

# Gene (transcript) sequence retrieval from NCBI database

**Buddhi Prakash Jain**  
**Department of Zoology**  
**Mahatma Gandhi Central University Motihari**

## What is NCBI database?

The NCBI database (<http://www.ncbi.nlm.nih.gov/>) is a collection of gene, genomic sequence, transcript and proteins. It is a repositories of *GenBank* (gene sequence database), Pubmed (biomedical literature, journals, books) protein databank PDB, EST (expressed sequence tags) database, RefSeq (non redundant, linked nucleotide and protein sequence) etc.

For any gene analysis, its sequence retrieval is the first and the important step.

In this lecture we will understand how a gene (transcript) sequence can be retrieved from NCBI database for its further analysis.

1. First open <https://www.ncbi.nlm.nih.gov/pubmed/> site.

The screenshot shows a web browser window with the address bar containing [ncbi.nlm.nih.gov/pubmed/](https://www.ncbi.nlm.nih.gov/pubmed/). The browser's top bar includes navigation icons and a list of open tabs: 'Apps', 'YouTube', 'Maps', 'Gmail', and 'SA How Did Insect Met...'. The website header features the NCBI logo, 'Resources' and 'How To' dropdown menus, and a 'Sign in to NCBI' link. The main navigation area includes the 'PubMed.gov' logo, the text 'US National Library of Medicine National Institutes of Health', a search dropdown menu set to 'PubMed', a search input field, a 'Search' button, and a 'Help' link. A prominent pink banner contains a warning icon and text: 'COVID-19 is an emerging, rapidly evolving situation. Get the latest public health information from CDC: <https://www.coronavirus.gov>. Get the latest research from NIH: <https://www.nih.gov/coronavirus>.' Below this is a blue promotional banner for 'New PubMed!' with icons for a magnifying glass, a lightbulb, a globe, and a download arrow. The text reads: 'Click here to try the New PubMed! An updated version of PubMed is now available. Come see the new improvements to the interface!'. The footer section features an image of a tablet displaying a search result on a bookshelf, followed by the heading 'PubMed' and the text: 'PubMed comprises more than 30 million citations for biomedical literature from MEDLINE, life science journals, and online books. Citations may include links to full-text content from PubMed Central and publisher web sites.'

2. On left upper side select gene. Type name of gene or gene symbol and organism name.

For example if we want to retrieve transcript sequence of p53 from *Homo sapiens*

The screenshot shows the NCBI PubMed website interface. At the top, the URL is [ncbi.nlm.nih.gov/pubmed/](https://ncbi.nlm.nih.gov/pubmed/). The page features the NCBI logo and navigation links for Resources and How To. A search bar contains the text "p53 homo sapiens" and a "Search" button. A dropdown menu is open over the search bar, listing various database categories under "Recent" and "All". The "Gene" option is highlighted in blue. Below the search bar, there is a red banner with COVID-19 information and a blue banner promoting the "New PubMed!" interface. At the bottom, a dark blue banner provides information about the PubMed database.

NCBI Resources How To Sign in to NCBI

PubMed.gov  
US National Library of Medicine  
National Institutes of Health

PubMed p53 homo sapiens Search

Recent  
Gene  
Nucleotide  
All Databases  
PubMed

All  
All Databases  
Assembly  
Biocollections  
BioProject  
BioSample  
BioSystems  
Books  
ClinVar  
Conserved Domains  
dbGaP  
dbVar  
Gene  
Genome  
GEO DataSets

COVID-19 is an emerging, rapidly evolving situation.  
Get the latest public health information from CDC: <https://www.coronavirus.gov>.  
Get the latest research from NIH: <https://www.nih.gov/coronavirus>.

Click here to try the  
**New PubMed!**

An updated version of PubMed is now available.  
Come see the new improvements to the interface!

**PubMed**  
PubMed comprises more than 30 million citations for biomedical literature from MEDLINE, life science journals, and online books. Citations may include links to full-text content from PubMed Central and publisher web sites.

### 3. Click on TP53 gene in the window.

ncbi.nlm.nih.gov/gene/?term=p53+homo+sapiens

Search results

Items: 1 to 20 of 2376

See also 73 discontinued or replaced items.

Name/Gene ID	Description	Location	Aliases	MIM
<input type="checkbox"/> <a href="#">TP53</a> ID: 7157	tumor protein p53 [ <i>Homo sapiens</i> (human)]	Chromosome 17, NC_000017.11 (7668402..7687550, complement)	BCC7, BMFS5, LFS1, P53, TRP53	191170
<input type="checkbox"/> <a href="#">MDM2</a> ID: 4193	MDM2 proto-oncogene [ <i>Homo sapiens</i> (human)]	Chromosome 12, NC_000012.12 (68808149..68850686)	ACTFS, HDMX, LSKB, hdm2	164785

Recent activity

- p53 homo sapiens AND (alive[prop]) (2376) Gene
- STRN3 striatin 3 [*Homo sapiens*] Gene



#### TP53 tumor protein p53 [ *Homo sapiens* (human) ]

Gene ID: 7157, updated on 29-Mar-2020

##### Summary

**Official Symbol** TP53 provided by [HGNC](#)

**Official Full Name** tumor protein p53 provided by [HGNC](#)

**Primary source** [HGNC:HGNC:11998](#)

**See related** [Ensembl:ENSG00000141510](#) [MIM:191170](#)

**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** P53; BCC7; LFS1; BMFS5; TRP53

**Summary** This gene encodes a tumor suppressor protein containing transcriptional activation, DNA binding, and oligomerization domains. The encoded protein responds to diverse cellular stresses to regulate expression of target genes, thereby inducing cell cycle arrest.

##### Table of contents

- Summary
- Genomic context
- Genomic regions, transcripts, ar
- Expression
- Bibliography
- Phenotypes
- Variation
- HIV-1 interactions
- Pathways from PubChem
- Interactions
- General gene information
  - Markers, Clone Names, Homol

4. In the window go down in the mRNA and proteins section where different isoforms of the gene are shown. In case of p53 there are 15 isoforms are present.

Now click on NM\_000546.6 to get nucleotide sequence and NP\_000537.3 for protein sequence.

We are analysing the nucleotide sequence of its first isoform only

*mRNA and Protein(s)*

1. NM\_000546.6 → NP\_000537.3 cellular tumor antigen p53 isoform a  
[See identical proteins and their annotated locations for NP\\_000537.3](#)

Status: REVIEWED

Description	Transcript Variant: This variant (1) can initiate translation from two in-frame AUG start codons. The isoform represented in this variant (a, also known as p53alpha) results from translation initiation at the upstream start codon. Both variants 1 and 2 encode isoform a, which is the longest isoform.	
Source sequence(s)	<a href="#">AK223026</a> , <a href="#">DA453049</a> , <a href="#">X02469</a>	
Consensus CDS	<a href="#">CCDS11118.1</a>	
UniProtKB/Swiss-Prot	<a href="#">P04637</a>	
UniProtKB/TrEMBL	<a href="#">K7PPA8</a> , <a href="#">Q53GA5</a>	
Related	<a href="#">ENSP00000269305.4</a> , <a href="#">ENST00000269305.8</a>	

Conserved Domains (3) [summary](#)

<a href="#">pfam00870</a> Location:95 → 289	P53; P53 DNA-binding domain
<a href="#">pfam07710</a> Location:319 → 358	P53_tetramer; P53 tetramerisation motif
<a href="#">pfam08563</a> Location:5 → 28	P53_TAD; P53 transactivation motif

2. NM\_001126112.2 → NP\_001119584.1 cellular tumor antigen p53 isoform a  
[See identical proteins and their annotated locations for NP\\_001119584.1](#)

Status: RFVIFWFD

5. After clicking on NM\_000546.6 the following window will open.

GenBank Send to: ▾

## Homo sapiens tumor protein p53 (TP53), transcript variant 1, mRNA

NCBI Reference Sequence: NM\_000546.6

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

---

Go to: ▾

LOCUS NM\_000546 2512 bp mRNA linear PRI 13-FEB-2020  
DEFINITION Homo sapiens tumor protein p53 (TP53), transcript variant 1, mRNA.  
ACCESSION NM\_000546  
VERSION NM\_000546.6  
KEYWORDS RefSeq; RefSeq Select.  
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 2512)  
AUTHORS Marcel V, Tran PL, Sagne C, Martel-Planche G, Vaslin L,  
Teulade-Fichou MP, Hall J, Mergny JL, Hainaut P and Van Dyck E.  
TITLE G-quadruplex structures in TP53 intron 3: role in alternative  
splicing and in production of p53 mRNA isoforms  
JOURNAL Carcinogenesis 32 (3), 271-278 (2011)

Change region shown ▾

Customize view ▾

---

Analyze this sequence ▾

Run BLAST

Pick Primers

Highlight Sequence Features

Find in this Sequence

Show in Genome Data Viewer

---

Articles about the TP53 gene ▾

COUP-TFII Overexpression Inhibits Cell Proliferation and Invasion [Anticancer Res. 2020]

Genetic polymorphisms of TP53 (rs1042522) and MDM2 (rs2279744) and colorectal c [Gene. 2020]

6. By clicking on FASTA, the sequence of the complete transcript (CDS, 5' UTR and 3' UTR) will open.

Select the sequence and copy paste it on a word file.

FASTA ▾

Send to: ▾

Change region shown ▾

Customize view ▾

### Homo sapiens tumor protein p53 (TP53), transcript variant 1, mRNA

NCBI Reference Sequence: NM\_000546.6

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

>NM\_000546.6 Homo sapiens tumor protein p53 (TP53), transcript variant 1, mRNA

```
CTCAAAAGTCTAGAGCCACCGTCCAGGGAGCAGGTAGCTGGGCTCCGGGACACTTTGCGTTCGGC
TGGGAGCGTGTCTTCCACGACGGTGACACGCTTCCC TGGATTGGCAGCCAGACTGCCTTCCGGGCACTG
CCATGGAGGAGCCGAGTCAAGTCC TAGCGTCCAGCCCCCTCTGAGTCAAGAAACATTTTCAGACCTATG
GAAACTACTTCTGAAAACAACGTTCTGTCCCCCTTCCCGTCCCAAGCAATGGATGATTTGATGCTGTCC
CAGGACGATATTGAACAATGGTTCACTGAAGACCCAGGTCCAGATGAAGCTCCAGAAATGCCAGAGGCTG
CTCCCCCGTGGCCCCGACCAAGCAGCTCTACACCGGCGGCCCTGCACCAAGCCCCCTCTGGCCCCCT
GTCATCTTCTGTCCCTTCCAGAAAACCTACCAAGGCGAGCTACGGTTTCCGCTCTGGGCTTCTTGCACTT
GGGACAGCAAGTCTGTACTGCACGTACTCCCTGCCCTCAACAAGATGTTTGCACCAAGTGGCCAAGA
CCTGCCCTGTGCAGCTGTGGGTTGATTCACACCCCCGCCCGCACCCGCGCTCCGCGCCATGGCCATCTA
CAAGCAGTACAGCAGTACGCGAGGTTGTGAGGCGCTGCCCCACCATGAGCGCTGCTCAGATAGCGAT
GGTCTGGCCCCCTCCCTCAGCATCTTATCCGAGTGGAAAGGAAATTTGCGTGTGGAGTATTTGGATGACAGAA
ACACTTTTCGACATAGTGTGGTGGTCCCTATGAGCGCCTGAGGTGGCTGACTGTACCACCATCCA
CTAAACACATGTGTAAACAGTTCTGATGGCGGCGATGAACCGGAGGCCATCTCACCATCATCACA
CTGGAAGACTCCAGTGGTAACTACTGGGACGGAACAGCTTGGAGTGGCTGTTTGTGCTGCTGGGA
GAGACCGGCGCACAGAGGAAGAGAACTCCGCAAGAAAGGGGAGCCTCACCAAGCTGCCCCAGGGAG
CACTAAGCGAGCACTGCCCAACAACACCAAGCTCCTCTCCCCAGCCAAGAAGAAACCATGGATGGAGAA
TATTTCAACCTTCAGATCCGTGGGCGTGGGCGCTTCGAGATGTTCCGAGAGCTGAATGAGGCCTTGGAA
TCAAGGATGCCCAGGCTGGGAAGGAGCCAGGGGGGAGCAGGGCTCACTCCAGCCACTGAAGTCAAAAA
GGGTGAGTCACTCCCGCCATAAAAACTCATGTTCAAGACAGAAGGGCCTGACTCAGACTGACATCT
CCACTTCTGTTCACCACTGACAGCTCCACCCCATCTCTCCCTCCCTGCCATTTTGGGTTTGGGT
CTTTGAACCTTGCTTGAATAGGTGTGCGTCAGAAGCACCCAGGACTTCCATTTGCTTTGCTCCCGGG
TCCACTGAAACAAGTTGGCTGCACTGTTTGTGTTGGGGAGGAGGATGGGGAGTAGGACATACCAGC
TTAGATTTAAGGTTTTTACTGTGAGGGATGTTGGGAGATGTAAGAAATGTTCTTCAGTTAAGGGTTA
GTTTACAATCAGCCACACTTCTAGGTAGGGGCCACTTCACCGTACTAACAGGGAAGCTGCCCCACTG
TTGAATTTCTCTAACTCAAGGCCATATCTGTGAAATGCTGGCATTGCACTACCTCACAGAGTGCA
TTGTGAGGGTTAATGAAATAATGTACTCTGGCCTTGAACCACTTTTATACATGGGGTCTAGAATT
GACCCCTTGAGGGTGTCTGTTCCCTCTCCCTGTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT
TGGTAGGTAGAGGGAGTTGTCAAGTCTCTGCTGGCCAGCCAAACCTGCTGACAACTCTTGGTGA
CCTTAGTACCTAAAAGGAAATCTCACCCATCCACACCCCTGGAGGATTTATCTCTTTGATATGATGAT
CTGGATCCACCAAGACTGTTTTATGCTCAGGGTCAATTTCTTTTTCTTTTTTTTTTTTTTTTTTTCTTT
TTCTTTGAGACTGGGCTCGCTTTGTTGCCAGGCTGGAGTGGAGTGGCGTATCTTGGCTACTGCAGC
CTTTGCCCTCCCCGGCTCAGCAGTCTGCTCAGCTCCGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGT
ATGGCCAGCAACTTTTGCATGTTTGTAGAGATGGGCTCTCACAGTGTGCCAGGCTGGTCTCAAAC
CTTGGGCTCAGGGGATCCACTGTCTCAGCTCCAGAGTGGGATTACAATTTGAGGACACACGCTC
CAGCTGGAAGGGTCAACTCTTTTACATCTGCAAGCAGTCTGCAATTTTCAACCCACCTTCCCTCT
CTCCCTTTTATATCCATTTTATATCGATCTTATTTTACAATAAACTTTGCTGCCA
```

**Analyze this sequence** ▾

- Run BLAST
- Pick Primers
- Highlight Sequence Features
- Find in this Sequence
- Show in Genome Data Viewer

**Articles about the TP53 gene** ▾

- COUP-TFII Overexpression Inhibits Cell Proliferation and Invasion [Anticancer Res. 2020]
- Genetic polymorphisms of TP53 (rs1042522) and MDM2 (rs2279744) and colorectal c [Gene. 2020]
- BRCA1/P53: Two strengths in cancer chemoprevention [Biochim Biophys Acta Rev Canc...

See all...

**Pathways for the TP53 gene** ▾

- Ferroptosis
- Mitophagy - animal
- Fluid shear stress and atherosclerosis

See all...

**Reference sequence information** ▾

- RefSeq alternative splicing
- See 15 reference mRNA sequence splice

## 7. Identify CDS (CoDing sequence), 5' UTR (untranslated region), and 3' UTR.

Go back in the step 5 window. Scroll down and click on CDS.

```
exon      /db_xref="MIM:191170"
1..114
/gene="TP53"
/gene_synonym="BCC7; BMFS5; LFS1; P53; TRP53"
/inference="alignment:Splign:2.1.0"
misc_feature 35..37
/gene="TP53"
/gene_synonym="BCC7; BMFS5; LFS1; P53; TRP53"
/note="upstream in-frame stop codon"
exon      115..216
/gene="TP53"
/gene_synonym="BCC7; BMFS5; LFS1; P53; TRP53"
/inference="alignment:Splign:2.1.0"
CDS       143..1324
/gene="TP53"
/gene_synonym="BCC7; BMFS5; LFS1; P53; TRP53"
/note="isoform a is encoded by transcript variant 1; tumor
protein 53; mutant tumor protein 53; cellular tumor
antigen p53; phosphoprotein p53; transformation-related
protein 53; p53 tumor suppressor; antigen NY-CO-13; tumor
suppressor p53"
/codon_start=1
/product="cellular tumor antigen p53 isoform a"
/protein_id="NP_000537.3"
/db_xref="CCDS:CCDS11118.1"
/db_xref="GeneID:7157"
/db_xref="HGNC:HGNC:11998"
/db_xref="MIM:191170"
/translation="MEEPQSDPSVEPPLSQETFSDLWKLLPENNVLSPLPSQAMDDL
LSPDDIEQWFTEDPGPDEAPRMPEAAPVAPAPAAPTAAAPAPAPSWPLSSVPSQKT
YQGSYGFRLGFLHSGTAKSVTCTYSPALNKMFCQLAKTQPVQLWVDSTPPPGTRVRAM
AIYKQSQHMTVEVRRCPHHERCSDSGLAPPQHLIRVEGNLRVEYLDDRRNTFRHSVVV
PYEPPEVGSDCITTIHYNMCMNSSCMGGMNRRPILTIITLEDSSGNLLGRNSFEVRVCA
CPGRDRRTEENLRKKGEPHHELPPGSTKRALPNNTSSSPQPKKPLDGEYFTLQIRG
RERFEMFRELNEALELKDAQAGKEPGGSSRAHSSHLKSKKGQSTSRHKKLMFKTEGPD
D"
misc_feature 143..1102
/gene="TP53"
/gene_synonym="BCC7; BMFS5; LFS1; P53; TRP53"
/experiment="experimental evidence, no additional details
recorded"
/note="annotated from UniProtKB/Swiss-Prot (P04637.4)";
```

8. After clicking on CDS, the CDS region in the complete transcript will be highlighted. The CDS starts with start codon (AUG) and ends with stop codon. The sequence before the CDS is called as 5' UTR and after CDS region is called as 3' UTR.

Now select the CDS region as highlighted and mark this sequence in the complete transcript sequence as pasted earlier in the word file.

Like this CDS, 5' UTR and 3' UTR sequence of any gene or transcript can be retrieved for further analysis, cloning and other purposes .

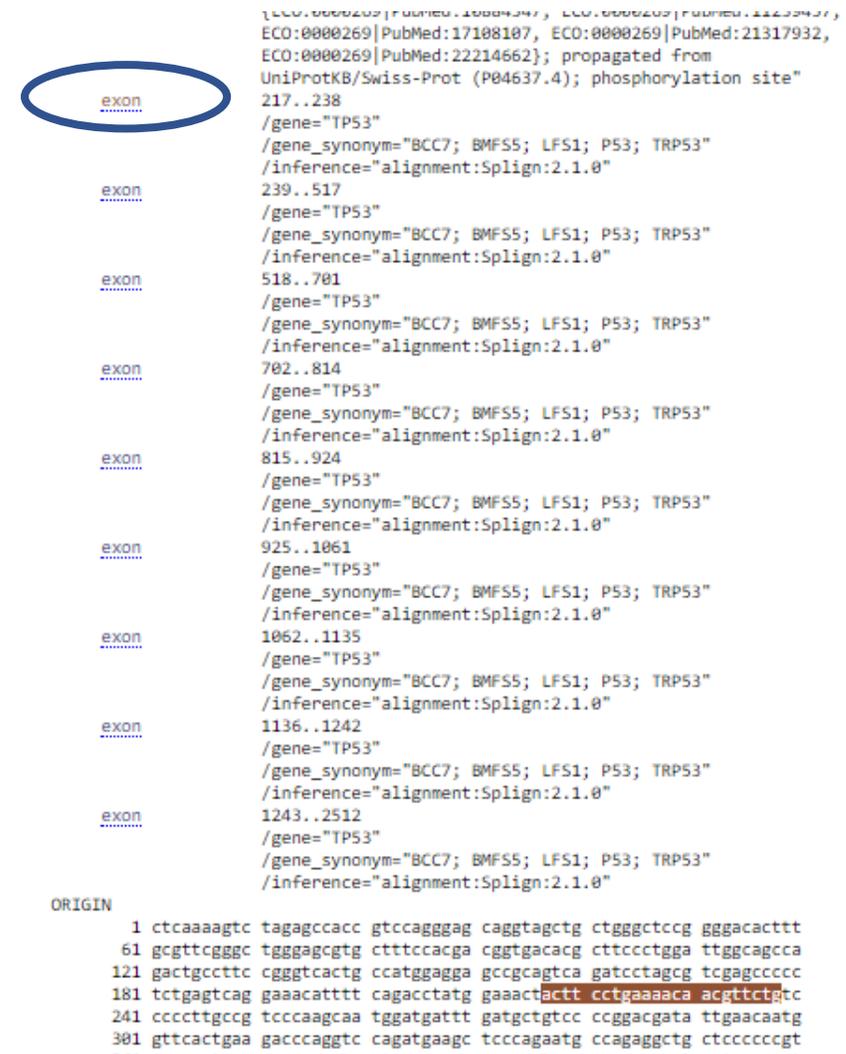
```
/gene_synonym="BCC7; BMF55; LFS1; P53; TRP53"  
/inference="alignment:Splign:2.1.0"  
  
ORIGIN  
1 ctcaaaagtc tagagccacc gttccaggag caggtagctg ctgggctccg gggacacttt  
61 gcgttcgggc tgggagcgtg ctttccacga cggtgacacg cttccctgga ttggcagcca  
121 cactgccttc cgggtcactg ccatggagga gccgcagtca gatcctagcg tcgagccccc  
181 tctgagtcag gaaacatttt cagacctatg gaaactactt cctgaaaaca acgttctgtc  
241 ccccttgcag tcccaagcaa tggatgattt gatgctgtcc ccgacgata ttgaaacaatg  
301 gttcactgaa gaccaggtc cagatgaagc tcccagaatg ccagaggctg ctccccctgt  
361 ggccccctga ccagcagctc ctacaccggc ggccccctga ccagccccct cctggcccc  
421 gtcactctct gtccttccc agaaaaccta ccagggcagc tacggtttcc gtcctggctt  
481 cttgcattct gggacagcca agtctgtgac ttgcacgtac tcccctgccc tcaacaagat  
541 gttttgccaa ctggccaaga cctgccctgt gcagctgtgg ttgattcca cacccccgcc  
601 cggcaccctc gtcctgcca tggccatcta caagcagtc cagcacatga cggaggttgt  
661 gagcgcctgc ccccaccatg agcctgctc agatagcagc ggtctgccc ctccctagca  
721 tcttatccga gtggaaggaa atttgcgtgt gtagtatttg gatgacagaa acactttctg  
781 acatagtgtg ttggtgccct atgagccgcc tgaggttggc tctgactgta ccaccatcca  
841 ctacaactac atgtgtaaca gttcctgcat gggcggcatg aaccggaggc ccactctcac  
901 catcatcaca ctggaagact ccagtgttaa tctactgga cggaaacagc ttgaggtcgg  
961 tgtttgtccc gtcctggga gagaccggc cacagaggaa gagaatctcc gcaagaaagg  
1021 ggagcctcac cagcagctgc ccccaggag cactaagcga gcactgccc acaaccacag  
1081 ctccctccc cagccaaga agaaacctt gtaggggaa tatttcacc ttcatagctg  
1141 tggcctgag cgttcgaga gttccgaga gctgaatgag gccttggaac tcaaggatgc  
1201 ccaggctggg aaggagccag gggggagcag ggtcactcc agccacctga agtccaaaaa  
1261 ggtcagctc acctcccgc ataaaaaact catgttaag acagaagggc ctgactcaga  
1321 ctgacattct ccacttttg ttccccactg acagcctccc acccccactc ctccctccc  
1381 tgccattttg ggtttgggt ctttgaacc ttgcttgcaa taggtgtcgc tcagaagcac  
1441 ccaggacttc catttgcttt gtcgggggc tccactgaac aagttggcct gcactgggtg  
1501 tttgtttggt gaggaggat ggggagtagg acataaccagc ttgatttta agtttttac  
1561 tgtgaggat gtttgggaga gtaagaaat gttcttcag ttaagggtta gtttacaatc  
1621 agccacattc taggtagggg cccacttcac cgtactaacc agggaaagctg tccctcactg  
1681 ttgaattttc tctaaactca aggccatata ctgtgaaatg ctggcatttg cacctacctc  
1741 acagagtgca ttgtgagggt taatgaaata atgtacatct ggccttgaaa caccctttta  
1801 ttacatgggg tctagaactt gacccccttg aggggtgctg ttccctctcc ctggttgctg  
1861 gttgggttgt agtttctaca gttgggcagc tgggttaggta gaggaggttg tcaagtctct  
1921 gctggcccca ccaaaccctg tctgacaacc tcttgggtgaa ccttagtacc taaaaggaaa  
1981 tctcaccctc tcccacacc tggaggattt catctctgt atatgatgat ctggatccac  
2041 caagacttgt tttatgctca ggttcaattt ctttttctt tttttttt tttttctt  
2101 ttctttgaga ctgggtctg ctttgttgc caggctggag tggagtgcc tgactctggc  
2161 ttactgcagc ctttgcctcc ccggctcag cagtcctgcc tcagcctcc gagtagctgg  
2221 gaccacaggt tcatgccacc atggccagcc aacttttga tgtttgtag agatgggtc  
2281 tcacagtgtt gccaggtctg gttctaaact cctgggctca ggcgatccac ctgtctcagc  
2341 ctcccagagt gctgggatta caattgtgag ccaccacgct cagctggaag ggtcaacatc  
2401 ttttacattc tgcaagcaca tctgcatttt caccccacc ttcccctct tttcccttt  
2461 tatatcccat ttttatatc atctcttatt ttacaataaa actttgtctc ca
```

## 9. Number of exons and their sequence can also be studied.

In case of first isoform of p53 (NM\_000546.6), there are nine exons are present.

The corresponding sequence will be highlighted when the specific exon is clicked.

These exon sequence can be marked in the original sequence pasted on word file.



```
ECO:0000269|PubMed:1688434, ECO:0000269|PubMed:1123437,
ECO:0000269|PubMed:17108107, ECO:0000269|PubMed:21317932,
ECO:0000269|PubMed:22214662}; propagated from
UniProtKB/Swiss-Prot (P04637.4); phosphorylation site"
217..238
/exon="TP53"
/gene_synonym="BCC7; BMFS5; LFS1; P53; TRP53"
/inference="alignment:Splign:2.1.0"
239..517
/exon="TP53"
/gene_synonym="BCC7; BMFS5; LFS1; P53; TRP53"
/inference="alignment:Splign:2.1.0"
518..701
/exon="TP53"
/gene_synonym="BCC7; BMFS5; LFS1; P53; TRP53"
/inference="alignment:Splign:2.1.0"
702..814
/exon="TP53"
/gene_synonym="BCC7; BMFS5; LFS1; P53; TRP53"
/inference="alignment:Splign:2.1.0"
815..924
/exon="TP53"
/gene_synonym="BCC7; BMFS5; LFS1; P53; TRP53"
/inference="alignment:Splign:2.1.0"
925..1061
/exon="TP53"
/gene_synonym="BCC7; BMFS5; LFS1; P53; TRP53"
/inference="alignment:Splign:2.1.0"
1062..1135
/exon="TP53"
/gene_synonym="BCC7; BMFS5; LFS1; P53; TRP53"
/inference="alignment:Splign:2.1.0"
1136..1242
/exon="TP53"
/gene_synonym="BCC7; BMFS5; LFS1; P53; TRP53"
/inference="alignment:Splign:2.1.0"
1243..2512
/exon="TP53"
/gene_synonym="BCC7; BMFS5; LFS1; P53; TRP53"
/inference="alignment:Splign:2.1.0"
ORIGIN
1 ctcaaaagtc tagagccacc gtccaggagg caggtagctg ctgggctccg gggacacttt
61 gcgttcgggc tgggagcgtg ctttccacga cggtgacacg cttccctgga ttggcagcca
121 gactgccttc cgggtcactg ccatggagga gccgcagtca gatcctagcg tcgagccccc
181 tctgagtcag gaaacatttt cagacctatg gaaactactt cctgaaaaaca acattcttctc
241 ccccttgccg tcccaagcaa tggatgattt gatgctgtcc ccggacgata ttgaacaatg
301 gttcactgaa gacccaggtc cagatgaagc tcccagaatg ccagaggctg ctccccccgt
```

# Summary

From NCBI database, the isoforms of any gene, their complete sequence can be retrieved for further study.

## Homo sapiens tumor protein p53 (TP53), transcript variant 1, mRNA

NCBI Reference Sequence: NM\_000546.6

[GenBank Graphics](#)

>NM\_000546.6 Homo sapiens ~~tumor~~ protein p53 (TP53), transcript variant 1, mRNA

```
CTCAAAAGCTAGAGCCACCGTCCAGGGACAGGTAGCTGCTGGGCTCCGGGACACTTTCCGTTGGGGC
TGGAGCGTGTCTTCCACGACGGTGCACAGCTTCCCTGGATTGGCAGCCAGACTGCCCTTCCGGGTCACTG
CCATGGAGGAGCCCGCAGTCAGATCCTTAGGCTCGAGCCCCCTCTGAGTCAGGAAACATTTTCAGACCTATG
GAAACTACTTCCGAAUACACCTCTCTGTCCTCCCTTGGCGTCCCAAGCAATGGATGATTTGATGCTGTC
CCGACGATATTGAACAATGGTTCACGAAAGACCCAGTCCAGATGAAGCTCCCAAAATGCCAGAGCTG
CTCCCCCGTGGCCCTGCACAGCAGCTCCTACACCGGGCGCCCTGCACAGCCCTCCCTGGCCCTT
GGCATCTCTGTCCTTCCCAAAUACCTACCAAGGACAGCTACGGTTTCCGCTGGGGCTTCTTGCATTC
GGCAGCCCAAGTCTGTGACTTGCACGTACTCCCTGCCCCCAACAAGATGTTTGGCAACTGGCCAGG
CCTGCCCTGTGCAGCTGTGGGTTGATTCACACCCCGCCCGGCAACCGGCGTCCGGCCATGGCCATCTA
CAAGCAGTCACAGCACAAGCCAGGTTGTGAGGCGCTGCCCCCACTAGCCCTGCTCAGATAGCGAT
GGCTGGCCCTCCCTCAGCATCTTACCGAGTGGAGGAAATTTGCGTGTGGAGTATTTGGATGACAGAA
ACACTTTTGACATAGTGTGGTGGTGCCTATGAGCCGCTGAGGTTGGCTGTGACTGTACACCATCCG
CTACAACACATGTGTAAACAGTTCCTGCAAGGGCCATGAACCGGAGGCCATCCCTCACCACATCAACA
CTGGAGACTCCAGTGGTAATCTACTGGGACGGAAACAGCTTGGAGTGGCTGTGTGTGCTGTCTCTGGGA
GAGACCGGCGACAGAGGAGAGAAATCCCGCAAGAAAGGGGAGCTCACACAGCTGCCCCACAGGG
CACTAAGCGAGCACTGCCAAACACACCACTCCCTCTCCCAAGCCAAAGAGAAACCACTGGATGGAGAA
TATTTCAACCTTCAGATCCGTGGGCGTGGAGGCTTCGAGATGTTCGAGAGCTGATGAGGCGTTGGAA
TCAAGATGCCAGGCTGGGAGGAGCCAGGGGGAGCAGGGCTCACTCCAGCCCTGAACTCCAAA
GGGTCAGTCTACTCCCGCCATAAAAAACTCATGTTCAAGACAGAAAGGGGCTGACTCAGACTGCAATTCT
CCACTCTGTGTTCCCCACTGCAGCCCTCCACCCCGCATCTCTCCCTCCCTGCCCATTGGGGTTTGGGT
CTTTGAACCCCTTGGTTCAGTGGGTGTGGCTCAGAGCACCCAGGACTTCCATTGGCTTGTGCTGGGGC
TCCACTGAAACAAGTTGGCTCCACTGGTGTTTTGTGTGGGAGGAGGATGGGGATAGGACATACCAAC
TTAGATTTAAGGTTTTTACTGTGAGGGATGTTGGGAGATGTAAGAAATGTTCTTCAGTTAAGGGTTA
GTTTACAATCAGCCACATCTAGGTAGGGGCCACTTCCACCTACCAACAGGGAGGCTGTCCCTCACTG
TTGAAITTTCTAATCTCAAGGCCATACTGTGAAATGCTGGCAITTGCACTTACCTCACAGATGCA
TTGTGAGGGTTAATGAAATAATGTACATCTGGCCCTGAAACACCTTTTATTACATGGGGCTCAGACTT
GACCCCTTGGAGGTTGCTTGTTCCTCTCCCTGTGGTGGGTTGGTAGTTCTACAGTTGGGCGAGC
TGGTTAGGTAGAGGGAGTTGTCAAGTCTCTGCTGGCCAGCCAAACCTGTCTGCAACCTCTTGGTGA
CCTTAGTACCTAAAAGGAAATCTCACCCCATCCACACCTGGAGGATTTGATCTCTTGTATATGATGAT
CTGGAATCCACDAAGACTTGTATATGCTCAGGGTCAAATTCCTTTTCTTTTCTTTTCTTTTCTTTT
TCTTTGAGACTGGGTTCTGGCTTGTGGCCAGGCTGGAGTGGAGTGGCGTGAATCTTGGCTTACTGCAAGC
CTTTGCTCCCGGCTGAGCAGTCTTGGCTCAGCTCCGGAGTAGCTGGGACCAAGGTTTCATGCTGATATGATGAT
ATGGCCAGCCAACTTTGCAATGTTTGTAGAGATGGGGTCTCACAAGTGTGGCCAGGCTGGTCTCAAACT
CCTGGGCTCAGGGGATCCACCTTGTCTCAGCTTCCAGAGTGTCTGGGATTAACAATGTGTAGGCCACCAAGTCTC
CAGCTGGAAAGGTTCAACATCTTTACATCTGCAAGCACAATCTGCAATTTTACCCCACTCTCCCTCTCT
TCTCCCTTTTATATCCCATTTTATATGATCTCTTATTTTACAATAAACTTGTCTGCA
```

## References:

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

THANK  
YOU