

Xenopus Genomes

Xenopus Bioinformatics Workshop 2018

Joshua Fortriede - Xenbase



Who creates the files?

- The assemblies and gene models are created by the JGI working group under the supervision of Richard Harland and Dan Rokshar
- Xenbase is responsible for housing the genomes, and maintaining the gene models
 - Involves helping to name the models
 - Feedback and contributions from the community are always appreciated

Current Versions – NCBI submitted

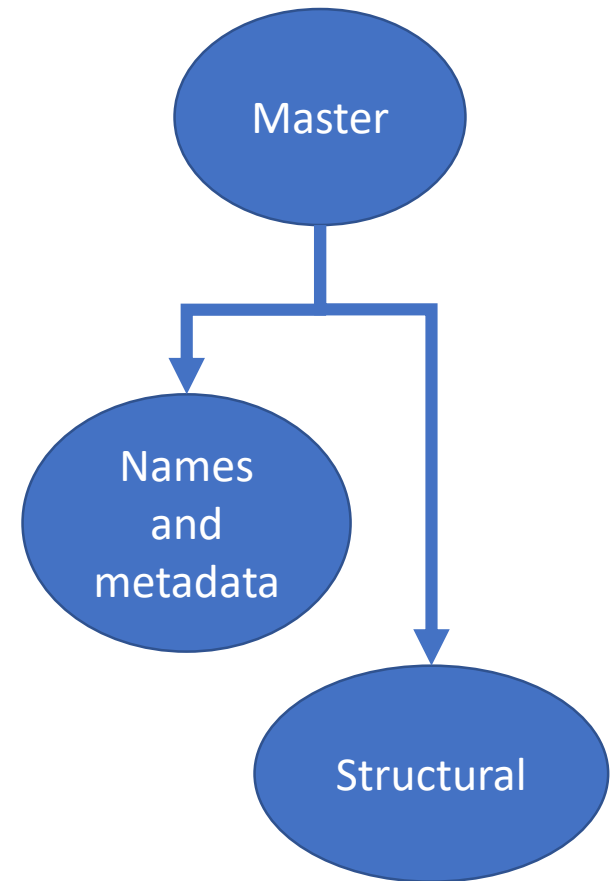
- *X. tropicalis* – v 9.1
 - Chromosome Level w/ <7000 scaffolds
 - 34,000 models w/ +18,000 named
- *X. laevis* – v 9.2
 - Chromosome Level w/ 100,000 scaffolds
 - 51,000 models w/ +22,000 named
- Download from Xenbase FTP and GitLab site
- View on Xenbase JBrowse

GitLab – GFF Files and Updates

- For the latest Genomes, both JGI and NCBI created separate, independent gene models
 - These models were merged together to create a “Master” Xenbase track
 - Some models present only in 1 of JGI or NCBI files
- Updates to the GFF files are now handled on [GitLab](#)
 - Allows ad hoc changes and robust tracking/versioning
 - Issue tracker allows for community feedback

GitLab – Three Branches

- Master
 - Official, released GFF File
- Names and metadata
 - Working section devoted to updating names and linking to Xenbase and Entrez genes
 - Will be regularly incorporated back into Master
- Structural
 - Working section devoted to fixing exons, adding, or removing models



GitLab – Contributing




- GitLab Wiki contains a [Contributions Guide](#) that outlines how to submit issue requests
 - Name Changes and adding Entrez/Xenbase links
 - Merging models
 - Creation of new models
 - Exon Editing
- A guide for large-scale contributions will be coming soon

GitLab – Getting Files

☰ | Xenbase > genomics > XENTR_9.1 > Repository

master ▾

XENTR_9.1 / + ▾

Name	Last update
 .gitignore	8 months ago
 README.md	a month ago
 XENTR_9.1_Xenbase.gff	8 months ago

 README.md

Xenopus tropicalis 9.1 genome assembly files

This project houses gene models (GFF) and other metadata files for the X. tropicalis 9.1 genome assembly.

☰ | Xenbase > genomics > XENTR_9.1 > Repository

master ▾

XENTR_9.1 / XENTR_9.1_Xenbase.gff

🔍 Find file

History

Permalink

 XENTR_9.1_Xenbase.gff 106 MB 



The source could not be displayed because it is larger than 10 MB. You can [download it](#) instead.

Future Work

- Improvement of gene models is of **top** priority
- [Synteny analysis](#) to help name models
 - Almost done with Human/Tropicalis synteny
 - Will run Tropicalis/Laevis synteny to carryover names
 - Finish with Human/Laevis synteny to catch anything that is missing from Tropicalis, or where tropicalis locus is incomplete
- Creation of new gene pages to house ortholog data, gene expression, NGS data, and link tropicalis and laevis genes

How else can we help?

- What additional files would you like in regards to the genome?
 - FASTA files – CDS, mRNA, protein
 - Model ID, current Symbol – aid in updating older results with newer annotations
 - List of problematic models?
 - Wrong start ATG
 - Missing ATG/Stop codon
 - Missing exons

Questions?