

Illumina reads
(100-150 bp)



Data Processing
(thinning, trimming, adaptor removal)



de novo Assembly
(CLC Genomics Workbench 4)



Contig sequences



Local tblastn for candidate genes
(CLC Genomics Workbench 4)



blastx against nr NCBI database
E-value $<10^{-5}$



Functional Annotation
(GO + PFAM)



Sequencing Depth
(Ortholog Hit Ratio)