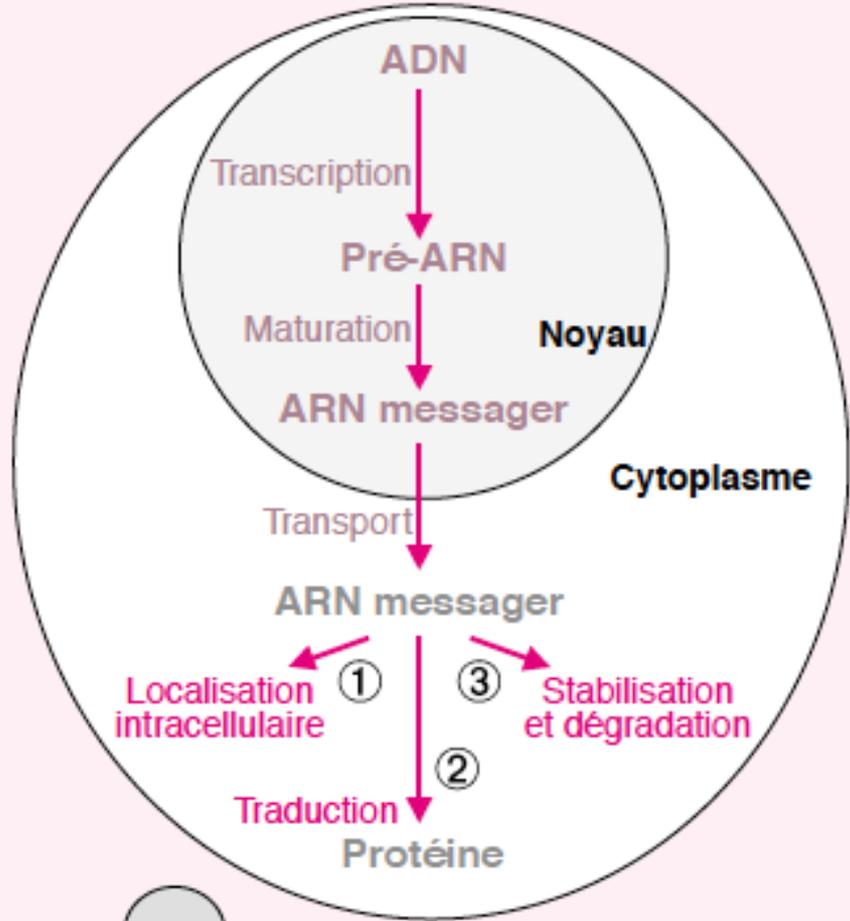
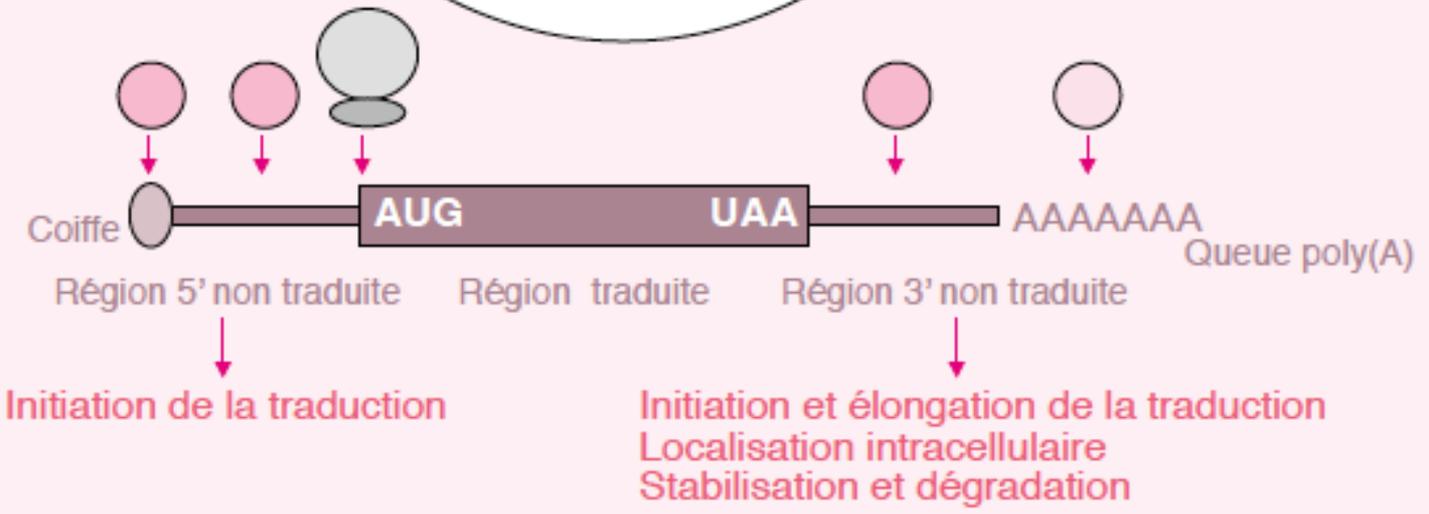


A**B**

National Center for
Biotechnology Information

All Databases

Search

NCBI Home

Resource List (A-Z)

All Resources

Chemicals & Bioassays

Data & Software

DNA & RNA

Domains & Structures

Genes & Expression

Genetics & Medicine

Genomes & Maps

Homology

Literature

Proteins

Sequence Analysis

Taxonomy

Training & Tutorials

Variation

Welcome to NCBI

The National Center for Biotechnology Information advances science and health by providing access to biomedical and genomic information.

[About the NCBI](#) | [Mission](#) | [Organization](#) | [NCBI News & Blog](#)

Submit

Deposit data or manuscripts
into NCBI databases



Download

Transfer NCBI data to your
computer



Learn

Find help documents, attend a
class or watch a tutorial



Develop

Use NCBI APIs and code
libraries to build applications



Analyze

Identify an NCBI tool for your
data analysis task



Research

Explore NCBI research and
collaborative projects



Popular Resources

[PubMed](#)[Bookshelf](#)[PubMed Central](#)[BLAST](#)[Nucleotide](#)[Genome](#)[SNP](#)[Gene](#)[Protein](#)[PubChem](#)

NCBI News & Blog

November 13 NCBI Minute: Resources
for next-gen sequence analysis

06 Nov 2019

On Wednesday, November 13, 2019 at
12 PM NCBI staff will present a webinar

NIH Biomedical Data Science Codeathon
in Pittsburgh, Jan 8-10

30 Oct 2019

NCBI is pleased to announce a
Biomedical Data Science Codeathon in

Search

- NCBI Home
- Resource List (A-Z)
- All Resources
- Chemicals & Bioassays
- Data & Software
- DNA & RNA
- Domains & Structures
- Genes & Expression
- Genetics & Medicine
- Genomes & Maps
- Homology
- Literature
- Proteins
- Sequence Analysis
- Taxonomy
- Training & Tutorials
- Variation

- All Databases
- Assembly
- Biocollections
- BioProject
- BioSample
- BioSystems
- Books
- ClinVar
- Conserved Domains
- dbGaP
- dbVar
- Gene
- Genome
- GEO DataSets
- GEO Profiles
- GTR
- HomoloGene
- Identical Protein Groups
- MedGen
- MeSH
- NCBI Web Site
- NLM Catalog
- Nucleotide
- OMIM
- PMC
- PopSet
- Probe
- Protein
- Protein Clusters
- PubChem BioAssay
- PubChem Compound
- PubChem Substance
- PubMed
- SNP
- Sparcle
- SRA
- Structure

to NCBI

Center for Biotechnology Information advances science and health by providing access to genomic information.

[NCBI](#) | [Mission](#) | [Organization](#) | [NCBI News & Blog](#)

Submit

Submit data or manuscripts to public databases



Download

Transfer NCBI data to your computer



Learn

Find help documents, attend a class or watch a tutorial



Develop

Use APIs and code to build applications



Analyze

Identify an NCBI tool for your data analysis task



Research

Explore NCBI research and collaborative projects



Popular Resources

[PubMed](#)

[Bookshelf](#)

[PubMed Central](#)

[BLAST](#)

[Nucleotide](#)

[Genome](#)

[SNP](#)

[Gene](#)

[Protein](#)

[PubChem](#)

NCBI News & Blog

November 13 NCBI Minute: Resources for next-gen sequence analysis

06 Nov 2019

On Wednesday, November 13, 2019 at 12 PM NCBI staff will present a webinar

NIH Biomedical Data Science Codeathon in Pittsburgh, Jan 8-10

30 Oct 2019

NCBI is pleased to announce a Biomedical Data Science Codeathon in

GenBank release 234 is available

Gene

Gene

Create RSS Save search Advanced

Gene sources

Genomic

Categories

Alternatively spliced

Annotated genes

Non-coding

Protein-coding

Sequence content

CCDS

Ensembl

RefSeq

RefSeqGene

Status

Current

Clear all

Show additional filters

Tabular 20 per page Sort by Relevance

Send to:

See [KHSRP \(KSRP\) KH-type splicing regulatory protein](#) in the Gene database
[ksrp](#) in [Homo sapiens](#) [Mus musculus](#) [Xenopus tropicalis](#) [All 5 Gene records](#)

Search results

Items: 1 to 20 of 60

<< First < Prev Page 1 of 3 Next > Last >>

See also 3 discontinued or replaced items.

clear

Name/Gene ID	Description	Location	Aliases	MIM
<input type="checkbox"/> KHSRP ID: 8570	KH-type splicing regulatory protein [<i>Homo sapiens</i> (human)]	Chromosome 19, NC_000019.10 (6413102..6424811, complement)	FBP2, FUBP2, KSRP, p75	603445
<input type="checkbox"/> Khsrp ID: 16549	KH-type splicing regulatory protein [<i>Mus musculus</i> (house mouse)]	Chromosome 17, NC_000083.6 (57021049..57031507, complement)	6330409F21Rik, Fbp2, Fubp2, Ksrp	
<input type="checkbox"/> khsrp ID: 100170586	KH-type splicing regulatory protein [<i>Xenopus tropicalis</i> (tropical clawed frog)]	Chromosome 3, NC_030679.1 (130039451..130057575, complement)	VgRBP71, fbp2, fubp2, ksrp	
<input type="checkbox"/> khsrp.S ID: 399189	KH-type splicing regulatory protein S homeolog [<i>Xenopus laevis</i> (African clawed frog)]		XELAEV_18000829mg, VgRBP71, fbp2, fubp2, khsrp-b, ksrp	
<input type="checkbox"/> khsrp.L	KH-type splicing	Chromosome 3L,	XELAEV_18019123mg,	

Gene

Gene

Advanced

Full Report Send to:

KHSRP KH-type splicing regulatory protein [*Homo sapiens* (human)]

Gene ID: 8570, updated on 3-Nov-2019

Summary

Official Symbol KHSRP provided by [HGNC](#)**Official Full Name** KH-type splicing regulatory protein provided by [HGNC](#)**Primary source** [HGNC:HGNC:6316](#)**See related** [Ensembl:ENSG00000088247](#) [MIM:603445](#)**Gene type** protein coding**RefSeq status** VALIDATED**Organism** [Homo sapiens](#)**Lineage** Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini; Catarrhini; Hominidae; Homo**Also known as** p75; FBP2; KSRP; FUBP2**Summary** The KHSRP gene encodes a multifunctional RNA-binding protein implicated in a variety of cellular processes, including transcription, alternative pre-mRNA splicing, and mRNA localization (Min et al., 1997 [PubMed 9136930]; Gherzi et al., 2004 [PubMed 15175153]).[supplied by OMIM, Apr 2010]**Expression** Ubiquitous expression in testis (RPKM 37.2), endometrium (RPKM 25.1) and 25 other tissues [See more](#)**Orthologs** [mouse](#) [all](#)

Genomic context

Nucleotide

Nucleotide ▾

Advanced

GenBank ▾

Send to: ▾

Homo sapiens chromosome 19, GRCh38.p13 Primary Assembly

NCBI Reference Sequence: NC_000019.10

[FASTA](#) [Graphics](#)

LOCUS NC_000019 11710 bp DNA linear CON 09-SEP-2019

DEFINITION Homo sapiens chromosome 19, GRCh38.p13 Primary Assembly.

ACCESSION [NC_000019](#) REGION: complement(6413102..6424811)

VERSION NC_000019.10

DBLINK BioProject: [PRJNA168](#)
 Assembly: [GCF_000001405.39](#)

KEYWORDS RefSeq.

SOURCE Homo sapiens (human)

ORGANISM [Homo sapiens](#)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini;
 Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 11710)

CONSRM International Human Genome Sequencing Consortium

TITLE Finishing the euchromatic sequence of the human genome

JOURNAL Nature 431 (7011), 931-945 (2004)

PUBMED [15496913](#)

REFERENCE 2 (bases 1 to 11710)

AUTHORS Grimwood,J., Gordon,L.A., Olsen,A., Terry,A., Schmutz,J.,
 Lamerdin,J., Hellsten,U., Goodstein,D., Couronne,O.,
 Tran-Gyamfi,M., Aerts,A., Altherr,M., Ashworth,L., Bajorek,E.,
 Black,S., Branscomb,E., Caenepeel,S., Carrano,A., Caoile,C.,
 Chan,Y.M., Christensen,M., Cleland,C.A., Copeland,A., Dalin,E.,
 Dehal,P., Denys,M., Detter,J.C., Escobar,J., Flowers,D.,
 Fotopulos,D., Garcia,C., Georgescu,A.M., Glavina,T., Gomez,M.,
 Gonzales,E., Groza,M., Hammon,N., Hawkins,T., Haydu,L., Ho,I.,
 Huang,W., Israni,S., Jett,J., Kadner,K., Kimball,H., Kobayashi,A.,
 Larionov,V., Leem,S.H., Lopez,F., Lou,Y., Lowry,S., Malfatti,S.,
 Martinez.D.. McCreadv.P.. Medina.C.. Morgan.J.. Nelson.K..

Nucleotide

Nucleotide

Search input field

Search

Advanced

Help

FASTA

Send to

Homo sapiens chromosome 19, GRCh38.p13 Primary Assembly

NCBI Reference Sequence: NC_000019.10

[GenBank](#) [Graphics](#)

>NC_000019.10:c6424811-6413102 Homo sapiens chromosome 19, GRCh38.p13 Primary Assembly

```
AGAGTGTCTCCGCGCCGTGTGGAGCGAGGCCCTTGTTCCTCCGCGTTGAGCCCGCCGCGCCGCGCCGCTCC
TCAGCTTCAGCCTCCGCGCCAGGCCCGGCCCGCCGCGCCATGTCGGACTACAGCACGGGAGGACCCCGG
CCCGGGCCCGCCGCGCCGCGCGGGGGGGGGGAGCCGGAGGCGCGGGGGAGGCCCTCCGCGGGGGC
CGCCAGGCGCGGGGACCGGGGCGCGCGGTCCCGCGCGCGCGCCCGGGCGGGGGTCCGGCGGGGG
CCCCCTCAGCACCCGCGGAGGCGGCCCGGAATCCGCAAGGACGCTTTCGCGACGCCGTGCAGCGG
CGCCGCCAGGTGAGGAGCCGAGGGCTGAGGGGCGCGCTCGCGCGCGCGCAGGCGGGGGAGGGGGCG
GGGGTCACGTGCGCGCGCGCGCAGGCTCGCGGAGGGGGCGGGGCCGTCGCGGGGTACGTGGGCGG
TAGGCGGGGCTGGCCCTCCGCTGCGCCTCGAGTGGGATCCCAACGGGTGACCCCTGACAGGAGTA
GGGCGCACTGGGGTACAGGAGCGCGCCTGGGTGTGACGGCGCCCGGAGCTCGGAGACGCACTTT
TGGGACCCCTTAGCCCCCGCCCTTAGGACATGCCTGGAGGAGGGTCCACCTTCCCGGGCGGAG
GGCCTTTACTCTCCCTCCATATCAGGGCAAAGGCGAGTGTTCGCCCCAACCCAGTCCCATTTCTGTCC
CCTCTCTAGGAAAAGGACATTTGTGGGGTGGGAGGACAGTCTTCTCTTCAGTCGGGAAAGTGGAT
GGGACCTGTCTAGGGCTGAGCACTTGGGCGAAATGGGAAAGACGAGCATGGGGACTGGCTCTTGGATGC
TTTGGTGGCTTAAAGCGTGGTCTGTCTTTTGGGGAAAGAAGGAGGGTCTCCCACCCCGTGTATTG
AGGCGTAATGACAAGGAAGTGTACATGCTGTAGGACTGACCTTCCAGTCCATTTGCAGCTCCCA
TGGCCTGGAGGGGTGGGGTGGTGTAGAAATGGGAGGACTTTGCCGCGACCGGCTCCTGGGCTGTCCCAAT
GCCTCCGGGCTAGAGACTGGGTCCTCCCAAGGTGGGAGGTCCCGTTAAGTGAACGCCTGTGTCTGGCC
CTCTGTCTCTCTCACCCCTTGGAGTGGGTGGGTCTAGCGTGCATTTGGTCTTATTGCCCATGGAAC
CGTCTGATTCTGGGACAGCTTGAAGGGAGGAAGCTGGTGTGTGAGTTGGTGGGACAGGCCCCCTGTGG
GCATCCCATGAGTTTTGGCTCTAGGGCAGCGTAGAAAAATGGGAGACTGGTGGAAAGGGAGATTCTTAACA
TGGTTTGGGTTTCTAAACCCAGACTCTGATCCTAAGTTTGTGGTGGGTTTCTTTAAAGTTGCTCT
TCACGTGTCTGAGATGTAACCTGTGTCCCATAGAGTCTTGGTTTTTGTGTTGTTTGTGTTTTTTGGAA
ACAGTCTCAATATCACCCAGGCAGGGGTGCAGTGGCACGATCTCAGCACACCACAGCCTCCACCTCTCCA
GTTGAAGTGTCTCTCTGCCTCAGCCTCCCGAGTAGCTGGGATTATAAGTGTGCACCACCACCCAGCT
```

Change region shown

Whole sequence
 Selected region

from: to:

[Update View](#)

Customize view

Display options

Show reverse complement

[Update View](#)

Analyze this sequence

[Run BLAST](#)

[Pick Primers](#)

[Highlight Sequence Features](#)

Related information

[Assembly](#)

[BioProject](#)

[Protein](#)

Primer-BLAST

A tool for finding specific primers

Finding primers specific to your PCR template (using Primer3 and BLAST).

[Reset page](#)[Save search parameters](#)[Retrieve recent results](#)[Publication](#)[Tips for finding specific primers](#)

PCR Template

Enter accession, gi, or FASTA sequence (A refseq record is preferred)  [Clear](#)

Range

From

To

Forward primer

[Clear](#)

Reverse primer

Or, upload FASTA file

aucun fichier sél.

Primer Parameters

Use my own forward primer
(5'→3' on plus strand)

[Clear](#)

Use my own reverse primer
(5'→3' on minus strand)

[Clear](#)

PCR product size

Max

of primers to return

Primer melting temperatures
(T_m)

Min

Opt

Max

Max T_m difference



Exon/intron selection

A refseq mRNA sequence as PCR template input is required for options in the section 

Exon junction span



Exon junction match

Exon at 5' side

Exon at 3' side

Minimal number of bases that must anneal to exons at the 5' or 3' side of the junction 

Intron inclusion

Primer pair must be separated by at least one intron on the corresponding genomic DNA 

Select the [Task](#) for primer selection

[Template masking](#) before primer design ([available species](#))

[Select species](#)

[Nucleotides to mask in 5' direction](#)

[Primer failure rate cutoff](#)

[Nucleotides to mask in 3' direction](#)

Paste source sequence below (5'→3', string of ACGTNacgtn -- other letters treated as N -- numbers and blanks ignored). FASTA format ok. Please N-out undesirable sequence (vector, ALUs, LINES, etc.) or use a [Mispriming Library \(repeat library\)](#)

Pick left primer,
or use left primer below

Pick hybridization probe (internal
oligo), or use oligo below

Pick right primer, or use right primer below
(5' to 3' on opposite strand)

Go to <http://genome.ucsc.edu/index.html> click on Genome Browser

UCSC Genome Bioinformatics

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

Genome
Browser

ENCODE

Neandertal

Blat

Table
Browser

Gene Sorter

In Silico PCR

Genome
Graphs

Galaxy

VisiGene

Utilities

Downloads

Release Log

Custom
Tracks

Microbial
Genomes

Mirrors

UCSC Genome Bioinformatics Site

Welcome to the UCSC Genome Browser website. This site contains the reference sequence and working draft assemblies for human and other vertebrate genomes. [Neandertal](#) projects.

We encourage you to explore these sequences with our tools. The [Genome Browser](#) zooms and scrolls over chromosomes, showing gene structure, homology and other information on groups of genes that can be related in many ways. [Blat](#) quickly maps your sequence to the database. [VisiGene](#) lets you browse through a large collection of *in situ* mouse and frog images to examine expression patterns.

The UCSC Genome Browser is developed and maintained by the Genome Bioinformatics Group, a cross-departmental team 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.

News

To receive announcements of new genome assembly releases, new software features, updates and training seminars by email, please [subscribe](#).

16 August 2012 - Announcing a Genome Browser for the Medium ground finch

We have released a browser for the Medium ground finch, *Geospiza fortis*, renowned as one of naturalist Charles Darwin's most important evolutionary studies, is one of a group of birds that evolved over a few million years from a single ancestral species into multiple species with diverse food resources. The phenotypic diversity of these birds contributed to Darwin's theory of evolution. The significance of this genome is being highlighted by the UCSC Center for Biomolecular Science and Engineering (CBSE).

The initial Medium ground finch genome assembly (GeoFor_1.0, UCSC version geoFor1) is the product of a collaboration between the UCSC Genome Browser and the University of California Santa Cruz. The sequence is the first to be released in the UCSC Genome Browser. For more information about this release, see the [Credits](#) page.

Bulk downloads of the sequence and annotation data are available via the Genome Browser [FTP server](#) or the [Downloads](#) page. See the [Credits](#) page for a detailed list of the organizations and individuals who contributed to this release. We'd like to thank the UCSC staff members who released this browser: Hiram Clawson and Greg Roe.

Begin typing the symbol of your gene into the "Search term" box. Select the most appropriate one from the autocomplete list

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	search term	
Mammal	Human	Feb. 2009 (GRCh37/hg19)	chr17:7571720-7590863	TP53	
				<input type="submit" value="submit"/>	

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

- TP53 (Homo sapiens tumor protein p53 (TP53), transcript variant 1, mRNA.)
- TP53AIP1 (Homo sapiens tumor protein p53 regulated apoptosis inducing protein 1 (TP53AIP1), mitochondrial protein, transcript variant 3, mRNA.)
- TP53BP1 (Homo sapiens tumor protein p53 binding protein 1 (TP53BP1), transcript variant 1, mRNA.)
- TP53BP2 (Homo sapiens tumor protein p53 binding protein, 2 (TP53BP2), transcript variant 1, mRNA.)
- TP53I11 (Homo sapiens tumor protein p53 inducible protein 11 (TP53I11), mRNA.)
- TP53I13 (Homo sapiens tumor protein p53 inducible protein 13 (TP53I13), mRNA.)
- TP53I3 (Homo sapiens tumor protein p53 inducible protein 3 (TP53I3), transcript variant 1, mRNA.)
- TP53INP1 (Homo sapiens tumor protein p53 inducible nuclear protein 1 (TP53INP1), transcript variant 1, mRNA.)
- TP53INP2 (Homo sapiens tumor protein p53 inducible nuclear protein 2 (TP53INP2), transcript variant 1, mRNA.)
- TP53RK (Homo sapiens TP53 regulating kinase (TP53RK), mRNA.)
- TP53TG1 (Homo sapiens TP53 target 1 (non-protein coding) (TP53TG1), non-coding RNA.)
- TP53TG3B (Homo sapiens TP53 target 3B (TP53TG3B), transcript variant 2, mRNA.)
- TP53TG3B (Homo sapiens TP53 target 3B (TP53TG3B), transcript variant 2, mRNA.)
- TP53TG3C (Homo sapiens TP53 target 3C (TP53TG3C), transcript variant 2, mRNA.)
- TP53TG5 (Homo sapiens TP53 target 5 (TP53TG5), mRNA.)

Reference sequence (GRCh37) was produced by the [Genome Reference Consortium](#). For more information see the [Genome Reference Consortium](#) website.

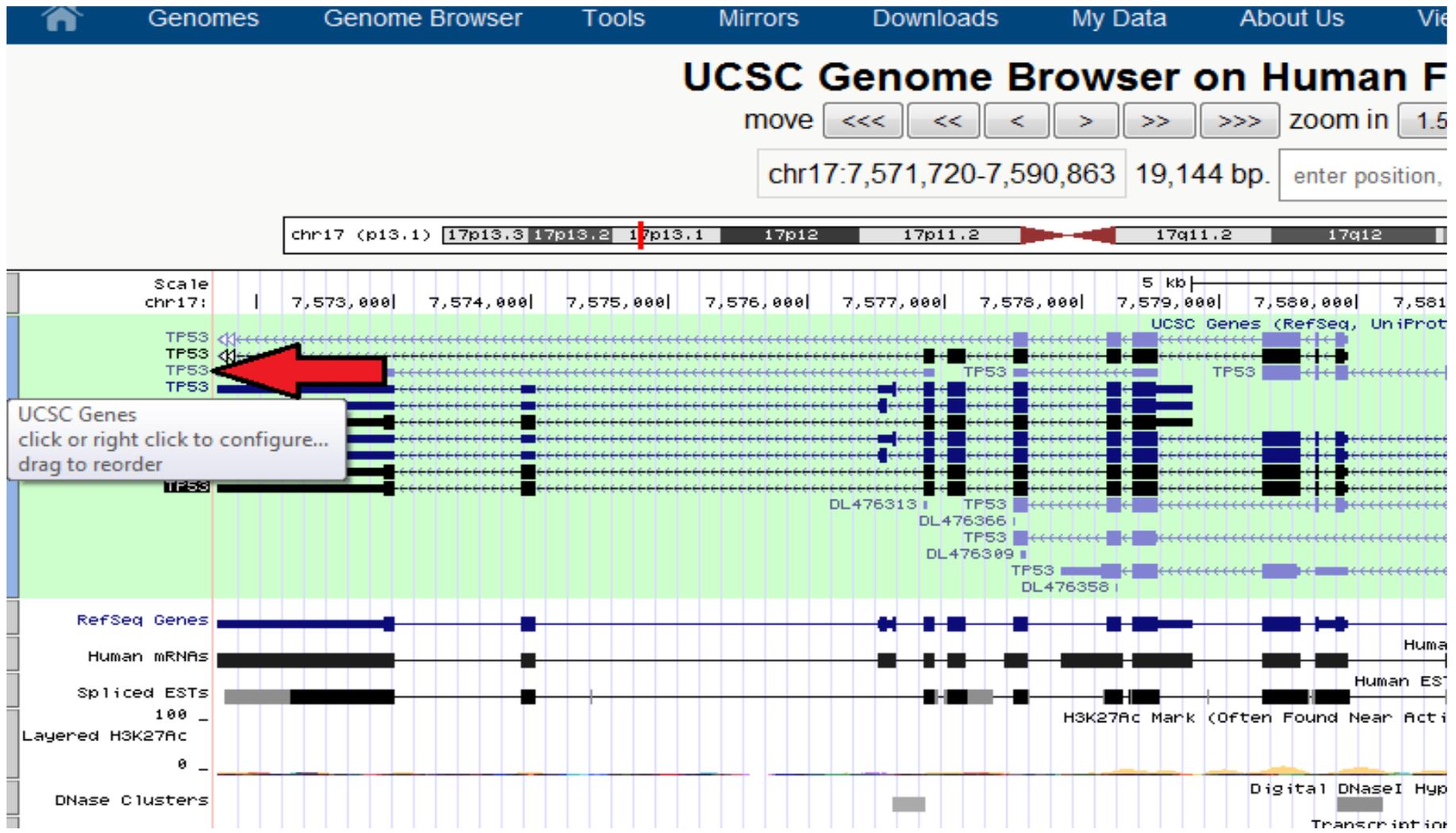
Reference sequence (GRCh37) was produced by the [Genome Reference Consortium](#). For more information see the [Genome Reference Consortium](#) website.

Identified by the accession number of a sequenced genomic clone, an mRNA or EST or STS marker, or a protein. The following list shows examples of valid position queries for the human genome. See the [UCSC Genome Browser](#) for more information.

UCSC Genome Browser Response:



On the genome viewer, click on the symbol of your gene in the “UCSC Genes” track, in this case, I clicked on the TP53. Any of those will take you to the next page. Please, note the direction of the arrows in the gene, since it indicates the direction of transcription. In this case, the 5’ end is to the right.



Click on the “Genomic sequence” link in the table called “sequence and Links to Tools and Databases”

Human Gene TP53 (uc010cng.1) Description and Page Index

Description: Homo sapiens tumor protein p53 (TP53), transcript variant 6, mRNA.

RefSeq Summary (NM_001126116): This gene encodes tumor protein p53, which responds to diverse cellular stimuli in metabolism. p53 protein is expressed at low level in normal cells and at a high level in a variety of transformed cells containing transcription activation, DNA-binding, and oligomerization domains. It is postulated to bind to a p53-binding site as a tumor suppressor. Mutants of p53 that frequently occur in a number of different human cancers fail to bind the consensus sequence not only as somatic mutations in human malignancies, but also as germline mutations in some cancer-prone families. Alternative splicing have been found. These variants encode distinct isoforms, which can regulate p53 transcriptional activity. [PubMed]

Transcription Chromosome: chr17 **Strand:** - **Size:** 7,092 **Start:** 7,571,719 **End:** 7,578,811 **Exon Count:** 6
Coding Size: 1,909 **Start:** 7,576,624 **End:** 7,578,533 **Exon Count:** 6

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

Data last updated: 2011-12-21

Sequence and Links to Tools and Databases

Genomic Sequence (chr17:7,571,720-7,578,811)	mRNA (may differ from genome)	Protein (209 aa)			
Gene Sorter	Genome Browser	Protein FASTA	Table Schema	BioGPS	CGAP
Ensembl	Entrez Gene	ExonPrimer	GeneCards	Gepis Tissue	H-INV
HGNC	Jackson Lab	MOPED	OMIM	PubMed	Reactome
Stanford SOURCE	UniProtKB	Wikipedia			

Select the "CDs in upper case, UTR in lower case" option.

Also (not indicated in the picture), you may want to select "One FASTA record per region (exon, intron, etc)" since it separates each region into different records.

Genomic Sequence Near Gene

Get Genomic Sequence Near Gene

Note: if you would prefer to get DNA for more than one feature of this track at a time, try

Sequence Retrieval Region Options:

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

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

Sequence Formatting Options:

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

submit

>hg19 knownGene uc010cng.1 range=chr17:7571720-7578811 5'pad=0 3'pad=0 strand=- repeatMasking=nor

cgagggccaggagatggaggctgcagtgagctgtgatcacaccaactgtgct
ccagcctgagtgacagagcaagaccctatctcaaaaaaaaaaaaaaaaaaa
gaaaagctcctgaggtgtagacgccaaactctctctagctcgctagtgggt
tgcaggaggtgcttacgcatgtttgtttctttgctgccgtcttccagttg
ctttatctgttcacttgtgccctgactttcaactctgtctccttctctt
cctacagtactcccctgccctcaacaagATGTTTGGCCAACTGGCCAAGA
CCTGCCCTGTGCAGCTGTGGGTTGATTCCACACCCCGCCGGCAGCCGC
GTCCGCGCCATGGCCATCTACAAGCAGTCACAGCACATGACGGAGGTTGT

GAGGCGCTGCCCCACCATGAGCGCTGCTCAGATAGCGATGGTCTGGCCC
CTCCTCAGCATCTTATCCGAGTGGAAGGAAATTTGCGTGTGGAGTATTTG
GATGACAGAAACACTTTTCGACATAGTGTGGTGGTGCCCTATGAGCCGCC
TGAGGTTGGCTCTGACTGTACCACCATCCACTACAACACTACATGTGTAACA
GTTCTGTCATGGGCGGCATGAACCGGAGGCCCATCCTCACCATCATCACA
CTGGAAGACTCCAGTGGTAATCTACTGGGACGGAACAGCTTTGAGGTGCG
TGTTTGTGCCTGTCCTGGGAGAGACCGGCGCACAGAGGAAGAGAATCTCC
GCAAGAAAGGGGAGCCTCACCACGAGCTGCCCCAGGGAGCACTAAGCGA
GCACTGCCCAACAACACCAGCTCCTCTCCCAGCCAAAGAAGAAACCACT
GGATGGAGAATATTTACCCTTCAGGACCAGACCAGCTTTCAAAAAGAAA