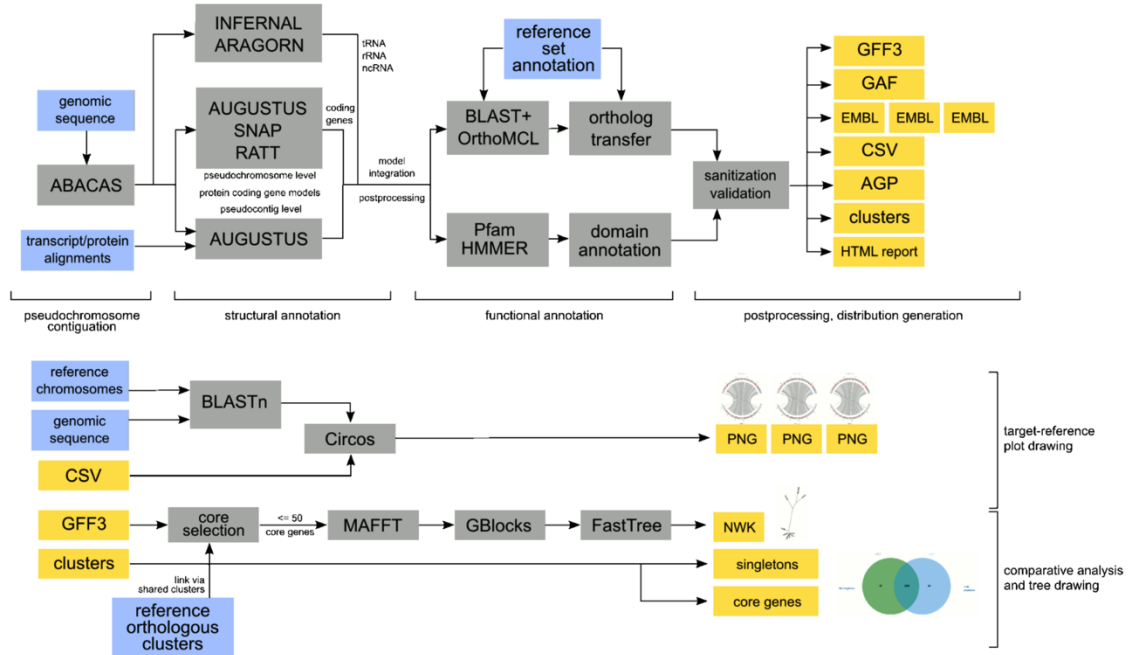


Genome Annotation with Companion (Part 2)

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 EuPathDB 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?)

[sanger COMPANION](#) [Submit job](#) [Getting started](#) [Example results](#) [FAQ](#) 1/2

P_coatneyi_Hackeri (pcoa)

Completed

This job was submitted 4 days ago and ran for about 4 hours, finally finishing at 2016-02-13 08:49:01 UTC.

Genome statistics | **Result files** | Orthology | Phylogeny | Synteny | Job parameters | Pipeline logs | Validator report

	Value
Number of annotated regions/sequences	15
Number of genes	5191
Gene density (genes/megabase)	185.11
Number of coding genes	5128
Number of genes with function	4949
Number of non-coding genes	65
Number of genes with multiple CDSs	2704
Overall GC%	39.66
Coding GC%	41.91

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

P_coatneyi_Hackeri (pcoa) Completed

This job was submitted 4 days ago and ran for about 4 hours, finally finishing at 2016-02-13 08:49:01 UTC.

Genome statistics **Result files** Orthology Phylogeny Synteny Job parameters Pipeline logs Validator report

	MD5	Size
Pseudochromosome level genomic sequence (FASTA)		7.81 MB
Pseudochromosome level genes (GFF3)		4.65 MB
Pseudochromosome layout (AGP)		32.6 KB
Scaffold level genomic sequence (FASTA)		7.82 MB
Scaffold level genes (GFF3)		7.95 MB
Scaffold layout (AGP)		165 KB
Pseudochromosome level sequence and annotation (EMBL)		13.5 MB
Functional GO annotation (GAF1)		1.56 MB
Protein sequences (FASTA)		3.56 MB

All files are provided in the following formats: GFF3 (format specification), FASTA (format specification), EMBL (format specification), GAF 1.0 (format specification) and AGP 2.0 (format specification).

- 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)?

P_coatneyi_Hackeri (pcoa) Completed

This job was submitted 4 days ago and ran for about 4 hours, finally finishing at 2016-02-13 08:49:01 UTC.

Genome statistics **Result files** **Orthology** Phylogeny Synteny Job parameters Pipeline logs Validator report

These Venn diagrams show shared and species-specific gene clusters in the target genome pcoa (left, green) and the Plasmodium knowlesi reference (right, blue). Singletons, i.e. genes without orthologs and paralogs, are shown outside the Venn diagram to the left and right.

Click on the numbers to view clusters and singleton gene lists.

Category	Count
pcoa singletons	181
Shared genes	4553
Pknowlesi singletons	554

[Cluster assignments as OrthoMCL 1.4 output file](#)

- What does the phylogeny tab represent? Does it make sense?

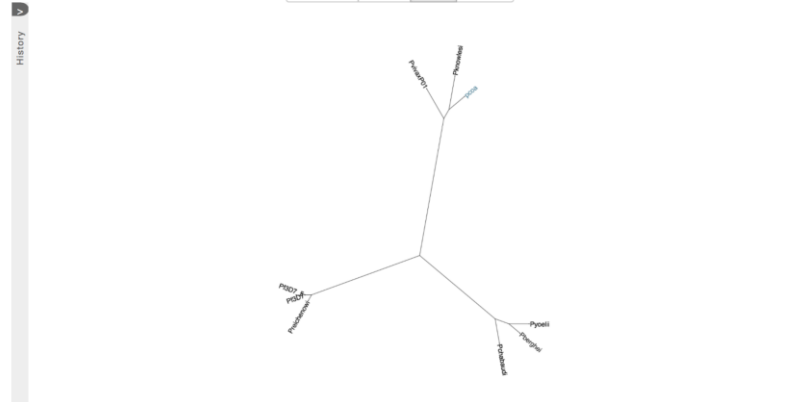
P_coatneyi_Hackeri (pcoa) Completed

This job was submitted 4 days ago and ran for about 4 hours, finally finishing at 2016-02-13 08:49:01 UTC.

[Genome statistics](#) [Result files](#) [Orthology](#) **Phylogeny** [Synteny](#) [Job parameters](#) [Pipeline logs](#) [Validator report](#)

Click and drag in the diagram below to pan around. Use the mouse wheel to zoom in and out. The newly annotated genome in this job is highlighted: pcoa.

Rectangular Circular Radial Diagonal



[Multiple sequence alignment for this tree \(FASTA\)](#)
[Core genes used to build this tree](#)

- Examine the Synteny tab – are these genomes syntenic?

P_coatneyi_Hackeri (pcoa) Completed

This job was submitted 4 days ago and ran for about 4 hours, finally finishing at 2016-02-13 08:49:01 UTC.

[Genome statistics](#) [Result files](#) [Orthology](#) **Synteny** [Phylogeny](#) [Job parameters](#) [Pipeline logs](#) [Validator report](#)

Each circle below represents a single target-reference pseudochromosome alignment. Click on the thumbnail to zoom in. [Download all 15 Images \(ZIP\)](#)

