Mol Genet Genomics (2017) 292:283–296
DOI 10.1007/s00438-016-1271-3

ORIGINAL ARTICLE

# Evolutionary dynamics of an at-rich satellite DNA and its contribution to karyotype differentiation in wild diploid Arachis species

**Sergio Sebastián Samoluk[1] · Germán Robledo[1,2] · David Bertioli[3,4] · José Guillermo Seijo[1,2]**

Received: 17 June 2016 / Accepted: 4 November 2016 / Published online: 12 November 2016


**Abstract** Satellite DNA (satDNA) is a major component of the heterochromatic regions of eukaryote genomes and usually shows a high evolutionary dynamic, even among closely related species. Section Arachis (genus *Arachis*) is composed of species belonging to six different genomes (A, B, D, F, G and K). The most distinguishing features among these genomes are the amount and distribution of the heterochromatin in the karyotypes. With the objective of gaining insight into the sequence composition and evolutionary dynamics of the heterochromatin fraction in *Arachis*, we investigated here the sequence diversity, genomic abundance, and chromosomal distribution of a satDNA family (ATR-2) among seven diploid species of section Arachis. All of the isolated sequences were AT-rich and highly conserved at both intraspecific and interspecific levels, without any species-specific polymorphism. Pairwise comparisons of isolated ATR-2 monomers revealed that most of the nucleotide sites were in the first two transitional stages of Strachan's model. However, the abundance of ATR-2 was significantly different among genomes according to the 'library hypothesis'. Fluorescent in situ hybridization revealed that ATR-2 is a main component of the DAPI$^+$ centromeric heterochromatin of the A, F, and K genomes. Thus, the evolution of the different heterochromatin patterns observed in *Arachis* genomes can be explained, at least in part, by the differential representation of ATR-2 among the different species or even among the chromosomes of the same complement. These findings are the first to demonstrate the participation of satDNA sequences in the karyotype diversification of wild diploid *Arachis* species.

**Keywords** Satellite DNA · Library hypothesis · Heterochromatin · Karyotype differentiation · *Arachis* species

Communicated by S. Hohmann.

**Electronic supplementary material** The online version of this article (doi:10.1007/s00438-016-1271-3) contains supplementary material, which is available to authorized users.

✉ Sergio Sebastián Samoluk
ssamoluk@agr.unne.edu.ar; samocarp@hotmail.com

Germán Robledo
grobledo@agr.unne.edu.ar

David Bertioli
davidbertioli@unb.br

José Guillermo Seijo
seijo@agr.unne.edu.ar

[1] Facultad de Ciencias Agrarias, Instituto de Botánica del Nordeste (UNNE-CONICET), Casilla de Correo 209, 3400 Corrientes, Argentina

[2] Facultad de Ciencias Exactas y Naturales y Agrimensura, Universidad Nacional del Nordeste, Av. Libertad 5460, 3400 Corrientes, Argentina

[3] Universidade de Brasília, Campus Universitário, 70910-900 Brasília, DF, Brazil

[4] Center for Applied Genetic Technologies, University of Georgia, Athens, GA, USA

## Introduction

Satellite DNA (satDNA), one of the major classes of repetitive sequences in almost all eukaryotic genomes, consists of tandem arrays of repetitive units preferentially located within heterochromatic regions of the karyotypes (Charlesworth et al. 1994; Schmidt and Heslop-Harrison 1998; Biscotti et al. 2015). Since its rapid evolution, this kind of sequence can vary in nucleotide composition, genome

representation, and chromosome distribution even among closely related species, leading to 'satDNA species-specific profiles' (Plohl et al. 2008). Therefore, comparative studies on satDNAs are useful for the investigation of evolutionary relationships between plant species (Menzel et al. 2008) as well as for understanding karyotype structure and genome evolution (Ugarkovic and Plohl 2002).

The repetitive units (monomers) of satDNAs have variable lengths and are tandemly arranged in blocks of up to 100 Mbp (Charlesworth et al. 1994; Schmidt and Heslop-Harrison 1998; Biscotti et al. 2015). The monomers typically range between 140 and 180 bp or 300 and 360 bp, which is equivalent to the length of DNA wrapped around one or two nucleosome cores, respectively (Henikoff et al. 2001; Macas et al. 2002; Plohl et al. 2008; Sharma and Raina 2005). These sequences are usually located in particular regions of the chromosomes, being a major component of centromeric (Hudakova et al. 2001; Gindullis et al. 2001; Urdampilleta et al. 2009), telomeric (Pich et al. 1996; Macas et al. 2000) and, less frequently, interstitial heterochromatin (Mukai et al. 1992; Plohl et al. 2008).

Families of satDNA are distinguished by nucleotide composition, sequence complexity, length of the repetitive unit, and copy number (Ugarkovic and Plohl 2002; Plohl et al. 2008). Each family may also have different evolutionary rates; therefore, while some satDNA families are species-specific (Bachmann and Sperlich 1993; Garrido-Ramos et al. 1999), others are more conserved among related species (Mravinac et al. 2002; Robles et al. 2004). A library of satDNA sequences may be shared by homologous and non-homologous chromosome regions (Lim et al. 2000; Hizume et al. 2002; Rosato et al. 2012) and by closely related species (Fry and Salser 1977; Ugarkovic and Plohl 2002; Plohl et al. 2008). However, the different members of each satDNA family may be differentially amplified (Ugarkovic and Plohl 2002). Thus, the dynamic of the evolutionary processes affecting satDNAs may lead to large structural chromosome changes (Kuhn et al. 2008; Koukalova et al. 2010; Ferree and Prasad 2012).

The establishment of species-specific patterns occurs, because some particular satDNA families tend to evolve concertedly (Heslop-Harrison et al. 2003; Rudd et al. 2006). It has been proposed that concerted evolution occurs through a two-step process (Dover 1982, 2002). This process initially includes the spreading of a new monomeric variant throughout a repetitive family by non-Mendelian molecular mechanisms (Stephan 1986; Dover 2002; Glinka et al. 2006; Plohl et al. 2008) followed by the fixation of the new variant within a population by sexual reproduction (Dover 1982, 2002; Ugarkovic and Plohl 2002). In this way, the whole process usually leads to the intraspecific homogenization of the satDNA family (Ugarkovic and Plohl 2002), while at the interspecific level leading to the fixation of species-specific polymorphisms (Ugarkovic and Plohl 2002). However, in many species, mainly in those with uniparental reproduction, satDNA families do not follow these evolutionary pathways and show high-sequence similarity among taxa diverged by long evolutionary periods (Plohl et al. 2010; Ugarkovic 2005; Vittorazzi et al. 2014).

The genus *Arachis* (Leguminosae) comprises approximately 80 species arranged in nine taxonomic sections (Fernandez and Krapovickas 1994; Krapovickas and Gregory 1994; Valls and Simpson 2005). Section Arachis is the most derived and is composed of 31 species, 29 wild diploids ($2n = 2x = 20$, 18), and 2 allotetraploids ($2n = 4x = 40$). Diploid species are assigned to different genomes (A, B, D, F, G, and K) and karyotype groups (Gregory and Gregory 1979; Stalker 1991; Robledo and Seijo 2008, 2010; Robledo et al. 2009; Silvestri et al. 2014), while the allotetraploids peanut (*Arachis hypogaea*) and its most closely related wild species, *Arachis monticola*, have an AABB genome constitution (Seijo et al. 2007; Grabiele et al. 2012). Moreover, three karyotype groups (Chiquitano, Pantanal, and La Plata River Basin) were identified among the A genome species (Robledo et al. 2009). The genome assignment in this group of species was based on cross-compatible assays (Gregory and Gregory 1979; Stalker et al. 1991), chromosome morphology (Smartt et al. 1978; Stalker et al. 1991; Fernandez and Krapovickas 1994), patterns of heterochromatin and rDNA loci (Robledo et al. 2009; Robledo and Seijo 2010), as well as on DNA phylogenies (Moretzsohn et al. 2013; Grabiele et al. 2012; Leal-Bertioli et al. 2015).

Among these characteristics, the most striking difference in *Arachis* genomes is the relative percentage of $DAPI^+$ heterochromatin and its distribution among karyotypes (Seijo et al. 2004; Robledo and Seijo 2008, 2010; Robledo et al. 2009). The species of the A, K, and D genomes have conspicuous centromeric bands in most (9–10 pairs) chromosomes; those of the F genome have tiny centromeric bands in six to seven chromosome pairs, while those of the B genome are deprived of conspicuous heterochromatin. The D genome species (*Arachis glandulifera*) has many distinctive interstitially located heterochromatic bands. This variation in heterochromatic patterns makes section Arachis a suitable model for studying the evolutionary dynamics of satDNAs and to infer the impact of this kind of sequence on karyotype evolution. Until now, only one satDNA sequence of 115 bp was isolated from a Cot-1 library of peanut, which is mainly distributed in the centromeric regions of the B genome complement of this allotetraploid (Zhang et al. 2012). However, this genome does not have conspicuous heterochromatic bands; therefore,

**Table 1** *Arachis* species analyzed, provenances, genomes, and karyotype groups

| Species | Provenance and collection number | Genome (Karyotype group) |
|---|---|---|
| *A. helodes* Martius ex Krapov. and Rigoni | Brasil, Mato Grosso, Cuiabá. K, G 30029 (CTES) | AA (Pantanal) |
| *A. duranensis* Krapov. and W. C. Gregory | Argentina, Salta, San Martín, Campo Durán. K 7988 (CTES) | AA (La Plata River Basin) |
| *A. cardenasii* Krapov. and W.C. Gregory | Bolivia, Santa Cruz, Chiquitos, Roboré. K, S, Sc 36015 (CTES) | AA (Chiquitano) |
| *A. ipaënsis* Krapov. and W. C. Gregory | Bolivia, Tarija, Gran Chaco, Ipa. K, G, B, P, Sc, S 30076 (CTES) | BB |
| *A. trinitensis* Krapov. and W.C. Gregory | Bolivia, Beni, Cercado, Trinidad. Wi 1117 (CTES) | FF |
| *A. batizocoi* Krapov. and W. C. Gregory | Bolivia, Santa Cruz, Cordillera, Parapeti. K 9484 (CTES) | KK |
| *A. glandulifera* Stalker | Bolivia, Santa Cruz, Velasco, San Ignacio. Se, Sn 3263 (CTES, LPB) | DD |

*B* D. Banks, *G* W. C. Gregory, *K* A. Krapovickas, *P* J. Pietrarelli, *S* C. E. Simpson, *Sc* A. Schinini, *Se* J. G. Seijo, *Sn* V. G. Solís Neffa, *Wi* D. E. Williams; *CTES* Herbarium of the Instituto de Botánica del Nordeste, Corrientes, Argentina, *FCA* plant introduction number of Facultad de Ciencias Agrarias, Universidad Nacional del Nordeste, Argentina, *LPB* National Herbarium of Bolivia, La Paz, Bolivia

this sequence cannot be associated with the heterochromatic fraction of the A complement of the cultigen.

In this context, the objective of this work was to gain insight into the sequence composition and the evolutionary dynamics of the heterochromatic fraction in species of section Arachis. For this purpose, in an initial survey for satDNAs in PeanutBase (http://peanutbase.org), a new family (named ATR-2) was identified as one of the most represented in the genome of *Arachis duranensis* (A genome). Here, we isolated this satDNA from seven wild diploid species ($x = 10$) belonging to the different genomes and karyotype groups of section Arachis. Furthermore, the sequence diversity, molecular evolution, chromosomal distribution, and genome representation were investigated for each species. The interspecific investigation performed in this study demonstrated that ATR-2 is conserved at the sequence level and is preferentially distributed in $DAPI^+$ heterochromatin blocks. However, quantitative analysis showed a different abundance of this sequence in all species analyzed. The results obtained here are the first to demonstrate that satDNA sequences have participated in the diversification of the heterochromatic patterns observed in the karyotypes of diploid species ($x = 10$) in section Arachis.

## Materials and methods

### Plant material

Leaf and root samples were obtained from diploid ($2n = 2x = 20$) species of the different genomes and karyotype groups of section *Arachis*. The seven species used in this study, their provenance, their genome constitution and their karyotype groups are listed in Table 1. All plant materials were obtained from the *Arachis* germplasm collection maintained at the Instituto de Botánica del Nordeste, Corrientes, Argentina.

### Genomic DNA extraction, amplification, and cloning

Genomic DNA was extracted from young leaves using double precipitation with cetyltrimethylammonium bromide (CTAB) according to Grattapaglia and Sederoff (1994). DNA concentrations were determined by spectrophotometry.

From the alignment of the putative monomers of the satDNA "ATR-2", a pair of specific primers was designed based on the most conserved regions of the motifs. Polymerase chain reaction (PCR) was performed in a final volume of 10 µl in the presence of 50 ng of total genomic DNA, 2 mM each primer, 0.2 mM dNTPs, 2.5 mM $MgCl_2$, and 0.5 U of Taq polymerase in the corresponding PCR buffer. After an initial denaturation step at 94 °C for 5 min, amplifications were conducted for 35 cycles at 94 °C for 30 s, 54 °C for 1 min, and 72 °C for 1 min, with a final elongation step at 72 °C for 5 min. The PCR products were analyzed by electrophoresis, and the bands were cut out of the gel, purified using the Wizard® SV Cleanup System (Promega), and ligated to pGEM®-T Easy Vector (Promega). The recombinant clones were sequenced by Macrogen Inc. Service (Korea).

### Sequence analysis

Alignments were performed using the Clustal W program (Thompson et al. 1994) and then manually edited. The ATR-2 satellite sequence was characterized by (a) the nucleotide composition; (b) the calculation of intraspecific variation, measured by the nucleotide diversity '$\pi$' as the average number of nucleotide differences per site between two sequences (Nei 1987); and (c) the calculation of interspecific variability (Dxy value; Nei 1987). These parameters were estimated using the program satDNA Analyzer version 1.2 (Navajas-Pérez et al. 2007). The analysis of molecular variance (AMOVA test, Excoffier et al. 1992)

was performed with the program ARLEQUIN 2.000 (Schneider et al. 2000).

The software satDNA Analyzer was also used to determine the pattern of variation at any given nucleotide position of ATR-2 (Strachan et al. 1985). This method compares pairs of sequences from different species at every nucleotide position independently. According to the variability observed, the method considers six different propagation stages (I–VI) of a new monomeric variant from its inception to its fixation on any family of sequences and species. Stage I represents the complete homogeneity of all monomeric variants compared between a given pair of species. Stage II represents the presence of a new mutation in a minority of monomers of one of the two species compared. In stage III, the frequency of the new mutation is approximately equal to the frequency of the nucleotide predecessor. Stage IV is characterized by a higher frequency of the new mutation than that of the ancestor nucleotide. Stage V represents a diagnostic site, where the new variant is completely homogenized and fixed in the monomers of one of the species. Finally, in stage VI, a new nucleotide variant appears, replacing a fixed mutation in stage V.

Distance analyses of the monomers isolated from all of the species studied here were performed with the software MEGA version 5 (Tamura et al. 2011). The trees were constructed using the neighbor-joining (NJ) algorithm (Saitou and Nei 1987) using a bootstrapping of 1000 replicates (Felsenstein 1985).

### Fluorescent in situ hybridization (FISH)

To analyze the chromosome distribution of the isolated satDNA, the AhyTR-2 c6 clone isolated from *A. hypogaea* was labelled with 1 mM biotin-11-dUTP by PCR using the same amplification conditions described earlier. The efficiency of the labelled probe was checked by a dot-blot analysis.

Hybridizations were performed according to Seijo et al. (2004). The hybridization mix contained DNA probe at a concentration of 2.5 ng/µl. The hybridization and post-hybridization washes were carried out with a stringency of 80%. The first and second sets of antibodies consisted of mouse anti-biotin and rabbit anti-mouse conjugated to tetramethyl-rhodamine isothiocyanate (TRITC), respectively (Sigma-Aldrich, St. Louis, Missouri, USA). Vectashield medium (Vector Laboratories) containing 2 mg/ml of 4′-6-diamidino-2-phenylindole (DAPI) was used to counterstain and mount the preparations. Pseudocolored images were obtained for each metaphase by overlaying the FISH hybridization signals (red) onto the DAPI fluorescence images (gray).

### Dot-blot analysis

To estimate the genomic representation of the ATR-2 sequences, different quantities of genomic DNA of the species analyzed were spotted onto Zeta-Probe® blotting membranes (Bio-Rad Laboratories) together with a dilution series of the AhyTR-2 c6 clone by duplicating each membrane. The membranes were hybridized with the probe (the same used for FISH) at a concentration of 25 ng/ml at 55 °C (overnight). Hybridization and post-hybridization washes were adjusted to detect sequences with at least 80% similarity. The membrane was incubated with CSPD substrate (Roche), exposed to a Hyperfilm ECL (Amersham) for approximately 1 h and then developed. The intensity of hybridization signals was determined using the Quantity One program version 4.6.1 (Bio-Rad). Once the linear relationship between the applied DNA amounts and the signal intensities was verified, the genomic representation of ATR-2 was estimated by comparing the signal intensities of the genomic DNA with those of the calibration curves. The percentage of ATR-2 sequences per haploid genome (Cx value) in the different species analyzed was calculated from three independent hybridization experiments considering the genome sizes estimated by Samoluk et al. (2015a).

## Results

### Isolation and sequence analysis

Fifty-six ATR-2 sequences were isolated from the genomic DNA of seven wild diploid *Arachis* species ($x = 10$) belonging to the different genomes and karyotype groups (Table 2). All of the species analyzed revealed a similar ladder pattern of electrophoretic bands, with lengths ranging from 300 to 1000 bp. The sequences were named with the letter 'A' (referring to the genus *Arachis*), followed by two identifying letters of the specific epithet of the species name from which they were isolated, 'TR-2' (by the name of the satDNA family ATR-2) and the letter 'c' (referring to the word clone), with a distinguishing number for each clone. For example, AduTR-2 c1 refers to clone number one of the ATR-2 satDNA family isolated from the species *A. duranensis*. The nucleotide sequences of all clones were submitted to public databases (GenBank KX685960–KX686015).

The alignment length of all of the isolated sequences was 339 bp. However, the average length of the isolated clones in different species ranged from 295 bp (*A. glandulifera*) to 300 bp (*Arachis trinitensis*). In the latter species, the clone AtrTR-2 c8 had the longest nucleotide sequence because of an internal duplication of 19 nucleotides

**Table 2** Sequence features of ATR-2 satDNA in seven wild diploid *Arachis* species ($2n = 2x = 20$)

| Species | $N^a$ | Length (bp)[b] | AT (%)[c] | Π[d] | Genome abundance (%)[e] | *N* (chromosome position)[f] |
|---|---|---|---|---|---|---|
| *A. duranensis* | 10 | 297 | 74.86 | 0.03 | 1.52 | 14 (C) |
| *A. helodes* | 8 | 295 | 74.57 | 0.09 | 1.36 | 16 (C) |
| *A. cardenasii* | 8 | 299 | 75.17 | 0.07 | 1.20 | 16 (C) |
| *A. ipaënsis* | 7 | 297 | 75.15 | 0.09 | 0.01 | 0 |
| *A. batizocoi* | 8 | 297 | 75.04 | 0.16 | 0.56 | 16 (C), 4 (I) |
| *A. trinitensis* | 10 | 300 | 75.55 | 0.06 | 0.48 | 14 (C), 4 (T) |
| *A. glandulifera* | 5 | 297 | 74.45 | 0.07 | 0.07 | 2 (C), 4 (I) |

The chromosome position is indicated between parentheses

*C* centromeric, *I* interstitial, *T* telomeric

[a] Number of sequenced clones

[b] Monomeric average length

[c] Percentage of nucleotides A and T

[d] Intraspecific variability: Π

[e] Genome abundance of ATR- 2 per haploid genome expressed as percentage

[f] Number of FISH signals per diploid complement ($2n = 2x = 20$)

**Table 3** Interspecific comparative analysis of ATR-2 sequences

| Pairwise comparisons | Genome | Dxy[a] | Transitional stages (Strachan et al. 1985)[b] | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | | 1 | 2 | 3 | 4 | 5 | 6 |
| *A. batizocoi–A. glandulifera* | K–D | 0.17 | 57.19 | 18.49 | 10.27 | 4.11 | 0.00 | 0.00 |
| *A. batizocoi–A. helodes* | K–A | 0.15 | 49.32 | 21.23 | 5.82 | 2.05 | 0.00 | 0.00 |
| *A. batizocoi–A. cardenasii* | K–A | 0.16 | 52.05 | 22.26 | 5.82 | 3.77 | 0.00 | 0.00 |
| *A. batizocoi–A. duranensis* | K–A | 0.15 | 55.14 | 18.49 | 5.82 | 4.11 | 0.00 | 0.00 |
| *A. batizocoi–A. trinitensis* | K–F | 0.14 | 43.49 | 23.97 | 5.47 | 3.77 | 0.00 | 0.00 |
| *A. batizocoi–A. ipaënsis* | K–F | 0.16 | 50.00 | 23.63 | 7.53 | 1.37 | 0.00 | 0.00 |
| *A. cardenasii–A. glandulifera* | A–D | 0.09 | 68.84 | 17.12 | 4.79 | 2.05 | 0.00 | 0.00 |
| *A. cardenasii–A. helodes* | A–A | 0.10 | 57.53 | 22.60 | 2.40 | 1.03 | 0.34 | 0.00 |
| *A. cardenasii–A. duranensis* | A–A | 0.05 | 68.15 | 19.18 | 0.00 | 0.00 | 0.00 | 0.00 |
| *A. cardenasii–A. trinitensis* | A–F | 0.09 | 53.08 | 25.00 | 1.71 | 1.37 | 0.34 | 0.34 |
| *A. cardenasii–A. ipaënsis* | A–B | 0.10 | 58.90 | 24.32 | 4.45 | 0.68 | 0.00 | 0.00 |
| *A. glandulifera–A. helodes* | D–A | 0.09 | 63.36 | 19.52 | 5.82 | 0.34 | 0.00 | 0.00 |
| *A. glandulifera–A. duranensis* | D–A | 0.07 | 75.00 | 9.25 | 4.45 | 1.37 | 0.00 | 0.00 |
| *A. glandulifera–A. trinitensis* | D–F | 0.08 | 62.67 | 14.38 | 5.14 | 1.37 | 0.00 | 0.00 |
| *A. glandulifera–A. ipaënsis* | D–B | 0.08 | 68.84 | 15.41 | 3.42 | 0.00 | 0.00 | 0.00 |
| *A. helodes–A. duranensis* | A–A | 0.08 | 59.24 | 20.21 | 2.40 | 1.37 | 0.00 | 0.00 |
| *A. helodes–A. trinitensis* | A–F | 0.08 | 48.63 | 25.00 | 1.71 | 0.34 | 0.00 | 0.00 |
| *A. helodes–A. ipaënsis* | A–B | 0.10 | 54.45 | 23.63 | 4.79 | 0.34 | 0.00 | 0.00 |
| *A. duranensis–A. trinitensis* | A–F | 0.07 | 57.53 | 18.49 | 1.71 | 1.71 | 0.00 | 0.34 |
| *A. duranensis–A. ipaënsis* | A–B | 0.08 | 63.01 | 18.49 | 4.45 | 0.34 | 0.00 | 0.00 |
| *A. trinitensis–A. ipaënsis* | F–B | 0.09 | 53.76 | 20.89 | 2.74 | 0.68 | 0.00 | 0.00 |

[a] Dxy, the average number of nucleotide substitutions per site between taxa (Nei 1987)

[b] Percentage of nucleotide positions falling into one of the six classes of transition stages proposed by Strachan et al. (1985)

(Electronic Supplemental Material 1). All sequences had a high percentage of A and T nucleotides, and the average value was highly conserved (approximately 75%) among species. The alignment analysis revealed the presence of the pentanucleotide motif CAAAA in all sequences, which occurred three times in the consensus sequence (nucleotide positions 184, 212, and 230). The intraspecific nucleotide diversity (π) of ATR-2 was low (Table 2) and ranged from 0.034 (*A. duranensis*) to 0.168 (*Arachis batizocoi*).

The interspecific sequence analysis (Table 3) showed that the average number of nucleotide substitutions per site between pairs of sequences isolated from different species (Dxy value) varied between 0.054 (*A. duranensis–Arachis cardenasii*) and 0.172 (*A. batizocoi–A. glandulifera*). These values were comparable with those observed in intraspecific comparisons. Statistical analysis of variation (AMOVA) showed that 80.92% of the total genetic variability derived from intraspecific variability, while the remaining 19.08% represented variation between the species. The phylogenetic tree using the NJ algorithm did not separate the clones of each species into particular clusters, and none of the groups recovered was statistically supported (Fig. 1).

The transition stages of the ATR-2 satDNA are shown in Table 3. In all pairwise comparisons, most of the nucleotide positions were classified into one of the six transitional stages proposed by Strachan et al. (1985). The majority of them were in stages 1 (from 43.49 to 75.00%) or 2 (from 9.25 to 25.00%), while the percentage of nucleotide positions in other stages was very low (from 0.00 to 10.27%). A similar pattern of transition stages was observed between any pair of species analyzed, regardless of their genome constitution.

### Chromosome localization of ATR-2 sequences

The chromosome distribution of ATR-2 in somatic chromosomes is shown in Figs. 2 and 3 and in Table 2. The hybridization signals of this satDNA showed a strong association with the centromeric DAPI$^+$ heterochromatin in most of the species analyzed. However, the species belonging to different genomes showed differences in the number, distribution, and size of the hybridization signals (Figs. 2, 3). Conspicuous ATR-2 signals were detected on most (14–16) of the centromeric bands of the A genome (*A. duranensis*, *Arachis helodes* and *A. cardenasii*) and K genome (*A. batizocoi*) species, while faint signals were observed in 14 centromeric bands of the F genome species (*A. trinitensis*). A completely different pattern of hybridization was found in *A. glandulifera* (D genome), in which only one pair of chromosomes (of seven pairs with centromeric heterochromatin) showed hybridization signals in the centromeric bands. In the B genome species (*Arachis ipaënsis*), no hybridization signals of ATR-2 were observed.

Few hybridization signals of ATR-2 were observed outside the centromeric bands, but all of them corresponded to DAPI$^+$ heterochromatic bands. Interstitial hybridization signals were observed in two chromosome pairs of *A. batizocoi* and in two pairs of *A. glandulifera*, while telomeric signals were observed in two pairs of *A. trinitensis*. In *A. batizocoi*, the interstitial signals were located in submetacentric chromosomes that also borne centromeric heterochromatic bands strongly hybridized with the ATR-2 probe. By contrast, the interstitial signals observed in *A. glandulifera* were located in two subtelocentric pairs, one of them with large centromeric bands of heterochromatin that did not hybridize with ATR-2. In *A. trinitensis*, one pair of telomeric signals was observed in chromosomes deprived of centromeric heterochromatin, while the other was located in the chromosomes having one of the largest centromeric bands hybridized with ATR-2. Differences in the signal intensities were also observed among the species belonging to different genomes. The A genome species showed the brightest hybridization signals of ATR-2 compared with those observed in the K and F genomes (Figs. 2, 3).

Within the A genome, although the number of FISH signals varied among species (14 in *A. duranensis* from the La Plata River Basin group, 16 in *A. helodes* from the Pantanal group, and 16 in *A. cardenasii* from the Chiquitano group), the hybridization pattern was conserved among them. In general, the size and intensity of the signals corresponded with the size of the heterochromatic bands, except for the 'A' chromosomes (after Husted 1936). This pair showed one of the smallest hybridization signals despite having the largest heterochromatic DAPI$^+$ bands.

### Genome representation

The ATR-2 genome abundance was estimated by dot-blot experiments using the probe derived from the AhyTR-2 c6 clone against known quantities of genomic DNAs of the seven species analyzed (Fig. 4). The chemiluminescent dots revealed that representatives of different genomes showed different amounts of ATR-2 sequences (Table 2). The highest genome representation of ATR-2 was found among the species of the A genome (from 1.20 to 1.52% of the haploid genome). The lowest abundance was observed in the B genome species (0.01% of the haploid genome). Between these values, the species showed different representations for this satDNA family, in agreement with the results observed in FISH experiments.

## Discussion

This work represents an initial study about the variability, evolution, genomic abundance, and chromosome distribution of satDNA sequences in species belonging to different genomes and karyotype groups of section *Arachis*. Here, a novel satDNA family (ATR-2) was isolated from the genomic DNA of seven diploid *Arachis* species. The isolated sequences were highly conserved and mainly distributed in the DAPI$^+$ pericentromeric heterochromatin of chromosomes. However, ATR-2 showed a differential representation among different species, following the predictions of the 'library' hypothesis (Ugarkovic and Plohl 2002;

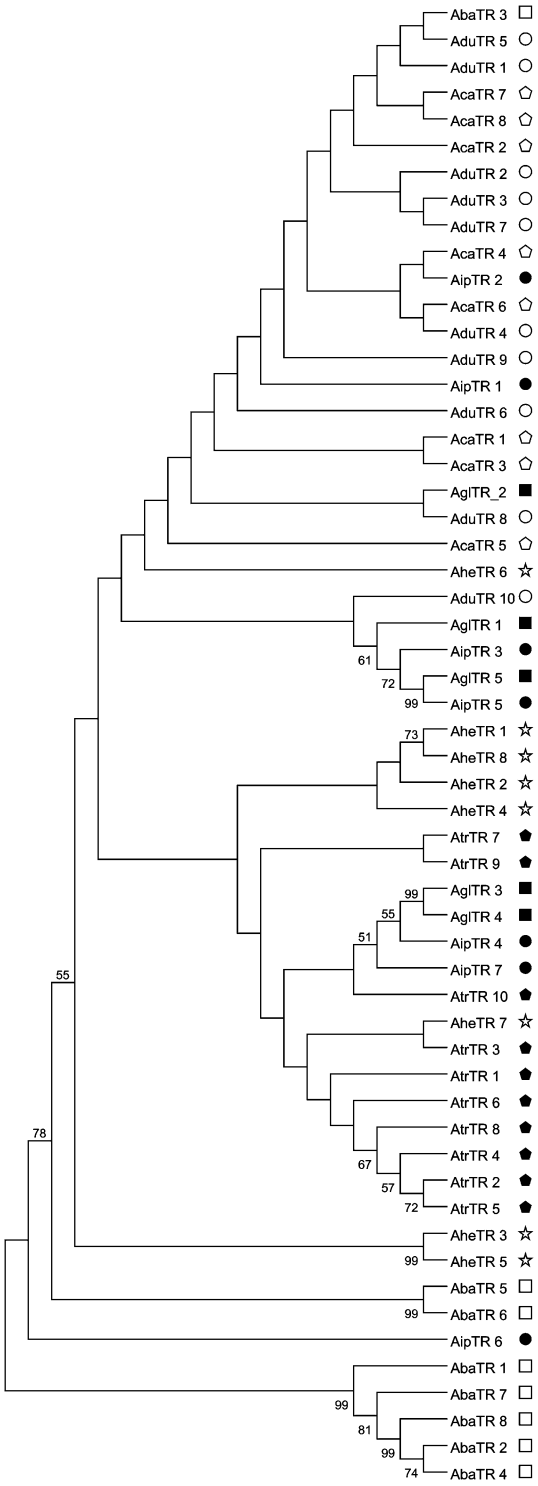


**Fig. 1** Dendrogram of ATR-2 sequences isolated from *Arachis* species. The NJ tree was constructed based on nucleotide sequences of ATR-2 monomeric variants isolated from wild diploid *Arachis* species belonging to different genomes and karyotype groups. The sequences isolated from the same species are indicated with the same symbol in the *right-hand column*: *A. cardenasii* (*open pentagon*), *A. duranensis* (*open circle*), *A. helodes* (*open star*), *A. batizocoi* (*open square*), *A. ipaënsis* (*closed circle*), *A. trinitensis* (*closed pentagon*), and *A. glandulifera* (*closed square*). The numbers adjacent to the branches indicate the bootstraps supporting a particular cluster

Plohl et al. 2008). This differential representation suggests that ATR-2 participated in the evolution and divergence of the heterochromatic fraction in diploid *Arachis* species.

## ATR-2 characteristics

Although the different satDNA families are highly heterogeneous in nucleotide composition, genome representation, and chromosome location, there are some particular features shared among them (Plohl et al. 2008). One of the typical features of plant satDNAs is the presence of the CAAAA motif, a pentanucleotide suspected to be involved in the breakage–reunion properties of these repeated sequences (Macas et al. 2002) and that can lead to sequence translocations and chromosomal rearrangements (Langdon et al. 2000). The presence of this motif at three conserved positions in the alignment of all ATR-2 clones isolated from *Arachis* confirms the satellite nature of these sequences.

Another common feature of satDNA sequences is the high content of A and T nucleotides (Flavell 1986; Kurbis et al. 1998; San Miguel and Bennetzen 1998; Macas et al. 2002; Urdampilleta et al. 2009). In this regard, the high percentage of A and T nucleotides (approximately 75%) observed in the ATR-2 family is one of the highest values recorded so far for plant satDNAs (Macas et al. 2000; Ansari et al. 2004; Mlinarec et al. 2009; Urdampilleta et al. 2009). It is not still clear why this base composition is prevalent in satellite sequences, but it has been related to the DNA bending properties required for highly packaged chromatin (Radic et al. 1987; Blattes et al. 2006).

In most plant species investigated, the length of the monomers is usually 150–180 or 300–360 bp (Sharma and Raina 2005). Consistently, the sequences of all the ATR-2 monomers isolated from *Arachis* were conserved in length, with an alignment length of 339 bp. This length has been correlated with the nucleosome periodicity and the packaging of tandemly organized repeat sequences. In this sense, the monomeric length of ATR-2 corresponds approximately to the length necessary for two turns of the DNA molecule around the histone core proteins (Henikoff et al. 2001; Macas et al. 2002; Plohl et al. 2008; Sharma and Raina 2005). Moreover, this length was associated with a requirement for interaction with specialized centromeric histones necessary for chromatin packaging (Zhang et al. 2013, Vittorazzi et al. 2014). The finding of ATR-2 sequences distributed mainly in the pericentromeric bands of DAPI$^+$ heterochromatin, together with the length of the monomers and the sequence characteristics, suggests a possible role of this satDNA family in heterochromatin packaging at the centromeres. However, further studies are necessary to properly understand the structural and functional roles of ATR-2 sequences in *Arachis* species.

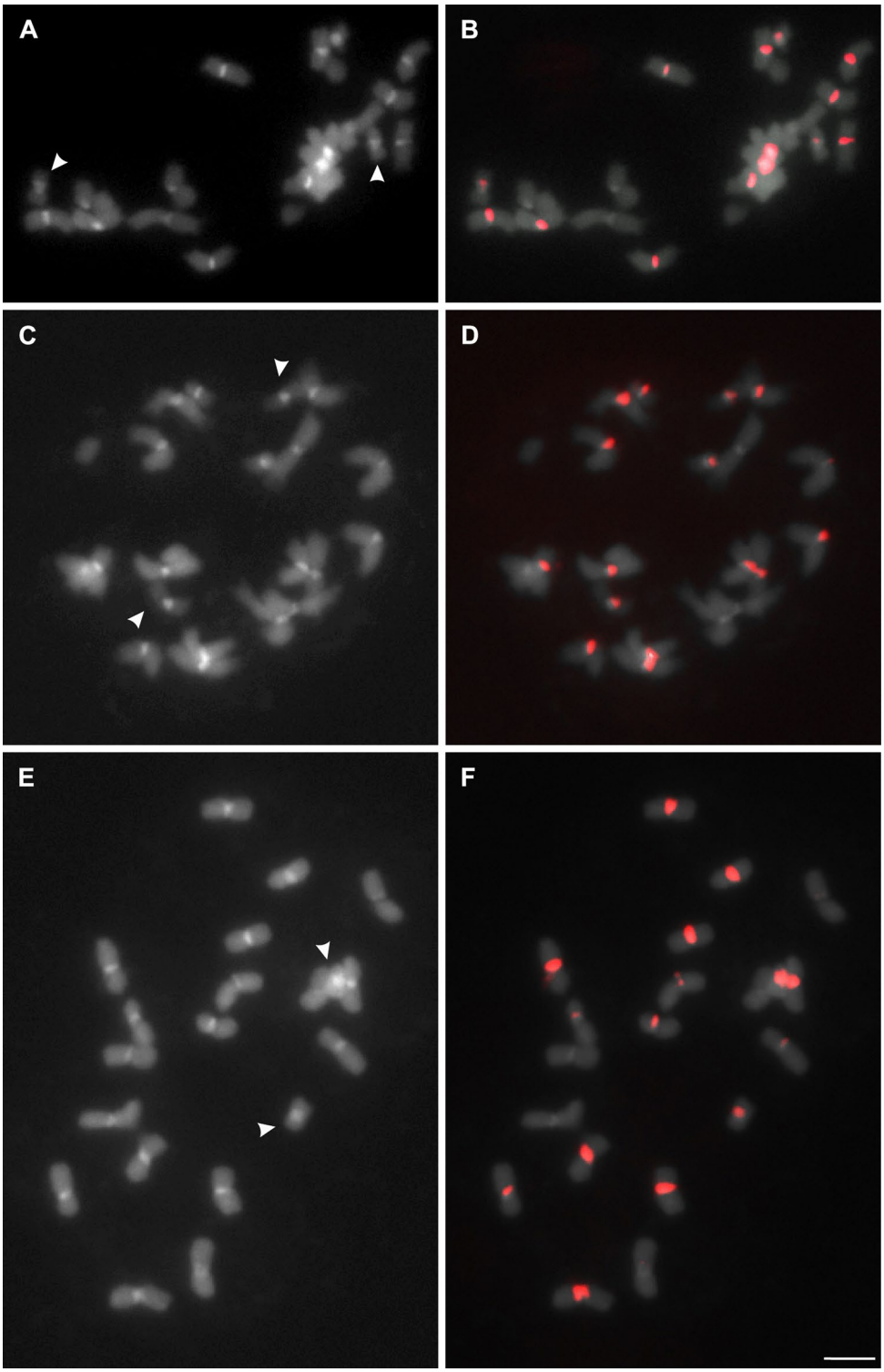


**Fig. 2** Chromosome distribution of ATR-2 sequences in diploid *Arachis* species with the A genome: *A. duranensis* (**a** and **b**), *A. cardenasii* (**c** and **d**), and *A. helodes* (**e** and **f**). **a**, **c,** and **e** Show the metaphases of the diploid species stained with DAPI (*in gray*). Pair A9 is indicated with *arrowheads*. **b**, **d,** and **f** Show the hybridization signals of the clone AhyTR-2 c6 (*in red*) located on the pericentromeric $DAPI^+$ bands of most chromosomes. *Scale bar* 3 µm (color figure online)

## High-sequence conservation of the library among genomes and lack of evidence for concerted evolution in ATR-2

SatDNAs are generally considered to be rapidly evolving (Plohl et al. 2012). However, some satDNAs show high-sequence conservation for extended evolutionary periods (Ugarkovic 2005; Robles et al. 2004; Meštrovic et al. 2006; Čížková et al. 2013; Melters et al. 2013). The overall variability profile of satDNA monomers in a group of species is a complex feature that depends on several factors, such as location, organization and copy number of repeats

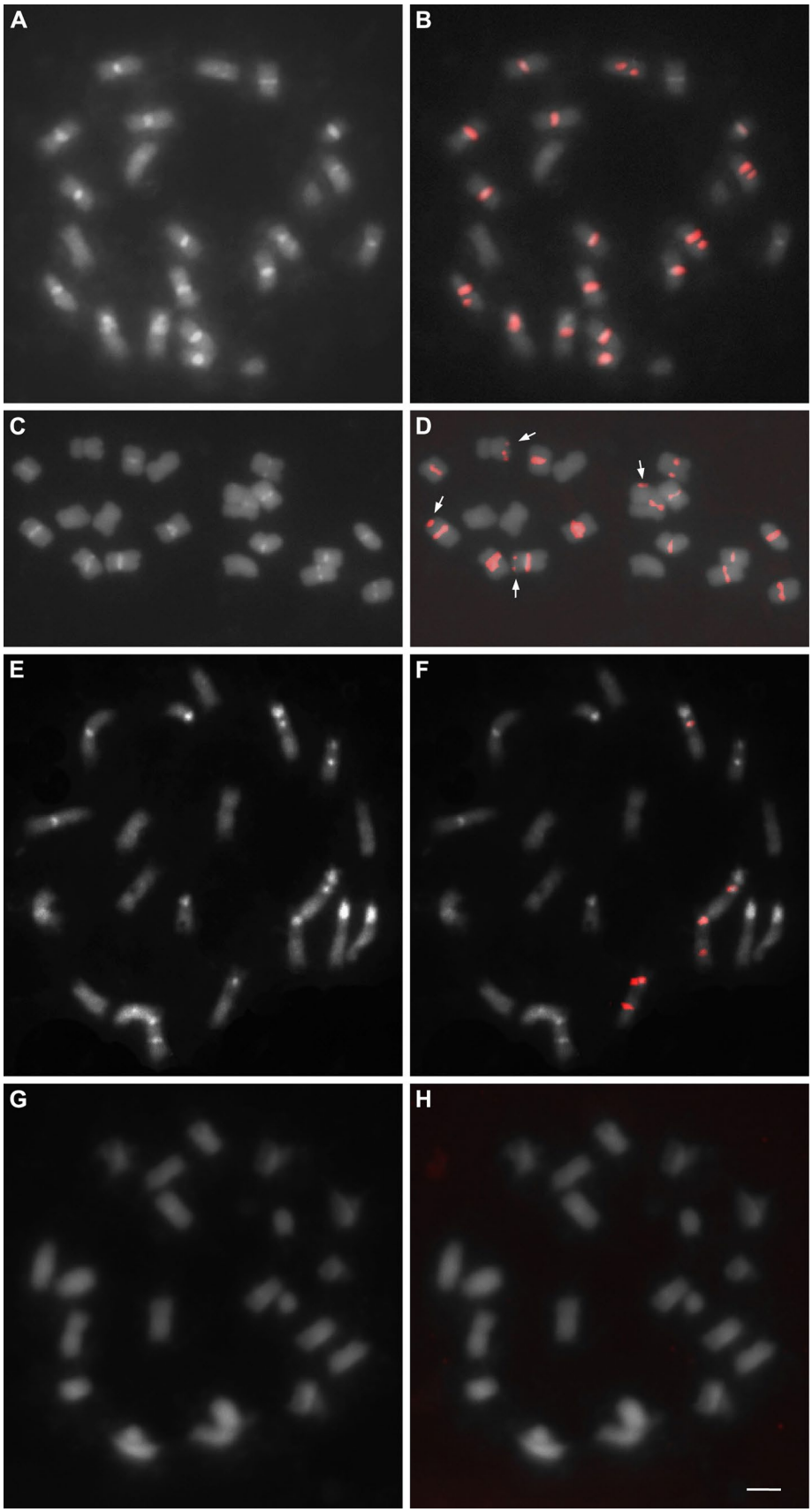


**Fig. 3** Chromosome distribution of ATR-2 sequences in *A. batizocoi* (**a** and **b**), *A. trinitensis* (**c** and **d**), *A. glandulifera* (**e** and **f**), and *A. ipaënsis* (**g** and **h**), belonging to the K, F, D, and B genomes, respectively. **a**, **c**, **e**, and **g** Show the metaphases of the diploid species stained with DAPI (*in gray*). **d**, **f,** and **h** Show the hybridization signals of the clone AhyTR-2 c6 (*in red*) located on the DAP $I^+$ bands of the chromosomes. Telomeric hybridization signals in *A. trinitensis* are indicated by *arrows*. *Scale bar* 3 µm (color figure online)

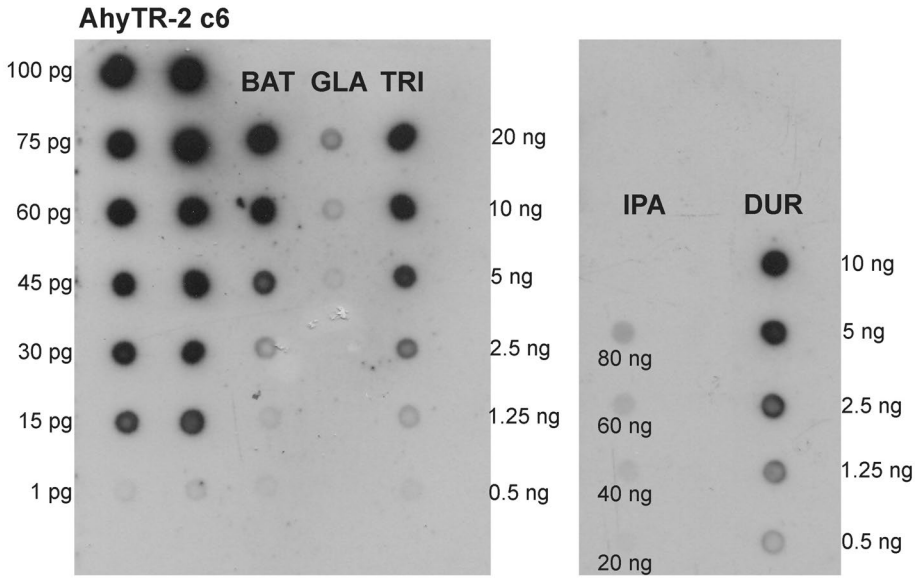


**Fig. 4** Estimation of the genome representation of ATR-2 sequences in different diploid species of section Arachis. **a** *Chemiluminograph dots* showing dilution series of the clone AhyTR-2 c6 isolated from *A. hypogaea* and the genomic DNAs from *Arachis* species representatives of different genome groups (DUR: *A. duranensis*, IPA: *A. ipaënsis*, TRI: *A. trinitensis*, BAT: *A. batizocoi*, and GLA: *A. glandulifera*) hybridized with a digoxigenin-labelled probe (AhyTR-2 c6). **b** *Calibration curve* based on signal intensities of the clone AhyTR-2 c6. The *curve* was constructed with the average values of the signal intensities from the individual *dots* of two dilution series

(Navajas-Pérez et al. 2009), divergence time (Pérez-Gutiérrez et al. 2012), biological factors (Luchetti et al. 2003, 2006; Robles et al. 2004; Suárez-Santiago et al. 2007), and functional constraints (Mravinac et al. 2005). In *Arachis* species, the facts that many monomeric variants of ATR-2 sequences were shared by different species (as observed in the distance tree) and that the genetic distances observed among the monomers isolated from different species were very low (similar or even lower than those observed among monomers of the same species) indicated that the molecular drive mechanisms proposed to operate on satDNAs were not strong enough to fix species-specific variants.

The analysis of the pattern of variation at any given nucleotide position of ATR-2 between two species (Strachan et al. 1985) was used to further evaluate the impact of concerted evolution. The similar pattern of transitional stages for almost all pairwise comparisons suggests that the evolution of ATR-2 sequences is not tightly associated with the divergence time. That is, the pattern of transitional stages of the pairwise *A. duranensis*–*A. ipaënsis* (divergence of the A and B genomes estimated to be approximately 2.88 mya, Moretzsohn et al. 2013) is similar to that of *A. duranensis*–*A. trinitensis* (divergence of the A and F genomes approximately 2.51 mya, Moretzsohn et al. 2013) and to that of *A. duranensis*–*A. batizocoi* (divergence of the A and K genomes approximately 2.31 mya, Moretzsohn et al. 2013). The transitional stages in the comparisons of the karyotype groups of the A genome were also very similar, although their divergence seems to have occurred more recently (Grabiele et al. 2012; Moretzsohn et al. 2013). Since the final outcome of concerted evolution is a higher repeat homogeneity within lineages than among them (Dover 1982; Rudd et al. 2006; Plohl et al. 2008), the conservation of the ATR-2 sequences suggests that their evolution in a concertedly way may be constrained.

Different hypotheses have been proposed to explain the conservation of particular satDNA families among taxa. One explanation considers that some monomer sequences might be positively selected over others either because of their functional potential or structural features (Plohl et al. 2008). Some examples supporting this hypothesis are the human α-satellite repeat present in the centromeres of old and new world monkey species (Horvath and Willard 2007) and the conserved centromeric satellite repeats isolated from several distantly related grass species (Lee et al. 2005; Zhang et al. 2013). In this regard and based on the localization of ATR-2 mainly at the centromeres, the sequence conservation of ATR-2 across taxa may be associated with selective pressures related to centromere function. However, the facts that the ATR-2 sequences had low abundance and were not detected in the centromeres of the B and D genomes (except in one pair of the latter) by FISH raise doubt on the general validity of this hypothesis.

An alternative hypothesis considers that variability can remain for long evolutionary periods by a reduced action of molecular mechanisms of non-reciprocal exchange (Plohl et al. 2008). As a consequence, the sequence variants of a satellite family persist in different species as a

library (Fry and Salser 1977; Ugarkovic and Plohl 2002; Mravinac et al. 2002; Meštrovic et al. 2006). Consequently, the lack of species-specific variants of ATR-2 in *Arachis* species suggests that this satDNA family was already present in the ancestral progenitor species and that the variability of this family remained more or less similar in all derived genomes. Therefore, it seems that the ATR-2 family was amplified/deleted without sequence fixation in a particular linage. Among the biological factors that constrain the effects of concerted evolution, the reproductive behavior of species is one of those with major effects (Luchetti et al. 2003). Since *Arachis* species are autogamous plants, new variants may persist in the population by direct inheritance in an ancestral-descendent pattern but may not be able to spread within the population and eventually be fixed. In these kinds of organisms without bisexual reproduction (such as autogamous and parthenogenetic organisms), similar variability is expected despite the level of taxonomic position (Mantovani 1998; Luchetti et al. 2003). Therefore, the lack of species-specific variants or polymorphism of ATR-2 in *Arachis* species may reflect the constraining effect of autogamy on concerted evolution.

### ATR-2 and heterochromatin in Arachis section

Our DAPI staining results showed a banding pattern similar to that reported previously for these *Arachis* species (Seijo et al. 2004; Robledo and Seijo 2008, 2010; Robledo et al. 2009). Accordingly, the number and distribution of bands and the total amount of heterochromatin were characteristics of the different genomes described for Arachis section. The chromosome distribution of ATR-2 sequences provides new markers for the identification of these genomes and further information to understand the nature of heterochromatin in *Arachis* species. The hybridization pattern of ATR-2 suggests that this satDNA is one of the major structural components of the centromeric heterochromatin in the A, K, and F genomes but not in the D genome.

Although ATR-2 was isolated from *A. ipaënsis*, the lack of FISH signals in its chromosome complement was expected, since the species of the B genome are deprived of conspicuous pericentromeric $DAPI^+$ bands. One of the variables affecting a FISH experiment is the minimum copy number of the target sequence at each individual chromosomal locus on mitotic chromosomes. In this sense, targets smaller than 2–5 kbp are hardly detected in plants (De Jong et al. 1999). Therefore, the absence of hybridization signals of ATR-2 sequences on mitotic chromosomes of *A. ipaënsis* may be a consequence of its low genome representation and their arrangements in tandem with few monomers.

### ATR-2, genomes, and species relationships in section Arachis

Given that one of the main features considered for distinguishing the genomes of section Arachis is the pattern of heterochromatin (Robledo and Seijo 2008, 2010; Robledo et al. 2009), the question of whether the heterochromatic bands observed in different genome types are homologous was unanswered. A detailed analysis of the FISH patterns of ATR-2 indicated that not all heterochromatic bands should be considered homologous and provided additional data for studying the relationship among the species and genomes of this section. The similar pattern (both in genome representation and in chromosome distribution of ATR-2 sequences) observed in species belonging to the three different karyotype groups of the A genome (*A. duranensis*, *A. cardenasii*, and *A. helodes*) showed a high homology in $DAPI^+$ heterochromatin, which supports their inclusion in the same genome.

*Arachis batizocoi* (K genome, after Robledo and Seijo 2010; formerly included within the B genome sensu lato) has a very similar pattern of heterochromatin to that observed in the A genome species. Here, we demonstrated that the heterochromatin of this species is mainly composed of ATR-2 satDNA, suggesting that the heterochromatic bands of the K genome may be considered homologous to those of the A genome species. This fact agrees with the close relationship of the A and K species as evidenced by phylogenetic analysis using NTS of 5S rDNA genes (Grabiele et al. 2012) and intron sequences from single-copy genes (Moretzsohn et al. 2013). However, it contrasts with data obtained in phylogenetic analyses based on ITS (Bechara et al. 2010; Friend et al. 2010), microsatellite (Moretzsohn et al. 2013) and plastid DNA (Friend et al. 2010; Grabiele et al. 2012), in which the K genome appeared more related to the B genome sensu *strict* (after Robledo and Seijo 2010). Moreover, in a recent GISH analysis of complex synthetic allotetraploids (hybrid between *A. hypogaea* 'AABB' and [*A. batizocoi* 'KK' × *A. stenosperma* 'AA']$^{4\times}$), it was suggested that the euchromatic fraction of the K genome chromosomes is more similar to that of the B genome than to that of the A genome chromosomes (Leal-Bertioli et al. 2015). The similar distribution of the heterochromatic bands mainly composed of ATR-2 in the K and A genomes adds a new source of evidence of the intermediate position of the K genome between the A and B genomes.

*Arachis trinitensis* (F genome, after Robledo and Seijo 2010; formerly included within the B genome sensu lato) also showed centromeric heterochromatin mainly composed of ATR-2. This fact revealed that the heterochromatic bands of the F genome may be considered homologous to those of the A and K genomes, although in most

phylogenetic analyses, the F genome appeared closely related to the B genome (Grabiele et al. 2012; Tallury et al. 2005).

According to phylogenetic (Tallury et al. 2005; Bechara et al. 2010; Grabiele et al. 2012) and cytogenetic (Stalker 1991; Fernandez and Krapovickas 1994) analyses, *A. glandulifera* (D genome, after Stalker 1991) is closely related to K genome species, particularly to *A. batizocoi*. However, the former species only showed one pair of centromeric bands with a strong hybridization of ATR-2. This fact suggests that if these two genomes were closely related, a rapid evolution of the satDNA fraction should have occurred, replacing most of the ATR-2 members by another satDNA family in almost all *A. glandulifera* chromosomes.

Several studies have revealed that qualitative and quantitative variations in the dispersed DNA fraction played an important role in the differentiation of the A and B genomes of *Arachis* (Nielen et al. 2010, 2012; Bertioli et al. 2013; Samoluk et al. 2015b). The results obtained here are the first to demonstrate that satDNA sequences have participated in the diversification of the heterochromatic fraction of the karyotypes of diploid species ($x = 10$) belonging to the different genomes described for section Arachis. Further studies comprising basal species of the genus *Arachis* will provide a broader view of evolutionary pathways followed by this satDNA.

#### Compliance with ethical standards

**Funding** This work was supported by the Agencia Nacional de Promoción Científica y Tecnológica (PICT 2012-1875 and PICTO 2011-0260) and the Consejo Nacional de Investigaciones Científicas y Técnicas (PIP 11220120100192CO), Argentina.

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

**Ethical approval** This article does not contain any studies with human participants or animals performed by any of the authors.

## References

Ansari HA, Ellison NW, Griffiths AG, Williams WM (2004) A lineage-specific centromeric satellite sequence in the genus *Trifolium*. Chromosome Res 12:357–367

Bachmann L, Sperlich D (1993) Gradual evolution of a specific satellite DNA family in *Drosophila ambigua*, *D. tristis* and *D. obscura*. Mol Biol Evol 10:647–659

Bechara MD, Moretzsohn MC, Palmieri DA, Monteiro JP, Bacci M Jr, Martins J Jr, Valls JFM, Lopes CR, Gimenes MA (2010) Phylogenetic relationships in genus *Arachis* based on ITS and 5.8S rDNA sequences. BMC Plant Biol 10:255

Bertioli DJ, Vidigal B, Nielen S, Ratnaparkhe MB, Lee T-H, Leal-Bertioli SCM, Kim C, Guimarães PM, Seijo G, Schwarzacher T, Paterson AH, Heslop-Harrison P, Araujo ACG (2013) The repetitive component of the A genome of peanut (*Arachis hypogaea*) and its role in remodelling intergenic sequence space since its evolutionary divergence from the B genome. Ann Bot 112:545–559

Biscotti MA, Olmo E, Heslop-Harrison JP (2015) Repetitive DNA in eukaryotic genomes. Chromosome Res 23:415–420

Blattes R, Monod C, Susbielle G, Cuvier O, Wu JH, Hsieh TS et al (2006) Displacement of D1, HP1 and topoisomerase II from satellite heterochromatin by a specific polyamide. EMBO J 25:2397–2408

Charlesworth B, Sniegowski P, Stephan W (1994) The evolutionary dynamics of repetitive DNA in eukaryotes. Nature 371:215–220

Čížková J, Hřibová E, Humplíková L, Christelová P, Suchánková P, Doležel J (2013) Molecular analysis and genomic organization of major DNA satellites in banana (*Musa* spp.). PLoS One 8:e54808

De Jong JH, Fransz P, Zabel P (1999) High resolution FISH in plants—techniques and applications. Trends Plant Sci 4:258–263

Dover G (1982) Molecular drive a cohesive mode of species evolution. Nature 299:111–117

Dover G (2002) Molecular drive. Trends Genet 18:587–589

Excoffier L, Smouse PE, Quattro JM (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. Genetics 131:479–491

Felsenstein J (1985) Confidence limits on phylogenies: an approach using the bootstrap. Evolution 39:783–791

Fernandez A, Krapovickas A (1994) Cromosomas y evolución en *Arachis* (Leguminosae). Bonplandia 8:187–220

Ferree PM, Prasad S (2012) How can satellite DNA divergence cause reproductive isolation? Let us count the chromosomal ways. Genet Res Int 2012:430136

Flavell RB (1986) Repetitive DNA and chromosome evolution in plants. Phil Trans R Soc Lond B 312:227–242

Friend SA, Quandt D, Tallury SP, Stalker HT, HiluK W (2010) Species, genomes, and section relationships in the genus *Arachis* (Fabaceae): a molecular phylogeny. Plant Syst Evol 290:185–199

Fry K, Salser W (1977) Nucleotide sequence of HS—a satellite DNA from kangaroo rat *Dipodomys ordii* and characterization of similar sequences in other rodents. Cell 12:1069–1084

Garrido-Ramos MA, de la Herran R, Ruiz-Rejon M, Ruiz- Rejon C (1999) A subtelomeric satellite DNA family isolated from the genome of the dioecious plant *Silene latifolia*. Genome 42:442–446

Gindullis F, Desel C, Galasso I, Schmidt T (2001) The large-scale organization of the centromeric region in *Beta* species. Genome Res 11:253–265

Glinka S, De Lorenzo D, Stephan W (2006) Evidence of gene conversion associated with a selective sweep in *Drosophila melanogaster*. Mol Biol Evol 23:1869–1878

Grabiele M, Chalup L, Robledo G, Seijo G (2012) Genetic and geographic origin of domesticated peanut as evidenced by 5S rDNA and chloroplast DNA sequences. Plant Syst Evol 298:1151–1165

Grattapaglia D, Sederoff R (1994) Genetic linkage Maps of *Eucalyptus grandis* and *Eucalyptus urophylla* using a pseudotestcross: mapping strategy and RAPD markers. Genetics 137:1121–1137

Gregory MP, Gregory WC (1979) Exotic germoplasm of *Arachis* L. interspecific hybrids. J Hered 70:185–193

Henikoff S, Ahmad K, Malik HS (2001) The centromere paradox: stable inheritance with rapidly evolving DNA. Science 293:1098–1102

Heslop-Harrison JS, Brandes A, Schwarzacher T (2003) Tandemly repeated DNA sequences and centromeric chromosomal regions of *Arabidopsis* species. Chromosome Res 11:241–253

Hizume M, Shibata F, Matsusaki Y, Garajova Z (2002) Chromosome identification and comparative karyotypic analyses of four *Pinus* species. Theor Appl Genet 105:491–497

Horvath JE, Willard HF (2007) Primate comparative genomics: lemur biology and evolution. Trends Genet 23:173–182

Hudakova S, Michalek W, Presting GG et al (2001) Sequence organization of barley centromeres. Nucleic Acids Res 29:5029–5035

Husted L (1936) Cytological studies of peanut, *Arachis*. II Chromosome number, morphology and behavior and their application to the origin of cultivated forms. Cytologia 7:396–423

Koukalova B, Moraes AP, Renny-Byfield S, Matyasek R, Leitch AR, Kovarik A (2010) Fall and rise of satellite repeats in allopolyploids of *Nicotiana* over c. 5 million years. New Phytol 186:148–160

Krapovickas A, Gregory WC (1994) Taxonomía del género *Arachis* (Leguminosae). Bonplandia 8:1–186

Kuhn GCS, Sene F, Moreira-Filho O, Schwarzacher T, Heslop-Harrison JS (2008) Sequence analysis, chromosomal distribution and long-range organization show that rapid turnover of new and old pBuM satellite DNA repeats leads to different patterns of variation in seven species of the *Drosophila buzzatii* cluster. Chromosome Res 16:307–324

Kurbis S, Schimdt T, Heslop Harrison JS (1998) Repetitive DNA elements as major componenet of plant genomes. Ann Bot 82:45–55

Langdon T, Seago C, Jones RN, Ougham H, Thomas H, Forster JW, Jenkins G (2000) De Novo evolution of satellite DNA on the rye B chromosome. Genetics 154:869–884

Leal-Bertioli SC, Santos SP, Dantas KM et al (2015) *Arachis batizocoi*: a study of its relationship to cultivated peanut (*A. hypogaea*) and its potential for introgression of wild genes into the peanut crop using induced allotetraploids. Ann Bot 115:237–249

Lee HR, Zhang WL, Langdon T et al (2005) Chromatin immunoprecipitation cloning reveals rapid evolutionary patterns of centromeric DNA in *Oryza* species. Proc Natl Acad Sci USA 102:11793–11798

Lim KY, Matyasek R, Lichtenstein CP, Leitch AR (2000) Molecular cytogenetic analyses and phylogenetic studies in the *Nicotiana* section Tomentosae. Chromosoma 109:245–258

Luchetti A, Cesari M, Carrara G, Cavicchi S, Passamonti M, Scali V, Mantovani B (2003) Unisexuality and molecular drive: bag320 sequence diversity in *Bacillus* taxa (Insecta Phasmatodea). J Mol Evol 56:587–596

Luchetti A, Marini M, Mantovani B (2006) Non-concerted evolution of the RET76 satellite DNA family in *Reticulitermes* taxa (Insecta, Isoptera). Genetica 128:123–132

Macas J, Pozrkovaa D, Navratilova A, Nouzova M, 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, Meszaros T, Nouzova M (2002) PlantSat: a specialized database for plant satellite repeats. Bioinformatics 18:28–35

Mantovani B (1998) Satellite sequence turnover in parthenogenetic systems: the apomictic triploid hybrid *Bacillus lynceorum* (Insecta, Phasmatodea). Mol Biol Evol 15:1288–1297

Melters DP, Bradnam KR, Young HA et al (2013) Comparative analysis of tandem repeats from hundreds of species reveals unique insights into centromere evolution. Genome Biol 14:R10

Menzel G, Dechyeva D, Wenke T, Holtgräwe D, Weisshaar B, Schmidt T (2008) Diversity of a complex centromeric satellite and molecular characterization of dispersed sequence families in sugar beet (*Beta vulgaris*). Ann Bot 102:521–530

Meštrovic N, Castagnone-Sereno P, Plohl M (2006) High conservation of the differentially amplified MPA2 satellite DNA family in parthenogenetic root-knot nematodes. Gene 376:260–267

Mlinarec J, Chester M, Siljak-Yakovlev S, Papeš D, Leitch AR, Besendorfer V (2009) Molecular structure and chromosome distribution of three repetitive DNA families in *Anemone hortensis* L. (Ranunculaceae). Chromosome Res 17:331–346

Moretzsohn MC, Gouvea EG, Inglis PW, Leal-Bertioli SCM, Valls JFM, Bertioli DJ (2013) A study of the relationships of cultivated peanut (*Arachis hypogaea*) and its most closely related wild species using intron sequences and microsatellite markers. Ann Bot 111:113–126

Mravinac B, Plohl M, Mestrovic N, Ugarkovic D (2002) Sequence of PRAT satellite DNA 'frozen' in some Coleopteran species. J Mol Evol 54:774–783

Mravinac B, Plohl M, Ugarkovic D (2005) Preservation and high sequence conservation of satellite DNA suggest functional constraints. J Mol Evol 61:542–550

Mukai Y, Friebe B, Gill BS (1992) Comparison of C-banding patterns and in situ hybridization sites using highly repetitive and total genomic rye DNA probes of 'Imperial' rye chromosomes added to 'Chinese Spring' wheat. Jpn J Genet 67:71–83

Navajas-Pérez R, Rubio-Escudero C, Aznarte JL, Rejón MR, Garrido-Ramos MA (2007) SatDNA Analyzer: a computing tool for satellite-DNA evolutionary analysis. Bioinformatics 23:767–768

Navajas-Pérez R, Del Bosque MQ, Garrido-Ramos MA (2009) Effect of location, organization, and repeat-copy number in satellite-DNA evolution. Mol Genet Genom 282:395–406

Nei M (1987) Molecular evolutionary genetics. Columbia University Press, New York

Nielen S, Campos-Fonseca F, Leal-Bertioli S, Guimaraes P, Seijo G, Town C, Arrial R, Bertioli D (2010) FIDEL—a retrovirus-like retrotransposon and its distinct evolutionary histories in the A and B-genome components of cultivated peanut. Chromosome Res 18:227–246

Nielen S, Vidigal B, Leal-Bertioli S, Ratnaparkhe M, Paterson A, Garsmeur O, D'Hont A, Guimaraes P, Bertioli D (2012) Matita, new retroelement from peanut: characterization and evolutionary context in the light of the Arachis A-B genome divergence. Mol Genet Genomics 287:21–38

Pérez-Gutiérrez MA, Suárez-Santiago VN, López-Flores I, Romero AT, Garrido-Ramos MA (2012) Concerted evolution of satellite DNA in *Sarcocapnos*: a matter of time. Plant Mol Biol 78:19–29

Pich U, Fritsch R, Schubert I (1996) Closely related *Allium* species (Alliaceae) share a very similar satellite sequence. Plant Syst Evol 202:255–264

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

Plohl M, Petrovic V, Luchetti A, Ricci A, Satovic E, Passamonti M, Mantovani B (2010) Long-term conservation vs high sequence divergence: the case of an extraordinarily old satellite DNA in bivalve mollusks. Heredity 104:543–551

Plohl M, Meštrović N, Mravinac B (2012) Satellite DNA evolution. Genome Dyn 7:126–152

Radic MZ, Lundgren K, Hamkalo BA (1987) Curvature of mouse satellite DNA and condensation of heterochromatin. Cell 50:1101–1108

Robledo G, Seijo JG (2008) Characterization of *Arachis* D genome using physical mapping of heterochromatic regions and rDNA loci by FISH. Genet Mol Biol 31:717–724

Robledo G, Seijo G (2010) Species relationships among the wild B genome of *Arachis* species (section Arachis) based on FISH mapping of rDNA loci and heterochromatin detection: a new proposal for genome arrangement. Theor Appl Genet 121:1033–1046

Robledo G, Lavia GI, Seijo G (2009) Species relations among wild *Arachis* species with the A genome as revealed by FISH mapping

of rDNA loci and heterochromatin detection. Theor Appl Genet 118:1295–1307

Robles F, de la Herran R, Ludwig A, Ruiz-Rejon C, Ruiz-Rejon M, Garrido-Ramos MA (2004) Evolution of ancient satellite DNAs in sturgeon genomes. Gene 338:133–142

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

Rudd MK, Wray GA, Willard HF (2006) The evolutionary dynamics of alpha-satellite. Genome Res 16:88–96

Saitou N, Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol 4:406–425

Samoluk SS, Chalup L, Robledo G, Seijo JG (2015a) Genome sizes in diploid and allopolyploid *Arachis* L. species (section Arachis). Genet Res Crop Evol 62:747–763

Samoluk SS, Robledo G, Podio M, Chalup L, Ortiz JPA, Pessino SC, Seijo JG (2015b) First insight into divergence, representation and chromosome distribution of reverse transcriptase fragments from L1 retrotransposons in peanut and wild relative species. Genetica 143:113–125

San Miguel P, Bennetzen JL (1998) Evidence that a recent increase in maize genome size was caused by the massive amplification of intergene retrotransposons. Ann Bot 82:37–44

Schmidt T, Heslop-Harrison JS (1998) Genomes, genes and junk: the large-scale organization of plant chromosomes. Trends Plant Sci 3:195–199

Schneider S, Roessli D, Excoffier L (2000) ARLEQUIN: a software for population genetics data analysis, version 2.000. Genetics and Biometry Laboratory, University of Geneva, Switzerland

Seijo G, Lavia GI, Fernández A, Krapovickas A, Ducasse D, Moscone EA (2004) Physical mapping of 5S and 18S-25S rRNA genes evidences that *Arachis duranensis* and *A. ipaensis* are the wild diploid species involved in the origin of *A. hypogaea* (Leguminosae). Am J Bot 91:2293–2303

Seijo G, Lavia GI, Fernández A, Krapovickas A, Ducasse D, Bertioli DJ, Moscone EA (2007) Genomic relationships between the cultivated peanut (*Arachis hypogaea*- Leguminosae) and its close relatives revealed by double GISH. Am J Bot 94:1963–1971

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

Silvestri MC, Ortiz AM, Lavia GI (2014) rDNA loci and heterochromatin positions support a distinct genome type for '$x = 9$ species' of section Arachis (*Arachis*, Leguminosae). Plant Syst Evol 301:555–562

Smartt J, Gregory WC, Gregory MP (1978) The genomes of *Arachis hypogaea*. 1. Cytogenetic studies of putative genome donors. Euphytica 27:665–675

Stalker HT (1991) A new species-section *Arachis* of peanuts with D genome. Am J Bot 78:630–637

Stalker HT, Dhesi JS, Parry DC, Hahn JH (1991) Cytological and interfertility relationships of *Arachis* section Arachis. Amer J Bot 78:238–246

Stephan W (1986) Recombination and the evolution of satellite DNA. Genet Res 47:167–174

Strachan T, Webb D, Dover G (1985) Transition stages of molecular drive in multiple-copy DNA families in *Drosophila*. EMBO J 4:1701–1708

Suárez-Santiago VN, Blanca G, Ruiz-Rejón M, Garrido-Ramos MA (2007) Satellite-DNA evolutionary patterns under a complex evolutionary scenario: the case of *Acrolophus* subgroup (*Centaurea* L., Compositae) from the western Mediterranean. Gene 404:80–92

Tallury SP, Hilu KW, Milla SR, Friend SA, Alsaghir M, Stalker HT, Quandt D (2005) Genomic affinities in *Arachis* section Arachis (Fabaceae): molecular and cytogenetic evidence. Theor Appl Genet 111:1229–1237

Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S (2011) MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol Biol Evol 28:2731–2739

Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice. Nucleic Acids Res 22:4673–4680

Ugarkovic D (2005) Functional elements residing within satellite DNAs. EMBO Rep 6:1035–1039

Ugarkovic D, Plohl M (2002) Variation in satellite DNA profiles-causes and effects. EMBO J 2:5955–5959

Urdampilleta JD, de Souza AP, Schneider DR, Vanzela AL, Ferrucci MS, Martins ER (2009) Molecular and cytogenetic characterization of an AT-rich satellite DNA family in *Urvillea chacoensis* Hunz. (Paullinieae, Sapindaceae). Genetica 136:171–177

Valls JFM, Simpson CE (2005) New species of *Arachis* (Leguminosae) from Brazil, Paraguay and Bolivia. Bonplandia 14:35–64

Vittorazzi SE, Lourenço LB, Recco-Pimentel SM (2014) Long-time evolution and highly dynamic satellite DNA in leptodactylid and hylodid frogs. BMC Genet 15:1–11

Zhang L, Xu C, Yu W (2012) Cloning and characterization of chromosomal markers from a Cot-1 library of peanut (*Arachis hypogaea* L.). Cytogenet Genome Res 137:31–41

Zhang T, Talbert PB, Zhang W, Wu Y, Yang Z, Henikoff JG et al (2013) The CentO satellite confers translational and rotational phasing on cenH3 nucleosomes in rice centromeres. Proc Natl Acad Sci 110:4875–4883