

Molecular cytogenetic mapping of *Cucumis sativus* and *C. melo* using highly repetitive DNA sequences

Dal-Hoe Koo · Young-Woo Nam · Doil Choi ·
Jae-Wook Bang · Hans de Jong · Yoonkang Hur

Received: 23 November 2009 / Accepted: 28 January 2010 / Published online: 3 March 2010
© Springer Science+Business Media B.V. 2010

Abstract Chromosomes often serve as one of the most important molecular aspects of studying the evolution of species. Indeed, most of the crucial mutations that led to differentiation of species during the evolution have occurred at the chromosomal level. Furthermore, the analysis of pachytene chromosomes appears to be an invaluable tool for the study of evolution due to its effectiveness in chromosome identification and precise physical gene mapping. By applying fluorescence in situ hybridization of 45S rDNA and CsCent1 probes to cucumber pachytene

chromosomes, here, we demonstrate that cucumber chromosomes 1 and 2 may have evolved from fusions of ancestral karyotype with chromosome number $n=12$. This conclusion is further supported by the centromeric sequence similarity between cucumber and melon, which suggests that these sequences evolved from a common ancestor. It may be after or during speciation that these sequences were specifically amplified, after which they diverged and specific sequence variants were homogenized. Additionally, a structural change on the centromeric region of

Responsible Editor: Herbert Macgregor.

Electronic supplementary material The online version of this article (doi:10.1007/s10577-010-9116-0) contains supplementary material, which is available to authorized users.

D.-H. Koo · J.-W. Bang · Y. Hur (✉)
School of Bioscience and Biotechnology,
Chungnam National University,
Daejeon 305-764, South Korea
e-mail: ykhur@cnu.ac.kr

Y.-W. Nam
Life Science, College of Natural Science,
Sogang University,
Seoul 100-611, South Korea

D. Choi
Department of Plant Sciences,
College of Agricultural and Life Sciences,
Seoul National University,
Seoul 151-742, South Korea

H. de Jong
Laboratory of Genetics,
Wageningen University and Research Center,
Wageningen, the Netherlands

D.-H. Koo (✉)
Department of Horticulture,
University of Wisconsin-Madison,
Madison, WI 53706, USA
e-mail: dkoo@wisc.edu

cucumber chromosome 4 was revealed by fiber-FISH using the mitochondrial-related repetitive sequences, BAC-E38 and CsCent1. These showed the former sequences being integrated into the latter in multiple regions. The data presented here are useful resources for comparative genomics and cytogenetics of *Cucumis* and, in particular, the ongoing genome sequencing project of cucumber.

Keywords *Cucumis* · centric fusion · fluorescence in situ hybridization · pachytene chromosome

Abbreviations

BAC	Bacterial artificial chromosome
CCD	Charge-coupled device
CTAB	Cetyltrimethylammonium bromide
DAPI	4',6-Diamidino-2-phenylindole
FISH	Fluorescence in situ hybridization
FITC	Fluorescein isothiocyanate
NOR	Nucleolar-organizing region
PI	Propidium iodide

Introduction

The genus *Cucumis* consists of about 30 species including *Cucumis sativus* L. (cucumber, $2n = 2x = 14$) and *Cucumis melo* L. (melon, $2n = 2x = 24$). *C. sativus* and *C. melo* are widely cultivated throughout the world. Cucumber and melon diverged from a common ancestor approximately nine million years ago (Schaefer et al. 2009). Recent comparative genetic analysis clearly showed that the five of the cucumber's seven chromosomes arose from fusions of ten ancestral chromosomes after divergence from *C. melo* (Huang et al. 2009). Comparative fluorescence in situ hybridization (FISH) mapping using the same sets of fosmid clones revealed different centromeric positions between two pairs of chromosomes from these two species (Han et al. 2009). Thus, chromosomal fusion and centromere repositioning occurred during the evolution of cucumber and melon chromosomes. However, the reduction of total cucumber chromosome number from 24 to the current count of 14 chromosomes remains to be examined in a further study. There is still uncertainty toward potential karyotypic evolution between the two species, and it will likely require more extensive evaluation and careful scrutiny before a

complete understanding of chromosome evolution between these species can take place.

One of the fundamental tools used for elucidating the origin of genome/chromosomal rearrangements between related taxa is the study of large repetitive DNA families (Hirai et al. 1995, 1996; Cerbah et al. 1998; Shan et al. 2003). The two most common families are the ubiquitous tandemly arrayed 5S and 45S ribosomal DNA family and the telomeric repeat. These repeats often reside on different chromosomes, and the 45S rDNA repeats are always located in nucleolus organizer regions (NORs). Other tandem DNA repeats are confined to closely related species and have been reported to play a role in the dynamics of genomic rearrangements in general, particularly in chromosomal changes associated with speciation (Wichman et al. 1991; Bradley and Wichman 1994; Garagna et al. 1997). Two satellite repeats in mice (Garagna et al. 2001), Muntjac deer (Li et al. 2000), and various regions of human short-arm acrocentrics, particularly satellite III DNA sequences, have been shown to be involved in Robertsonian translocations (Sullivan et al. 1996; Page et al. 1996). Moreover, the use of satellite DNA as landmarks of chromosomal evolution has been discussed (Ugarkovic and Plohl 2002; Plohl et al. 2008). The 45S rDNA was also used as a cytogenetic marker to track chromosomal evolution in different *Triticum* species, including polyploid wheat (Jiang and Gill 1994). Although several repetitive elements have been isolated from cucumber (Ganal et al. 1986; Ganal and Hemleben 1988; Koo et al. 2005; Han et al. 2008) and melon (Brennicke and Hemleben 1983), the chromosomal organization and evolutionary aspects of these repetitive elements are largely unknown.

FISH, the most widespread technology for mapping chromosomal repeats, allows for the simultaneous detection of different probes in the same or in successive treatments on the same cells (Pinkel et al. 1986; Cheng et al. 2001; Kato et al. 2004). The technology is versatile and can be used on highly condensed metaphase complements, interphase nuclei, and even extended DNA fibers (Fransz et al. 1996; de Jong et al. 1999; Jackson et al. 1998; Jiang and Gill 2006). A major breakthrough in plant cytogenetics has been the development of high-resolution FISH on pachytene chromosomes (Fransz et al. 1998, 2000; Peterson et al. 1999; Cheng et al. 2001). Pachytene chromosomes are favorite targets in

FISH, especially in diploid plant species with low chromosome numbers and low- to medium-sized genomes. This is due to a ten to 50 greater spatial resolution than that for mitotic chromosomes and a well-differentiated pattern of heterochromatin. FISH mapping on pachytene chromosomes allows for centromeres, NORs, and heterochromatin blocks to be distinguished as landmarks. These regions are characterized by high numbers of tandem and dispersed repeats and low numbers of genes and are morphologically recognizable in PI- or 4',6-diamidino-2-phenylindole (DAPI)-stained cell spreads. Thus, pachytene chromosomes provide cytological marks for chromosome identification and an essential tool for studying molecular and biochemical features associated with euchromatin and/or heterochromatin (Koo and Jiang 2009).

The centromeres are composed of long arrays of tandem DNA repeats such as the 171-bp α -satellite (Wevrick and Willard 1989), the 180-bp pAL1 repeat (Maluszynska and Heslop-Harrison 1991), and the 155-bp CentO repeat (Cheng et al. 2002), but they are not conserved. In contrast, proteins associated with centromeric sequences can be conserved. The most conserved of all known centromere proteins is centromere-specific histone H3, CENH3 (Henikoff et al. 2001). However, the N-terminal domain of CENH3 is in general highly variable between species. For example, the CENH3s are known variously as CENP-A (human), Cse4 (*Saccharomyces cerevisiae*), Cnp1 (*Schizosaccharomyces pombe*), Cid (fly), and HTR12 (*Arabidopsis*) (Choo 2001; Talbert et al. 2002; Jiang et al. 2003).

In this study, we isolated and sequenced several repetitive DNA clones derived from the pericentromeric regions of cucumber and melon chromosomes. We found that the CsCent1 satellite and CsCent2 retrotransposon in cucumber and the CmCent satellite in melon are the primary DNA components of cucumber and melon centromeres, respectively. Moreover, by localizing CsCent1 and 45S rDNA on cucumber pachytene chromosomes, we found an unusual distribution of the FISH signal which appears to have evolved through a chromosomal fusion in the ancestors of cucumber. Fiber-FISH localization of mitochondrial DNA-related sequence revealed that structural rearrangements were detected on the centromeric region of cucumber chromosome 4.

Materials and methods

Plant materials

Cucumber (*C. sativus* L. cv. Winter Long) and melon (*C. melo* L. cv. Romance) seeds were obtained from Dongbu Hannong Seeds Co., Korea. Plants were grown under greenhouse conditions. Young leaves were collected for DNA isolation, and young fast-growing root tips and anthers at meiotic prophase I were selected for chromosome preparations.

Cloning of repetitive DNA elements

An overview of all repeats used in this study is presented in Table 1. The CsCent1 and CsPc pericentromere repeats from the repeat-rich B68 and B17 bacterial artificial chromosome (BAC) clones were selected. Sequencing the ends of BAC-B68 revealed a retrotransposon-like DNA, which was used to design the oligonucleotide primer 5'-TAGAGGATCCATTAACCTAT-3' for amplifying the 2,107-bp CsCent2 fragment; this fragment was then subcloned in the pGEM-T-Easy vector (Promega). A second pericentromere repeat, CmCent, was isolated by polymerase chain reaction (PCR) amplification of *C. melo* genomic DNA using oligonucleotide primer pair 5'-AGCTTCGGC CATCTTTTGGGA-3' and 5'-TCCAACGAGTGGC GAACGCC-3'. These primers were designed on the basis of the DNA satellite sequence from the NCBI database (acc. no. X97847). The resulting 354-bp PCR product was subcloned in pGEM-T-Easy.

Genomic DNA and BAC-DNA isolation

Genomic DNA was extracted from young leaves using the CTAB method (Rogers and Bendich 1988). The cucumber BAC libraries (Nam et al. 2005) were screened by filter hybridization with a genomic fragment amplified from a long 3' exon of a 1-aminocyclopropane-1-carboxylic acid (ACC) synthase gene as a probe. Of the 125 positively identified clones, two BACs from the *Bam*HI library, B18F10 and B48F02, and one BAC from the *Eco*RI library, E36H12, were isolated, renamed as B17, B68, and E38, respectively, and used for pilot BAC FISH experiments. To this end, BAC-DNA was prepared with an alkaline lysis method (Sambrook et al. 1989) and purified by CsCl ultracentrifugation.

Table 1 Summary of repetitive DNA sequences used in this study

Name	Type	Length (bp)	Source
CsCent1	Satellite DNA	166–167	BAC-B68
CsCent2	Retrotransposon	2,107	BAC-B68
BAC-E38	Mitochondrial DNA	85,887	BAC-E38
CsPc	Dispersed repeat	429	BAC-B17
CmCent	Satellite DNA	354	Brennicke and Hemleben (1983)
45S rDNA	Satellite DNA	9,100	Gerlach and Bedbrook (1979)

Sequencing of BAC-E38

The sequence of the BAC-E38 was determined essentially as described by Jo et al. (2009). A shotgun sequencing library was constructed in the pUC118 vector for average insert sizes of 0.7–1.5 kb. BigDye Terminator chemistry v3.1 (ABI, Foster City, CA, USA) was used for the sequencing reactions. The sequences were analyzed using an ABI3730XL automatic DNA sequencer (ABI, Foster City, CA, USA). Individual sequences were assembled into contigs using the Sequencher 4.1.4 (Gene Codes Corp.). The gaps were filled by directed sequencing procedures using the BAC-DNA. The consensus sequence was generated by a final editing process that included manual visual confirmation of the original chromatograms using the Sequencher program. The sequences reported in this paper have been deposited in the GenBank with the accession numbers (FJ216936).

Chromosome preparation

Seeds were germinated on moistened filter paper in a Petri dish at 28°C for 48 h. Root tips roughly 1 cm long were collected, pretreated in saturated 1-bromonaphthalene at 20°C for 2 h, and then fixed in 1:3 (v/v) acetic acid/ethanol for 1 to 2 days. Fixed root tips were washed in distilled water for 10 min. The meristematic portions of the root tips were excised and incubated in a mixture of 2% cellulase (Onozuka, cat. no. 201064), 1.5% macerozyme (Yakult, cat. no. 202025), and 1% pectolyase (Sigma) in 1 mM EDTA at 37°C for 40 min, followed by washing in distilled water for 20 min. Root tips were then transferred to a clean glass slide and macerated in a drop of ethanol/acetic acid (3:1, v/v) using a fine-pointed forceps. After air-drying, the slides were stored at –20°C.

For pollen mother cell spread preparations, the method of Fransz et al. (1998) was used with some modifications. Immature flower buds were fixed in ethanol/acetic acid (3:1) for 2 h and stored at 4°C. The flower buds were rinsed in distilled water and incubated in an enzyme mix (0.3% pectolyase, 0.3% cytohelicase, and 0.3% cellulase) in citrate buffer (10 mM sodium citrate, pH4.5) for 2 h. Each bud was softened in 60% acetic acid on an uncoated and ethanol-cleaned microscopic slide kept at 45°C on a hot plate. The cells were spread over the slide, fixed with ice-cold ethanol/acetic acid (3:1), and air-dried.

Extended DNA fibers

Leaf nuclei were prepared as described by Jackson et al. (1998). A suspension of nuclei was deposited at one end of a poly-L-lysine-coated slide (Sigma) and left to air dry for 10 min. Eight-microliter STE lysis buffer was pipetted on top of the nuclei, and the slide was incubated at room temperature for 4 min. A clean coverslip was used to slowly drag the contents along the glass surface. The preparation was air-dried, fixed in 3:1 ethanol/glacial acetic acid for 2 min, and baked at 60°C for 30 min.

Fluorescence in situ hybridization

Repetitive DNA sequences of *Cucumis* species used as probes are listed in Table 1. Probes were labeled with digoxigenin-11-dUTP or biotin-16-dUTP by nick translation or PCR according to the manufacturer's protocols (Roche). The FISH procedure applied to both mitotic and meiotic chromosomes was essentially the same as reported in Koo et al. (2005). High-quality pachytene chromosome preparations were first selected under the phase-contrast microscope. After the first round of probing and image capture, slides were

soaked in a 4× SSC/0.2% Tween 20 to remove the coverslips. These were then air-dried, denatured in 70% formamide at 70°C for 2 min, dehydrated in cold ethanol series, and incubated with another probe for FISH analysis. Biotin- and digoxigenin-labeled probes were detected using a fluorescein-isothiocyanate-conjugated antibiotin antibody (Vector Laboratories) and a rhodamine-conjugated antidigoxigenin antibody (Roche), respectively. Chromosomes were counterstained with DAPI in Vectashield antifade (Vector Laboratories).

For extended DNA fiber preparations, the hybridization mix was covered with a 22×40 mm coverslip and sealed with rubber cement. The slide was placed in an 80°C oven in direct contact with a heated surface for 3 min, transferred to a wet chamber, which was prewarmed in an 80°C oven for 2 min, and transferred to 37°C for overnight incubation. Posthybridization washing stringency was the same as for FISH of the chromosome spreads. Signal detection was carried out according to Koo et al. (2005).

Image analysis

Images were captured with a CoolSNAP CCD camera (Photometrics) attached to a Leica epifluorescence microscope. Image processing and digital measurements were done with MetaMorph (Meta Imaging Series, v. 4.6) software. Global image adjustments for contrast, brightness, and color saturation were done with Adobe Photoshop v7.0.

Results and discussion

Satellite DNAs are a major component of *Cucumis* centromeres

The results from BAC library screening showed that the number of positively identified clones for an ACC synthase gene approximately corresponded with the predicted genome coverage of the two cucumber libraries (Nam et al. 2005). Subsequent fingerprinting experiments with these clones, however, revealed an extremely uneven clone distribution in the presumptive physical contigs. Instead of approximately ten clones per contig, the number of clones varied from two to 55 for the 12 contigs prepared (data not

shown). These results raised interesting questions about the putative presence of repetitive sequences and their roles in cucumber genome structure.

Pilot BAC FISH experiments carried out with some of the ACC synthase-positive BACs revealed two BACs with specific signals on the centromeres and pericentromeres of cucumber chromosomes. We isolated a 170-bp *Sau3AI* fragment from BAC-B68, referred to as CsCent1 (GenBank ac no. FJ377876-FJ377881), which produced major FISH signals on all centromeres (Fig. 1a) and minor foci in the intercalary region of the long arm of chromosome 1 (arrowheads) and the heterochromatic knob of the short arm of chromosome 2 (arrows). Sequence analysis revealed retrotransposon-like motifs which show only 70% homology with the cucumber Type III satellite (Ganal et al. 1986). A PCR amplification of the BAC end sequencing of BAC-B68 using the primer 5'-TAGAGGATCCATTAACCTA-3' gave a 2,107-bp fragment (GenBank ac no. FJ377881) with homology to the *Ty1-copia* retrotransposons of rice and *Arabidopsis* (Supplemental Figure 1). FISH analysis showed that the 2,107-bp DNA fragment hybridized only with *C. sativus* chromosomes. It was similar to CsCent1 and was designated as CsCent2 (Fig. 1b). We also found a 429-bp *Sau3AI* repeat from BAC-B17 that displayed pericentromere specificity (Fig. 1c). This repeat, which we refer to as CsPc (GenBank ac no. FJ377882), gave signals in all chromosomes except chromosome 3 (arrows). The CsCent1, CsCent2, and CsPc repeats could not be detected in *C. melo* (data not shown). On the other hand, FISH with a 354-bp DNA fragment (GenBank ac no. FJ10941-FJ10943) amplified from melon using genomic DNA as a probe (Brennicke and Hemleben 1983) were detected in the centromeric regions of melon chromosomes (Fig. 1e) and designated as CmCent. A sequence comparison between CsCent1 and CmCent showed that the sequences of CsCent1 from 1 to 166 bp had 51.80% identity to CmCent sequences from 82 to 260 bp (Supplemental Figure 2). The CsCent1 and CmCent satellite DNAs in this study shared many features with *Arabidopsis* pAL1, human alphoid DNA, and rice CentO satellite DNAs. The CsCent1 and CmCent families are both tandemly arranged in large arrays (Fig. 1f and g). As in the pAL1 family, the CsCent1 sequence is disrupted by other cucumber centromeric DNAs, including CsCent2 and mitochondria-related DNA.

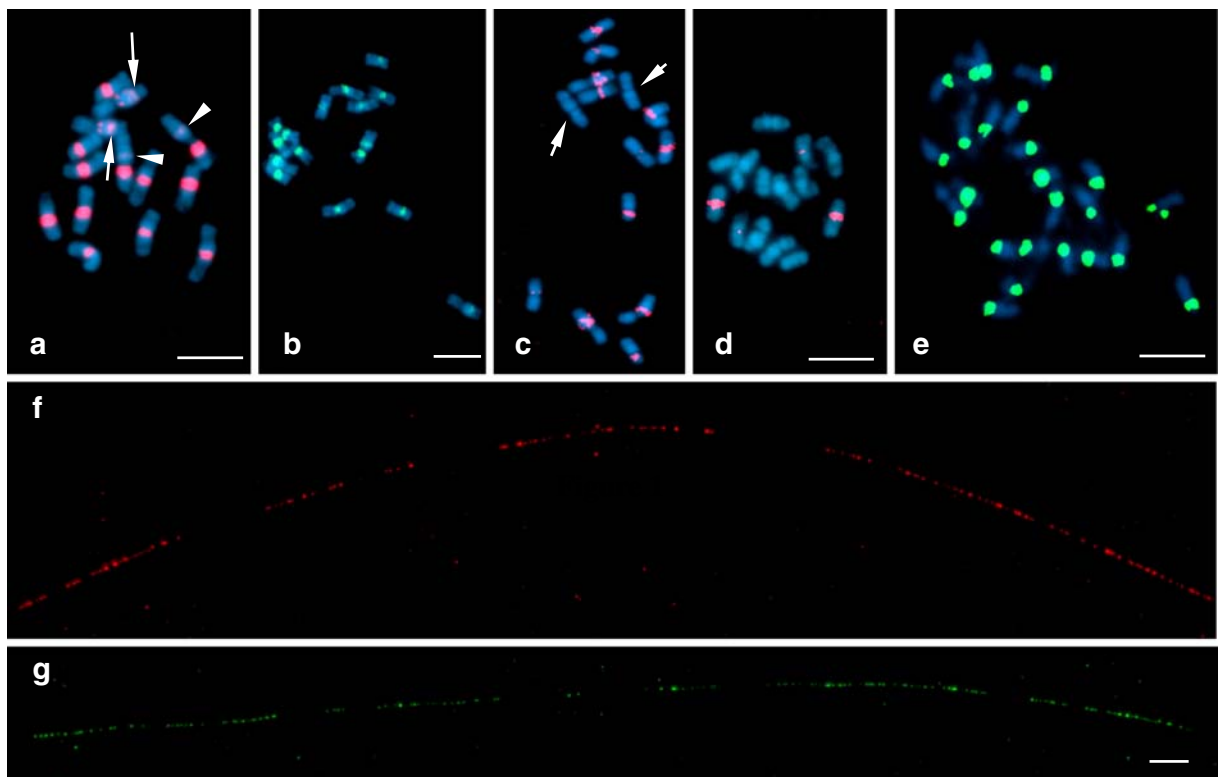


Fig. 1 FISH mapping of repetitive DNAs on cucumber and melon mitotic metaphase chromosomes. **a** CsCent1; **b** CsCent2; **c** CsPc; **d** BAC-E38; **e** CmCent. Fiber-FISH analysis using CsCent1 (**f**) and CmCent (**g**). Bars, 5 μ m

The possible origin of cucumber chromosomes 1 and 2

Pachytene complements in cucumber measure between 73.18 and 129.08 μ m and thus are 25 times longer than their mitotic metaphase counterparts (Koo et al. 2005). They display a distinct pattern of heterochromatin blocks at the distal ends, the pericentromeres, and a few intercalary knobs in the long-arm euchromatin (Koo et al. 2005, Fig. 2). Centromeres are recognized as short, weakly stained constrictions often seen in the middle of the long or short pericentromeric heterochromatin blocks. Interestingly, there is a short unstained gap in the short-arm pericentromere of chromosomes 1 and 2. FISH with the CsCent1 probe gave clear signals in the centromeres, the flanking parts of the pericentromeric blocks, and in the intercalary heterochromatic knobs of chromosomes 1 and 2 (Fig. 2a). In metaphase I complement, the homologous centromeres face opposite poles due to the pulling forces of microtubules. The CsCent1 FISH signals covered the most poleward-directed parts

of the chromosome; this pattern supports the observation that this repeat covers the centromeres (Fig. 2, a1). As the anti-CENH3 antibody is unavailable in cucumber, we cannot identify its functional centromere region.

Using high-resolution pachytene-FISH with 45S rDNA as a probe, we report unusual chromosome morphology in cucumber pachytene chromosomes. This is seen in chromosomes 1 and 2, which appear to have evolved by chromosomal fusion of the progenitor of cucumber (Fig. 3). FISH analysis supported this idea of chromosomal fusion by showing that 45S rDNA was interposed between the CsCent1 regions of chromosomes 1 and 2. The most intriguing finding is the existence of 45S rDNA sequences at both ends of the centromeric sequences, as revealed by the presence of CsCent1 in cucumber chromosome 2 (Fig. 3a). Here, we propose the possible model for evolution of chromosomes 1 and 2, in which two nonhomologous NOR-bearing acrocentric chromosomes (chromosomes that are not members of the same pair and have centromeres near their ends) break

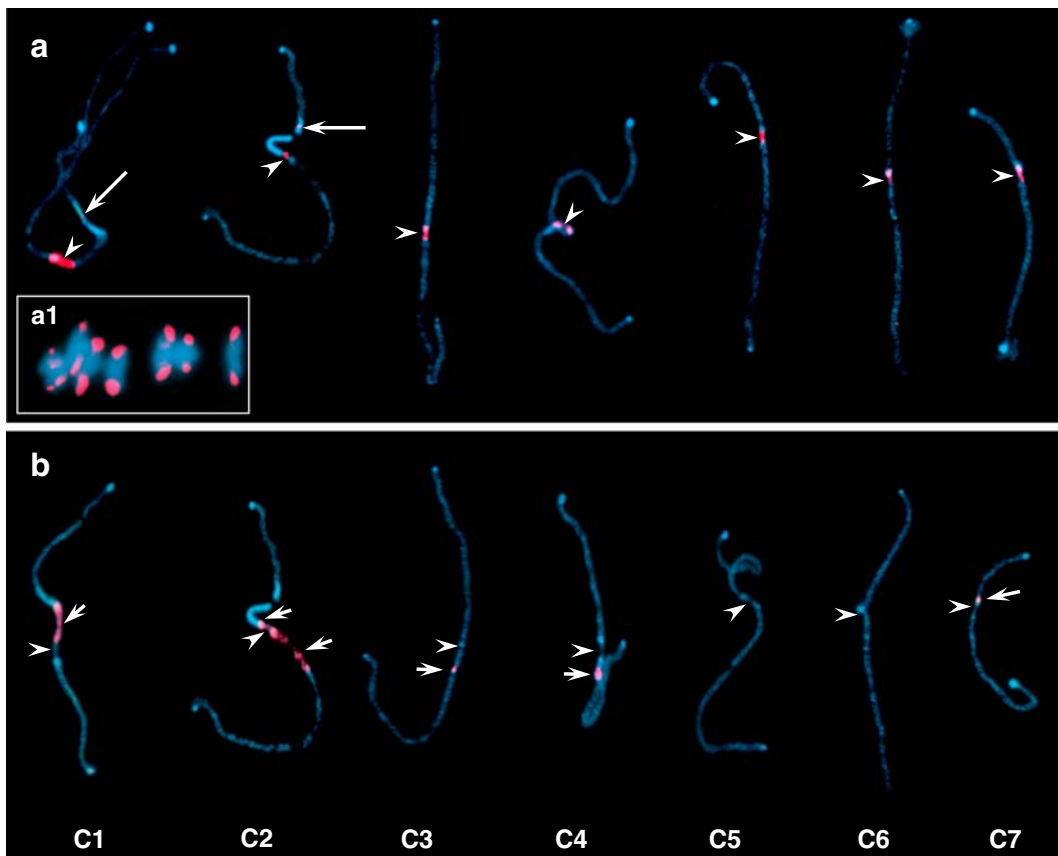


Fig. 2 Fine mapping of repetitive DNAs on cucumber pachytene chromosomes by FISH using CsCent1 (**a**) (arrowheads indicate centromere and arrows indicate minor FISH signal) and

45S rDNA (**b**) (arrows indicate 45S rDNA loci). Chromosomes were ordered according to Koo et al. (2005)

at their centromeres or NORs. The short arms of two chromosomes are lost and the long arms joined end to end at the centromeres or NOR regions (Fig. 3b). Blocks of satellite DNA or rDNA sequences are often assumed to constitute preferential breaking points, particularly in the rDNA clusters (Hall and Parker 1995). In humans, the most common structural rearrangements are Robertsonian translocations involving the NOR-bearing chromosomes (Stahl et al. 1983; Choo et al. 1988).

In the present findings, FISH using CsCent1 and 45S rDNA supports the possibility of chromosomal fusion of the progenitor of cucumber, which indicates that the genome of cucumber is highly likely to have originated from fusions of ancestral chromosomes after divergence from *C. melo* (Huang et al. 2009). This conclusion was further demonstrated by the sequence comparison between the two centromeric

satellite repeats (Supplemental Figure 2). Significant sequence similarity observed between these centromeric satellite repeats implies that both types of satellite sequences were present in a common progenitor. It may be that, after or during speciation, these sequences were specifically amplified, after which they diverged and specific sequence variants were homogenized (Fry and Salser 1977).

Trace mitochondrial DNA in the cucumber centromeric region is interspersed by repetitive DNA

A centromere-specific sequence was found in BAC-E38, but this repeat occurred only on chromosome 4 of the cucumber genome (Fig. 1d). The full sequencing of this BAC (8× coverage) gave 85,887 bp (GenBank ac no. FJ216936). Repeats were confirmed by a dot plot self-alignment (Supplemental Figure 3).

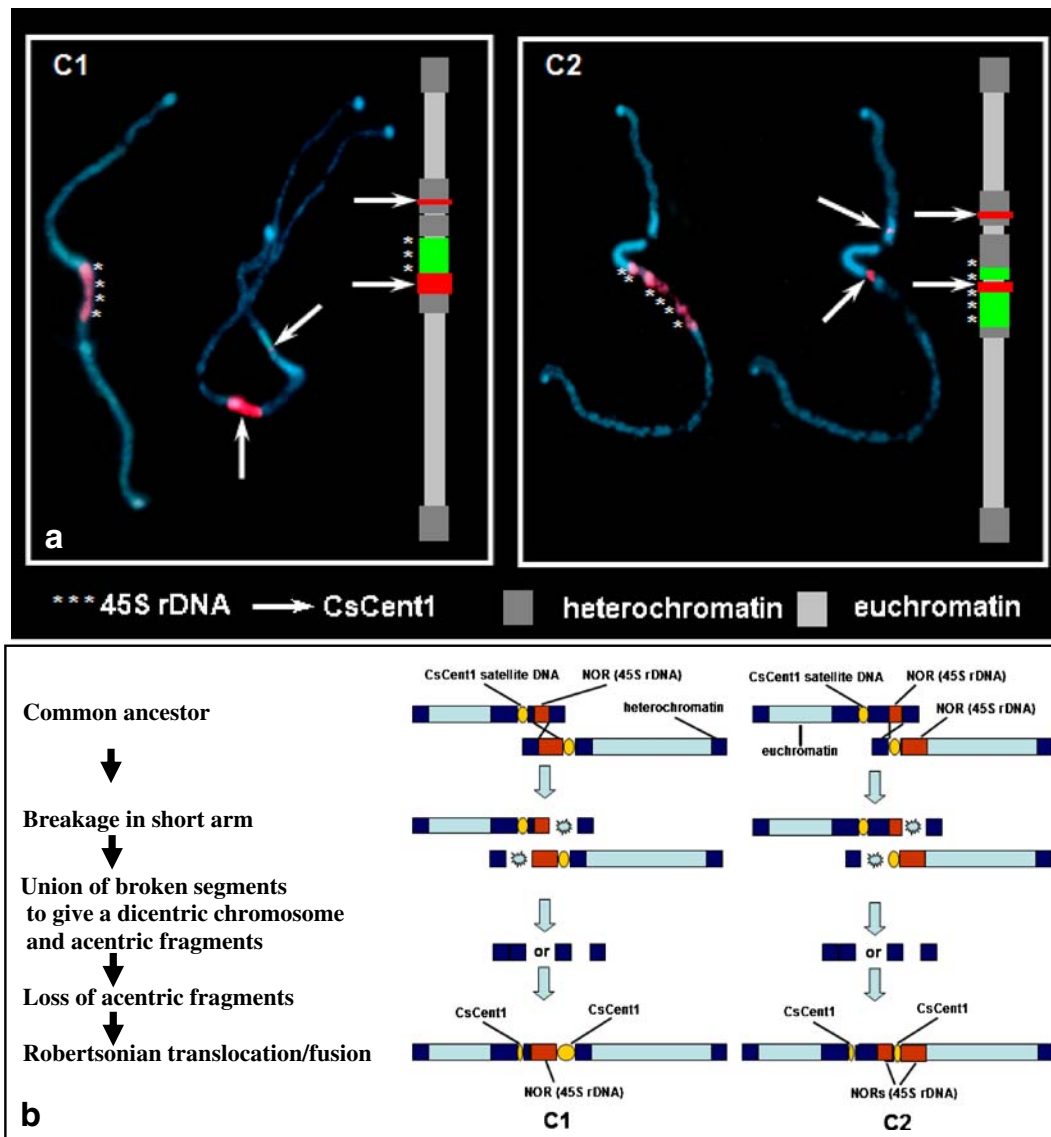


Fig. 3 An ideogram of the cucumber karyotype (**a**); locations of rDNA are indicated by *asterisks* and CsCent1 (centromeric satellite DNA) by *arrows*. The *dark gray regions* represent heterochromatin (DAPI-bright regions). A possible model for formation of Robertsonian translocations in cucumber chromosomes 1 and 2 (**b**); Robertsonian translocations are formed from

a break in the short arms of two ancestor acrocentric chromosomes each of which have a nucleolus organizer region (NOR). The resultant chromosome contains the long arms of each chromosome, both centromeres, and a fragment of each short arm. The remainder of both short arms is lost. C1 and C2 indicate cucumber chromosomes 1 and 2, respectively

This analysis revealed one repeat between 51 and 37,846 bp and a second between 48,089 and 85,887 bp. Both showed a 99.4% sequence identity (data not shown). RepeatMasker (<http://www.repeatmasker.org/>) revealed various AT-rich regions varying from 21 to 114 bp and consisting of two simple repeats; [TAAAA] n and [TA] n with n ranging from 25 to 33 (Supplemental Table 1). Using Miropeats

(Parsons 1995) and Tandem Repeat Finder software programs (Benson 1999), we detected five tandem and inverted repeats in the region between 51 and 37,846 bp (Supplemental Figure 4 and Supplemental Table 2). Three tandem and two inverted repeats in the region between 48,089 and 85,887 bp (Supplemental Table 3) were also detected. Further, dot plot analysis revealed high homology between 14 repeti-

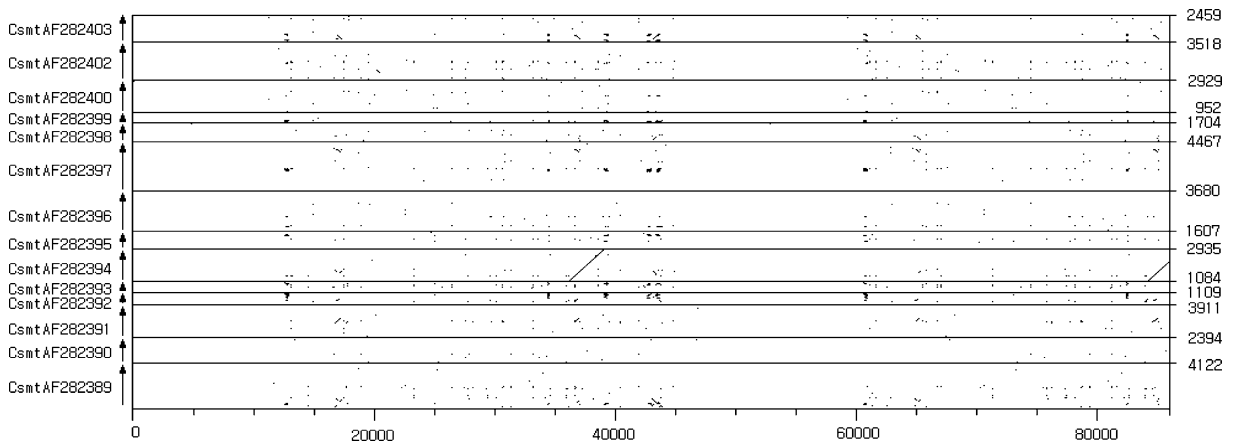
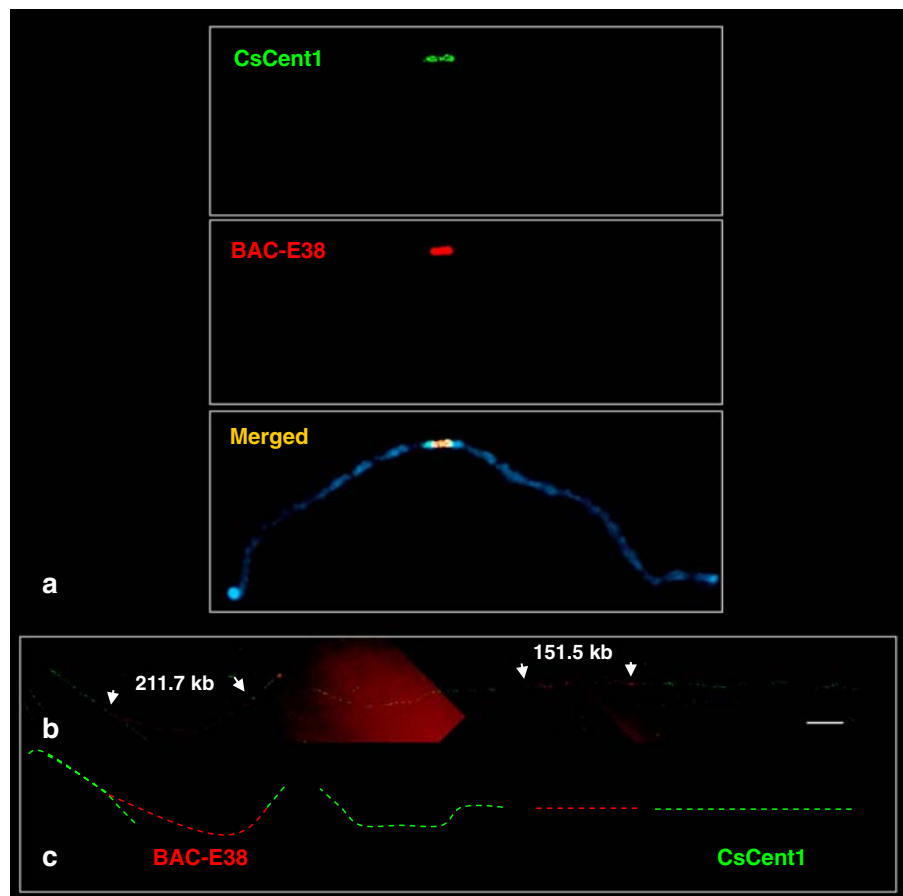


Fig. 4 Dot plot for cucumber repetitive mitochondrial sequences homologous to mitochondrial DNA sequences to BAC-E38 by Dotter program (window size=25). We used dotter program to compare 86-kb sequence from BAC clone to a set of ~1–4 kb sequenced mitochondrial clone. For 13 of the sequences,

we did not find stretches of conserved sequences over 50 bp. However, we did detect two small regions from CsmtAF282394 that are 100% similar to part of BAC-E38 sequence

Fig. 5 Pachytene- and fiber-FISH analysis using CsCent1 (green) and BAC-E38 (red). On highly extended pachytene chromosomes, signal for BAC-E38 was restricted to the centromeric region (a). Long BAC-E38 signals (211.7 and 151.5 kb) inserted into CsCent1 satellite DNA are detected within ~1.5 Mb on centromeric region of chromosome 4 (b). A diagram based on the fiber-FISH signal pattern in b (c). Three digital images of the fiber-FISH signals were collected and measured using FISH imaging system (Meta Morph). The lengths of the fiber-FISH signals (in micrometers) were converted to kilobases using a 2.87 kb/ μ m conversion rate (Jackson et al. 1998). Bar, 20 μ m



tive DNA mitochondrial clones (Lilly and Havey 2001) and the repeats in this BAC (Fig. 4). We also identified nine open reading frames, designated as G01 to G09, and one tRNA gene (Supplemental Table 4 and Supplemental Figure 5).

The distributions of CsCent1 and BAC-E38 sequences in the centromeres of cucumber chromosome 4 were analyzed by pachytene- and fiber-FISH (Fig. 5). FISH analysis of extended DNA fibers showed that BAC-E38 localized within the CsCent1 repeats. CsCent1 could be observed as a larger stretch of green signals on either side of the BAC-E38 sequence. The BAC integration sites were 211.7 ± 41 and 151.5 ± 36 kb (Fig. 5b).

One feature of organellar genomes is that they easily transfer to the nuclear genome. Many organellar genes have been transferred to the nucleus, especially those of mitochondrial origin (Blanchard and Lynch 2000). Among plant species, members of the Cucurbitaceae possess the largest known mitochondrial genome. The sizes of the cucumber and melon mitochondrial genomes are 1,500 and 2,400 kb, respectively (Ward et al. 1981). Lilly and Havey (2001) reported that short repetitive DNA motifs significantly contribute to the enlargement of cucumber mitochondrial DNAs. A dot plot comparison of the 14 mitochondrial repetitive DNA clones with BAC-E38 confirmed that the latter contained mitochondrial DNA (Fig. 4).

The 86-kb BAC clone containing the cucumber mitochondrial genome perhaps represents a nuclear insertion that occurred during the course of evolution. In *Arabidopsis thaliana*, an insertion of 270-kb mtDNA was found in the pericentromeric region on the short arm of chromosome 2 (Lin et al. 1999). DNA fiber-based FISH analyses revealed that the mtDNA insert is 620 kb, 2.3 times greater than that determined by contig assembly and sequencing analysis (Stupar et al. 2001). This may be due to duplication of mitochondrial sequences. The use of BAC-E38 and CsCent1 as probes in FISH analysis on extended DNA fibers showed the former sequences to be integrated into the latter (Fig. 5). Moreover, the signals do not consist of the entire sequence of the 86-kb BAC-E38 FISH signal; instead, its integration in CsCent1 involved three or four copies (Fig. 5b). However, we do not provide any insight at this time as to whether the multiplication occurred in the mitochondrial genome before nuclear integration or in the nuclear genome after the integration.

Acknowledgements We are grateful to Dr. Jiming Jiang for his valuable comments on this manuscript. This work was supported by a grant from the Korea Research Foundation (KRF-2000-DP0401) to YH.

References

- Benson G (1999) Tandem repeats finder: a program to analyze DNA sequences. *Nucl Acids Res* 27:573–580
- Blanchard JL, Lynch M (2000) Why do organellar genes end up in the nucleus? *Trends Genet* 16:315–320
- Bradley RD, Wichman HA (1994) Rapidly evolving repetitive DNAs in a conservative genome: a test of factors that affect chromosomal evolution. *Chromosome Res* 2:354–360
- Brennicke A, Hemleben V (1983) Sequence analysis of the cloned *Cucumis melo* highly repetitive satellite DNA. *Z Naturforsch* 38:1062–1065
- Cerbah M, Coulaud J, Siljak-Yakovlev S (1998) rDNA organization and evolutionary relationships in the genus *Hypochoeris* (Asteraceae). *J Hered* 89:312–318
- Cheng Z, Robin-Buell C, Wing RA, Gu M, Jiang JM (2001) Toward a cytological characterization of the rice genome. *Genome Res* 11:2133–2141
- Cheng Z, Dong F, Langdon T, Ouyang S, Robin-Buell C, Gu M, Blattner FR, Jiang JM (2002) Functional rice centromeres are marked by a satellite repeat and a centromere-specific retrotransposon. *Plant Cell* 14:1691–1704
- Choo KH (2001) Domain organization at the centromere and neocentromere. *Dev Cell* 1:165–177
- Choo KH, Vissel B, Brown R, Filby RG, Earle E (1988) Homologous alpha satellite sequences on human acrocentric chromosomes with selectivity for chromosomes 13, 14, and 21: implications for recombination between non homologs and Robertsonian translocations. *Nucl Acids Res* 16:1273–1284
- de Jong JH, Fransz P, Zabel P (1999) High resolution FISH in plants—techniques and applications. *Trends Plant Sci* 4:258–263
- Fransz PF, Alonso-Blanco C, Liharska TB, Peeters AJM, Zabel P, de Jong JH (1996) High resolution physical mapping in *Arabidopsis thaliana* and tomato by fluorescence *in situ* hybridization to extended DNA fibers. *Plant J* 9:421–430
- Fransz PF, Armstrong S, Alonso-Blanco C, Fischer TC, Torres-Ruiz RA, Jones GH (1998) Cytogenetics for the model system *Arabidopsis thaliana*. *Plant J* 13:867–876
- Fransz PF, Armstrong S, de Jong JH, Parnell LD, van Drunen C, Dean C, Zabel P, Bisseling T, Jones GH (2000) Integrated cytogenetic map of chromosome arm 4 S of *A. thaliana*: structural organization of heterochromatic knob and centromere region. *Cell* 100:367–376
- Fry K, Salser W (1977) Nucleotide sequences of HS-alpha satellite DNA from kangaroo rat *Dipodomys ordii* and characterization of similar sequences in other rodents. *Cell* 12:1069–1084
- Ganal M, Hemleben V (1988) Insertion and amplification of a DNA sequence in satellite DNA of *Cucumis sativus* L. (Cucumber). *Theor Appl Genet* 75:357–361

- Ganal M, Riede I, Hemleben V (1986) Organization and sequence analysis of two related satellite DNAs in cucumber (*Cucumis sativus* L.). *J Mol Evol* 23:23–30
- Garagna S, Zuccotti M, Redi CA, Capanna E (1997) Trapping speciation. *Nature* 390:241–242
- Garagna S, Marziliano N, Zuccotti M, Searle JB, Capanna E, Redi CA (2001) Pericentromeric organization at the fusion point of mouse Robertsonian translocation chromosomes. *Proc Natl Acad Sci U S A* 98:171–175
- Gerlach WL, Bedbrook JR (1979) Cloning and characterization of ribosomal RNA genes from wheat and barley. *Nucleic Acids Res* 7:1869–1885
- Hall KJ, Parker JS (1995) Stable chromosome fission associated with rDNA mobility. *Chromosome Res* 3:417–422
- Han YH, Zhang Z, Liu JH, Lu JY, Huang SW, Jin WW (2008) Distribution of the tandem repeat sequences and karyotyping in cucumber (*Cucumis sativus* L.) by fluorescence *in situ* hybridization. *Cytogenet Genome Res* 122:90–98
- Han YH, Zhang Z, Liu C, Huang S, Jiang JM, Jin WW (2009) Centromere repositioning in cucurbit species: Implication of the genomic impact from centromere activation and inactivation. *Proc Natl Acad Sci U S A* 106:14937–14941
- Henikoff S, Ahmad K, Malik HS (2001) The centromere paradox: stable inheritance with rapidly evolving DNA. *Science* 293:1098–1102
- Hirai H, Yamamoto MT, Ogura K, Satta Y, Yamada M, Taylor RW, Imai HT (1995) Multiplication of 28S rDNA and NOR activity in chromosome evolution among ants of the *Myrmecia pilosula* species complex. *Chromosoma* 103:171–178
- Hirai H, Yamamoto MT, Taylor RW, Imai HT (1996) Genomic dispersion of 28S rDNA during karyotypic evolution in the ant genus *Myrmecia* (Formicidae). *Chromosoma* 105:190–196
- Huang S, Li R, Zhang Z et al (2009) The genome of the cucumber, *Cucumis sativus* L. *Nat Genet* 41:1275–1281
- Jackson SA, Wang MI, Goodman HM, Jiang JM (1998) Application of fiber-FISH in physical mapping of *Arabidopsis thaliana*. *Genome* 41:566–572
- Jiang JM, Gill BS (1994) New 18S, 26S ribosomal RNA gene loci: chromosomal landmarks for the evolution of polyploid wheats. *Chromosoma* 103:179–185
- Jiang JM, Gill BS (2006) Current status and the future of fluorescence *in situ* hybridization (FISH) in plant genome research. *Genome* 49:1057–1068
- Jiang JM, Birchler J, Parrott W, Dawe R (2003) A molecular view of plant centromeres. *Trends Plant Sci* 8:570–575
- Jo SH, Koo DH, Kim JF, Hur CG, Lee S, Yang TJ, Kwon SY, Choi D (2009) Evolution of ribosomal DNA-derived satellite repeat in tomato genome. *BMC Plant Bio* 9:42
- Kato A, Lamb JC, Birchler JA (2004) Chromosome painting using repetitive DNA sequences as probes for somatic chromosome identification in maize. *Proc Natl Acad Sci U S A* 101:13554–13559
- Koo DH, Jiang JM (2009) Super-stretched pachytene chromosomes for fluorescence *in situ* hybridization mapping and immunodetection of DNA methylation. *Plant J* 59:509–516
- Koo DH, Choi HW, Cho J, Hur Y, Bang JW (2005) A high-resolution karyotype of cucumber (*Cucumis sativus* L. 'Winter Long') revealed by C-banding, pachytene analysis, and RAPD-aided fluorescence *in situ* hybridization. *Genome* 48:534–540
- Li YC, Lee C, Sanoudou D, Hsu TH, Li SY, Lin CC (2000) Interstitial colocalization of two cervid satellite DNAs involved in the genesis of the Indian muntjac karyotype. *Chromosome Res* 8:363–373
- Lilly JW, Havey MJ (2001) Small, repetitive DNAs contribute significantly to the expanded mitochondrial genome of cucumber. *Genetics* 159:317–328
- Lin X, Kaul S, Rounsley S, Shea TP et al (1999) Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*. *Nature* 402:761–768
- Maluszynska J, Heslop-Harrison JS (1991) Localization of tandemly repeated DNA sequences in *Arabidopsis thaliana*. *Plant J* 1:159–166
- Nam YW, Lee JR, Song KH, Lee MK, Robbins MD, Chung SM, Staub JE, Zhang HB (2005) Construction of two BAC libraries from cucumber (*Cucumis sativus* L.) and identification of clones linked to yield component quantitative trait loci. *Theor Appl Genet* 111:150–161
- Page SL, Shin JC, Han JY, Choo KHA, Shaffer LG (1996) Breakpoint diversity illustrates distinct mechanisms for Robertsonian translocation formation. *Hum Mol Genet* 5:1279–1288
- Parsons JD (1995) Miropoints: graphical DNA sequence comparisons. *Comput Appl Biosci* 11:615–619
- Peterson DG, Lapitan NLV, Stack SM (1999) Localization of single- and low-copy sequences on tomato synaptonemal complex spreads using fluorescence *in situ* hybridization (FISH). *Genetics* 152:427–439
- Pinkel D, Straume T, Gray JW (1986) Cytogenetic analysis using quantitative, high-sensitivity, fluorescence hybridization. *Proc Natl Acad Sci U S A* 83:2934–2938
- Plohl M, Luchetti A, Mestrovic 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
- Rogers SO, Bendich AJ (1988) Extraction of milligram amounts of DNA from fresh, herbarium and mummified plant tissue. *Plant Mol Biol Rep* 5:69–76
- Sambrook J, Fritsch EF, Maniatis T (1989) *Molecular cloning: a laboratory manual*, 2nd edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor
- Schaefer H, Heibl C, Renner SS (2009) Gourds afloat: A dated phylogeny reveals an Asian origin of the gourd family (Cucurbitaceae) and numerous overseas dispersal events. *Proc R Soc B* 276:843–851
- Shan F, Yan G, Plummer A (2003) Cyto-evolution of *Boronia* genomes revealed by fluorescence *in situ* hybridization with rDNA probes. *Genome* 46:507–513
- Stahl A, Luciani JM, Hartung M, Devictor M, Berge Lefranc JL, Guichaoua M (1983) Structural basis for Robertsonian translocations in man: association of ribosomal genes in the nucleolar fibrillar center in meiotic spermatocytes and oocytes. *Proc Natl Acad Sci U S A* 80:5946–5950
- Stupar RM, Lilly JW, Town CD, Cheng Z, Kaul S, Buell CR, Jiang JM (2001) Complex mtDNA constitutes and approximate 620 kb insertion on *Arabidopsis thaliana* chromosome 2: implication of potential sequencing errors

- caused by large-unit repeats. *Proc Natl Acad Sci U S A* 98:5099–5103
- Sullivan BA, Jenkins LS, Karson EM, Leana-Cox J, Schwartz S (1996) Evidence for structural heterogeneity from molecular cytogenetic analysis of dicentric Robertsonian translocations. *Am J Hum Genet* 59:167–175
- Talbert PB, Masuelli R, Tyagi AP, Comai L, Henikoff S (2002) Centromeric localization and adaptive evolution of an *Arabidopsis* histone H3 variant. *Plant Cell* 14:1053–1066
- Ugarkovic D, Plohl M (2002) Variation in satellite DNA profiles—causes and effects. *EMBO J* 21:5955–5959
- Ward BL, Anderson RS, Bendich AJ (1981) The mitochondrial genome is large and variable in a family of plants (Cucurbitaceae). *Cell* 25:793–803
- Wevrick R, Willard HF (1989) Long-range organization of tandem arrays of a satellite DNA at the centromeres of human chromosomes: high frequency array-length polymorphism and meiotic stability. *Proc Natl Acad Sci U S A* 86:9394–9398
- Wichman HA, Payne CT, Ryder OA, Hamilton MJ, Maltbie M, Baker RJ (1991) Genomic distribution of heterochromatic sequences in equids: implications to rapid chromosomal evolution. *J Hered* 82:369–377