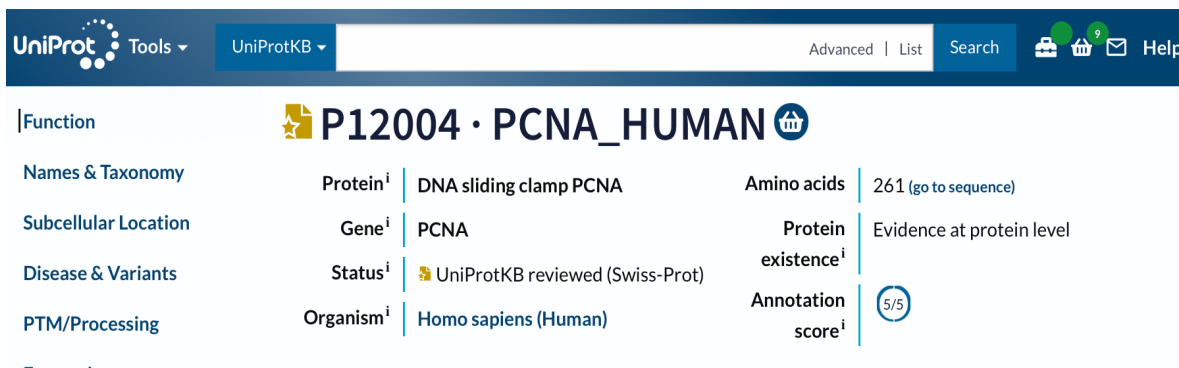


# NALOGA 1 - UniProt

1. V UniProt vpišeš PCNA, pod filtri označiš Homo sapiens in kot prvi zadetek dobiš:

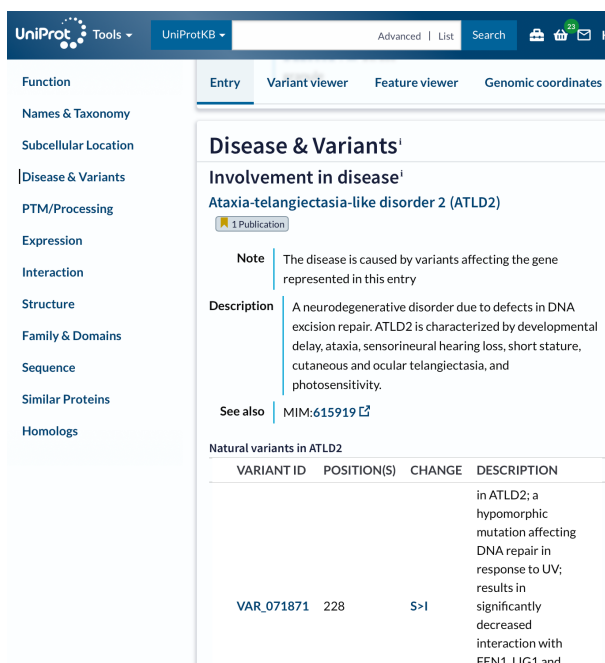


The screenshot shows the UniProt entry for P12004 · PCNA\_HUMAN. The header includes the UniProt logo, navigation tools, and a search bar. The main content is organized into a table with categories on the left and details on the right.

Function	P12004 · PCNA_HUMAN	
Names & Taxonomy	Protein <sup>i</sup>   DNA sliding clamp PCNA	Amino acids   261 (go to sequence)
Subcellular Location	Gene <sup>i</sup>   PCNA	Protein existence <sup>i</sup>   Evidence at protein level
Disease & Variants	Status <sup>i</sup>   UniProtKB reviewed (Swiss-Prot)	Annotation score <sup>i</sup>   5/5
PTM/Processing	Organism <sup>i</sup>   Homo sapiens (Human)	

KODA: P12004

2. PCNA vsebuje 261 aminokislin, molekulska masa je 28.769 Da (oba podatka pišeta pod razdelkom Sequence)
3. Vsebuje interakcijsko regijo: Interaction with NUDT15, ki vsebuje kar nekaj modificiranih aminokislinskih ostankov. (Regijo najdeš pod Family & Domains, potem pa glede na lokacijo regije pogledaš pod PTM in preveriš modifikacije.)
4. Bolezen, ki je neposredno povezana s proteinom PCNA je ATLD2. To je nevrodegenerativna motnja zaradi okvar v popraviljanju DNA z izrezovanjem. Zanj so značilni razvojni zaostanek, ataksija, senzorinevralna izguba sluha, nizka rast, kožne in očesne teleangiektazije ter občutljivost na svetlobo.



The screenshot shows the 'Disease & Variants' section of the UniProt entry for P12004 · PCNA\_HUMAN. It includes a table of natural variants in ATLD2.

**Disease & Variants<sup>1</sup>**  
**Involvement in disease<sup>1</sup>**  
**Ataxia-telangiectasia-like disorder 2 (ATLD2)**  
1 Publication

**Note** | The disease is caused by variants affecting the gene represented in this entry

**Description** | A neurodegenerative disorder due to defects in DNA excision repair. ATLD2 is characterized by developmental delay, ataxia, sensorineural hearing loss, short stature, cutaneous and ocular telangiectasia, and photosensitivity.

**See also** | MIM:615919 [↗](#)

**Natural variants in ATLD2**

VARIANT ID	POSITION(S)	CHANGE	DESCRIPTION
VAR_071871	228	S>I	in ATLD2; a hypomorphic mutation affecting DNA repair in response to UV; results in significantly decreased interaction with FEN1, IIC1 and

## NALOGA 2 – GenBank

1. V GenBank vpišeš ime proteina in Homo Sapiens ter pod zadetki najdeš transkriptno varianto 2.

NM\_182649.2

2. 6 (Pod FEATURES prešteješ vse eksone)
- 3.

```

16-1271          BC000491.2          1-1256
1272-1276        BU626265.1          18-22          c
FEATURES
  source          1..1276
                  /organism="Homo sapiens"
                  /mol_type="mRNA"
                  /db_xref="taxon:9606"
                  /chromosome="20"
                  /map="20p12.3"
  gene            1..1276
                  /gene="PCNA"
                  /gene_synonym="ATLD2"
                  /note="proliferating cell nuclear antigen"
                  /db_xref="GeneID:5111"
                  /db_xref="HGNC:HGNC:8729"
                  /db_xref="MIM:176740"
  exon            1..381
                  /gene="PCNA"
                  /gene_synonym="ATLD2"
                  /inference="alignment:Splign:2.1.0"
  misc_feature    134..136
                  /gene="PCNA"
                  /gene_synonym="ATLD2"
                  /note="upstream in-frame stop codon"
  CDS             161..946
                  /gene="PCNA"
                  /gene_synonym="ATLD2"

```

Na 20. kromosomu.

4. Deluje kot homotrimer. Njegovi psevdogeni so bili opisani na kromosomu 4 in kromosomu X.

Psevdogeni so nefunkcionalne kopije ali ostanki nekdanjih genov na kromosomih, ki so ob evlucijskih mutacijah izgubili zmožnost izražanja.

```

TITLE            Structure of the human gene for the proliferating cell nuclear
                antigen
JOURNAL          J Biol Chem 264 (13), 7466-7472 (1989)
PUBMED           2565339
COMMENT          REVIEWED REFSEQ: This record has been curated by NCBI staff. The
                reference sequence was derived from AK313286.1, BC000491.2 and
                BU626265.1.
                This sequence is a reference standard in the RefSeqGene project.
                On Nov 28, 2018 this sequence version replaced NM_182649.1.
                Summary: The protein encoded by this gene is found in the nucleus
                and is a cofactor of DNA polymerase delta. The encoded protein acts
                as a homotrimer and helps increase the processivity of leading
                strand synthesis during DNA replication. In response to DNA damage,
                this protein is ubiquitinated and is involved in the RAD6-dependent
                DNA repair pathway. Two transcript variants encoding the same
                protein have been found for this gene. Pseudogenes of this gene
                have been described on chromosome 4 and on the X chromosome.
                [provided by RefSeq, Jul 2008].
                Transcript Variant: This variant (2) differs in the 5' UTR compared
                to variant 1. Variants 1 and 2 encode the same protein.
                Publication Note: This RefSeq record includes a subset of the
                publications that are available for this gene. Please see the Gene
                record to access additional publications.
                ##Evidence-Data-START##
                Transcript exon combination :: SRR3476690.537031.1,
                SRR1803611.24070.1 [ECO:0000332]
                RNAseq introns              :: single sample supports all introns
                SAMEA1968540, SAMEA2147920
                [ECO:0000348]
                ##Evidence-Data-END##
                ##RefSeq-Attributes-START##
                MANE Ensembl match          :: ENST00000379143.10/ ENSP00
                RefSeq Select criteria      :: based on conservation, exp
                longest protein
                1249..1254
                /regulatory_class="polyA_sig
                /gene="PCNA"
                /gene_synonym="ATLD2"
                /note="hexamer: AATAAA"

```

# NALOGA 3 – BLAST

1. Kopiraš ak zaporedje človeškega PCNA iz UniProt, prilepiš v pblast in nastaviš iskanje po nr zbirki proteinov. Označiš tudi iskanje le po organizmih Archaea.

BLAST Search Results for protein sequence DGUNXJUV016. The search was limited to the Archaea database. The results table below shows sequences producing significant alignments.

Description	Scientific Name	Max Score	Total Score	Query Cover %	E value	Par. Ident %	Acc. Len	Accession
proliferating cell nuclear antigen (pcna) [archaeon]	archaeon	327	327	99%	6e-111	57.85%	263	M0E2434378.1
proliferating cell nuclear antigen (pcna) [Promethesarchaeia archaeon]	Promethear...	310	310	98%	2e-102	56.25%	408	MG81680798.1
proliferating cell nuclear antigen (pcna) [archaeon]	archaeon	206	206	84%	4e-64	44.09%	222	M0C02815406.1
MAG_annotating cell nuclear antigen (pcna) [archaeon]	archaeon	157	157	49%	4e-46	51.18%	138	RY088154.1
MAG_hypothetical protein AC481_02760 [miscellaneous Crenarchaeota group archa...	miscellaneo...	142	142	99%	9e-39	32.18%	251	K0N828062.1
proliferating cell nuclear antigen (pcna) [Candidatus Methanomethylicus sp.]	Candidatus...	142	142	99%	1e-38	30.23%	250	HR954213.1
proliferating cell nuclear antigen (pcna) [Thermococcus archaeon]	Thermocoot...	142	142	98%	2e-38	32.03%	247	MCC6018166.1

PCNA iz organizma *archaeon*, 6e-111

2. Identičnost je najpogosteje okoli 27 do 35 % . To je kar nizko, a ko pomislimo, da sta liniji ločeni več kot 2 milijardi let, je to zelo visoka podobnost . Ohranjenost kaže, da je struktura "obročna", evolucijsko nepogrešljiva.

3. (Ponoviš enako iskanje kot prej, le da sedaj pri organizmih označiš Bacteria)

Homologa ne najdemo.

Beta objemalka (beta-clamp) - dnaN



*P. furiosus* se nahaja zunaj vseh evkariontov kot outgroup – kar je pričakovano, saj so arheje in evkarioti ločene domene življenja. Drevo to jasno pokaže z dolgo vejno ločitvijo.

- Ja, saj lahko opazimo, da so sesalci (*Mus musculus in Human*) skupaj, opazimo tudi ločene vretenčarje (sesalci + *X. laevis* + *D. rerio*). Vidimo, da je rastlina *A. thaliana* oddaljena od živali in da je arheja *P. furiosus* outgroup (zunanja skupina)

#### 4. Ohranjen je le G

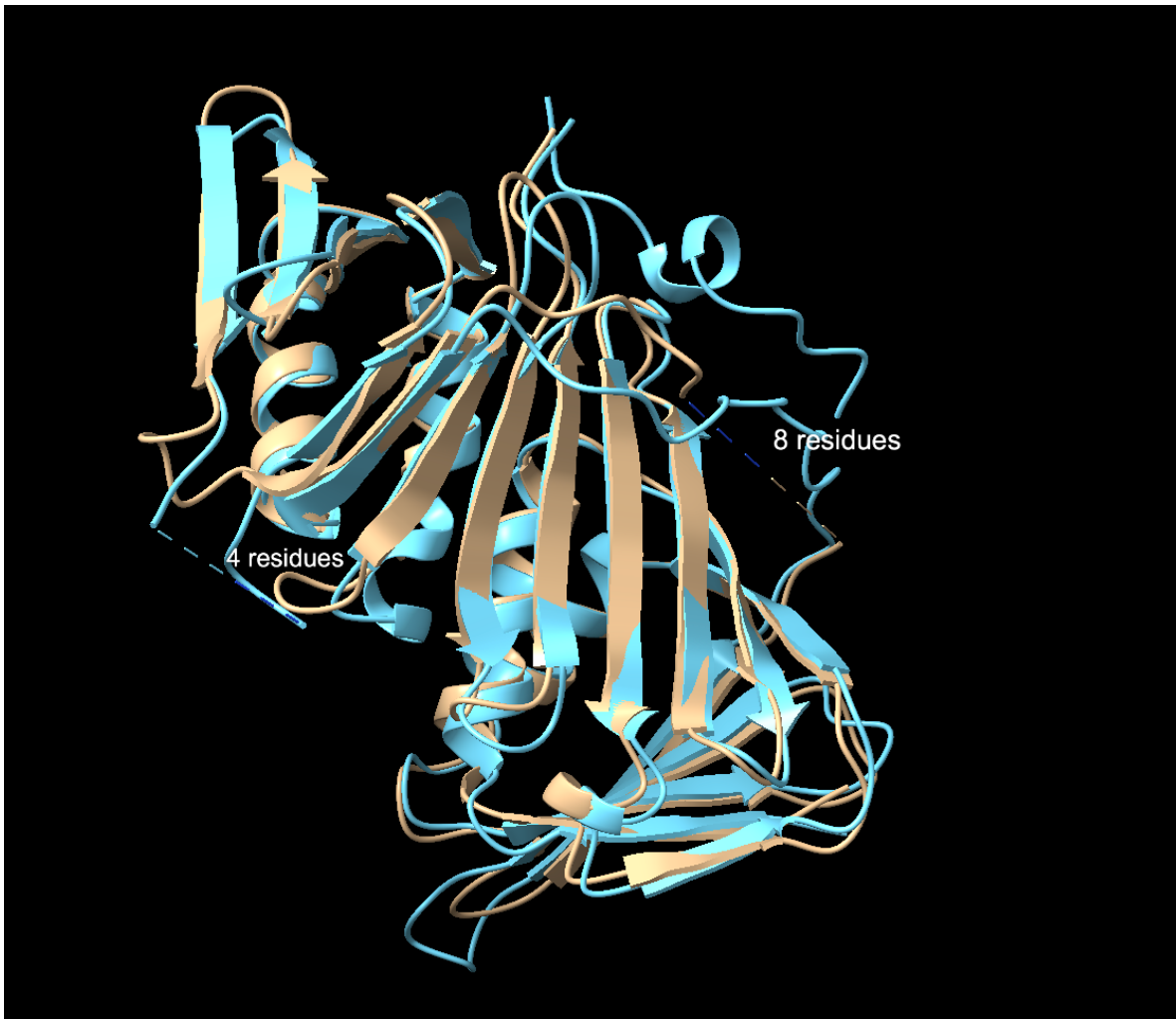
(Narediš enako kor pri 1. nalogi, le da v košarico dodaš še dnaN iz E. coli K12 in ponovno narediš poravnavo)

<input type="checkbox"/> sp P0A988 DPO3B_ECOLI	- - - - CS - - MP I G Q S L P S H S V I V P R K G V	20
<input type="checkbox"/> sp O73947 PCNA_PYRFU	S K Y E V V E P E T I G V N M D H L K K I L K R G K A	85
<input type="checkbox"/> sp P15873 PCNA_YEAST	Q E Y R C D H P V T L G M D L T S L S K I L R C G N N	84
<input type="checkbox"/> sp Q9M7Q7 PCNA1_ARATH	E H Y R C D R N L S M G M N L G N M S K M L K C A G N	84
<input type="checkbox"/> sp P17917 PCNA_DROME	D K F R C D R N L S M G M N L G S M A K I L K C A N N	84
<input type="checkbox"/> sp P18248 PCNA_XENLA	D T Y R C D R N Q S I G V K M S S M S K I L K C A A S	84
<input type="checkbox"/> sp Q9PTP1 PCNA_DANRE	D S Y R C D R N L A M G V N L S S M S K I L K C A G N	84
<input type="checkbox"/> sp P12004 PCNA_HUMAN	D T Y R C D R N L A M G V N L T S M S K I L K C A G N	84
<input type="checkbox"/> sp P17918 PCNA_MOUSE	D T Y R C D R N L A M G V N L T S M S K I L K C A G N	84

#### P12004:DNA binding

<input type="checkbox"/> sp P0A988 DPO3B_ECOLI	I E L M R M L D G G D N P L R V Q I G S N N I R A H V	22
<input type="checkbox"/> sp O73947 PCNA_PYRFU	K D T L I L K K G - E E N F L E - - - - - - - - - -	10
<input type="checkbox"/> sp P15873 PCNA_YEAST	T D T L T L I A D - N T P D S I I - - - - - - - - - -	10
<input type="checkbox"/> sp Q9M7Q7 PCNA1_ARATH	D D I I T I K A D - D G G D T V T - - - - - - - - - -	10
<input type="checkbox"/> sp P17917 PCNA_DROME	E D N V T M K A Q - D N A D T V T - - - - - - - - - -	10
<input type="checkbox"/> sp P18248 PCNA_XENLA	D D I I T L R A E - D N A D T V T - - - - - - - - - -	10
<input type="checkbox"/> sp P12004 PCNA_HUMAN	D T Y R C D R N L A M G V N L T S M S K I L K C A G N	84

- V UniProt pri obeh proteinih pod razdelkom Sequence najdeš PDB kodo za prikaz le A verige. Pri človeku je to koda 1U7B, pri *Pyrococcus furiosus* pa 1GE8. Ko imaš kodo odpreš ChimeraX in napišeš open 1U7B in nato open 1GE8. Nato napišeš matchmake #1 to #2 in dobiš superpozicijo.



modro: 1U7B

bež: 1GE8

Razlik med strukturama proteinov skoraj ni. Skoraj vse  $\alpha$ -vijačnice in  $\beta$ -ploskve se zelo dobro prekrivajo. Opazimo lahko le zelo kratko alfa vijačnico (zgoraj desno), ki jo ima le človeški protein. Manjše neidentičnosti (dolžina) so tudi pri neurejenih regijah.

# NALOGA 5 – FIZIKALNOKEMIJSKE LASTNOSTI

1. pl: 4,57

nestabilnostni indeks: 45,15

	(X)	0	0.0%
	<b>Total number of negatively charged residues (Asp + Glu): 41</b>		
	<b>Total number of positively charged residues (Arg + Lys): 24</b>		
	<b>Atomic composition:</b>		
	Carbon	C	1257
	Hydrogen	H	2020
	Nitrogen	N	328
	Oxygen	O	408
	Sulfur	S	16
	<b>Formula:</b> C <sub>1257</sub> H <sub>2020</sub> N <sub>328</sub> O <sub>408</sub> S <sub>16</sub>		
	<b>Total number of atoms:</b> 4029		
	<b>Extinction coefficients:</b>		
	Extinction coefficients are in units of M <sup>-1</sup> cm <sup>-1</sup> , at 280 nm measured in water.		
	Ext. coefficient		16305
	Abs 0.1% (=1 g/l)	0.567,	assuming all pairs of Cys residues form cystines
	Ext. coefficient	15930	
	Abs 0.1% (=1 g/l)	0.554,	assuming all Cys residues are reduced
	<b>Estimated half-life:</b>		
	The N-terminal of the sequence considered is M (Met).		
	The estimated half-life is: 30 hours (mammalian reticulocytes, in vitro).		
		>20 hours	(yeast, in vivo).
		>10 hours	(Escherichia coli, in vivo).
	<b>Instability index:</b>		
	The instability index (II) is computed to be 45.15		
	This classifies the protein as unstable.		
	<b>Aliphatic index:</b> 94.87		
	<b>Grand average of hydropathicity (GRAVY):</b> -0.095		
Number of amino acids:	261		
Theoretical pI:	4.57		
Molecular weight:	28768.78		
Amino acid composition:	<a href="#">CSV format</a>		
Ala (A)	19	7.3%	
Arg (R)	8	3.1%	
Asn (N)	12	4.6%	
Asp (D)	18	6.9%	
Cys (C)	6	2.3%	
Gln (Q)	8	3.1%	
Glu (E)	23	8.8%	
Gly (G)	14	5.4%	
His (H)	3	1.1%	
Ile (I)	14	5.4%	
Leu (L)	29	11.1%	
Lys (K)	16	6.1%	
Met (M)	10	3.8%	
Phe (F)	8	3.1%	
Pro (P)	7	2.7%	
Ser (S)	25	9.6%	
Thr (T)	12	4.6%	
Trp (W)	1	0.4%	
Tyr (Y)	7	2.7%	
Val (V)	21	8.0%	
Pyl (O)	0	0.0%	
Sec (U)	0	0.0%	
(B)	0	0.0%	
(Z)	0	0.0%	
(X)	0	0.0%	
Total number of negatively charged residues (Asp + Glu):	41		
Total number of positively charged residues (Arg + Lys):	24		

2. Protein računalniško predvideno ni stabilen, a PCNA je v resnici zelo stabilen protein. To kaže na omejitev orodja, da instability index ni vedno zanesljiv, najpogosteje za proteine, ki svojo stabilnost dosežejo šele ko se sestavijo v kompleks (trimer v našem primeru).

3. Pri fiziološkem pH je protein negativno nabit.

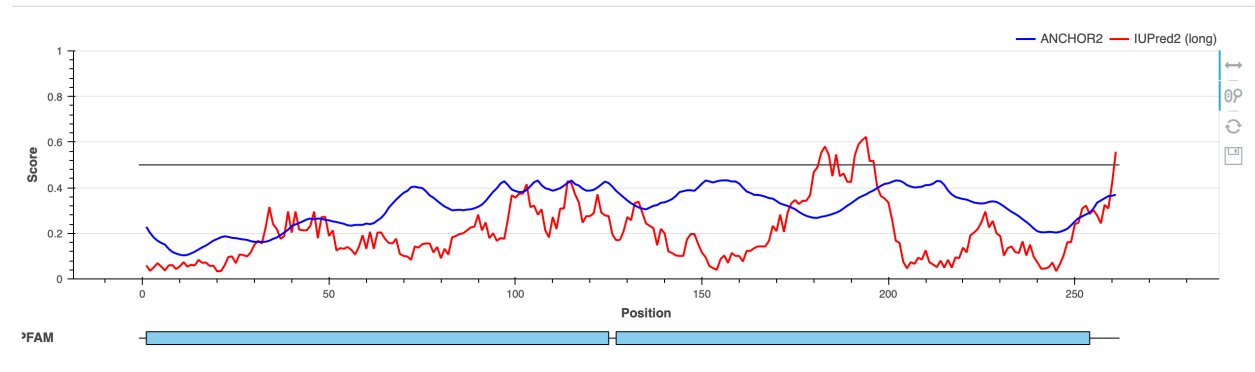
4. Ker piše, da smo pri izolaciji uporabili pufer z dodatkom DTT, so cisteini v reducirani obliki, zato izberemo ekstinkcijski koeficient 15930 M<sup>-1</sup>cm<sup>-1</sup>.

$$A = -\log(0,355) = 0,4498$$

$$c = \frac{A}{\epsilon \cdot l} = \frac{0,4498}{15930 \text{ M}^{-1}\text{cm}^{-1} \cdot 1 \text{ cm}} = 2,82 \times 10^{-5} \text{ M} = 28,2 \text{ } \mu\text{M}$$

## NALOGA 6 - IUPred2

Greš na spletno stran <https://iupred2a.elte.hu> in prilepiš človeško PCNA aminokislinsko zaporedje. Odpre se ti graf:



1. PCNA je pretežno urejen protein, saj ima skoraj celotno zaporedje IUPred2 vrednosti pod 0,5 (pod črno horizontalno črto).
2. Vrednosti nad 0,5 se pojavijo pri C-terminalnem repu in okoli pozicije 180–200.
3. ANCHOR2 napoveduje mesta ki so neurejena, ampak ob vezavi pridobijo urejenejšo strukturo. C-končni rep je torej tisto mesto kjer se vežejo drugi proteini. To je klasičen primer "disorder-to-order transition" ob vezavi.

# NALOGA 7 – Human Protein Atlas

Na spletni strani poiščeš PCNA protein. Ko se ti odprejo rezultati imaš zgoraj različne razdelke.

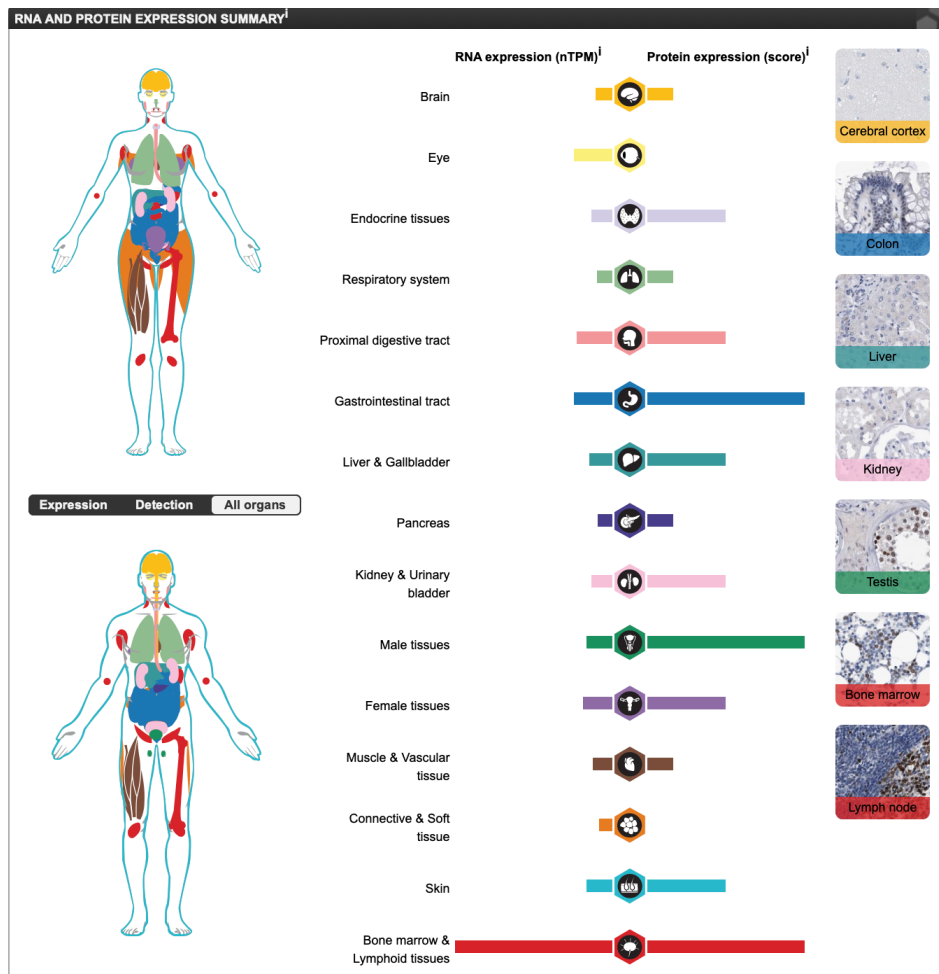
**THE HUMAN PROTEIN ATLAS** | RESOURCES | ABOUT | NEWS | LEARN | DATA | HELP

Search result (15 genes): **PCNA** | PCLAF | RFC4 ...

Navigation: SUMMARY | TISSUE | BRAIN | SINGLE CELL | SUBCELL | CANCER | BLOOD | CELL LINE | STRUCTURE | INTERACTION

PCNA INFORMATION	Value
Protein <sup>1</sup>	Proliferating cell nuclear antigen
Gene name <sup>1</sup>	PCNA
Protein class <sup>1</sup>	Cancer-related genes Disease related genes Essential proteins Human disease related genes Plasma proteins
Protein evidence	Evidence at protein level (all genes)
Number of transcripts <sup>1</sup>	2
Protein interactions	Interacting with 34 proteins

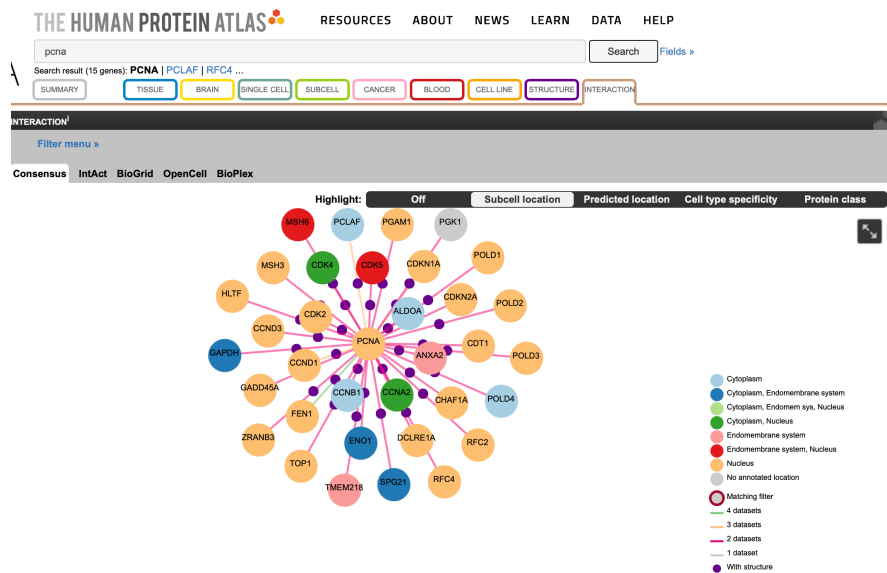
1. Izbereš razdelek TISSUE in odpre se ti to:



Najmočnejše izražanje: kostni mozeg in limfoidno tkivo, moška tkiva, prebavni trakt.

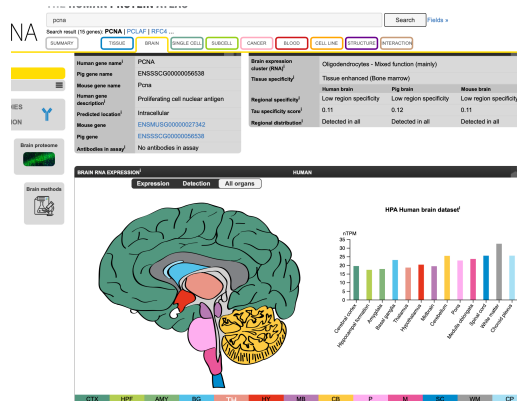
Najšibkejšo izražanje: oči, vezivno in mehko tkivo, mišično in žilno tkivo, trebušna slinavka, dihalni sistem.

2. Klikneš razdelek INTERACTION in se pomakneš malo dol po strani, dokler ne opaziš diagrama interakcij. Kjer piše Highlight označiš Subcell location.



Nahajajo se v jedru, endomembranskem sistemu in citoplazmi.

3. Izbereš razdelek BRAIN in si ogledaš sliko



Povsod in to približno enakomerno.