

**APPENDIX S3.** Summary statistics derived from the genome (left) and transcriptome (right) assemblies. Two *k*-mer sizes (k56 and k64) were used in the genome assemblies. Species analyzed are abbreviated as follows: Ath = *Arabidopsis thaliana* (*Arabidopsis*); Lpur = *Lablab purpureus* (*lablab*); Osat = *Oryza sativa* (*rice*); Gmax = *Glycine max* (*soybean*); Slyc = *Solanum lycopersicum* (*tomato*). Four different depths of sequencing were used (2M, 5M, 10M, 20M reads).

