Survey of extrachromosomal circular DNA derived from plant satellite repeats

Alice Navrátilová, Andrea Koblížková and Jiří Macas*

Address: Biology Centre ASCR, Institute of Plant Molecular Biology, Branišovská 31, Ústí nad Labem, Czech Republic

Email: Alice Navrátilová - navratil@umbr.cas.cz; Andrea Koblížková - andrea@umbr.cas.cz; Jiří Macas* - macas@umbr.cas.cz

* Corresponding author

Abstract

Background: Satellite repeats represent one of the most dynamic components of higher plant genomes, undergoing rapid evolutionary changes of their nucleotide sequences and abundance in a genome. However, the exact molecular mechanisms driving these changes and their eventual regulation are mostly unknown. It has been proposed that amplification and homogenization of satellite DNA could be facilitated by extrachromosomal circular DNA (eccDNA) molecules originated by recombination-based excision from satellite repeat arrays. While the models including eccDNA are attractive for their potential to explain rapid turnover of satellite DNA, the existence of satellite repeat-derived eccDNA has not yet been systematically studied in a wider range of plant genomes.

Results: We performed a survey of eccDNA corresponding to nine different families and three subfamilies of satellite repeats in ten species from various genera of higher plants (Arabidopsis, Oryza, Pisum, Secale, Triticum and Vicia). The repeats selected for this study differed in their monomer length, abundance, and chromosomal localization in individual species. Using two-dimensional agarose gel electrophoresis followed by Southern blotting, eccDNA molecules corresponding to all examined satellites were detected. EccDNA occurred in the form of nicked circles ranging from hundreds to over eight thousand nucleotides in size. Within this range the circular molecules occurred preferentially in discrete size intervals corresponding to multiples of monomer or higher-order repeat lengths.

Conclusion: This work demonstrated that satellite repeat-derived eccDNA is common in plant genomes and thus it can be seriously considered as a potential intermediate in processes driving satellite repeat evolution. The observed size distribution of circular molecules suggests that they are most likely generated by molecular mechanisms based on homologous recombination requiring long stretches of sequence similarity.

Background

Higher plant genomes contain considerable amounts of satellite repeats which make up to 20% of nuclear DNA in some species [1]. Satellite repeats occur in a genome as continuous arrays of tandemly arranged basic repeated units (monomers). Although the monomers are usually only tens to hundreds of nucleotides long, they can accumulate into millions of copies, forming megabase-sized clusters distinguishable as heterochromatic regions on mitotic chromosomes or in interphase nuclei. Satellite...
repeats undergo rapid evolutionary changes of their sequences and abundance, leading to the frequent occurrence of genus- or species-specific families of satellite DNA [2-5]. Contrary to this diversification observed between various taxa, the repeat monomers are usually well-homogenized within a species. This process of intra-specific sequence homogenization, generally referred to as ‘concerted evolution’, is supposed to arise from a concurrent action of various molecular mechanisms including unequal crossing-over and gene conversion [6,7]. Although some of these mechanisms have been characterized in detail, their overall contribution to satDNA evolution remains elusive. Moreover, theoretical models and computer simulations suggest that these mechanisms alone cannot account for efficient amplification and long-term persistence of satellites within genomes [8-10]. Therefore, it is supposed that other processes capable of efficient sequence amplification probably act on satDNA. It has been proposed that they involve extrachromosomal circular DNA (eccDNA) molecules, arising from intra-strand recombination between monomers within satellite arrays and subsequently serving as a template for rolling-circle replication. This process would result in the synthesis of linear DNA fragments composed of multiple copies of the circular template molecules and their reintegration into the genome, thus providing an efficient mechanism for amplification and eventual sequence homogenization of satDNA.

Although eccDNA has been reported from a wide range of eukaryotic organisms including yeast, Drosophila, Xenopus, mouse and human [11-14], there are only a few studies focusing on its formation from satellite repeats [13-15]. In plants, eccDNA has been derived from centromeric repeats in Arabidopsis [16] and repetitive element Bdm29 in Brachycome dichromosomatica [17] have been detected. In spite of this progress in eccDNA research, its formation from a wider range of plant satellite repeats and different species has not been studied so far. Consequently, there is only a little known about the structure of circular DNA molecules in plant genomes and mechanisms of their formation. There is also an interesting question concerning the role of eccDNA in the evolution of monomer size of satellite repeats. Similar to other groups of eukaryotes, plant satellites show a clear preference for monomer sizes in ranges between 135 – 195 bp and their multiples [18]. Although the correspondence of this length with the length of DNA wrapped around nucleosome particles has been pointed out [19,20], there is no mechanism known to explain this phenomenon. It has been demonstrated that nucleosomes constrain accessibility of enzymatic apparatus to certain regions of associated DNA [21]. Thus, recombination-based sequence homogenization or excision of eccDNA may be more frequent in more accessible regions (e.g. nucleosome linkers), leading to the emergence of the nucleosome-sized repeated units.

In this study, we addressed some of the questions raised above by investigating the occurrence and properties of eccDNA molecules derived from satellite repeats in a range of species from three genera of higher plants (Fabaceae, Poaceae, Brassicaceae). The repeats to be studied were selected based on their various monomer lengths and eventual presence of higher-order repeats, in order to follow the importance of these properties for formation and size of eccDNA molecules. Our results demonstrated that formation of eccDNA from satellite repeats is a common phenomenon in higher plants, and that it is strongly dependent on sequence similarity.

**Results**

We employed two-dimensional (2-D) agarose gel electrophoresis [22] followed by Southern blot hybridization to examine the presence of eccDNA in extracts of total genomic DNA from selected plant species. In addition to separating DNA molecules based on their size, 2-D electrophoresis also allows resolution based on their structure, resulting in formation of separated arcs on the gel representing linear and various conformations of circular DNA molecules (Fig. 1A). In our initial experiments we performed 2-D electrophoresis in 0.4% agarose in the first dimension, and in 1% agarose supplemented with ethidium bromide in the second dimension, which provided efficient resolution of linear and circular molecules from over 9 kb down to 3 kb as assessed from migration of linear and supercoiled or open circular marker molecules (not shown). These conditions were then used to analyze DNA samples of Vicia faba, a species containing highly abundant satellite repeat FokI [23,24] (Table 1). A strong arc of linear DNA and a weaker one corresponding to open circular molecules were revealed after hybridization with FokI probe. We used this species further for investigating FokI eccDNA levels in different tissues including young/developing or old leaves and root meristems (Fig. 1B–D), and also in leaf tissues stressed by wounding (Fig. 1E–F). However, we did not observe significant differences in eccDNA signals between these samples, suggesting that there are comparable levels of circular molecules derived from FokI satellite in all investigated V. faba tissues and that these levels are not significantly affected by stress conditions. As only open circular molecules were detected and there was no hybridization signal corresponding to covalently closed circular DNA, we tested whether this could be a result of damaging closed circles during DNA isolation. However, including supercoiled control plasmid DNA into tissue samples at the beginning of the procedure revealed that the isolation does not lead to significant nicking or other degradation of this DNA,
Detection of eccDNA molecules derived from Vicia faba FokI repeats in different tissues and under stress conditions

Figure 1
Detection of eccDNA molecules derived from Vicia faba FokI repeats in different tissues and under stress conditions. (A) Schematic outline of the migration patterns of linear and circular DNA forms on 2-D gel electrophoresis. (B-F) Detection of eccDNA in genomic DNA samples of V. faba separated on 2-D electrophoresis (0.4% and 1% agarose), Southern blotted and hybridized with FokI-derived probe. The DNA was isolated from root meristems (B), young (C) or old leaves (D), and from wounded leaves one (E) and two (F) days after the treatment.
and thus that open circles are the predominant form of satellite eccDNA in the examined tissues.

Following these initial experiments, we performed a detailed survey of the presence of eccDNA derived from nine satellite repeats and three subfamilies in ten plant species (Table 1). However, as eccDNA was found to occur at relatively low levels, we improved our isolation protocol by including treatment with Plasmid-safe ATP-dependent DNase, which specifically degrades linear DNA while leaving covalently closed and nicked circular molecules intact. Removing the majority of linear frag-

| Satellite repeat family | Species                  | Monomer length (bp) | Location(a) | Copies/haploid genome (1C) | References | EccDNA size | Figure |
|-------------------------|--------------------------|---------------------|-------------|-----------------------------|------------|-------------|--------|
| Afa                     | Triticum aestivum        | 340                 | P, I, T     | 30,000                      | [49-51]    | (340)n      | 3C     |
| AT-180                  | Arabidopsis thaliana     | 179                 | C, P        | > 55,000                    | [52,53]    | (179)n      |        |
| IGS-like                | Vicia sativa             | 173                 | I           | 10,000 – 100,000            | [47]       | (173)n      | 3D     |
| CentO                   | Oryza sativa             | 155                 | C, P        | 44,400                      | [54,55]    | (155)n      | 3E     |
| ScI19                   | Secale cereale           | 118                 | I, T        | 1,500,000                   | [56]       | (118)n      |        |
| VicTR-A                 | Vicia narbonensis        | 69                  | I, T        | 100,000 – 1,000,000         | [3,57]     | (138)n      |        |
| VicTR-A_c666            | Vicia pannonica          | 69/138              | n.d.(b)     | n.d.(b)                     | this publication | continuous arc | 2L     |
| VicTR-A_c653            | Vicia pannonica          | 180                 | n.d.(b)     | n.d.(b)                     | this publication | (180)n | 3K     |
| FokI                    | Vicia faba               | 59                  | I           | 5,400,000 – 21,000,000      | [24]       | continuous arc | 3F     |
| PisTR-B                 | Pisum sativum            | 50                  | P, T        | 380,000                     | [58,59]    | continuous arc | 3G     |
| VicTR-B                 | Vicia grandiflora        | 38                  | I, T        | 1,000,000 – 5,000,000       | [3]        | continuous arc | 3H     |
| VicTR-B_VG-V            | Vicia grandiflora        | 186                 | T           | n.d.                        | [26]       | (186)n      |        |
| VicTR-B                 | Vicia sativa             | 38                  | I           | 1,000,000 – 5,000,000       | [3]        | continuous arc |        |

(a) Positions of signals on chromosomes. C, centromeric; P, pericentromeric; I, intercalary; T, (sub-)telomeric.
(b) VicTR-A family in V. pannonica is located in (sub-)telomeric chromosome regions [3], but chromosome localization and copy numbers of its specific sequence variants represented by the clones c666 and c653 were not determined.

Figure 2
Improved detection and size determination of VicTR-A repeats in Vicia narbonensis. EccDNA detection in V. narbonensis genomic DNA sample pre-treated with Plasmid-safe DNase and mung bean nuclease and resolved on 2-D electrophoresis using 0.7% and 2% agarose gels. Prior to electrophoresis, the sample was mixed with a set of circular DNA markers. The Southern blot of the gel was hybridized with VicTR-A probe (A), and then reprobed with the lambda DNA used to visualize DNA markers (B). Panel C shows superposition of signals from VicTR-A nad marker probes and gives the length of the open circle (OC), supercoiled (SC) and linear (L) markers (the lengths are in base pairs). VicTR-A signals on panel C are enhanced by longer exposure compared to panel A in order to reveal spots of the shortest circular molecules.
ments by this treatment allowed loading of equivalents of up to 240 μg of original undigested genomic DNA onto the gel. We also increased the discriminatory power of the technique towards efficient resolution of short circular molecules in order to exactly determine size differences of circular molecules derived from satellites with shorter monomers. This was achieved by increasing agarose concentrations to 0.7% and 2% in the first and second dimension, respectively, which led to improved resolution of circular DNA from linear fragments (down to 500 bp) and to resolving circular molecules differing by at least 90 bp.

Size estimations of eccDNA molecules were done using a mixture of specifically designed markers including open circular molecules produced by Cre/lox recombination and nickase treatment of larger plasmid templates (see Methods for details). This is demonstrated in Figure 2, showing detection of VicTR-A satellite in *V. narbonensis*. Although the size of VicTR-A monomers is 69 bp, its eccDNA was found to occur in size steps of about 140 bp, thus corresponding to multiples of 138 bp. Moreover, there was an alternating pattern of stronger and weaker spots, suggesting more frequent excision of circular molecules differing in multiples of 276 bp.

Extrachromosomal circular molecules were detected for all nine satellite families and three subfamilies tested, and in all cases they occurred as open circles (Fig. 3). The size of the eccDNA molecules ranged from over 8 kb down to 500 bp and in most cases their hybridization signals formed discrete spots corresponding to multiples of monomer or higher-order unit lengths (Table 1). However, there was often a continuous smear underlying these spots, as shown on Fig. 3A for Sc119 satellite of *S. cereale*. Treating the samples with mung bean nuclease, which specifically degrades single-stranded but not double-stranded or nicked DNA almost completely removed the smear, resulting in more round and discrete spots (Fig. 3B), thus suggesting that the smear was composed of partially single-stranded circular molecules. In addition to the arcs representing circular and linear DNA molecules, in some cases we encountered hybridization signals forming a relatively fuzzy arc that could not be assigned to linear, covalently closed supercoiled or open circular molecules by co-migration with the respective markers. The signal was reproducibly detected for satellites from *Secale cereale*, *Vicia sativa* and *Oryza sativa* (Fig. 3B, D, E). As it was present even in mung bean nuclease-treated samples, its composition from single-stranded DNA or RNA could be ruled out. Thus, we tested whether it could be affected by treatment with RNaseH or RNaseH followed by mung bean nuclease. However, these treatments had no effect, implying that the extra arc did not include hybrid DNA:RNA molecules (data not shown).

The VicTR-A satellite of *V. narbonensis* was found to produce eccDNA in size steps corresponding to dimers (138 bp) and more frequently tetramers (276 bp) of the basic repeated unit of 69 bp (Fig. 2). In the related species *V. pannonica*, initial investigation of the VicTR-A repeats, using probe amplified in a PCR reaction with genomic DNA as a template, revealed a different hybridization pattern for their eccDNA, consisting of the smear with spots spaced by about 180 bp which could not be derived from multiples of the monomer length (Fig. 3I). As there was only limited information available about sequence variability of VicTR-A satellites in *V. narbonensis* and *V. pannonica*, we constructed whole genome shotgun libraries from these species and screened them for VicTR-A clones which were subsequently sequenced and analyzed. The analysis of 29 *V. narbonensis* VicTR-A clones (19,006 bp in total) using nucleotide autocorrelation functions [26,27] revealed that their sequences are well-homogenized with a basic periodicity of 69 bp. However, some preference for homogenization of the tetramer-based units was also revealed by the increased height of the 276 bp peak (Fig. 4A). This is consistent with the observed preferential formation of the tetramer-based eccDNA, although it is not clear why similar correlation was not observed for the dimer periodicity, which was detected in eccDNA sizes but not by the sequence analysis.
Figure 3
Survey of eccDNA derived from various plant satellites. Samples of genomic DNA enriched for circular DNA were separated on 0.7% and 2% 2-D gels, blotted and hybridized with satellite-specific probes. The arcs corresponding to linear (L) and open circle (OC) DNA are indicated with black arrowheads and positions of the longest (1,674 bp) and the shortest (558 bp) OC markers are shown as blue spots. The extra arc of unknown origin is indicated with black arrow on panels A, B, D and E. (A-B) Detection of Sc119 repeats in Secale cereale, demonstrating the effect of single-stranded DNA degradation by mung bean nuclease (B) compared to untreated control (A). Figure insets contain magnified regions marked by the rectangles showing the effect of the nuclease treatment on shape of the spots. All other samples shown on panels C-L were also treated with mung bean nuclease. (C) Afa repeats in Triticum aestivum, (D) IGS-like in Vicia sativa, (E) CentO in Oryza sativa, (F) FokI in V. faba, (G) PisTR-B in Pisum sativum (this sample was resolved on 1.5% and 2% 2-D gels), (H) VicTR-B in V. grandiflora, (I) VG-V subfamily of VicTR-B in V. grandiflora, (J) VicTR-A in V. pannonica detected using PCR probe including a mixture of genomic VicTR-A sequences, (K-L) VicTR-A in V. pannonica detected using specific clones c653 (K) and c666 (L).
In contrast to *V. narbonensis*, analysis of 26 *V. pannonica* clones (16,234 bp in total) showed that they are considerably less homogenized, as revealed by much lower peak heights (Fig. 4B). Their basic periodicity was also 69 bp, however, there was one *V. pannonica* clone (c653) with periodicity of 180 bp and high sequence similarity (>97%) between the 180 bp repeated units. Comparative analysis with other VicTR-A sequences (Fig. 5A,B) revealed that this repeat most probably originated by recombination between AT-rich regions resembling the CAAAA motif, which is supposed to be involved in breakage-reunion of repeated sequences [28,29]. Using this clone as a probe confirmed that this VicTR-A subfamily gives rise to the 180 bp-spaced signals on the blots of extrachromosomal circular molecules (Fig. 3K), while the other subfamily with 69/138 bp periodicity (clone c666) produces weak signals of continuous smear (Fig. 3L). Southern blots of restriction enzyme-digested genomic DNA hybridized with the probes differentiating the two subfamilies confirmed that the 180 bp subfamily occurs in the *V. pannonica* genome along with the previously reported VicTR-A repeats with 69/138 bp repeated units [3] (Fig. 5C).

**Discussion**

Our results show that all the plant satellite repeats that we investigated are prone to the formation of eccDNA. These results complement similar findings described for insects and animals [13-15,30], and significantly broaden our knowledge about plant satellite repeats [16,17] by detection of satellite-derived eccDNA in a total of ten species. Investigated satellites differed in their monomer length, proportion in the genome and chromosomal localization (Table 1). Since eccDNA was detected for all of them, it can be concluded that these features do not have a crucial impact on the formation of circular molecules. A common feature of all investigated satellites was that their eccDNA occurred in the form of open circles, the double-stranded circular molecules relaxed due to the presence of single strand nicks. The possibility that these findings resulted from DNA damage during sample preparation was excluded by detection of intact supercoiled control plasmid added to the samples at the beginning of the isolation procedure. Similar control was used for eccDNA isolation from *Xenopus* embryos which was also found to occur as open circles [13].

In yeast, eccDNA formation requires chromosomal replication as it originates from stalled replication forks [31]. On the other hand, eccDNA production in *Xenopus* is supposed to be uncoupled from DNA replication, although some synthesis requiring replicative polymerases was detected on the newly formed eccDNA [13,32]. The experiments that focused on FokI repeats in *V. faba* revealed comparable levels of eccDNA in mature leaves and young (growing) leaves or root meristems, thus indicating that FokI eccDNA formation is also not tightly linked to DNA replication. We did not observe a significant increase of eccDNA concentration in mechanically damaged leaves, suggesting that its formation is not induced by this sort of stress or DNA degradation processes. It should be noted, however, that due to generally low levels of eccDNA in the investigated tissues and the only semi-quantitative nature of the assay, we could not detect subtle changes in eccDNA concentration.

Our experimental results, together with previous reports [14,17,32], support the hypothesis that eccDNA is produced by homologous intra-strand recombination between satellite repeat units [11,33]. This process is supposed to result in eccDNA molecule sizes corresponding to the multiples of monomer length, which is consistent with our observations. Moreover, the eccDNA patterns detected for different VicTR-A and VicTR-B subfamilies or higher-order repeats suggest that relatively long regions (tens to hundreds of nucleotides) of high sequence similarity are required for efficient recombination. It has been reported that efficiency of homologous recombination depends on similarity of involved sequences [34-36] and is proportional to the length of the similarity [37]. If the length or degree of sequence similarity is decreased, the rate of recombination is reduced rapidly. In yeast, a divergence of only 1% between 350 bp substrates caused a 5-23-fold reduction of mitotic and meiotic recombination, and divergence of 15% led to a 700-fold reduction [38]. In plants, 1.6% and 1.9% sequence heterogeneity was found to decrease the frequency of intrachromosomal recombination by 3.6 [36] and 9.6-fold [39], respectively, and the recombination between 585 bp inverted repeat substrates was reduced by about 4–20 fold when the level of divergence increased from 0.5% to 9% [40]. A similar adverse effect of decreasing similarity on the efficiency of eccDNA formation was also evident in the case of HOR units of the subfamily VG-V of *V. grandiflora* VicTR-B repeats. The 186 bp HORs have significantly higher average similarity (89%) than the individual monomers from which they are composed (78%) [26]. Consequently, the eccDNA was detected only in sizes corresponding to multiples of the HOR lengths but not in multiples of the 38 bp monomers (Fig. 3I). Other VicTR-B subfamilies from the same species that are homogenized at the level of the monomers produced eccDNA with corresponding size distributions (Fig. 3H). Thus, homologous recombination seems to be the major mechanism of eccDNA origin from plant satellite repeats. Other possible recombination mechanisms such as the nonhomologous end joining (NHEJ) repair pathway probably do not contribute significantly to eccDNA production [16]. This conclusion is also consistent with previous studies on yeast, which demonstrated the requirement for RAD52-dependent homologous
Periodicity analysis of VicTR-A repeats in *Vicia narbonensis* and *Vicia pannonica*. The nucleotide autocorrelation analysis of VicTR-A sequences from *V. narbonensis* (A) and *V. pannonica* (B), measuring the excess of pairs of identical nucleotides at various distances (2–500 base pairs). Sequence periodicity was calculated for a distance of k base pairs and nucleotide X as a difference $C_{XX}(k) = p_{XX}(k)-p_Xp_X$, where $p_{XX}$ is the observed frequency of identical nucleotides X and $p_X$ is the proportion of nucleotide X in the sequence [26,27]. The periodicity is revealed as regularly spaced peaks. The results combining periodicities of all four nucleotides are shown.

**Figure 4**

Periodicity analysis of VicTR-A repeats in *Vicia narbonensis* and *Vicia pannonica*. The nucleotide autocorrelation analysis of VicTR-A sequences from *V. narbonensis* (A) and *V. pannonica* (B), measuring the excess of pairs of identical nucleotides at various distances (2–500 base pairs). Sequence periodicity was calculated for a distance of k base pairs and nucleotide X as a difference $C_{XX}(k) = p_{XX}(k)-p_Xp_X$, where $p_{XX}$ is the observed frequency of identical nucleotides X and $p_X$ is the proportion of nucleotide X in the sequence [26,27]. The periodicity is revealed as regularly spaced peaks. The results combining periodicities of all four nucleotides are shown.
recombination in the formation of eccDNA from rDNA repeats [41].

As the eccDNA size distributions of satellites corresponded to multiples of monomer or higher-order repeat lengths we did not find any direct evidence of eccDNA production reflecting the periodicity of nucleosomal structure of chromatin. Nevertheless, it cannot be ruled out that a minor subpopulations of circles arise via this recombination pathway and were not detected under our experimental conditions. Such a mechanism could explain the origin of the VG-V subfamily, which represents a clear example of evolutionary shift from 38 bp monomers to the pentamer-based 186 bp HORs. Moreover, the recombination-based elimination of specific sequence regions was likely involved in the formation of the 180 bp VicTR-A subfamily in *V. pannonica* (Fig. 5A,B) which could represent another case of evolution towards the nucleosome-sized monomers. The constraints imposed on the formation of eccDNA by chromatin could
also explain some discrepancies observed when comparing sequence periodicity of *V. narbonensis* VicTR-A satellite (Fig. 4A) to the size distribution of its eccDNA spots detected on 2-D blots (Fig. 2). While the periodicity of the sequenced VicTR-A clones was found to be based on a monomer-sized repeated units, the eccDNA was found to occur in multiples of dimer and preferentially tetramer units. Interestingly, the tetramer peak (276 bp) on the periodicity plot is higher that the one corresponding to the trimer (207 bp), suggesting emergent tetramer-based periodicity. Thus, this satellite might be in a transition stage towards HOR periodicity driven by preferential formation of dimer/tetramer-based eccDNA. On the other hand, we can not exclude that the observed eccDNA pattern could arise from some unknown, less abundant subfamily of VicTR-A repeats with already developed HOR periodicity which we did not detect among the cloned sequences.

The observed common occurrence of eccDNA is important for proving its role in satellite repeat evolution, but it remains to be investigated to what extent it participates in the processes of satellite repeat amplification and sequence homogenization. In the simplest case, recombination-based excision of eccDNA may only represent a deletion mechanism reducing copy numbers of satellite repeats in the genome. Alternatively, open circular molecules can be further utilized as a replication template, leading to production of long linear stretches on newly synthesized DNA fragments composed from multiple copies of the original circular sequence. Thus, this mechanism would provide both amplification as well as sequence homogenization of satellite DNA. In *Drosophila*, specific circle-with-tail structures of tandemly arranged genes corresponding to *Stellate, Suppressor of Stellate* and histone genes were observed on 2-D gels, suggesting the occurrence of rolling circle replication of these eccDNA. Such rolling circle intermediates (RCIs) of satellite eccDNA molecules were not, however, observed due to the methodological constraints of 2-D electrophoresis [42]. Nor did we find RCIs of satellite eccDNA in our experimental system. This obstacle could be overcome in future experiments by visualizing the content of samples under the electron microscope, as already successfully done for RCIs of mitochondrial plasmid mp1 in *Chenopodium album* [43] and rDNA rolling circles in *Xenopus* [44].

**Conclusion**

This work demonstrated the existence of eccDNA molecules derived from various plant satellite repeats, providing strong support for theoretical models predicting eccDNA as an intermediate in satellite DNA evolution. However, it is yet to be seen to what extent and how the eccDNA is utilized in these processes. Future detailed examination of the molecular basis for recombination events and analysis of replication intermediates should provide a better understanding of the biological principles and constraints involved in these processes.

**Methods**

**Plant material and genomic DNA isolation**

Seeds of plants used in this study were obtained from Osiva Boršov, Czech Republic (V. faba cv. Merkur, V. panmonica cv. Detenická panonská), IPK Gatersleben, Germany (V. grandiflora Scop. var. grandiflora, V. narbonensis L.), NASC, Loughborough, UK (Arabidopsis thaliana Columbia), the Breeding Station at Slapy u Tábora, Czech Republic (*Pisum sativum* cv. Carrera), and the Crop Research Institute, Prague, Czech Republic (*Secale cereale* cv. Dankovské, Triticum aestivum cv. Saxana). Seeds of rice (*Oryza sativa* ssp. japonica var. Nipponbare) were kindly provided by Prof. J. Jiang (University of Wisconsin, Madison, USA). Total genomic DNA was extracted from leaves pooled from several plants as described by Dellaporta et al. [45]. In *V. faba*, DNA was isolated from young (developing) leaves, mature (one-month-old) leaves or 0.5 cm long root tips in order to compare eccDNA levels in various tissues. The DNA concentration measurements were performed using PicoGreen dye (Invitrogen, USA) according to the manufacturer’s recommendations.

**Preparation of circular DNA size markers**

Plasmid-based open circle markers were prepared by cloning *Bs*II-digested lambda DNA (Fermentas International Inc., Canada) into plasmid vector pBluescript II SK+ (Stratagene, USA). Selected plasmid clones of different sizes were isolated and converted from a supercoiled to a linear form by nicking activity of DNaseI (Boehringer Mannheim, Germany). The reaction was performed in a mixture consisting of 25 pg of DNA, 50 mM Tris-HCl pH 7.5, and 10 mM MgCl2 (total volume of 15 μl), for 15 min at 37°C. Open circle markers of small sizes (558, 930, 1,302, 1,674 bp) were designed using LoxP-directed cloning [46]. Complementary oligonucleotides (5’-GAT CTA TAA CIT CGT ATA ATG TAT GCT ATG CGA AGT TAT G-3’, 5’-AAT TCA TAA CIT CGT ATA GCA TAC ATT ATA CGA AGT TAT A-3’) were annealed to form a linear double-stranded fragment harboring the LoxP site and single-stranded overhangs compatible with BamHI and EcoRI restriction sites. The fragment was cloned into BamHI/EcoRI-digested plasmid vector and this construct was further modified by incubation with Cre recombinase (New England BioLabs, USA) and the linear LoxP fragment (in a 20 μl reaction mixture containing 2 μg LoxP vector, 0.5 μM annealed LoxP oligonucleotides, 1 × ligase buffer (Fermentas International Inc.), 10 U Cre recombinase, at 37°C for 30 min and heat-inactivated at 70°C for 10 min), producing linearized plasmid carrying one LoxP
sequence terminated with a BamHI and EcoRI overhang at each end, respectively. This vector was used for cloning BamHI/EcoRI-digested PCR fragments of various lengths amplified from the lambda DNA template using a forward primer (5Ј-TTG CTG AGG ATC CTG TAC CGG CTG TCT GGT ATG TAT G-3Ј) in combination with one of the following reverse primers (5Ј-TTG CTG AGA ATT CTG CTG CGA TCC TTC CTT C-3Ј, 5Ј-TTG CTG AGA ATT CAT CCG CAG GGT GAT CGG CGC-3Ј, 5Ј-TTG CTG AGA ATT CTG GAA CTG CGG AGC CAT C-3Ј, 5Ј-TTG CTG AGA ATT CGC GGC TCC AAG CGC AAC-3Ј). The final constructs thus contained LoxP-PCR fragment-LoxP cassettes which were subsequently released by Cre-mediated recombination in the form of covalently closed circular molecules. The length of these circular DNA markers was 558, 930, 1,302 or 1,674 bp. They were treated with a nicking endonuclease Nt.AlwI (New England BioLabs) to convert them into open circles, and purified by agarose-gel electrophoresis. 50 pg of each of these markers were added into genomic DNA samples prior to 2-D electrophoresis and their positions on the gel were determined using Southern hybridization with the lambda DNA probe.

**EccDNA analysis on two-dimensional agarose gel electrophoresis**

Neutral-neutral 2-D agarose gel electrophoresis was performed as described by Cohen and Lavi [22] with the following modifications. Samples of up to 20 μg of genomic DNA were analyzed (in the case of comparative analysis of eccDNA content in different tissues or in stressed plants, the same DNA amounts were always loaded). The DNA was separated on 0.4% agarose in 1 × TBE buffer for 18 hr and the lanes with samples were excised and stained in 1 × TBE buffer containing 0.3 μg/ml of ethidium bromide for 2 hr. Stained lanes were placed on a gel support at 90° orientation to the direction of electrophoresis and embedded by 1% agarose supplemented with 0.3 μg/ml of ethidium bromide. The second dimension was run in 1 × TBE buffer, 0.3 μg/ml of ethidium bromide at 4 V/cm for 4 hr. Alternatively, the electrophoresis was run on 0.7% and 2% agarose gels (in some cases, 1.5% and 2% was used) for 21 and 8 hours, respectively, in order to improve resolution of small DNA molecules. In addition, the sensitivity of the assay was increased by treating the samples with Plasmid-safe ATP-dependent DNase (Epicentre Biotechnologies), which selectively degraded linear DNA fragments, thus allowing equivalents of 80 – 240 μl of undigested genomic DNA to be loaded. The Plasmid-safe ATP-dependent DNase treatment was preceded by passing high molecular weight genomic DNA through a hypodermic needle (Omnican 100, 0.3 mm in diameter, B. Braun Petzold GmbH, Germany), resulting in its slight shearing which promoted linear DNA degradation. The treatment was performed using 160 – 480 U of Plasmid-safe ATP-dependent DNase in 800 – 2400 μl of reaction buffer overnight at 37°C and stopped by incubation at 70°C for 30 min. Short fragments of degraded nucleic acids were removed using molecular weight cut-off columns (Microcon-30 and Microcon-100, Millipore, USA) and the digestion and purification was repeated once more.

The presence of single-stranded DNA was tested by dividing purified samples into halves and incubating them with or without 5 – 50 U of mung bean nuclease (MBN, Takara Bio Inc., Japan) in 1 × MBN buffer for 10 min at 37°C. Alternatively, RNaseH (Ambion, USA) directed degradation or a combination of RNaseH followed by MBN treatment was performed to check for the presence of DNA-RNA hybrid molecules. RNaseH treatment was carried out in 100 μl of 1 × RNaseH reaction mixture and 25 U of RNaseH for 1 hr at 37°C. Finally, to remove any contamination that could affect subsequent electrophoretic separation, the samples were purified using Wizard SV Gel and PCR Clean-Up System columns (Promega, USA).

**Southern hybridization**

Following electrophoresis, DNA was transferred onto Hybond-N+ membranes (Amersham Biosciences, USA) by capillary transfer. Hybridization probes for satellite repeats were derived from fragments amplified by PCR using specific primers and genomic DNA as a template (AT_180: 5Ј-ACC TTC TTC TTC GTG CTT TTC AAA G-3Ј, 5Ј-GTT GGT TAG TGT TTG GGA TCT G-3Ј; CentO: 5Ј-AAA ACA TGA TTT TTG GAC ATA TTG G-3Ј, 5Ј-TGA CAA AAG TIC GCC GCC-3Ј; PisTR-B: 5Ј-ACC CAT GAA ATT TGA TTG-3Ј, 5Ј-CAA ATA TTT TTG GAC ATA TTG G-3Ј; FokI: 5Ј-CAT TAT GGA AGG TAG TCT GTG GTC G-3Ј; VicTR-A: 5Ј-TAC ATA AAA GTC AYG AAG TT-3Ј, 5Ј-TGA TAA GTC TTC ARA AAA T-3Ј, 5Ј-GAA GAC TTA TAT TCA CIT-3Ј). The probe for IGS-like satellite was prepared by insert amplification from clone S12 [47] using T3/T7 primers and removing surrounding polylinker sequences by restriction digestion and gel purification. The same procedure was used for probe preparation for the VG-V subfamily of VicTR-B repeats (using clone c609 [GenBank:DO139394]) and for VicTR-A clones c653 [GenBank:EU568805] and c666 [GenBank:EU568818]. Fragment labeling and hybridization were done using the AlkPhos Direct Kit (Amersham Biosciences), according to the manufacturer's recommendations (hybridization and washing temperatures varied between 50°C and 61°C according to probe AT/GC content). Hybridization specificity for the VG-V subfamily was verified in hybridization using clone c609 as a probe and membranes with clones...
c605 [GenBank:DQ139381], c788 [GenBank:DQ139383], c606 [GenBank:DQ139368] and c610 [GenBank:DQ139388] [26] representing other VicTR-B subfamilies. The specificity of VicTR-A clones c653 and c666 was confirmed in hybridization using sequences of clones c651 [GenBank:EU1568803], c652 [GenBank:EU1568804], c654 [GenBank:EU1568806], c763 [GenBank:EU1568830] and c768 [GenBank:EU1568835] as negative controls. To detect signals, blots were incubated with chemiluminescent substrate (CDP-Star, Amersham Biosciences) and exposed to X-ray film for up to 60 hr.

Cloning and sequence analysis of VicTR-A repeats
The preparation of V. pannonica and V. narbonensis shotgun genomic libraries and screening and sequencing of their clones was performed as described by [26] except that VicTR-A clone P5 [3] was used as a probe to screen the libraries by colony hybridization. The sequences were deposited in GenBank under the accession nos. EU568802–EU568868. Sequence periodicity analysis based on the concept of nucleotide autocorrelation functions [27] was performed as described previously [26]. Dot-plot sequence comparisons were done using a dotter program [48].

Authors’ contributions
AN and JM designed the study, and AN carried out most of the experimental work. AK constructed and screened genomic DNA libraries and participated in eccDNA detection experiments. JM performed bioinformatic analysis of the newly sequenced satellite repeats. All authors contributed to the manuscript preparation and approved its final version.

Acknowledgements
We thank Ms. J. Látalová and Ms. H. Štepanicková for their excellent technical assistance. This work was supported by grants GA204/06/P360 from the Czech Science Foundation, AVOZ50510513 from the Academy of Sciences of the Czech Republic, and LC06004 from the Ministry of Education, Youth and Sports of the Czech Republic.

References
1. Ingham LD, Hanna WW, Baier JW, Hannah LC: Origin of the main class of repetitive DNA within selected Pennisetum species. Mol Gen Genet 1993, 238(3):350-356.
2. Dechyeva D, Gindullis F, Schmidt T: Divergence of satellite DNA and interspersion of dispersed repeats in the genome of the wild beet Beta campestris. Chromosome Res 2003, 11(1):3-21.
3. Macas J, Požárková D, Navrátilová A, Nouzová M, Neumann P: Two new families of tandem repeats isolated from genus Vicia using genomic self-priming PCR. Mol Gen Genet 2000, 263(5):741-751.
4. Nouzová M, Kubálková M, Doleželová M, Kubišková A, Neumann P, Doležel J, Macas J: Cloning and characterization of new repetitive sequences in field bean (Vicia faba L.). Ann Bot 1999, 83(5):535-542.
5. Tek AL, Song JQ, Macas J, Jiang JM: Sobo, a recently amplified satellite repeat of potato, and its implications for the origin of tandemly repeated sequences. Genetics 2005, 170(3):1231-1238.
6. Dover G: Molecular drive. Trends in Genetics 2002, 18(11):587-589.
7. Stephan W: Recombination and the evolution of satellite DNA. Genet Res 1986, 47(3):167-174.
8. Charlesworth B, Sniegowski P, Stephan W: The evolutionary dynamics of repetitive DNA in eukaryotes. Nature 1994, 371(6494):215-220.
9. Stephan W, Cho S: Possible role of natural selection in the formation of tandem repetitive noncoding DNA. Genetics 1994, 136(1):333-341.
10. Walsh JB: Persistence of tandem arrays: implications for satellite and simple-sequence DNAs. Genetics 1987, 115(3):553-567.
11. Gaubatz JW: Extrachromosomal circular DNAs and genomic sequence plasticity in eukaryotic cells. Mutat Res 1990, 237(5-6):271-292.
12. Cohen S, Regev A, Lavi S: Small polydispersed circular DNA (spcDNA) in human cells: association with genomic instability. Oncogene 1997, 14(8):977-985.
13. Cohen S, Menut S, Méchali M: Regulated formation of extrachromosomal circular DNA molecules during development in Xenopus laevis. Mol Cell Biol 1999, 19(10):6682-6689.
14. Colan S, Yacobi K, Segal D: Extrachromosomal circular DNA of tandemly repeated genomic sequences in Drosophila. Genome Res 2003, 13(6A):1133-1145.
15. Cohen Z, Bacharach E, Lavi S: Mouse major satellite DNA is prone to eccDNA formation via DNA Ligase IV-dependent pathway. Oncogene 2006, 25(33):4515-4524.
16. Zellinger B, Akimcheva S, Puizinova J, Schirato M, Riha K: Ku suppresses formation of telomeric circles and alternative telomere lengthening in Arabidopsis. Mol Cell 2007, 27(1):163-169.
17. Cohen S, Houben A, Segal D: Extrachromosomal circular DNA derived from tandemly repeated genomic sequences in plants. Plant J 2008, 53(4):1027-34.
18. Macas J, Mezáros T, Nouzová M: PlantSat: a specialized database for plant satellite repeats. Bioinformatics 2002, 18(1):28-35.
19. Schmidt T, Heslop-Harrison JS: Genomes, genes and junk: the large-scale organization of plant chromosomes. Trends Plant Sci 1998, 3(5):195-199.
20. Henikoff S, Ahmad K, Malik HS: The centromere paradox: stable inheritance with rapidly evolving DNA. Science 2001, 293(5532):1099-1102.
21. Luger K: Dynamic nucleosomes. Chromosoma Res 2006, 14(1):5-16.
22. Cohen S, Lavi S: Induction of circles of heterogeneous sizes in carcinogen-treated cells: two-dimensional gel analysis of circular DNA molecules. Mol Cell Biol 1996, 16(5):2505-2520.
23. Kato A, Yakura K, Tanifuji S: Sequence analysis of Vicia faba repeated DNA, the FokI repeat element. Nucleic Acids Res 1984, 12(16):6415-6426.
24. Maggini F, D'Ovidio R, Gelati MT, Frediani M, Cremonini R, Ceccarelli M, Minelli S, Cionini PG: FDOK DNA repeats in the genome of Vicia faba: species specificity, structure, redundancy modulation, and nuclear organization. Genome 1995, 38(6):1255-1261.
25. Lee HR, Neumann P, Macas J, Jiang JM: Transcription and evolutionary dynamics of the centromeric satellite repeat CentO in rice. Mol Biol Evol 2006, 23(12):2505-2520.
26. Macas J, Navrátilová A, Kubišková A: Sequence homogenization and chromosomal localization of VicTR-B satellites differ between closely related Vicia species. Chromosoma 2006, 115(6):437-447.
27. Herzel H, Weiss O, Trifonov EN: 10–11 bp periodicities in complete genomes reflect protein structure and DNA folding. Bioinformatics 1999, 15(3):187-193.
28. Appels R, Moran LB, Gustafson JP: Rye heterochromatin. I: Studies on clusters of the major repeating sequence and the identification of a new dispersed repetitive sequence element. Can J Genet Cytol 1986, 28(5):645-657.
29. Katsiotis A, Hagidimitriou M, Douka A, Hatzopoulos P: Genomic organization, sequence interrelationship, and physical localization using in situ hybridization of two tandemly repeated DNA sequences in the genus Olea. Genome 1998, 41(4):527-534.
30. Ohki R, Oishi M, Kiyama R: Preference of the recombination sites involved in the formation of extrachromosomal copies.
of the human alphoid Sau3A repeat family. Nucleic Acids Res 1995, 23(24):4971-4977.

31. Blakely Y, Himmels T, Kobayashi T. Transcription-dependent recombination and the role of fork collision in yeast rDNA. Genes Dev 2003, 17(12):1497-1506.

32. Cohen S, Mächali M. A novel cell-free system reveals a mechanism of circular DNA formation from tandem repeats. Nucleic Acids Res 2001, 29(12):2542-2548.

33. Smith CA, Vinograd J. Small polydisperse circular DNA of HeLa cells. J Mol Biol 1972, 69(2):163-178.

34. Kawasaki I, Bae YS, Eki T, Kim Y, Ikeda H. Homologous recombination of monkey alpha-satellite repeats in an in vitro simian virus 40 replication system: possible association of recombination with DNA-replication. Mol Cell Biol 1994, 14(6):4173-4182.

35. Modrich P, Lahue R. Mismatch repair in replication fidelity, genetic recombination, and cancer biology. Annu Rev Biochem 1996, 65(1):93-133.

36. Opperman R, Emmanuel E, Levy AA. The effect of sequence divergence on recombination between direct repeats in Arabidopsis. Genetics 2004, 168(4):2207-2215.

37. Rubnitz J, Subramani S. The minimum amount of homology required for homologous recombination in mammalian cells. Mol Cell Biol 1984, 4(11):2253-2258.

38. Chen W, Jinks-Robertson S. The role of the mismatch repair machinery in regulating mitotic and meiotic recombination between diverged sequences in yeast. Genetics 1999, 151(4):1299-1313.

39. Li L, Santerre-Ayotte S, Boivin EB, Jean M, Belzile F. A novel reporter for intrachromosomal homoeologous recombination in Arabidopsis thaliana. Plant J 2004, 40(6):1007-1015.

40. Li L, Jean M, Belzile F. The impact of sequence divergence and DNA mismatch repair on homologous recombination in Arabidopsis. Plant J 2006, 45(6):908-916.

41. Park NU, Defosse PA, Guarente L. Effects of mutations in DNA repair genes on formation of ribosomal DNA circles and life span in Saccharomyces cerevisiae. Mol Cell Biol 1999, 19(5):3848-3856.

42. Cohen S, Agmon N, Yacobi K, Mislovati M, Segal D. Evidence for rolling circle replication of tandem genes in Drosophila. Nucleic Acids Res 2005, 33(14):4519-4526.

43. Elion R. R-loops and homologous recombinational repair. J Mol Biol 1996, 254(2):163-180.

44. Hourcade D, Dressler D, Wolfson J. Amplification of ribosomal-RNA genes involves a rolling circle intermediate. Proc Natl Acad Sci USA 1973, 70(10):2926-2930.

45. Dellaporta SL, Wood J, Hicks JB. A plant DNA minipreparation: A new model for DNA concatemer resolution by mitochondrial plasmid mpl. EMBO J 2002, 21(12):3128-3136.

46. Buchholz F, Bishopp M. LoXP-directed cloning: use of Cre recombinase as a universal restriction enzyme. Biotechniques 2001, 31(4):906-908.

47. Neumann P, Nouzová M, Macas J. Molecular and cytogenetic analysis of repetitive DNA in pea (Pisum sativum L.). Genome 2001, 44(4):716-728.

48. Macas J, Neumann P, Navrátilová A. Repetitive DNA in the pea (Pisum sativum L.) genome: comprehensive characterization using 454 sequencing and comparison to soybean and Medicago truncatula. BMC Genomics 2007, 8:427.

49. Dong F, Miller JT, Jackson SA, Wang GL, Ronald PC, Jiang J. Rice (Oryza sativa) centromeric regions consist of complex DNA. Proc Natl Acad Sci USA 1998, 95(14):8135-8140.

50. Chenz Z, Dong F, Langdon T, Ouyang S, Buell CR, Gu M, Blattner FR, Jiang J. Functional rice centromeres are marked by a satellite repeat and a centromere-specific retrotransposon. Plant Cell 2002, 14(8):1691-1704.

51. Bedbrook JR, Jones J, O’Dell M, Thompson RD, Flavell RB. A molecular description of telomeric heterochromatin in Secale species. Cell 1980, 19(2):545-560.

52. Navrátilová A, Neumann P, Macas J. Karyotype analysis of four Vicia species using in situ hybridization with repetitive sequences. Annu Bot 2003, 91(7):921-926.

53. Neumann P, Nouzová M, Macas J. A plant DNA minipreparation: A universal restriction enzyme. BioTechniques 1983, 19:151-158.

54. Sonnhammer EL, Durbin R. A dot-matrix program with dynamic threshold control suited for genomic DNA and protein sequence analysis. Gene 1995, 167(1-2):GCl-10.

55. Ananthawat-Jansson K, Heslop-Harrison JS. Isolation and characterization of genome-specific DNA sequences in Triticaceae species. Mol Gen Genet 1993, 240(2):151-158.

56. Vereshchagina A, Svitasev S, Gummesson PO, Salomon B, von Bothmer R, Bryngelson T. Characterization of a family of tandemly repeated DNA sequences in Triticaceae. Theor Appl Genet 1994, 89(2):217-225.

57. Heslop-Harrison JS. Comparative genome organization in plants: from sequence and markers to chromatins and chromosomes. Plant Cell 2000, 12(5):617-636.

58. Murata M, Ogura Y, Motoyoshi F. Centromeric repetitive sequences in Arabidopsis thaliana. Jpn J Genet 1994, 69(4):261-370.

59. Kumekawa N, Hoshouchi T, Tsuruoka H, Kotani H. The size and sequence organization of the centromeric region of Arabidopsis thaliana chromosome 4. DNA Res 2001, 8(6):285-290.