



Oct 13, 2020

Protocol to UCSC Genome Browser & BLAST

Forked from [Protocols for Bioinformatic Tools](#)

1

¹UCSC**1** *Works for me* This protocol is published without a DOI.

UCSC BME 22L



ABSTRACT

In this lab, students will work through BLAST and the UCSC Genome browser to find and analyze information about their genes of interest.

PROTOCOL CITATION

2020. Protocol to UCSC Genome Browser & BLAST . **protocols.io**
<https://protocols.io/view/protocol-to-ucsc-genome-browser-amp-blast-bnc3mayn>

FORK FROM

Forked from [Protocols for Bioinformatic Tools](#), Alyssa Ayala

LICENSE

This is an open access protocol distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited

CREATED

Oct 13, 2020

LAST MODIFIED

Oct 13, 2020

PROTOCOL INTEGER ID

43131

DISCLAIMER:

DISCLAIMER – FOR INFORMATIONAL PURPOSES ONLY; USE AT YOUR OWN RISK

The protocol content here is for informational purposes only and does not constitute legal, medical, clinical, or safety advice, or otherwise; content added to [protocols.io](#) is not peer reviewed and may not have undergone a formal approval of any kind. Information presented in this protocol should not substitute for independent professional judgment, advice, diagnosis, or treatment. Any action you take or refrain from taking using or relying upon the information presented here is strictly at your own risk. You agree that neither the Company nor any of the authors, contributors, administrators, or anyone else associated with [protocols.io](#), can be held responsible for your use of the information contained in or linked to this protocol or any of our Sites/Apps and Services.

ABSTRACT

In this lab, students will work through BLAST and the UCSC Genome browser to find and analyze information about their genes of interest.


UCSC Genome Browser

1 **Genome Browser**

You will now have the opportunity to utilize these tools yourself to investigate a gene that you find interesting. Use

these tools to answer the questions below and then relay these in the Lab Results of this Lab Notebook.


Find a gene of interest and type it onto the genome browser. Use the human GRCh37/hg19 Assembly. In your own words, give a brief description of this gene's function.



UNIVERSITY OF CALIFORNIA

SANTA CRUZ

Genomics Institute










Genome Browser Gateway

[Home](#)
[Genomes](#)
[Genome Browser](#)
[Tools](#)
[Mirrors](#)
[Downloads](#)
[My Data](#)
[Projects](#)
[Help](#)
[About Us](#)

Browse/Select Species

POPULAR SPECIES

 Human
  Mouse
  Rat
  Zebrafish
  Fruitfly
  Worm
  Yeast

REPRESENTED SPECIES

Find Position

Human Assembly

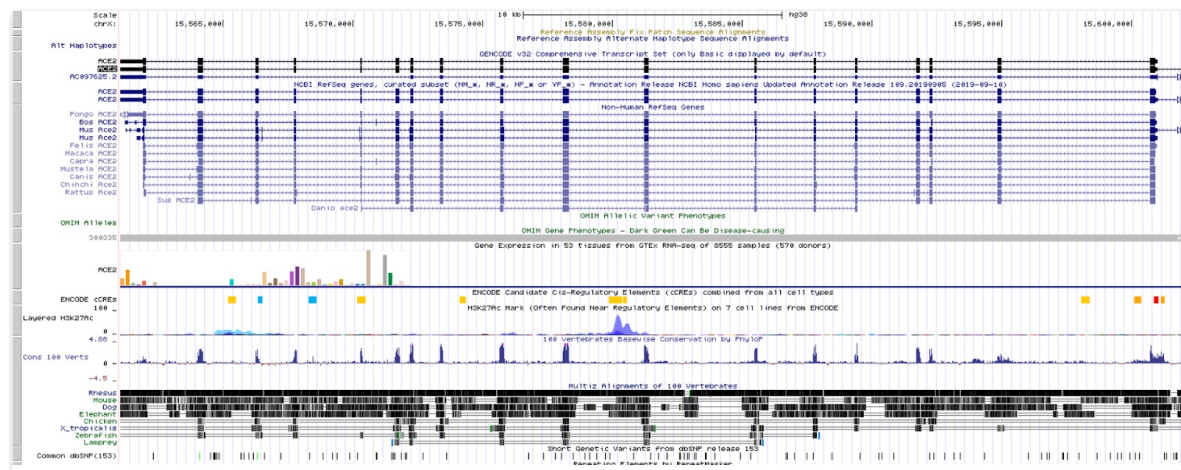
Dec. 2013 (GRCh38/hg38)

Position/Search Term

ACE2

GO

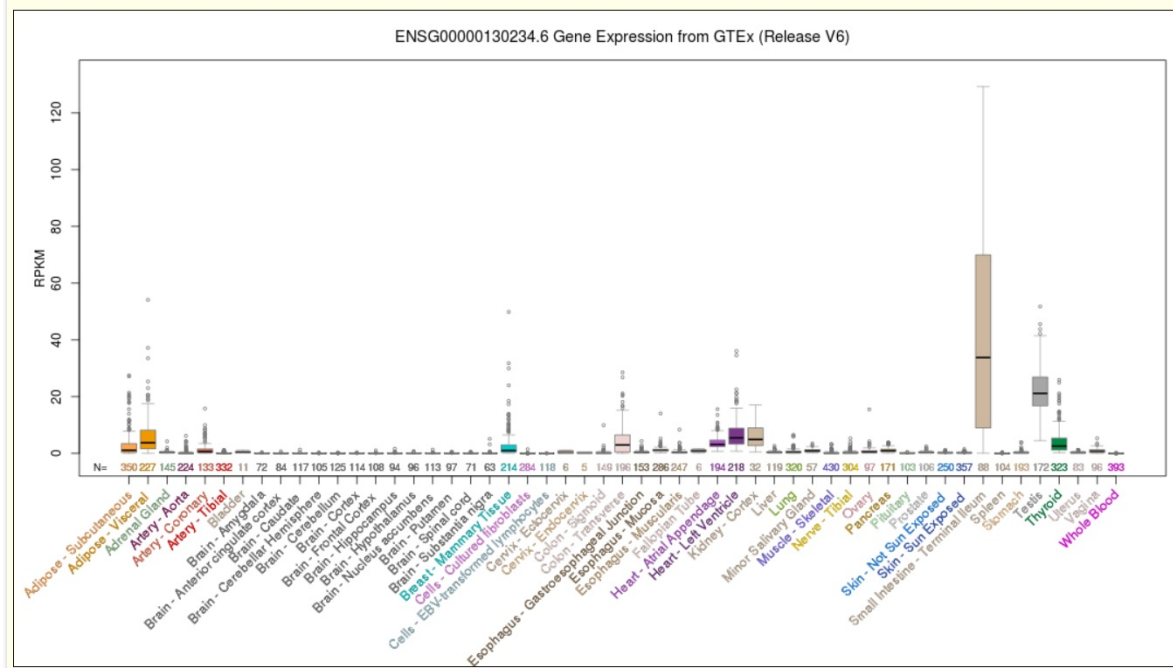
Current position: chrX:15,561,033-15,602,069



- Look through the tracks given on the displays and give the name of track which is used to give gene expression for different tissue types. What tissue do we observe the highest expression of RNA? (use track other than NCBI RefSeq)

RNA-Seq Expression Data from GTEx (53 Tissues, 570 Donors)

Highest median expression: 33.78 RPKM in Small Intestine - Terminal Ileum
Total median expression: 90.25 RPKM



3 What orthologs does this gene have? What is the significance of having these as an ortholog?

Orthologous Genes in Other Species

Orthologies between human, mouse, and rat are computed by taking the best BLASTP hit, and filtering out non-synthetic hits. For more distant species reciprocal-best BLASTP hits are used. Note that the absence of an ortholog in the table below may reflect incomplete annotations in the other species rather than a true absence of the orthologous gene.

Mouse	Rat	Zebrafish	D. melanogaster	C. elegans	S. cerevisiae
Genome Browser	Genome Browser	No ortholog	No ortholog	No ortholog	No ortholog
Gene Details					
Gene Sorter					
MGI	RGD	Ensembl			
Protein Sequence	Protein Sequence				
Alignment	Alignment				

4 What Biochemical and Signaling Pathways is this gene involved in? Again, what is the significance?

Biochemical and Signaling Pathways

KEGG - Kyoto Encyclopedia of Genes and Genomes

[hsa04614](#) - Renin-angiotensin system

BioCarta from NCI Cancer Genome Anatomy Project

[h_ace2Pathway](#) - Angiotensin-converting enzyme 2 regulates heart function

Reactome (by CSHL, EBI, and GO)

Protein Q9BYF1 ([Reactome details](#)) participates in the following event(s):

[R-HSA-2022378](#) ACE2 hydrolyzes Angiotensin-(1-10) to Angiotensin-(1-9)

[R-HSA-2022379](#) ACE2 hydrolyzes Angiotensin-(1-8) to Angiotensin-(1-7)

[R-HSA-2022377](#) Metabolism of Angiotensinogen to Angiotensins

[R-HSA-2980736](#) Peptide hormone metabolism

[R-HSA-392499](#) Metabolism of proteins

5 What diseases are associated with this gene?

MalaCards Disease Associations

MalaCards Gene Search: [ACE2](#)

Diseases sorted by gene-association score: [severe acute respiratory syndrome](#) (28), [neurogenic hypertension](#) (18), [hartnup disorder](#) (14), [tetanus neonatorum](#) (9), [internal hemorrhoid](#) (9), [intracranial aneurysm](#) (8), [posterior urethral valves](#) (6), [hypertension, essential](#) (2), [myocardial infarction](#) (1)

BLAST

6 BLAST

Do a BLAST search of a gene of interest.

Enter accession number(s), gi(s), or FASTA sequence(s) [?](#) [Clear](#) [Query subrange](#) [?](#)

From

To

```

hg38_knownGene_ENST00000427411.1 range=chrX:15561033-15602069 5'pad=0
3'pad=0 strand=- repeatMasking=none
GTTGTTGTGATCCCATGGCTACAGAGGATCAGGAGTTGACATAGATACTC
TTTGGATTTTCATACCATGTGGAGGCTTTCTTACTTCCACGTGACCTTGAC
TGAGTTTGAATAGGTAAAGTGAAGGAGAAGGAGGCACCTCAAGAAAGTCAG
CCACAGAACCAAGTGTGAGAAATAGGAAATGAGCTTTTTTAAGTTTTGCAA
AGGCAGATCAGGAGAGTTGACCTGTGGAGTGGAGAGTAGTCATAATTTTA
AAAAATGGCCATGGAAATTAAGAACTGATCAGAAATGGCTGGGCACAGTGG
CTCACGCCCTGTAATCCTAGCACTTTGGGAGGCCGAGTTGGGCAGATCACA
AGGTCAGGAGATAGAGACCGTCTGACTAACACGGTGAAACCCCGTCTCT
ATTAAAAATACAAAACCTTAGCTGGGCGTGGTGGTGGGCACCTGTAGTCC
CAGCTACTCGGAGGCTGAGGCAGGAGAAATGGCGTGAACCTGGGAGGCAG
AGCTTGCAGTGAAGCCAGATCGCCCACTGTGCTCAGCCTGGGCGAGAG

```

Or, upload file No file chosen [?](#)

Job Title [?](#)
Enter a descriptive title for your BLAST search [?](#)

☐ Align two or more sequences [?](#)

Choose Search Set

Database ☒ Standard databases (nr etc.): ☐ rRNA/ITS databases ☐ Genomic + transcript databases ☐ Betacoronavirus

Organism [Optional](#) [?](#)

Exclude [Optional](#) ☐ exclude [+](#)

Limit to [Optional](#) ☐ Models (XM/XP) ☐ Uncultured/environmental sample sequences

Entrez Query [Optional](#) [?](#)
Enter organism common name, binomial, or tax id. Only 20 top taxa will be shown [?](#)

☐ Sequences from type material

[?](#) [YouTube](#) [Create custom database](#)

Program Selection

Optimize for ☒ Highly similar sequences (megablast) ☐ More dissimilar sequences (discontiguous megablast) ☐ Somewhat similar sequences (blastn) [?](#)
Choose a BLAST algorithm [?](#)

BLAST Search **database** Nucleotide collection (nr/nt) using Megablast (Optimize for highly similar sequences) ☐ Show results in a new window

- 7 Look at the BLAST results. Give the:
1. Top 5 scores
 2. Top 5 species this gene is found in
 3. Lowest E-Value (if any)

[< Edit Search](#)
[Save Search](#)
[Search Summary](#)

[How to read this report?](#)
[BLAST Help Videos](#)
[Back to Traditional Results Page](#)

Job Title hg38_knownGene_ENST00000427411.1 range=chrX:15561033-
 RID [NG26BYXJ016](#) Search expires on 09-10 06:42 am [Download All](#)
Program BLASTN [Citation](#)
Database nt [See details](#)
Query ID lcl|Query_16745
 Description hg38_knownGene_ENST00000427411.1 range=chrX:1556103 ...
 Molecule type dna
 Query Length 41037
 Other reports [Distance tree of results](#) [MSA viewer](#)

Filter Results

Organism only top 20 will appear ☐ exclude

[+ Add organism](#)

Percent Identity to
E value to
Query Coverage to

[Filter](#)
[Reset](#)

Descriptions
Graphic Summary
Alignments
Taxonomy

Sequences producing significant alignments
[Download](#)
[Manage Columns](#)
Show

☒ select all 100 sequences selected

[GenBank](#)
[Graphics](#)
[Distance tree of results](#)

	Description	Max Score	Total Score	Query Cover	E value	Per. Ident	Accession
<input checked="" type="checkbox"/>	Homo sapiens angiotensin I converting enzyme 2 (ACE2), RefSeqGene on chromosome X	75782	80038	100%	0.0	100.00%	NG_012575.2
<input checked="" type="checkbox"/>	Homo sapiens X BAC RP11-478H11 (Roswell Park Cancer Institute Human BAC Library) complete sequence	75782	86411	100%	0.0	100.00%	AC097625.11
<input checked="" type="checkbox"/>	Homo sapiens Xp22 BAC GS-594A7 (Genome Systems Human BAC library) contains Bmx gene, complete sequence	75702	81772	100%	0.0	99.97%	AC003669.1
<input checked="" type="checkbox"/>	Homo sapiens angiotensin I converting enzyme (peptidyl-dipeptidase A) 2 (ACE2), gene, complete cds	75041	76758	99%	0.0	99.90%	AY217547.1
<input checked="" type="checkbox"/>	Pan troglodytes BAC clone CH251-585H11 from chromosome x, complete sequence	72687	81004	100%	0.0	98.65%	AC193855.3
<input checked="" type="checkbox"/>	Homo sapiens ACE2 promoter region (LOC117134593) on chromosome X	2326	2326	3%	0.0	100.00%	NG_068141.1