TITLE CHAPTER: CYTOLOGICAL FEATURES OF PEANUT GENOME

AUTHORS: Guillermo Seijo[1,2], María C. Silvestri[1], Sergio S. Samoluk[1], Laura Chalup[1], Alejandra V. García[1,2], Alejandra M. Ortiz[1,2], Germán Robledo[1,2], Graciela I. Lavia[1,2]

[1]Instituto de Botánica del Nordeste (CONICET-UNNE, Fac. Cs. Agrarias), Sargento Cabral 2131, C.C. 209, 3400 Corrientes, Argentina.
[2]Facultad de Ciencias Exactas y Naturales y Agrimensura, Universidad Nacional del Nordeste, Av. Libertad 5460, 3400 Corrientes, Argentina

E-mails

Guillermo Seijo	seijo@agr.unne.edu.ar
Celeste Silvestri	celestesilvestri@gmail.com
Sergio S. Samoluk	samocarp@hotmail.com
Laura Chalup	laurachalup@gmail.com
Alejandra V. García	alegarcia_89@hotmail.com
Alejandra M. Ortiz	ortizalejandr@gmail.com
Germán Robledo	germanrobledo71@gmail.com
Graciela I. Lavia	graciela.lavia@yahoo.com.ar

Corresponding author: Graciela I. Lavia graciela.lavia@yahoo.com.ar

Telephone number of all authors: +54 379 4427589

General introduction

The genus *Arachis* is composed of 82 species (Krapovickas and Gregory 1994; Valls and Simpson 2005; Valls et al 2013; Santana and Valls 2015) distributed within a large region of South America, which extends from the eastern foothills of the Andes Mountains in Bolivia and northern Argentina to the Atlantic coast in Brazil, and from the southern limit of the Amazonian rainforest towards the northern coast of La Plata River in Uruguay.

Based on morphology, cross-compatibility, viability of the hybrids, geographic distribution, and cytogenetics, the *Arachis* species have been arranged in nine taxonomic sections: *Trierectoides*, *Erectoides, Procumbentes, Rhizomatosae, Heteranthae, Caulorrhizae, Extranervosae*, *Triseminatae* and *Arachis* (Krapovickas and Gregory 1994; Fernández and Krapovickas 1994; Lavia 1996; Valls and Simpson 2005). Cross compatibility, karyotypic and meiotic analysis also allowed the identification and description of six different genomes within the section *Arachis*: namely A, B, D, F, K and G (Smartt et al 1978; Stalker 1991; Robledo and Seijo 2008, 2010; Robledo et al 2009; Silvestri et al 2015). The genomic constitution of the remaining species of the genus, in the absence of comprehensive cytogenetic and molecular analysis, is less precise and have been traditionally assigned on the basis of the subgeneric divisions, that is: Am (*Heteranthae*), C (*Caulorrhizae*), E (*Trierectoides*, *Erectoides* and *Procumbentes*), Ex (*Extranervosae*), T (*Triseminatae*) and R (*Rhizomatosae*) (Smartt and Stalker 1982).

Classical and modern molecular cytogenetics revealed a huge variability within and among species of different sections. These studies provided important information about the complexity of the genome of peanut, but they were also very useful to unravel the taxonomy of the genus and to establish relationships among the wild species with the cultivated peanut. Here we present an update of the cytological information on *Arachis* species, and some examples in which the use of chromosome markers were decisive to understand critical and long lasting problems in the genus.

General chromosome features

The available information evidences that the genus *Arachis* is dibasic, with a great predominance of the species with x=10 and only four species with x=9 (Fernández and

Krapovickas 1994; Lavia 1996, 1998; Peñaloza et al 1996; Peñaloza and Valls 2005). Most of the species of the genus are diploid except the cultivated peanut and *A. monticola* (the most probable wild tetraploid ancestor of cultigen) from section Arachis, three species of section *Rhizomatosae* and one triploid accession of *A. pintoi* from the *Caulorrhizae* section.

Chromosome identification in peanut begun with Husted (1933), who distinguished one pair of chromosomes that borne a secondary constriction (B chromosomes) and another one that are conspicuously smaller than any of the other of the complement (A chromosomes). The occurrence of only a single pair of these chromosomes was considered an indicative of cytological differentiation between the two genomes, A genome and B genome, present in peanut (Smartt and Gregory 1967). The karyotypes of peanut and *A. monticola* are highly symmetric with a predominance of metacentric chromosomes. The most common haploid karyotype formulae are 19 m + 1 sm or 18 m + 2 sm and the average size of chromosomes is 1.88 µm (reviewed in Fernández and Krapovickas 1994; Lavia and Fernández 2004; Seijo et al 2004). Only one (very rarely two) pair of chromosomes with satellites (SAT chromosomes) has been usually distinguished in all the varieties of peanut and in *A. monticola*. In all cases the satellites have usually been observed far from the proximal segment of the chromosome arm in prometaphases. Different types of SAT chromosomes (type 3, 5 and 6) have been observed among the land races according to the Fernández and Krapovickas (1994) classification (Lavia and Fernández 2004). In general, the chromosomes belonging to the A genome are indistinguishable from those of the B genome by classical techniques, except the A chromosomes.

The three tetraploid species of section *Rhizomatosae* have their complements mainly constituted by metacentric (m) chromosomes of similar size and SAT chromosomes type

3A. Among them, *A. pseudovillosa* has a karyotype formula of 20 m, whereas both varieties of *A. glabrata* and *A. nitida* have 19 m + 1 sm (Fernández and Krapovickas 1994, Peñaloza and Valls 2005, Ortiz et al 2014).

Among the diploid species, only those from the sections Arachis, Caulorrhizae, Heteranthae and Rhizomatosae were comprehensively studied. Within species with 2*n*=20 of section Arachis, 15 are characterised by the presence of A chromosomes and have symmetric karyotypes mainly composed of metacentric chromosomes (Fernández and Krapovickas 1994; Lavia 1996, 2000; Peñaloza and Valls 2005; Robledo et al 2009). The most frequent haploid karyotype formula is 9 m + 1 sm. The remaining species with 2*n*=20 (genomes B, F and K, after Robledo and Seijo 2010) have symmetric karyotypes but lack the A chromosomes (Smartt et al 1978; Smart and Stalker 1982; Fernández and Krapovickas 1994; Lavia 1996; Peñaloza and Valls 2005). Morphological analysis of the karyotypes showed that all the studied species have their complements completely or mainly composed of metacentric chromosomes; the most frequent haploid karyotype formula is 10m. However, one to four pairs of submetacentric chromosomes were observed in different species (Robledo and Seijo 2010). *Arachis glandulifera* (D genome) is the only species (2*n*=20) that has an asymmetric karyotype with a karyotype formula of 4 m + 2 sm + 4 st (Stalker 1991; Robledo and Seijo 2008). Diploid species with 2*n*=18 (G genome, after Silvestri et al 2015) have symmetric karyotypes and lack the A chromosomes (Lavia 1996, 1998; Peñaloza and Valls 1997). Almost all species of the section have only one pair of secondary constrictions localized on the long arms of pair 10. The exceptions are *A. krapovickasii* and *A. gregoryi*, which have two pairs of secondary constrictions (Robledo and Seijo 2008, 2010; Robledo et al 2009). The species with 2*n*=18 have one pair of secondary constrictions in the long arms of pair 9 (Silvestri et al 2015).

Species of section *Heteranthae* present karyotype formula 9 m + 1 sm and type 2 of the SAT chromosomes (*A. dardani*, *A. pusilla*, and *A. interrupta*), or 8 m + 2 sm and with SAT chromosomes type 10 (*A. sylvestris* and *A. giacomettii*) (Silva et al 2010). Within section *Caulorrhizae* the diploid *A. pintoi* shows 9 m + 1 sm and SAT chromosomes type 2, and *A. repens* equal karyotype formula, but the SAT chromosomes are type 3 (Pucciariello et al 2013). The only diploid so far known within of section *Rhizomatosae* has karyotype formula of 10 m and SAT chromosomes type 8 (Ortiz et al 2013).

Heterochromatin

Chromatin has several classes of proteins complexed with its DNA which are responsible for the major levels of its compaction. Of these levels of chromatin compaction, one of the most conspicuous is the fraction that remains highly condensed throughout the cell cycle, usually defined as constitutive heterochromatin. Although heterochromatin is largely the gene-poor fraction, nowadays it is consider that it plays important roles in the modulation of gene expression, in chromosome structure and in speciation and evolution of eukaryotes (review in Grewal and Moazed 2003). Accordingly, the analyses of heterochromatin in *Arachis* species revealed extremely different patterns among species and provided a huge amount of chromosome and genome markers (Seijo et al 2004; Robledo and Seijo 2008, 2010; Robledo et al 2009; Custodio et al 2013; Silvestri et al 2015).

In the complements of peanut and *A. monticola*, half of the chromosomes—those belonging to the A genome—have centromeric DAPI^{+} bands (AT-rich), while the remainder (those of the B genome) lack detectable centromeric heterochromatin. The bands differ in size, with the most conspicuous being those borne by the A chromosomes (A9

pair). All the *A. hypogaea* varieties and *A. monticola* have a similar distribution and amount of heterochromatin, which accounts for about 7% of the karyotype length (Seijo et al 2004).

Among the diploid species of section Arachis, one of the most variable characters is the amount and distribution of heterochromatin. The A genome species are characterized by having conspicuous DAPI$^+$ centromeric bands in all (or nine) of the chromosome pairs and a total amount of heterochromatin of around 12% of the karyotype length. The size of the bands is around the 10% of the chromosome length, except in the A9 pair, in which the band size ranged from 25 - 46%, and the submetacentric A10 pair, in which the band size ranged from 14.50 - 20.12% among species (Robledo et al 2009).

The pattern of DAPI$^+$ heterochromatin in K genome species is characterized by conspicuous centromeric bands in nine chromosome pairs (Robledo and Seijo 2010). The total amount of centromeric heterochromatin per haploid complement is around 12%. Pairs 9 and 10 have the largest bands (16 - 20%), and the remaining chromosomes have the smallest ones (around 10%) in relation to the chromosome length.

The karyotypes of the F genome species characteristically had small and faint DAPI$^+$ centromeric bands in only seven or eight chromosome pairs (Robledo and Seijo 2010). All the bands had a similar size (around 8% of the chromosome length) and the total amount of heterochromatin per haploid complement was almost half (around 6.5%) of that observed in the species of the A and K genomes.

The B genome species had karyotypes completely devoid of centromeric bands. However, few species had one small interstitial or distal band in the short arms of pair 3, which usually covers less than 10% of the chromosome length (Robledo and Seijo 2010).

The three *x*=9 species of section *Arachis* have centromeric DAPI$^+$ bands with the same brightness, position, and size in all chromosome pairs, except *A. palustris*, which lacks these bands in one chromosome pair (Silvestri et al 2015).

Aside the species of section Arachis, karyotypes with centromeric DAPI$^+$ bands in all (or most) of the chromosome pairs are found in all the sections so far analysed –*Erectoides*, *Heteranthae*, *Procumbentes*, *Triseminatae*, *Caulorrhizae* and *Rhizomatosae* (Raina and Mukai 1999; Lavia et al 2011; Pucciarello et al 2013; Ortiz et al unpublished), constituting probably a primitive character for the genus. The karyotypes without DAPI$^+$ bands may have been originated independently at least three times in the evolution of the genus *Arachis* since it appeared in three phylogenetically distant groups (Ortiz et al unpublished).

In section *Arachis*, *Caulorrhizae* and *Heteranthae* the CMA$^+$/DAPI$^-$ (GC-rich) heterochromatin are restricted to the secondary constrictions of SAT chromosomes (see below). Exception are *A. dardani* and *A. giacomettii* that present a centromeric CMA$^+$ band in one additional chromosome and *A. pusilla* that has centromeric GC rich heterochromatin in the centromeres of all the chromosomes (Silva et al 2010).

rDNA loci

Localization of the 5S and 18S-26S rRNA genes on the chromosomes by fluorescent in situ hybridization (FISH) was initially applied to a small set of species from different sections of *Arachis* (Raina and Mukai 1999) revealing their usefulness for the characterization of the species. Chromosome mapping of these loci was later used to analyze in detail the karyotypes of all the species included in section *Arachis* (Seijo et al 2004; Robledo and Seijo 2008; Robledo et al 2009, 2010; Robledo and Seijo 2010, Lavia et al 2011; Custodio et al 2013; Silvestri et al 2015).

Physical mapping of the rDNA loci in the six botanical varieties of *A. hypogaea* and in *A. monticola* revealed two pairs of 5S and five pairs of 18S–26S rDNA sites. In both species, the 5S loci are proximally located in short arms (pairs A3 and B3), while the 18S–26S rDNA loci are proximally (pairs A2, A10, B3 and B10) or subterminally placed (B7). One 5S site is syntenic to an 18S–26S site (B3). The high degree of homeology detected between *A. monticola* and *A. hypogaea* strongly evidences that they are very close related taxa. The mapping of the rDNA loci, together with the heterochromatin analysis, provided the first chromosome map for peanut (Seijo et al 2004).

Concerning diploid species of section Arachis, those within the A genome, have only one pair of interstitial (or rarely proximal) 5S rDNA loci located on the A3 pair. However, the number, size and chromosomal localization of the 18S–26S rDNA loci are variable among the species (Robledo et al 2009). The number of these gene clusters ranges from two to four pairs of variable size. In general, the largest loci are located in A10 pair, the A2 pair carries loci of intermediate size, while the smallest and faintest signals (in the cases that the species have more than two loci) are borne by the A7 and A4 chromosomes. According to the pattern of rDNA loci and heterochromatic bands, the A genome species have been further arranged into three karyotype groups (Robledo et al 2009): Chiquitano, Pantanal and La Plata river basin.

Arachis glandulifera (D genome) has only one 5S rDNA loci subterminally located in long arms of D5 and ten loci for the 18S–26S rDNA located in different positions and arms in five chromosome pairs. The markers so far identified were enough to the precise identification of all the chromosome pairs of the karyotype and the construction of the first wholly-resolved idiogram for an *Arachis* species (Robledo and Seijo 2008).

All the species without A chromosomes and symmetric karyotypes had one pair of 5S rDNA loci localized in proximal or interstitial position on the short arms of the metacentric pair 3. Exceptions are the species of the K genome that had two additional pairs of loci located in K8 and K10 pairs. The number of 18S–26S rDNA loci ranged from two (in *A. gregoryi* and *A. trinitensis*) to four pairs (in *A. magna* and *A. valida*). Most of them were located in pericentromeric or interstitial position on the long arms. In most species, the largest and brightest 18S–26S rDNA loci corresponded to the clusters located in the secondary constrictions of the SAT chromosomes (pair 10), while the remaining ones were small and pale. In same species one 18S–26S rDNA and one 5S rDNA loci are localized on the same chromosome. In the species of the K genome, these loci colocalized on the long arms of pair 10, while in *A. benensis* and *A. magna* they colocalized on the short arms of pair 3. In *A. ipaënsis*, both rDNA loci mapped to pair 3, but on different arms (Robledo and Seijo 2010).

The three species with x=9 have only one pair of 18S–26S rDNA sites in the proximal region of the long arm of the SAT chromosomes (pair 9), and one pair of 5S rDNA sites in the short arm of the pair 6 (Silvestri et al 2015).

Repetitive sequences

Plant genomes are composed of single copy sequences, with one or few copies (gene sequences), and repetitive sequences, with a higher copy number. The latter can be found as dispersed repeated sequences (transposons) and tandemly repeated ones (satellite DNA) (Schmidt and Heslop- Harrison 1998). One of the most important features of the repetitive genome component is its rapid evolution both at sequence level and genome representation (Schmidt and Heslop- Harrison 1998), which makes it a useful tool for the study of

evolutionary relationships between plant species (Dechyeva et al 2003; Navajas- Pérez et al 2009; Nielen et al 2010, 2012; Samoluk et al 2015a). Moreover, many of these sequences when probed onto chromosomes provide useful markers for genome characterization, for the establishment of homeologies and for the construction of chromosome maps. In *Arachis*, there has been an increasing interest in the study of repetitive elements and some of these element provided useful markers for genome and chromosome identification.

Dispersed sequences

Studies based on retrotransposons at a genomic scale are scarce in *Arachis*, and have been focused on peanut and its wild diploid progenitors (*A. ipaënsis* and *A. duranensis*). To date, Ty3-gypsy (Nielen et al 2010) and Ty1copia (Nielen et al 2012) LTR-retrotransposons have been characterized and quantified in these three species. A significant differential representation of Ty3-gypsy retrotransposons, but not of Ty1-copia retrotransposons, was described in the two analysed diploid species. The element, named FIDEL (Fairly long Inter-Dispersed Euchromatic LTR retrotransposon), is more frequent in the A- than in the B-genome, with copy numbers of about 3,000 (±950, *A. duranensis*), 820 (±480, *A. ipaënsis*), and 3,900 (±1,500, *A. hypogaea*) per haploid genome. Phylogenetic analysis of reverse transcriptase sequences showed distinct evolution of FIDEL in the ancestor species. Fluorescent in situ hybridization revealed a disperse distribution in euchromatin and absence from centromeres, telomeric regions, and the nucleolar organizer region. The distribution of FIDEL onto the chromosomes highly reflects the pattern of GISH using genomic probes of the diploid progenitors onto the chromosomes of peanut.

By contrast, Ty1copia retrotransposon from the Bianca lineage (named it Matita) is a moderate copy number element (Nielen et al 2012). This element is almost equally

represented in the A and B genomes in relatively low copy. FISH experiments showed that Matita is mainly located on the distal regions of chromosome arms and its chromosome-specific hybridization pattern aided in the identification of some individual chromosomes. By probing BAC libraries with overgos probes for Matita, it was demonstrated that this element is not randomly distributed in the genome but exhibits a significant tendency of being more abundant near resistance gene homologues than near single-copy genes.

A more recent study comparing 1.26 Mb of homeologous A and B genomes BAC clones evidenced the existence of a diverse group of complete and truncated copies of the LTR retrotransposons fraction that covered more than 40% of the sequences analysed (Bertioli et al 2013). BAC-FISH using 27 *A. duranensis* BAC clones as probes gave dispersed and repetitive DNA characteristic signals, predominantly in interstitial regions of the peanut A chromosomes. In general, the sequences of 14 BAC clones revealed that a substantial proportion of the highly repetitive component of the peanut A genome appears to be accounted for by relatively few LTR retrotransposons and their truncated copies or solo LTRs.

Recently, the representation of the L1, LINEs was analysed in the five genomes and karyotype groups of section *Arachis* (Samoluk et al 2015a). In spite that the copy number of L1 sequences was higher than the expected one for plants and directly related to genome size, the differences in copy number among genomes were significant only for the B genome. Moreover, the phylogenetic analysis of the reverse transcriptase of these elements showed that the lineages were distributed independently of the genome or karyotype groups. FISH experiments revealed a dispersed pattern with hybridization signals mainly located on the euchromatin of interstitial and distal regions of most chromosome arms in all the genome types analysed in that study.

The sum of available data evidences that retroelements have a dispersed pattern in all the genomes analysed, although with different representation among them. In spite that some of these elements may be used for the identification of different chromosome complements in hybrids and allopolyploids, their usefulness as cytogenetic markers for individual chromosome identification is limited.

Clustered sequences

Satellite DNA constitutes a significative portion of eukaryotes genomes. It is formed by repetitive units of variable length (140 - 180 bp or 300 - 360 bp) tandemly arranged in blocks of up to 100 Mpb (Charlesworth et al 1994; Schmidt and Heslop-Harrison 1998; Plohl et al 2008). These sequences usually show particular chromosome locations, being a major component of the 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). Because of this they have become a useful tool to study the karyotype evolution in different groups of species (Lanfredi et al 2001; Slamovits et al 2001).

One of the most striking differences among peanut A and B genomes is the relative percentage of the $DAPI^+$ heterochromatin (Seijo et al 2004). This is also true for the other species of section *Arachis* (Robledo et al 2009; Robledo and Seijo 2008, 2010). The analysis of a satellite sequence named ATR-2 in seven representative diploid species (x=10) of different genomes and karyotype groups revealed an infraspecific and interspecific conservation of these sequences, with a low spreading of new monomeric variants in the six species analysed (Smoluk et al unpublish.). However, the quantitative analyses revealed a different abundance of this satellite DNA among them, according to the

predictions of the “library hypothesis” (Salser et al 1976). FISH analyses revealed that ATR-2 is exclusively distributed at the DAPI$^+$ centromeric heterochromatin; however, it may not be the only sequence that conform this genomic fraction. Hence, despite the sequence conservation of this satellite, our findings suggest a key role of the ATR-2 in remodeling the karyotypes of the *Arachis* species.

More recently, other satellite sequences were cloned from a Cot-1of *A. hypogaea*. Some of them were probed into peanut chromosomes and they were useful for a further identification of different chromosomes of the A and B genomes (Zang et al 2012).

Genome size

Genome size is a useful descriptor for characterization of plant genetic resources (Ozias Akins and Jarret 1994; Rayburn et al 1997; Hendrix and Stewart 2005; Loureiro et al 2007). Until recently nuclear DNA contents were studied in a very limited number of *Arachis* species and the available genome size estimations were controversial. Most of these determinations were made by the Feulgen densitometry method (Dhillon et al 1980; Resslar et al 1981; Singh et al 1996; Lavia and Fernández 2008). However, measurements in *A. duranensis* and *A. hypogaea* by flow cytometry evidenced that the data obtained by Feulgen densitometry [except in Dhillon et al (1980) for *A. hypogaea*] overestimated the genome sizes of *Arachis* species by two (Temsch and Greilhuber 2000, 2001).

In a recent contribution, 26 diploid species of the section *Arachis* were studied by flow cytometry and the 2C values ranged from 2.55 to 3.22 pg among the diploid species (Samoluk et al 2015b). The annual species belonging to different genomes (sensu Robledo et al 2009; Robledo and Seijo 2010) tend to have different genome sizes. However, the 2C values of the perennial species of the A genome were distributed almost along the whole

range of genome sizes of diploid species. The comparisons of 2C values with karyotype parameters suggests that changes in DNA content have been proportionally distributed among the chromosome arms, and that the heterochromatic fraction was not directly involved in that changes. Within the A genome, the annual species has lower DNA content than the perennial ones which is in accordance with the predictions of the nucleotype hypothesis (Bennet 1982). However, the lack of significant relationships with geoclimatic variables suggests that there are many intrinsic factors determining particular ecological roles of the DNA content in the different lineages of section Arachis. A critical review of the DNA content of other species of *Arachis* is still needed to address the direction of the genome change during evolution of the genus as a whole.

Measurements of the genome size of the AABB species of *Arachis* showed that those polyploids are among the few number of species in which the genome size of the allopolyploids are very close to the sum of their diploid progenitors, like in tobacco (Leitch and Bennett 2004), *Hordeum* (Jakob et al 2004) and AD *Gossypium* species (Wendel et al 2002). The genome sizes estimated for *A. monticola* (5.70 pg) and for the cultivated peanut (5.60 pg) did not showed statistically significant departure from the expected 2C value estimated from the sum of the genome sizes of their parental species (*A. duranensis*, 2C = 2.55 pg; and *A. ipaënsis*, 2C = 3.19 pg) (Samoluk et al 2015b). The constancy in the Cx values suggests that the hybridization and chromosome doubling events that occurred during the origin of the cultivated peanut did not induced significant changes in genome size.

Usefulness of chromosome markers

The use of chromosome markers in *Arachis* species came to complete and extend a large number of studies from different contributors from the taxonomic, classical cytogenetics, cross hybridization, molecular markers and geographical points of view. Here we will describe two cases in which the use of these chromosome markers contribute to shed light in long debates among researches, one is about the origin of peanut and other tetraploids of *Arachis*, and the other deals with the genome characterization of the *Arachis* species.

Inferences on peanut origin

The origin of peanut has been a matter of study for several decades and has long been assessed from different points of view. The identification of two pairs of marker chromosomes (a pair of distinctly small chromosomes and a pair with secondary constriction and a satellite) in *A. hypogaea* in the pioneering report of Husted (1933, 1936), led to the proposal that the peanut is an allotetraploid species with $2n=4x=40$ and with an AABB genome constitution (Smartt et al 1978). This has been confirmed by studies on interspecific hybridization among the cultivated peanut and different wild diploid species (Gibbons and Tupley 1967; Smartt and Gregory 1967; Stalker and Wynne 1979; Singh 1986).

However, the diploid species that have participated in the origin of peanut was until recently under debate. Before the development of chromosome markers, more than eight wild diploid species having different genome types were considered as involved in the origin of peanut (reviewed in Seijo et al 2007; Grabiele et al 2012). Molecular markers studies showed that diverse species of the A genome could be considered as the most probable ancestor of peanut. Restriction fragment length polymorphism (RFLP) revealed *A.*

duranensis (Kochert et al 1991, 1996) whereas randomly amplified polymorphic DNA (RAPD) and inter-simple sequence repeat (ISSR) analysis showed *A. villosa* as the best candidates (Raina et al 2001). On the other hand, PCR amplified fragment length polymorphism (AFLP) data have shown that at least three diploid species with the A genome have small genetic distance compared with the cultigen (Milla et al 2005). Similarly, microsatellite markers have revealed that, although *A. duranensis* is closely related to the cultigen, a small group of other species having the A genome could also be possible genome donors (Moretzsohn et al 2004). Classical cytogenetics provided the first approach for the identification of the diploid progenitors of peanuts. However, only when chromosome banding and molecular cytogenetic techniques where applied massively to *Arachis* species by determining the patterns of heterochromatin and the number (Raina and Mukai 1999) and position (Seijo et al 2004; Robledo et al 2009, Robledo and Seijo 2010; Custodio et al 2013) of the rRNA genes by FISH the wild diploid progenitors were identified. The analysis of the rDNA loci distribution showed that the species included in the La Plata River Basin group of the A genome are the most related to the A genome of the tetraploids and that *A. ipaënsis*, from the B genome *s.s.* group, is the most probable B genome donor. Only when double GISH experiments using genomic DNA of the diploid *Arachis* species (2*n*=2*x*=20) was used, *A. duranensis* was undoubtedly identified as the A-genome donor of *A. hypogaea* (Seijo et al 2007).

Arachis monticola is the only wild allotetraploid within section *Arachis* and it is currently known only from three very close locations in NW Argentina. In all the dendrograms built using molecular markers, the peanut and *A. monticola* lay together, with very low or null genetic distance (Gimenez et al 2002; Grabiele et al 2012; Moretzohn et al 2012). They are known to be interfertile, with no apparent sterility of the F1 hybrids (Kirti

et al 1983). Therefore, *A. monticola* is regarded by different authors either as the direct progenitor of the peanut or as an introgressive derivative between the peanut and wild species (see Stalker and Moss 1987; Grabiele et al 2012). However, the latter hypothesis has to be discarded in the light of the molecular cytogenetic data, since *A. monticola* has exactly the same chromosome complements than *A. hypogaea* (Seijo et al 2004, 2007). That is, since *A. hypogaea* is an AABB tetraploid, if any cross with diploid species of the A or B genomes through reduced gametes (n) had occurred, it would have produced an infertile or very poorly fertile triploid with the AAB or ABB genome constitutions, respectively. Alternatively, if crosses between *A. hypogaea* and diploid species of the A or B genomes had occurred via unreduced gametes (2n) of the wild species, then the AAAB or ABBB genome constitutions, respectively, would have been expected for *A. monticola*. Neither of these scenarios would have given rise to an allotetraploid with the genome constitution compatible with that of *A. monticola* (AABB). The fact that the amphidiploid that resulted from the artificial resynthesis from *A. ipaënsis* and *A. duranensis* (Fávero et al 2006) is morphologically very similar to *A. monticola*, and that can hybridize with all the varieties of the cultigen producing fertile offspring supports the model advanced. Therefore, all the available evidences support *A. monticola* as the direct progenitor of *A. hypogaea*. However, the ability of *A. monticola* to persist as natural populations, unlike the cultivated peanut, and the fruit structure as found in all wild species (wherein each seed has its own shell separated by an isthmus) support its maintenance as a separate taxonomic species.

That both subspecies and all the botanical varieties of the cultigen had identical patterns of molecular cytogenetic markers and genomic hybridization (GISH) suggests that the same wild species participated in their origin. Moreover, this finding implies that all

presently known varieties and subspecies of peanut arose from a single, unique allotetraploid plant population (Seijo et al 2004; Seijo et al 2007). The common ancestry of all infraspecific taxa of *A. hypogaea* is supported by the low genetic variability so far detected with most molecular markers in the cultivated peanut (Halward et al 1991; Kochert et al 1996; Herselman 2003; Grabiele et al 2012; Moretzohn et al 2012).

In this scenario, after the origin of the wild allotetraploid (which probably had larger seeds than any of the progenitors as a result of the gigas effect in polyploids, like in *A. monticola*), *A. hypogaea* may have arisen through domestication. Therefore, the large morphological, ecological, phenological and chemical variability accounted in the many landraces of peanut (Krapovickas et al 2009, 2013) would have mainly result from artificial selection, under high selective pressure in different agroecological environments (Krapovickas and Gregory 1994; Grabiele et al 2012).

Origin of other tetraploid Arachis

Section *Rhizomatosae* is currently defined exclusively on morphological features, mainly because all (four) the species have rhizomes and can be asexually propagated. Among this group of taxa, *A. burkartii* is the only diploid with $2n=2x=20$, while the remainders, *A. pseudovillosa, A. nitida* and *A. glabrata* are tetraploids with $2n=4x=40$ (Gregory et al 1973; Fernández and Krapovickas 1994; Peñaloza and Valls 2005). Interestingly, the species with different ploidy level of this section occur in different geographical areas which only overlap in very narrow stretch (Krapovickas and Gregory 1994; Valls and Simpson 2005).

If this section was monophyletic, the expectation would be that the *A. burkartii* would be at least one of the genome donors of the rhizomatous tetraploids being them either auto

or allotetraploids. However, classical cytogenetic and molecular data argued the monophyly of the section and the nature of the polyploids. Therefore, an extensive analysis of the meiotic chromosome pairing and the comparative analysis of chromosomal markers revealed by fluorescent banding and FISH were recently done in these species (Ortiz et al 2011; Ortiz et al unpublished). The meiotic analysis of *A. nitida*, revealed that the chromosomes are arranged in 20II in most (65%) of the cells, but up to four multivalents (trivalents and quadrivalents) were observed in lower frequencies (Ortiz et al 2011). The molecular cytogenetic analysis revealed a similar pattern of centromeric heterochromatic bands and the same pattern of 5S rDNA loci in the four chromosome sets, but an uneven number of chromosome pairs with 18S–26S rDNA. The difference in the latter loci was interpreted as partial homology of the complements and, thus, a segmental allopolyploid origin for *A. nitida* was suggested.

Controversial hypotheses have been put forward for the nature of *A. glabrata*. Some authors proposed that this species is a true autopolyploid (Singh and Simpson 1994; Ortiz et al 2011) while others suggested that it may be an allopolyploid with the EERR genome constitution (Gregory and Gregory 1979; Bechara et al 2010). Meiotic analyses of this species showed that the frequency in which the chromosomes paired in 20II differed significantly (from 20% to 81 %) among accessions (Ortiz et al 2011). Also the frequency of trivalents (from 1 to 3III) and quadrivalents (from 1 to 8IV) was very variable among them. FISH analysis evidenced that the four chromosome sets of *A. glabrata* have a very similar pattern of CMA-DAPI bands and of the 5S and 18S–26S rDNA loci. Based on the detection of up to eight quadrivalents in meiotic cells (Ortiz et al 2011) and the similar patterns of chromosome markers in somatic metaphases, it was suggested that this species

may be either a true autopolyploid, but with different degrees of diploidyzation, or, alternatively, a segmental allopolyploid (Ortiz et al unpublished).

Concerning the genomic constitution of rhizomatous tetraploid species, the cytogenetic evidences suggest that the three species may have at least one common diploid ancestor (Ortiz et al unpublished). In this sense, all the species of the section *Rhizomatosae* (2x and 4x) have been traditionally assigned to R genome (Smartt and Stalker, 1982) assuming that the section was monophyletic and the diploid rhizomatous species, *A. burkartii,* was the natural ancestor of the tetraploids. However, the fact that *A. burkartii* has an haploid karyotype formula of 10 m with a SAT chromosome type 8, lacks DAPI^{+} heterochromatic bands, and presents four pairs of 18S–26S rDNA loci (located on three chromosome pairs) and one pair of interstitial 5S rDNA sites (that co-localized with of 18S–26S rDNA loci on the smallest pair of chromosome complement) evidenced that this diploid species is unlikely the genome donor of the rhizomatous tetraploids. Moreover, the complements of these tetraploids showed high homology with those of the *Erectoides* and *Procumbentes* section. Thus, the tetraploid species should be excluded from the R genome, which remain exclusively for *A. burkartii* (Ortiz *et al* unpublished).

Molecular markers strongly support the conclusion obtained from the cytogenetic data from the rhizomatous species. RAPD (Nobile et al 2004) and SSR (Angelici et al 2008) analyses that included the four species of section *Rhizomatosae* evidenced that the tetraploid species were clustered together, but *A. burkartii* appeared separated distantly in the tree. Furthermore, the AFLP analysis including representatives of seven different sections revealed a close association of *A. glabrata* with *A. major* and *A. paraguariensis* (sect. *Erectoides*), while *A. burkartii* was associated with the two species of the section

Caulorrhizae (Gimenez et al 2002). In addition, molecular phylogenies based on chloroplast and nuclear DNA sequences (Bechara et al 2010, Friend et al, 2010), which only included *A. burkartii* and *A. glabrata*, have shown that the diploid taxon is found in an individual and isolated clade, while *A. glabrata* grouped in a distant clade with members of sections *Erectoides* and *Procumbentes.*

The aforementioned cytogenetic and molecular marker data do not support the origin of rhizomatous tetraploids from *A. burkartii* and evidences that the section *Rhizomatosae* is not monophyletic and that rhizomes may have arisen more than once in the genus *Arachis* (Nóbile et al 2004, Angelici et al 2008, Bechara et al, 2010; Friend et al, 2010, Ortiz et al unpublished).

Genome arrangement of wild Arachis species

Diploid species of section *Arachis* with 2*n* = 20 and symmetric karyotypes have been traditionally arranged first into two different genome groups (A and non A genomes) on the basis of the presence of the small A chromosomes (firstly observed in *A. hypogaea*) in their karyotypes. These two groups of species show strong reproductive isolation manifested as low hybrid production, low chromosome pairing in interspecific hybrids and low pollen stainability (Gregory and Gregory 1979; Stalker et al 1991; Krapovickas and Gregory 1994; Tallury et al 2005). The only species with 2*n*=20 and asymmetric karyotype (*A. glandulifera*) has been considered as having the D genome (Stalker 1991).

The development of chromosome markers by molecular cytogenetic revealed a high degree of homogeneity in the karyotypes among the species with A chromosomes. However, variation in number and positions of $DAPI^+$ bands and major 18S–26S rDNA sites among species was used to establish three subgroups of karyotype homeologies

(Robledo et al 2009) having in mind that closeness of taxa is usually correlated with the similarity of their heterochromatin and rDNA FISH patterns (Hizume et al 2002; Liu et al 2003). Since the groups that resulted from the homeology analysis included species that tend to be more closely distributed geographically than those belonging to different groups, they were named using a geographical reference (Robledo et al 2009). The Chiquitano group comprised the species (*A. cardenasii*, *A. herzogii*, and *A. kempff-mercadoi*) that grow in the southern and western portion of the Chiquitanía biogeographic region in Santa Cruz Department of Bolivia. The Pantanal group includes the species (*A. diogoi*, *A. kuhlmannii*, *A. helodes*, *A. simpsonii* and *A. stenosperma*) which are distributed in the Pantanal biogeographic region, in western Brazil, northern Paraguay and eastern Bolivia. This group may also include *A. linearifolia*. The La Plata River Basin group corresponds to the species (*A. duranensis*, *A. schininii*, *A. correntina* and *A. villosa;* probably *A. microsperma*) that are distributed along the La Plata River basin (except for the region that comprises the upper stream of the Paraguay River in the Pantanal).

Most molecular marker studies including a large set of A genome species support the Pantanal group (Kochert et al 1991; Raina et al 2001; Milla et al 2005, Moretzon et al 2012). The Chiquitano group is the least represented in molecular analyses, but whenever *A. herzogii* and *A. kempff-mercadoi* have been included, they always clustered together (Milla et al 2005; Tallury et al 2005). In a recent microsatellite based phylogenetic analysis *A. cardenasii* and A *kempff-mercadoi* were intermixed with the species that belong to the Pantanal group (Moretzohn et al 2013), but only two accession of *A. cardenasii* and none of *A. herzogii* were included in that study. The grouping of the species included in the La Plata River basin is generally observed in the molecular markers analyses, although with different degrees of support (Milla et al 2005, Moretzohn et al 2013).

The non A genome with 2*n*=20 has been segregated into three genomes based on the different patterns of chromosome markers that they present in their karyotypes (Robledo and Seijo 2010). One of these karyotype groups, named as K genome, includes *A. batizocoi*, *A. cruziana*, and *A. krapovickasii.* This genome is characterized by having conspicuous heterochromatic bands in nine chromosome pairs and three 5S rDNA loci. The second group was named F genome and it is integrated by *A. benensis* and *A. trinitensis.* The karyotypes of these species have small and faint bands in seven or eight chromosome pairs and only one 5S rDNA locus. The remaining species have karyotypes which characteristically lack pericentromeric heterochromatin. This group of species includes *A. ipaënsis*, *A. gregoryi*, *A. magna, A. valida* and *A. williamsii,* and retained the B genome *sensu stricto* designation (Robledo and Seijo 2010). This is because it is the group that showed the highest homeology with the B genome of *A. hypogaea*. *Arachis glandulifera* has asymmetric karyotype (Stalker 1991; Fernandez and Krapovickas 1994), a unique pattern of heterochromatin distribution and 10 pairs of 18S–26S rDNA loci. On this basis, this species was confirmed as having the D genome (Robledo and Seijo 2008), as first considered by Stalker (1991).

The proposed genome arrangement is supported by species crossability, pollen stainability; morphological characters and geographical distribution of the species (Smartt et al 1978; Stalker 1991; Krapovickas and Gregory 1994; Tallury et al 2005; Burrow et al 2009). The position of the K, D and F genomes with respect to the A and B genomes is controversial. The analysis of chloroplast markers revealed that the D, K and F genomes are closer to the B genome, but the NTS of the 5S rDNA genes and AFLP showed that the F genome is indeed closer to the B genome, the K and D ones clustered as sister groups of the A genome (Tallury et al 2005; Grabiele et al 2012). The phylogenetic relationships of

species of section *Arachis* analysed based on DNA sequence information of three single-copy gene introns was consistent with the current genome classification, since clades contained species with the same genome types (Moretzshon et al 2013). In that analysis, the species with D, F and K genome species were closer to the A genome species, but the microsatellite analysis done in the same report showed that those genomes are closer to the B genome than to the A genome. In a more recent analysis using nine intron sequences and GISH, it was shown that the K genome would be closer to the B genome than to the A genome of *A. hypogaea* (Leal-Bertioli et al 2015). All the results showed that the K, F and D genomes are in an intermediate position between the A and the B genome, and share a set of markers (chromosomes, AFLP, 5S rDNA, and some introns) with the former, while others (chloroplasts, other introns) with the latter.

Geographically, the species included within each of the genomes segregated from the non A group showed a strong tendency to be co-distributed. The species of the K genome are distributed in the NW of the Chacoan Boreal region, while the species of the F genomes are restricted to the lowland savannas of Beni department in Bolivia. The species of the B genome are more widely distributed in semi-deciduous forests and savannas of the cerrado associated with the chiquitano planalto and west pantanal. *Arachis ipaënsis*, known only from one population, was collected at the top of the sand banks of streams in an ecotone between the tucumano-oranence deciduous forest and the chacoan xerophytic forest (Robledo and Seijo 2010).

More recently, the three species with x=18 chromosomes of section *Arachis* were analysed. The comparative analyses with the x=10 species of section *Arachis* indicate that x=9 species are chromosomally most closely related to A genome species, because of their similarities in centromeric heterochromatic bands, the interchromosomal asymmetry index

and the same number and position of 5S rDNA sites. However, the sum of karyotype markers together with the reproductive isolation of this species with any other of section *Arachis* led to the proposal of a different genome type to this group of species, that is the G genome (Silvestri et al 2015). Analyses of molecular datasets have revealed that in spite that the three x=9 species form a compact clade different from any other in the section *Arachis* (Bechara et al 2010; Friend et al 2010; Moretzshon et al 2013), their relationship with other species of the section is still unclear. The analysis of microsatellites (Moretzshon et al 2004, 2013), single-copy gene sequences (Moretzshon et al 2013), trnT-F cpDNA marker (Tallury et al 2005), and ITS and 5.8S of the nuclear rDNA (Bechara et al 2010) suggest that these taxa are more closely related to some non A genome species; whereas, RAPD (Creste et al 2005) and AFLP (Milla et al 2005) analyses suggest that they are genetically more similar to A genome species.

Therefore, section *Arachis* is composed of diploid species that showed six different arrangements of heterochromatin and rDNA loci, that have been assigned to different genomes (A, B, D, F, G, and K), each one showing different degrees of reproductive isolation and with differential potential for introgression of valuable agronomic traits.

References

Angelici C, Hoshino AA, Nóbile PM, Palmieri DA, Valls JFM, Gimenes MA, Lopes CR (2008) Genetic diversity in Rhizomatosae section of the genus *Arachis* (Fabaceae) based on microsatellite markers. Genet Mol Biol 31:79-88

Bechara MD, Moretzsohn MDC, Palmieri DA, Monteiro JP, Bacci Jr M, Martins Jr J, 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

Bennet MD (1982) Nucleotypic basis of the spatial ordering of chromosomes in eukaryotes and the implications of the order for genome evolution and phenotypic variation. In: Dover GA, Flavel RB (eds) Genome evolution, Academic Press, New York, p 239-261

Bertioli DJ, Vidigal BS, Nielen S, Ratnaparkhe MB, Lee T, 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

Burow MD, Simpson CE, Faries MW, Starr JL, Paterson AH (2009) Molecular biogeographic study of recently described B- and A-genome *Arachis* species, also providing new insights into the origins of cultivated peanut. Genome 52:107-119

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

Creste S, Yulmann Neto A, Figueira A (2001) Detection of single sequence repeat polymorphisms in denaturing polyacrylamide sequencing gels by silver staining. Plant Mol Biol Rep 19:299–306

Custódio AR, Seijo G, Valls JFM (2013) Characterization of Brazilian accessions of wild *Arachis* species of section Arachis (Fabaceae) using heterochromatin detection and fluorescence in situ hybridization (FISH). Genet Mol Biol 36:364-370

Dechyeva D, Gindullis F, Schmidt T (2003) Divergence of satellite DNA and interspersion of dispersed repeats in the genome of the wild beet *B. procumbens*. Chrom Res 11:3–21

Dhillon SS, Rake AV, Miksche JP (1980) Reassociation kinetics and cytophotometric characterisation of peanut (*Arachis hypogaea* L.) DNA. Plant Physiol 65:1121–1127

Fávero AP, Simpson CE, Valls JFM, Vello NA (2006) Study of the evolution of cultivated peanut through crossability studies among *Arachis ipaensis*, *A. duranensis* and *A. hypogaea*. Crop Sci 46:1546–1552

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

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

Gimenes MA, Lopes Romero C, Galgaro ML, Valls JFM (2002) RFLP analysis of genetic variation in species of section Arachis, genus *Arachis* (Leguminosae). Euphytica 123:421-429

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

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

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

Gregory WC, Gregory MP, Krapovickas A, Smith BW, Yarbrough JA (1973) Structure and genetic resources of peanuts. In: Wilson CT (ed) Peanut culture and uses, Amer. Peanut Res. and Educ. Asso, Stillwater, p 47–133

Grewal SIS, Moazed D (2003) Heterochromatin and Epigenetic Control of Gene Expression. Science 301:798-802

Halward TM, Stalker HT, Larue EA, Kochert G (1991) Genetic variation detectable with molecular markers among unadapted germplasm resources of cultivated peanut and wild species. Genome 34:1013–1020

Hendrix B, Stewart JM (2005) Estimation of the nuclear DNA content of *Gossypium* species. Ann Bot 95:789–797

Herselman L (2003) Genetic variation among Southern African cultivated peanut (*Arachis hypogaea* L.) genotypes as revealed by AFLP analysis. Euphytica 133:319–327

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.

Hudakova S, Michalek W, Presting GG. ten Hoopen R, dos Santos K, Jasencakova Z, Schubert I (2001) Sequence organization of barley centromeres. Nucleic Acids Res. 29:5029-5035

Husted L (1933) Cytological studies on the peanut, *Arachis*. I. Chromosome number and morphology. Cytologia (Tokyo) 5:109-117

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

Jakob SS, Meister A, Blattner FR (2004) The considerable genome size variation of *Hordeum* species (Poaceae) is linked to phylogeny, life form, ecology, and speciation rates. Mol Biol Evol 21:860–869

Kirti PB, Bharathi M, Murty UR, Rao NGP (1983) Chromosome morphology in three diploid species of *Arachis* and its bearing on the genomes of groundnut (*Arachis hypogaea* L.). Cytologia 48:139–151

Kochert G, Branch WD, Simpson CE (1991) RFLP variability in peanut (*Arachis hypogaea* L.) cultivars and wild species. TAG Theor Appl Genet 81:565-570

Kochert G, Stalker HT, Gimenes MA, Galgaro ML, Lopes Romero C, Moore K (1996) RFLP and cytogenetic evidence on the origen and evolution of allotetraploid domesticated peanut, *Arachis hypogaea* (Leguminosae). Am J Bot 83:1282-1291

Krapovickas A (1973) Evolution of the genus *Arachis*. In: Moav R (ed), Agricultural Genetics Selected Topics, National Council for Research and Development, Jerusalem, p 135-151

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

Krapovickas A, Vanni RO, Pietrarelli JR, Simpson CE (2013) Las razas de maní de Perú. Bonplandia 22:19-90.

Krapovickas A, Vanni RO, Pietrarelli JR, Williams DE, Simpson CE (2009) Las razas de maní de Bolivia. Bonplandia 18:95-189.

Lanfredi M, Congiu L, Garrido-Ramos MA, de la Herrán R, Leis M, Chicca M, Rossi R, Tagliavini J, Ruiz Rejón C, Ruiz Rejón M, Fontana F (2001) Chromosomal location and evolution of a satellite DNA familyin seven sturgeon species. Chromosome Res 9:47-52.

Lavia GI (1996) Estudios cromosómicos en *Arachis* (Leguminosae). Bonplandia 9:111-120

Lavia GI (1998) Karyotypes of *Arachis palustris* and *A. praecox* (Section Arachis), Two Species with Basic Chromosome Number x=9. Cytologia (Tokyo) 63:177-181

Lavia GI, Fernández A (2004) Karyotypic studies in *Arachis hypogaea* L. varieties. Caryologia 57:353-359

Lavia GI, Fernández A (2008) Genome size in wild and cultivated peanut germplasm. Plant Syst Evol 272:1-10

Lavia GI, Ortiz AM,, Robledo G, Fernández A, Seijo G (2011) Origin of triploid *Arachis pintoi* (Leguminosae) by autopolyploidy evidenced by FISH and meiotic behavior. Ann Bot 108:103–111

Leal-Bertioli SCM, Santos SP, Dantas KM, Inglis PW, Nielen S, Araujo ACG, Silva JP, Cavalcante U, Guimarães PM, Brasileiro ACM, Carrasquilla-Garcia N, Varma

Penmetsa R, Cook D, Moretzsohn MC, Bertioli DJ (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

Leitch IJ, Bennett MD (2004) Genome downsizing in polyploidy plants. Biol J Linn Soc 82:651–663

Liu ZL, Zhang D, Hong DY, Wang XR (2003) Chromosomal localization of 5S and 18S-5.8S-25S ribosomal DNA sites in five Asian pines using Fluorescence in situ hybridization. Theor Appl Genet 106:198-204.

Loureiro J, Rodriguez E, Costa A, Santos C (2007) Nuclear DNA content estimations in wild olive (*Olea europaea* L. ssp. *europaea* var. *sylvestris* Brot.) and Portuguese cultivars of *O. europaea* using flow cytometry. Genet Resour Crop Evol 54:21–25

Macas J, Požárková D, Navrátilová A, Nouzová 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

Milla R, Isleib TG, Stalker HT (2005) Taxonomic relationships among Arachis sect. Arachis species as revealed by AFLP markers. Genome 48:1-11

Moretzsohn MDC, Hopkins MS, Mitchell SE, Kresovich S, Valls JFM, Ferreira ME (2004) Genetic diversity of peanut (*Arachis hypogaea* L.) and its wild relatives based on the analysis of hypervariable regions of the genome. BMC Plant Biol 4:11

Moretzhon 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

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, Del Bosque MQ, Garrido-Ramos MA (2009) Effect of location, organization, and repeat-copy number in satellite-DNA evolution. Mol Gen Genomics 282:395-406

Nielen S, Campos-Fonseca F, Leal-Bertioli SCM, Guimarães 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 BS, Leal-Bertioli SCM, Ratnaparkhe M, Paterson AH, Garsmeur O, D'Hont A, Guimarães PM, Bertioli DJ (2012) Matita, a new retroelement from peanut: characterization and evolutionary context in the light of the *Arachis* A-B genome divergence. Mol Genet Genomics 287:21-38

Nóbile PM, Gimenes MA, Valls JFM, Lopes CR (2004) Genetic variation within and among species of genus *Arachis,* section *Rhizomatosae*. Genet Resour Crop Evol 51:299-307

Ortiz AM, Seijo JG, Fernández A, Lavia GI (2011) Meiotic behavior and pollen viability of tetraploid *Arachis glabrata* and *A. nitida* species (Section Rhizomatosae,

Leguminosae): implications concerning their polyploidy nature and seed set production. Plant Syst Evol 292:73–83

Ortiz AM, Silvestri MC, Lavia GI (2013) Karyotypic studies in wild species of *Arachis* (Leguminosae) belonging to sections Erectoides, Procumbentes and Rhizomatosae. Bol Soc Argent Bot 48:295-300

Ortiz AM, Seijo G, Lavia GI (2014) *Arachis.* In: Marhold, K (ed), IAPT/IOPB chromosome data 17. Taxon 63:1153

Ozias-Akins P; Jarret RL (1994) Nuclear DNA content and ploidy levels in the genus *Ipomoea*. J Am Soc Hortic Sci 119:110-115

Peñaloza APS, Valls JFM (2005) Chromosome number and satellited chromosome morphology of eleven species of *Arachis* (leguminosae). Bonplandia 14:1-8

Peñaloza A, Pozzobon MT, Valls JFM (1996) Cytogenetic findings in wild species of *Arachis* (Leguminosae). In: Sociedade Brasileira de Genética (ed) Rev Bras Genét, Ribeirão Preto

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, Meštrović N, Mantovani B (2008) Satellite DNAs between selfishness and functionality: Structure, genomics and evolution of tandem repeats in centromeric heterochromatin. Gene 409:72–82

Pucciariello O, Ortiz AM, Fernández A, Lavia GI (2013) Análisis cromosómico del híbrido *Arachis pintoi* x *A. repens* (Leguminosae) mediante citogenética clásica y molecular. Bol. Soc. Argent. Bot. 48:111-119

Raina SN, Mukai Y (1999) Detection of a variable number of 18S – 5 . 8S – 26S and 5S ribosomal DNA loci by fluorescent in situ hybridization in diploid and tetraploid *Arachis* species. Genome 59:52-59

Raina SN, Rani V, Kojima T, Ogihara Y, Singh KP, Devarumath RM (2001) RAPD and ISSR fingerprints as useful genetic markers for analysis of genetic diversity, varietal identification, and phylogenetic relationships in peanut (*Arachis hypogaea*) cultivars and wild species. Genome 44:763-772

Rayburn AL, Biradar DP, Bullock DG, Nelson RL, Gourmet C, Wetzel JB (1997) Nuclear DNA content diversity in Chinese soybean introductions. Ann Bot 80:321–325

Resslar PM, Stucky JM, Miksche JP (1981) Cytophometric determination of the amount of DNA in *Arachis* (Leguminosae). Amer J Bot 68:149–153

Robledo G, Seijo G (2008) Characterization of the *Arachis* (Leguminosae) D genome using fluorescence in situ hybridization (FISH) chromosome markers and total genome DNA hybridization. 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

Robledo G, Lavia GI, Seijo G (2010) Genome re-assignment of *Arachis trinitensis* (Sect. Arachis, Leguminosae) and its implications for the genetic origin of cultivated peanut. Genet Mol Biol 33:714-718

Salser W, Bowen S, Browne D, El-Adli F, Federoff N, Fry K, Heindall H, Paddock G, Poon R, Wallace B, Whitcomb P (1976) Investigation of the organization of mammalian chromosomes at the DNA sequence level. Fed Proc 35:23–35

Samoluk SS, Robledo G, Podio M, Chalup L, Ortiz JPA, Pessino SC, Seijo JG (2015a) 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

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

Santana SH, Valls JFM (2015) *Arachis veigae* (Fabaceae), the most dispersed wild species of the genus, and yet taxonomically overloocked. Bonplandia 24:139-150

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

Seijo G, Lavia GI, Fernández A, Krapovickas A, Ducasse D, Moscone E (2004) Physical mapping of the 5S and 18S–25S rRNA genes by FISH as evidence that *Arachis duranensis* and *A. ipaensis* are the wild diploid progenitors of *A. hypogaea* (Leguminosae). Am J Bot 91:1294-1303

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

Silva SC, Martins MIG, Santos RC, Peñaloza APS, Melo Filho PA, Benko-Iseppon AM, Valls JFM, Carvalho R (2010) Karyological features and banding patterns in *Arachis* species belonging to the Heteranthae section. Plant Syst Evol 285:201–207

Silvestri MC, Ortiz AM, Lavia GI (2015) 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

Singh AK (1986) Utilization of wild relatives in the genetic improvement of *Arachis hypogaea* L. 8. Synthetic amphidiploids and their importance in interspecific breeding. Theor Appl Genet 72:433-439

Singh AK, Moss JP (1982) Utilization of wild relatives in genetic improvement of *Arachis hypogaea* L. Theor Appl Genet 61:305-314

Singh AK, Simpson CE (1994) Biosystematics and genetic resources. In: Smartt J (ed) The Groundnut Crop, a Scientific Basis for Improvement, Chapman and Hall, London, p 96–138

Singh AK, Smartt J (1998) The genome donors of the groundnut / peanut (*Arachis hypogaea* L.) revisited. Genet Resour Crop Evol 45:113-118

Singh AK, Raina SN, Singh AK (1996) Variation in chromosomal DNA associated with the evolution of *Arachis* species. Genome 39:890-897

Slamovits CH, Cook JA, Lessa EP, Rossi MS (2001) Recurrent amplifications and deletions of satellite DNA accompanied chromosomal diversification in South American Tuco-tucos (Genus *Ctenomys*, Rodentia: Octodontidae): a phylogenetic approach. Mol Biol Evol 18:1708–1719

Smartt J, Gregory WC (1967) Interspecific cross-compatibility between the cultivated peanut *Arachis hypogaea* L. and other members of the genus *Arachis*. Oléagineux 22:455-459

Smartt J, Stalker HT (1982) Speciation and cytogenetics of *Arachis*. In: Pattee HE, Young CT (eds), Peanut science and technology, Peanut Res and Educ Soc, Yoakum, p 21–49

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

Stalker HT (1991) A new species in section Arachis of peanuts with a D genome. Amer J Bot 78:630-637

Stalker HT, Moss JP (1987) Speciation, cytogenetics, and utilization of *Arachis* species. Adv Agron 41:1–40

Stalker HT, Wynne JC (1979) Cytology of interspecific hybrids in section Arachis of peanuts. Peanut Sci 6:110–114

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

Tallury SP, Hilu KW, Milla R, 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

Temsch EM, Greilhuber J (2000) Genome size variation in *Arachis hypogaea* and *A. monticola* re-evaluated. Genome 43:449-451. http://www.ncbi.nlm.nih.gov/pubmed/10902707

Temsch EM, Greilhuber J (2001) Genome size in *Arachis* duranensis: a critical study. Genome 44:826-830

Urdampilleta JD, Pereira de Souza A, Schneider DRS, Vanzela ALL, Ferrucci MS, Martins ERF (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 14:35–63

Valls JFM, Costa LC, Custodio AR (2013) A novel trifoliolate species of *Arachis* (Fabaceae) and further comments on the taxonomic section Trierectoides. Bonplandia 22:91-97

Wendel JF, Cronn RC, Johnston JS, Price HJ (2002) Feast and famine in plant genomes. Genetica 115:37–47

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