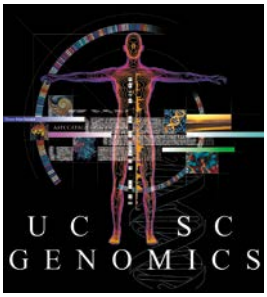


UCSC Genome Browser Workshop Practical
Variant Effect Prediction Course 2018
Johor, Malaysia

This document presents two exercises designed to demonstrate features of the Genome Browser and the Table Browser. We will use the Table Browser filter function, the Saved Sessions feature, and the Data Integrator. The goal will be to see how to upload and intersect your variants with OMIM Allelic variants, then to use the Data Annotation Integrator to assess biochemical consequences of your variants. The instructions are deliberately not comprehensive so that you will think about each step and look carefully at the Browser as you work.

#1. Load Custom track in pgSNP format.

This exercise shows how to load a pgSNP custom track from a file or by copy/paste. Note: VCF format is also supported, but it is too big to upload and must be hosted on a server accessible to http: protocol.

OUTLINE

1. **Set up the Genome Browser to show the UCSC Genes track on hg19.**
2. **Find file, pgSnpTrack.txt, and copy either the URL or data.**
3. **Load Custom Track into Browser.**
4. **View the SNPs in the Browser.**

DETAILS

1. Set up the Genome Browser to show the UCSC Genes track on hg19.

- At the Genome Browser (http://genome-asia.ucsc.edu), click “Genome Browser” and on the following page, use “Genome Browser” menu to “Reset All User Settings,” then go the human hg19 genome assembly.
- Find the [hide all] button below the graphic and click it.
- Turn on the UCSC Genes track to “pack” (look for it in the Genes and Gene Prediction blue-bar group below the Browser graphic). [refresh]
- Zoom out 10x to see adjacent genes.

2. Find file, pgSnpTrack.txt, and copy either the URL or data.

- In a separate web browser window or tab, navigate to http://bit.ly/ucscMalaysia2018
- Click the filename, pgSnpTrack.txt
- Copy either the URL or the data.
- NOTE about the pgSNP format: The last two fields, alleleFreq and alleleScores can 0,0, as long as the number of items matches alleleCount.

#chrom	chromStart	chromEnd	name	alleleCount	alleleFreq	alleleScores
chr6	31691111	31691112	G/A	2	30,60	90,70
chr6	31691277	31691278	T/G/A	3	9,60,7	80,80,30

can be:

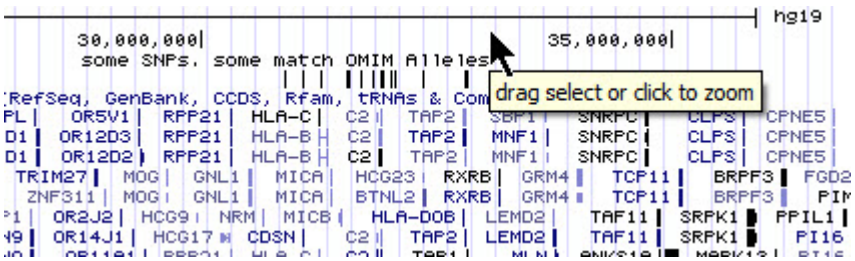
chr6	31691111	31691112	G/A	2	0,0	0,0
chr6	31691277	31691278	T/G/A	3	0,0,0	0,0,0

3. Load Custom Track into Browser.

- Back at the Genome Browser, pull down the menu, My Data > Custom Tracks
- Paste the URL or data into the upper text box. [Submit]
  If you are pasting the data, you may edit the top line to rename the track. Substitute your name as, e.g.:
  description= "Johan's data" or whatever
  (quotes are required if there is a space in the title)…. This will appear in the Browser graphic above the data in the display.

4. View the SNPs in the Browser.

- Click on the item in the "Pos" column: chr6.
- Zoom out 10x.
- Zoom out 100x three times (total of 10^7 zoomout!).
- You can see the cluster or variants in the middle of the screen. Zoom into the region with the SNPs by putting your mouse at the top of the coordinate window and dragging right or left to highlight the region:



- Turn on the OMIM Alleles track to dense and drag it up next to your data track. You can see that most of these match your data.

#2. Check Your Custom Track for OMIM Allelic Variant SNPs.

This exercise shows how to use the Data Integrator to check your list of SNPs against OMIM Allelic Variants to see which of them have a known phenotype.

OUTLINE

1. **Choose the OMIM Alleles and your table in the Data Integrator.**
2. **Select relevant data fields for display.**
3. **Export list.**

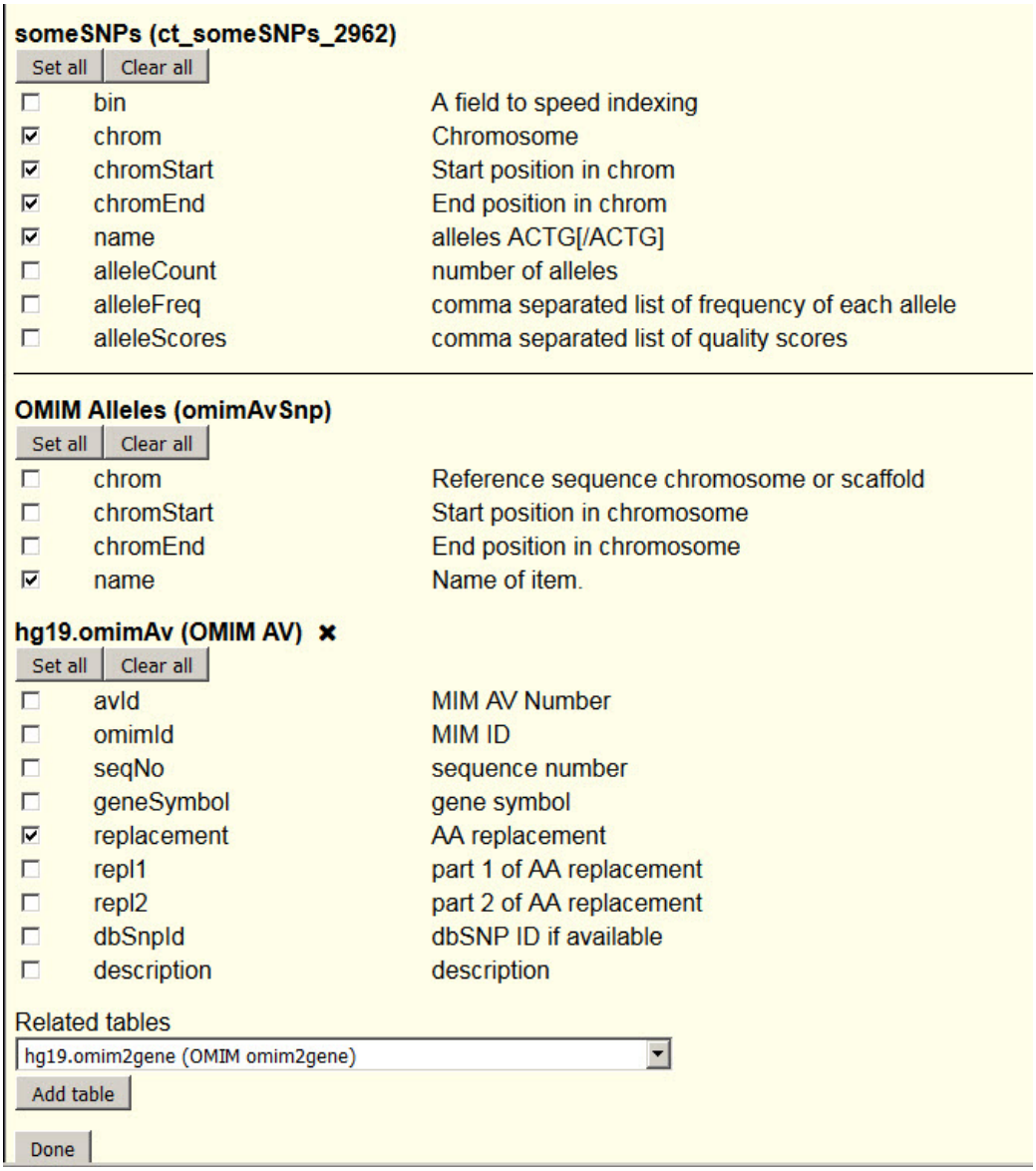
DETAILS

1. Choose the OMIM Alleles and your table in the Data Integrator.

- In the pulldown menus at the top of the Browser, go to Tools > Data Integrator.
- In "region to annotate," type "chr6" and Enter.
- In the "Add Data Source" section, choose your Custom Track and add to list of tables. It should be already set, but if not, look for it in "track group" custom tracks and select "someSNPs." [Add]
- Add a second table from track group "Phenotype and Literature": OMIM Alleles. [Add]

2. Select relevant data fields for display.

- Under "Output Options," [choose fields….]
- From the someSNPs table, [clear all], then select "chrom", "chromStart", "chromEnd" and "name"
- From the OMIM Alleles table, [clear all], then select "name"
- Under Related tables, "hg19.omimAv" [add table]
- Select "replacement." [Done]



- Add one more table, omimPhenotype and select the data field, phenotype. [done]

- Back at the Data Integrator, [get output]
- Wherever a variant in your file matches a location already known to OMIM Allelic Variants, the result will show it, along with the gene and amino-acid substitution, if applicable. Not that your variant may not change the amino acid in the same way as the change annotated by OMIM at the same location.

```
# hgIntegrator: database=hg19 region=chr6:1-171115067 Tue Nov  1 04:57:02 2016
#ct_someSNPs_2962.chrom ct_someSNPs_2962.chromStart     ct_someSNPs_2962.chromEnd       ct_someSNPs_2962.name
omimAvSnp.name  omimAvSnp.omimAv_replacement
chr6    30886637        30886638        C
chr6    30889752        30889753        GG      612802.0002     VARS2, ALA626ASP
chr6    31084645        31084646        A       602593.0004     CDSN, 1-BP DEL, 746G
chr6    31084803        31084804        G
chr6    31084848        31084849        C
chr6    31084967        31084968        T       602593.0006     CDSN, GLY142TER
chr6    31085216        31085217        AC      602593.0003     CDSN, LYS59TER
chr6    31085224        31085225        C       602593.0005     CDSN, 4-BP DUP, 164GCCT
chr6    31274383        31274384        G
chr6    31540240        31540241        G
chr6    31540312        31540313        G       153440.0002     LTA, 252A-G
chr6    31540783        31540784        -       153440.0001     LTA, THR26ASN
chr6    31542481        31542482        T       191160.0005     TNF, -850C-T
chr6    31544937        31544938        A
chr6    31560909        31560910        T       611550.0001     NCR3, -412G-C, 5-PRIME UTR
chr6    31691111        31691112        G/A
chr6    31691111        31691112        C
chr6    31691111        31691112        A
chr6    31691167        31691168        T/C
chr6    31691167        31691168        C
```

#3. Use Variant Annotation Integrator to assess biochemical consequences of your SNPs.

This exercise shows how to use the Variant Annotation Integrator to obtain the molecular consequences of your variants. The VA will accept data in pgSNP or VCF formats.

OUTLINE

1. **Set up Variant Annotation Integrator.**
2. **Export results.**

DETAILS

1. Set up Variant Annotation Integrator.

- In the Tools pulldown menu, select Variant Annotation Integrator.
- In the Select Variants section, choose your custom track, "some SNPs…" or "Johans' data" or whatever you called it.
- Under Select Genes, choose RefSeq Genes.
- Note the many options for analyzing your data. For now, accept the defaults, and also add the HGVS g. terms, c. terms and p. terms (= genomic, cDNA and protein coordinates).

2. Export results.

- Configure output: Choose Variant Effect Predictor (HTML) to make the results easier to read onscreen. If you wish to save to a file, then the (tab-separated text) option will be better. [get results]

#4. Bonus question: Extract list of genes and Common SNPs from a region.

This exercise shows how to use the Data Integrator to pull a list of genes and SNPs in those genes from a genomic interval. The Data Integrator allows you to pull data from up to five datasets.

OUTLINE

1. **Navigate to cytoband 4q13.1 (or any region you may be interested in).**
2. **Using the Data Integrator, add UCSC Genes, kgXref, and Common SNPs tables to the list.**
3. **Choose data fields.**
4. **Export list.**

DETAILS

1. Navigate to cytoband 4q13.1.

- Genome: hg19 [Hide All]
- Turn on UCSC Genes
- Type "4q13.1" into position/search box. (And coordinate interval can be used.)

2. Using the Data Integrator, add UCSC Genes, kgXref, and Common SNPs tables to the list.

- Go to "Tools, Data Integrator."
- Note that position is pre-selected.
- Select tables:
  group, Genes and Gene Prediction; track, UCSC Genes [add]
  group, Variation; track, Common SNPs; [add]

3. Choose data fields.

- [choose fields]
- In the UCSC Genes table, [clear all]
- Then check the box at: chrom
- In the Related tables: kgXref [add table]
- In the kgXref table, check the box at: geneSymbol
- In the Common SNPs 147 table, [clear all]
- Then check the boxes: chromStart, name, func [done]

4. Export list.

- [get output]
- The result is a list of all the genes in the region, all the Common SNPs in the gene footprint and the functional effect of each SNP.