

Supplemental figures Kourelis et al: *A homology-guided, genome-based proteome for improved proteomics in the allopoloid Nicotiana benthamiana*

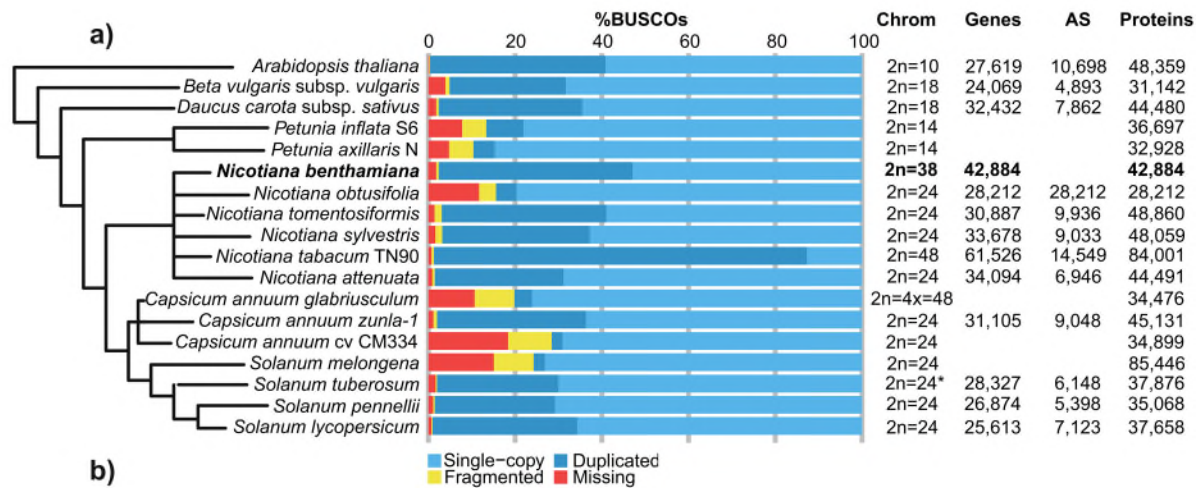
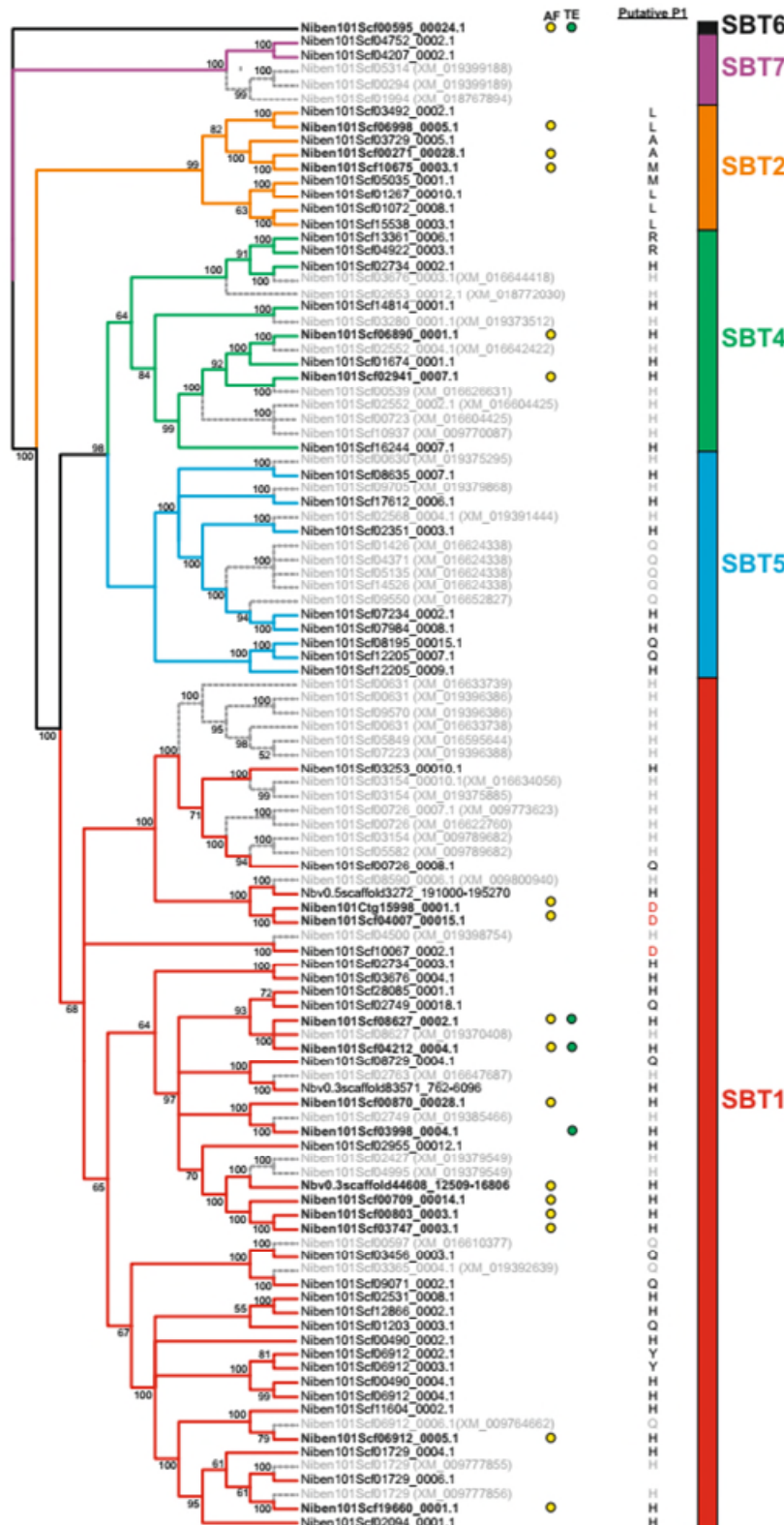


Figure S1: Comparison of Solanaceae proteomes. **a)** Completeness of the predicted proteomes from sequenced Solanaceae genomes was estimated using BUSCO v3 with the embryophyta database. *Certain *Solanum tuberosum* species are polyploid. **b)** Violin and boxplot g of log₁₀ protein length distribution of each predicted proteome. Jittered dots show the raw underlying data.



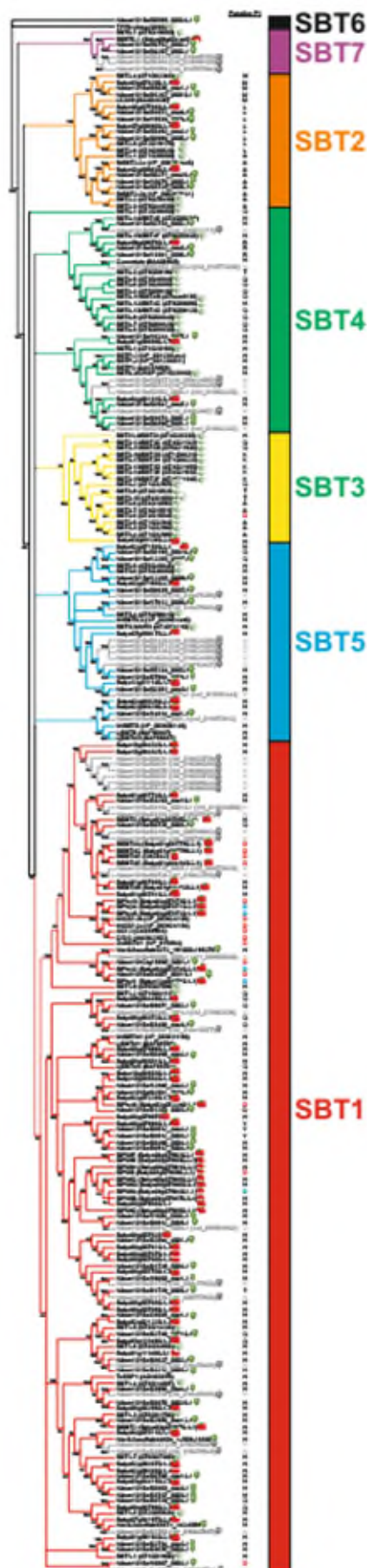


Figure S3: Phylogenetic analysis of the subtilisin gene-family of tomato and Arabidopsis and including other previously characterized subtilisins. The evolutionary history of the subtilase gene family was inferred by using the Maximum Likelihood method based on the Whelan and Goldman model. The bootstrap consensus tree inferred from 250 replicates is taken to represent the evolutionary history of the taxa analysed. Putative pseudogenes are indicated in grey. Putative P1 based on residue at the prodomain junction is indicated. Putative phytaspases are indicated in red, confirmed phytaspases in blue. Naming of subtilase clades according to (Taylor and Qiu, 2017).