, MO, USA) (Ribeiro et al., 2017), using pretreated roots with 2 mM 8-hydroxyquinoline for 24 h at 4°C, fixed in ethanol: acetic acid 3:1 (v/v) for 2 h and stored at −20°C. The slides with the highest number of cells in metaphases and spread chromosomes were selected for FISH after screening with 1 μg ml^{-1} DAPI in glycerol 75% and destained using 3:1 ethanol:acetic acid for 30 min and 100% ethanol for 1 h. FISH was performed as described by Pedrosa et al. (2002), except for JefSAT2-180, for which we also performed FISH at a lower stringency (~40%; Leitch et al., 1994), with a mix containing 10% dextran sulfate, 6× SSC, 10 ng probe and washing six times with 6× SSC at room temperature. The slides were counterstained with 2 μg ml^{-1} DAPI in Vectashield (Vector Laboratories, Inc., Newark, CA, USA) mounting medium. The images were captured using a BX61 epifluorescence microscope (Olympus, Hamburg, Germany) equipped with a cooled CCD camera (Orca ER, Hamamatsu; Digital Imaging Systems Ltd., Buckinghamshire, UK), a DM5500 epifluorescence microscope (Leica Microsystems GmbH, Wetzlar, Germany) or a super-resolution spatial structured illumination microscopy (3D-SIM) using a 63×/1.40 Oil Plan-Apochromat objective of an Elyra PS.1 microscope system (Carl Zeiss GmbH, Oberkochen, Germany). Overlapping, brightness and contrast adjustments were performed in Adobe Photoshop® CS3.

Chromatin immunoprecipitation sequencing

The *J. effusus* CENH3 (*Je*CENH3) gene was identified in the assembled genome of *J. effusus* var. *spiralis* (Hofstatter et al., 2022), and the N-terminal amino acid sequence VRTKHFSSRPAGSGRPRKR of the complete *Je*CENH3 protein was used as a peptide to produce polyclonal antibodies in rabbits using the service of LifeTein, LLC (Somerset, NJ, USA). ChIP experiments were performed following Reimer and Turck (2010). *J. effusus* var. *spiralis* 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 into 300 bp fragments. Then, 40 ng of sonicated chromatin was incubated with 2 ng of *Je*CENH3 antibody overnight. In addition, immunoprecipitation experiments were carried out with rabbit anti-H3K4me3 (2 ng; ab8580, Abcam, Cambridge, UK), and mouse anti-H3K9me2 (2 ng, ab1220; Abcam). Recombinant rabbit IgG (Abcam; ab172730) and no-antibody inputs were used as controls. Two experimental replications were performed for all combinations. After overnight incubation of chromatin with antibody, protein beads (anti-mouse: Protein G Sepharose 4 Fast Flow and anti-rabbit: rProtein A Sepharose Fast Flow) were added to the chromatin-antibody mixture. The bound chromatin was finally eluted, de-crosslinked, precipitated, and quality-controlled using the NGS-assay on a FEMTO-pulse (Agilent, Waldbronn, Germany). An Illumina-compatible library was prepared with the Ovation Ultralow V2 DNA-Seq library preparation kit (Tecan Genomics, Männedorf, Switzerland) and single-end 1× 150-bp sequenced on a HiSeq 3000 (Illumina) device. For each library, an average of 20 million reads were obtained.

For ChIP-seq analysis, the raw sequencing reads were trimmed by Cutadapt (Martin, 2011) to remove low-quality nucleotides (with quality score less than 20) and adapters. Trimmed ChIPed 150-bp single-end reads were mapped to *J. effusus* var. *spiralis* reference genome with bowtie2 (Langmead & Salzberg, 2012) with default parameters. All mapping duplicates were removed using the command line *samtools view -q 10 -F 4 -hb -@8 input.bam > output.bam* and only the single best matching read was kept on the final alignment BAM file, thereby preventing mapping biases in highly identical regions. BAM files were converted into BIGWIG coverage tracks using the bamCoverage tool from deeptools using the command line *bamCoverage -b1 ChIP.bam -b2 Control.bam -o output.bw -of bigwig -p 20 --normalizeUsing RPKM --scaleFactorsMethod None --operation log2* (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). Individual repeat annotation were converted to Browser Extensible Data (BED) and used as input track for genome-wide overview with ShinyCircos (Yu et al., 2018). Additionally, files for CENH3, anti-H3K4me3 (Abcam, ab8580) and anti-H3K9me2 (Abcam; ab1220) were converted to BIGWIG format and used as input track for chromatin visualization with pyGenomeTracks (Lopez-Delisle et al., 2021). The assembled genome of *J. effusus* var. *spiralis* used as input was obtained in the following available database https://data.cyverse.org.

To assess the enrichment of repeats associated with CENH3-containing nucleosomes, CENH3-ChIPseq single-end and input reads were filtered for quality using the 'Processing of FASTQ reads' tool implemented in Galaxy-based RepeatExplorer. ChIP-Seq mapper (Novák et al., 2020) was used to map ChIP and input reads into RepeatExplorer contig sequences of repeat elements using the default parameters. Additionally, metaplots for all ChIPseq treatment files were plotted with plotHeatmap available from deeptools to calculate their enrichment on gene bodies, centromeric repeats, and TEs with computeMatrix (Ramírez et al., 2016).

Tandem pairwise sequence identity heatmaps

To generate pairwise sequence identity heatmaps, we extracted the centromere and the respective surrounding regions from all 21 *J. effusus* chromosomes. The extracted regions were subjected to an analysis with ModDotPlot (https://github.com/marbl/ModDotPlot) using default parameters (command: *moddotplot -i INPUT_FASTA_FILE -o output*). ModDotPlot is a novel dot plot visualization tool used to view tandem repeats, similar to StainedGlass. ModDotPlot utilizes modimizers to compute the Jaccard coefficient to estimate sequence identity.

Indirect immunostaining and immuno-FISH

Mitotic preparations for both *J. effusus* accessions were made from root meristems fixed in 4% paraformaldehyde in Tris buffer (10 mM Tris, 10 mM EDTA, 100 mM NaCl, 0.1% Triton, pH 7.5) for 40 min on ice in a vacuum and for another 20 min only on ice. After washing twice in 1× PBS for 10 min, the roots were digested in a cellulase-pectinase (2%/20% v/v) solution containing 1× PBS buffer for 1 h and squashed in 1× PBS. The coverslips were removed in liquid nitrogen and the slides were air-dried and stained in DAPI:Vectashield for slide selection under the epifluorescence microscope. The slides with the highest number of cells in division and spread chromosomes were incubated in 3% (w/v) bovine serum albumin containing 0.1% Triton X-100 in PBS. Immunostaining was performed using the primary antibodies: rabbit anti-JeCENH3 (diluted 1:300) and mouse anti-alpha-tubulin (clone DM 1A; Sigma, diluted 1:200). As secondary antibodies, a Cy3-conjugated anti-rabbit IgG (Dianova, Hamburg, Germany) and a FITC-conjugated anti-mouse Alexa488 antibody (Molecular Probes, Eugene, Oregon, USA) were used in a 1:500 dilution each. Slides were incubated overnight at 4°C, washed three times in 1× PBS and then the secondary antibodies were applied for 2 h at

room temperature. Immuno-FISH was performed following Houben et al. (2007). The immunostained slides were washed with PBS for 15 min, post-fixed in 4% paraformaldehyde in PBS for 5 min, and then hybridized with the satellite JefSAT1-155. Labeling and detection of satDNA repeats were performed as described above.

AUTHOR CONTRIBUTIONS

YD performed low-coverage repeat analyses, FISH experiments and drafted the first version of the manuscript; AM, LC and YM-S performed *in silico* analysis of repeats in the assembled genome; MB analyzed low-coverage repeat data; YM-S performed immuno-FISH experiments; GT and YM-S performed and analyzed ChIP experiments; AH, AM and AP-H discussed the data, provided resources and laboratory structure and supervised the work; AP-H designed the study. All authors read and approved the final version of the manuscript. The authors declare that there are no competing interests.

ACKNOWLEDGEMENTS

We thank Prof. Dr. Marcelo Guerra (UFPE) and Prof. Dr. Leonardo Pessoa Felix (UFPB) for kindly providing the plant material and for the previous study that instigated us to start this work. We thank Dr. Veit Schubert (IPK, Germany) for taking the high-resolution photos in Figure 2d; and Dr. Petr Novák (Institute of Plant Molecular Biology, Czech Republic) for the help with Tidecluster annotation. We thank Bruno Huettel from the Max Planck Genome Center Cologne for the preparation of ChIPseq libraries and sequencing. We thank the ELIXIR-CZ Research Infrastructure Project (LM2015047) for providing computational resources for RepeatExplorer analysis. We thank Christina Philipp for the excellent technical support. This research was financially supported by CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior, finance Code 001; grant number 495995/2020-00 to Y.M.-S.; PROBRAL CAPES/DAAD project number 88881.144086/2017-01 to A.P.-H. and A.H.), CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnologico; grant number 141725/2018-4 to Y.D.) and FACEPE (grant number BFP-0040-2.03/22 to L.C.). A.M. is financially supported by the Max Planck Society, DFG (grant MA 9363/3-1) and by the European Union (European Research Council Starting Grant, HoloRECOMB, grant no. 101114879). Open Access funding enabled and organized by Projekt DEAL.

CONFLICT OF INTEREST

The authors declare that they have no competing interests.

DATA AVAILABILITY STATEMENT

All sequence data used in this work are deposited in the NCBI Sequence Read Archive under BioProject accession numbers SRR17245910 and SRR26090479. Information on primers and CENH3 protein for probe generation are available in the text or in the supplementary information. Additional data related to this study is available from the authors upon request.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article.

Figure S1. Phylogenetic relationships between genera within the cyperid clade confirmed as having holocentric (orange circle) or monocentric (blue circle) chromosomes. Genera without clear centromere information are indicated with a gray circle. Phylogeny was taken from Báez et al. (2020).

Figure S2. Comparative *J. effusus* subsp. *effusus* (Jeffu) and *J. effusus* var. *spiralis* (Jspir) clustering analysis using RepeatExplorer pipeline. The size of the rectangle is proportional to the number of reads in that cluster for each species. The colors of the rectangle correspond to the different repetitive sequence types. Differences between both accessions are mostly associated with the 35S rDNA abundance.

Figure S3. Graphs and alignment of the CL10 and CL18 clusters obtained by the TAREAN tool in Repeat Explorer. (a) 5235 bp CL10 showing similarity to the LTR/Ty1-copia-Bianca element, including the presence of protein domains in its structure and a non-tandem distribution in the genome. (b) The plot of the 1595 bp CL18 cluster shows high similarity to organellar elements and their presence in the chloroplast of *Juncus effusus* (GenBank NC_059754.1).

Figure S4. Satellite DNAs from *Juncus effusus*. (a) Dot plot showing similarities among tandem repeat clusters, forming families and superfamilies. (b) Sequence logo of the main *Juncus effusus* satellite families.

Figure S5. Immunostaining using the *Je*CENH3 (green) antibody developed for *Juncus effusus* in prometaphase of *J. effusus* subsp. *effusus* and *J. effusus* var. *spiralis*, showing 42 CENH3 signals (top cell).

Figure S6. Distribution of CENH3-interacting sequences, high-copy satellite repeats and retrotransposons along *Juncus effusus* var. *spiralis* chromosomes. The centromeric satellite arrays match with the enriched sites of CENH3, gene-poor and H3K9 dimethylated sites. The transposable elements, Ty3/Athila and Ty1/Angela show scattered distribution in the (peri)centromeric regions.

Figure S7. ChIP-seq data for different histone modifications in *Juncus effusus*. (a) Genomic proportions and normalized enrichment in CENH3-ChIPseq of the three main satellite families using ChIP-seq mapper. Metaplots of the enrichment of CENH3 (b), H3K4me3 (c), and H3K9me2 (d) from the start and end of different types of sequences: genes (gray line), TEs (brown) and repeats (green). ChIP-seq signals are shown as $\log_2$ (normalized RPKM ChIP/input).

Figure S8. Chromosomes with a type 1 centromere configuration based on the distribution of CENH3, satDNAs and retrotransposons.

Figure S9. Chromosomes with a type 2 centromere configuration based on the distribution of CENH3, satDNAs and retrotransposons.

Figure S10. Distribution of DNA methylation (CpG, CHG, CHH) along *Juncus effusus* var. *spiralis* centromeres. Reduction of CGH methylation is observed along JefSAT1 tracks, enriched or not for CENH3, and is less evident in JefSAT2 and JefSAT3 arrays and along transposable elements.

Figure S11. Pairwise sequence identity of the 21 centromeres of *Juncus effusus*. Each centromere has a mirror organization consisting of evolutionarily distinct layers. The size of each centromeric region is defined by the centromeric repeat matrices JefSAT1 (purple) and JefSAT2 (orange) along with CENH3 as $\log_2$ enrichment (normalized RPKM ChIP/input in rose). Sequence identity heat maps are plotted with ModDotPlot on the same color scale.

Figure S12. TE elements erroneously annotated in satellite-rich regions.

Figure S13. Example of distribution of full and truncated Class I and Class II transposable elements.

Table S1. Summary of 1C genome size measurement replicates of *Juncus effusus*. CV = coefficient of variation; pg, picograms; Mbp, megabase pair.

Table S2. List of consensus sequences of the satellite DNAs from *Juncus effusus*.

Table S3. List of probes and primer sequences used for repeat amplification in *Juncus effusus*.

Table S4. De novo TE annotation using E.D.T.A. pipeline in the genome of *Juncus effusus* var. *spiralis*.

Table S5. De novo TE annotation of intact Class II elements using E.D.T.A. pipeline in the genome of *Juncus effusus* var. *spiralis*.

Table S6. De novo TE and tandem repeat annotation using DANTE-LTR and TideCluster pipelines in the genome of *Juncus effusus* var. *spiralis*.

Data S1. Results: Full annotation of the repetitive fraction of the *Juncus effusus* var. *spiralis* genome using the Extensive De-novo TE Annotator (E.D.T.A.) and DANTE-LTR tools.

REFERENCES

Aldrup-MacDonald, M.E. & **Sullivan, B.A.** (2014) The past, present, and future of human centromere genomics. *Genes*, **5**, 33–50. Available from: https://doi.org/10.3390/genes5010033

Ambrožová, K., **Mandáková, T.**, **Bureš, P.**, **Neumann, P.**, **Leitch, I.J.**, **Koblížková, A.** ***et al.*** (2011) Diverse retrotransposon families and an AT-rich satellite DNA revealed in giant genomes of *Fritillaria* lilies. *Annals of Botany*, **107**, 255–268. Available from: https://doi.org/10.1093/aob/mcq235

Ávila Robledillo, L., **Koblížková, A.**, **Novák, P.**, **Böttinger, K.**, **Vrbová, I.**, **Neumann, P.** ***et al.*** (2018) Satellite DNA in *Vicia faba* is characterized by remarkable diversity in its sequence composition, association with centromeres, and replication timing. *Scientific Reports*, **8**, 5838. Available from: https://doi.org/10.1038/s41598-018-24196-3

Báez, M., **Kuo, Y.T.**, **Dias, Y.**, **Souza, T.**, **Boudichevskaia, A.**, **Fuchs, J.** ***et al.*** (2020) Analysis of the small chromosomal *Prionium serratum* (Cyperid) demonstrates the importance of a reliable method to differentiate between mono- and holocentricity. *Chromosoma*, **129**, 285–297. Available from: https://doi.org/10.1007/s00412-020-00745-6

Báez, M., **Vaio, M.**, **Dreissig, S.**, **Schubert, V.**, **Houben, A.** & **Pedrosa-Harand, A.** (2019) Together but different: the subgenomes of the bimodal *Eleutherine* karyotypes are differentially organized. *Frontiers in Plant Science*, **10**, 1170. Available from: https://doi.org/10.3389/fpls.2019.01170

Cheng, Z., **Dong, F.**, **Langdon, T.**, **Ouyang, S.**, **Buell, C.R.**, **Gu, M.** ***et al.*** (2002) Functional rice centromeres are marked by a satellite repeat and a centromere-specific retrotransposon. *The Plant Cell*, **14**, 1691–1704. Available from: https://doi.org/10.1105/tpc.003079

Cuacos, M., **Franklin, F.C.** & **Heckmann, S.** (2015) Atypical centromeres in plants—what they can tell us. *Frontiers in Plant Science*, **6**, 913. Available from: https://doi.org/10.3389/fpls.2015.00913

De Souza, T.B., **Chaluvadi, S.R.**, **Johnen, L.**, **Marques, A.**, **González-Elizondo, M.S.**, **Bennetzen, J.L.** ***et al.*** (2018) Analysis of retrotransposon abundance, diversity and distribution in holocentric *Eleocharis* (Cyperaceae) genomes. *Annals of Botany*, **122**, 279–290. Available from: https://doi.org/10.1093/aob/mcy066

Despot-Slade, E., **Mravinac, B.**, **Širca, S.**, **Castagnone-Sereno, P.**, **Plohl, M.** & **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. *Molecular Biology and Evolution*, **38**, 1943–1965. Available from: https://doi.org/10.1093/molbev/msaa336

Doležel, J., **Sgorbati, S.** & **Lucretti, S.** (1992) Comparison of three DNA fluorochromes for flow cytometric estimation of nuclear DNA content in plants. *Physiologia Plantarum*, **85**, 625–631. Available from: https://doi.org/10.1111/j.1399-3054.1992.tb04764.x

Dong, F., **Miller, J.T.**, **Jackson, S.A.**, **Wang, G.L.**, **Ronald, P.C.** & **Jiang, J.** (1998) Rice (*Oryza sativa*) centromeric regions consist of complex DNA. *Proceedings of the National Academy of Sciences of the United States of America*, **95**, 8135–8140. Available from: https://doi.org/10.1073/pnas.95.14.8135

Drinnenberg, I.A., **deYoung, D.**, **Henikoff, S.** & **Malik, H.S.** (2014) Recurrent loss of CENH3 is associated with independent transitions to holocentricity in insects. *eLife*, **3**, e03676. Available from: https://doi.org/10.7554/eLife.03676

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

Gindullis, F., **Desel, C.**, **Galasso, I.** & **Schmidt, T.** (2001) The large-scale organization of the centromeric region in *Beta* species. *Genome Research*, **11**, 253–265. Available from: https://doi.org/10.1101/gr.162301

Gong, Z., **Wu, Y.**, **Koblížková, A.**, **Torres, G.A.**, **Wang, K.**, **Iovene, M.** ***et al.*** (2012) Repeatless and repeat-based centromeres in potato: implications for centromere evolution. *The Plant Cell*, **24**, 3559–3574. Available from: https://doi.org/10.1105/tpc.112.100511

Greilhuber, J. (1995) Chromosomes of the monocotyledons (general aspects). In: Rudall, P.J., Cribb, P.J., Cutler, D.F. & Humphries, C.J. (Eds.) *Monocotyledons: systematics and evolution*. Richmond: Royal Botanic Gardens Kew, pp. 379–414.

Gržan, T., **Despot-Slade, E.**, **Meštrović, N.**, **Plohl, M.** & **Mravinac, B.** (2020) CenH3 distribution reveals extended centromeres in the model beetle *Tribolium castaneum*. *PLoS Genetics*, **16**, e1009115. Available from: https://doi.org/10.1371/journal.pgen.1009115

Guerra, M., **Ribeiro, T.** & **Felix, L.P.** (2019) Monocentric chromosomes in *Juncus* (Juncaceae) and implications for the chromosome evolution of the family. *Botanical Journal of the Linnean Society*, **191**, 475–483. Available from: https://doi.org/10.1093/botlinnean/boz065

Hartley, G. & **O'Neill, R.J.** (2019) Centromere repeats: hidden gems of the genome. *Genes*, **10**, 223. Available from: https://doi.org/10.3390/genes10030223

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. *The Plant Journal*, **73**, 555–565. Available from: https://doi.org/10.1111/tpj.12054

Hofstatter, P.G., **Thangavel, G.**, **Lux, T.**, **Neumann, P.**, **Vondrak, T.**, **Novak, P.** ***et al.*** (2022) Repeat-based holocentromeres influence genome architecture and karyotype evolution. *Cell*, **185**, 3153–3168. Available from: https://doi.org/10.1016/j.cell.2022.06.045

Houben, A., **Schroeder-Reiter, E.**, **Nagaki, K.**, **Nasuda, S.**, **Wanner, G.**, **Murata, M.** ***et al.*** (2007) CENH3 interacts with the centromeric retrotransposon cereba and GC-rich satellites and locates to centromeric substructures in barley. *Chromosoma*, **116**, 275–283. Available from: https://doi.org/10.1007/s00412-007-0102-z

Huang, Y.C., **Lee, C.C.**, **Kao, C.Y.**, **Chang, N.C.**, **Lin, C.C.**, **Shoemaker, D.** ***et al.*** (2016) Evolution of long centromeres in fire ants. *BMC Evolutionary Biology*, **16**, 1–14. Available from: https://doi.org/10.1186/s12862-016-0760-7

Ibiapino, A., **García, M.A.**, **Amorim, B.**, **Báez, M.**, **Costea, M.**, **Stefanović, S.** ***et al.*** (2022) The evolution of cytogenetic traits in *Cuscuta* (Convolvulaceae), the genus with the most diverse chromosomes in angiosperms. *Frontiers in Plant Science*, **13**, 842260. Available from: https://doi.org/10.3389/fpls.2022.842260

Iwata, A., **Tek, A.L.**, **Richard, M.M.**, **Abernathy, B.**, **Fonseca, A.**, **Schmutz, J.** ***et al.*** (2013) Identification and characterization of functional centromeres of the common bean. *The Plant Journal*, **76**, 47–60. Available from: https://doi.org/10.1111/tpj.12269

Jin, W., **Melo, J.R.**, **Nagaki, K.**, **Talbert, P.B.**, **Henikoff, S.**, **Dawe, R.K.** ***et al.*** (2004) Maize centromeres: organization and functional adaptation in the genetic background of oat. *The Plant Cell*, **16**, 571–581. Available from: https://doi.org/10.1105/tpc.018937

Junichi, S., **Nagano, K.** & **Hoshi, Y.** (2011) A chromosome study of two centromere differentiating *Drosera* species, *D. arcturid* and *D. regia*. *Caryologia*, **64**, 453–463. Available from: https://doi.org/10.1080/00087114.2011.10589813

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

Král, J., **Kováč, L.U.**, **Šťáhlavský, F.**, **Lonský, P.** & **L'uptáčik, P.** (2008) The first karyotype study in palpigrades, a primitive order of arachnids (Arachnida: Palpigradi). *Genetica*, **134**, 79–87.

Kuo, Y.T., **Câmara, A.S.**, **Schubert, V.**, **Neumann, P.**, **Macas, J.**, **Melzer, M.** ***et al.*** (2023) Holocentromeres can consist of merely a few megabase-sized satellite arrays. *Nature Communications*, **14**, 3502. Available from: https://doi.org/10.1038/s41467-023-38922-7

Langmead, B. & **Salzberg, S.L.** (2012) Fast gapped-read alignment with bowtie 2. *Nature Methods*, **9**, 357–359. Available from: https://doi.org/10.1038/nmeth.1923

Leitch, A.R., **Schwarzacher, T.**, **Jackson, D.** & **Leitch, I.J.** (1994) *In situ hybridization: a practical guide*. Oxford: Bios Scientific Publishers Ltd.

Liao, Y.J., **Zhai, H.F.**, **Zhang, B.**, **Duan, T.X.** & **Huang, J.M.** (2011) Anxiolytic and sedative effects of dehydroeffusol from *Juncus effusus* in mice. *Planta Medica*, **77**, 416–420. Available from: https://doi.org/10.1055/s-0030-1250517

Liu, Y., **Yi, C.**, **Fan, C.**, **Liu, Q.**, **Liu, S.**, **Shen, L.** ***et al.*** (2023) Pan-centromere reveals widespread centromere repositioning of soybean genomes. *Proceedings of the National Academy of Sciences of the United States of America*, **120**, e2310177120. Available from: https://doi.org/10.1073/pnas.2310177120

Logsdon, G.A., **Vollger, M.R.**, **Hsieh, P.**, **Mao, Y.**, **Liskovykh, M.A.**, **Koren, S.** ***et al.*** (2021) The structure, function and evolution of a complete human chromosome 8. *Nature*, **593**, 101–107. Available from: https://doi.org/10.1038/s41586-021-03420-7

Lopez-Delisle, L., **Rabbani, L.**, **Wolff, J.**, **Bhardwaj, V.**, **Backofen, R.**, **Grüning, B.** ***et al.*** (2021) pyGenomeTracks: reproducible plots for multivariate genomic data sets. *Bioinformatics*, **37**, 422–423. Available from: https://doi.org/10.1093/bioinformatics/btaa692

Loureiro, J., **Rodriguez, E.**, **Doležel, J.** & **Santos, C.** (2007) Two new nuclear isolation buffers for plant DNA flow cytometry: a test with 37 species. *Annals of Botany*, **100**, 875–888. Available from: https://doi.org/10.1093/aob/mcm152

Macas, J., **Ávila Robledillo, L.**, **Kreplak, J.**, **Novák, P.**, **Koblížková, A.**, **Vrbová, I.** ***et al.*** (2023) Assembly of the 81.6 Mb centromere of pea chromosome 6 elucidates the structure and evolution of metapolycentric chromosomes. *PLoS Genetics*, **19**, e1010633. Available from: https://doi.org/10.1371/journal.pgen.1010633

Macas, J., **Novák, P.**, **Pellicer, J.**, **Čížková, J.**, **Koblížková, A.**, **Neumann, P.** ***et al.*** (2015) In depth characterization of repetitive DNA in 23 plant genomes reveals sources of genome size variation in the legume tribe Fabeae. *PLoS One*, **10**, e0143424. Available from: https://doi.org/10.1371/journal.pone.0143424

Marchler-Bauer, A., **Bo, Y.**, **Han, L.**, **He, J.**, **Lanczycki, C.J.**, **Lu, S.** ***et al.*** (2017) CDD/SPARCLE: functional classification of proteins via subfamily domain architectures. *Nucleic Acids Research*, **45**, D200–D203. Available from: https://doi.org/10.1093/nar/gkw1129

Marques, A., **Ribeiro, T.**, **Neumann, P.**, **Macas, J.**, **Novák, P.**, **Schubert, V.** ***et al.*** (2015) Holocentromeres in *Rhynchospora* are associated with genome-wide centromere-specific repeat arrays interspersed among euchromatin. *Proceedings of the National Academy of Sciences of the United States of America*, **112**, 13633–13638. Available from: https://doi.org/10.1073/pnas.1512255112

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

Márquez-Corro, J.I., **Escudero, M.** & **Luceño, 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 Research*, **26**, 139–152. Available from: https://doi.org/10.1007/s10577-018-9572-z

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

Mata-Sucre, Y., **Matzenauer, W.**, **Castro, N.**, **Huettel, B.**, **Pedrosa-Harand, A.**, **Marques, A.** ***et al.*** (2023) Repeat-based phylogenomics shed light on unclear relationships in the monocentric genus *Juncus* L.(Juncaceae). *Molecular Phylogenetics and Evolution*, **189**, 107930. Available from: https://doi.org/10.1016/j.ympev.2023.107930

McKinley, K.L. & **Cheeseman, I.M.** (2015) The molecular basis for centromere identity and function. *Nature Reviews Molecular Cell Biology*, **17**, 16–29. Available from: https://doi.org/10.1038/nrm.2015.5

Melters, D.P., **Bradnam, K.R.**, **Young, H.A.**, **Telis, N.**, **May, M.R.**, **Ruby, J.G.** ***et al.*** (2013) Comparative analysis of tandem repeats from hundreds of species reveals unique insights into centromere evolution. *Genome Biology*, **14**, 1–20. Available from: https://doi.org/10.1186/gb-2013-14-1-r10

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

Moore, G., **Aragón-Alcaide, L.**, **Roberts, M.**, **Reader, S.**, **Miller, T.** & **Foote, T.** (1997) Are rice chromosomes components of a holocentric chromosome ancestor? *Plant Molecular Biology*, **35**, 17–23. Available from: https://doi.org/10.1023/A:1005849912205

Nagaki, K., **Kashihara, K.** & **Murata, M.** (2005) Visualization of diffuse centromeres with centromere-specific histone h3 in the holocentric plant *Luzula nivea*. *The Plant Cell*, **17**, 1886–1893. Available from: https://doi.org/10.1105/tpc.105.032961

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

Neumann, P., **Navrátilová, A.**, **Koblížková, A.**, **Kejnovský, E.**, **Hřibová, E.**, **Hobza, R.** ***et al.*** (2011) Plant centromeric retrotransposons: a structural and cytogenetic perspective. *Mobile DNA*, **2**, 1–16. Available from: https://doi.org/10.1186/1759-8753-2-4

Neumann, P., **Navratilova, A.**, **Schroeder-Reiter, E.**, **Koblížková, A.**, **Steinbauerova, V.**, **Chocholova, E.** ***et al.*** (2012) Stretching the rules: monocentric chromosomes with multiple centromere domains. *PLoS Genetics*, **8**, e1002777. Available from: https://doi.org/10.1371/journal.pgen.1002777

Neumann, P., **Novák, P.**, **Hoštáková, N.** & **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**, 1–17. Available from: https://doi.org/10.1186/s13100-018-0144-1

Neumann, P., **Oliveira, L.**, **Čížková, J.**, **Jang, T.S.**, **Klemme, S.**, **Novák, P.** ***et al.*** (2021) Impact of parasitic lifestyle and different types of centromere organization on chromosome and genome evolution in the plant genus *Cuscuta*. *New Phytologist*, **229**, 2365–2377. Available from: https://doi.org/10.1111/nph.17003

Neumann, P., **Oliveira, L.**, **Jang, T.S.**, **Novák, P.**, **Koblížková, A.**, **Schubert, V.** ***et al.*** (2023) Disruption of the standard kinetochore in holocentric *Cuscuta* species. *Proceedings of the National Academy of Sciences of the United States of America*, **120**, e2300877120. Available from: https://doi.org/10.1073/pnas.2300877120

Novák, P., **Ávila Robledillo, L.**, **Koblížková, A.**, **Vrbová, I.**, **Neumann, P.** & **Macas, J.** (2017) TAREAN: a computational tool for identification and characterization of satellite DNA from unassembled short reads. *Nucleic Acids Research*, **45**, e111. Available from: https://doi.org/10.1093/nar/gkx257

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

Novák, P., **Neumann, P.**, **Pech, J.**, **Steinhaisl, J.** & **Macas, J.** (2013) RepeatExplorer: a galaxy-based web server for genome-wide characterization of eukaryotic repetitive elements from next-generation sequence reads. *Bioinformatics*, **29**, 792–793. Available from: https://doi.org/10.1093/bioinformatics/btt054

Ou, S., **Su, W.**, **Liao, Y.**, **Chougule, K.**, **Agda, J.R.**, **Hellinga, A.J.** ***et al.*** (2019) Benchmarking transposable element annotation methods for creation of a streamlined, comprehensive pipeline. *Genome Biology*, **20**, 1–18. Available from: https://doi.org/10.1186/s13059-019-1905-y

Ourari, M., **Coriton, O.**, **Martin, G.**, **Huteau, V.**, **Keller, J.**, **Ainouche, M.L.** ***et al.*** (2020) Screening diversity and distribution of Copia retrotransposons reveals a specific amplification of BARE1 elements in genomes of the polyploid *Hordeum murinum* complex. *Genetica*, **148**, 109–123. Available from: https://doi.org/10.1007/s10709-020-00094-3

Pazy, B. & **Plitmann, U.** (1994) Holocentric chromosome behavior in *Cuscuta* (Cuscutaceae). *Plant Systematics and Evolution*, **191**, 105–109. Available from: https://doi.org/10.1007/BF00985345

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

Planta, J., **Liang, Y.Y.**, **Xin, H.**, **Chansler, M.T.**, **Prather, L.A.**, **Jiang, N.** ***et al.*** (2022) Chromosome-scale genome assemblies and annotations for Poales species *Carex cristatella*, *Carex scoparia*, *Juncus effusus*, and *Juncus inflexus*. *G3 (Bethesda)*, **12**, jkac211. Available from: https://doi.org/10.1093/g3journal/jkac211

POWO. (2023) *Plants of the world online*. Facilitated by the Royal Botanic Gardens, Kew. Available from: http://www.plantsoftheworldonline.org/ [Accessed 2nd September 2023].

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

Reimer, J.J. & **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. In: Kovalchuk, I. & Zemp, F. (Eds.) *Epigenetics: methods in molecular biology*. Totowa, NJ: Humana Press, pp. 139–160. Available from: https://doi.org/10.1007/978-1-60761-646-7_12

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

Ribeiro, T., **Vasconcelos, E.**, **Dos Santos, K.G.**, **Vaio, M.**, **Brasileiro-Vidal, A.C.** & **Pedrosa-Harand, A.** (2020) Diversity of repetitive sequences within compact genomes of *Phaseolus* L. beans and allied genera *Cajanus* L. and *Vigna* Savi. *Chromosome Research*, **28**, 139–153. Available from: https://doi.org/10.1007/s10577-019-09618-w

Roalson, E.H. (2005) Phylogenetic relationships in the Juncaceae inferred from nuclear ribosomal DNA internal transcribed spacer sequence data. *International Journal of Plant Sciences*, **166**, 397–413. Available from: https://doi.org/10.1086/428757

Ruiz-Ruano, F.J., **Castillo-Martínez, J.**, **Cabrero, J.**, **Gómez, R.**, **Camacho, J.P.M.** & **López-León, M.D.** (2018) High-throughput analysis of satellite DNA in the grasshopper *Pyrgomorpha conica* reveals abundance of homologous and heterologous higher-order repeats. *Chromosoma*, **127**, 323–340. Available from: https://doi.org/10.1007/s00412-018-0666-9

Ruiz-Ruano, F.J., **López-León, M.D.**, **Cabrero, J.** & **Camacho, J.P.M.** (2016) High-throughput analysis of the satellitome illuminates satellite DNA evolution. *Scientific Reports*, **6**, 28333. Available from: https://doi.org/10.1038/srep28333

Sader, M., **Vaio, M.**, **Cauz-Santos, L.A.**, **Dornelas, M.C.**, **Vieira, M.L.C.**, **Melo, N.** ***et al.*** (2021) Large vs small genomes in *Passiflora*: the influence of the mobilome and the satellitome. *Planta*, **253**, 1–18. Available from: https://doi.org/10.1007/s00425-021-03598-0

Šatović-Vukšić, E. & **Plohl, M.** (2023) Satellite DNAs—from localized to highly dispersed genome components. *Genes*, **14**, 742. Available from: https://doi.org/10.3390/genes14030742

Schnable, P.S., **Ware, D.**, **Fulton, R.S.**, **Stein, J.C.**, **Wei, F.**, **Pasternak, S.** ***et al.*** (2009) The B73 maize genome: complexity, diversity, and dynamics. *Science*, **326**, 1112–1115. Available from: https://doi.org/10.1126/science.1178534

Schubert, V., **Neumann, P.**, **Marques, A.**, **Heckmann, S.**, **Macas, J.**, **Pedrosa-Harand, A.** ***et al.*** (2020) Super-resolution microscopy reveals diversity of plant centromere architecture. *International Journal of Molecular Sciences*, **21**, 3488. Available from: https://doi.org/10.3390/ijms21103488

Senaratne, A.P., **Cortes-Silva, N.** & **Drinnenberg, I.A.** (2022) Evolution of holocentric chromosomes: drivers, diversity, and deterrents. *Seminars in Cell & Developmental Biology*, **127**, 90–99. Available from: https://doi.org/10.1016/j.semcdb.2022.01.003

Shirakawa, J., **Hoshi, Y.** & **Kondo, K.** (2011) Chromosome differentiation and genome organization in carnivorous plant family Droseraceae. *Chromosome Botany*, **6**, 111–119. Available from: https://doi.org/10.3199/iscb.6.111

Šmarda, P., **Bureš, P.**, **Horová, L.**, **Leitch, I.J.**, **Mucina, L.**, **Pacini, E.** ***et al.*** (2014) Ecological and evolutionary significance of genomic GC content diversity in monocots. *Proceedings of the National Academy of Sciences of the United States of America*, **111**, E4096–E4102. Available from: https://doi.org/10.1073/pnas.1321152111

Sonnhammer, E.L. & **Durbin, R.** (1995) A dot-matrix program with dynamic threshold control suited for genomic DNA and protein sequence analysis. *Gene*, **167**, GC1–GC10. Available from: https://doi.org/10.1016/0378-1119(95)00714-8

Sullivan, L.L. & **Sullivan, B.A.** (2020) Genomic and functional variation of human centromeres. *Experimental Cell Research*, **389**, 111896. Available from: https://doi.org/10.1016/j.yexcr.2020.111896

Untergasser, A., **Cutcutache, I.**, **Koressaar, T.**, **Ye, J.**, **Faircloth, B.C.**, **Remm, M.** ***et al.*** (2012) Primer3—new capabilities and interfaces. *Nucleic Acids Research*, **40**, e115. Available from: https://doi.org/10.1093/nar/gks596

Van-Lume, B., **Mata-Sucre, Y.**, **Báez, M.**, **Ribeiro, T.**, **Huettel, B.**, **Gagnon, E.** ***et al.*** (2019) Evolutionary convergence or homology? Comparative cytogenomics of *Caesalpinia* group species (Leguminosae) reveals diversification in the pericentromeric heterochromatic composition. *Planta*, **250**, 2173–2186. Available from: https://doi.org/10.1007/s00425-019-03287-z

Vymazal, J. (2014) Constructed wetlands for treatment of industrial wastewaters: a review. *Ecological Engineering*, **73**, 724–751. Available from: https://doi.org/10.1016/j.ecoleng.2014.09.034

Wang, B., **Yang, X.**, **Jia, Y.**, **Xu, Y.**, **Jia, P.**, **Dang, N.** ***et al.*** (2022) High-quality *Arabidopsis thaliana* genome assembly with nanopore and hifi long reads. *Genomics, Proteomics & Bioinformatics*, **20**, 4–13. Available from: https://doi.org/10.1016/j.gpb.2021.08.003

Wlodzimierz, P., **Rabanal, F.A.**, **Burns, R.**, **Naish, M.**, **Primetis, E.**, **Scott, A.** ***et al.*** (2023) Cycles of satellite and transposon evolution in *Arabidopsis* centromeres. *Nature*, **618**, 557–565. Available from: https://doi.org/10.1038/s41586-023-06062-z

Wolfgruber, T.K., **Sharma, A.**, **Schneider, K.L.**, **Albert, P.S.**, **Koo, D.H.**, **Shi, J.** ***et al.*** (2009) Maize centromere structure and evolution: sequence analysis of centromeres 2 and 5 reveals dynamic loci shaped primarily by retrotransposons. *PLoS Genetics*, **5**, e1000743. Available from: https://doi.org/10.1371/journal.pgen.1000743

Yu, Y., **Ouyang, Y.** & **Yao, W.** (2018) shinyCircos: an R/shiny application for interactive creation of Circos plot. *Bioinformatics*, **34**, 1229–1231. Available from: https://doi.org/10.1093/bioinformatics/btx763

Záveská Drábková, L. (2013) A survey of karyological phenomena in the Juncaceae with emphasis on chromosome number variation and evolution. *The Botanical Review*, **79**, 401–446. Available from: https://doi.org/10.1007/s12229-013-9127-6

Zedek, F., **Šmerda, J.**, **Veselý, P.**, **Horová, L.**, **Kocmanová, J.** & **Bureš, P.** (2021) Elevation-dependent endopolyploid response suggests that plants with holocentric chromosomes are less stressed by UV-B. *Botanical Journal of the Linnean Society*, **195**, 106–113. Available from: https://doi.org/10.1093/botlinnean/boaa114