Genome finishing (closing gaps in scaffolds)

Repeat library

Masked genome sequence of *Phialocephala subalpina*

Gene predictors

Training dataset ("gold standard")

EST alignments

Trans-alignments

~ 1,000 manually annotated full-length cDNAs

~ 27,500 454 ESTs ~ 28,000 454 sESTs

Trans-alignments of protein data

GeneMark

FGENESH (2 matrices)

Augustus

G

G

G

Selection of best gene model by jigsaw

Manual QC of gene calls in Apollo

Functional annotation in PEDANT

G = GenomeThreader (spliced alignments of ESTs, cDNA and proteins using splice site models trained for *P. subalpina*)