

The *Saccharomyces* Genome Database: Exploring Genome Features and Their Annotations

J. Michael Cherry¹

Department of Genetics, Stanford University School of Medicine, Stanford, California 94305-5120

Genomic-scale assays result in data that provide information over the entire genome. Such base pair resolution data cannot be summarized easily except via a graphical viewer. A genome browser is a tool that displays genomic data and experimental results as horizontal tracks. Genome browsers allow searches for a chromosomal coordinate or a feature, such as a gene name, but they do not allow searches by function or upstream binding site. Entry into a genome browser requires that you identify the gene name or chromosomal coordinates for a region of interest. A track provides a representation for genomic results and is displayed as a row of data shown as line segments to indicate regions of the chromosome with a feature. Another type of track presents a graph or wiggle plot that indicates the processed signal intensity computed for a particular experiment or set of experiments. Wiggle plots are typical for genomic assays such as the various next-generation sequencing methods (e.g., chromatin immunoprecipitation [ChIP]-seq or RNA-seq), where it represents a peak of DNA binding, histone modification, or the mapping of an RNA sequence. Here we explore the browser that has been built into the *Saccharomyces* Genome Database (SGD).



MATERIALS

Equipment

Internet-connected computer with web browser

METHOD

At any step in this protocol, support is available via the comprehensive Help documents maintained by SGD. These can be accessed via the help button at the top of each page. The drop-down Help menu at the top of the gbrowse page leads you directly to extensive help and to video tutorials created by SGD for the use of gbrowse.

The SGD is continually updated; therefore, specific items presented in this protocol may not appear on the SGD website exactly as described.

1. Open the URL, www.yeastgenome.org, in any modern web browser (e.g., Firefox, Chrome, or Safari).
2. Enter “PIG1” in the Search box and press return to go directly to the *PIG1* Locus Summary page. In the Sequence section, click on “View in: GBrowse”. This opens the graphical browser page (<http://browse.yeastgenome.org>).

¹Correspondence: cherry@stanford.edu

On the left side of the browser page, above the graphic is a search box labeled “Landmark or Region”; it should currently say “chrXII:689,083..691,029,” which are the genomic coordinates for the *PIG1* gene (Fig. 1). The yeast chromosomes are labeled with roman numerals, and thus *PIG1* is on chromosome 12 from base 689,083 to 691,029. The *PIG1* open-reading frame (ORF) is on the reverse strand, going right to left, but that is not apparent from the coordinates. The ORF name syntax informs you of this orientation. ORFs ending in *W* for Watson are on the left to right (forward) strand and those ending in *C* for Crick are on the other (reverse) strand. Above the red arrow labeled “*PIG1*, Verified, Putative targeting subunit for . . .” is the *PIG1* ORF name, YLR273C, indicating *PIG1* is on the reverse strand. In addition, the red arrow points to the left and indicates ORF orientation. More information about the naming conventions of yeast chromosomal features can be found at SGD, <http://www.yeastgenome.org/help/community/nomenclature-conventions>. The *gbrowse* view provides a convenient way to see which genes are near each other, including tRNAs and transposon long terminal repeats (LTRs). The feature downstream from *PIG1*, labeled *tV(AAC)L*, is the tRNA for valine (*V*) with the anticodon AAC; the *L* indicates its location on chromosome 12, *L* being the 12th letter of the alphabet. There are two LTRs associated with *Ty1* downstream from *PIG1*. These are labeled YLRCdelta16 (delta element number 16 on chromosome 12, right of the centromere on the complement or Crick strand) and YLRWdelta17 (delta element number 17 on chromosome 12, right of the centromere on the Watson strand).

3. Locate the scrolling and zooming buttons at the top of the page and navigate along the chromosome.

The “>” moves the view a little to the right and the “>>” moves the view a lot to the right. The yellow “+” zooms in and the “-” zooms out. A drop down is provided to quick zoom in or out to a specific window size for the chromosome. The flip button changes the orientation of the view.

4. There are two metered regions on the upper two chromosome tracks that are highlighted in light purple. Drag the solid purple region to scroll from side to side. Click and drag in the shaded region to define a different-sized region, or a new region. These movements only work in the metered region.

There are several hundred tracks available for display, but it is best to only show a few at a time. By default the “All Annotated Sequence Features” track will be shown; however, *gbrowse* remembers the last set of tracks viewed.

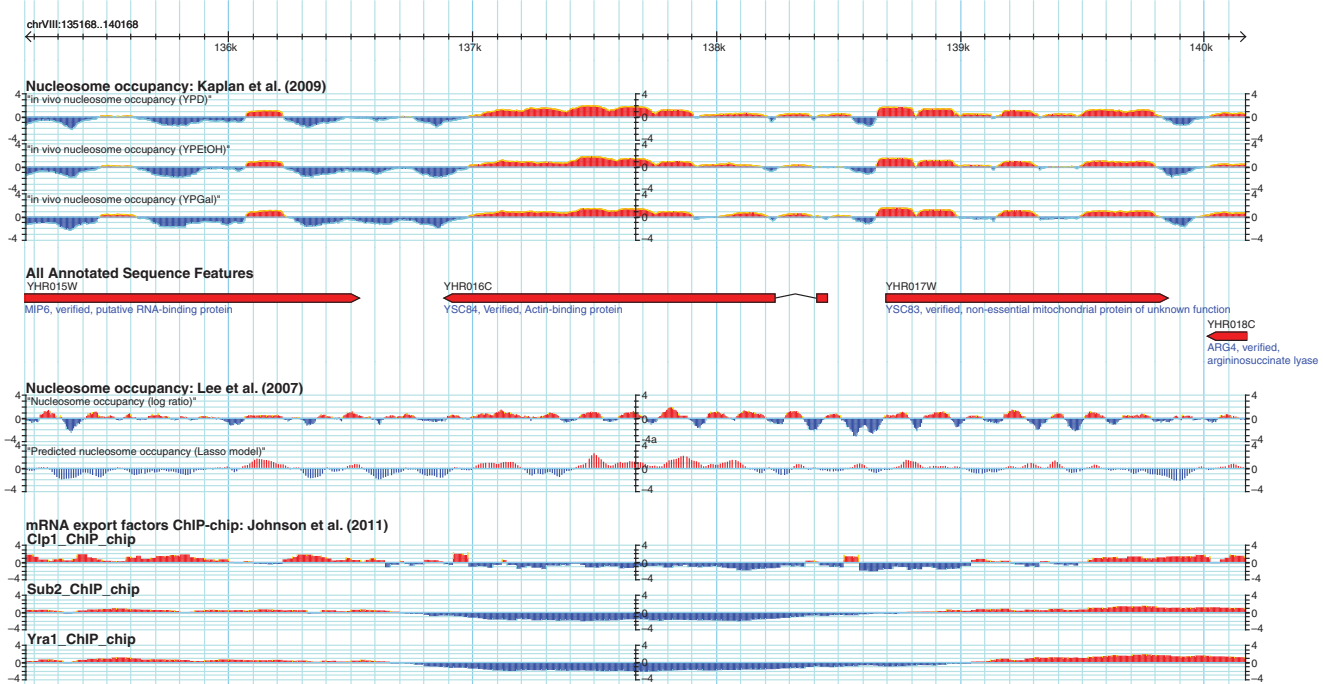


FIGURE 1. View of the genome around the *PIG1* locus. This display is created by the GBrowse software and shows horizontal tracks providing chromosomal feature annotations and results of genomic experiments. The top three tracks show the experimental results for the nucleosome occupancy from Kaplan et al. (2009). The first two tracks below the feature annotations provide nucleosome occupancy results from Lee et al. (2007), and at the bottom tracks show ChIP-chip results from Johnson et al. (2011) for Clp1, Sub2, and Yra1.

5. Scroll to the bottom of the page and click on the “Select Tracks” button. A large list of available tracks from published genomic studies will appear.
6. Click on the “All on” button next to Histone modifications, then the “Back to Browser” button located in the upper left.

You will see the results from five publications. All of these publications have multiple tracks that correspond to multiple data sets.
7. Mouse over the little buttons to the right of the star to learn about their functions for modifying a track. To move a track up or down in the view click on the publication authors name and drag the group of tracks.
8. Return to the “Select Tracks” page. Click on the “All off” button next to Histone modification. To turn on tracks from a specific study mouse over the citation and click when it turns dark. A check mark will appear to its left to indicate that track is now on.
9. Select a subset of tracks from a study by clicking on the blue text that includes “subtracks selected.”
10. Click on a star to denote the track as a favorite (clicking on a star does not turn the track on). View all favorites by selecting “Show Favorites Only” at the top of the “Select Tracks” page.
11. Return to the Browser page and click on an ORF. This takes you to the Locus Summary page for that gene.
12. Click on a track. This takes you to the paper for that study.
13. Click on Preferences. This allows you to control the image created by gbrowse.

This can be useful for creating a screenshot for a presentation.
14. Pull down the File menu at the top of the gbrowse page, just below the horizontal SGD option bar.

The File menu is used to export the image in PNG or SVG graphic format, or as GFF (genome feature format) or FASTA (sequence) files.

ACKNOWLEDGMENTS

I am grateful to all the present and past staff of the *Saccharomyces* Genome Database project for their dedication to accuracy and service to life science educators and researchers. I also want to thank the yeast research community for their support and suggestions. This work was supported by the National Human Genome Research Institute (grant number U41 HG001315), and Funding for open access charge was provided by the National Institutes of Health. The content is solely the responsibility of the author and does not necessarily represent the official views of the National Human Genome Research Institute or the National Institutes of Health.

REFERENCES

- | | |
|---|---|
| <p>Kaplan N, Moore IK, Fondufe-Mittendorf Y, Gossett AJ, Tillo D, Field Y, LeProust EM, Hughes TR, Lieb JD, Widom J, et al. 2009. The DNA-encoded nucleosome organization of a eukaryotic genome. <i>Nature</i> 458: 362–366.</p> <p>Lee W, Tillo D, Bray N, Morse RH, Davis RW, Hughes TR, Nislow C. 2007. A high-resolution atlas of nucleosome occupancy in yeast. <i>Nat Genet</i> 39: 1235–1244.</p> | <p>Johnson SA, Kim H, Erickson B, Bentley DL. 2011. The export factor Yra1 modulates mRNA 3' end processing. <i>Nat Struct Mol Biol</i> 18: 1164–1171.</p> |
|---|---|

