

# Hierarchically aligning 10 legume genomes establishes a family-level genomics platform

## Supplemental text

### Supplemental Methods

#### *Inference of genomic homology*

As mentioned in the main context, homologous gene dotplotting is crucial to deconvolute the genomic complexity. We always use a reference genome that is well sequenced and assembled to understand genomes affected by extra polyploidizations. Here, for example, to decipher the homoeologous duplicated regions in barrel medic, we adopted grape genome as a reference. Therefore, a comparative dotplot was produced by using putative homologous genes between the two genomes (inferred by performing BLASTP searching at E-value =  $1e-5$ ). Due to the eudicot-common ECH, and the legume-specific LCT, we anticipated that a grape region would have 2 best-matched or orthologous barrel medic regions, and 4 outparalogous regions, which were produced by the ECH (**Supplemental Fig. S11**). In the grape-barrel medic dotplot, orthologous and outparalogous blocks can be inferred without much difficulty. A grape chromosomal region is often much more similar to its barrel medic orthologous regions than to the outparalogous regions. Chromosomal similarity was measured by colinear gene number. Some outparalogous blocks can have few homologous gene dots and can only be inferred by transitively using paralogy between grape chromosomes (**Supplemental Fig. S11**).

Here, let us show an example using the ECH produced grape chromosome triplets, being homoeologous (paralogous) chromosomes to one another: Vv9, Vv11, and part of Vv4, the latter of which acquired an extra segment from other non-homologous chromosome. Vv4 has two best-matched or orthologous regions in barrel medic, one Mt1 and the other in Mt3, with homologous gene dotting regions circled out by solid-line rectangles numbered by 1 and 2 in red circles in **Supplemental Fig. S11**. The dotting regions contain 220, 270 collinear genes inferred by ColinearScan. The Vv4's orthologous regions in Mt1 and Mt3 are each outparalogous to Vv9 and Vv11, and the expected regions in **Supplemental Fig. S11** were circled out by broken-lined rectangles numbered by 1-4 in blue circles. Few dots,

(15 collinear genes identified between Vv9 and Mt1, 27 collinear genes identified between Vv11 and Mt1, 41 collinear genes identified between Vv9 and Mt3, and 49 collinear genes identified between Vv11 and Mt3), could be found in the outparalogous blocks. However, the homoeology between Vv4, Vv9 and Vv11 provide transitive information to help identify outparalogy between grape and barrel medic chromosomes.

In similar strategy, for Vv9 and Vv11, we identified their respective orthologous regions in barrel medic, and these grape homoeologous chromosomes have different orthologous regions (**Supplemental Fig. S11**).

Although a grape chromosome would have 2 orthologous corresponding regions, they were often broken into pieces by chromosomal rearrangements. A complementary pattern of broken segments helps infer their being derived from the same ancestral chromosome. For example, Vv9 has two orthologous counterparts in barrel medic (**Supplemental Fig. S11**). One can be identified in Mt2 but the other one was broken into two pieces, being complement to one another, identified in Mt4 and Mt8. Similarly, Vv11 has two barrel medic counterparts, which were broken into pieces, however, most of them could be identified by using complement patterns (**Supplemental Fig. S11**).

### *Synonymous substitutions*

Synonymous nucleotide substitutions on synonymous sites (Ks) were estimated by using the Nei-Gojobori approach (Nei & Gojobori, 1986) implemented by using the Bioperl Statistical module.

### *Modeling gene losses*

Using the grape, barrel medic and common bean genomes as references, we counted genes in the reference genome that did not have colinear orthologs in each legume chromosomal region. The number of consecutive genes in the referenced genome that had no legume match was also counted. This ‘run length’ – that is, the number of consecutive genes in the reference genome with no matching legume gene – should follow a geometric distribution if gene loss/translocation had occurred in a random manner. For a probability  $p$  that a gene was lost at random, the probability of loss of a consecutive run of  $n - 1$  genes can be estimated by the formula:

$$y = f(x = k / p) = (1 - p)^{n-1} p$$

Fitting procedures were performed by implementing home-made Python scripts to run fitting functions curve-fit and numpy.phyfit and perform *F-tests* using R function var.test.

### ***Kernel function analysis of Ks***

Distributions of synonymous nucleotide substitutions on synonymous sites (Ks) of homologous genes could reflect multiple and overlapping genomic duplications, and speciations if the homoeologs are from different genomes. We adopted a kernel function analysis of Ks distribution of colinear homoeologs from within a legume genome, or between different genomes. A Ks distribution was viewed as a mix of multiple normal distributions. We used the kernel smoothing density function **ksdensity** (width is generally set to 0.05) in Matlab to estimate the probability density of each Ks list and obtain the density distribution curve. Then we performed the Gaussian multi-peak fitting of the curve by using the gaussian approximation function **Gaussian** in the fitting toolbox **cftool**. We set *R-squared*, a parameter to evaluate the fitting goodness, to be at least 95%, used the smallest number of normal distributions to represent the complex Ks distribution, and the principle one was used to represent the corresponding evolutionary event.

### ***Evolutionary dating correction***

By aligning the peaks of soybean-specific tetraploidization (SST) from different Ks distributions to the corresponding location in the soybean Ks distribution, we performed evolutionary rate correction. We suppose that the soybean peak appears at  $\mu_{gm}$ , and for the other legumes, supposing that the peak appears at  $\mu_i$ , the relative evolutionary rate of legume *i* can be described with

$$r = (\mu_i - \mu_{gm}) / \mu_{gm}.$$

Then we perform rate correction to find the corrected rate  $\mu_{i-correction}$  of the legume *i* relative to  $\mu_{gm}$ :

- ( 1 ) For a specific legume *i*, for the Ks between its duplicates, we can define a correction coefficient  $C_i$  as,

$$\frac{\mu_{i\text{-correction}}}{\mu_i} = \frac{\mu_{gm}}{\mu_i} = C_i$$

therefore, we get

$$\mu_{i\text{-correction}} = \frac{\mu_{gm}}{\mu_i} \times \mu_i = \frac{1}{1+r} \times \mu_i$$

and ,  $C_i = \frac{1}{1+r}$  ;

( 2 ) For Ks between homologous genes from two non-soybean legumes  $i, j$  , if the peak was located at  $\mu_{ij}$  , we used the arithmetic mean of two correction coefficients in two genomes :  $C_{ij} = (C_i + C_j)/2$  , then we calculated a corrected evolutionary rate

$$\mu_{ij\text{-correction}} = C_{ij} \times \mu_{ij} ;$$

( 3 ) For Ks between homologous genes from soybean and another legume  $i$  , if the peak was located at  $\mu_{igm}$  , supposing the correction coefficient  $C_i$  in the legume  $i$  , then we calculated a corrected evolutionary rate

$$\mu_{igm\text{-correction}} = C_i \times \mu_{igm} .$$

### ***Gene ontology analysis***

In order to get a functional overview of those genes generated by legume common tetraploidy (LCT) and soybean specific tetraploidy (SST), InterProScan 5(Jones *et al.*, 2014) was used to determine the GO classification of each gene. All records are derived from literature-based annotations and protein domain-based electronic annotations. All numbers in each Gene Ontology function category were standardized by log function.

### ***Polyploidization and NBS-LRR genes***

Legume polyploidizations may have also contributed to the expansion of NBS-LRR resistance genes, which was further subjected to a birth-death process due to ectopic

recombination. In well assembled legume genomes such as soybean, barrel medic, peanut and common bean, often more than 30-79% of NBS-LRR genes with CC domains were located in homologous regions supported by gene colinearity (**Supplemental Tables S19-24**). In soybean, 20%, 59%, and 72% of NBS-LRR genes with CC domains could be related to the ECH, LCT, and SST events, respectively; in peanut, 11-15% and 30% and in barrel medic 8% and 37% were related to the ECH and LCT events, respectively.

NBS-LRR genes were much more affected by tandem and/proximal gene duplication (**Supplemental Fig. S9 and S10**) than nodulation and oil-synthesis genes. NBS-LRR gene production hotspots on chromosomes 3, 6, 7, 15, 16, 18, and 19 were indicated based on 263 young duplicates with  $K_s < 0.15$ . Actually, ~38% (99) of them have  $K_s < 0.05$ , indicating very recent origin. In other legumes, these resistance genes show a very similar pattern of proximal localization and young origin.

## Reference

- Wang, X., Wang, J., Jin, D., Guo, H., Lee, T.H., Liu, T., and Paterson, A.H.** (2015a). Genome Alignment Spanning Major Poaceae Lineages Reveals Heterogeneous Evolutionary Rates and Alters Inferred Dates for Key Evolutionary Events. *Mol Plant* **8**, 885-898.
- Wang, X., Wang, J., Guo, H., Lee, T., Liu, T., Jin, D., and Paterson, A.H.** (2015b). Genome Alignment Spanning Major Poaceae Lineages Reveals Heterogeneous Evolutionary Rates and Alters Inferred Dates for Key Evolutionary Events. *Molecular plant* **8**, 14.
- Maere, S., De Bodt, S., Raes, J., Casneuf, T., Van Montagu, M., Kuiper, M., and Van de Peer, Y.** (2005). Modeling gene and genome duplications in eukaryotes. *Proc Natl Acad Sci U S A* **102**, 5454-5459.
- Tang, H., Bowers, J.E., Wang, X., and Paterson, A.H.** (2010). Angiosperm genome comparisons reveal early polyploidy in the monocot lineage. *Proc Natl Acad Sci U S A* **107**, 472-477.
- Bowers, J.E., Chapman, B.A., Rong, J., and Paterson, A.H.** (2003). Unravelling angiosperm genome evolution by phylogenetic analysis of chromosomal duplication events. *Nature* **422**, 433-438.
- Jaillon, O., Aury, J.M., Noel, B., Policriti, A., Clepet, C., Casagrande, A., Choisne, N., Aubourg, S., Vitulo, N., Jubin, C., Vezzi, A., Legeai, F., Huguene, P., Dasilva, C., Horner, D., Mica, E., Jublot, D., Poulain, J., Bruyere, C., Billault, A., Segurens, B., Gouyvenoux, M., Ugarte, E., Cattonaro, F., Anthouard, V., Vico, V., Del Fabbro, C., Alaux, M., Di Gaspero, G., Dumas, V., Felice, N., Paillard, S., Juman, I., Moroldo, M., Scalabrin, S., Canaguier, A., Le Clainche, I., Malacrida, G., Durand, E., Pesole, G., Laucou, V., Chatelet, P., Merdinoglu, D., Delledonne, M., Pezzotti, M., Lecharny, A., Scarpelli, C., Artiguenave, F., Pe, M.E., Valle, G., Morgante, M., Caboche, M., Adam-Blondon, A.F., Weissenbach, J., Quetier, F., Wincker, P., and French-Italian Public Consortium for Grapevine Genome, C.** (2007). The grapevine genome sequence suggests ancestral hexaploidization in major angiosperm phyla. *Nature* **449**, 463-467.
- Nei, M., and Gojobori, T.** (1986). Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. *Molecular biology and evolution* **3**, 418-426.
- Jones, P., Binns, D., Chang, H.Y., Fraser, M., Li, W.Z., McAnulla, C., McWilliam, H., Maslen, J., Mitchell, A., Nuka, G., Pesseat, S., Quinn, A.F., Sangrador-Vegas, A., Scheremetjew, M., Yong, S.Y., Lopez, R., and Hunter, S.** (2014). InterProScan 5: genome-scale protein function classification.

Bioinformatics **30**, 1236-1240.

## Supplemental tables and figure legends

### Supplemental Tables

**Supplemental Table S1.** Number of homologous blocks and gene pairs within a genome or between genomes.

Homologous Blocks within and among genome	BL <sup>a</sup> > 4	BL > 10	BL > 20	BL > 50	ACGP <sup>b</sup>	LDB <sup>c</sup>	LDB on chromosomes
<i>Vitis vinifera</i>	3,435/312	2,351/126	1,248/47	111/2	11.01, 18.66, 26.55, 55.5	61	VV05-VV07
<i>Arachis duranensis</i>	4,037/492	1,819/85	1,076/28	409/6	8.21, 21.40, 38.41, 68.17	96	AD03-AD08
<i>Arachis ipaensis</i>	4,187/534	1,808/89	1,038/29	386/6	7.84, 20.31, 35.79, 64.33	84	AI03-AI08
<i>Lotus japonicus</i>	2,822/317	1,373/63	829/23	256/4	8.90, 21.79, 36.04, 64.00-	86	LJ04-LJ06
<i>Medicago truncatula</i>	4,391/493	2,133/84	1,519/36	660/9	8.91, 25.39, 42.19, 73.33	144	MT01-MT07
<i>Cicer arietinum</i>	4,376/545	1,927/91	1,141/31	416/5	8.03, 21.18, 36.81, 83.20	145	CA07-CA08
<i>Glycine max</i>	25,302/1294	20,365/434	17,578/224	13,191/87	19.55, 46.92, 78.47, 151.62	824	GM10-GM20
<i>Vigna angularis</i>	1,956/272	656/42	241/9	-/-	7.19, 15.62, 26.78, -	39	VU07-VU11
<i>Vigna radiat</i>	4,546/609	1,686/79	980/24	403/5	7.46, 21.34, 40.83, 80.60	107	VR07-VR11
<i>Cajanus cajan</i>	2,086/317	574/39	105/4	-/-	6.58, 14.72, 26.25, -	34	CC08-CC11
<i>Phaseolus vulgaris</i>	5,336/561	2,920/124	1,888/45	903/12	9.51, 23.55, 41.96, 75.25	114	PV01-PV09
<i>V. vinifera</i> vs <i>A. duranensis</i>	14,443/1,286	9,397/377	6,589/164	2,719/37	11.23, 24.93, 40.18, 73.49	145	VV18-AD08
<i>V. vinifera</i> vs <i>A. ipaensis</i>	14,204/1,248	9,258/364	6,530/158	2,944/40	11.38, 25.43, 41.33, 73.60	145	VV18-AI07
<i>V. vinifera</i> vs <i>L. japonicus</i>	11,825/1,015	7,907/324	5,362/135	2,300/35	11.65, 24.40, 39.72, 65.71	89	VV14-LJ01
<i>V. vinifera</i> vs <i>M. truncatula</i>	14,300/1,059	10,415/389	7,445/166	3,910/52	13.50, 26.77, 44.85, 75.19	168	VV18-MT01
<i>V. vinifera</i> vs <i>C. arietinum</i>	15,765/1,518	9,639/367	6,942/168	3,237/44	10.39, 26.26, 41.32, 73.57	149	VV18-CA04
<i>V. vinifera</i> vs <i>G. max</i>	25,609/1,726	19,806/747	14,485/351	6,771/91	14.8, 26.51, 46.27, 74.41	161	VV18-GM04
<i>V. vinifera</i> vs <i>V. angularis</i>	9,201/1,153	4,178/238	1,892/62	276/5	7.98, 17.55, 30.52, 55.20	63	VV18-VU05
<i>V. vinifera</i> vs <i>V. radiata</i>	14,710/1,475	8,586/339	5,821/137	2,720/34	9.97, 25.33, 42.49, 80.00	151	VV18-VR05
<i>V. vinifera</i> vs <i>C. cajan</i>	11,583/1,344	5,787/297	3,321/107	654/11	8.62, 19.48, 31.04, 59.45	70	VV06-CC10
<i>V. vinifera</i> vs <i>P. vulgaris</i>	17,530/1,531	11,561/442	8,274/199	3,828/50	11.45, 26.16, 41.58, 76.56	172	VV09-PV10
<i>A. duranensis</i> vs <i>A. ipaensis</i>	29,116/1,889	20,753/281	18,701/122	16,484/50	15.41, 73.85, 153.29, 329.68	1,213	AD10-AI10

<i>A. duranensis</i> vs <i>L. japonicus</i>	14,965/1,029	10,900/287	91,12/152	6,197/57	14.54, 37.98, 59.95, 108.72	396	AD03-LJ03
<i>A. duranensis</i> vs <i>M. truncatula</i>	18,739/1,164	14,138/340	12,049/182	8,665/76	16.10, 41.58, 66.20, 114.01	437	AD07-MT05
<i>A. duranensis</i> vs <i>C. arietinum</i>	21,183/1,797	13,472/354	11,089/174	7,950/71	11.79, 38.06, 63.73, 111.97	416	AD07-CA02
<i>A. duranensis</i> vs <i>G. max</i>	34,488/1,927	27,248/669	23,113/359	16,423/151	17.90, 40.73, 64.38, 108.76	377	AD04-GM18
<i>A. duranensis</i> vs <i>V. angularis</i>	10,767/1,202	5,543/257	3,377/94	1,289/18	8.96, 21.57, 35.93, 71.61	107	AD05-VU11
<i>A. duranensis</i> vs <i>V. radiata</i>	19,021/1,690	11,874/356	9,312/160	5,922/50	11.26, 33.35, 58.2, 118.44	414	AD03-VR02
<i>A. duranensis</i> vs <i>C. cajan</i>	13,794/1,390	7,957/340	5,234/140	2,166/29	9.92, 23.40, 37.39, 74.69	130	CC02-AD07
<i>A. duranensis</i> vs <i>P. vulgaris</i>	23,292/1,780	15,867/412	13,084/203	9,151/75	13.09, 38.51, 64.45, 122.01	476	AD03-PV11
<i>A. ipaensis</i> vs <i>L. japonicus</i>	14,905/1,019	10,893/290	9,191/158	6,111/56	14.63, 37.56, 58.17, 109.13	398	AI03-LJ03
<i>A. ipaensis</i> vs <i>M. truncatula</i>	18,970/1,196	14,324/365	11,992/189	8,498/76	15.86, 39.24, 63.45, 111.82	373	AI09-MT01
<i>A. ipaensis</i> vs <i>C. arietinum</i>	20,169/1,626	13,371/360	10,987/182	7,402/65	12.40, 37.14, 60.37, 113.88	372	AI09-CA04
<i>A. ipaensis</i> vs <i>G. max</i>	34,801/1,985	27,278/671	23,276/368	16,490/151	17.53, 40.65, 63.25, 109.2	355	AI08-GM13
<i>A. ipaensis</i> vs <i>V. angularis</i>	10,440/1,146	5,542/265	3,328/94	1,171/15	9.11, 20.91, 35.40, 78.07	131	AI01-VU11
<i>A. ipaensis</i> vs <i>V. radiata</i>	18,349/1,543	11,921/353	9,310/156	6,098/52	11.89, 33.77, 59.68, 117.27	422	AI03-VR02
<i>A. ipaensis</i> vs <i>C. cajan</i>	13,331/1,320	7,917/339	5,275/140	2,091/28	10.10, 23.35, 37.68, 74.68	146	AI08-CC10
<i>A. ipaensis</i> vs <i>P. vulgaris</i>	23,021/1,723	15,976/423	13,187/210	9,206/76	13.36, 37.77, 62.80, 121.13	491	AI03-PV11
<i>L. japonicus</i> vs <i>M. truncatula</i>	19,336/1,154	14,494/280	12,811/145	10,287/63	16.76, 51.76, 88.35, 163.29	1,048	LJ05-MT01
<i>L. japonicus</i> vs <i>C. arietinum</i>	15,759/1,038	11,369/242	9,802/126	7,481/46	15.18, 46.98, 77.79, 162.63	874	LJ05-CA04
<i>L. japonicus</i> vs <i>G. max</i>	35,076/1,890	27,583/557	24,413/315	19,037/147	18.56, 49.52, 77.50, 129.50	626	LJ01-GM19
<i>L. japonicus</i> vs <i>V. angularis</i>	10,417/1,074	5,852/249	3,858/99	1,321/16	9.70, 23.50, 38.97, 82.56	131	LJ05-VU06
<i>L. japonicus</i> vs <i>V. radiata</i>	13,917/964	10,192/281	8,293/139	5,502/46	14.44, 36.27, 59.66, 119.61	349	LJ02-VR11
<i>L. japonicus</i> vs <i>C. cajan</i>	10,680/952	7,039/295	4,745/125	1,891/27	11.22, 23.86, 37.96, 70.04	125	LJ05-CC02
<i>L. japonicus</i> vs <i>P. vulgaris</i>	20,181/1,330	14,804/345	12,654/176	9,443/67	15.17, 42.91, 71.90, 140.94	597	LJ05-PV07
<i>M. truncatula</i> vs <i>C. arietinum</i>	20,631/1,023	16,337/246	14,682/121	12,367/46	20.17, 66.41, 121.34, 268.85	1,316	MT07-CA03
<i>M. truncatula</i> vs <i>G. max</i>	50,672/2,824	38,904/717	34,404/369	28,193/175	17.94, 54.26, 93.24, 161.10	950	MT07-GM19
<i>M. truncatula</i> vs <i>V. angularis</i>	12,541/1,238	7,160/284	4,978/120	2,147/26	10.13, 25.21, 41.48, 82.58	180	MT04-VU11
<i>M. truncatula</i> vs <i>V. radiata</i>	16,859/1,063	12,663/320	10,510/155	7,489/58	15.86, 39.57, 67.81, 129.12	579	MT05-VR11
<i>M. truncatula</i> vs <i>C. cajan</i>	13,083/1,115	8,808/367	5,935/152	2,300/31	11.73, 24.00, 39.05, 74.19	134	MT04-CC10
<i>M. truncatula</i> vs <i>P. vulgaris</i>	27,309/1,734	20,196/445	17,259/219	13,311/89	15.75, 45.38, 78.81, 149.56	673	MT01-PV07
<i>C. arietinum</i> vs <i>G. max</i>	35,268/1,840	28,387/635	24,025/308	18,754/135	19.17, 44.70, 78.00, 138.92	822	CA03-GM19



<i>C.arietinum</i> vs <i>V.angularis</i>	12,568/1,460	5,959/256	3,845/98	1,546/20	8.61, 23.28, 39.23, 77.30	150	CA07-VU11
<i>C.arietinum</i> vs <i>V.radiata</i>	22,835/2,252	12,701/346	10,151/152	7,300/58	10.14, 36.71, 66.78, 125.86	558	CA08-VR11
<i>C.arietinum</i> vs <i>C.cajan</i>	15,189/1,576	8,372/327	5,854/139	2,780/33	9.64, 25.60, 42.12, 84.24	198	CA02-CC02
<i>C.arietinum</i> vs <i>P. vulgaris</i>	24,912/1,958	16,528/401	13,939/200	10,042/69	12.72, 41.22, 69.69, 145.54	648	CA03-PV01
<i>G. max</i> vs <i>V.angularis</i>	23,214/2,047	15,306/635	10,350/265	3,900/46	11.34, 24.10, 39.06, 84.78	185	GM17-VU11
<i>G. max</i> vs <i>V.radiata</i>	34,129/1,909	26,966/678	22,407/337	15,888/130	17.88, 39.77, 66.49, 122.22	563	GM06-VR05
<i>G. max</i> vs <i>C. cajan</i>	25,823/2,058	18,148/727	12,729/317	5,305/67	12.55, 24.96, 40.15, 79.18	162	GM14-CC02
<i>G. max</i> vs <i>P. vulgaris</i>	52,702/2,787	41,818/860	36,555/457	28,453/194	18.91, 48.63, 79.99, 146.66	759	GM03-PV01
<i>V.angularis</i> vs <i>V.radiata</i>	13,660/1,510	7,046/280	5,016/122	2,123/25	9.05, 25.16, 41.11, 84.92	193	VU11-VR07
<i>V.angularis</i> vs <i>C. cajan</i>	9,104/1,216	3,450/192	1,645/55	241/4	7.49, 17.97, 29.91, 60.25	72	VU06-CC02
<i>V.angularis</i> vs <i>P. vulgaris</i>	15,154/1,502	8,782/340	6,212/144	2,773/32	10.09, 25.83, 43.14, 86.66	276	VU11-PV03
<i>V.radiata</i> vs <i>C. cajan</i>	15,656/1,687	8,243/324	5,717/131	2,631/30	9.28, 25.44, 43.64, 87.70	249	VR06-CC02
<i>V.radiata</i> vs <i>P. vulgaris</i>	24,221/1,741	16,781/366	14,296/178	11,175/76	13.91, 45.85, 80.31, 147.04	602	VR02-PV11
<i>C. cajan</i> vs <i>P. vulgaris</i>	17,425/1,625	10,716/415	7,688/184	3,596/46	10.72, 25.82, 41.78, 78.17	162	CC10-PV05

<sup>a</sup> BL: block\_length; <sup>b</sup> ACGP: average colinear gene pairs respectively per block; <sup>c</sup> LDB: number of colinear gene pairs reside in longest duplicated block

**Supplemental Table S2.** Number of homologous genes within a genome or between genomes.

Homologous Blocks within and among genome	BL <sup>a</sup> > 4	BL> 1,0	BL> 20	BL> 50	LDB <sup>b</sup>	LDB on chromosomes
<i>Vitis vinifera</i>	4,122	3,030	2,070	222	61	VV05-VV07
<i>Arachis duranensis</i>	5,937	3,405	2,143	818	96	AD03-AD08
<i>Arachis ipaensis</i>	6,155	3,239	1,969	771	84	AI03-AI08
<i>Lotus japonicus</i>	4,437	2,410	1,445	412	86	LJ04-LJ06
<i>Medicago truncatula</i>	7,079	4,050	3,027	1320	144	MT01-MT07
<i>Cicer arietinum</i>	5,676	3,302	2,182	832	145	CA07-CA08
<i>Glycine max</i>	29,352	27,398	25,782	22,624	824	GM10-GM20
<i>Vigna angularis</i>	3,256	1,257	482	0	39	VU07-VU11
<i>Vigna radiata</i>	5,591	2,912	1,908	805	107	VR07-VR11
<i>Cajanus cajan</i>	3,072	1,000	210	0	34	CC08-CC11
<i>Phaseolus vulgaris</i>	7,739	5,225	3,628	1,805	114	PV01-PV09
<i>V. vinifera</i> vs <i>A. duranensis</i>	8,762 vs 10,236	6,993 vs 7,985	5,430 vs 6,218	2,454 vs 2,665	145	VV18-AD08
<i>V. vinifera</i> vs <i>A. ipaensis</i>	8,713 vs 10,273	6,886 vs 7,890	5,379 vs 6,102	2,601 vs 2,892	145	VV18-AI07
<i>V. vinifera</i> vs <i>L. japonicus</i>	8,167 vs 9,089	6,390 vs 7,018	4,741 vs 5,163	2,147 vs 2,257	89	VV14-LJ01
<i>V. vinifera</i> vs <i>M. truncatula</i>	9,018 vs 11,479	7,549 vs 9,193	6,046 vs 7,017	3,461 vs 3,908	168	VV18-MT01
<i>V. vinifera</i> vs <i>C. arietinum</i>	9,181 vs 10,062	7,113 vs 7,905	5,632 vs 6,272	2,895 vs 3,181	149	VV18-CA04
<i>V. vinifera</i> vs <i>G. max</i>	9,236 vs 20,713	8,048 vs 17,208	6,648 vs 13,455	3,601 vs 6,756	161	VV18-GM04
<i>V. vinifera</i> vs <i>V. angularis</i>	6,184 vs 6,464	3,540 vs 3,713	1,723 vs 1,799	268 vs 254	63	VV18-VU05
<i>V. vinifera</i> vs <i>V. radiata</i>	8,482 vs 9,299	6,336 vs 7,109	4,692 vs 5,242	2,309 vs 2,617	151	VV18-VR05
<i>V. vinifera</i> vs <i>C. cajan</i>	7,570 vs 7,580	4,884 vs 5,015	3,021 vs 3,187	640 vs 654	70	VV06-CC10
<i>V. viniferavsP. vulgaris</i>	9,889 vs 12,475	8,050 vs 9,714	6,382 vs 7,505	3,317 vs 3,771	172	VV09-PV10
<i>A. duranensis</i> vs <i>A. ipaensis</i>	18,598 vs 18,993	16,724 vs 16,812	16,339 vs 16,382	15,741 vs 15,756	1,213	AD10-AI10
<i>A. duranensis</i> vs <i>L. japonicus</i>	10,924 vs 10,865	9,314 vs 9,227	8,135 vs 8,073	5,874 vs 5,810	396	AD03-LJ03
<i>A. duranensis</i> vs <i>M. truncatula</i>	12,540 vs 13,879	11,182 vs 11,854	10,133 vs 10,609	7,942 vs 8,103	437	AD07-MT05

<i>A. duranensis</i> vs <i>C. arietinum</i>	12,863 vs 11,910	10,603 vs 10,321	9,399 vs 9,186	7,275 vs 7,226	416	AD07-CA02
<i>A. duranensis</i> vs <i>G. max</i>	12,748 vs 25,489	11,557 vs 22,507	10,585 vs 20,119	8,665 vs 15,280	377	AD04-GM18
<i>A. duranensis</i> vs <i>V. angularis</i>	7,859 vs 7,243	4,905 vs 4,790	3,169 vs 3,146	1,269 vs 1,267	107	AD05-VU11
<i>A. duranensis</i> vs <i>V. radiata</i>	11,712 vs 10,667	9,396 vs 9,036	7,926 vs 7,750	5,428 vs 5,416	414	AD03-VR02
<i>A. duranensis</i> vs <i>C. cajan</i>	9,924 vs 8,475	6,993 vs 6,415	4,894 vs 4,664	2,074 vs 2,031	130	AD07-CC02
<i>A. duranensis</i> vs <i>P. vulgaris</i>	13,543 vs 14,717	11,732 vs 12,493	10,549 vs 11,011	8,002 vs 8,247	476	AD03-PV11
<i>A. ipaensis</i> vs <i>L. japonicus</i>	10,988 vs 10,948	9,287 vs 9,212	8,145 vs 8,109	5,912 vs 5,785	398	AI03-LJ03
<i>A. ipaensis</i> vs <i>M. truncatula</i>	12,722 vs 14,111	11,274 vs 11,974	10,176 vs 10,626	7,819 vs 7,986	373	AI09-MT01
<i>A. ipaensis</i> vs <i>C. arietinum</i>	12,708 vs 11,787	10,639 vs 10,292	9,408 vs 9,129	6,917 vs 6,840	372	AI09-CA04
<i>A. ipaensis</i> vs <i>G. max</i>	13,027 vs 25,828	11,737 vs 22,687	10,780 vs 20,280	8,908 vs 15,435	355	AI08-GM13
<i>A. ipaensis</i> vs <i>V. angularis</i>	7,819 vs 7,176	4,925 vs 4,801	3,115 vs 3,108	1,171 vs 1,142	131	AI01-VU11
<i>A. ipaensis</i> vs <i>V. radiata</i>	11,638 vs 10,653	9,458 vs 9,094	7,946 vs 7,818	5,598 vs 5,528	422	AI03-VR02
<i>A. ipaensis</i> vs <i>C. cajan</i>	9,801 vs 8,339	7,007 vs 6,381	4,891 vs 4,675	2,021 vs 1,999	146	AI08-CC10
<i>A. ipaensis</i> vs <i>P. vulgaris</i>	13,665 vs 14,766	11,887 vs 12,597	10,620 vs 11,028	8,056 vs 8,317	491	AI03-PV11
<i>L. japonicas</i> vs <i>M. truncatula</i>	15,138 vs 13,665	12,722 vs 12,025	11,696 vs 11,195	9,870 vs 9,630	1,048	LJ05-MT01
<i>L. japonicas</i> vs <i>C. arietinum</i>	11,609 vs 11,142	9,743 vs 9,630	8,778 vs 8,725	7,164 vs 7,123	874	LJ05-CA04
<i>L. japonicas</i> vs <i>G. max</i>	27,639 vs 14,713	24,079 vs 13,476	22,191 vs 12,741	18,131 vs 11,084	626	LJ01-GM19
<i>L. japonicus</i> vs <i>V. angularis</i>	8,031 vs 7,647	5,221 vs 5,212	3,625 vs 3,618	1,300 vs 1,304	131	LJ05-VU06
<i>L. japonicus</i> vs <i>V. radiata</i>	10,289 vs 9,761	8,593 vs 8,485	7,400 vs 7,361	5,178 vs 5,168	349	LJ02-VR11
<i>L. japonicus</i> vs <i>C. cajan</i>	8,662 vs 7,718	6,435 vs 6,060	4,553 vs 4,352	1,882 vs 1,886	125	LJ05-CC02
<i>L. japonicus</i> vs <i>P. vulgaris</i>	13,458 vs 14,566	11,775 vs 12,461	10,703 vs 11,151	8,573 vs 8,820	597	LJ05-PV07
<i>M. truncatula</i> vs <i>C. arietinum</i>	15,653 vs 13,449	13,795 vs 12,781	13,002 vs 12,340	11,715 vs 11,471	1,316	MT07-CA03
<i>M. truncatula</i> vs <i>G. max</i>	21,103 vs 34,822	18,802 vs 3,1088	17,765 vs 29,257	15,938 vs 25,430	950	MT07-GM19
<i>M. truncatula</i> vs <i>V. angularis</i>	9,947 vs 8,591	6,485 vs 6,173	4,680 vs 4,545	2,114 vs 2,053	180	MT04-VU11
<i>M. truncatula</i> vs <i>V. radiata</i>	12,658 vs 10,945	10,645 vs 9,890	9,198 vs 8,750	7,010 vs 6,780	579	MT05-VR11
<i>M. truncatula</i> vs <i>C. cajan</i>	10,667 vs 8,806	8,034 vs 7,176	5,694 vs 5,299	2,252 vs 2,218	134	MT04-CC10

<i>M. truncatula</i> vs <i>P. vulgaris</i>	18,295 vs 17,269	15,849 vs 15,610	14,545 vs 14,420	1,193 vs 11,875	673	MT01-PV07
<i>C. arietinum</i> vs <i>G. max</i>	12,450 vs 26,401	11,654 vs 23,436	10,798 vs 21,018	9,532 vs 17,510	822	CA03-GM19
<i>C. arietinum</i> vs <i>V. angularis</i>	8,217 vs 7,898	5,133 vs 5,122	3,533 vs 3,538	1,528 vs 1,479	150	CA07-VU11
<i>C. arietinum</i> vs <i>V. radiata</i>	11,931 vs 11,554	9,713 vs 9,659	8,515 vs 8,544	6,611 vs 6,596	558	CA08-VR11
<i>C. arietinum</i> vs <i>C. cajan</i>	9,997 vs 8,818	7,211 vs 6,843	5,461 vs 5,197	2,713 vs 2,736	198	CA02-CC02
<i>C. arietinum</i> vs <i>P. vulgaris</i>	13,163 vs 15,443	11,871 vs 13,044	10,979 vs 11,737	8,845 vs 9,161	648	CA03-PV01
<i>G. max</i> vs <i>V. angularis</i>	18,557 vs 10,249	13,679 vs 8,350	9,746 vs 6,321	3,870 vs 2,741	185	GM17-VU11
<i>G. max</i> vs <i>V. radiata</i>	25,359 vs 11,979	22,368 vs 11,293	19,642 vs 10,399	14,815 vs 8,133	563	GM06-VR05
<i>G. max</i> vs <i>C. cajan</i>	20,982 vs 9,674	16,541 vs 8,323	12,178 vs 6,609	5,213 vs 3,124	162	GM14-CC02
<i>G. max</i> vs <i>P. vulgaris</i>	35,374 vs 19,151	32,186 vs 18,090	30,166 vs 17,260	25,496 vs 14,694	759	GM03-PV01
<i>V. angularis</i> vs <i>V. radiata</i>	8,455 vs 8,655	5,980 vs 6,029	4,647 vs 4,664	2,067 vs 2,098	193	VU11-VR07
<i>V. angularis</i> vs <i>C. cajan</i>	6,385 vs 6,034	3,207 vs 3,080	1,602 vs 1,562	241 vs 241	72	VU06-CC02
<i>V. angularis</i> vs <i>P. vulgaris</i>	9,532 vs 10,932	7,219 vs 7,694	5,618 vs 5,827	2,673 vs 2,745	276	VU11-PV03
<i>V. radiata</i> vs <i>C. cajan</i>	9,743 vs 8,792	7,042 vs 6,598	5,277 vs 5,063	2,589 vs 2,550	249	VR06-CC02
<i>V. radiata</i> vs <i>P. vulgaris</i>	13,103 vs 15,501	12,244 vs 13,367	11,606 vs 12,255	9,812 vs 10,178	602	VR02-PV11
<i>C. cajan</i> vs <i>P. vulgaris</i>	9,996 vs 12,576	8,200 vs 9,311	6,622 vs 7,183	3,390 vs 3,481	162	CC10-PV05

<sup>a</sup>BL: block\_length; <sup>b</sup>LDB: number of colinear gene pairs reside in longest duplicated block

1 **Supplemental Table S3.** Number of paralogous, orthologous and out-paralogous gene pairs within a genome or between genomes.

Species	<i>V. vinifera</i>	<i>M. truncatula</i>	<i>C. arietium</i>	<i>A. duranensis</i>	<i>A. ipaensis</i>	<i>V. radiata</i>	<i>V. angularis</i>	<i>P. vulgaris</i>	<i>L. japonicus</i>	<i>C. cajan</i>	<i>G. max</i>
<i>V. vinifera</i>	2,423	8,055	9,533	7,514	7,349	5,842	4,240	8,782	6903	3,841	15,787
<i>M. truncatula</i>	6,665	4,703	10,909	8,324	9,133	7,600	5,580	12,665	10,946	10,909	24,465
<i>C. arietium</i>	8,367	4,522	4,416	14,792	14,648	11,965	8,543	18,283	15,144	7,653	34,027
<i>A. duranensis</i>	5,907	4,537	10,286	1,389	15,913	10,336	7,239	15,650	13,177	6,478	30,303
<i>A. ipaensis</i>	5,919	5,123	10,445	14,985	3,139	10,377	7,255	15,394	13,240	6,339	30,033
<i>V. radiata</i>	4,776	7,083	8,977	6,237	6,121	2,104	5,490	11,807	10,322	5,400	25,133
<i>V. angularis</i>	2,946	4,530	4,645	3,260	3,458	3,446	1,263	6,721	8,442	4,117	19,113
<i>P. vulgaris</i>	6,754	11,220	13,617	8,717	9,146	8,052	4,523	4,294	16,866	8,319	42,659
<i>L. japonicus</i>	5,480	5,076	9,671	6,081	6,044	5,653	3,496	8,734	2,731	6,522	33,362
<i>C. cajan</i>	2,920	4,522	4,606	2,846	3,261	3,188	1,569	3,697	2,790	685	15,896
<i>G. max</i>	14,729	15,251	25,238	20,310	20,190	17,916	9,042	28,770	20,893	8,178	34,575

2 Numbers on the main diagonal colored with dark sea green denote paralogous gene pairs within a genome, numbers above the diagonal colored in golden rod denote  
3 orthologous gene pairs between two genome, and numbers below the diagonal and colored in paleviolet red denote out-paralogous gene pairs between two genome.

4  
5  
6

7 **Supplemental Table S4.** Number of paralogous, orthologous and out-paralogous genes within a genome or between genomes.

Species	<i>V. vinifera</i>	<i>M. truncatula</i>	<i>C. arietium</i>	<i>A. duranensis</i>	<i>A. ipaensis</i>	<i>V. radiata</i>	<i>V. angularis</i>	<i>P. vulgaris</i>	<i>L. japonicus</i>	<i>C. cajan</i>	<i>G. max</i>
<i>V. vinifera</i>	3,851	6,698/7,928	7,344/8,295	6,242/7,277	6,119/7,112	4,904/5,335	3,540/3,944	6,671/8,186	5,955/6,713	3,422/3,518	7,496/15,194
<i>M. truncatula</i>	5,093/4,530	6,860	10,790/10,707	8,247/8,262	9,017/9,010	7,503/7,463	5,437/5,374	12,459/12,414	10,735/10,739	10,790/10,707	16,247/23,906
<i>C. arietium</i>	5,784/5,508	3,645/4,257	5,705	10,685/10,888	10,673/1,0888	9,014/8,500	6,477/6,529	12,063/13,620	10,525/11,701	6,419/5,567	12,188/26,796
<i>A. duranensis</i>	4,335/4,221	3,868/4,314	7,381/7,257	2,658	15,848/1,5863	8,373/7,829	5,683/5,763	11,178/12,287	9,590/10,610	5,669/5,004	11,597/24,486
<i>A. ipaensis</i>	4,344/4,192	4,239/4,759	7,524/7,328	11,491/11,294	4,625	8,439/7,975	5,745/5,799	10,920/12,028	9,646/10,643	5,555/4,871	11,628/24,356
<i>V. radiata</i>	3,349/3,537	4,687/5,663	6,121/6,521	4,464/4,785	4,309/4,740	3,157	5,313/5,319	11,508/11,591	7,366/8,731	4,741/4,459	9,690/21,052
<i>V. angularis</i>	2,249/2,273	3,356/3,759	3,494/3,612	2,551/2,616	2,726/2,776	2,742/2,607	2,077	6,504/6,508	6,399/6,743	3,667/3,263	8,573/15,473
<i>P. vulgaris</i>	5,014/4,543	7,978/8,477	9,813/8,921	6,686/6,241	6,829/6,339	6,142/5,365	3,681/3,416	6,157	12,461/12,471	7,331/5,953	17,288/31,949
<i>L. japonicus</i>	4,081/4,019	4,263/4,502	7,351/7,123	4,771/4,620	4,700/4,488	4,464/4,139	2,896/2,790	6,564/6,738	4,115	5,860/4,796	13,955/25,698
<i>C. cajan</i>	2,055/2,462	3,645/4,257	3,271/3,955	2,116/2,497	2,367/2,826	2,418/2,710	1,272/1,402	2,689/3,296	2,119/2,519	1,181	6,239/14,308
<i>G. max</i>	10,987/5,850	12,360/8,789	19,188/9,746	8,490/15,748	15,611/8,441	14,028/7,240	7,516/4,640	20,826/12,458	16,770/10,049	7,282/3,578	3,4441

8 See the legends of Table S3.

9 In non-diagonal cells, gene numbers in two corresponding species were shown, from vertical and horizontal lists respectively.

10

11

12

13

14  
15

**Supplemental Table S5.** Number of paralogous, orthologous and out-paralogous blocks within a genome or between genomes.

Species	<i>V. vinifera</i>	<i>M. truncatula</i>	<i>C. arietium</i>	<i>A. duranensis</i>	<i>A. ipaensis</i>	<i>V. radiata</i>	<i>V. angularis</i>	<i>P. vulgaris</i>	<i>L. japonicus</i>	<i>C. cajan</i>	<i>G. max</i>
<i>V. vinifera</i>	86	304	565	319	317	218	234	274	322	210	632
<i>M. truncatula</i>	316	420	152	174	194	199	275	224	191	212	330
<i>C. arietium</i>	422	189	438	459	434	383	363	439	390	328	733
<i>A. duranensis</i>	334	176	448	185	69	314	317	344	344	300	662
<i>A. ipaensis</i>	313	198	416	394	183	318	323	355	344	288	642
<i>V. radiata</i>	240	298	392	327	326	134	262	285	287	206	554
<i>V. angularis</i>	208	264	286	218	228	214	72	334	331	223	660
<i>P. vulgaris</i>	308	434	471	401	389	355	271	192	342	295	671
<i>L. japonicus</i>	310	200	366	292	305	275	206	347	135	250	645
<i>C. cajan</i>	192	143	271	210	220	191	125	235	173	47	585
<i>G. max</i>	637	381	726	682	700	578	474	805	629	439	741

16 See the legends of Table S3.  
17 In non-diagonal cells, gene numbers in two corresponding species were shown, from vertical and horizontal lists respectively.  
18  
19  
20  
21

22

23 **Supplemental Table S6.** Legume gene loss rates and gene translocation with grape as reference genome.

Species	Grape Chr#	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
	Genes	1,326	1,236	999	1,637	1,747	1,778	1,408	1,866	1,220	631	1,106	1,480	1,328	1,728	560	646	1,167	1,885	1,134
	Loss in Para#																			
<i>Arachis duranensis</i>	para1	0.72	0.86	0.75	0.80	0.77	0.79	0.70	0.77	0.85	0.73	0.78	0.86	0.77	0.85	0.80	0.63	0.78	0.81	0.80
	para2	0.75	0.81	0.79	0.77	0.70	0.76	0.77	0.80	0.84	0.78	0.75	0.82	0.76	0.81	0.72	0.90	0.74	0.74	0.78
	all para.	0.68	0.79	0.74	0.69	0.72	0.71	0.64	0.69	0.80	0.66	0.72	0.80	0.73	0.75	0.73	0.82	0.73	0.65	0.71
<i>Arachis ipaensis</i>	para1	0.72	0.85	0.68	0.75	0.80	0.77	0.73	0.70	0.79	0.77	0.75	0.83	0.77	0.86	0.81	0.73	1.00	0.75	0.79
	para2	0.76	0.78	0.80	0.76	0.70	0.74	0.73	0.79	0.84	0.78	0.72	0.80	0.76	0.80	0.78	0.88	0.76	0.72	0.84
	all para.	0.71	0.78	0.71	0.67	0.73	0.68	0.63	0.66	0.76	0.70	0.70	0.79	0.73	0.73	0.78	0.83	0.76	0.66	0.74
<i>Cicer arietium</i>	para1	0.64	0.76	0.69	0.77	0.74	0.74	0.59	0.71	0.79	0.79	0.72	0.74	0.65	0.79	0.75	0.77	0.63	0.69	0.78
	para2	0.69	0.67	0.71	0.72	0.78	0.68	0.68	0.79	0.90	0.72	0.66	0.71	0.53	0.67	0.79	0.80	0.78	0.67	0.76
	all para.	0.61	0.69	0.67	0.64	0.68	0.66	0.52	0.67	0.78	0.69	0.69	0.68	0.56	0.64	0.70	0.73	0.62	0.57	0.72
<i>Cajanus cajan</i>	para1	0.92	0.88	0.87	0.89	0.85	0.88	0.91	0.94	0.89	0.99	0.86	0.95	0.93	0.94	0.94	0.99	0.95	0.95	0.96
	para2	0.94	0.95	0.91	0.97	0.92	0.92	0.98	0.97	0.96	0.97	0.90	0.93	0.98	0.98	0.96	0.96	0.87	0.88	0.95
	all para.	0.85	0.84	0.80	0.87	0.79	0.83	0.89	0.91	0.87	0.96	0.77	0.89	0.91	0.92	0.90	0.95	0.82	0.84	0.91
<i>Glycine max</i>	para11	0.78	0.83	0.73	0.83	0.74	0.76	0.78	0.67	0.82	0.78	0.76	0.83	0.81	0.83	0.90	0.70	0.72	0.74	0.80
	para12	0.79	0.77	0.74	0.68	0.81	0.72	0.76	0.74	0.85	0.80	0.77	0.86	0.77	0.77	0.85	0.85	0.76	0.82	0.78
	para21	0.77	0.79	0.81	0.72	0.73	0.73	0.78	0.79	0.88	0.85	0.77	0.93	0.82	0.82	0.82	1.00	0.78	0.75	0.70



<i>Lotus japonicus</i>	para22	0.78	0.79	0.81	0.79	0.71	0.82	0.75	0.82	0.86	0.81	0.83	0.76	0.81	0.78	0.90	0.89	0.78	0.70	0.72
	all para.	0.69	0.75	0.68	0.65	0.73	0.73	0.66	0.63	0.75	0.72	0.67	0.78	0.72	0.66	0.74	0.81	0.62	0.60	0.68
	para1	0.71	0.84	0.69	0.81	0.76	0.75	0.76	0.73	1.00	0.76	0.66	0.92	0.71	0.81	0.86	0.77	0.85	0.89	0.82
<i>Medicago</i>	para2	0.76	0.71	0.83	0.76	0.79	0.84	0.75	0.80	0.88	0.75	0.80	0.65	0.84	0.79	0.78	0.89	0.72	0.76	0.78
	all para.	0.71	0.75	0.69	0.70	0.73	0.74	0.65	0.68	0.88	0.65	0.62	0.80	0.73	0.70	0.77	0.80	0.66	0.76	0.75
	para1	0.75	0.82	0.70	0.73	0.69	0.73	0.71	0.72	0.80	0.75	0.72	0.84	0.78	0.81	0.73	0.75	0.73	0.76	0.80
<i>Phaseolus vulgaris</i>	para2	0.77	0.79	0.77	0.76	0.74	0.77	0.72	0.77	0.88	0.78	0.72	0.67	0.79	0.81	0.75	0.87	0.75	0.73	0.76
	all para.	0.70	0.72	0.67	0.63	0.63	0.63	0.62	0.64	0.76	0.70	0.63	0.74	0.74	0.70	0.71	0.73	0.65	0.62	0.74
	para1	0.94	0.83	0.83	0.81	0.82	0.83	0.74	0.73	0.83	0.82	0.72	0.88	0.91	0.80	0.89	0.89	0.86	0.83	0.87
<i>Vigna radiata</i>	para2	0.89	0.84	0.92	0.80	0.90	0.87	0.77	0.83	0.91	0.87	0.92	0.87	0.93	0.87	0.81	0.96	0.76	0.78	0.91
	all para.	0.83	0.73	0.80	0.68	0.76	0.77	0.60	0.64	0.78	0.75	0.67	0.78	0.88	0.72	0.73	0.86	0.67	0.68	0.81
	para1	0.87	0.88	0.96	0.86	0.85	0.94	0.79	0.84	0.87	0.84	0.92	0.94	0.88	0.85	0.95	0.97	0.94	0.86	0.90
<i>Vigna angularis</i>	para2	0.96	0.91	0.85	0.82	1.00	0.88	0.80	0.93	0.91	0.98	0.87	0.93	0.91	0.87	0.92	0.97	0.93	0.80	0.91
	all para.	0.85	0.81	0.83	0.74	0.85	0.84	0.67	0.80	0.81	0.83	0.82	0.86	0.83	0.76	0.88	0.95	0.87	0.72	0.84
	para1	0.93	0.93	0.98	0.91	0.89	0.97	0.82	0.88	0.99	0.89	0.84	0.95	0.91	0.95	0.97	0.97	0.93	0.90	0.97
<i>Vigna angularis</i>	para2	1.00	0.89	0.92	0.87	0.94	0.96	0.90	0.90	1.00	0.97	0.88	0.95	1.00	0.88	0.93	0.98	0.85	0.91	0.95
	all para.	0.92	0.84	0.91	0.80	0.85	0.94	0.74	0.81	0.99	0.86	0.75	0.90	0.91	0.84	0.90	0.96	0.81	0.82	0.93

24

25

26

27 **Supplemental Table S7.** Legume gene loss and gene translocation rates with medicago as reference genome.

	Medicago Chr#	1	2	3	4	5	6	7	8
Species	Genes	8,386	7,245	8,481	9,132	7,072	4,399	7,586	7,269
	Loss in Para#								
<i>Arachis duranensis</i>	all para.	0.82	0.79	0.80	0.79	0.81	0.87	0.81	0.83
<i>Arachis ipaensis</i>	all para.	0.80	0.79	0.79	0.79	0.82	0.87	0.80	0.80
<i>Cicer arietium</i>	all para.	0.75	0.77	0.78	0.73	0.76	0.80	0.74	0.77
<i>Cajanus cajan</i>	all para.	0.79	0.86	0.82	0.78	0.81	0.88	0.74	0.76
	para1	0.74	0.75	0.72	0.70	0.77	0.83	0.75	0.75
<i>Glycine max</i>	para2	0.73	0.76	0.72	0.71	0.75	0.82	0.73	0.73
	all para.	0.69	0.70	0.67	0.66	0.71	0.80	0.70	0.71
<i>Lotus japonicus</i>	all para.	0.73	0.76	0.77	0.73	0.77	0.80	0.77	0.71
<i>Phaseolus vulgaris</i>	all para.	0.70	0.75	0.70	0.72	0.76	0.82	0.75	0.75
<i>Vigna radiata</i>	all para.	0.76	0.81	0.82	0.87	0.81	0.86	0.81	0.82
<i>Vigna angularis</i>	all para.	0.79	0.85	0.83	0.80	0.86	0.78	0.84	0.84

28  
29  
30  
31  
32

33  
34  
35 **Supplemental Table S8.** Legume gene loss and gene translocation rates with common bean as reference genome.

Common bean Chr#		1	2	3	4	5	6	7	8	9	10	11
Species	Genes	3,164	3,898	3,467	2,030	2,180	2,561	3,266	3,404	3,094	1,930	2,510
Loss in Para#												
<i>Arachis duranensis</i>	all para.	0.59	0.56	0.63	0.72	0.67	0.58	0.60	0.66	0.57	0.67	0.65
<i>Arachis ipaensis</i>	all para.	0.59	0.59	0.62	0.74	0.59	0.59	0.61	0.66	0.58	0.59	0.66
<i>Cicer arietium</i>	all para.	0.55	0.56	0.61	0.69	0.61	0.61	0.52	0.67	0.55	0.72	0.69
<i>Cajanus cajan</i>	all para.	0.61	0.61	0.57	0.80	0.72	0.61	0.59	0.72	0.75	0.65	0.67
	para1	0.46	0.50	0.46	0.58	0.56	0.48	0.46	0.55	0.48	0.54	0.63
<i>Glycine max</i>	para2	0.47	0.49	0.47	0.61	0.57	0.45	0.44	0.53	0.48	0.51	0.55
	all para.	0.38	0.42	0.39	0.53	0.49	0.38	0.39	0.48	0.40	0.48	0.54
<i>Lotus japonicus</i>	all para.	0.57	0.50	0.57	0.67	0.60	0.56	0.51	0.65	0.57	0.63	0.62
<i>Medicago</i>	all para.	0.52	0.52	0.61	0.70	0.73	0.55	0.70	0.60	0.61	0.59	0.67
<i>Vigna radiata</i>	all para.	0.49	0.51	0.47	0.64	0.67	0.60	0.51	0.59	0.50	0.66	0.61
<i>Vigna angularis</i>	all para.	0.67	0.68	0.65	0.68	0.83	0.67	0.63	0.68	0.67	0.84	0.78

36  
37  
38

**Supplemental Table S9.** The observed distribution of gene loss and translocation numbers fitted by using different density curves of geometry distribution.

Reference genome	Genomes	Parameter of geometry distribution	Fitness (R-square)	P-value (F-test)
<i>Vitis Vinifera</i>	<i>Arachis duranensis</i>	0.3091	0.9911	0.9048
	<i>Arachis ipaensis</i>	0.3085	0.9926	0.9052
	<i>Cicer arietium</i>	0.3274	0.9959	0.924
	<i>Cajanus cajan</i>	0.3456	0.9933	0.9099
	<i>Glycine max</i>	0.2997	0.9936	0.9136
	<i>Lotus japonicus</i>	0.3144	0.9938	0.9037
	<i>Medicago truncatula</i>	0.3147	0.9914	0.9084
	<i>Phaseolus vulgaris</i>	0.3302	0.9951	0.9174
	<i>Vigna radiata</i>	0.3111	0.9934	0.9093
	<i>Vigna angularis</i>	0.2197	0.9911	0.92
<i>Medicago truncatula</i>	<i>Arachis ipaensis</i>	0.31	0.9837	0.8931
	<i>Arachis duranensis</i>	0.3039	0.9876	0.9024
	<i>Cicer arietium</i>	0.3711	0.994	0.9174
	<i>Cajanus cajan</i>	0.3853	0.9958	0.9186
	<i>Glycine max</i>	0.351	0.9873	0.8973
	<i>Lotus japonicus</i>	0.359	0.9869	0.8895
	<i>Phaseolus vulgaris</i>	0.3755	0.9906	0.9077
	<i>Vigna radiata</i>	0.319	0.9914	0.9012
	<i>Vigna angularis</i>	0.1951	0.9673	0.9665
<i>Phaseolus vulgaris</i>	<i>Arachis ipaensis</i>	0.5086	0.9948	0.94
	<i>Arachis duranensis</i>	0.4966	0.997	0.9474
	<i>Cicer arietium</i>	0.5517	0.9948	0.9542

	<i>Cajanus cajan</i>	0.6797	0.9972	0.9574
	<i>Glycine max</i>	0.546	0.9978	0.9549
	<i>Lotus japonicus</i>	0.4939	0.9969	0.934
	<i>Medicago truncatula</i>	0.5147	0.9964	0.943
	<i>Vigna radiata</i>	0.61752	0.997	0.9602
	<i>Vigna angularis</i>	0.3518	0.9966	0.9617

41  
42  
43  
44

**Supplemental Table S10.** Gene retention in soybean duplicated chromosomes.

Barrel medic chr #	1	2	3	4	5	6	7	8	Total
Retained genes in soybean homoeologous region 1	692	494	553	1,132	470	295	500	599	4,735
Retained genes in soybean homoeologous region 2	619	457	447	728	397	134	437	511	3,730
No expected homologs in soybean homoeologous region 1 and 2	5,767	5,112	6,088	6,221	5,017	3,720	5,488	5,318	42,731
Retained genes in both homoeologous regions	1,309	1,183	1,394	1,052	1,189	251	1,162	842	8,382
Chisq-test p-value	4.38E-02	2.30E-01	8.02E-04	7.43E-21	1.32E-02	7.66E-15	3.96E-02	8.26E-03	8.92E-28

**Supplemental Table S11.** Kernel function analysis of Ks distribution related to duplication events within each genome and between selected legumes (before evolutionary rate correction).

Genomes	Weight coefficient related to LCT or speciation	Peak of Ks distribution ( $\mu_1$ ) related to LCT or speciation	Deviation ( $\sigma_1$ )	Weight coefficient related to ECH	Peak of Ks distribution ( $\mu_2$ ) related to ECH	Deviation ( $\sigma_2$ )	Weight coefficient related to SST	Peak of Ks distribution ( $\mu_3$ ) related to SST	Deviation ( $\sigma_3$ )
<i>G. max</i>	0.352	0.627	0.105	0.498	1.31	0.197	0.146	0.164	0.069
<i>A. ipaensis</i>	0.459	0.771	0.080	0.545	1.235	0.243	---	---	---
<i>V. vinifera</i>	---	---	---	0.916	1.108	0.136	---	---	---
<i>M. truncatula</i>	0.252	0.745	0.071	0.748	1.321	0.173	---	---	---
<i>C. arietium</i>	0.396	0.761	0.091	0.606	1.251	0.150	---	---	---
<i>L. japonicus</i>	0.632	0.588	0.124	0.381	1.167	0.211	---	---	---
<i>P. vulgaris</i>	0.443	0.746	0.082	0.394	1.429	0.099	---	---	---
<i>V. radiata</i>	0.389	0.761	0.083	0.402	1.357	0.128	---	---	---
<i>V. angularis</i>	0.552	0.757	0.146	0.225	1.40	0.128	---	---	---
<i>C. cajan</i>	0.390	0.653	0.120	0.619	1.287	0.266	---	---	---
<i>V. radiata vs P. vulgaris</i>	0.709	0.207	0.065	---	---	---	---	---	---
<i>V. radiata vs G. max</i>	0.806	0.346	0.070	---	---	---	---	---	---
<i>V. angularis vs V. radiata</i>	0.554	0.121	0.050	---	---	---	---	---	---
<i>P. vulgaris vs G. max</i>	0.757	0.316	0.069	---	---	---	---	---	---
<i>P. vulgaris vs C. cajan</i>	0.544	0.366	0.072	---	---	---	---	---	---
<i>M. truncatula vs G. max</i>	0.899	0.619	0.102	---	---	---	---	---	---
<i>M. truncatula vs A. ipaensis</i>	0.913	0.726	0.095	---	---	---	---	---	---
<i>A. duranensis vs A. ipaensis</i>	0.307	0.075	0.025	---	---	---	---	---	---
<i>L. japonicus vs M. truncatula</i>	0.818	0.535	0.092	---	---	---	---	---	---

<i>C. arietium</i> vs <i>M. truncatula</i>	0.724	0.365	0.072	---	---	---	---	---	---
-----------------------------------------------	-------	-------	-------	-----	-----	-----	-----	-----	-----

---

**Supplemental Table S12.** Kernel function analysis of Ks distribution related to duplication events within each genome and between selected legumes (after evolutionary rate correction).

Genomes	Weight coefficient related to LCT or speciation	Peak of Ks distribution ( $\mu$ 1) related to LCT or speciation	Deviation ( $\sigma$ 1)	Weight coefficient related to ECH	Peak of Ks distribution ( $\mu$ 2) related to ECH	Deviation ( $\sigma$ 2)	Weight coefficient related to SST	Peak of Ks distribution ( $\mu$ 3) related to SST	Deviation ( $\sigma$ 3)
<i>G. max</i>	0.352	0.627	0.105	0.498	1.31	0.197	0.146	0.164	0.069
<i>A. ipaensis</i>	0.459	0.627	0.072	0.545	1.01	0.199	---	---	---
<i>V. vinifera</i>	---	---	---	0.916	1.108	0.136	---	---	---
<i>M. truncatula</i>	0.249	0.627	0.065	0.782	1.118	0.155	---	---	---
<i>C. arietium</i>	0.398	0.627	0.080	0.604	1.031	0.127	---	---	---
<i>L. japonicus</i>	0.665	0.627	0.131	0.383	1.243	0.225	---	---	---
<i>P. vulgaris</i>	0.499	0.627	0.086	0.461	1.186	0.105	---	---	---
<i>V. radiata</i>	0.436	0.627	0.084	0.473	1.105	0.129	---	---	---
<i>V. angularis</i>	0.499	0.627	0.117	0.257	1.175	0.128	---	---	---
<i>C. cajan</i>	0.423	0.627	0.120	0.662	1.237	0.255	---	---	---
<i>V. radiata vs P. vulgaris</i>	0.719	0.174	0.062	---	---	---	---	---	---
<i>V. radiata vs G. max</i>	0.828	0.288	0.067	---	---	---	---	---	---
<i>V. angularis vs V. radiata</i>	0.470	0.115	0.042	---	---	---	---	---	---
<i>P. vulgaris vs G. max</i>	0.787	0.275	0.064	---	---	---	---	---	---
<i>P. vulgaris vs C. cajan</i>	0.890	0.332	0.074	---	---	---	---	---	---
<i>M. truncatula vs G. max</i>	0.860	0.522	0.092	---	---	---	---	---	---
<i>M. truncatula vs A. ipaensis</i>	0.919	0.601	0.084	---	---	---	---	---	---
<i>A. duranensis vs A. ipaensis</i>	0.280	0.064	0.022	---	---	---	---	---	---
<i>L. japonicus vs M. truncatula</i>	0.777	0.511	0.090	---	---	---	---	---	---



<i>C. arietium</i> vs <i>M. truncatula</i>	0.740	0.307	0.069	---	---	---	---	---	---
-----------------------------------------------	-------	-------	-------	-----	-----	-----	-----	-----	-----

---

**Supplemental Table 15.** Nodulation genes related to duplication events in each legume genome.

Genomes	ECH-related	LCT-related	SST-related	All genes in colinear regions (#%)	All nodulation genes in each genome
<i>Arachis duranensis</i>	237 (20%)	474 (40%)		528 (44%)	1,196
<i>Arachis ipaensis</i>	217 (17%)	473 (36%)		513 (39%)	1,312
<i>Cicer arietinum</i>	381 (42%)	671 (73%)		699 (76%)	916
<i>Cajanus cajan</i>	66 (9.7%)	73 (11%)		118 (17%)	681
<i>Glycine max</i>	411 (24%)	1,003 (59%)	1,209 (71%)	1,324 (78%)	1,702
<i>Lotus japonicus</i>	86 (7%)	245 (19%)		280 (22%)	1,262
<i>Medicago truncatula</i>	284 (17%)	852 (51%)		908 (54%)	1,675
<i>Phaseolus vulgaris</i>	315 (24%)	671 (50%)		711 (53%)	1,329
<i>Vigna radiata</i>	193 (24%)	279 (35%)		340 (42%)	804
<i>Vigna angularis</i>	64 (11%)	154 (26%)		187 (32%)	583
<i>Vitis vinifera</i>	325 (30%)			325 (30%)	1,077

**Supplemental Table S16.** Nodulation 7 subfamily genes related to duplication events in soybean genome.

Subfamily	1	2	3	4	5	6	7
<b>Total</b>	41	324	958	48	51	268	12
<b>All genes in colinear region (#%)</b>	39(95%)	288(89%)	849(89%)	44(92%)	49(96%)	241(90%)	10(83%)
<b>ECH-related</b>	29(71%)	198(61%)	627(65%)	36(75%)	28(55%)	188(70%)	10(83%)
<b>LCT-related</b>	38(93%)	230(71%)	698(73%)	34(71%)	39(76%)	211(79%)	7(58%)
<b>SST-related</b>	35(85%)	247(76%)	737(77%)	40(83%)	44(86%)	218(81%)	9(75%)

**Supplemental Table S17.** Oil 9 subfamily genes related to duplication events in soybean genome.

Subfamily	DSL&SFA <sup>a</sup>	SML in ES <sup>b</sup>	FAE&W&CM <sup>c</sup>	LS <sup>d</sup>	MAL in M <sup>e</sup>	ML <sup>f</sup>	MAL in M7 <sup>g</sup>	S&S8 <sup>h</sup>	SFA in P9 <sup>i</sup>
<b>Total</b>	104	132	330	355	97	313	57	55	85
<b>All genes in colinear regions(#%)</b>	93(89%)	126(95%)	289(88%)	338(95%)	94(97%)	288(92%)	54(95%)	51(93%)	80(94%)
<b>ECH-related</b>	26(25%)	37(28%)	86(26%)	141(40%)	30(31%)	91(29%)	15(26%)	29(53%)	25(29%)
<b>LCT-related</b>	66(63%)	88(67%)	196(59%)	258(73%)	70(72%)	201(64%)	44(77%)	37(67%)	56(66%)
<b>SST-related</b>	68(65%)	102(77%)	245(74%)	281(79%)	72(74%)	245(78%)	37(65%)	42(76%)	60(71%)

<sup>a</sup> **DSL&SFA:** Degradation of storage lipids and straight fatty acids.

<sup>b</sup> **SML in ES**: Synthesis of membrane lipids in endomembrane system.

<sup>c</sup> **FAE&W&CM**: Fatty acid elongation and wax and cutin metabolism.

<sup>d</sup> **LS**: Lipid signalling.

<sup>e</sup> **MAL in M**: Metabolism of acyl lipids in mitochondria.

<sup>f</sup> **ML**: Miscellaneous.

<sup>g</sup> **MAL in M7**: Metabolism of acyl lipids in mitochondria<sup>7</sup>.

<sup>h</sup> **S&S8**: Synthesis and storage of oil<sup>8</sup>.

<sup>i</sup> **SFA in P9**: Synthesis of fatty acids in plastids<sup>9</sup>.

**Supplemental Table S18.** Oil genes related to duplication events in each legume genome.

Genomes	ECH-related	LCT-related	SST-related	All genes in colinear regions (#%)	All oil genes in each genome
<i>A. duranensis</i>	189 (22%)	366 (42%)		411 (48%)	865
<i>A. ipaensis</i>	165 (19%)	358 (41%)		398 (46%)	874
<i>C. arietinum</i>	346 (43%)	574 (72%)		610 (76%)	801
<i>C. cajan</i>	58 (10%)	74 (13%)		109 (19%)	572
<i>G max</i>	392 (26%)	879 (58%)	996 (65%)	1,121 (73%)	1,528
<i>L. japonicus</i>	110 (13%)	199 (23%)		244 (28%)	860
<i>M. truncatula</i>	178 (18%)	581 (58%)		603 (60%)	1,002
<i>P. vulgaris</i>	243 (29%)	436 (51%)		480 (57%)	847
<i>V. radiata</i>	151 (30%)	199 (39%)		235 (46%)	510
<i>V. angularis</i>	63 (14%)	124 (28%)		156 (35%)	447
<i>V. vinifera</i>	228 (32%)			228 (32%)	716

**Supplemental Table S19.** NBS-CC genes related to duplication events in each legume genome.

Genome	ECH-related	LCT-related	SST-related	All genes in colinear region(%)	All CC genes in each genome
<i>Arachis duranensis</i>	112 (15%)	234 (30%)		269 (35%)	771
<i>Arachis ipaensis</i>	85 (11%)	241 (30%)		255 (32%)	798
<i>Cicer arietinum</i>	186 (36%)	329 (64%)		343 (66%)	516
<i>Cajanus cajan</i>	19 (3%)	34 (5%)		50 (7%)	690
<i>Glycine max</i>	146 (20%)	431 (59%)	528 (72%)	579 (79%)	732
<i>Lotus japonicus</i>	39 (6%)	109 (17%)		124 (20%)	630
<i>Medicago truncatula</i>	87 (8%)	389 (37%)		409 (38%)	1,064
<i>Phaseolus vulgaris</i>	104 (15%)	253 (36%)		272 (38%)	710
<i>Vigna radiata</i>	52 (13%)	106 (27%)		122 (31%)	392
<i>Vigna angularis</i>	16 (4%)	55 (13%)		63 (15%)	425
<i>Vitis angularis</i>	110 (13%)			110 (13%)	827

**Supplemental Table S20.** NBS-TIR genes related to duplication events in each legume genome.

Genomes	ECH-related	LCT-related	SST-related	All genes in colinear region(%)	All TIR genes in each genome
<i>Arachis duranensis</i>	41 (18%)	34 (15%)		54 (23%)	234
<i>Arachis ipaensis</i>	24 (9%)	47 (17%)		48 (18%)	273
<i>Cicer arietinum</i>	19 (30%)	26 (40%)		30 (47%)	64

<i>Cajanus cajan</i>	4 (3%)	2 (1%)		5 (3%)	152
<i>Glycine max</i>	27 (14%)	92 (49%)	114 (60%)	133 (70%)	189
<i>Lotus japonicus</i>	6 (2%)	22 (9%)		33 (13%)	252
<i>Medicago truncatula</i>	31 (7%)	163 (36%)		167 (37%)	450
<i>Phaseolus vulgaris</i>	22 (13%)	38 (22%)		48 (28%)	173
<i>Vigna radiata</i>	8 (17%)	11 (24%)		12 (26%)	46
<i>Vigna angularis</i>	4 (7%)	4 (7%)		7 (11%)	61
<i>Vitis vinifera</i>	5 (2%)			5 (2%)	206

**Supplemental Table S21.** NBS-TNL genes related to duplication events in each legume genome.

Genome	ECH-related	LCT-related	SST-related	All genes in colinear region(%)	All TNL genes in each genome
<i>Arachis duranensis</i>	47 (13%)	76 (21%)		96 (27%)	355
<i>Arachis ipaensis</i>	32 (8%)	93 (22%)		94 (22%)	424
<i>Cicer arietinum</i>	27 (26%)	44 (43%)		49 (48%)	103
<i>Cajanus cajan</i>	6 (3%)	4 (2%)		9 (4%)	239
<i>Glycine max</i>	33 (11%)	155 (53%)	192 (65%)	218 (74%)	295
<i>Lotus japonicus</i>	7 (2%)	28 (9%)		29 (10%)	296
<i>Medicago truncatula</i>	41 (6%)	218 (32%)		225 (33%)	684
<i>Phaseolus vulgaris</i>	39 (12%)	70 (21%)		86 (26%)	331

<i>Vigna radiata</i>	9 (14%)	16 (25%)	17 (26%)	65
<i>Vigna angularis</i>	4 (4%)	6 (7%)	9 (10%)	89
<i>Vitis vinifera</i>	28 (7%)		28 (7%)	429

**Supplemental Table S22.** NBS-TNx genes related to duplication events in each legume genome.

Genomes	ECH-related	LCT-related	SST-related	All genes in colinear region(%)	All TNx genes in each genome
<i>Arachis duranensis</i>	36 (12%)	62 (21%)		80 (27%)	291
<i>Arachis ipaensis</i>	26 (8%)	75 (22%)		76 (22%)	342
<i>Cicer arietinum</i>	16 (24%)	30 (44%)		33 (49%)	68
<i>Cajanus cajan</i>	5 (2%)	3 (1%)		8 (4%)	211
<i>Glycine max</i>	25 (10%)	128 (51%)	164 (66%)	182 (73%)	249
<i>Lotus japonicus</i>	7 (3%)	23 (11%)		24 (11%)	214
<i>Medicago truncatula</i>	32 (5%)	188 (31%)		195 (32%)	603
<i>Phaseolus vulgaris</i>	31 (10%)	62 (21%)		75 (25%)	300
<i>Vigna radiata</i>	5 (11%)	9 (20%)		9 (20%)	46
<i>Vigna angularis</i>	2 (3%)	4 (6%)		6 (9%)	67
<i>Vitis vinifera</i>	24 (6%)			24 (6%)	378

**Supplemental Table S23.** NBS-xNL genes related to duplication events in each genome.

Genome	ECH-related	LCT-related	SST-related	All genes in colinear region(%)	All xNL genes in each genome
<i>Arachis duranensis</i>	122 (15%)	243 (30%)		282 (34%)	819
<i>Arachis ipaensis</i>	87 (10%)	250 (30%)		264 (31%)	843
<i>Cicer arietinum</i>	185 (35%)	328 (63%)		345 (66%)	523
<i>Cajanus cajan</i>	18 (3%)	31 (4%)		48 (7%)	714
<i>Glycine max</i>	143 (19%)	441 (58%)	545 (72%)	597 (79%)	760
<i>Lotus japonicus</i>	40 (6%)	117 (18%)		131 (20%)	665
<i>Medicago truncatula</i>	88 (8%)	410 (37%)		431 (38%)	1,122
<i>Phaseolus vulgaris</i>	104 (14%)	259 (34%)		280 (37%)	762
<i>Vigna radiata</i>	50 (13%)	106 (27%)		120 (30%)	394
<i>Vigna angularis</i>	15 (3%)	56 (13%)		63 (15%)	431
<i>Vitis vinifera</i>	109 (13%)			109 (13%)	838

**Supplemental Table S24.** NBS-xNx genes related to duplication events in each genome.

Genome	ECH-related	LCT-related	SST-related	All genes in colinear region(%)	All xNx genes in each genome
<i>Arachis duranensis</i>	35 (11%)	71 (22%)		90 (27%)	329
<i>Arachis ipaensis</i>	28 (7%)	90 (24%)		90 (24%)	382
<i>Cicer arietinum</i>	26 (28%)	41 (44%)		44 (47%)	93



<i>Cajanus cajan</i>	6 (3%)	6 (3%)		10 (5%)	205
<i>Glycine max</i>	23 (10%)	129 (53%)	160 (67%)	178 (74%)	240
<i>Lotus japonicus</i>	6 (3%)	19 (11%)		21 (12%)	175
<i>Medicago truncatula</i>	36 (6%)	202 (31%)		210 (32%)	649
<i>Phaseolus vulgaris</i>	37 (10%)	76 (21%)		87 (25%)	355
<i>Vigna radiata</i>	5 (8%)	12 (19%)		13 (21%)	62
<i>Vigna angularis</i>	1 (1.5%)	3 (4%)		4 (6%)	67
<i>Vitis vinifera</i>	29 (7%)			29 (7%)	399

**Supplemental Table S25.** Bidirectional BLAST searched against all annotated genes between grape and legume.

Legume genomes	Grape(30,434) as reference						
	Grape search Legume	Blast best hit				Bidirectional best hit	Supporting by collinearity
		Supporting by collinearity	Legume search Grape	Supporting by collinearity			
barrel medic(51,135) <sup>a</sup>	14,466(47.53%)	5,948(41.12%)	15,311(29.94%)	5,511(35.99%)	7,494	3,723(49.68%)	
soybean(56,044)	14,795(48.61%)	6,243(42.20%)	31,611(56.40%)	9,680(30.62%)	8,292	4,230(51.01%)	
lotus(39,371)	12,984(42.66%)	4,830(37.20%)	17,013(43.21%)	4,839(28.44%)	6,667	2,969(44.53%)	
chickpea(28,269)	14,220(46.72%)	6,041(42.28%)	12,577(44.49%)	6,397(50.86%)	7,205	3,717(51.59%)	
pigeon pea(48,680)	14,887(48.92%)	5,014(33.68%)	15,806(32.47%)	4,843(30.64%)	7,651	3,085(40.32%)	
mung bean(22,368)	12,750(41.89%)	5,176(40.60%)	11,459(51.23%)	5,792(50.55%)	6,315	3,195(50.59%)	
adzuki bean(27,028)	13,273(43.61%)	3,719(28.02%)	12,241(45.29%)	3,246(26.52%)	6,564	2,116(32.24%)	
common bean(27,203)	15,111(49.65%)	6,790(44.93%)	13,605(50.01%)	7,029(51.66%)	7,882	4,262(54.07%)	
peanut A(36,734)	13,175(43.29%)	5,451(41.37%)	13,552(36.89%)	6,332(46.72%)	6,496	3,324(51.17%)	
peanut B(41,840)	12,722(41.80%)	5,143(40.43%)	14,494(34.64%)	6,344(43.77%)	6,266	3,143(50.16%)	

Alignment of identity  $\geq 60\%$ , and coverage  $\geq 50\%$ ; <sup>a</sup> Gene number

**Supplemental Table S26.** Bidirectional BLAST searched against all annotated genes between barrel medic and other legumes.

Legume genome	Barrel medic(51,135) as reference					
	barrel search other legumes	Supporting by collinearity	Blast best hit other legumes search barrel	Supporting by collinearity	Bidirectional best hit	Supporting by collinearity
soybean(56,044) <sup>a</sup>	21,488(42.02%)	11,171(51.97%)	38,055(67.90%)	16,362(43.00%)	11,596	7,476(64.47%)
lotus(39,371)	18,816(36.80%)	5,262(27.97%)	22,491(57.13%)	8,658(38.50%)	9,633	3,231(33.54%)
chickpea(28,269)	23,746(46.44%)	10,539(44.38%)	17,855(63.16%)	11,305(63.32%)	12,036	7,028(58.39%)
pigeon pea(48,680)	21,279(41.61%)	6,541(30.74%)	20,186(41.47%)	6,820(33.79%)	10,802	4,143(38.35%)
mung bean(22,368)	18,038(35.28%)	7,161(39.70%)	15,019(67.15%)	8,386(55.84%)	9,265	4,540(49.00%)
adzuki bean(27,028)	19,053(37.26%)	5,164(27.10%)	16,156(59.78%)	4,789(29.54%)	9,699	3,011(31.04%)
common bean(27,203)	22,010(43.04%)	10,916(49.60%)	17,475(64.24%)	11,240(64.32%)	11,543	6,967(60.36%)
peanut A(36,734)	17,900(35.01%)	7,557(42.22%)	17,198(46.82%)	9,287(54.00%)	8,850	4,742(53.58%)
peanut B(41,840)	17,215(33.67%)	7,183(41.73%)	18,497(44.21%)	9,363(50.62%)	8,439	4,442(52.64%)

Alignment of identity  $\geq 60\%$ , and coverage  $\geq 50\%$ ; <sup>a</sup> Gene number

**Supplemental Table S27.** Barrel medic soybean, and lotus genes against their respective EST sequences (Alignment of coverage  $\geq 30\%$ ).

Genome	Barrel medic			Soybean			Lotus		
	Total genes	Confirmed	Non-confirmed	Total genes	Confirmed	Non-confirmed	Total genes	Confirmed	Non-confirmed
Identity $\geq 40\%$		23,139(45.25%)	27,996(54.75%)		27,721(49.46%)	28,323(50.54%)		17,469(44.37%)	21,902(55.63%)
Identity $\geq 60$		23,139(45.25%)	27,996(54.75%)		27,721(49.46%)	28,323(50.54%)		17,469(44.37%)	21,902(55.63%)
Identity $\geq 70$	51,135	23,139(45.25%)	27,996(54.75%)	56,044	27,721(49.46%)	28,323(50.54%)	39,371	17,469(44.37%)	21,902(55.63%)
Identity $\geq 80$		23,122(45.22%)	28,013(54.78%)		27,744(49.50%)	28,300(50.50%)		17,398(44.19%)	21,973(55.81%)
Identity $\geq 90$		21,143(41.35%)	29,992(58.65%)		27,530(49.12%)	28,514(50.88%)		15,437(39.21%)	23,934(60.79%)

Alignment of coverage  $\geq 30\%$

**Supplemental Table S28.** Barrel medic soybean, and lotus genes against their respective EST sequences (Alignment of coverage  $\geq 30\%$ ).

Genome	Barrel medic			Soybean			Lotus		
	Total genes	Confirmed	Non-confirmed	Total genes	Confirmed	Non-confirmed	Total genes	Confirmed	Non-confirmed
Identity $\geq 40\%$		20,290(39.68%)	30,845(60.32%)		24,700(44.07%)	31,344(55.93%)		14,982(38.05%)	24,389(61.95%)
Identity $\geq 60$		20,290(39.68%)	30,845(60.32%)		24,700(44.07%)	31,344(55.93%)		14,982(38.05%)	24,389(61.95%)
Identity $\geq 70$	51,135	20,290(39.68%)	30,845(60.32%)	56,044	24,700(44.07%)	31,344(55.93%)	39,371	14,982(38.05%)	24,389(61.95%)
Identity $\geq 80$		20,260(39.62%)	30,875(60.38%)		24,698(44.07%)	31,346(55.93%)		14,939(37.94%)	24,432(62.06%)
Identity $\geq 90$		18,913(36.99%)	32,222(63.01%)		24,650(43.98%)	31,394(56.02%)		13,510(34.31%)	25,861(65.69%)

**Supplemental Table S29.** Information of original data material.

Order	Species name	Common name	Version	Data source	Journal	Reference
1	<i>Vitis vinifera</i>	Grape vine	Genoscope.12 X	JGI ( <a href="https://phytozome.jgi.doe.gov/pz/portal.html">https://phytozome.jgi.doe.gov/pz/portal.html</a> )	Nature	Jaillon O, <i>et al</i> , 2007 <sup>[1]</sup>
2	<i>Arachis duranensis</i>	Peanut	v1.0	Peanut Base ( <a href="http://peanutbase.org/">http://peanutbase.org/</a> )	Nature Genetics	Bertioli D J, <i>et al</i> , 2016 <sup>[2]</sup>
3	<i>Arachis ipaensis</i>	Peanut	v1.0	Peanut Base ( <a href="http://peanutbase.org/">http://peanutbase.org/</a> )	Nature Genetics	Bertioli D J, <i>et al</i> , 2016 <sup>[2]</sup>
4	<i>Lotus japonicus</i>	Lotus	v2.5	PGDD ( <a href="http://chibba.agtec.uga.edu/duplication">http://chibba.agtec.uga.edu/duplication</a> )	DNA Research	Sato S, <i>et al</i> , 2008 <sup>[3]</sup>
5	<i>Cicer arietinum</i>	Chickpea	v1.0	PGDD ( <a href="http://chibba.agtec.uga.edu/duplication">http://chibba.agtec.uga.edu/duplication</a> )	Nature Biotechnology	Varshney R K, <i>et al</i> , 2013 <sup>[4]</sup>
6	<i>Medicago truncatula</i>	Barrel medic	Mt4.0v1	JGI ( <a href="https://phytozome.jgi.doe.gov/pz/portal.html">https://phytozome.jgi.doe.gov/pz/portal.html</a> )	Nature	Young N D, <i>et al</i> , 2011 <sup>[5]</sup>
7	<i>Cajanus cajan</i>	Pigeonpea	Nov 2011	PGDD ( <a href="http://chibba.agtec.uga.edu/duplication">http://chibba.agtec.uga.edu/duplication</a> )	Nature Biotechnology	Varshney R K, 2012 <sup>[6]</sup>
8	<i>Glycine max</i>	Soybean	Wm82.a2.v1	JGI ( <a href="https://phytozome.jgi.doe.gov/pz/portal.html">https://phytozome.jgi.doe.gov/pz/portal.html</a> )	Nature	Schmutz J, <i>et al</i> , 2010 <sup>[7]</sup>
9	<i>Vigna angularis</i>	Adzuki bean	adzuki_ver3	Crop Genomics Lab. ( <a href="http://plantgenomics.snu.ac.kr/">http://plantgenomics.snu.ac.kr/</a> )	Proceedings of the National Academy of Sciences	Yang K, <i>et al</i> , 2015 <sup>[8]</sup>
10	<i>Vigna radiata</i>	Mung bean	Vradi.ver6	Crop Genomics Lab. ( <a href="http://plantgenomics.snu.ac.kr/">http://plantgenomics.snu.ac.kr/</a> )	Nature Communications	Kang Y J, <i>et al</i> , 2014 <sup>[9]</sup>
11	<i>Phaseolus vulgaris</i>	Common bean	v1.0	JGI ( <a href="https://phytozome.jgi.doe.gov/pz/portal.html">https://phytozome.jgi.doe.gov/pz/portal.html</a> )	Genome Biology	Vlasova A, <i>et al</i> , 2016 <sup>[10]</sup>

## Supplemental figure legends

**Supplemental Figure S1.** Homologous dotplot between *Vitis vinifera* and *Medicago truncatula* genomes. The best, secondary, and other matched homologous gene pairs output by Blast were dotplotted by red, blue, and gray colors in this figure, respectively. The *Vitis vinifera* 19 chromosomes colored by the 7 eudicot ancestral chromosomes, as described in the main text. Genomic orthology and outparalogy chromosome regions were identified and distinguished by solid and dashed squares, respectively. Bi-directional arrows link complement correspondence showing likely chromosome breakages during evolution.

**Supplemental Figure S2.** Homologous dotplot between *Vitis vinifera* and *Arachis duranensis* genomes. The best, secondary, and other matched homologous gene pairs output by Blast were dotplotted by red, blue, and gray colors in this figure, respectively. The *Vitis vinifera* 19 chromosomes colored by the 7 eudicot ancestral chromosomes, as described in the main text. Genomic orthology and outparalogy chromosome regions were identified and distinguished by solid and dashed squares, respectively. Bi-directional arrows link complement correspondence showing likely chromosome breakages during evolution.

**Supplemental Figure S3.** Homologous dotplot between *Vitis vinifera* and *Arachis ipaensis* genomes. The best, secondary, and other matched homologous gene pairs output by Blast were dotplotted by red, blue, and gray colors in this figure, respectively. The *Vitis vinifera* 19 chromosomes colored by the 7 eudicot ancestral chromosomes, as described in the main text. Genomic orthology and outparalogy chromosome regions were identified and distinguished by solid and dashed squares, respectively. Bi-directional arrows link complement correspondence showing likely chromosome breakages during evolution.

**Supplemental Figure S4.** Homologous dotplot between *Medicago (Medicago truncatula)* and Soybean (*Glycine max*) genomes. The best, secondarily best, and other matched homologous gene pairs output by BLAST were dotplotted by red, blue, and gray colors in this figure, respectively. Colinear gene blocks were inferred and circled out. We used median synonymous substitution rates or Ks to help distinguish the blocks from different events. Based chromosomal breakpoints and positional closeness, a set of soybean chromosomal regions making up one copy of a *medicago* chromosome are circled in red, and the other set making up to another copy of the same *medicago* chromosome are circled in black.

**Supplemental Figure S5.** Homologous alignments of 10 legume genomes with *Medicago truncatula* as reference. With *Medicago truncatula* as reference genome, genomic paralogy, orthology and outparalogy informations within and among 10 legumes were displayed in 22 circles.: The curved lines within inner circle, colored by eudicot ancestral 7 chromosomes , as

described in the main text. Linked paralog pairs on *Medicago truncatula* 8 chromosomes produced by LCT. The short lines forming the innermost circles represents all predicted genes in *Medicago truncatula*, which have one paralogous regions, forming another circle. Each of the two sets of *Medicago truncatula* paralogous chromosomal regions have one orthologous copies in a legume with exception of soybean, which would have 2. Therefore, 10 genome will result in 22  $((9+1 \times 2) \times 2)$  circles in the figure. Each circle is colored as to its source plant corresponding to the color scheme in Figure 1. Homologous genes denoted by shot lines standing on a chromosome circle, and colored as to its chromosome number in the source plant shown in the inset legend.

**Supplemental Figure S6.** GO analysis distribution of *Glycine max* retention genes produced by ECH, LCT and SST. The comparison results of retention *Glycine max* genes and GO database are shown in the figure. Here, those genes are divided into four categories: non-colinear (gray), LCT reserved (blue), SST reserved (green), ECH reserved (yellow). Possible gene function pathways are divided into three major categories and 45 pathways. The numbers of those genes were standardized by log function.

**Supplemental Figure S7.** GO analysis distribution of *Glycine max* lost genes in ECH, LCT, SST and LCT-SST. The comparison results of *glycine max* lost genes and GO database are shown in the figure. Here, those genes are divided into six categories: non-colinear (gray), ECH lost (green), LCT lost (blue), SST lost (purple), LCT-SST lost (yellow) and both retention after LCT and SST (orange). Possible gene function pathways are divided into three major categories and 45 pathways. The numbers of those genes were standardized by log function.

**Supplemental Figure S8.** Oil genes amplification model related to gene duplication events in soybean. (A) The curved lines within inner circle, colored by green, linked paralog pairs on soybean 20 chromosomes produced by SST, (B) LCT and (C) ECH. Oil subfamily genes with lightgreen (DSL&SFA), lightblue (SML in ES), lightsalmon (FAE&W&CM), green (LS), grey (MAL in M), yellow (ML), black (MAL in M7), blue (S&S8) and red (SFA in P9). Colored curved lines linked nodulation gene pairs with  $K_s < 0.15$ .

**Supplemental Figure S9.** NBS-class genes amplification model related to gene duplication events in soybean. (A) The curved lines within inner circle, colored by green, linked paralog pairs on soybean 20 chromosomes produced by SST, (B) LCT and (C) ECH. NBS-class subfamily genes with darkblue (NBS-CC) and darkorange (NBS-TIR). Colored curved lines linked nodulation gene pairs with  $K_s < 0.15$ .

**Supplemental Figure S10.** NBS-domains genes amplification model related to gene duplication events in soybean. (A) The curved lines within inner circle, colored by green, linked paralog pairs

on soybean 20 chromosomes produced by SST, **(B)** LCT and **(C)** ECH. NBS-domains subfamily genes with darkgreen (NBS-TNL), crimson (NBS-TNx), darkorange" (NBS-xNL) and navy (NBS-xNx). Colored curved lines linked nodulation gene pairs with  $K_s < 0.15$ .

**Supplemental Figure S11.** Homologous dotplot between grape and barrel medic chromosomes.

The grape chromosomes 9, 11, and part of 4, being homoeologous triplets produced by the eudicot-common hexaploidy, and their matched barrel medic chromosomes were aligned in horizontal and vertical directions, respectively. Red, blue, and gray dots were used to show the best, secondary, and other matched homologous genes respectively. The grape 4, 9 and 11 chromosomes colored by the 7 eudicot ancestral chromosomes, as described in the main text. Best-matched or orthologous regions were marked out by solid-line rectangles numbered by 1 and 2 in red circles; outparalogous regions or secondary-matched were marked out by broken-lined rectangles numbered by 1-4 in blue circles. Arrows show complement correspondence produced by chromosome breakages during evolution.

## References

- [1] Jaillon O, Aury J M, Noel B, et al. The grapevine genome sequence suggests ancestral hexaploidization in major angiosperm phyla[J]. *Nature*, 2007, 449(7161): 463-7.
- [2] Bertioli D J, Cannon S B, Froenicke L, et al. The genome sequences of *Arachis duranensis* and *Arachis ipaensis*, the diploid ancestors of cultivated peanut[J]. *Nat Genet*, 2016, 48(4): 438-46.
- [3] Sato S, Nakamura Y, Kaneko T, et al. Genome structure of the legume, *Lotus japonicus*[J]. *DNA Res*, 2008, 15(4): 227-39.
- [4] Varshney R K, Song C, Saxena R K, et al. Draft genome sequence of chickpea (*Cicer arietinum*) provides a resource for trait improvement[J]. *Nat Biotechnol*, 2013, 31(3): 240-6.
- [5] Young N D, Debelle F, Oldroyd G E, et al. The Medicago genome provides insight into the evolution of rhizobial symbioses[J]. *Nature*, 2011, 480(7378): 520-4.
- [6] Varshney R K, Chen W, Li Y, et al. Draft genome sequence of pigeonpea (*Cajanus cajan*), an orphan legume crop of resource-poor farmers[J]. *Nat Biotechnol*, 2012, 30(1): 83-9.
- [7] Schmutz J, Cannon S B, Schlueter J, et al. Genome sequence of the palaeopolyploid soybean[J]. *Nature*, 2010, 463(7278): 178-83.
- [8] Yang K, Tian Z, Chen C, et al. Genome sequencing of adzuki bean (*Vigna angularis*) provides insight into high starch and low fat accumulation and domestication[J]. *Proc Natl Acad Sci U S A*, 2015, 112(43): 13213-8.
- [9] Kang Y J, Kim S K, Kim M Y, et al. Genome sequence of mungbean and insights into evolution within *Vigna* species[J]. *Nat Commun*, 2014, 5: 5443.
- [10] Vlasova A, Capella-Gutierrez S, Rendon-Anaya M, et al. Genome and transcriptome analysis of the Mesoamerican common bean and the role of gene duplications in establishing tissue and temporal specialization of genes[J]. *Genome Biol*, 2016, 17: 32.