# Supplementary Figures and Notes

**Supplemental Figure S1.** Sequence logos of conserved motifs in PDH1 protein sequences identified by Multiple Em for Motif Elicitation (MEME, https://meme-suite.org/meme/tools/meme). Motif colours in legend correspond to conserved blocks in Figure 1.

****

**Supplemental Figure S2:** Example CpG plot generated by EMBOSS for *GmPDH1* genomic context (50kb upstream and downstream of transcriptional start site). The ‘island’ shown at the bottom is within *GmSCL15*, additional information available at Supplemental Table S4.

****

**Supplemental Figure S3:** Timeline of black-eyed pea pod and seed development from late embryogenesis to maturity, with key pod development stages studied in Figure 3 indicated at the bottom (P6/P10/P16). Image taken directly from Yao et al. 2016 unaltered, additional information available at DOI: 10.1111/tpj.13279.

**Appendix S1.** All non-*Glycine* phaseoloids possess a second *SCL15* gene annotated elsewhere in the genome, which is adjacent to a ‘*HVA22*-like protein’ (*HVA22*) in all cases except *V. radiata* (Table SX). *G. max* possesses paralogs of the *SCL15-HVA22* dyad as a result of the *Glycine* specific whole genome duplication. The *SCL15*-*HVA22* appears to have been lost in other legume clades including the IRLC (*P. sativum, M. trunculata*) and genistoids (*L. albus*) for which no matching annotation was present.

**Appendix S2.** The only other pod specific (<100 FPKM in every non-pod tissues), ultra-high expressed gene in mature pod shells (16dap), Vigun09g047400, is annotated as an ‘Unknown protein’. When comparing the genic sequence of Vigun09g047400, the only BLAST hit (E-value = 5.40e-62 and E-value = 8.91e-70 for both exons) in the other legume genomes was LOC108318781, a ‘Repetitive proline-rich cell wall protein 1-like’ gene on an unplaced pangenome scaffold (NW\_016152098.1) in *V. angularis*.