# SUPPLEMENTARY DATA

**Fig. S1.** Dot matrix plots showing the presence of a 36 kb inversion in *Lupinus luteus* plastome. (A) Comparison of *Lupinus luteus* to *Glycine max* (Fabaceae) plastome revealing the presence of a 36 kb inversion. (B) Comparison of *Lupinus luteus* to *Cucumis sativus* (Fabaceae outgroup) revealing the presence of a 36 kb inversion embedded in the 50 kb inversion present in most Papilionideae (Fabaceae).

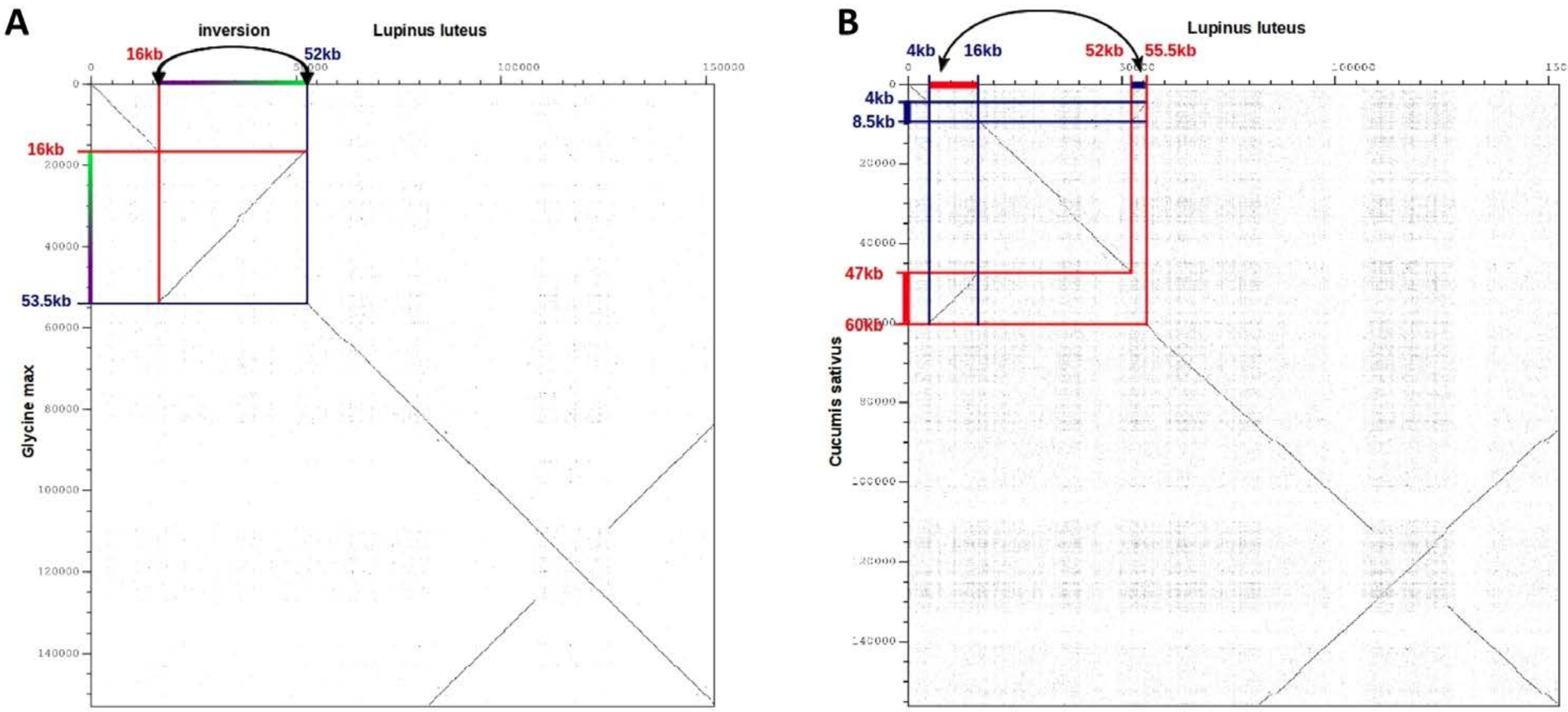

**Fig. S2.** Comparative plastomic maps showing the presence of a 36 kb inversion in *Lupinus luteus* in comparison to other Papilionoideae. The genes present and directly surrounding this inversion are shown. Only four representatives of the Fabaceae species having their plastomes fully sequenced (*Glycine max*, *Phaseolus vulgaris*, *Lotus japonicus* and *Medicago truncatula*) are presented here and the phylogenetic relationships between the different Papilionoideae represented are redrawn from Cardoso et al. (2012). The blue circle on a branch indicates the origin of the 36kb inversion event. The partial plastomic maps are drawn to scale.

**Fig. S3.** Phylogenetic analysis of plastidic and nuclear *rpl22* protein sequences. Presented is a Neighbor Joining tree obtained using Jukes–Cantor model and rooted using the eubacteria *Mycoplasma*. Bootstrap values obtained from 10,000 replicates are shown above the branches. The scale bar denotes substitutions per site along the branches. GenBank accession numbers of the *rpl22* chloroplastic sequences are as follows: *Cyanophora paradoxa* (NC_001675), *Gracilaria tenuistipitata* (NC_006137), *Marchantia polymorpha* (NC_001319) *Mycoplasma* (M74770), *Nicotiana tabacum* (NC_001879), *Oryza sativa indica* (NC_008155), *Pelargonium* x. *hortorum* (NC_008454), *Spinacia oleracea* (NC_002202) and *Zea mays* (NC_001666). GenBank accession numbers of the *rpl22* genes functionally transferred from the chloroplast to the nucleus in a common ancestor of all flowering plants (Gantt *et al.*, 1991) are: *Lupinus mariae josephi* nuc (sequence available upon request) and *Pisum sativum* nuc (M60951).

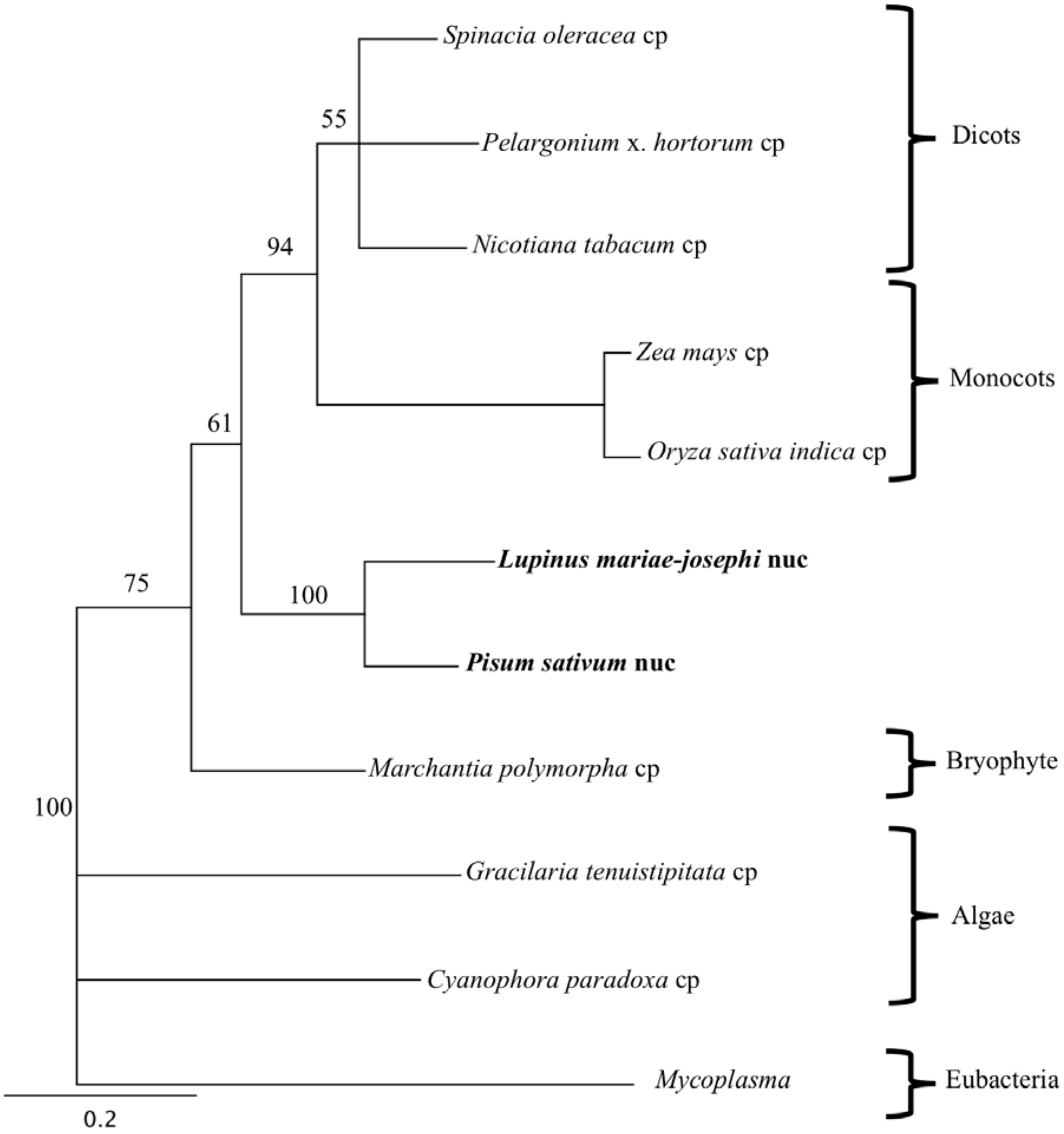

**Fig. S4.** Pairwise distance between *Lupinus luteus* and other Fabaceae orthologous plastomic regions. Pairwise distance was calculated using K2p model (Kimura, 1980) for introns and intergenic spacers, whereas sequence divergence of protein-coding genes (exons) were calculated using the synonymous (Ks) nucleotide substitution rates. The protein-coding, intronic or intergenic regions presenting a higher evolutionary rate than those previously used in Fabaceae evolutionary studies (such as *rbcL* gene, *trnK-UUU* intron and *trnL-trnT* IGS) and a minimum size of 300 bp are highlighted in red (exons), green (introns) and blue (IGS). The size (in bp) of each highlighted region in *Lupinus* is indicated between brackets. Only one IR region is represented on each plastomic map. The endpoints of the 50 kb and 36 kb inversions are indicated by red and green arrows, respectively.

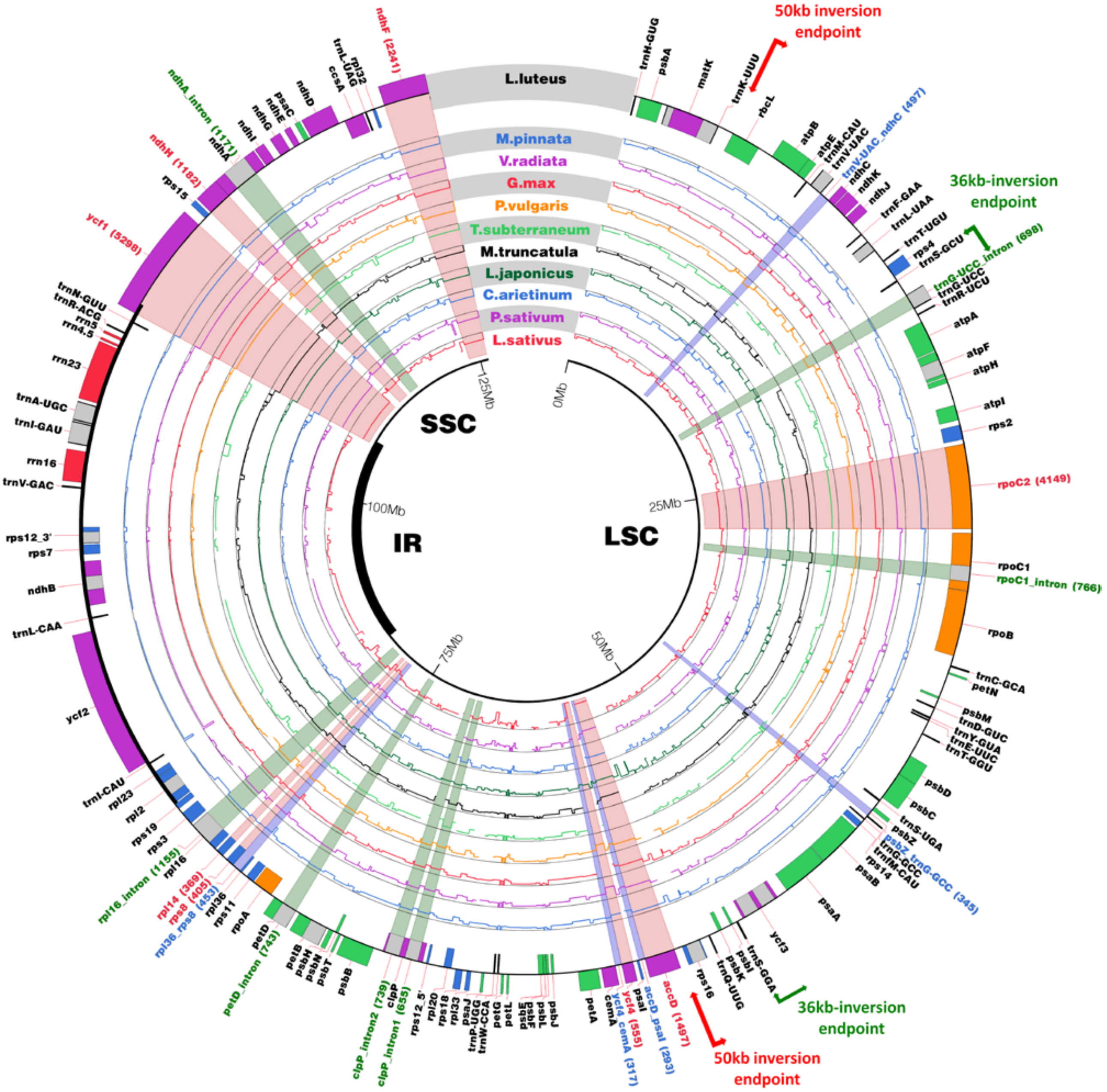

**Fig. S5.** Synonymous and non-synonymous divergence in Legume *ycf4* gene. Shown are *dN* (left) and *dS* (right) phylogenetic trees obtained from a codon based maximum likelihood method (GTR + G + I model). The trees, obtained using MEGA 5.0 (Tamura, 2011), were rooted with the *Populus alba* sequence. Bootstrap values obtained from 1,000 replicates are shown on each branch. The asterisk indicates the branch in which rate acceleration is first seen (Magee *et al.*, 2010). The scale bar denotes substitutions per site along the branches. The legume *ycf4* sequences were obtained from Stefanovic *et al.* (2009) and Magee *et al.* (2010).