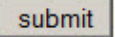# UCSC Genome Browser Introductory Workshop
# UC Davis
# Davis, CA

We will be using human assembly hg19. These problems will take you through a variety of resources at the UCSC Genome Browser. You will learn how to use the Table Browser, Custom Tracks, Saved Sessions and some navigation processes. Some steps may seem a bit cryptic or terse. That is by design, so you will think about things as you go … and so you will experiment. YOU WILL NOT BREAK ANYTHING.

In this document, if you see something in square brackets, e.g., [submit], look for a button with that label: submit .

At any point in this workshop, you can login to the Sessions tool and save your session. Go to the pulldown menu, "My Data…" and choose "Sessions." Login using:

**username:** davis2015
**password:** genome

This will be a shared login! Please use your initials and/or some other unique text string and do not overwrite the sessions of anyone else.

## 1a. Visit a session and observe the Publications track.
This track was created by aligning sequences appearing in millions of publications to the genome.

- Use "Sessions" or "My data, Sessions" pulldown to get to the sessions page.
- Under "Restore Settings" type
   user:   example
   session:  hg19_GNRH1    [Submit]
- In the "Updated Session" section, click the "Browser" link.
- Observe this gene:  2 isoforms, one with intron in 5'-UTR
- Put your mouse over the items in the Publications track. This data track is made by aligning sequences from the literature to the genome. What do the alignments in this track say about the two isoforms in the genes track above?
- Find Khan2007 or Khan2008 papers in the middle of the screen (or really, any paper you see there). Hover your mouse over it to see the tooltip, then click into it to see details.

## 1b.  Zoom to the exon in the middle of the screen.
- Place your cursor over the large exon in the middle of the screen.  Note that the two isoforms have different numbers of exons, depending on the splicing pattern.
- Place your cursor in the Base Position section of the Browser graphic (at the top, where the scale bar is).
- Drag it to the right or left to highlight the area of the exon and release when you have the exon (and the Khan papers) centered.

## 1c.  Zoom to codon numbering.
- Drag-and-zoom again to a window about 5 amino acids wide.
- Turn on the codon numbering by clicking the little minibutton to the left of the UCSC Genes track, then checking the "show codon numbering" checkbox.

## 1d.  Jump to the end of the exon.
- Position your cursor over the double-headed arrows to the right and then the left of the two isoforms of GNRH1.
- Click the right arrowheads in the upper isoform and observe the start codon.

## 1e.  Log in and save a session with the full exon.
- Zoom out using 10x button.
- Move the display to the right, using click-and-hold in the graphic and dragging the image, centering the exon.
- In "My data, sessions" log yourself in:
      user:   davis2015
      pwd:   genome
- In "Save settings"
       name:  hg19_XX_gnrh1Exon   (or something else of your choosing)
     where "XX" is your initials (if someone else has used your initials, use "hg19_XX2_ gnrh1Exon2 or something else)
- Throughout this exercise, feel free to save any sessions you like or to visit the sessions created by others (click the "Browser" link on the right end of the line naming the session).

## 2.  Find pathway information about metabolism.
The vacuolar ATPase plays a key role in energy conversion

- Type the gene name "ATP6V0A1" into the position/search box (that's a zero) and select it from the suggestions that appear.  [ go ]
- Click into the highlighted gene.
- Find the "Pathways" link in the top box of links and click.
- Visit the Kyoto Encyclopedia of Genes and Genomes (KEGG) link to oxidative phosphorylation.
- Note that the gene you clicked is highlighted in red type.

## 3a.  Go to gene NR5A1 and turn on the "OMIM AV SNPs" track to see the biological function.

- [ hide all ] (below the Browser graphic).
- Type the gene name into the position/search box and select it from the suggestions that appear. [ go ]
- Turn on the "OMIM AV SNPs" track from the "Phenotype…" group
- Mouseover the SNPs in this track to see the phenotype.
- Click into one or more of them to read details to learn the function of the gene.
- Click into "OMIM Allelic Variant" on the details page to see the record at OMIM.

## 3b.  Jump to mouse homologue of  NR5A1.

- Use "View,   In Other Genomes" and go to mouse mm10.
- Is the gene named the same in mouse?  Many are not.
- Open the "Other RefSeq" track (switch visibility to pack using the right-mouse button) and see if you can find the human homologue aligned to the mouse.
- Use the "Genomes" pulldown to get back to human hg19.  Which gene are you at? The "Genomes" pulldown remembers your last location at a gene assembly.

## 3c.  Check out the coordinates in hg38.

- Note the hg19 coordinates in the window to the nearest one million bases.
- Use "Genomes" pulldown and go to human hg38, the most recent release of the reference assembly.  Which gene are you at?  Not the same gene, is it?
- To get to the same gene in hg38, you need to use the "View,   In Other Genomes," which uses homology, so go back to hg19 via "Genomes," then navigate to hg38 via "View,   In Other Genomes"
- Are the coordinates the same as on hg19?  Nope.  They never are (almost never) !  It is important to know what assembly someone is using when they give you coordinates!

## 4.  Back at hg19, find out where in the genome is the amino acid sequence FPVALSCRCGPCRRSTSDCGG.

- Use Blat from the "Tools" pulldown and select Human, hg19.
- Is it unique in the genome?
- Check out both the "Browser" and "details" links.
- From which strand is it transcribed?  Hint:  check out the amino acids in the UCSC Genes track.  You can also use the minibutton next to the Base Position track at the top to turn on all three reading frames (switch from "dense" to "full").  And use the little arrow in the upper left of the graphic image to flip strands.
- If you still have the OMIM AV SNPs track turned on, hover your mouse over the SNP 152780.0003 in the display.  Note that the AA change in that SNP is GLY to SER.  Is that consistent with everything else you see?

## 5.  View the probes in the commercial microarray CytoScan probeset from Affymetrix.

- Click to the [ track search ]  button below the Browser graphic.
- Type in "cytoscan" and [ search ]
- Click into the "Affy CytoScan" link to read about the probeset.
- Set the track to "dense" then [ view in browser ].
- Zoom in and out and check  out some genes to observe the probe coverage across the genome.
- Now scroll down the page to the "Variation" track group and note that the "SNP/CNV Arrays" track is set to dense.
- Click into the link to above the track controls to see other probesets that are available.
- Click into the "schema" link to see how many probes there is (rowcount in the table).

## 6a.  Look at the conservation signal for the IRF6 gene across many species.

- [hide all]
- In the search box, type IRF6
- In the Comparative Genomics group, turn on the Conservation track to full.
- Use the minibutton to left of the track to configure the track:  turn on more species, esp those far from human (e.g., fish) and [ submit ]
- Zoom into exons.
- Click into the image in the Conservation track.
- Use toggle at top of page to switch to showing only bases that differ.
- Turn on the UCSC Genes track and click into one of the genes (the darker ones are the more reliable annotations).  Read about the function of the gene.  Is it reasonable that it be conserved through evolution in the pattern you see?
- Turn on the OMIM Genes track (Phenotype and Literature section) and mouse-over the item.
- If you do not know what van der Woude syndrome is click into the gene in the OMIM Genes track and read about it.  You may want to click over to OMIM by clicking their link at the top of the section.  There is abundant information about many genes at OMIM.

## 6b.  Make a Session from this view to share with others.

- Click back to Genome Browser in the top blue bar
- Using "My data…  Sessions" pulldown, name your session.
    session:  hg19_<yourInitials>_IRF6
- Note that we are sharing this sessions login.  You may make your own as well, though it requires a confirmation email.

## 7. Set a filter to display only the SNPs from snp142 that change an amino acid in NR5A1.

- Type the gene name into the s box and select it from the pulldown. Submit.
- Turn on "All SNPs (142)" track to "pack." (Find it in the "Variation" group at the bottom of the screen)

- Use minibutton on left to access configuration options.
- Scroll far down the page to the section that describes the color conventions used on the page. Note which SNP types are considered "non-synonymous."
- Scroll back up the page and open the Filtering Options.
- Select the classes of SNPs you would classify as non-synonymous.

## 8. Find a list of genes involved in luteinization and pathway information about some of them.

- Type the word "luteinizing" into the search box on the Genome Browser.
- Find and click on the FSHB gene in the list (or another)
- Click into the gene in the Browser display.
- Click the "Pathways" link in the top section of the gene details page.
- Click into both of the links in the KEGG section and also click the Biocarta link.

## 9. Get a list of all missense SNPs in NR5A1 from the Table Browser into a Custom Track.

- In Table Browser ("Tools" pulldown menu),
      Choose group "Variation and…" track All SNPs(142)
- Type NR5A1 into "Position" box, then "lookup" and choose from list.
- Output format "selected fields from…"   [ Get output ]
- Choose checkboxes "name" and "func" and the coordinates.  [ Get output ]
- To test that you have retrieved SNPs from the right place, grab an rs# and copy/paste or type into "position and search" box on the Genome Browser.  Still in the same gene?
- Repeat this with new output format: "custom track."   [ Get output ]
- Make up meaningful names at the top. "name" = shortLabel (becomes link above track controls); "description" = longLabel (displayed in graphic)

## 10. Find all the genes in the chromosome cytological band, 4p16.3, their coordinates and a brief description.

- Using the Table Browser, find the table, knownCanonical (associated with UCSC Genes track in Genes group).  (note:  if you use the table, knownGene, you get info about all isoforms).
- Use "position" box and "lookup" button to get coords (type in "4p16.3").
- output format:  selected fields from ….  [ get output ]
- Select chrom, chromStart, chromEnd from knownCanonical and geneSymbol and description from kgXref.  [ get output ] [ get output ]

## 11. Get coordinates from the Table Browser, using a list of genes as input.

- Go to http://bit.ly/ucDavis2015 and select the genelist file.
- Copy the contents of the file.
- Go to the Table Browser and select "UCSC Genes" track and "knownGene" table. [ paste list ]
- Use output format, "selected fields from…" to get just the coordinates (chrom, txStart, txEnd), and from the kgXref table, the geneSymbol and description.

## 12.  Load a group of Custom Tracks from a URL.

- From the Browser graphical viewer, navigate to "add custom tracks"  (or "manage custom tracks" first).
- In a second Firefox (or Chrome, or…) window, go to http://bit.ly/ucDavis2015 and then click into ctExamples.txt.
- Copy the URL of this file into the upper text box (or you could copy the data) [ Submit ].
- Click the top "chr6" link the "Pos" column.
- Compare the data on the .txt page and the display in the Genome Browser to see how the text determines the properties of the tracks.
- If you are attending the Advanced workshop, you will want to save your session at this point so you can reuse it.

## 13.  Load a session to observe the reads in a BAM file.

A BAM file can be loaded into the Browser, but only from a location already on the web somewhere.  The BAM and its associated .bai file must be together in the same directory on your server.  Read more about loading BAM files: by go to Custom Tracks page and click the "BAM" link at the top of the page.

- Load session:
   user: example
   session: hg18_bamSnps
- Note the individual reads, display of mismatches and coloring.
- Check out the track controls to alter display options.
- Zoom out to view larger region.

## 14.  Compare disease-associated variation resources.

Many databases contain information about variants that or do not cause pathology.   The following session shows them all together so you can observe overlap, and click over to the original resource.  Note that you do not need to be logged in to view some else's session. Login is only required if you wish to save your own.

- Load session:
   user: example
   session: hg19_variants4
- Navigate to your favorite gene (or grab one from the genelist at http://bit.ly/ucDavis2015) and observe the SNPs at that location.  Save the session under user: davis2015.

## 15.  Compare normal variation.

Many sequencing projects have released variation data from apparently normal individuals.

- Load session:
   user: example
   session: hg19_variants3normal

## 16.  Make primers that will amplify a fragment from mRNA, but not from DNA, using the large exon and one neighboring exon in the middle of the gene, BRCA1.

- Type the gene name into the gene box.
- Zoom in to include the large exon + a neighbor or two.
- Go to "View…   DNA" in the blue bar.
- Click "Use extended case/color options."  To highlight the exons,
     choose  "lower case" for default.
     in UCSC Genes row, check "toggle case," and "bold."
     type 255 in your color of choice.
- Open a second window or tab and go to "Tools,  In-silico PCR"
- Copy/Paste ~25-30 bp primers from the exons onto PCR page in the separate window.
  Choose "Target:  UCSC Genes" and remember to "Flip reverse primer"
- Try the primers on Mouse assembly mm10.


## 17.  Find predicted sites for common restriction enzymes in the gene, NRP1.

- Type gene name in gene box in Genome Browser.
- In Mapping and Sequencing section below the graphic, click on link, "Restr Enzymes"
- Type "EcoRI, BamHI, HindIII," and turn the pulldown to "pack" [ Submit ].
- Click into your BamHI site in the graphic to see details.
- From the details page, export a list of all BamHI sites in the region using the link "just BamHI"
- Zoom into any BamHI item down to the base level.