**Comparative analyses of sex chromosomes and autosomes in *Rumex acetosa***

Wojciech Jesionek1, Radim Cegan1, Janka Puterova1,2, Boris Vyskot1, Jan Vrana3, Jan Safar3, Roman Hobza1,3,\*

*1Department of Plant Developmental Genetics, Institute of Biophysics, Academy of Sciences of the Czech Republic, Královopolská 135, 61200 Brno, Czech Republic*

*2Department of Information Systems, Faculty of Information Technology, Brno University of Technology, Bozetechova 2, 61200 Brno, Czech Republic*

*3Institute of Experimental Botany AS CR, Centre of the Region Hana for Biotechnological and Agricultural Research, Olomouc-Holice, Czech Republic*

\* Corresponding author: Roman Hobza, Department of Plant Developmental Genetics, Institute of Biophysics, Kralovopolska 135, 612 00 Brno, Czech Republic, phone: +420 541517203, [www.ibp.cz](http://www.ibp.cz/), Email: hobza@ibp.cz

**Abstract**

The presence of dioecious species with well-established sex chromosomes is very rare in the plant kingdom. Common sorrel (*Rumex acetosa*) is a dioecious plant with chromosomal sex determination. Although the evolution of sex chromosomes has been the subject of numerous studies, a global view of sex chromosome structure is still missing in this species. We have flow-sorted and separately sequenced sex chromosomes and autosomes in *R. acetosa*. We have identified and quantified individual repeats and compared their distribution within individual chromosomes. We focused on the role of various repetitive elements in the process of X and Y chromosome evolution. Our data demonstrate that the *R. acetosa* genome was formed by the expansion of several repetitive elements with a specific pattern of distribution in the case of the sex chromosomes. We show that some tandem repeats and retroelements are ubiquitous in the *R. acetosa* genome but surprisingly absent on X and/or Y chromosomes. We provide first chromosome specific repeatome analysis in plant species possessing sex chromosomes.

**Key words**: sex chromosomes; genome dynamics; transposable elements; satellites

**Introduction**

In contrast to animals, only 5% of the flowering plant species (Charlesworth 2016) contain individuals with separate sexes, and only a few of these have morphologically recognizable sex chromosomes. Some, such as *Silene latifolia* (white campion) or *Coccinia grandis* (Sousa et al. 2016), are model plants for studying sex chromosome evolution, contain two X chromosomes in a female, and one X and one Y chromosome in a male. In these species, the sex is determined by the dominant Y chromosome. On the other hand, *Rumex acetosa* (common garden sorrel) has six pairs of autosomes and two (XX, female) or three (XY1Y2, male) sex chromosomes (Kihara and Ono 1923). In this species with aneuploid sex chromosomes, sex determination is based on the ratio between the number of X chromosomes and autosomes (X:A) (Ainsworth et al. 1998).

Most dioecious plants have homomorphic sex chromosomes. Microscopically distinguishable sex chromosomes were reported in 19 species (Ming. et al. 2011; Renner 2014). The most studied plant species with heteromorphic sex chromosomes are *Cannabis sativa* (hemp) (Sakamoto et al. 2000), *Hippophae rhamnoides* (seabuckthorn) (Puterova et al. 2017), *Coccinia grandis* (Sousa et al. 2016), and classical models *R. acetosa* (sorrel) and *Silene latifolia* (white campion) (Vyskot et al. 2004). Heteromorphic sex chromosomes evolved from a pair of autosomes acquiring sex-determining genes. Because of recombination suppression in sex determining locus, transposons and satellites started to accumulate along with gene degeneration and loss. The different degeneration level of plant Y chromosomes can be caused by recent and/or repeated origin of sex chromosomes. It was hypothesized that suppression of recombination in *R. acetosa* spread to the whole chromosomes causing the change of the sex determination system from the dominant Y to the X-to-autosome ratio (Ming et al. 2011). *R. acetosa* Y chromosomes are large, full of repetitive sequences, and cytologically heterochromatic (Shibata et al. 1999; Mariotti et al. 2006; Steflova et al. 2013). The small number of chromosomes and their heteromorphic character makes *R. acetosa* an ideal model for cytogenetic studies. Accumulation of repetitive DNA on the Y chromosome is a complex process which is modified by a great number of genetic and epigenetic factors (Hobza et al. 2015). This phenomenon can be seen e.g. in *S. latifolia*, in which the Y chromosome has expanded in size and has accumulated satellite sequences but remained euchromatic (Puterova et al. 2018). Accumulation of satellite DNA frequently corresponds with heterochromatization of the Y chromosome (Lengerova and Vyskot 2001). Dynamic changes of repetitive DNA may also affect X chromosome causing its expansion (Gschwend et al. 2012).

Sex chromosomes in plants differ in size (Parker 1990; Ainsworth 2000).  In various species, the DNA amount and chromosome length differ between male and female haploid genome and varies from 3 to 6% (Costich et al. 1991; Veuskens et al. 1992; Vagera et al. 1994; Matsunaga et al. 1994; Doležel and Gohde 1995; Grabowska-Joachimiak and Joachimiak 2002; Grabowska-Joachimiak et al. 2005; Błocka-Wandas et al. 2007). Male to female genome size difference in *R. acetosa* is 7.11% at the diploid level because of a small number of autosomes and exceptionally big and heteromorphic sex chromosomes (Błocka-Wandas et al. 2007). *R. acetosa* sex chromosomes are the biggest chromosomes in the genome and contain a large number of repetitive sequences. Many repetitive DNA sequences that have accumulated and/or have a specific pattern on *R. acetosa* sex chromosomes were described previously (Shibata et al. 1999; Shibata et al. 2000; Navajas-Perez et al. 2005; Mariotti et al. 2009; Steflova et al. 2013). Also, a repeat composition of the male and female genome was estimated (Steflova et al. 2013). These studies have shed light on the content of sex chromosomes in *R. acetosa*. However, previous analysis was not sufficient to fully estimate and describe the repetitive fraction of sex chromosomes.

So far, none of the plant sex chromosomes were individually sequenced and analysed in detail in terms of the repetitive DNA characterization. We used FISHIS - Fluorescence *In Situ* Hybridization in Suspension method to sort autosomes, X and Y chromosomes separately for subsequent whole chromosome sequencing experiments. Here, we describe structural differences between sex chromosomes and autosomes in *R. acetosa*. Using direct sequencing and subsequent repeatome assembly of X, Ys and autosomes, we quantified the proportion of individual repeats within the chromosomes. Based on this approach we raise the following questions: What kind of repeats accumulate on Y chromosomes? How do X and Y chromosomes differ structurally from each other and from the rest of the genome? What is the extent of the individual categories of repeats (transposable elements, tandem repeats) responsible for the potential structural differences?

**Materials and methods**

**Chromosome sorting using Fluorescence In Situ Hybridization in Suspension (FISHIS)**

The samples for flow cytometric experiments were prepared from *R. acetosa* root tips according to Vrána et al. (2016). Seeds of *R. acetosa* were germinated in a Petri dish immersed in water at 25 °C for 2 days until the optimal length of roots was achieved (ca. 1 cm). The root cells were synchronized by treatment with 2 mM hydroxyurea at 25 °C for 18 h. Accumulation of metaphases was achieved using 10 μM oryzalin solution at 25 °C for 2 h. Approximately 200 root tips were required to prepare 1 ml of sample. The chromosomes were released from the root tips by mechanical homogenization using a Polytron PT1200 homogenizer (Kinematica AG, Littau, Switzerland) at 18,000 rpm for 13 s. The crude suspension was filtered. For better discrimination of Y chromosomes, we performed FISHIS using Chromosome Flow Sorting (Giorgi et al. 2013). Briefly, 1 ml of crude suspension was used. 10 M NaOH was used to reach pH in the range 12.8 to 13.3. The suspension was incubated for 15 min on ice, then the pH was adjusted in the range of 8.5 to 9.1 using Tris·Cl. 5′-FITC-(CAA)10 probe solution (1 ng/μl) was added to the final concentration 180 ng/ml and the suspension was incubated for 1 h in the dark at room temperature and finally was put on ice and kept until the flow cytometric analysis. The samples were counterstained with DAPI (2 μg/ml final concentration). All flow cytometric experiments were performed on FACSAria II SORP flow cytometer (BD Biosciences, San José, Calif., USA). Chromosomes were sorted out by the relative DNA content (DAPI signal) and (CAA)10 microsatellite abundance (FITC signal) for proper discrimination of the Ys from other chromosomes. We obtained 6 chromosomal fractions – 4 autosomal, X fraction and Y1Y2 fraction. Quality was checked by microscopy of each sample. Purity was estimated at 95%. We used about 1 milion chromosomes (100 ng of DNA) which were preamplified with Phi29 polymerase (Šimková et al. 2008). Chromosome sorting and amplification were repeated three times to reduce the amplification bias.

**Illumina Sequencing**

We performed one run of paired-end Illumina MiSeq sequencing generating 301 bp long reads for autosomes and 251 bp long reads for X chromosome and Y chromosomes (accession number PRJEB23612) separately. We obtained 25,672,002 raw paired-end reads from autosomes, 4,591,591 raw paired-end reads from X chromosome and 2,731,018 raw paired-end reads from Y chromosomes. Sequencing reads were checked for quality using FastQC tool (available at: <http://www.bioinformatics.babraham.ac.uk/projects/fastqc>). Reads were preprocessed based on quality with subsequent adaptor trimming, filtering out short or unpaired sequences and trimming all reads to a uniform length of 235 nucleotides by Trimmomatic tools (Bolger et al. 2014) using the Galaxy platform (Afgan et al. 2016).

We estimated the coverage of the male genome using the chromosome length as described in Lengerova and Vyskot (2001). The genome size of *R. acetosa* was previously reported to be 7.0 pg for the female and 7.5 pg for the male genome (2C) (Blocka-Wandas et al. 2007). For further analysis, we randomly sampled the sequencing data proportionally to reflect the male genome and ended up with 1,702,340 reads from autosomes, 287,234 from X chromosome, and 376,276 from Y chromosomes what equals ~0.074x coverage of the male genome. Such coverage is sufficient for the assembly of highly and moderately repetitive sequences (Macas et al. 2015).

**Identification of repetitive sequences**

To identify repetitive DNA in autosomes and X and Y chromosomes of *R. acetosa* we carried out comparative analysis using RepeatExplorer tool (Novak et al. 2010; Novak et al. 2013). This tool performs graph-based clustering of sequences based on their similarity. Clusters were annotated manually using Geneious software version 7.1.9 (Drummond et al. 2013) and automatically using RepeatExplorer output. We screened the clustering results to find sequences which were reported previously. Clusters containing unknown sequences were investigated for typical transposon protein domains by CDD (Marchler-Bauer A et al. 2017). Monomers of satellite DNA were detected by TRF 4.09 (Benson 1999). Finally, we manually created a library of repeats using sequences derived from clusters.

**FISH analysis**

Specific primers were designed for contigs from selected clusters. For the transposons, primers were designed for the LTRs and/or the transposon domains (for instance reverse transcriptase). Monomers of the satellite DNA were chosen for primer design. In the first step, template DNA was amplified using PCR with a mix containing 1 complete PCR buffer, 0.1 mM dNTPs, 0.1 mM primers, 0.5 U Taq polymerase (Top Bio), and 10–15 ng of template DNA. Reaction conditions were as follows: 95 °C/4 min 34✕ (95 °C/50s + 55 °C/50 s + 72 °C/1 min) + 72 °C/10 min. PCR products were checked by gel electrophoresis, cleaned using the PCR purification kit (Qiagen), cloned into a pDrive vector (Qiagen), and transformed into *Escherichia coli*. Clones were sequenced to verify the presence of a specific product. Selected clones were then used for probes preparation for fluorescence in situ hybridization (FISH) by PCR and labeling using Nick Translation Kit (Roche).

FISH was performed on mitotic metaphase chromosomes prepared from root tip cells. The hybridization mix contained 50% formamide, 2✕ SSC, and 10% dextran sulphate. The labeled DNA (1 - 5 ng/μl) was denatured, added to the slide, and hybridized at 37 °C for 18 h. Slides were then washed with medium stringency: 250 s in 2✕ SSC at 42 °C, 250 s in 0.1 SSC at 42 °C, 250 s in 2✕ SSC at 42 °C, 50 s in 2✕ SSC at room temperature, 70 s in 4✕ SSC + 1% Tween, and finally washed in 1✕ PBS. The chromosomes were counterstained with DAPI and mounted in vectashield, examined under an Olympus AX70 fluorescent microscope, scanned by CCD camera, and analyzed by ISIS software.

**BAC library construction and screening**

The BAC library was constructed from *R. acetosa* male high molecular weight genomic DNA. Briefly, DNA was digested with *Hind*III enzyme and inserted into pIndigoBAC-5 vector. Clones were then grid in duplicate on Hybond N+ (Amersham, Biosciences) nitrocellulose membrane filters in a 4 x 4 pattern that allowed us to identify well positions and plate numbers of each clone. The filters were incubated and processed as described in Bouzidi et al. (2006). The *R. acetosa* BAC library (total of 72,000 colonies) was arrayed on six nylon filters with 18,432 colonies each, and one nylon filter containing 9,216 clones. The average insert-size of the library is 128 kb. Based on nuclear size data, we have estimated that coverage of the *R. acetosa* BAC library is 2.84 complements of the male haploid genome. The screening was performed by radioactive hybridization with α32P using Prime-It II Random Primer Labelling Kit (Stratagene) according to the manufacturer's protocol. Probes were prepared by PCR amplification of the different sequences derived from contigs. We selected clones showing strong hybridization with the probe, and only those that were confirmed by PCR with probe-derived primers were used in further analyses. Clones were sequenced using Illumina MiSeq 300 nt paired-end sequencing. Raw data processing, sequence assembling, alignment, and annotation were done with Geneious software (Drummond et al. 2013) and Edena v3 assembler (Hernandez et al. 2008).

**Results**

**Repeat assembly, annotation and quantification**

Using a RepeatExplorer pipeline, we identified the main groups of repetitive DNA in *R. acetosa* genome. We estimated the proportion of the main repeat families in *R. acetosa* autosomes, X and Y chromosomes separately. For the further analyses, we used 319 clusters representing at least 0.01% of the genome which totally makes up 57.62% of autosomes, 68.07% of X and 73.75% of Y chromosome reads. We counted proportions and described the main types of repetitive DNA. 39 out of 319 studied clusters were annotated as satellites and 123 clusters as retrotransposons. One cluster corresponded to 5S rDNA (CL285) and one to 45S rDNA (CL165). Since we used flow-sorted chromosomes, none of the analyzing contigs contained chloroplast DNA, although cpDNA were found in smaller clusters probably because of the nuclear cpDNA insertions (Steflova et al. 2014). Four clusters (CL54, CL66, CL77, CL115) were considered to be bacterial contamination.

**Chromosome specific comparative analysis revealed new satellites**

For each identified satellite, we reconstructed a monomer and described its size. Using sequences available in NCBI, we identified known *R. acetosa* satellite DNA sequences: RAYSI, RAYSII, RAYSIII, RAE180, RAE730, RA160, RA690. All of the newly described satellite sequences occur mostly on Y chromosomes (see complete list of satellites in Table 1). We named newly discovered sequence according to the genome of origin (RA) and monomer size of the satellite, or in case of the Y specific satellites, we continued naming repeats with RAYS (***R****umex* ***a****cetosa* **Y** **s**pecific as in Shibata et al. (1999)) prefix. To determine character of distribution of individual repeats we performed FISH experiments (Fig. 3). We used RAYSI satellite as a Y chromosomes marker - the signal is localized in four spots on each arm of the Y1 chromosome and in two spots on each arm of Y2 chromosome.

RAE173 tandem repeat is localized prefentially on both Y chromosomes and together with RAE180 paint the Y chromosomes almost completely (Fig. 3a). RAE180 and RA173 make together almost 7.5% of the Y chromosomes. RA173 satellite, contrary to RAE180, has lower proportin on X chromosomes and autosomes (see Table 1).

RAYSIV has a specific pattern on Y chromosomes. FISH signals is localized in two loci on Y1 - one on every chromosome arm, and in three loci on the Y2 chromosome - two on the p arm and one on the q arm (Fig. 3b).

While 0.012% of autosomes and 0.028% of Y chromosomes consist of RAYSV satellite, FISH revealed only one locus on the Y1 chromosome and no signal on other sex chromosomes and autosomes (Fig. 3c). To explain the discrepancy between the sequencing data and FISH results, we screened *R. acetosa* BAC library with RAYSV derived probe. We chose six BACs with the strongest signal and sequenced them. Sequencing data revealed that this satellite sequence is highly variable and individual monomers differ significantly from each other. Although FISH analysis revealed distinct and specific signal on the Y1 chromosome, sequencing data suggest that this satellite is also present on autosomes and completely missing on the X.

Although RAYSVI has similar FISH pattern as RAYSI there is no sequence similarity. FISH reveals 1 minor and 2 major loci at the p arm of the Y1 chromosome. In the case of the Y2 chromosome FISH shows 3 minor signals on the p arm and 1 strong signal on the q arm (Fig. 3d).

RAYSVII is localized at the end of longer arm of the Y1 chromosome (Fig. 3e). Contrary to the variable manner of RAYSV monomers, we found only one BAC containing RAYSVII sequence. This tandem repeat makes 0.211% of Y chromosomes and less than 0.002% of autosomes and the X chromosome. According to the proportion of the repeat and BAC sequencing results, we assume that the visible localization of the RAYSV satellite is the only locus of this repeat in the *R. acetosa*’s genome.

Contrary to some newly described satellites (RAYSIV, RAYSVI, RAYSVII), which create discrete tandemly arranged loci in the genome, RAE244 is localized in multiple loci on both Y chromosomes and two pairs of autosomes. (Fig. 3f).

Based on the clustering data, RAE93 is attached to the Ogre/Tat element. FISH using RA Ogre/TAT derived probe paints entire Y chromosome with minor addional signals in rest of the genome (Fig. 3h). FISH with RAE93 satellite shows similar pattern of distribution as the Ogre/Tat element with lower number of hybridizing loci (Fig. 3g). Since transposons are usually not tandemly arranged, but rather widespread through the genome, we assume that the transposon containing satellite spread mostly on Y chromosomes while the same transposon without tandem repeat is dominant in the rest of the genome.

One of the tandemly arranged sequences has a different nature than satellites. It contains *gag* protein like domain of retrotransposon *Ty1/Copia* AleII and domains similar to DnaJ protein, PPT, PBS, and LTR. This element lacks reverse transcriptase and other *pol* genes. RAAleII makes 0.066% of autosomes, 0.337% of X and 0.031% of Y chromosomes. FISH image shows signal at the distal part of the shorter arm of the Y1 and on the X chromosome (Fig. 3i). A tandem nature of this repeat was confirmed by BAC sequencing.

**Transposable elements – variable distribution on sex chromosomes**

We classified and calculated the abundance of the transposable elements in *R. acetosa* genome. Most of the *Ty1*/*Copia* transposons were classified as a Maximus/SIRE family. The differences in the percentage of the Copia-like elements among chromosomes vary between 23.03% on autosomes and 27.83% on X chromosome (Fig. 1).

A much more evident difference in the transposon abundance was apparent in *Ty3/Gypsy* group of transposons. Autosomal data consisted of 15.39% of the Gypsy-like elements, while X and Y chromosomes have about 26% (Fig. 1). Closer analysis of the RT domains of elements falling into the Gypsy family revealed three main subfamilies that are the most abundant in *R. acetosa* genome (Fig. 2). While there is no apparent difference in the total percentage of Gypsy transposons in all 3 groups of chromosomes, by comparing individual subfamilies we can detect significant variations. The most common was Athila which abundance varies between 9.36% in autosomes up to 12.50% in Ys. A similar situation appears in the Chromoviruses - the smallest percentage of this element is on the autosomes (5.09%) and X chromosome (7.02%). Y chromosomes have the biggest percentage of this elements – 12.06%. The greatest discrepancy is visible in Ogre/Tat subfamily, where these elements constitute 8.57% of X chromosomes while in autosomes and Y chromosomes their proportion is much smaller and does not exceed 2% (Fig. 2).

DNA transposons are mostly abundant on the autosomes (11.26%). X and Y chromosomes have a similar proportion of this repeat type (5.5% and 5.96% respectively)

**Discussion**

This study represents a direct approach to analyse and quantify individual repetitive elements on the sex chromosomes and autosomes. Using direct sorting and sequencing of individual chromosomes, we demonstrate the differences between autosomes, X and Y chromosomes in *R. acetosa*. We present the first quantitative analysis of repetitive sequences in plant sex chromosomes. Sex chromosomes as a non-recombining part of a genome frequently accumulate various types of repetitive DNA sequences. Surprisingly, in some species the regulation of retroelement activity in males supressess the spread of TEs on the Y chromosomes (Kubat et al. 2014). Therefore, sex chromosomes quickly diverge from the rest of the genome. Those processes are monitored either by cytogenetic methods (e.g. visualization of heterochromatic regions and/or FISH experiments with selected probes) or by analysis of whole genome sequences. Previous studies were based on either description of the differences between male and female genomes (Steflova et al. 2013) or only local aspects of sex chromosome divergence were studied (Shibata et al. 1999; Mariotti et al. 2009; Shibata et al. 2000; Navajas-Perez et al. 2005; Mariotti et al. 2013; Steflova et al. 2013).

**Satellite sequences – the key players of the Y chromosome expansion?**

Although it was already shown that *R. acetosa* Y chromosomes encompass a greater percentage of satellite sequences than the X and autosomes, we have identified and described six novel satellites (Table 1) which make along with previously published tandem repeats (RAYSI, RAYSII, RAYSIII, RAE180, RAE730, RA160, RA690) more than 13% of Y chromosomes. Two of them (RAE180 and RAE173 - see Table 1) make up half of this number. Our chromosome-based approach has extended and improved genome assembly at the repeatome level and has enabled the identifcation of new sex enriched repeats. Out of six newly identified satellites, four are sex-chromosome specific. Surprisingly, previous studies focused on the description of major genome repeats (Shibata et al. 1999; Shibata et al. 2000; Navajas-Perez et al. 2005; Mariotti et al. 2009; Mariotti et al. 2013; Steflova et al. 2013) did not identify satellites RAE173, RAE244, RAYSIV, RAYSV, RAYSVI, RAYSVII (5.32% of Y chromosomes). Such an elevated number of the different satellite families is very unusual in plant sex chromosomes. Recently, Puterova et al. (2017) revealed a similar expansion of different satellites in seabuckthorn.

**Retrotransposons and sex chromosome evolution**

LTR retrotransposons in many plant species are the primary genome components (Macas et al. 2015). In most cases, only one or a few families of either Copia or Gypsy retrotransposons are dominant. Our data show that LTR retrotransposons play a crucial role in the genome size evolution in *R. acetosa* and their distribution differs on autosomes, X and Y chromosomes. Autosomes consist of 38.42% of TEs, wile X and Y chromosomes 54.32% and 52.68% respectively. Almost all Copia retrotransposons (98.75%) found in *R. acetosa* belong to Maximus/SIRE lineage with a quite high sequence similarity among transposons. Such high sequence conservation was reported in other species and it has been suggested by Macas et al. (2015) that it might be a general feature of Copia elements. Also, a high sequence similarity between Maximus/SIRE transposons implies that they may have evolved from one transposon spreading through the genome after Y chromosomes evolved. It resulted in the absence of these elements on Y chromosomes (Steflova et al. 2013). Another explanation was suggested by Kubat et al. (2014) who reported that the absence of transposons on Y chromosomes in *S. latifolia* was probably caused by epigenetic silencing in the sperm and embryo.

There is almost no difference between Copia TEs proportion in X and Y chromosomes (27.83% and 26.37%, respectively). Autosomes consist of smaller percentage of Copia TEs (23.03%) than the sex chromosomes.

In contrast to Copia elements, there are three main subfamilies of Gypsy LTR retrotransposons in *R. acetosa* - Athila, Chromovirus and Ogre/TAT. The proportion of Athila elements differ among autosomes, X and Y chromosomes (9.36%, 10.89% and 12.50%, respectively). The Chromovirus and Ogre/Tat subfamilies show bigger differences in their abundance than Athila. Chromovirus is relatively abundant on the Y chromosomes (12.06%) but it is less dominant on autosomes and X chromosomes (5.09% and 7.02%, respectively). Ogre/TAT makes 8.57% of X chromosome, but only 0.94 % of autosomes and 1.75% of Y chromosomes. Our data show that Athila and Chromovirus elements are highly diverse and mostly nonfunctional - without domain(s) or functional LTRs. Ogre/Tat, on the other hand, have complete ORF and both LTRs. This finding suggests they are evolutionary young and still active. Sequences of Ogre/Tat elements during clustering of the sequencing reads were not divided into separate clusters in most cases. Their high abundance on the X chromosome may occur because of the limited recombination of the X chromosome in males.

Surprisingly, DNA transposons differ in their distribution, both from the satellites and retrotransposons. They are preferentially distributed and accumulated on the autosomes (11.26% of the autosomes, 5.50% of X and 5.96% of Y chromosomes).

**Sex chromosome formation – combination of various effects**

There are many processes that influence Y chromosome formation during evolution (Hobza et al. 2015). The accumulation of microsatellites is one of the first steps and microsatellite arrays have been suggested as being targets for TE insertions (Kejnovsky et al. 2013). In case of *R. acetosa*, this process probably has a limited impact on Y chromosome formation at the current stage of Y chromosome evolution since we did not observe any major microsatellite clusters. On the other hand, suppressed recombination between Y chromosomes reduced the rate of concerted evolution and led to the diversification of specific satellites such as RAYSV, as has been suggested by Navajas-Pérez et al. (2006). The non-recombining parts of sex chromosomes can lead to the evolution of new forms of repetitive sequences that can later spread into the rest of the genome. Although the widely accepted hypothesis predicts accumulation of repeats in the non-recombining region of Y chromosome (Charlesworth 1991), many repeats tend to have the opposite pattern of distribution. Our study supports observation by Kubat et al. (2014) suggesting sex specific regulation of retrotransposon spread. In the case of *R. acetosa* Ogre element, the absence of this TE on the Y chromosomes accompanied with accumulation of this element on the X chromosome resembles Ogre element proliferation as described in *S. latifolia* (Kubat et al. 2014; Hobza et al. 2017; Puterova et al. 2018). Male-specific silencing of TEs resulting in such a pattern of distribution has been shown to slow Y chromosome size expansion and genetic and epigenetic degeneration of Y-linked genes, and it has been suggested that it increases the size of X chromosome in *S. latifolia*.

In contrast to euchromatic X chromosome, both Y1 and Y2 have a heterochromatic nature. On the other hand, recent study showed the presence of functional genes on Y chromosomes in *R. acetosa* (Michalovova et al. 2015). Still are missing information about the localisation of potential gene regions on Y chromosomes and pseudo autosomal region (PAR). Such information could help to answer questions about the origin of Y1 and Y2 chromosomes. Farooq et al. (2013) reported that during meiosis sex chromosomes of *R. acetosa* form chain- or circle-shape trivalent (Y1–X–Y2). Our data support this observation since RAAleII retrotransposon (Fig. 3i) sequence (which is highly conservative) occur uniquely on the ends of the X and Y1 chromosomes. Contrary to RAAleII retrotransposon, TatCL11 element is spread through all autosomes, X chromosomes and only terminal parts of the Y chromosomes (Steflova et al. 2013). These results suggest that PAR is localised in the distal parts of the sex chromosomes. So far this is the first report of a shared part of Y and X chromosomes in this species.

In general, the repat composition of *R. acetosa* Y chromosomes differs significantly from the rest of the genome. This divergence is in accordance with the fact that Y chromosome is a special part of the genome because of its limited recombination and transmission only in the male lineage. Surprisingly, in the case of some elements, X chromosome is even more specific and diverged from other chromosomes. The higher proportion of Ogre/TAT retroelements on the X chromosome clearly distinguishes this chromosome from Ys and autosomes. Autosomes contain a similar percentage of satellite sequences as the X chromosome, but a higher proportion of DNA transposons. Y chromosomes contain an elevated percentage of tandem repeats and Athila and Chromovirus transposons. These results show that both types of sex chromosomes - X and Y - are unique in their structure and are different from autosomes.

**Acknowledgments**

This research was supported by the Czech Science Foundation (grants 18-06147S and 16-08698S).

This work was supported by grants of Brno University of Technology [FIT-S-17-3964].

We would like to thank Francesco Muto for English corrections

**Fig. 1**

**Repeat composition of autosomes, X and Y chromosomes of *Rumex acetosa* genome**

****

**Fig. 2**

**Abundance of subfamilies of Gypsy-like transposable elements on autosomes, X and Y chromosomes in *Rumex acetosa* genome**

****

**Fig. 3**

**Localisation of satellite DNA and transposable elements on metaphase chromosomes of *R. acetosa* using FISH. Bar = 10 µm**

**Figure 4. Chromosomal localization of repetitive DNA in *Silene latifolia* (a-c) and *Rumex acetosa* (d-f) determined by fluorescence *in situ* hybridization.** (a) Microsatellite (CA)15 is accumulated on the Y chromosome (red signal), (b) satellite STAR (red signal) is present in the centromeres of all autosomes and the X chromosome, and satellite X-43.1 (green signal) is gathered in both subtelomeres of all chromosomes with the exception of the Y chromosome possessing only one subtelomeric signal, (c) *Ogre* LTR retrotransposon (red signal) is ubiquitous on all chromosomes except the Y chromosome, (d) mixture of all mono-, di- and trinucleotide microsatellites (red signal) shows strong accumulation along both Y chromosomes, (e) (TA)15 microsatellite (red signal) gives a signal at several discrete loci on both Y chromosomes, and RAYSI satellite (green signal) is present at the distal regions of both Y chromosomes, (f) Maximus/SIRE LTR retrotransposon (red signal) covers all autosomes and the X chromosome except telomeres/subtelomeres but is absent from both Y chromosomes, RAYSI satellite is localized at several discrete loci on both Y chromosomes (green signal). Bars indicate 10µm.

****

**Table 1 Estimation satellite sequences of *Rumex acetosa* genome. Estimated from**

**Illumina Sequencing Data**

|  |
| --- |
| **Satellite sequences** |
| **Repeat name** | **Localisation** | **Reference** | **Proportion on chromosomes** |
| **A** | **X** | **Y** |
| RAE180 | Mostly on Y | Shibata et al. 2000 | 1.68% | 0.23% | 2.94% |
| RAYSI | Y specific | Shibata et al. 1999 | 0.08% | 0.05% | 1.39% |
| RAYSII | Y specific | Mariotti et al. 2009 | 0.01% | 0.01% | 0.03% |
| RAYSIII | Y specific | Mariotti et al. 2009 | 0.07% | 0.01% | 0.20% |
| RA160 | Y, X and 2 A | Steflova et el. 2013 | 0.11% | 0.00% | 0.01% |
| RA690 | Y, X and 2 A | Steflova et el. 2013 | 0.23% | 0.14% | 1.24% |
| RAE730 | Y and 1 A | Shibata et al. 2000 | 0.08% | 0.02% | 0.46% |
| **Novel Satellite Sequences** |
| **Repeat name** | **Localisation** | **Putative monomer length** | **Proportion on chromosomes** |
| **A** | **X** | **Y** |
| RAE173 | Mostly on Y | 173 | 0.09% | 0.17% | 4.51% |
| RAE244 | Mostly on Y | 244 | 0.01% | 0.00% | 0.09% |
| RAYSIV | Y specific | 175 | 0.01% | 0.01% | 0.26% |
| RAYSV | Y1 specific | 468 | 0.00% | 0.00% | 0.03% |
| RAYSVI | Y specific | 445 | 0.02% | 0.01% | 0.22% |
| RAYSVII | Y1 specific | 164 | 0.00% | 0.00% | 0.21% |

**References**

Afgan E, Baker D, van den Beek M, et al (2016) The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2016 update. Nucleic Acids Res 44:W3–W10. doi: 10.1093/nar/gkw343

Ainsworth C (2000) Boys and Girls Come Out to Play: The Molecular Biology of Dioecious Plants. Ann Bot 86:211–221. doi: 10.1006/anbo.2000.1201

Ainsworth C, Parker J, Buchanan-Wollaston V (1998) Sex determination by X: autosome dosage: Rumex acetosa (sorrel). In: Ainsworth C (ed) Sex Determination in Plants. Curr Top Dev Biol 38:167–223.

Benson G (1999) Tandem repeats finder: a program to analyze DNA sequences. Nucleic Acids Res 27:573–80.

Błocka-Wandas M, Sliwinska E, Grabowska-Joachimiak A, et al (2007) Male gametophyte development and two different DNA classes of pollen grains in Rumex acetosa L., a plant with an XX/XY1Y2 sex chromosome system and a female-biased sex ratio. Sex Plant Reprod 20:171–180. doi: 10.1007/s00497-007-0053-9

Bolger AM, Lohse M, Usadel B (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30:2114–2120. doi: 10.1093/bioinformatics/btu170

Bouzidi MF, Franchel J, Tao Q, et al (2006) A sunflower BAC library suitable for PCR screening and physical mapping of targeted genomic regions. Theor Appl Genet 113:81–89. doi: 10.1007/s00122-006-0274-6

Charlesworth B (1991) The evolution of sex chromosomes. Science 251:1030–3. doi: 10.1126/SCIENCE.1998119

Charlesworth D (2016) Plant Sex Chromosomes. Annu Rev Plant Biol 67:397–420. doi: 10.1146/annurev-arplant-043015-111911

Costich DE, Meagher TR, Yurkow EJ (1991) A rapid means of sex identification inSilene latifolia by use of flow cytometry. Plant Mol Biol Report 9:359–370. doi: 10.1007/BF02672012

Doležel J, Göhde W (1995) Sex determination in dioecious plants Melandrium album and M. rubrum using high-resolution flow cytometry. Cytometry 19:103–106. doi: 10.1002/cyto.990190203

Farooq U, Lovleen, Saggoo MIS (2014) Male meiosis and behaviour of sex chromosomes in different populations of Rumex acetosa L. from the Western Himalayas, India. Plant Syst Evol 300:287–294. doi: 10.1007/s00606-013-0881-z

Giorgi D, Farina A, Grosso V, et al (2013) FISHIS: Fluorescence In Situ Hybridization in Suspension and Chromosome Flow Sorting Made Easy. PLoS One 8:e57994. doi: 10.1371/journal.pone.0057994

Grabowska-Joachimiak A, Joachimiak A (2002) C-banded karyotypes of two *Silene* species with heteromorphic sex chromosomes. Genome 45:243–252. doi: 10.1139/g01-143

Grabowska-Joachimiak A, Kula A, Książczyk T, et al (2015) Chromosome landmarks and autosome-sex chromosome translocations in Rumex hastatulus, a plant with XX/XY1Y2 sex chromosome system. Chromosom Res 23:187–197. doi: 10.1007/s10577-014-9446-4

Gschwend AR, Yu Q, Tong EJ, et al (2012) Rapid divergence and expansion of the X chromosome in papaya. Proc Natl Acad Sci U S A 109:13716–21. doi: 10.1073/pnas.1121096109

Hernandez D, François P, Farinelli L, et al (2008) De novo bacterial genome sequencing: millions of very short reads assembled on a desktop computer. Genome Res 18:802–9. doi: 10.1101/gr.072033.107

Hobza R, Cegan R, Jesionek W, et al (2017) Impact of repetitive elements on the Y chromosome formation in plants. Genes (Basel). doi: 10.3390/genes8110302

Hobza R, Kubat Z, Cegan R, et al (2015) Impact of repetitive DNA on sex chromosome evolution in plants. Chromosom Res 23:561–570. doi: 10.1007/s10577-015-9496-2

Kejnovský E, Michalovova M, Steflova P, et al (2013) Expansion of Microsatellites on Evolutionary Young Y Chromosome. PLoS One 8:e45519. doi: 10.1371/journal.pone.0045519

Kihara H, Ono T (1923) Cytological Studies on Rumes L. I. Chromosomes of Rumes Acetosa L. Shokubutsugaku Zasshi 37:84–90. doi: 10.15281/jplantres1887.37.435\_84

Kubat Z, Zluvova J, Vogel I, et al (2014) Possible mechanisms responsible for absence of a retrotransposon family on a plant Y chromosome. New Phytol 202:662–678. doi: 10.1111/nph.12669

Lengerova M, Vyskot B (2001) Sex chromatin and nucleolar analyses in Rumex acetosa L. Protoplasma 217:147–153. doi: 10.1007/BF01283395

Macas J, Novák P, Pellicer J, 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. doi: 10.1371/journal.pone.0143424

Marchler-Bauer A, Bo Y, Han L, et al (2017) CDD/SPARCLE: functional classification of proteins via subfamily domain architectures. Nucleic Acids Res 45:D200–D203. doi: 10.1093/nar/gkw1129

Mariotti B, Manzano S, Kejnovský E, et al (2009) Accumulation of Y-specific satellite DNAs during the evolution of Rumex acetosa sex chromosomes. Mol Genet Genomics 281:249–259. doi: 10.1007/s00438-008-0405-7

Mariotti B, Navajas-Pérez R, Lozano R, et al (2006) Cloning and characterization of dispersed repetitive DNA derived from microdissected sex chromosomes of Rumex acetosa. Genome 49:114–121. doi: 10.1139/G05-089

Matsunaga S, Hizume M, Kawano S, Kuroiwa T (1994) Cytological Analyses in Melandrium album: Genome Size, Chromosome Size and Fluorescence in situ Hybridization. Cytologia (Tokyo) 59:135–141. doi: 10.1508/cytologia.59.135

Michalovova M, Kubat Z, Hobza R, et al (2015) Fully automated pipeline for detection of sex linked genes using RNA-Seq data. BMC Bioinformatics 16:78. doi: 10.1186/s12859-015-0509-0

Ming R, Bendahmane A, Renner SS (2011) Sex Chromosomes in Land Plants. Annu Rev Plant Biol 62:485–514. doi: 10.1146/annurev-arplant-042110-103914

Navajas-Pérez R, la Herrán R de, Jamilena M, et al (2005) Reduced Rates of Sequence Evolution of Y-Linked Satellite DNA in Rumex (Polygonaceae). J Mol Evol 60:391–399. doi: 10.1007/s00239-004-0199-0

Navajas-Pérez R, Schwarzacher T, de la Herrán R, et al (2006) The origin and evolution of the variability in a Y-specific satellite-DNA of Rumex acetosa and its relatives. Gene 368:61–71. doi: 10.1016/J.GENE.2005.10.013

Novák P, Neumann P, Macas J (2010) Graph-based clustering and characterization of repetitive sequences in next-generation sequencing data. BMC Bioinformatics 11:378. doi: 10.1186/1471-2105-11-378

Novak P, Neumann P, Pech J, et al (2013) RepeatExplorer: a Galaxy-based web server for genome-wide characterization of eukaryotic repetitive elements from next-generation sequence reads. Bioinformatics 29:792–793. doi: 10.1093/bioinformatics/btt054

Parker, S. J (1990) Sex-chromosome and sexual differentiation in flowering plants. Chromosom Today 10:187–198.

Puterova J, Kubat Z, Kejnovsky E, et al (2018) The slowdown of Y chromosome expansion in dioecious Silene latifolia due to DNA loss and male-specific silencing of retrotransposons. BMC Genomics 19:153. doi: 10.1186/s12864-018-4547-7

Puterova J, Razumova O, Martinek T, et al (2017) Satellite DNA and Transposable Elements in Seabuckthorn ( *Hippophae rhamnoides* ), a Dioecious Plant with Small Y and Large X Chromosomes. Genome Biol Evol 9:evw303. doi: 10.1093/gbe/evw303

Renner SS (2014) The relative and absolute frequencies of angiosperm sexual systems: Dioecy, monoecy, gynodioecy, and an updated online database. Am J Bot 101:1588–1596. doi: 10.3732/ajb.1400196

Sakamoto K, Ohmido N, Fukui K, et al (2000) Site-specific accumulation of a LINE-like retrotransposon in a sex chromosome of the dioecious plant Cannabis sativa. Plant Mol Biol 44:723–732. doi: 10.1023/A:1026574405717

Shibata F, Hizume M, Kuroki Y (1999) Chromosome painting of Y chromosomes and isolation of a Y chromosome-specific repetitive sequence in the dioecious plant Rumex acetosa. Chromosoma 108:266–270. doi: 10.1007/s004120050377

Shibata F, Hizume M, Kuroki Y (2000) Differentiation and the polymorphic nature of the Y chromosomes revealed by repetitive sequences in the dioecious plant, Rumex acetosa. Chromosom Res 8:229–236. doi: 10.1023/A:1009252913344

Šimková H, Svensson JT, Condamine P, et al (2008) Coupling amplified DNA from flow-sorted chromosomes to high-density SNP mapping in barley. BMC Genomics 9:294. doi: 10.1186/1471-2164-9-294

Sousa A, Bellot S, Fuchs J, et al (2016) Analysis of transposable elements and organellar DNA in male and female genomes of a species with a huge Y chromosome reveals distinct Y centromeres. Plant J 88:387–396. doi: 10.1111/tpj.13254

Steflova P, Hobza R, Vyskot B, Kejnovsky E (2014) Strong accumulation of chloroplast DNA in the Y chromosomes of Rumex acetosa and Silene latifolia. Cytogenet Genome Res 142:59–65. doi: 10.1159/000355212

Steflova P, Tokan V, Vogel I, et al (2013) Contrasting patterns of transposable element and satellite distribution on sex chromosomes (XY1Y2) in the dioecious plant Rumex acetosa. Genome Biol Evol 5:769–782. doi: 10.1093/gbe/evt049

Vagera J, Paulíková D, Doležel J (1994) The Development of Male and Female Regenerants by In Vitro Androgenesis in Dioecious Plant Melandrium album. Ann Bot 73:455–459. doi: 10.1006/anbo.1994.1056

Veuskens J, Ye D, Oliveira M, et al (1992) Sex determination in the dioecious *Melandrium album* : androgenic embryogenesis requires the presence of the X chromosome. Genome 35:8–16. doi: 10.1139/g92-002

Vrána J, Cápal P, Šimková H, et al (2016) Flow Analysis and Sorting of Plant Chromosomes. In: Current Protocols in Cytometry. John Wiley & Sons, Inc., Hoboken, NJ, USA, p 5.3.1-5.3.43

Vyskot B, Hobza R (2004) Gender in plants: sex chromosomes are emerging from the fog. Trends Genet 20:432–438. doi: 10.1016/J.TIG.2004.06.006