Xenopus genomic data and browser resources

Peter D. Vize[a] and **Aaron M. Zorn**[b]

[a]Departments of Biological Science and Computer Science, University of Calgary, Alberta, Canada

[b]Division of Developmental Biology, Cincinnati Children's Research Foundation, Cincinnati, OH, USA

Abstract

The two species of *Xenopus* most commonly used in biomedical research are the diploid *Xenopus (Silurana) tropicalis* and the tetraploid *Xenopus laevis*. The *X. tropicalis* genome sequence has been available since 2010 and this year the *X. laevis,* genome from two distinct genetic backgrounds has been published. Multiple genome assemblies available for both species and transcriptomic and epigenetic data sets are growing rapidly, all of which are available from a variety of web resources. This review describes the contents of these resources, how to locate and download genomic data, and also how to view and manipulate these data on various public genome browsers, with an emphasis on Xenbase, the *Xenopus* model organism database.

Introduction

Genomic data is central to all modern experimental design and interpretation. Access to these data, and knowing how to use them efficiently, is critical to biomedical research, from designing morpholinos and gene-editing CRISPRs to analyzing gene expression with RNA-seq and ChIP-seq experiments. There are an enormous variety of tools available to access and interrogate genomic data, and no one site can host and maintain all of them. The purpose of this brief review is to guide users to the various resources supporting *Xenopus* genomic data, from where to download the genomes themselves, to tools for visualizing and analyzing genomes, and how to perform visualization on your own desktop computer. Not all available sites are covered and the emphasis is on the most recent *Xenopus laevis* genome resources as these are the newest and less well known. Sites hosting servers providing major *Xenopus* genomics resources and their URLs are listed in Table 1.

DNA sources used in generating *Xenopus* genome sequences

The large scale *Xenopus* sequencing projects by the Joint Genome Institute (JGI), succeeded by the *Xenopus* Genome Project Consortium (XGPC), used inbred frogs to minimize

Correspondence to: Peter D. Vize.

polymorphisms in both the *X. laevis* and *X. tropicalis* genome projects, and in each case a single female was used as the source of both DNA for the genome and RNA for supporting cDNA/RNA-seq data. Female DNA sources were chosen as in *Xenopus* this is the heterogametic sex. The XGPC *X. laevis* project used a deeply inbred strain called J-strain, originally developed in Japan for immunological research (see Izutsu and Maeno. 2005). The XGPC project has released a number of draft builds over the past four years and the most recent, version 9.1, has the vast majority of the genome assembled into chromosome scale scaffolds (Session et al. submitted). Table 2 lists the major *X. laevis* genome builds available on various genome browsers hosted by a variety of sites.

The DNA used by the JGI/XGCP for the *X. tropicalis* genome builds followed a similar strategy; a single female from an inbred line (Hellsten, et al. 2010). In this case the frog selected was from a seventh generation inbreed Nigerian line generated by the Grainger laboratory (Grainger, 2012). A number of more recent assemblies of this *X. tropicalis* genome have been produced by the XGPC by incorporating new genetic map data and Dovetail reads to improve long range linkage to the chromosome level (A. Session, pers.comm.).

Finally the genome sequence of an outbred wild-type *X. laevis* female with an independent genome assembly from the Michalak group (Virginia Tech.) was publicly released in 2015 (Table 2). This animal was obtained from one of the common suppliers of *Xenopus* frogs for research, *Xenopus* Express.

Data availability

The genomic data for the various genome releases are available from a number of public web sites (Table 1). All genome browser resources also provide links to the source data, so users can get content from these or from the genome data repository at Xenbase. Xenbase (Bowes, et al. 2008; Karpinka et al., 2015; James-Zorn et al., 2015) stores all available genome builds and their supporting files, so is the most comprehensive. Typically the scaffold sequences themselves are in FASTA format, as are the set of transcripts and proteins predicted from the scaffold data. The gene models that show exon structure, translation start and end points etc. are usually provided in a GFF3 file, but some genome browsers, e.g. the UCSC browser (Karolchik, et al. 2003), use the alternate BED file format for this data. There is also an annotation file that associates gene models with the official gene symbols, which are linked to stable gene IDs from Xenbase.

The FASTA, GFF3, BED and annotation files are available from the Xenbase ftp repository for all of the major genome builds of both species (http://www.xenbase.org/other/static/ftpDatafiles.jsp). For more recent builds Xenbase offers both GFF3 and BED files for gene models. Other sites also have subsets of these data including the JGI, the Crick Institute, ENSEMBL NCBI, UCSC, University of Texas XenMIne and the *Xenopus laevis* Genome Project (Japan). Moreover Xenbase provides pre-publication versions of the genome assemblies, which are typically available several years before they are supported by more general resources such as Genbank, NCBI Genomes, UCSC and Ensembl. Table 2 shows the genome assemblies and sources data for both *Xenopus* species that are available at these

different resources. All of these sites host genome browsers where these data can be visualized and the various browsers and genome versions are also available in Table 2.

Using *Xenopus* genome data

There are many different approaches to finding a gene or chromosomal region of interest in one of the *Xenopus* genomes. The search functions differ in each browser. In GBrowse one can go via the various Xenbase search tools, and links to the region of interest on all available genome builds will be present on individual gene pages. Xenbase GBrowse also has its own search tool towards the upper left of the screen window enabling genome navigation in the browser environment. Either a chromosomal position or a gene model annotation (a gene model number or a gene symbol) can be entered into this box and searched. This tool allows wildcards, so if uncertain of a gene annotation enter a symbol with a wildcard, e.g. if a user searched for 'pax*', all genes with 'pax' in their annotations will be returned on a disambiguation page. GBrowse does not have an AJAX style autocomplete suggestions feature that presents options as you type, so wildcards are very useful.

The other, and often more useful, approach is to use BLAST. Go to the Xenbase BLAST page, select the genome version, paste in a block of sequence and run. The hits will have hyperlinks to the matched region in GBrowse. This method avoids any incomplete or inaccurate annotations and is very fast, with searches only taking a second or two. This option allows you to take many different paths to a snapshot in a genome browser. For example you could search for genes expressed in the heart, select a hyperlink to go to a gene page, then select the link to a genome in GBrowse and view the ChIP-seq data associated with the heart gene's promoter.

Other sites have alternative BLAST and/or BLAT options. All sites hosting the UCSC browser utilize its integrated BLAT option, including the Crick Institute, UCSC itself, and UCSC at Xenbase. The *Xenopus laevis* Genome Project (Japan) offers both BLAST and BLAT gateways to the *X. laevis* genome. BLAST results can also be displayed in JBrowse, but to our knowledge no *Xenopus* resource has this feature available yet.

The latest XGPC genomes (*X. tropicalis v9.0* and *X. laevis v9.1*) have much higher numbers of gene models with more complete and higher quality annotations than the outbred *X. laevis* genome (Table 3), but this later genome provides an extremely valuable comparative resource for genome-scale nucleotide polymorphisms. DNA polymorphisms within different *Xenopus* lines have been known for some time to result in suboptimal morpholino design, and cautious researchers PCR amplify and sequence the target regions of interest from their own local *Xenopus* stocks before designing morpholinos or CRISPR guide RNAs for genome editing. While this is still recommended, if it is not feasible a sensible alternative would be to identify the gene model of interest from the inbred J-strain v9.1 assembly and BLAST the target region of interest against the outbred *X. laevis* genome in Xenbase. If the target area shows no sequence differences, then this region is likely to be a good candidate for morpholino or CRISPR design. Alternatively, many investigators obtain the inbred Nigerian *X. tropicalis* or J-strain *X. laevis* from the National Xenopus stock centers in the

US (NXR), UK (EXRC) or Japan (NBRP) to ensure that their colony has the same DNA sequence as the published genomes.

Genome metrics

Table 3 shows metrics for the most recent *X. laevis* and *X. tropicalis* genome releases. The allotetraploid genome of *X.laevis* is twice the size of the *X. tropicalis* genome, at 2.8 GB versus 1.4 GB. For comparative purposes, the human genome is 3.2 GB. The number of primary gene models is 1.7 fold greater in *X. laevis*. For a detailed comparison of the *X. laevis* and *X. tropicalis* genomes we refer readers to Session et al., (submitted). The genome metrics for assembly are also summarized in the 'readme' information found with each genome at the Xenbase data repository.

Browsing *Xenopus* genomes in Xenbase

Multiple genome browsers support *Xenopus* data, and these plus the genomes supported are listed in Table 2. Xenbase, which supports GBrowse (Stein et al. 2002) has the most extensive *Xenopus* genome support to date and will be the focus of this section. GBrowse, like other genome bowsers, loads a default set of tracks on launch that is set by the host site (Figure 1). The view in GBrowse is highly customizable. Users can move along the scaffold by clicking on a track and dragging to the left or right. Different tracks can be reordered by clicking on a track header and dragging it up or down. At any time the sequence currently being viewed can be downloaded by clicking on the small floppy disc icon on the track header (FASTA or GFF3). Users can also load their own tracks either for private viewing or sharing with the community using the "Custom Tracks" tab.

If the view in GBrowse is not what was expected, this may be due to GBrowse storing recent views in cache, and may have loaded your most recent work window rather than defaults. Should you ever get stuck in GBrowse, flushing the cache is a good first problem solving step, and this can be done using the Preferences tab by unchecking the "Cache tracks" button.

There is a large amount of public Next Generation Sequencing (NGS) data supporting *Xenopus* genomes (e.g. Bogdanovic, et al. 2012; van Heeringen, et al. 2014), including RNA-seq, RNA polymerase II and p300 binding ChIP-seq, and over 150 histone modification tracks, e.g. see http://gbrowse.xenbase.org/fgb2/gbrowse/xt7_1/? Unfortunately these data are at the moment mostly only mapped against earlier genome assemblies used by the submitting group. Xenbase is presently working towards remapping these data and other public data from the Short Read Archive (SRA) to the most recent genome builds to making browsing and locating up to date content easier. Xenbase also supports NGS tracks on a UCSC browser track hub (discussed below), and links to XenMine (Table 1) and via other custom data set specific resources such as Tan et al. (2013), Collart et al. (2014), Owens et al. (2015) and Peshkin et al. (2015), where FPKM and custom expression data are available. To view the different tracks including RNA-seq and epigenetic histone marks from various developmental stages to the Xenbase Gbrowse page and select the tab marked "Select Tracks" (Figure 1).

Some of the more recent genome releases are also available on JBrowse (Westesson, et al. 2013). Both XenMine and UTexas support the recent *X. laevis* genome in JBrowse, and Xenbase is also currently adding this browser. JBrowse is a modern browser that works in a similar manner to Google maps, so users will find the interface intuitive. It is also very fast and runs largely on the users' desktop computer, which greatly decreases the lag time experienced in GBrowse as that browser is constantly send large volumes of data between the desktop computer and the server.

For users who prefer the UCSC browser (Karolchik, et al. 2003) a number of options are available The main UCSC site itself hosts *Xenopus*, but only older versions of the *X. tropicalis* genome v4.2 and v7.1 are currently available. However, two other sites run custom UCSC instances with more recent genomes for both *X. laevis* and *X. tropicalis*. These are the Crick Institute hosted by the Gilchrist lab and Xenbase (Table 1). These instances host genome, track and NGS data, but do not have all of the content users are used to on the UCSC main site- such as other genomes for comparative purposes, or all of the support data available on Xenbase. Users in Europe may find the Crick Institute browser provides better performance, while those in North America will likely find the Xenbase instance faster.

A valid question to ask is why are more recent genome builds not available on popular browsers such as UCSC, Ensembl and NCBI? Each of these organizations has its own data processing pipelines and data quality metrics. Only when all of the genome data and metadata passes these quality controls and the genomes are official submitted to Genbank do their systems process the genomes. These conditions were not met for some intermediate genome assemblies. However with the recent genome submission of *X. laevis* 9.1 to the NCBI this situation should be amended and hopefully these excellent resources will soon also host *Xenopus* builds along with their current content.

Data hubs

Data hubs allow you to load custom genomic and NGS data onto the UCSC browser then use the massive computing resources at that site, even when UCSC does not contain the content you wish to view- such as the latest *Xenopus* genomes. The hub can be anywhere that is web accessible via http or ftp. UCSC's main site is directed to the hub and it loads the remote tracks. While this sounds like it would be a slow process with genome scale data, the system only loads the portion of the genome being viewed so it works very quickly. For example, loading the *X. tropicalis* genome plus NGS tracks from the Xenbase hub hosted in Canada takes UCSC only about 5 seconds. Once the hub has loaded, opening a different genome and its associated NGS tracks also only takes another 5 or so seconds.

There are two track hubs currently hosting *Xenopus* data, the Radboud University site hosted by G. Veenstra and Xenbase. Both work in the same way and at the moment have similar content, but these will diverge over time. Xenbase plans to host both bigwig files for viewing wiggle type data and also bam files. Bam files contain raw read alignments so are quite large, usually about a GB. As discussed above they load via a hub quite quickly but the display is a little more cumbersome than is the bigwig version. Bam files are used for bioinformatic analysis, while bigwig are not, which is why both are hosted.

To access the Xenbase hub go to http://www.xenbase.org and under the top navigation banner select Genomes>UCSC Track Hub while at Radboud University go to http://veenstra.ncmls.nl/trackhub.htm and follow the simple instructions (Table 1). These will provide a URL that you copy, and a link to take you to the USCS browser. Once on UCSC you click on the "track hubs" button then select the "My Hubs" tab. Paste the provided URL into the window and click on "Add Hub". Once the hub loads a default set of tracks will be displayed, select the desired tracks and reload to view the data you require. Available tracks are listed directly below the scaffold visualization window in the UCSC browser.

Over 150 NGS tracks are currently available on the Radboud and Xenbase hubs, mostly for *X. tropicalis*, but there are some RNA-seq and ChIP-seq data for *X. laevis* on Xenbase. Hundreds of new tracks are currently being processed from public data on NCBI's SRA and these will be available for both species on the Xenbase hub soon.

Build your own tracks: run your own browser

Building your own tracks and running a browser locally on your desktop is a good option if you want maximal control, for example you want to align a result against a different genome build, or privately analyze your own NGS data prior to publication, or find track upload times to GBrowse are too slow, and it is not complicated. We use the Galaxy workflow manager (Giardine, et al. 2005) to perform this task. Once you have set up a workflow in Galaxy you can use it over and over again very easily. There are multiple Galaxy resources around the world and for details on how to implement this approach we direct readers to www.usegalaxy.org. Should none of your available Galaxy resources work at a reasonable speed (this does happen) contact Xenbase and we can give you temporary 7 day access to our server. This time limit is necessary as Galaxy bloats rapidly and hard drive space needs to be cleared regularly.

The NGS data to be viewed, in FASTQ format, and gzip compressed, is uploaded via the Upload File tool in the left tool column near the top of the Galaxy interface. Set the "Type" dropdown tool to FASTQ- as using autodetect does not always work. Once upload is complete Galaxy will automatically decompress the file. Then also upload your target genome in FASTA format, again gzip compressed. Next, scroll down the tool list to the NGS Toolbox section. Select "FASTQ Parallel Groomer". In the drop downs select your file then click on "Execute". Once the Groomer has processed your FASTQ file you now have all you need to align the file to the genome. To perform the alignment, select Bowtie2 from the same NGS Toolkit section on the left. Select the NGS data in the FASTQ options box, and under "Will you select a reference genome...?" set to the second option "Use a genome from the history" then find your genome FASTA. Select "Execute", and once the run is complete (1–2 hours) you have your new track. Make sure you download both the bam file and the bam index file (available via the small floppy disc icon) as you will need both. There are many workshops held to train users in more advanced features of Galaxy, and the range of tools available is enormous. We recommend the bioinformatics workshop hosted by the NXR, the *Xenopus* resource center at the MBL in Woods Hole (http://www.mbl.edu/xenopus/).

To view your new track you can upload it to a remote server such as GBrowse or a track hub such as UCSC, or you can run a desktop genome browser. Using a hub to load onto a remote browser is described above. In this section we briefly cover using the Integrated Genomics Viewer (IGV) (Thorvaldsdóttir et al., 2013) on your desktop. The IGV produced by the Broad Institute offers a very easy to use desktop browser for Macintosh and Unix variants. Windows users can also use this software via Java, and instructions on how to install and configure Java on windows is provided at the IGV website (Thorvaldsdóttir et al., 2013). For Macintosh users it is as simple as downloading and launching the installer.

To load a genome and the bam file generated by Galaxy follow these steps in the IGV:

Genomes>Load Genome from File

- Locate and load the genome FASTA

File>Load from File...

- Locate and load the gene model GFF3 file. If you want to, also load a bam file with NGS read data

File>Load from File...

- Locate and load both the bam and bam index files. The will need the same name but different extensions, e.g. stage12_RNA.bam and stage12_RNA.bai. Screenshots of the IGV browser showing RNA-seq aligned to *X. laevis* build 9.1 using this workflow are shown in figure X.

Genome nomenclature

Xenopus gene names are based on human gene names and are administered by the *Xenopus* Gene Nomenclature Committee. The current Xenopus gene nomenclature guidelines can be found at http://www.xenbase.org/gene/static/geneNomenclature.jsp. Xenbase serves as the clearing house for *Xenopus* gene names and manages both individual and genome scale nomenclature issues. Many issues are detected during the annotation of papers describing work performed using Xenopus and others are submitted directly by the research community. In response to a nomenclature issue Xenbase staff perform an analysis of the gene in question using BLAST, synteny and phylogeny. They then liaise with the Human Gene Nomenclature Committee if necessary to ensure compliance and compatibility before promoting a new gene symbol or name. Wherever possible *Xenopus* names and symbols are exactly the same as those for their human gene orthologs.

The allotetraploid nature of the *X. laevis* genome required changes to gene symbols for this species. In the past (now obsolete) the two alloalleles (also referred to as homeologs) in this species were distinguished by appending '-a' or '-b' to gene symbols. As there was no way to know what chromosome each gene was derived from, the gene-a and gene-b assignments were essentially random. Xenbase used the arbitrary assignments of Hellsten et al. (2007) to determine -a and -b tags in older releases. With the chromosome level builds of *X. laevis* in version 9.1 having over 91% of all genes assigned to a specific chromosome, a systematic assignment of gene symbols was made possible. This genome contains two individual sets of chromosomes from its two ancestral species, each of which has unique characteristics, such

as transposon sequences, and sizes. With this information, the *Xenopus* Genome Project Consortium in collaboration with the *Xenopus* Gene Nomenclature Committee chose to change the alloallele tags to a .L and a .S, with the "L" indicating the long chromosome (from ancestral species 1), and the "S" indicating the short (from ancestral species 2) (Matsuda et al 2015). Both *X. laevis* alloalleles, for example hhex.L and hhex.S have their own unique identifiers in Xenbase, and are also linked to a top level name and identifier, in this example hhex, for linking to the *X. tropicalis* gene and to external resources and other model organisms. How Xenbase deals with diploid and tetraploid data is discussed elsewhere (Vize et al., 2015).

Conclusions

A variety of sites host the raw genomic data and visualization tools for viewing and downloading *Xenopus* data. Multiple genome assembly versions and accompanying NGS content is available on multiple browsers and track hubs allow loading of data to browsers that do not already host them. While ongoing efforts to improve the *Xenopus* genomes will result in improved scaffolds and better annotation will improve the utility of these resources there is still no coordinated effort to fully curate these genomes. Incorrect annotations and errors in machine predicted gene models reduce the value of these extraordinary resources. While additional computer generated models and annotations will improve this situation, this will still fall short of what is really required- a manual annotation project similar to those that have been performed for all other major model organism genomes.

Acknowledgments

References

Bogdanovic O, Fernandez-Minan A, Tena JJ, de la Calle-Mustienes E, Hidalgo C, van Kruysbergen I, van Heeringen SJ, Veenstra GJ, Gomez-Skarmeta JL. Dynamics of Enhancer Chromatin Signatures Mark the Transition from Pluripotency to Cell Specification during Embryogenesis. Genome Res. 2012; 22:2043–2053. [PubMed: 22593555]

Bowes JB, Snyder KA, Segerdell E, Gibb R, Jarabek C, Noumen E, Pollet N, Vize PD. Xenbase: A Xenopus Biology and Genomics Resource. Nucleic Acids Res. 2008; 36:D761–776. [PubMed: 17984085]

Collart C, Owens ND, Bhaw-Rosun L, Cooper B, De Domenico E, Patrushev I, Sesay AK, Smith JN, Smith JC, Gilchrist MJ. High-resolution analysis of gene activity during the Xenopus mid-blastula transition. Development. 2014; 141:1927–1939. [PubMed: 24757007]

Giardine B, Riemer C, Hardison RC, Burhans R, Elnitski L, Shah P, Zhang Y, Blankenberg D, Albert I, Taylor J, Miller W, Kent WJ, Nekrutenko A. Galaxy: A Platform for Interactive Large-Scale Genome Analysis. Genome Res. 2005; 15:1451–1455. [PubMed: 16169926]

Grainger RM. Xenopus Tropicalis as a Model Organism for Genetics and Genomics: Past, Present, and Future. Methods Mol Biol. 2012; 917:3–15. [PubMed: 22956079]

Hellsten U, Khokha MK, Grammer TC, Harland RM, Richardson P, Rokhsar DS. Accelerated gene evolution and subfractionalization in the pseudotetraploid frog Xenopus laevis. BMC Biol. 2007; 5:31. [PubMed: 17651506]

Hellsten U, Harland RM, Gilchrist MJ, Hendrix D, Jurka J, Kapitonov V, Ovcharenko I, Putnam NH, Shu S, Taher L, Blitz IL, Blumberg B, Dichmann DS, Dubchak I, Amaya E, Detter JC, Fletcher R, Gerhard DS, Goodstein D, Graves T, Grigoriev IV, Grimwood J, Kawashima T, Lindquist E, Lucas SM, Mead PE, Mitros T, Ogino H, Ohta Y, Poliakov AV, Pollet N, Robert J, Salamov A, Sater AK, Schmutz J, Terry A, Vize PD, Warren WC, Wells D, Wills A, Wilson RK, Zimmerman LB, Zorn AM, Grainger R, Grammer T, Khokha MK, Richardson PM, Rokhsar DS. The Genome of the Western Clawed Frog Xenopus Tropicalis. Science. 2010; 328:633–636. [PubMed: 20431018]

Izutsu Y, Maeno M. Analyses of Immune Responses to Ontogeny-Specific Antigens using an Inbred Strain of Xenopus Laevis (J Strain). Methods Mol Med. 2005; 105:149–158. [PubMed: 15492394]

James-Zorn C, Ponferrada VG, Burns KA, Fortriede JD, Lotay VS, Liu Y, Karpinka JB, Karimi K, Zorn AM, Vize PD. Xenbase: Core features, data acquisition, and data processing. Genesis. 2015; 53:486–497. [PubMed: 26150211]

Karolchik D, Baertsch R, Diekhans M, Furey TS, Hinrichs A, Lu YT, Roskin KM, Schwartz M, Sugnet CW, Thomas DJ, Weber RJ, Haussler D, Kent WJ. University of California Santa Cruz. The UCSC Genome Browser Database. Nucleic Acids Res. 2003; 31:51–54. [PubMed: 12519945]

Karpinka JB, Fortriede JD, Burns KA, James-Zorn C, Ponferrada VG, Lee J, Karimi K, Zorn AM, Vize PD. Xenbase, the Xenopus model organism database; new virtualized system, data types and genomes. Nucleic Acids Res. 2015; 43:D756–763. [PubMed: 25313157]

Matsuda Y, Uno Y, Kondo M, Gilchrist MJ, Zorn AM, Rokhsar DS, Schmid M, Taira M. A new nomenclature of *Xenopus laevis* chromosomes based on the phylogenetic relationship to *Silurana/ Xenopus tropicalis*. Cytogenet Genome Res. 2015; 145:187–191. [PubMed: 25871511]

Owens ND, Blitz IL, Lane MA, Patrushev I, Overton JD, Gilchrist MJ, Cho KW, Khokha MK. Measuring absolute RNA copy numbers at high temporal resolution reveals transcriptome kinetics in development. Cell Rep. 2016; 14:632–647. [PubMed: 26774488]

Peshkin L, Wühr M, Pearl E, Haas W, Freeman RM, Gerhart JC, Klein AM, Horb M, Gygi SP, Kirschner MW. On the relationship of protein and mRNA dynamics in vertebrate embryonic development. Dev Cell. 2015; 35:383–394. [PubMed: 26555057]

Stein LD, Mungall C, Shu S, Caudy M, Mangone M, Day A, Nickerson E, Stajich JE, Harris TW, Arva A, Lewis S. The Generic Genome Browser: A Building Block for a Model Organism System Database. Genome Res. 2002; 12:1599–1610. [PubMed: 12368253]

Tan MH, Au KF, Yablonovitch AL, Wills AE, Chuang J, Baker JC, Wong WH, Li JB. RNA sequencing reveals a diverse and dynamic repertoire of the Xenopus tropicalis transcriptome over development. Genome Res. 2013; 23:201–216. [PubMed: 22960373]

Thorvaldsdóttir H, Robinson JT, Mesirov JP. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. Brief Bioinform. 2013; 14:178–192. [PubMed: 22517427]

van Heeringen SJ, Akkers RC, van Kruijsbergen I, Arif MA, Hanssen LL, Sharifi N, Veenstra GJ. Principles of Nucleation of H3K27 Methylation during Embryonic Development. Genome Res. 2014; 24:401–410. [PubMed: 24336765]

Vize PD, Liu Y, Karimi K. Database and informati challenges in representing both diploid and tetraploid Xenopus species in Xenbase. Cytogenet. Genome Res. 2015; 145:278–282.

Westesson O, Skinner M, Holmes I. Visualizing Next-Generation Sequencing Data with JBrowse. Brief Bioinform. 2013; 14:172–177. [PubMed: 22411711]

Highlights

- Multiple genome browsers support both X. laevis and X. tropicalis genomes

- Chip-seq and RNA-seq support for both species is available

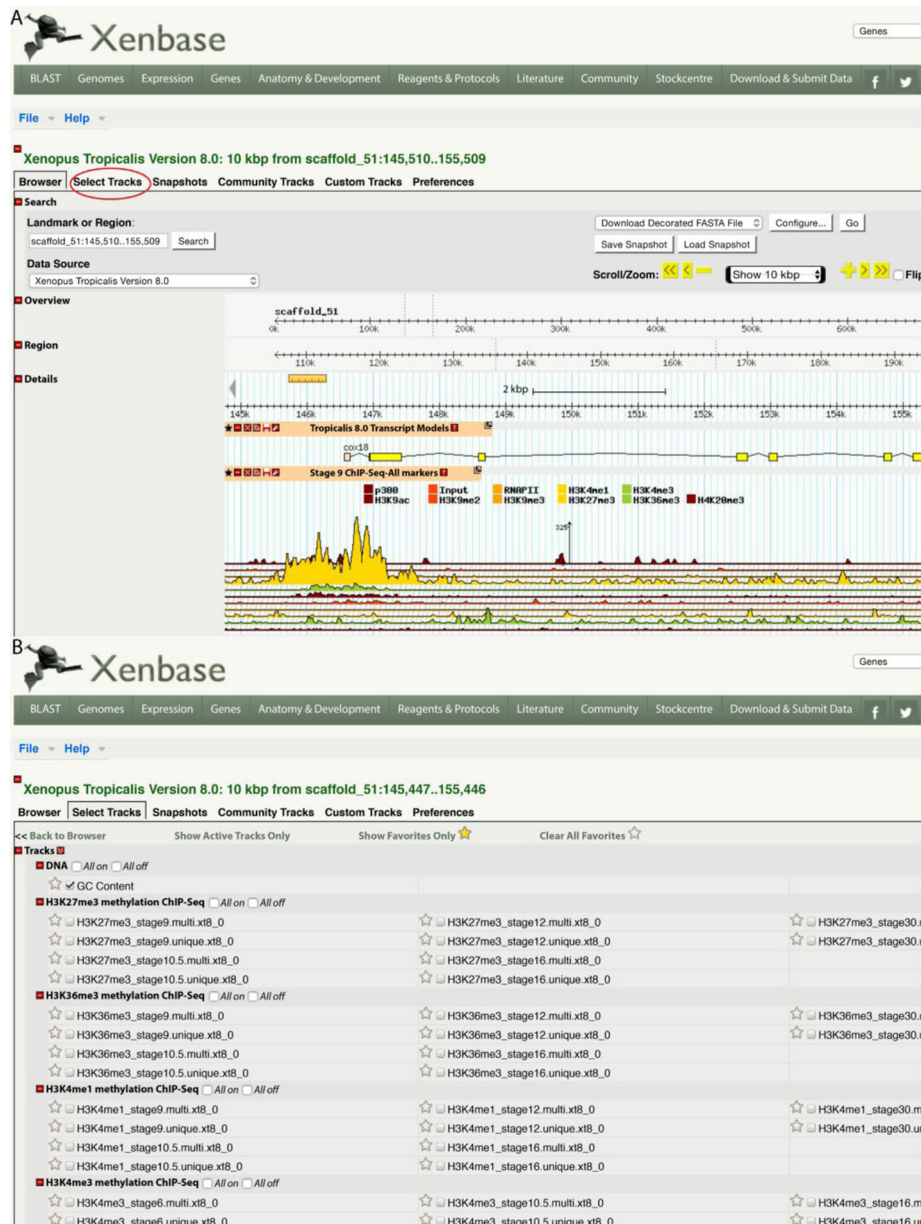- Instructions for using resources and for running desktop browsers is provided

Figure 1.

Xenbase GBrowse with NGS data. A. The default view of the cox18 gene in Xenbase, with multiple NGS tracks displayed in a stacked format using a tool called Topoview, developed by FlyBase (http://flybase.org/static_pages/docs/software/topoview.html). Note the large histone H3 lysine 4 methylation (H3K4me1) peak at stage 9 over the transcription start site for this gene. To view additional/alternative tracks, select the "Select Tracks" tab circled in red. B. The Select Tracks view offers an enormous number of supporting NGS data including histone methylation and acetylation, p300 and RNA polymerase binding sites and RNA-seq data. As noted in the text, the content currently varies with the genome build, so

exploring other genome versions is worthwhile. To go back to the browser view click on the "Browser" tab on the top left or the "Back to Browser" button at the bottom of the page.

Table 1

Major web resource sites for *Xenopus* genomic data.

Institute	Tool	URL
Broad Institute	Integrative Genomics Viewer	http://www.broadinstitute.org/software/igv/
Crick Institute	UCSC browser	http://genomes.crick.ac.uk
ENSEMBL	Ensembl browser	http://www.ensembl.org/Xenopus_tropicalis/
Joint Genome Institute	JGI browser	http://genome.jgi.doe.gov/Xentr4/Xentr4.info.html
NCBI	NCBI browser	http://www.ncbi.nlm.nih.gov/genome/?term=xenopus+tropicalis
Radboud University	UCSC track hub	http://veenstra.ncmls.nl/trackhub.htm
UCSC	UCSC browser	https://genome.ucsc.edu/
University of Texas	JBrowse	http://daudin.icmb.utexas.edu/XENLA_JGIv91/
Xenbase	GBrowse, JBrowse, Genomes Repository, UCSC track hub	http://www.xenbase.org
XenMine	InterMine for Xenopus, JBrowse	http://www.xenmine.org
Xenopus laevis Genome Project (Japan)	GBrowse	http://xenopus.lab.nig.ac.jp

Table 2

X. laevis (XL) genome build versions and most recent *X. tropicalis* (XT) version available at various sites.

	GBrowse XL	JBrowse XL	UCSC XL	UCSC hub XL	Ensembl XL	latest tropicalis (XT)
Crick Inst.			9.1			9
NCBI						7
UCSC/Ens						4.2
UTexas		9.1				9
Xenbase	6, 7.2, 9.1, O1 *	pending	7.2, 9.1	7.2, 9.1		9
Xenmine		7.1				8.3
Xenopus laevis Genome Project	Soap 2					9

*
O indicates outbred genome

Table 3

Metrics for most recent genome builds for both inbred (XGPC) and outbred *X. laevis* lines.

	XGPC X. laevis v9.1	VT X. laevis v1.0	XGPC X. tropicalis v9
Scaffolds	402K	872K	6823
Scaffold N50	136M	6300	135M
total length	2.765G	2.709G	1.441G
ungapped length	2.45G	2.705G	1.37G
primary gene models	45099	29454	26550