**ReadMe\_UCSC Genome Browser**

**1. *Xiphophorus***

**2. UCSC genome browser**

**3. Navigate the genome**

**4. Data visualization**

**5. Data download**

**1. *Xiphophorus***

If you are interested in *Xiphophorus*, please visit the *Xiphophorus* Genetic Stock Center website (www.xiphophorus.txstate.edu/) for more details.

**2. UCSC genome browser**

UCSC genome browser implemented at *Xiphophorus* genetic stock center is a mirror of UCSC genome browser (<https://genome.ucsc.edu/>; the human genome browser at UCSC. Genome Res. 2002 Jun;12(6):996-1006.). For a full tutorial of this platform, please refer to UCSC genome browsers website.

**3. Navigate the genome**

The genome database allow four types of input to navigate the genome:

-1 Ensembl transcript ID: ENSXMAT\* (exp. ENSXMAT00000003110)

-2 Ensembl gene ID: ENSXMAG\* (exp. ENSXMAG00000003084)

-3 Ensembl external gene name: Name:\* (exp. Name:esr1), note the “Name:” prefix is required.

-4 chromosome name and coordinates: chr\* (exp. chr5) or chr\*(exp. chr5:9,450,000-15,890,000). Note the “,” is required.

**4. Data** **visualization**

The genome browser is launched with gene model (EnsGTF) and 3 types of data (multiz: *Xiphophorus* species sequence alignments); Xmx0: RNA-Seq gene expression profile data; and Short Match: short sequence match. Users can choose to show or hide these data by changing the drop-down list in “Mapping and Sequencing Tracks”. To get sequences of homologs loci among *Xiphophorus* species, click the “multiz” track, the browser will show sequence alignment of loci of interest.

**---Import data from workstation or user custom data server:**

To load your own data for visualization, you need to prepare your own data on your local workstation or data server. UCSC genome browser accepts multiple file formats, including bedGraph, GTF, PSL, BED, bigBed, WIG, bigGenePred, bigWig, BAM, VCF, MAF, BED detail, Personal Genome SNP, broadPeak, narrowPeak, and microarray (BED15). Here is an example of visualizing transcriptome .bam file (for data generated using different toolkits, please refer to their instruction).

1. Alignment using tophat2 (https://ccb.jhu.edu/software/tophat/index.shtml):

a. Building genome index: bowtiw2-build genome.fa database\_index

b. Genome alignment: tophat2 -o output database\_index pairend1.fastq pairend2.fastq 2. Indexing alignment file: samtools index output/accepted.bam

To load the alignment file, click “My Data”, then “Custom Track”, select ”Choose File” and navigate to the folder/path that stores your data. Data can also provided via URL. If users have a data server/data repository, providing the URL to the data will directly pull the data to genome browser.

**---Import data from XGSC data server:**

*Xiphophorus* Genetic Stock Center shares data with the research community. Currently, we share a few tissue specific RNA-Seq data. These data are available for download at https://viewer.xgsc.txstate.edu/data/SampleData/UCSC. To load sample data directly, click "My Data"->"Custom Track", and paste the URL of the data of interest to the dialog box, the data will appear as a track in genome browser. We recommend you use this way to visualize data instead of downloading them to your workstation or data server, unless you have special tasks to perform with the data.

We plan to host *Xiphophorus* genetic data from the whole community in the future. However, due to hardware limitations, we can only host selected datasets to share with the community. If you are interested in sharing your data, especially RNA-Seq, ChIP-Seq, non-coding RNA data, please contact XGSC staff/faculty.

We are migrating to cloud based data processing and data sharing services. Upon server availability, we will provide bioinformatics toolkit and dedicated servers to carry out users' task.

**5. Data download**

*X. helleri* chromosome assembly genome, *X. couchianus* chromosome assembly, *X. maculatus* chromosome assembly and contig assembly (.fasta) sequences and genome annotation (.gtf) are all available in “viewer.xgsc.txstate.edu/data” for download. The genome alignment data is also available in the /data directory. The multiple genome alignment was performed using lastz aligner and multiz-tba package. You can perform genome alignments between other species and *X. maculatus* using your favorite pipeline or you request that we align your favorite genomes to *Xiphophorus*.