Planta (2017) 246:749–761
DOI 10.1007/s00425-017-2716-6

ORIGINAL ARTICLE

Organization and evolution of four differentially amplified tandem repeats in the *Cucumis hystrix* genome

Shuqiong Yang[1] · Xiaodong Qin[1] · Chunyan Cheng[1] · Ziang Li[1] · Qunfeng Lou[1] · Ji Li[1] · Jinfeng Chen[1]

Received: 26 January 2017 / Accepted: 29 May 2017 / Published online: 1 July 2017

Abstract

Main conclusion **Three subtelomeric satellites and one interstitial 5S rDNA were characterized in *Cucumis hystrix*, and the pericentromeric signals of two *C. hystrix* subtelomeric satellites along *C. sativus* chromosomes supported the hypothesis of chromosome fusion in *Cucumis*.**

Tandem repeats are chromosome structural fractions consisting of highly repetitive sequences organized in large tandem arrays in most eukaryotes. Differentiation of tandem repeats directly affects the chromosome structure, which contributes to species formation and evolution. *Cucumis hystrix* ($2n = 2x = 24$) is the only wild *Cucumis* species grouped into the same subgenus with *C. sativus* ($2n = 2x = 14$), hence its phylogenetic position confers a vital role for *C. hystrix* to understand the chromosome evolution in *Cucumis*. However, our knowledge of *C. hystrix* tandem repeats is insufficient for a detailed understanding of the chromosome evolution in *Cucumis*. Based on de novo tandem repeat characterization using bioinformatics and in situ hybridization (ISH), we identified and characterized four differentially amplified tandem repeats, *Cucumis hystrix* satellite 1–3 (CuhySat1–CuhySat3) located at the subtelomeric regions of all chromosomes, and *Cucumis hystrix* 5S (Cuhy5S) located at the interstitial regions of one single chromosome pair. Comparative ISH mapping using CuhySat1–3 and Cuhy5S revealed high homology of tandem repeats between *C. hystrix* and *C. sativus*. Intriguingly, we found signal distribution variations of CuhySat2 and CuhySat3 on *C. sativus* chromosomes. In comparison to their subtelomeric signal distribution on *C. hystrix* chromosomes, CuhySat3 showed a pericentromeric signal distribution and CuhySat2 showed both subtelomeric and pericentromeric signal distributions on *C. sativus* chromosomes. This detailed characterization of four *C. hystrix* tandem repeats significantly widens our knowledge of the *C. hystrix* chromosome structure, and the observed signal distribution variations will be helpful for understanding the chromosome evolution of *Cucumis*.

Keywords Tandem repeats · Chromosome structure · *Cucumis hystrix* · Chromosome evolution in *Cucumis*

Abbreviations

ISH	In situ hybridization
FISH	Fluorescence in situ hybridization
GISH	Genome in situ hybridization
CuhySat	*Cucumis hystrix* satellite
Cuhy5S	*Cucumis hystrix* 5S
rDNA	Ribosomal DNA

✉ Jinfeng Chen
jfchen@njau.edu.cn

Shuqiong Yang
2014204012@njau.edu.cn

Xiaodong Qin
2014204011@njau.edu.cn

Chunyan Cheng
2010204010@njau.edu.cn

Ziang Li
2014104071@njau.edu.cn

Qunfeng Lou
qflou@njau.edu.cn

Ji Li
liji1981@njau.edu.cn

[1] State Key Lab of Crop Genetics and Germplasm Enhancement, College of Horticulture, Nanjing Agricultural University, Nanjing 210095, China

Introduction

Repetitive DNAs make up a significant portion of every eukaryotic genome, and among them, a large fraction comprises sequences that are repeated in tandem arrays known as tandem repeats. Traditionally, tandem repeats are categorized according to their repeat unit length and array size into microsatellites with 2–5 bp monomers and an array size order of 10–100 units, minisatellites with 6–100 bp (usually ~15 bp) monomers and a 0.5–30 kb array size, and satellite DNAs with a variable AT-rich repeat monomer with a length ranging from 150 to 400 bp, often forming an array size up to 100 Mb (Sharma and Raina 2005). Moreover, tandem repeats also include telomeric repeats and ribosomal DNA (rDNA) with defined functions that differ from typical repetitive sequences.

Microsatellites are often referred to as simple sequence repeats (SSRs), which are used in genetic linkage analysis and marker-assisted selection to map a gene or a mutation responsible for a given trait or disease by plant geneticists. Minisatellites are prominent in the centromeres and telomeres of chromosomes, the latter protect the chromosomes from damage. Satellite DNAs are portions of the genome that can be isolated using the methods including density gradient ultracentrifugation, isolation of prominent bands from restriction enzyme-digested genomic DNAs fractionated by gel electrophoresis, and genomic self-priming PCR (GSP-PCR), etc. (Beridze 1986; Macas et al. 2000). Satellite DNAs often form families that differ in sequence, repeat monomer length, abundance, and chromosome location. These families experience high rates of genomic change and replacement, and they are usually species specific or shared by a group of related species (Garrido-Ramos 2015). Telomeres are complex nucleoprotein structures that form the physical ends of linear eukaryotic chromosomes and are considered as a part of the cellular clock machinery that determines the life span of the cell (Dvorácková et al. 2015). In most plant species, telomeric DNA is composed of short tandem repeats of TTTAGGG, but some sequence variants have also been found in some genera (Fajkus et al. 2016; Peška et al. 2015). Ribosomal DNAs contain two different families, 45S and 5S rDNA, with multiple copies of the coding sequences and intergenic spacers as tandem repeats in the genome (Zhang et al. 2016). Chromosomal fluorescence ISH (FISH) mapping of rDNAs have provided a useful approach for identifying chromosomes and studying chromosome structure and evolution.

Specific tandem repeats have been used to investigate their cytogenetic positions among related species to reveal chromosome evolution (Grzywacz et al. 2014; Lim et al. 2006; Rosato et al. 2012). Cytogenetic evolution studies using specific tandem repeats such as rDNAs have been applied to analyze the relationships among the *Cucumis* species. Koo et al. (2010) found an unusual location of 45S rDNA interposed with a centromeric domain when they comparatively mapped 45S rDNA and CsCent1 in *C. sativus* and *C. melo*. They thought that *C. sativus* chromosomes 1 and 2 might have evolved from fusions of the ancestral karyotype $2n = 2x = 24$. A molecular cytogenetic analysis of five wild *Cucumis* species from Southern Africa (*C. africanus*, *C. anguria*, *C. myriocarpus*, *C. zeyheri*, and *C. heptadactylus*) using FISH with 45S and 5S rDNA suggested that at least 2 steps of chromosomal rearrangements might have occurred during the evolution of the tetraploid *C. heptadactylus* (Yagi et al. 2015). Zhang et al. (2016) found that 45S rDNA showed a trend toward an increasing site number but a relatively conserved location among 20 *Cucumis* accessions, whereas the 5S rDNA maintained a conserved site number but showed a polyploidization-related tendency towards the terminal location from an interstitial location. A comparison of the distribution of Type I/II, Type III, Type IV and 45S rDNA in three variants of *C. sativus* revealed that the relationship between *C. sativus* var. *xishuangbannesis* and *C. sativus* var. *hardwickii* was no closer than that between the cucumber cultivar and *C. sativus* var. *xishuangbannesis* (Zhao et al. 2011). Comparative cytogenetic mapping of six specific tandem repeats (45S rDNA, Type I/II, Type III, Type IV, CentM and telomeric repeat) among five *Cucumis* species (*C. sativus*, *C. hystrix*, *C. melo*, *C. anguria* and *C. metuliferus*) revealed a significant differentiation of chromosome structure during the formation of these species with the nucleotide sequence and copy number as the main changing parameters while the evolution of specific repeats (Zhang et al. 2015).

Among the 40 *Cucumis* species whose chromosome number have been counted, *C. sativus* is the only species with a basic chromosome number of $x = 7$, while all other species have a chromosome basic number of $x = 12$, providing an interesting model for studying chromosome number evolution in the genus *Cucumis* (Zhang et al. 2015; Yang et al. 2014). Of all wild *Cucumis* species, *C. hystrix* is the only one that is grouped into the same subgenus together with *C. sativus*, and the only species that is cross-compatible with *C. sativus* (Chen et al. 1997; Sebastian et al. 2010). Therefore, the chromosome structure analysis of *C. hystrix* is crucial for understanding chromosome evolution in the genus *Cucumis*. However, to date, only a few studies of *C. hystrix* repetitive DNAs have been reported. The locations of 45S and 5S rDNA loci were confirmed via FISH mapping, with three pairs of 45S rDNA signals located at the subtelomeric regions of *C. hystrix* chromosomes and 5S rDNA to produce one pair of interstitial signals (Chen et al. 1999; Zhang et al. 2016).

Self-genomic ISH (GISH) of *C. hystrix* produced very strong signals at two ends or one end of the chromosomes, which indicated a preferential location of repeats at the distal regions of the chromosomes (Zhang et al. 2015). FISH mapping of *C. hystrix* 24 Ty1-*copia* retrotransposons showed that they were widely dispersed over all *C. hystrix* chromosomes, clustering at terminal heterochromatin regions (Jiang et al. 2010). For studying chromosome structure and evolution, our knowledge of the repetitive sequences of *C. hystrix* remains insufficient, especially in terms of tandem repeats.

Due to the introduction of next-generation sequencing technologies and corresponding bioinformatics approaches, much progress has been made toward conducting a detailed characterization of repetitive DNAs in plant genomes (Weiss-Schneeweiss et al. 2015). One of these novel approaches, graph-based read clustering of low coverage genome sequencing reads (Novák et al. 2010), has been proved to be particularly efficient for repetitive DNA identification and characterization in many plant genomes (He et al. 2015; Heitkam et al. 2015; Torres et al. 2011; Cai et al. 2014). In this study, 4.53G of paired-end sequencing reads was used to analyze the repetitive elements of *C. hystrix*. We found four types of tandem repeats in the *C. hystrix* genome, and their nucleotide sequences and chromosomal locations were identified and characterized in detail. Comparative ISH technology was applied to investigate the homology and evolution of repeats between *C. hystrix* and *C. sativus*. Our work delivers chromosomal insights into *C. hystrix* chromosomes and provides some useful information for further studies of chromosome evolution in the genus *Cucumis*.

Materials and methods

Plant materials

Cucumis hystrix Chakr. ($2n = 2x = 24$) was used for Illumina HiSeq genome sequencing and cytogenetic studies. *Cucumis sativus* var. *sativus* ($2n = 2x = 14$) was used for comparative ISH mapping of the tandem repeats. Samples were grown in the greenhouse of Nanjing Agricultural University under 14 h of light and 10 h of the dark.

Genomic DNA isolation and Illumina sequencing

Total genomic DNA was extracted from young leaves of seedlings using the cetyltrimethyl-ammonium bromide (CTAB) based on method described by Murray and Thompson (1980). DNA was treated with DNase-free-RNase A for 30 min at 37 °C to remove RNA and purified by phenol/chloroform precipitation. Illumina sequencing of randomly sheared DNA was performed by HiSeq2000 platform (Berry Genomics, Beijing, China). One ninety bp paired-end reads library was obtained from the results.

Bioinformatics identification of tandemly repeated sequences in the *C. hystrix* genome

To gain an overview of the repetitive elements of the *C. hystrix* genome and to facilitate the identification of tandem repeats, we performed a graph-based clustering approach implemented in the RepeatExplorer software (Novák et al. 2013). After removing the linker/primer contaminations and artificially duplicated reads, a set of 4.53 G whole genome Illumina paired-end reads (average length of reads was 90 bp), representing an approximately 10× genome equivalent of *C. hystrix*, was used for clustering analysis. This analysis was conducted using a read similarity cutoff of 90% over at least 55% of the shorter sequence length. Reads within individual clusters were assembled into contigs, which were used for sequence-similarity searches against the Repbase library using RepeatMasker (Smit et al. 2014), and the conserved protein domains were detected using RPS-Blast (Altschul et al. 1997) to identify the repeat type and family. Clusters corresponding to putative mitochondrial and plastid contaminations were eliminated after identification by searching GenBank (Clark et al. 2016). The fraction of reads in clusters versus the total number of reads used for the analysis provides an estimate of the size of the repetitive portion of the genome.

Clusters with a tandem repeat typical star-like and circular graphical representation have been selected for further analysis. RepeatExplorer derived contigs have been aligned and assembled, and putative monomers have been detected using Tandem Repeats Finder (Benson 1999) and Phobos, a tandem repeat detection tool integrated within Geneious 5.4 (https://www.geneious.com, Kearse et al. 2012).

Probe preparation

DNA probes used for *C. hystrix* tandem repeats were prepared by polymerase chain reaction (PCR) amplification using specific primers (Table 1) designed according to the monomer sequences (Fig. 2). PCR reactions consisting of 50 ng genomic DNA template were performed in a 25-μl volume that contained 2.5 μl 10 x PCR buffer (Mg^{2+} plus), 2 μl 2.5 mM dNTP mixture, 0.625 U r-Taq (Takara, Japan), 1 μl each of two primers and 17.375 μl of ddH_2O. Touchdown PCR program was employed for all primer sets and consisted of a 3-min initial denaturation at 95 °C; six cycles of 45 s at 94 °C, 5 min at 68 °C, and 1 min at 72 °C, with a reduction of the annealing temperature by 2 °C per cycle; eight cycles of 45 s at 94 °C, 2 min at

Table 1 Primer sequences for amplification of tandem repeat-specific probes

Primer	Repeat family	Sequence (5′–3′)
CuhySatl_F	CuhySatl	TCAAAATGAGGCCACCCAGT
CuhySatl_R	CuhySatl	GCCCCCAAAAACGTTGGAAA
CuhySat2_F	CuhySat2	ATTGCACAAAATGAGGCCACC
CuhySat2_R	CuhySat2	CGTTGGTGGTGTACTACGGT
CuhySat3_F	CuhySat3	CACCCAACATAACAGCCGA
CuhySat3_R	CuhySat3	AAAGAGGAGAGGATCCCCAC
Cuhy5S_F	Cuhy5S	CGGCTTTACTGGTCGCA
Cuhy5S_R	Cuhy5S	ACTGTTTTCTTCAATCCCTCC

F forward primer, *R* reverse primer

58 °C, and 1 min at 72 °C, with a reduction of the annealing temperature by 1 °C per cycle; and a final 25 cycles of 45 s at 94 °C, 2 min at 50 °C and 1 min at 72 °C (Weng et al. 2005).

The 9-kb pTa71 plasmid from wheat (Gerlach and Bedbrook 1979), provided by Dr. Jiang Jiming (University of Wisconsin-Madison), was used as the 45S rDNA probe. Type III repeat with a size of 177 bp (GenBank Accession No. 18287) was used to identify the centromeres of *C. sativus*. The *Arabidopsis*-type telomeric DNA was generated by PCR method in the absence of template using primers $(TTTAGGG)_4$ and $(CCCTAAA)_4$ according to Ijdo et al. (1991). All DNA probes were labeled with either biotin-dUTP or digoxigenin-dUTP (Roche, http://www.roche-applied-science.com) using a standard nick translation reaction.

Preparation of chromosome spreads

Mitotic metaphase and meiotic pachytene chromosomes were prepared following the published protocols (Lou et al. 2013) with some modifications. Seeds of all species were geminated on moistened filter paper at 28 °C. Approximately 1-cm-long lateral root tips were collected after induction by cutting away the main root tips, pretreated for 2 h in 0.002 M 8-hydroxyquinoline, and fixed in Carnoy's fixative solution (ethanol:Oglacial acetic acid, 3:1, by vol.) for at least one day. To obtain the mitotic metaphase chromosome preparations, the fixed root tips were digested with enzyme mixtures containing 4% cellulose R-10 (Yakult, http://www.yakult.co.jp) and 2% pectinase (Sigma-Aldrich, http://www.sigmaaldrich.com) in 1x PBS (phosphate-buffered saline, pH 5.5) buffer at 37 °C for 50 min. The enzyme solution was replaced with deionized water and maintained on ice for approximately 10 min. The digested root tips were then fixed in Carnoy's fixative solution. Slides with well-spread metaphase chromosomes were obtained using the 'flame dried' method according to the published protocol (Iovene et al. 2008). For meiotic pachytene chromosome preparation, young flower buds of *C. sativus* were harvested and fixed in 3:1 Carnoy's fixative solution (methanol:glacial acetic acid, 3:1, by vol.) for at least one day. The anthers at the pachytene stage were digested with the enzyme mixtures mentioned above at 37 °C for 1.5 h. The digested anthers were then fixed in 3:1 ethanol:glacial acetic acid (v/v) for at least 1 h. Slides with well-spreed pachytene chromosomes were obtained by the 'smear' method (Lysak et al. 2006).

In situ hybridization and signal detection

In situ hybridization was conducted according to Lou et al. (2010). A total 20 µl hybridization solution containing denatured probes, 50% formamide, and 10% dextran sulfate in 2× SSC (saline sodium citrate) was applied to the denatured slide. The slides were then incubated overnight at 37 °C. Chromosomes were counterstained with DAPI (4′,6-diamidino-2-phenylindole) (Vector Laboratories. Burlingame, USA). The fluorescence signals were detected using a fluorescein isothiocyanate-conjugated antibiotin antibody and a rhodamine-conjugated anti-digoxigenin antibody (Roche, http://www.roche-appliedscience.com). Images were captured with a SENSYS (http://www.photometrics.com) CCD color chromatic camera attached to an Olympus BX 51 microscope (http://www.olympus-global.com) controlled by Applied Spectral Imaging FISH view 5.5 software (Applied Spectral Imaging Inc, http://www.spectral-imaging.com). The photographic contrast was adjusted using Adobe Photoshop 6.0 (Adobe Systems, http://www.adobe.com). The chromosome numbering in the figures was performed according to Yang et al. (2012, 2014).

Results

Identification and characterization of tandem repeats in the *C. hystrix* genome

To evaluate the global repeat composition of the *C. hystrix* genome, a combined dataset of 487,102 randomly selected reads was analyzed using the RepeatExplorer pipeline. One hundred sixty-six clusters were generated with a cluster size threshold of 0.01%, and clusters of annotated putative mitochondrial and plastid contaminations were removed. Finally, 125 clusters remained and were used to calculate the genome proportions (Table 2). In the *C. hystrix* genome, approximately 19% of the genome consisted of repetitive DNAs. Retrotransposons, occupying 9.9% of the genome, made up the most abundant repetitive DNAs. Tandem repeats, including rDNA and satellites, made up

Table 2 Repeat elements and their genome proportions in the *C. hystrix* genome

Class	Order	Superfamily	GP (genome proportion, %)
Retrotransposon			9.901
	LTR		6.786
		Gypsy	3.262
		Copia	2.734
		Caulimovirus	0.739
		Unknown	0.051
	LINE		3.115
		L1	3.115
DNA transposon			0.233
	MULE.MuDR		0.15
	CMC.EnSpm		0.083
rDNA			3.809
Satellite			0.271
Unclassified			5.436
Total			19.065

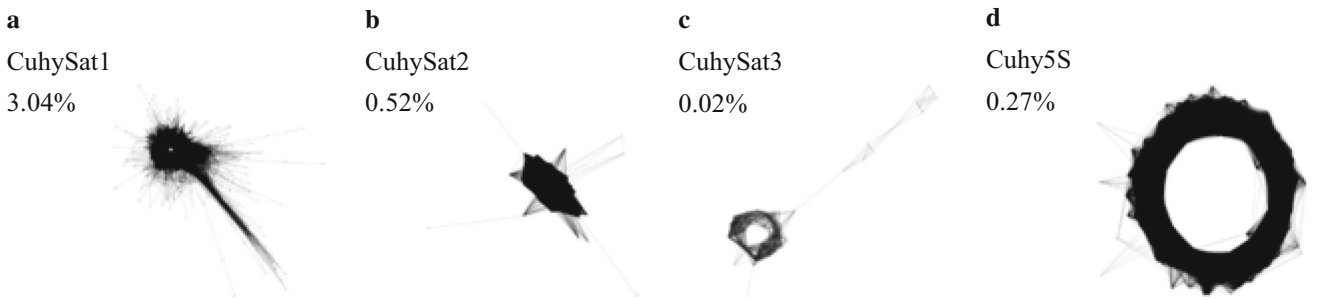

Fig. 1 Analysis of four major tandem repeats in the *C. hystrix* genome. *Star*-like or *circular graph* layouts of read clusters indicative of *C. hystrix* tandem repeats. The graph layouts corresponding to read clusters of **a** CuhySat1, **b** CuhySat2, **c** CuhySat3, and **d** Cuhy5S are represented. The *percentage* indicates the genome proportion of each cluster

4.08% of the genome, whereas MULE.MuDR and CMC.EnSpm transposons contributed to only 0.150 and 0.083%. In addition to 45S rDNA, four clusters generated by RepeatExplorer, which were represented by star-like or circular graphs typical for tandem repeats, have been selected for further chromosome structural and evolutionary analyses (Fig. 1).

The first three clusters (Fig. 1a–c) represent noncoding satellite DNAs and have been designated *Cucumis hystrix* satellite 1–3 (CuhySat1–CuhySat3) in decreasing abundance. The consensus sequences of their putative monomers are reconstructed and represented in Fig. 2. Comprising nearly 3.04% of the *C. hystrix* genome, CuhySat1 is a tandem repeat with a 197-bp-length monomer. A search of the GenBank databases revealed a high similarity to the *C. melo* satellite Csst1-Cm sequence (GenBank accession No. FJ410944, ~88% identity over 194 bp). CuhySat2, a satellite making up 0.52% of the genome, has a monomer length of 361 bp. The GenBank and PlantSat (Macas et al. 2002) databases-based BLAST analysis showed CuhySat2 shared a high similarity with a *C. sativus* subtelomeric satellite DNA, Type IV (GenBank accession No. X69163, ~81% identity over 281 bp) (Ganal and Hemleben 1998). Occupying a genome proportion of 0.02%, CuhySat3 corresponds to a moderately amplified satellite. Its monomer sequence is 342 bp in length. Unlike CuhySat1 and CuhySat2, CuhySat3 revealed no homology with any other known repeats. The consensus sequence of the fourth cluster, which was designated *Cucumis hystrix* 5S (Cuhy5S), displayed a high homology with plant 5S ribosomal RNA (rRNA) genes. An alignment of well-described plant 5S rRNA genes (Waminal et al. 2014; Cloix et al. 2000) served to delimit the boundaries of the 5S rDNA coding region and allowed the identification of the 94-bp Cuhy5S gene (Fig. 3a). With a length of 323 bp, the Cuhy5S monomer contains a 199-bp intergenic spacer flanking the gene. When compared with the *C. sativus* 5S rDNA (Cusa5S) monomer (Waminal et al. 2014), a sequence identity of 73% in the coding region and 59% in the intergenic spacer with two Indels and numerous

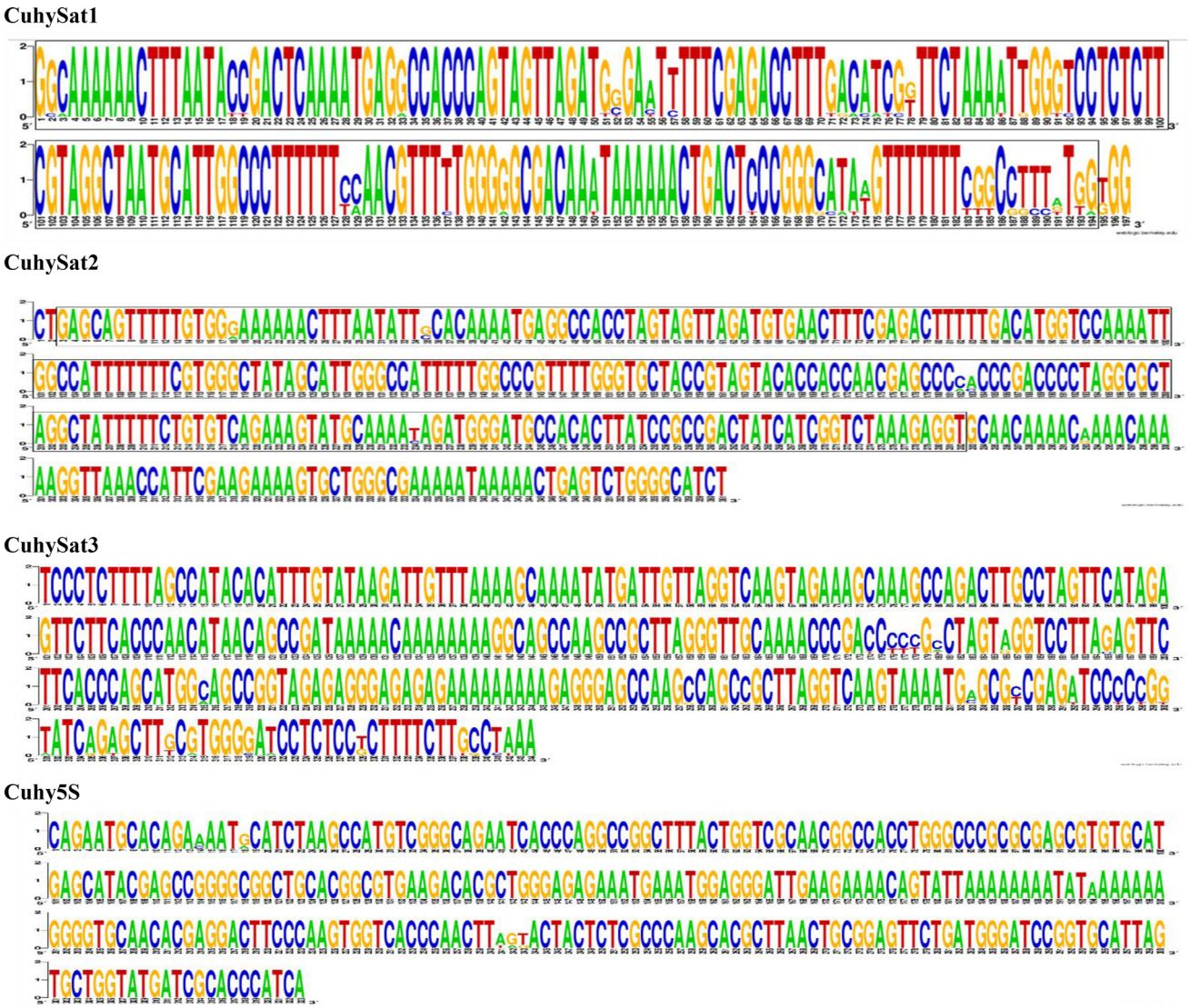

Fig. 2 Consensus sequences of CuhySat1, CuhySat2, CuhySat3, and Cuhy5S. The *black rectangle* outlines the positions in the sequences of CuhySat1 and CuhySat2 that share similarities with the *C. melo* satellite Csst1-Cm and *C. sativus* type IV, respectively

nucleotide polymorphism loci were observed (Fig. 3b). All four tandem repeat sequences were registered in NCBI GenBank with the accession numbers KY368163–KY368166.

Chromosomal distribution of tandem repeats revealed by FISH

FISH analysis revealed that all three CuhySat satellites were located exclusively at the distal ends of all *C. hystrix* mitotic metaphase chromosomes (Fig. 4b–d). A significant difference of satellite site size was observed in all three CuhySat satellites based on the FISH signal intensity (indicated by arrows in Fig. 4b–d). The subtelomeric locations of these three CuhySat satellites were confirmed by co-hybridization with the *Arabidopsis*-type telomeric DNA probe (Fig. 4f–h). When the three CuhySat satellites were comapped on the same mitotic metaphase cell, they showed the same hybridization signal pattern on *C. hystrix* mitotic metaphase chromosomes (Fig. 4i–o). Each CuhySat satellite produced a total of 28 subtelomeric signals, with 20 hybridization signals at ten pairs of chromosomes' one arm and eight signals at the both arms of two pairs of chromosomes. Self-GISH using *C. hystrix* genomic DNA produced the same signal pattern on each chromosome with these CuhySat satellites (Fig. 4a). Cuhy5S-specific signals were located at the interstitial regions of a single chromosome pair's one arm (indicated by arrows in Fig. 4e) together with one pair of terminal 45S rDNA signals.

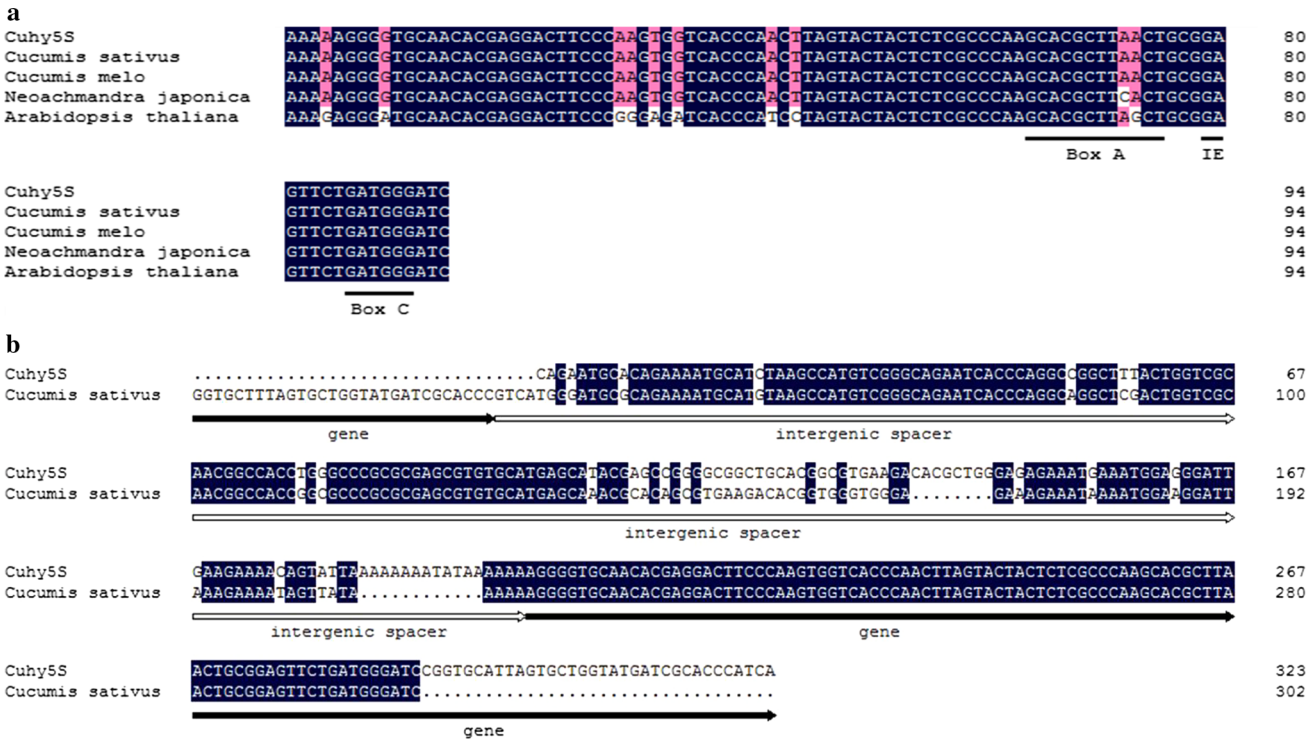

Fig. 3 Alignment of Cuhy5S to other plant 5S rRNAs. Conserved residues in all compared sequences are shaded. **a** The Cuhy5S gene boundaries were derived with the help of an alignment against four characterized plant 5S rRNA genes. The plant sequences included stem from *C. sativus* (KF543346), *C. melo* (KF543344), *Neoachmandra japonica* (KF543352), and *Arabidopsis thaliana* (AF198185). The region of the intragenic promotor, defined by three elements named *Box A*, Intermediate Element (IE), and *Box C* (Cloix et al. 2000) are marked. **b** The Cuhy5S monomer consensus has been aligned with the 5S rRNA monomer of *C. sativus* (KF543346). The coding region and the intergenic spacer are marked

Homology and evolution of tandem repeats between *C. hystrix* and *C. sativus*

In order to investigate the homology and evolution of tandem repeats between *C. hystrix* and *C. sativus*, comparative ISH was employed to probe the signals of four *C. hystrix* tandem repeats on *C. sativus* chromosomes. When CuhySat1 was used as a comparative ISH probe, it produced strong signals at the subtelomeric regions of all *C. sativus* mitotic metaphase chromosomes, including five pairs of chromosomes with signals at both arms and two pairs of chromosomes bearing signals at one arm (Fig. 5b). The subtelomeric locations of CuhySat1 were confirmed by co-hybridization with the *Arabidopsis*-type telomeric DNA probe (Fig. 5h, i). CuhySat2 not only produced strong subtelomeric signals located at all *C. sativus* mitotic metaphase chromosomes (Fig. 5c, f), but also with six clear pericentromeric signals on three pairs of chromosomes (indicated by arrowheads in Fig. 5c, f). When CuhySat1 and CuhySat2 were comapped on *C. sativus* mitotic metaphase and meiotic pachytene chromosomes, we observed that all the CuhySat1 subtelomeric signals were more close to telomere than the CuhySat2 subtelomeric signals (Fig. 5j, k). The *C. hystrix*-specific repeat CuhySat3 showed strong signals at the pericentromeric regions of all *C. sativus* mitotic metaphase chromosomes (Fig. 5d, g). The pericentromeric locations of CuhySat2 and CuhySat3 were confirmed by co-hybridization with the *C. sativus* centromeric Type III probe on *C. sativus* mitotic metaphase chromosomes. All of the CuhySat2 pericentromeric signals partially or completely overlapped with the centromeric Type III signals (Fig. 5f). When CuhySat3 and the centromeric Type III were comapped on *C. sativus* mitotic metaphase chromosomes, we observed three types of co-hybridization signal distribution patterns, overlapping, flanking and nested distribution patterns (Fig. 5g). When CuhySat2 and CuhySat3 were comapped on *C. sativus* mitotic metaphase and meiotic pachytene chromosomes, we observed that the pericentromeric signals of CuhySat3 flanked the pericentromeric signals of CuhySat2 (Fig. 5l, m). Comparative GISH showed that the *C. hystrix* genomic DNA also produced pericentromeric signals on some *C. sativus* chromosomes (indicated by arrowheads in Fig. 5a), excluding obvious subtelomeric signals on all *C. sativus* mitotic metaphase chromosomes (Fig. 5a). Two interstitial Cyhy5S signals were detected on one pair of *C. sativus*

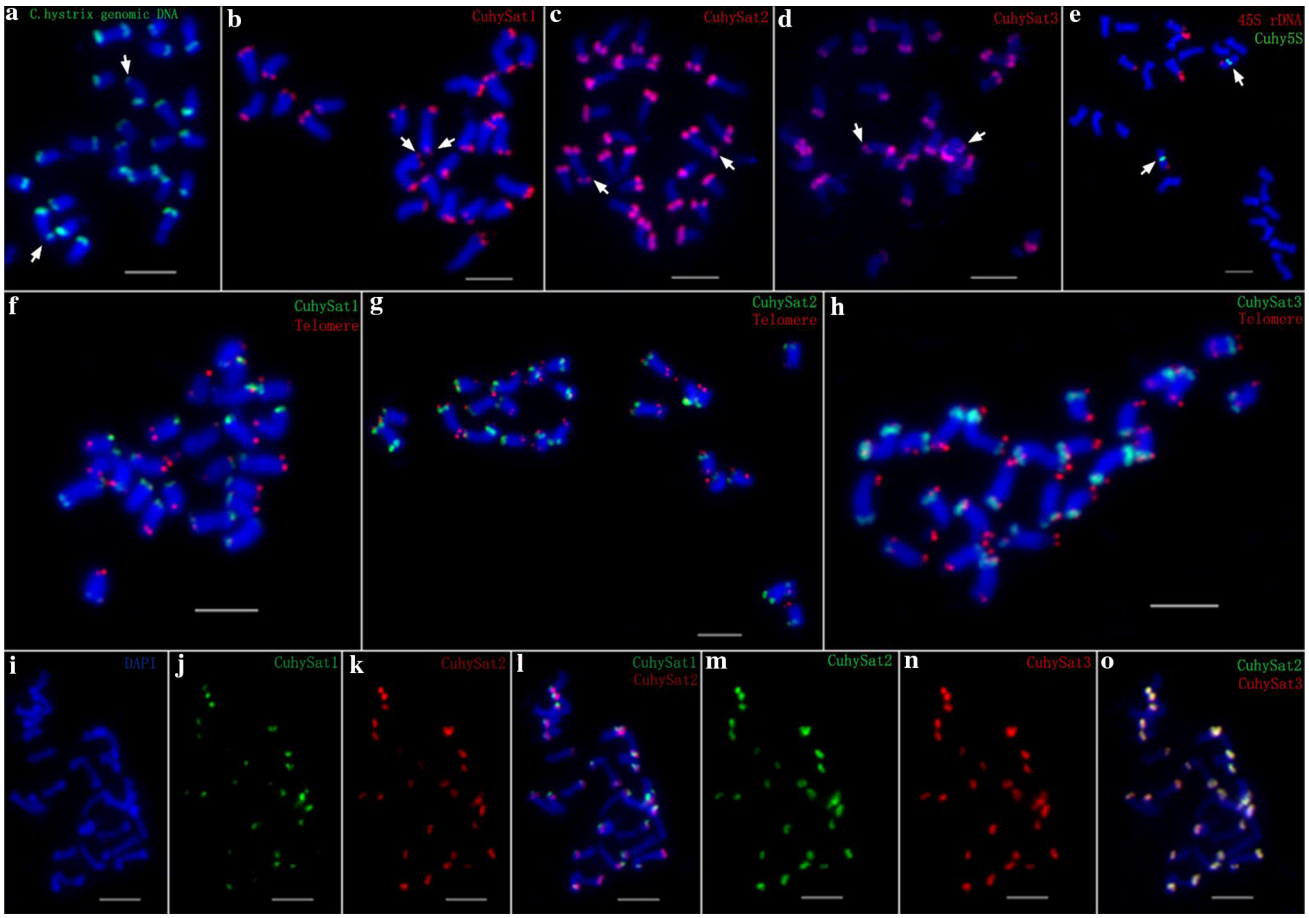

Fig. 4 Chromosomal location of *C. hystrix* genomic DNA, CuhySat1-3, Cuhy5S and *Arabidopsis*-type telomeric DNA on *C. hystrix* mitotic metaphase chromosomes. **a** Chromosomal location of *C. hystrix* genomic DNA (*green*) on its mitotic metaphase chromosomes. **b–d** Chromosomal location of CuhySat1-3 (*red*) on *C. hystrix* mitotic metaphase chromosomes. **e** Chromosomal location of Cuhy5S (*green*) and 45S rDNA (*red*) on *C. hystrix* mitotic metaphase chromosomes. **f** Chromosomal location of CuhySat1 (*green*) and telomeric DNA (*red*) on *C. hystrix* mitotic metaphase chromosomes. **g** Chromosomal location of CuhySat2 (*green*) and telomeric DNA (*red*) on *C. hystrix* mitotic metaphase chromosomes. **h** Chromosomal location of CuhySat3 (*green*) and telomeric DNA (*red*) on *C. hystrix* mitotic metaphase chromosomes. **i** DAPI-stained *C. hystrix* mitotic metaphase chromosomes. **j** Signals of CuhySat1 (*green*) on *C. hystrix* mitotic metaphase chromosomes. **k** Signals of CuhySat2 (*red*) on *C. hystrix* mitotic metaphase chromosomes. **l** Merged picture from **i, j** and **k**. **m** Signals of CuhySat2 (*green*) on *C. hystrix* mitotic metaphase chromosomes. **n** Signals of CuhySat3 (*red*) on *C. hystrix* mitotic metaphase chromosomes. **o** Merged picture from **i, m** and **n**. Arrows in **a–d** point to these weak subtelomeric signals derived from the *C. hystrix* genomic DNA and CuhySat1-3 probes, respectively. The *arrows* in **e** indicate the two interstitial signals derived from the Cuhy5S probe. *Scale bars* 5 μm

mitotic metaphase chromosomes, in agreement with our previous study (Fig. 5e) (Zhang et al. 2016).

The inferred evolution events resulting in the pericentromeric locations of two *C. hystrix* subtelomeric satellites on *C. sativus* chromosomes

Yang et al. (2014) chromosome synteny research between *C. hystrix* and *C. sativus* showed that *C. sativus* chromosome 7 (CS7) was highly conserved with *C. hystrix* chromosome 1 (CH1) with the occurrence of one pericentric inversion and one paracentric inversion. And *C. sativus* chromosome 5 (CS5) was probably originated from fusion of *C. hystrix* chromosome 9 (CH9) and *C. hystrix* chromosome 10 (CH10), followed by nine inversions. Combining these chromosome synteny, rearrangements with our comparative ISH results, we drew the idiograms to infer the possible evolution events resulting in the pericentromeric signals produced by subtelomeric CuhySat2 and CuhySat3 on *C. sativus* chromosomes (Fig. 6). On CS5 and CS7, we deduced that the pericentromeric location of CuhySat3 might be the result of a paracentric inversion during chromosome evolution (Fig. 6). On CS5, we inferred that some CuhySat2 sequences were left at the pericentromeric region of newly fused chromosome and formed pericentromeric

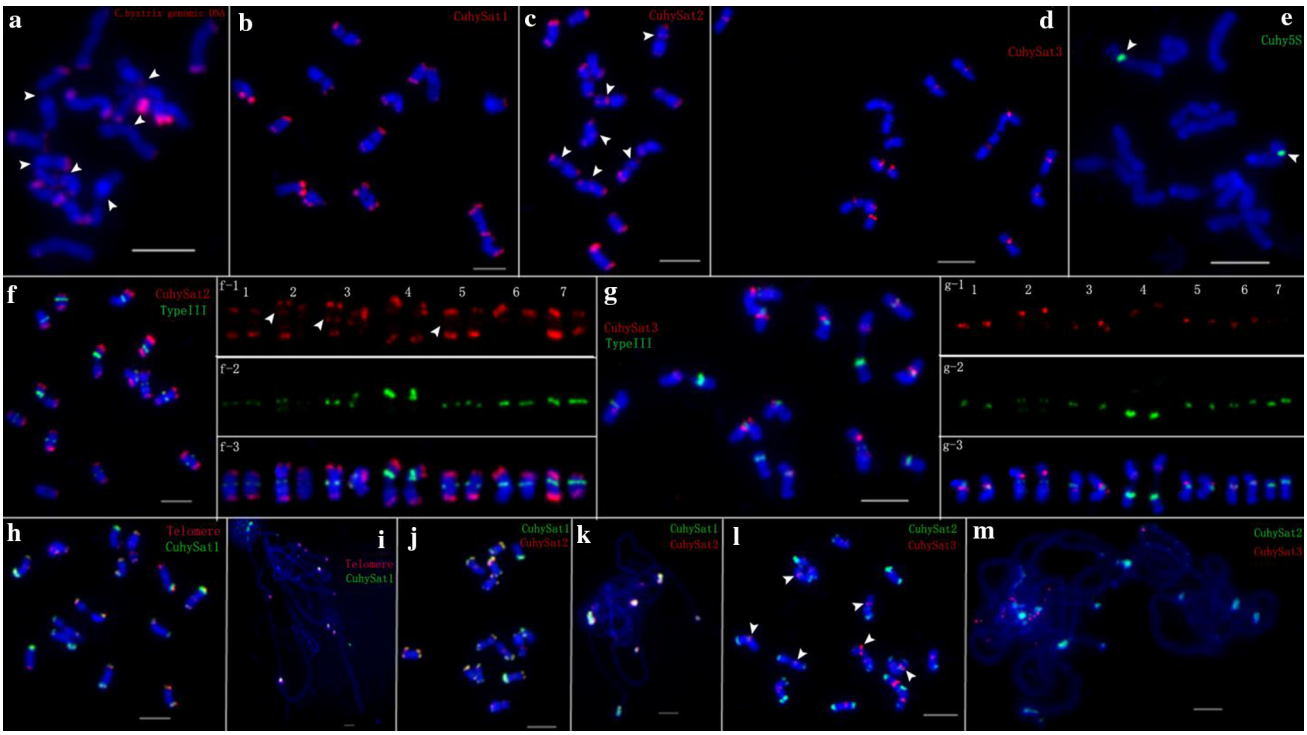

Fig. 5 FISH mapping of *C. hystrix* genomic DNA, CuhySat1-3, Cuhy5S, *Arabidopsis*-type telomeric DNA and *C. sativus* centromeric Type III probe on *C. sativus* chromosomes. **a** FISH mapping of *C. hystrix* genomic DNA (*red*) on *C. sativus* mitotic metaphase chromosomes. **b** FISH mapping of CuhySat1 (*red*) on *C. sativus* mitotic metaphase chromosomes. **c** FISH mapping of CuhySat2 (*red*) on *C. sativus* mitotic metaphase chromosomes. **d** FISH mapping of CuhySat3 (*red*) on *C. sativus* mitotic metaphase chromosomes. **e** FISH mapping of Cuhy5S (*green*) on *C. sativus* mitotic metaphase chromosomes. **f** FISH mapping of CuhySat2 (*red*) and the centromeric Type III probe (*green*) on *C. sativus* mitotic metaphase chromosomes. **g** FISH mapping of CuhySat3 (*red*) and the centromeric Type III probe (*green*) on *C. sativus* mitotic metaphase chromosomes. **h** FISH mapping of telomeric DNA (*red*) and CuhySat1 (*green*) on *C. sativus* mitotic metaphase chromosomes. **i** FISH mapping of telomeric DNA (*red*) and CuhySat1 (*green*) on *C. sativus* meiotic pachytene chromosomes. **j** FISH mapping of CuhySat1 (*green*) and CuhySat2 (*red*) on *C. sativus* mitotic metaphase chromosomes. **k** FISH mapping of CuhySat1 (*green*) and CuhySat2 (*red*) on *C. sativus* meiotic pachytene chromosomes. **l** FISH mapping of CuhySat2 (*green*) and CuhySat3 (*red*) on *C. sativus* mitotic metaphase chromosomes. **m** FISH mapping of CuhySat2 (*green*) and CuhySat3 (*red*) on *C. sativus* meiotic pachytene chromosomes. *Arrowheads* in **a**, **c** and **f–l** point to these clear pericentromeric signals derived from the *C. hystrix* genomic DNA and CuhySat2 probes, respectively. *Arrowheads* in **l** point to these pericentromeric signals of CuhySat3 that flanked the pericentromeric signals of CuhySat2. The *arrowheads* in **e** indicate the two interstitial signals derived from the Cuhy5S probe. *Scale bars* 5 μm

satellite sites after sequence amplification and accumulation (Fig. 6b).

Discussion

Tandem repeats are a critical class of repetitive DNAs in higher eukaryotes. Long stretches of tandem repeats may reside in a special chromosomal locus as a hallmark, functional unit, or a driving force for chromosome evolution (Plohl et al. 2008). In this study, graph-based clustering enabled the identification of four major clusters as tandem repeats by sequence-similarity searches and graphed structural analyses, including three satellite DNAs and one 5S rDNA repeat. Satellite DNA sequences are considered as a rapidly evolving component of eukaryotic genomes, comprising highly repetitive, tandemly arrayed and highly conserved monomer sequences. However, different families of satellite DNA are characterized by a huge variety of sequence compositions and great diversities in repeat monomer length and abundance (Garrido-Ramos 2015). In most plants and animals, satellite DNAs display a variable AT-rich repeat unit with monomer length ranging from 150 to 400 bp, usually forming arrays up to 100 Mb (Mehrotra and Goyal 2014). Based on this classical view of satellites, CuhySat1-3 are standard satellite DNAs that display AT-rich features with an AT content of 56.9, 56.2, and 53.5%, respectively. They have a repeat monomer length of 197, 361, and 344 bp, making up approximately 13.13 Mb, 2.25 Mb, and 88.47 kb of the *C. hystrix* genome.

Satellite DNA families accumulate in heterochromatin at different regions of eukaryotic chromosomes, including mainly subtelomeric and pericentromeric regions (Garrido-Ramos 2015). Subtelomeric regions are one of the most

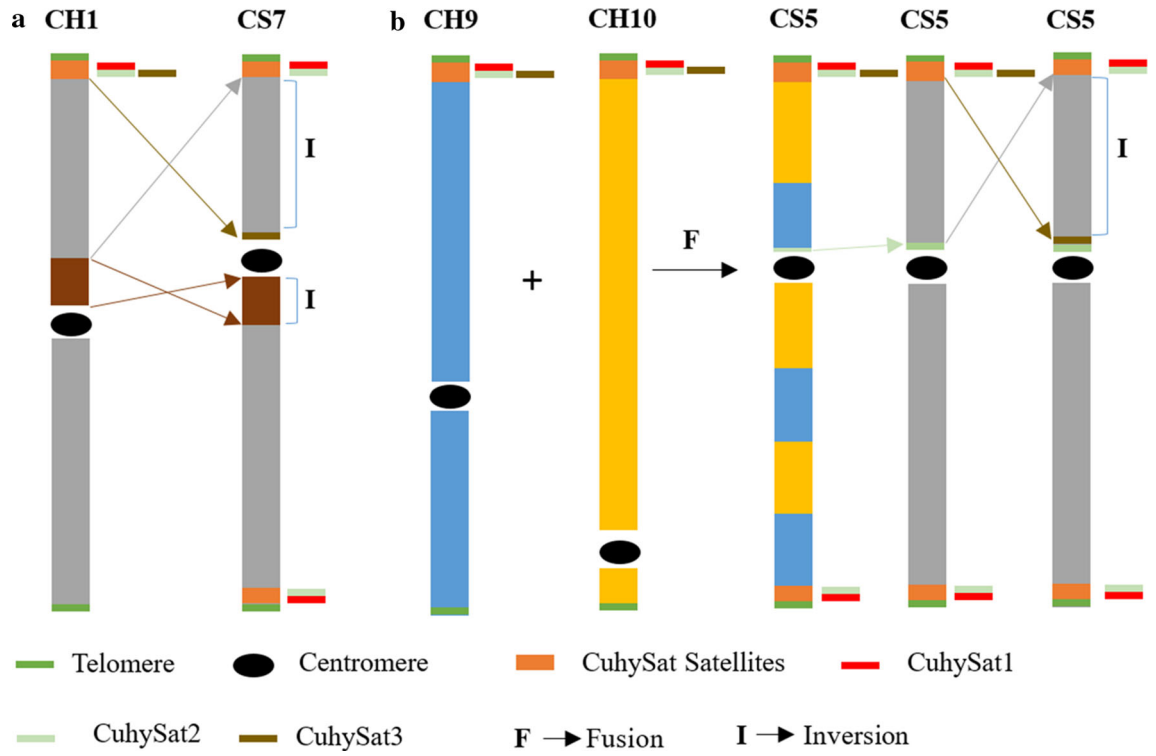

Fig. 6 Graphical representation of inferred evolution events resulting in the pericentromeric locations of two *C. hystrix* subtelomeric satellites on *C. sativus* chromosomes. **a** One pericentric inversion and another paracentric inversion occurred during *C. sativus* chromosome 7 (CS7) evolved from *C. hystrix* chromosome 1 (CH7), and the pericentromeric location of CuhySat3 might be the result of the paracentric inversion. **b** Idiogram of inferred chromosomal fusion and rearrangements that resulted in the pericentromeric locations of CuhySat2 and CuhySat3 on *C. sativus* chromosome 5 (CS5) which was probably originated from fusion of *C. hystrix* chromosome 9 (CH9) and *C. hystrix* chromosome 10 (CH10)

dynamic and rapidly evolving regions, it is common to find several different subtelomeric satellites within the same species (Richard et al. 2013; Torres et al. 2011). Navajas-Pérez et al. (2009) supposed that the sequence relationship between different satellite DNA families found within the same genome was indicative of common origins of currently highly differentiated satellite DNAs. A comparative analysis of CuhySat1-3 revealed that there was no homology between them so that they might not be differentiated from a common ancestor. However, their relatively conserved distribution patterns remain an interesting topic for further studies on investigating the evolutionary process of these satellites.

The 5S rDNA is present universally in plants, with multiple copies of coding sequences and intergenic spacers as tandem repeats in the genome (Katsiotis et al. 2000). Zhang et al. (2016) found that Cusa5S exhibited a high level of related-species homogeneity when were hybridized as a FISH probe onto chromosomes of different *Cucumis* species. It produced one pair of interstitial signals on both *C. sativus* and *C. hystrix* mitotic metaphase chromosomes. Our FISH results showed that Cuhy5S produced the same signal pattern with Cusa5S on one single chromosome pair of *C. sativus* and *C. hystrix*. The sequence-similarity alignment indicated that Cuhy5S and Cusa5S shared a conserved coding gene and a variable spacer region. Therefore, we inferred it might be the conserved coding gene that resulted in the same FISH signaling pattern. Because Cuhy5S is the first 5S rDNA to be detailed identified among wild species of *Cucumis*, whether the coding gene is the decisive factor for the highly related-species homogeneity of Cusa5S among *Cucumis* species will require the analysis of more 5S rDNA information from other wild species.

Many molecular phylogenetic studies suggest that $n = 12$ is ancestral in the genus *Cucumis*. A postulated fusion hypothesis indicated that $n = 7$ was derived from $n = 12$ via unequal translocations, inversions and fusions of non-homologous chromosomes (Ghebretinsae et al. 2007; Renner et al. 2007; Sebastian et al. 2010; Yang et al. 2014). Based on the chromosome synteny and collinearity analysis between *C. hystrix* and *C. sativus*, Yang et al. (2014) inferred that at least 59 chromosome rearrangement events have occurred during the dysploid chromosome reduction, including 50 inversions, five fusions and four translocations. Our comparative FISH results showed that

CuhySat2 not only maintained its subtelomeric positions on all *C. sativus* chromosomes, but also produced pericentromeric signals on some chromosomes. Minor pericentromeric signals might be the result of dynamic rearrangements or sequence exchanges between non-homologous chromosomes, which could lead to the formation and amplification of new satellite families or loci (Macas et al. 2006). Two kinds of signals produced by CuhySat2 revealed that some chromosomal rearrangements might have occurred after chromosome fusion. Some CuhySat2 sequences were left at the pericentromeric regions of some newly fused chromosomes and formed new pericentromeric satellite sites after sequence amplification and accumulation.

Unlike CuhySat2, CuhySat3 located all its signals at the pericentromeric regions of all *C. sativus* chromosomes. Different comparative FISH signal patterns revealed that they might have undergone different evolutionary processes. Comparative FISH mapping and sequence alignments revealed a one-to-one whole chromosome synteny of CS7 with CH1, with one pericentric inversion and one paracentric inversion occurred during *C. sativus* domestication (Yang et al. 2014). On CS7, the pericentromeric location of CuhySat3 might be the result of a paracentric inversion during chromosome evolution. Co-localization of CuhySat3 with the *C. sativus* centromeric Type III probe showed its pericentromeric signals flanking or overlapping the centromeric signals on all *C. sativus* mitotic metaphase chromosomes. Therefore, we inferred that a paracentric inversion might also have occurred on all fused chromosomes after chromosome fusion. However, the whole arm did not participate in the paracentric inversion event. The majority of the chromosome arm ends remained and were involved in the stabilization of the chromosome arms, thus supporting a telomere function.

In this study, the de novo identification of tandem repeats has been proved to be an efficient method for characterizing repetitive DNAs in the *C. hystrix* genome. This is the first study to show three different CuhySat satellites from one species with the same chromosomal distribution pattern. In addition, the pericentromeric signals of two subtelomeric CuhySat satellites along *C. sativus* chromosomes provide some useful information supporting the underlying chromosomal fusion hypothesis. Altogether, the results presented herein provide a fundamental understanding of the *C. hystrix* chromosome structure and should be helpful for further studies of chromosome evolution in the genus *Cucumis*.

Author contribution statement SQY and JFC conceived and designed research. SQY, XDQ, CYC and ZAL conducted experiments. SQY, QFL and JL analyzed data. JFC and SQY wrote the manuscript. All authors read and approved the final manuscript.

Acknowledgements This research was funded by the Natural Science Foundation of China (Nos. 31430075 and 31471872) and the National Key Research and Development Program of China (Nos. 2016YFD0101705 and 2016YFD0100204-25).

Compliance with ethical standards

Conflict of interest The authors declare that they have no competing interests.

References

Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 25:3389–3402

Benson G (1999) Tandem repeats finder: a program to analyze DNA sequences. Nucleic Acids Res 27:573–580

Beridze T (1986) Satellite DNA. Springer, Berlin

Cai Z, Liu H, He Q, Pu M, Chen J, Lai J, Li X, Jin W (2014) Differential genome evolution and speciation of *Coix lacryma-jobi* L. and *Coix aquatica* Roxb. hybrid guangxi revealed by repetitive sequence analysis and fine karyotyping. BMC Genom 15:1025

Chen JF, Staub JE, Tashiro Y, Isshiki S, Miyazaki S (1997) Successful interspecific hybridization between *Cucumis sativus* L. and *C. hystrix* Chakr. Euphytica 96:413–419

Chen JF, Staub JE, Adelberg JW, Jiang JM (1999) Physical mapping of 45S rRNA genes in *Cucumis* species by fluorescence in situ hybridization. Can J Bot 77:389–393

Clark K, Karsch-Mizrachi I, Lipman DJ, Ostell J, Sayers EW (2016) GenBank. Nucleic Acids Res 44:D67–72

Cloix C, Tutois S, Mathieu O, Cuvillier C, Espagnol MC, Picard G, Tourmente S (2000) Analysis of 5S rDNA arrays in *Arabidopsis thaliana*: physical mapping and chromosome-specific polymorphisms. Genome Res 10:679–690

Dvořáčková M, Fojtová M, Fajkus J (2015) Chromatin dynamics of plant telomeres and ribosomal genes. Plant J 83:18–37

Fajkus P, Peška V, Sitová Z, Fulnečková J, Dvořáčková M, Gogela R, Sýkorová E, Hapala J, Fajkus J (2016) *Allium* telomeres unmasked: the unusual telomeric sequence (CTCGGTTATGGG)*n* is synthesized by telomerase. Plant J 85:337–347

Ganal M, Hemleben V (1998) Insertion and amplification of a DNA sequence in satellite DNA of *Cucumis sativus* L. (cucumber). Theor Appl Genet 75:357–361

Garrido-Ramos MA (2015) Satellite DNA in Plants: more than Just Rubbish. Cytogenet Genome Res 146:153–170

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

Ghebretinsae AG, Thulin M, Barber JC (2007) Relationships of cucumbers and melons unraveled: molecular phylogenetics of Cucumis and related genera (Benincaseae, Cucurbitaceae). Am J Bot 94:1256–1266

Grzywacz B, Chobanov DP, Maryańska-Nadachowska A, Karamysheva TV, Heller KG, Warchalowska-Śliwa E (2014) A comparative study of genome organization and inferences for the systematics of two large bush-cricket genera of the tribe Barbitistini (Orthoptera: Tettigoniidae: Phaneropterinae). BMC Evol Biol 14:48

He QY, Cai ZX, Hu TH, Liu HJ, Bao CL, Mao WH, Jin WW (2015) Repetitive sequence analysis and karyotyping reveals centromere-associated DNA sequences in radish (*Raphanus sativus* L.). BMC Plant Biol 15:105

Heitkam T, Petrasch S, Zakrzewski F, Kögler A, Wenke T, Wanke S, Schmidt T (2015) Next-generation sequencing reveals differentially amplified tandem repeats as a major genome component of Northern Europe's oldest *Camellia japonica*. Chromosome Res 23:791–806

Ijdo JW, Wells RA, Baldini A, Reeders ST (1991) Improved telomere detection using a telomere repeat probe (TTAGGG)n generated by PCR. Nucleic Acids Res 19:4780

Iovene M, Wielgus SM, Simon PW, Buell CR, Jiang JM (2008) Chromatin structure and physical mapping of chromosome 6 of potato and comparative analyses with tomato. Genetics 180:1307–1317

Jiang B, Wu ZM, Lou QF, Wang D, Zhang WP, Chen JF (2010) Genetic diversity of Ty1-copia retrotransposons in a wild species of *Cucumis* (*C. hystrix*). Sci Hortic 127:46–53

Katsiotis A, Loukas M, Heslop-Harrison JS (2000) Repetitive DNA, genome and species relationships in *Avena* and *Arrhenatherum* (Poaceae). Ann Bot 86:1135–1142

Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C, Thierer T, Ashton B, Meintjes P, Drummond A (2012) Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics 28:1647–1649

Koo DH, Nam YW, Choi D, Bang JW, de Jong H, Hur Y (2010) Molecular cytogenetic mapping of *Cucumis sativus* and *C. melo* using highly repetitive DNA sequences. Chromosome Res 18:325–336

Lim KY, Kovarik A, Matyasek R, Chase MW, Knapp S, McCarthy E, Clarkson JJ, Leitch AR (2006) Comparative genomics and repetitive sequence divergence in the species of diploid *Nicotiana* section Alatae. Plant J 48:907–919

Lou QF, Iovene M, Spooner DM, Buell CR, Jiang JM (2010) Evolution of chromosome 6 of *Solanum* species revealed by comparative fluorescence in situ hybridization mapping. Chromosoma 119:435–442

Lou QF, He YH, Cheng CY, Zhang ZH, Li J, Huang SW, Chen JF (2013) Integration of high-resolution physical and genetic map reveals differential recombination frequency between chromosomes and the genome assembling quality in cucumber. PLoS One 8:e62676

Lysak MA, Berr A, Pecinka A, Schmidt R, McBreen K, Schubert I (2006) Mechanisms of chromosome number reduction in *Arabidopsis thaliana* and related Brassicaceae species. Proc Natl Acad Sci USA 103:5224–5229

Macas J, Požárková D, Navrátilová A, Neumann P (2000) Two new families of tandem repeats isolated from genus *Vicia* using genomic self-priming PCR. Mol Gen Genet 263:741–751

Macas J, Mészáros T, Nouzová M (2002) PlantSat: a specialized database for plant satellite repeats. Bioinformatics 18:28–35

Macas J, Navrátilová A, Koblížková A (2006) Sequence homogenization and chromosomal localization of VicTR-B satellites differ between closely related *Vicia* species. Chromosoma 115:437–447

Mehrotra S, Goyal V (2014) Repetitive sequences in plant nuclear DNA: types, distribution, evolution and function. Genom Proteom Bioinform 12:164–171

Murray MG, Thompson WF (1980) Rapid isolation of high molecular weight plant DNA. Nucleic Acids Res 8:4321–4325

Navajas-Pérez R, Schwarzacher T, Ruiz Rejón M, Garrido-Ramos MA (2009) Characterization of RUSI, a telomere-associated satellite DNA, in the genus *Rumex* (Polygonaceae). Cytogenet Genome Res 124:81–89

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

Novák P, Neumann P, Pech J, Steinhaisl J, Macas J (2013) RepeatExplorer: a galaxy-based web server for genome-wide characterization of eukaryotic repetitive elements from next-generation sequence reads. Bioinformatics 29:792–793

Peška V, Fajkusm P, Fojtová M, Dvořáčková M, Hapala J, Dvořáček V, Polanská P, Leitch A, Sýkorová E, Fajkus J (2015) Characterisation of an unusual telomere motif (TTTTTTAGGG)n in the plant *Cestrum elegans* (Solanaceae), a species with a large genome. Plant J 82:644–654

Plohl M, Luchetti A, Meštrović N, Mantovani B (2008) Satellite DNAs between selfishness and functionality: structure, genomics and evolution of tandem repeats in centromeric (hetero)chromatin. Gene 409:72–82

Renner SS, Schaefer H, Kocyan A (2007) Phylogenetics of *Cucumis* (Cucurbitaceae): cucumber (*C. sativus*) belongs in an Asian/Australian clade far from melon (*C. melo*). BMC Evol Biol 7:58

Richard MM, Chen NW, Thareau V, Pflieger S, Blanchet S, Pedrosa-Harand A, Iwata A, Chavarro C, Jackson SA, Geffroy V (2013) The Subtelomeric *khipu* satellite repeat from *Phaseolus vulgaris*: lessons learned from the genome analysis of the andean genotype G19833. Front Plant Sci 4:109

Rosato M, Galian JA, Rossello JA (2012) Amplification, contraction and genomic spread of a satellite DNA family (E180) in *Medicago* (Fabaceae) and allied genera. Ann Bot 109:773–782

Sebastian P, Schaefer H, Telford IR, Renner SS (2010) Cucumber (*Cucumis sativus*) and melon (*C. melo*) have numerous wild relatives in Asia and Australia, and the sister species of melon is from Australia. Proc Natl Acad Sci USA 107:14269–14273

Sharma S, Raina SN (2005) Organization and evolution of highly repeated satellite DNA sequences in plant chromosomes. Cytogenet Genome Res 109:15–26

Smit AFA, Hubley R, Green P (2014) RepeatMasker open-4.0.5 (RMLib: 20140131 & Dfam: 1.3). http://www.repeatmasker.org

Torres GA, Gong Z, Iovene M, Hirsch CD, Buell CR, Bryan GJ, Novák P, Macas J, Jiang JM (2011) Organization and evolution of subtelomeric satellite repeats in the potato genome. Genes Genom Genet 1:85–92

Waminal NE, Ryu KB, Park BR, Kim HH (2014) Phylogeny of Cucurbitaceae species in Korea based on 5S rDNA non-transcribed spacer. Genes Genom 36:57–64

Weiss-Schneeweiss H, Leitch AR, McCann J, Jang T-S, Macas J (2015) Employing next generation sequencing to explore the repeat landscape of the plant genome. In: Hörandl E, Appelhans MS (eds) Next-generation sequencing in plant systematics. International Association for Plant Taxonomy Publishing, Germany, pp 1–25

Weng Y, Li W, Devkota RN, Rudd JC (2005) Microsatellite markers associated with two *Aegilops tauschii*-derived greenbug resistance loci in wheat. Theor Appl Genet 110:462–469

Yagi K, Pawelkowicz M, Osipowski P, Siedlecka E, Przybecki Z, Tagashira N, Hoshi Y, Malepszy S, Plader W (2015) Molecular cytogenetic analysis of *Cucumis* wild species distributed in Southern Africa: physical mapping of 5S and 45S rDNA with DAPI. Cytogenet Genome Res 146:80–87

Yang LM, Koo DH, Li YH, Zhang XJ, Luan FS, Havey MJ, Jiang JF, Weng YQ (2012) Chromosome rearrangements during domestication of cucumber as revealed by high-density genetic mapping and draft genome assembly. Plant J 71:895–906

Yang LM, Koo DH, Li DW, Zhang T, Jiang JM, Luan FS, Renner SS, Hénaff E, Sanseverino W, Garcia-Mas J, Casacuberta J, Senalik DA, Simon PW, Chen JF, Weng YQ (2014) Next-generation sequencing, FISH mapping and synteny-based modeling reveal mechanisms of decreasing dysploidy in *Cucumis*. Plant J 77:16–30

Zhang YX, Cheng CY, Li J, Yang SQ, Wang YZ, Li ZA, Chen JF, Lou QF (2015) Chromosomal structures and repetitive sequences divergence in *Cucumis* species revealed by comparative cytogenetic mapping. BMC Genomics 16:730

Zhang ZT, Yang SQ, Li ZA, Zhang YX, Wang YZ, Cheng CY, Li J, Chen JF, Lou QF (2016) Comparative chromosomal localization of 45S and 5S rDNAs and implications for genome evolution in *Cucumis*. Genome 59:449–457

Zhao X, Lu J, Zhang Z, Hu J, Huang S, Jin W (2011) Comparison of the distribution of the repetitive DNA sequences in three variants of *Cucumis sativus* reveals their phylogenetic relationships. J Genet Genomics 38:39–45