



Topic 5. Genes and Genomes

Retrieving gene information from the NCBI Gene & Gene Cards. Browsing genome information at Ensembl & the UCSC genome browser. Population genomics: HapMap & the 1000 Genomes Browser. Comparative genomics.

The Genotype to Phenotype challenge

Topic 5 - Genes & Genomes

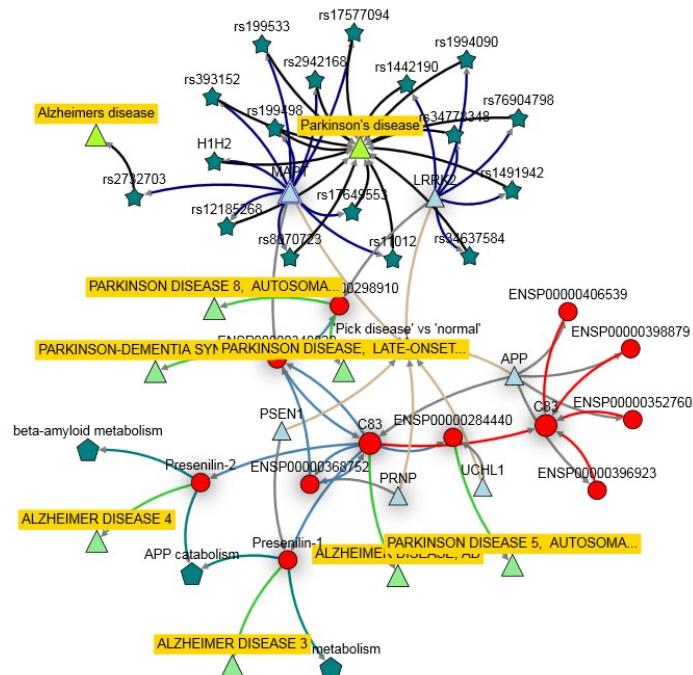
Topic 6 - Functional genomic

Topic 7 - Networks & Pathways

Topic 8 - Phenotypes & Diseases



Genotype
GWAS, QTL



Biological Knowledge Discovery

Data selection, processing, transformation, Alzheimer, Parkinson integration, interpretation



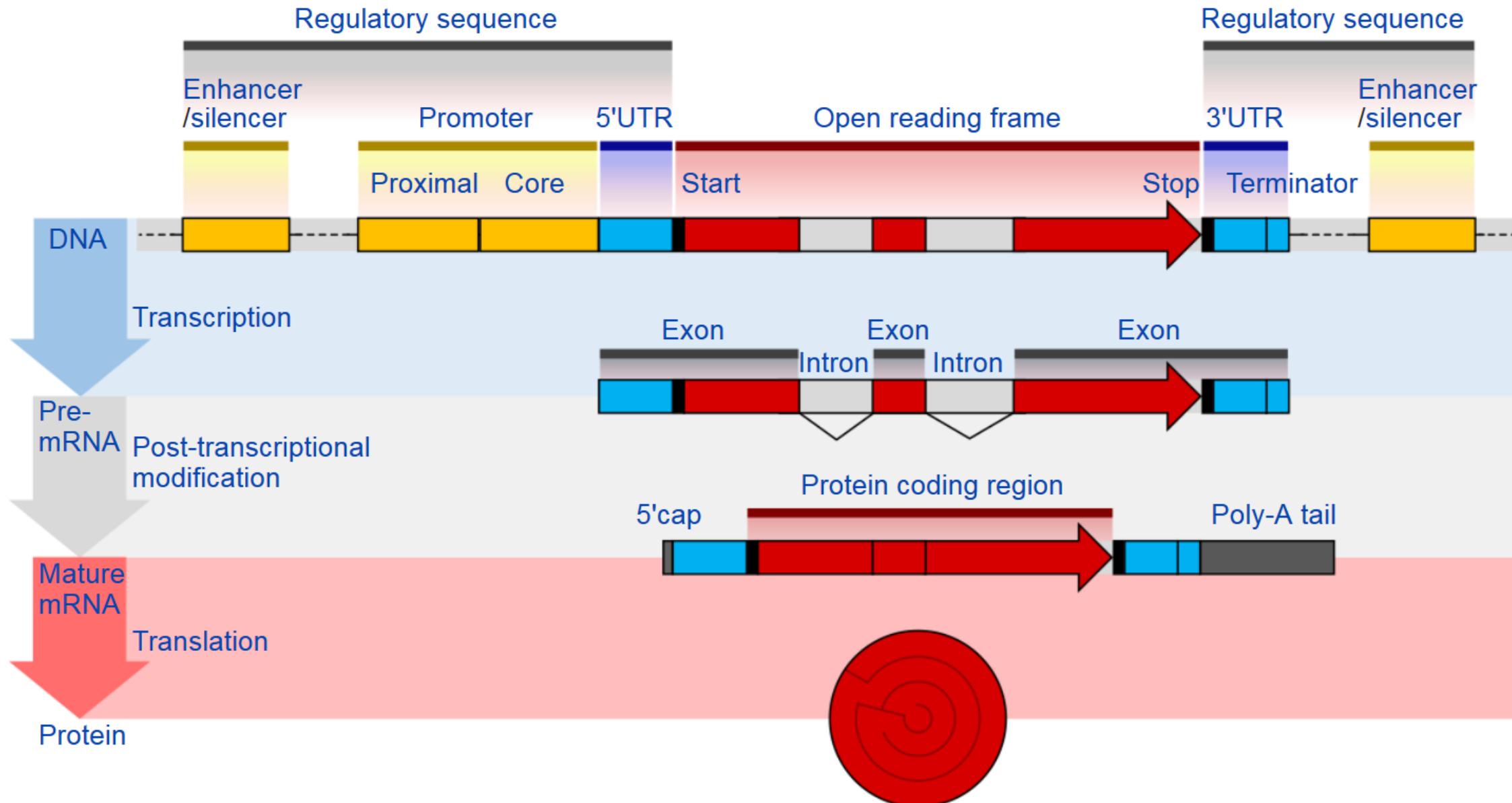
Phenotype

Precision medicine

Drugs



Genes as the carriers of protein function



NCBI Gene

NCBI Resources How To Sign in to NCBI

Gene Gene MAPT Advanced Search Help



Gene

Gene integrates information from a wide range of species. A record may include nomenclature, Reference Sequences (RefSeqs), maps, pathways, variations, phenotypes, and links to genome-, phenotype-, and locus-specific resources worldwide.

Using Gene

- [Gene Quick Start](#)
- [FAQ](#)
- [Download/FTP](#)
- [RefSeq Mailing List](#)
- [Gene News](#) 
- [Factsheet](#)

Gene Tools

- [Submit GeneRIFs](#)
- [Submit Correction](#)
- [Statistics](#)
- [BLAST](#)
- [Genome Workbench](#)
- [Splign](#)

Other Resources

- [HomoloGene](#)
- [OMIM](#)
- [RefSeq](#)
- [RefSeqGene](#)
- [UniGene](#)
- [Protein Clusters](#)

<https://www.ncbi.nlm.nih.gov/gene/>

NCBI Gene

NCBI Resources How To

Gene Gene Advanced

Full Report

MAPT microtubule associated protein tau [*Homo sapiens* (human)]

Gene ID: 4137, updated on 15-Oct-2017

Summary

Official Symbol MAPT provided by HGNC

Official Full Name microtubule associated protein tau provided by HGNC

Primary source HGNC:HGNC:6893

See related Ensembl:ENSG00000186868 MIM:157140; Vega:OTTHUMG00000168833

Gene type protein coding

RefSeq status REVIEWED

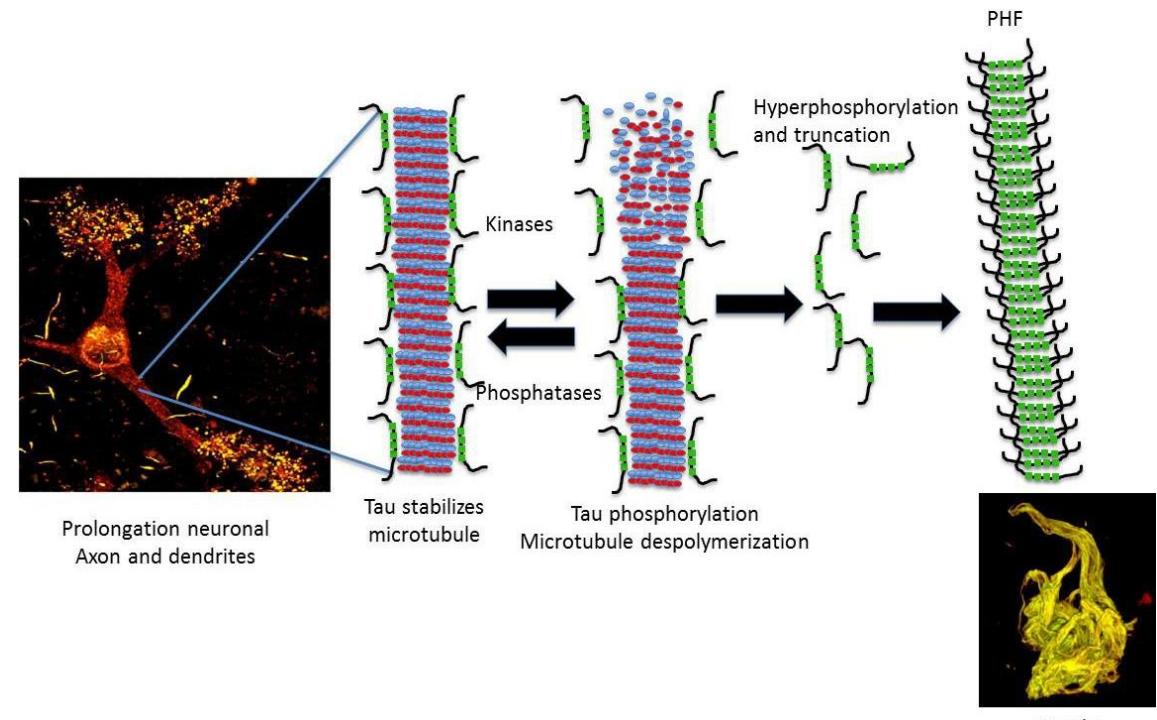
Organism *Homo sapiens*

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

Also known as TAU; MSTD; PPND; DDPAC; MAPTL; MTBT1; MTBT2; FTDP-17; PPP1R103

Summary This gene encodes the microtubule-associated protein tau (MAPT) whose transcript undergoes complex, regulated alternative splicing, giving rise to several mRNA species. MAPT transcripts are differentially expressed in the nervous system, depending on stage of neuronal maturation and neuron type. MAPT gene mutations have been associated with several neurodegenerative disorders such as Alzheimer's disease, Pick's disease, frontotemporal dementia, cortico-basal degeneration and progressive supranuclear palsy. [provided by RefSeq, Jul 2008]

Orthologs [mouse](#) [all](#)



<https://www.ncbi.nlm.nih.gov/gene/4137>

Gene Cards

GeneCardsSuite GeneCards MalaCards LifeMap Discovery PathCards TGex VarElect GeneAnalytics GeneALaCart GenesLikeMe

Free for academic non-profit institutions. Other users need a [Commercial license](#). WEIZMANN INSTITUTE OF SCIENCE LifeMap SCIENCES

GeneCards® HUMAN GENE DATABASE

Keywords ▾ Search Term Advanced

Home User Guide Analysis Tools ▾ News And Views About ▾ My Genes Log In / Sign Up

GeneCards®: The Human Gene Database

GeneCards is a searchable, integrative database that provides comprehensive, user-friendly information on all annotated and predicted human genes. It automatically integrates gene-centric data from ~125 web sources, including genomic, transcriptomic, proteomic, genetic, clinical and functional information.



Explore a Gene

MAPT



Jump to section for this gene:

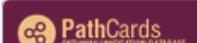
Aliases Disorders Domains Drugs Expression Function Genomics Localization Orthologs
Paralogs Pathways Products Proteins Publications Sources Summaries Transcripts Variants

GeneCardsSuite

NGS Analysis Tools



Affiliated Databases



Analysis Tools



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

Gene Cards

[GeneCardsSuite](#)[GeneCards](#)[MalaCards](#)[LifeMap Discovery](#)[PathCards](#)[TGex](#)[VarElect](#)[GeneAnalytics](#)[GeneALaCart](#)[GenesLikeMe](#)Free for academic non-profit institutions. Other users need a [Commercial license](#)WEIZMANN
INSTITUTE
OF SCIENCELifeMap
SCIENCES

Keywords ▾

Search Term

[Advanced](#)[Home](#)[User Guide](#)[Analysis Tools ▾](#)[News And Views](#)[About ▾](#)[My Genes](#)[Log In / Sign Up](#)

MAPT Gene (Protein Coding) ★

Microtubule Associated Protein Tau

GCID: GC17P045894 [?](#)GIFTs: 74 [?](#)Genes
ParticipantsUDN
Undiagnosed
Diseases NetworkJump to
section[Aliases](#) [Disorders](#) [Domains](#) [Drugs](#) [Expression](#) [Function](#) [Genomics](#) [Localization](#) [Orthologs](#)
[Paralogs](#) [Pathways](#) [Products](#) [Proteins](#) [Publications](#) [Sources](#) [Summaries](#) [Transcripts](#) [Variants](#)[Genes](#) [Peptides](#) [Proteins](#) [CRISPR](#)[Proteins](#) [Antibodies](#) [Assays](#) [Genes](#)
[shRNA](#) [Primers](#) [CRISPR](#)[Genes \(adenoviral\)](#)
[Genes \(lentiviral\)](#) [miRNA](#)
[shRNA \(AAV\)](#)

Aliases for MAPT Gene



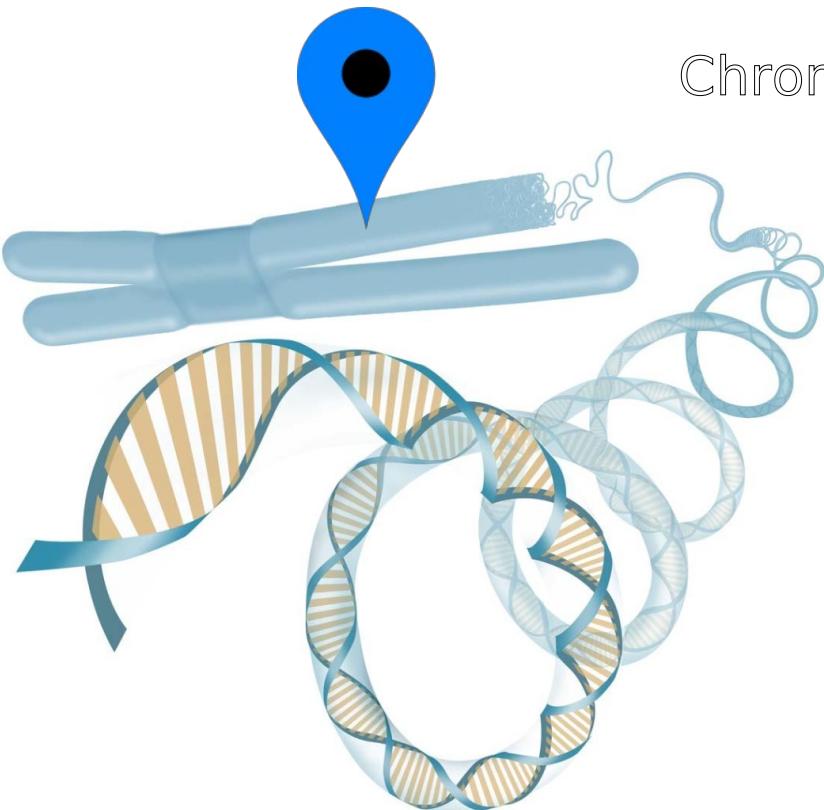
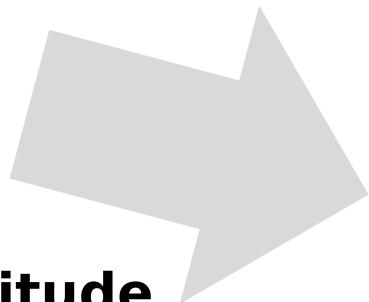
Aliases for MAPT Gene

[Microtubule Associated Protein Tau ^{2 3 5}](#)[G Protein Beta1/Gamma2 Subunit-Interacting Factor 1 ^{2 3}](#)[Protein Phosphatase 1, Regulatory Subunit 103 ^{2 3}](#)[Neurofibrillary Tangle Protein ^{3 4}](#)[Paired Helical Filament-Tau ^{3 4}](#)[PHF-Tau ^{3 4}](#)[MAPTL ^{3 4}](#)[MTBT1 ^{3 4}](#)[TAU ^{3 4}](#)[Microtubule-Associated Protein Tau, Isoform 4 ²](#)[Microtubule-Associated Protein Tau ³](#)[PPP1R103 ³](#)[FTDP-17 ³](#)[MTBT2 ³](#)[DDPAC ³](#)[MSTD ³](#)[PPND ³](#)<http://www.genecards.org/cgi-bin/carddisp.pl?gene=MAPT>

Genome browsers



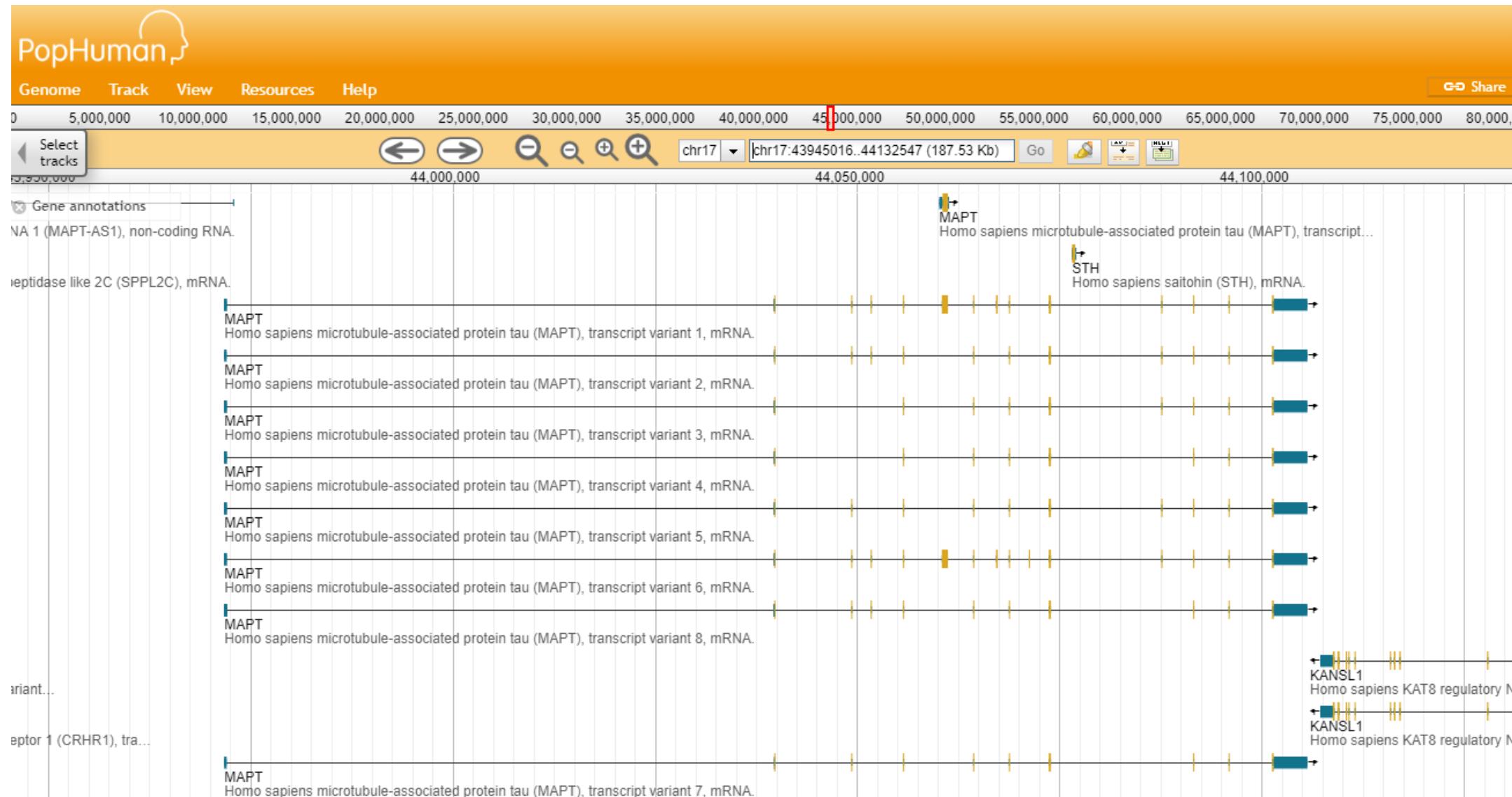
**MAP
GPS**
Latitude + Longitude
41.5011, 2.1097



**GENOME
BROWSERS**

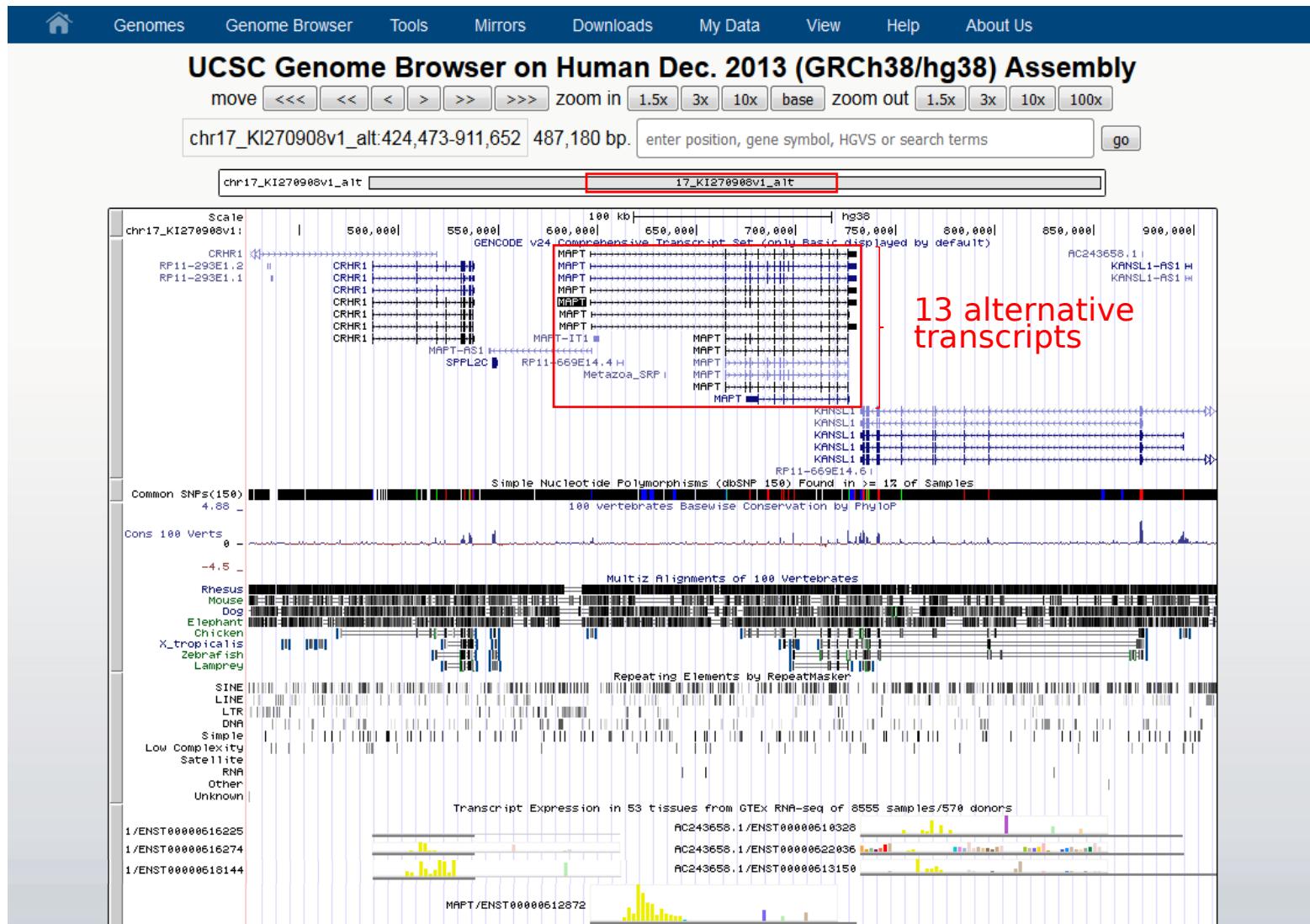
Chromosome + Position
chr3:4544933

PopHuman



<https://pophuman.uab.cat/?loc=chr17%3A43945016..44132547>

Genome browsers: UCSC



https://genome.ucsc.edu/cgi-bin/hgTracks?db=hg38&position=chr17_KI270908v1_alt%3A424473%2D911652



Tools

BioMart >

[All tools](#)

Export custom datasets from Ensembl with this data-mining tool

BLAST/BLAT >

Search our genomes for your DNA or protein sequence

Variant Effect Predictor >

Analyse your own variants and predict the functional consequences of known and unknown variants

Ensembl is a genome browser for vertebrate genomes that supports research in comparative genomics, evolution, sequence variation and transcriptional regulation. Ensembl annotate genes, computes multiple alignments, predicts regulatory function and collects disease data. Ensembl tools include BLAST, BLAT, BioMart and the Variant Effect Predictor (VEP) for all supported species.

Ensembl Release 110 (July 2023)

- The human genome assembly has been updated to the latest patch release GRCh38.p14
- Regulatory annotation annotation of open chromatin regions and promoters in pig, chicken, Atlantic salmon, turbot, and European seabass has been added (a collaboration with the GENE-SWiCH and AQUA-FAANG consortia)
- Three rat strains have been added: SHR/Utx RGD_8142385, WKY/Bbb RGD_1581635 and SHRSP/BbbUtx
- MANE and Ensembl canonical transcript attributes to be added as tags in GFF3 files

[More release news](#) on our blog

Search

All species for

MAPT ←

e.g. BRCA2 or rat 5:62797383-63627669 or rs699 or coronary heart disease

All genomes

-- Select a species --

Pig breeds
Pig reference genome and 12 additional breeds

[View full list of all species](#)

Favourite genomes

Human
GRCh38.p14
[Still using GRCh37?](#)

Mouse
GRCm39

Zebrafish
GRCz11

<https://www.ensembl.org/index.html>

Ensembl Rapid Release

New assemblies with gene and protein annotation every two weeks.

Note: species that already exist on this site will continue to be updated with the full range of annotations.

[Go](#)

The Ensembl Rapid Release website provides annotation for recently produced, publicly available vertebrate and non-vertebrate genomes from biodiversity initiatives such as Darwin Tree of Life, the Vertebrate Genomes Project and the Earth BioGenome Project.

[Rapid Release news](#) on our blog

Ensembl



BLAST/BLAT | VEP | Tools | BioMart | Downloads | Help & Docs | Blog

Login/Register

Search Human...



Human (GRCh38.p14) ▾

Location: 17:45,894,382-46,028,334

Gene: MAPT

Gene-based displays

Summary

- Splice variants
- Transcript comparison
- Gene alleles

Sequence

- Secondary Structure

Comparative Genomics

- Genomic alignments
- Gene tree
- Gene gain/loss tree
- Orthologues
- Paralogues

Ontologies

- GO: Molecular function
- GO: Biological process
- GO: Cellular component

Phenotypes

Genetic Variation

- Variant table
- Variant image
- Structural variants
- Gene expression
- Pathway
- Molecular interactions
- Regulation
- External references
- Supporting evidence

ID History

- Gene history

Configure this page

Custom tracks

Export data

Share this page

Gene: MAPT ENSG00000186868

Description

microtubule associated protein tau [Source:HGNC Symbol;Acc:[HGNC:6893](#)]

Gene Synonyms

DDPAC, FLJ31424, FTDP-17, MAPTL, MGC138549, MSTD, MTBT1, MTBT2, PPND, PPP1R103, TAU, TAU-40, TAU-PHF6

Location

[Chromosome 17: 45,894,527-46,028,334](#) forward strand.

GRCh38:CM000679.2

[View alleles](#) of this gene on alternative sequences

About this gene

This gene has 30 transcripts ([splice variants](#)), 1 gene allele, [226 orthologues](#), [1 parologue](#) and is associated with [13 phenotypes](#).

Transcripts

[Hide transcript table](#)

Show/hide columns (1 hidden)

Filter

Transcript ID	Name	bp	Protein	Biotype	CCDS	UniProt Match	RefSeq Match	Flags
ENST00000262410.10	MAPT-201	6815	833aa	Protein coding	CCDS92345	A0A7I2PJZ2	NM_001377265.1	MANE Select Ensembl Canonical GENCODE basic APPRIS ALT2 TSL:1
ENST00000351559.10	MAPT-204	5639	441aa	Protein coding	CCDS11499	P10636-8	-	GENCODE basic APPRIS ALT2 TSL:1
ENST00000446361.7	MAPT-208	5345	383aa	Protein coding	CCDS11500	P10636-6	-	GENCODE basic APPRIS ALT2 TSL:1
ENST00000344290.10	MAPT-203	2609	736aa	Protein coding	CCDS92346	A0A7I2PLE3	-	GENCODE basic TSL:1
ENST00000415613.6	MAPT-205	2331	776aa	Protein coding	CCDS45715	P10636-9	-	GENCODE basic APPRIS ALT2 TSL:5
ENST00000571987.5	MAPT-212	2277	758aa	Protein coding	CCDS11501	P10636-1	-	GENCODE basic APPRIS ALT2 TSL:5
ENST00000680674.1	MAPT-218	1936	424aa	Protein coding		A0A7P0T936	-	GENCODE basic
ENST00000420682.7	MAPT-206	1871	412aa	Protein coding	CCDS45716	P10636-7	-	GENCODE basic APPRIS P3 TSL:1
ENST00000535772.6	MAPT-209	1353	381aa	Protein coding	CCDS92347	P10636-4	-	GENCODE basic TSL:5
ENST00000574436.5	MAPT-214	1326	441aa	Protein coding	CCDS11499	P10636-8	-	GENCODE basic APPRIS ALT2 TSL:5
ENST00000431008.7	MAPT-207	1233	410aa	Protein coding	CCDS56033	P10636-5	-	GENCODE basic TSL:5
ENST00000334239.12	MAPT-202	1107	352aa	Protein coding	CCDS11502	P10636-2	-	GENCODE basic TSL:1
ENST00000680542.1	MAPT-217	2988	412aa	Nonsense mediated decay	CCDS45716	P10636-7	-	-
ENST00000703924.1	MAPT-221	2971	412aa	Nonsense mediated decay	CCDS45716	-	-	-

https://www.ensembl.org/Homo_sapiens/Gene/Summary?db=core;g=ENSG00000186868;r=17:45894382-46028334



Gene-based displays

Summary

- Splice variants
- Transcript comparison
- Gene alleles
- Sequences

Secondary Structure

Comparative Genomics

- Genomic alignments
- Gene tree
- Gene gain/loss tree
- Orthologues
- Paralogues

Ontologies

- GO: Molecular function
- GO: Biological process
- GO: Cellular component

Phenotypes

Genetic Variation

- Variant table
- Variant image
- Structural variants
- Gene expression
- Pathway
- Molecular interactions
- Regulation
- External references
- Supporting evidence

ID History

- Gene history

Configure this page

Custom tracks

Export data

Share this page

Gene: MAPT ENSG00000186868
Summary & Structure

Description

microtubule associated protein tau [Source:HGNC Symbol;Acc:HGNC:6893]

Gene Synonyms

DDPAC, FLJ31424, FTDP-17, MAPTL, MGC138549, MSTD, MTBT1, MTBT2, PPND, PPP1R103, TAU, TAU-40, TAU-PHF6

Location

Chromosome 17: 45,894,527-46,028,334 forward strand.

GRCh38:CM000679.2

View [alleles](#) of this gene on alternative sequences

About this gene

This gene has 30 transcripts ([splice variants](#)), 1 gene allele, 226 orthologues, 1 parologue and is associated with 13 phenotypes.

Transcripts

[Hide transcript table](#)

Transcript table

Show/hide columns (1 hidden)

Filter



Transcript ID	Name	bp	Protein	Biotype	CCDS	UniProt Match	RefSeq Match	Flags
ENST00000262410.10	MAPT-201	6815	833aa	Protein coding	CCDS92345	A0A7I2PJZ2	NM_001377265.1	MANE Select Ensembl Canonical GENCODE basic APPRIS ALT2 TSL
ENST00000351559.10	MAPT-204	5639	441aa	Protein coding	CCDS11499	P10636-8	-	GENCODE basic APPRIS ALT2 TSL:1
ENST00000446361.7	MAPT-208	5345	383aa	Protein coding	CCDS11500	P10636-6	-	GENCODE basic APPRIS ALT2 TSL:1
ENST00000344290.10	MAPT-203	2609	736aa	Protein coding	CCDS92346	A0A7I2PLE3	-	GENCODE basic TSL:1
ENST00000415613.6	MAPT-205	2331	776aa	Protein coding	CCDS45715	P10636-9	-	GENCODE basic APPRIS ALT2 TSL:5
ENST00000571987.5	MAPT-212	2277	758aa	Protein coding	CCDS11501	P10636-1	-	GENCODE basic APPRIS ALT2 TSL:5
ENST00000680674.1	MAPT-218	1936	424aa	Protein coding		A0A7P0T936	-	GENCODE basic
ENST00000420682.7	MAPT-206	1871	412aa	Protein coding	CCDS45716	P10636-7	-	GENCODE basic APPRIS P3 TSL:1
ENST00000535772.6	MAPT-209	1353	381aa	Protein coding	CCDS92347	P10636-4	-	GENCODE basic TSL:5
ENST00000574436.5	MAPT-214	1326	441aa	Protein coding	CCDS11499	P10636-8	-	GENCODE basic APPRIS ALT2 TSL:5
ENST00000431008.7	MAPT-207	1233	410aa	Protein coding	CCDS56033	P10636-5	-	GENCODE basic TSL:5
ENST00000334239.12	MAPT-202	1107	352aa	Protein coding	CCDS11502	P10636-2	-	GENCODE basic TSL:1
ENST00000680542.1	MAPT-217	2988	412aa	Nonsense mediated decay	CCDS45716	P10636-7	-	-
ENST00000703924.1	MAPT-221	2971	412aa	Nonsense mediated decay	CCDS45716	-	-	-

https://www.ensembl.org/Homo_sapiens/Gene/Summary?db=core;g=ENSG00000186868;r=17:45894382-46028334



Gene-based displays

Summary

- Splice variants
- Transcript comparison
- Gene alleles
- Sequences

Secondary Structure

Comparative Genomics

Genomic alignments

Gene tree

Gene gain/loss tree

Orthologues

Paralogues

Ontologies

GO: Molecular function

GO: Biological process

GO: Cellular component

Phenotypes

Genetic Variation

Variant table

Variant image

Structural variants

Gene expression

Pathway

Molecular interactions

Regulation

External references

Supporting evidence

ID History

Gene history

Configure this page

Custom tracks

Export data

Share this page

Gene: MAPT ENSG00000186868
Summary & Structure

Description

microtubule associated protein tau [Source:HGNC Symbol;Acc:HGNC:6893]

Gene Synonyms

DDPAC, FLJ31424, FTDP-17, MAPTL, MGC138549, MSTD, MTBT1, MTBT2, PPND, PPP1R103, TAU, TAU-40, TAU-PHF6

Location

Chromosome 17: 45,894,527-46,028,334 forward strand.

GRCh38:CM000679.2

View alleles of this gene on alternative sequences

Comparative genomics

About this gene

This gene has 30 transcripts (splice variants), 1 gene allele, 226 orthologues, 1 parologue and is associated with 13 phenotypes.

Transcripts

Hide transcript table

Transcript table

Show/hide columns (1 hidden)

Filter

Transcript ID	Name	bp	Protein	Biotype	CCDS	UniProt Match	RefSeq Match	Flags
ENST00000262410.10	MAPT-201	6815	833aa	Protein coding	CCDS92345	A0A7I2PJZ2	NM_001377265.1	MANE Select Ensembl Canonical GENCODE basic APPRIS ALT2 TSL
ENST00000351559.10	MAPT-204	5639	441aa	Protein coding	CCDS11499	P10636-8	-	GENCODE basic APPRIS ALT2 TSL:1
ENST00000446361.7	MAPT-208	5345	383aa	Protein coding	CCDS11500	P10636-6	-	GENCODE basic APPRIS ALT2 TSL:1
ENST00000344290.10	MAPT-203	2609	736aa	Protein coding	CCDS92346	A0A7I2PLE3	-	GENCODE basic TSL:1
ENST00000415613.6	MAPT-205	2331	776aa	Protein coding	CCDS45715	P10636-9	-	GENCODE basic APPRIS ALT2 TSL:5
ENST00000571987.5	MAPT-212	2277	758aa	Protein coding	CCDS11501	P10636-1	-	GENCODE basic APPRIS ALT2 TSL:5
ENST00000680674.1	MAPT-218	1936	424aa	Protein coding		A0A7P0T936	-	GENCODE basic
ENST00000420682.7	MAPT-206	1871	412aa	Protein coding	CCDS45716	P10636-7	-	GENCODE basic APPRIS P3 TSL:1
ENST00000535772.6	MAPT-209	1353	381aa	Protein coding	CCDS92347	P10636-4	-	GENCODE basic TSL:5
ENST00000574436.5	MAPT-214	1326	441aa	Protein coding	CCDS11499	P10636-8	-	GENCODE basic APPRIS ALT2 TSL:5
ENST00000431008.7	MAPT-207	1233	410aa	Protein coding	CCDS56033	P10636-5	-	GENCODE basic TSL:5
ENST00000334239.12	MAPT-202	1107	352aa	Protein coding	CCDS11502	P10636-2	-	GENCODE basic TSL:1
ENST00000680542.1	MAPT-217	2988	412aa	Nonsense mediated decay	CCDS45716	P10636-7	-	
ENST00000703924.1	MAPT-221	2971	412aa	Nonsense mediated decay	CCDS45716	-	-	

https://www.ensembl.org/Homo_sapiens/Gene/Summary?db=core;g=ENSG00000186868;r=17:45894382-46028334



Gene-based displays

- Summary
 - Splice variants
 - Transcript comparison
 - Gene alleles
 - Sequences
- Secondary Structure

- Comparative Genomics
 - Genomic alignments
 - Gene tree
 - Gene gain/loss tree
 - Orthologues
 - Paralogues

- Ontologies
 - GO: Molecular function
 - GO: Biological process
 - GO: Cellular component

- Phenotypes
- Genetic Variation
 - Variant table
 - Variant image
 - Structural variants
 - Gene expression
 - Pathway
 - Molecular interactions
 - Regulation
 - External references
 - Supporting evidence
- ID History
 - Gene history

[Configure this page](#)[Custom tracks](#)[Export data](#)[Share this page](#)

Gene: MAPT ENSG00000186868

Summary & Structure

Description

microtubule associated protein tau [Source:HGNC Symbol;Acc:HGNC:6893]

Gene Synonyms

DDPAC, FLJ31424, FTDP-17, MAPTL, MGC138549, MSTD, MTBT1, MTBT2, PPND, PPP1R103, TAU, TAU-40, TAU-PHF6

Location

Chromosome 17: 45,894,527-46,028,334 forward strand.

GRCh38:CM000679.2

View [alleles](#) of this gene on alternative sequences

Comparative genomics

[About this gene](#)

Transcripts

[Hide transcript table](#)

Ontologies & Phenotype

[Show/hide columns \(1 hidden\)](#)[Filter](#)

Transcript table

Transcript ID	Name	bp	Protein	Biotype	CCDS	UniProt Match	RefSeq Match	Flags
ENST00000262410.10	MAPT-201	6815	833aa	Protein coding	CCDS92345	A0A7I2PJZ2	NM_001377265.1	MANE Select Ensembl Canonical GENCODE basic APPRIS ALT2 TSL
ENST00000351559.10	MAPT-204	5639	441aa	Protein coding	CCDS11499	P10636-8	-	GENCODE basic APPRIS ALT2 TSL:1
ENST00000446361.7	MAPT-208	5345	383aa	Protein coding	CCDS11500	P10636-6	-	GENCODE basic APPRIS ALT2 TSL:1
ENST00000344290.10	MAPT-203	2609	736aa	Protein coding	CCDS92346	A0A7I2PLE3	-	GENCODE basic TSL:1
ENST00000415613.6	MAPT-205	2331	776aa	Protein coding	CCDS45715	P10636-9	-	GENCODE basic APPRIS ALT2 TSL:5
ENST00000571987.5	MAPT-212	2277	758aa	Protein coding	CCDS11501	P10636-1	-	GENCODE basic APPRIS ALT2 TSL:5
ENST00000680674.1	MAPT-218	1936	424aa	Protein coding		A0A7P0T936	-	GENCODE basic
ENST00000420682.7	MAPT-206	1871	412aa	Protein coding	CCDS45716	P10636-7	-	GENCODE basic APPRIS P3 TSL:1
ENST00000535772.6	MAPT-209	1353	381aa	Protein coding	CCDS92347	P10636-4	-	GENCODE basic TSL:5
ENST00000574436.5	MAPT-214	1326	441aa	Protein coding	CCDS11499	P10636-8	-	GENCODE basic APPRIS ALT2 TSL:5
ENST00000431008.7	MAPT-207	1233	410aa	Protein coding	CCDS56033	P10636-5	-	GENCODE basic TSL:5
ENST00000334239.12	MAPT-202	1107	352aa	Protein coding	CCDS11502	P10636-2	-	GENCODE basic TSL:1
ENST00000680542.1	MAPT-217	2988	412aa	Nonsense mediated decay	CCDS45716	P10636-7	-	-
ENST00000703924.1	MAPT-221	2971	412aa	Nonsense mediated decay	CCDS45716	-	-	-

https://www.ensembl.org/Homo_sapiens/Gene/Summary?db=core;g=ENSG00000186868;r=17:45894382-46028334



Location: 17:45,894,382-46,028,334

Gene: MAPT

Gene-based displays

- Summary
 - Splice variants
 - Transcript comparison
 - Gene alleles
 - Sequences
- Secondary Structure

- Comparative Genomics
 - Genomic alignments
 - Gene tree
 - Gene gain/loss tree
 - Orthologues
 - Paralogues

- Ontologies
 - GO: Molecular function
 - GO: Biological process
 - GO: Cellular component

- Phenotypes
- Genetic Variation
 - Variant table
 - Variant image
 - Structure viewer

- Gene expression
- Pathway
- Molecular interactions
- Regulation

- External references
- Supporting evidence
- ID History
 - Gene history

 Configure this page Custom tracks Export data Share this pageGene: MAPT ENSG00000186868
Summary & Structure

Description

microtubule associated protein tau [Source:HGNC Symbol;Acc:HGNC:6893]

Gene Synonyms

DDPAC, FLJ31424, FTDP-17, MAPTL, MGC138549, MSTD, MTBT1, MTBT2, PPND, PPP1R103, TAU, TAU-40, TAU-PHF6

Location

Chromosome 17: 45,894,527-46,028,334 forward strand.

GRCh38:CM000679.2

View [alleles](#) of this gene on alternative sequences

Comparative genomics

About this gene

Transcripts

[Hide transcript table](#)

Transcript table

Ontologies & Phenotype

[Show/hide columns \(1 hidden\)](#)

Filter

Genetic variation

Expression & Regulation

Transcript ID	Name	bp	Protein	Biotype	CCDS	UniProt Match	RefSeq Match	Flags
ENST00000262410.10	MAPT-201	6815	833aa	Protein coding	CCDS92345	A0A7I2PJZ2	NM_001377265.1	MANE Select Ensembl Canonical GENCODE basic APPRIS ALT2 TSL:1
ENST0000031559.10	MAPT-204	5639	441aa	Protein coding	CCDS11499	P10636-8	-	GENCODE basic APPRIS ALT2 TSL:1
ENST00000446361.7	MAPT-208	5345	383aa	Protein coding	CCDS11500	P10636-6	-	GENCODE basic APPRIS ALT2 TSL:1
ENST00000446361.8	MAPT-209	5345	386aa	Protein coding	CCDS92346	A0A7I2PLE3	-	GENCODE basic TSL:1
ENST00000415613.6	MAPT-205	2331	776aa	Protein coding	CCDS45715	P10636-9	-	GENCODE basic APPRIS ALT2 TSL:5
ENST00000571987.5	MAPT-212	2277	758aa	Protein coding	CCDS11501	P10636-1	-	GENCODE basic APPRIS ALT2 TSL:5
ENST00000680674.1	MAPT-218	1936	424aa	Protein coding		A0A7P0T936	-	GENCODE basic
ENST00000420682.7	MAPT-206	1871	412aa	Protein coding	CCDS45716	P10636-7	-	GENCODE basic APPRIS P3 TSL:1
ENST00000535772.6	MAPT-209	1353	381aa	Protein coding	CCDS92347	P10636-4	-	GENCODE basic TSL:5
ENST00000574436.5	MAPT-214	1326	441aa	Protein coding	CCDS11499	P10636-8	-	GENCODE basic APPRIS ALT2 TSL:5
ENST00000431008.7	MAPT-207	1233	410aa	Protein coding	CCDS56033	P10636-5	-	GENCODE basic TSL:5
ENST00000334239.12	MAPT-202	1107	352aa	Protein coding	CCDS11502	P10636-2	-	GENCODE basic TSL:1
ENST00000680542.1	MAPT-217	2988	412aa	Nonsense mediated decay	CCDS45716	P10636-7	-	-
ENST00000703924.1	MAPT-221	2971	412aa	Nonsense mediated decay	CCDS45716	-	-	-

https://www.ensembl.org/Homo_sapiens/Gene/Summary?db=core;g=ENSG00000186868;r=17:45894382-46028334

Genome browsers: Ensembl



BLAST/BLAT | VEP | Tools | BioMart | Downloads | Help & Docs | Blog

Login/Register

Search Human...



Human (GRCh38.p14) ▾

Location: 17:45,894,382-46,028,334

Gene: MAPT

Location-based displays

- Whole genome
- Chromosome summary
- Region overview
- Region in detail

- Comparative Genomics
 - Synteny
 - Alignments (image)
 - Alignments (text)
 - Region Comparison

- Genetic Variation

- Variant table
- Resequencing
- Linkage Data

- Markers

- Other genome browsers
 - UCSC
 - NCBI
 - Ensembl GRCh37

Configure this page

Custom tracks

Export data

Share this page

Bookmark this page

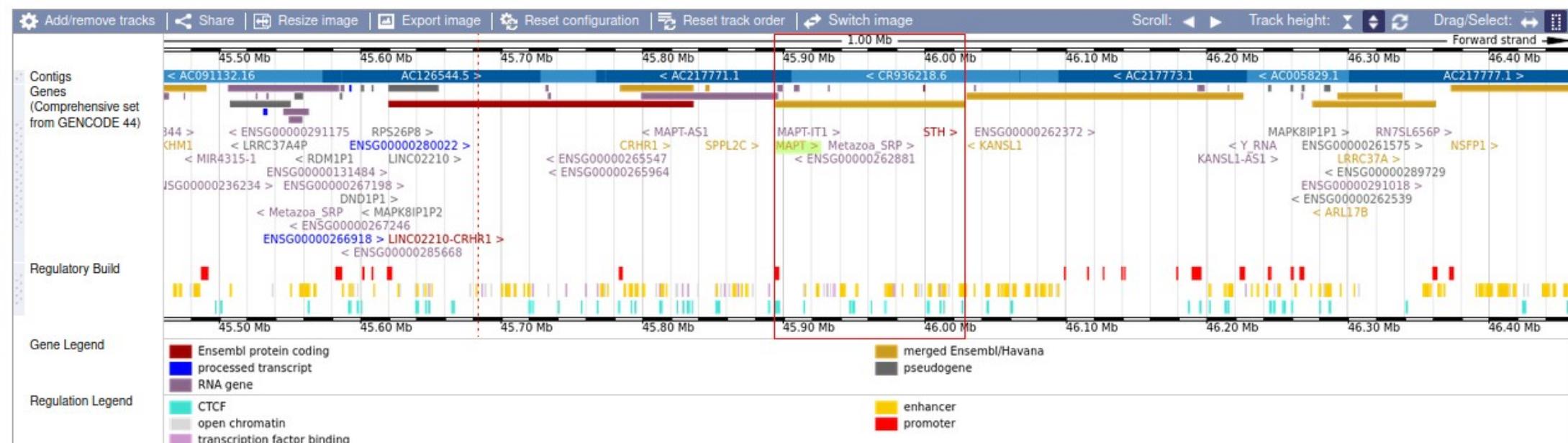
Chromosome 17: 45,894,382-46,028,334

Add/remove tracks | Share | Export image | Reset configuration |

Chr. 17

p13.3 p13.2 p13.1 p12 p11.2 p11.1 p11.2 q11.2 q12 q21.2 q21.3 q21.32 q21.33 q22 q23.2 q24.2 q24.3 q25.1 q25.3

Region in detail

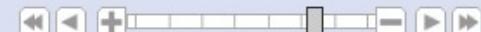


Location: 17:45894382-46028334

Go

Gene:

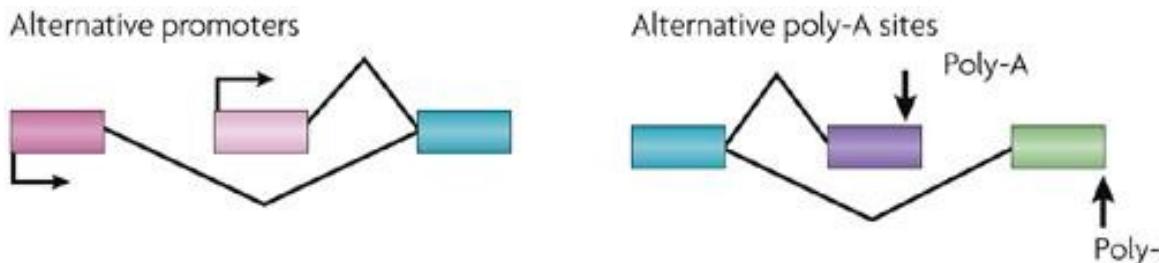
Go



https://www.ensembl.org/Homo_sapiens/Location/View?db=core;g=ENSG00000186868;r=17:45894382-46028334

Alternative splicing

Initial/Final exons



Internal exons

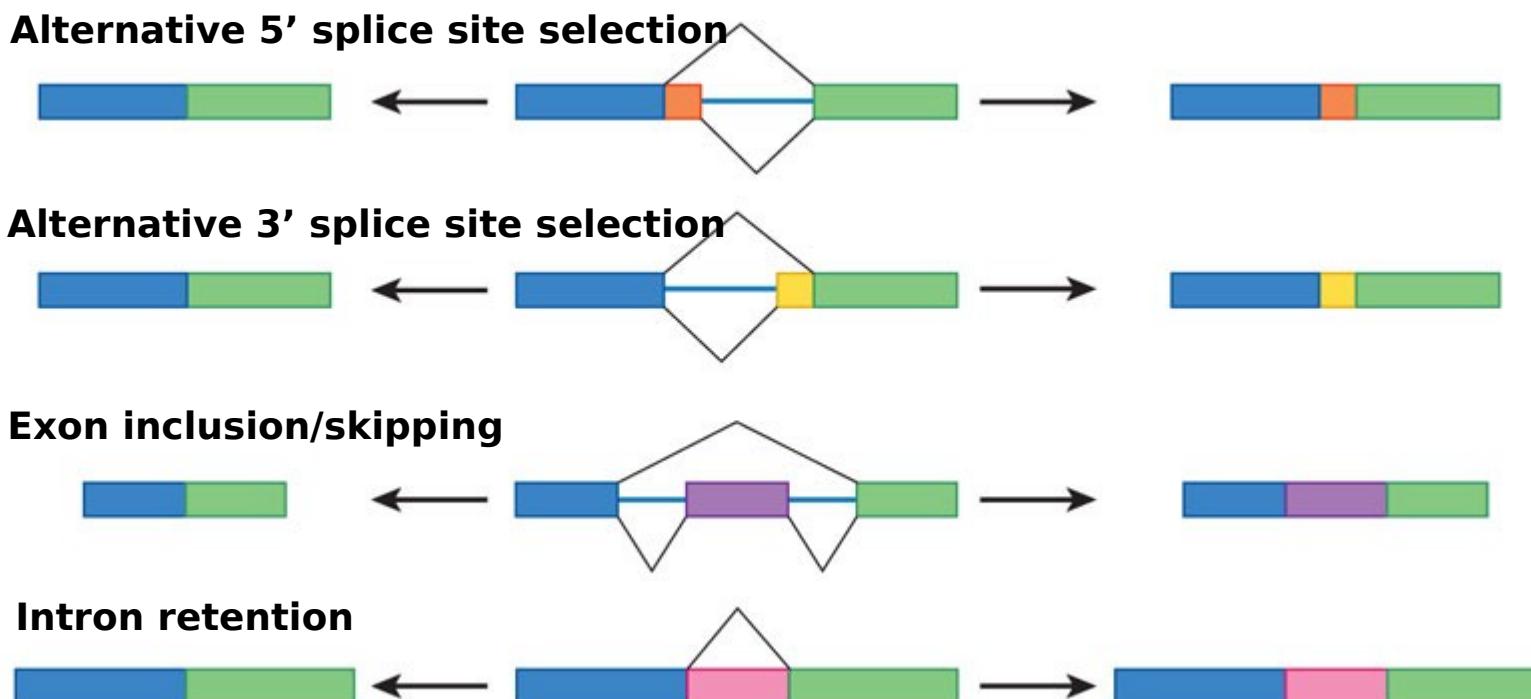


Figure 1. Nielsen and Graveley (2010) *Nature* 463: 457-463

Figure 1. Li et al. (2007) *Nature Reviews Neuroscience* 8: 819-831.

Example: alpha-tropomiosin

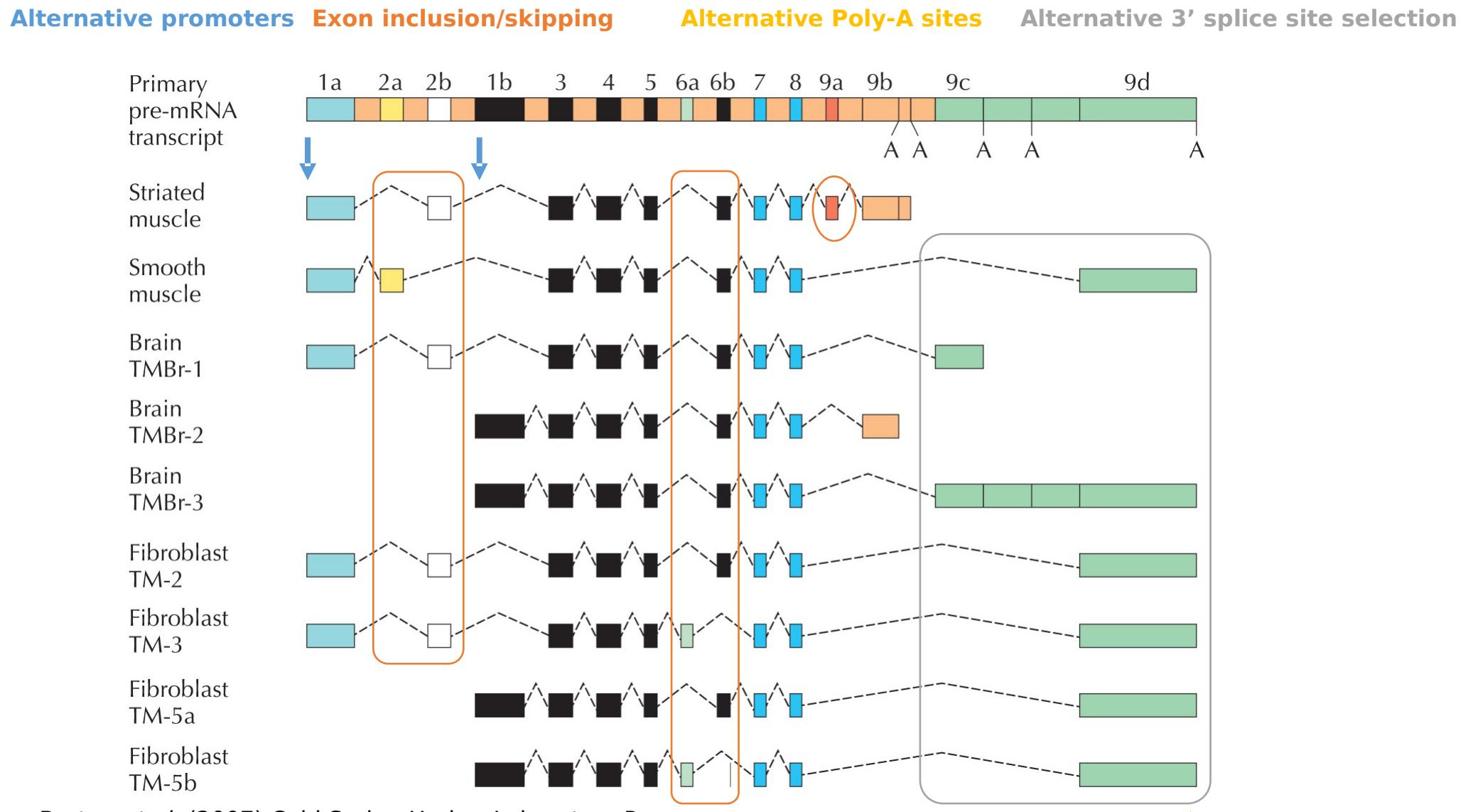


Figure 8.22. Evolution. Barton et al. (2007) Cold Spring Harbor Laboratory Press



Other examples of alternative splicing

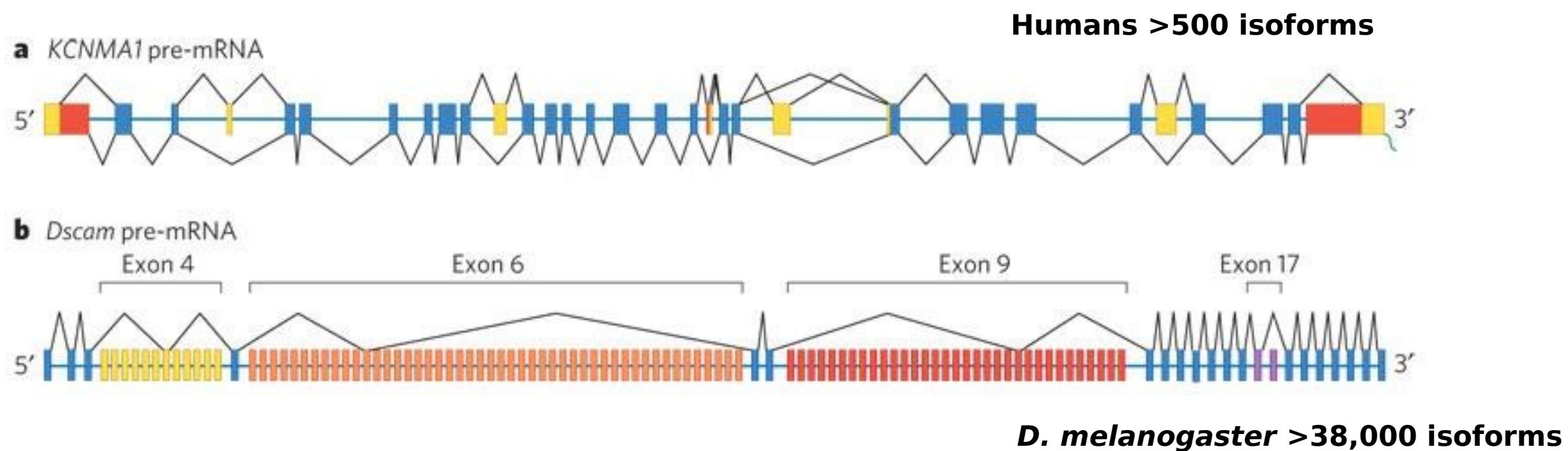
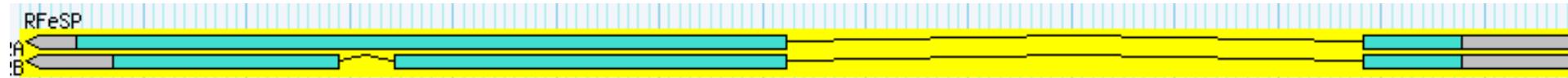


Figure 2. Nielsen and Graveley (2010) *Nature* 463: 457-463

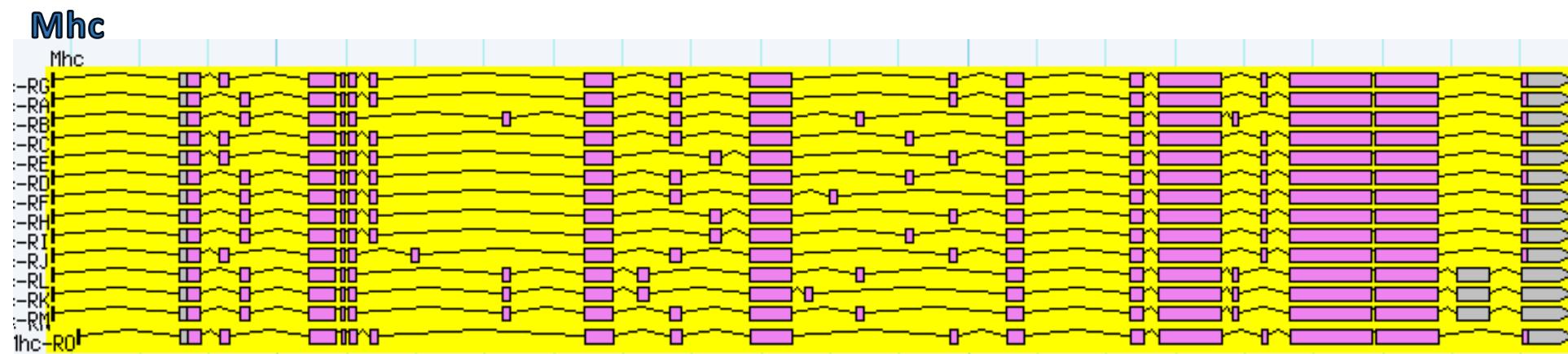
Quiz



RFeSP



Quiz



Quiz



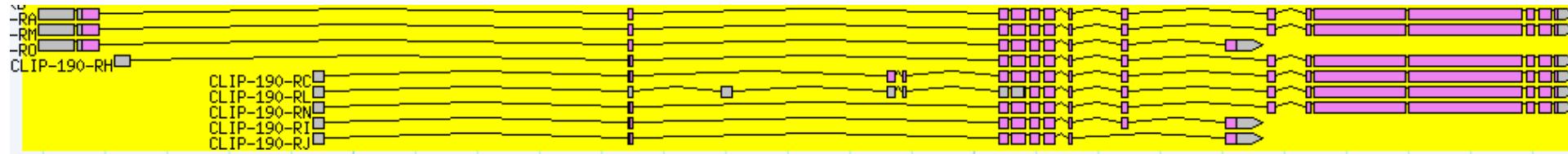
mdy



Quiz



CLIP-190



Population genomics: catalogs of genetic variation in humans



Phase III

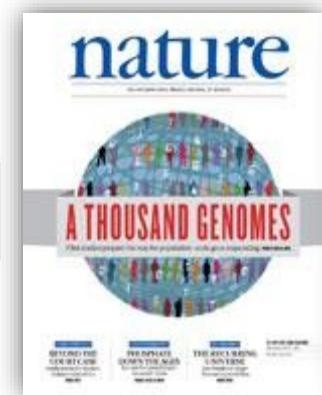
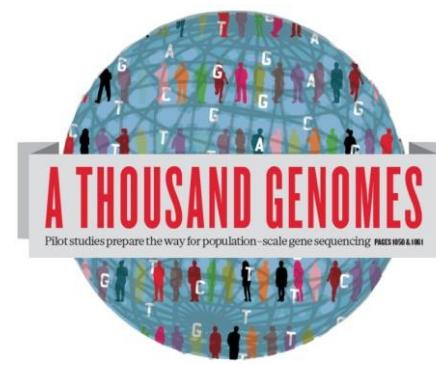
Nature 2010

1,184 individuals
11 populations
>4 million SNPs

"create a resource that would accelerate the identification of genetic factors that influence medical traits"

The International HapMap Cons. (2010) *Nature* 467: 52-58.

"provide a deep characterization of human genome sequence variation as a foundation for investigating the role of genetic variation in human history, evolution and disease."



Pilot phase

Nature 2010

179 individuals
4 populations
15 million SNPs
1 million small indels
20,000 SVs

Phase I

Nature 2012

1,092 individuals
14 populations
38 million SNPs
1.4 million small indels
14,000 deletions

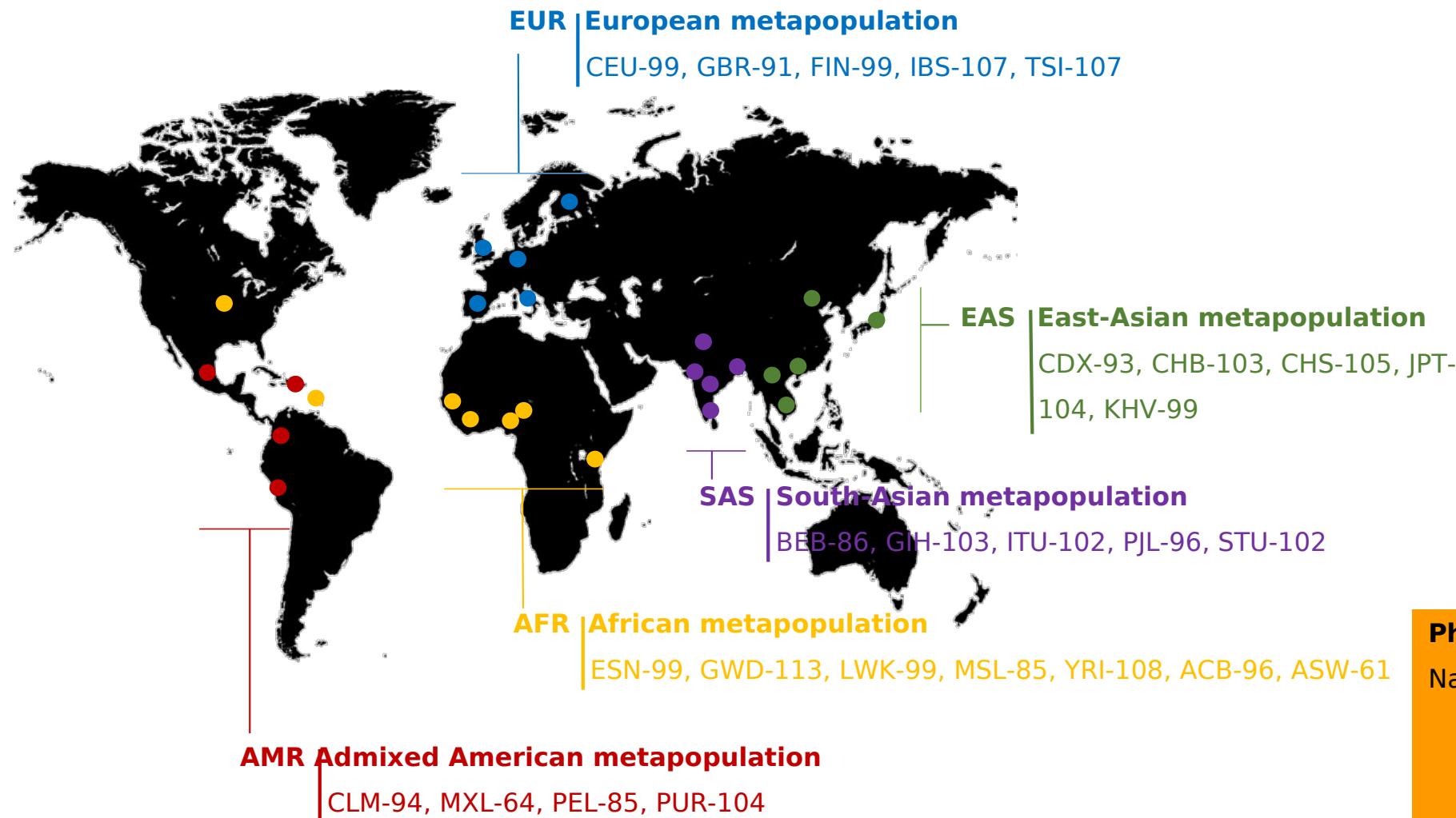
Phase III

Nature 2015

2,504 individuals
26 populations
84.7 million SNPs
3.6 million small indels
60,000 SVs

The 1000 Genomes Project Cons. (2015) *Nature* 526: 68-74.

Population genomics: the 1000 Genomes Project



Phase III	2,504 individuals
	26 populations
	84.7 million SNPs
	3.6 million small indels
	60,000 SVs

The 1000 Genomes Project Cons. (2015) *Nature* 526: 68-74.

Population genomics: Ensembl

Genetic Variation

- Variant table
- Variant image
- Structural variants
- Gene expression
- Pathway
- Molecular interactions
- Regulation
- External references
- Supporting evidence

ID History

- Gene history

Configure this page

Custom tracks

Export data

Share this page

Bookmark this page

There are 35,291 variants for this Gene, which is too many to display in this page, so **only exonic variants** are displayed. Please use [BioMart](#) to extract all data.

Filter Global MAF: 0.05 - 0.5 SIFT: All PolyPhen: All Consequences: All Filter Other Columns Show/hide columns Search... Variant ID Chr: bp Alleles Global MAF Class Source Evidence Clin. Sig. Conseq. Type AA AA co-ord SIFT Poly-Phen CADD REVEL Metal R Mutation Assessor Transcript rs62056779 17:45894571 C/A 0.086 (A) SNP dbSNP + 5 prime UTR variant - - - - - - - - - - ENST0000026241 0.10 rs11575896 17:45894776 G/A 0.086 (A) SNP dbSNP - intron variant - - - - - - - - - - ENST0000026241 0.10 rs3744456 17:45894810 C/G 0.110 (G) SNP dbSNP + intron variant - - - - - - - - - - ENST0000026241 0.10

Population genomics: Ensembl

There are 35,291 variants for this Gene, which is too many to display in this page, so **only exonic variants** are displayed. Please use [BioMart](#) to extract all data.

Filter Global MAF: 0.05 - 0.5 SIFT: All PolyPhen: All Consequences: All Filter Other Columns

Show/hide columns Search...

Variant ID	Chr: bp	Alleles	Global MAF	Class	Source	Evidence	Clin. Sig.	Conseq. Type	AA	AA co-ord	SIFT	Poly-Phen	CADD	REVEL	Metal R	Mutati on Asses sor	Transcript	
rs62056779	17:45894571	C/A	0.086 (A)	SNP	dbSNP	      		5 prime UTR variant	-	-	-	-	-	-	-	-	-	ENST0000026241 0.10
rs11575896	17:45894776	G/A	0.086 (A)	SNP	dbSNP	      		intron variant	-	-	-	-	-	-	-	-	ENST0000026241 0.10	
rs3744456	17:45894810	C/G	0.110 (G)	SNP	dbSNP	      		intron variant	-	-	-	-	-	-	-	-	ENST0000026241 0.10	

Most severe consequence

Alleles

Change tolerance

Location

Evidence status 

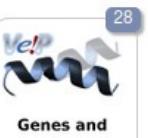
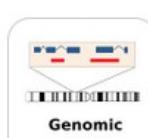
HGVs names

Synonyms

Original source

About this variant

Explore this variant 



ATT CATT
CGG **S**TG
TCAT GCT

Flanking sequence



Population genetics



2504

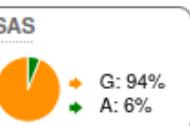
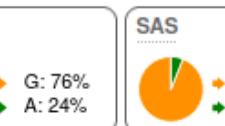
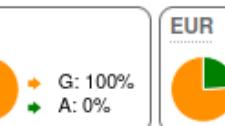
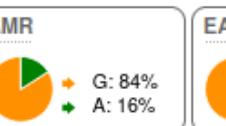
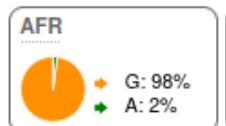
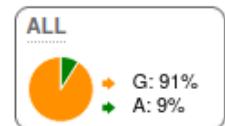
T	C		
T	T	T	C
C	T	C	C

Linkage disequilibrium

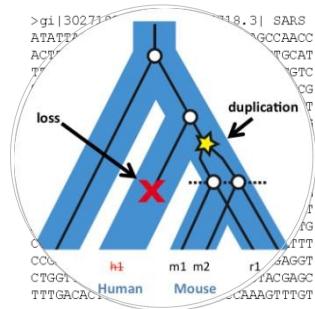
Phylogenetic context

Population genetics

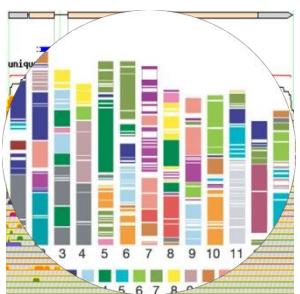
1000 Genomes Project Phase 3 allele frequencies



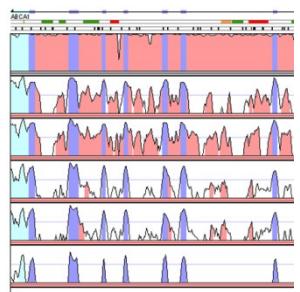
Comparative genomics



Orthologues, Paralogues & Genetrees

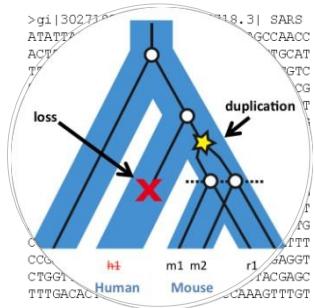


Genomic alignments & Syntenic blocks

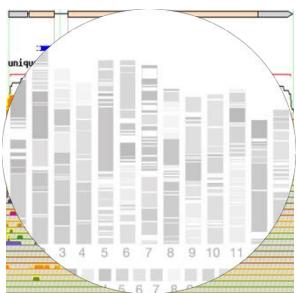


Sequence conservation & Functional constraints

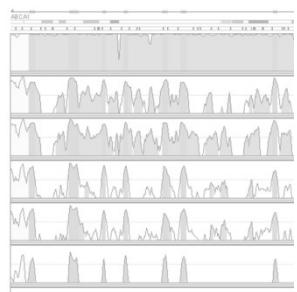
Comparative genomics



Orthologues, Paralogues & Genetrees

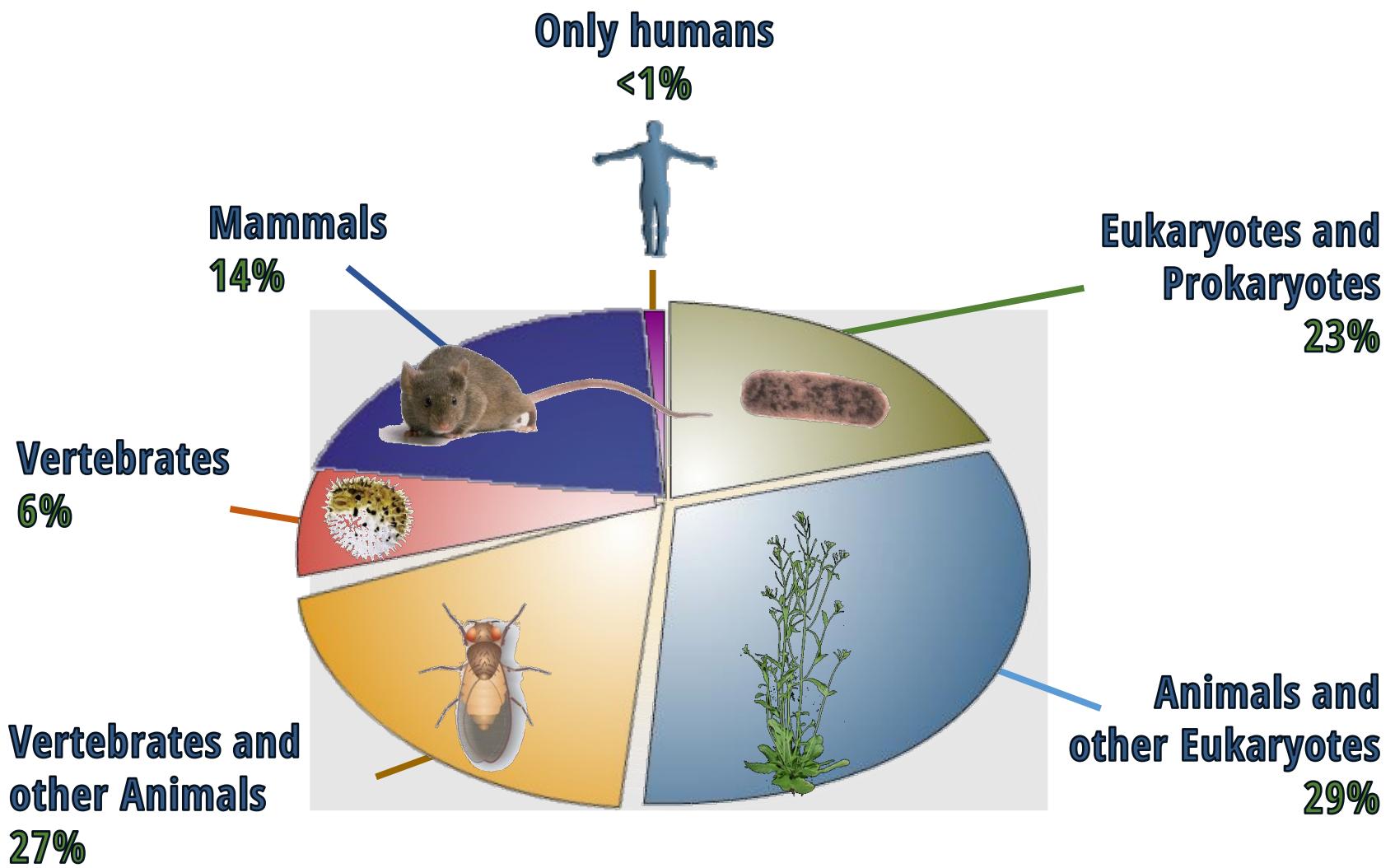
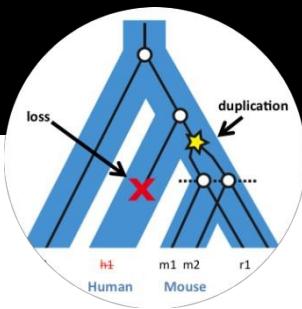


Genomic alignments & Syntenic blocks

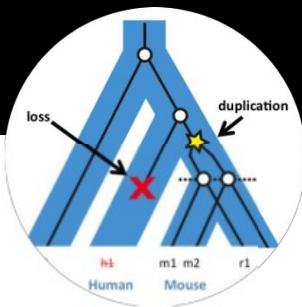


Sequence conservation & Functional constraints

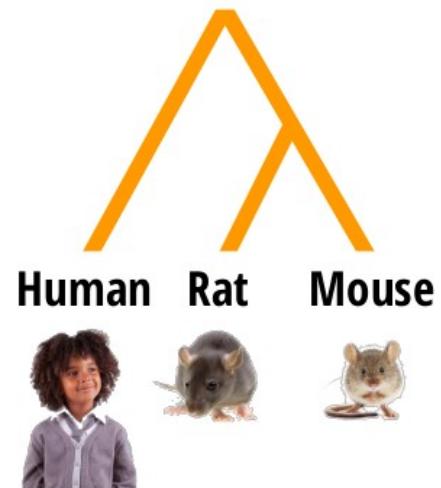
Comparative genomics: orthologues, paralogues and genetrees



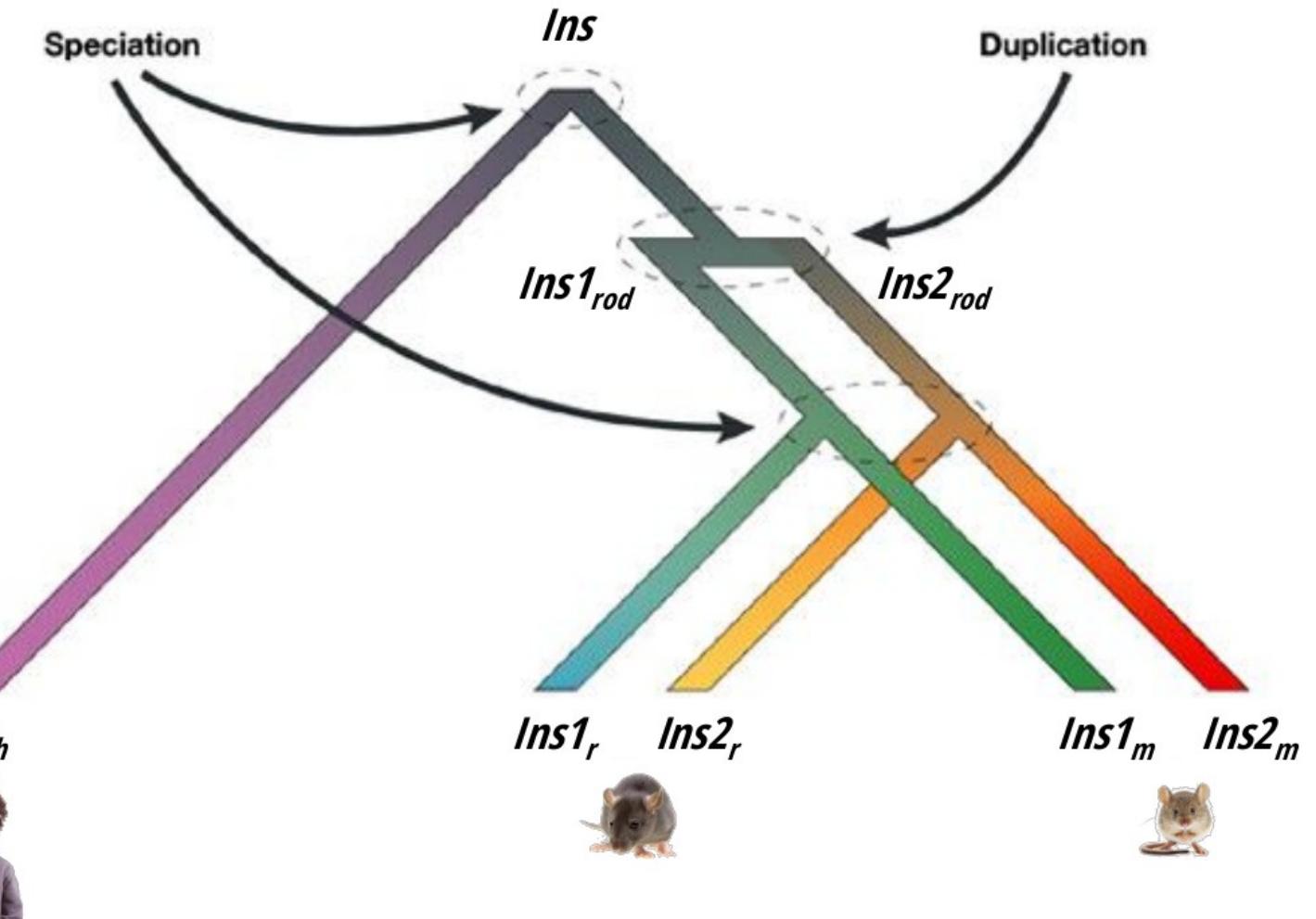
Comparative genomics: orthologues, paralogues and genetrees



SPECIES TREE



GENE TREE



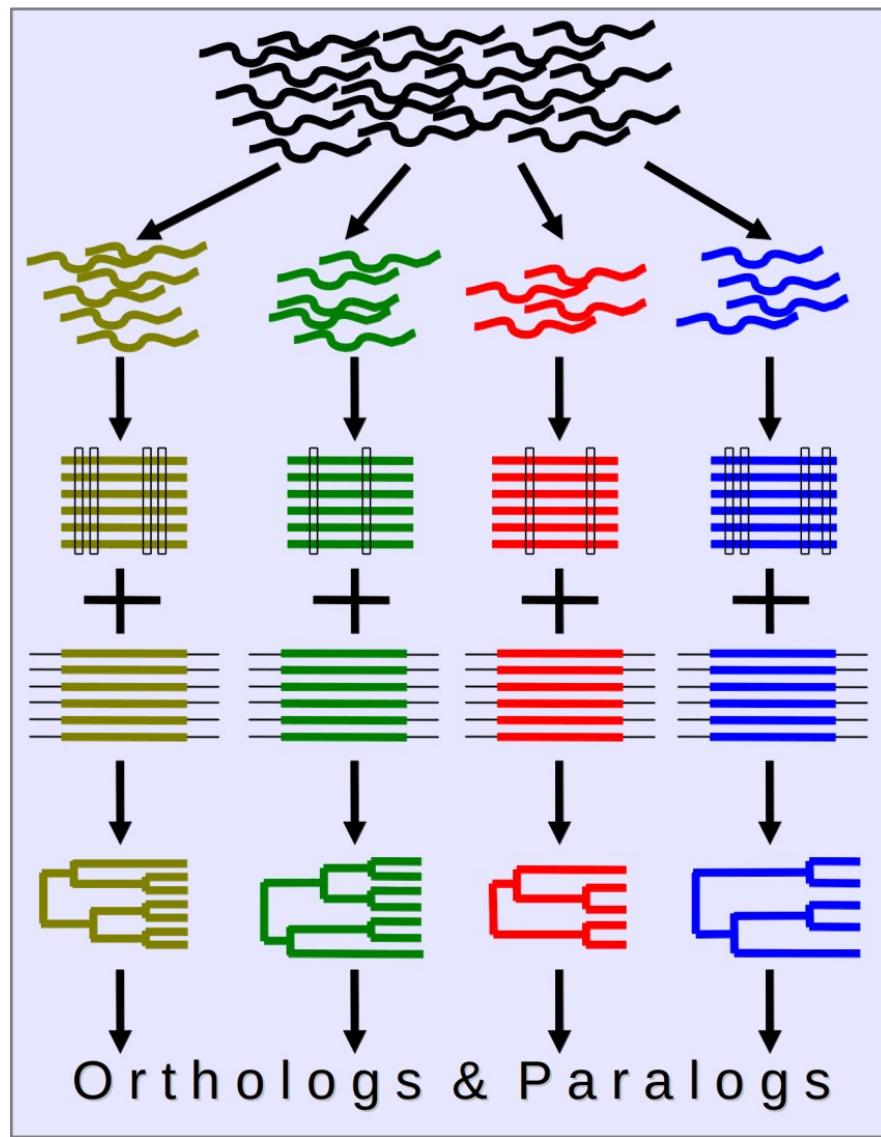
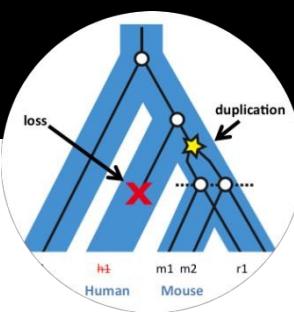
Orthologues = Result of a **SPECIATION**

Ins_1_r vs Ins_1_m ; Ins_2_r vs Ins_2_m

Paralogues = Result of a **DUPLICATION**

Ins_1_r vs Ins_2_r ; Ins_1_m vs Ins_2_m

Comparative genomics: orthologues, paralogues and genetrees



Longest translation for each *Ensembl* gene

WU-BLASTP all-versus-all hcluster_sg

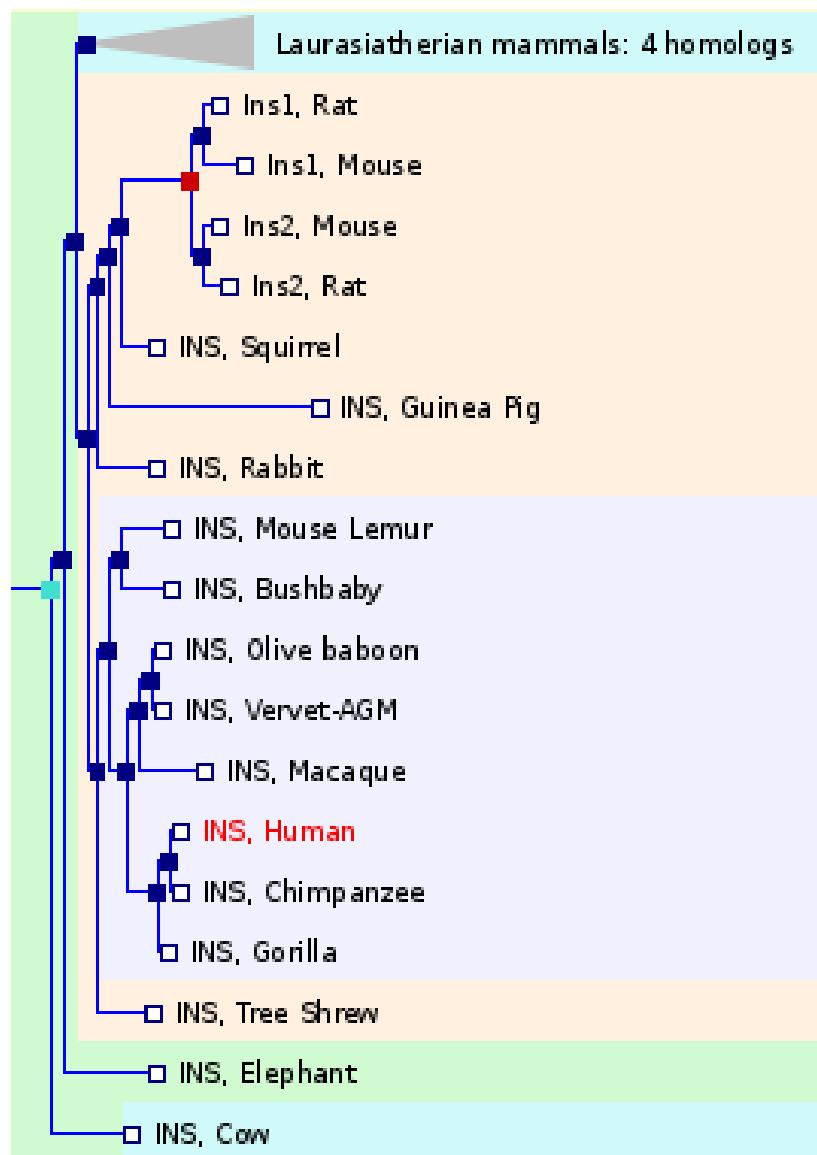
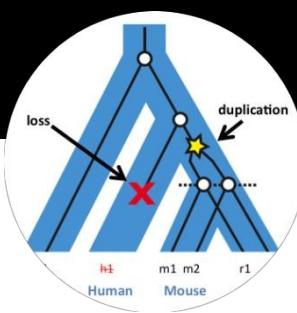
M-Coffee

TreeBeST
reconciliation with species tree

Infer orthologs and paralogs



Comparative genomics: orthologues, paralogues and genetrees



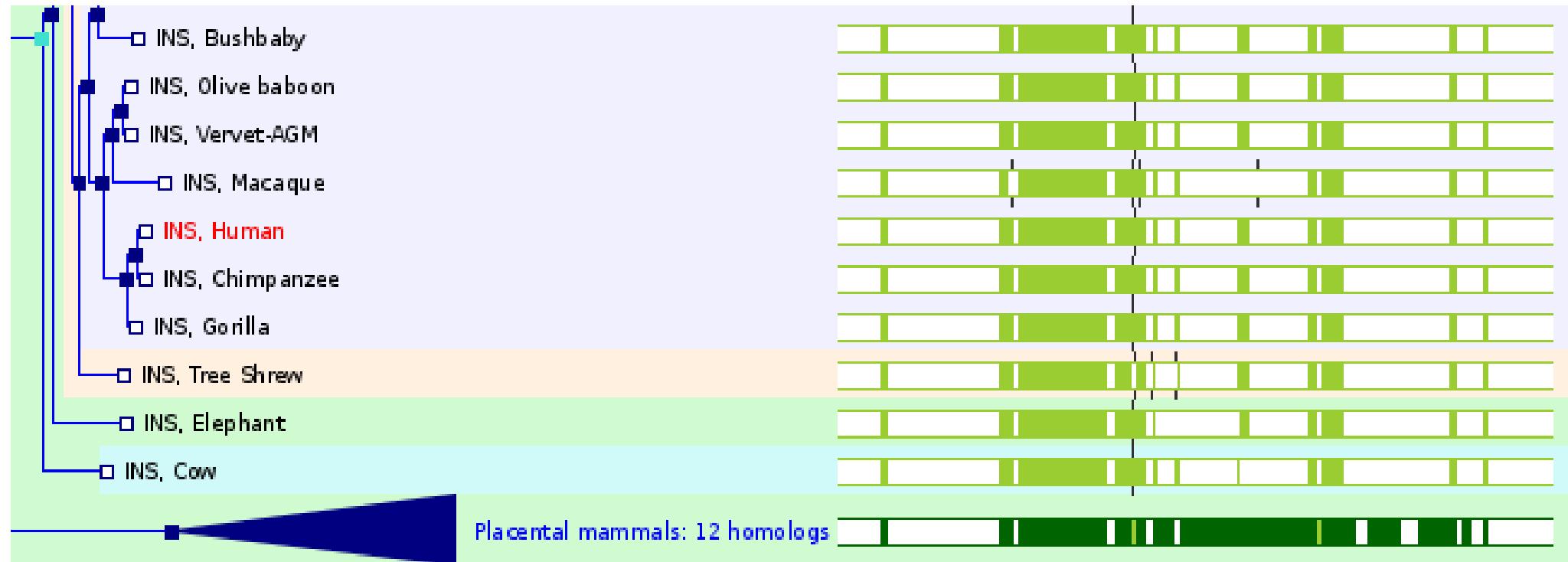
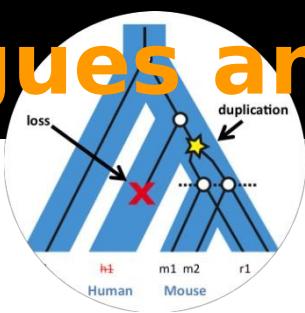
INS (insulin precursor)

**A blue square is a
speciation event
(orthologues)**

**A red square is a
duplication event
(paralogues)**



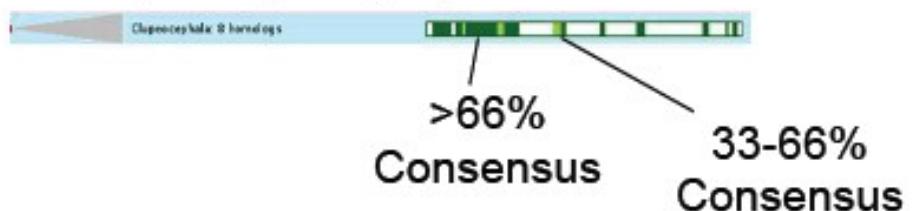
Comparative genomics: orthologues, paralogues and homologous recombination



Expanded nodes:

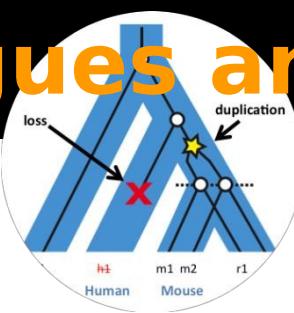


Collapsed (triangle) nodes:



e!Ensembl

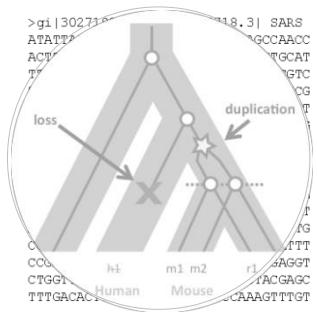
Comparative genomics: orthologues, paralogues and homologues



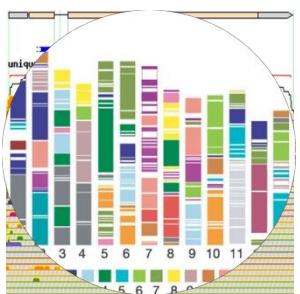
Species	Type	dN/dS	Ensembl identifier & gene name	Compare
Guinea Pig (<i>Cavia porcellus</i>)	1-to-1	n/a	ENSCPOG00000008198 INS Cavia porcellus insulin (Ins), mRNA. [Source:RefSeq mRNA;Acc:NM_001172891]	<ul style="list-style-type: none">Region ComparisonAlignment (protein)Alignment (cDNA)Gene Tree (image)
Mouse (<i>Mus musculus</i>)	1-to-many	0.12182	ENSMUSG00000000215 Ins2 insulin II [Source:MGI Symbol;Acc:MGI:96573]	<ul style="list-style-type: none">Region ComparisonAlignment (protein)Alignment (cDNA)Gene Tree (image)
Mouse (<i>Mus musculus</i>)	1-to-many	0.12202	ENSMUSG00000035804 Ins1 insulin I [Source:MGI Symbol;Acc:MGI:96572]	<ul style="list-style-type: none">Region ComparisonAlignment (protein)Alignment (cDNA)Gene Tree (image)
Rabbit (<i>Oryctolagus cuniculus</i>)	1-to-1	n/a	ENSOCUG00000000243 INS insulin [Source:HGNC Symbol;Acc:HGNC:6081]	<ul style="list-style-type: none">Region ComparisonAlignment (protein)Alignment (cDNA)Gene Tree (image)
Rat (<i>Rattus norvegicus</i>)	1-to-many	n/a	ENSRNOG00000012052 Ins1 insulin 1 [Source:RGD Symbol;Acc:2915]	<ul style="list-style-type: none">Region ComparisonAlignment (protein)Alignment (cDNA)Gene Tree (image)
Rat (<i>Rattus norvegicus</i>)	1-to-many	0.10477	ENSRNOG00000020405 Ins2 insulin 2 [Source:RGD Symbol;Acc:2916]	<ul style="list-style-type: none">Region ComparisonAlignment (protein)Alignment (cDNA)Gene Tree (image)

e!Ensembl

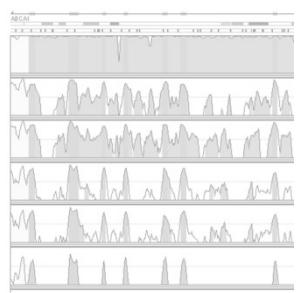
Comparative genomics



Orthologues, Paralogues & Genetrees

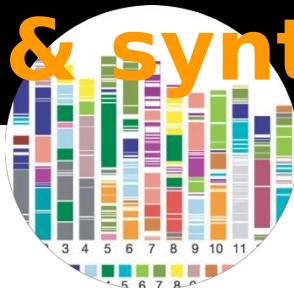


Genomic alignments & Syntenic blocks

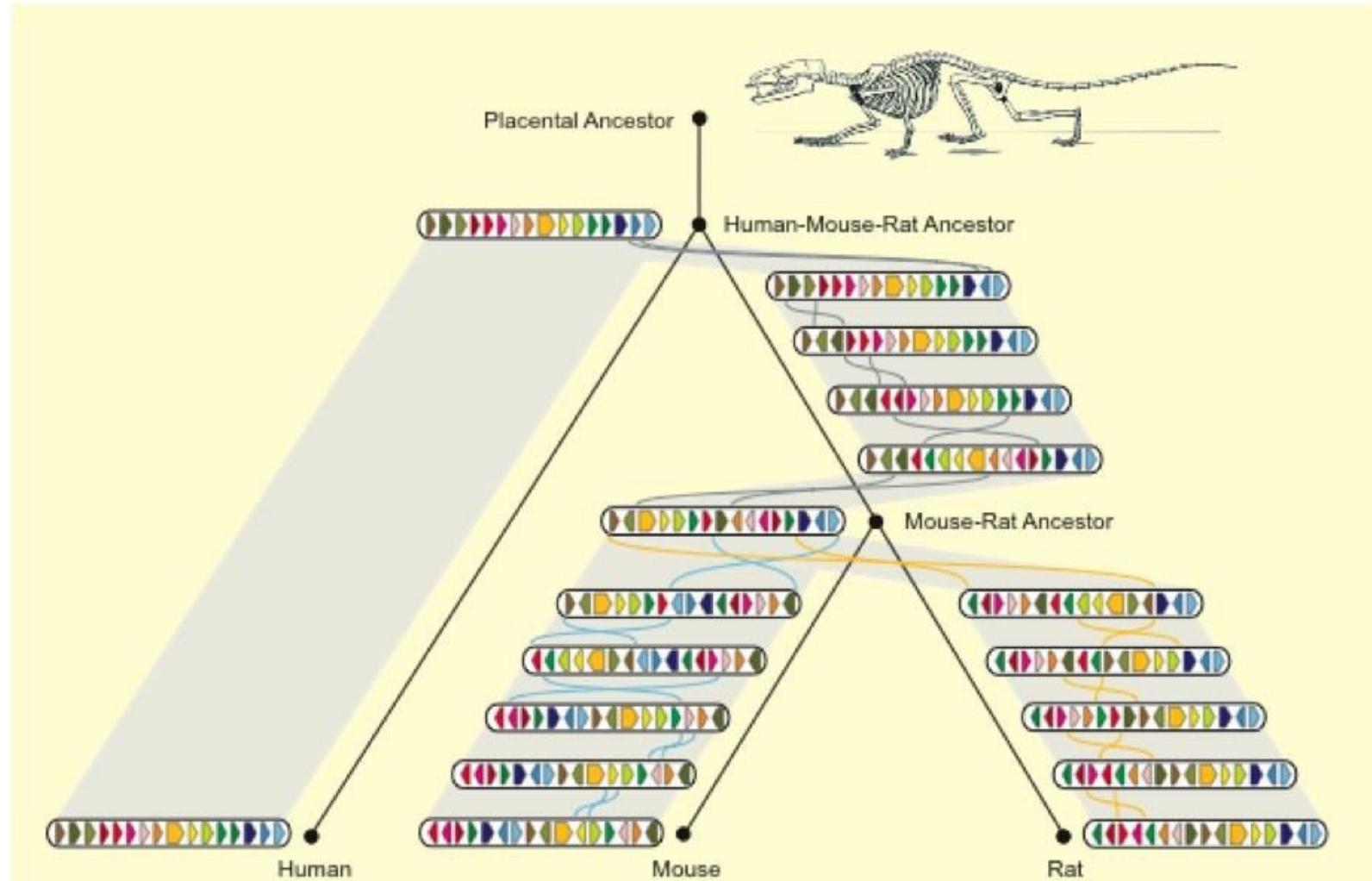


Sequence conservation & Functional constraints

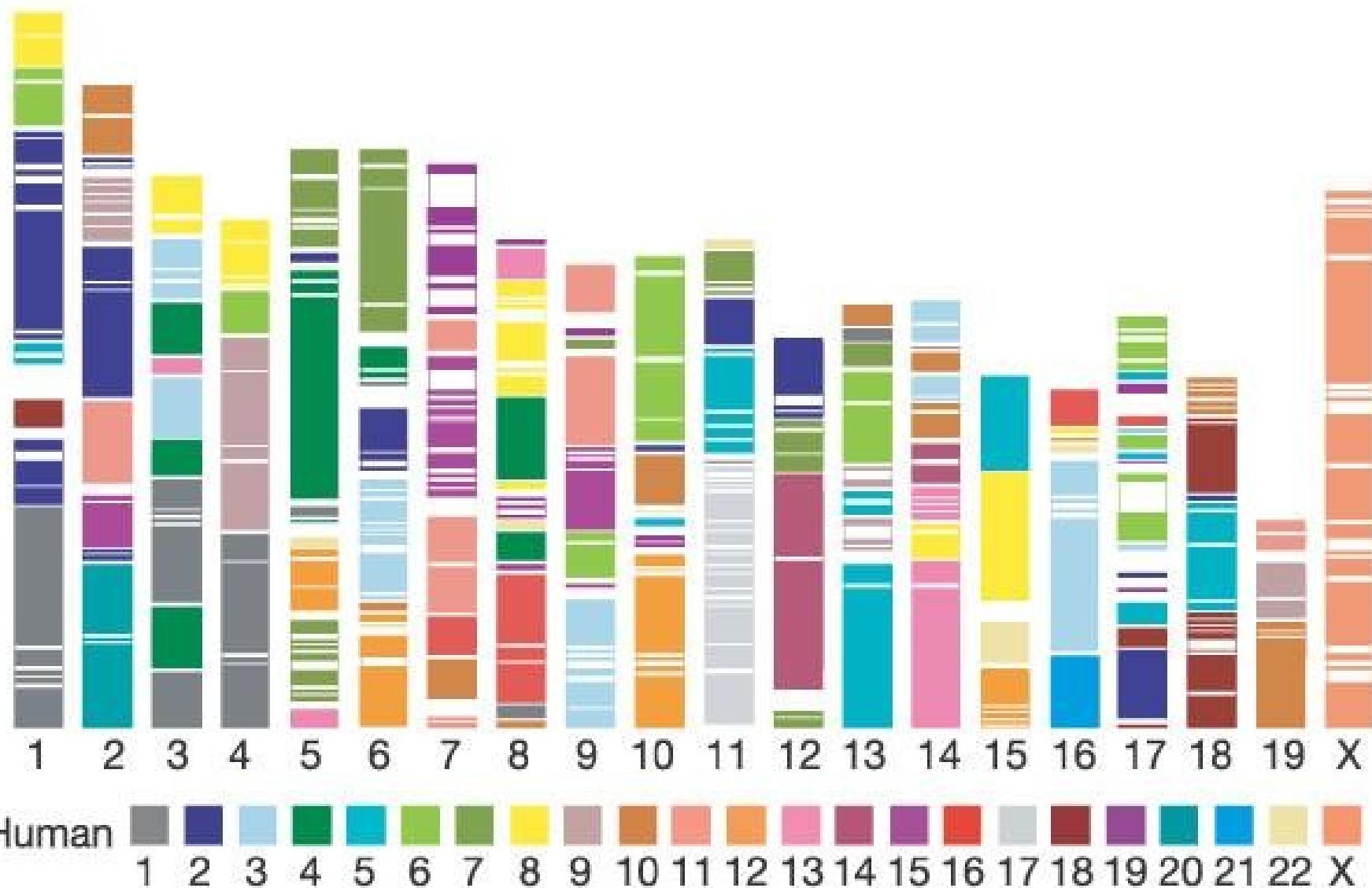
Comparative genomics: genomic alignments & synteny



Relative position of genes changes across evolution: **syntenic blocks**



Comparative genomics: genomic alignments & synteny blocks

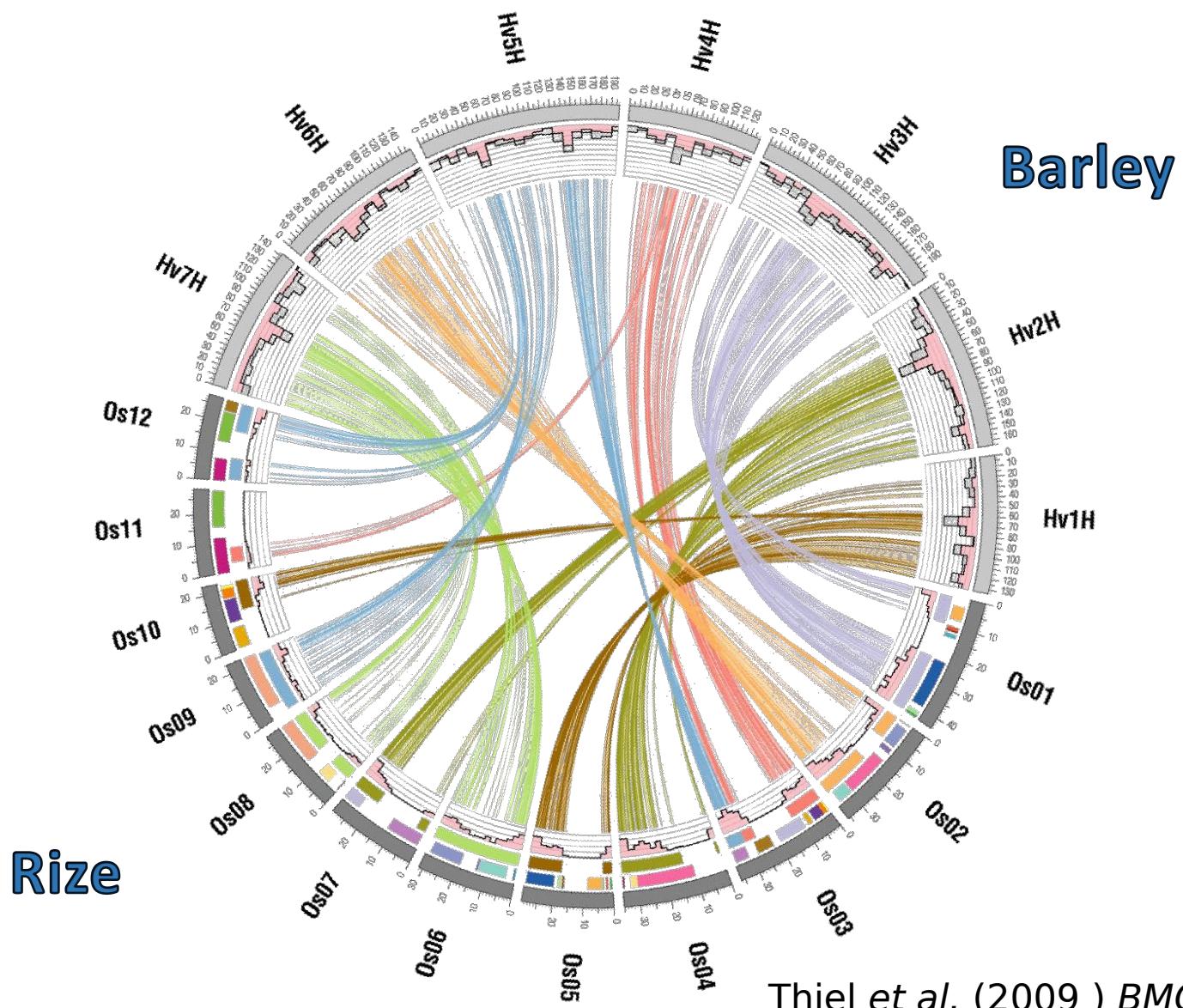


Human 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 X

Figure 3. Mouse Genome Sequencing Consortium (2002) *Nature* 420: 520-562

342 segments >300 kb
conserved between
the **human** and the
mouse genomes (avg.
6.9 Mb) that imply
295 chromosome
rearrangements

Comparative genomics: genomic alignments & syntentic blocks

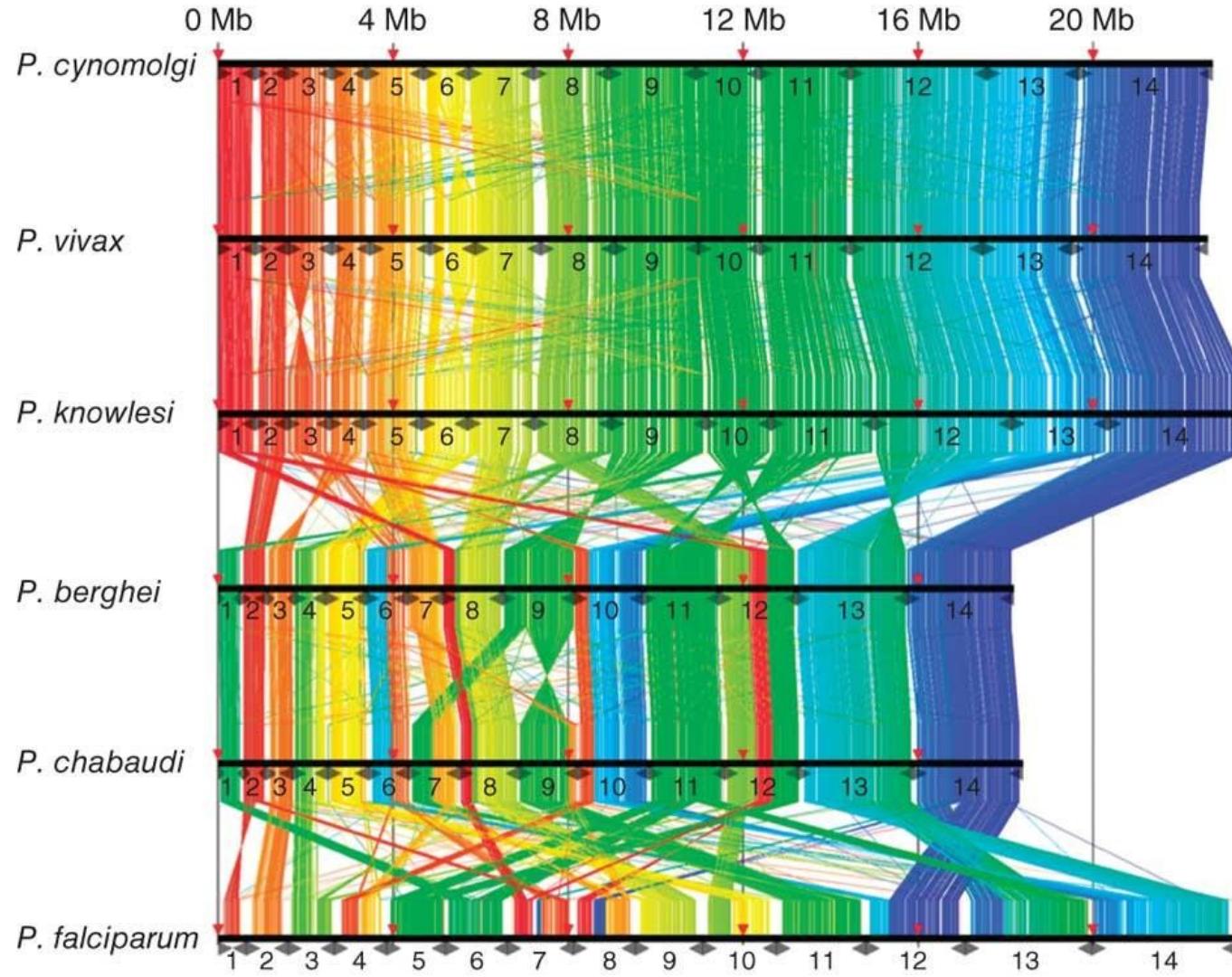


Thiel et al. (2009) *BMC Evolutionary Biology*

Comparative genomics: genomic alignments & synteny blocks



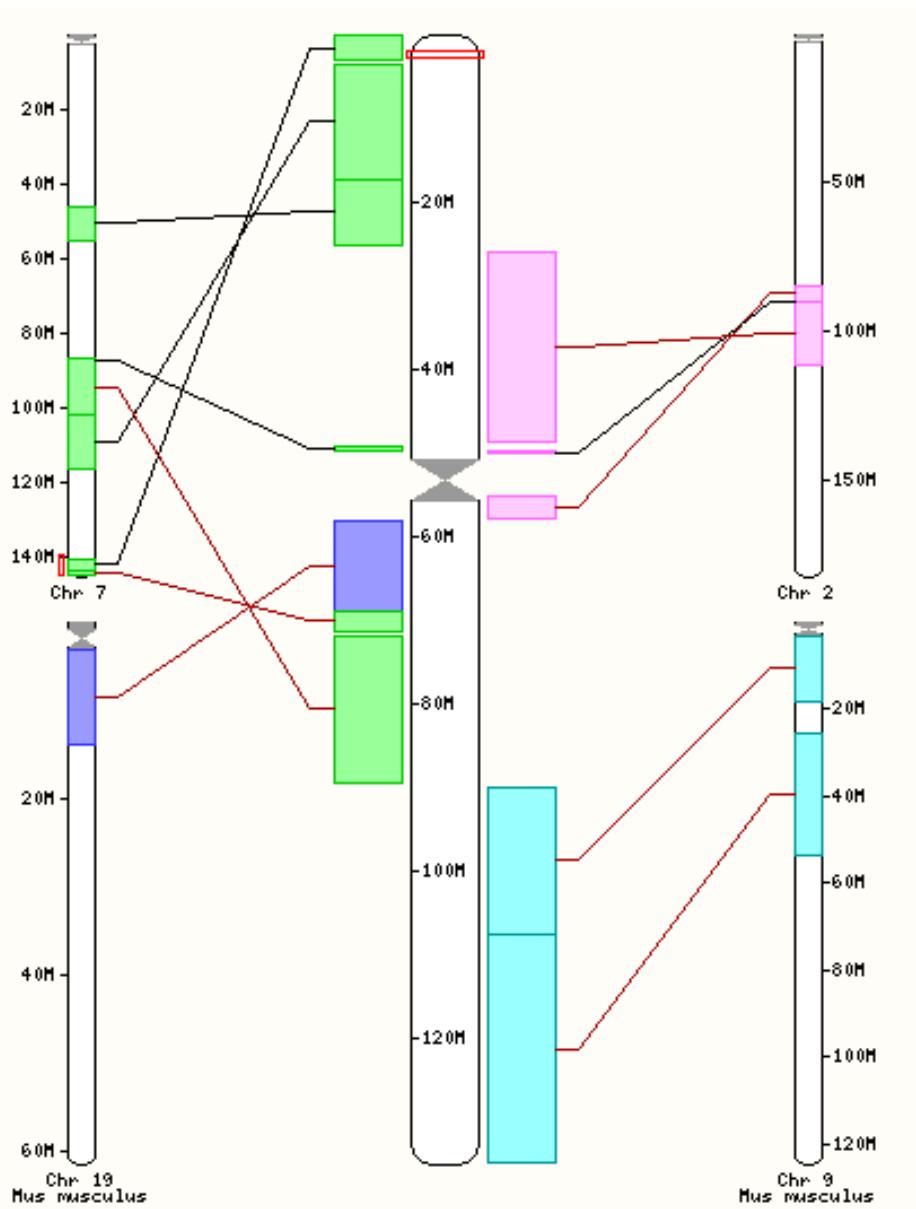
6 *Plasmodium* species



Shin-Ichiro Tachibana et al. (2012) *Nature Genetics* 44:1051-1055

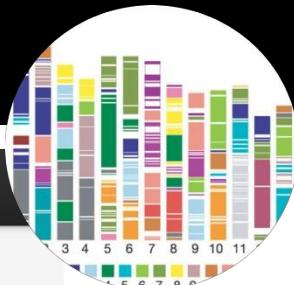


Synteny between Human chromosome 11 and Mouse chromosomes in Ensembl



e!Ensembl

Synteny Portal



SYNTENY PORTAL

SynCircos

SynBrowser

SynSearcher

SynBuilder

Documentation

Welcome to SYNTENY PORTAL!

SYNTENY PORTAL is a web application portal for visualizing, browsing, searching and building synteny blocks.

SynCircos

SynCircos draws the interactive Circos plot by using selected species and chromosomes.

Reference	Human	Assembly	Hg19 (GRCh37)	Chromosomes >>	1,3,5,6,8,12,15,19
Target 01	Mouse	mm10	Chromosomes >>	1,3,7,17,18	
Target 02	Cow	bovTau7	Chromosomes >>	3,9,10,14,17,22	

Add a target Delete a target

Resolution (bp) 150,000 Name type Species name Submit Reset

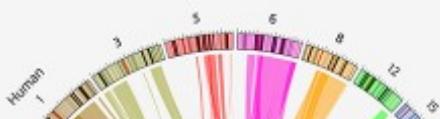
Image format SVG Download Show job status

If more than 20 chromosomes are selected, the Circos plot will be shown in a job status page.

It may take several minutes depending on selected numbers of species and chromosomes.

The Circos plot is created by using the Circos software package.

The Circos plot is dynamically highlighted by putting a mouse on a reference chromosome number, track, or ribbon.



SynCircos provides the high-quality Circos plot of multiple species synteny blocks

SynBrowser

SynBrowser shows synteny relationships between two chosen species with annotated genes of a reference species.

Users can easily navigate the reference chromosomes by using coordinates or gene symbols.

Reference	Human (hg19)	Chromosome	shrt	Target	Mouse (mm10)	Resolution (bp)	160,000	Submit
-----------	--------------	------------	------	--------	--------------	-----------------	---------	--------

Image format SVG Download Load an example Clear

Gene/Protein search in the reference Query Search

Navigate to the reference chromosome shrt start and Go

Users can also click on chromosome numbers in the color legend below to change a reference chromosome.

Gene annotation browser is displayed when a synteny block in the image below is clicked on.

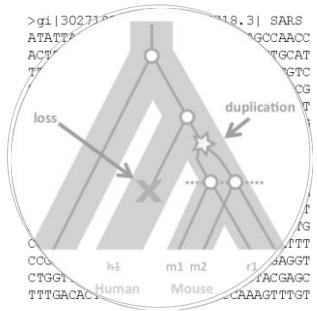
The coordinates of synteny blocks are shown when a mouse pointer is over a synteny block in the image below.



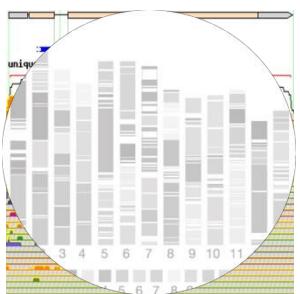
SynBrowser provides the chromosome level view and searching service of synteny blocks

http://bioinfo.konkuk.ac.kr/synteny_portal/

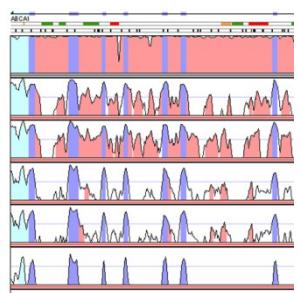
Comparative genomics



Orthologues, Paralogues & Genetrees

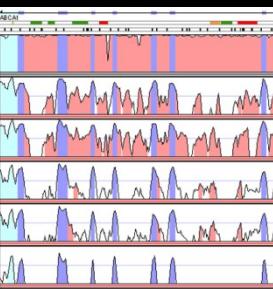


Genomic alignments & Syntenic blocks

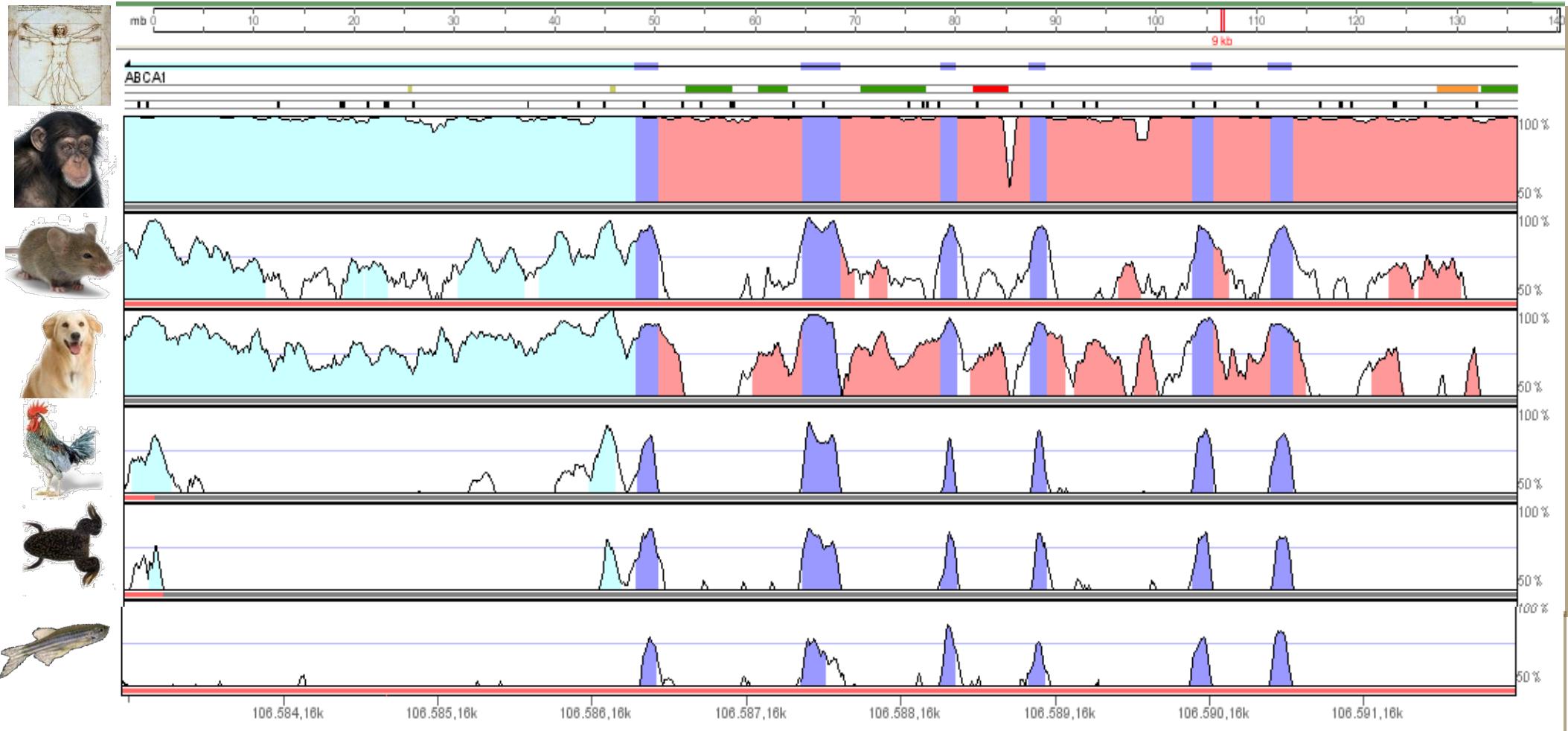


Sequence conservation & Functional constraints

Comparative genomics: sequence conservation & functional constraint

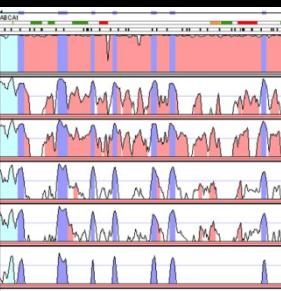


Functional sequences are **conserved** across evolution:



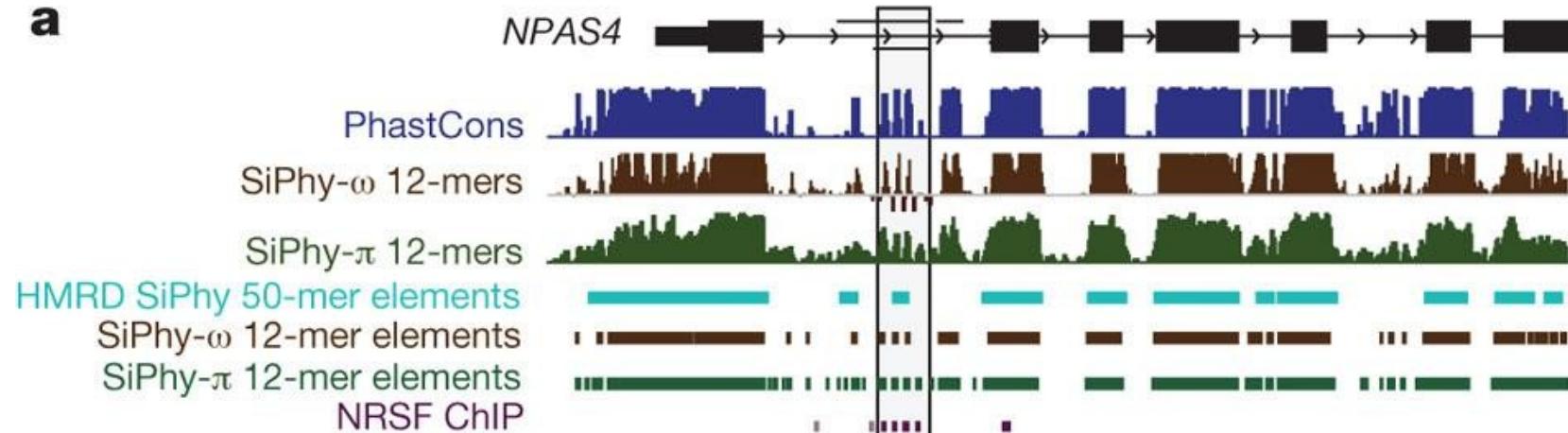
<http://pipeline.lbl.gov/>

Comparative genomics: sequence conservation & functional constraint

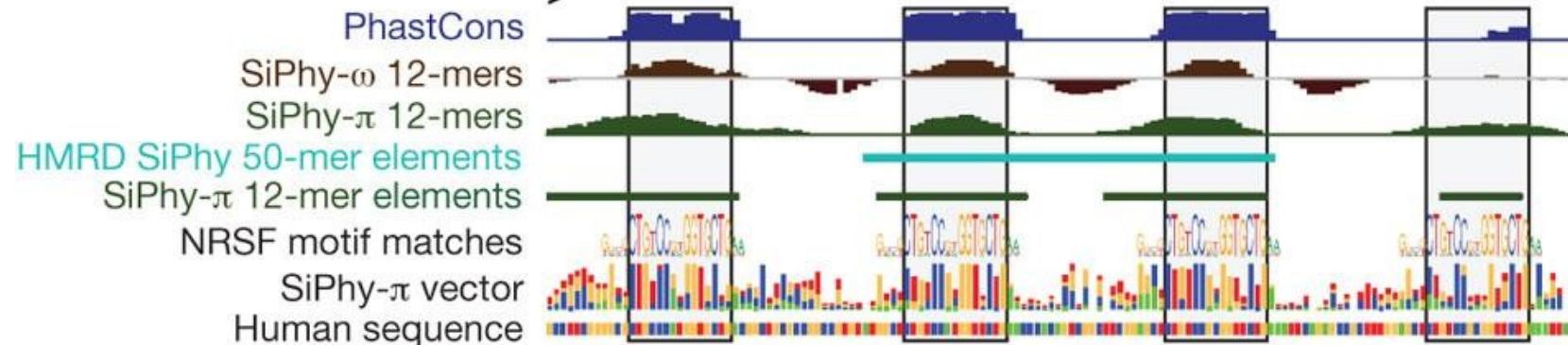


Functional sequences are **conserved** across evolution:

a



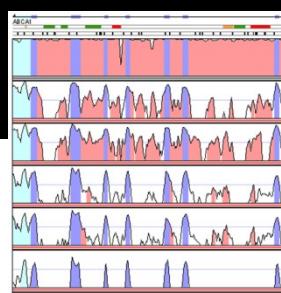
b



Conserved sequences in 29 mammalian genomes:

- ▶ ~5% of the human genome evolves under purifying selection (~1.5% are exons)
- ▶ 3.6 million conserved elements (avg. 36 bp)

Lindblad-Toh et al. 2011



Pairwise (two species)

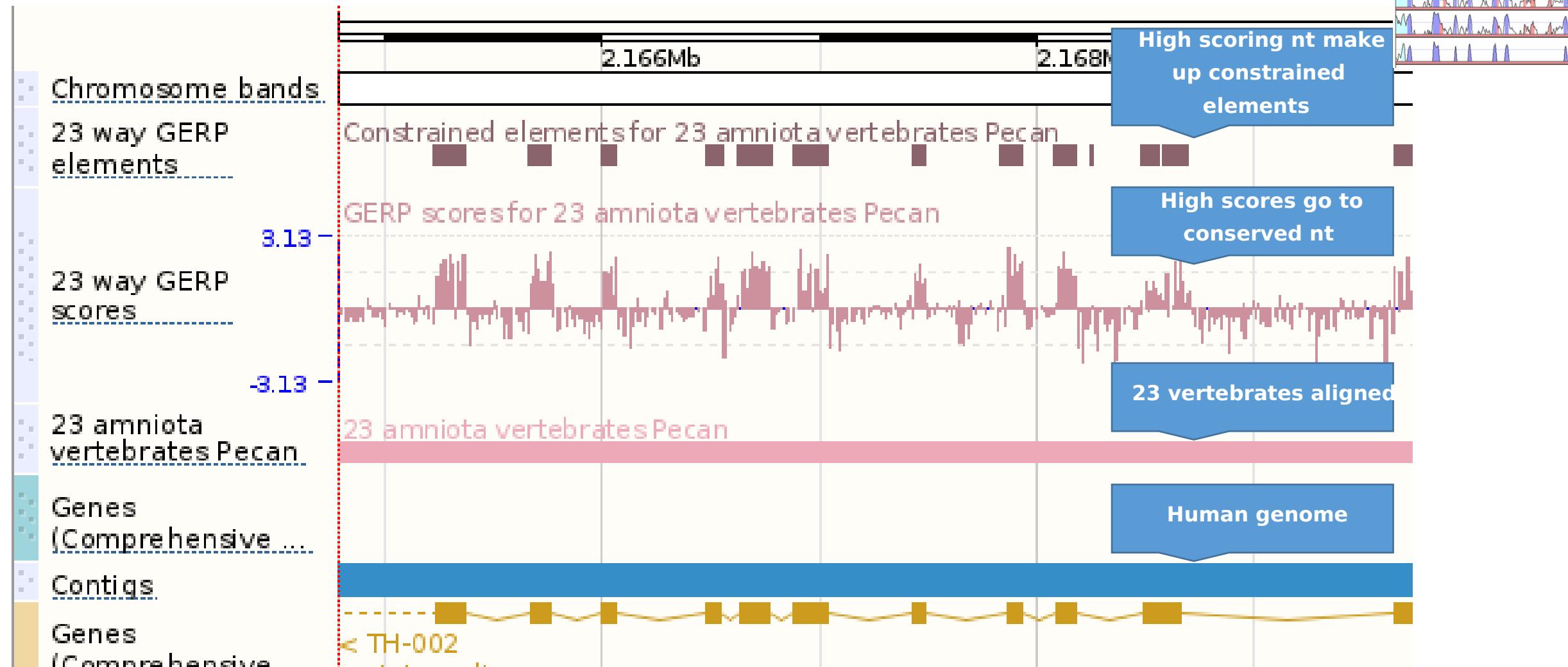
- Nucleotide alignment: BLASTZ/LASTZ-net  closer species e.g. human – mouse
- Amino acid alignment: Translated BLAT  more distant species, e.g. human – zebrafish

Multi-species (more than two species)

- Nucleotide alignments: EPO/PECAN  selected sets (primates, fish, birds, mammals, vertebrates)



Comparative genomics: sequence conservation & functional constraint



e!Ensembl

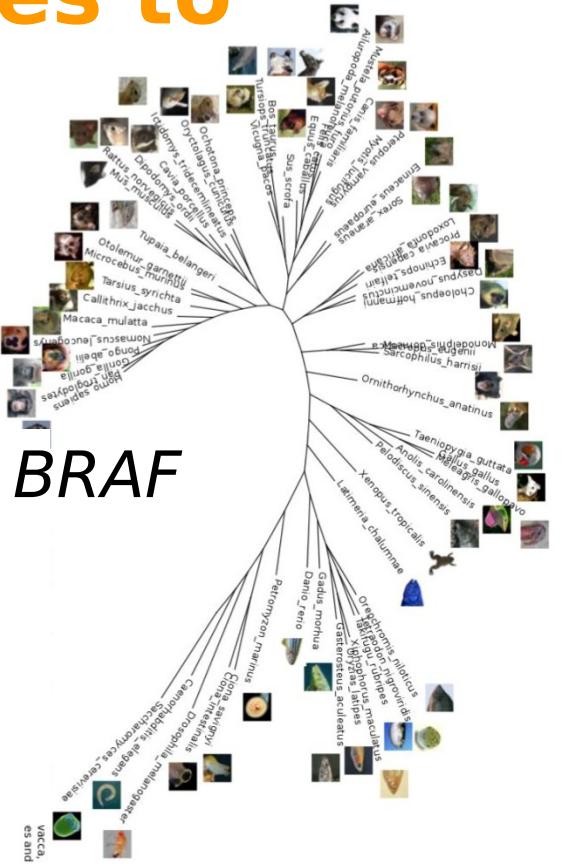
Practical work

Comparative genomics: from genetrees to whole genome alignments



<https://stepik.org/lesson/59030/step/1?unit=36598>

1. Orthologues, paralogues and protein families, human *BRAF*
 2. Sequence conservation and constraint, human *RHO*
 3. Synteny between human and horse, *RHO*
 4. Whole genome alignments, human *BRCA2*



 Ensembl