

How to create a personalized syndrome description

This tutorial has 3 parts:

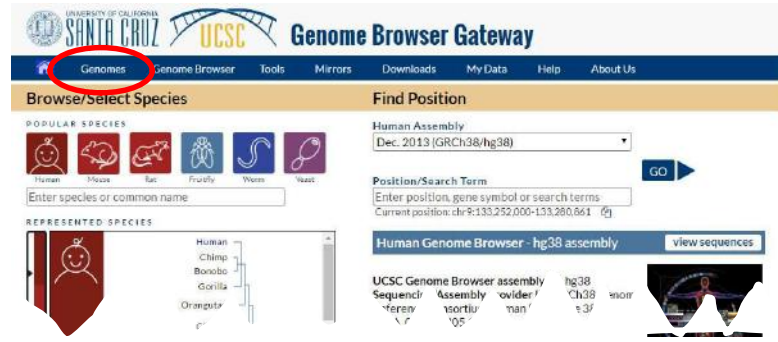
1. How to use the UCSC Genome Browser
2. How to look up individual genes or phenotypes
3. How to compile gene dosage information for a region of interest

How to use the UCSC Genome Browser

1. Go to the website for the University of California at Santa Cruz Genome Browser at: genome.ucsc.edu

2. You will see the home page that looks like this.

3. Click on "Genomes" and select : Human GRCh37/hg19



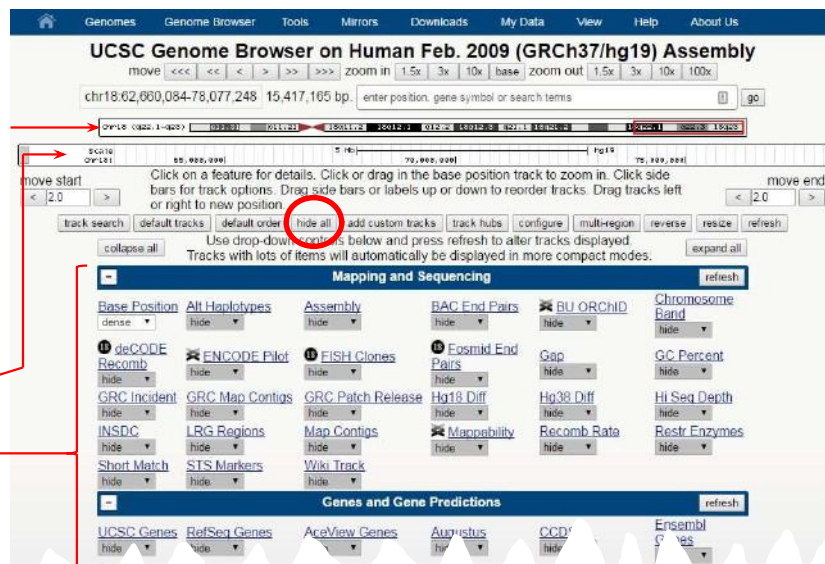
4. Under "assembly" select the assembly used in the genomic lab report that diagnosed the chromosome 18 abnormality; hg18 or hg19.

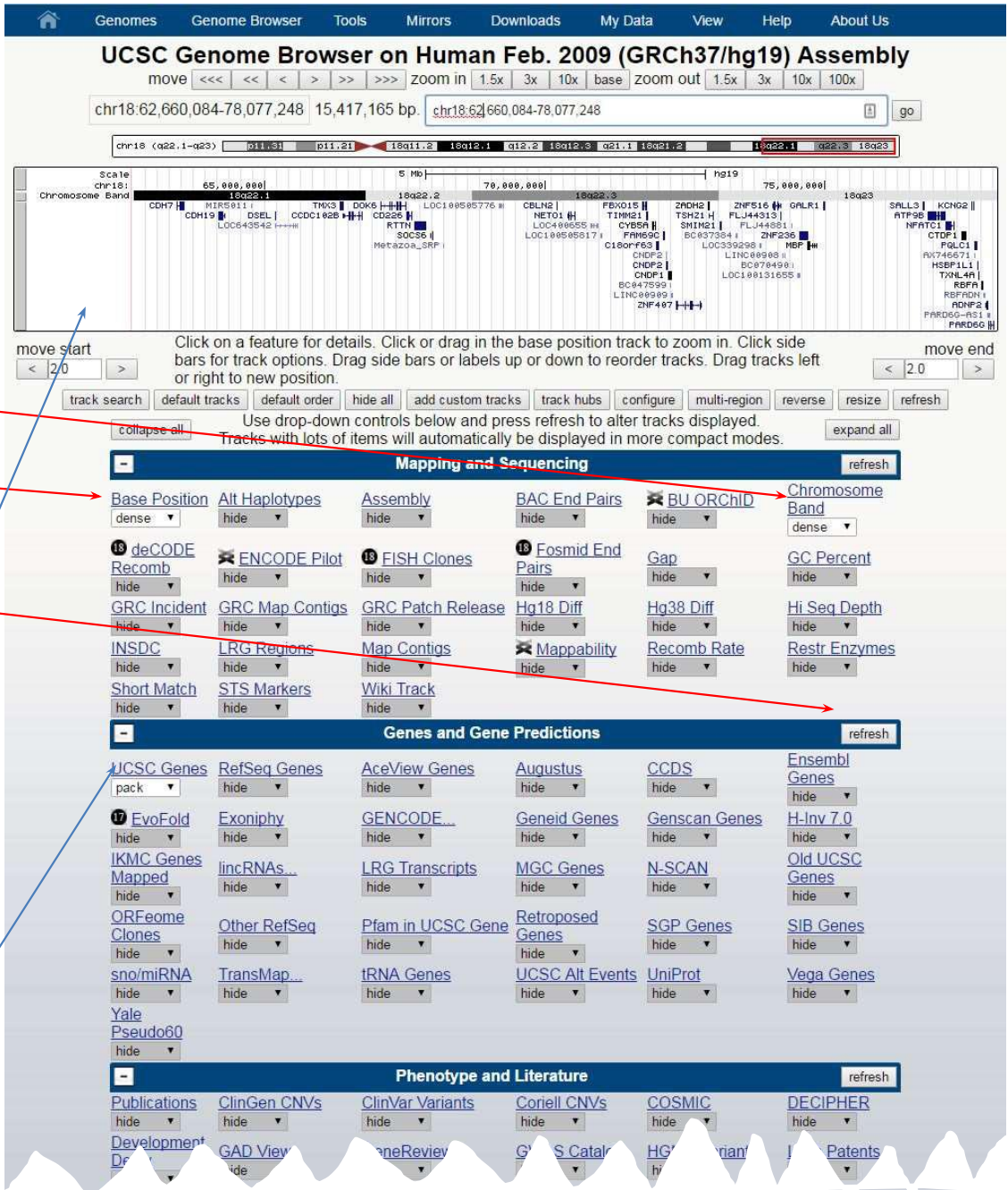
5. Under "search term" type the coordinates of the persons' deletion or duplication. For example, someone with 18q- and a breakpoint within the Reference Group region might have a breakpoint such as this:
chr18:62,660,084-78,077,248
Then click "submit"



7. The next screen may not look exactly like this, so click on "hide all." This step will remove all the information you don't need. In the next steps you can add just the information you want to see.

8. This is the main page you will work from. The layout of this page includes an ideogram of the selected chromosome with a red box around the part of the chromosome shown in the viewing area. The bottom portion of the screen includes about 150 drop-down boxes for selecting "tracks" that include different types of genome annotation data. You will not need to use the majority of these tracks





9. The next step is to select only the genome annotation tracks that you want to see. By clicking on any of the track names you will go to a page with an explanation of the data in that track.

10. Select under "Chromosome band" "dense".

11. Select under "UCSC Genes" "pack."

12. Click on any "refresh" button on the right.

13. Your genome viewing window should look like the one pictured here – with the scale bar and base position at the top. The chromosome bands shown in the next row and the genes shown beneath. In this particular view I have clicked on the "UCSC Genes" and on the subsidiary page de-selected "splice variants" because I do not want to see all the possible variations for each gene.

14. Now you are ready to explore the genetic content of a deletion or duplication. Click on any gene in the viewing window to learn more.

15. In addition you can determine if particular genes are duplicated or deleted in a control population. Such a gene is unlikely to be a cause of disability. These data are collated by the Database of Genomic Variation and can be found by scrolling down to the section of genome annotation tracks labeled "Variation" and then selecting the drop-down box labeled "DGV Struct Var." Then click on a "refresh" button and areas of structural variation will be displayed in the viewing window.

16. Information on individual genes can be investigated by clicking on the gene of interest. For example, click on *TSHZ1*.

17. Clicking on the gene takes you to this page. From this page you can access many databases with information on the gene. A short summary is also included here.

Human Gene TSHZ1 (uc002ily.4) Description and Page Index

Description: Homo sapiens teashirt zinc finger homeobox 1 (TSHZ1), mRNA.
RefSeq Summary (NM_005786): This gene encodes a colon cancer antigen that was defined by serological analysis of recombinant cDNA expression libraries. The encoded protein is a member of the teashirt C2H2-type zinc-finger protein family and may be involved in transcriptional regulation of developmental processes. Mutations in this gene may be associated with congenital aural atresia syndrome. [provided by RefSeq, Jan 2012].
Transcript (including UTRs)
Position: chr18:72,922,710-73,001,905 **Size:** 79,196 **Total Exon Count:** 2 **Strand:** +
Coding Region
Position: chr18:72,997,498-73,000,596 **Size:** 3,099 **Coding Exon Count:** 1

Page Index	Sequence and Links	UniProtKB Comments	Genetic Associations	MalaCards	CTD
Gene Alleles	RNA Structure	Protein Structure	Other Species	GO Annotations	mRNA Descriptions
Other Names	Model Information	Methods			

Data last updated: 2013-06-14

Sequence and Links to Tools and Databases

Genomic Sequence (chr18:72,922,710-73,001,905)	mRNA (may differ from genome)	Protein (1032 aa)
Gene Sorter	Ensembl	Entrez Gene
CGAP	Ensembl	Entrez Gene
Gepis Tissue	GTEX	H-INV
MGI	MOPED	neXtProt
UniProtKB		

Comments and Description Text from UniProtKB

ID: TSH1_HUMAN
DESCRIPTION: RecName: Full=Teashirt homolog 1; AltName: Full=Antigen NY-CO-33; AltName: Full=Serologically defined colon cancer antigen 33;
FUNCTION: Probable transcriptional regulator involved in developmental processes. May act as a transcriptional repressor (Potential).
SUBUNIT: Interacts (via homeobox domain) with APBB1 (via PID domain 1).
SUBCELLULAR LOCATION: Nucleus
TISSUE SPECIFICITY: Expressed in brain; strongly reduced in post-mortem elderly subjects with Alzheimer disease.
DISEASE: Aural atresia, congenital (CAA) [MIM:607842]: A rare anomaly of the ear that involves some degree of failure of the development of the external auditory canal. The malformation can also involve the tympanic membrane, ossicles and middle ear space. The inner ear development is most often normal. Different CAA forms are known. CAA type I is characterized by bony or fibrous atresia of the lateral part of the external auditory canal and an almost normal medial part and middle ear. CAA type II is the most frequent type and is characterized by partial or total aplasia of the external auditory canal. CAA type IIA involves an external auditory canal with either complete bony atresia of the medial part or partial aplasia that ends blindly in a fistula leading to a rudimentary tympanic membrane. CAA type IIB is characterized by bony stenosis of the total length of the external auditory canal. CAA type III involves bony atresia of the external auditory canal and a very small or absent middle-ear cavity. Note=The disease is caused by mutations affecting the gene represented in this entry.
SIMILARITY: Belongs to the teashirt C2H2-type zinc-finger protein family.
SIMILARITY: Contains 5 C2H2-type zinc fingers.
SIMILARITY: Contains 1 homeobox DNA-binding domain.
SEQUENCE CAUTION: Sequence=AAC18047.1; Type=Frameshift; Positions=304, 1048; Evidence=; Sequence=BAE06124.1; Type=Erroneous initiation; Note=Translation N-terminally shortened.; Evidence=;

Genetic Association Studies of Complex Diseases and Disorders

Genetic Association Database: TSHZ1
CDC HuGE Published Literature: TSHZ1
Positive Disease Associations: Glucose, Heart Rate
Related Studies:

18. The databases with potential clinical significance are OMIM (Online Mendelian Inheritance in Man), PubMed and MGI (Mouse Genome Institute).

In addition, The Chromosome 18 Clinical Research Center continuously curates emerging data as it relates to potential gene dosage effects, for all the genes on chromosome 18. These data are accessible through custom tracks on the UCSC Genome Browser.

How to compile gene dosage information for a region of interest

Custom tracks created for investigating chromosome 18 gene dosage effects and visualized using the UCSC Genome Browser are explained on our website at:

<http://www.pediatrics.uthscsa.edu/centers/chromosome18/dosage.asp>

There are 2 reasons for using our Gene Dosage Maps:

1. To **investigate individual genes or phenotypes** to determine what is known related to the gene dosage effects of that gene or phenotype.
2. To **investigate the clinical consequences of a chromosome deletion or duplication** region. This is the most straightforward approach to investigate a patient's deletion.

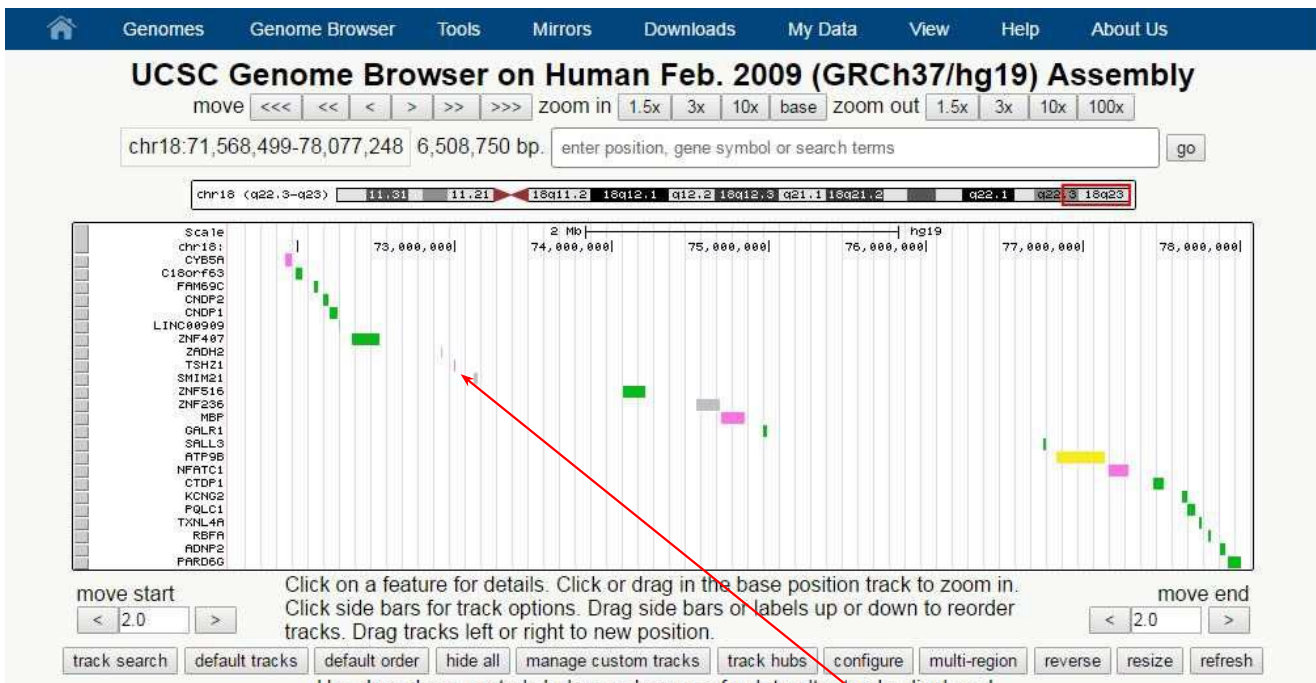
Investigating individual genes and phenotypes:

There are 2 types of data and therefore 2 custom track data sets;

1. The annotated genes (Gene Dosage Map) includes information related to the gene dosage effects for each gene on chromosome 18 and are color coded as shown below.
2. The annotated phenotype regions (Phenotype Map) indicates the region of chromosome 18 linked to a specific phenotype for which a gene has not yet been attributed. When these data come from linkage or GWAS studies they are agnostic with regard to mechanism. This means they may or may not be relevant in a gene dosage context. Phenotype regions are also derived from critical regions data identified by genotype / phenotype mapping of people with copy number changes which are of by definition relevant to an abnormal gene dosage mechanism. The phenotype regions are color coded with regard to mechanism as shown below.

Green	This gene is unlikely to cause a phenotype when there is a copy number change.
Pink	This gene is dosage sensitive. There can be either high or low penetrance of the abnormal phenotype.
Yellow	A copy number change in this gene ONLY results in a phenotype in the presence of a second event (eg. drug exposure or a second genetic change).
Red	This gene is thought to be haplolethal.
Grey	The consequences of a copy number change in this gene are unknown.

On our website: (<http://www.pediatrics.uthscsa.edu/centers/chromosome18/dosage.asp>) if you click on the link to the "Gene Dosage Map" and select the region "chr18:71,568,499-78,077,248" you will see a screen that looks like this:



Here all the genes are color coded as described above. To learn more about a specific gene with regard to gene dosage, click on that gene. For example, *TSHZ1*.

In addition, you can add any of the other standard data tracks to you browser window as explained previously.

clicking on the *TSHZ1* gene in the Gene Dosage Map takes you to this page.

Next click on TSHZ1 Outside Link.

Custom Track: TSHZ1

TSHZ1

Outside Link: [TSHZ1](#)

Item: TSHZ1

Score: 0

Position: [chr18:72922732-73001901](#)

Band: 18q22.3

Genomic Size: 79170

Strand: +

[View DNA for this feature \(hg38/Human\)](#)

[Go to TSHZ1 track controls](#)

Data last updated: 2015-10-21

You are then directed to this details page with a brief description of the data and the references.

By collating the data on the pink and yellow coded genes within a region of an individual's deletion a genetic basis can be compiled. In addition, Determining which critical regions and dosage sensitive genes are NOT within a person's deletion can eliminate certain phenotypes as irrelevant to this individual.

Gene Symbol

TSHZ1

Dosage Sensitivity Class

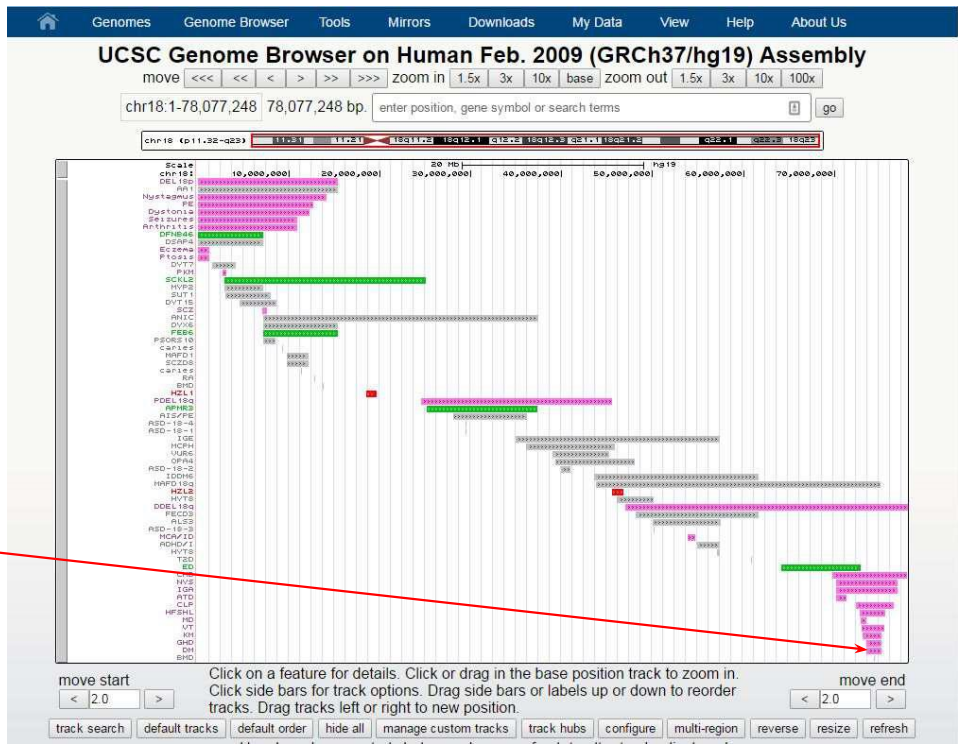
This gene is dosage sensitive. There can be either high or low penetrance of the abnormal phenotype.

Comments

Recently, TSHZ1 (teeshirt family zinc finger 1) has been linked with the aural atresia phenotype (1). Four individuals in two families were identified with aural atresia and small interstitial deletions of 18q. The common overlapping region of their deletion included a single gene: TSHZ1. They proceeded to identify mutations in this gene in four additional individuals from two families with isolated aural atresia. Aural atresia without microtia is a very rare finding and because it is easily observable and highly penetrant, it is the most characteristic clinical feature of distal 18q-. The heterozygous knock-out mouse shares this phenotype as well (2). This gene also plays a role in olfaction in the heterozygous mouse with a deletion of 1 copy of the e gene and in humans with mutations in this gene (3). This gene was also implicated in orofacial clefts (4).

Key References

(1) Feenstra I, Vissers LELM, Pennings RJE, Nillessen W, Pfundt R, Kunst HP, Admiraal R, Veltman JA, van Ravenswaaij-Arts, Brunner HG, Cremers CWRJ (2011) Disruption of teeshirt zinc finger homeobox 1 is associated with congenital aural atresia in humans. *J Hum Genet* 89:P19. doi: 10.1007/s12267-011-9116-1 (2) Cai T, Xie C, et al. (2011) Disruption of Tshz1 in mice leads to aural atresia and microtia. *BMC Med* 9:117. doi: 10.1186/1745-7189-9-117 (3) ... (4) ...



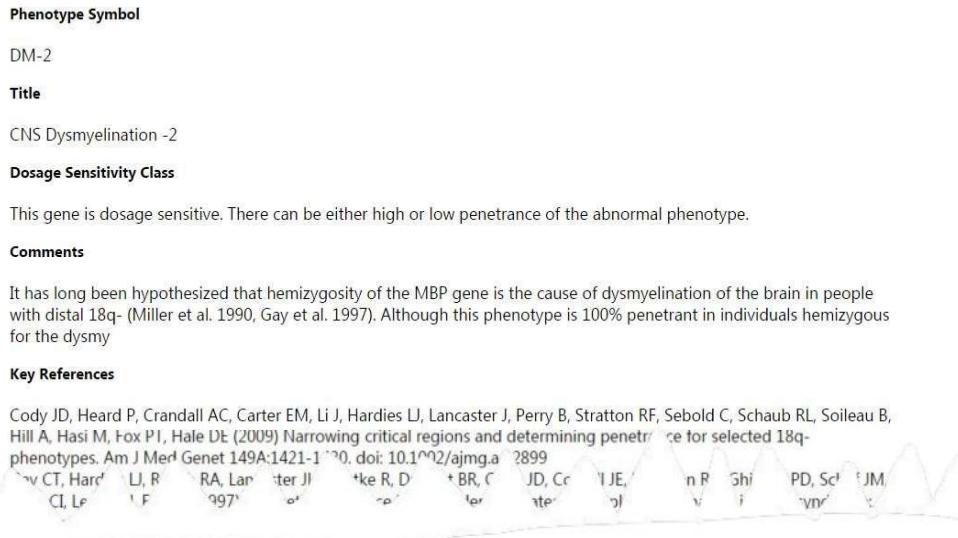
To look up a specific phenotype the process is similar. Go to our website and click on the “Phenotype Map.” To the right is the phenotype map for the entire chromosome. To learn more about the data used to identify the region, click on a region of interest. For example, DM (dysmyelination).

This will take you to a browser details page on which you have to click on the “Outside Link”



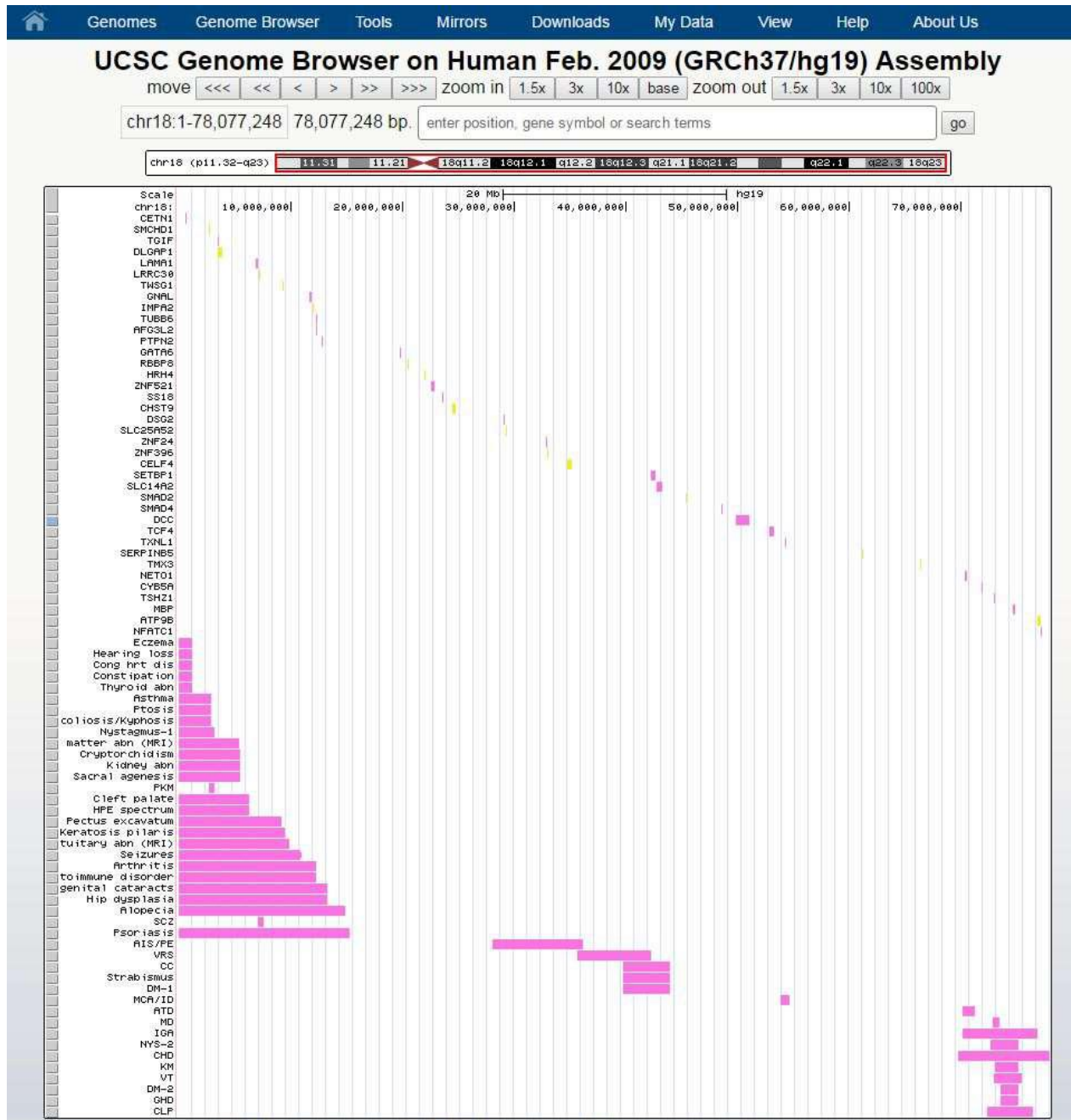
You are then directed to this details page with a brief description of the data and the references.

By collating the data on the pink coded phenotypes within a region of an individual’s deletion a list of potential phenotypes can be compiled. In addition, determining which phenotype regions are NOT within a person’s deletion can eliminate certain phenotypes as relevant to this individual.



Investigating the clinical consequences of a chromosome deletion or duplication

The Gene Dosage Map and the Phenotype map just described include the information on all the genes on chromosome 18 and all of the chromosome 18 localized phenotypes. Since most of the genes and many of the phenotypes are not thought to be dosage relevant, we have created a combined custom track with only the dosage relevant information. These versions of the gene dosage maps only include information on genes that have known clinical relevance (a small proportion of the genes) and those phenotypes known to be caused by gene dosage abnormalities. These are the genes and phenotypes color coded pink or yellow in the tracks described above. This is therefore a condensed version allowing a focus only on clinically relevant information. It can be used/viewed in the same way as the other custom tracks; select the region of interest, then click on the gene or phenotypes in order to get to more detailed information.



Note: When you are viewing these custom tracks within the Genome Browser you can also navigate within the browser and you can add additional tracks in the same way as in the previous set of directions