

UCSC GENOME BROWSER AND BIOMART

Dr. Garrett Dancik

UCSC Genome Browser

- The University of California, Santa Cruz (UCSC) Genome Browser (<http://genome.ucsc.edu>) contains a variety of tools
- The **Genome Browser** allows for interactive visualization of genomic data, which is visualized based on tracks:
 - NCBI RefSeq information
 - OMIM alleles
 - Clinical variants and SNPs, etc
- The **Table Browser** allows you to download sequence or feature data from all (selected) genes
 - Can download promoter regions, introns, exons, etc

UCSC Genome Browser examples

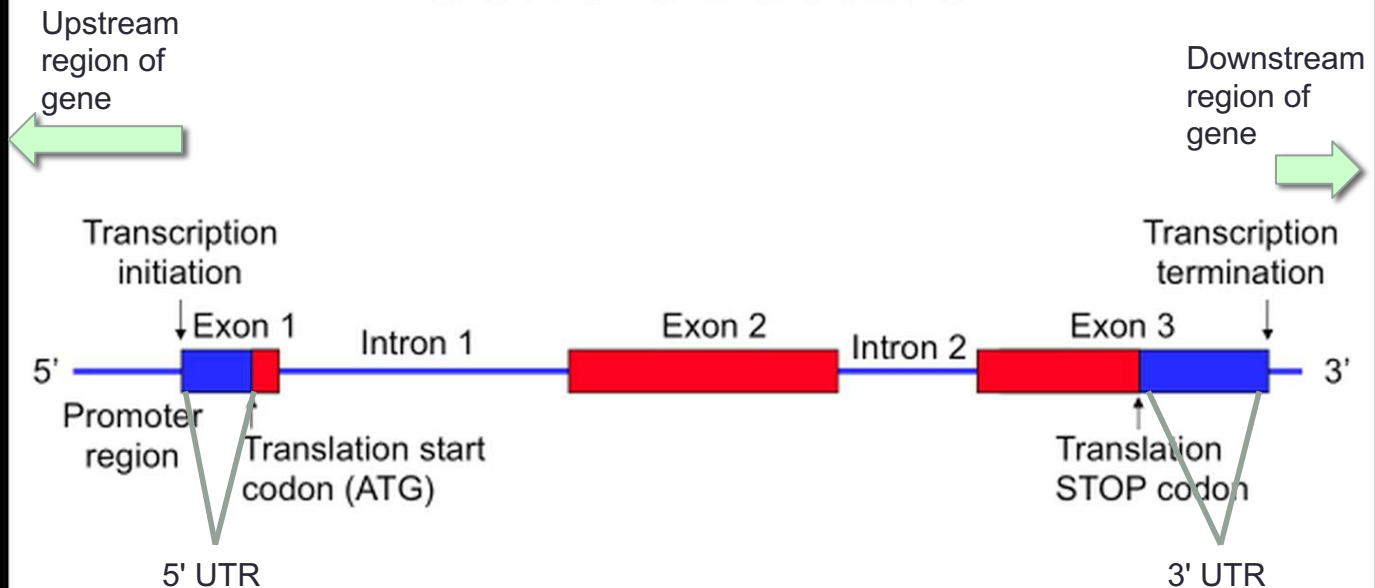
- Using the Genome Browser, let's look at the gene MAOA
 - Let's hide all tracks, then show the following (use pack view):
 - GenCode v32
 - NCBI RefSeq
 - Human mRNAs
 - OMIM alleles (display as “full”)
 - All SNPs (144)

UCSC Table Browser examples

- Using the Table Browser – to get selected information about genes/proteins
- Let's look at all NCBI RefSeq genes in humans
 - Select clade (organism type) and
 - For humans, select “Mammal” for *clade* and “Human” for *genome*
 - For *group*, select “Genes and Gene Predictions”
 - For *track* we will use “NCBI RefSeq”
 - For *region* we will use “genome”
 - We can enter specific genes (identifiers), if we want
 - For output format, select “selected fields from ..”
 - We will select name, chrom, exonCount, name2
 - Click “Get output” and select features

Overview of Gene Structure

Gene Structure



© Joanne Lind

Original image and video: <https://www.youtube.com/watch?v=h5HI2OqOJA0>

UCSC Table Browser examples

- Using the Table Browser – to get sequence information
 - Specify organism, database, and genes as before
 - For output format, select “sequence” and click “Get output”
 - You will then need to specify what parts of the sequence to get
- Can you get the following?
 - Get the CDS exon and intron sequences for BRCA2
 - Get the 5' UTR exon sequences for all human genes

BioMart

- BioMart (<http://www.ensembl.org/biomart/martview/>) is a database system providing access to genomic feature and sequence data for specific genes or proteins
- Ensemble is a genome browser and a joint project of the European Bioinformatics Institute (EBI) of the European Molecular Biology Laboratory (EMBL), and the Wellcome Trust Sanger Institute
- Using BioMart
 - Select a Dataset, such as Ensemble Genes → Human genes
 - Select your Filters (this is the query step, where you can enter the genes of interest, for example)
 - Select the Attributes (the desired output)
 - Click on Results

BioMart ortholog example

- I have recently received data from a mouse experiment
- High expression of the following genes (MGI symbols) were associated with resistance to radiation treatment in mice with bladder cancer:
 - Arl4a
 - Birc6
 - Car2
 - Oog3
 - Zfp110
- The next step is to evaluate these genes in human bladder cancer patients – what are the orthologs of these genes? We want to identify human orthologs with high confidence.