Tutorial 3: Bioinformatics for Physiology

Finding Protein Sequence(s) of Your Interest from Public Databases

Kenta Nakai and Riu Yamashita

University of Tokyo

In The Post-Genomics Era,

- Classical Repository databases, such as GenBank, may be obsolete.. (Use RefSeq)
- Results of many useful programs, such as BLAST, are already presented in DBs
- Many databases are inter-linked and so it may not be important where to start
- Genome Browsers can be a good portal site

Genome Browsers

- NCBI
  - http://www.ncbi.nlm.nih.gov/Genomes/
- Ensembl
  - http://www.ensembl.org/index.html
- UCSC Genome browser
  - http://genome.ucsc.edu/

UCSC Genome Bioinformatics

UCSC Genome Browser

http://genome.ucsc.edu/

UCSC Genome Bioinformatics

Genomes - Blat - Tables - Gene Sorter - PCR - VisiGene - Proteome - Session - FAQ - Help

Genome Browser
ENCODE
Blat
Table Browser
Gene Sorter
In Silico PCR
Genome Graphs
Galaxy
VisiGene
Proteome Browser
Utilities
Downloads
Release Log

About the UCSC Genome Bioinformatics Site

Welcome to the UCSC Genome Browser website. This site contains the reference sequence and working draft assemblies for a large collection of genomes. It also provides a portal to the ENCODE project.

We encourage you to explore these sequences with our tools. The Genome Browser zooms and scrolls over chromosomes, showing the work of annotators worldwide. The Gene Sorter shows expression, homology and other information on groups of genes that can be related in many ways. Blat quickly maps your sequence to the genome. The Table Browser provides convenient access to the underlying database. VisiGene lets you browse through a large collection of *in situ* mouse and frog images to examine expression patterns. Genome Graphs allows you to upload and display genome-wide data sets.

The UCSC Genome Browser is developed and maintained by the Genome Bioinformatics Group, a cross-departmental team within the Center for Biomolecular Science and Engineering (CBSE) at the University of California Santa Cruz (UCSC). If you have feedback or questions concerning the tools or data on this website, feel free to contact us on our public mailing list. To view the results of the Genome Browser users' survey we conducted in May 2007, click here.

News

News Archives ▶

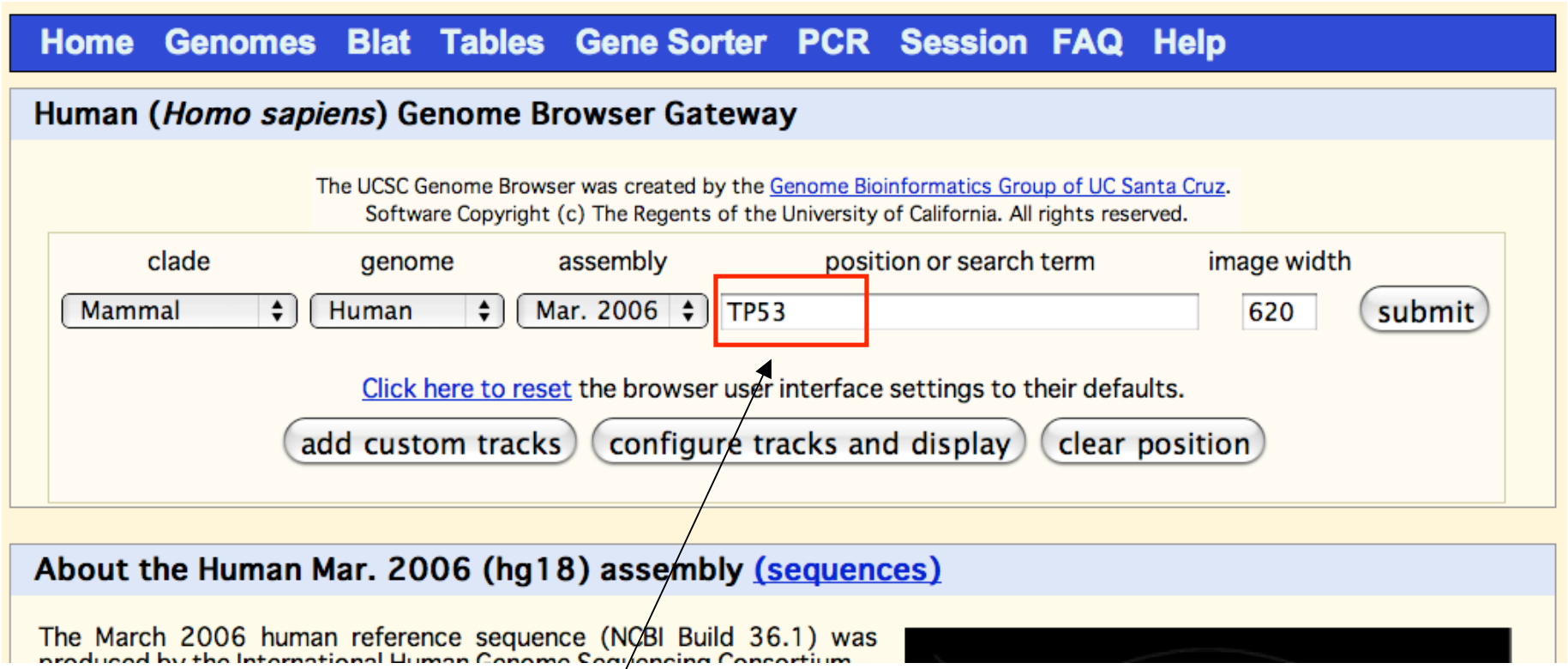
To receive announcements of new genome assembly releases, new software features, updates and training seminars by email, subscribe to the genome-announce mailing list.

23 Apr. 2008 - Marmoset Browser Released

We'd like to announce the release of a Genome Browser and Blat server for the marmoset genome (*Callithrix jacchus*). The June 2007 assembly -- WUSTL version Callithrix jacchus-2.0.2, UCSC version calJac1 -- was produced by Washington University St. Louis (WUSTL) School of Medicine Genome Sequencing Center in St. Louis, MO, USA.

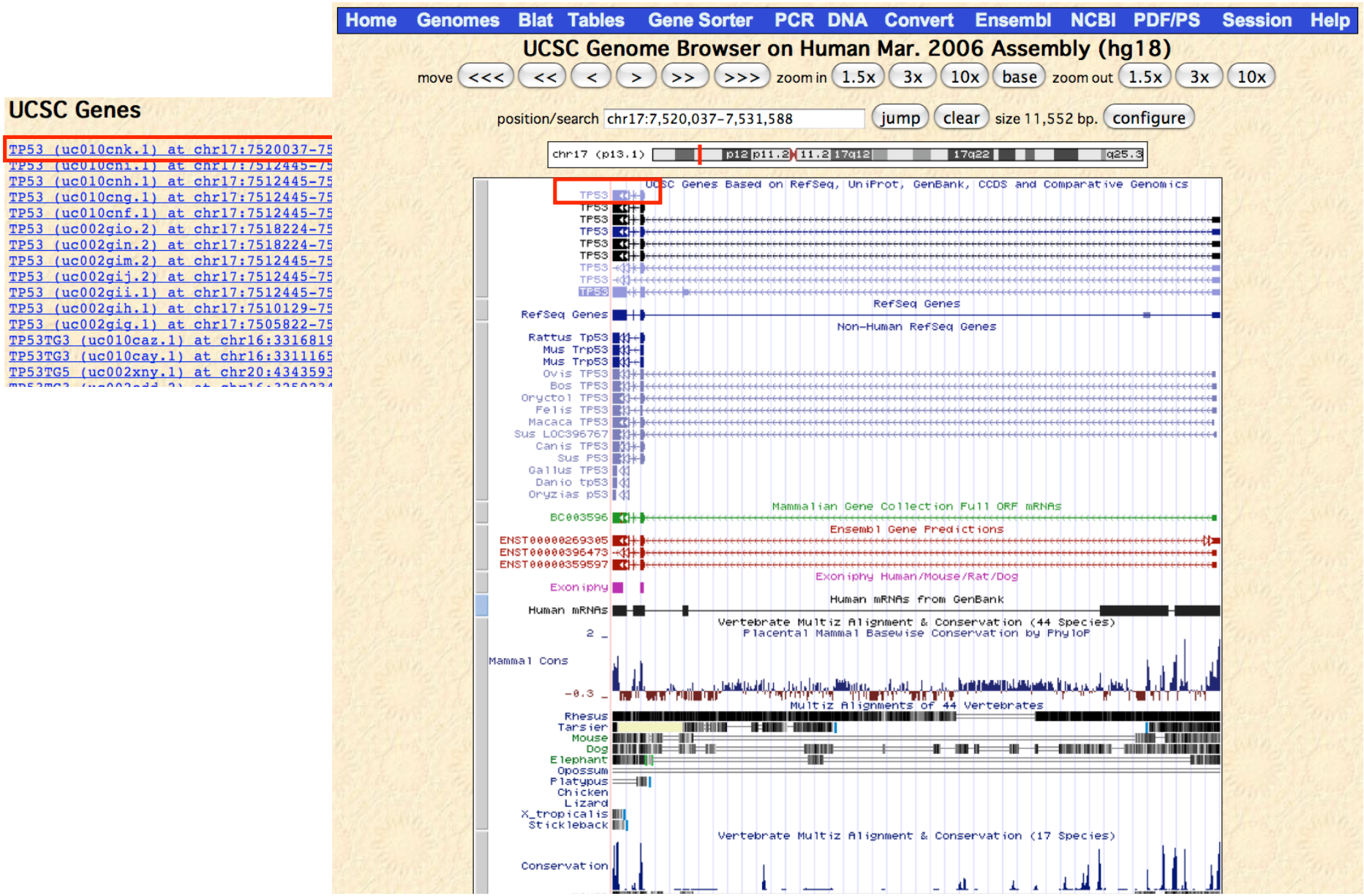
C. jacchus, a member of the New World monkey clade, is the most widely studied marmoset. It is a popular non-human primate model due to its small body size and unique biological features, and has contributed to the study of brain function, immunity, reproductive biology and drug toxicity. *C. jacchus* marmosets typically give birth to

Input

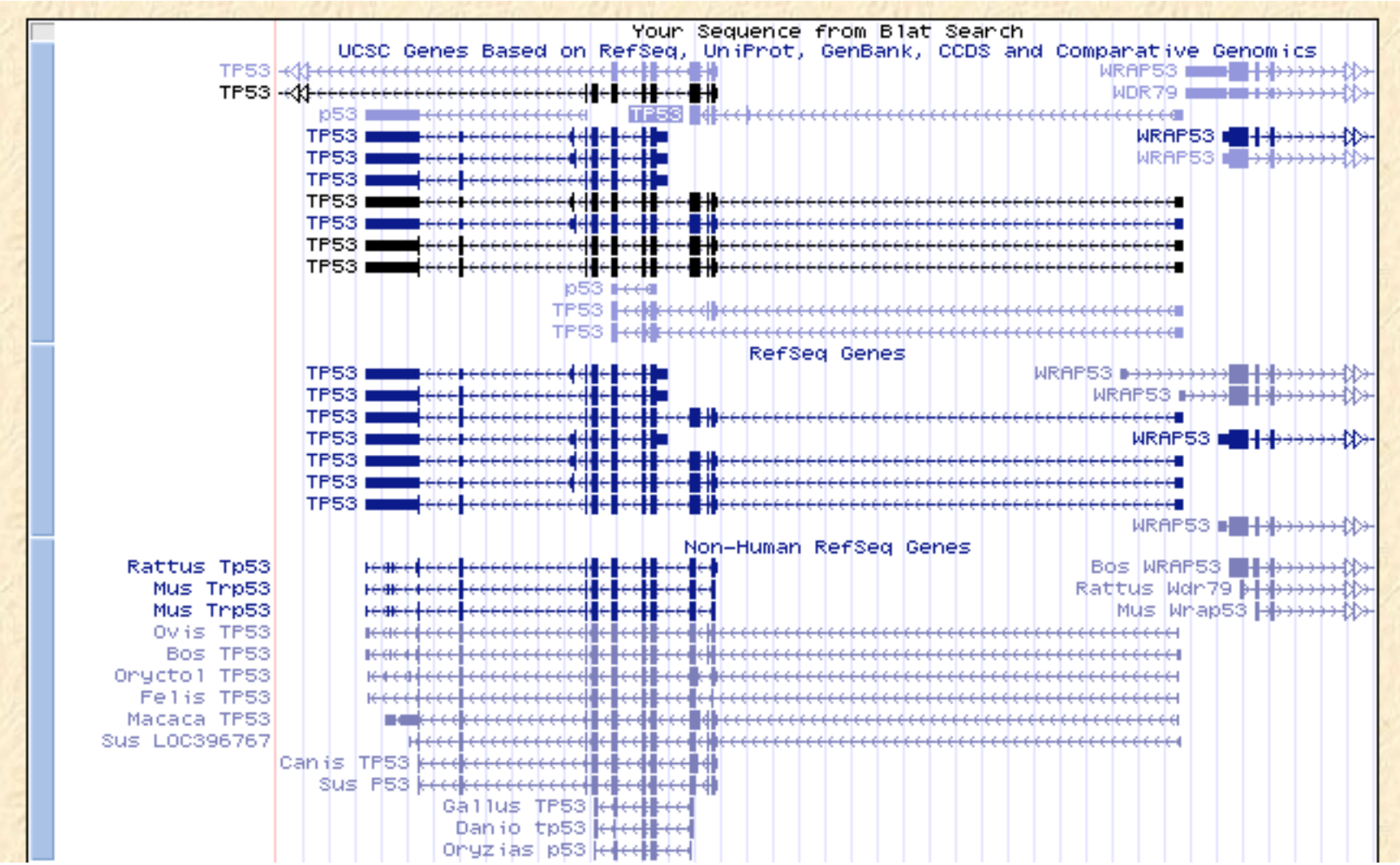



```
chr3:1-1000000,
NM_000984….
```

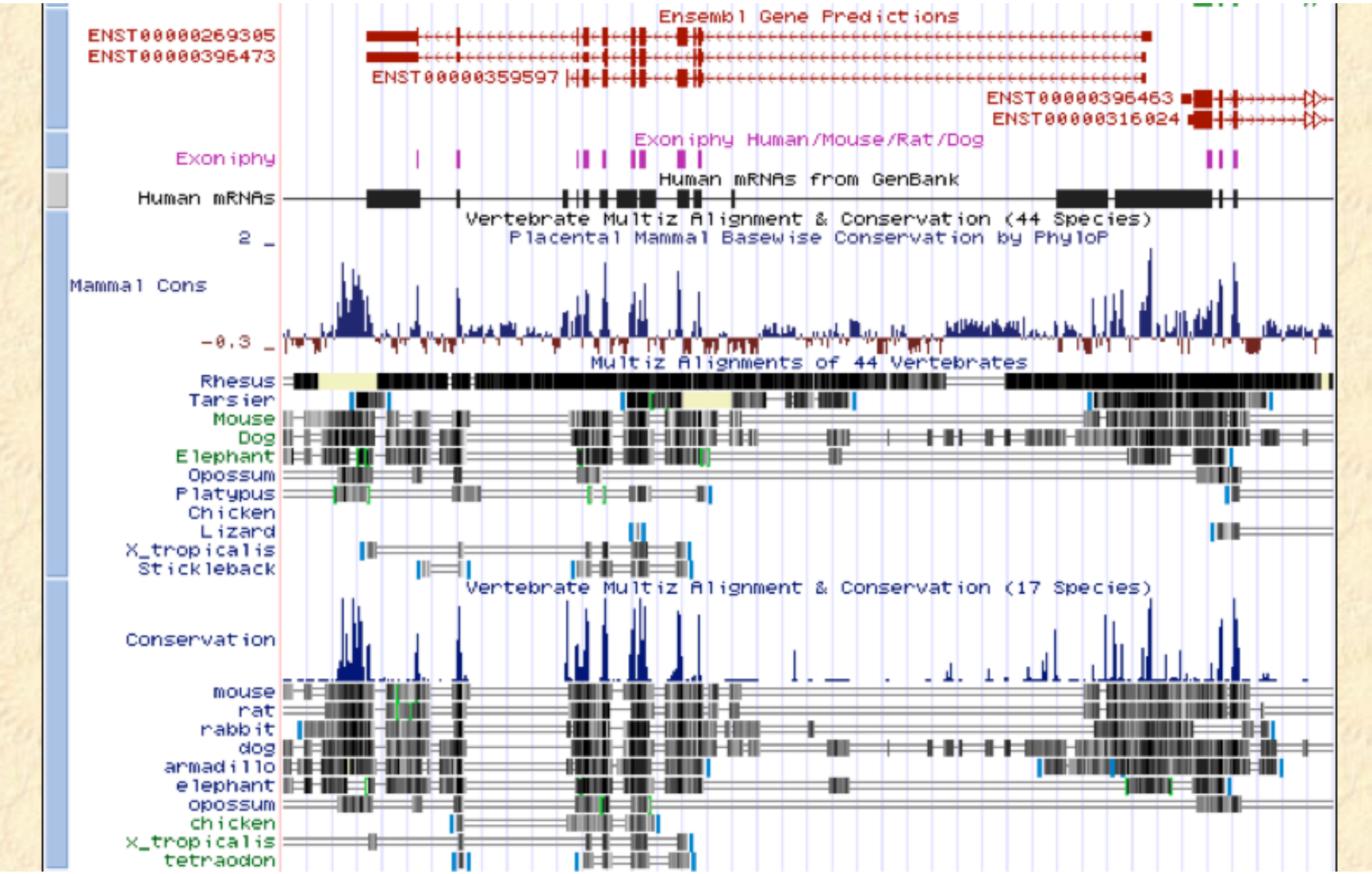
Genome View


UCSC Genes
TP53 (uc010cnk.1) at chr17:7520037-75
TP53 (uc010cnh.1) at chr17:7512445-75
TP53 (uc010cng.1) at chr17:7512445-75
TP53 (uc010cnf.1) at chr17:7512445-75
TP53 (uc002gio.2) at chr17:7518224-75
TP53 (uc002gin.2) at chr17:7518224-75
TP53 (uc002gim.2) at chr17:7512445-75
TP53 (uc002gij.2) at chr17:7512445-75
TP53 (uc002gii.1) at chr17:7512445-75
TP53 (uc002gih.1) at chr17:7510129-75
TP53 (uc002gig.1) at chr17:7505822-75
TP53TG3 (uc010caz.1) at chr16:3316819
TP53TG3 (uc010cay.1) at chr16:3311165
TP53TG5 (uc002xny.1) at chr20:4343593
Home Genomes Blat Tables Gene Sorter PCR DNA Convert Ensembl NCBI PDF/PS Session Help
UCSC Genome Browser on Human Mar. 2006 Assembly (hg18)
move <<< << < > >> >>> zoom in 1.5x 3x 10x base zoom out 1.5x 3x 10x
position/search chr17:7,520,037-7,531,588 jump clear size 11,552 bp. configure
chr17 (p13.1)
UCSC Genes Based on RefSeq, UniProt, GenBank, CCDS and Comparative Genomics
RefSeq Genes
Non-Human RefSeq Genes
Mammalian Gene Collection Full ORF mRNAs
Ensembl Gene Predictions
Exoniphy Human/Mouse/Rat/Dog
Human mRNAs from GenBank
Vertebrate Multiz Alignment & Conservation (44 Species)
Placental Mammal Basewise Conservation by PhyloP
Mammal Cons
Multiz Alignments of 44 Vertebrates
Vertebrate Multiz Alignment & Conservation (17 Species)
Conservation

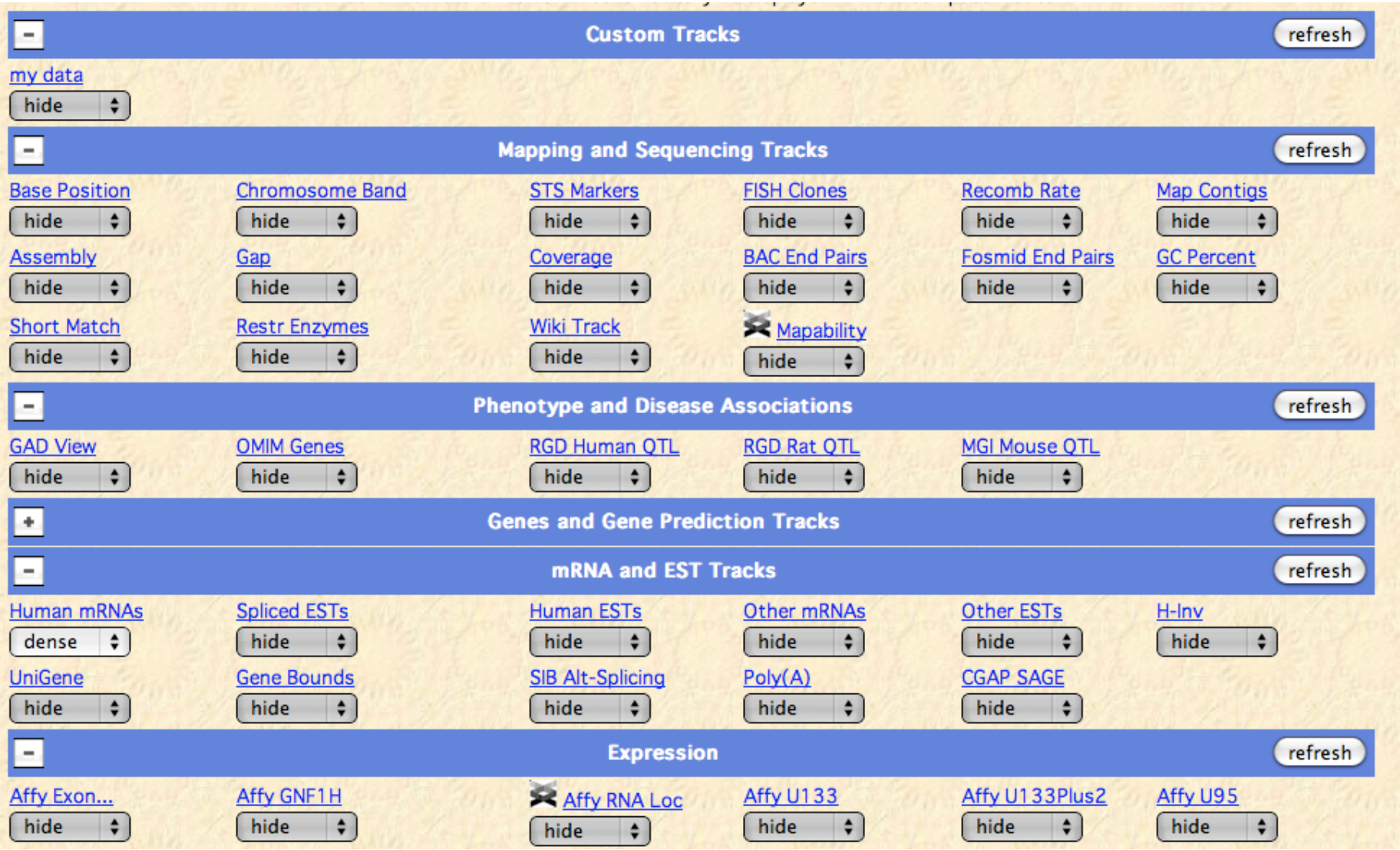
UCSC genes and Reference Sequences



Conservation



Customization of The Display



Information of the Sequence

Home Genomes Genome Browser Blat Tables Gene Sorter PCR Session FAQ Help

Human Gene TP53 (uc002gig.1) Description and Page Index

Description: tumor protein p53 isoform b
RefSeq Summary (NM_001126114): This gene encodes tumor protein p53, which responds to diverse cellular stresses to regulate target genes that induce cell cycle arrest, apoptosis, senescence, DNA repair, or changes in metabolism. p53 protein is expressed at low level in normal cells and at a high level in a variety of transformed cell lines, where it's believed to contribute to transformation and malignancy. p53 is a DNA-binding protein containing transcription activation, DNA-binding, and oligomerization domains. It is postulated to bind to a p53-binding site and activate expression of downstream genes that inhibit growth and/or invasion, and thus function as a tumor suppressor. Mutants of p53 that frequently occur in a number of different human cancers fail to bind the consensus DNA binding site, and hence cause the loss of tumor suppressor activity. Alterations of this gene occur not only as somatic mutations in human malignancies, but also as germline mutations in some cancer-prone families with Li-Fraumeni syndrome. Multiple p53 variants due to alternative promoters and multiple alternative splicing have been found. These variants encode distinct isoforms, which can regulate p53 transcriptional activity. [provided by RefSeq].
Strand: - **Genomic Size:** 14841 **Exon Count:** 7 **Coding Exon Count:** 7

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

− Sequence and Links to Tools and Databases

Genomic Sequence (chr17:7,505,822-7,520,662)			mRNA (may differ from genome)		Protein (285 aa)
Gene Sorter	Genome Browser	Protein FASTA	Proteome Browser	Table Schema	CGAP
Ensembl	Entrez Gene	ExonPrimer	GeneCards	GeneTests	Gepis Tissue
H-INV	HGNC	HuGE	Jackson Lab	OMIM	PubMed
Stanford SOURCE	Treefam	UniProtKB	User annotations		

Amino Acid Sequence

Genomic Sequence (chr17:7,505,822-7,520,662)			mRNA (may differ from genome)		Protein (285 aa)
Gene Sorter	Genome Browser	Protein FASTA	Proteome Browser	Table Schema	CGAP
Ensembl	Entrez Gene	ExonPrimer	GeneCards	GeneTests	Gepis Tissue
H-INV	HGNC	HuGE	Jackson Lab	OMIM	PubMed
Stanford SOURCE	Treefam	UniProtKB	User annotations		

Species selection: + − Defaults

Primate + −

☐ chimp ☐ gorilla ☐ orangutan ☑ rhesus ☐ marmoset
☑ tarsier ☐ mouse lemur ☐ bushbaby

Placental Mammal + −

☐ treeShrew ☑ mouse ☐ rat ☐ kangaroo rat ☐ guinea Pig
☐ squirrel ☐ rabbit ☐ pika ☐ alpaca ☐ dolphin
☐ cow ☐ horse ☐ cat ☑ dog ☐ microbat
☐ megabat ☐ hedgehog ☐ shrew ☑ elephant ☐ rock hyrax
☐ tenrec ☐ armadillo ☐ sloth

Vertebrate + −

☑ opossum ☑ platypus ☑ chicken ☐ zebra finch ☑ lizard
☑ x. tropicalis ☐ tetraodon ☐ fugu ☑ stickleback ☐ medaka
☐ zebrafish ☐ lamprey

Submit

For information about output data format see the User's Guide

```
>uc002gig.1_hg18 286 chr17:7505982-7520637-
MEEPQSDPSVEPPLSQETFSDLWKLLPENNVLSPLPSQAMDDLMLSPDDIEQWFTEDPGPDEAPRMPEAAPPVAPAPAAPTPAAPAPAPSWPLSSSVPSQKTYQGSYGFRLGFLHSGTAKSVTCTYSPALNKMFCQ
>uc002gig.1_rheMac2 286 chr16:7392311-7407511-
MEEPQSDPSIEPPLSQETFSDLWKLLPENNVLSPLPSQAVDDLMLSPDDLAQWLTEDPGPDEAPRMSEAAPPMAPTPAAPTPAAPAPAPSWPLSSSVPSQKTYHGSYGFRLGFLHSGTAKSVTCTYSPDLNKMFCQ
>uc002gig.1_tarSyr1 286 scaffold_121145:134-1367-
----------SQKTYHGSYGFRLGFLHSGTAKSVTCTYSPALNKMFCQ
>uc002gig.1_mm9 286 chr11:69400402-69402752+
MEESQSDISLELPLSQETFSGLWKLLPPEDIL---PSPRMDDLLLP--DVEEFF-E---PSEALRVSGAPAAQDPVTETPGPVAPAPATPWPLSSFVPSQKTYQGNYGFHLGFLQSGTAKSVMCTYSPPLNKLFCQ
>uc002gig.1_canFam2 286 chr5:35558952-35560756-
MQEPQSELNIDPPLSQETFSELWNLLPENNVLSSELCPAVDELLLPE--VVNWLDE---SDDAPRMPATSA------TAPGPAP-----SWPLSSSVPSPKTYPGTYGFRLGFLHSGTAKSVTWTYSPLLNKLFCQ
>uc002gig.1_loxAfr2 286 scaffold_28406:25564-27374-
MK-PQSDLSTELPLSQETFSYLWELLPENPVLSPTLPPAMDDLL--SEDTANWL-----QVEAQGMSTTPA------ATPTPVAPAPATSWTLSSSVPSQKTYPGTYGFRLGFLHSGTAKSVTCTYSPDLNKLFCQ
>uc002gig.1_ornAna1 286 Contig14337:7594-11525-
MEESLTE---EPPMSQETFLDLWQ----------YSPLLNKLFCQ
>uc002gig.1_xenTro2 286 scaffold_440:76586-78543-
----------PLAPDMTVL--RAYVPCPPZ--SSAVPSTEDYAGSYGLKLEFQQNGTAKSVTCTYSTDLNKLFCQ
>uc002gig.1_gasAcu1 286 chrVIII:17664060-17671051+
----------E---P--ASRTEAP--PPAPTTTTT--TAPSPAPPSTPSPAIPSNTDYAGTHAFDVSFQQSSTAKSATWTYSTDLKKLYCQ
```

```
>uc002gig.1 (TP53) length=285
MEEPQSDPSVEPPLSQETFSDLWKLLPENNVLSPLPSQAMDDLMLSPDDIEQWFTEDPGP
DEAPRMPEAAPPVAPAPAAPTPAAPAPAPSWPLSSSVPSQKTYQGSYGFRLGFLHSGTAK
SVTCTYSPALNKMFCQLAKTCPVQLWVDSTPPPGTRVRAMAIYKQSQHMTEVVRRCPHHE
RCSDSDGLAPPQHLIRVEGNLRVEYLDDRNTFRHSVVVPYEPPEVGSDCTTIHYNYMCNS
SCMGGMNRRPILTIITLEDSRPRQVDHLRSGVQDSLANIAKSHLY
```

Proteome Browser

Genomic Sequence (chr17:7,505,822-7,520,662)			mRNA (may differ from genome)		Protein (285 aa)
Gene Sorter	Genome Browser	Protein FASTA	Proteome Browser	Table Schema	CGAP
Ensembl	Entrez Gene	ExonPrimer	GeneCards	GeneTests	Gepis Tissue

Human protein: Q1MSW9 Cellular tumor antigen p53 (Fragment).

Move <<< << < > >> >>> Current scale: FULL Rescale to 1/6 1/2 FULL

AA Scale 50 100

AA Sequence EPQSDPSVEPPLSQETFSDLWKLLPENNVLSPLPSQAMDDLMLSPDDIEQWFTEDPGPDEAPRMPEAAPPVAPAPAAPTPAAPAPAPSWPLSSSVPSQKTYQGSYGFRLGF

Genome Browser Previous position in UCSC Genome Browser: chr17:7,518,519-7,530,070

Polarity

Hydrophobicity

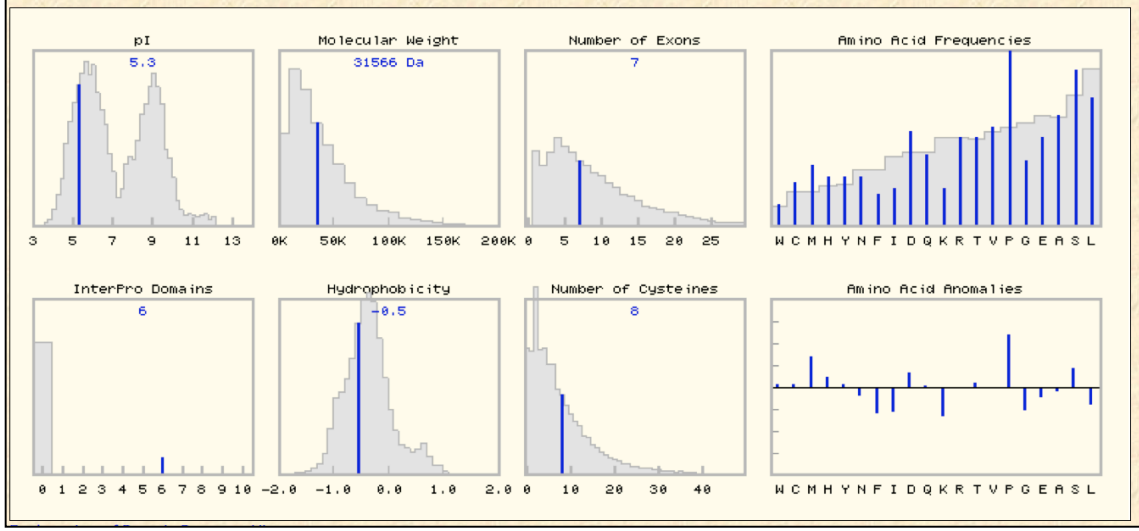
Cysteines

Predicted Glycosylation

AA Anomalies

AA Scale 50 100

Explanation of Protein Tracks



UCSC Links:

- Genome Browser - uc002gig.1
- Gene Details Page - uc002gig.1
- Gene Sorter - uc002gig.1

InterPro Domains: Graphical view of domain structure

- IPR002117 - p53
- IPR011615 - p53_DNA_bd
- IPR012346 - p53_RUNT_DNA_bd
- IPR013872 - p53_TAD
- IPR015551 - Trp53

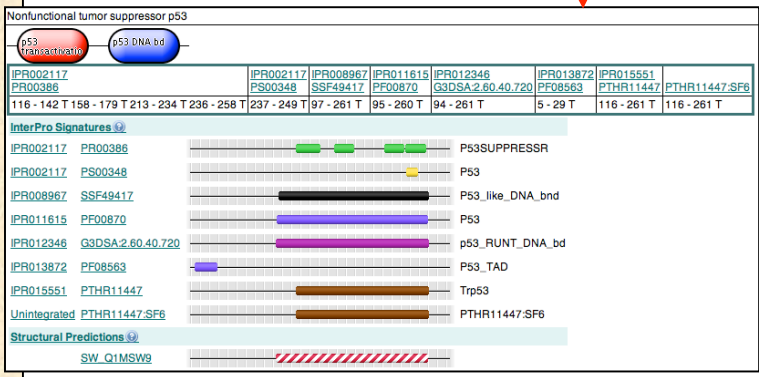
Pfam Domains:

PF00870 - P53 DNA-binding domain

PF08563 - P53 transactivation domain (TAD)

SCOP Domains:

49417 - p53-like transcription factors



Gene Ontology

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

http://www.geneontology.org/

Gene Ontology (GO) Annotations with Structured Vocabulary

Molecular Function:
GO:0003677 DNA binding
GO:0003700 transcription factor activity
GO:0008270 zinc ion binding
GO:0046872 metal ion binding

Biological Process:
GO:0006350 transcription
GO:0006355 regulation of transcription, DNA-dependent
GO:0006915 apoptosis
GO:0007049 cell cycle
GO:0045449 regulation of transcription
GO:0045786 negative regulation of cell cycle

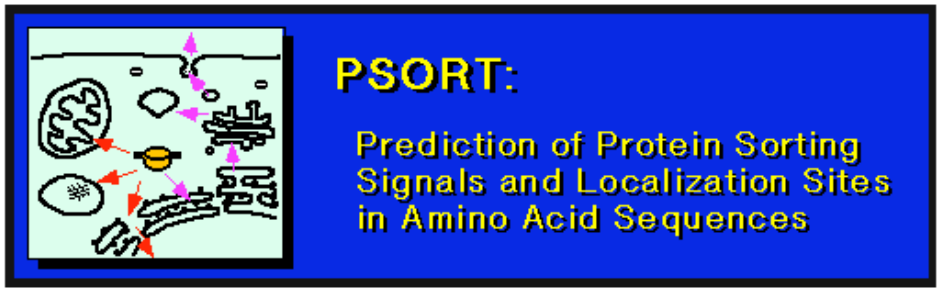
Cellular Component:
GO:0005634 nucleus
GO:0005737 cytoplasm

- Hierarchical structure of biological terms
- These terms are used for gene annotation

GO:0003674 : molecular_function [171023 gene products]
GO:0005488 : binding [46219 gene products]
GO:0003676 : nucleic acid binding [14830 gene products]
GO:0003677 : DNA binding [9865 gene products]
GO:0000497 : base pairing with DNA [5 gene products]
GO:0003684 : damaged DNA binding [136 gene products]
GO:0008301 : DNA bending activity [43 gene products]
GO:0003689 : DNA clamp loader activity [26 gene products]
GO:0003916 : DNA topoisomerase activity [158 gene products]
GO:0015616 : DNA translocase activity [13 gene products]
GO:0010844 : recombination hotspot binding [2 gene products]
GO:0043565 : sequence-specific DNA binding [1078 gene products]
GO:0043566 : structure-specific DNA binding [584 gene products]
GO:0003700 : transcription factor activity [5927 gene products]

Subcellular Localization (PSORT)

http://psort.ims.u-tokyo.ac.jp/



Input: Amino acid sequence (FASTA format)

CONTENTS

WoLF PSORT (an update of PSORT II for fungi/animal/plant sequences)

WoLF PSORT Prediction

PSORT II (Recommended for animal/yeast sequences)

PSORT II Users' Manual
PSORT II Prediction

PSORT (Old version; for bacterial/plant sequences)

PSORT Users' Manual (WWW version)
PSORT Prediction

WoLF PSORT

WoLF PSORT

Protein Subcellular Localization Prediction

about WoLF PSORT WoLF PSORTについて links Example Output

Please select an organism type:
- (•) Animal
- ( ) Plant
- ( ) Fungi

Please select input method:
- (•) From Text Area
- ( ) From File

Input Filename:
ファイルを選択 ファイルが選...ていません

Text Area: Enter multifasta format protein sequence(s) here.

```
>uc002gig.1 (TP53) length=285
MEEPQSDPSVEPPLSQETFSDLWKLLPENNVLSPLPSQAMDDLMLSPDDIEQWFTEDPGP
DEAPRMPEAAPPVAPAPAAPTPAAPAPAPSWPLSSSVPSQKTYQGSYGFRLGFLHSGTAK
SVTCTYSPALNKMFCQLAKTCPVQLWVDSTPPPGTRVRAMAIYKQSQHMTEVVRRCPHHE
RCSDSDGLAPPQHLIRVEGNLRVEYLDDRNTFRHSVVVPYEPPEVGSDCTTIHYNYMCNS
SCMGGMNRRPILTIITLEDSRPRQVDHLRSGVQDSLANIAKSHL
```

送信 clear

(↑ Select organism type to activate the submit button)

Input the query sequence

Results for uc002gig.1

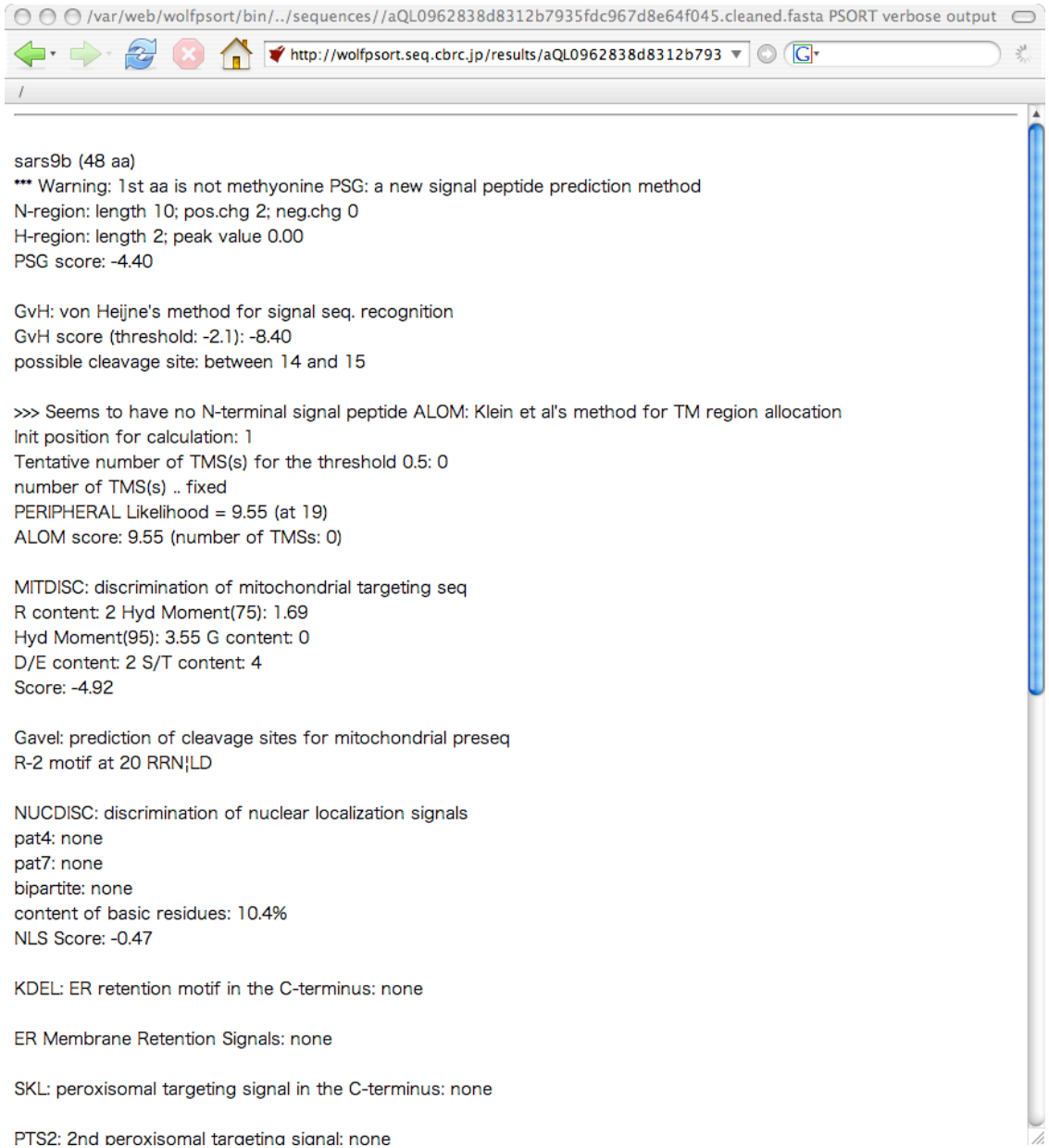
uc002gig.1 WoLFPSORT prediction nucl: 30.5, cyto_nucl: 16.5

PSORT features and traditional PSORTII prediction

32 Nearest Neighbors

id	site	distance	identity	comments
MYF5_BOVIN	nucl	37.8	14%	[Uniprot] SWISS-PROT45:Nuclear.
MYF5_MOUSE	nucl	39.5	13%	[Uniprot] SWISS-PROT45:Nuclear.
MYOG_CHICK	nucl	41.9	14%	[Uniprot] SWISS-PROT45:Nuclear.
MYF5_CHICK	nucl	42.5	14%	[Uniprot] SWISS-PROT45:Nuclear.
MYF5_HUMAN	nucl	43.8	14%	[Uniprot] SWISS-PROT45:Nuclear.
MYF5_NOTVI	nucl	46.0	15%	[Uniprot] SWISS-PROT45:Nuclear.
MYOG_HUMAN	nucl	47.1	16%	[Uniprot] SWISS-PROT45:Nuclear.
MYOG_MOUSE	nucl	47.3	16%	[Uniprot] SWISS-PROT45:Nuclear.
PIX3_XENLA	nucl	51.9	11%	[Uniprot] SWISS-PROT45:Nuclear.
MYOD_BRARE	nucl	51.9	14%	[Uniprot] SWISS-PROT45:Nuclear.
COT1_MOUSE	nucl	52.1	12%	[Uniprot] SWISS-PROT45:Nuclear.
COT2_BOVIN	nucl	52.2	12%	[Uniprot] SWISS-PROT45:Nuclear.
COT2_HUMAN	nucl	53.0	13%	[Uniprot] SWISS-PROT45:Nuclear. GO:0005634; C:nucleus; Evidence:TAS.
COT2_MOUSE	nucl	53.0	13%	[Uniprot] SWISS-PROT45:Nuclear.
MYOG_PIG	nucl	53.7	15%	[Uniprot] SWISS-PROT45:Nuclear.
MYOD_MOUSE	nucl	56.0	14%	[Uniprot] SWISS-PROT45:Nuclear. GO:0005634; C:nucleus; Evidence:IDA.
COT2_RAT	nucl	57.6	13%	[Uniprot] SWISS-PROT45:Nuclear.
MYF5_XENLA	nucl	59.2	15%	[Uniprot] SWISS-PROT45:Nuclear.
MYO2_ONCMY	nucl	63.7	10%	[Uniprot] SWISS-PROT45:Nuclear.
MYOG_RAT	nucl	67.8	10%	[Uniprot] SWISS-PROT45:Nuclear.
LPAX_HUMAN	cyto	71.9	15%	[Uniprot] SWISS-PROT45:Cytoplasmic.
HXB1_HUMAN	nucl	72.7	16%	[Uniprot] SWISS-PROT45:Nuclear. GO:0005634; C:nucleus; Evidence:NAS.
ELT2_CAEEL	nucl	73.7	13%	[Uniprot] SWISS-PROT45:Nuclear.
ESC_DROME	nucl	74.0	14%	[Uniprot] SWISS-PROT45:Nuclear. GO:0005634; C:nucleus; Evidence:IDA.
HXA1_HUMAN	nucl	75.2	11%	[Uniprot] SWISS-PROT45:Nuclear.
COT2_CHICK	nucl	75.9	13%	[Uniprot] SWISS-PROT45:Nuclear.
HXB1_MOUSE	nucl	77.1	15%	[Uniprot] SWISS-PROT45:Nuclear.
P53_HUMAN	cyto_nucl	78.5	68%	[Uniprot] SWISS-PROT45:Cytoplasmic and nuclear. GO:0005634; C:nucleus; Evidence:TAS.
MYOD_SHEEP	nucl	78.5	13%	[Uniprot] SWISS-PROT45:Nuclear.
PAX6_ORYLA	nucl	79.4	14%	[Uniprot] SWISS-PROT45:Nuclear.
ASC2_MOUSE	nucl	79.6	13%	[Uniprot] SWISS-PROT45:Nuclear.
HXA1_MOUSE	nucl	80.0	11%	[Uniprot] SWISS-PROT45:Nuclear.

Output of PSORT II



- Prediction results for
  - Signal peptides
  - Transmembrane regions
  - Other sorting signals

GeneCards

GeneCards® an academic web site of the WEIZMANN INSTITUTE OF SCIENCE in association with XENNEX

Gene Search (GeneCards Home) | GeneCards Guide | User Feedback | Terms of Use | Notice about third-party sites

This service is provided free to academic non-profit institutions. ALL other users require a Commercial License from XenneX, Inc.

TP53 Gene
protein-coding GIFtS: 84
GC17M007512

tumor protein p53
Symbol approved by the HUGO Gene Nomenclature Committee (HGNC) database

invitrogen Gene Antibodies / RNAi / Pathways | SIGMA Life Science Proteins Small Molecules / Antibodies | Services Jump to Section... | Antibodies / Pathways Proteins & Enzymes Assays/Kits/cDNA/RNAi

Aliases & Descriptions for TP53

(According to [1]HGNC, [2]Entrez Gene, [3]UniProtKB/Swiss-Prot, [4]UniProtKB/TrEMBL, [5]OMIM, [6]GeneLoc, and/or [7]Ensembl)
About This Section

Jump to Section...

Aliases	Descriptions	External Ids
FLJ92943 [2]	Antigen NY-CO-13 [3]	HGNC: 11998[1]
LFS1 [1, 2, 5]	Phosphoprotein p53 [2, 3]	Entrez Gene: 7157[2]
P53 [3, 5]	Tumor suppressor p53 [3]	UniProtKB: P04637[3]
TRP53 [2]	p53 antigen [2]	Ensembl: ENSG00000141510[7]
p53 [1, 2]	p53 transformation suppressor [2]	
	p53 tumor suppressor [2]	
	transformation-related protein 53 [2]	
	tumor protein p53 [1, 2]	

Search outside databases for aliases for TP53 gene

Previous GC identifers: GC17P008026 GC17M008311 GC17M007514 GC17M007772

Summaries for TP53

(According to Entrez Gene, Wikipedia's Gene Wiki, UniProtKB/Swiss-Prot, and/or UniProtKB/TrEMBL)
About This Section

Jump to Section...

EntrezGene summary for TP53:

This gene encodes tumor protein p53, which responds to diverse cellular stresses to regulate target genes that induce cell cycle arrest, apoptosis, senescence, DNA repair, or changes in metabolism. p53 protein is expressed at low level in normal cells and at a high level in a variety of transformed cell lines, where it's believed to contribute to transformation and malignancy. p53 is a DNA-binding protein containing transcription activation, DNA-binding, and oligomerization domains. It is postulated to bind to a p53-binding site and activate expression of downstream genes that inhibit growth and/or invasion, and thus function as a tumor suppressor. Mutants of p53 that frequently occur in a number of different human cancers fail to bind the consensus DNA binding site, and hence cause the loss of tumor suppressor activity. Alterations of this gene occur not only as somatic mutations in human malignancies, but also as germline mutations in some cancer-prone families with Li-Fraumeni syndrome. Multiple p53 variants due to alternative promoters and multiple alternative splicing have been found. These variants encode distinct isoforms, which can regulate p53 transcriptional activity. [provided by RefSeq]

Add Your Own Data to Browser

Human (*Homo sapiens*) Genome Browser Gateway

The UCSC Genome Browser was created by the Genome Bioinformatics Group of UC Santa Cruz.

clade	genome	assembly	position or search term	image width	
Mammal	Human	Mar. 2006	chr17:7,505,937-7,536,236	620	submit

Click here to reset the browser user interface settings to their defaults.

add custom tracks | configure tracks and display | clear position

Add Custom Tracks

clade Mammal genome Human assembly Mar. 2006 [hg18]

Display your own data as custom annotation tracks in the browser. Data must be formatted in BED, bigBed, BEDG
Publicly available custom tracks are listed here. Examples are here.

Paste URLs or data: Or upload: ファイルを選択 ファイルが選...ていません Submit

```
track type=wiggle_0 name="my data"
chr17   7521037   7521137   2
chr17   7524037   7524137   0
chr17   7526934   7527034   5
chr17   7528167   7528267   2
```

Clear

Manage Custom Tracks

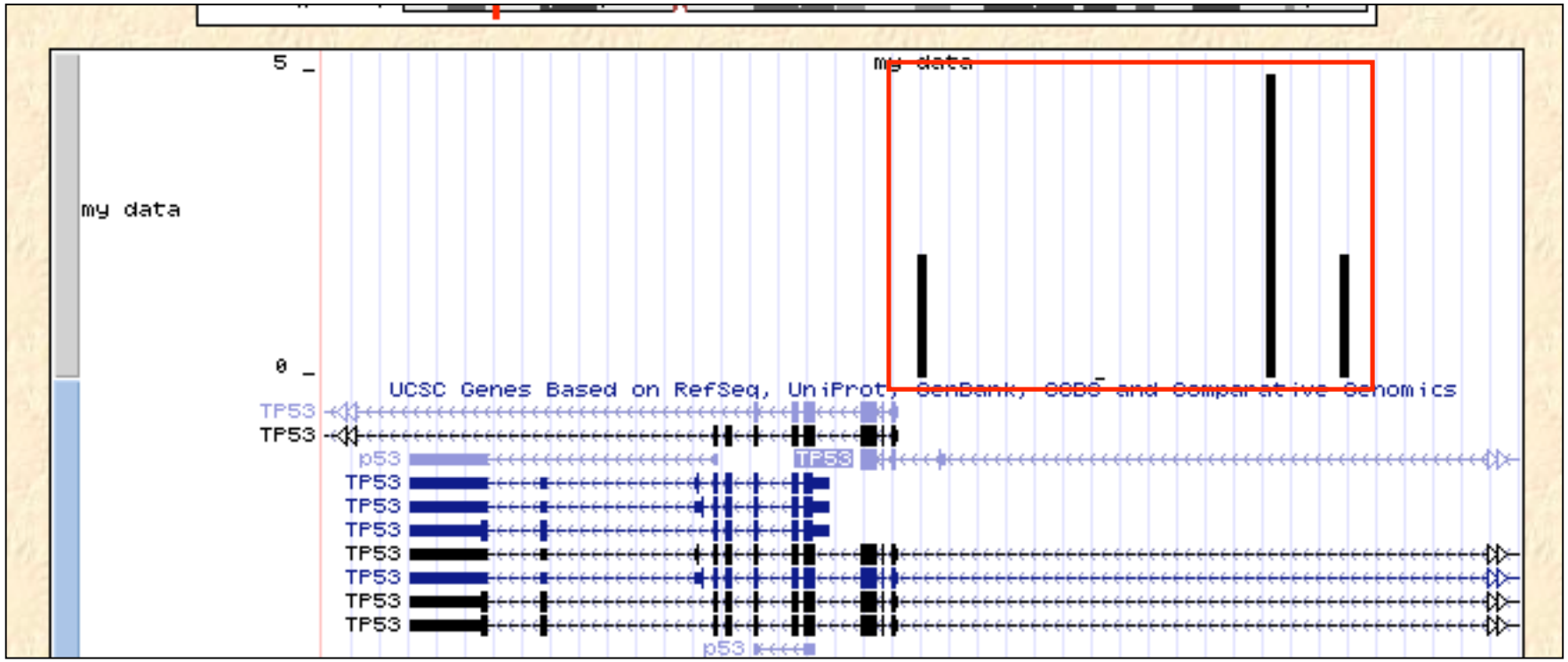
genome: Human **assembly:** Mar. 2006 [hg18]

Replaced: my data

Name	Description	Type	Doc	delete
my data	my data	wiggle_0		☐

add custom tracks
go to genome browser
go to table browser

Your Data on the Browser



http://genome.ucsc.edu/cgi-bin/hgTracks?hgsid=137297224&Submit=go+to+genome+browser

You can access to your data at

http://genome.ucsc.edu/cgi-bin/hgTracks?hgsid=137297224

within 48 hours after the last access

Inside of Genome browser: download

Human Genome

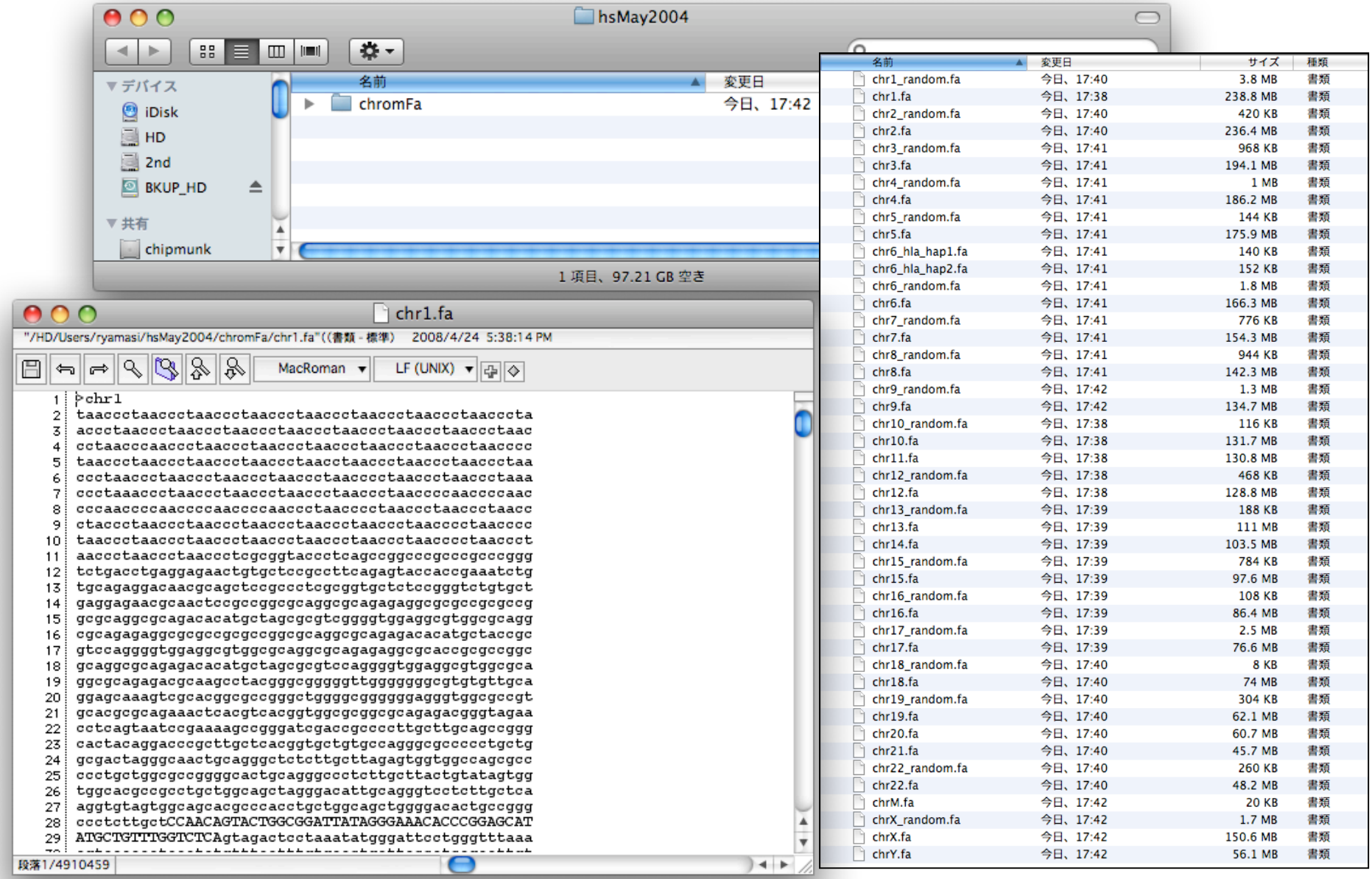
Mar. 2006 (hg18)

- Full data set
- Data set by chromosome
- Annotation database
- Affymetrix transcriptome phase 3 data
- Regulatory potential data
- GC percent data
- Whole-Genome ENCODE wiggle data
- Protein database for hg18
- SNP-masked FASTA files
- LiftOver files
- Pairwise Alignments
  - Human self alignments
  - Human/Chimp (panTro2)
  - Human/Chimp (panTro1)
  - Human/Rhesus (rheMac2)
  - Human/Mouse (mm9)
  - Human/Mouse (mm8)
  - Human/Mouse (mm7)
  - Human/Rat (rn4)
  - Human/Cat (felCat3)
  - Human/Dog (canFam2)
  - Human/Horse (equCab1)

Name	Last modified	Size	Description
Parent Directory		-	
chromAgp.zip	06-Apr-2006 09:45	537K	
chromFa.zip	03-Feb-2006 11:05	899M	
chromFaMasked.zip	03-Feb-2006 11:12	487M	
chromOut.zip	03-Feb-2006 10:51	153M	
chromTrf.zip			
contigAgp.zip			
contigFa.zip			
contigFaMasked.zip			
contigOut.zip			
contigTrf.zip			
est.fa.gz			
hg18.2bit			
liftAll.zip			
md5sum.txt			
mrna.fa.gz			
refMrna.fa.gz			
upstream1000.zip			
upstream2000.zip			
upstream5000.zip			
xenoMrna.fa.gz			

phastConsElements28way.txt.gz
phastConsElements28wayPlacMammal.sql
phastConsElements28wayPlacMammal.txt.gz
polyaDb.sql
polyaDb.txt.gz
polyaPredict.sql
polyaPredict.txt.gz
productName.sql
productName.txt.gz
recombRate.sql
recombRate.txt.gz
refFlat.sql
refFlat.txt.gz
refGene.sql
refGene.txt.gz
refLink.sql

The Genome Sequences



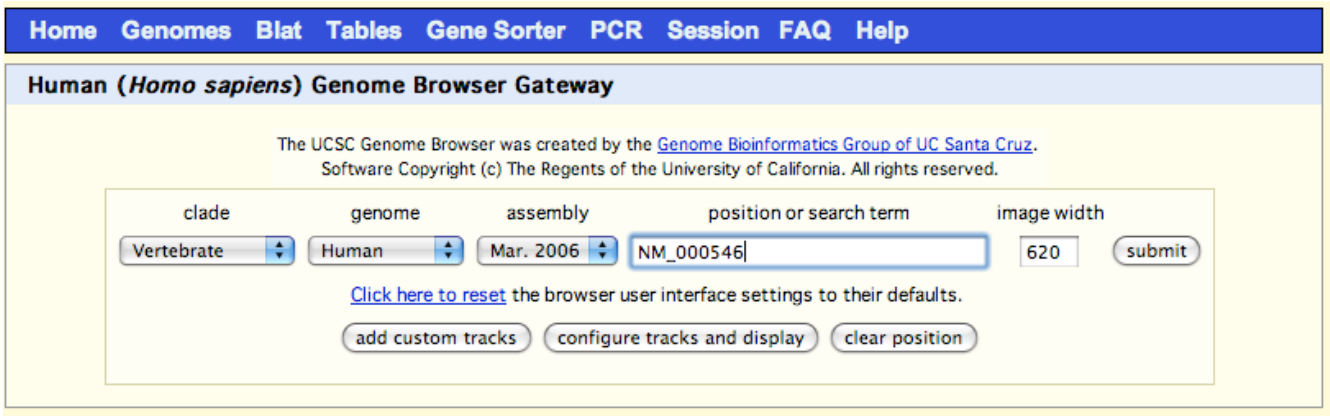
名前	変更日	サイズ	種類
chr1_random.fa	今日、17:40	3.8 MB	書類
chr1.fa	今日、17:38	238.8 MB	書類
chr2_random.fa	今日、17:40	420 KB	書類
chr2.fa	今日、17:40	236.4 MB	書類
chr3_random.fa	今日、17:41	968 KB	書類
chr3.fa	今日、17:41	194.1 MB	書類
chr4_random.fa	今日、17:41	1 MB	書類
chr4.fa	今日、17:41	186.2 MB	書類
chr5_random.fa	今日、17:41	144 KB	書類
chr5.fa	今日、17:41	175.9 MB	書類
chr6_hla_hap1.fa	今日、17:41	140 KB	書類
chr6_hla_hap2.fa	今日、17:41	152 KB	書類
chr6_random.fa	今日、17:41	1.8 MB	書類
chr6.fa	今日、17:41	166.3 MB	書類
chr7_random.fa	今日、17:41	776 KB	書類
chr7.fa	今日、17:41	154.3 MB	書類
chr8_random.fa	今日、17:41	944 KB	書類
chr8.fa	今日、17:41	142.3 MB	書類
chr9_random.fa	今日、17:42	1.3 MB	書類
chr9.fa	今日、17:42	134.7 MB	書類
chr10_random.fa	今日、17:38	116 KB	書類
chr10.fa	今日、17:38	131.7 MB	書類
chr11.fa	今日、17:38	130.8 MB	書類
chr12_random.fa	今日、17:38	468 KB	書類
chr12.fa	今日、17:38	128.8 MB	書類
chr13_random.fa	今日、17:39	188 KB	書類
chr13.fa	今日、17:39	111 MB	書類
chr14.fa	今日、17:39	103.5 MB	書類
chr15_random.fa	今日、17:39	784 KB	書類
chr15.fa	今日、17:39	97.6 MB	書類
chr16_random.fa	今日、17:39	108 KB	書類
chr16.fa	今日、17:39	86.4 MB	書類
chr17_random.fa	今日、17:39	2.5 MB	書類
chr17.fa	今日、17:39	76.6 MB	書類
chr18_random.fa	今日、17:40	8 KB	書類
chr18.fa	今日、17:40	74 MB	書類
chr19_random.fa	今日、17:40	304 KB	書類
chr19.fa	今日、17:40	62.1 MB	書類
chr20.fa	今日、17:40	60.7 MB	書類
chr21.fa	今日、17:40	45.7 MB	書類
chr22_random.fa	今日、17:40	260 KB	書類
chr22.fa	今日、17:40	48.2 MB	書類
chrM.fa	今日、17:42	20 KB	書類
chrX_random.fa	今日、17:42	1.7 MB	書類
chrX.fa	今日、17:42	150.6 MB	書類
chrY.fa	今日、17:42	56.1 MB	書類

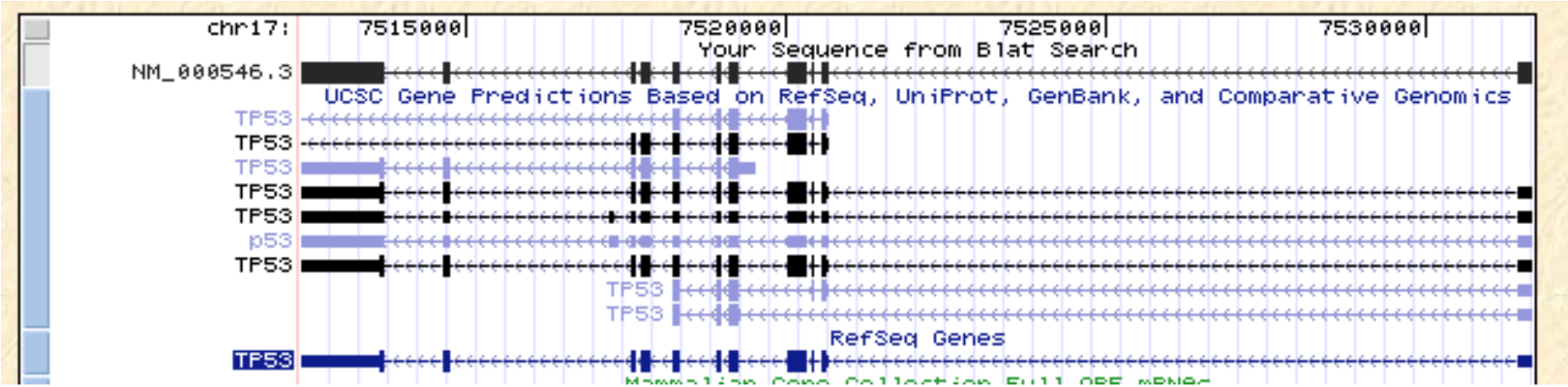
refGene.txt

ブック1

	A	B	C	D	E	F	G	H	I	J	K
1		name	chromosome	strand	mRNA start	mRNA end	CDS start	CDS end	exon	exon start	exon end
2	1643	NM_016459	chr5	-	138751155	138753504	138751352	138753444	4	138751155,138751608,138752048,138753267	138751509,138751719,138752173,138753504,
3	1643	NM_152686	chr5	-	138775278	138803038	138777735	138803033	8	138775278,138783640,138786295,138788592	138777860,138783813,138786405,138788702,138789
4	1644	NM_001077693	chr5	-	138812143	138822504	138817312	138822439	5	138812143,138817312,138817525,138817907	138812257,138817373,138817576,138817952,138822
5	1644	NM_198282	chr5	-	138835732	138842476	138836029	138841473	8	138835732,138837097,138838038,138840558	138836223,138837284,138838277,138840667,138841
6	1143	NM_022170	chr7	+	73226641	73249365	73226649	73247574	7	73226641,73239876,73241938,73242087,7324	73226708,73240064,73242003,73242184,73242572,73
7	1142	NM_001081755	chr7	+	73080362	73122172	73080453	73120966	32	73080362,73087629,73088820,73089972,7309	73080535,73087680,73088850,73090005,73093517,73
8											
9	642	NM_000546	chr17	-	7512444	7531642	7513651	7520637	11	7512444,7514651,7517577,7517743,7518223,	7513733,7514758,7517651,7517880,7518333,7519014
10											

26250 rows





Exercises

Let's access

http://www.hgc.jp/~ryamasi/others/2009IUPS_lecture/

and find the exercises.