Satellite DNA and Transposable Elements in Seabuckthorn (*Hippophae rhamnoides*), a Dioecious Plant with Small Y and Large X Chromosomes

Janka Puterova[1,2], Olga Razumova[3], Tomas Martinek[2], Oleg Alexandrov[3], Mikhail Divashuk[3], Zdenek Kubat[1], Roman Hobza[1,4], Gennady Karlov[3,5], and Eduard Kejnovsky[1,*]

[1]Department of Plant Developmental Genetics, Institute of Biophysics, Academy of Sciences of the Czech Republic, Brno, Czech Republic

[2]Department of Information Systems, Faculty of Information Technology, Brno University of Technology, Brno, Czech Republic

[3]Centre for Molecular Biotechnology, Russian State Agrarian University – Moscow Timiryazev Agricultural Academy, Moscow, Russia

[4]Institute of Experimental Botany, Center of the Region Haná for Biotechnological and Agricultural Research, Olomouc, Czech Republic

[5]All-Russia Research Institute of Agricultural Biotechnology, Moscow, Russia

*Corresponding author: E-mail: kejnovsk@ibp.cz.

Accepted: January 3, 2017

Data deposition: This project has been deposited in European Nucleotide Archive (ENA) under study accession PRJEB14480. http://www.ebi.ac.uk/ena/data/view/PRJEB14480.

Abstract

Seabuckthorn (*Hippophae rhamnoides*) is a dioecious shrub commonly used in the pharmaceutical, cosmetic, and environmental industry as a source of oil, minerals and vitamins. In this study, we analyzed the transposable elements and satellites in its genome. We carried out Illumina DNA sequencing and reconstructed the main repetitive DNA sequences. For data analysis, we developed a new bioinformatics approach for advanced satellite DNA analysis and showed that about 25% of the genome consists of satellite DNA and about 24% is formed of transposable elements, dominated by Ty3/*Gypsy* and Ty1/*Copia* LTR retrotransposons. FISH mapping revealed X chromosome-accumulated, Y chromosome-specific or both sex chromosomes-accumulated satellites but most satellites were found on autosomes. Transposable elements were located mostly in the subtelomeres of all chromosomes. The 5S rDNA and 45S rDNA were localized on one autosomal locus each. Although we demonstrated the small size of the Y chromosome of the seabuckthorn and accumulated satellite DNA there, we were unable to estimate the age and extent of the Y chromosome degeneration. Analysis of dioecious relatives such as Shepherdia would shed more light on the evolution of these sex chromosomes.

Key words: sex chromosomes, genome composition, chromosomal localization, repetitive DNA.

Introduction

Seabuckthorn (*Hippophae rhamnoides*) is a hardy, deciduous dioecious shrub belonging to the Elaeagnaceae family with a natural habitat extending widely across Europe and Asia. It is used in traditional Chinese, Tibetan and Siberian medicine and has special characteristics exploitable in biotechnology, pharmaceutical and cosmetic sciences, as a source of oil, minerals and vitamins. The size of seabuckthorn genome is ~2.55 Gbp/2C (Zhou et al. 2010) but there is a dearth of information on its composition. The ribosomal DNA ITS regions were compared among *H. rhamnoides* ssp chinensis from different geographical areas of China and showed distinct genetic variation (Chen et al 2010). RAPD markers (Sharma et al. 2010) were identified with the aim of determining the sex of individuals. Cytogenetic analysis is represented only by the older works of Shchapov (1979) and Rousi and Arohonka (1980) who both determined the diploid chromosome number $2n = 24$. Shchapov (1979) revealed the small Y and large X chromosomes. Seabuckthorn transcriptome has been analyzed recently providing a resource for gene discovery and development of molecular markers (Ghangal et al. 2013).

Sex chromosomes have evolved repeatedly and independently in the plant kingdom with different age and degree of degeneration shown in various dioecious species (Ming et al.

2011; Hobza and Vyskot 2015; Charlesworth 2016). The evolution of the Y chromosomes is characterized by gene erosion/loss and accumulation of repetitive DNA (Kejnovsky et al. 2009). The most studied dioecious model species with heteromorphic sex chromosomes are white campion (*Silene latifolia*, Kejnovsky and Vyskot 2010), sorrel (*Rumex acetosa*, Steflova et al. 2013; *R. hastatulus*, Hough et al. 2014), ivy gourd (*Coccinia grandis*, Sousa et al. 2013), and members of the Cannabaceae family (*Humulus lupulus*, Divashuk et al. 2011; *H. japonicus*, Alexandrov et al. 2012; *Cannabis sativa*, Divashuk et al. 2014).

The majority of large plant genomes are formed of repetitive DNA, mostly by transposable elements and tandem repeats (satellite DNA). The processes of repetitive DNA amplification and elimination are only partially understood. Turnover of repeats is high and corresponds only to million of years (Lim et al. 2007). The localization of repetitive DNA on sex chromosomes is different from that of autosomes, reflecting different repeat dynamics, especially on the nonrecombining regions of the Y chromosomes (Kejnovsky et al. 2009). Satellite DNA has mostly discrete localization in the genome and some satellites are thus Y chromosome-specific (Mariotti et al. 2009). In contrast, transposable elements have more homogenous distribution and are only slightly enriched on the Y chromosome (Charlesworth 1991; Cermak et al. 2008) or alternatively absent on the Y chromosome as shown in *Silene latifolia* (Cermak et al. 2008; Kubat et al. 2014) and *Rumex acetosa* (Steflova et al. 2013) despite their presence in the rest of genome. The striking example is the large Y chromosome of the dioecious plant *Coccinia grandis* showing accumulation of transposable elements, satellites, and organellar DNA (Souza et al. 2016). One review published recently discusses the role of repetitive DNA in the evolution of sex chromosomes and includes a database of transposable elements of dioecious plants (Li et al. 2016a, 2016b).

In this study, we analyzed the transposable elements and satellites in the seabuckthorn genome and determined the chromosomal localization of these repeats. We showed that seabuckthorn has an XY system with large X and small Y chromosomes.

Materials and Methods

Illumina Sequencing

DNA isolation from male (Pollinator 1) and female (cv "Botanicheskaya lyubitelskaya") plants was carried out according to Doyle and Doyle (1990). One Illumina MiSeq sequencing run was performed for each male and female genomic DNA. The voucher specimen of the plants used in the study was kept for record in the herbarium (AT) of Department of Botany and Breeding of Horticultural Crops of the Russian State Agrarian University – MTAA (Voucher No.5470). Sequencing reads were analyzed by quality control tool FastQC (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/; last accessed January 4, 2017) followed by quality filtering based on the sequence quality score, adaptors trimming, filtering out short or unpaired sequences and trimming all reads to lengths of 230 nucleotides using the Trimmomatic tool (Bolger et al. 2014), leading to 1,848,543 male and 1,863,670 female paired-end reads. Quality-filtered reads were randomly sampled to 415,650 paired-end reads for both male and female individuals and the reads were merged together (totally 1,662,600 reads). As the nuclear DNA content of *H. rhamnoides* reported in Zhou et al. (2010) was determined to be ~2.61/2C pg (without detailed specification of male or female) we converted it to genome size (in bp) using following formula (Doležel et al. 2003): $g = \text{DNA content (pg)} \times (0.978 \times 10^9)$, resulting into ~2.55 Gbp/2C, our samples represent ~30% of haploid genome. Genome coverage was calculated as follow: $cov = (r \times l)/g$, where r corresponds to number of reads used in our analysis, l to read length and g to haploid genome size of *H. rhamnoides*.

Repeat Identification and Annotation

In order to identify repetitive sequences in the *H. rhamnoides* genome we employed comparative graph-based clustering analysis of sequenced reads by RepeatExplorer pipeline (Novak et al. 2013). Only clusters containing at least 0.01% of all clustered reads were considered and they corresponded to 58.5% of the genome. These were further manually characterized based on the similarity search results from RepeatMasker (http://www.repeatmasker.org; last accessed January 4, 2017) against Viridiplantae database and blastn and blastx (Altschul et al. 1990) against GenBank nr (Benson et al. 2009), which are part of the RepeatExplorer output. Cluster shapes were also used for repeat identification as tandem repeats with monomer longer than read length have typical donut-shaped clusters (Novak et al. 2010). Additionally, advanced analysis of satellite sequences, described in the section Satellite DNA sequences analysis, was used in the manual annotation of clusters.

Structural Annotation of LTR Retrotransposons

We reconstructed several Ty3/*Gypsy* and Ty1/*Copia* retrotransposons. The reconstruction comprised several steps. First, clusters belonging to particular element were visualized in SeqGrapheR (https://cran.r-project.org/web/packages/SeqGrapheR/index.html; last accessed January 4, 2017) program and contigs which together covered the whole elements were selected. These contigs were searched for occurrences of protein domains (GAG, RT, RH, AP, INT) by querying them to CDD (Marchler-Bauer et al. 2015). We then did multiple sequence alignment to create a consensus sequence of these contigs using progressive pairwise alignment implemented in Geneious 8.1.7 (http://www.geneious.com; last accessed January 4, 2017, Kearse et al. 2012). If necessary, resulting alignments were manually modified with respect to the order

of domains for particular type of transposable element. The consensus sequence of reconstructed elements was then searched for the structural motif characteristics (ORFs and LTRs). Possible ORFs were detected by ORF Finder (https://www.ncbi.nlm.nih.gov/orffinder/; last accessed January 4, 2017). LTRs were determined on the basis of shape of a cluster and the element's coverage. Male and female coverage of reconstructed elements was determined by mapping reads which formed a current element to its consensus sequence using BowTie2 tool (Langmead and Salzberg 2012). Structural features and male and female coverage of reconstructed elements were visualized by custom R script and graph layouts of reconstructed elements were depicted by SeqGrapheR.

Phylogeny and Classification

Firstly, we created custom databases of plant LTR retrotransposon RT domains from sequences available in TREP (Wicker et al. 2002) and GyDB (Llorens et al. 2011) databases, independently for Ty3/*Gypsy* and Ty1/*Copia* retrotransposons. Contigs corresponding to retrotransposons were examined for the presence of a reverse transcriptase domain and Ty3/*Gypsy* and Ty1/*Copia* cores of RT domains were trimmed from these contigs based on the exact localization designated by CDD (Marchler-Bauer et al. 2015). Cores of RT domains were aligned by MUSCLE algorithm (Edgar 2004) together with our custom-made database of RT domains, and the resulting multiple sequence alignment was used as an input to create Neighbor-Joining tree (Saitou and Nei 1987) with Jukes-Cantor distance model using Geneious 8.1.7 (http://www.geneious.com; last accessed January 4, 2017, Kearse et al. 2012).

Preparation of Chromosomes and Probes and Fluorescence *In Situ* Hybridization

For chromosome preparations vegetatively propagated for commercial use, male ("Pollinator 1" and "Pollinator 3") and female (cv "Lomonosovskaya" and cv "Botanicheskaya ljubitelskaya") plants were used. Plant material was kindly provided by Dr G. Boyko, Lomonosov Moscow State University. The root tips were harvested separately from the individual male and female plants grown in pots. The harvested root tips were immediately pre-treated with a 2 mM aqueous solution of 8-hydroxyquinoline for 6 h at 20 °C. A 3:1 ethanol/glacial acetic acid (v/v) mix was used for fixation. Meristems 2 mm long were cut from the fixed root tips and digested in 10 µl enzyme solution [0.5% cellulase Onozuka R-10 (Serva, Germany) and 0.5% pectolyase Y-23 (Seishin Corp., Japan)] in 10 mM citrate buffer (pH = 4.9) for 2.5 h at 37 °C. The suspended cells were used for chromosome preparation as described by Kirov et al. (2014). The quality of spreads was assessed microscopically using phase-contrast and only preparations with at least 20 well-spread metaphases were used.

Probes for fluorescence *in situ* hybridization were generated using PCR-DIG Labeling Mix PLUS (Roche Diagnostics Gmbh) or by Biotin-11-dUTP 1/3 PCR labeling Mix (ZAO Sileks, Moscow). Primers for RT domain of selected transposable elements and determined monomer sequence of satellites were designed by Primer3 tool (Untergasser et al. 2012), were synthesized by ZAO "Syntol" (Moscow). These are available in supplementary table S1, Supplementary Material online. The pTa71 (45S rDNA) and pCT4.2 (5S rDNA) clones labeled by DIG-Nick translation kit were also used (Gerlach and Bedbrook 1979; Campell et al. 1992).

FISH experiments were performed as described in Alexandrov and Karlov (2016). For digoxigenin and biotin detection, slides were incubated with anti-DIG-FITC conjugate (Roche) and/or streptavidin-Cy3 conjugate (Sigma). The chromosomes were counterstained with DAPI (2 µg/ml) and mounted in Vectashield (Vector). An AxioImager M1 fluorescent microscope (Zeiss) was used to observe metaphase plates with fluorescent signals that were photographed with a monochrome AxioCam MRm CCD camera and visualized using Axiovision software (Zeiss).

Satellite DNA Sequences Analysis

As the seabuckthorn genome is abundant in satellite DNA and manual inspection would be exhaustive, we developed a custom bioinformatics approach which extended the basic analysis of RepeatExplorer tool. As an input the satellite clusters identified by RepeatExplorer are required. It is highly recommended to do manual inspection of these clusters and verify their structure and interaction with other clusters based on similarities among other clusters and pair-end reads connections. Our approach consisted of three basic steps.

(i) *Detection of satellite monomers*: First, assembled contigs of selected clusters were extracted from RepeatExplorer output and for each contig the monomer length was estimated from the distribution of distances between the same k-mers. The resulting monomer sequence was then extracted from the most covered part of the contig of previously determined length. Only the monomers with clearly distinguishable length, longer than 100 bp and reaching average coverage 50x and more were taken into account.

(ii) *Estimation of satellite families composition in genome and their annotation*: First, all to all monomer similarity was calculated. In order to do alignment of tandemly repeated monomers correctly (offsets between monomers are not known) we used one monomer as a subject and two copies in a row of the second monomer as a query. The similarity between monomers was then determined based on semiglobal alignment. To estimate the composition of satellite families in the genome, we clustered the monomer's similarity matrix using UPGMA method. The resulting dendrogram was then inspected by the user and cut off at the level that best discriminated the individual

families (usually 70-85% of monomer identity). Identified families were visualized by the algorithm described by Fruchterman and Reingold (1991) implemented in igraph library and only connections that exceeded specified cut-off were considered and depicted. Secondly, to annotate identified families, all monomers were searched for similarity hits with sequences in the public nucleotide database and PlantSat database (Macas et al. 2002) using blastn (Altschul et al. 1990) with word size set to 11. Only results with an e-value lower than 10^{-20} were considered as significant. Finally, to depict satellite diversity inside the family, we chose the most covered monomer as a reference and mapped all reads belonging to the family onto its reference using BWA-MEM mapping tool (Li 2013). Conservation of different parts of the monomer was depicted using sequence logo created by WebLogo (Crooks et al. 2004) tool.

(iii) *Visualization of satellite families homogeneity:* First, the relative abundance of male and female reads was calculated in each tandem repeat family. This enabled us to predict their presence in sex chromosomes. We visualized the satellite homogeneity using the following procedure: reads from each identified family were merged together and sampled randomly to limit the maximum number of reads to speed up the following analysis. Similarity of sampled reads from all families was calculated using the megablast tool (Camacho et al. 2009) that performed all against all sequence comparison. Pairs of reads that met specific similarity threshold (70% sequence identity over at least 55% of sequence length) were further used for graph construction and visualization. Male and female reads were distinguished by color (male—blue, female—red), tandem repeat families were highlighted by different colors and the algorithm by Fruchterman and Reingold (1991) was used to depict the results. Additionally, graphs for selected families were refined with similarity thresholds ranging from 70% to 95% sequence identity to show satellite composition more clearly. Each satellite falling within individual satellite family was marked by a different color.

Results

Genomic Composition

We performed one Illumina MiSeq platform sequencing run for each male and female genomic DNA followed by graph-based clustering of reads and characterization of repetitive sequences by RepeatExplorer (Novak et al. 2013). All 223 clusters (with more than 167 reads) contained 973,049 reads corresponding to 58.5% of genome (fig. 1) and their identification showed that dominant (first) clusters corresponded to satellite DNA followed by Ty3/*Gypsy* and Ty1/*Copia* LTR retrotransposons. One cluster (CL97) corresponded to 5S rDNA, two clusters (CL40, CL71) to 45S rDNA and 15 clusters to chloroplast DNA (cpDNA). Although the majority of chloroplast DNA reads probably originated from contaminating cpDNA, some proportion could come from nuclear cpDNA insertions (NUPTs).

We identified main types of repetitive DNA and their genome proportions in male and female individuals (table 1). All transposable elements represented together 24% of male and 23% of female genome. Ty1/*Copia* retrotransposons formed 12%, Ty3/*Gypsy* retrotransposons 11% and DNA transposons 1.5% of male genome. The most abundant among Ty1/*Copia* retrotransposons were Angela/Tork and Ale/Retrofit, among Ty3/*Gypsy* retrotransposons Athila and chromoviruses dominated. No LINE elements were found in the whole seabuckthorn genome. Satellites together comprised about 27% of male and 24% of female genomes. The 45S rDNA formed 0.7% of both male and female genomes and 5S rDNA represented 0.2% of both male and female genomes.

Transposable Elements

To determine the phylogenetic relationships of Ty1/*Copia* and Ty3/*Gypsy* retrotransposons, we aligned their reverse transcriptase (RT) domains from individual clusters and constructed the phylogenetic trees. Both Ty3/*Gypsy* (fig. 2*A*) and Ty1/*Copia* (fig. 2*B*) trees contained families identified in our clusters (in red) mixed with representatives of known subfamilies of Ty1/*Copia* or Ty3/*Gypsy* from other plant species (in black). Among Ty3/*Gypsy* retrotransposons, we identified five clusters containing Athila subfamilies, one CRM subfamily, one Galadriel, one Reina and one Tat/Ogre subfamily (fig. 2*A*). Among Ty1/*Copia* retrotransposons, we found four subfamilies of Ale/Retrofit, four Angela/Tork subfamilies, one Maximus/SIRE subfamily, two TAR subfamilies and two Ivana/Oryco subfamilies (fig. 2*B*). The Angela/Tork and Ale/Retrofit subfamilies showed higher variability while Athila subfamilies were homogenous. Highest homogeneity were shown by chromoviruses where all reads were assembled into a single cluster for CRM, Galadriel and Reina families (fig. 2*A*).

We reconstructed the structure of the main Ty3/*Gypsy* and Ty1/*Copia* subfamilies (fig. 3) and identified all main features such as *gag* and *pol* genes (with all domains) and long terminal repeats (LTRs). In some retrotransposons (CL6, CL16) LTR regions were assembled into one long terminal repeat while in other clusters (CL7, CL27) right and left LTR were distinguished. This may be a consequence of lower or higher mutual diversity of LTRs in one element, and could correspond to age differences of elements. Graph layouts (right part of fig. 3) show the variability of specific parts of elements as well as alternative variants of elements, e.g., potential spliced variant (Novak et al. 2010). The similar coverage of elements by male and female reads indicates that elements are present on all chromosomes without accumulation/absence on

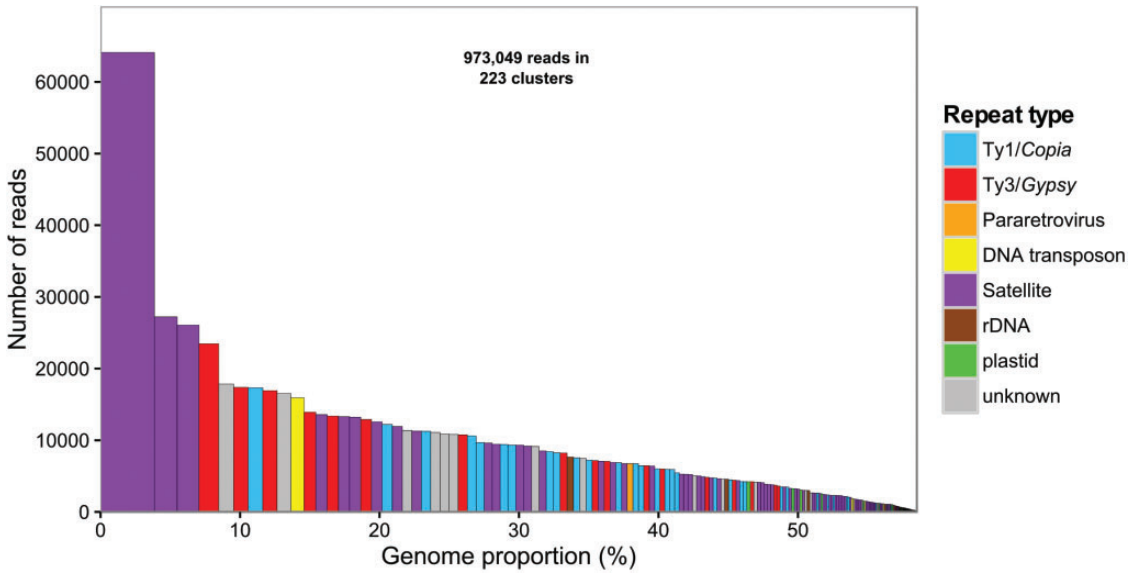


Fig. 1.—Repeat composition of clusters and their genomic proportions. Each column corresponds to one cluster and repeat types are distinguished by colors. The height of columns represents number of reads in each cluster, the width of column indicate genomic proportion of cluster.

Table 1
Repeat Composition in *Hippophae rhamnoides* Genome

Classification			Genome Proportion (%)	
Repeat Type	**Super Family**	**Family**	**Male**	**Female**
LTR retroelements	Ty1/Copia	Angela/Tork	4.83	4.90
		Ale/Retrofit	4.93	4.38
		TAR	1.34	1.06
		Maximus/SIRE	0.44	0.57
		Ivana/Oryco	0.25	0.23
		Total Ty1/*Copia*	11.79	11.15
	Ty3/Gypsy	Athila	6.39	5.36
		Chromovirus—CRM	2.98	3.58
		Chromovirus—Galadriel	1.28	0.80
		Chromovirus—others	0.27	0.31
		Chromovirus—Reina	0.06	0.04
		Tat/Ogre	0.05	0.05
		Total Ty3/*Gypsy*	11.04	10.15
DNA transposons			1.52	1.46
Total transposable elements			24.35	22.76
Pararetrovirus			0.48	0.59
rDNA	45S		0.77	0.69
	5S		0.20	0.16
Satellites			26.92	23.74
All repetitive elements			52.72	47.94
Unclassified			6.96	11.39
Low/single copy			38.96	39.50
Plastids			1.36	1.17

Note.—Types of repetitive DNA and their genome proportions.

the X or Y chromosome. Some elements had uninterrupted ORF corresponding to *gag* and *pol* (CL7, CL27, and CL43) and hence they can be active. Interruption of ORFs in other elements may have been caused by assembling errors during reconstruction (CL6, CL16, and CL37).

Satellite DNA

We developed a new bioinformatics approach for detailed analysis of satellite DNA in genomes. This method includes: (i) identification of satellite monomers based on distribution of distances of k-mers in assembled contigs, (ii) clustering of monomers allowing identification and annotation of satellite families in genome, and (iii) visualization of satellites homogeneity and male/female composition allowing better prediction of their localization with respect to sex chromosomes. Detailed description of the whole procedure is available in the section Materials and Methods and in supplementary figure S4, Supplementary Material online.

We utilized this approach for analysis of the seabuckthorn genome, but it is generally applicable in genomic studies of other species as well. As an input we used the 38 largest manually inspected satellite clusters from RepeatExplorer output extended by five smaller clusters with potentially interesting chromosomal localization (X, Y chromosomes). All clusters were grouped into 12 main superclusters that correspond to the 12 main families of satellite DNA in the seabuckthorn genome. Satellites were named HRTR1-HRTR12 (supplementary fig. S1, Supplementary Material online and table 2). Copy number of individual satellite families was determined based on following formula: $cn = [(s \times l)/m]/cov$, where s represents number of reads of individual satellite family, l corresponds to

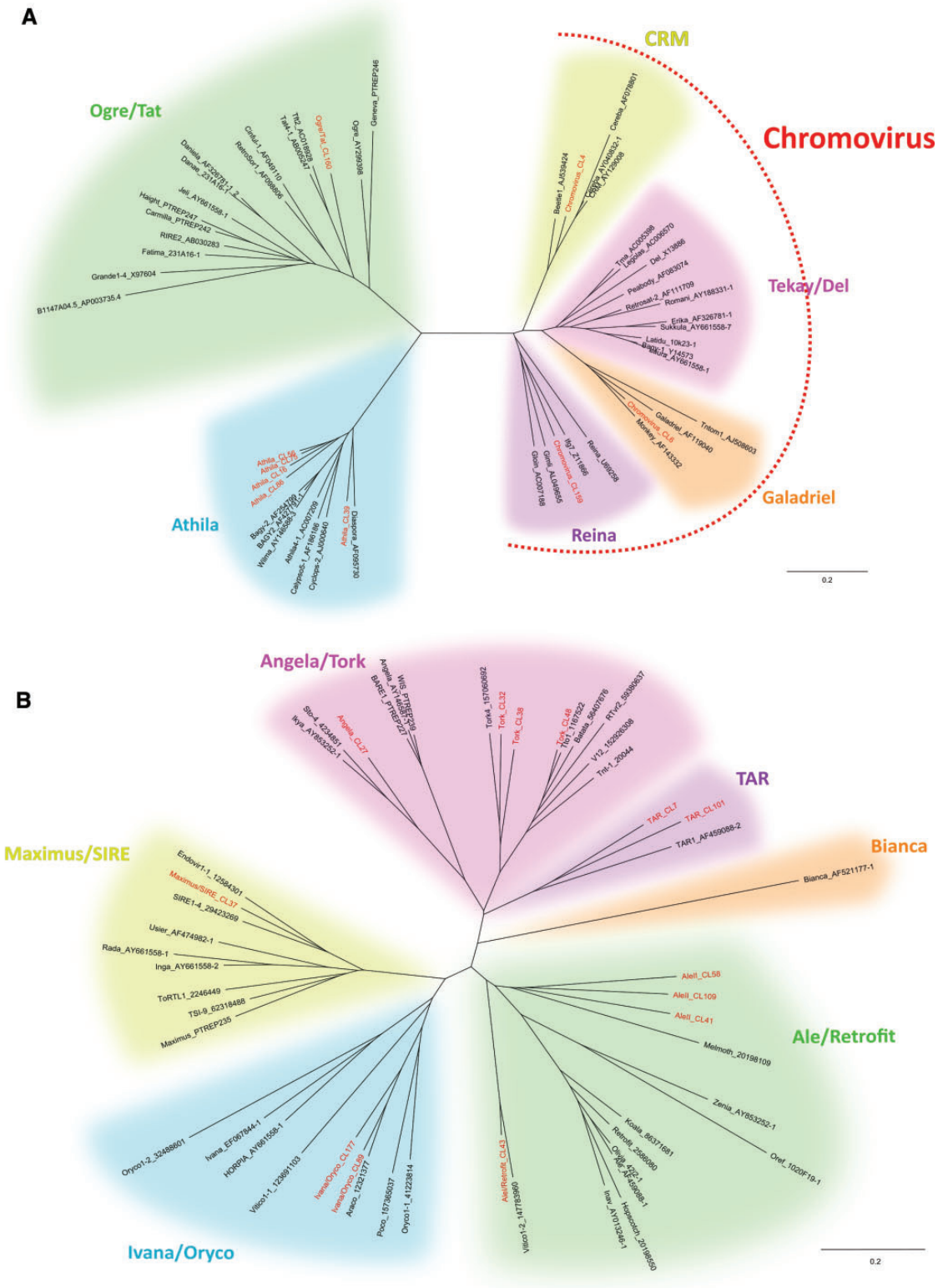


Fig. 2.—Phylogenetic trees of *Hippophae rhamnoides* Ty3/*Gypsy* (*A*) and Ty1/*Copia* (*B*) retrotransposons based on reverse transcriptase sequences. RT domains of retrotransposons reconstructed from Illumina reads in this study are in red, representative RT domains of retrotransposons from other plant species (from TREP and GyDB) are in black. Individual families are highlighted by different colors.

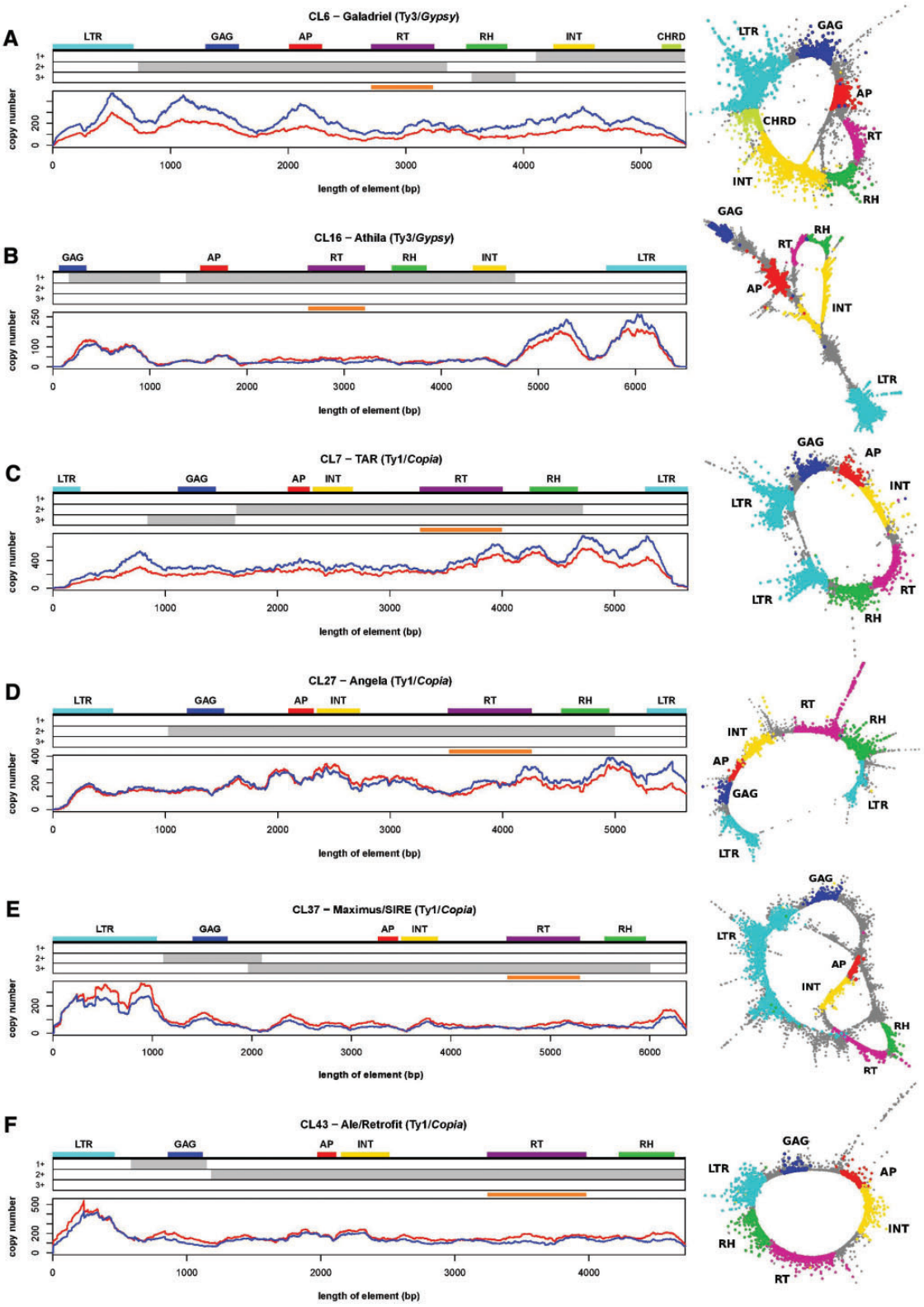


Fig. 3.—Comparison of structure of selected retrotransposon families in *Hippophae rhamnoides*. Graphs of coverage by male (in blue) and female (in red) genomic reads are showed under the structure of Ty3/*Gypsy* (*A*, *B*) and Ty1/*Copia* (*C*–*F*) elements shown in phylogenetic tree (fig. 2). Graph layouts on the right are visualized by SeqGrapheR program (http://cran.rproject.org/web/packages/SeqGrapheR/index.html). Protein domains and possible LTRs are distinguished by colors, found possible different three ORFs are marked by grey rectangles and orange line represents sequence for probes used for FISH.

Table 2
Main Satellite Families in *Hippophae rhamnoides* Genome

Name	Number of Reads	Localization	Monomer Length	M (%)	F (%)	Copy Number
HRTR1	129843	Strong signal on six pairs of small autosomes and weak signal on one pair of small autosomes	363	59.90	40.10	82270
HRTR2	60455	X and Y chromosome and weak signal on one pair of large and one pair of small autosomes	541	43.03	56.97	25702
HRTR3	46881	Dispersed signal on two large autosomal pairs	656	49.60	50.40	16437
HRTR4	27219	One pair of large and one pair of small autosomes	720	51.30	48.70	8695
HRTR5	23060	One pair of small autosomes	819	57.61	42.39	6476
HRTR6	19415	Three pairs of small autosomes	198[a]	53.67	46.33	5784[b]
HRTR7	14861	One pair of large autosomes and one pair of small autosomes	493[a]	68.38	31.62	4828[b]
HRTR8	12570	X chromosome and weak signal on one pair of small autosomes	826	35.06	64.94	3500
HRTR9	11155	One pair of small autosomes	354	69.52	30.48	7248
HRTR10	7476	Centromere of one pair of small autosomes	940	49.80	50.20	1829
HRTR11	4088	One pair of small autosomes	643	66.78	33.22	1462
HRTR12	1718	Y chromosome	257	100.00	0.00	1538

Note.—Names, monomer lengths, copy numbers, chromosomal localizations, and genome proportions.
[a]Shared length of the monomer in the family.
[b]Estimated based on average monomer length. 772 bp for HRTR6 and 708 bp for HRTR7.

read length, *m* represents estimated monomer length for satellite family and *cov* is genome coverage. Sequence logos show the monomer sequences of the main satellites and the sequence variability (supplementary fig. S2*A–L*, Supplementary Material online). Only HRTR1 and HRTR12 showed significant similarity hits with blast nucleotide (nr/nt) database (to previously deposited microsatellite markers of *H. rhamnoides*). There were no significant hits with PlantSat database for all satellite groups.

Based on our detailed analysis of HRTR6 and HRTR7, sharing small part of monomers (supplementary fig. S3*C*, Supplementary Material online), we decided to retain them as two separate tandem repeat families instead of one. These two families were very divergent and each showed variability in monomer' length (HRTR6: 730–810 bp, HRTR7: 475–830 bp). Monomers in each family had a common sequence (HRTR6: 198 bp, HRTR7: 493 bp) while other parts of monomers were significantly different from each other. For this reason, we only created sequence logos for the shared part of monomers for each family (supplementary fig. S2*F* and *G*, Supplementary Material online).

Male versus Female Comparison

To compare male and female genomes and to predict which repetitive DNA is specific for or accumulated on the X and Y chromosomes, we plotted the numbers of male versus female reads corresponding to individual clusters (fig. 4). This analysis involved all 223 clusters. The majority of clusters was located on the diagonal and these corresponded to transposable elements, rDNA and some satellites. However, some clusters containing satellites were enriched or even specific for males and represented potential Y-specific repeats. Other repeats, mostly satellites, were more abundant in females which could reflect their enrichment or specific localization on the X chromosome.

The greatest differences in composition of male and female reads were observed in satellites (five clusters located in the left; fig. 4). Detailed analysis showed that one of these (CL123—HRTR12) formed an isolated family composed of male reads only which suggests its localization only on the Y chromosome (fig. 5). The other four male biased satellites represented either a variant of a specific widespread cluster with Y chromosome presence (CL99 and CL144—HRTR2) or a satellite with a minor presence on the Y chromosome (CL150—HRTR1 and CL132—HRTR3). Eight satellites contained more female than male reads (2:1) indicating its localization on the X chromosome (female has two X chromosomes, male only one). HRTR2 satellite also contained more female than male reads but the ratio was 1.3 to 1 which could be explained by the localization on both sex chromosomes with greater abundance on the X than on the Y chromosome (fig. 5). Most other satellites had similar abundance of male and female reads, suggesting their localization (at least mostly) on autosomes.

Chromosomal Localization of Transposable Elements and Satellites

For determination of the chromosomal localization of transposable elements and satellites in seabuckthorn, we prepared probes representing reverse transcriptase region of individual TE families or part of a satellite monomer (supplementary fig.

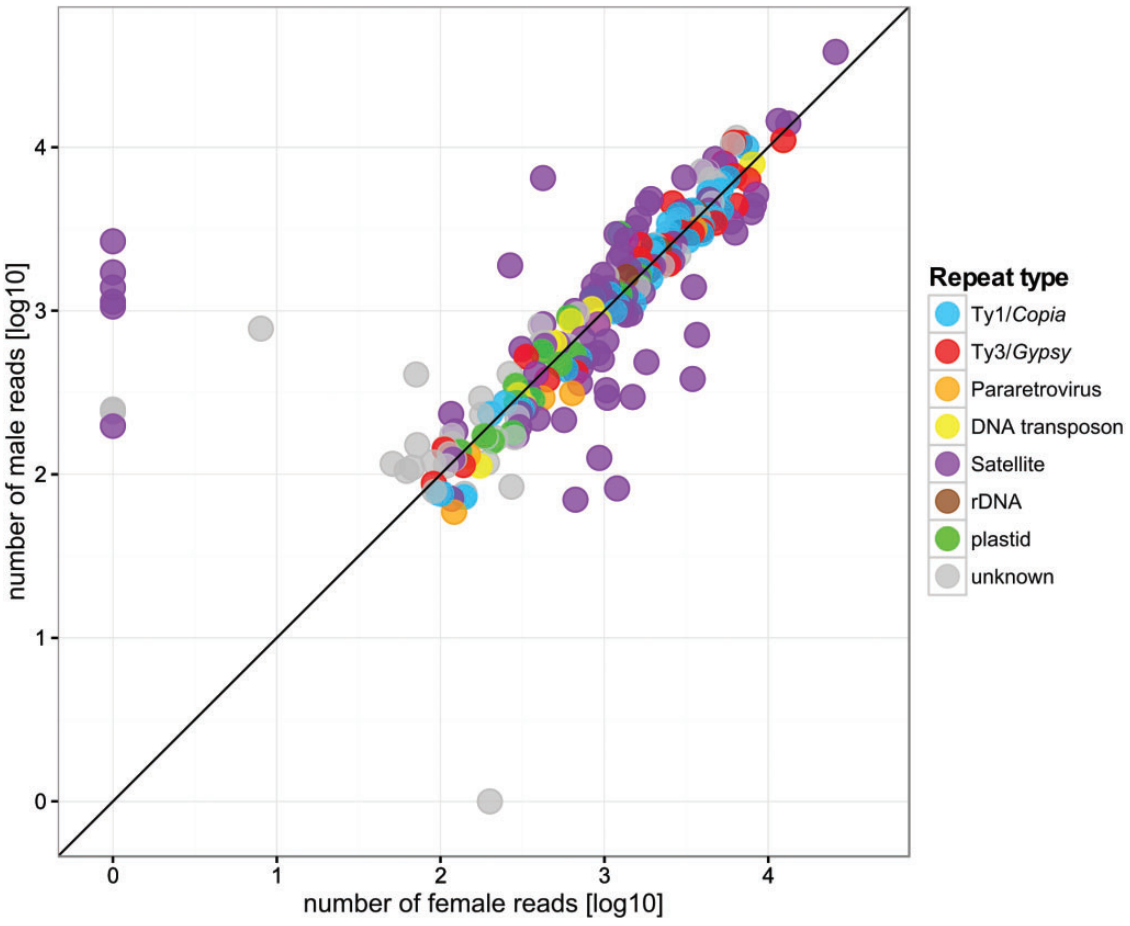


Fig. 4.—Comparison of repeats in male and female of *Hippophae rhamnoides*. Number of male versus female reads corresponding to individual clusters. Each circle in plot represents one cluster. Repeat types are marked by different color. Clusters in left upper part of graph are enriched (or specific) for males and thus potentially located on the Y chromosome while clusters in the right bottom part are enriched in female and thus potentially located on the X chromosome.

S1, Supplementary Material online) and used them for fluorescence *in situ* hybridization (FISH). In all FISH experiments we used both male (Pollinator 1, Leningradskaya region) and female (cv "Botanicheskaya lyubitelskaya") metaphases from plants that was used for sequencing. FISH experiments were also expanded to male ("Pollinator 3" Kaliningrad region) and female (cv "Lomonosovskaya"). In all ecotypes, we got the same results with X and Y.

FISH with satellite DNA showed various localization patterns on metaphase chromosomes of *H. rhamnoides* (fig. 6). The HRTR2, HRTR8 and HRTR12 show the sex specific or accumulation pattern of hybridization, while for HRTR3, HRTR4, HRTR5, HRTR6, HRTR7, HRTR9, HRTR10, and HRTR11 the hybridization patterns was the same for male as well as for female. The HRTR1 satellite hybridized mainly to heterochromatic arms of six pairs of small autosomes and weakly on one more pair of small autosomes (fig. 6*A* and *B*). In addition, a weak signal was detected distal to centromere on one arm of one large chromosome (chromosome X) in male (fig. 6*A*) and two large chromosomes in female (fig. 6*B*). The HRTR2 satellite gave a strong FISH signal on one large chromosome (chromosome X) and on one small chromosome (chromosome Y) in male (fig. 6*C*) and a strong FISH signal on two large chromosomes (chromosome X) in female (fig. 6*D*). Also a weak signal on the centromeric region of a pair of large and a pair of small autosomes was detected in both sexes. The HRTR3 satellite was localized on two large autosomal pairs with the FISH signal dispersed along these chromosomes (fig. 6*E*). The HRTR4 localized on one pair of large and on one pair of small autosomes (fig. 6*F*). The HRTR5 signal was detected on one pair of small autosomes only (fig. 6*G*). HRTR6 gave a strong signal on one autosomal pair and a weaker signals on two autosomal pairs (fig. 6*H*). The HRTR7 showed two sites of hybridization on one arm of a pair of large autosomes and on the centromeric region of a pair of small autosomes (fig. 6*I*). The HRTR8 hybridized mainly to the one large chromosome

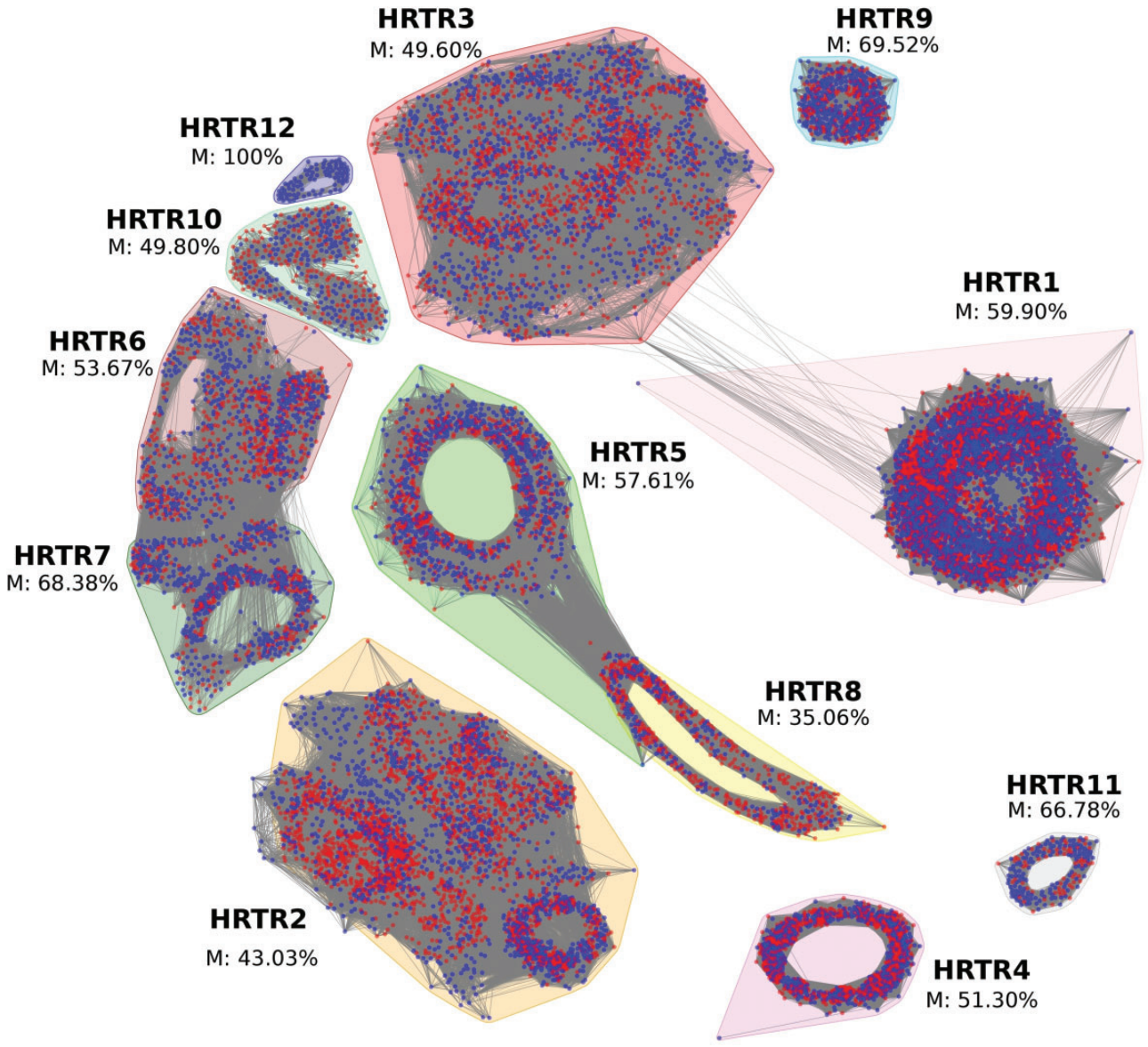


Fig. 5.—Visualization of male/female reads homogeneity in satellite families. Graph nodes correspond to sequenced reads and edges connect overlapping reads with more than 70% of sequence identity over at least 55% read length. Distances between reads are inversely proportional to their sequence similarity. Male reads are labeled by blue and female reads by red color. Individual families are highlighted by different colors. Please note HRTR12 family that is composed of male reads only assuming to be Y-specific.

(chromosome X) in male (Fig. 6*J*) and to the two large chromosomes (chromosomes X) in female (fig. 6*K*). A weak signal was also detected on one pair of small autosomes. The HRTR9, HRTR10, and HRTR11 were localized on one pair of small autosomes each (fig. 6*L–N*). The HRTR12 hybridized specifically to the small chromosome (Y chromosome) (fig. 6*O*) in male and no signal was detected in female (fig. 6*D*). The FISH signal intensity from HRTRs on X chromosomes varied depending on genotype.

Localization of the HRTR1 and the Y-specific (HRTR12), X-accumulated (HRTR8) and X and Y-accumulated (HRTR2) satellites on sex chromosomes was demonstrated by bicolor FISH using combinations of these probes and is summarized in a scheme (fig. 7). This together with specific or enriched representation of clusters in male and female (figs. 4 and 5), clearly demonstrates that *H. rhamnoides* has heteromorphic sex chromosomes (XY system) with large X and the small Y chromosomes.

We also mapped ribosomal genes. 45S rDNA was localized on one pair of small autosomes (fig. 8*A*) and 5S rDNA was localized on another pair of autosomes (fig. 8*B*). FISH with probes derived from transposable elements showed that

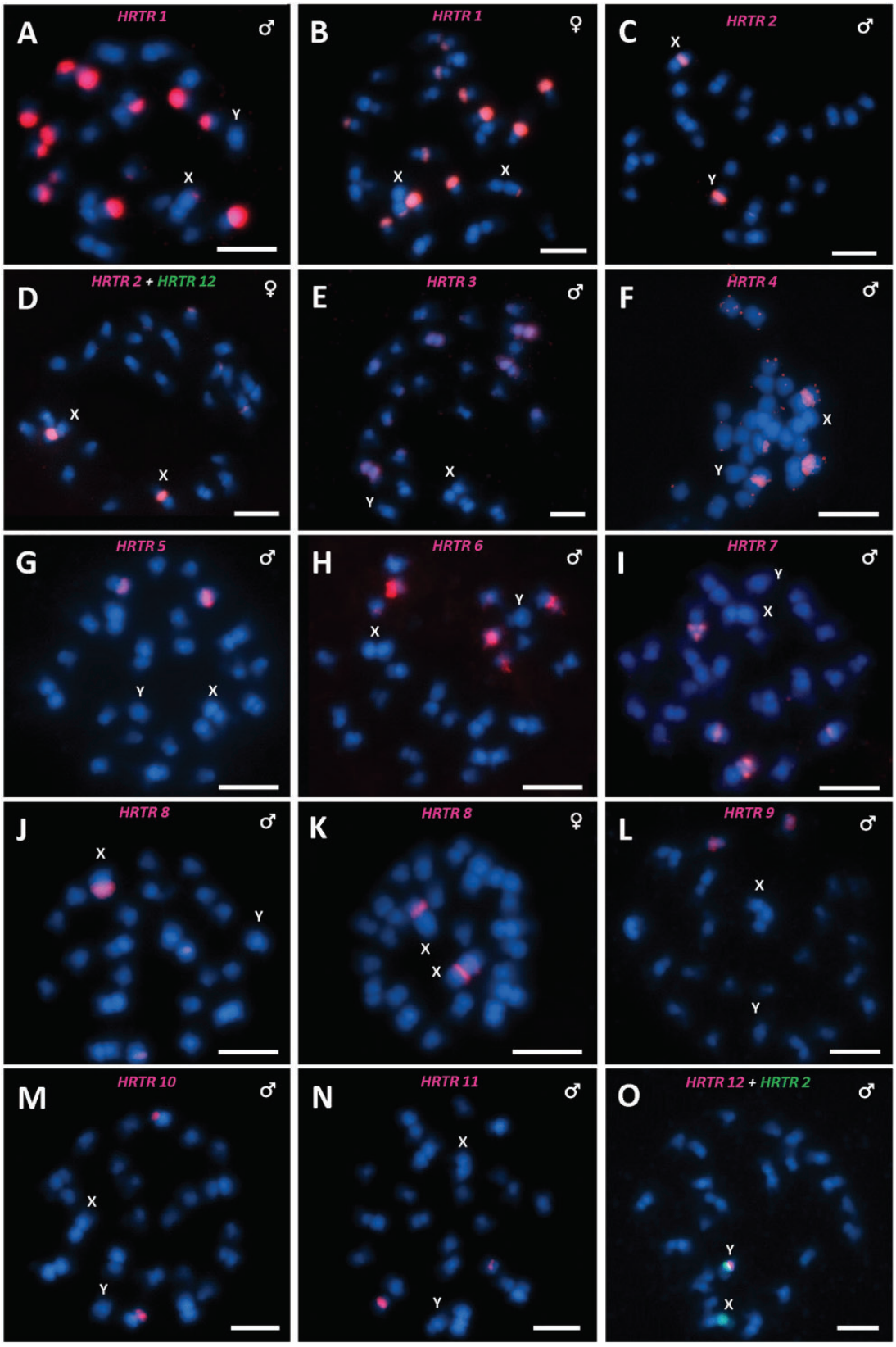
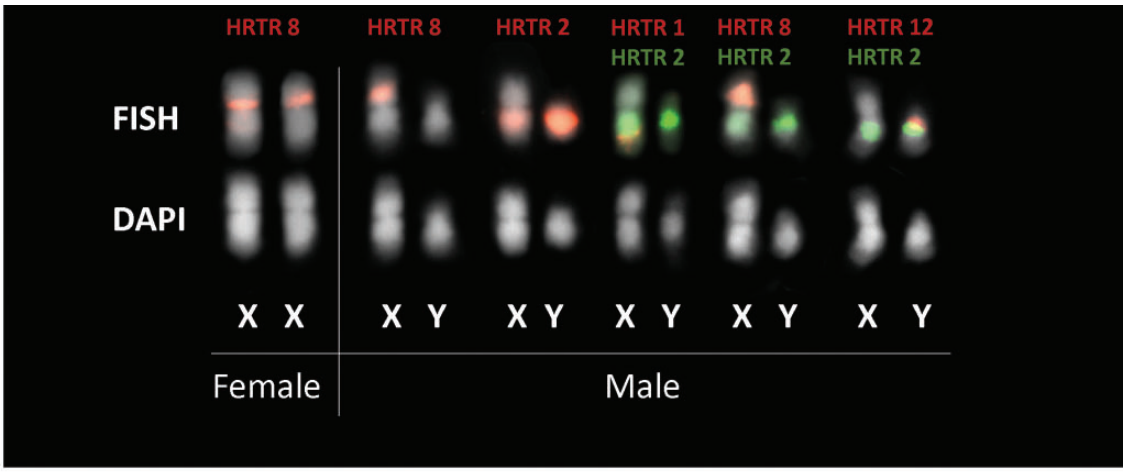


Fig. 6.—Localization of main satellite families on metaphase chromosomes of *Hippophae rhamnoides* using fluorescence *in situ* hybridization. The name of satellite family and sex of individual are indicated inside each figure. Blue are DAPI stained chromosomes, red and green signals show chromosomal localization of satellite families. Bar indicates 5 μm.



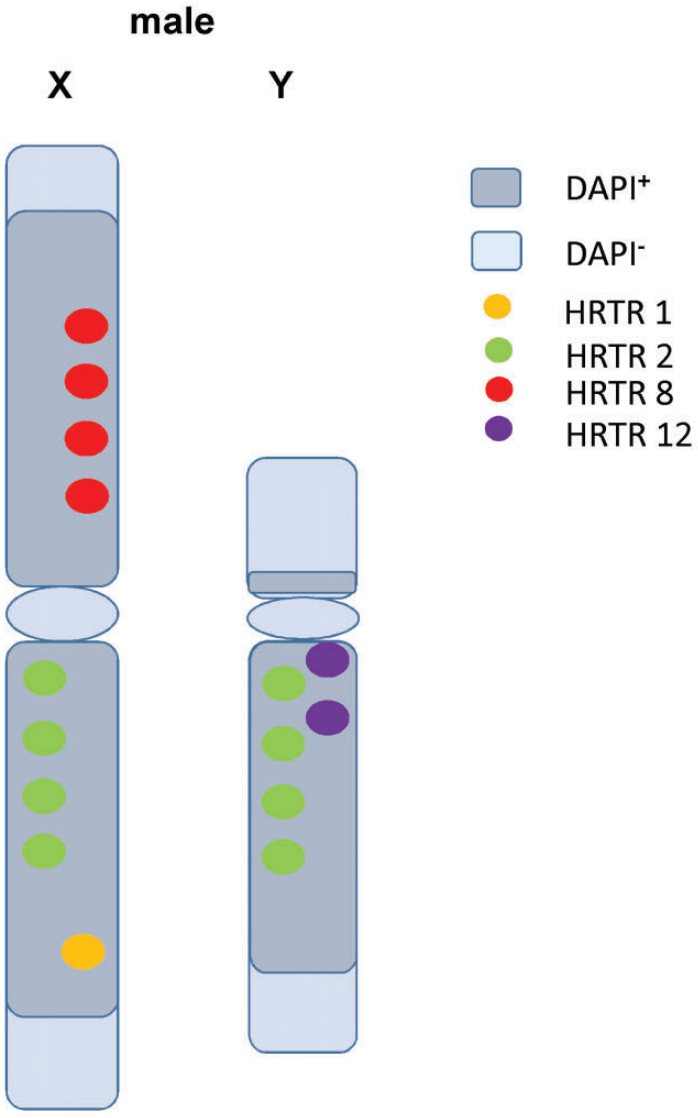


Fig. 7.—FISH and scheme of four satellites on sex chromosomes. The HRTR1, Y-specific HRTR12, X-accumulated HRTR8, sex chromosome-accumulated HRTR2.

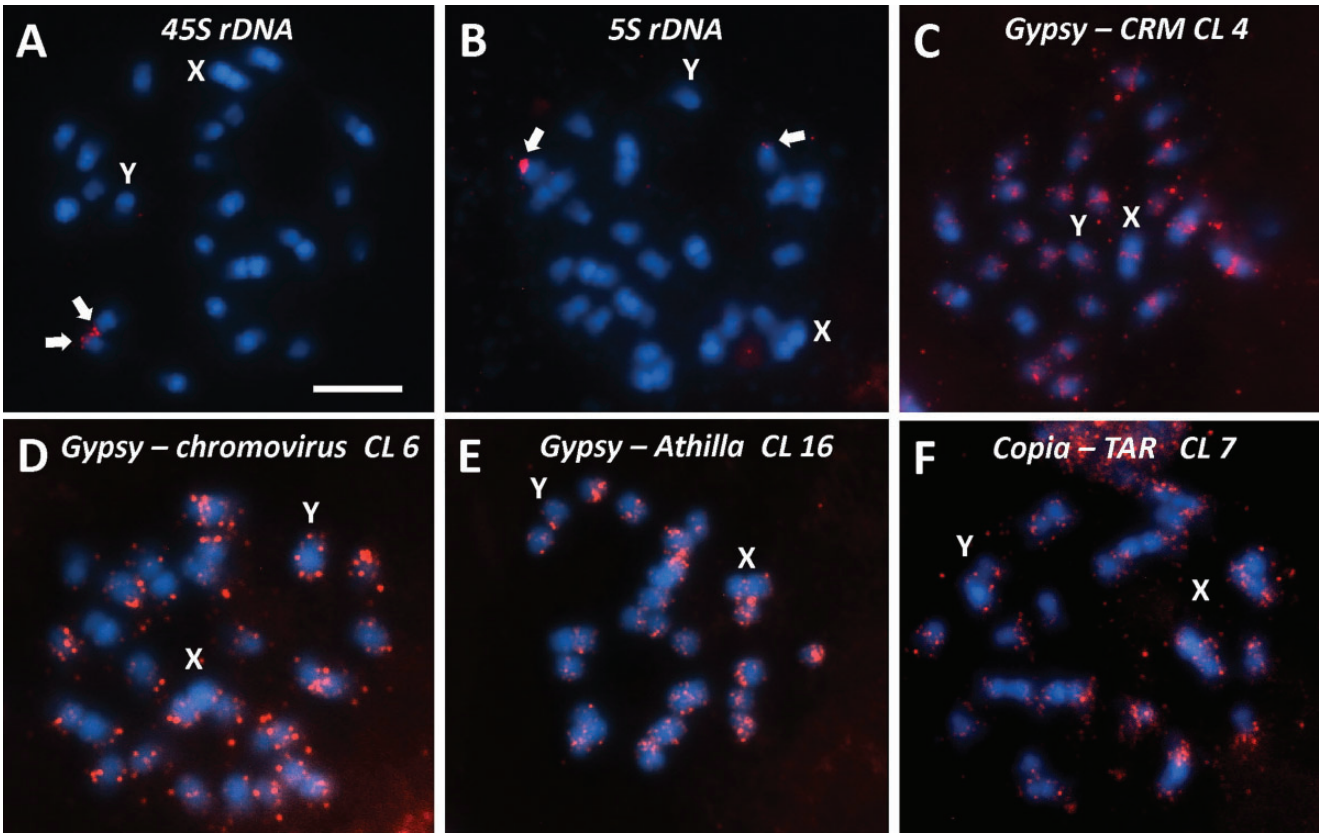


Fig. 8.—Localization of transposable elements and rDNA on metaphase chromosomes of *Hippophae rhamnoides* using fluorescence *in situ* hybridization. The name of transposable element family (together with the number of corresponding cluster) or type of rDNA cluster is inside each figure. Blue are DAPI stained chromosomes, red signal shows chromosomal localization of selected transposable elements and 45S and 5S rDNA. Bar indicates 5 μm.

three of four studied groups of TEs are present mainly in subtelomeres of all chromosomes (fig. 8*D–F*) and only the CRM retroelements (CL4) that was localized in the centromeric region of all chromosomes (fig. 8*C*).

Discussion

We present the first comprehensive analysis of seabuckthorn (*H. rhamnoides*) genome. We found that about one quarter of the genome is composed of TEs and another quarter of satellite DNA which is comparable to other plant genomes. Nevertheless, the seabuckthorn genome contains an unusually large number of different satellites (table 2, 12 main tandem repeats) compared with most other plant genomes (Mehrotra and Goyal 2014). Moreover, some satellites evolve rapidly into new variants. In particular, HRTR2 and HRTR3 satellite superclusters are comprised of a number of smaller clusters where each cluster represents an individual satellite (supplementary fig. S3, Supplementary Material online). Thus, the number of different satellites may be even higher if more strict criteria were used for tandem repeat classification. Transposable elements are represented by all main families of both Ty3/*Gypsy* and Ty1/*Copia* retrotransposons (fig. 2) with chromoviruses (CRM and Galadriel) and TAR families dominating (table 1). Most transposable element families are represented by only one or two clusters indicating their long term presence without changes in sequence or structure. Only Athila, Angela, Tork and Ale/Retrofit retrotransposons are found in multiple clusters (data not shown) suggesting higher divergence. Well preserved long ORFs in some TEs indicate the recent amplification/younger age and low level of degeneration of these elements. All in all, high variability of some satellites and TE families indicate high tempo of their diversification in the seabuckthorn genome, while other repeats remain relatively conserved. Nevertheless, this conclusion should be verified by comparative analysis of at least two closely related species. Recent analysis by Macas et al. (2015) showed that it is not transposable elements but satellites that are the most variable repeats among closely related species of Fabae genus.

Comparison of numbers of male and female reads constituting satellite superclusters, enabled us to predict satellites localized on the Y chromosome, X chromosome, on both sex chromosomes or on autosomes as each specific ratio of

abundance of male and female reads in a cluster corresponded to specific chromosomal distribution. Our FISH results showed that this prediction works well in most cases as verified by satellites accumulated on the X chromosome (HRTR8) and both X and Y chromosomes, and specific for the Y chromosome (HRTR12) and for autosomes (HRTR1, 3, 4, 5, 6, and 10). It is a question whether or not the higher number of different satellites in the seabuckthorn genome than in the majority of plant genomes (Mehrotra and Goyal 2014) somehow correlates with the presence of sex chromosomes representing a specific genomic context, each shaped by different evolutionary forces.

The localization of satellites is remarkable and shows that satellites are gathered not only on the nonrecombining region of the Y chromosome but some are specific for the X chromosome or for both sex chromosomes. They are gathered in heterochromatic parts of sex chromosomes what can reflect possible role of satellites in heterochromatinization. The list of chromosomal localization of satellites and TEs in dioecious plants was recently presented by Li et al. (2016a). Although Y chromosome divergence and specific repeat composition is a generally accepted feature, an accumulation of X-specific repeats during plant sex chromosome evolution has been suggested only by limited number of studies (Hobza et al. 2004). As satellites localized on either X or Y chromosomes are mutually different, we prefer the explanation that these satellites originated and expanded on the sex chromosomes long after the X–Y divergence. Therefore, it would be interesting to compare X and Y-linked variants of HRTR2 satellite and, if present, to assess the extent of X- and Y-linked satellite divergence.

The localization of transposable elements mainly in subtelomeres is a feature characteristic of the seabuckthorn genome. However, transposable elements are accumulated in subtelomeres in other plant species too (Zhang and Wessler 2004), and, among dioecious plants, subtelomeric localization was shown in *Retand* retrotransposon in *Silene latifolia* (Kejnovsky et al. 2006). Retrotransposons are found in or around centromeres as well (Miller et al. 1998; Neumann et al. 2011).

Our results clearly confirm the existence of the XY system in seabuckthorn found by Shchapov (1979) and they show that the Y chromosome is small and the X chromosome large. We mention in passing the work of Truta et al. (2011) who initially found a large Y chromosomes and small X chromosome in three Romanian seabuckthorn genotypes that later investigation of Romanian genotypes failed to confirm (Dr. Elena Truta, Institute of Biological Research Iasi, Romania, personal communication, June 15, 2016). Another cytogenetic study on seabuckthorn using C-banding that unfortunately showed only female karyotype without marking sex chromosomes (Rousi and Arohonka 1980).

Estimation of the age of sex chromosomes is not yet possible in this species because no X- and Y-linked genes are known. It remains a question whether the large size difference between X and Y chromosomes, the small size of the Y chromosome and accumulation of different satellites on both sex chromosomes indicates greater age of these sex chromosomes or not. It is remarkable that another genus of the Elaeagnaceae family—Shepherdia (Elaeagnaceae contains three genera—Elaeagnus, Hippophae, and Shepherdia) contains only three species that are all dioecious (Veldkamp 1986). Moreover, the Elaeagnaceae family belongs to the order of Rosales containing other plants with heteromorphic sex chromosomes like Humulus and Cannabis. Although karyotypes were described in Elaeagnus ($2n = 28$ in *E. angustifolia*) and Shepherdia ($2n = 26$ in *S. argentea* and $2n = 22$ in *S. canadensis*), the sex chromosomes were not revealed (Rousi and Arohonka 1980). Therefore, it is not possible to draw conclusions about the formation or age of sex chromosomes during phylogeny.

The small Y chromosome containing several satellite DNA and a large X chromosome revealed in seabuckthorn resemble the mammalian sex chromosomal system. To the best of our knowledge, such a system is very rare among plants. Sex chromosomes in plants are mostly evolutionarily young—e.g., *Silene latifolia* (6 Ma, Kubat et al. 2014), *Rumex acetosa* (12–13 Ma, Navajas-Perez et al. 2005), or *Coccinia grandis* (3 Ma, Sousa et al. 2013)—and only sex chromosomes of *Marchantia polymorpha* are thought to be older (Yamato et al. 2007). A small Y chromosome and the large X chromosome were revealed in *Humulus lupulus* (Shephard et al. 2000; Karlov et al. 2003) and also in gymnosperm species *Cycas revoluta* (Segawa et al. 1971). The small size of the seabuckthorn Y chromosome may be caused by the loss of DNA which indicates that the Y chromosome could be in a shrinkage phase of evolution [reviewed in Hobza et al. (2015)] and thus could represent a rare example of an evolutionarily old plant sex chromosome. This assumption is supported by the FISH results which indicate that the large part of the Y chromosome arm that is homologous to the arm of the X chromosome, carrying HRTR8, was lost (fig. 7).

In this study, we developed and used a new bioinformatics approach for analysis of satellite DNA allowing prediction of satellite monomers, their grouping into clusters corresponding to main satellite families in the genome and visualization of their male/female homogeneity. This enabled prediction of satellite localization with respect to the sex determination system in species studied.

Supplementary Material

Supplementary data are available at *Genome Biology and Evolution* online.

Acknowledgments

This work was supported by the Czech Science Foundation (grant P501/12/G090). Access to computing and storage

facilities owned by parties and projects contributing to the National Grid Infrastructure MetaCentrum, provided under the program "Projects of Large Infrastructure for Research, Development, and Innovations" (LM2010005), is greatly appreciated. The work of JP was supported by the Research and Application of Advanced Methods in ICT project (FIT-S-14-2299; http://www.fit.vutbr.cz/). The work of T.M. was supported from IT4Innovations excellence in science project (LQ1602). O.R., O.A., M.D., and G.K. were supported by Government Program supporting of the Leading Scientific Schools of Russian Federation (grant SS-8315.2016.11).

Literature Cited

Alexandrov OS, Divashuk MG, Yakovin NA, Karlov GI. 2012. Sex chromosome differentiation in *Humulus japonicus* Siebold & Zuccarini, 1846 (Cannabaceae) revealed by fluorescence in situ hybridization of subtelomeric repeat. Comp Cytogenet. 47:239–247.

Alexandrov OS, Karlov GI. 2016. Molecular cytogenetic analysis and genomic organization of major DNA repeats in castor bean (*Ricinus communis* L.). Mol Genet Genomics 291:775–787.

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. J Mol Biol. 215:403–410.

Benson DA, Karsch-Mizrachi I, Lipman DJ, Ostell J, Sayers EW. 2009. GenBank. Nucleic Acids Res. 38:D46–D51.

Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30:2114–2120.

Camacho C, et al. 2009. BLAST+: architecture and applications. BMC Bioinformatics 10:421.

Campell BR, Song Y, Posch TE, Cullis CA, Town CD. 1992. Sequence and organization of 5S ribosomal RNA-encoding genes of Arabidopsis thaliana. Gene 112:225–228.

Cermak T, et al. 2008. Survey of repetitive sequences in *Silene latifolia* with respect to their distribution on sex chromosomes. Chromosome Res. 16:961–976.

Charlesworth B. 1991. The evolution of sex chromosomes. Science 251:1030–1033.

Charlesworth D. 2016. Plant sex chromosomes. Annu Rev Plant Biol. 67:397–420.

Chen L, Yu Z, Jin H. 2010. Comparison of ribosomal DNA ITS regions among *Hippophae rhamnoides* ssp. sinensis from different geographical area in China. Plant Mol Biol Rep. 28:635–645.

Crooks G, Hon G, Chandonia J, Brenner S. 2004. WebLogo: a sequence logo generator. Genome Res. 14:1188–1190.

Divashuk MG, Alexandrov OS, Kroupin PY, Karlov GI. 2011. Molecular cytogenetic mapping of *Humulus lupulus* sex chromosomes. Cytogenet Genome Res. 134:213–219.

Divashuk MG, Alexandrov OS, Razumova OV, Kirov IV, Karlov GI. 2014. Molecular cytogenetic characterization of the dioecious *Cannabis sativa* with an XY chromosome sex determination system. PLoS One 9:e85118.

Doležel J, Bartoš J, Voglmayr H, Greilhuber J. 2003. Nuclear DNA content and genome size of trout and human. Cytom Part A 51:127–128.

Doyle JJ, Doyle JL. 1990. Isolation of plant DNA from fresh tissue. Focus 12:13–15.

Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 32:1792–1797.

Fruchterman TMJ, Reingold EM. 1991. Graph drawing by force-directed placement. Softw Pract Exp. 21:1129–1164.

Ghangal R, Chaudhary S, Jain M, Purty RS, Sharma PC. 2013. Optimization of de novo short read assembly of seabuckthorn (*Hippophae rhamnoides* L.) transcriptome. PLoS One 8:e72516.

Gerlach WL, Bedbrook JR. 1979. Cloning and characterization of ribosomal RNA genes from wheat and barley. Nucleic Acids Res. 7:1869–1885.

Hobza R, Lengerova M, Cernohorska H, Rubes J, Vyskot B. 2004. FAST-FISH with laser beam microdissected DOP-PCR probe distinguishes the sex chromosomes of *Silene latifolia*. Chromosome Res. 12:245–250.

Hobza R, Vyskot B. 2015. The genomics of plant sex chromosomes. Plant Sci. 236:126–135.

Hobza R, et al. 2015. Impact of repetitive DNA on sex chromosome evolution in plants. Chromosome Res. 23:561–570.

Hough J, Hollister JD, Wang W, Barrett SCH, Wright SI. 2014. Genetic degeneration of old and young Y chromosomes in the flowering plant *Rumex hastatulus*. Proc Natl Acad Sci U S A. 111:7713–7718.

Kearse M, et al. 2012. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics 28:1647–1649.

Karlov GI, Danilova TV, Horlemann C, Weber G. 2003. Molecular cytogenetic in hop (*Humulus lupulus* L.) and identification of sex chromosomes by DAPI banding. Euphytica 132:185–190.

Kejnovsky E, et al. 2006. Retand: A novel family of gypsy-like retrotransposon harboring an amplified tandem repeat. Mol Genet Genomics 276:254–263.

Kejnovsky E, Hobza R, Kubat Z, Cermak T, Vyskot B. 2009. The role of repetitive DNA in structure and evolution of sex chromosomes in plants. Heredity 102:533–541.

Kejnovsky E, Vyskot B. 2010. *Silene latifolia*: the classical model to study heteromorphic sex chromosomes. Cytogenet Genome Res. 129:250–262.

Kirov I, Divashuk M, Van Laere K, Soloviev A, Khrustaleva L. 2014. An easy "SteamDrop" method for high quality plant chromosome preparation. Mol Cytogenet. 7:21.

Kubat Z, et al. 2014. Possible mechanisms responsible for absence of retrotransposon family on a plant Y chromosome. New Phytol. 202:662–678.

Langmead B, Salzberg S. 2012. Fast gapped-read alignment with Bowtie 2. Nat Methods 9:357–359.

Li H. 2013. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. arXiv 1303.3997.

Li SF, Zhang GJ, Yuan JH, Deng CL, Gao WJ. 2016a. Repetitive sequences and epigenetic modification: inseparable partners play important roles in the evolution of plant sex chromosomes. Planta 243:1083–1095.

Li SF, et al. 2016b. DPTEdb, an integrative database of transposable elements in dioecious plants. Database 2016:1–10.

Lim KY, et al. 2007. Sequence of events leading to near-complete genome turnover in allopolyploid Nicotiana within five million years. New Phytol. 175:756–763.

Llorens C, et al. 2011. The Gypsy database (GyDB) of mobile genetic elements: release 2.0. Nucleic Acids Res. 39:D70–D74.

Macas J, Meszaros T, Nouzova M. 2002. PlantSat: a specialized database for plant satellite repeats. Bioinformatics 18:28–35.

Macas 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.

Marchler-Bauer A, et al. 2015. CDD: NCBI's conserved domain database. Nucleic Acids Res. 43:222–226.

Mariotti B, Manzano S, Kejnovsky E, Vyskot B, Jamilena M. 2009. Accumulation of Y-specific satellite DNAs during the evolution of *Rumex acetosa* sex chromosomes. Mol Genet Genomics 281:249–259.

Mehrotra S, Goyal V. 2014. Repetitive sequences in plant nuclear DNA: types, distribution, evolution and function. Genomics Proteomics Bioinformatics 12:164–171.

Miller JT, Dong F, Jackson SA, Song J, Jiang J. 1998. Retrotransposon-related DNA sequences in the centromeres of grass chromosomes. Genetics 150:1615–1623.

Ming R, Bendahmane A, Renner SS. 2011. Sex chromosomes in land plants. Annu Rev Plant Biol. 62:485–514.
Navajas-Perez R, et al. 2005. The evolution of reproductive systems and sex-determining mechanisms within rumex (polygonaceae) inferred from nuclear and chloroplastidial sequence data. Mol Biol Evol. 22:1929–1939.
Neumann P, et al. 2011. Plant centromeric retrotransposons: a structural and cytogenetic perspective. Mobile DNA 2:4.
Novak P, Neumann P, Macas J. 2010. Graph-based clustering and characterization of repetitive sequences in next-generation sequencing data. BMC Bioinformatics 11:378–389.
Novak 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.
Rousi A, Arohonka T. 1980. C-band and ploidy level of *Hippophae rhamnoides*. Hereditas 92:327–330.
Saitou N, Nei M. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol. 4:406–425.
Segawa M, Kishi S, Tatuno S. 1971. Sex chromosomes of *Cycas revoluta*. Jpn J Genet. 46:33–39.
Sharma A, Zinta G, Rana S, Shirko P. 2010. Molecular identification of sex in *Hippophae rhamnoides* L. using isozyme and RAPD markers. For Stud China 12:62–66.
Shchapov NS. 1979. On the karyology of *Hippophaë rhamnoides* L. Tsitol Genet. 13:45–47.
Shephard HL, Parker JS, Darby P, Ainsworth CC. 2000. Sexual development and sex chromosomes in hop. New Phytol. 148:397–411.
Sousa A, Fuchs J, Renner SS. 2013. Molecular cytogenetics (FISH, GISH) of *Coccinia grandis*: a ca. 3 myr-old species of Cucurbitaceae with the largest Y/autosome divergence in flowering plants. Cytogenet Genome Res. 139:107–118.
Souza A, Bellot S, Fuchs J, Houben A, Renner SS. 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.
Steflova P, et al. 2013. Contrasting patterns of transposable element and satellite distribution on sex chromosomes (XY_1Y_2) in the dioecious plant *Rumex acetosa*. Genome Biol Evol. 5:769–782.
Truta E, et al. 2011. Morphometric pattern of somatic chromosomes in three Romanian seabuckthorn genotypes. Caryologia 64:189–196.
Untergasser A, et al. 2012. Primer3–new capabilities and interfaces. Nucleic Acids Res. 40:1–12.
Veldkamp JF. 1986. Elaeagnaceae. In: Van Steenis CGGJ, de Wilde WJJO, editors . Flora Malesiana 10 (2), Martinus Nijhoff. Boston, London: The Hague, p. 151–156.
Wicker T, Matthews DE, Keller B. 2002. TREP: A database for Triticeae repetitive elements. Trends Plant Sci. 7:561–562.
Yamato KT, et al. 2007. Gene organization of the liverwort Y chromosome reveals distinct sex chromosome evolution in a haploid system. Proc Natl Acad Sci U S A. 104:6472–6477.
Zhang X, Wessler SR. 2004. Genome-wide comparative analysis of the transposable elements in the related species *Arabidopsis thaliana* and *Brassica oleracea*. Proc Natl Acad Sci U S A. 101:5589–5594.
Zhou X, et al. 2010. Genome size of the diploid hybrid species *Hippophae goniocarpa* and its parental species, *H. rhamnoides* ssp. sinensis and *H. neurocarpa* ssp. neurocarpa (Elaeagnaceae). Acta Biol Cracoviensia Ser Bot. 52:12–16.

Associate editor: Ellen Pritham