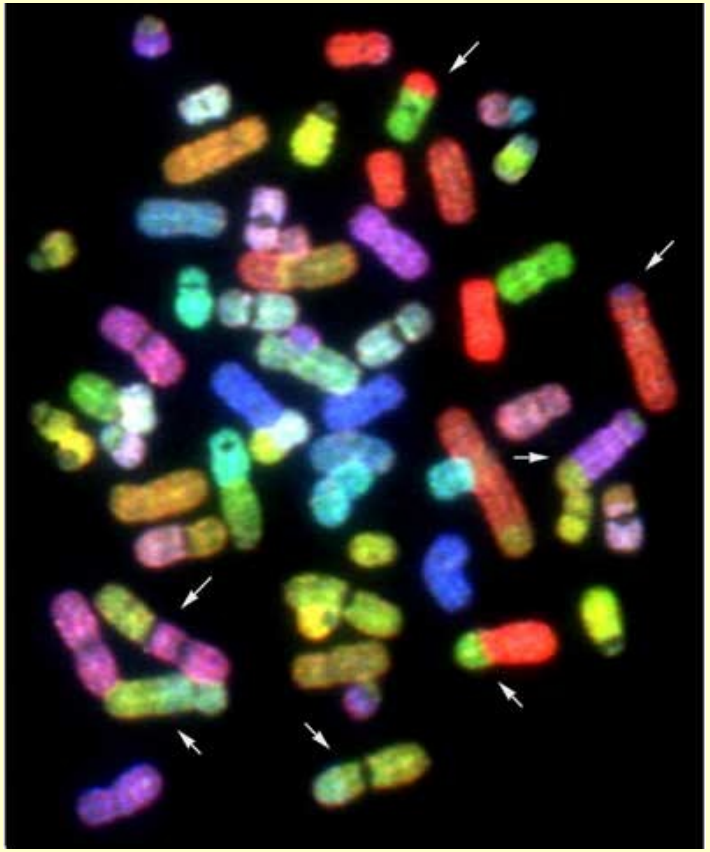
Computational Molecular Biology

Biochem 218 – BioMedical Informatics 231

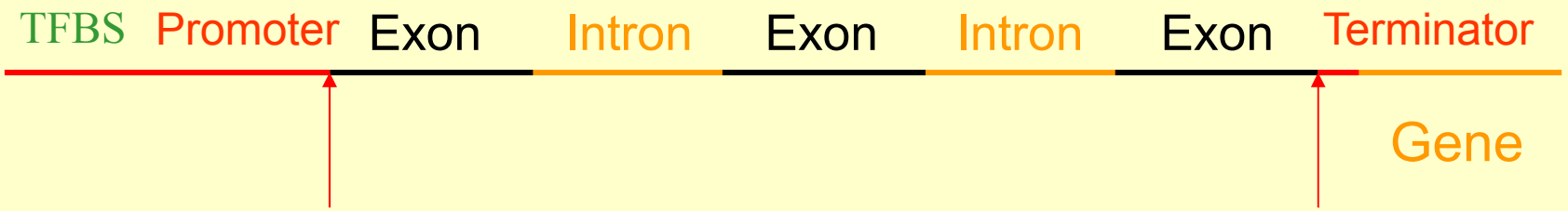
http://biochem218.stanford.edu/

Genome Databases



Doug Brutlag
Professor Emeritus
Biochemistry & Medicine (by courtesy)

Components of a Typical Human Gene

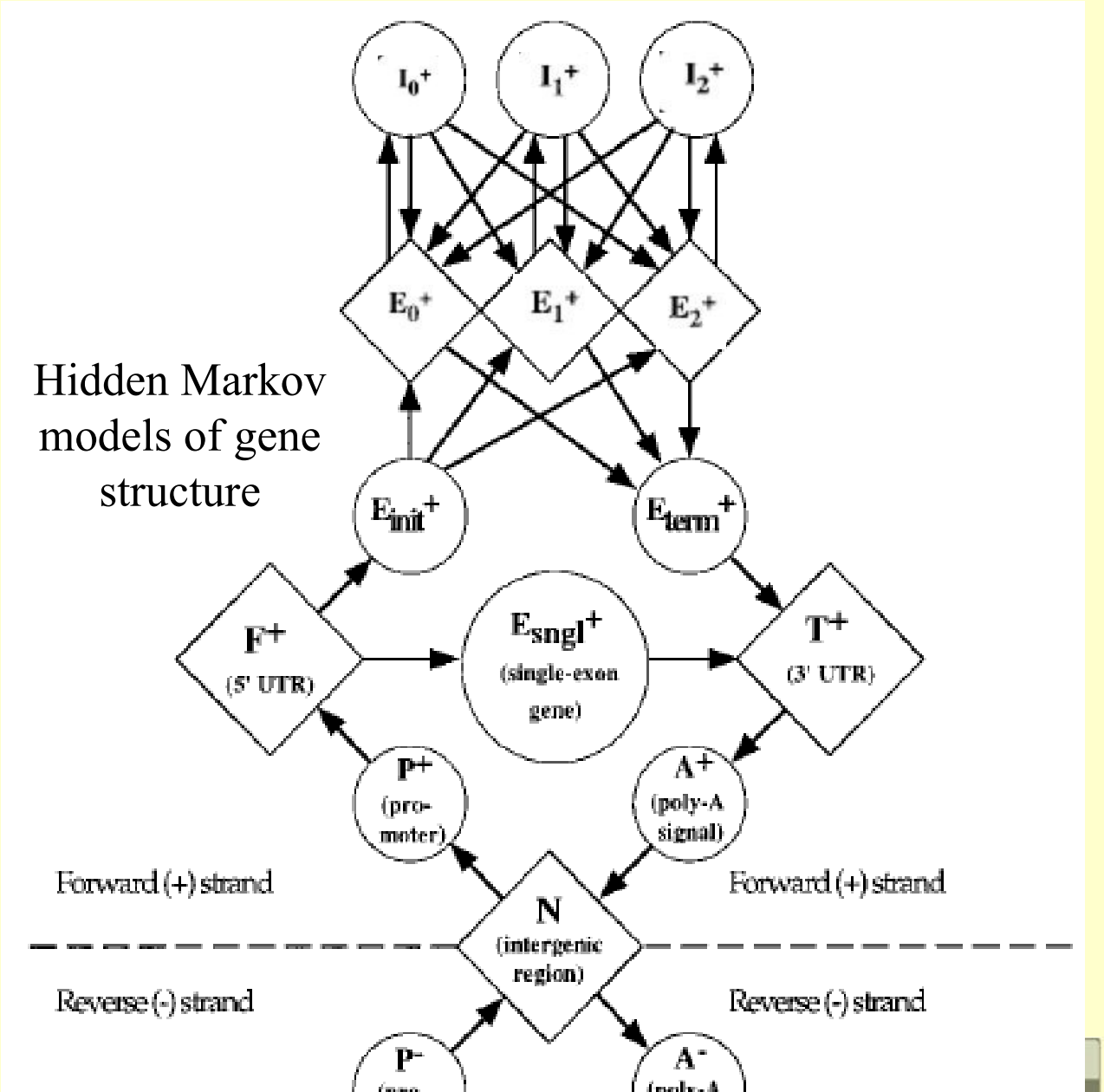


GENSCAN Gene Model

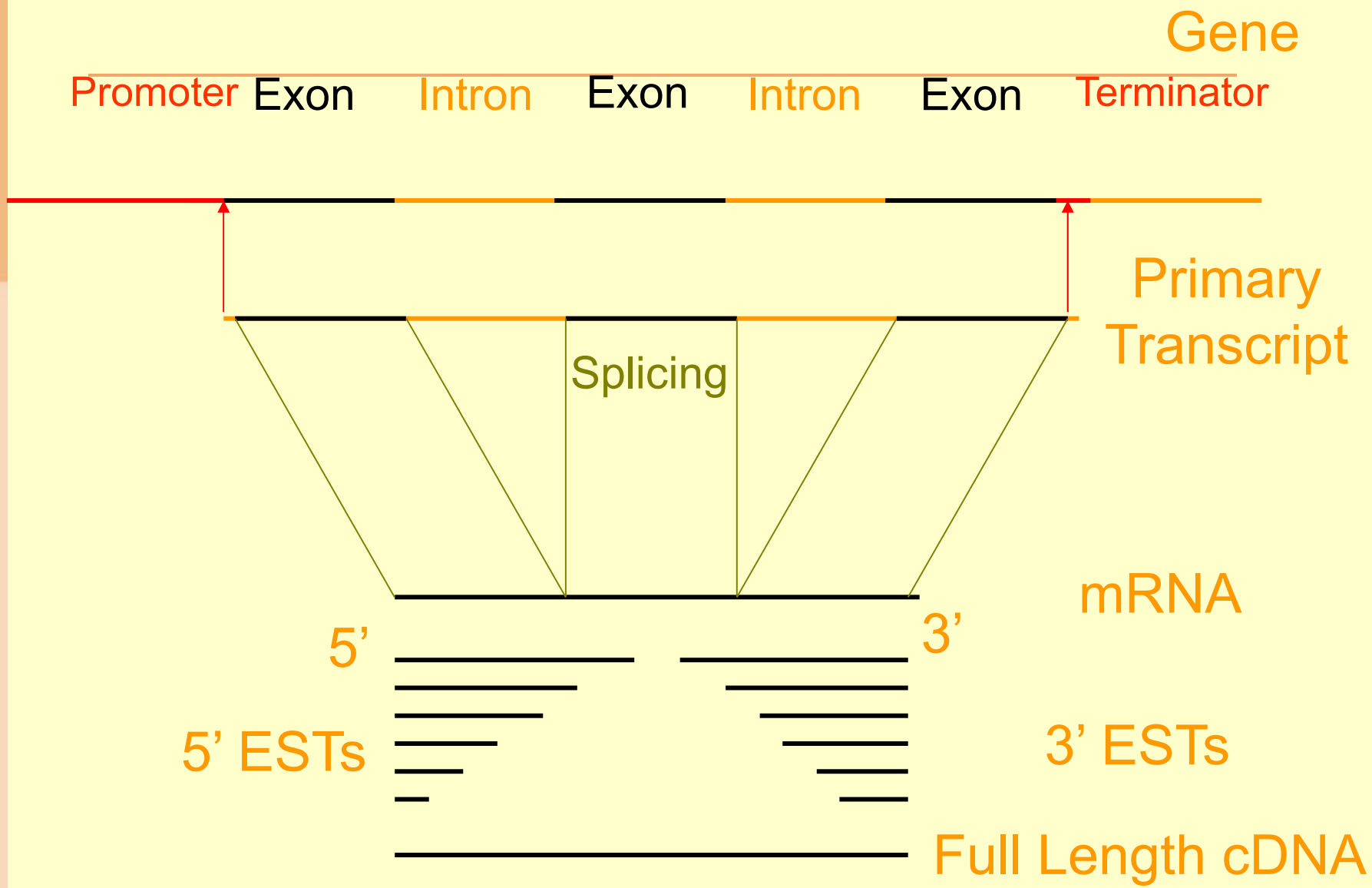
http://genes.mit.edu/GENSCAN.html

Hidden Markov Models of Gene Structure

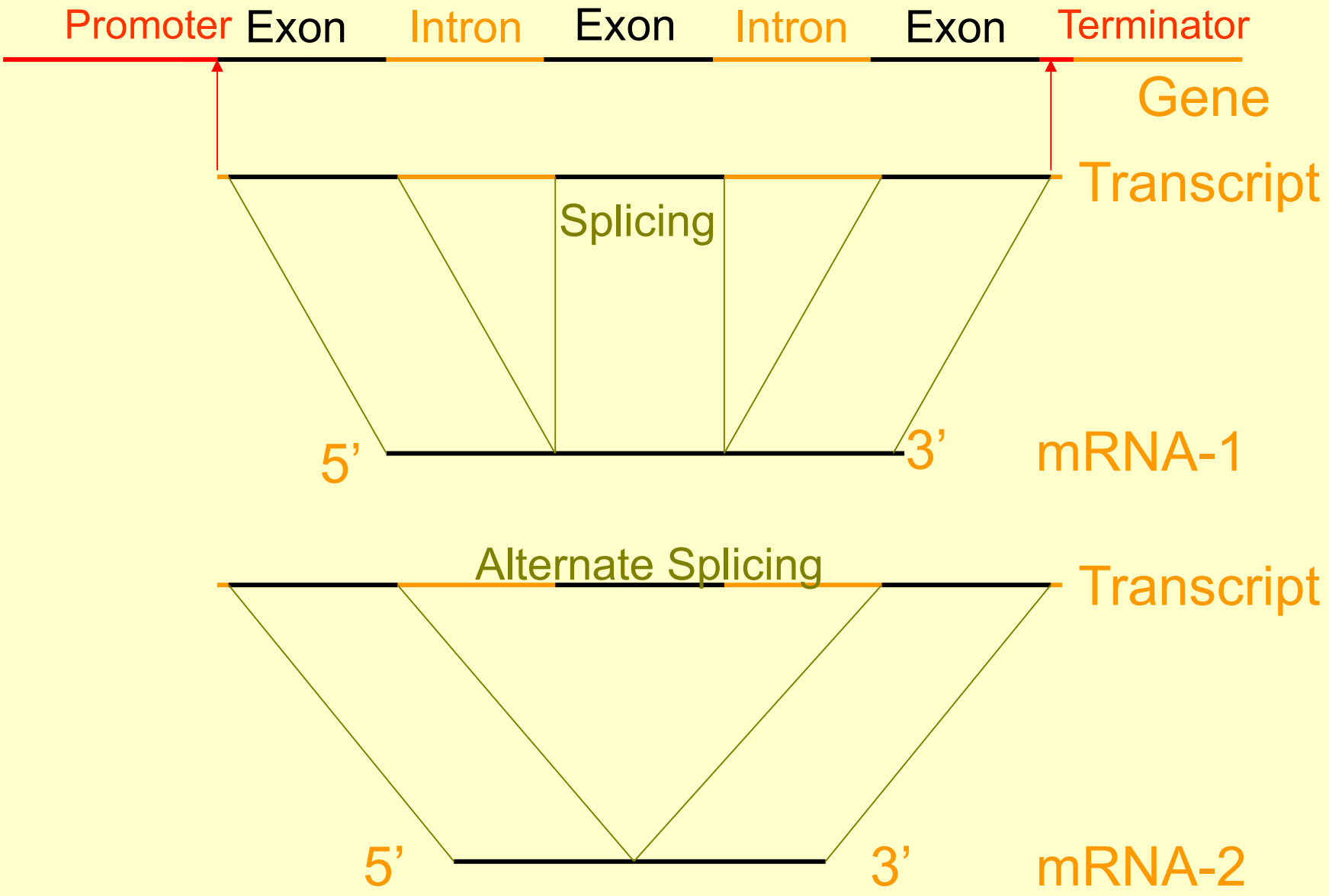
Hidden Markov models of gene structure



ESTs, Full Length cDNA

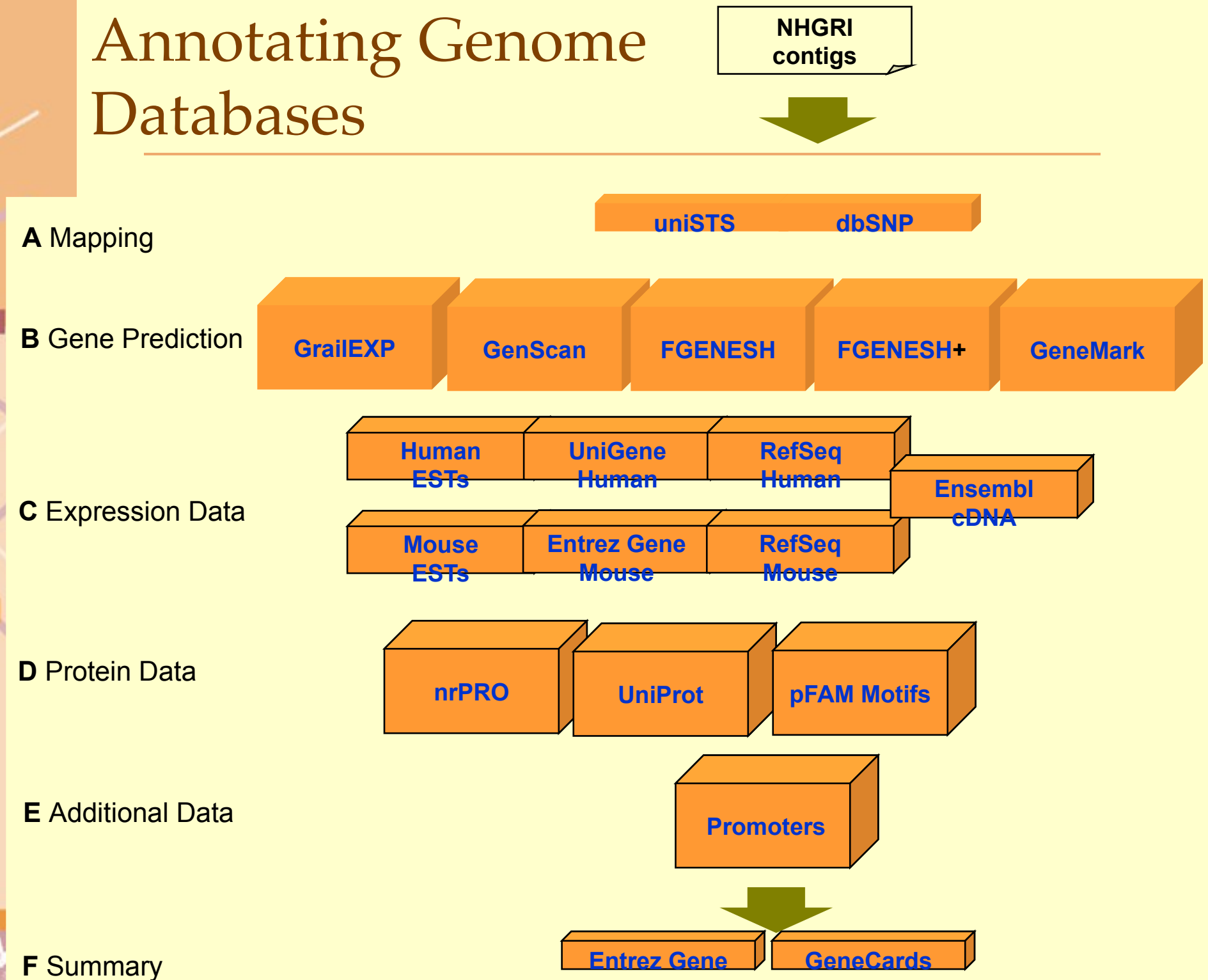


Alternative Splicing Generates Distinct Proteins in Different Tissues



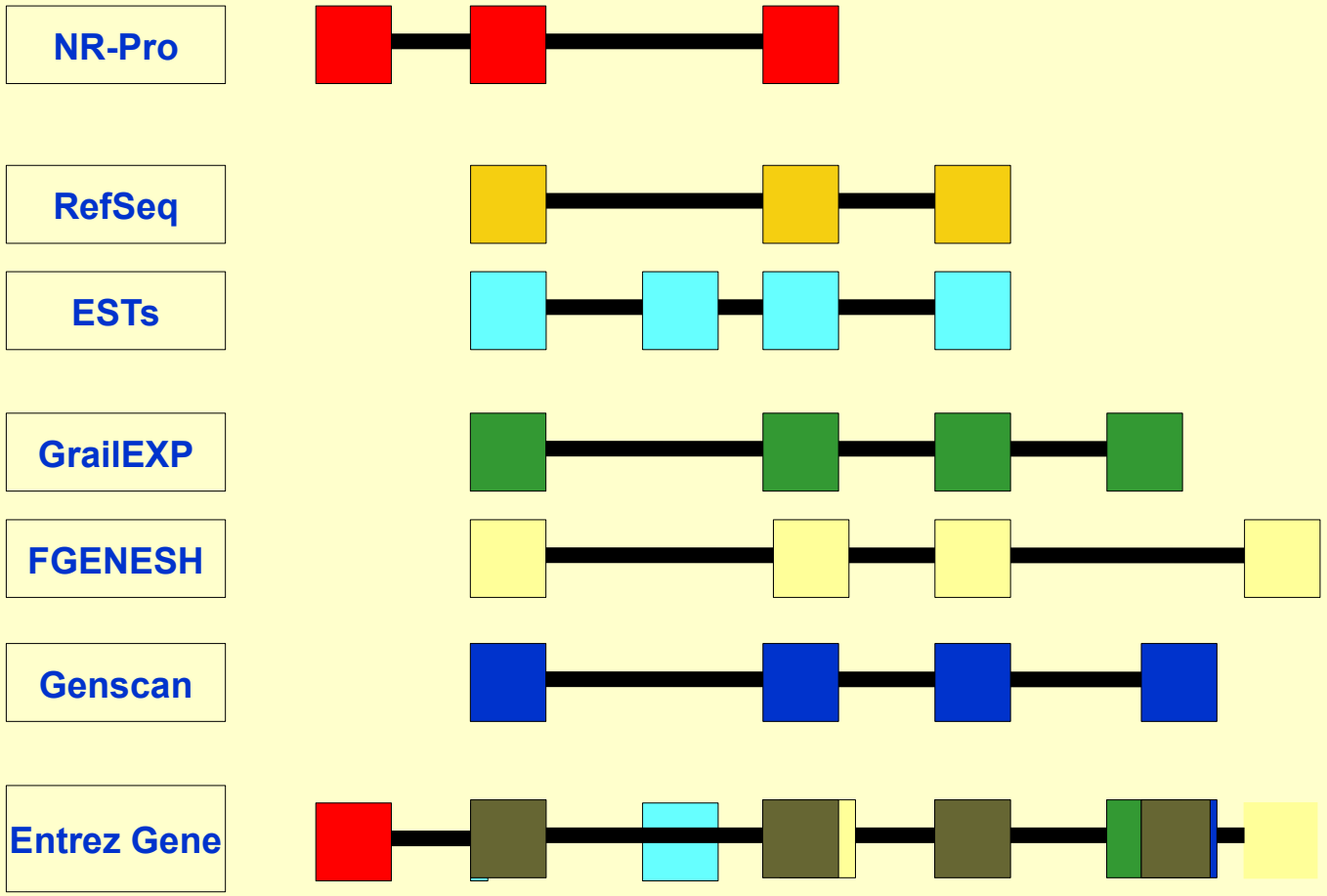
Alternative Splicing Detected in EST Libraries

Annotating Genome Databases



Entrez Gene Loci

http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=gene



NCBI Genome Page

http://www.ncbi.nlm.nih.gov/sites/entrez?db=genome

NCBI

Entrez Genome

My NCBI
Welcome brutlag. [Sign Out]

All Databases | PubMed | Nucleotide | Protein | Genome | Structure | OMIM | PMC | Journals | Books

Search Genome for [ ] Go Clear

Limits | Preview/Index | History | Clipboard | Details

About Entrez

Entrez Genome Help

Submiting Genome Project Genome sequence

Microbial Genome Projects

Genomic BLAST Microbial Eukaryotic

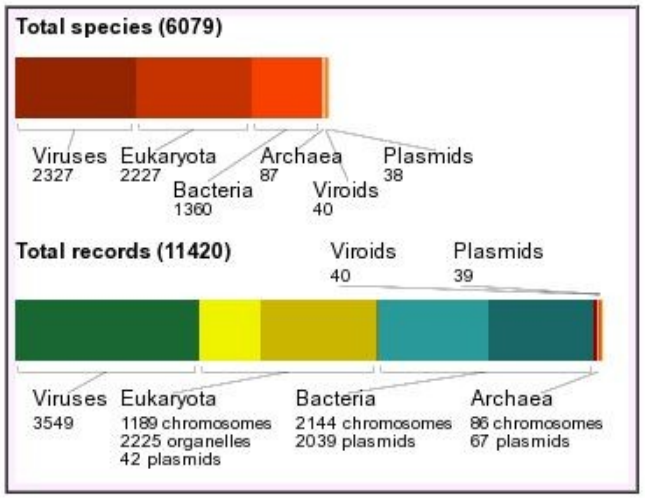
Archaea Chromosome Plasmid DraftAssembly

Bacteria Chromosome Plasmid DraftAssembly

Eukaryota Chromosome Plasmid Organelles

Viruses Phages

The Genome database provides views for a variety of genomes, complete chromosomes, sequence maps with contigs, and integrated genetic and physical maps. The database is organized in six major organism groups: Archaea, Bacteria, Eukaryotae, Viruses, Viroids, and Plasmids and includes complete chromosomes, organelles and plasmids as well as draft genome assemblies.



Genome Sequencing Milestone Reached! There are now 1000 complete Prokaryotic Genomes available in Entrez Genome. See the full list of complete bacterial and archaeal genomes. Microbial Resources are available for search, retrieval, and analysis of all genomes.

Related resources

Entrez Genome Project complete and incomplete large-scale sequencing projects

Entrez Protein Clusters a collection of related protein sequences

Eukaryotic genome projects and sequences

Genomes of Bacillus anthracis reference genome and related sequences

Influenza Virus Resource sequence database and analyses

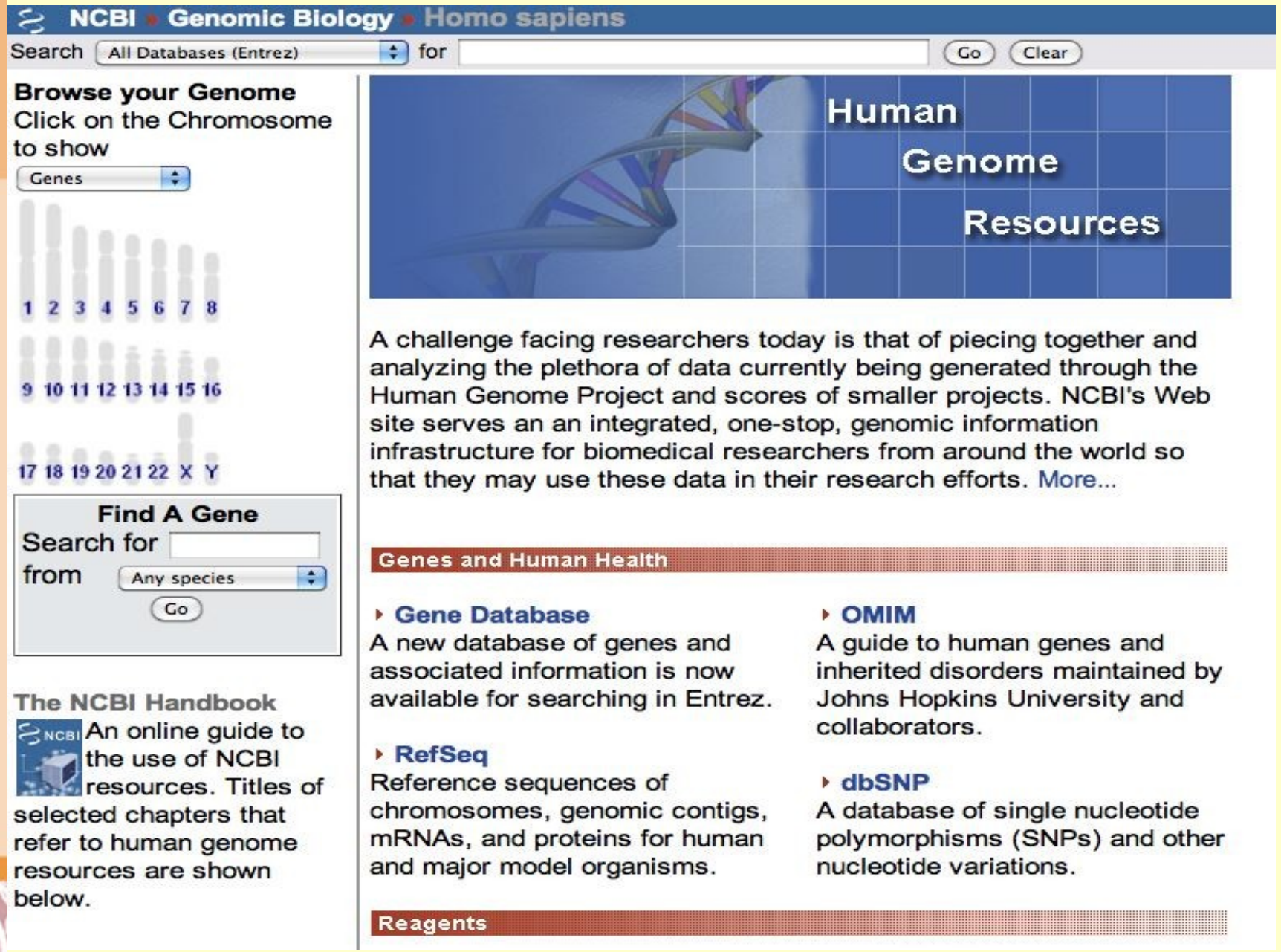
Microbial Genomes reference sequences and resources

Organelle reference sequences and tools

Plant Genomes Central major plant genome

Human Genome Resources

http://www.ncbi.nlm.nih.gov/projects/genome/guide/human/



Entrez Human Genome Projects

http://www.ncbi.nlm.nih.gov/sites/entrez?Db=genomeprj&cmd=ShowDet

NCBI ENTREZ Genome Project connection information discovery

My NCBI
Welcome brutlag. [Sign Out]

All Databases | PubMed | Nucleotide | Protein | Genome | Structure | OMIM | PMC | Journals | Books

Search Genome Project for Go Clear

Limits | Preview/Index | History | Clipboard | Details

Display Overview Show 20 Send to

All: 1 | Environmental: 0 | Eukaryotes: 1 | Prokaryotes: 0

Links

Genome Project > *Homo sapiens* (human)

Resource Links

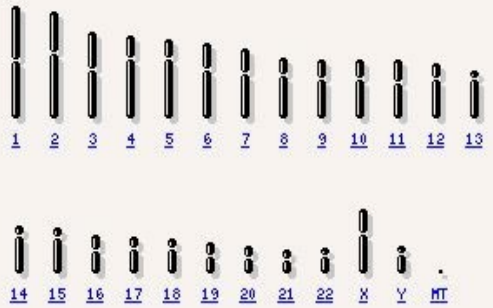
NCBI Resources

- MapViewer
- BLAST genome
- GRC
- RefSeq
- Whole Genome Association (WGA)
- Human Genome Resources
- Consensus CoDing Sequence (CCDS)
- NCBI Handbook

Human genome projects have generated an unprecedented amount of knowledge about human genetics and health.

Project data

Lineage: *Eukaryota*; *Metazoa*; *Chordata*; *Craniata*; *Vertebrata*; *Euteleostomi*; *Mammalia*; *Eutheria*; *Euarchontoglires*; *Primates*; *Haplorrhini*; *Catarrhini*; *Hominidae*; *Homo*; *Homo sapiens*



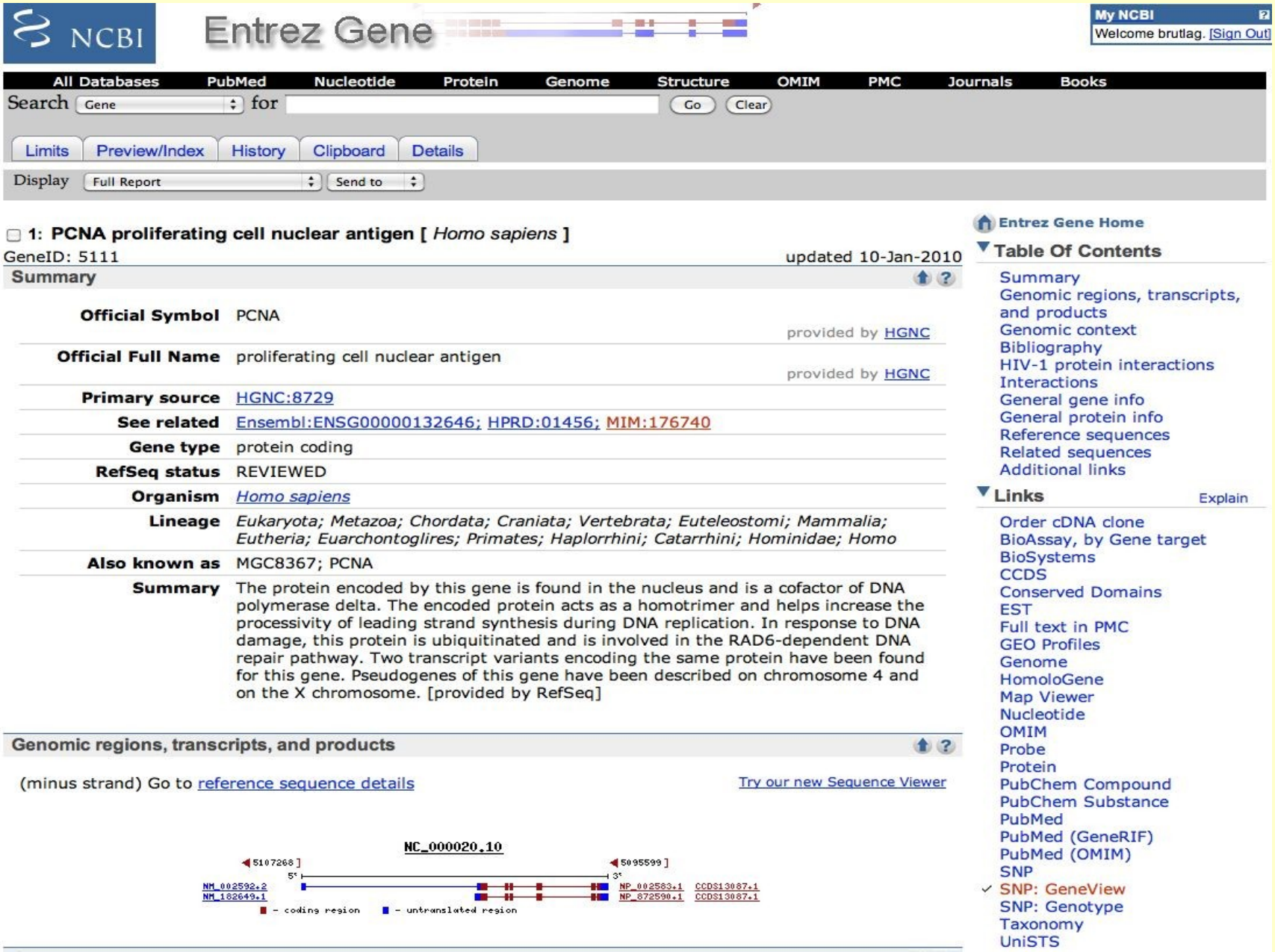
Search Map Viewer for Find Advanced Search

Available maps: View Available Assemblies

Sequence Maps	Cytogenetic maps	Genetic maps	RH maps
27 maps	4 maps	3 maps	7 maps

PCNA Gene in Entrez Gene

http://www.ncbi.nlm.nih.gov/gene/5111



MapViewer of PCNA Gene

NCBI NCBI Map Viewer

PubMed | Entrez | BLAST | OMIM | Taxonomy | Structure

Search [ ] Find | Find in This View | Advanced Search

Human genome overview page (Build 37.1)
Human genome overview page (Build 36.3)
Map Viewer Home
Map Viewer Help
Human Maps Help
FTP
Data As Table View
Maps & Options
Compress Map
Region Shown:
4,800K
5,400K
Go
out
zoom in
You are here:
Ideogram
default
master

Homo sapiens (human) Build 37.1 (Current)

BLAST The Human Genome

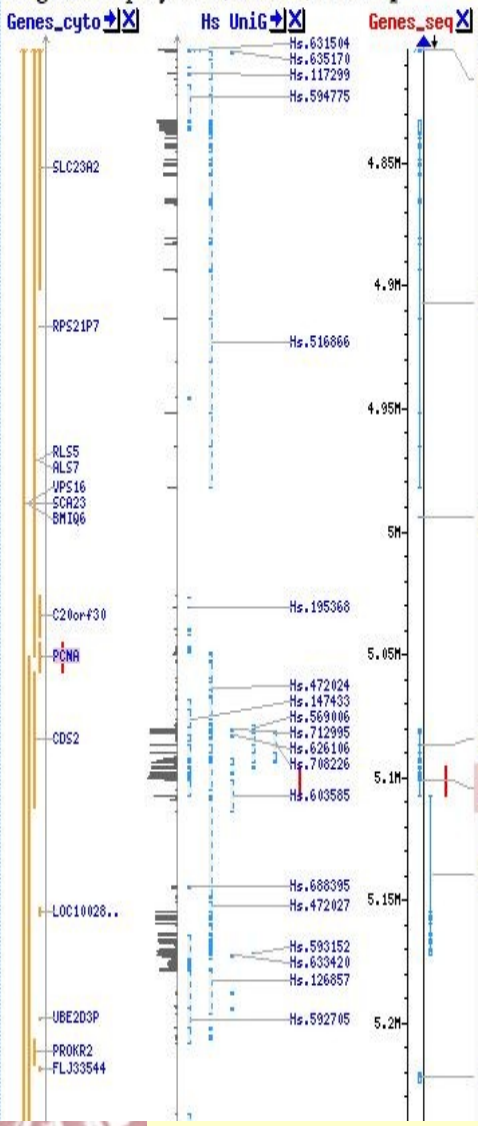
Chromosome: 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 [20] 21 22 X Y MT

Query: PCNA [clear]

Master Map: Genes On Sequence | Summary of Maps | Maps & Options

Region Displayed: 4,800K-5,400K bp

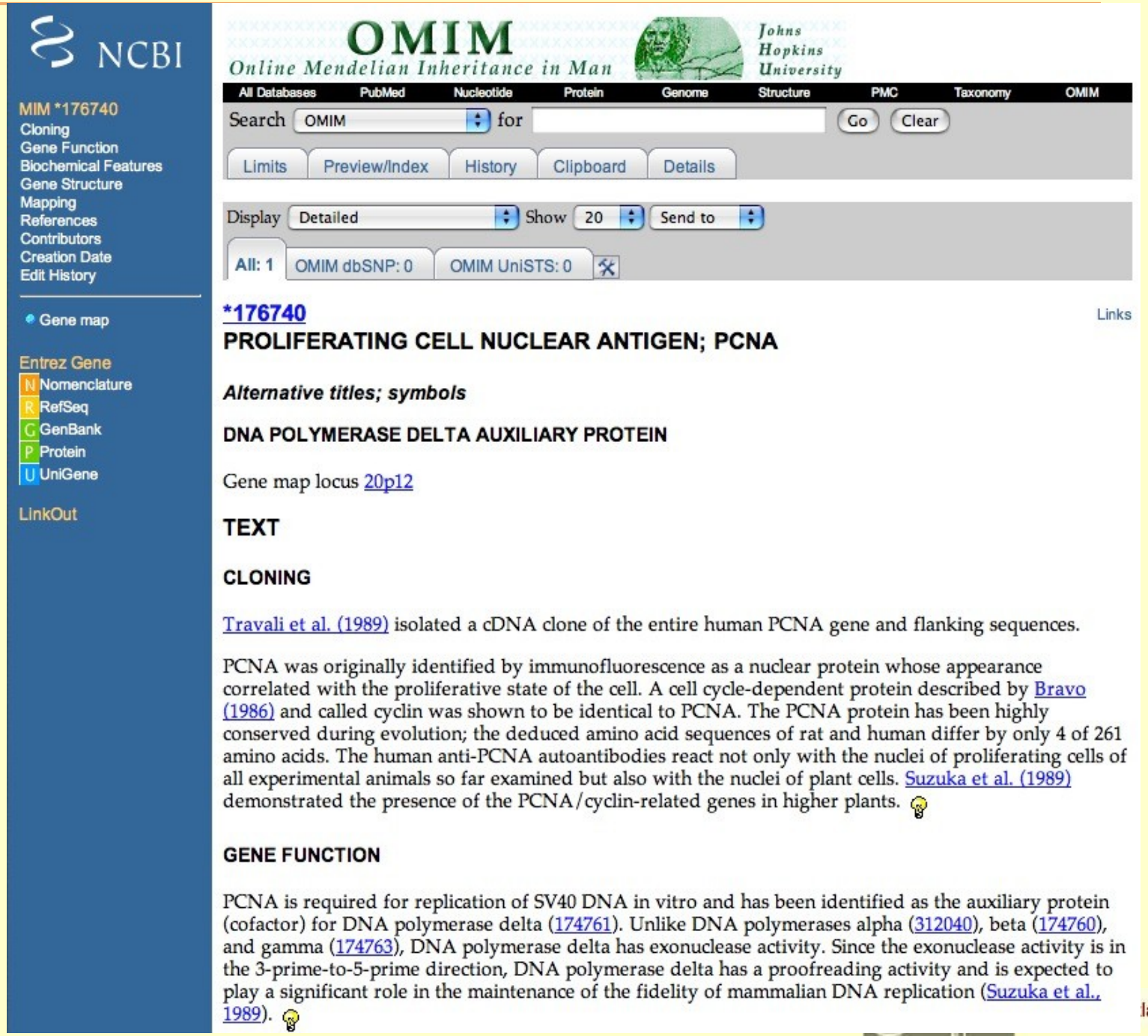
Download/View Sequence/Evidence



Symbol	O	Links	E	Cyto	Description
RASSF2	↑	OMIM HGNC sv pr dl ev mm hm sts CCDS SNP	best RefSeq	20pter-p12.1	Ras association (RalGDS/AF-6) domain fa
SLC23A2	↑	OMIM HGNC sv pr dl ev mm hm sts CCDS SNP	best RefSeq	20p13	solute carrier family 23 (nucleobase transp
RPS21P7	↑	HGNC sv dl ev mm	best RefSeq	20p13	ribosomal protein S21 pseudogene 7
C20orf30	↑	HGNC sv pr dl ev mm hm sts CCDS SNP	best RefSeq	20p13	chromosome 20 open reading frame 30
PCNA	↑	OMIM HGNC sv pr dl ev mm hm sts CCDS SNP	best RefSeq	20pter-p12	proliferating cell nuclear antigen
CDS2	↓	OMIM HGNC sv pr dl ev mm hm sts CCDS SNP	best RefSeq	20p13	CDP-diacylglycerol synthase (phosphatida
LOC100288854	↑	sv dl ev mm	mRNA	20	similar to lactation elevated 1 (predicted)

OMIM Entry for PCNA

http://www.ncbi.nlm.nih.gov/entrez/dispomim.cgi?id=176740



NCBI

MIM *176740
Cloning
Gene Function
Biochemical Features
Gene Structure
Mapping
References
Contributors
Creation Date
Edit History

Gene map

Entrez Gene
Nomenclature
RefSeq
GenBank
Protein
UniGene

LinkOut

OMIM
Online Mendelian Inheritance in Man
Johns Hopkins University

All Databases | PubMed | Nucleotide | Protein | Genome | Structure | PMC | Taxonomy | OMIM

Search OMIM for [ ] Go Clear

Limits | Preview/Index | History | Clipboard | Details

Display Detailed Show 20 Send to

All: 1 | OMIM dbSNP: 0 | OMIM UniSTS: 0

*176740

Links

PROLIFERATING CELL NUCLEAR ANTIGEN; PCNA

Alternative titles; symbols

DNA POLYMERASE DELTA AUXILIARY PROTEIN

Gene map locus 20p12

TEXT

CLONING

Travali et al. (1989) isolated a cDNA clone of the entire human PCNA gene and flanking sequences.

PCNA was originally identified by immunofluorescence as a nuclear protein whose appearance correlated with the proliferative state of the cell. A cell cycle-dependent protein described by Bravo (1986) and called cyclin was shown to be identical to PCNA. The PCNA protein has been highly conserved during evolution; the deduced amino acid sequences of rat and human differ by only 4 of 261 amino acids. The human anti-PCNA autoantibodies react not only with the nuclei of proliferating cells of all experimental animals so far examined but also with the nuclei of plant cells. Suzuka et al. (1989) demonstrated the presence of the PCNA/cyclin-related genes in higher plants.

GENE FUNCTION

PCNA is required for replication of SV40 DNA in vitro and has been identified as the auxiliary protein (cofactor) for DNA polymerase delta (174761). Unlike DNA polymerases alpha (312040), beta (174760), and gamma (174763), DNA polymerase delta has exonuclease activity. Since the exonuclease activity is in the 3-prime-to-5-prime direction, DNA polymerase delta has a proofreading activity and is expected to play a significant role in the maintenance of the fidelity of mammalian DNA replication (Suzuka et al., 1989).

Hugo Nomenclature for PCNA

http://www.genenames.org/data/hgnc_data.php?hgnc_id=8729

Symbol Report: PCNA

About HGNC | Gene Search | Guidelines | Gene Submission | Downloads | Home

Giving unique and meaningful names to every human gene

Quick Gene Search

Core Data	
Approved Symbol +	PCNA
Approved Name +	proliferating cell nuclear antigen
HGNC ID +	HGNC:8729
Status +	Approved
Chromosome +	20pter-p12
Previous Symbols +	
Previous Names +	
Aliases +	
Name Aliases +	
Locus Type +	gene with protein product

Gene Symbol Links

GENATLAS GeneCards GeneClinics/GeneTests GoPubmed

HCOP H-InvDB Treefam wikigenes

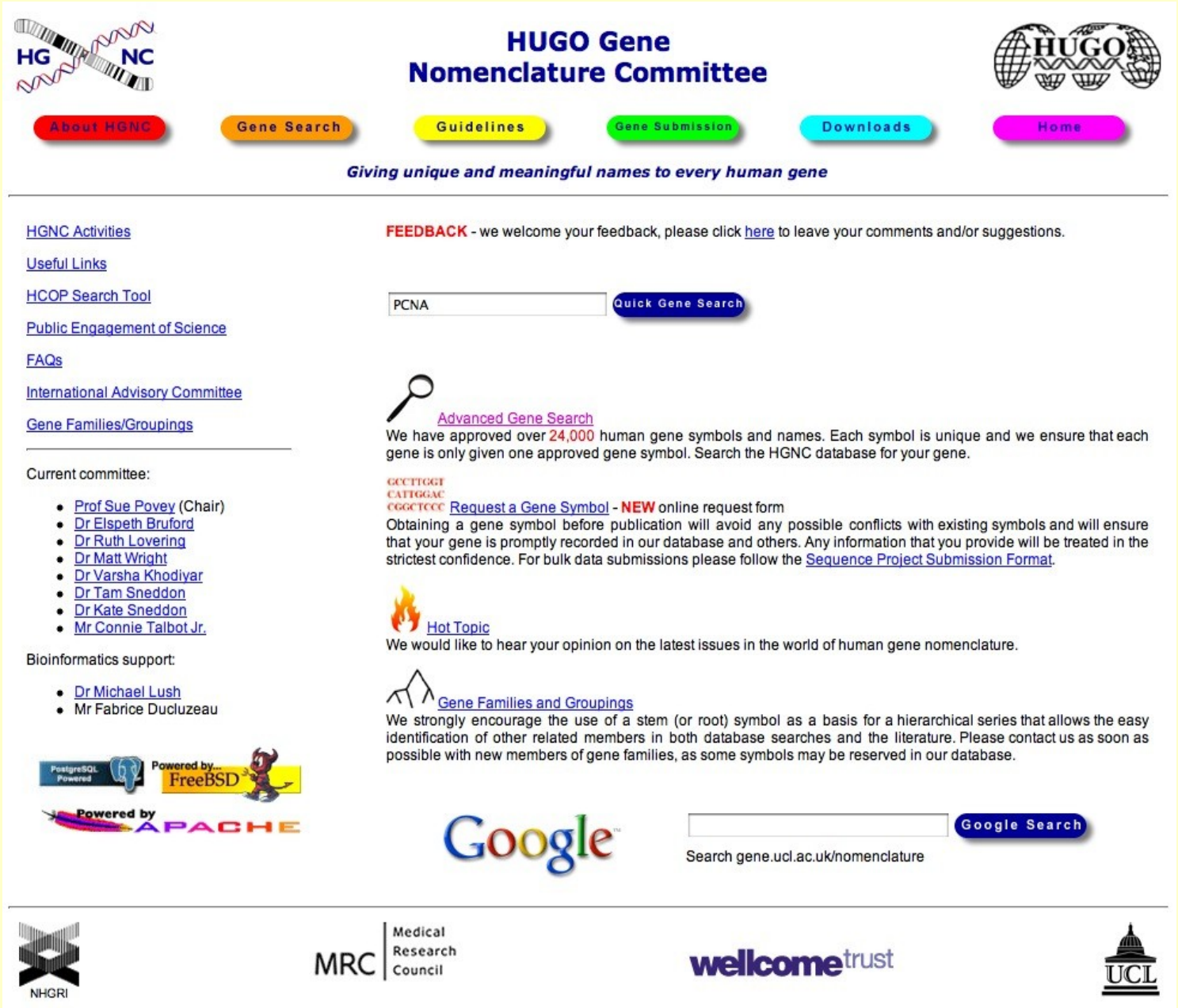
Specialist Database Links

COSMIC

Database Links				
Accession Numbers +				
J04718	GenBank	EMBL	DDBJ	UCSC
Mouse Genome Database ID +				
MGI:97503	MGD ID			
Rat Genome Database ID (mapped data supplied by RGD) +				
RGD:3269	RGD ID			
CCDS IDs +				
CCDS13087.1	CCDS ID			
Pubmed IDs +				
2565339	PMID		CiteXplore	
VEGA IDs +				
OTTHUMG00000031798	VEGA GeneView			
Ensembl ID (mapped data supplied by Ensembl) +				
ENSG00000132646	Ensembl GeneView		UCSC	
Entrez Gene ID (mapped data supplied by NCBI) +				
5111	Gene		Map Viewer	
RefSeq (mapped data supplied by NCBI) +				
NM_182649	GenBank	EMBL	DDBJ	UCSC
OMIM ID (mapped data supplied by NCBI) +				
176740	OMIM			
UCSC ID (mapped data supplied by UCSC) +				
uc002wlp.2	UCSC Index			
UniProt ID (mapped data supplied by UniProt) +				
P12004	UniProt		UCSC	

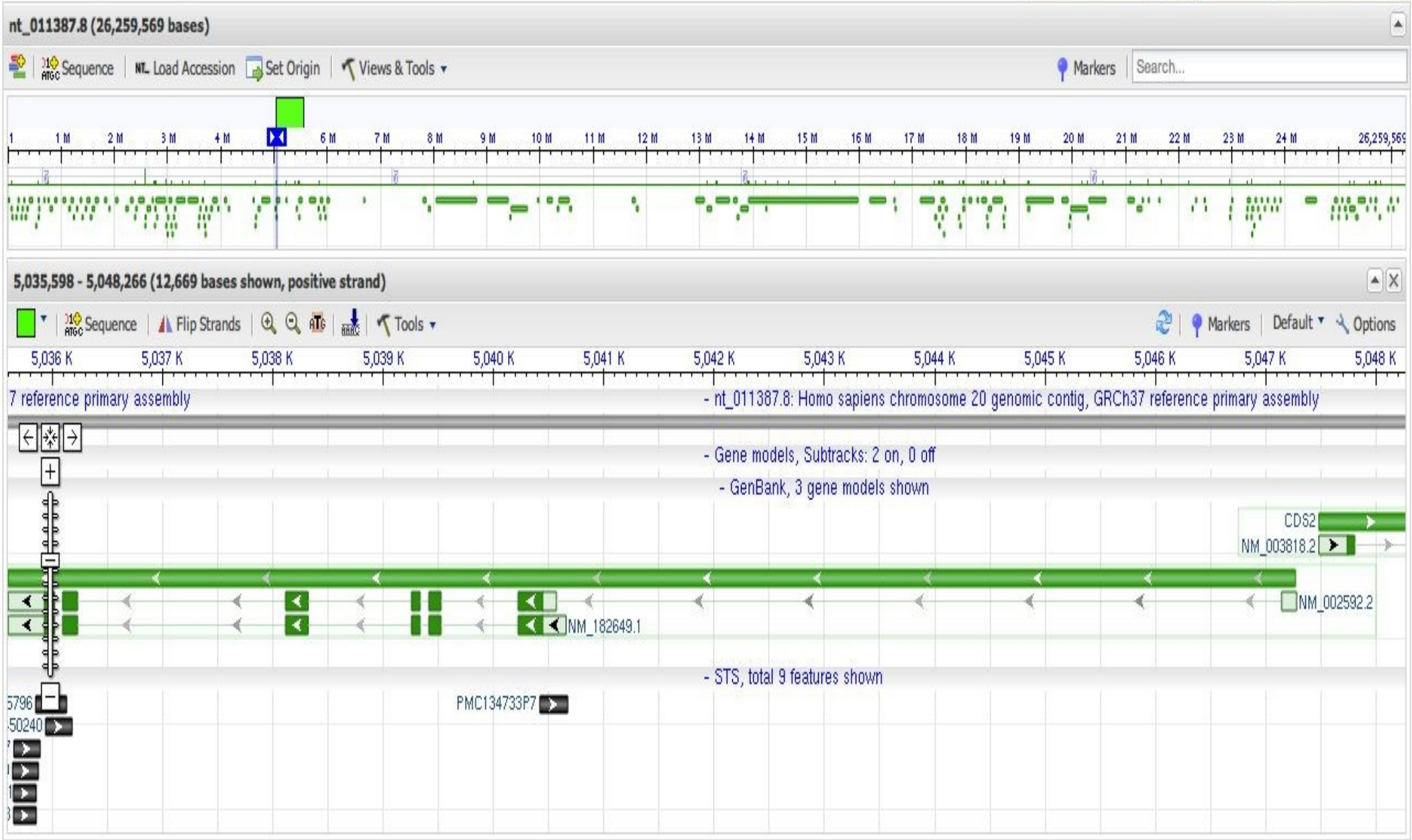
HUGO Home Page

http://www.genenames.org/



Homo sapiens chromosome 20 genomic contig, GRCh37 reference primary assembly

gi|27501067|ref|NT_011387.8|



Evidence Viewer for PCNA

http://www.ncbi.nlm.nih.gov/sutils/evv.cgi?taxid=9606&contig=NT_011387.8&gene=PCN

Key for display of mRNAs aligning in this region: MapView Evidence Viewer Help

- Genomic sequence (C)
- model exons, single (M)
- mRNA exons, single (G, R)
- model exons, overlapping (M)
- mRNA exons, overlapping (G, R)

C = contig; M = model mRNA; R = RefSeq mRNA; G = GenBank mRNA

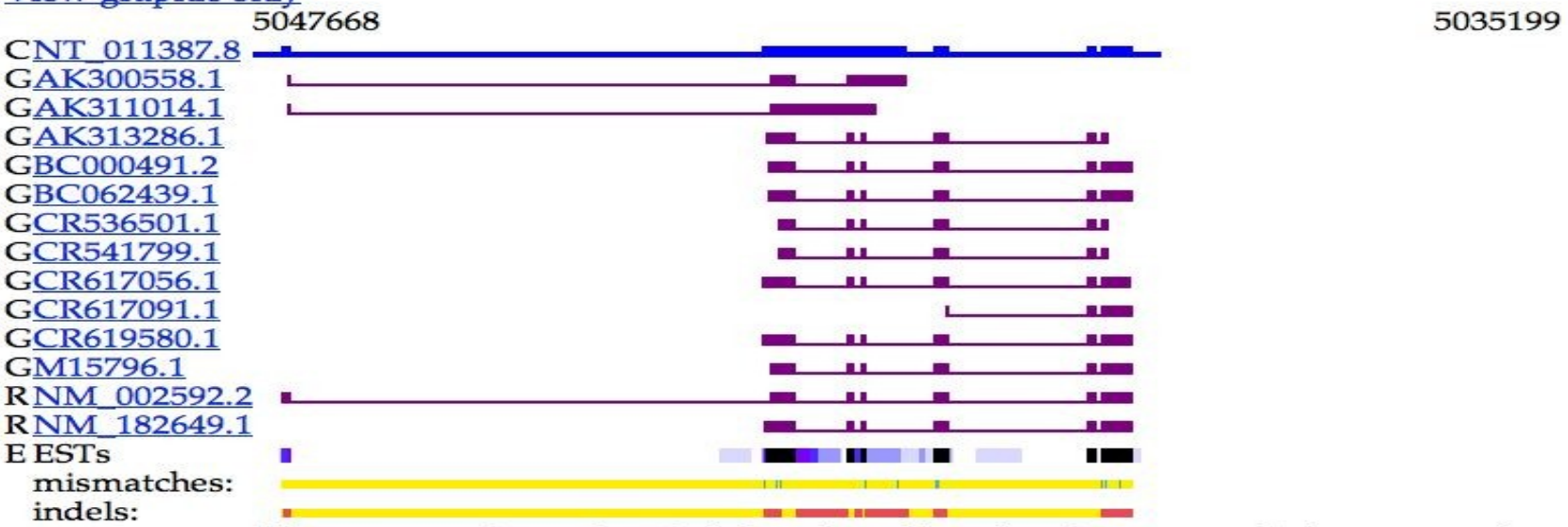
R = new since last genome build; R = updated since last genome build

EST density key (E):

- 1 EST
- 2-5 ESTs
- 6-20 ESTs
- 21-99 ESTs
- >100 ESTs

5 exons and 1 gene found in this genomic region spanning 12470 bp.

View graphic only



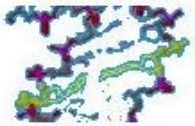
Mouse over mismatches, indels and unaligned regions to see their exon number.

SNP Polymorphisms in PCNA

http://www.ncbi.nlm.nih.gov/SNP/snp_ref.cgi?locusId=5111

NCBI

Single Nucleotide Polymorphism



PubMed Nucleotide Protein Genome Structure PopSet Taxonomy OMIM Books SNP

Search for SNP on NCBI Reference Assembly

Search Entrez SNP for Go

BUILD 130

Have a question about dbSNP? Try searching the SNP FAQ Archive!

Go

GENERAL
HUMAN VARIATION
Search, Annotate, Submit NEW
Annotate and Submit Batch Data with Clinical Impact NEW
SNP SUBMISSION
DOCUMENTATION
SEARCH
RELATED SITES

SNP linked to Gene PCNA(geneID:5111) Via Contig Annotation

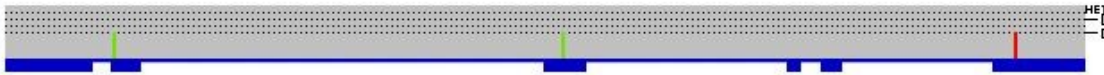
Send rs# on all gene models to Batch Query Download all rs# to file. Genotype

Gene Model (mRNA alignment) information from genome sequence

Total gene model (contig mRNA transcript):				2		
mrna	transcript	protein	mrna orientation	Contig	Contig Label	List SNP
NM_182649.1	minus strand	NP_872590.1	reverse	NT_011387.8	reference	<- currently shown
NM_002592.2	minus strand	NP_002583.1	reverse	NT_011387.8	reference	View snp on GeneModel

Include clinically associated in gene region cSNP has frequency double hit refresh

gene model (contig mRNA transcript):	Contig Label	Contig	mrna	protein	mrna orientation	transcript	snp count
	reference	NT_011387.8	NM_182649.1	NP_872590.1	reverse	minus strand	3, coding



Color Legend

Region	Contig position	mRNA pos	dbSNP rs# cluster id	Hetero-zygosity	Validation	3D	Clinically Associated	Function	dbSNP allele	Protein residue	Codon pos	Amino acid pos	PubMed
exon_5	5036099	941	rs17353	0.004		Yes		synonymous	G	Pro [P]	3	234	
								contig reference	C	Pro [P]	3	234	
exon_4	5038203	734	rs17350	N.D.		Yes		synonymous	T	Asp [D]	3	165	
								contig reference	C	Asp [D]	3	165	
exon_1	5040328	356	rs1050525	N.D.	H			missense	A	Arg [R]	3	39	
								contig reference	C	Ser [S]	3	39	
exon_1	5040444	204						start codon				1	

Entrez Gene

NCBI

NCBI

Entrez Gene

My NCBI
Welcome brutlag. [Sign Out]

All Databases | PubMed | Nucleotide | Protein | Genome | Structure | OMIM | PMC | Journals | Books

Search Gene for Go Clear

Limits | Preview/Index | History | Clipboard | Details

Entrez Gene
- Home
- About
- FAQ
- Help
- Gene Handbook
- Statistics
- Downloads (FTP)

Mailing Lists
- Gene
- RefSeq

Feedback
- Help Desk
- Corrections
- About GeneRIFs

Related Sites
- BLAST
- Entrez Genome
- Genome Projects
- Genomic Biology
- GEO
- HomoloGene
- Map Viewer
- OMIM
- Probe
- RefSeq

Entrez Gene is a searchable database of genes, from RefSeq genomes, and defined by sequence and/or located in the NCBI Map Viewer

News Links to Ensembl; New rnatype Properties News archives...

Sample Searches

Find genes by...	Search text
free text	human muscular dystrophy
partial name and multiple species	transporter[title] AND ("Drosophila melanogaster"[orgn] OR "Mus musculus"[orgn])
chromosome and symbol	(II[chr] OR 2[chr]) AND adh*[sym]
associated sequence accession number	M11313[accn]
gene name (symbol)	BRCA1[sym]
publication (PubMed ID)	11331580[PMID]
Gene Ontology (GO) terms or identifiers	"cell adhesion"[GO] 10030[GO]
Genes with variants of clinical significance (under development)	gene_snp_clin[filter]
chromosome and species	Y[CHR] AND human[ORGN]
Enzyme Commission (EC) numbers	1.9.3.1[EC]

more ways to search...

About Entrez Gene

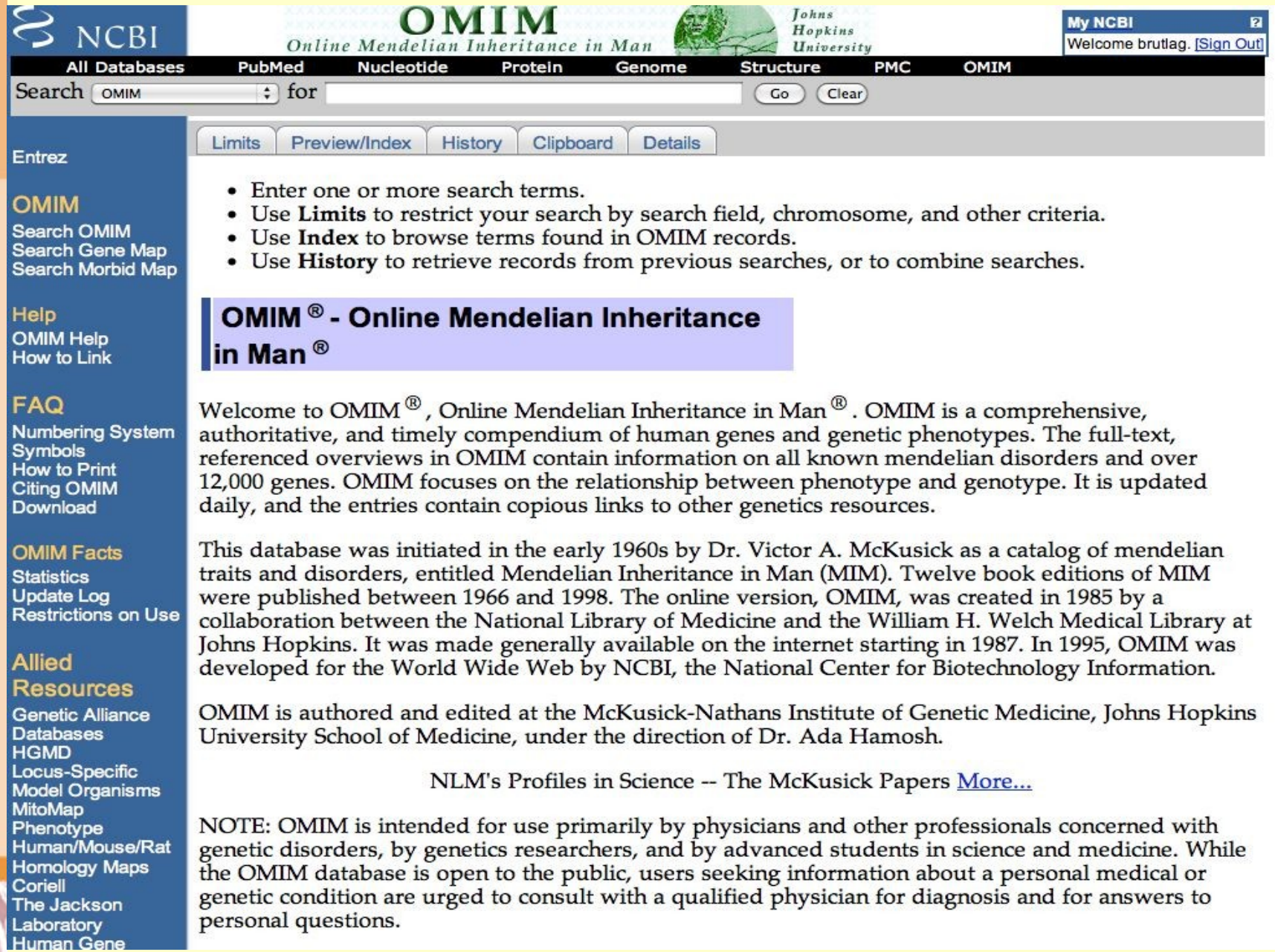
- Entrez Gene Help
- Frequently Asked Questions
- Entrez Gene: gene-centered information at NCBI. *Nucleic Acids Res.* 2005 Jan 1;33:D54-8.
- General help on the Entrez search and retrieval system
- NCBI Handbook Chapter on Entrez Gene (download PDF)
- Download data via FTP
- View statistics for Entrez Gene

Corrections • Additions • Feedback

- Report a new gene
- Report a new splice variant
- How to add information about function (GeneRIF)
- Correct or update a Gene record
- Correct or update a reference sequence
- Report a publication or GeneRIF error
- Report a search or display problem
- Report an FTP problem
- Make a suggestion for Entrez Gene

OMIM Home Page

http://www.ncbi.nlm.nih.gov/omim



OMIM Coverage

http://www.ncbi.nlm.nih.gov/Omim/mimstats.html

OMIM Online Mendelian Inheritance in Man — Johns Hopkins University

PubMed | Nucleotide | Protein | Genome | Structure | PMC | Taxonomy | OMIM

OMIM Statistics for January 11, 2010

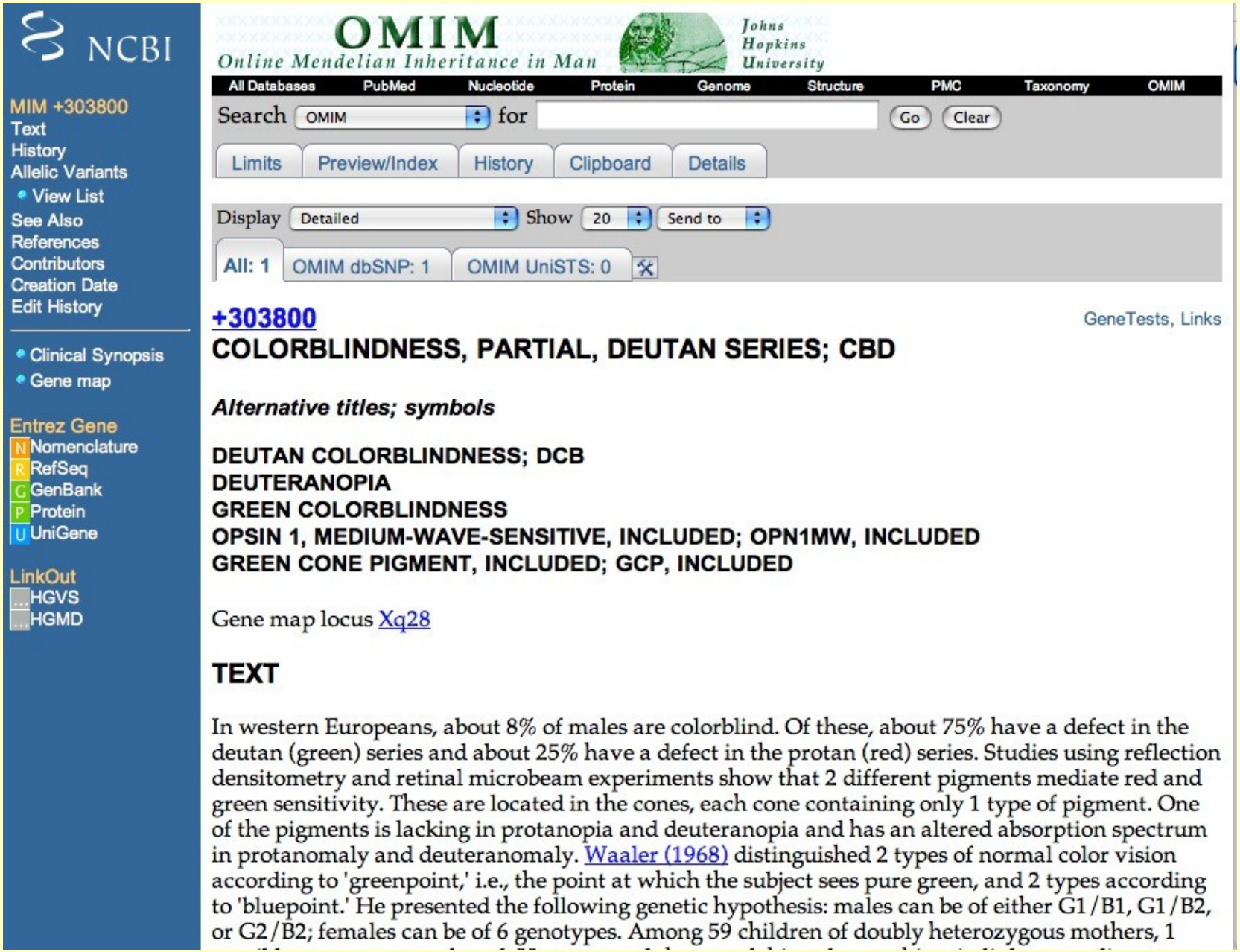
Number of Entries

	Autosomal	X-Linked	Y-Linked	Mitochondrial	Total
* Gene with known sequence	12327	609	48	35	13019
+ Gene with known sequence and phenotype	325	19	0	2	346
# Phenotype description, molecular basis known	2438	216	4	26	2684
% Mendelian phenotype or locus, molecular basis unknown	1641	142	5	0	1788
Other, mainly phenotypes with suspected mendelian basis	1866	138	2	0	2006
Total	18597	1124	59	63	**19843**

Colorblindness in OMIM

Colorblindness in OMIM

http://www.ncbi.nlm.nih.gov/entrez/dispomim.cgi?id=303800



NCBI

MIM +303800
Text
History
Allelic Variants
View List
See Also
References
Contributors
Creation Date
Edit History

Clinical Synopsis
Gene map

Entrez Gene
Nomenclature
RefSeq
GenBank
Protein
UniGene

LinkOut
HGVS
HGMD

OMIM
Online Mendelian Inheritance in Man
Johns Hopkins University

All Databases | PubMed | Nucleotide | Protein | Genome | Structure | PMC | Taxonomy | OMIM

Search OMIM for [ ] Go Clear

Limits | Preview/Index | History | Clipboard | Details

Display Detailed Show 20 Send to

All: 1 | OMIM dbSNP: 1 | OMIM UniSTS: 0

+303800

GeneTests, Links

COLORBLINDNESS, PARTIAL, DEUTAN SERIES; CBD

Alternative titles; symbols

DEUTAN COLORBLINDNESS; DCB
DEUTERANOPIA
GREEN COLORBLINDNESS
OPSIN 1, MEDIUM-WAVE-SENSITIVE, INCLUDED; OPN1MW, INCLUDED
GREEN CONE PIGMENT, INCLUDED; GCP, INCLUDED

Gene map locus Xq28

TEXT

In western Europeans, about 8% of males are colorblind. Of these, about 75% have a defect in the deutan (green) series and about 25% have a defect in the protan (red) series. Studies using reflection densitometry and retinal microbeam experiments show that 2 different pigments mediate red and green sensitivity. These are located in the cones, each cone containing only 1 type of pigment. One of the pigments is lacking in protanopia and deuteranopia and has an altered absorption spectrum in protanomaly and deuteranomaly. Waaler (1968) distinguished 2 types of normal color vision according to 'greenpoint,' i.e., the point at which the subject sees pure green, and 2 types according to 'bluepoint.' He presented the following genetic hypothesis: males can be of either G1/B1, G1/B2, or G2/B2; females can be of 6 genotypes. Among 59 children of doubly heterozygous mothers, 1

Human Gene Mutation Database

http://www.hgmd.cf.ac.uk/ac/gene.php?gene=OPN1MW

The Human Gene Mutation Database at the Institute of Medical Genetics in Cardiff

BIOBASE BIOLOGICAL DATABASES

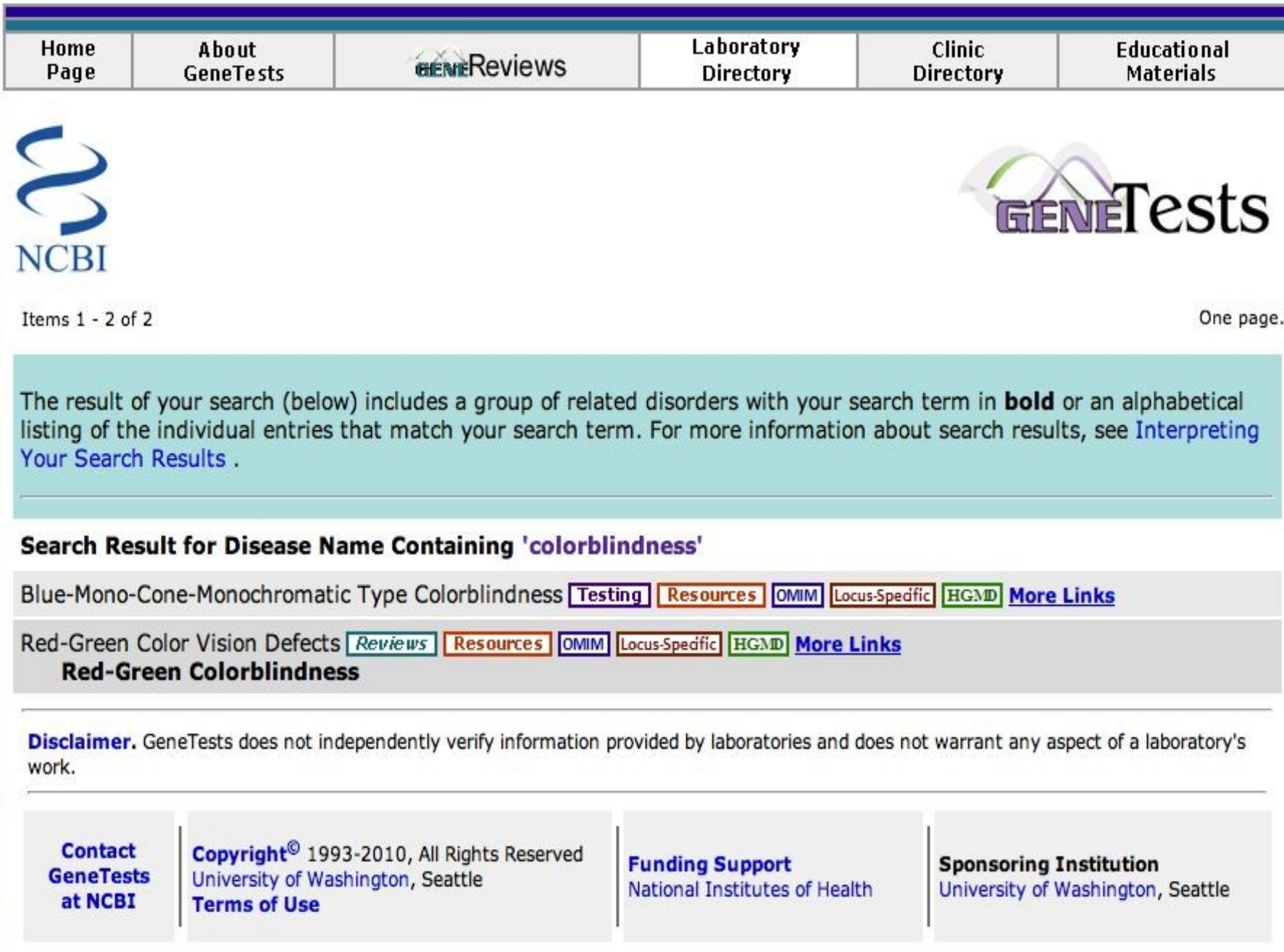
Gene Symbol	Chromosomal location	Gene name	cDNA sequence	Extended cDNA	Splice junctions	Mutation map
OPN1MW	Xp28	Opsin 1 (cone pigments), medium-wave-sensitive (color blindness, deutan) GCP	Get cDNA	Not available	Not available	Mutation map

Mutation type	Number of mutations	Mutation data by type (register) or (login)
Missense/nonsense	3	Get mutations
Splicing	0	No mutations
Regulatory	1	Get mutations
Small deletions	0	No mutations
Small insertions	0	No mutations
Small indels	0	No mutations
Gross deletions	2	Get mutations
Gross insertions	0	No mutations
Complex rearrangements	4	Get mutations
Repeat variations	0	No mutations
Public total (HGMD Professional 6.4 total)	**10** (11)	

Disease/phenotype	Number of mutations	Mutation data by disease/phenotype
Deuteranopia	5	BIOBASE Feature available to subscribers
Visual dichromacy	2	BIOBASE Feature available to subscribers
Blue cone monochromatism	1	BIOBASE Feature available to subscribers
Deutan color-vision deficiency, association with	1	BIOBASE Feature available to subscribers
Trichromacy, deutan	1	BIOBASE Feature available to subscribers

GeneTests

http://www.genetests.org/



GeneReviews on Red-Green Color Vision Defects

http://www.ncbi.nlm.nih.gov/bookshelf/br.fcgi?book=gene&part=rgcb

Red-Green Color Vision Defects

[*Includes: Red-Green Color Blindness*]

Samir S Deeb, PhD

Research Professor, Medicine and Genome Sciences
University of Washington
Seattle, W
sdeeb@u.washington.edu

Arno G Motulsky, MD

Professor Emeritus (Active), Medicine and Genome Sciences
University of Washington
Seattle, WA
agmot@u.washington.edu

Initial Posting: September 19, 2005.

All GeneReviews
Disable Glossary Links
Print this GeneReview

In this GeneReview

Summary
Diagnosis
Clinical Description
Differential Diagnosis
Management
Genetic Counseling
Molecular Genetics
Resources
References
Chapter Notes

Related to this GeneReview

Consumer Resources
OMIM
Gene

Summary

Disease characteristics. Hereditary red-green color vision defects are manifest in early infancy, mostly in males; the condition is not accompanied by ophthalmologic or other associated clinical abnormalities. Most individuals with protanomalous and deuteranomalous color vision defects (i.e., anomalous trichromats) have no problems in naming colors; some males with mildly defective red-green color vision may not be aware of it until

Genes & Disease

http://www.ncbi.nlm.nih.gov/bookshelf/br.fcgi?book=gnd&part=gnd_book_info

Genes and Disease NCBI

Search This book Bookshelf PubMed > Search

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 X Y

Click on a Chromosome | View Contents | Download PDF | Download PDA | Browse more Books

Navigation

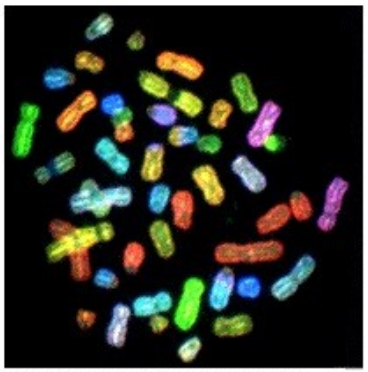
- *About this book*
- Blood and Lymph Diseases
- Cancers
- The Digestive System
- Ear, Nose, and Throat
- Diseases of the Eye
- Female-Specific Diseases
- Glands and Hormones
- The Heart and Blood Vessels
- Diseases of the Immune System
- Male-Specific Diseases
- Muscle and Bone
- Neonatal Diseases
- The Nervous System
- Nutritional and Metabolic Diseases
- Respiratory Diseases
- Skin and Connective Tissue
- Chromosome Map

Introduction to Genes and Disease

Genes and Disease is a collection of articles that discuss genes and the diseases that they cause. These genetic disorders are organized by the parts of the body that they affect. As some diseases affect various body systems, they appear in more than one chapter.

With each genetic disorder, the underlying mutation(s) is discussed, along with clinical features and links to key websites. You can browse through the articles online, and you can also download a printable file (PDF) of each chapter.

From *Genes and Disease* you can delve into many online related resources with free and full access. For example, you can visit the human genome to see the location of the genes implicated in each disorder. You can also find related gene sequences in different organisms. And for the very latest information, you can search for complete research articles, and look in other books in the NCBI Bookshelf.



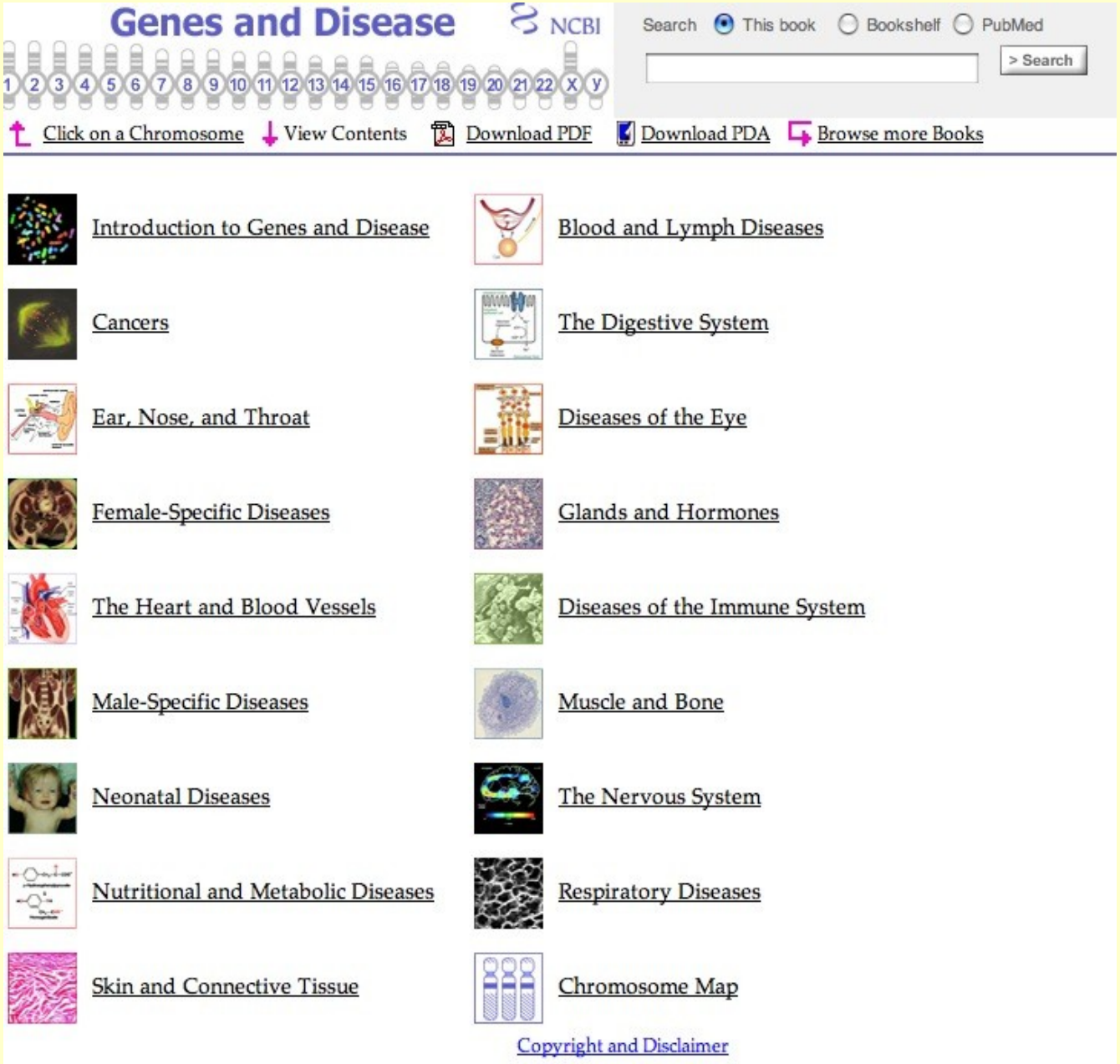
This photograph shows a complete set of chromosomes from an acute promyelocytic leukemia (APL) patient. A new technique called chromosome painting allows visual distinction between chromsomes and can be used to show the chromosome translocations that frequently occur in human cancers. In the case of APL, chromosome 13 is lost, there is a translocation between chromosomes 7 and 15, translocation between chromosomes 11, 15, 17, and between chromosomes 9 and 18. (Look for chromosomes painted with more than one color.) With thanks to Thomas Ried, National Human Genome Research Institute, NIH, for supplying the picture.

Currently over 80 genetic disorders have been summarized, and the content of *Genes and Disease* is continually growing. Your ideas and suggestions are welcome. You can contact us at: info@ncbi.nlm.nih.gov.

Copyright and Disclaimer

Genes & Disease Table of Contents

http://www.ncbi.nlm.nih.gov/bookshelf/br.fcgi?book=gnd

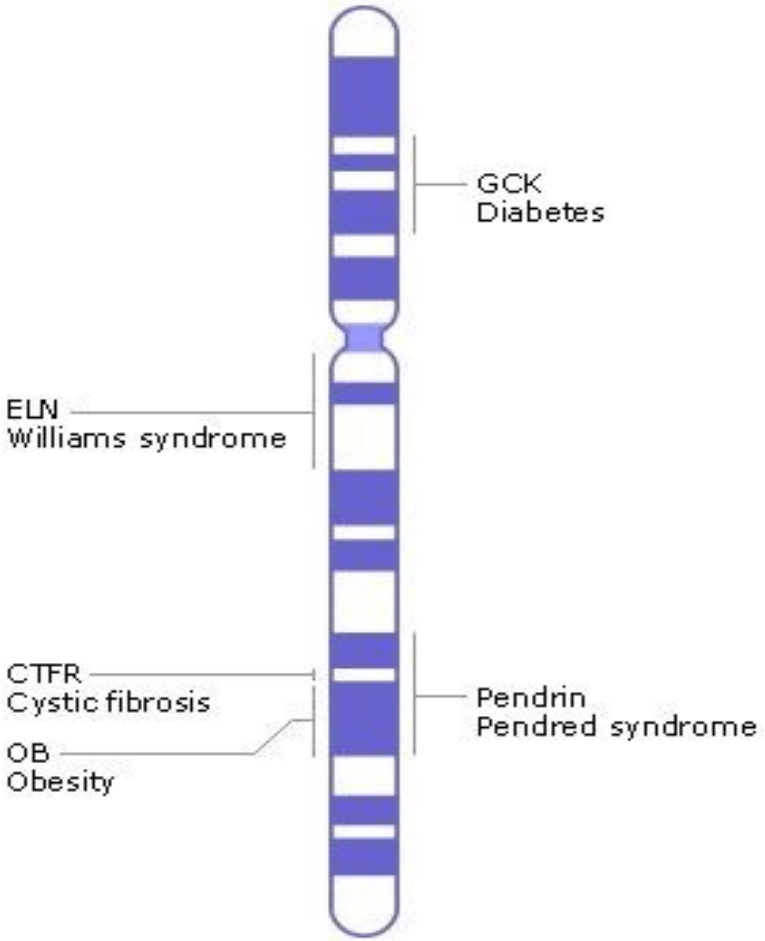


Chromosome 7

http://www.ncbi.nlm.nih.gov/bookshelf/br.fcgi?book=gnd&part=A272#A279

Chromosome 7

- Contains approximately 1800 genes
- Contains over 150 million base pairs, of which over 95% have been determined
- See the diseases associated with chromosome 7 in the MapViewer.



UCSC Genome Browser

UCSC Genome Bioinformatics

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.

http://genome.ucsc.edu/

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.

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.

14 Dec. 2009 - New job posting: Biological Data Technician

The UCSC Genome Browser project is looking for a bioinformatician, biologist, or software engineer with a strong biology background to collect and import data into the UCSC Genome Browser database and website. This person will work closely with external research laboratories to capture their experimental results and methods and with internal software developers and database testing staff to make the data accessible to the worldwide scientific community.

UCSC Genome

UCSC Genome Browser

Home | Genomes | Blat | Tables | Gene Sorter | PCR | Session | FAQ | Help

Human (*Homo sapiens*) Genome Browser Gateway

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

clade: Mammal | genome: Human | assembly: GRCh37 | position or search term: chr21:33,031,597-33,041,570 | image width: 800 | submit

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

add custom tracks | configure tracks and display | clear position

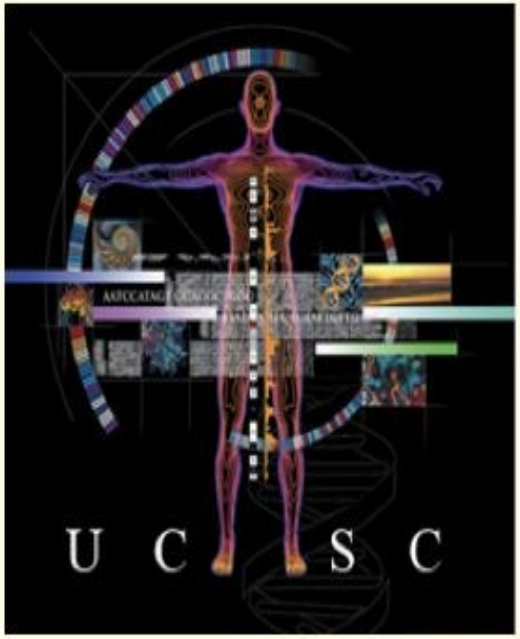
About the Human GRCh37 (hg19) assembly (sequences)

The February 2009 human reference sequence (GRCh37) was produced by the Genome Reference Consortium.

Sample position queries

A genome position can be specified by the accession number of a sequenced genomic clone, an mRNA or EST or STS marker, a chromosomal coordinate range, or keywords from the GenBank description of an mRNA. The following list shows examples of valid position queries for the human genome. See the User's Guide for more information.

Request:	Genome Browser Response:
chr7	Displays all of chromosome 7
chrUn_gl000212	Displays all of the unplaced contig gl000212
chr3:1-1000000	Displays first million bases of chr 3, counting from p-arm telomere

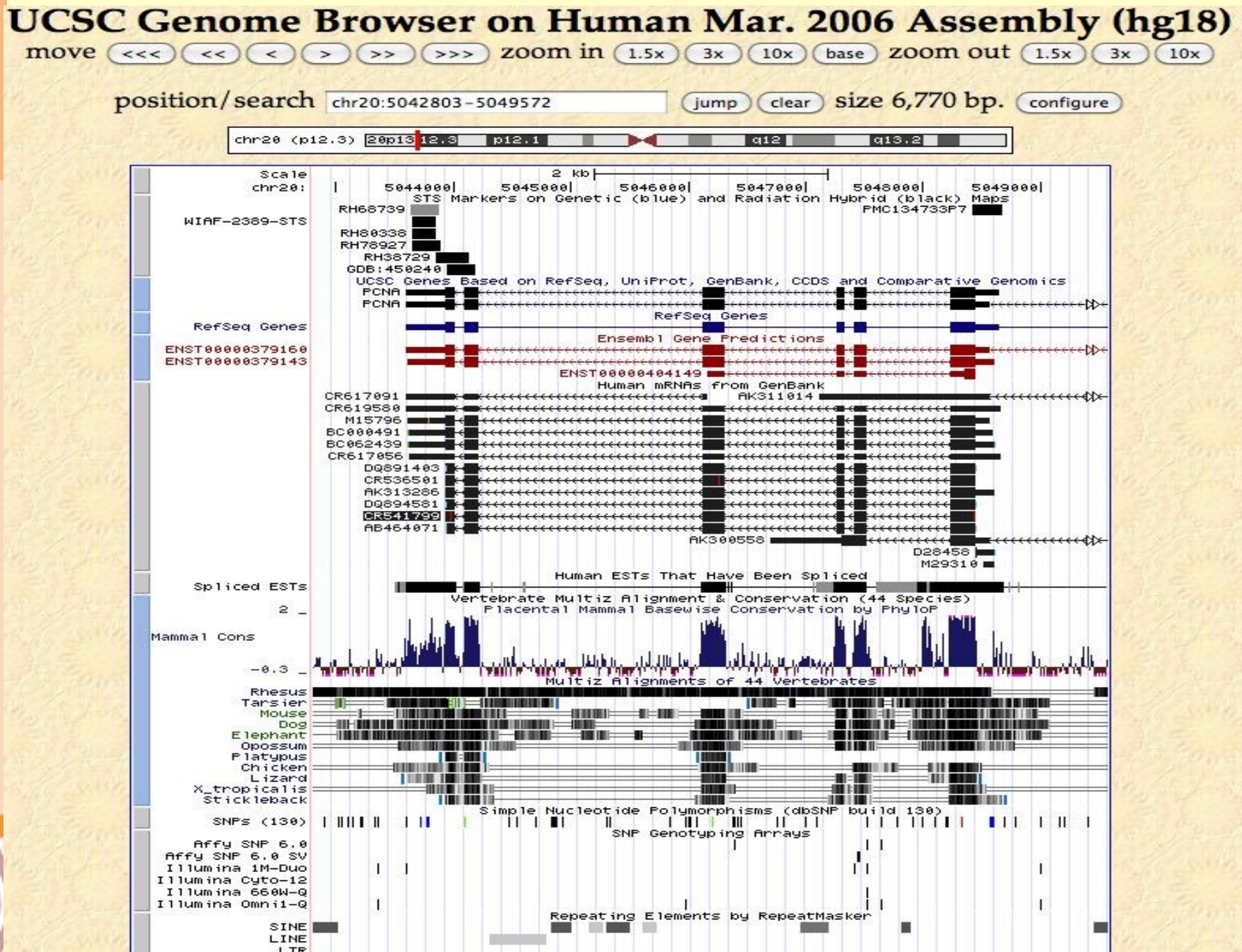


Homo sapiens
(Graphic courtesy of CBSE)

UCSC Genome

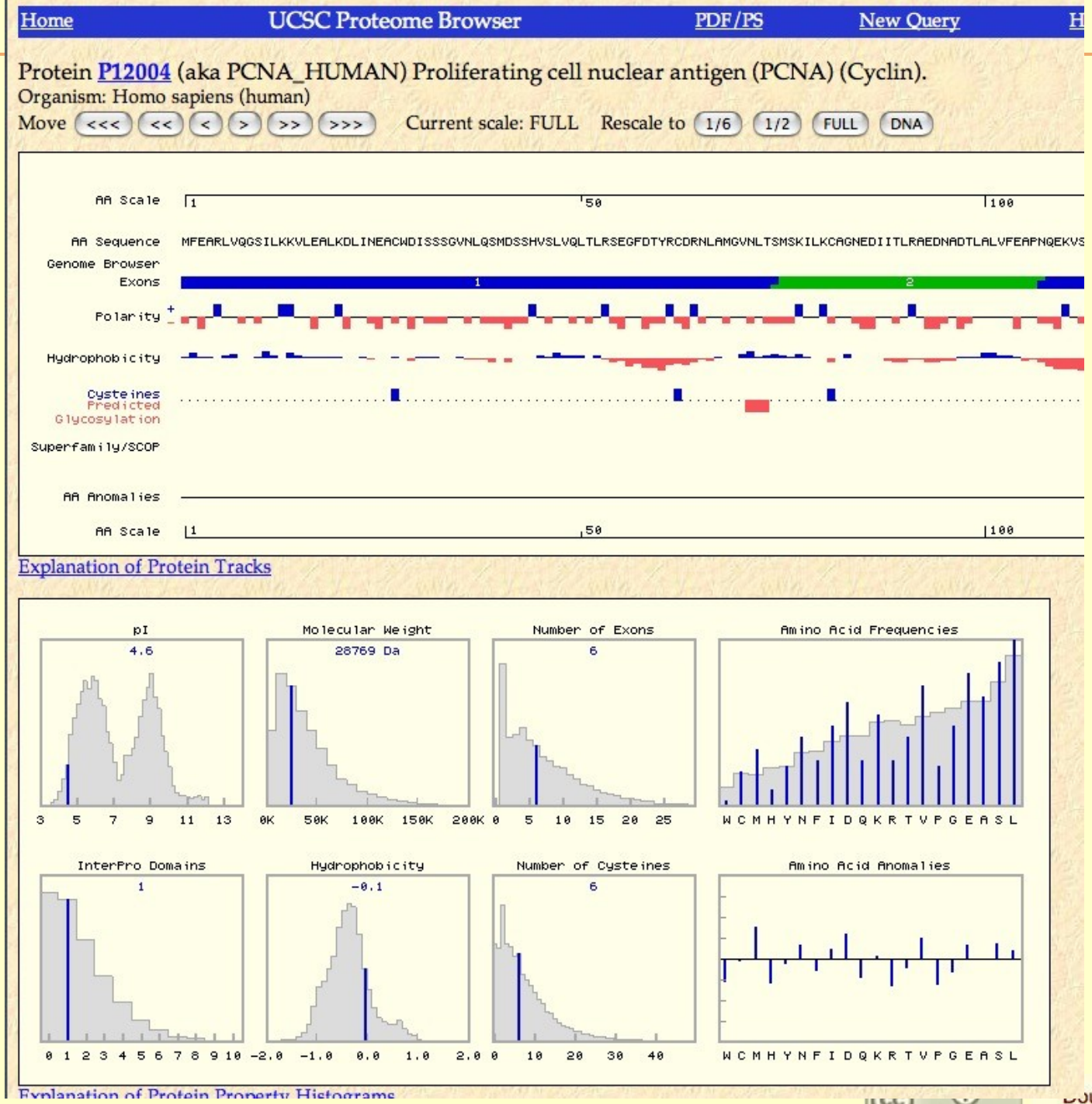
PCNA Entry at UCSC

http://genome.cse.ucsc.edu/cgi-bin/hgTracks?insideX=115&revCmplDisp=0&hgsid=149552355&hgt.out1=1.5x&position=chr20%3A5043932-5048



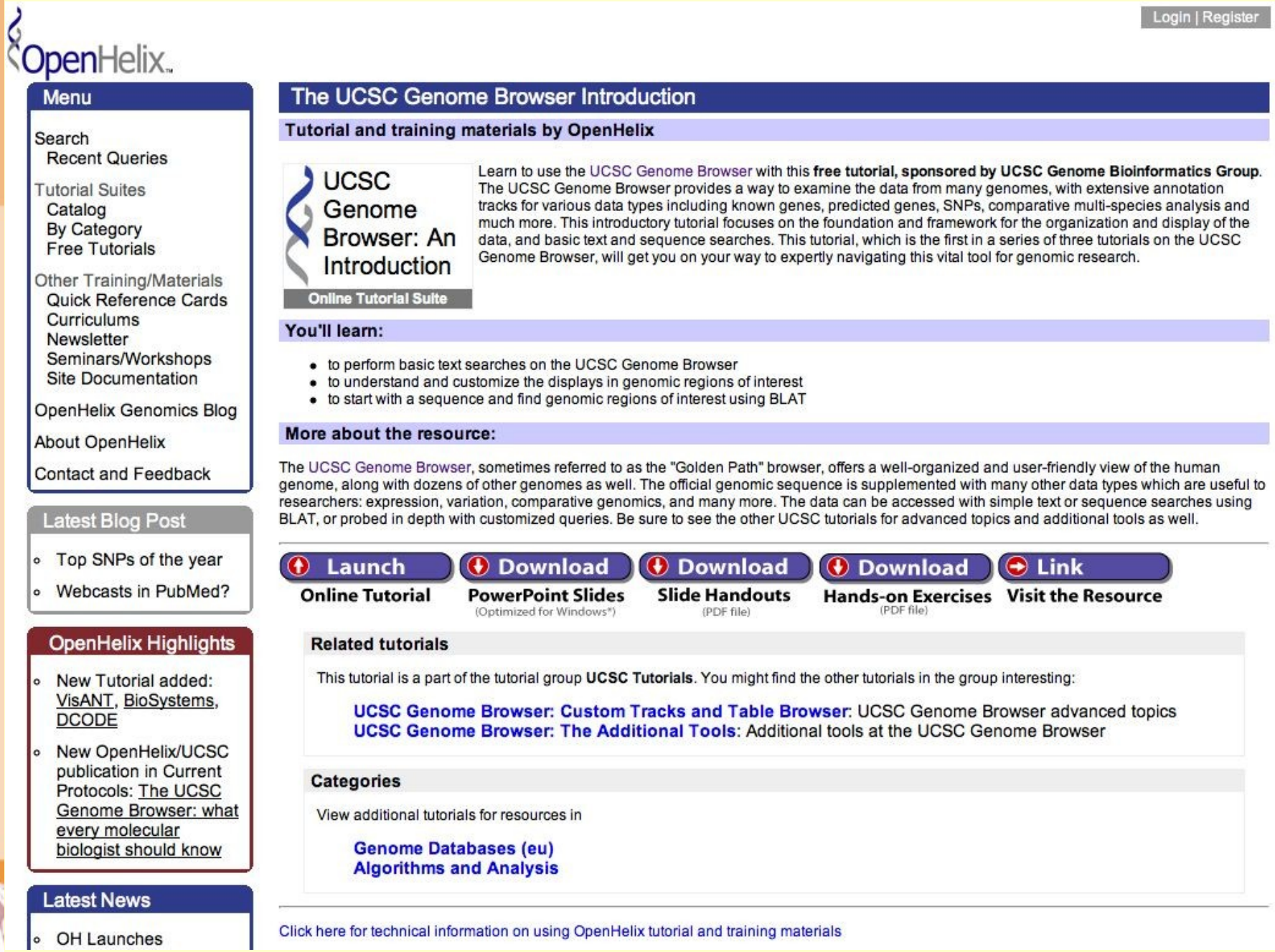
UCSC Proteome Browser

http://genome.ucsc.edu/cgi-bin/pbGlobal?proteinID=P12004



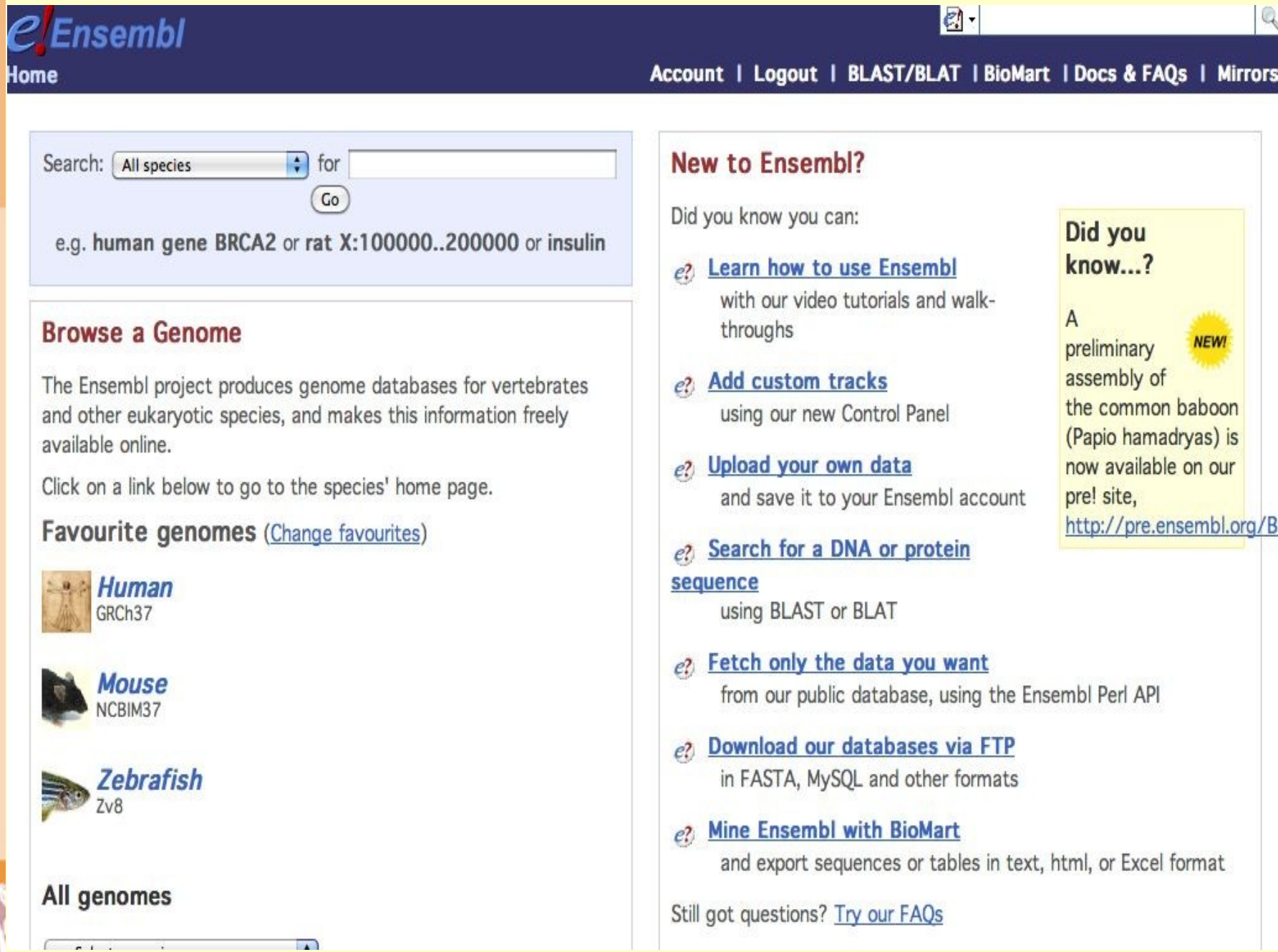
UCSC Tutorials

http://www.openhelix.com/downloads/ucsc/ucsc_home.shtml



Ensembl Genomes Home Page

http://www.ensembl.org/



Ensembl Human Genome

Ensembl

Home > Human [GRCh37]

Account | Logout | BLAST/BLAT | BioMart | Docs & FAQs | Mirrors

About this species

- Description
- Genome Statistics
  - Assembly and Genebuild
  - Top 40 InterPro hits
  - Top 500 InterPro hits
- What's New
- Sample entry points
  - Karyotype
  - Location (6:133017695-
  - Gene (BRCA2)
  - Transcript (FOXP2-203)
  - Variation (rs1333049)

- Configure this page
- Manage your data
- Export data
- Bookmark this page

Search Ensembl Human

Search for: PCNA Go

e.g. **gene BRCA2** or **6:133017695-133161157** or **muscular dystrophy**

Description

Assembly and Genebuild »

Human (*Homo sapiens*)

Assembly

This site provides a data set based on the February 2009 *Homo sapiens* high coverage assembly from the Genome Reference Consortium. The data set consists of gene models built from the genewise alignments of the human proteome as well as from alignments of human cDNAs using the cDNA2genome model of exonerate

This release of the assembly has the following properties:

- 27478 contigs.
- contig length total 3.2 Gb.
- chromosome length total 3.1 Gb.

It also includes nine haplotypic regions, mainly in the MHC region of chromosome 6.

To convert your old data from Human assembly NCBI36 to GRCh37, click on 'Manage your data' on any human page and select 'Assembly converter' from the left-hand menu.

Ensembl DNA View of Human PCNA

http://www.ensembl.org/Homo_sapiens/Gene/Summary?g=ENSG00000132646

Ensembl

Home > Human [GRCh37]

Account | Logout | BLAST/BLAT | BioMart | Docs & FAQs | Mirrors

Location: 20:5,095,599-5,107,272 | Gene: PCNA

Gene-based displays

- Gene summary
- Splice variants (2)
- Supporting evidence
- Sequence
- External references (3)
- Regulation
- Comparative Genomics
  - Genomic alignments (51)
  - Gene Tree (image)
    - Gene Tree (text)
    - Gene Tree (alignment)
  - Orthologues (47)
  - Paralogues (0)
  - Protein families (1)
- Genetic Variation
  - Variation Table
  - Variation Image
- External Data
  - Personal annotation
- ID History
  - Gene history

- Configure this page
- Manage your data
- Export data
- Bookmark this page

Gene: PCNA (ENSG00000132646)

Proliferating cell nuclear antigen (PCNA)(Cyclin) Source: UniProtKB/Swiss-Prot P12004

Location Chromosome 20: 5,095,599-5,107,272 reverse strand.

Transcripts There are 2 transcripts in this gene: hide transcripts

Name	Transcript ID	Protein ID	Description
PCNA-001	ENST00000379143	ENSP00000368438	protein_coding
PCNA-002	ENST00000379160	ENSP00000368458	protein_coding

Transcript and Gene level displays

In Ensembl a gene is made up of one or more transcripts. We provide displays at two levels:

- Transcript views which provide information specific to an individual transcript such as the cDNA and CDS sequences and protein domain annotation.
- Gene views which provide displays for data associated at the gene level such as orthologues and paralogues, regulatory regions and splice variants.

This view is a gene level view. To access the transcript level displays select a Transcript ID in the table above and then navigate to the information you want using the menu at the left hand side of the page. To return to viewing gene level information click on the Gene tab in the menu bar at the top of the page.

Gene summary help

Splice variants »

Name PCNA (HGNC (curated))

CCDS This gene is a member of the Human CCDS set: CCDS13087

Gene type Known protein coding

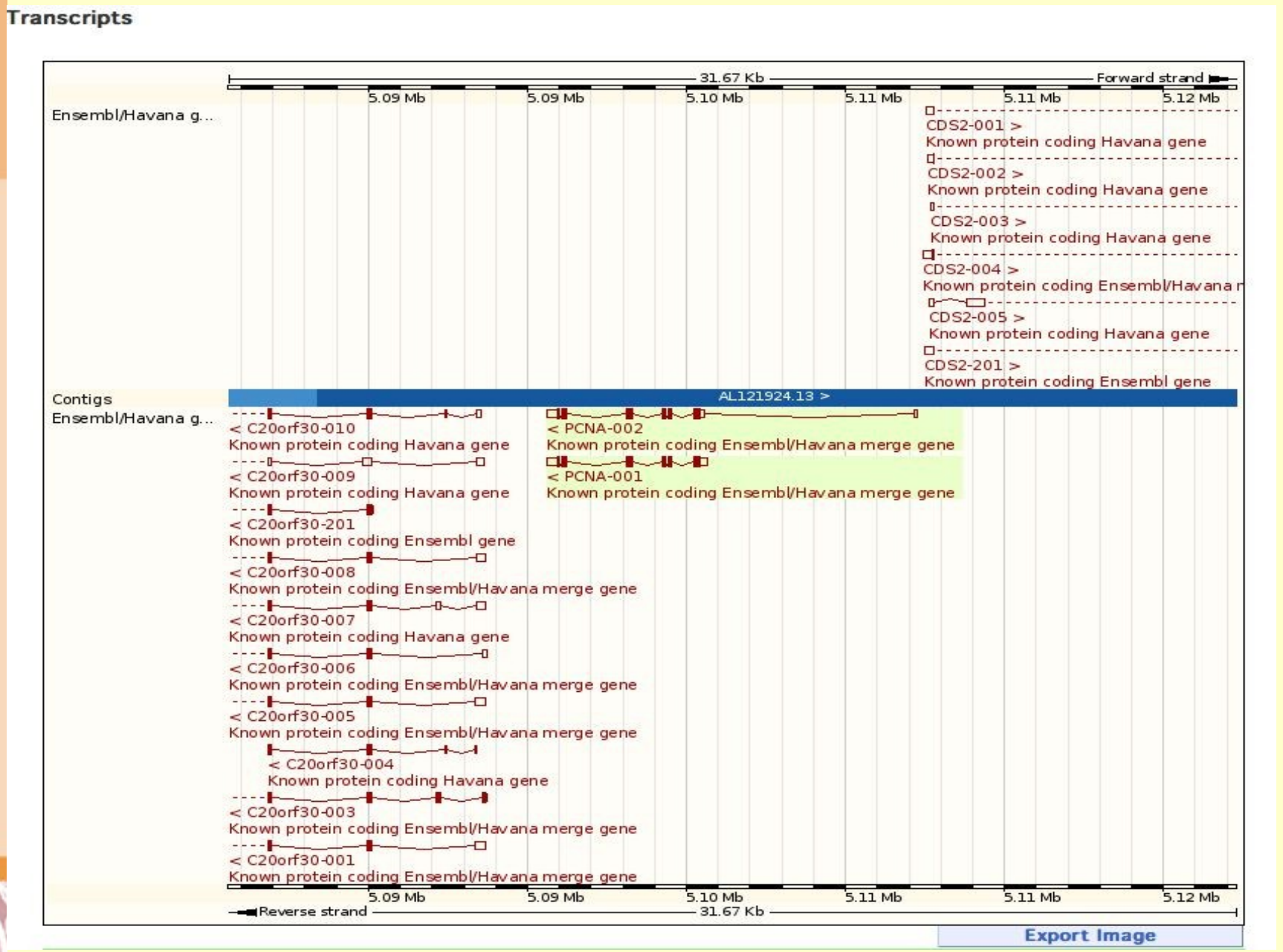
Prediction Method Gene containing both Ensembl genebuild transcripts and Havana manual curation, see article.

Alternative genes **This Known protein coding entry corresponds to the following database identifiers:**

Havana Gene: OTTHUMG00000031798 [view all locations]

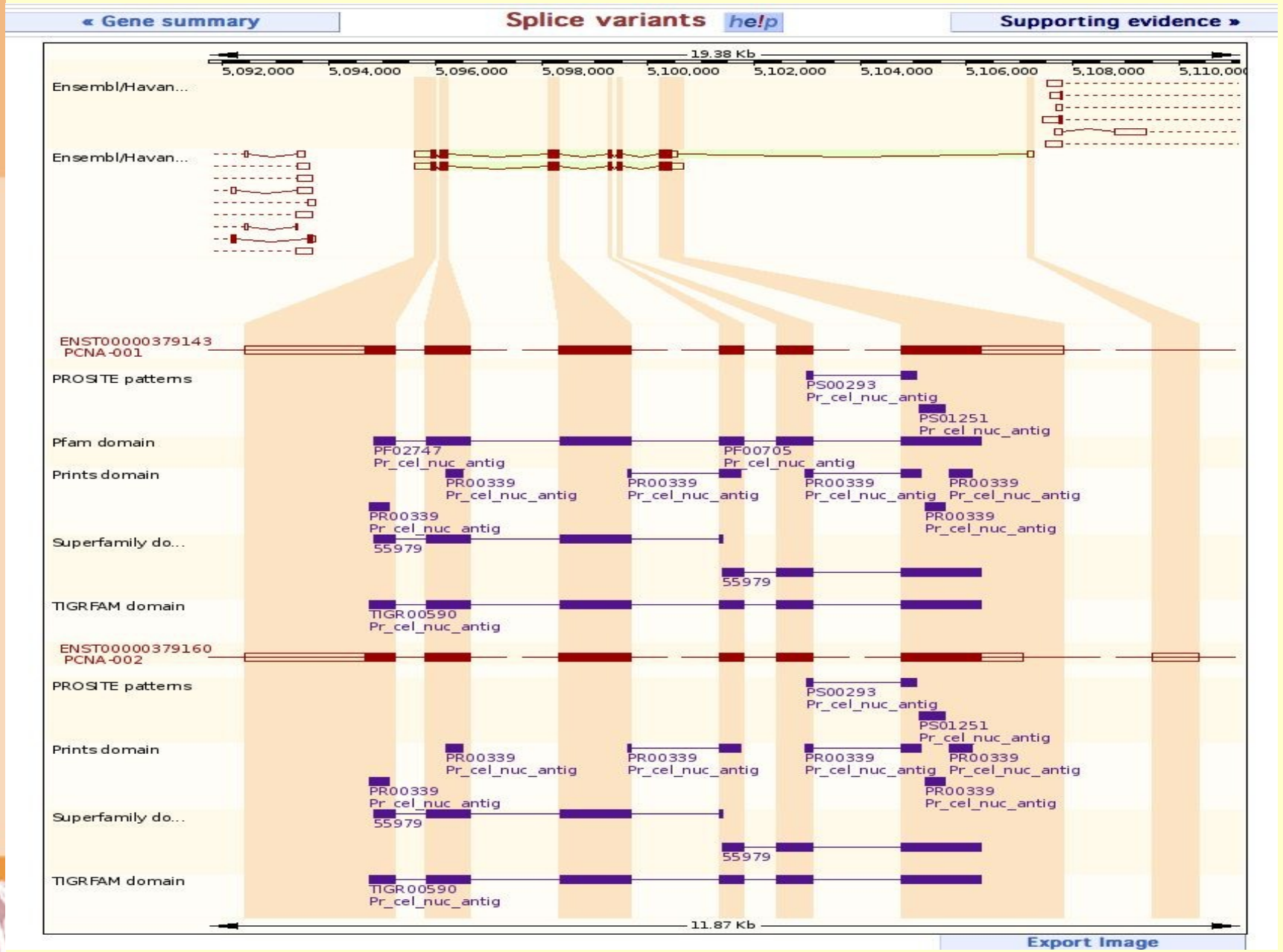
Ensembl PCNA Transcripts

http://www.ensembl.org/Homo_sapiens/Gene/Summary?g=ENSG00000132646



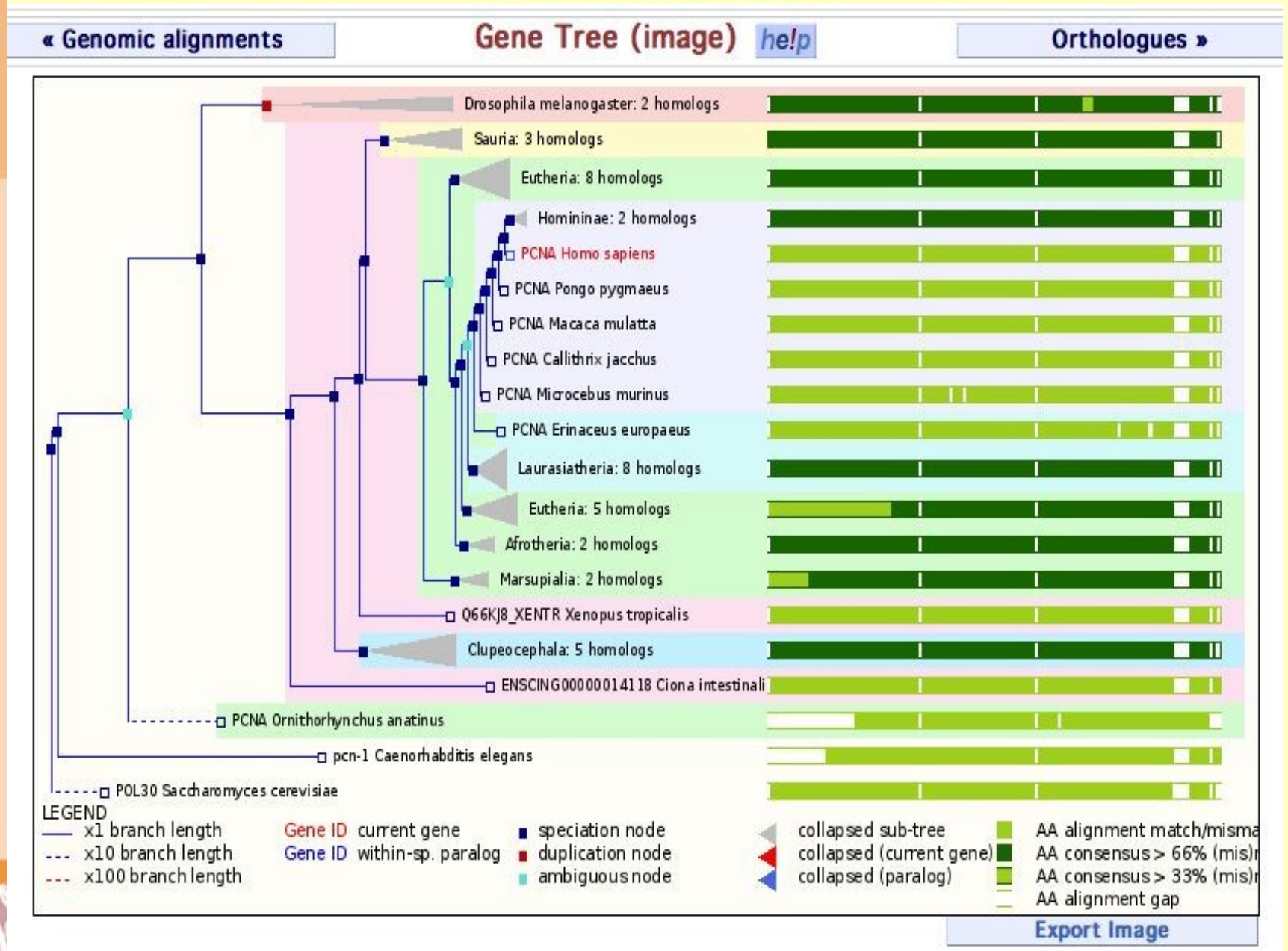
Ensembl PCNA Splice Variants

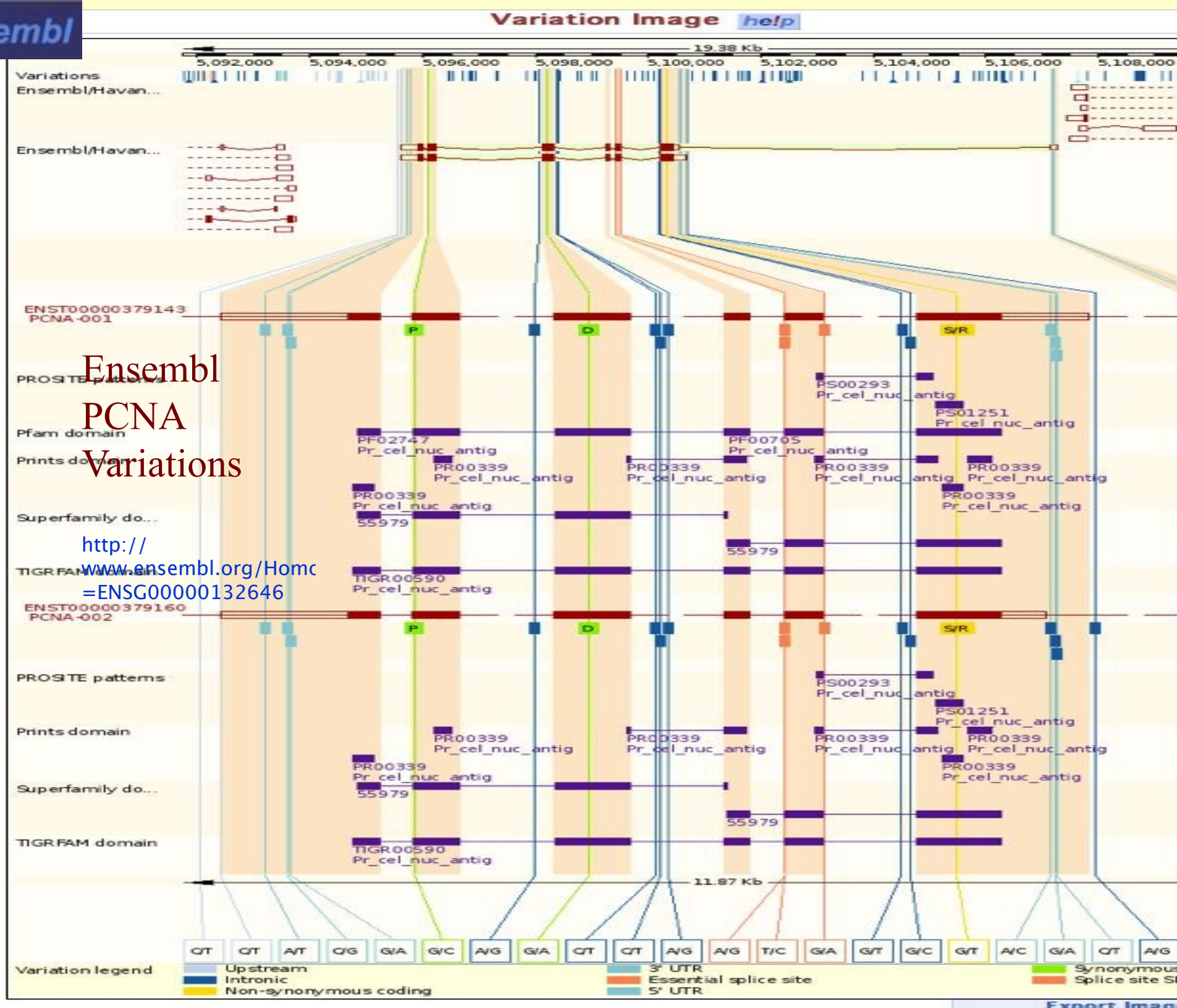
http://www.ensembl.org/Homo_sapiens/Gene/Summary?g=ENSG00000132646



Ensembl PCNA Gene Tree (Pecan)

http://www.ensembl.org/Homo_sapiens/Gene/Summary?g=ENSG00000132646

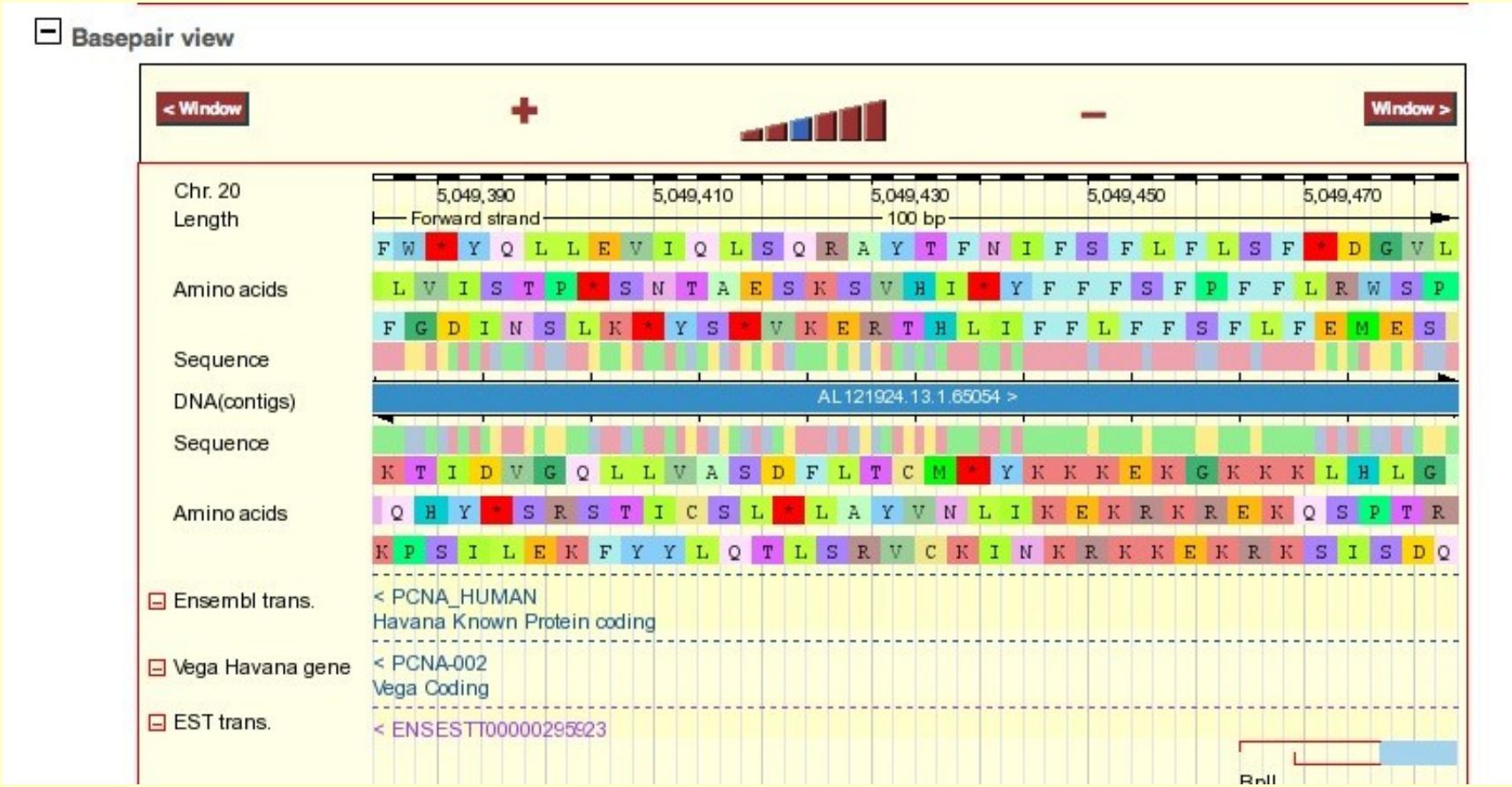




Ensembl
PCNA
Variations

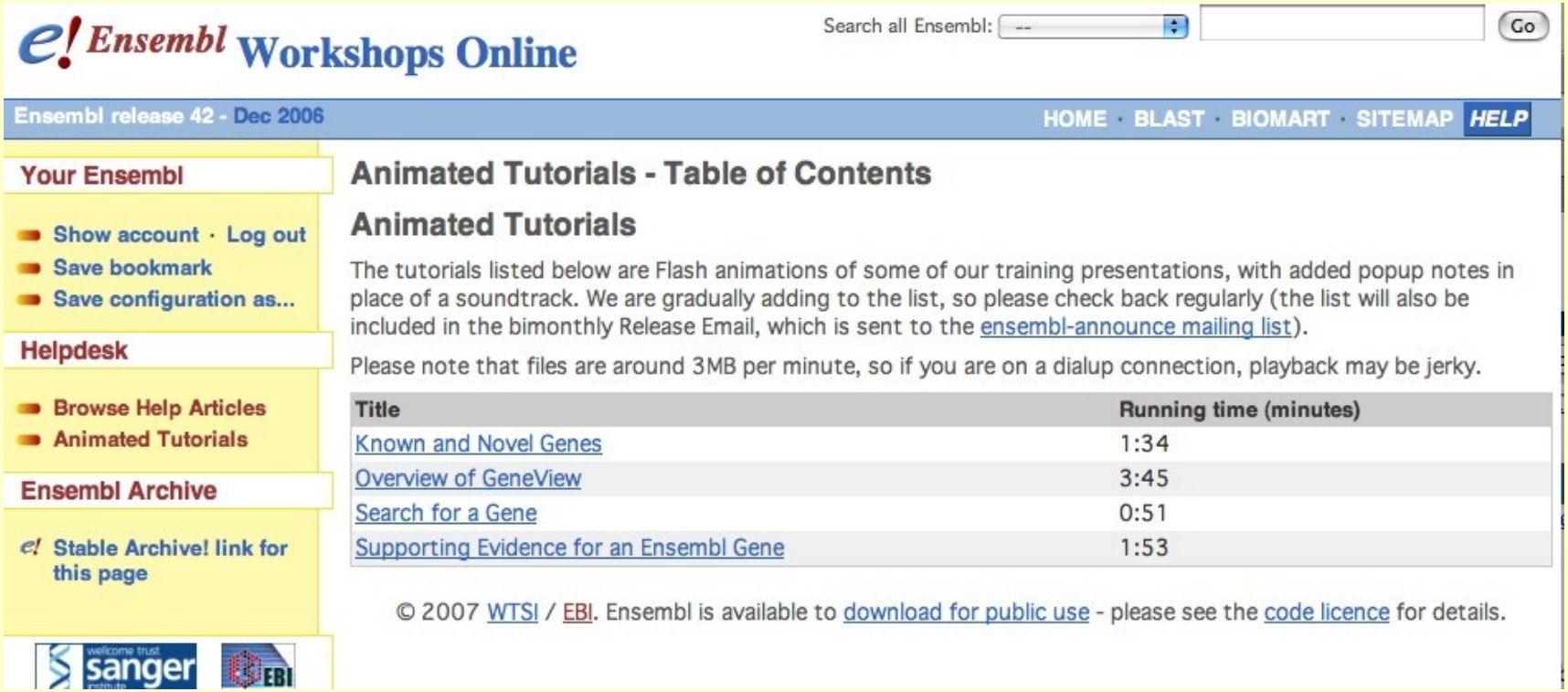
http://
www.ensembl.org/Homo
=ENSG00000132646

Ensembl Protein View of PCNA

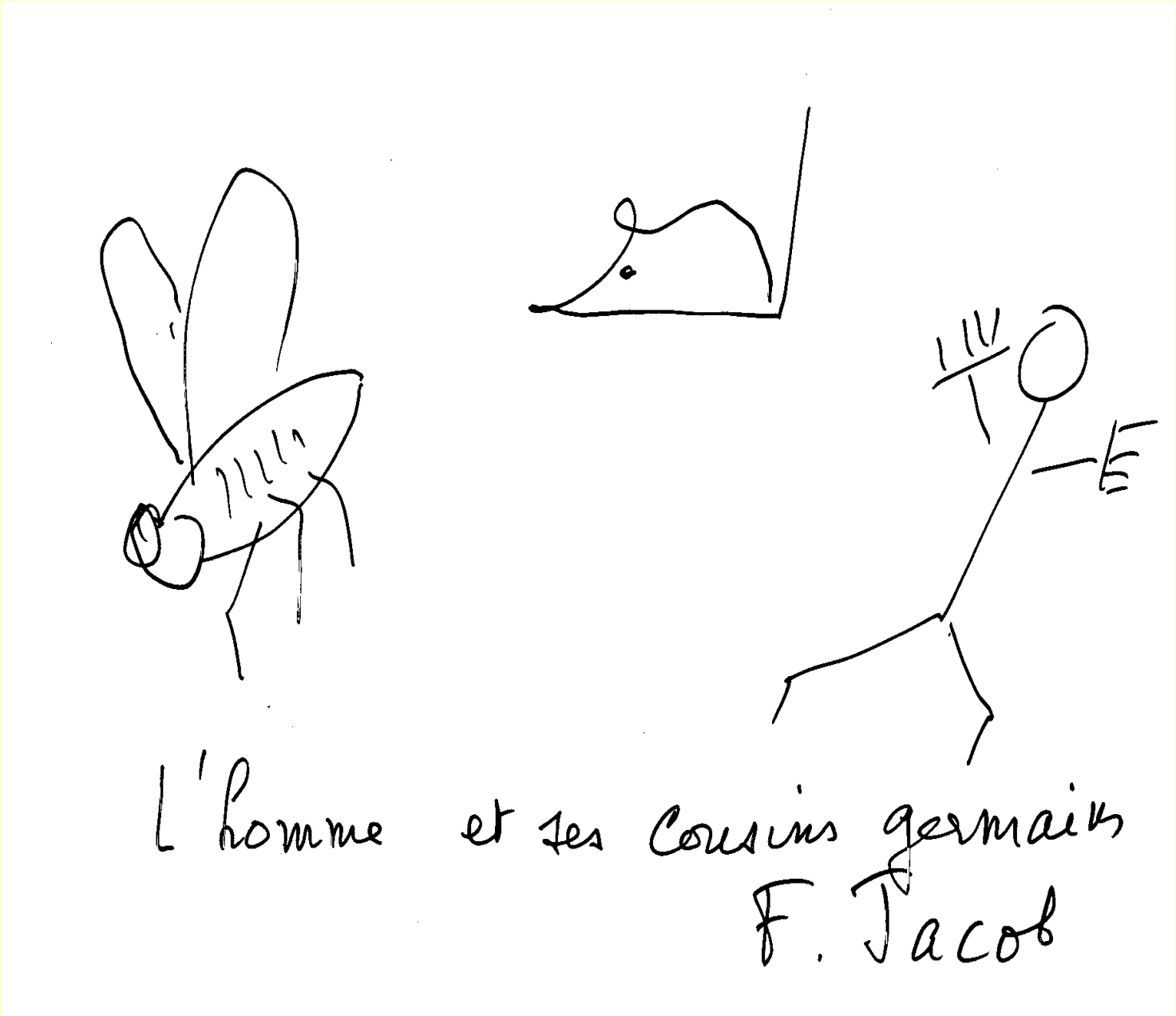


Ensembl Tutorials

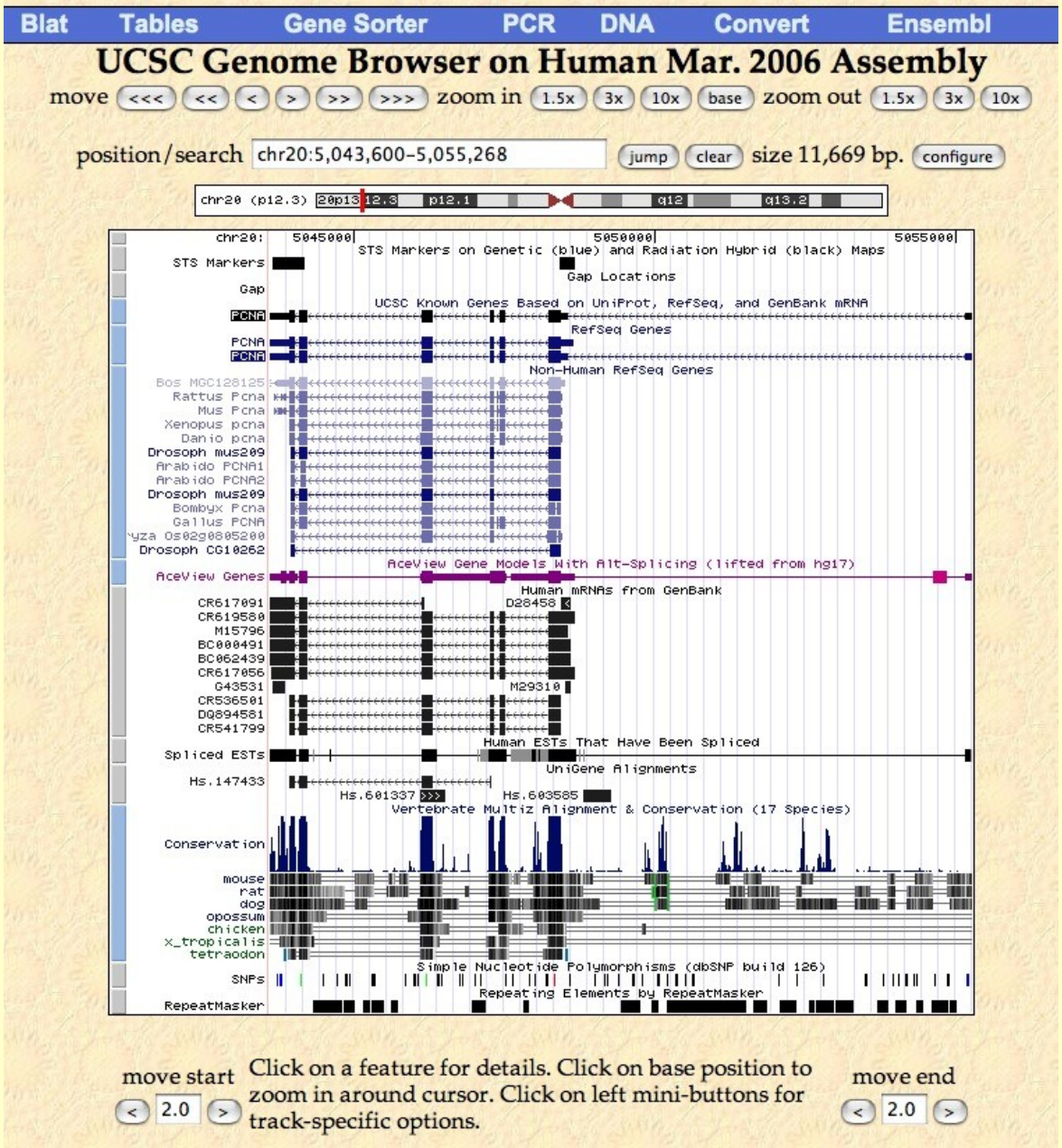
http://www.ensembl.org/common/Workshops_Online



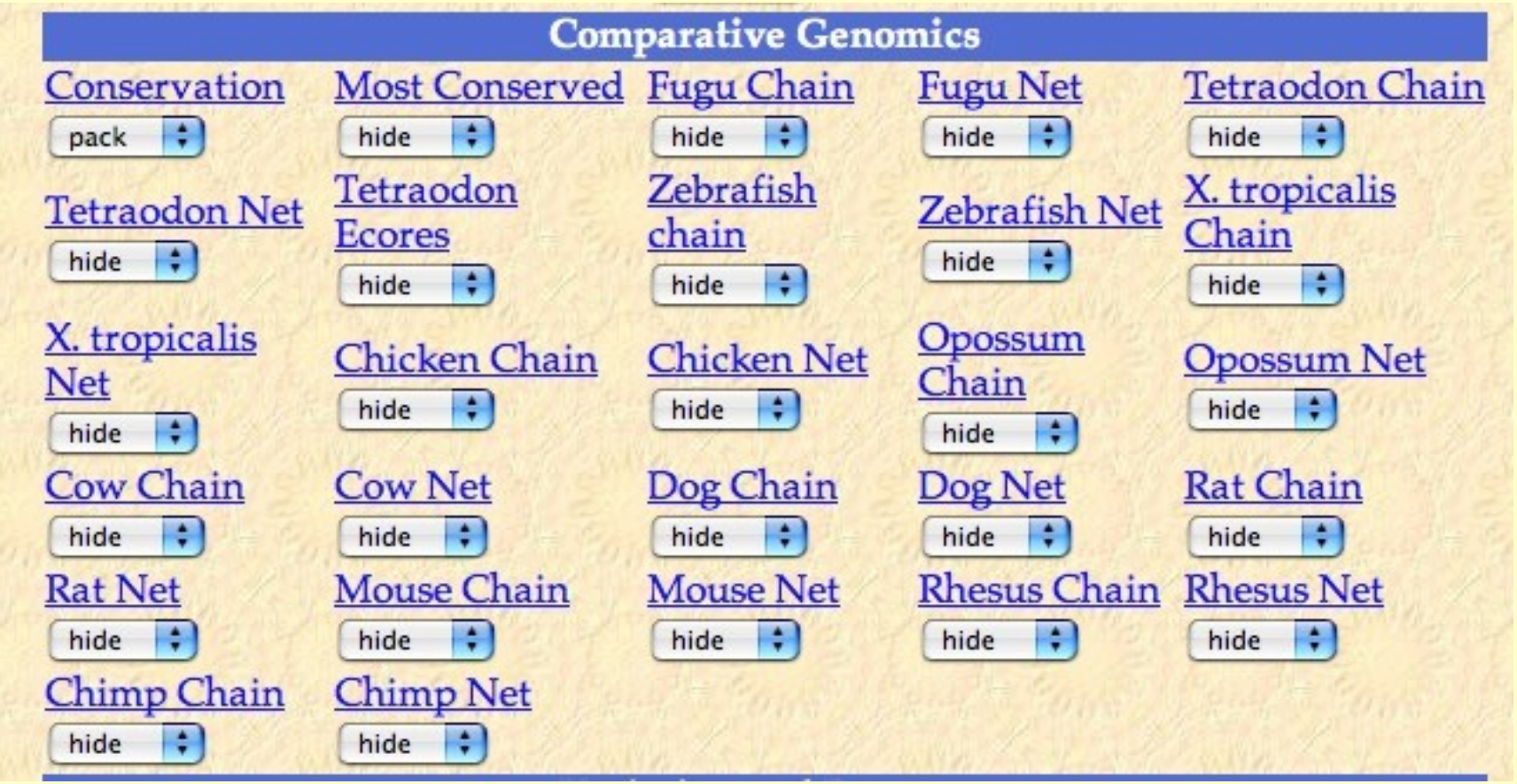
Comparative Genomics



UCSC PCNA Entry



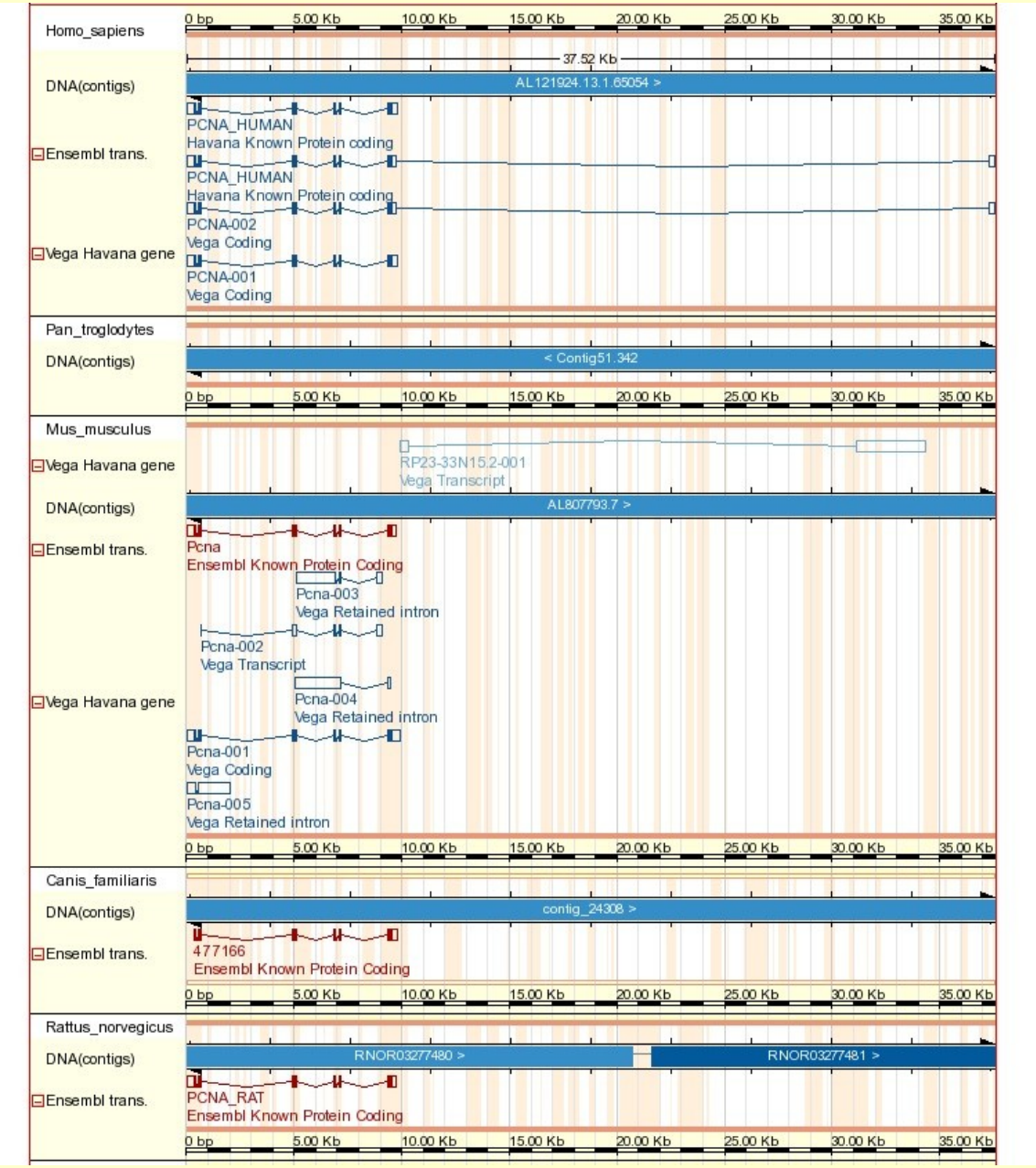
UCSC Comparative Genomics



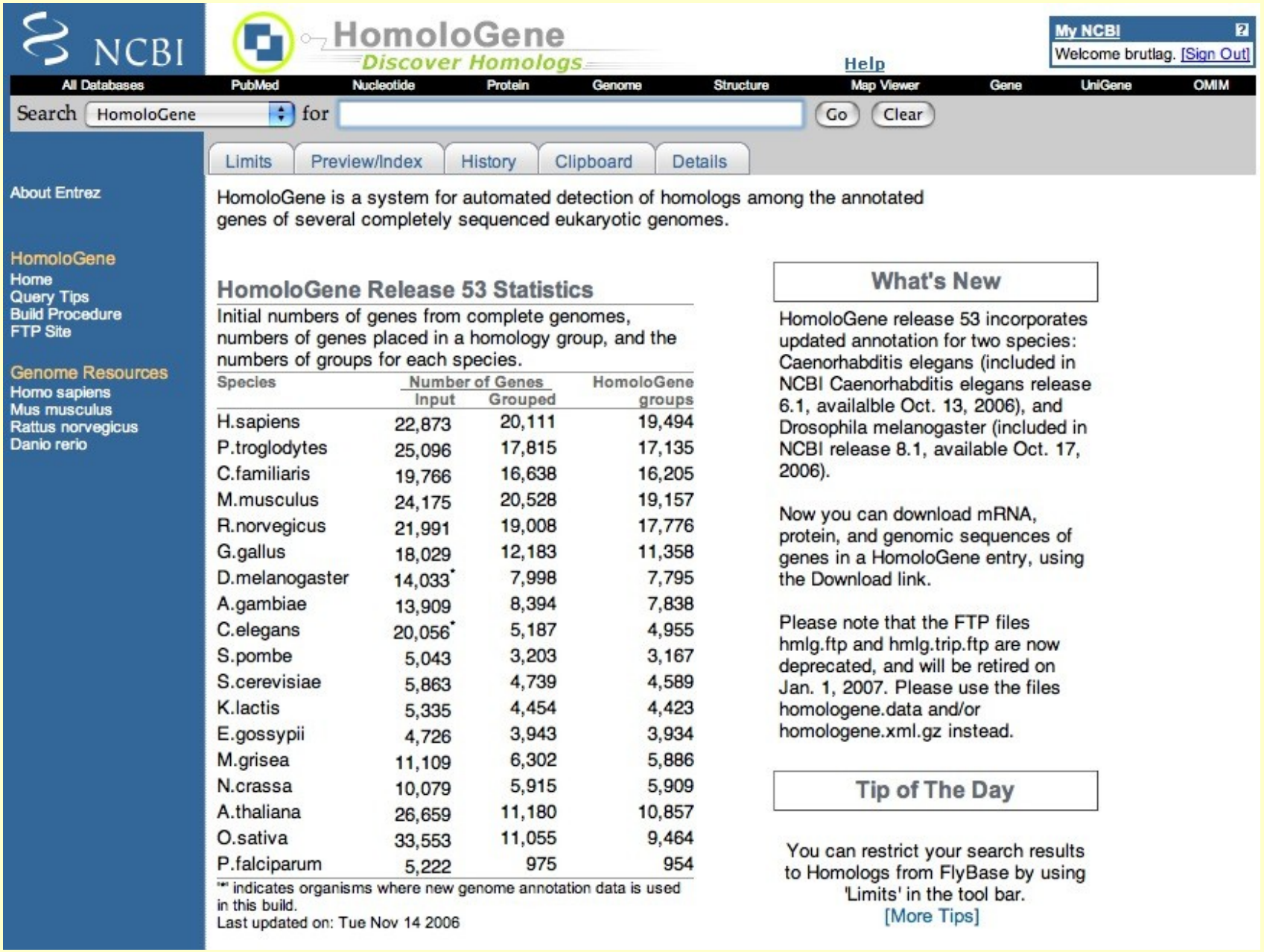
Ensembl Genome Alignments

Alignments	**This gene can be viewed in genomic alignment with other species** view genomic alignment with **7 eutherian mammals Pecan** view genomic alignment with **9 amniota vertebrates Pecan** view genomic alignment with **Rattus norvegicus** view genomic alignment with **Mus musculus** view genomic alignment with **Bos taurus** view genomic alignment with **Monodelphis domestica** view genomic alignment with **Macaca mulatta** view genomic alignment with **Loxodonta africana** view genomic alignment with **Echinops telfairi** view genomic alignment with **Oryctolagus cuniculus** view genomic alignment with **Dasypus novemcinctus** view genomic alignment with **Canis familiaris** view genomic alignment with **Pan troglodytes** view genomic alignment with **Gallus gallus** view genomic alignment with **Ornithorhynchus anatinus**
Orthologue	The following gene(s) have been identified as putative orthologues:

Ensembl Genome Alignments (PCNA)



NCBI Homologene Database



Homologene PCNA Entries