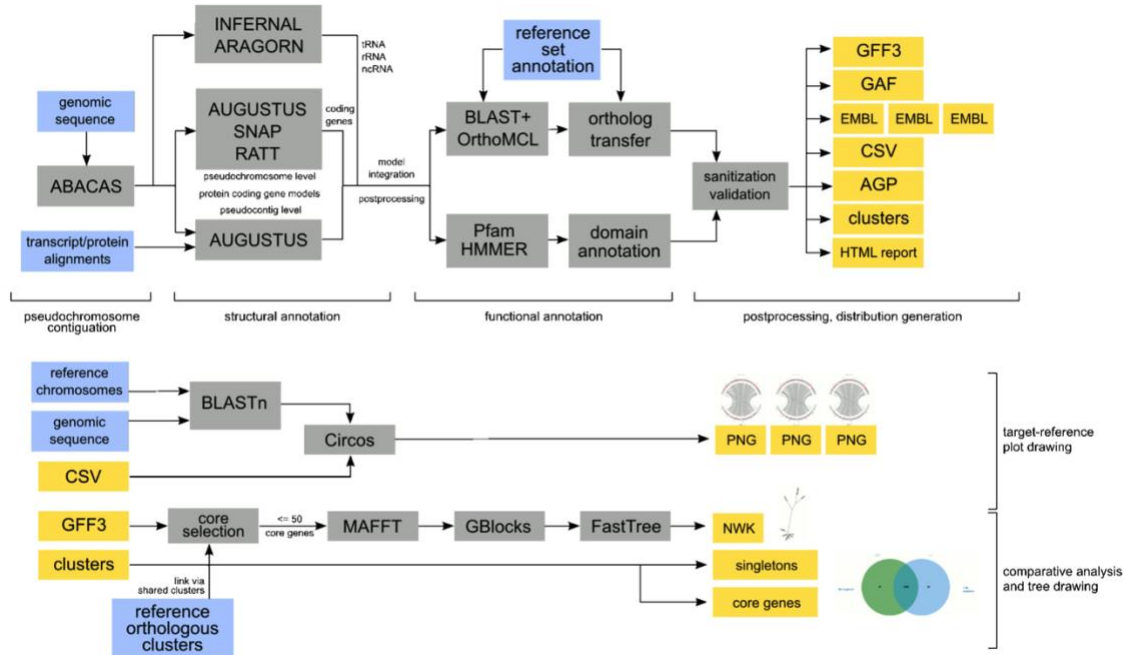


Genome Annotation with Companion (Part 2)

You should have gotten an email indicating the status of your annotation (ie. job started and job complete). The email contains a link to the annotation output.



- Explore your results with your group and discuss the annotation findings:
 - What does the genome statics tab tell you about your annotation? Are the results surprising? You can explore the reference genome you used in VEuPathDB to help you assess the results. (For example, are you getting a reasonable number of genes? What about the GC content? Number of non-coding genes?)










Pcoa-Pkno (PCOA)

Completed

This job was submitted 6 days ago and ran for about 3 hours, finally finishing at 2020-09-09 15:33:07 UTC.

Genome statistics	Result files	Orthology	Phylogeny	Synteny	Job parameters	Pipeline logs	Validator report
							Value
Number of annotated regions/sequences							14
Number of genes							5011
Gene density (genes/megabase)							178.54
Number of coding genes							4943
Number of pseudogenes							533
Number of genes with function							4793
Number of pseudogenes with function							528
Number of non-coding genes							68
Number of genes with multiple CDSs							2693
Overall GC%							39.65
Coding GC%							41.89

- What does the “Result files” tab contain? What is an AGP file? What is a GFF3 file?

Genome statistics	Result files	Orthology	Phylogeny	Synteny	Job parameters	Pipeline logs	Validator report			
								Format	MD5	Size
Pseudochromosome level genomic sequence								FASTA		7.79 MB
Pseudochromosome level gene annotations								GFF3		5.44 MB
Pseudochromosome layout								AGP		618 Bytes
Scaffold level genomic sequence								FASTA		7.79 MB
Scaffold level gene annotations								GFF3		5.47 MB
Scaffold layout								AGP		825 Bytes
Pseudochromosome level sequence and annotation								EMBL		13.4 MB
Gene Ontology function assignments								GAF1		1.65 MB
Protein sequences								FASTA		3.99 MB

- What does the “orthology” tab display? How many predicted proteins from your new genome are in common with ones from the reference genome? How many are unique to yours? What do singletons represent (click on the singleton number to see what these genes are)?

