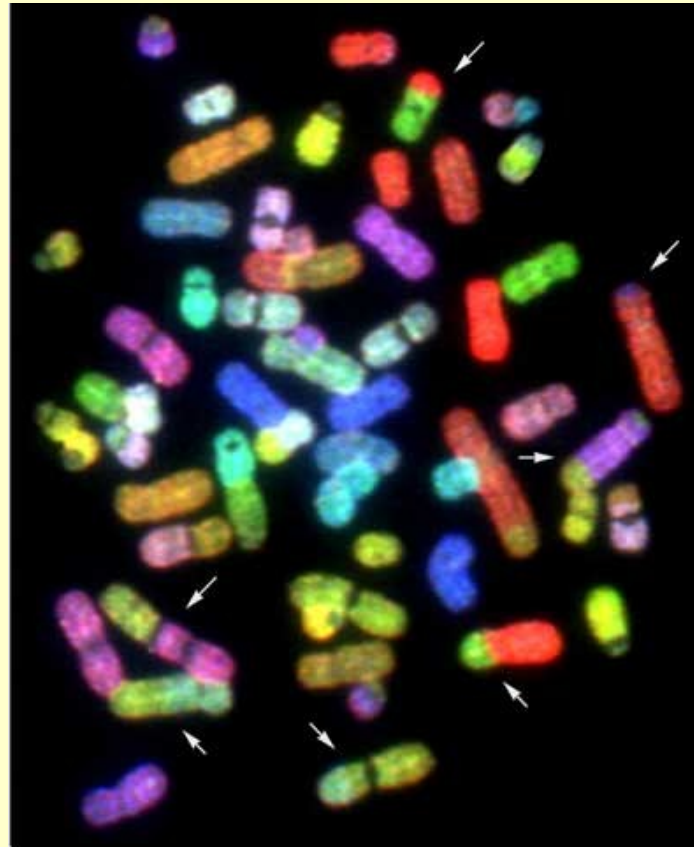


Computational Molecular Biology

Biochem 218 – BioMedical Informatics 231

<http://biochem218.stanford.edu/>

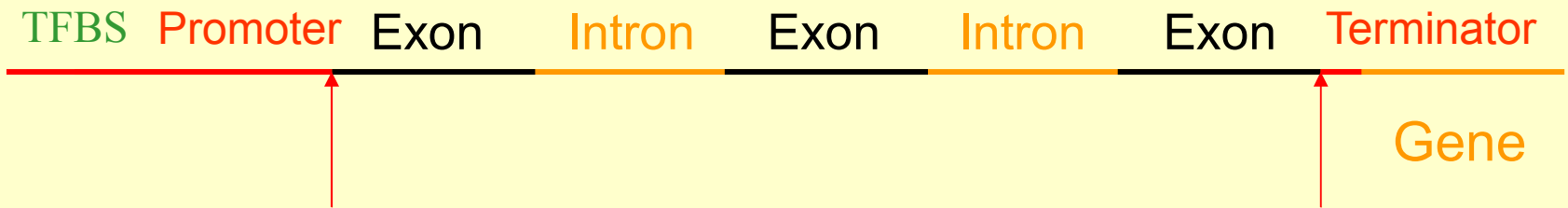
Genome Databases



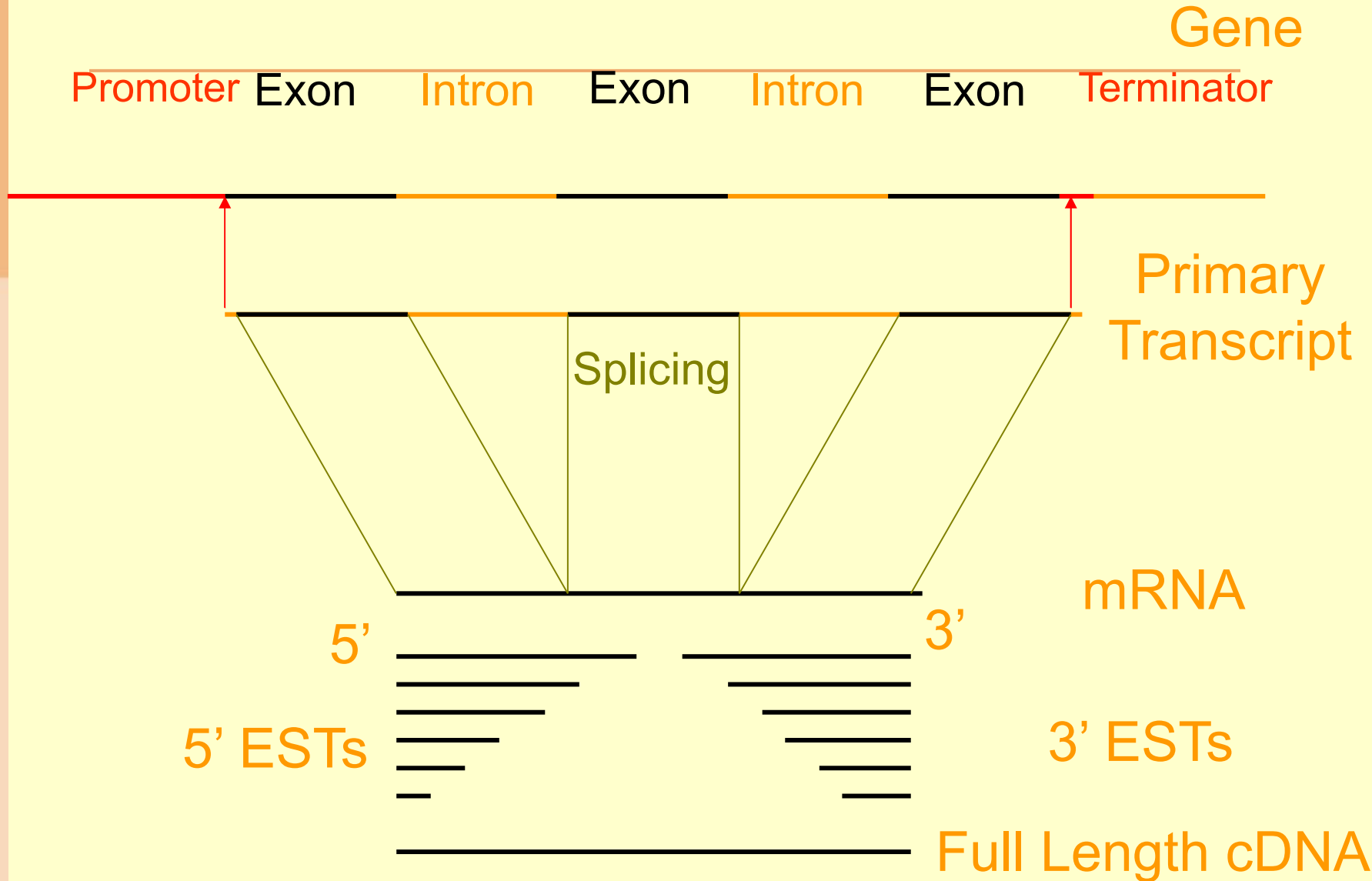
Doug Brutlag
Professor Emeritus
Biochemistry & Medicine (by courtesy)



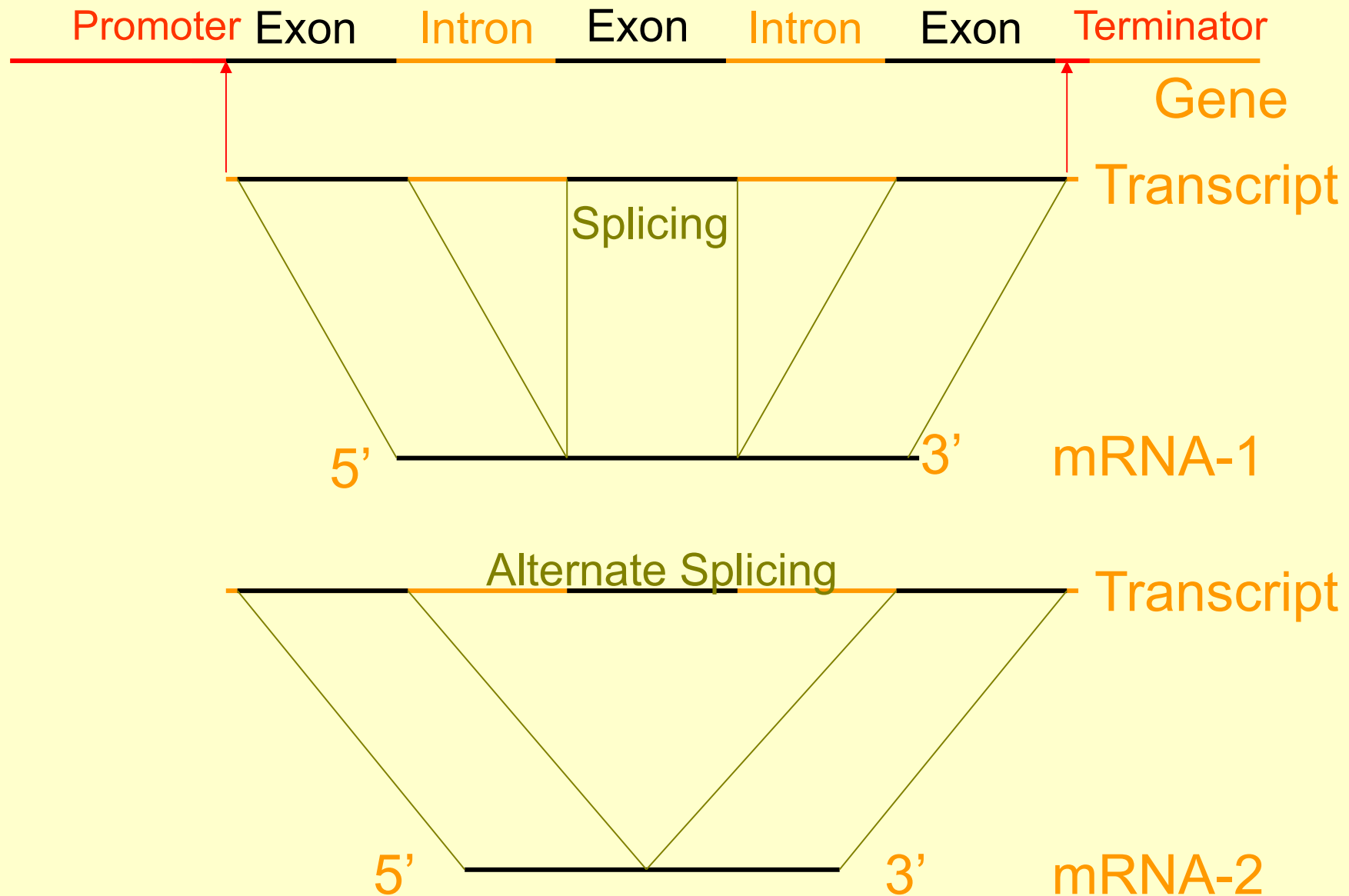
Components of a Typical Human Gene



ESTs, Full Length cDNA



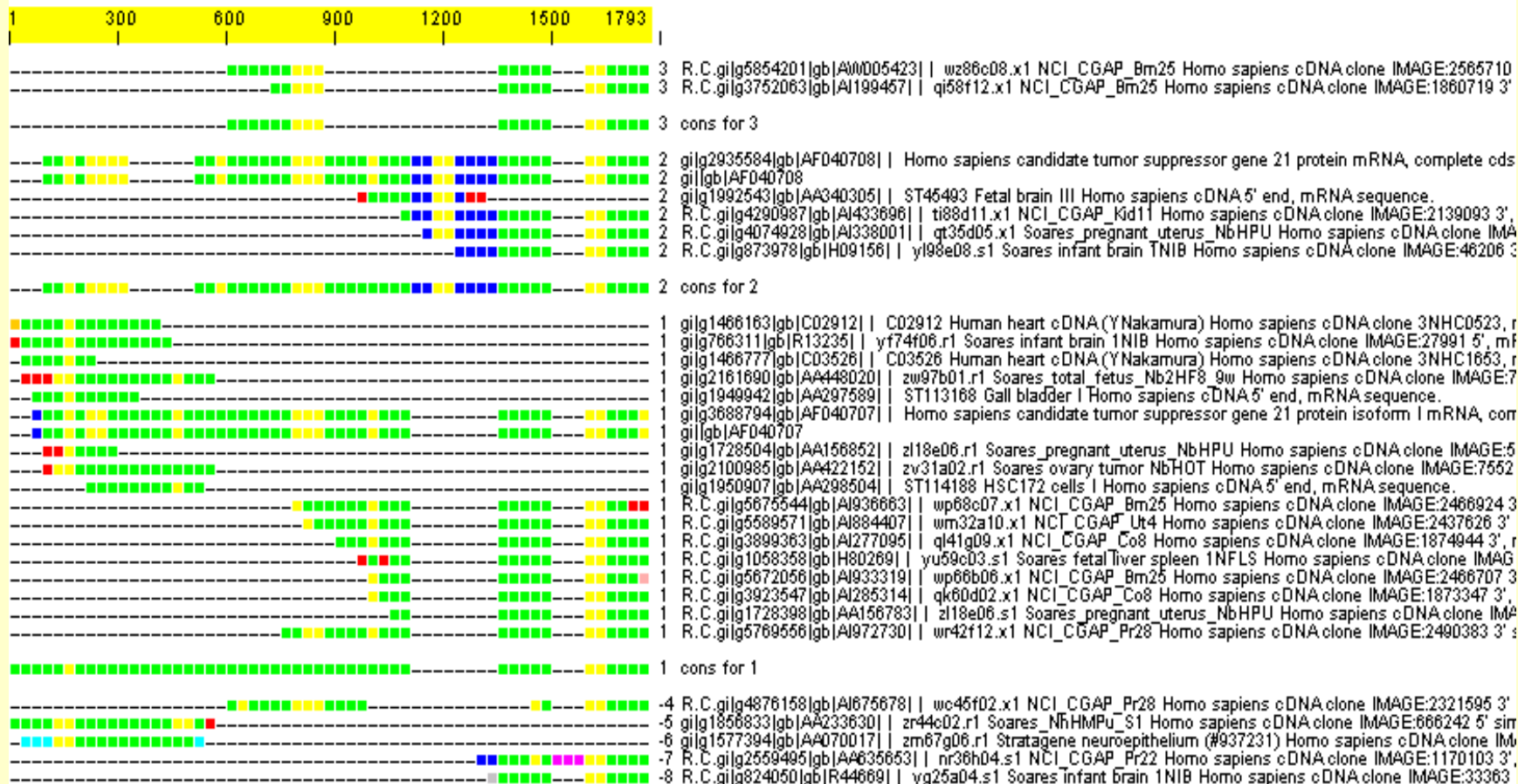
Alternative Splicing Generates Distinct Proteins in Different Tissues



Alternative Splicing Detected in EST Libraries

One position equals 30 bases.

- if more than 3 bases disagree with consensus sequences.
- if more than 15 positions are unknown.
- _ if more than 15 positions are gap characters.



Annotating Genome Databases

NHGRI contigs



A Mapping

uniSTS

dbSNP

B Gene Prediction

GrailEXP

GenScan

FGENESH

FGENESH+

GeneMark

C Expression Data

Human ESTs

UniGene Human

RefSeq Human

Ensembl cDNA

Mouse ESTs

Entrez Gene Mouse

RefSeq Mouse

D Protein Data

nrPRO

UniProt

pFAM Motifs

E Additional Data

Promoters

F Summary

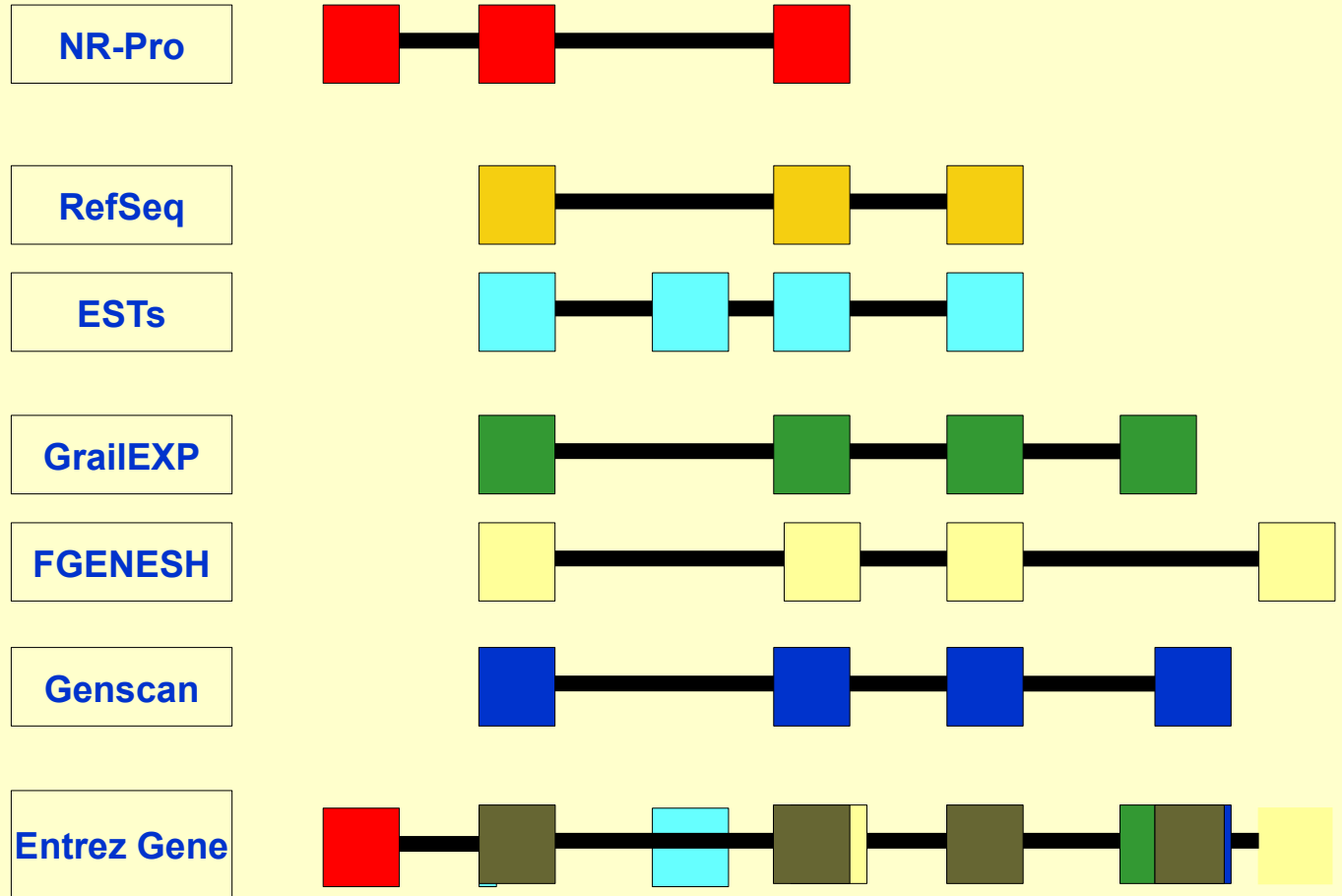
Entrez Gene

GeneCards



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>

[All Databases](#)
[PubMed](#)
[Nucleotide](#)
[Protein](#)
[Genome](#)
[Structure](#)
[OMIM](#)
[PMC](#)
[Journals](#)
[Books](#)

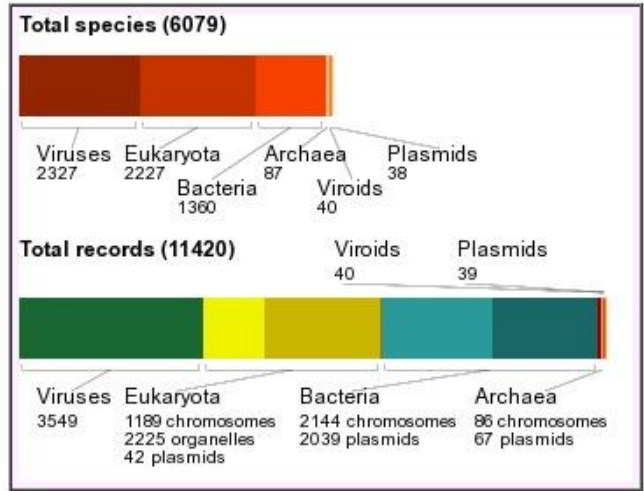
Search for

- About Entrez
- Entrez Genome Help
- Submitting 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

[Limits](#)
[Preview/Index](#)
[History](#)
[Clipboard](#)
[Details](#)

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.


- **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



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.

Human Genome Resources

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



NCBI • Genomic Biology • Homo sapiens


Search for

Browse your Genome
Click on the Chromosome to show

Genes

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

Find A Gene
Search for
from

The NCBI Handbook
 An online guide to the use of NCBI resources. Titles of selected chapters that refer to human genome resources are shown below.

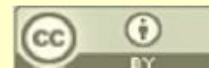
Human Genome Resources

A challenge facing researchers today is that of piecing together and analyzing the plethora of data currently being generated through the Human Genome Project and scores of smaller projects. NCBI's Web site serves as an integrated, one-stop, genomic information infrastructure for biomedical researchers from around the world so that they may use these data in their research efforts. [More...](#)

Genes and Human Health

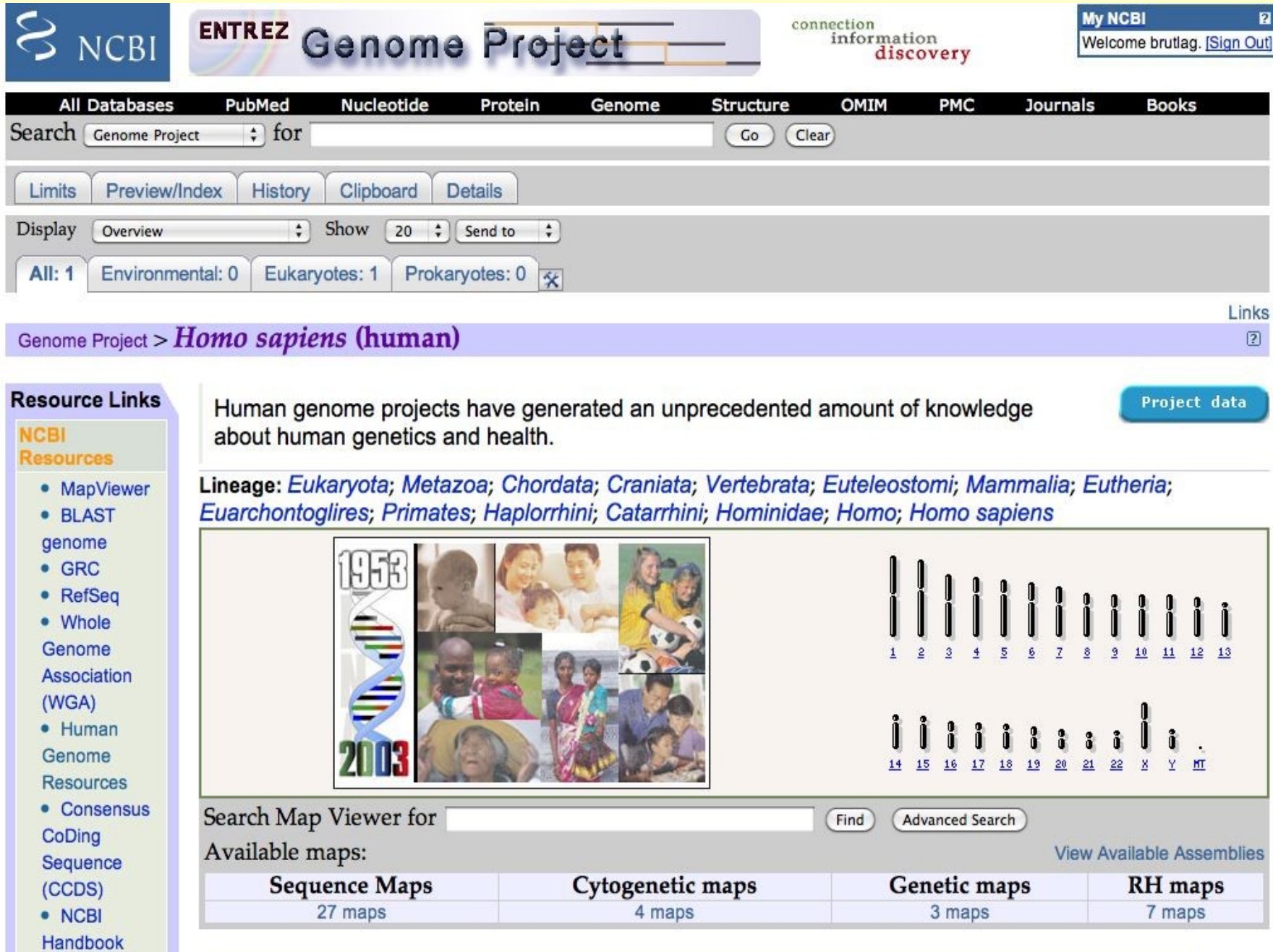
- ▶ **Gene Database**
A new database of genes and associated information is now available for searching in Entrez.
- ▶ **OMIM**
A guide to human genes and inherited disorders maintained by Johns Hopkins University and collaborators.
- ▶ **RefSeq**
Reference sequences of chromosomes, genomic contigs, mRNAs, and proteins for human and major model organisms.
- ▶ **dbSNP**
A database of single nucleotide polymorphisms (SNPs) and other nucleotide variations.

Reagents



Entrez Human Genome Projects

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



The image shows a screenshot of the NCBI Entrez Human Genome Project website. The page features a navigation bar with tabs for 'All Databases', 'PubMed', 'Nucleotide', 'Protein', 'Genome', 'Structure', 'OMIM', 'PMC', 'Journals', and 'Books'. A search bar is present with a dropdown menu set to 'Genome Project' and a search button. Below the search bar are buttons for 'Limits', 'Preview/Index', 'History', 'Clipboard', and 'Details'. A 'Display' section shows 'Overview' selected, 'Show' set to 20, and a 'Send to' dropdown. A summary bar indicates 'All: 1', 'Environmental: 0', 'Eukaryotes: 1', and 'Prokaryotes: 0'. The main content area is titled 'Genome Project > *Homo sapiens* (human)'. A text block states: 'Human genome projects have generated an unprecedented amount of knowledge about human genetics and health.' Below this is a 'Lineage' section listing taxonomic groups: *Eukaryota*; *Metazoa*; *Chordata*; *Craniata*; *Vertebrata*; *Euteleostomi*; *Mammalia*; *Eutheria*; *Euarchontoglires*; *Primates*; *Haplorrhini*; *Catarrhini*; *Hominidae*; *Homo*; *Homo sapiens*. A central image shows a DNA double helix with the years '1953' and '2003' and a collage of diverse people. To the right is a karyotype showing 22 numbered chromosomes, X, Y, and MT. A 'Resource Links' sidebar on the left lists various tools and databases. At the bottom, there is a 'Search Map Viewer for' section with a search bar and buttons for 'Find' and 'Advanced Search'. Below this is a table of available maps.

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

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*

1953 2003

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

Search Map Viewer for [] Find Advanced Search

Available maps:


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

View Available Assemblies




PCNA Gene in Entrez Gene

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



Entrez Gene



My NCBI
Welcome brutlag. [\[Sign Out\]](#)

All Databases
PubMed
Nucleotide
Protein
Genome
Structure
OMIM
PMC
Journals
Books

Search for Go Clear

Limits
Preview/Index
History
Clipboard
Details

Display Send to

1: **PCNA proliferating cell nuclear antigen** [*Homo sapiens*]
 GeneID: 5111 updated 10-Jan-2010

Summary

Official Symbol	PCNA	provided by HGNC
Official Full Name	proliferating cell nuclear antigen	provided by HGNC
Primary source	HGNC:8729	
See related	Ensembl:ENSG00000132646 ; HPRD:01456 ; MIM:176740	
Gene type	protein coding	
RefSeq status	REVIEWED	
Organism	Homo sapiens	
Lineage	<i>Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini; Catarrhini; Hominidae; Homo</i>	
Also known as	MGC8367; PCNA	
Summary	The protein encoded by this gene is found in the nucleus and is a cofactor of DNA polymerase delta. The encoded protein acts as a homotrimer and helps increase the processivity of leading strand synthesis during DNA replication. In response to DNA damage, this protein is ubiquitinated and is involved in the RAD6-dependent DNA repair pathway. Two transcript variants encoding the same protein have been found for this gene. Pseudogenes of this gene have been described on chromosome 4 and on the X chromosome. [provided by RefSeq]	

[Entrez Gene Home](#)

Table Of Contents

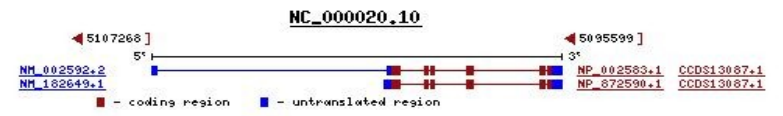
- Summary
- Genomic regions, transcripts, and products
- Genomic context
- Bibliography
- HIV-1 protein interactions
- Interactions
- General gene info
- General protein info
- Reference sequences
- Related sequences
- Additional links

Links Explain

- Order cDNA clone
- BioAssay, by Gene target
- BioSystems
- CCDS
- Conserved Domains
- EST
- Full text in PMC
- GEO Profiles
- Genome
- HomoloGene
- Map Viewer
- Nucleotide
- OMIM
- Probe
- Protein
- PubChem Compound
- PubChem Substance
- PubMed
- PubMed (GeneRIF)
- PubMed (OMIM)
- SNP
- ✓ [SNP: GeneView](#)
- SNP: Genotype
- Taxonomy
- UniSTS

Genomic regions, transcripts, and products ↑ ?

(minus strand) Go to [reference sequence details](#) [Try our new Sequence Viewer](#)



■ - coding region ■ - untranslated region

MapViewer of PCNA Gene

Search Find Find in This View Advanced Search

[BLAST The Human Genome](#)

***Homo sapiens (human)* Build 37.1 (Current)**

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](#)

[Download/View Sequence/Evidence](#)

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

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](#)

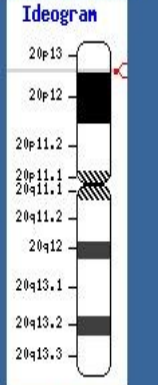
[Compress Map](#)

Region Shown:

 [Go](#)

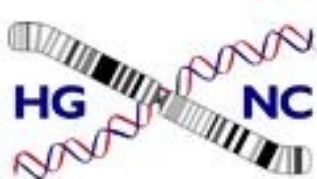
out
 200n
 in

You are here:



default
 master

Genes_cyto	Hs UniG	Genes_seq	Symbol	O	Links	E	Cyto	Description
	Hs.631504 Hs.635170 Hs.117299 Hs.594775		RASSF2	+	OMIM HGNC sv pr d lev mm hm sts CCDSSNP	best RefSeq	20pter-p12.1	Ras association (RalGDS/AF-6) domain fam
SLC23A2			SLC23A2	+	OMIM HGNC sv pr d lev mm hm sts CCDSSNP	best RefSeq	20p13	solute carrier family 23 (nucleobase transp
RPS21P7	Hs.516866		RPS21P7	+	HGNC sv d lev mm	best RefSeq	20p13	ribosomal protein S21 pseudogene 7
C20orf30	Hs.195368		C20orf30	+	HGNC sv pr d lev mm hm sts CCDSSNP	best RefSeq	20p13	chromosome 20 open reading frame 30
PCNA	Hs.472024 Hs.147433 Hs.569006 Hs.712995 Hs.626106 Hs.708226		PCNA	+	OMIM HGNC sv pr d lev mm hm sts CCDSSNP	best RefSeq	20pter-p12	proliferating cell nuclear antigen
CDS2	Hs.603585		CDS2	+	OMIM HGNC sv pr d lev mm hm sts CCDSSNP	best RefSeq	20p13	CDP-diacylglycerol synthase (phosphatida
LOC10028854	Hs.688395 Hs.472027		LOC10028854	+	sv d lev mm	mRNA	20	similar to lactation elevated 1 (predicted)



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](#)

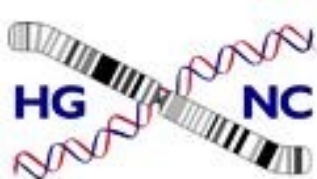
[Downloads](#)

[Home](#)

Giving unique and meaningful names to every human gene

[Quick Gene Search](#)

Core Data		Database Links			
Approved Symbol +	PCNA	Accession Numbers +			
Approved Name +	proliferating cell nuclear antigen	J04718	GenBank	EMBL	DDBJ UCSC
HGNC ID +	HGNC:8729	Mouse Genome Database ID +			
Status +	Approved	MGI:97503	MGD ID		
Chromosome +	20pter-p12	Rat Genome Database ID (mapped data supplied by RGD) +			
Previous Symbols +		RGD:3269	RGD ID		
Previous Names +		CCDS IDs +			
Aliases +		CCDS13087.1	CCDS ID		
Name Aliases +		Pubmed IDs +			
Locus Type +	gene with protein product	2565339	PMID	CiteXplore	
Gene Symbol Links		VEGA IDs +			
GENATLAS GeneCards GeneClinics/GeneTests GoPubmed		OTTHUMG00000031798	VEGA GeneView		
HCOP H-InvDB Treefam wikigenes		Ensembl ID (mapped data supplied by Ensembl) +			
		ENSG00000132646	Ensembl GeneView	UCSC	
		Entrez Gene ID (mapped data supplied by NCBI) +			
		5111	Gene	Map Viewer	
Specialist Database Links		RefSeq (mapped data supplied by NCBI) +			
COSMIC		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/>



HUGO Gene Nomenclature Committee



About HGNC

Gene Search

Guidelines

Gene Submission

Downloads

Home

Giving unique and meaningful names to every human gene

[HGNC Activities](#)

[Useful Links](#)

[HCOP Search Tool](#)

[Public Engagement of Science](#)

[FAQs](#)

[International Advisory Committee](#)

[Gene Families/Groupings](#)

Current committee:

- [Prof Sue Povey](#) (Chair)
- [Dr Elspeth Bruford](#)
- [Dr Ruth Lovering](#)
- [Dr Matt Wright](#)
- [Dr Varsha Khodiyar](#)
- [Dr Tam Sneddon](#)
- [Dr Kate Sneddon](#)
- [Mr Connie Talbot Jr.](#)

Bioinformatics support:

- [Dr Michael Lush](#)
- Mr Fabrice Ducluzeau



FEEDBACK - we welcome your feedback, please click [here](#) to leave your comments and/or suggestions.

PCNA



[Advanced Gene Search](#)

We have approved over **24,000** human gene symbols and names. Each symbol is unique and we ensure that each gene is only given one approved gene symbol. Search the HGNC database for your gene.

GCCTGGT
CATTGGAC
CGGCTCC [Request a Gene Symbol](#) - **NEW** online request form

Obtaining a gene symbol before publication will avoid any possible conflicts with existing symbols and will ensure that your gene is promptly recorded in our database and others. Any information that you provide will be treated in the strictest confidence. For bulk data submissions please follow the [Sequence Project Submission Format](#).



[Hot Topic](#)

We would like to hear your opinion on the latest issues in the world of human gene nomenclature.

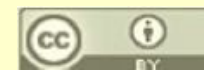


[Gene Families and Groupings](#)

We strongly encourage the use of a stem (or root) symbol as a basis for a hierarchical series that allows the easy identification of other related members in both database searches and the literature. Please contact us as soon as possible with new members of gene families, as some symbols may be reserved in our database.



Search gene.ucl.ac.uk/nomenclature

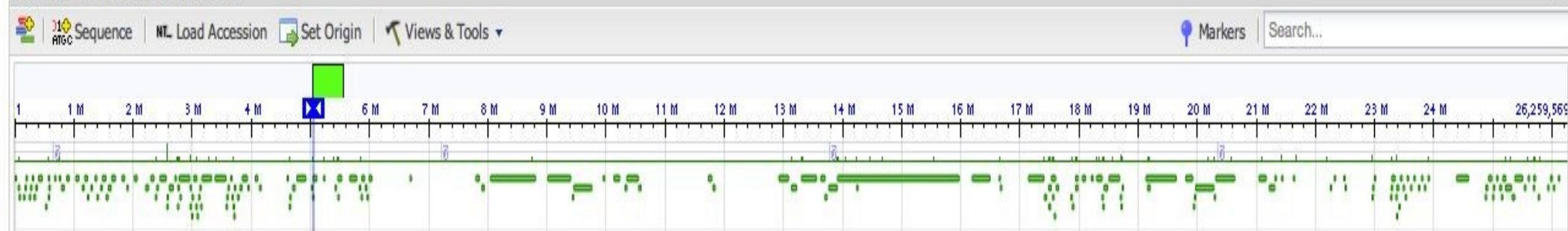


Homo sapiens chromosome 20 genomic contig, GRCh37 reference primary assembly

gi|27501067|ref|NT_011387.8

[Link To This Page](#) | [Help](#) | [Feedback](#) | [Printer-Friendly Page](#)

nt_011387.8 (26,259,569 bases)

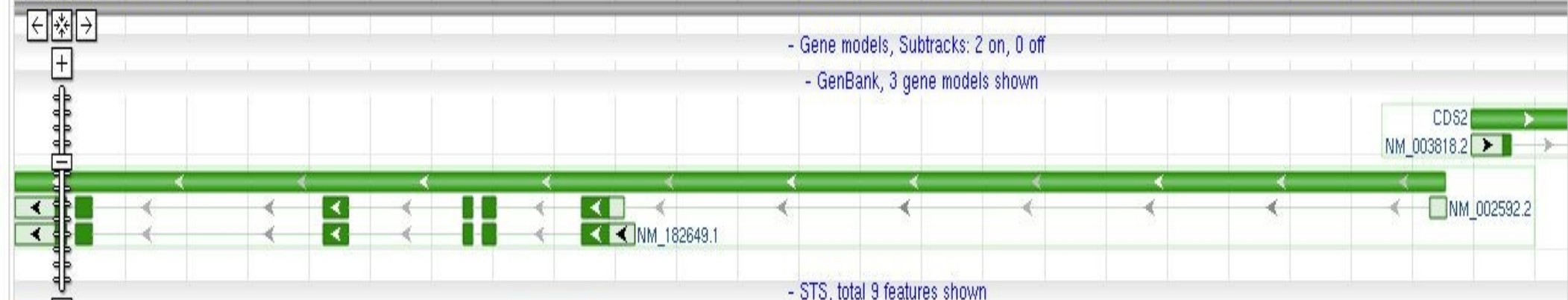


5,035,598 - 5,048,266 (12,669 bases shown, positive strand)



7 reference primary assembly

- nt_011387.8: Homo sapiens chromosome 20 genomic contig, GRCh37 reference primary assembly



5796 PMC134733P7

50240

Evidence Viewer for PCNA

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




Evidence Viewer
Homo sapiens
PCNA

PubMed
Nucleotide
Protein
OMIM
Genome
Taxonomy
PopSet

Key for display of mRNAs aligning in this region: [MapView](#)
[Evidence Viewer](#)
[Help](#)

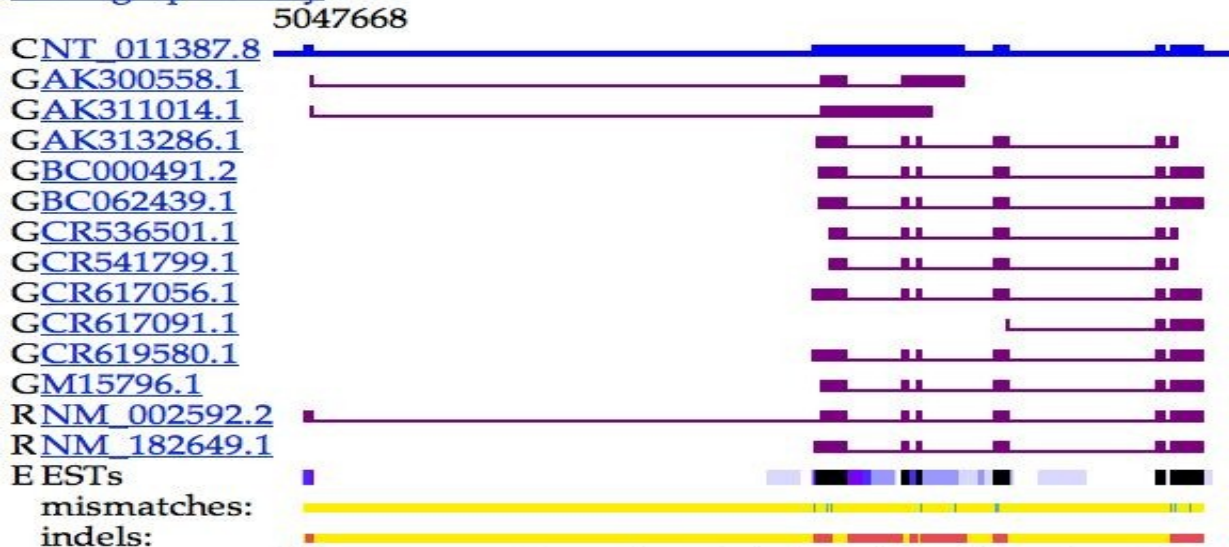
- Genomic sequence (C)
 - model exons, single (M) ■ mRNA exons, single (G, R)
 - model exons, overlapping (M) ■ mRNA exons, overlapping (G, R)
- C = contig; M = model mRNA; R = RefSeq mRNA; G = GenBank mRNA
R = new since last genome build; R = updated since last genome build

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](#)

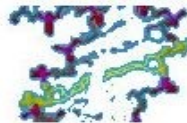


Mouse over mismatches, indels and unaligned regions to see their exon number.

SNP Polymorphisms in PCNA

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

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. 





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

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



Region	Contig position	mRNA pos	dbSNP rs# cluster id	Heterozygosity	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	

Search for

[Limits](#) [Preview/Index](#) [History](#) [Clipboard](#) [Details](#)

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...

- free text
- partial name and multiple species
- chromosome and symbol
- associated sequence accession number
- gene name (symbol)
- publication (PubMed ID)
- Gene Ontology (GO) terms or identifiers
- Genes with variants of clinical significance (under development)
- chromosome and species
- Enzyme Commission (EC) numbers

Search text

- [human muscular dystrophy](#)
- [transporter\[title\] AND \("Drosophila melanogaster"\[orgn\] OR "Mus musculus"\[orgn\]\)](#)
- [\(II\[chr\] OR 2\[chr\]\) AND adh*\[sym\]](#)
- [M11313\[accn\]](#)
- [BRCA1\[sym\]](#)
- [11331580\[PMID\]](#)
- ["cell adhesion"\[GO\]](#)
[10030\[GO\]](#)
- [gene_snp_clin\[filter\]](#)
- [Y\[CHR\] AND human\[ORGN\]](#)
- [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

Entrez Gene

- [Home](#)
- [About](#)
- [FAQ](#)
- [Help](#)
- [Gene Handbook](#)
- [Statistics](#)
- [Downloads \(FTP\)](#)

Mailing Lists

- [Gene](#)
- [RefSeq](#)

Feedback

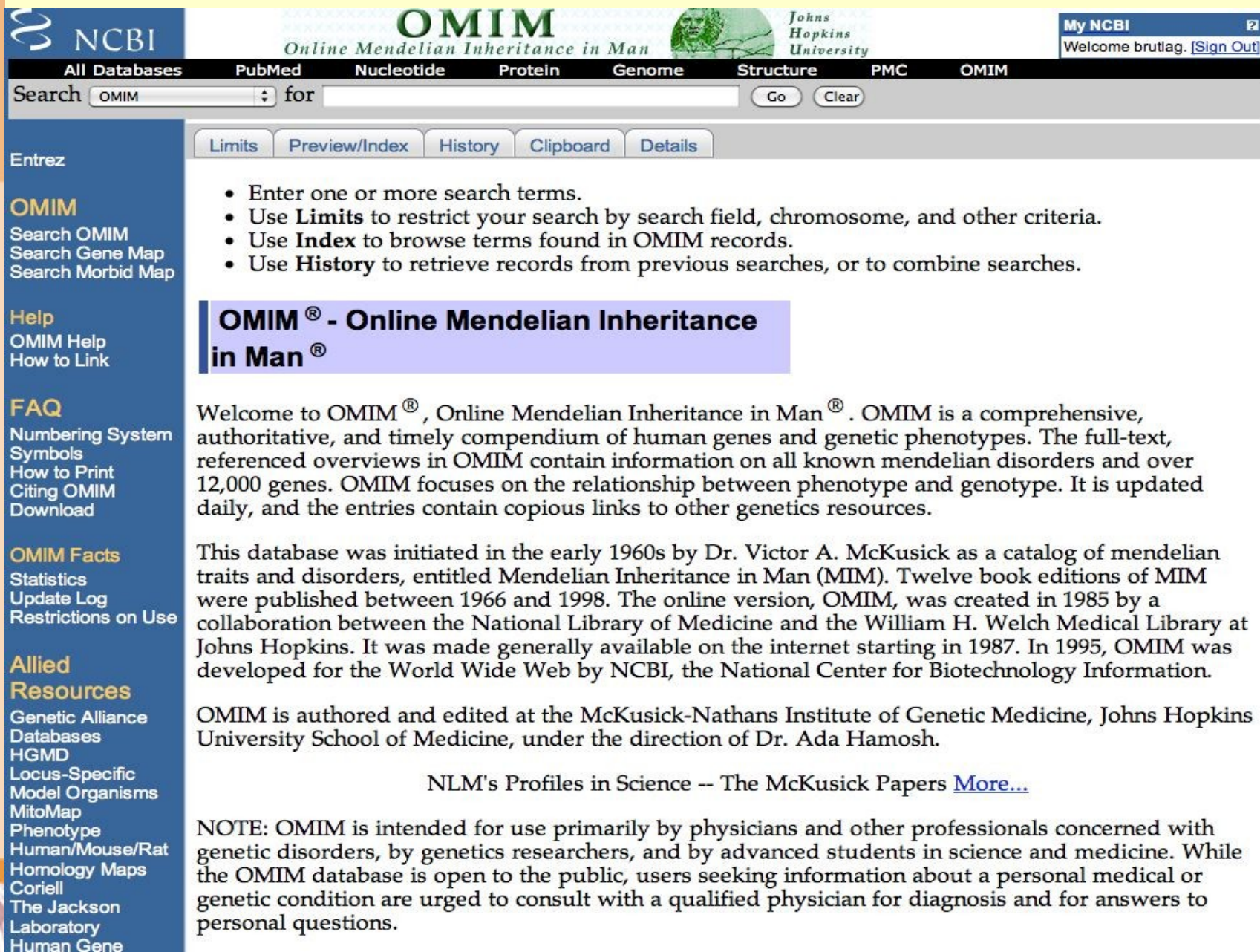
- [Help Desk](#)
- [Corrections](#)
- [About GeneRIFs](#)

Related Sites

- [BLAST](#)
- [Entrez Genome](#)
- [Genome Projects](#)
- [Genomic Biology](#)
- [GEO](#)
- [HomoloGene](#)
- [Map Viewer](#)
- [OMIM](#)
- [Probe](#)
- [RefSeq](#)

OMIM Home Page

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



The screenshot shows the OMIM (Online Mendelian Inheritance in Man) website interface. At the top, there is a navigation bar with the NCBI logo, the OMIM title, and the Johns Hopkins University logo. Below this is a search bar with a dropdown menu set to 'OMIM' and a search field. The main content area is divided into a left sidebar and a main text area. The sidebar contains links for Entrez, OMIM (Search OMIM, Search Gene Map, Search Morbid Map), Help (OMIM Help, How to Link), FAQ (Numbering System, Symbols, How to Print, Citing OMIM, Download), OMIM Facts (Statistics, Update Log, Restrictions on Use), and Allied Resources (Genetic Alliance Databases, HGMD, Locus-Specific Model Organisms, MitoMap, Phenotype, Human/Mouse/Rat Homology Maps, Coriell, The Jackson Laboratory, Human Gene).

OMIM
 Search OMIM
 Search Gene Map
 Search Morbid Map

Help
 OMIM Help
 How to Link

FAQ
 Numbering System
 Symbols
 How to Print
 Citing OMIM
 Download

OMIM Facts
 Statistics
 Update Log
 Restrictions on Use

Allied Resources
 Genetic Alliance Databases
 HGMD
 Locus-Specific Model Organisms
 MitoMap
 Phenotype
 Human/Mouse/Rat Homology Maps
 Coriell
 The Jackson Laboratory
 Human Gene

OMIM[®] - Online Mendelian Inheritance in Man[®]

Welcome to OMIM[®], Online Mendelian Inheritance in Man[®]. OMIM is a comprehensive, authoritative, and timely compendium of human genes and genetic phenotypes. The full-text, referenced overviews in OMIM contain information on all known mendelian disorders and over 12,000 genes. OMIM focuses on the relationship between phenotype and genotype. It is updated daily, and the entries contain copious links to other genetics resources.

This database was initiated in the early 1960s by Dr. Victor A. McKusick as a catalog of mendelian traits and disorders, entitled Mendelian Inheritance in Man (MIM). Twelve book editions of MIM were published between 1966 and 1998. The online version, OMIM, was created in 1985 by a collaboration between the National Library of Medicine and the William H. Welch Medical Library at Johns Hopkins. It was made generally available on the internet starting in 1987. In 1995, OMIM was developed for the World Wide Web by NCBI, the National Center for Biotechnology Information.

OMIM is authored and edited at the McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University School of Medicine, under the direction of Dr. Ada Hamosh.

NLM's Profiles in Science -- The McKusick Papers [More...](#)

NOTE: OMIM is intended for use primarily by physicians and other professionals concerned with genetic disorders, by genetics researchers, and by advanced students in science and medicine. While the OMIM database is open to the public, users seeking information about a personal medical or genetic condition are urged to consult with a qualified physician for diagnosis and for answers to personal questions.

OMIM Coverage

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



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




OMIM
Online Mendelian Inheritance in Man

Johns Hopkins University
 My NCBI
 Welcome brutlag. [\[Sign Out\]](#)

All Databases PubMed Nucleotide Protein Genome Structure PMC Taxonomy OMIM

Search for [Save Search](#)

History has expired due to inactivity.
 Did you mean: [color blindness](#) (46 items)

Display Show Send to

All: 47

Items 1 - 20 of 47 Page of 3 Next


- 1: +303800** GeneTests, Links
 COLORBLINDNESS, PARTIAL, DEUTAN SERIES; CBD
 OPSIN 1, MEDIUM-WAVE-SENSITIVE, INCLUDED; OPN1MW, INCLUDED
 Gene map locus [Xq28](#)
- 2: +190900** Links
 TRITANOPIA
 OPSIN 1, SHORT-WAVE-SENSITIVE, INCLUDED; OPN1SW, INCLUDED
 Gene map locus [7q31.3-q32](#)
- 3: +303900** GeneTests, Links
 COLORBLINDNESS, PARTIAL, PROTAN SERIES; CBP
 OPSIN1, LONG-WAVE-SENSITIVE, INCLUDED; OPN1LW, INCLUDED
 Gene map locus [Xq28](#)
- 4: 304000** Links
 COLORBLINDNESS, PARTIAL TRITANOMALY
- 5: %303700** GeneTests, Links
 COLORBLINDNESS, BLUE-MONO-CONE-MONOCHROMATIC TYPE; CBBM
 RED AND GREEN PIGMENT GENES, CONTROLLER OF, INCLUDED
 Gene map locus [Xq28, Xq28](#)

<http://www.omim.org/entry/303800>



Colorblindness in OMIM

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





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

• Clinical Synopsis
 • Gene map

Entrez Gene
N Nomenclature
R RefSeq
G GenBank
P Protein
U UniGene

LinkOut
... HGVS
... HGMD

Online Mendelian Inheritance in Man

All Databases
PubMed
Nucleotide
Protein
Genome
Structure
PMC
Taxonomy
OMIM

Search for

Display Show Send to

All: 1

[+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. [Waalder \(1968\)](#) 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

Copyright © Cardiff University 2006. All Rights Reserved



[Search help](#)
[Statistics](#)
[New genes](#)
[What's new](#)
[Background](#)
[Publications](#)
[Contact us](#)
[Register](#)
[Log in](#)
[Mutation submission](#)
[Locus-specific databases](#)
[Other useful links](#)

Symbol:

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	<input type="button" value="Get cDNA"/>	Not available	Not available	Mutation map

Mutation type	Number of mutations	Mutation data by type (register) or (login)
Missense/nonsense	3	<input type="button" value="Get mutations"/>
Splicing	0	No mutations
Regulatory	1	<input type="button" value="Get mutations"/>
Small deletions	0	No mutations
Small insertions	0	No mutations
Small indels	0	No mutations
Gross deletions	2	<input type="button" value="Get mutations"/>
Gross insertions	0	No mutations
Complex rearrangements	4	<input type="button" value="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



Home Page	About GeneTests	GENEReviews	Laboratory Directory	Clinic Directory	Educational Materials
---------------------------	---------------------------------	-----------------------------	--------------------------------------	----------------------------------	---------------------------------------



Items 1 - 2 of 2

One page.

The result of your search (below) includes a group of related disorders with your search term in **bold** or an alphabetical listing of the individual entries that match your search term. For more information about search results, see [Interpreting Your Search Results](#).

Search Result for Disease Name Containing 'colorblindness'

Blue-Mono-Cone-Monochromatic Type Colorblindness [Testing](#) [Resources](#) [OMIM](#) [Locus-Specific](#) [HGMD](#) [More Links](#)

Red-Green Color Vision Defects [Reviews](#) [Resources](#) [OMIM](#) [Locus-Specific](#) [HGMD](#) [More Links](#)

Red-Green Colorblindness

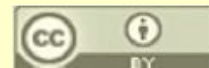
Disclaimer. GeneTests does not independently verify information provided by laboratories and does not warrant any aspect of a laboratory's work.

Contact GeneTests at NCBI

Copyright © 1993-2010, All Rights Reserved
University of Washington, Seattle
Terms of Use

Funding Support
National Institutes of Health

Sponsoring Institution
University of Washington, Seattle



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.

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

[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

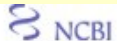
[Consumer Resources](#)
[OMIM](#)
[Gene](#)

Genes & Disease

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

Genes and Disease





Search
 This book
 Bookshelf
 PubMed

> Search

[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

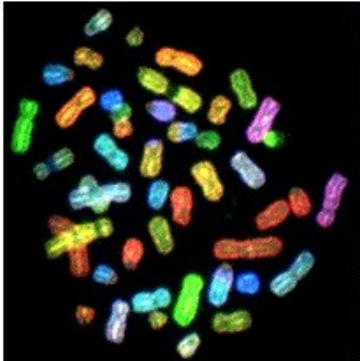
Genes and Disease is a collection of articles that discuss genes and the diseases that they cause. These genetic disorders are organized by the parts of the body that they affect. As some diseases affect various body systems, they appear in more than one chapter.

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.

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](#)




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 chromosomes 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.




Genes & Disease Table of Contents

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



















Genes and Disease





Search This book Bookshelf PubMed

[↑ Click on a Chromosome](#)
[↓ View Contents](#)
[Download PDF](#)
[Download PDA](#)
[Browse more Books](#)

 <p><u>Introduction to Genes and Disease</u></p>	 <p><u>Blood and Lymph Diseases</u></p>
 <p><u>Cancers</u></p>	 <p><u>The Digestive System</u></p>
 <p><u>Ear, Nose, and Throat</u></p>	 <p><u>Diseases of the Eye</u></p>
 <p><u>Female-Specific Diseases</u></p>	 <p><u>Glands and Hormones</u></p>
 <p><u>The Heart and Blood Vessels</u></p>	 <p><u>Diseases of the Immune System</u></p>
 <p><u>Male-Specific Diseases</u></p>	 <p><u>Muscle and Bone</u></p>
 <p><u>Neonatal Diseases</u></p>	 <p><u>The Nervous System</u></p>
 <p><u>Nutritional and Metabolic Diseases</u></p>	 <p><u>Respiratory Diseases</u></p>
 <p><u>Skin and Connective Tissue</u></p>	 <p><u>Chromosome Map</u></p>

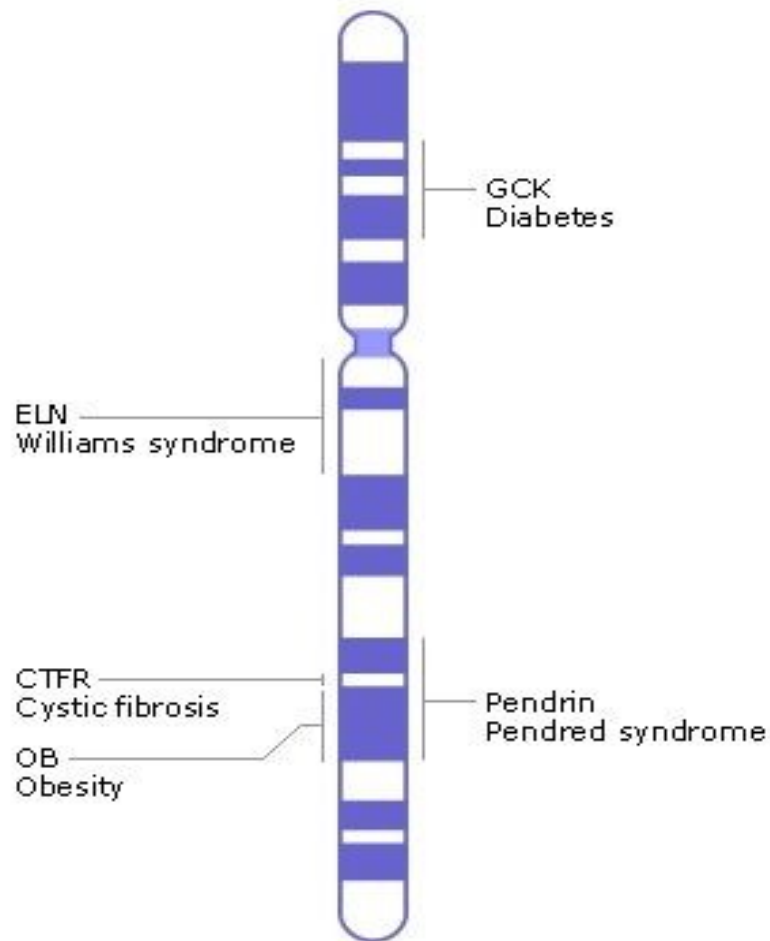
[Copyright and Disclaimer](#)

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](#).



Genome
Browser

ENCODE

Blat

Table
Browser

Gene Sorter

In Silico
PCRGenome
Graphs

Galaxy

VisiGene

Proteome
Browser

Utilities

Downloads

Release Log

Custom
Tracks

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.

Human (*Homo sapiens*) 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.

clade	genome	assembly	position or search term	image width	
<input type="text" value="Mammal"/>	<input type="text" value="Human"/>	<input type="text" value="GRCh37"/>	<input type="text" value="chr21:33,031,597-33,041,570"/>	<input type="text" value="800"/>	<input type="button" value="submit"/>

[Click here to reset](#) the browser user interface settings to their defaults.

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

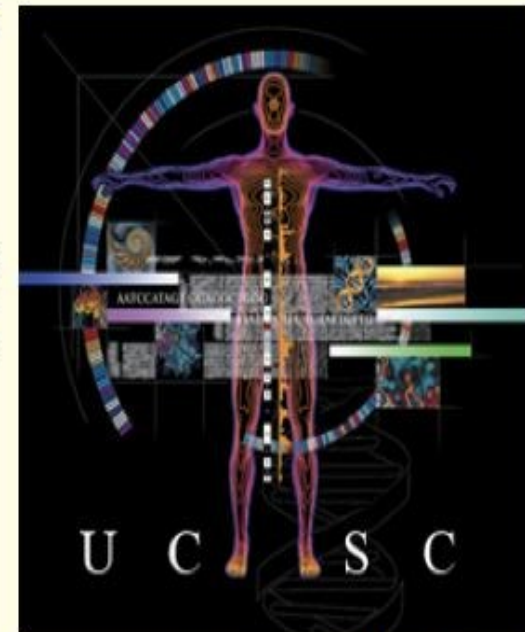
Displays all of chromosome 7

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)

PCNA Entry at UCSC

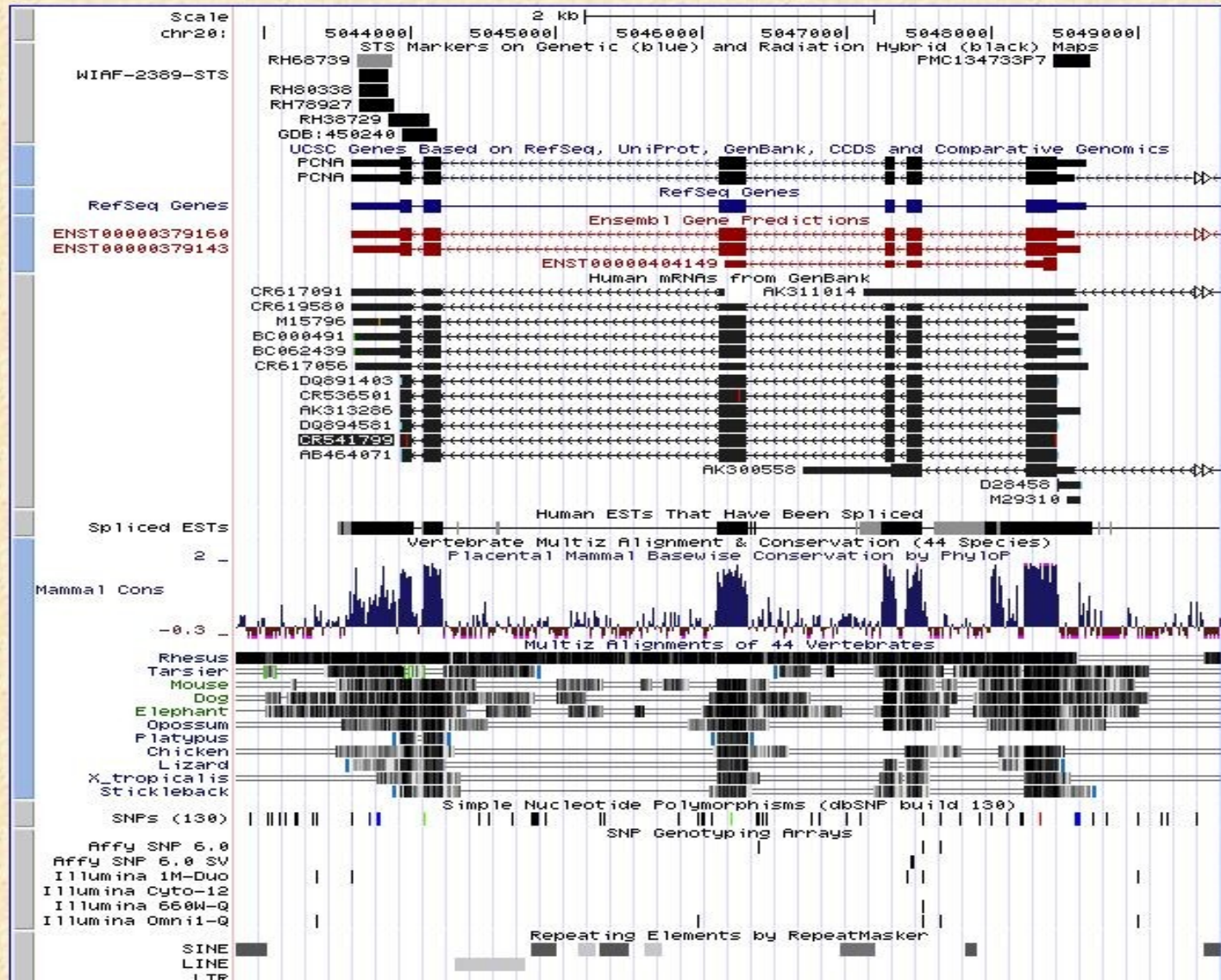
<http://genome.cse.ucsc.edu/cgi-bin/hgTracks?insideX=115&revCmplDisp=0&hgsid=149552355&hgt.out1=1.5x&position=chr20%3A5043932-5049572>

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 chr20:5042803-5049572 jump clear size 6,770 bp. configure

chr20 (p12.3) 20p13.12.3 p12.1 q12 q13.2



UCSC Proteome Browser

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

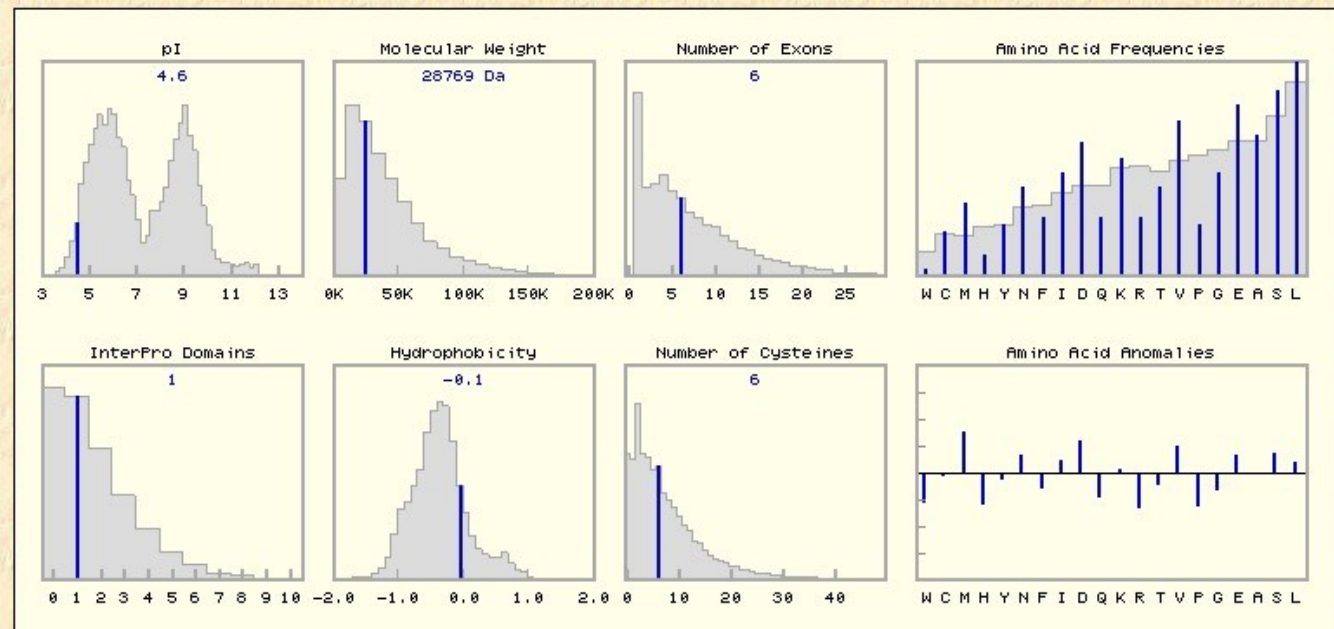
[Home](#)
[UCSC Proteome Browser](#)
[PDF/PS](#)
[New Query](#)
[H](#)

Protein **P12004** (aka PCNA_HUMAN) Proliferating cell nuclear antigen (PCNA) (Cyclin).
 Organism: Homo sapiens (human)

Move
 Current scale: FULL Rescale to

AA Scale | 1 | 50 | 100
 AA Sequence MFEARLVQGSILKKVLEALKDLINERACHDISSSGVNLQSMDSHSVSLVQLTLRSEGFDTYRCDRNLAMGVNLTSMKILKCGAGNEDIITLRAEDNADTLALVFEAFNQEKVS
 Genome Browser
 Exons 1 2
 Polarity +
 Hydrophobicity
 Cysteines
 Predicted Glycosylation
 Superfamily/SCOP
 AA Anomalies
 AA Scale | 1 | 50 | 100

[Explanation of Protein Tracks](#)



[Explanation of Protein Property Histograms](#)

UCSC Tutorials

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

[Login](#) | [Register](#)

 OpenHelix™

Menu

Search

[Recent Queries](#)

Tutorial Suites

[Catalog](#)

[By Category](#)

[Free Tutorials](#)

Other Training/Materials

[Quick Reference Cards](#)

[Curriculums](#)

[Newsletter](#)

[Seminars/Workshops](#)

[Site Documentation](#)

OpenHelix Genomics Blog

About OpenHelix

Contact and Feedback

Latest Blog Post

- [Top SNPs of the year](#)
- [Webcasts in PubMed?](#)

OpenHelix Highlights

- New Tutorial added: [VisANT](#), [BioSystems](#), [DCODE](#)
- New OpenHelix/UCSC publication in Current Protocols: [The UCSC Genome Browser: what every molecular biologist should know](#)

Latest News

- [OH Launches](#)

The UCSC Genome Browser Introduction

Tutorial and training materials by OpenHelix



Online Tutorial Suite

Learn to use the [UCSC Genome Browser](#) with this **free tutorial**, sponsored by **UCSC Genome Bioinformatics Group**. The UCSC Genome Browser provides a way to examine the data from many genomes, with extensive annotation tracks for various data types including known genes, predicted genes, SNPs, comparative multi-species analysis and much more. This introductory tutorial focuses on the foundation and framework for the organization and display of the data, and basic text and sequence searches. This tutorial, which is the first in a series of three tutorials on the UCSC Genome Browser, will get you on your way to expertly navigating this vital tool for genomic research.

You'll learn:

- to perform basic text searches on the UCSC Genome Browser
- to understand and customize the displays in genomic regions of interest
- to start with a sequence and find genomic regions of interest using BLAT

More about the resource:

The [UCSC Genome Browser](#), sometimes referred to as the "Golden Path" browser, offers a well-organized and user-friendly view of the human genome, along with dozens of other genomes as well. The official genomic sequence is supplemented with many other data types which are useful to researchers: expression, variation, comparative genomics, and many more. The data can be accessed with simple text or sequence searches using BLAT, or probed in depth with customized queries. Be sure to see the other UCSC tutorials for advanced topics and additional tools as well.

 Launch	 Download	 Download	 Download	 Link
Online Tutorial	PowerPoint Slides <small>(Optimized for Windows*)</small>	Slide Handouts <small>(PDF file)</small>	Hands-on Exercises <small>(PDF file)</small>	Visit the Resource

Related tutorials

This tutorial is a part of the tutorial group **UCSC Tutorials**. You might find the other tutorials in the group interesting:

[UCSC Genome Browser: Custom Tracks and Table Browser](#): UCSC Genome Browser advanced topics
[UCSC Genome Browser: The Additional Tools](#): Additional tools at the UCSC Genome Browser

Categories

View additional tutorials for resources in

[Genome Databases \(eu\)](#)
[Algorithms and Analysis](#)

[Click here for technical information on using OpenHelix tutorial and training materials](#)



Ensembl Genomes Home Page

<http://www.ensembl.org/>

Search: for

e.g. human gene BRCA2 or rat X:100000..200000 or insulin

Browse a Genome

The Ensembl project produces genome databases for vertebrates and other eukaryotic species, and makes this information freely available online.

Click on a link below to go to the species' home page.

Favourite genomes [\(Change favourites\)](#)



Human
GRCh37



Mouse
NCBIM37










Zebrafish
Zv8

All genomes


New to Ensembl?

Did you know you can:

-  [Learn how to use Ensembl](#)
with our video tutorials and walk-throughs
-  [Add custom tracks](#)
using our new Control Panel
-  [Upload your own data](#)
and save it to your Ensembl account
-  [Search for a DNA or protein sequence](#)
using BLAST or BLAT
-  [Fetch only the data you want](#)
from our public database, using the Ensembl Perl API
-  [Download our databases via FTP](#)
in FASTA, MySQL and other formats
-  [Mine Ensembl with BioMart](#)
and export sequences or tables in text, html, or Excel format

Did you know...?

A preliminary assembly of the common baboon (*Papio hamadryas*) is now available on our pre! site, <http://pre.ensembl.org/B>



Still got questions? [Try our FAQs](#)

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 (6:133017695-133161157)
 - Gene (BRCA2)
 - Transcript (FOXP2-203)
 - Variation (rs1333049)

- [Configure this page](#)
- [Manage your data](#)
- [Export data](#)
- [Bookmark this page](#)

Search Ensembl Human

Search for:

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

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](#).

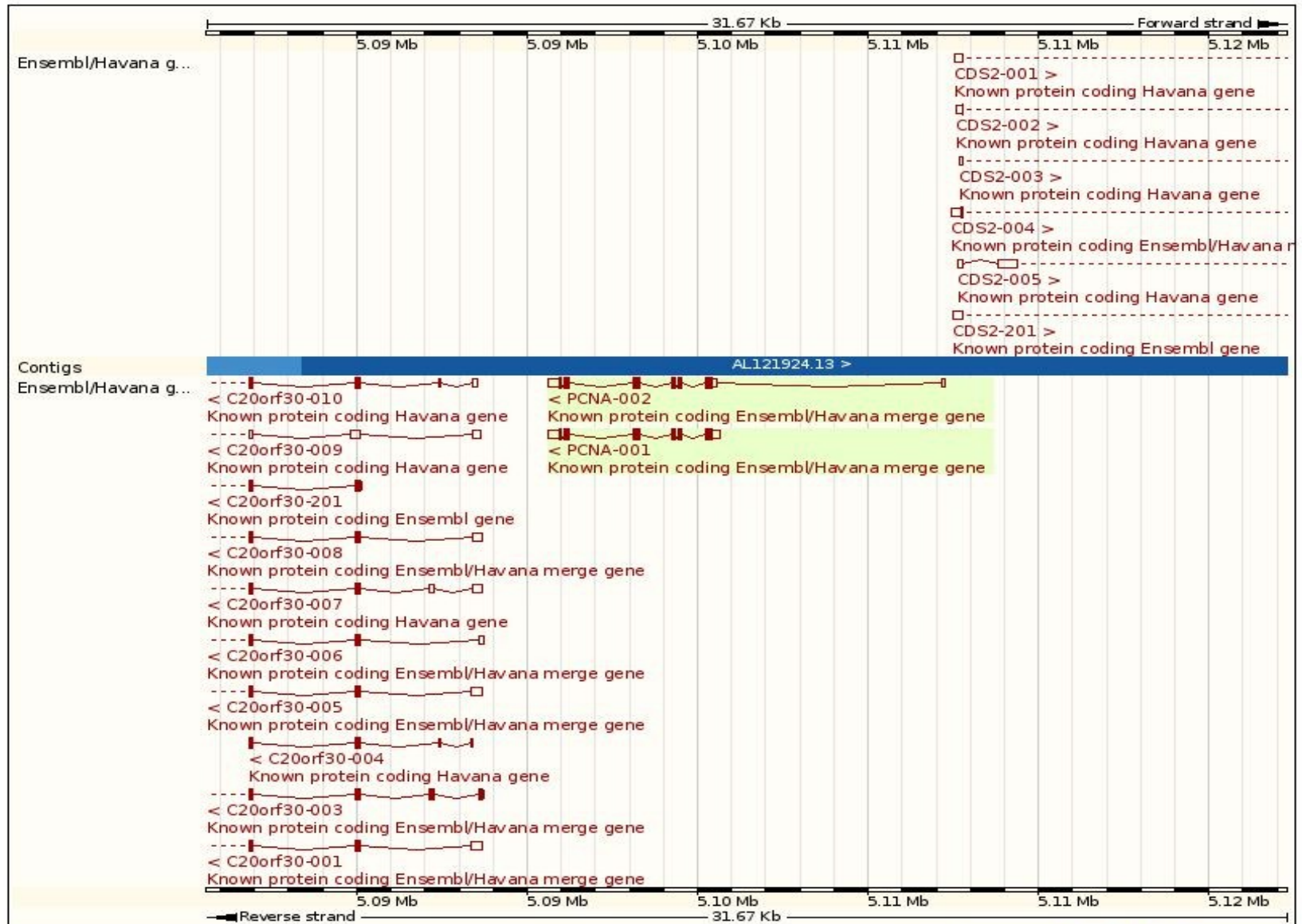
Alternative genes This Known protein coding entry corresponds to the following database identifiers:

Havana Gene: [OTTHUMG00000031798](#) [\[view all locations\]](#)

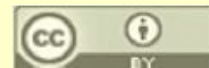
Ensembl PCNA Transcripts

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

Transcripts

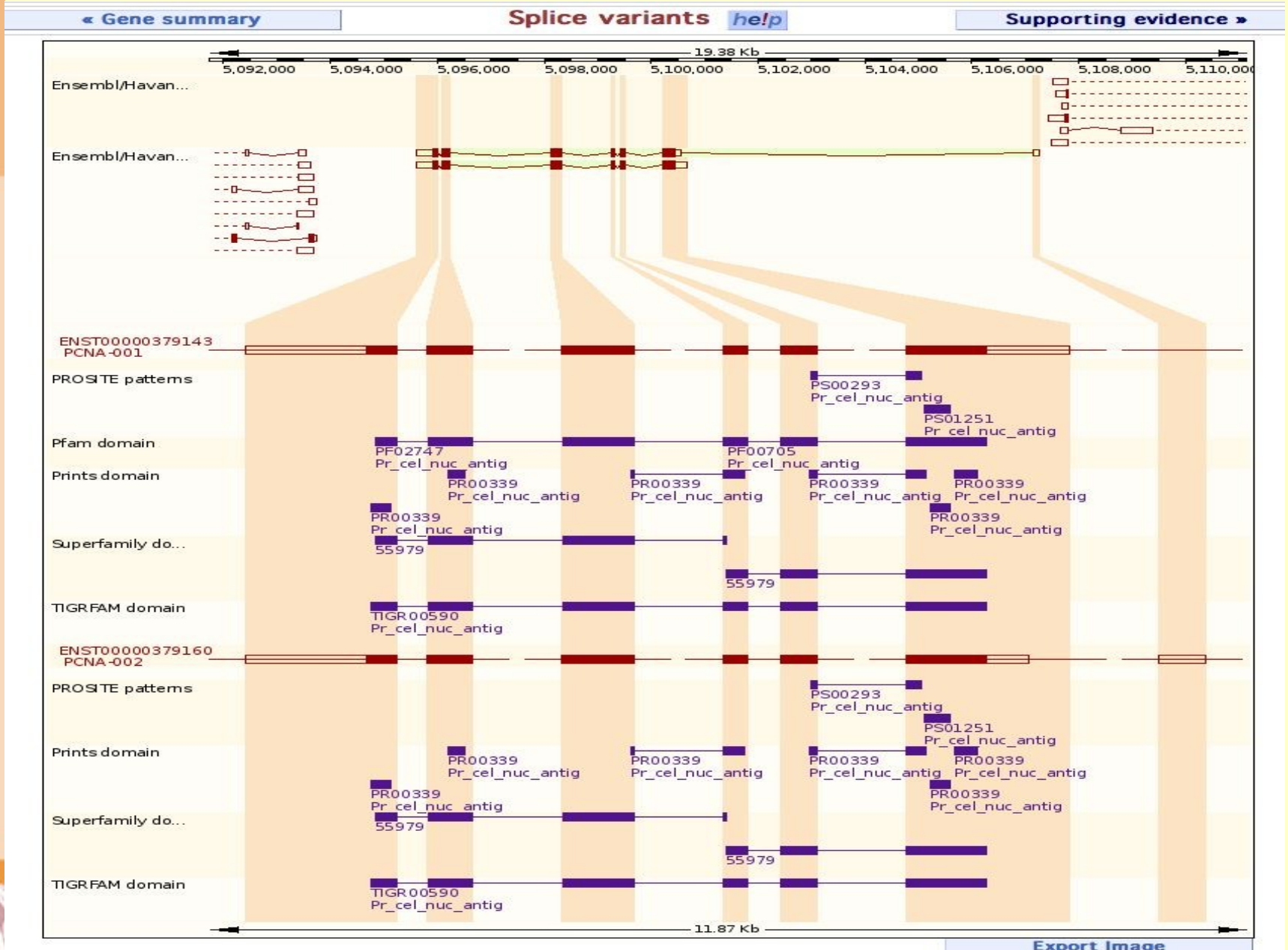


Export Image



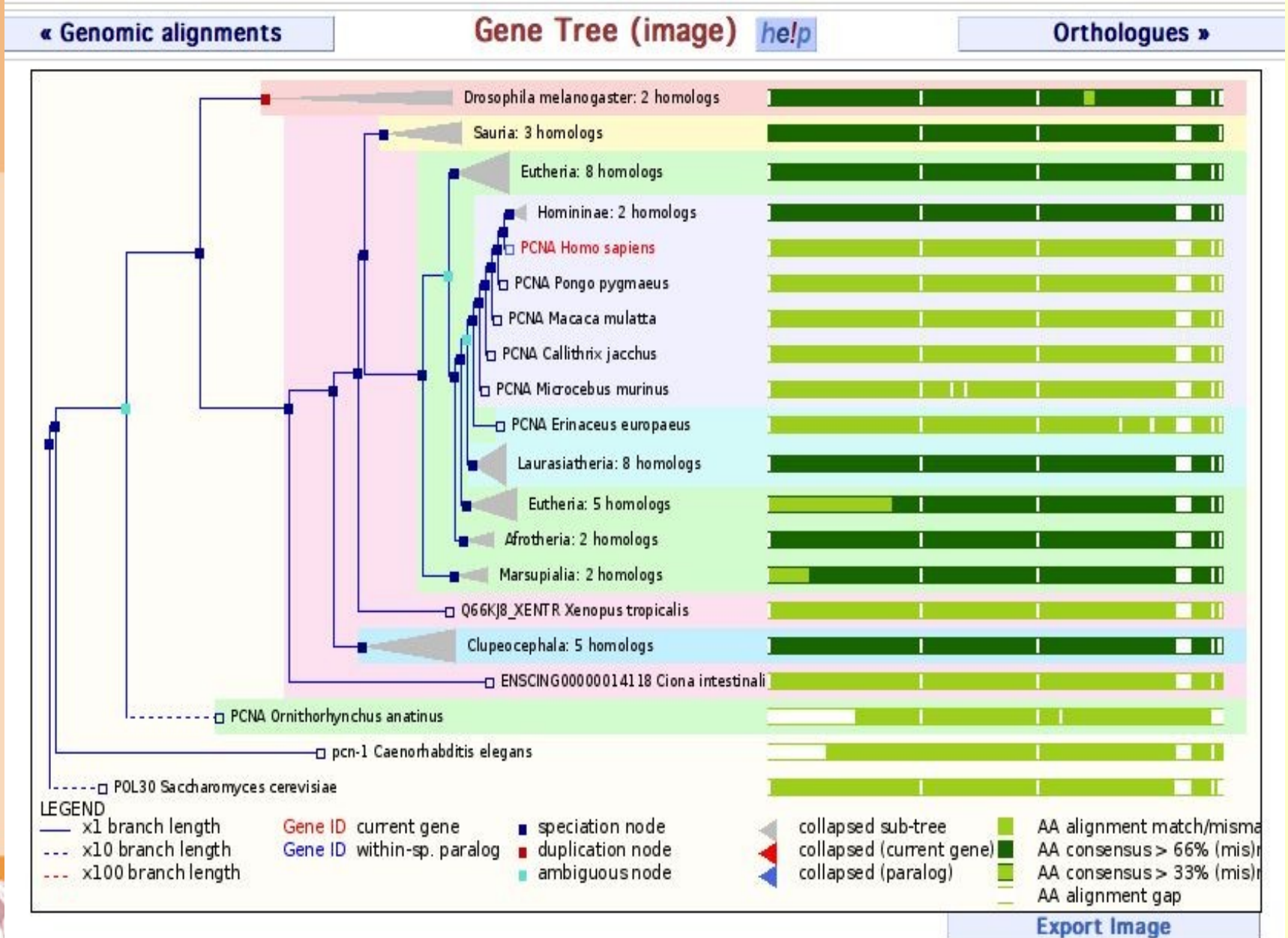
Ensembl PCNA Splice Variants

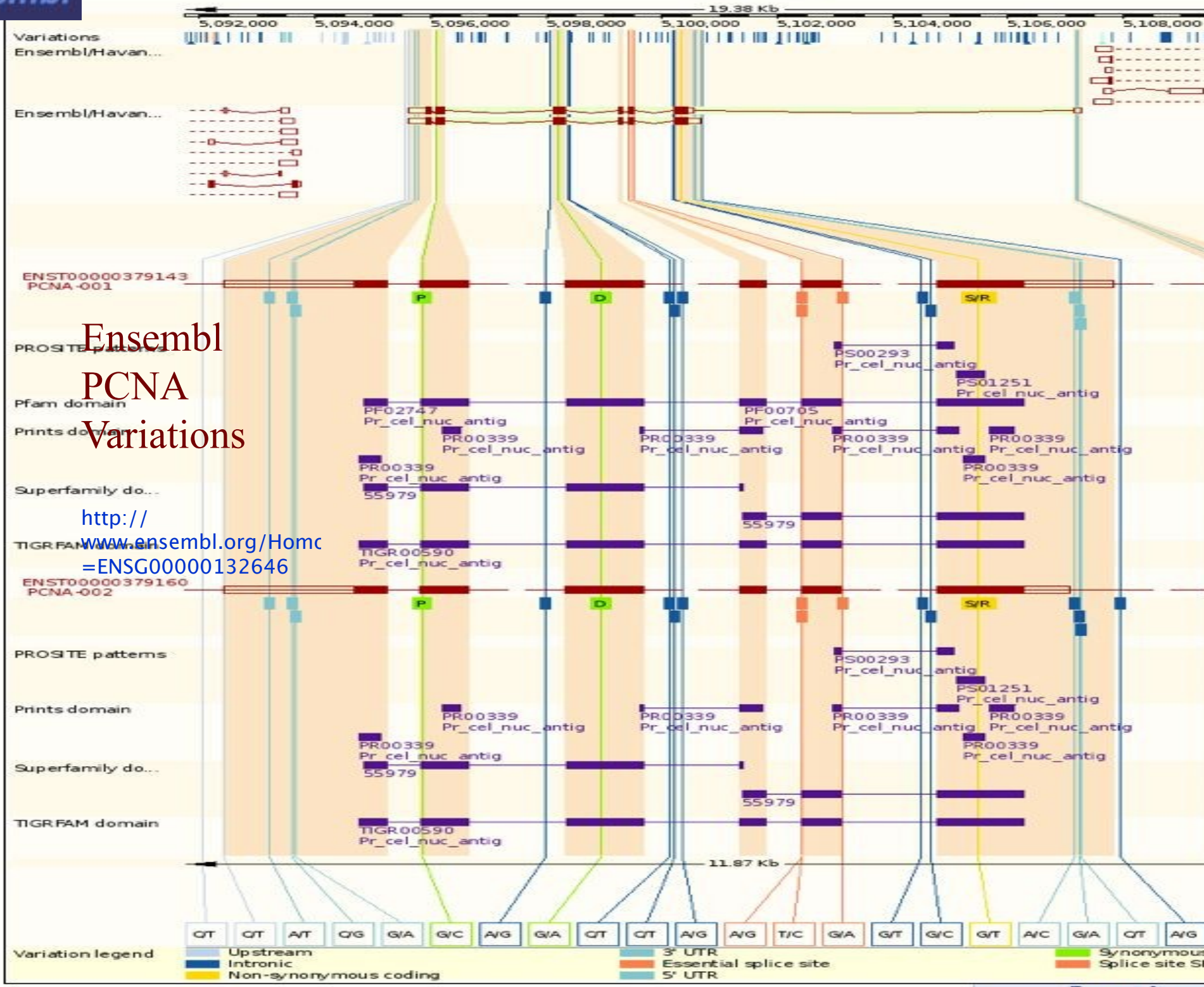
http://www.ensembl.org/Homo_sapiens/Gene/Summary?g=ENSG00000132646



Ensembl PCNA Gene Tree (Pecan)

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

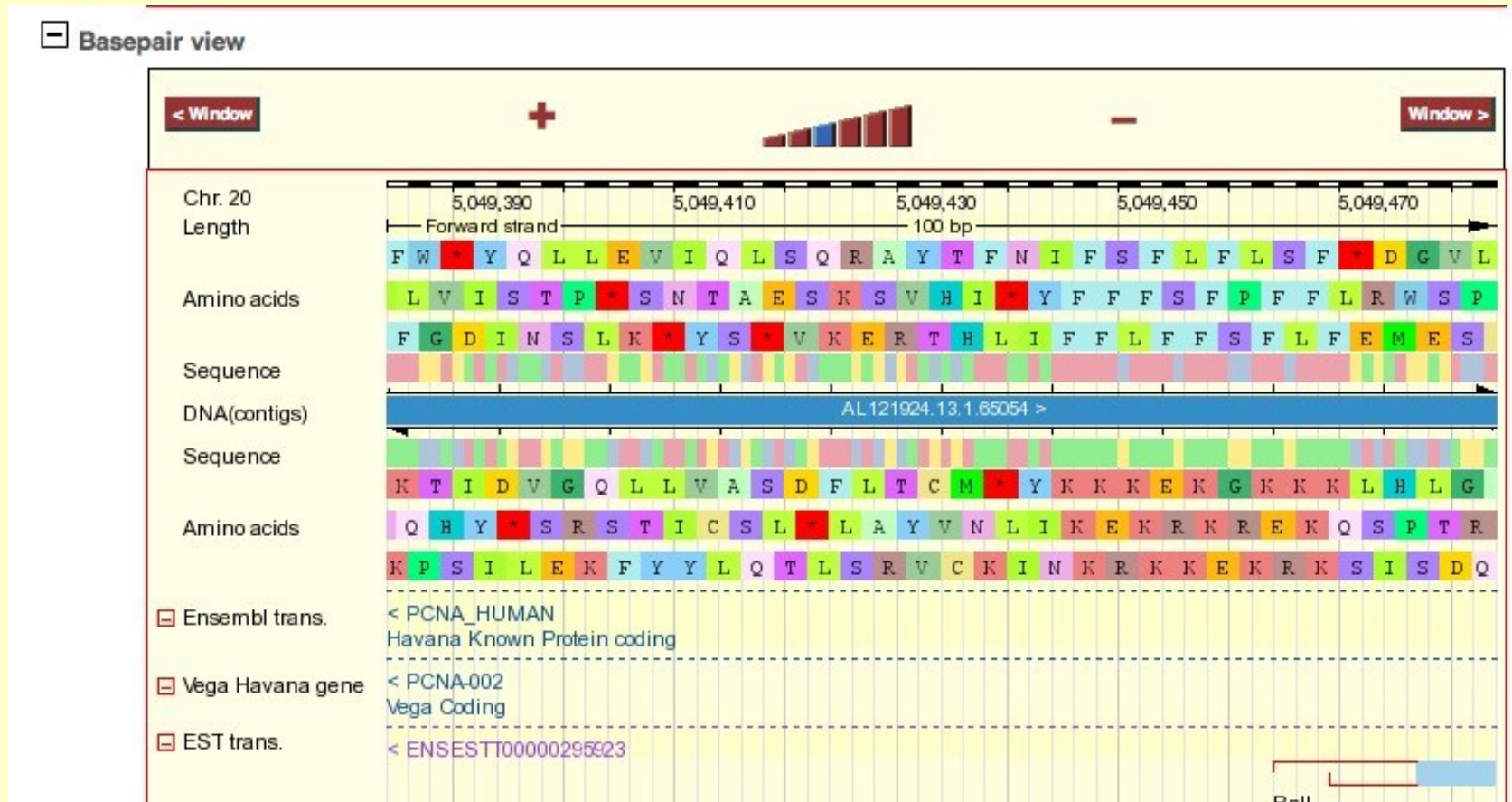




Ensembl PCNA Variations


<http://www.ensembl.org/Homc>
=ENSG00000132646


Ensembl Protein View of PCNA



Ensembl Tutorials

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





Search all Ensembl:

Ensembl release 42 - Dec 2006
HOME · BLAST · BIOMART · SITEMAP **HELP**

Your Ensembl

- [Show account](#) · [Log out](#)
- [Save bookmark](#)
- [Save configuration as...](#)

Helpdesk

- [Browse Help Articles](#)
- [Animated Tutorials](#)

Ensembl Archive

[Stable Archive! link for this page](#)

Animated Tutorials - Table of Contents



Animated Tutorials

The tutorials listed below are Flash animations of some of our training presentations, with added popup notes in place of a soundtrack. We are gradually adding to the list, so please check back regularly (the list will also be included in the bimonthly Release Email, which is sent to the [ensembl-announce mailing list](#)).

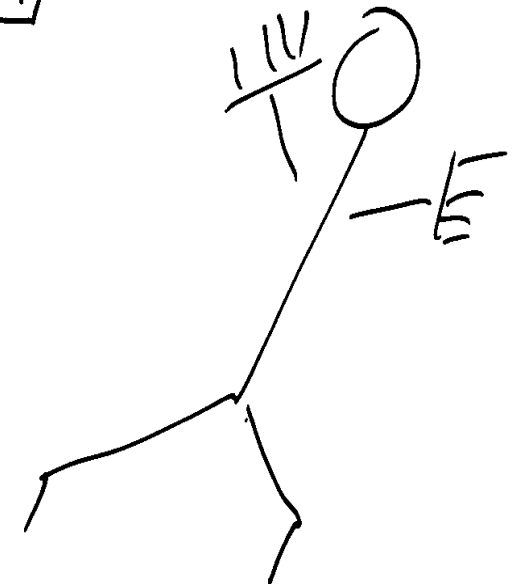
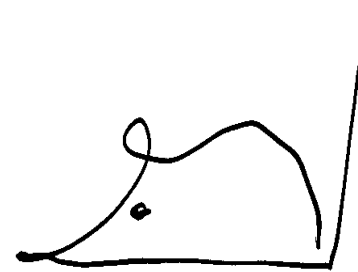
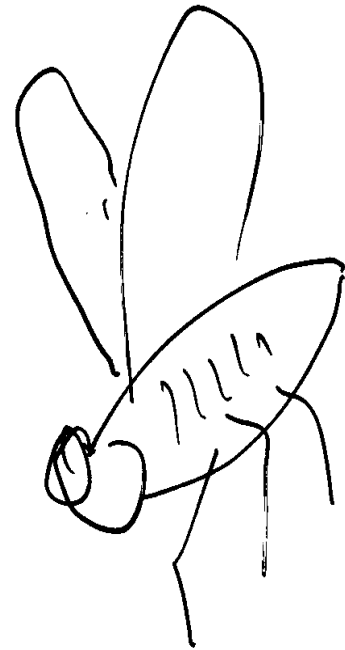
Please note that files are around 3MB per minute, so if you are on a dialup connection, playback may be jerky.

Title	Running time (minutes)
Known and Novel Genes	1:34
Overview of GeneView	3:45
Search for a Gene	0:51
Supporting Evidence for an Ensembl Gene	1:53

© 2007 [WTSI](#) / [EBI](#). Ensembl is available to [download for public use](#) - please see the [code licence](#) for details.

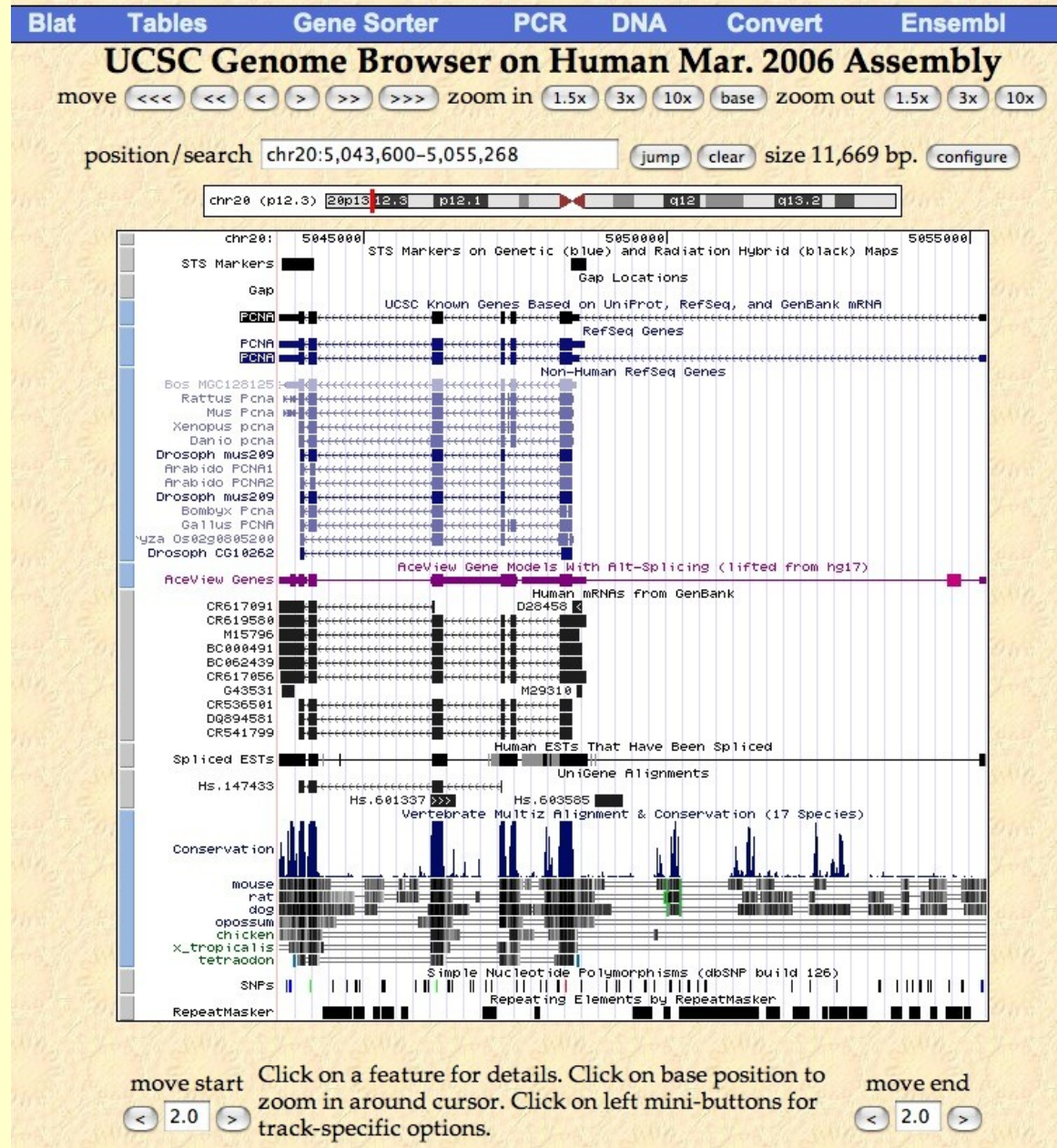



Comparative Genomics



L'homme et ses cousins germains
F. Jacob

UCSC PCNA Entry



UCSC Comparative Genomics



Comparative Genomics

Conservation <input type="button" value="pack"/>	Most Conserved <input type="button" value="hide"/>	Fugu Chain <input type="button" value="hide"/>	Fugu Net <input type="button" value="hide"/>	Tetraodon Chain <input type="button" value="hide"/>
Tetraodon Net <input type="button" value="hide"/>	Tetraodon Ecores <input type="button" value="hide"/>	Zebrafish chain <input type="button" value="hide"/>	Zebrafish Net <input type="button" value="hide"/>	X. tropicalis Chain <input type="button" value="hide"/>
X. tropicalis Net <input type="button" value="hide"/>	Chicken Chain <input type="button" value="hide"/>	Chicken Net <input type="button" value="hide"/>	Opossum Chain <input type="button" value="hide"/>	Opossum Net <input type="button" value="hide"/>
Cow Chain <input type="button" value="hide"/>	Cow Net <input type="button" value="hide"/>	Dog Chain <input type="button" value="hide"/>	Dog Net <input type="button" value="hide"/>	Rat Chain <input type="button" value="hide"/>
Rat Net <input type="button" value="hide"/>	Mouse Chain <input type="button" value="hide"/>	Mouse Net <input type="button" value="hide"/>	Rhesus Chain <input type="button" value="hide"/>	Rhesus Net <input type="button" value="hide"/>
Chimp Chain <input type="button" value="hide"/>	Chimp Net <input type="button" value="hide"/>			



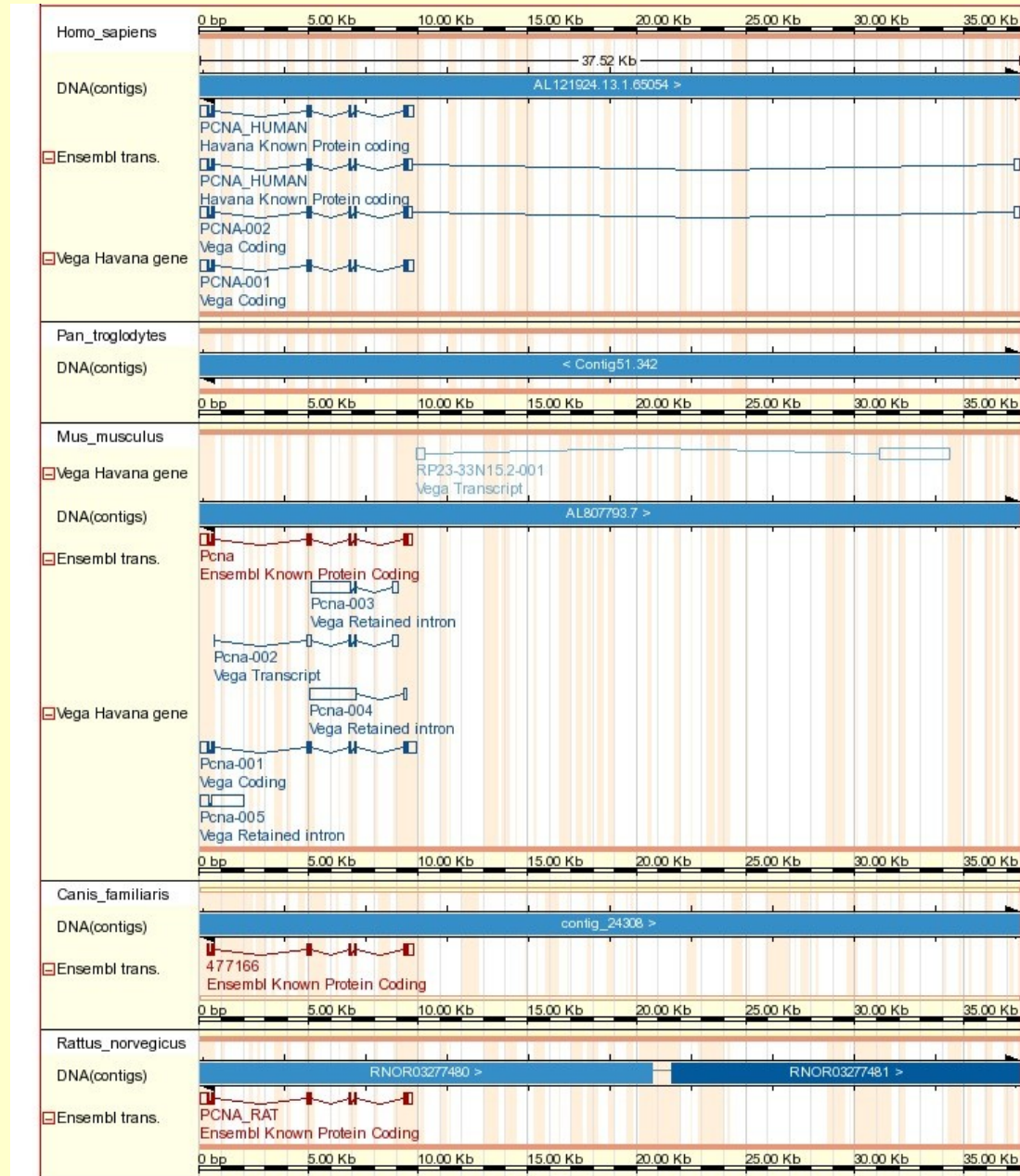
Ensembl Genome Alignments




<input type="checkbox"/> Alignments	<p>This gene can be viewed in genomic alignment with other species</p> <ul style="list-style-type: none"> view genomic alignment with 7 eutherian mammals Pecan view genomic alignment with 9 amniota vertebrates Pecan view genomic alignment with <i>Rattus norvegicus</i> view genomic alignment with <i>Mus musculus</i> view genomic alignment with <i>Bos taurus</i> view genomic alignment with <i>Monodelphis domestica</i> view genomic alignment with <i>Macaca mulatta</i> view genomic alignment with <i>Loxodonta africana</i> view genomic alignment with <i>Echinops telfairi</i> view genomic alignment with <i>Oryctolagus cuniculus</i> view genomic alignment with <i>Dasyopus novemcinctus</i> view genomic alignment with <i>Canis familiaris</i> view genomic alignment with <i>Pan troglodytes</i> view genomic alignment with <i>Gallus gallus</i> view genomic alignment with <i>Ornithorhynchus anatinus</i>
<input type="checkbox"/> Orthologue	<p>The following gene(s) have been identified as putative orthologues:</p>




Ensembl Genome Alignments (PCNA)




NCBI Homologene Database





HomoloGene
Discover Homologs

[Help](#)
My NCBI 
Welcome brutlag. [Sign Out](#)

All Databases
PubMed
Nucleotide
Protein
Genome
Structure
Map Viewer
Gene
UniGene
OMIM

Search for

About Entrez

HomoloGene

Home
Query Tips
Build Procedure
FTP Site

Genome Resources

Homo sapiens
Mus musculus
Rattus norvegicus
Danio rerio

HomoloGene is a system for automated detection of homologs among the annotated genes of several completely sequenced eukaryotic genomes.

HomoloGene Release 53 Statistics

Initial numbers of genes from complete genomes, numbers of genes placed in a homology group, and the numbers of groups for each species.

Species	Number of Genes		HomoloGene groups
	Input	Grouped	
H.sapiens	22,873	20,111	19,494
P.troglodytes	25,096	17,815	17,135
C.familiaris	19,766	16,638	16,205
M.musculus	24,175	20,528	19,157
R.norvegicus	21,991	19,008	17,776
G.gallus	18,029	12,183	11,358
D.melanogaster	14,033*	7,998	7,795
A.gambiae	13,909	8,394	7,838
C.elegans	20,056*	5,187	4,955
S.pombe	5,043	3,203	3,167
S.cerevisiae	5,863	4,739	4,589
K.lactis	5,335	4,454	4,423
E.gossypii	4,726	3,943	3,934
M.grisea	11,109	6,302	5,886
N.crassa	10,079	5,915	5,909
A.thaliana	26,659	11,180	10,857
O.sativa	33,553	11,055	9,464
P.falciparum	5,222	975	954

* indicates organisms where new genome annotation data is used in this build.
Last updated on: Tue Nov 14 2006

What's New

HomoloGene release 53 incorporates updated annotation for two species: *Caenorhabditis elegans* (included in NCBI *Caenorhabditis elegans* release 6.1, available Oct. 13, 2006), and *Drosophila melanogaster* (included in NCBI release 8.1, available Oct. 17, 2006).

Now you can download mRNA, protein, and genomic sequences of genes in a HomoloGene entry, using the Download link.

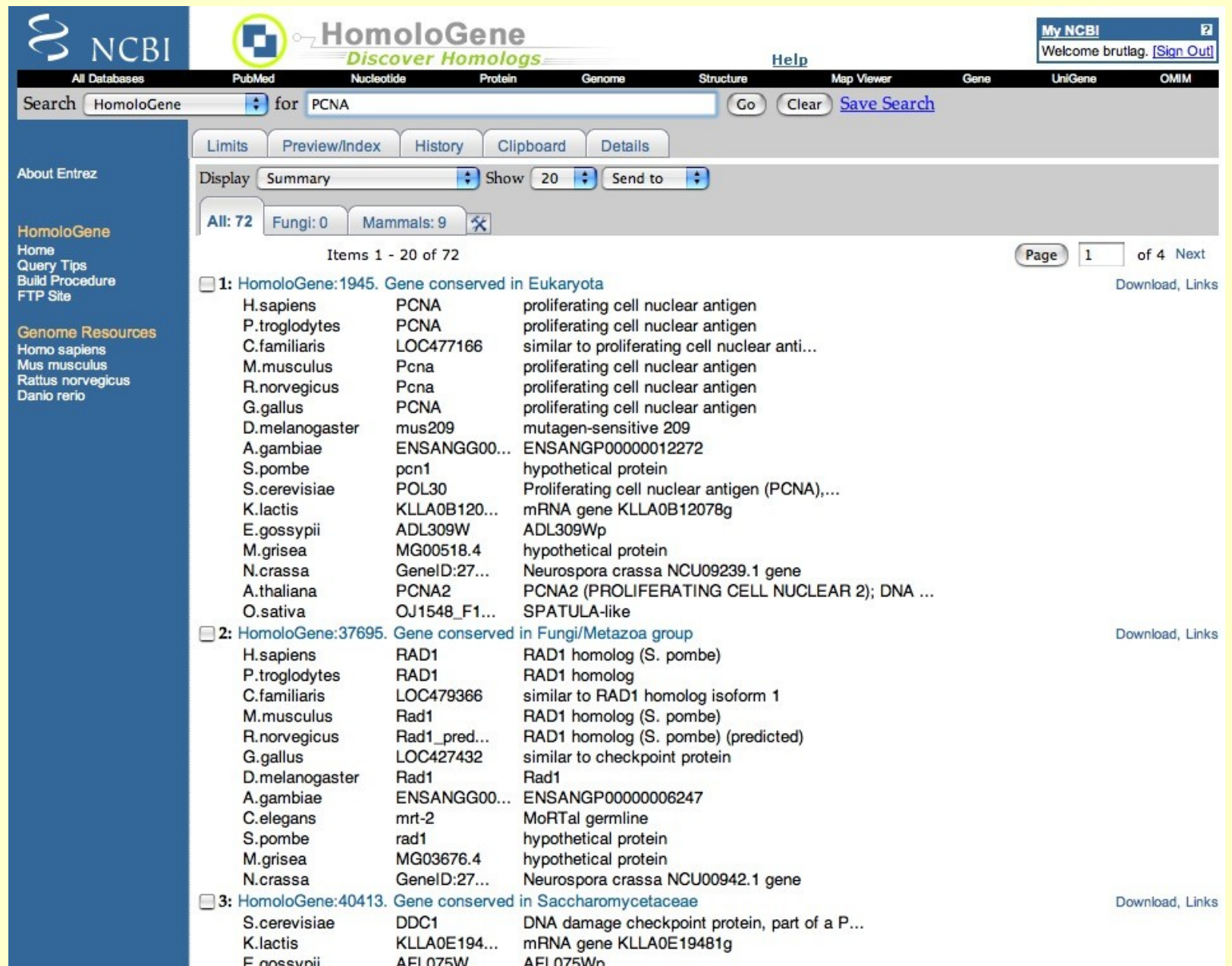
Please note that the FTP files `hmlg.ftp` and `hmlg.trip.ftp` are now deprecated, and will be retired on Jan. 1, 2007. Please use the files `homologene.data` and/or `homologene.xml.gz` instead.

Tip of The Day

You can restrict your search results to Homologs from FlyBase by using 'Limits' in the tool bar.

[\[More Tips\]](#)

Homologene PCNA Entries



NCBI HomoloGene
Discover Homologs

Search: HomoloGene for PCNA [Go] [Clear] [Save Search]

Display: Summary Show 20 Send to

All: 72 Fungi: 0 Mammals: 9

Items 1 - 20 of 72

Page 1 of 4 Next

1: HomoloGene:1945. Gene conserved in Eukaryota [Download, Links]

H.sapiens	PCNA	proliferating cell nuclear antigen
P.troglodytes	PCNA	proliferating cell nuclear antigen
C.familiaris	LOC477166	similar to proliferating cell nuclear anti...
M.musculus	Pcna	proliferating cell nuclear antigen
R.norvegicus	Pcna	proliferating cell nuclear antigen
G.gallus	PCNA	proliferating cell nuclear antigen
D.melanogaster	mus209	mutagen-sensitive 209
A.gambiae	ENSANGG00...	ENSANGP00000012272
S.pombe	pcn1	hypothetical protein
S.cerevisiae	POL30	Proliferating cell nuclear antigen (PCNA),...
K.lactis	KLLA0B120...	mRNA gene KLLA0B12078g
E.gossypii	ADL309W	ADL309Wp
M.grisea	MG00518.4	hypothetical protein
N.crassa	GeneID:27...	Neurospora crassa NCU09239.1 gene
A.thaliana	PCNA2	PCNA2 (PROLIFERATING CELL NUCLEAR 2); DNA ...
O.sativa	OJ1548_F1...	SPATULA-like

2: HomoloGene:37695. Gene conserved in Fungi/Metazoa group [Download, Links]

H.sapiens	RAD1	RAD1 homolog (S. pombe)
P.troglodytes	RAD1	RAD1 homolog
C.familiaris	LOC479366	similar to RAD1 homolog isoform 1
M.musculus	Rad1	RAD1 homolog (S. pombe)
R.norvegicus	Rad1_pred...	RAD1 homolog (S. pombe) (predicted)
G.gallus	LOC427432	similar to checkpoint protein
D.melanogaster	Rad1	Rad1
A.gambiae	ENSANGG00...	ENSANGP00000006247
C.elegans	mrt-2	MoRTal germline
S.pombe	rad1	hypothetical protein
M.grisea	MG03676.4	hypothetical protein
N.crassa	GeneID:27...	Neurospora crassa NCU00942.1 gene

3: HomoloGene:40413. Gene conserved in Saccharomycetaceae [Download, Links]

S.cerevisiae	DDC1	DNA damage checkpoint protein, part of a P...
K.lactis	KLLA0E194...	mRNA gene KLLA0E19481g
E.gossypii	AFL075W	AFL075Wp