Supplemental Information: Asymmetric evolution of protein domains in the leucine-rich repeat receptor-like kinase (LRR-RLK) family of plant signaling proteins

Jarrett Man, Thomas Harrington, Kyra Lally, Madelaine E. Bartlett.

Supplemental Figures

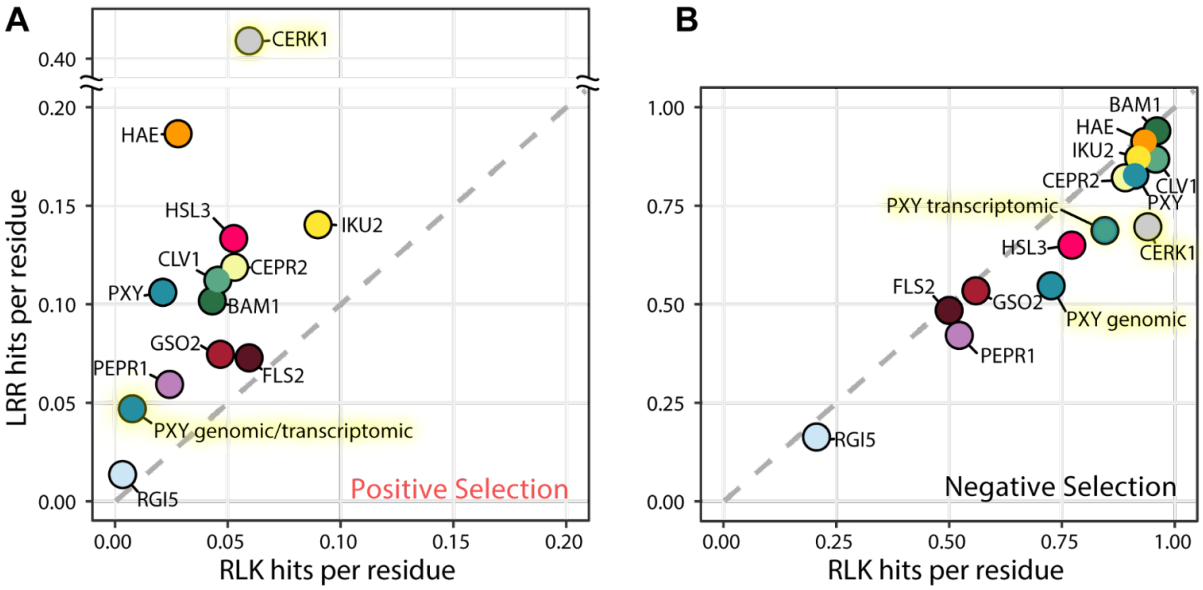
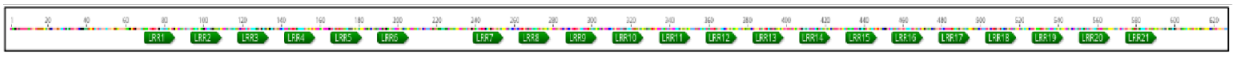


FIG. S1. Transcriptome sequencing bias likely doesn't underlie higher positive selection acting on LRR domains. *(A)* Positive selection detection rate for *CERK1*, *PXY* from genomic and *PXY* from transcriptomic datasets are also shown in addition to the other LRR-RLK clades. The two new *PXY* data points are identical. *(B)* Negative selection detection rate for *CERK1*, *PXY* from genomic and *PXY* from transcriptomic datasets are also shown in addition to the other LRR-RLK clades. Gray dashed line defines the region of no ratiometric bias between domains.

Amborella trichopoda BAM1 (Not in alignment)



BAM1 clade alignment without *Amborella trichopoda* BAM1

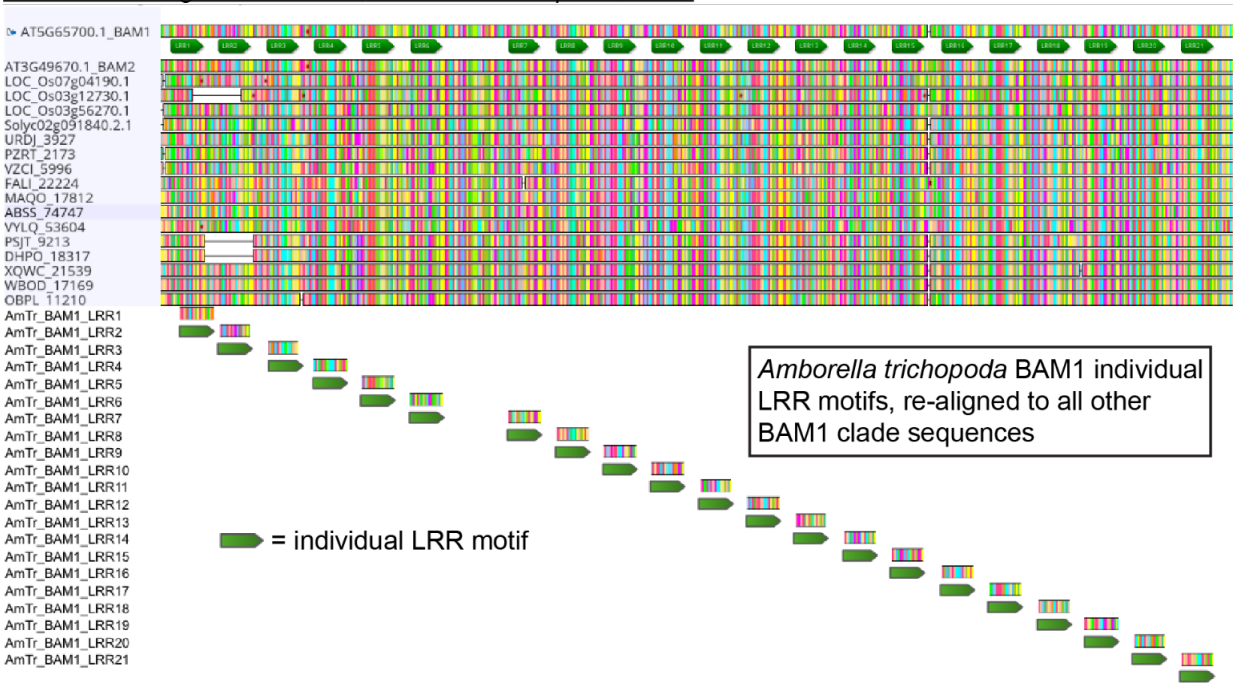


FIG. S2. LRR motif misalignment is unlikely to explain higher positive selection in LRR domains. Individual LRR motifs extracted from *Amborella trichopoda* BAM1 (top panel), aligned to all other BAM1 sequences, align correctly (bottom panel).

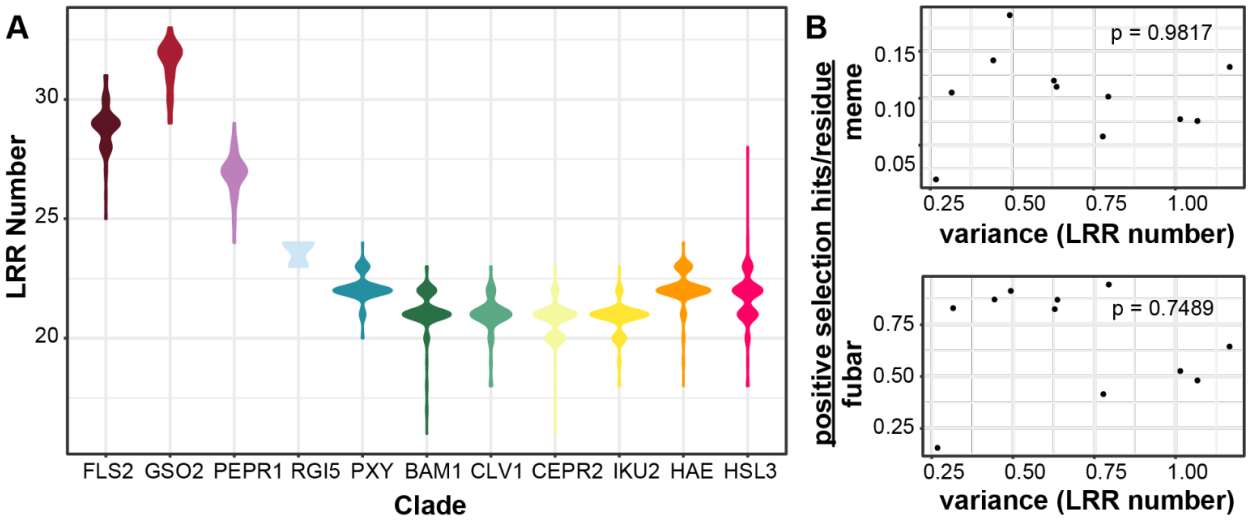


FIG. S3. Variation in LRR number cannot account for variation in positive selection. (*A*) LRR motif counts per clade. (*B*) The number of positive selection hits per LRR domain residue is not related to variance in LRR number per clade (p-values from Pearson correlation test).

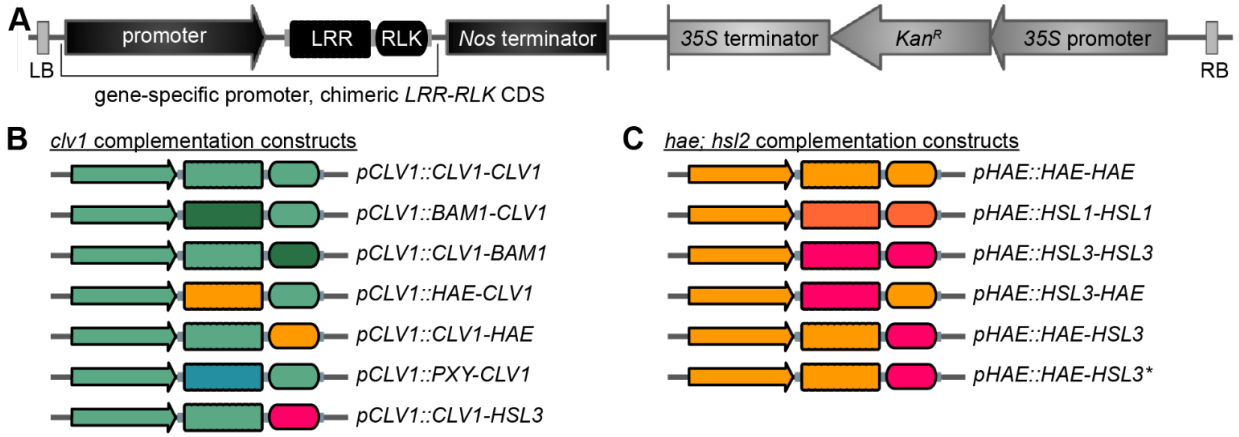


FIG. S4. Complementation constructs. (*A*) generalized T-DNA structure for all constructs. (*B*) Promoter and CDS sequences used in *clv1* complementation experiments. (*C*) Promoter and CDS sequences used in *hae; hsl2* complementation experiments. *HSL3** = coding sequence for enzyme-dead HSL3. UTRs derived from cognate *LRR*- or *RLK*-encoding genes in all cases.

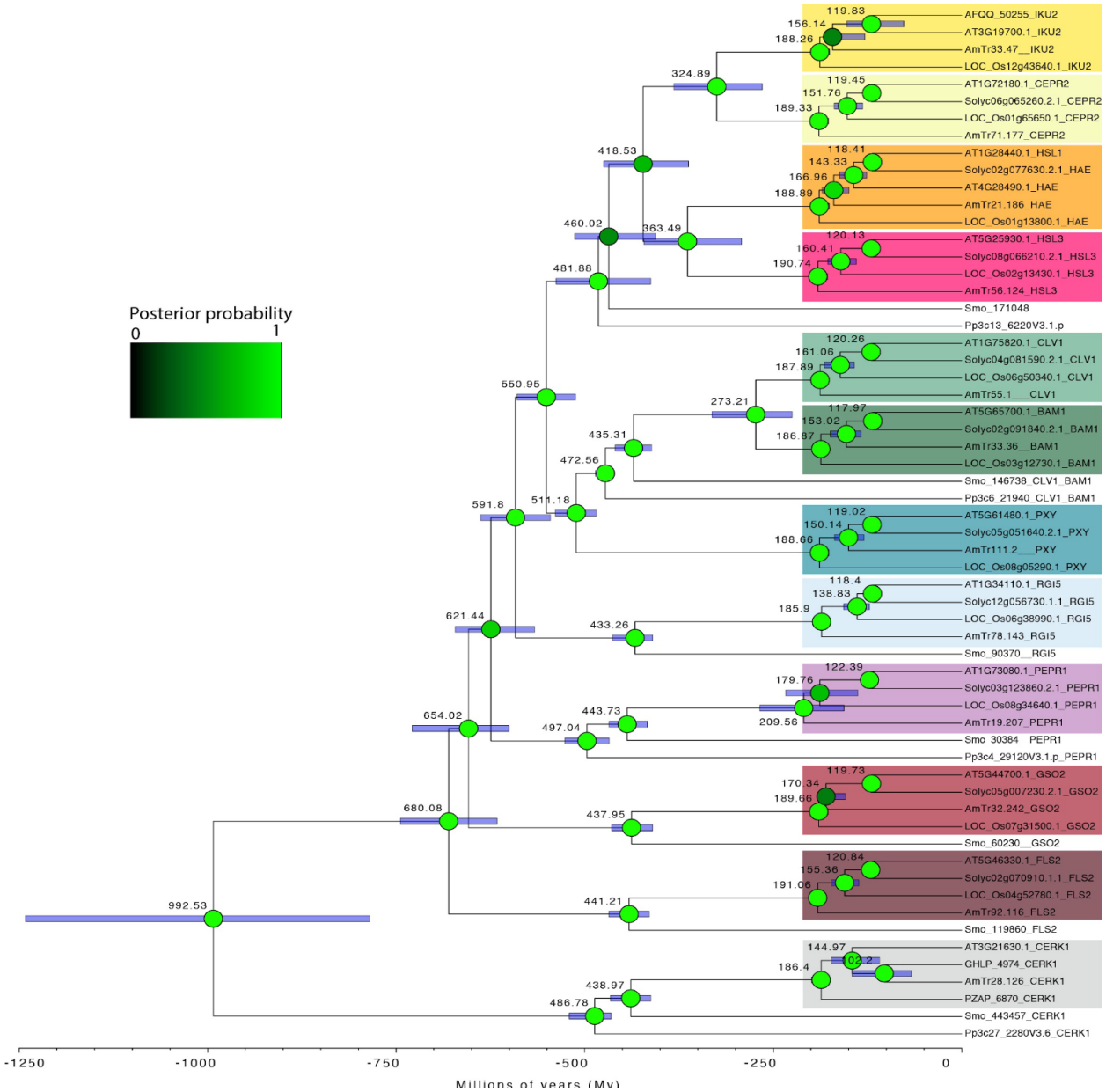


FIG. S5. Chronogram for inference of gene clade divergence times and tree topology. 4 angiosperm representatives from 11 LRR-RLK clades, and from *CERK1*, as well as well-supported representatives from *S. moellendorffii* and *P. patens* sister to those clades, were used to infer a time-calibrated gene divergence tree by Bayesian analysis. Blue bars over nodes represent 95% distribution probability range, while circles over nodes are colored according to posterior probability. Number next to each node is the estimated mean divergence time.

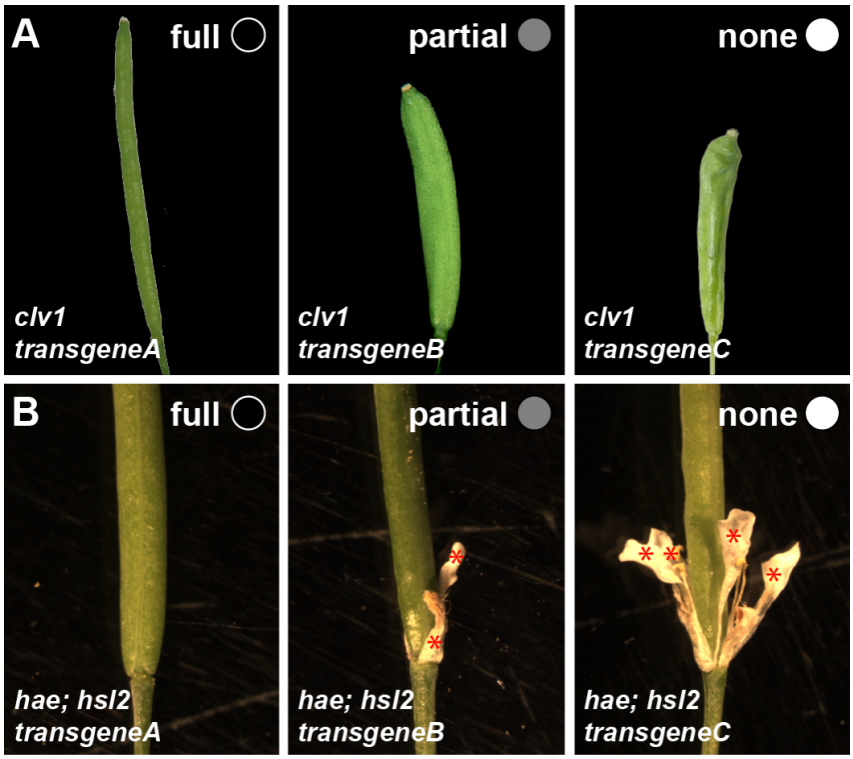


FIG. S6. Representative siliques from full, partial, and no complementation of mutants. (*A*) *clv1* complementation (*B) hae; hsl2* complementation, * = retained floral organ

Supplemental Tables

Table S1. Positive and negative selection acting on select proteins in the Papaveraceae

	BAM1	**CEPR2**	**HAE**	**CERK1**
MEME				
LRR residues under positive selection	11	17	16	15
RLK residues under positive selection	4	6	4	3
Chi^2 p-val	4.88E-01	3.24E-01	1.32E-01	2.03E-04
FUBAR				
LRR residues under negative selection	383	263	337	49
RLK residues under negative selection	228	156	204	106
Chi^2 p-val	2.76E-01	5.88E-01	4.59E-01	2.62E-02