

Yale University

From the Selected Works of Rolando Garcia-Milian

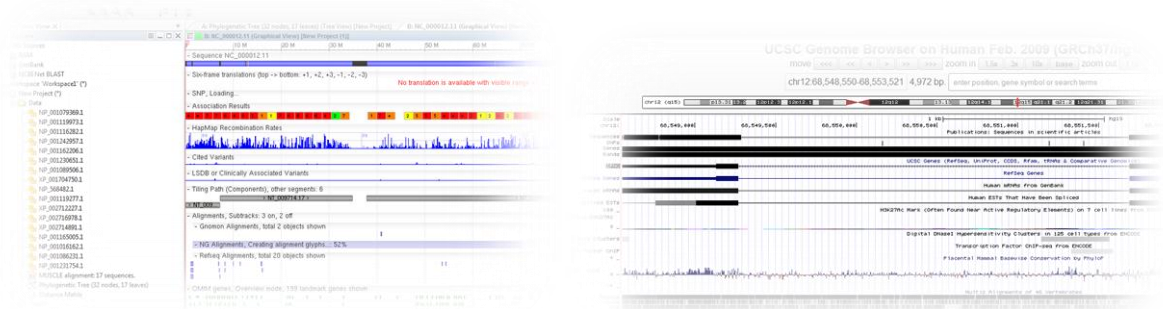
2014

Introduction to Genome Browsers

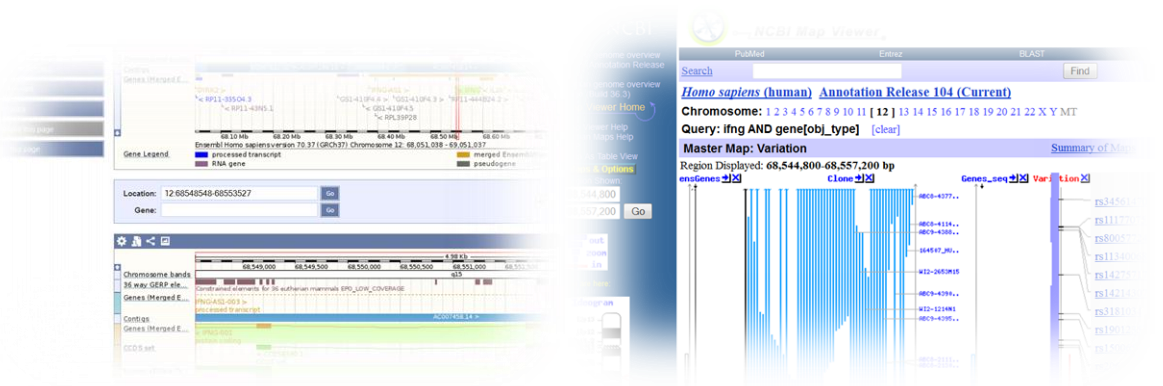
Rolando Garcia-Milian, *Yale University*



Available at: https://works.bepress.com/rolando_garciamilian/3/



Introduction to Genome Browsers



Rolando Garcia-Milian, MLS, AHIP

Rolando.milian@yale.edu



Curriculum and Research Support Department
Yale Medical Library

In this workshop we will learn how to navigate the UCSC Genome Browser, Ensembl, NCBI's Map Viewer and Genome Workbench. These browsers are valuable tools when identifying, localizing genes, and looking at their information in the genomic context. By using concrete examples, it will be shown how to locate a human gene, download a gene sequence and its upstream sequence, locate Single Nucleotide Polymorphisms (SNPs) and conserved regions, and use the browsers to download results in a batch. This handout contains screen shots of the demonstration taken on November, 2013

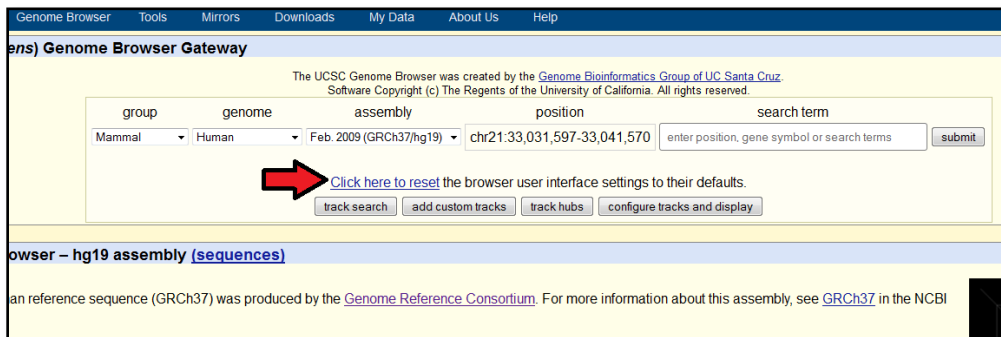
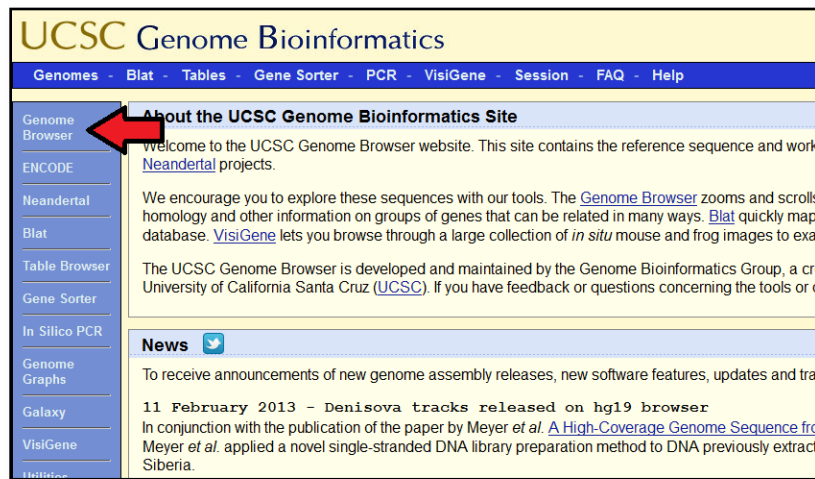
Contents

UCSC Genome Browser.....	3
Accessing the gene region and downloading its genomic sequence.....	3
Viewing SNPs annotated on the region	7
Downloading SNPs in the conserved region	8
Ensembl.....	12
Searching for the IFNG gene and retrieving the cDNA sequence with its amino acid translation.	12
Downloading upstream sequences from multiple human genes in a batch mode	16
Map Viewer.....	21
Identifying gene homologs in other organisms	21
NCBI's Genome Workbench.....	22
Retrieving a sequence from NCBI public database, loading it into genome workbench and displaying it graphically	22
Retrieving a sequence from NCBI and loading it into genome workbench. (if you DO NOT know the accession number)	24
Loading a genome.....	29
References	31

UCSC Genome Browser

Accessing the gene region and downloading its genomic sequence

Open the UCSC Genome Browser <http://genome.ucsc.edu/> in a web browser and click on “Genome Browser” link. Click on “Click here to reset” link to reset the browser user interface setting to their defaults. Under “assembly”, select the preferred assembly. We will use the default one “Feb. 2009 (GRCh37/hg19)”. Click on the “configure tracks and display” button. Under “Configure Tracks on UCSC Genome Browser” heading click on the “hide all” button and then, click on the “submit” button.



Genome Browser Gateway

The UCSC Genome Browser was created by the [Genome Bioinformatics Group of UC Santa Cruz](#).
Software Copyright (c) The Regents of the University of California. All rights reserved.

group	genome	assembly	position	search term
Mammal	Human	Feb. 2009 (GRCh37/hg19)	chr21:33,031,597-33,041,570	enter position, gene symbol or search terms

[track search](#)
[add custom tracks](#)
[track hubs](#)
[configure tracks and display](#)

ser – hg19 assembly (sequences)

reference sequence (GRCh37) was produced by the [Genome Reference Consortium](#). For more information about this assembly, see

Genome Browser Gateway

The UCSC Genome Browser was created by the [Genome Bioinformatics Group of UC Santa Cruz](#).
Software Copyright (c) The Regents of the University of California. All rights reserved.

group	genome	assembly	position	search term
Mammal	Human	Feb. 2009 (GRCh37/hg19)	chr21:33,031,597-33,041,570	enter position, gene symbol or search terms

[track search](#)
[add custom tracks](#)
[track hubs](#)
[configure tracks and display](#)

– hg19 assembly (sequences)

Configure Image

[submit](#)

image width: 1563 pixels
label area width: 17 characters
text size: 8

☒ Display chromosome ideogram above main graphic
☒ Show light blue vertical guidelines
☒ Display labels to the left of items in tracks
☒ Display description above each track
☒ Show track controls under main graphic
☐ Next/previous item navigation
☒ Next/previous exon navigation

Configure Tracks on UCSC Genome Browser: Human Feb. 2009 (GRCh37/hg19)

Tracks: [track search](#) [hide all](#) [show all](#) [default](#) Groups: [collapse all](#) [expand all](#)

Control track and group visibility more selectively below.

Mapping and Sequencing Tracks		
Base Position	dense	Chromosome position in bases. (Clicks here zoom in 3x)
Chromosome Band	hide	Chromosome Bands Localized by FISH Mapping Clones

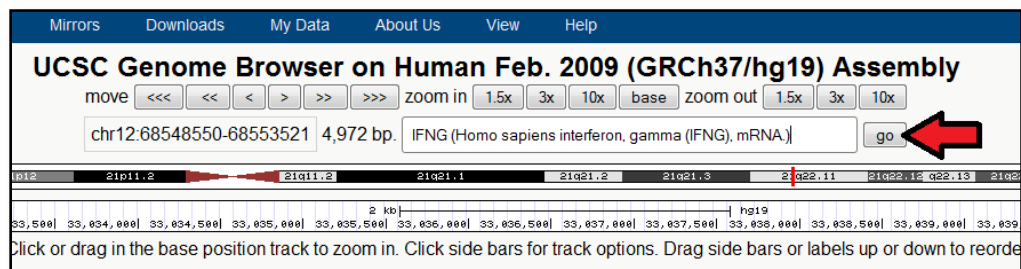
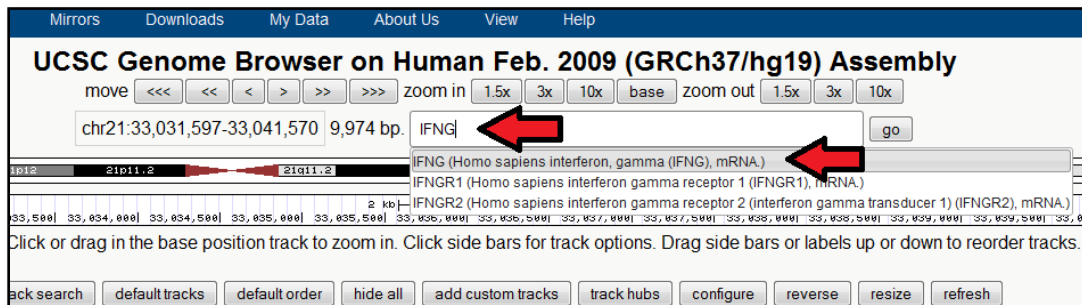
Configure Tracks on UCSC Genome Browser: Human Feb. 2009 (GRCh37/hg19)

Tracks: [track search](#) [hide all](#) [show all](#) [default](#) Groups: [collapse all](#) [expand all](#)

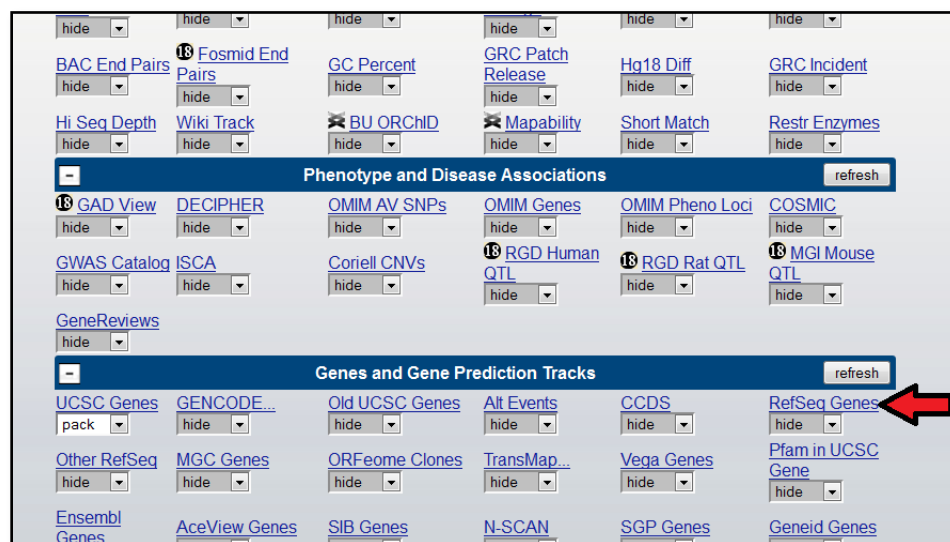
Control track and group visibility more selectively below.

Mapping and Sequencing Tracks		
Base Position	dense	Chromosome position in bases. (Clicks here zoom in 3x)
Chromosome Band	hide	Chromosome Bands Localized by FISH Mapping Clones
STS Markers	hide	STS Markers on Genetic (blue) and Radiation Hybrid (black) Maps
FISH Clones	hide	Clones Placed on Cytogenetic Map Using FISH
Recomb Rate	hide	Recombination Rate from deCODE, Marshfield, or Genethon Maps (deCODE default)

Enter your query in the search box (e.g. position, gene symbol or search term); in this example we will type IFNG (Homo sapiens interferon gamma). Click on the IFNG mRNA from the autocomplete feature.




The region corresponding to the IFNG gene on chromosome 12 (chr12: 68,548,550-68,553,521) is displayed. Under the “Genes and Gene Prediction Tracks”, click on the “RefSeq Genes” link.



From the “Display mode” select “pack” from the pull down menu, and change the “Color tracks by codons” to “genomic codons”. Check the box “Label” by gene symbol and click on the “Submit” button.

RefSeq Genes Track Settings

RefSeq Genes (▲All Genes and Gene Prediction Tracks)

Display mode: pack 

Label: ☒ gene ☐ accession ☐ OMIM ID

Hide non-coding genes: ☐

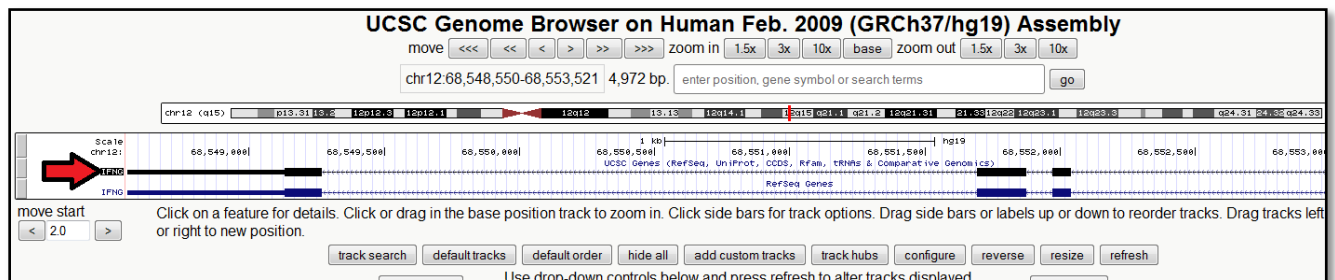
Color track by codons: genomic codons [Help on codon coloring](#)

Show codon numbering: ☐

[View table schema](#)

Data last updated: 2013-02-12

The gene is located on the reverse strand as indicated from the direction of arrows on the gene model. From the TrackMap click on the IFNG symbol label to the left of the UCSC Genes track.



Under “Sequence and Links to Tools and Databases” click on the “Genomic Sequence” Link to download the genomic sequence and/or the upstream and downstream sequence.

Sequence and Links to Tools and Databases					
Genomic Sequence (chr12:68,548,550-68,553,521)	mRNA (may differ from genome)	Protein (166 aa)			
Gene Sorter	Genome Browser	Protein FASTA	VisiGene	Table Schema	BioGPS
CGAP	Ensembl	Entrez Gene	ExonPrimer	GeneCards	GeneNetwork
Gepis Tissue	HGNC	HPRD	Jackson Lab	MOPED	OMIM
PubMed	Reactome	Stanford SOURCE	Treefam	UniProtKB	Wikipedia

Select your options for retrieving the sequence and click on “Submit” button.

Genomic Sequence Near Gene

Get Genomic Sequence Near Gene

Note: if you would prefer to get DNA for more than one feature of this track


Sequence Retrieval Region Options:

☐ Promoter/Upstream by 1000 bases
☒ 5' UTR Exons
☒ CDS Exons
☒ 3' UTR Exons
☒ Introns
☐ Downstream by 1000 bases
☒ One FASTA record per gene.
☐ One FASTA record per region (exon, intron, etc.) with 0 extra
☐ Split UTR and CDS parts of an exon into separate FASTA records

Note: if a feature is close to the beginning or end of a chromosome and

Sequence Formatting Options:

☒ Exons in upper case, everything else in lower case.
☐ CDS in upper case, UTR in lower case.
☐ All upper case.
☐ All lower case.
☐ Mask repeats: ☒ to lower case ☐ to N

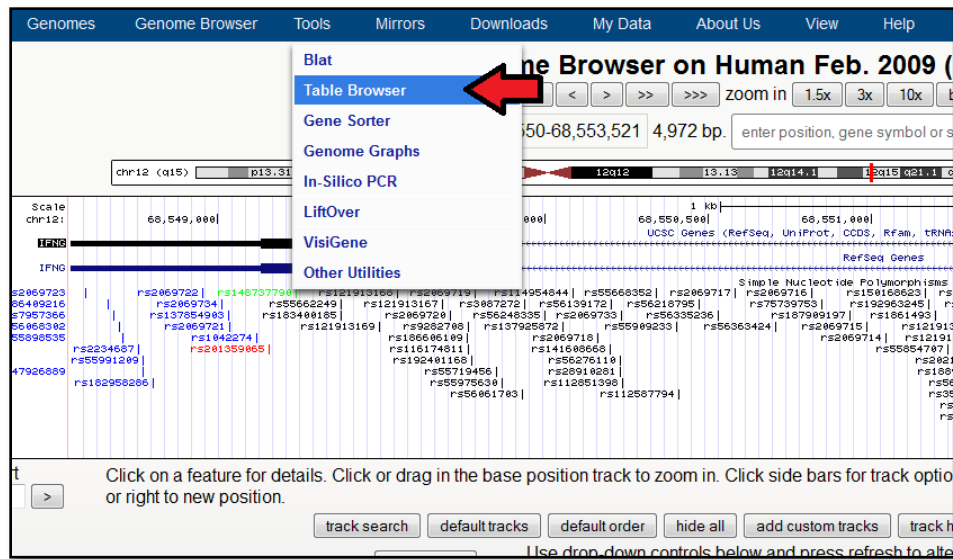
submit 

```

http://genome.ucsc.edu/cgi-bin/hgC?hgid=325792871&g=htcDnaNearGene&i=uc001stw.1&c=chr12&d=68548549&r=68553521&o=knownGene&b=
genome.ucsc.edu
Most Visited Getting Started Latest Headlines
>hg19_knownGene_uc001stw.1 range=chr12:68548550-68553521 5'pad=0 3'pad=0 strand=- repeatMasking=none
CACATTGTTCTGATCATCTGAAGATCAGCTATTAGAAAGAGAAATCAGT
TAAGTCCTTTGACCTGATCAGCTTGATCAAGAACTACTGATTCAACT
TCTTTGGCTTAATTTCTCGGAAACGATGAAATATACAAGTTATATCTTG
GCTTTTCAGCTCTGCATCGTTTTGGGTTCTTTGGCTGTTACTGCCAGGA
CCCATATGTAAGGAAGCAGAAACCTTAAGAAATATTTTgaagtatga
cttttaaatagtaactgtttgtgtgtgaaatgactgaatcgaacttgc
tgtagcatctctgtaggtgtctctctgttaggcagtcattttgagat
ttgtgtttattttgttaattattgactagatgagttccttgactaaataa
tctagatattgttttaacctctgtcagtttgtatagagacttaaaagg
gatttatgaattttccaaaagatgggcatatagggatgaagcataat
gatgttaaatattgtgtgtggaactcattcagttgtgtagtcaagga
gtatgcagattgaaaaaatgattgtttattgttttgacttctcagac
tcaaggtcaagattagcattaaaaagtttaaggaattgtttacaatt
aaagtcaaaaaggtccttaaaagcttgggttaaaaaaatcaactgatagg
gattttctccaaaagtgatttcaacattctgcttctctatctattac
ttgtgaagtattccgggaactctgtgtcactgggattttggaagaatta
tgattctggctaaaggaatttttaaaatttaagtgaatttttgagttt
cttttaaaattttattgattgttttaaggaattttacatttttaaatat
ttcattattgttttaaaacttagctgttataattatagctgtcataataa
tattcagacattcacaattgattttatttacaacacaaaatcaaatca
cacacacacacacacacacacacacacacacacacacacacacacac
ttaaagctogtataataataacccacaggaaggcacagtagatgtaata
aaacotgtacattggggggcagttattatagtggtgtgtgtgtgtgt
ttttgtttttgtatttttttagctagctgaaaaaactttcttttagctt
actatagtttttgggacotttgggagatcagctttgttgagctcatttgt
gacattgcaatttaagtgttatattgggaataaaaaagctaaaaagaaca
taatagtctttgtctatctcacataagccttttgggaactatttgt
tagaactaagcagaagagttgaaaaggaaatcagtgaaattgtgcacac
tgagttcaatgaaactgaaatatatttttaaggaattttatgggcaat
tgtaaaccaatttttttttttttttttttagAATGCAGGTCATTGAGAT
GTAGCGGATAATGGAACCTTTTCTTAGGCATTTGAAGAATTGGAAGA
  
```

Viewing SNPs annotated on the region

Under the track “Variation and Repeats”, click on the All SNPs(137) link. At the top of the “All SNPs(137) Track Setting”, change the display mode option to “pack”. Expand “Coloring Options” and under “SNP Feature for Color Specification”, change “Splice Site” from red to blue and click on the “Submit” button



Select the “Variation and Repeats tracks from “group”, and select the “position” radio button. Click on the “create” button next to “intersection”.

Table Browser

Use this program to retrieve the data associated with a track in text format, to calculate intersection, to calculate the overlap between two tracks, to calculate the overlap between two tracks, to calculate the overlap between two tracks. For help in using this application see [Using the Table Browser](#) for a description of the queries, and the OpenHelix Table Browser [tutorial](#) for a narrated presentation of the software. [Galaxy](#) or our [public MySQL server](#). To examine the biological function of your set through a [page](#) for the list of contributors and usage restrictions associated with these data. All tables [Downloads](#) page.

clade: Mammal genome: Human assembly: Feb. 2009 (GRCh37/hg19)

group: Variation **track:** Common SNPs(138) [add custom tracks](#)

table: snp138Common [describe table schema](#)

region: ☐ genome ☐ ENCODE Pilot regions ☒ position chr8:128750227-128751312 [look up](#)

identifiers (names/accessions): [paste list](#) [upload list](#)

filter: [create](#)

intersection: [create](#)

correlation: [create](#)

output format: all fields from selected table [Send output to](#) ☐ [Galaxy](#) ☐ [Google](#)

output file: (leave blank to keep output in browser)

file type returned: ☒ plain text ☐ gzip compressed

From “group”, select “Comparative Genomics”; and “Conservation” from “track”. Click on the “Submit” button.

Intersect with Common SNPs(137)

Select a group, track and table to intersect with:

group: Comparative Genomics **track:** Conservation

table: Primate Cons (phyloP46wayPrimates)

Intersect Common SNPs(137) items with bases covered by Conservation:

These combinations will maintain the names and gene/alignment structure (if any) of

- ☒ All Common SNPs(137) records that have any overlap with Conservation
- ☐ All Common SNPs(137) records that have no overlap with Conservation
- ☐ All Common SNPs(137) records that have at least 80 % overlap with Conserv
- ☐ All Common SNPs(137) records that have at most 80 % overlap with Conserv

Intersect bases covered by Common SNPs(137) and/or Conservation:

These combinations will discard the names and gene/alignment structure (if any) of C

- ☐ Base-pair-wise intersection (AND) of Common SNPs(137) and Conservation
- ☐ Base-pair-wise union (OR) of Common SNPs(137) and Conservation

Check the following boxes to complement one or both tables. To complement a table

- ☐ Complement Common SNPs(137) before base-pair-wise intersection/union
- ☐ Complement Conservation before base-pair-wise intersection/union

submit

From “output format”, select “BED – browser extensible data” and click on the “get output” button.

Table Browser

Use this program to retrieve the data associated with a track in text format. For more complex queries, you may want to use [Galaxy](#) or our [public MySQL](#) contributors and usage restrictions associated with these data. All tables

clade: Mammal **genome:** Human **assembly:**

group: Variation and Repeats **track:** Common SNPs

table: snp137Common **describe table schema**

region: ☐ genome ☐ ENCODE Pilot regions ☒ position chr12:68

identifiers (names/accessions): **paste list** **upload list**

filter: **create**

intersection with phyloP46wayPrimates: **edit** **clear**

correlation: **create**

output format: BED - browser extensible data **Send output to** ☐ **Get**

output file: (leave blank to keep output file name)

file type returned: ☒ plain text ☐ gzip compressed

Note: The all fields and selected fields output formats are not available

get output **statistics**

To reset all user cart settings (including custom tracks), [click here](#).

Click on “get BED” button. Your results will open in a new window.

Genomes Genome Browser Tools Mirrors

Output snp137Common as BED

☐ Include [custom track header](#):

name=

description=

visibility=

url=


Create one BED record per:

☒ Whole Gene

☐ Upstream by bases

☐ Downstream by bases

Note: if a feature is close to the beginning or end of a chromosome



Mozilla Firefox

File Edit View History Bookmarks Tools Help

http://genome.ucsc.edu/cgi-bin/hgTables?hgid=325821889&boolshad.hgta_printCustomTrackHeader

genome.ucsc.edu

Most Visited Getting Started Latest Headlines

chr12	68548593	68548594	rs2069723	0	-
chr12	68548952	68548953	rs2069722	0	-
chr12	68549009	68549010	rs2069734	0	-
chr12	68549785	68549786	rs2069719	0	-
chr12	68550161	68550162	rs2069718	0	-
chr12	68550231	68550231	rs56276110	0	+
chr12	68550232	68550232	rs2069733	0	-
chr12	68550580	68550581	rs2069717	0	-
chr12	68550814	68550815	rs2069716	0	-
chr12	68550985	68550986	rs2069715	0	-
chr12	68551043	68551044	rs2069714	0	-
chr12	68551195	68551196	rs1861493	0	+
chr12	68551408	68551409	rs1861494	0	+
chr12	68552475	68552476	rs2069713	0	-
chr12	68552521	68552522	rs2430561	0	+

Ensembl

Searching for the IFNG gene and retrieving the cDNA sequence with its amino acid translation.

Access the Ensembl genome browser page (<http://useast.ensembl.org/index.html>). Select “Human” genome under “Popular genomes”. Type IFNG in the search box and click on the “Go” button. Under “By Feature type”, click on “Gene” and “Human”. Click on the IFNG link or its Ensembl Gene ID link (ENSG00000111537).



e!Ensembl east BLAST/BLAT | BioMart | Tools | Downloads | Help & Documentation | Blog |

Human (GRCh37) ▾

Search Ensembl
 ▾ New Search

⚙️ Configure this page

📄 Add your data

📄 Export data

🔖 Bookmark this page

🔗 Share this page

Results Summary

Your search of Human with 'IFNG' returned the following results

By Feature type	
Total	54
▼ Gene	2
Human (2)	
▶ Marker	2
▶ Somatic mutation	4
▶ Transcript	4
▶ Variation	42

e!Ensembl east BLAST/BLAT | BioMart | Tools | Downloads | Help & Documentation | Blog | M

Human (GRCh37) ▾

Search Ensembl
 ▾ New Search

⚙️ Configure this page

📄 Add your data

📄 Export data

🔖 Bookmark this page

🔗 Share this page

Result in Detail

2 Genes match your query ('IFNG') in Human

IFNG

Description: interferon, gamma [Source:HGNC Symbol;Acc:5438]
 Gene ID: [ENSG00000111537](#)
 Location: [12:68548548-68553527:-1](#)
 Variations: [Variation Table](#)
 Source: e70

IFNG-AS1

Description: IFNG antisense RNA 1 [Source:HGNC Symbol;Acc:5439]
 Gene ID: [ENSG00000255733](#)
 Location: [12:68383225-68628466:1](#)

Click on the transcript ID link, it will take you the transcript tab (IFNG-001). Under the “Transcript-based displays” menu (left side of the screen), click on the cDNA link.

e!Ensembl east BLAST/BLAT | BioMart | Tools | Downloads | Help & Documentation | Blog | Mirrors

Human (GRCh37) ▾ Location: 12:68,548,548-68,553,527 Gene: IFNG Transcript: IFNG-001

Gene-based displays

- Gene summary
- Splice variants (1)
- Supporting evidence
- Sequence
- External references
- Regulation
- Comparative Genomics
 - Genomic alignments
 - Gene tree (image)
 - Gene tree (text)
 - Gene tree (alignment)
 - Gene gain/loss tree
- Orthologues (41)
- Paralogues
- Protein families (1)
- Phenotype
- Genetic Variation
 - Variation table

Gene: IFNG ENSG00000111537

Description: interferon, gamma [Source:HGNC Symbol;Acc:5438]
 Location: [Chromosome 12: 68,548,548-68,553,527](#) reverse strand.
 INSDC coordinates: chromosome:GRCh37:CM000674.1:68548548:68553527:1
 Transcripts: This gene has 1 transcript

Name	Transcript ID	Length (bp)	Protein ID	Length (aa)	Biotype	CCDS
IFNG-001	ENST00000229135	1218	ENSP00000229135	166	Protein coding	CCDS8980


Transcript and protein level displays

In Ensembl we provide displays at two levels:

Ensembl FAST BLAST/BLAT | BioMart | Tools | Downloads | Help & Documentation | Blog | Mirrors

Human (GRCh37) Location: 12:68,548,548-68,553,527 Gene: IFNG Transcript: IFNG-001

Transcript-based displays

- Transcript summary
- Supporting evidence (30)
 - Sequence
 - Exons (4)
 - cDNA (4) 
 - Protein
- External References
 - General identifiers (39)
 - Oligo probes (20)
- Ontology
 - Ontology graph (70)
 - Ontology table (70)
- Genetic Variation
 - Variation table
 - Variation image
 - Population comparison
 - Comparison image
- Protein Information

You've been redirected to your nearest mirror - useast.ensembl.org

Take me back to www.ensembl.org

Transcript: IFNG-001 ENST00000229135


Description: interferon, gamma [Source:HGNC Symbol;Acc:5438]

Location: [Chromosome 12: 68,548,548-68,553,527](#) reverse strand.

Gene: This transcript is a product of gene [ENSG00000111537](#) - This gene has 1 t

Name	Transcript ID	Length (bp)	Protein ID	Length (aa)	Biotype	CCDS
IFNG-001	ENST00000229135	1218	ENSP00000229135	166	Protein coding	CCDS8980

Use “Configure this page” link to change the display settings of the sequence. This will open a pop-up window. Once you have made changes to the settings, just close the window by clicking on the checkmark in the upper right corner.

Configure this page 

Domains & features (4)

- Variations (17)
- External data
- Personal annotation
- ID History
- Transcript history
- Protein history

Configure this page

Add your data


Export data

Bookmark this page

Share this page

Download view as RTF

BLAST this sequence

cDNA sequence 

Key

Codons: Alternating codons Alternating codons

Exons: Alternating exons Alternating exons

Variations: 3 prime UTR 5 prime UTR Missense Synonymous

Other features: UTR

1 GTGCAGCACATTGTTCTGATCATCTGAAGATCAGCTATTAGAAGAGAAAGATCAGTTAAG

61 TCCTTTGGACCTGATCAGCTTGATCAAGAACTACTGATTTCAACTTCTTTGGCTTAATT

121 CTCTCGGAACGATGAAATATCAAGTTATATCTTGGCTTTTCAGCTCTGCATCGTTTGG

181 GGTTCCTCTGGCTGTTACTGCGAGGACCCATAGTAAAGAGAGCAAAAACCTTAAGAAA

49 GGTTCCTCTGGCTGTTACTGCGAGGACCCATATGTAAGAGAGCAAAAACCTTAAGAAA

237 Location: 12:68,548,548-68,553,527 Gene: IFNG Transcript: IFNG-001

Configure Page **Manage Configurations** **Custom Data**

Display options

Save as:

Load configuration:

Reset configuration:

Display options

Show exons: Yes

Show codons: Yes

Show UTR: Yes

Show coding sequence: Yes

Show protein sequence: Yes

Show RNA features: Yes

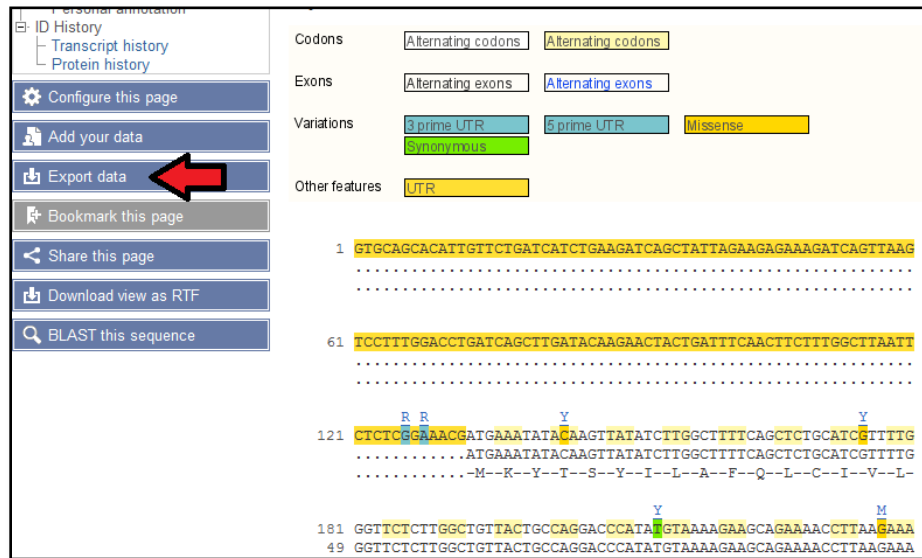
Show variations: Yes

Filter variations by consequence type: No filter

Line numbering: Yes



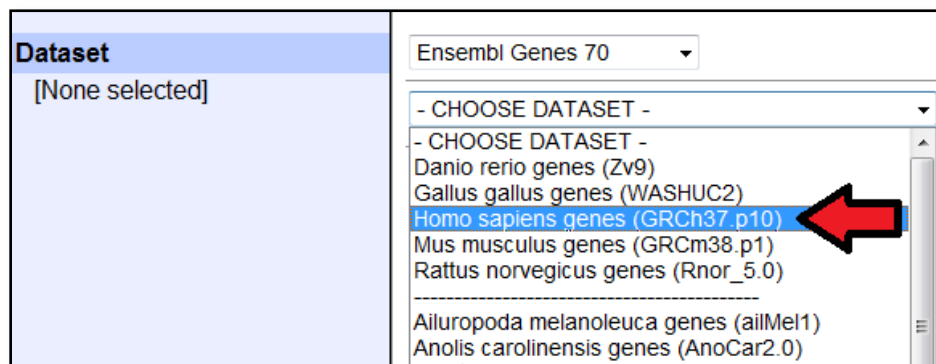
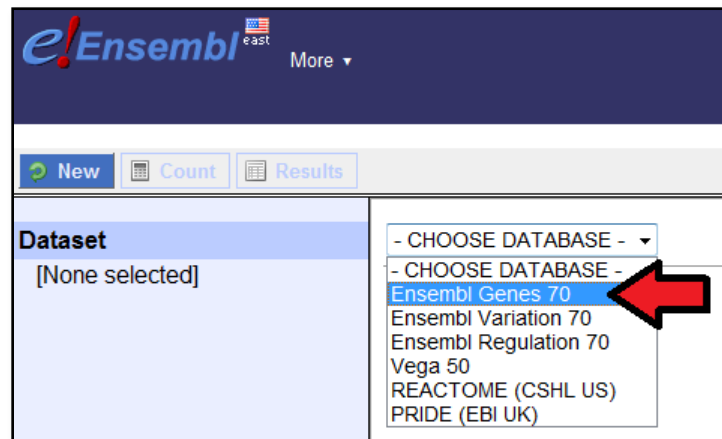
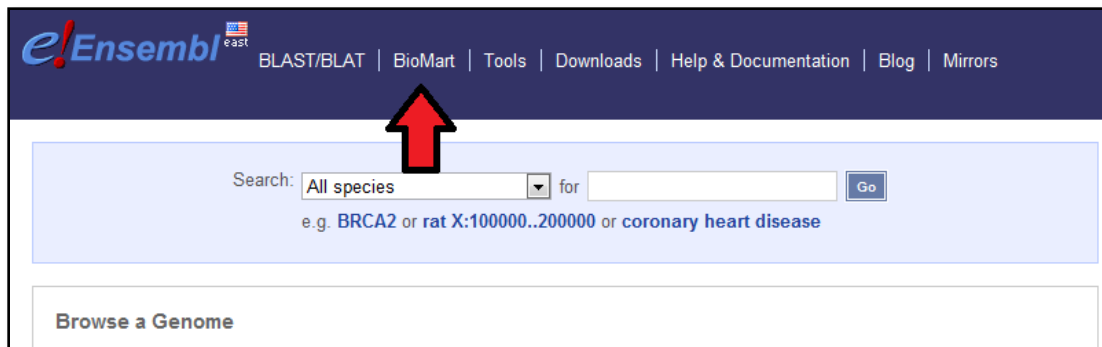
Data can be exported by clicking on the “Export data” link. This opens a pop up window that allows you to change the configuration of the data you will export. Click on the “Next” button to apply your changes. Select your preferred output format for your export. In this example we will click on the HTML link.



The screenshot shows the main interface with a sidebar on the left containing several buttons. A red arrow points to the 'Export data' button. The main area displays sequence information with various features highlighted in different colors: Alternating codons (yellow), Alternating exons (blue), 3' prime UTR (light blue), 5' prime UTR (light blue), Missense (yellow), Synonymous (green), and UTR (yellow). The sequence is shown in a multi-line format with line numbers 1, 61, 121, 181, and 49.



The screenshot shows the 'Export data' configuration window. The 'Export Configuration - Feature List' section is visible. The 'Transcript to export' is set to 'ENST00000229135 (IFNG-001)'. The 'Output' is set to 'FASTA sequence'. The 'Strand' is set to 'Feature strand'. The '5' Flanking sequence (upstream)' and '3' Flanking sequence (downstream)' are both set to '0'. The 'Next >' button is highlighted with a red arrow. Below the configuration section, the 'Options for FASTA sequence' section is visible, with 'Genomic' set to 'None', 'Select/deselect all:' checked, and 'cDNA:' checked.



Click on the “Filters” menu. Open the GENE criteria by clicking on the + sign next to it. Check the “ID list limit [Max 500]” box and select the ID type from the dropdown menu. Type the gene symbols of your query in the box below. In this example we will use IFNG, HRAS, BRCA1, RB1. Click on the “Count” button (top of left side) to see how many genes BioMart can find.

e!Ensembl east [More](#) ▾

[New](#) [Count](#) [Results](#)

Dataset	Ensembl Genes 70 ▾
Homo sapiens genes (GRCh37.p10)	Homo sapiens genes (GRCh37.p10) ▾
Filters	
[None selected]	
Attributes	
Ensembl Gene ID	
Ensembl Transcript ID	

[New](#) [Count](#) [Results](#) [URL](#) [XML](#) [Perl](#) [Help](#)

Dataset

Homo sapiens genes (GRCh37.p10)

Filters

[None selected]

Attributes

Ensembl Gene ID

Ensembl Transcript ID

Please restrict your query using criteria below

☐ REGION:

☐ GENE:

☐ TRANSCRIPT EVENT:

☐ GENE ONTOLOGY:

☐ EXPRESSION:

[New](#) [Count](#) [Results](#) [URL](#) [XML](#) [Perl](#) [Help](#)

Dataset **4**

Homo sapiens genes (GRCh37.p10)

Filters

HGNC symbol(s) [e.g. ZFY]: [ID-list specified]

Attributes

Ensembl Gene ID

Ensembl Transcript ID

Please restrict your query using criteria below

☐ REGION:

☐ GENE:

☐ Limit to genes ...

☒ ID list limit [Max 500] **1**

☐ with ArrayExpress ID(s)

2 HGNC symbol(s) [e.g. ZFY] **3**

IFNG
HRAS
BRCA1
RB1

[Browse...](#)

[New](#) [Count](#) [Results](#) [URL](#) [XML](#) [Perl](#) [Help](#)

Dataset 4 / 62380 Genes Dataset 4 / 62380 Genes

Homo sapiens genes (GRCh37.p10)

Filters

HGNC symbol(s) [e.g. ZFY]: [ID-list specified]

Attributes

Ensembl Gene ID

Ensembl Transcript ID

Please restrict your query using criteria below

☐ REGION:

☐ GENE:

☐ Limit to genes ...

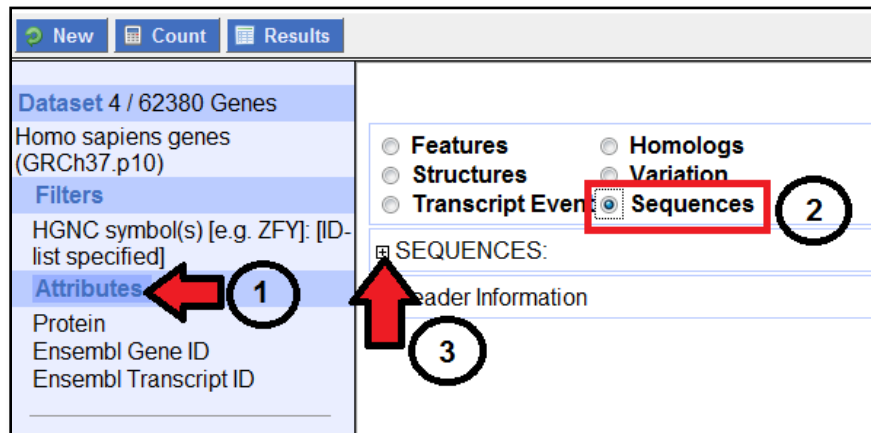
☒ ID list limit [Max 500]

☐ with ArrayExpress ID(s)

HGNC symbol(s)

IFNG
HRAS
BRCA1
RB1

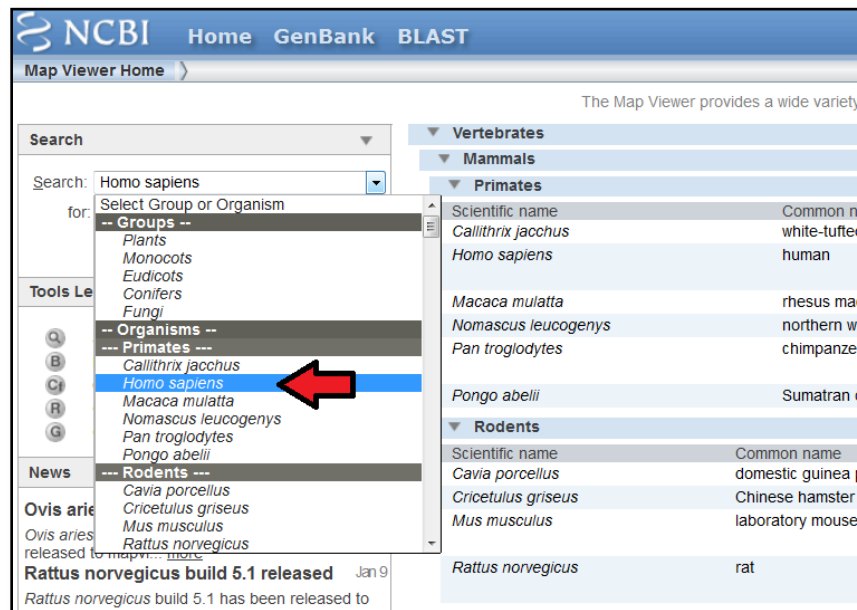
Click on the “Attributes” option and check the “SEQUENCES” radio button. Open the “SEQUENCES” menu by clicking on the + sign. Check the “Flank (Gene)” radio button. Select the “Upstream flank” and enter the number of nucleotides desired (1000 in this example). Open the “Header information” menu by clicking on the + sign. Select your preferred option such as “Associated Gene Name”, “Description”, etc.



Map Viewer

Identifying gene homologs in other organisms

Open NCBI's MapViewer on your internet browser (<http://www.ncbi.nlm.nih.gov/mapview/>). Select *Homo sapiens* from "Select Group or Organism" drop-down menu. Type "IFNG" in the search box and click on the "Go" button. Select "Gene" from the "Quick Filter" and click on the "Filter" button. Results will show the "reference" assembly.



Match	Map element	Type	Maps
all matches			
H.sapiens mRNA for IFN-gamma (pKC-0).	X62468.1	TRANSCRIPT	Hs RNA
Human mRNA for HuIFN-gamma interferon	X01992.1	TRANSCRIPT	Hs RNA
H.sapiens mRNA for IFN-gamma (pKC-10L)	X62469.1	TRANSCRIPT	Hs RNA
H.sapiens mRNA for IFN-gamma (pKC-11)	X62473.1	TRANSCRIPT	Hs RNA
H.sapiens mRNA for IFN-gamma (pKC-11L).	X62474.1	TRANSCRIPT	Hs RNA
H.sapiens mRNA for IFN-gamma (pKC-14L).	X62472.1	TRANSCRIPT	Hs RNA
H.sapiens mRNA for IFN-gamma (pKC-19L)	X62471.1	TRANSCRIPT	Hs RNA
Human mRNA for interferon IFN-gamma	X13274.1	TRANSCRIPT	Hs RNA

Quick Filter

☒ Gene

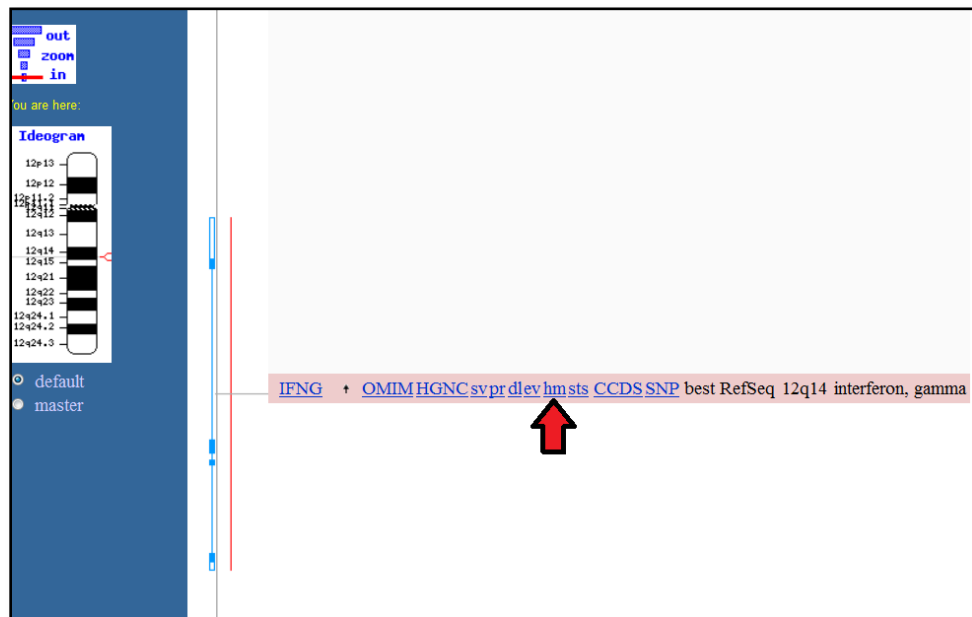
Transcript : ☐ all ☐ RefSeq ☐ STS

Filter

Under "Maps", click on the "Genes_seq" link from the "reference" assembly.

Chr	Assembly	Match	Map element	Type	Maps
12	reference	all matches			
		IFNG : interferon, gamma	IFNG	GENE	Genes_cyto Genes_seq
		IFNG	IFNG	GENE	ensGenes
12	CHM1_1.0-Primary Assembly	IFNG : interferon, gamma	IFNG	GENE	Genes_seq
12	HuRef-Primary Assembly	IFNG : interferon, gamma	IFNG	GENE	Genes_seq

Click on the “hm” link to access the IFGN “HomoloGene” page.



1: HomoloGene:55526. Gene conserved in Eutheria	
Genes <i>Genes identified as putative homologs of one another during the construction of HomoloGene.</i>	Proteins <i>Proteins used in sequence comparisons and their conserved domain architectures.</i>
IFNG, <i>H.sapiens</i> interferon, gamma	NP_000610.2 166 aa
IFNG, <i>P.troglodytes</i> interferon, gamma	NP_001180594.1 166 aa
IFNG, <i>M.mulatta</i> interferon, gamma	NP_001028077.1 165 aa
IFNG, <i>C.lupus</i> interferon gamma	NP_001003174.1 166 aa
IFNG, <i>B.taurus</i> interferon, gamma	NP_776511.1 166 aa
Ifng, <i>M.musculus</i> interferon gamma	NP_032363.1 155 aa
Ifng, <i>R.norvegicus</i> interferon gamma	NP_620235.1 156 aa

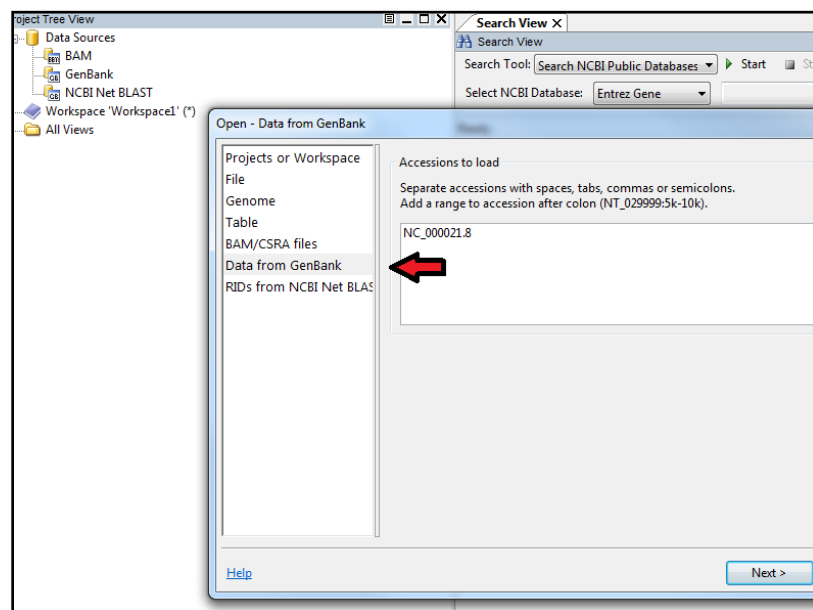
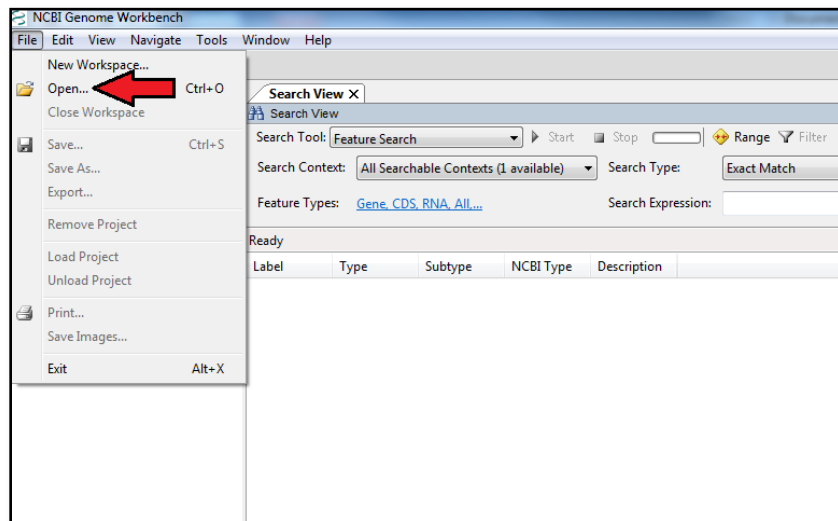
NCBI's Genome Workbench

Retrieving a sequence from NCBI public database, loading it into genome workbench and displaying it graphically

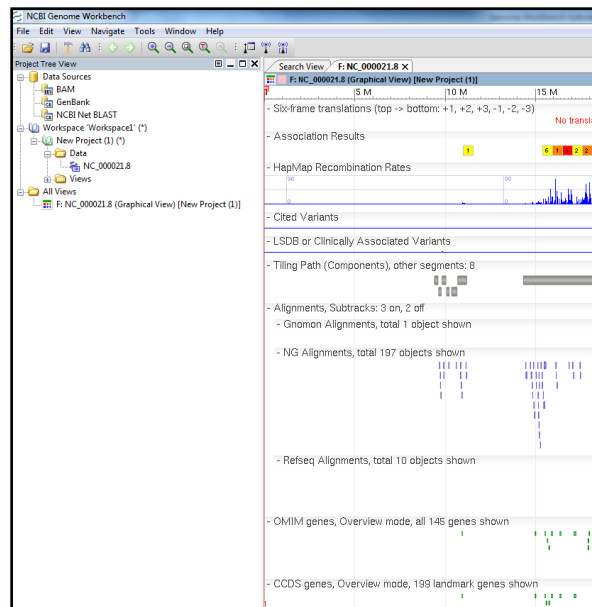
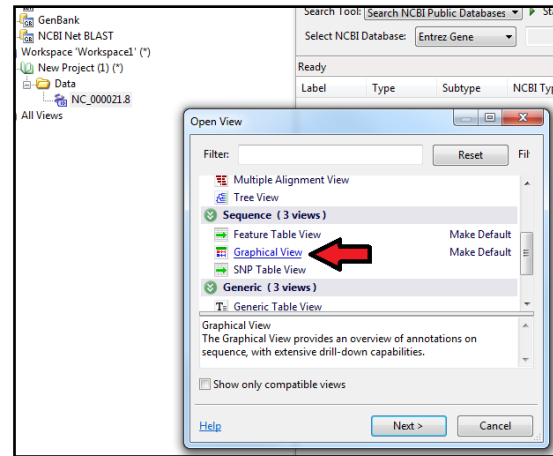
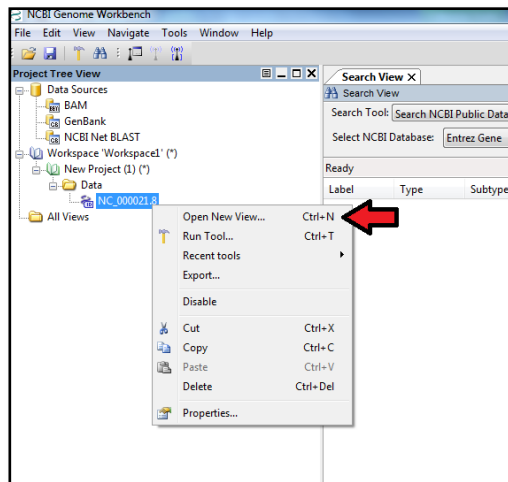
Download and install the Genome Workbench from

<http://www.ncbi.nlm.nih.gov/tools/gbench/> . Open the Genome Workbench and click on File ->

Open -> Data from GenBank. Enter your accession number in the box, in this example: NC_000021.8. Click on “Next” and “Finish”.

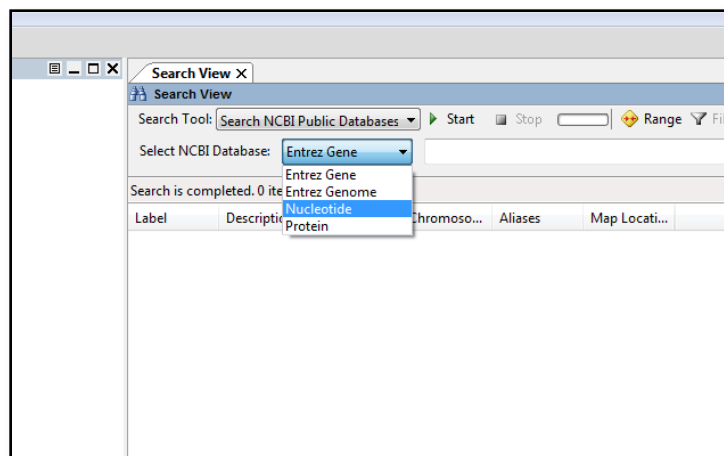
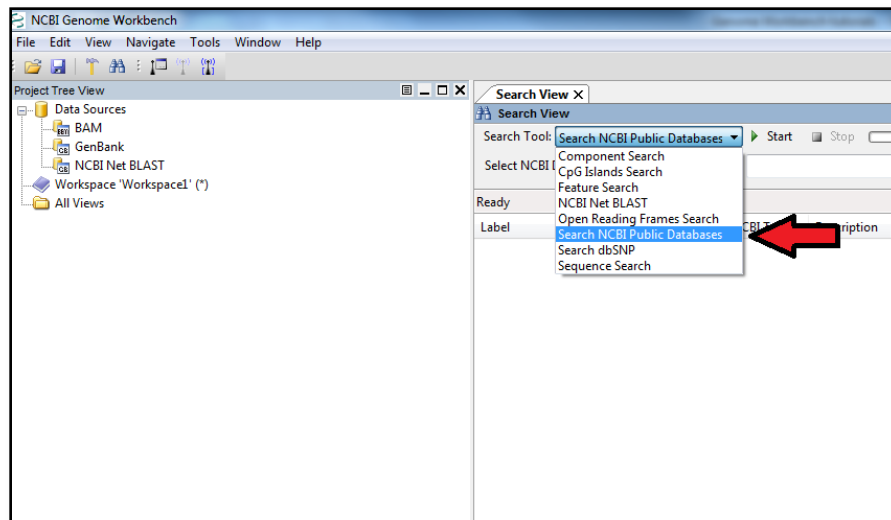


Select your data from the “Data” folder and right-click, and then select “Open New View”. Click on “Graphical View” in the pop-up window. A new window should appear showing the available tracks for your data.



Retrieving a sequence from NCBI and loading it into genome workbench. (if you DO NOT know the accession number)

On the “Search View” window, select “Search NCBI Public Databases” from the “Search Tools” drop-down menu. Select one of the NCBI databases from the “Select NCBI Database” drop-down menu; in this example, we will use the “Nucleotide” database.



Enter your search query. You can make sophisticated Entrez queries in order to obtain a set of sequences through the Genome Workbench. In this example: *tp53[ti] AND human[orgn] AND "refseq"[Filter]*.

In the search results, double-click on your sequence (e.g. NG_017013), decide the project to place the new item (e.g. add to an existing project, create a new project, decide later) and click on the "OK" button.

Search View

Search Tool: Search NCBI Public Databases Start Stop Range Filter Form

Select NCBI Database: Nucleotide tp53[ti] AND human[orgn] AND *refseq[Filter]

Search is completed. 36 items found.

Label	Description	FASTA IDs
NM_001126117	Homo sapiens tumor protein p53 (TP53), transcript variant 7, mRNA	gi 187830908 refseq NM_001126117.1
NM_018081	Homo sapiens WD repeat containing, antisense to TP53 (WRAP53), transcript variant 1, mRNA	gi 221136853 refseq NM_018081.1
NM_001143990	Homo sapiens WD repeat containing, antisense to TP53 (WRAP53), transcript variant 2, mRNA	gi 221136857 refseq NM_001143990.1
NM_001143991	Homo sapiens WD repeat containing, antisense to TP53 (WRAP53), transcript variant 3, mRNA	gi 221136861 refseq NM_001143991.1
NM_001143992	Homo sapiens WD repeat containing, antisense to TP53 (WRAP53), transcript variant 4, mRNA	gi 221136865 refseq NM_001143992.1
NG_009456	Homo sapiens PERP, TP53 apoptosis effector pseudogene (LOC100131159) on chromosome 6	gi 222080051 refseq NG_009456.1
NG_009455	Homo sapiens PERP, TP53 apoptosis effector pseudogene (LOC100129725) on chromosome 5	gi 222080051 refseq NG_009455.1
NG_022121	Homo sapiens PERP, TP53 apoptosis effector (PERP), mRNA	gi 222080101 refseq NG_022121.1
NG_009454	Homo sapiens PERP, TP53 apoptosis effector pseudogene (LOC400446) on chromosome 15	gi 222418691 refseq NG_009454.1
NM_015713	Homo sapiens ribonucleotide reductase M2 B (TP53 inducible) (RRM2B), transcript variant 1, ...	gi 289177072 refseq NM_015713.1
NM_001172477	Homo sapiens ribonucleotide reductase M2 B (TP53 inducible) (RRM2B), transcript variant 2, ...	gi 289177073 refseq NM_001172477.1
NM_001172478	Homo sapiens ribonucleotide reductase M2 B (TP53 inducible) (RRM2B), transcript variant 3, ...	gi 289177075 refseq NM_001172478.1
NG_016617	Homo sapiens ribonucleotide reductase M2 B (TP53 inducible) (RRM2B), RefSeqGene on chr...	gi 289802980 refseq NG_016617.1
NG_027980	Homo sapiens TP53 regulated inhibitor of apoptosis 1 pseudogene 1 (TRIAPI1) on chromos...	gi 313760610 refseq NG_027980.1
NG_028245	Homo sapiens WD repeat containing, antisense to TP53 (WRAP53), RefSeqGene (LRG_375) on...	gi 326807026 refseq NG_028245.1
NM_001205259	Homo sapiens TP53 target 3C (TP53TG3C), transcript variant 1, mRNA	gi 327532729 refseq NM_001205259.1
NM_001205260	Homo sapiens TP53 target 3C (TP53TG3C), transcript variant 2, mRNA	gi 327532732 refseq NM_001205260.1
NM_001099687	Homo sapiens TP53 target 3B (TP53TG3B), transcript variant 1, mRNA	gi 327532754 refseq NM_001099687.1
NM_001205265	Homo sapiens TP53 target 3B (TP53TG3B), transcript variant 2, mRNA	gi 327532757 refseq NM_001205265.1
NM_001243722	Homo sapiens TP53 target 3D (TP53TG3D), mRNA	gi 343962649 refseq NM_001243722.1
NM_000546	Homo sapiens tumor protein p53 (TP53), transcript variant 1, mRNA	gi 371502114 refseq NM_000546.1
NM_001126112	Homo sapiens tumor protein p53 (TP53), transcript variant 2, mRNA	gi 371502115 refseq NM_001126112.1
NM_001126113	Homo sapiens tumor protein p53 (TP53), transcript variant 4, mRNA	gi 371502116 refseq NM_001126113.1
NM_001126114	Homo sapiens tumor protein p53 (TP53), transcript variant 3, mRNA	gi 371502117 refseq NM_001126114.1
NM_001126118	Homo sapiens tumor protein p53 (TP53), transcript variant 8, mRNA	gi 371502118 refseq NM_001126118.1
NG_017013	Homo sapiens tumor protein p53 (TP53), RefSeqGene on chromosome 17	gi 383209646 refseq NG_017013.1
NM_016399	Homo sapiens TP53 regulated inhibitor of apoptosis 1 (TRIAPI1), mRNA	gi 38679910 refseq NM_016399.1
NM_016212	Homo sapiens TP53 target 3 (TP53TG3), transcript variant 1, mRNA	gi 392513673 refseq NM_016212.1
NM_001267813	Homo sapiens TP53 target 3 (TP53TG3), transcript variant 2, mRNA	gi 392513676 refseq NM_001267813.1

Add to Project

Project items to be created

Label	Type	Subtype	NCBI Type	Description
NG_017013	Sequence ID	DNA	Seq-id	Homo sapi...

Project to place new items

☒ Create a new Project

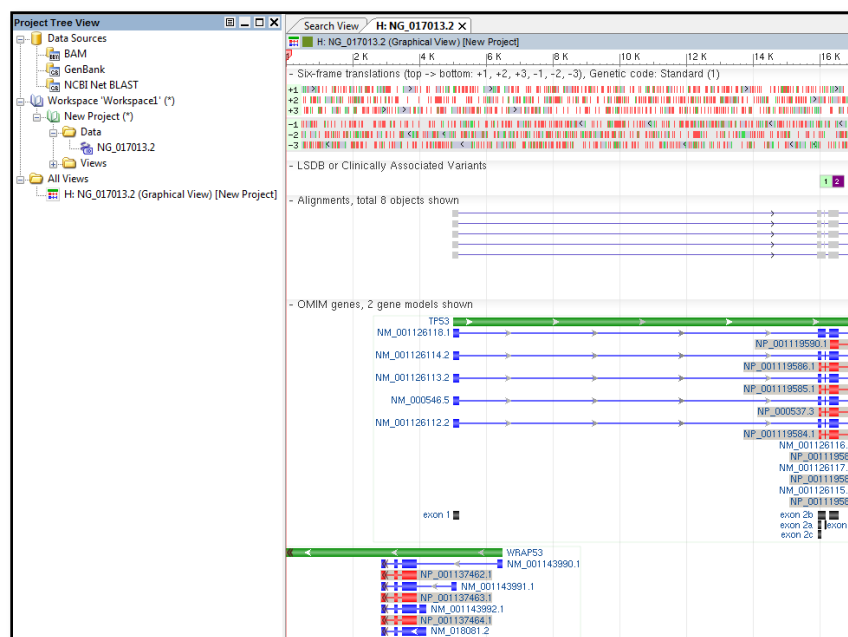
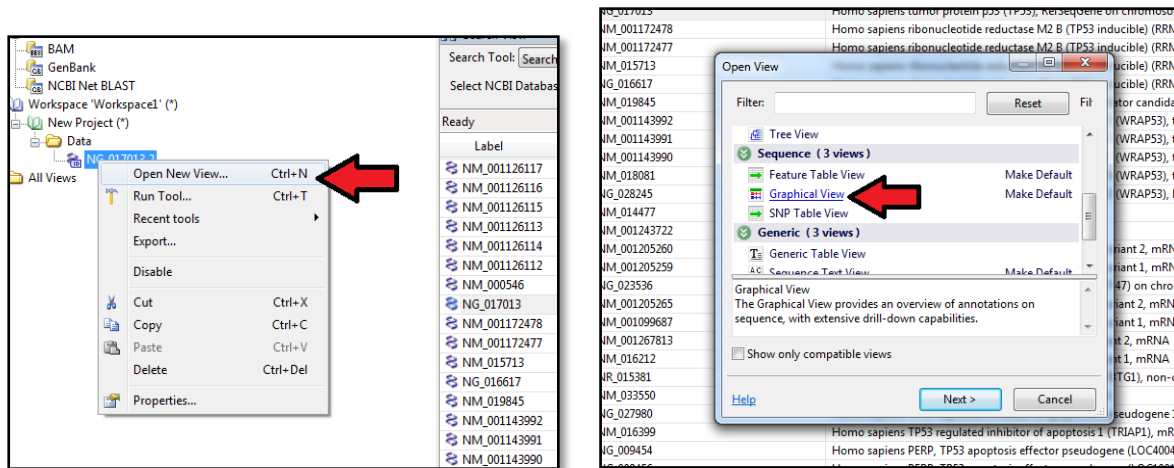
☐ Create a separate Project for every item

☐ Decide Later

☐ Add to an existing Project

OK Cancel

Select your data from the “Data” folder and right-click, and then select “Open New View”. Click on “Graphical View” in the new window. A new view window should appear showing the available tracks for your data.

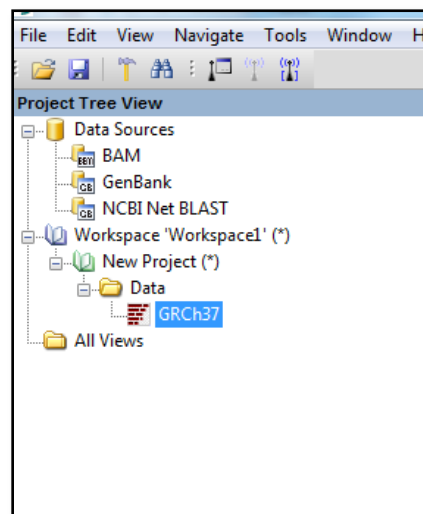
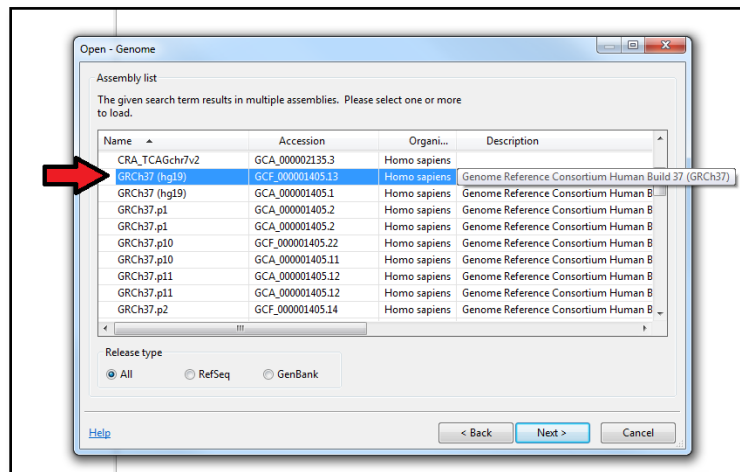


You can also import your data into the Genome Browser.

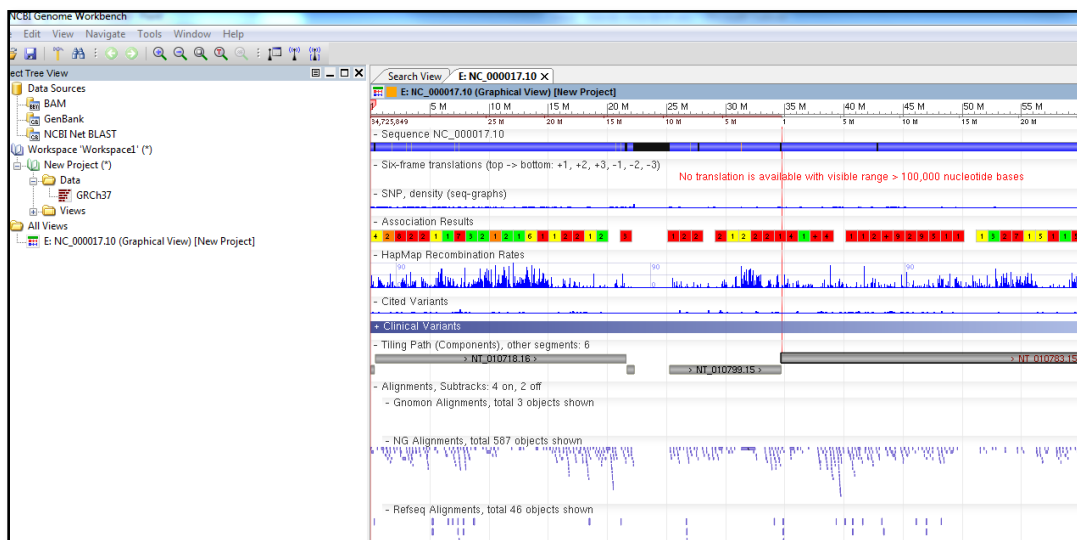
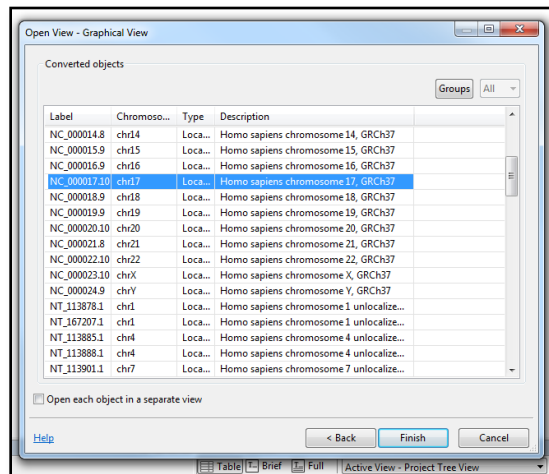
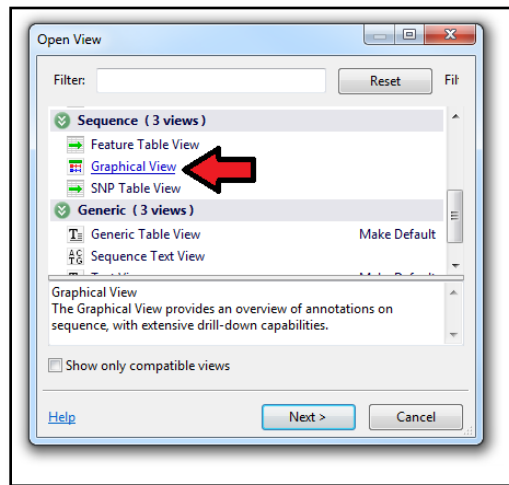
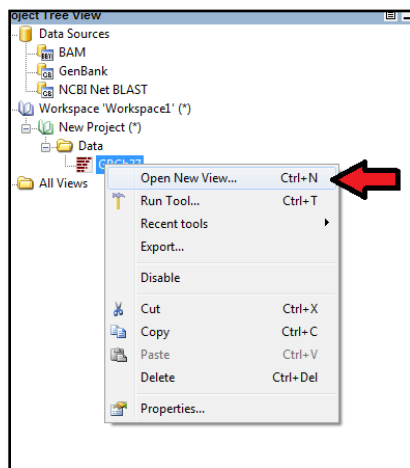
Click on File -> Open. Select "File" from the "Open- File" pop-up window, and then select the File Format, in this example FASTA files. Select and open the file of interest. Click on "Next"

Loading a genome

Go to File -> Open. Select “Genome” and type your search term in the search box (e.g. type “human”). Click on “Next” and select one or more assemblies from the table. Select “GRCh37 (hg19)” (Genome Reference Consortium Built 37) and Click on “Next”. Select “Create a new Project” and Click on “Finish”. The human genome build should be visible under a New Project.



Once the genome is loaded, you can use the Genome workbench tools with it. In this example we will visualize Chromosome 17. Right-click on the GRCh37 and select “Open New View”. Select “Graphical View”. A new window will open; select the chromosome you would like to view (NC_000017.10) and click “Finish”



References

Ensembl tutorials and worked examples.

<http://www.ensembl.org/info/website/tutorials/index.html>

Fernandez-Suarez XM, Schuster MK (2010) Unit 1.15 Using the Ensembl Genome Server to Browse Genomic Sequence Data

<http://onlinelibrary.wiley.com/doi/10.1002/0471250953.bi0115s30/pdf>

Furey, TS. (2006) Comparison of human (and other) genome browsers. Human Genomics 2(4): 266–270.

<http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2834533/>

Karolchik D, Hinrichs AS, Kent J (2012) Unit 1.4 The UCSC genome browser. Current Protocols in Bioinformatics.

<http://onlinelibrary.wiley.com/doi/10.1002/0471250953.bi0104s40/abstract>

Karolchik D, Hinrichs AS, Kent J (2011) The UCSC Genome Browser. Unit 18.6 The UCSC genome browser. Current Protocols in Human Genetics.

NCBI, Entrez Map Viewer help document

<http://www.ncbi.nlm.nih.gov/projects/mapview/static/MapViewHelp.html>

NCBI, Genome Workbench basic operation

<http://www.ncbi.nlm.nih.gov/tools/gbench/tutorial1/>

UCSC genome browser user guide

<http://genome.ucsc.edu/goldenPath/help/hgTracksHelp.html>

Human Interferon gamma P01579 (IFNG_HUMAN) Reviewed, UniProtKB/Swiss-Prot

<http://www.uniprot.org/uniprot/P01579>



Curriculum and Research Support Department
Yale Medical Library

