Comparative cytogenomics reveals genome reshuffling and centromere repositioning in the legume tribe Phaseoleae

Claudio Montenegro[1], Lívia do Vale Martins[2,4], Fernanda de Oliveira Bustamante[3], Ana Christina Brasileiro-Vidal[2], Andrea Pedrosa-Harand[1]✉

[1]Laboratory of Plant Cytogenetics and Evolution, Department of Botany, Federal University of Pernambuco, Recife, PE, Brazil.

[2]Laboratory of Plant Genetics and Biotechnology, Department of Genetics, Federal University of Pernambuco, Recife, PE, Brazil.

[3]Minas Gerais State University, Divinópolis Unity, Divinópolis, MG, Brazil.

[4]Department of Biology, Federal University of Piauí, Teresina, PI, Brazil

Claudio Montenegro - ORCID: 0000-0003-2089-1608

Lívia do Vale Martins - ORCID: 0000-0003-4645-9055

Fernanda de Oliveira Bustamante - ORCID: 0000-0002-2826-5217

Ana Christina Brasileiro-Vidal - ORCID: 0000-0002-9704-5509

Andrea Pedrosa-Harand - ORCID: 0000-0001-5213-4770

✉ For correspondence (e-mail: andrea.harand@ufpe.br)

Andrea Pedrosa-Harand

Laboratório de Citogenética e Evolução Vegetal, Departamento de Botânica, Universidade Federal de Pernambuco – UFPE

R. Prof. Moraes Rego, s/n, CDU, 50670-420, Recife, PE, Brazil

Tel. number: + 55 81 2126 8846

Fax number: + 55 81 2126 8358

E-mail: andrea.harand@ufpe.br

SUMMARY

The tribe Phaseoleae (Leguminosae; Papilionoideae) includes several legume crops with assembled genomes. Comparative genomic studies indicated the preservation of large genomic blocks in legumes. However, the chromosome dynamics along its evolution was not investigated in the tribe. We conducted a comparative genomic analysis using CoGe Synmap platform to define a useful genomic block (GB) system and to reconstruct the ancestral Phaseoleae karyotype (APK). We defined the GBs based on orthologous genes between *Phaseolus vulgaris* and *Vigna unguiculata* genomes (n = 11), then searched for these GBs in different genome species belonging to the Phaseolinae (*P. lunatus*, n = 11) and Glycininae (*Amphicarpaea edgeworthii*, n = 11 and *Spatholobus suberectus*, n = 9) subtribes, and in the outgroup (*Medicago truncaluta*, n = 8). To support our *in silico* analysis, we used oligo-FISH probes of *P. vulgaris* chromosomes 2 and 3 to paint the orthologous chromosomes of the non-sequenced Phaseolinae species (*Macroptilium atropurpureum* and *Lablab purpureusi*, n = 11). We inferred the APK with n = 11, 22 GBs (A to V) and 60 sub-GBs. We hypothesized that the main rearrangements within Phaseolinae involved nine APK chromosomes, with extensive centromere repositioning resulting from evolutionary new centromeres (ENC) in the *Phaseolus* lineage. We demonstrated that the *A. edgeworthii* genome is more reshuffled than the dysploid *S. suberectus* genome, in which we could reconstructed the main events responsible for the chromosome number reduction. The development of the GB system and the proposed APK provide useful tools for future comparative genomic analyses of legume species.

Keywords: Genomic Blocks; Ancestral Karyotype; Leguminosae; Oligo-FISH; Dysploidy; Comparative genomics; Genome structure and evolution

Significance statement: We developed a genomic block system and proposed an Ancestral Phaseoleae Karyotype based on available genome assemblies of these legume crops. These tools enabled to reconstruct the main chromosomal rearrangements responsible for the genome reshuffling among the diploid investigated taxa. The analyses also revealed centromere repositioning for all chromosomes, despite conservation of chromosome number.

INTRODUCTION

Since the first plant genome was sequenced and assembled (The Arabidopsis Genome Initiative, 2000), genome sequencing technologies improved and became more accessible, increasing the genomic data available for economic and evolutionary important plant species (e.g., Li *et al*., 2018; Chen *et al*., 2018). Sequencing genomes is essential for functional and comparative genomics, playing a fundamental role in understanding the plant biology and chromosomal evolutionary dynamics, such as genome reshuffling (Pavy *et al*., 2012; Cheng *et al*., 2014), genome size variation (Wendel *et al*., 2016; Pellicer *et al*., 2018; Kreplak *et al*., 2019; Zhang *et al*., 2020), polyploidy (Jiao *et al*., 2011; Soltis *et al*., 2015; Geiser *et al*., 2016; Ruprecht *et al*., 2017; Wu *et al*., 2020), dysploidy (Lysak *et al*., 2006; Yang *et al*., 2014; Mandáková and Lysak, 2018), and other mechanisms for genome and species diversification.

Cytogenomic comparisons of related species may provide important evolutionary insights. In Brassicaceae for instance, chromosome painting based on *A. thaliana* BACs (Bacterial Artificial Chromosomes) as FISH (Fluorescent *in situ* Hybridization) probes (Lysak *et al*., 2001), together with a genomic block system (Schranz *et al*., 2006), elucidated the karyotype evolution within this family. These studies allowed to infer the ancestral crucifer karyotype (ACK), revealing chromosomal rearrangements related to the decreasing dysploidy in *A. thaliana* (Lysak *et al*., 2006), chromosomal reshuffling after whole genome triplication (WGT) in *Brassica* species (Cheng *et al*., 2014), and centromere repositioning across the family (Willing *et al*., 2015; Lysak *et al*., 2016; Mandáková *et al*., 2020). In *Cucumis* L., genomic blocks combined with FISH maps revealed mechanisms of genome reshuffling, centromere repositioning and decreasing dysploidy in cucumber (*C. sativus* L.) based on ancestral karyotype as reference (Yang *et al*., 2014). However, due to the complex genome structure and the repetitive DNA content in most plant genomes, chromosome painting probes are available for inter-specific comparison only for a small number of species.

The tribe Phaseoleae (Leguminosae; Papilionoideae) comprises the most economically important legumes, including the paleopolyploid soybean ($2n = 4x = 40$, *Glycine max* L.), and the diploids common bean ($2n = 2x = 22$, *Phaseolus vulgaris* L.) and the cowpea [$2n = 2x = 22$, *Vigna unguiculata* (L.) Walp. (Moussa *et al*., 2011; Brookes and Barfoot, 2014; Myers and Kmiecik, 2017)], with genome sizes of ~ 1.1 Gb, ~587 Mb and ~640 Mb, respectively (Schmutz *et al*., 2010; 2014; Lonardi *et al*., 2019). Analyses of soybean genome revealed a legume-common tetraploidization (LCT) around 60 million years ago (Mya), and a soybean-specific tetraploidization (SST) in the *Glycine* lineage ~12 Mya (Schmutz *et al*., 2010). More recently, multi-alignment analyses for ten legume genomes [*Arachis duranensis* Krapov. & W. C. Greg, *A. ipaensis* Krapov. & W. C. Greg, *Cajanus cajan* (L.) Millsp., *Cicer arietium* L., *G. max*, *Lotus japonicus* (Regel) K. Larsen, *P. vulgaris*, *Vigna angularis* (Willd.) Ohwi & H. Ohashi and *V. radiata* (L.) R. Wilczek] revealed insights into the ancestral polyploidization of legumes and the specific autopolyploidization of *Glycine*, suggesting a tendency of gene loss after polyploidization and extensive chromosome reshuffling (Wang *et al*., 2017). Nevertheless, the analyses also suggested high levels of synteny among these genomes, with large conserved genomic blocks (Wang *et al*., 2017).

Within the genus *Phaseolus*, BAC-FISH using *P. vulgaris* probes demonstrated conserved synteny among three species of different clades (Bonifácio *et al.*, 2012; Fonsêca and Pedrosa-Harand, 2013). On the other hand, similar comparative cytogenetic mapping of two species of the Leptostachyus clade from the same genus showed extensive genome reshuffling associated with descending dysploidy involving a nested chromosome fusion between chromosomes 10 and 11 (Fonsêca *et al.*, 2016; Ferraz *et al.*, 2020). Furthermore, comparative cytogenetics and sequence alignment between *Vigna* species and *P. vulgaris* revealed a high degree of synteny with five chromosomes involved in synteny breaks (Vasconcelos *et al.*, 2015; Lonardi *et al.*, 2019; Oliveira *et al.*, 2020; do Vale Martins *et al.*, 2021). A detailed analysis of chromosomes 2 and 3 of *V. angularis*, *V. unguiculata* and *P. vulgaris* based on sequence alignment and oligo-FISH painting integrative approaches, identified additional macro- and micro inversions, translocations, and intergeneric centromere repositioning (do Vale Martins *et al.*, 2021). Centromere repositioning was also detected on chromosomes 5, 7, and 9 by oligo-FISH barcode combined with genome sequence data in *V. unguiculata* and *P. vulgaris* (Bustamante *et al.*, 2021). In addition, the authors detected the involvement of chromosome 5 in the translocation complex 1, 5 and 8, a paracentric inversion on chromosome 10, and detailed a pericentric inversion on chromosome 4. The direction and time of these rearrangements were, however, not determined.

To understand the dynamics of genome reshuffling among diploid Phaseoleae legumes, we constructed a genomic block (GB) system based on comparative cytogenomic data. We compared *P. vulgaris* and *V. unguiculata* genomes to define the GB system using *P. vulgaris* as reference and applied the GBs to four other species with genome assembled: *Phaseolus lunatus* L. ($2n = 2x = 22$), also from the Phaseolinae subtribe; *Spatholobus suberectus* Dunn ($2n = 2x = 18$) and *Amphicarpaea edgeworthii* Benth. ($2n = 2x = 22$), both belonging to the Glycininae subtribe; and *Medicago truncatula* Gaertn ($2n = 2x = 16$, tribe Trifolieae) as an outgroup. Moreover, we performed oligo-FISH using specific probes for *P. vulgaris* chromosomes 2 and 3 to visualize the orthologous chromosomes in two non-sequenced Phaseolinae species with intermediate phylogenetic positions, *Macroptilium atropurpureum* DC. Urb. ($2n = 2x = 22$) and *Lablab purpureus* L. ($2n = 2x = 22$). We hypothesized the Ancestral Phaseolae Karyotype (APK) and inferred the main chromosomal rearrangements related to the evolution and diversification of these legumes. Our results indicate extensive genome reshuffling in particular lineages and centromere repositioning across the Phaseolinae and Glycininae subtribes at diploid level, with centromere repositioning for all 11 chromosomes of *P. vulgaris/V. unguiculata*, despite their conserved karyotypes. Our GB system and the proposed APK are promising tools for future comparative genomics analyses when further genome assemblies become available.

RESULTS

Genomic blocks and the inferred Ancestral Phaseoleae Karyotype

We aligned *P. vulgaris* (*Pv*) and *V. unguiculata* (*Vu*) genomes based on the collinear arrangement of orthologous genes in dotplots (Supplementary Table 1). Twenty-two genomic blocks (GBs) were defined based on the synteny breaks between both genomes. Using *P. vulgaris* as a reference

karyotype, these GBs were named (A to V), ordered and oriented according to their distribution in the 11 *P. vulgaris* chromosome pairs, with each chromosome containing at least one GB (Supplementary Table 2). Moreover, we divided the 22 GBs into sub-GBs, designed by letters followed by numbers, based on collinearity breaks within GBs. According to the position, orientation, and distribution of the sub-GBs, we performed comparative genomic analyses across the Phaseoleae tribe, detecting the GBs and sub-GBs in *V. unguiculata* (*Vu*), *P. lunatus* (*Pl*), *Spatholobus suberectus* (*Ss*), *Amphicarpaea edgeworthii* (*Ae*) and *M. truncatula* (*Mt*) genomes. In general, all the 22 generated GBs were found in all analysed species. However, due to rearrangements between *P. vulgaris* and other Phaseoleae species, different numbers of sub-GBs were found for each species, with subdivisions within the sub-GBs represented by lowercase letters after numbers (Supplementary Figures 1-5). When sub-GBs could not be detected by the standard SynMap analysis, we conducted a manual search using blastn on CoGe to find these sub-GBs (asterisks on Supplementary Table 2).

To establish the ancestral karyotype of Phaseoleae (APK) based on our GB comparisons, we selected the most frequent GB associations, particularly those also shared with *M. truncatula*, considering the phylogenetic relationships among species, as described by Li *et al*. (2013), and checking if they shared same or similar breaks points. We proposed the APK with $n = 11$ (most common chromosome number within the tribe), 22 GBs and 60 sub-GBs (Supplementary Table 2). The number of APK chromosomes was chosen to maximize the chromosome orthology within the Phaseolinae species (Figure 1A and Supplementary Table 2). The centromere positions in APK chromosomes were hypothesized based on the frequency of associations between GBs and centromeres in the analysed species. However, as the centromere of APK6 (in M2-R8a) was observed only in the Phaseolinae species, it might not represent the ancestral state.

Chromosomal rearrangements and centromere repositioning in Phaseolinae subtribe in relation to the APK

Six APK chromosomes displayed full synteny with at least one chromosome of the three analysed Phaseolinae species: APK2 (*Vu*2), APK3 (*Vu*3), APK4 (*Pv*4, *Pl*4, *Vu*4), APK5 (*Pv*5, *Pl*5), APK7 (*Vu*7, *Pv*7, *Pl*7) and APK8 (*Vu*8). We propose that the main chromosomal rearrangements common to all Phaseolinae species (Figure 1.B1) involved a complex translocation between APK1, 6 and 9, and a reciprocal translocation between APK10 and 11, resulting in the ancestral Phaseolinae karyotype (APnK) chromosomes 1, 6, 9, 10 and 11 (Figure 1.B1). Our data confirmed exclusive rearrangements for *Vigna unguiculata* and *Phaseolus* species (Figure 1.B2 and 1.B3, respectively). In *V. unguiculata* (*Vu*), we observed a reciprocal translocation between APnK1 and APK5, resulting in chromosomes *Vu*1 (K+B) and *Vu*5 (P+L), in addition to a large pericentric inversion comprising most of APK4 (I and J) generating *Vu*4 (Figure 1.B2). Furthermore, in the ancestral *Phaseolus* karyotype (APsK) two reciprocal translocations occurred: between APnK1 and 8, resulting in APsK1 (A+B) and APsK8 (P+Q); and between APK2 and 3, generating APsK2 (C+E+C+D+F) and APsK 3 (G+H), followed by inversions and intrachromosomal translocations on APsK2 (C and E) and APsK3 (H) (Figure 1.B3). We also detected previously described inversions between *P. vulgaris* and *P. lunatus* (Bonifácio *et al.*, 2012; Garcia *et al*.,

2021) in Chr1 (in B); Chr2 (C, D and E); Chr3 (within H4); Chr7 (in N and O); Chr9 (in R), and Chr10 (in S and T). The complex, multiple inversions in chromosome 2 and 7, also involving intrachromosomal translocations, occurred independently in *P. vulgaris* and *P. lunatus*, while the intrachromosomal translocations on chromosomes 1 and inversions on chromosomes 1, 3, 9 and 10 occurred in *P. lunatus* (*Pl*1) and *P. vulgaris* (*Pv*1, 3, 9 and 10) lineages, respectively (Figure 1.B4 and 1.B5).

Our analyses indicated centromere repositioning for all centromeres of APK in at least one of the Phaseolinae species (Figure 2): centAPK1: P2 (*Vu*: P2-L); centAPK2: G (*Pl*: C4-D; *Pv*: C3); centAPK3: H1-H5 (*Vu*: F1-D; *Pl*, *Pv*: H3-H4); centAPK4: I4 (*Vu*: I3-I2); centAPK5: K (*Pl*, *Pv*: L1-L2), centAPK6: M2-R8a (*Pv*: M1-M2), centAPK7: O7 (*Vu*: O5-O3; *Pl*: N2-N1; *Pv*: O1-N2), centAPK8: A (*Pl*, *Pv*: B1), centAPK9: R8b (*Vu*, *Pl*: R8a-R7; *Pv*: R1), centAPK10: U (*Vu*: S1; *Pl*: S3; *Pv*: T2) and centAPK11: V (*Pl*, *Pv*: U1). Within *Phaseolus* species, centromeres of chromosomes 2, 6, 7, 9 and 10 were repositioned. Only the centromere of *Pl*6 was maintained from APK (Figure 1.A4 and 2), while centromeric repositioning in Chrs. 2, 7, 9 and 10 may have occurred independently in both species. Centromeric repositioning was observed for chromosome 1, 3, 4, 7 and 10 in the *Vigna* lineage, while repositioning in Chr. 9 was shared with *P. lunatus* and probably occurred in the Phaseolinae ancestral (Figures 1.A2 and 2). Overall, we could hypothesize that the major events of centromere repositioning were derived from Evolutionary New Centromere (ENC) events (Figure 2).

The main rearrangements between Phaseolinae, Glycininae and *M. truncatula* as inferred from comparison with the APK

All the GBs were conserved in *S. suberectus*, *A. edgeworthii* (Glycininae) and *M. truncatula* (Figure 1A), with a possible exception of J in *S. suberectus*, which was detected in scaffolds not in the pseudomolecules. Three GB associations were shared between *S. suberectus* and *M. truncatula*: B3b+M3a (APK6: *Ss*9, *Mt*3), B3a+M3b+R8c (APK1: *Ss*4, *Mt*7) and M3c+R8b (APK9: *Ss*1, *Mt*2), R8b putatively centromeric. However, these associations were not observed within the Phaseolinae species. The N+O association (APK7: *Pv*7, *Pl*7, *Vu*7, *Ss*7, *Mt*1) and sub-GBs in Q (APK8: *Pv*8, *Pl*8, *Vu*8, *Ss*6, *Ae*3, *Mt*5) were highly syntenic between the subtribes, with the sub-GBs in Q highly collinear among the analysed species, with an exclusive inversion on Q2 in *Phaseolus* (Figure 1A). The N+O association was also maintained among species, except for *A. edgeworthii* (N: *Ae*1 and O: *Ae*8), with O7 centromeric in *Ae*8, *Mt*1 and *Ss*7. Furthermore, some centromeric GBs of APK were conserved between Phaseolinae and Glycininae and even in *M. truncatula*: A (APK8: *Ae*6, *Ss*6, *Vu*8), G (APK2: *Ae*11, *Mt*5, *Vu*2), I4 (APK4: *Pv*4, *Pl*4 and *Ss*9), K (APK5: *Ae*1, *Ss*8, eventually *Vu*1), M2-R8a (APK6: *Ae*10, *Vu*8 and *Ss*1), P2 (APK1: *Ae*9, *Ss*4 and *Pv*8), and V (APK11: *Ae*4, *Mt*8, *Ss*2 and *Vu*11) (Figure 2).

Comparison between Glycininae and Phaseolinae subtribes showed that almost all chromosomes were involved in breaks of synteny and/or collinearity, which led to a higher number of sub-GBs, especially in *A. edgeworthii*. Although this species maintained the ancestral chromosome number $n = 11$, several rearrangements lead to complex GB associations (Figure 1.A7). Nevertheless, we were able to find APK associations that helped to unveil the main translocations in this genome (Supplementary Figure 6). On the other hand, despite the descending dysploidy to $n = 9$, the *S. suberectus* genome showed fewer rearrangements compared to APK than the Phaseolinae species (Figure 1.A6). All the 22 GBs were

detected in *M. truncatula* genome (diverged from *P. vulgaris* ~50 Mya), with *Mt*1 (N+O) and *Mt*6 (I+J) highly syntenic to Phaseolinae chromosomes 7 and 4, respectively (Figure 1A).

Chromosome number reduction in *S. suberectus* compared to the APK

Based on the APK, we proposed the main chromosomal rearrangements leading to the descending dysploidy in *S. suberectus* (n = 11 to n = 9). They involved six APK chromosomes (APK2, APK4, APK5, APK6, APK9 and APK11), resulting in four *S. suberectus* chromosomes (*Ss*1, *Ss*2, *Ss*8 and *Ss*9) (Figure 1.B6). APK 4, 5, 6 and 9 were involved in a complex translocation, originating *Ss*1, *Ss*9 and *Ss*8. The whole APK2 and APK11 were combined by a translocation with terminal breakpoints, resulting in *Ss*2, followed by centromere loss in the G block. Additional reciprocal translocations have occurred between APK1 and APK10 generating *Ss*3 and *Ss*4, and between APK3 and APK8, resulting in *Ss*5 and *Ss*6. Only the APK7 is conserved in *S. suberectus*.

Oligo-FISH in the two non-sequenced species *M. atropurpureum* and *L. purpureus*

To further investigate chromosome evolution within the Phaseolinae subtribe, we selected *M. atropurpureum* (*Ma*) and *L. purpureus* (*Lp*), two species with no assembled genome. We hybridized two oligopaiting probes from *P. vulgaris* chromosomes, *Pv*2 (green) and *Pv*3 (red) to *M. atropurpureum* and *L. purpureus* metaphase cells. The oligo-FISH painting did not reflect the patterns that would be expected for APK (Figure 2). *Macroptilium atropurpureum* (Figure 2c) and *L. purpureus* (Figure 2d) showed oligo-FISH signals more similar to *V. unguiculata*, with translocations between chromosomes 2 and 3 (Figure 2b), corroborating APK predictions. *Macroptilium atropurpureum* ortholog chromosome 2 (*Ma*2) presented the short arm in red (as for *Pv*3) and almost the entire long arm in green (as for *Pv*2), except for an interstitial red region close to the pericentromere. However, *Ma*3 presented the short arm and around half of the long arm in red (ortholog to *Pv*3). The distal long arm region was painted in green (ortholog to *Pv*2), while the *Vu*3 short arm was painted in green, and almost the entire long arm red. In *L. purpureus* (Figure 2d), one arm of chromosome 2 (*Lp*2) showed small green (*Pv*2) signals intermingled with red (*Pv*3) signals, while the opposite arm was all painted in green (*Pv*2). On *Lp*3, the oligo-FISH signals were similar to *Vu*3, with small intermingled green signal in the long red arm (Figure 3b).

Our data support the exclusivity of the translocation event between APK2 and APK3 for the genus *Phaseolus* since *Lablab* and *Macroptilium* chromosomes 2 and 3 resemble those of *Vigna* and are closer to the APK. However, gaps in the pericentromeric regions of *M. atropurpureum* and *L. purpureus* orthologous chromosomes 2 and 3, different chromosome arm sizes in *Ma*3 and intermingled oligo-FISH signals (represented by green and red arrows in Figure 3), may indicate independent rearrangements and small breaks of collinearity, possibly related to inversions.

DISCUSSION

Here we established a GB/sub-GB system for comparative chromosome analyses of Phaseoleae legumes. Our system detected several chromosome rearrangements, most of them described for the first time, as well as frequent centromere repositioning, especially in the *Phaseolus* lineage. The GB system was also applicable for *Medicago*, from distantly related tribe, suggesting its use to unveil chromosome evolution in a wide range of legumes with sequenced genomes. The sub-GBs revealed further rearrangements inside the GBs as independent events during evolution. In Brassicaceae, the identification of independent rearrangements inside the GBs were essential for understanding phylogenetic relationships in different taxa, such as *Aethionema arabicum* (L.) Andrz. ex DC. (Walden *et al.*, 2020), *Arabis alpina* L. (Willing *et al.*, 2015) and *Brassica oleracea* L. (Parkin *et al.*, 2014). The proposed GB system might be useful in future phylogenetic analyses in the Leguminosae family.

Based on the GBs analyses, we reconstructed an ancestral Phaseoleae karyotype (APK) with 2*n* = 22 chromosomes. Despite the chromosome number variation inside the tribe (2*n* = 18 to 2*n* = 84; Rice *et al.*, 2015), genomic, cytogenetic and phylogenetic evidence suggest as ancestral chromosome number *n* = 11 (Li *et al.*, 2013; Rice *et al.*, 2015; Wang *et al.*, 2017). Reconstruction of ancestral karyotypes are essential for comparative genomic analyses. For grasses, cucurbits, crucifers and other flowering plants, ancestral karyotype models contributed greatly to discuss chromosome number variation, genome reshuffling and recombination hotspots (Murat *et al.*, 2010; Lysak *et al.*, 2016; Murat *et al.*, 2017; Xie *et al.*, 2019). More recently, an ancestral karyotype of *Cucumis* was inferred by comparative oligo-painting (COP) in different species of African and Asian clades, indicating constant genome reshuffling caused by large-scale inversions, centromere repositioning, and other rearrangements (Zhao *et al.*, 2021). Our analyses showed similar results, with highly conserved macrosynteny in Phaseoleae tribe, and particular rearrangements in each clade.

Five APK chromosomes (APK2, APK3, APK4, APK5 and APK7) showed high conservation of synteny within the tribe, as observed in previous studies (Schmutz *et al.*, 2010; McConnell *et al.*, 2010; Wang *et al.*, 2017; Ho *et al.*, 2017; Lonardi *et al.*, 2019). Overall, APK7 is the most conserved chromosome. Its "N+O" GB association was conserved in most analysed genomes, with only intrachromosomal rearrangements, except for *A. edgeworthii*, which displayed, in general, a higher number of rearrangements. Chromosome 7 also showed high conservation of synteny when compared to non-Phaseoleae species, such as *Mt*1, *Ah*9 (*Arachis hypogaea* L. chromosome 9) and *Lj*5 [*Lotus japonicus* (Regel) K.Larsen chromosome 5] (Bertioli *et al.*, 2009), as well as to soybean *Gm*10 and *Gm*20 chromosomes (Schmutz *et al.*, 2014; Wang *et al.*, 2017). The meaning of this conservation is not yet clear.

Few translocations and a large number of inversions were detected within Phaseolinae, some of which were previously identified by BAC-FISH, oligo-FISH and comparative genomics in *P. vulgaris*, *P. lunatus* and *V. unguiculata* (Bonifácio *et al.*, 2012; Vasconcelos *et al.*, 2015; Lonardi *et al.*, 2019; Oliveira *et al.*, 2020; do Vale Martins *et al.*, 2021; Garcia *et al.*, 2021; Bustamante *et al.*, 2021). Based on our APK and oligo-FISH approaches, we can propose the direction of these rearrangements in a

phylogenetic context. The reciprocal translocation between chromosomes 1 and 8, 2 and 3, and inversions on chromosomes 2 and 3 were exclusive events of *Phaseolus* genus, while the translocation between chromosomes 1 and 5, and the inversions on chromosome 4 were exclusive of *Vigna*. Independent intrachromosomal translocations and inversions in *P. vulgaris* (*Pv*1, 2, 3, 7, 9 and 10) and *P. lunatus* (*Pl*1, 2 and 7) were also detected. Translocations and inversions are important drivers of speciation, reducing meiotic recombination (Noor *et al.*, 2001; Rieseberg, 2001; Faria and Navarro, 2010; Feulner and De-Kayne, 2017), and observed as the main chromosomal rearrangement during *Musa* spp. evolution (Martin *et al.*, 2020). Chromosomal inversions in *Drosophila persimilis* and *D. pseudoobscura*, for instance, were associated with reproductive isolation (Fuller *et al.*, 2018).

Several translocations and inversions were detected in the Glycininae subtribe. Although the ancestral chromosome number was conserved in *A. edgeworthii*, its karyotype was the most rearranged among the analysed species. Whole genome analysis indicated large-scale ectopic recombination and reduction of Long Terminal Repeats (LTR) retrotransposons during evolution, compacting the genome but preserving important genes (Liu *et al*., 2020). The genomic analysis also indicated that remarkable genome reshuffling was not a consequence of polyploidy (Liu *et al*., 2020). Indeed, none of the GBs were duplicated, but multiple GBs were subdivided due to the extensive rearrangements. On the other hand, the dysploid *S. suberectus* species presented the most conserved karyotype. We proposed two major translocations events combining APK6 and 9 into *Ss*1, and APK2 and APK11 into *Ss*2, leading to the descending dysploidy in *S. suberectus* (from $n = 11$ to $n = 9$). Centromeres of APK6 and 9 were probably fused in a Robertsonian translocation, while the centromere of APK2 was probably eliminated. Descending dysploidy events were also unveiled by ancestral karyotype models in Brassicaceae (Lysak *et al.*, 2006; Schranz *et al.*, 2006; Lysak *et al.*, 2016) and in the wild *C. sativus* var. *hardwickii*, from $n = 12$ to $n = 7$ (Yang *et al.*, 2014).

Despite the uncertain centromere position of *S. suberectus* pseudomolecules in the genome assembly, several GB-centromeric associations were preserved among Phaseoleae and in other legume species as observed for the GBs A (APK8), G (APK2), I4 (APK4), K (APK5), M2-R8a (APK6), P2 (APK1) and V (APK11). Nevertheless, our GBs system revealed that centromere repositioning was not associated with other rearrangements or chromosome number reduction, especially in Phaseolinae. We confirmed previously inferred centromere repositioning for five chromosome pairs between *P. vulgaris* and *V. unguiculata*: chromosomes 2, 3, 5, 7 and 9 (do Vale Martins *et al*., 2021; Bustamante *et al.*, 2021). Additionally, we identified centromere repositioning for the remaining six chromosome pairs. Thus, our results indicated repositioning for all centromeres between these two closely related species, separated by ~10 Mya (Li *et al.*, 2013). Apart from the one putative event in the ancestral of both genera (in Chr. 9), nine repositioning events occurred after *P. vulgaris* separation, and five in the *Vigna unguiculata* lineage.

P. vulgaris and *V. unguiculata* have different centromeric satDNA families and more than one centromeric tandem repeat sequence in each species (Iwata *et al.*, 2013; Iwata-Otsubo *et al.*, 2016). In addition, repetitive sequences showed different amplification dynamics within *Phaseolus* and between *Phaseolus*, *Vigna* and *Cajanus*, indicating fast evolution of centromere repeats in Phaseoleae (Iwata *et al*., 2013; Iwata-Otsubo *et al.*, 2016; Ribeiro *et al*., 2017; Ribeiro *et al.*, 2020). As demonstrated by our GB

analyses, repositioned centromeres were not associated to changes in macro-collinearity. These changes may have occurred via invasion of retroelements and tandem repeats and recruitment of the centromeric protein CENH3 to a new position, forming a functional 'Evolutionary New Centromere' - ENC (Schubert, 2018; Talbert and Henikoff, 2020), as observed in *Solanum* L. (Gong *et al.*, 2012; Zhang *et al.*, 2014), *Oryza* L. (Liao *et al.*, 2018), *Aubrieta* Adans. and *Draba* L. (Mandáková *et al.*, 2020). In maize (*Zea mays* L.), for example, centromere repositioning occurred after DBS in centromeric satellite sequences that led to loss of *CentC*, and resulted in neocentromere formation at linked genes (selected during domestication), which facilitate the emergence of ENC (Schneider *et al.*, 2016; Zhao *et al.*, 2017).

While centromere repositioning in Phaseoleae (especially in Phaseolinae) seems to result from ENC events, chromosome rearrangements, such as inversions and translocations, might have resulted in centromere repositioning within same GB (Schubert, 2018; Talbert and Henikoff, 2020). In *S. suberectus*, two ancestral centromeres seemed involved in the dysploidy event through a Robertsonian translocation (Figure 1). The association of chromosome rearrangements and centromere repositioning was observed in *Arabis alpina* (Willing *et al.*, 2015), which preserved some associations of centromere-GBs from ACK chromosomes. Altogether, our findings suggest that *de novo* centromere formation is the main mechanism responsible for the frequent centromere repositioning between GBs observed in Phaseoleae legumes. Assembly of whole centromeres after long read sequencing will be necessary to confirm the additional role of chromosome rearrangements within GBs and/or repetitive sequence dynamics in explaining small differences not considered in the present analysis.

Our GB system enabled us to reconstruct the ancestral karyotype of the Phaseoleae tribe and to determine the direction of the main chromosomal evolutionary events related to diversification and speciation of closely and distantly related legume species. We identified genome reshuffling in legumes not associated to polyploidy. Moreover, we observed frequent centromere repositioning in this group, especially in *Phaseolus* genus, despite its karyotype stability. These repositioning likely involved the emergence of ENC. Altogether, our results open a bright perspective for future genomic analysis within the Phaseoleae tribe and other legume species. With an increasing number in legume genome sequencing, this genomic tool provides new perspectives for understanding the events that shaped the actual legume genomes, as well as the role of centromere repositioning for plant chromosome evolution.

EXPERIMENTAL PROCEDURES

Genomic data sets

For genomic analysis, we selected the reference genomes of *P. vulgaris* 'G19833' (GenBank ID: 8715468, Schmutz *et al.*, 2014), *P. lunatus* 'G27455' (GenBank ID: 20288068, Garcia *et al.*, 2021), *V. unguiculata* 'IT97K-499-35' (GenBank ID: 8372728, Lonardi *et al.*, 2019), *S. suberectus* 'SS-2018' (GenBank ID: 8715468, Qin *et al.*, 2019) and *A. edgeworthii* 'Qianfo Mountain' (GenBank ID: 22470258, Liu *et al.*, 2020). Moreover, we used the *M. truncatula* 'Jemalong A17' (GenBank ID: 7445598, Pecrix *et al.*, 2018) genome as an outgroup.

Genomic blocks definition based on *P. vulgaris* and *V. unguiculata* genome comparison

To find the syntenic blocks between the *P. vulgaris* and *V. unguiculata* genomes, we used the CoGe SynMap platform (https://genomevolution.org/coge/SynMap.pl) (Lyons *et al.*, 2008), following the steps and parameters described by Walden *et al.* (2020): 1) the orthologs genes were identified by using BlastZ tool; 2) synteny analysis was performed by using DAGChainer, with 25 genes as the maximum distance between two matches (-D), and 20 genes as the minimum number of aligned pairs (-A); 3) Quota Align Merge was used to merge the syntenic blocks, with 50 genes as the maximum distance between them; and 4) the ortholog and paralog blocks were differentiated based on the synonymous substitution rate (Ks) by CodeML (where 2 was the maximum value of log10) and different dotplot colours.

The syntenic blocks were defined using the 'Final syntenic gene-set output with GEvo link' (Supplementary Table 1), with start and end of the blocks represented by the gene IDs of each species. Each generated genomic block (GB) contained at least 20 genes, and one of the two following main criteria were used to separate adjacent blocks: 1) presence of translocations break points; and/or 2) genic regions on the same chromosome separated by at least 10 Mb away.

Based on the gene orthology between *P. vulgaris* and *V. unguiculata*, we compared the *P. vulgaris* genome with *P. lunatus*, *S. suberectus*, *A. edgeworthii*, and *M. truncatula* genomes following the parameters described above. These further comparisons generated sub-genomic blocks (sub-GBs) based on the following criteria: 1) segments of the same block split at least 3 Mb apart; 2) breaks of collinearity inside the GBs (inversions); and 3) breaks of collinearity or synteny involving association to other GB (inter and intra-chromosomal translocations). Some sub-GBs were not found with standard dotplot settings. In these cases, we used CoGe blastn for finding these blocks (https://genomevolution.org/CoGe/CoGeBlast.pl), with default parameters: 1e-5 for E-value and 1, -2 for nucleotides Match/Mismatch scores.

The centromeric regions were defined based on the centromeric data available in the genome assemblies (Schmutz *et al.*, 2014; Pecrix *et al.*, 2018; Lonardi *et al.*, 2019; Liu *et al.*, 2020; Garcia *et al.*, 2021). Because the centromere positions for *S. suberecuts* were not indicated in the genome assembly, we hypothesized the centromere region according to the peaks of TE accumulations along the chromosomes (Qin *et al.*, 2019). Finally, we standardize the centromere regions using the mid-point of each centromeric region, and considering that each centromere represents ~ 2Mb.

Plant material and chromosome preparation

For cytogenetics analyses, we used *P. vulgaris* 'BAT 93' (Embrapa Recursos Genéticos e Biotecnologia - Cenargen, Brasília, Distrito Federal, Brazil), *V. unguiculata* 'BR14 Mulato' (Embrapa Meio-Norte, Teresina, Piauí, Brazil), *M. atropurpureum* (International Center for Tropical Agriculture, CIAT 4413) and *L. purpureus* (UFP87699). Root tips from germinated seeds were collected and pre-treated with 2 mM 8-hydroxyquinoline for 5 h at 18º C, fixed in methanol or ethanol: acetic acid (3:1 v/v)

for 2-24 h at room temperature and stored at -20 ºC until use. For chromosome preparation, the roots were washed twice with distilled water, digested with an enzymatic solution containing 2% pectolyase (Sigma-Aldrich), 4% cellulase (Onozuka or Sigma-Aldrich) and 20% pectinase (Sigma-Aldrich) for 1-2 h at 37ºC (humid chamber). Slides were prepared following the air dry protocol (Carvalho and Saraiva, 1993) with minor modifications.

Oligo-FISH, image acquisition and data processing

The design, synthesis and labelling of the *Pv*2 and *Pv*3 oligo probes was described by do Vale Martins *et al*. (2021). Oligo-FISH was carried out according to Han et al. (2015), with minor changes. The hybridization mixture consisted of 50% formamide, 10% dextran sulphate, 2× saline sodium citrate (SSC), 350 ng of the biotin-labelled probe (*Pv*2, green) and 300 ng of the digoxigenin-labelled probe (*Pv*3, red), in a total volume of 10 µL per slide. The hybridization mix was applied to the slides for 5 min at 75°C and hybridized for 2-3 days at 37°C. *Pv*2 and *Pv*3 oligo probes were detected with anti-biotin fluorescein (Vector Laboratories) and anti-digoxigenin rhodamine (Roche), respectively, both diluted in 1× TNB (1M Tris HCl pH 7.5, 3 M NaCl-blocking reagent, Sigma-Aldrich) with posterior incubation for 1 h at 37ºC. Chromosomes were counterstained with 2 µg/mL DAPI in Vectashield antifade solution (Vector Laboratories). Oligo-FISH images were captured with a Hamamatsu CCD camera attached to an Olympus BX51 epifluorescence microscope or with Leica DM5500B fluorescence microscope. The images were uniformly adjusted and optimized for brightness and contrast using Adobe Photoshop CC (2019). *Phaseolus vulgaris* and *V. unguiculata* idiograms were assembled based on Vasconcelos et al. (2015).

ACCESSION NUMBERS

Phaseolus vulgaris 'G19833' (GenBank ID: 8715468, Schmutz *et al*., 2014)

Phaseolus lunatus 'G27455' (GenBank ID: 20288068, Garcia *et al*., 2021),

Vigna unguiculata 'IT97K-499-35' (GenBank ID: 8372728, Lonardi *et al*., 2019),

Spatholobus suberectus 'SS-2018' (GenBank ID: 8715468, Qin *et al*., 2019)

Amphicarpaea edgeworthii ' Qianfo Mountain' (GenBank ID: 22470258, Liu *et al*., 2020)

Medicago truncatula 'Jemalong A17' (GenBank ID: 7445598, Pecrix *et al*., 2018)

ACKNOWLEDGEMENTS

We thank Embrapa Meio-Norte (Teresina, Piauí, Brazil), Embrapa Cenargen (Brasília, Distrito Federal, Brazil), CIAT (International Center for Tropical Agriculture) and Prof. Marcelo Guerra (UFPE) for providing the *V. unguiculata*, *P. vulgaris*, *M. atropurpureum* and *L. purpureus* seeds, respectively. We

thank Ingo Schubert (IPK) and André Marques (MPIPZ) for critical early review for the manuscript. We also thank FACEPE (Fundação de Amparo à Ciência e Tecnologia do Estado de Pernambuco, grant no. IBPG-1520-2.03/18) and CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico, grant no. 310804/2017-5 and 313944/2020-2) for financial support.

AUTHOR CONTRIBUTIONS

C.M.: conducted the comparison between genomes, defined the blocks, conduced the oligo-FISH painting experiments in *M. atropurpureum* and *L. purpureus*, prepared the images and wrote the original draft of the manuscript. L.V.M: synthesized the oligo painting probes, conduced the oligo-FISH in *P. vulgaris* and *V. unguiculata* and constructed the oligo-FISH image. F.O.B: provided the resources for this research and discussed the data. A.C.B.V: co-supervised the experiments and contributed with data analyses and discussion. A.P.H: conceptualized and supervised the experiments and provided resources for this research. All authors reviewed the manuscript.

CONFLICT OF INTEREST

The authors declare no conflicts of interest.

ETHICS APPROVAL

Not applicable

DATA AVABIALITY

All data generated or analysed during this study are included as supplementary materials.

SUPPORTING INFORMATION

Supplementary Figure 1. Dotplot of genome comparison between *P. vulgaris* and *V. unguiculata* with the indication of each GB coloured based on the APK.

Supplementary Figure 2. Dotplot of genome comparison between *P. vulgaris* and *P. lunatus* with the indication of each GB coloured based on the APK.

Supplementary Figure 3. Dotplot of genome comparison between *P. vulgaris* and *S. subectrus* with the indication of each GB coloured based on the APK.

Supplementary Figure 4. Dotplot of genome comparison between *P. vulgaris* and *A. edgeworthii* with the indication of each GB coloured based on the APK.

Supplementary Figure 5. Dotplot of genome comparison between *P. vulgaris* and *M. truncatula* with the indication of each GB coloured based on the APK.

Supplementary figure 6. Schematic representation of the most conserved GBs associations of *A. edgeworthii* karyotype as inferred from comparison with the APK. Despite extensive genome reshuffling,

the main GB associations involved in the formation of each *Ae* chromosome are indicated by dotted lines in the corresponding APK chromosome colours.

Supplementary Table 1. CoGe SynMap pipeline output of final syntenic gene set (paralogous and orthologous) with GEvo links.

Supplementary Table 2. Genomic blocks in APK, *P. vulgaris*, *P. lunatus*, *V. unguiculata*, *S. suberectus*, *A. edgeworthii* and *M. truncatula*. **Chr**: Chromosome; **GB**: Genomic Block; **SGB**: Sub Genomic Block; **GP**: Genome Position in Mbp; **CP**: Centromere position in Mpb; **CB**: Centromere block; ↑: Inverted orientation based on APK; *: GBs detected by CoGe BLAST

REFERENCES

Bonifácio, E.M., Fonsêca, A., Almeida, C., Santos, K.G.B. dos and Pedrosa-Harand, A. (2012) Comparative cytogenetic mapping between the lima bean (*Phaseolus lunatus* L.) and the common bean (*P. vulgari*s L.). *Theor Appl Genet*, **124**, 1513–1520. https://doi.org/10.1007/s00122-012-1806-x

Bustamante, F.O., Nascimento, T.H. do, Montenegro, C., Dias, S., do Vale Martins, L., Braz, G.T., Benko-Iseppon, A.M., Jiang, J., Pedrosa-Harand, A. and Brasileiro-Vidal, A.C. (2021) Oligo-FISH barcode in beans: a new chromosome identification system. *Theor Appl Genet*, **in press.** https://doi.org/10.1007/s00122-021-03921-z

Brookes, G. and Barfoot, P. (2014) Economic impact of GM crops. *GM Crops & Food*, **5**, 65–75. https://doi.org/10.4161/gmcr.28098x

Carvalho, C.R. and Saraiva, L.S. (1993) An Air Drying Technique for Maize Chromosomes without Enzymatic Maceration. *Biotechnic & Histochemistry*, **68**, 142–145. https://doi.org/10.3109/10520299309104684

Chen, F., Dong, W., Zhang, J., Guo, X., Chen, J., Wang, Z., Lin, Z., Tang, H. and Zhang, L. (2018) The Sequenced Angiosperm Genomes and Genome Databases. *Front. Plant Sci.*, **9**, 418. https://doi.org/10.3389/fpls.2018.00418

Cheng, F., Wu, J. and Wang, X. (2014) Genome triplication drove the diversification of *Brassica* plants. *Horticulture Research*, **1**, 1–8. https://doi.org/10.1038/hortres.2014.24

do Vale Martins, L. do V., Oliveira Bustamante, F. de, Silva Oliveira, A.R. da, et al. (2021) BAC- and oligo-FISH mapping reveals chromosome evolution among *Vigna angularis*, *V. unguiculata*, and *Phaseolus vulgaris*. *Chromosoma*. https://doi.org/10.1007/s00412-021-00758-9

Faria, R. and Navarro, A. (2010) Chromosomal speciation revisited: rearranging theory with pieces of evidence. *Trends in Ecology & Evolution*, **25**, 660–669. https://doi.org/10.1016/j.tree.2010.07.008

Ferraz, M.E., Fonsêca, A. and Pedrosa-Harand, A. (2020) Multiple and independent rearrangements revealed by comparative cytogenetic mapping in the dysploid Leptostachyus group (*Phaseolus* L., Leguminosae). *Chromosome Res*. https://doi.org/10.1007/s10577-020-09644-z

Feulner, P. and De-Kayne, R. (2017) Genome evolution, structural rearrangements and speciation. *Journal of Evolutionary Biology*, **30**, 1488–1490. https://doi.org/10.1111/jeb.13101

Fonsêca, A., Ferraz, M.E. and Pedrosa-Harand, A. (2016) Speeding up chromosome evolution in *Phaseolus*: multiple rearrangements associated with a one-step descending dysploidy. *Chromosoma*, **125**, 413–421. https://doi.org/10.1007/s00412-015-0548-3

Fonsêca, A. and Pedrosa-Harand, A. (2013) Karyotype stability in the genus *Phaseolus* evidenced by the comparative mapping of the wild species *Phaseolus microcarpus*. *Genome*, **56**, 9. https://doi.org/10.1139/gen-2013-0025

Fuller, Z.L., Leonard, C.J., Young, R.E., Schaeffer, S.W. and Phadnis, N. (2018) Ancestral polymorphisms explain the role of chromosomal inversions in speciation. *PLOS Genetics*, **14**, e1007526. https://doi.org/10.1371/journal.pgen.1007526

Garcia, T., Duitama, J., Zullo, S.S., et al. (2021) Comprehensive genomic resources related to domestication and crop improvement traits in Lima bean. *Nature Communications*, **12**, 702. https://doi.org/10.1038/s41467-021-20921-1

Geiser, C., Mandáková, T., Arrigo, N., Lysak, M.A. and Parisod, C. (2016) Repeated Whole-Genome Duplication, Karyotype Reshuffling, and Biased Retention of Stress-Responding Genes in Buckler Mustard. *The Plant Cell*, **28**, 17–27. https://doi.org/10.1105/tpc.15.00791

Gong, Z., Wu, Y., Koblízková, A., et al. (2012) Repeatless and repeat-based centromeres in potato: implications for centromere evolution. *Plant Cell*, **24**, 3559–3574. https://doi.org/10.1105/tpc.112.100511

Ho, W.K., Chai, H.H., Kendabie, P., Ahmad, N.S., Jani, J., Massawe, F., Kilian, A. and Mayes, S. (2017) Integrating genetic maps in bambara groundnut [*Vigna subterranea* (L) Verdc.] and their syntenic relationships among closely related legumes. *BMC Genomics*, **18**, 192. https://doi.org/10.1186/s12864-016-3393-8

Iwata, A., Tek, A.L., Richard, M.M.S., et al. (2013) Identification and characterization of functional centromeres of the common bean. *The Plant Journal*, **76**, 47-60. http://doi.wiley.com/10.1111/tpj.12269

Iwata-Otsubo, A., Lin, J.-Y., Gill, N. and Jackson, S.A. (2016) Highly distinct chromosomal structures in cowpea (Vigna unguiculata), as revealed by molecular cytogenetic analysis. *Chromosome Res*, **24**, 197–216. https://doi.org/10.1007/s10577-015-9515-3

Jiao, Y., Wickett, N.J., Ayyampalayam, S., et al. (2011) Ancestral polyploidy in seed plants and angiosperms. *Nature*, **473**, 97–100. https://doi.org/10.1038/nature09916

Kreplak, J., Madoui, M.-A., Cápal, P., et al. (2019) A reference genome for pea provides insight into legume genome evolution. *Nature Genetics*, **51**, 1411–1422. https://doi.org/10.1038/s41588-019-0480-1

Li, C., Lin, F., An, D., Wang, W. and Huang, R. (2018) Genome Sequencing and Assembly by Long Reads in Plants. *Genes*, **9**, 6. https://doi.org/10.3390/genes9010006

Li, H., Wang, W., Lin, L., Zhu, Xiangyun, Li, J., Zhu, Xinyu and Chen, Z. (2013) Diversification of the phaseoloid legumes: effects of climate change, range expansion and habit shift. *Frontiers in Plant Science*, 9. https://doi.org/10.3389/fpls.2013.00386

Liao, Y., Zhang, X., Li, B., et al. (2018) Comparison of *Oryza sativa* and *Oryza brachyantha* Genomes Reveals Selection-Driven Gene Escape from the Centromeric Regions. *Plant Cell*, **30**, 1729–1744. https://doi.org/10.1105/tpc.18.00163

Liu, Y., Zhang, X., Han, K., et al. (2020) Insights into amphicarpy from the compact genome of the legume *Amphicarpaea edgeworthii*. *Plant Biotechnol J*, pbi.13520. https://doi.org/10.1111/pbi.13520

Lonardi, S., Muñoz□Amatriaín, M., Liang, Q., et al. (2019) The genome of cowpea (*Vigna unguiculata* [L.] Walp.).The Plant Journal, **98**, 767–782. https://doi.org/10.1111/tpj.14349

Lyons, E., Pedersen, B., Kane, J., et al. (2008) Finding and Comparing Syntenic Regions among *Arabidopsis* and the Outgroups Papaya, Poplar, and Grape: CoGe with Rosids. *Plant Physiol.*, **148**, 1772–1781. https://doi.org/10.1104/pp.108.124867

Lysak, M.A., Berr, A., Pecinka, A., Schmidt, R., McBreen, K. and Schubert, I. (2006) Mechanisms of chromosome number reduction in Arabidopsis thaliana and related Brassicaceae species. *Proceedings of the National Academy of Sciences*, **103**, 5224–5229. https://doi.org/10.1073/pnas.0510791103

Lysak, M.A., Fransz, P.F., Ali, H.B.M. and Schubert, I. (2001) Chromosome painting in Arabidopsis thaliana: Chromosome painting in *Arabidopsis*. *The Plant Journal*, **28**, 689–697. http://doi.wiley.com/10.1046/j.1365-313x.2001.01194.x

Lysak, M.A., Mandáková, T. and Schranz, M.E. (2016) Comparative paleogenomics of crucifers: ancestral genomic blocks revisited. *Current Opinion in Plant Biology*, **30**, 108–115. https://doi.org/10.1016/j.pbi.2016.02.001

Mandáková, T., Hloušková, P., Koch, M.A. and Lysak, M.A. (2020) Genome Evolution in Arabideae Was Marked by Frequent Centromere Repositioning. *Plant Cell*, **32**, 650–665. https://doi.org/10.1105/tpc.19.00557

Mandáková, T. and Lysak, M.A. (2018) Post-polyploid diploidization and diversification through dysploid changes. *Current Opinion in Plant Biology*, **42**, 55–65. https://doi.org/10.1016/j.pbi.2018.03.001

Martin, G., Baurens, F.-C., Hervouet, C., et al. (2020) Chromosome reciprocal translocations have accompanied subspecies evolution in bananas. *The Plant Journal*, **104**, 1698–1711. https://doi.org/10.1111/tpj.15031

McConnell, M., Mamidi, S., Lee, R., Chikara, S., Rossi, M., Papa, R. and McClean, P. (2010) Syntenic relationships among legumes revealed using a gene-based genetic linkage map of common bean (*Phaseolus vulgaris* L.). *Theor Appl Genet*, **121**, 1103–1116. https://doi.org/10.1007/s00122-010-1375-9

Moussa, B., Lowenberg-DeBoer, J., Fulton, J. and Boys, K. (2011) The economic impact of cowpea research in West and Central Africa: A regional impact assessment of improved cowpea storage technologies. *Journal of Stored Products Research*, **47**, 147–156. https://doi.org/10.1016/j.jspr.2011.02.001

Murat, F., Armero, A., Pont, C., Klopp, C. and Salse, J. (2017) Reconstructing the genome of the most recent common ancestor of flowering plants. *Nature Genetics*, **49**, 490–496. https://doi.org/10.1038/ng.3813

Murat, F., Xu, J.-H., Tannier, E., Abrouk, M., Guilhot, N., Pont, C., Messing, J. and Salse, J. (2010) Ancestral grass karyotype reconstruction unravels new mechanisms of genome shuffling as a source of plant evolution. *Genome Res.*, **20**, 1545–1557. https://doi.org/10.1101/gr.109744.110

Myers, J.R. and Kmiecik, K. (2017) Common Bean: Economic Importance and Relevance to Biological Science Research. In M. Pérez de la Vega, M. Santalla, and F. Marsolais, eds. *The Common Bean Genome*. Compendium of Plant Genomes. Cham: Springer International Publishing, pp. 1–20. https://doi.org/10.1007/978-3-319-63526-2

Noor, M.A.F., Grams, K.L., Bertucci, L.A. and Reiland, J. (2001) Chromosomal inversions and the reproductive isolation of species. *Proceedings of the National Academy of Sciences*, **98**, 12084–12088. https://doi.org/10.1073/pnas.221274498

Oliveira, A.R. da S., do Vale Martins, L., Bustamante, F. de O., Muñoz-Amatriaín, M., Close, T., Costa, A.F. da, Benko-Iseppon, A.M., Pedrosa-Harand, A. and Brasileiro-Vidal, A.C. (2020) Breaks of macrosynteny and collinearity among moth bean (*Vigna aconitifolia*), cowpea (*V. unguiculata*), and common bean (*Phaseolus vulgaris*). *Chromosome Res*. https://doi.org/10.1007/s10577-020-09635-0

Parkin, I.A., Koh, C., Tang, H., et al. (2014) Transcriptome and methylome profiling reveals relics of genome dominance in the mesopolyploid *Brassica oleracea*. *Genome Biol*, **15**, R77. https://doi.org/10.1186/gb-2014-15-6-r77

Pavy, N., Pelgas, B., Laroche, J., Rigault, P., Isabel, N. and Bousquet, J. (2012) A spruce gene map infers ancient plant genome reshuffling and subsequent slow evolution in the gymnosperm lineage leading to extant conifers. *BMC Biology*, **10**, 84. https://doi.org/10.1186/1741-7007-10-84

Pecrix, Y., Staton, S.E., Sallet, E., et al. (2018) Whole-genome landscape of *Medicago truncatula* symbiotic genes. *Nature Plants*, **4**, 1017–1025. https://doi.org/10.1038/s41477-018-0286-7

Pellicer, J., Hidalgo, O., Dodsworth, S. and Leitch, I.J. (2018) Genome Size Diversity and Its Impact on the Evolution of Land Plants. *Genes*, **9**, 88. https://doi.org/10.3390/genes9020088

Qin, S., Wu, L., Wei, K., et al. (2019) A draft genome for *Spatholobus suberectus*. *Sci Data*, **6**, 113. https://doi.org/10.1038/s41597-019-0110-x

Ribeiro, T., Dos Santos, K.G.B., Richard, M.M.S., Sévignac, M., Thareau, V., Geffroy, V. and Pedrosa-Harand, A. (2017) Evolutionary dynamics of satellite DNA repeats from *Phaseolus* beans. *Protoplasma*, **254**, 791–801. https://doi.org/10.1007/s00709-016-0993-8

Ribeiro, T., Vasconcelos, E., Santos, K.G.B. dos, Vaio, M., Brasileiro-Vidal, A.C. and Pedrosa-Harand, A. (2020) Diversity of repetitive sequences within compact genomes of *Phaseolus* L. beans and allied genera *Cajanus* L. and *Vigna* Savi. *Chromosome Res*, **28**, 139–153. https://doi.org/10.1007/s10577-019-09618-w.

Rice, A., Glick, L., Abadi, S., Einhorn, M., Kopelman, N.M., Salman□Minkov, A., Mayzel, J., Chay, O. and Mayrose, I. (2015) The Chromosome Counts Database (CCDB) – a community resource of plant chromosome numbers. *New Phytol*, **206**, 19–26. https://doi.org/10.1111/nph.13191

Rieseberg, L.H. (2001) Chromosomal rearrangements and speciation. *Trends in Ecology & Evolution*, **16**, 351–358. https://doi.org/10.1016/S0169-5347(01)02187-5

Ruprecht, C., Lohaus, R., Vanneste, K., Mutwil, M., Nikoloski, Z., Peer, Y.V. de and Persson, S. (2017) Revisiting ancestral polyploidy in plants. *Science Advances*, **3**, e1603195. https://doi.org/10.1126/sciadv.1603195

Schmutz, J., Cannon, S.B., Schlueter, J., et al. (2010) Genome sequence of the palaeopolyploid soybean. *Nature*, **463**, 178–183. https://doi.org/10.1038/nature08670

Schmutz, J., McClean, P.E., Mamidi, S., et al. (2014) A reference genome for common bean and genome-wide analysis of dual domestications. *Nat Genet*, **46**, 707–713. https://doi.org/10.1038/ng.3008

Schneider, K.L., Xie, Z., Wolfgruber, T.K. and Presting, G.G. (2016) Inbreeding drives maize centromere evolution. *Proc Natl Acad Sci USA*, **113**, 987–996.

Schranz, M., Lysak, M. and Mitchellolds, T. (2006) The ABC's of comparative genomics in the Brassicaceae: building blocks of crucifer genomes. *Trends in Plant Science*, **11**, 535–542. https://doi.org/10.1016/j.tplants.2006.09.002

Schubert, I. (2018) What is behind "centromere repositioning"? *Chromosoma*, **127**, 229–234. https://doi.org/10.1007/s00412-018-0672-y

Soltis, P.S., Marchant, D.B., Van de Peer, Y. and Soltis, D.E. (2015) Polyploidy and genome evolution in plants. *Current Opinion in Genetics & Development*, **35**, 119–125. https://doi.org/10.1016/j.gde.2015.11.003

Talbert, P.B. and Henikoff, S. (2020) What makes a centromere? *Experimental Cell Research*, **389**, 111895. https://doi.org/10.1016/j.yexcr.2020.111895

The Arabidopsis Genome Initiative (2000) Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature*, **408**, 796–815. https://doi.org/10.1038/35048692

Vasconcelos, E.V., Andrade Fonsêca, A.F. de, Pedrosa-Harand, A., Andrade Bortoleti, K.C. de, Benko-Iseppon, A.M., Costa, A.F. da and Brasileiro-Vidal, A.C. (2015) Intra- and interchromosomal rearrangements between cowpea [*Vigna unguiculata* (L.) Walp.] and common bean (*Phaseolus vulgaris* L.) revealed by BAC-FISH. *Chromosome Res*, **23**, 253–266. http://doi.org/10.1007/s10577-014-9464-2

Walden, N., Nguyen, T.-P., Mandáková, T., Lysak, M.A. and Schranz, M.E. (2020) Genomic Blocks in Aethionema arabicum Support Arabideae as Next Diverging Clade in Brassicaceae. *Front. Plant Sci.*, **11**, 719. https://doi.org/10.3389/fpls.2020.00719

Wang, J., Sun, P., Li, Y., et al. (2017) Hierarchically Aligning 10 Legume Genomes Establishes a Family-Level Genomics Platform. *Plant Physiol.*, **174**, 284–300. https://doi.org/10.1104/pp.16.01981

Wendel, J.F., Jackson, S.A., Meyers, B.C. and Wing, R.A. (2016) Evolution of plant genome architecture. *Genome Biology*, **17**, 37. https://doi.org/10.1186/s13059-016-0908-1

Willing, E.-M., Rawat, V., Mandáková, T., et al. (2015) Genome expansion of Arabis alpina linked with retrotransposition and reduced symmetric DNA methylation. *Nature Plants*, **1**, 14023. https://doi.org/10.1038/nplants.2014.23

Wu, S., Han, B. and Jiao, Y. (2020) Genetic Contribution of Paleopolyploidy to Adaptive Evolution in Angiosperms. *Molecular Plant*, **13**, 59–71. https://doi.org/10.1016/j.molp.2019.10.012

Xie, D., Xu, Y., Wang, J., et al. (2019) The wax gourd genomes offer insights into the genetic diversity and ancestral cucurbit karyotype. *Nature Communications*, **10**, 5158. https://doi.org/10.1038/s41467-019-13185-3

Yang, L., Koo, D.-H., Li, D., et al. (2014) Next-generation sequencing, FISH mapping and synteny-based modeling reveal mechanisms of decreasing dysploidy in *Cucumis*.The Plant Journal, **77**, 16–30. http://doi.wiley.com/10.1111/tpj.12355

Zhang, H., Koblížková, A., Wang, K., et al. (2014) Boom-Bust Turnovers of Megabase-Sized Centromeric DNA in *Solanum* Species: Rapid Evolution of DNA Sequences Associated with Centromeres. *The Plant Cell*, **26**, 1436–1447. https://doi.org/10.1105/tpc.114.123877

Zhang, S.-J., Liu, L., Yang, R. and Wang, X. (2020) Genome Size Evolution Mediated by Gypsy Retrotransposons in Brassicaceae. *Genomics, Proteomics & Bioinformatics*, **18**, 321–332. https://doi.org/10.1016/j.gpb.2018.07.009

Zhao, H., Zeng, Z., Koo, D.-H., Gill, B.S., Birchler, J.A. and Jiang, J. (2017) Recurrent establishment of de novo centromeres in the pericentromeric region of maize chromosome 3. *Chromosome Res*, **25**, 299–311. https://doi.org/10.1007/s10577-017-9564-x

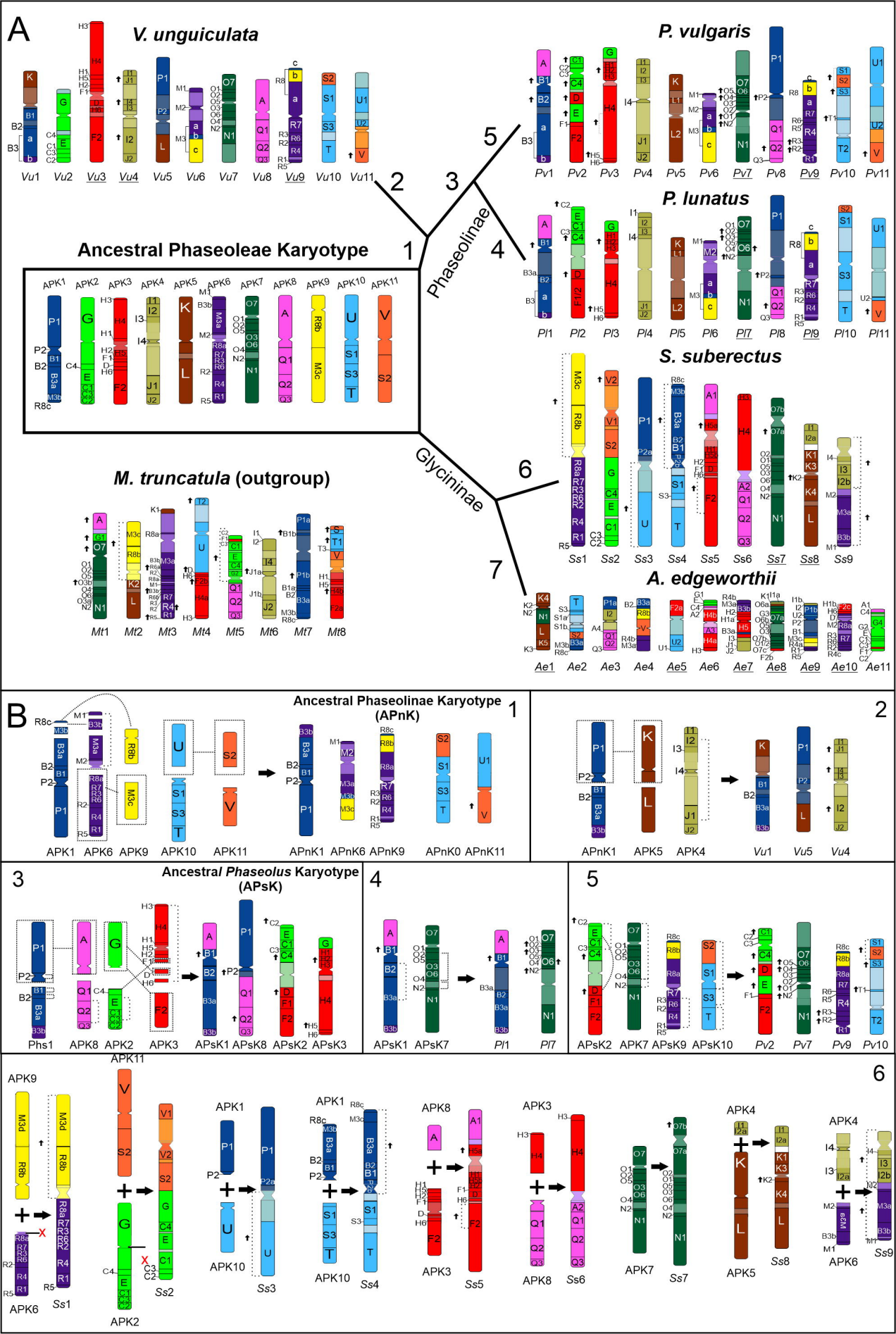
Zhao, Q., Meng, Y., Wang, P., et al. (2021) Reconstruction of ancestral karyotype illuminates chromosome evolution in the genus *Cucumis*. *The Plant Journal*. https://doi.org/10.1111/tpj.15381

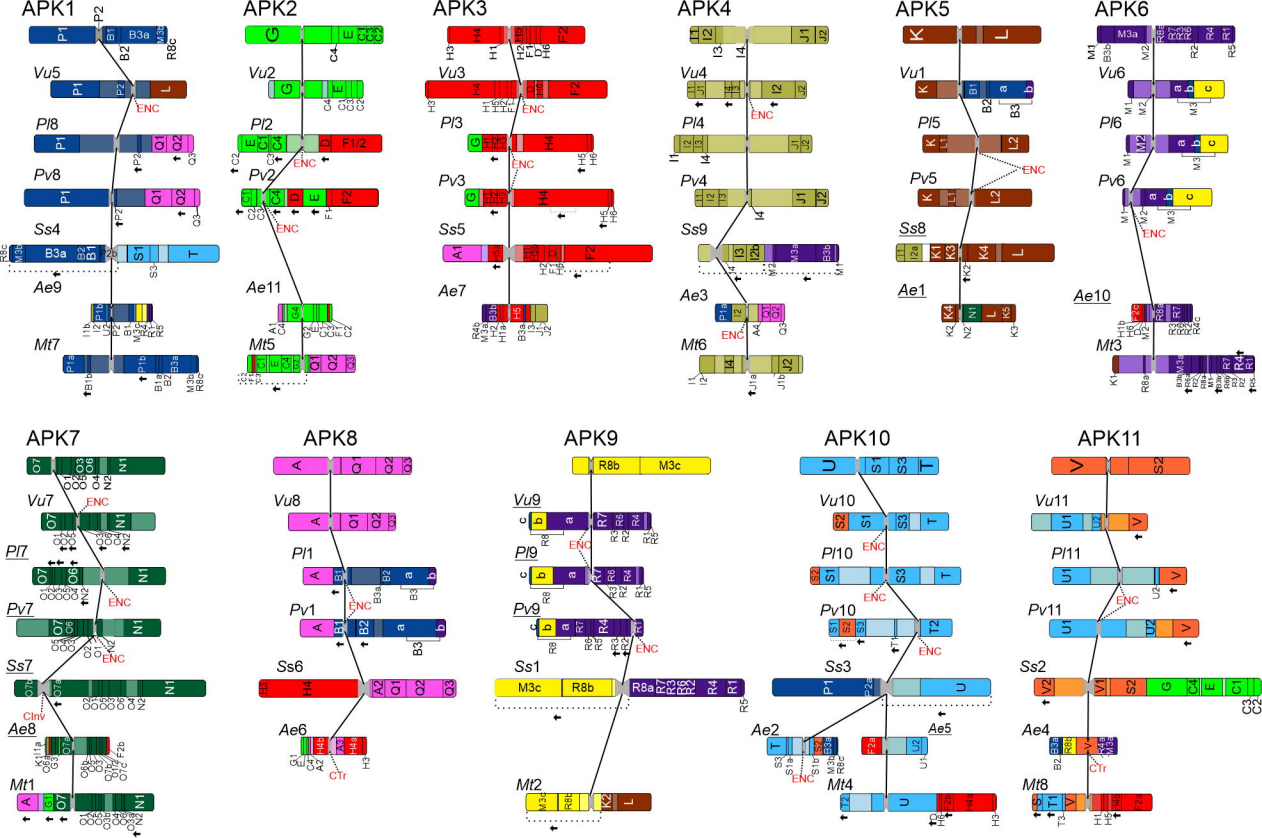
Figure legends

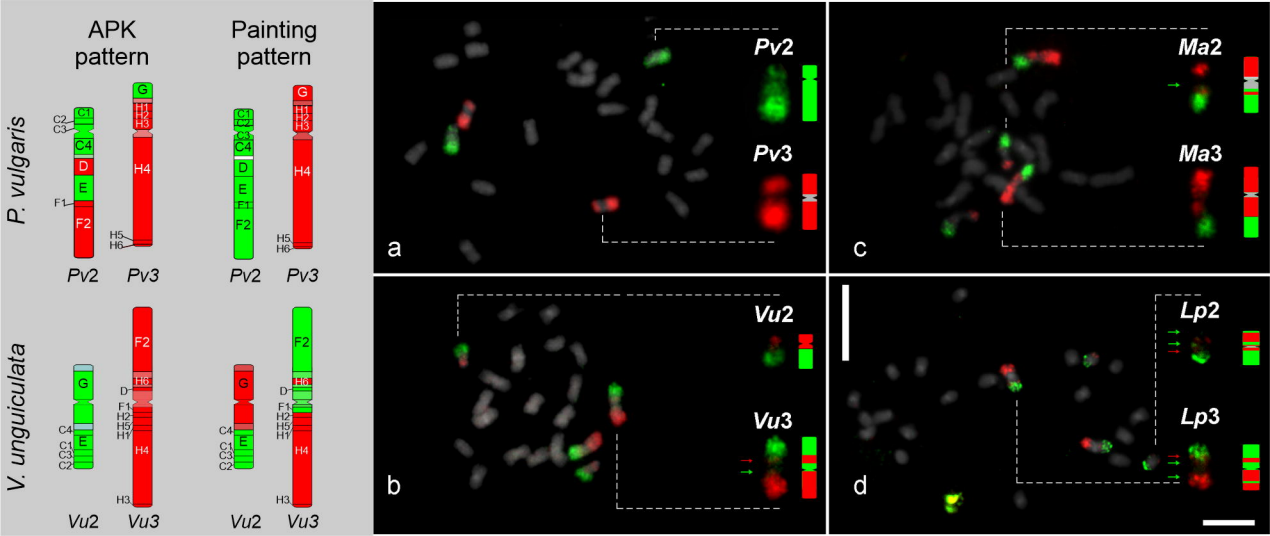
Figure 1. Overview of genomic comparisons between Phaseoleae species using *Medicago truncatula* as outgroup. The genomic block system and the hypothetical Ancestral Phaseolae Karyotype (APK) are indicated. Arrows in **A** indicate change in the block orientation. Some entire chromosome orientations were inverted (underlined chromosomes) to better demonstrate the orthology between APK and other karyotypes. **A**) Species are grouped according to the phylogenetic relationships proposed by Li *et al*. (2013). **1** to **5**) Phaseolinae subtribe species: *Vigna unguiculata*, *Phaseolus lunatus* and *P. vulgaris*; **6** and **7**) Glycininae subtribe species: *Spatholobus suberectus* and *Amphicarpea edgeworthii*. **B**) Main chromosomal rearrangements: **1**) Translocations and pericentric inversion resulting in the Ancestral Phaseolinae Karyotype (APnK) chromosomes 1, 6, 9, 10 and 11; **2**) Reciprocal translocation and pericentric inversions resulting in *V. unguiculata* (*Vu*) chromosomes 1, 4 and 5; **3**) Reciprocal translocation and inversions resulting in the Ancestral *Phaseolus* Karyotype (APsK) chromosomes 1, 2, 3 and 8; **4**) Inversions and intrachromosomal translocations resulting in *P. lunatus* (*Pl*) chromosomes 1 and 7; **5**) Inversions and intrachromosomal translocation resulting in *P. vulgaris* (*Pv*) chromosomes 2, 7, 9 and 10; **6**) Descending dysploidy and genome reorganization scenario for *S. suberectus* (*Ss*) karyotype.

Figure 2. Centromere repositioning in Phaseoleae tribe and *M. truncatula* based on the Ancestral Phaseolae Karyotype (APK) established. ENC – Evolutionary New Centromere.

Figure 3. Oligo-FISH using *Phaseolus vulgaris* chromosomes-specific 2 (green) and 3 (red) probes hybridized to other Phaseolinae species. In the left, differences between the APK and oligo-FISH painting signals of chromosomes 2 and 3. In the right, the painted mitotic metaphases of *P. vulgaris* (**a**), *Vigna unguiculata* (**b**), *Macroptilium atropurpureum* (**c**) and *Lablab purpureus* (**d**). Green and red arrows indicate small differences in the painting patterns, corresponding to specific regions of *Pv*2 (green) or *Pv*3 (red) probes, respectively. For each species, orthologous chromosomes of *P. vulgaris* chromosomes 2 and 3 are detailed in the insets (right) of each metaphase cell and represented in the lateral of each inset. Vertical bar = 5 µm.


A
V. unguiculata
Vu1 Vu2 Vu3 Vu4 Vu5 Vu6 Vu7 Vu8 Vu9 Vu10 Vu11
P. vulgaris
Pv1 Pv2 Pv3 Pv4 Pv5 Pv6 Pv7 Pv8 Pv9 Pv10 Pv11
P. lunatus
Pl1 Pl2 Pl3 Pl4 Pl5 Pl6 Pl7 Pl8 Pl9 Pl10 Pl11
S. suberectus
Ss1 Ss2 Ss3 Ss4 Ss5 Ss6 Ss7 Ss8 Ss9
A. edgeworthii
Ae1 Ae2 Ae3 Ae4 Ae5 Ae6 Ae7 Ae8 Ae9 Ae10 Ae11
Ancestral Phaseoleae Karyotype
APK1 APK2 APK3 APK4 APK5 APK6 APK7 APK8 APK9 APK10 APK11
M. truncatula (outgrowth)
Mt1 Mt2 Mt3 Mt4 Mt5 Mt6 Mt7 Mt8
1
2
3
4
5
6
7
Phaseolinae
Glycininae
B
Ancestral Phaseolinae Karyotype (APnK)
APK1 APK6 APK9 APK10 APK11 APnK1 APnK6 APnK9 APnK0 APnK11
APnK1 APK5 APK4 Vu1 Vu5 Vu4
Ancestral Phaseolus Karyotype (APsK)
Phs1 APK8 APK2 APK3 APsK1 APsK8 APsK2 APsK3
APsK1 APsK7 Pl1 Pl7
APsK2 APK7 APsK9 APsK10 Pv2 Pv7 Pv9 Pv10
APK9 APK6 Ss1
APK11 APK2 Ss2
APK1 APK10 Ss3
APK1 APK10 Ss4
APK8 APK3 Ss5
APK3 APK8 Ss6
APK7 Ss7
APK4 APK5 Ss8
APK4 APK6 Ss9


APK1
APK2
APK3
APK4
APK5
APK6
APK7
APK8
APK9
APK10
APK11


APK pattern
Painting pattern
P. vulgaris
V. unguiculata
Pv2
Pv3
Vu2
Vu3
a
b
c
d
Pv2
Pv3
Vu2
Vu3
Ma2
Ma3
Lp2
Lp3