

#S12-rešitve

- datum izdelave: 2026-05-07

- koda seminarja: S12

a)

1. V iskalniku UniProt pod zadetki Cas9 dobimo naslednje:

UniProtKB 81,503 results or search "cas9" as a Protein Name, Gene Name, Protein family, or Strain

Tools Download (82k) Add View: Cards Table Customize columns Share

Entry	Entry Name	Protein Names	Gene Names	Organism	Length
G3ECR1	CAS9_STRTR	CRISPR-associated endonuclease Cas9[...]	cas9, csn1	Streptococcus thermophilus	1,409 AA
J7RUA5	CAS9_STAAU	CRISPR-associated endonuclease Cas9[...]	cas9	Staphylococcus aureus	1,053 AA
A0Q5Y3	CAS9_FRATN	CRISPR-associated endonuclease Cas9[...]	cas9, FTN_0757	Francisella tularensis subsp. novicida (strain ATCC 15482 / CCUG 33449 / U112)	1,629 AA

Drugi zadetek je tisti, ki ga iščemo, njegov ID pa je takoj za zlatim znakom.

ID: J7RUA5

J7RUA5 · CAS9_STAAU

Proteinⁱ | CRISPR-associated endonuclease Cas9
Geneⁱ | cas9
Statusⁱ | UniProtKB reviewed (Swiss-Prot)
Organismⁱ | Staphylococcus aureus

2. V razdelku »Sequence« poiščemo »Length«, ta nam daje dolžino proteina.

1053

MKRNYILGLD

Function Entry Variant viewer Feature viewer Genomic coordinates Publications External links History

Sequenceⁱ

Sequence statusⁱ | Complete

See also sequence in UniParc or sequence clusters in UniRef

Tools Download Add Highlight Copy sequence

Length 1,053
Mass (Da) 123,949
Last updated 2012-10-31 v1
MD5 Checksumⁱ 7A4C65E2396E4551015E4662F91763E2

MKRNYILGLD 10 IGITSVGYGI 20 IDYETRDVID 30 AGVRLFKEAN 40 VENNENRRSK 50 RGARRLKRKR 60 RHRIQRVKKL 70 LFDYNLLTDH 80 SELSGINPYE 90 ARVKGLSQKL 100 SEEFSAALL 110

HLAKRRGVHN 120 VNEVEEDTGN 130 ELSTKEQISR 140 NSKALEEKYV 150 AELQLERLKK 160 DGEVRGSINR 170 FKTSYVYKEA 180 KQLLVQKAY 190 HQLDQSFIDT 200 YIDLLETART 210 YYEGPGE6SP 220

FGWKDIKEWY 230 EMLMGHCTYF 240 PEELRSVKYA 250 YNADLYNALN 260 DLNLLVITRD 270 ENKLEYVEK 280 FQIIEVFKQ 290 KKKPTLKQIA 300 KEILVNEEDI 310 KGYRVTSYTK 320 PEFTNLKVVH 330

DIKIDITARKE 340 IIEAELLDQ 350 IAKILTIYQS 360 SEDIQEELTN 370 LNSELTQEII 380 EQISNLKGYT 390 GTHNLSLKAI 400 NLILDELWHT 410 NDNQIAIFNR 420 LKLVPKVDL 430 SQQKEIPTTL 440

3. Pod razdelkom »Names & Taxonomy« najdemo »Recommended name« kot tudi »Alternative name«:

CRISPR-associated endonuclease Cas9 (priporočljivo)

SaCas9 (alternativno)

Names & Taxonomy

Subcellular Location

Phenotypes & Variants

PTM/Processing

Expression

Interaction

Structure

Names & Taxonomyⁱ

Protein namesⁱ

Recommended name	CRISPR-associated endonuclease Cas9 UniRule Annotation
EC number	EC:3.1.-- (UniProtKB ENZYME ↗ Rhea ↗) UniRule Annotation
Alternative name	SaCas9 1 Publication

4. Pod razdelkom »Structure« poiščemo »Method«

X-ray

Names & Taxonomy

Subcellular Location

Phenotypes & Variants

PTM/Processing

Expression

Interaction

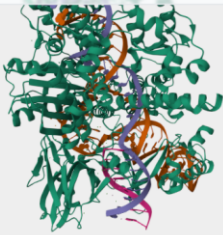
Structure

Family & Domains

Sequence

Similar Proteins

Homologs



SOURCE	IDENTIFIER	METHOD	RESOLUTION	CHAIN	POSITIONS	LINKS
-- Select --	-- Select --	-- Select --				
PDB	5AXW	X-ray	2.70 Å	A	1-1053	PDB · RCSB-PDB · PDBj · PDBsum Download Foldseek
PDB	5CZZ	X-ray	2.60 Å	A	1-1053	PDB · RCSB-PDB · PDBj · PDBsum Download Foldseek

5. Pod razdelkom »Function« poiščemo o ligandih »Ligand«:

Mangan

Function

Names & Taxonomy

Subcellular Location

Phenotypes & Variants

PTM/Processing

Expression

Interaction

Structure

Family & Domains

Sequence

Similar Proteins

Entry

Variant viewer

Feature viewer

Genomic coordinates

Publications

External links

Access the complete set of GO annotations on QuickGO [↗](#)

Keywordsⁱ

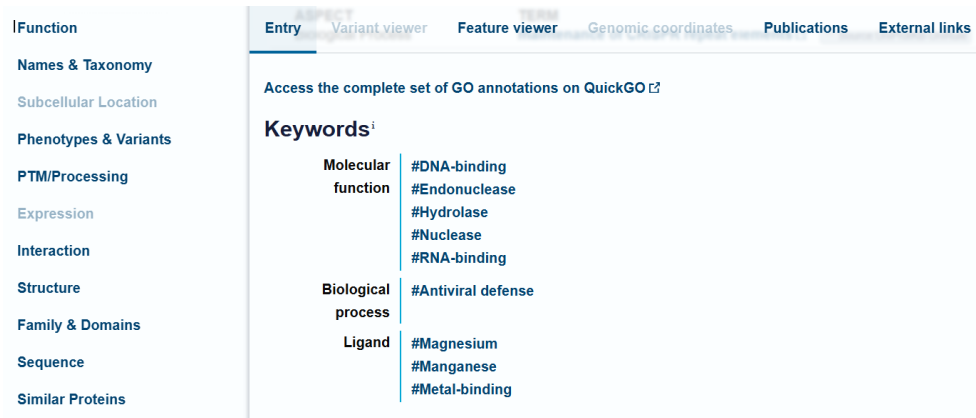
Molecular function	#DNA-binding #Endonuclease #Hydrolase #Nuclease #RNA-binding
Biological process	#Antiviral defense
Ligand	#Magnesium #Manganese #Metal-binding

Kovinski ioni (predvsem Mg^{2+}) so nujni za katalitično aktivnost Cas9, ker stabilizirajo negativne fosfatne skupine DNA in omogočajo pravilno orientacijo substrata ter aktivnih ostankov v katalitičnem centru. Zato se brez njih Cas9 še vedno lahko veže DNA, vendar ne more učinkovito izvesti cepitve, kar pojasni zmanjšano ali odsotno aktivnost v reakciji.

6. Pod razdelkom »Function« poiščemo »Molecular function«

Da, ima tudi nukleazno aktivnost, saj je to anotirano v UniProt.

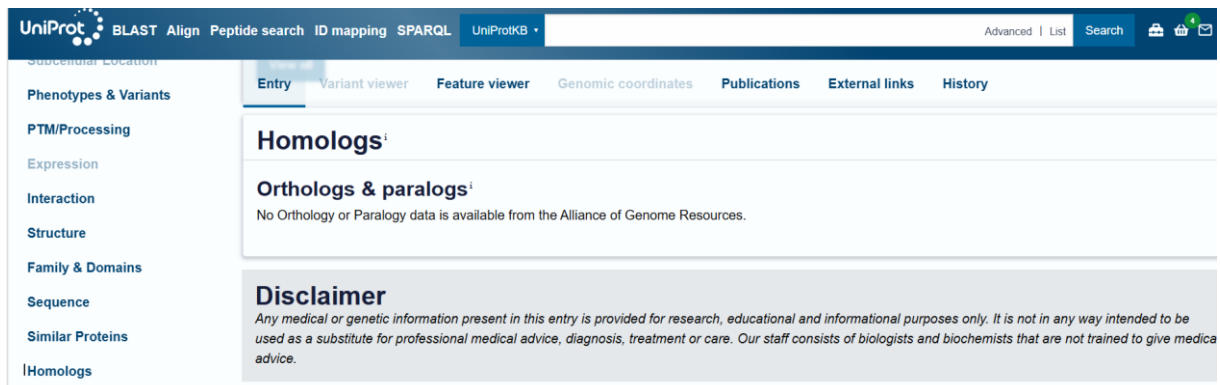
Podatek najdemo v razdelku »Function« med »Molecular function«



The screenshot shows the UniProt 'Function' section. On the left is a navigation menu with categories like 'Names & Taxonomy', 'Subcellular Location', 'Phenotypes & Variants', 'PTM/Processing', 'Expression', 'Interaction', 'Structure', 'Family & Domains', 'Sequence', and 'Similar Proteins'. The main content area has tabs for 'Entry', 'Variant viewer', 'Feature viewer', 'Genomic coordinates', 'Publications', and 'External links'. Below the tabs, there is a link to 'Access the complete set of GO annotations on QuickGO'. The 'Keywords' section is expanded to show 'Molecular function' with the following terms: #DNA-binding, #Endonuclease, #Hydrolase, #Nuclease, and #RNA-binding. Below that, 'Biological process' includes #Antiviral defense, and 'Ligand' includes #Magnesium, #Manganese, and #Metal-binding.

7. Pod razdelkom »Homologs« poiščemo ali je karkoli anotirano.

Ni paralogov, vsaj UniProt nam tega ne podaja.



The screenshot shows the UniProt 'Homologs' section. The top navigation bar includes 'UniProt', 'BLAST', 'Align', 'Peptide search', 'ID mapping', 'SPARQL', 'UniProtKB', and a search bar. The main content area has tabs for 'Entry', 'Variant viewer', 'Feature viewer', 'Genomic coordinates', 'Publications', 'External links', and 'History'. The 'Homologs' section is expanded to show 'Orthologs & paralog data' with the message: 'No Orthology or Paralogy data is available from the Alliance of Genome Resources.' Below this is a 'Disclaimer' section with the text: 'Any medical or genetic information present in this entry is provided for research, educational and informational purposes only. It is not in any way intended to be used as a substitute for professional medical advice, diagnosis, treatment or care. Our staff consists of biologists and biochemists that are not trained to give medical advice.'

b) Najprej poiščemo zadetke, ko se Cas9 nahaja v naslovu:

V iskalniku PubMed med Advanced iskanjem izberemo article review, nato kliknemo AND in dodamo tudi opcijo Title, kjer zapišemo Cas9, saj nas to zanima. Nato damo search in izberemo tudi časovni razpon (2013-2015, 2016-2018, 2019-2021, 2022-2024).

Enako naredimo tudi, ko se Cas9 nahaja le v besedilu, le da ne izberemo še Cas9[Title]

The screenshot shows the PubMed search interface. The search bar contains the query: `(article review[Publication Type]) AND (Cas9[Title])`. Below the search bar, there are buttons for 'Save', 'Email', and 'Send to', and a 'Sort by' dropdown set to 'Best match'. On the left, there is a 'MY CUSTOM FILTERS' section with 'Edit custom filters' and a 'RESULTS BY YEAR' bar chart for the years 2013-2015. The main results area shows 84 results, with the first two items:

- Genome modification by CRISPR/Cas9.**
1 Ma Y, Zhang L, Huang X.
Cite FEBS J. 2014 Dec;281(23):5186-93. doi: 10.1111/febs.13110. Epub 2014 Nov 7.
PMID: 25315507 **Free article.** Review.
- Genome editing. The new frontier of genome engineering with CRISPR-Cas9.**
2 Doudna JA, Charpentier E.
Cite Science. 2014 Nov 28;346(6213):1258096. doi: 10.1126/science.1258096.
PMID: 25430774 Review.

Ko poiščemo 2013-2015, da se Cas9 nahaja v naslovu: 84 rezultatov

The screenshot shows the PubMed search interface. The search bar contains the query: `(article review[Publication Type]) AND Cas9`. Below the search bar, there are buttons for 'Save', 'Email', and 'Send to', and a 'Sort by' dropdown set to 'Best match'. On the left, there is a 'MY CUSTOM FILTERS' section with 'Edit custom filters' and a 'RESULTS BY YEAR' bar chart for the years 2013-2015. The main results area shows 219 results, with the first item:

- Genome modification by CRISPR/Cas9.**
1 Ma Y, Zhang L, Huang X.
Cite FEBS J. 2014 Dec;281(23):5186-93. doi: 10.1111/febs.13110. Epub 2014 Nov 7.
PMID: 25315507 **Free article.** Review.
One major concern about this system is its specificity; although CRISPR-Cas9-mediated off-target mutation has been broadly studied, more efforts are required to further improve the specificity of CRISPR-Cas9. We will also discuss the potential applications of CRISPR ...

Ko poiščemo 2013-2015, da se Cas9 NE nahaja v naslovu: 219 rezultatov

DELEŽ : 84/219= 38,36%

PubMed®

(article review[Publication Type]) AND (Cas9[Title])

Advanced Create alert Create RSS User Guide

Save Email Send to Sort by: Best match Display options

MY CUSTOM FILTERS Edit custom filters

RESULTS BY YEAR

2016-2018

329 results

1 **CRISPR-Cas9 Structures and Mechanisms.**
1 Jiang F, Doudna JA.
Cite Annu Rev Biophys. 2017 May 22;46:505-529. doi: 10.1146/annurev-biophys-062215-010822. Epub 2017 Mar 30.
PMID: 28375731 Review.

2 **CRISPR/Cas9 Immune System as a Tool for Genome Engineering.**
2 Hryhorowicz M, Lipiński D, Zeyland J, Słomski R.

Ko poiščemo 2016-2018, da se Cas9 nahaja v naslovu: 329 rezultatov

PubMed®

(article review[Publication Type]) AND Cas9

Advanced Create alert Create RSS User Guide

Save Email Send to Sort by: Best match Display options

MY CUSTOM FILTERS Edit custom filters

RESULTS BY YEAR

2016-2018

1,045 results

1 **CRISPR-Cas9 Structures and Mechanisms.**
1 Jiang F, Doudna JA.
Cite Annu Rev Biophys. 2017 May 22;46:505-529. doi: 10.1146/annurev-biophys-062215-010822. Epub 2017 Mar 30.
PMID: 28375731 Review.
Target recognition strictly requires the presence of a short protospacer adjacent motif (PAM) flanking the target site, and subsequent R-loop formation and strand scission are driven by complementary base pairing between the guide RNA and target DNA. **Cas9**-DNA interactions ...

Ko poiščemo 2016-2018, da se Cas9 NE nahaja v naslovu: 1045 rezultatov

DELEŽ: 329/1045= 31,48%

NIH National Library of Medicine National Center for Biotechnology Information Log in

PubMed®

(article review[Publication Type]) AND (Cas9[Title])

Advanced Create alert Create RSS User Guide

Save Email Send to Sort by: Best match Display options

MY CUSTOM FILTERS Edit custom filters

RESULTS BY YEAR

2019-2021

339 results

1 **CRISPR-Cas9 for cancer therapy: Opportunities and challenges.**
1 Chen M, Mao A, Xu M, Weng Q, Mao J, Ji J.
Cite Cancer Lett. 2019 Apr 10;447:48-55. doi: 10.1016/j.canlet.2019.01.017. Epub 2019 Jan 23.
PMID: 30684591 Review.

2 **CRISPR/Cas9 gene editing for curing sickle cell disease.**
2 Park SH, Bao G.
Cite Transfus Apher Sci. 2021 Feb;60(1):103060. doi: 10.1016/j.transci.2021.103060. Epub 2021 Jan 10.
PMID: 33455878 **Free PMC article.** Review.

Ko poiščemo 2019-2021, da se Cas9 nahaja v naslovu: 319 rezultatov

NIH National Library of Medicine
National Center for Biotechnology Information

Log in

PubMed®

(article review[Publication Type]) AND Cas9

Advanced Create alert Create RSS User Guide

Save Email Send to Sort by: Best match Display options

MY CUSTOM FILTERS
Edit custom filters

RESULTS BY YEAR

2019-2021

1,427 results

Page 1 of 143

CRISPR-Cas9 for cancer therapy: Opportunities and challenges.

1 Chen M, Mao A, Xu M, Weng Q, Mao J, Ji J.
Cancer Lett. 2019 Apr 10;447:48-55. doi: 10.1016/j.canlet.2019.01.017. Epub 2019 Jan 23.
Cite PMID: 30684591 Review.
Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR)-Cas9-mediated genome editing technology has been extensively applied in various cell types and organisms, both in vitro and in vivo, for efficient gene disruption and gene modification. ...Here, we present ...

Ko poiščemo 2019-2021, da se Cas9 NE nahaja v naslovu: 1427 rezultatov

DELEŽ: $319/1427 = 22,35\%$

NIH National Library of Medicine
National Center for Biotechnology Information

Log in

PubMed®

(article review[Publication Type]) AND (Cas9[Title])

Advanced Create alert Create RSS User Guide

Save Email Send to Sort by: Best match Display options

MY CUSTOM FILTERS
Edit custom filters

RESULTS BY YEAR

2022-2024

400 results

Page 1 of 40

Current applications and future perspective of CRISPR/Cas9 gene editing in cancer.

1 Wang SW, Gao C, Zheng YM, Yi L, Lu JC, Huang XY, Cai JB, Zhang PF, Cui YH, Ke AW.
Mol Cancer. 2022 Feb 21;21(1):57. doi: 10.1186/s12943-022-01518-8.
Cite PMID: 35189910 Free PMC article. Review.

CRISPR-Cas9: A History of Its Discovery and Ethical Considerations of Its Use in Genome Editing.

2

Ko poiščemo 2022-2024, da se Cas9 nahaja v naslovu: 400 rezultatov

NIH National Library of Medicine
National Center for Biotechnology Information

Log in

PubMed®

(article review[Publication Type]) AND Cas9

Advanced Create alert Create RSS User Guide

Save Email Send to Sort by: Best match Display options

MY CUSTOM FILTERS
Edit custom filters

RESULTS BY YEAR

2022-2024

1,757 results

Page 1 of 176

Current applications and future perspective of CRISPR/Cas9 gene editing in cancer.

1 Wang SW, Gao C, Zheng YM, Yi L, Lu JC, Huang XY, Cai JB, Zhang PF, Cui YH, Ke AW.
Mol Cancer. 2022 Feb 21;21(1):57. doi: 10.1186/s12943-022-01518-8.
Cite PMID: 35189910 Free PMC article. Review.
Here, we review current status of CRISPR/Cas9 gene editing technology in oncological research. We first explain the basic principles of CRISPR/Cas9 gene editing and introduce several new CRISPR-based gene editing modes. We next detail the rapid progress of CRISPR sc ...

Ko poiščemo 2022-2024, da se Cas9 NE nahaja v naslovu: 1757 rezultatov

DELEŽ: 400/1757= 22,77%

Rezultati kažejo, da je bil delež review člankov, kjer je Cas9 omenjen v naslovu glede na vse omembe, najvišji v obdobju 2013–2015, nato pa postopno upada do 2019–2021, kjer doseže minimum, in se v zadnjem obdobju 2022–2024 rahlo stabilizira oziroma rahlo poveča.

Največji relativni padec deleža po letu 2015 lahko razložimo s tem, da je bil Cas9 v zgodnjem obdobju CRISPR raziskav še osrednja novost področja, zato se je pogosto pojavljal že v naslovih review člankov kot glavni fokus. Kasneje, z eksplozijo CRISPR raziskav po letu 2016, se je področje razširilo na številne Cas variante, aplikacije in širše genomske tehnologije, zato Cas9 ni bil več vedno glavni poudarek v naslovu, temveč pogosto le ena od komponent širših pregledov.

V obdobju 2022–2024 se delež rahlo stabilizira, kar lahko kaže na to, da je Cas9 postal standardizirano orodje, zato se ponovno pogosteje pojavlja kot ključna referenčna točka v naslovih preglednih člankov, zlasti v kontekstu kliničnih aplikacij in optimizacij sistema.

c) Podobno kot smo v točki a) poiskali naš Cas9, poiščemo v UniProt tudi Cas12a. Pazimo, da je organizem res tisti, ki ga iščemo. Zadetek se sicer nahaja takoj na začetku vseh zadetkov, tako da je verjetnosti, da se zmotimo res malo.

Ko najdemo protein Cas12a, enako kot smo poiskali v točki a) za Cas9 poiščemo med »Sequence« aminokislinsko zaporedje in ga kopiramo.

S tem zaporedjem gremo v EMBOSS NEEDLE, kamor pripnemo aminokislinsko zaporedje proteina Cas9 in v naslednje okno še zaporedje proteina Cas12a.

EMBOSS NEEDLE
Pairwise Sequence Alignment (PSA)

Home Help & Privacy Recent Jobs Input Form Feedback

Welcome to the Job Dispatcher website! If you need assistance or have feedback, please [contact us](#).

EMBOSS Needle reads two input sequences and writes their optimal global sequence alignment to file.

Input sequence ⓘ

Sequence type
 Protein DNA

Paste your **first** sequence here - or use the example sequence

```
YETFKKHLNLAKGKGRISKTKKEYLLEERDJNRFVQKDFINRNLVUTRYATRGLMNL
RSYFRVNNLDVVKVKSINGGFTSLRRKWKFKKERNKGKHAEDALIANADIFKEWKK
LDKAKKVMENQMFEEKQAESMPEIETEQEYKEIFITPHQIKHIKDFKDYKSHRVDKKN
RELINDLYSTRKDDKGNLTVNNGLYDKDNDKLLINKSPEKLLMYHDPQTYQKL
KLIMEQYGDENPLYKYEEETGNLYTKYSKKNPVIKIKIYGNKLNALDITDDYPNS
RNKVVKLSLKPYPFDVYLDNGVYKFTVTKNLDVIKKENYEVNSKCYEAKLKKISNQ
EFIASFYNNDLIKINGELYRIGVNNLLNRIEVNMDITYREYLENMNDKRPPIIKTI
ASKTQSIKKYSTDILGNLYEVKSKKHPQIIKKG
```


V poravnavi si v idealnem primeru želimo največ simbola “*”, saj ta označuje identične aminokislino v obeh zaporedjih in zato neposredno kaže na najvišjo stopnjo ohranjenosti.

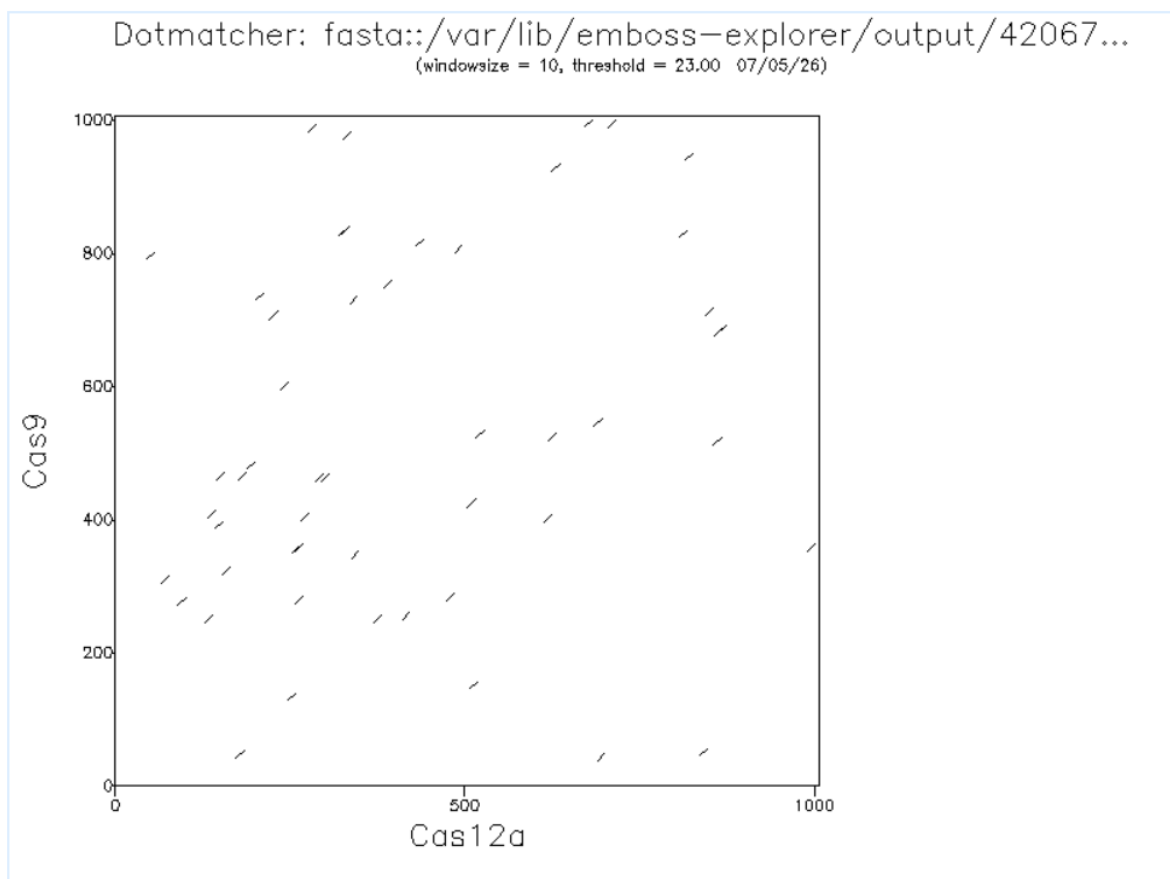
Pri primerjavi Cas9 in Cpf1 (Cas12a) ta pogoj ni v veliki meri izpolnjen, saj je delež identičnosti relativno nizek (okoli 15–21 %).

Takšen vzorec poravnave nastane zaradi velike evlucijske razdalje med proteinoma, saj sta se Cas9 in Cas12a razvila iz različnih evlucijskih linij CRISPR sistemov in sta kljub enaki funkciji (cepitev DNA) razvila različne aminokislinske sekvence in domenske organizacije. Posledično se ohranjajo le ključni funkcionalni motivi, medtem ko so ostali deli močno divergentni, kar se v poravnavi kaže kot veliko število neskladij in vrzeli.

d) Ko izvajamo poravnavo v dotmatcher nam Google takoj najde ta program, sam dotmatcher pa nas lepo vodi kam pripravimo aminokislinsko zaporedje. Parametra za threshold pri tej nalogi ne spreminjamo.

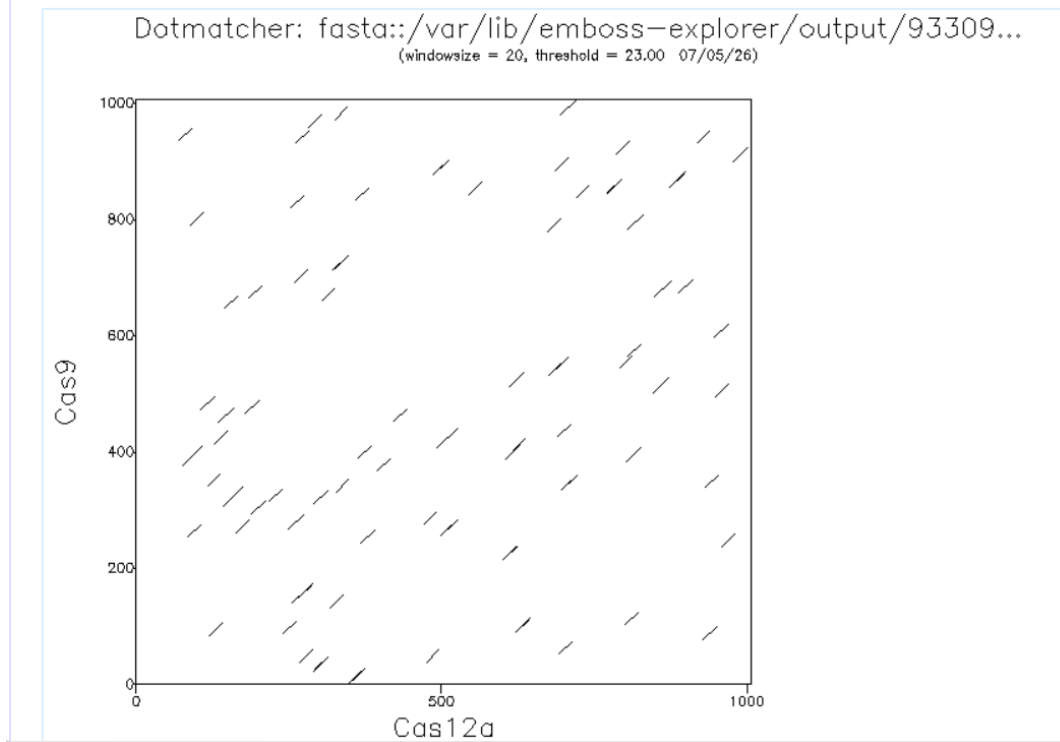
Ko nastavimo okno 10:

IMAGE FILE [dotmatcher.1.png](#)



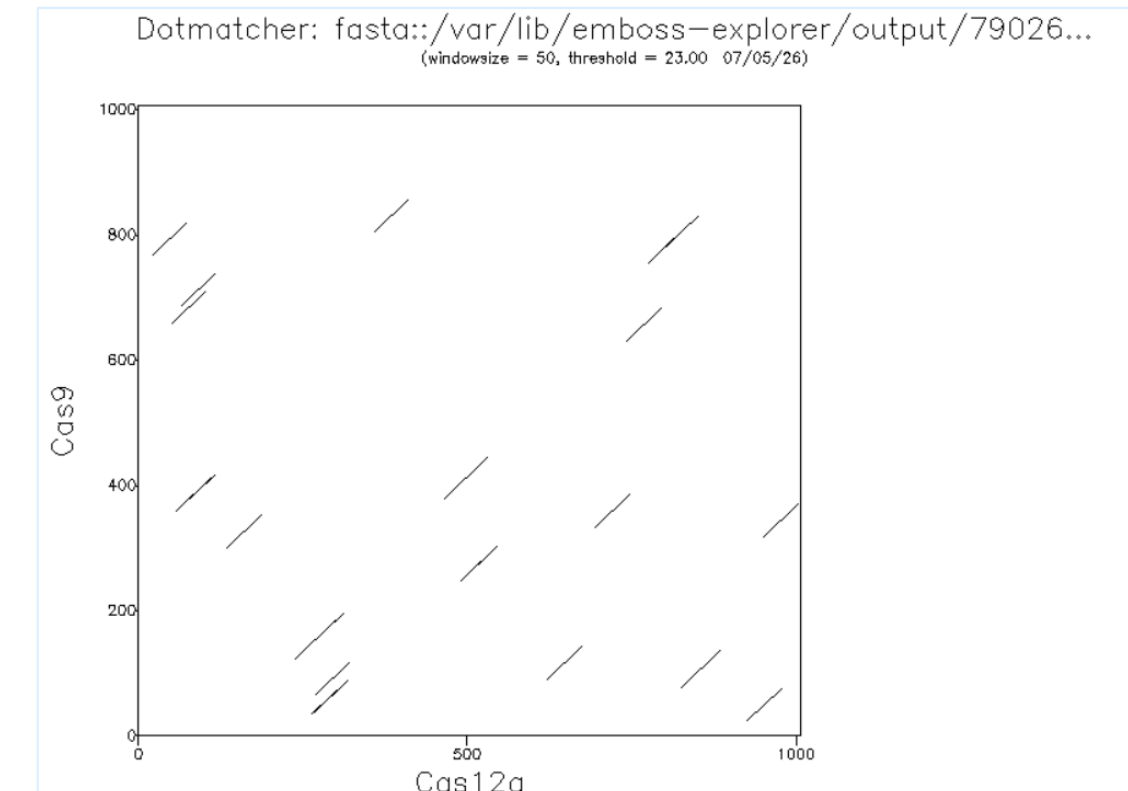
Ko nastavimo okno 20:

IMAGE FILE [dotmatcher.1.png](#)



Ko nastavimo okno 50:

IMAGE FILE [dotmatcher.1.png](#)



Pri manjšem oknu zaznamo več kratkih lokalnih podobnosti, vendar se poveča tudi količina šuma, medtem ko večje okno prikaže predvsem najbolj ohranjene funkcionalne regije proteinov.

Pri izbranih nastavitvah analize opazimo več kratkih diagonalnih segmentov, vendar brez ene kontinuirane diagonale preko celotnega grafa. To pomeni, da med proteinoma Cas9 in Cas12a ni visoke globalne podobnosti, temveč si delita le posamezne lokalno ohranjene regije.

Ugotavljanje pomena substitucijske matrike:

PAM250 je bolj "tolerantna" matrika, ker dovoljuje več aminokislinskih zamenjav, ki so nastale skozi daljšo evolucijo. Zato zazna tudi šibkejše in oddaljene podobnosti, ki jih EBLOSUM62 (strožja matrika, temeljena na bolj konzerviranih zaporedjih) pogosto ne zazna.

Ker PAM250 uporablja nižji prag za sprejem podobnosti med aminokisljinami, zato sprejme več "šibkih" ujemanj in poveča občutljivost analize. Kot posledica se poleg pravih homolognih regij pojavijo tudi naključna ujemanja, kar vizualno izgleda kot več pikic.

Za Cas9 in Cas12a je primernejša EBLOSUM62 za osnovno interpretacijo, ker daje bolj zanesljive (manj lažno pozitivne) rezultate in jasno pokaže samo močno ohranjene funkcionalne regije.

e) Primerjamo rezultate poravnav iz točke c) in d)

Rezultati dot plot analize se dobro ujemajo z rezultati globalne (Needle) in lokalne (Water) poravnave. Needle poravnava pokaže nizko globalno podobnost in visoko število vrzeli, kar potrjuje, da si proteina Cas9 in Cas12a) nista tesno evolucijsko sorodna po celotni dolžini zaporedja.

Water poravnava po drugi strani zazna lokalno ohranjene regije, kar se ujema z kratkimi diagonalnimi segmenti v dot plot analizi, ki predstavljajo funkcionalno pomembne motive.

Dot plot vizualno potrdi oba rezultata: odsotnost ene kontinuirane diagonale (kot bi bila pri tesno sorodnih proteinih) in prisotnost le fragmentiranih podobnosti kaže na evolucijsko oddaljenost proteinov.

Skupaj vse tri metode konsistentno kažejo, da imata Cas9 in Cas12a podobno funkcijo, vendar različen evolucijski izvor, kar pomeni funkcionalno konvergenco ob strukturni divergenci.

f) Ponovno se vrnemo v UniProt med razdelek »Family & Domains«

Lokalne podobnosti v dotplotu odražajo funkcionalno ekvivalentne katalitične regije, predvsem nukleazne domene (RuvC-like in HNH pri Cas9 ter RuvC-like jedro pri Cas12a), ki omogočajo isto biološko funkcijo (cepitev DNA), kljub evolucijski razliki v strukturi proteinov.

Rezultati iskanja pri Cas9:

Names & Taxonomy

Subcellular Location

Phenotypes & Variants

PTM/Processing

Expression

Interaction

Structure

Family & Domains

Sequence

Similar Proteins

Homologs

TYPE	ID	POSITION(S)	DESCRIPTION	
Region		1-41	RuvC-I	
Region		41-426	Recognition lobe	
Region		435-481	RuvC-II	
Domain		480-646	HNH Cas9-type	
Region		650-775	RuvC-III	
Region		882-889	PAM substrate-binding	
Region		910-1053	PAM-interacting domain (PI)	
Region		985-993	PAM substrate-binding	

Rezultati iskanja pri Cas12a:

Function

Names & Taxonomy

Subcellular Location

Phenotypes & Variants

PTM/Processing

Expression

Interaction

Structure

Family & Domains

Sequence

Similar Proteins

Homologs

Entry Variant viewer Feature viewer Genomic coordinates Publications External links History

TYPE	ID	POSITION(S)	DESCRIPTION	
Region		1-35	WED-I (OBD-I)	
Region		36-320	REC1 (helical-I)	
Coiled coil		74-106		
Region		321-526	WED-II (helical-II)	
Region		527-598	WED-III (OBD-I)	
Region		599-718	PI (LHD)	
Region		719-884	WED-III (OBD-III)	

Preverimo še njuno molekularno funkcijo, da lahko potem zaključimo glede njune podobnosti. Ponovno poiščemo v UniProt med »Function«

Rezultati za Cas9:

Names & Taxonomy

Subcellular Location

Phenotypes & Variants

PTM/Processing

Expression

Interaction

Structure

Family & Domains

Sequence

Similar Proteins

Homologs

Keywords:

Molecular function	#DNA-binding #Endonuclease #Hydrolase #Nuclease #RNA-binding
Biological process	#Antiviral defense
Ligand	#Magnesium #Manganese #Metal-binding

Enzyme and pathway databases

ENZYME | Search...

Names & Taxonomy:

Protein names:

Recommended	CRISPR-associated endonuclease Cas9
-------------	-------------------------------------

Rezultati za Cas12a:

The screenshot shows the UniProt entry for Cas12a. The left sidebar lists various categories: Function, Names & Taxonomy, Subcellular Location, Phenotypes & Variants, PTM/Processing, Expression, Interaction, Structure, Family & Domains, Sequence, Similar Proteins, and Homologs. The main content area is divided into several sections:

- Function:** Molecular function includes #DNA-binding, #Endonuclease, #Hydrolase, #Lyase, #Nuclease, and #RNA-binding. Biological process includes #Antiviral defense. Ligand includes #Magnesium.
- Enzyme and pathway databases:** ENZYME search links.
- Names & Taxonomy:** Protein names section showing the recommended name: CRISPR-associated endonuclease Cas12a (1 Publication).

ZAKLJUČIMO LAHKO, da sta Cas9 in Cas12a različna CRISPR tipa (II vs V), evlucijsko sta se razvila ločeno in uporabljata različno arhitekturo, a podobno funkcijo.

g) V UniProt v razdelku »Structure« poiščemo, kje najdemo AlphaFold

The screenshot shows the UniProt Structure section for Cas12a. It features a 3D ribbon diagram of the protein structure. Below the diagram is a table with columns: SOURCE, IDENTIFIER, METHOD, RESOLUTION, CHAIN, and POSITIONS LINKS. The table lists several entries, with the AlphaFold DB entry highlighted by a red box:

SOURCE	IDENTIFIER	METHOD	RESOLUTION	CHAIN	POSITIONS LINKS
-- Select --	-- Select --	-- Select --			
PDB	8YWH	EM	2.97 Å	A	1-1053 PDB · PDBj · PDBsum Foldseek
AlphaFold DB	AF-J7RUA5-F1	Predicted			1-1053 AlphaFold DB Foldseek
SWISS-MODEL	J7RUA5_1-1053:7eni.1.A	Modeling		A	1-1053 SWISS-MODEL Foldseek
SWISS-MODEL	J7RUA5_3-1052:7ef1.1.A	Modeling		A	3-1052 SWISS-MODEL Foldseek
SWISS-MODEL	J7RUA5_1-	Modeling		R	1-1053 SWISS-MODEL Foldseek

Below the table, there is a section for 3D structure databases.

1 struktura

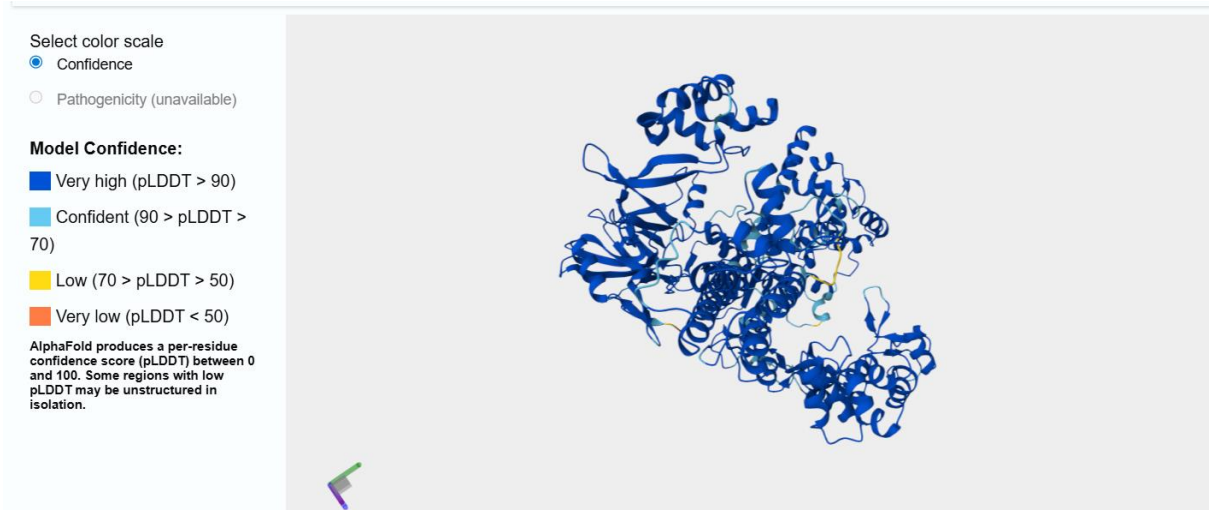
Zanesljivost je večinoma visoka, kar prepoznamo po pLDDT vrednostih:

modre regije (>90): zelo visoka zanesljivost

zelene regije (70-90): dobra zanesljivost

rumene/rdeče (<70): nizka zanesljivost

Model, ki je predstavljen za Cas9 ima prevladujoče modre regije, torej zelo visoka zanesljivost.



Regije z nizkim pLDDT predstavljajo običajno linkerji med domenami, fleksibilne zanke in gibljivi deli. Njihova funkcionalna vloga je predvsem omogočanje konformacijskih sprememb ob vezavi DNA, aktivacijo encima in premik katalitičnih domen.

Da, Cas9 je bila sicer tudi eksperimentalno določena.

Npr. X-ray je bila ena od eksperimentalnih metod, EM.

The screenshot shows the UniProt database entry for the Cas9 protein structure. The 'Feature viewer' tab is selected, displaying a table of structural features. The 'METHOD' column is highlighted with a red box, showing 'X-ray' for all four entries. The table includes columns for SOURCE, IDENTIFIER, METHOD, RESOLUTION, CHAIN, POSITIONS, and LINKS.

SOURCE	IDENTIFIER	METHOD	RESOLUTION	CHAIN	POSITIONS	LINKS
PDB	7EL1	X-ray	2.22 Å	A	1-1053	PDB · RCSB-PDB · PDBj · PDBsum · Foldseek
PDB	7ENH	X-ray	2.10 Å	A	481-646	PDB · RCSB-PDB · PDBj · PDBsum · Foldseek
PDB	7ENI	X-ray	2.63 Å	A	1-1053	PDB · RCSB-PDB · PDBj · PDBsum · Foldseek
PDB	7ENR	X-ray	4.21 Å	A	1-1053	PDB · RCSB-PDB · PDBj · PDBsum · Foldseek

3D structure databases

AlphaFoldDB | J7RUA5 [↗](#) SMR | J7RUA5 [↗](#)

Anotacija Position 1-68 pomeni, da je prikazan le N-terminalni del proteina Cas9. Torej le prvih 68 aminokislinskih ostankov. Ta del je zato pogosto bolj fleksibilen, manj strukturiran, lahko pa je ključen za regulacijo ali začetno zvijanje proteina.

The screenshot shows the UniProt entry page for J7RUA5. The '3D structure databases' section is visible, listing several models. The first entry, 'AlphaFold DB' with identifier 'AF-J7RUA5-F1' and method 'Predicted', is highlighted in yellow. The position range '1-1053' is circled in red. Other entries include 'SWISS-MODEL' models for 'J7RUA5_1-1053:7eni.1.A', 'J7RUA5_3-1052:7ei1.1.A', and 'J7RUA5_1-1053:9mb6.1.B'.

SOURCE	IDENTIFIER	METHOD	RESOLUTION CHAIN	POSITIONS	LINKS
AlphaFold DB	AF-J7RUA5-F1	Predicted		1-1053	AlphaFold DB
SWISS-MODEL	J7RUA5_1-1053:7eni.1.A	Modeling	A	1-1053	SWISS-MODEL
SWISS-MODEL	J7RUA5_3-1052:7ei1.1.A	Modeling	A	3-1052	SWISS-MODEL
SWISS-MODEL	J7RUA5_1-1053:9mb6.1.B	Modeling	B	1-1053	SWISS-MODEL

h) Da najdemo povezavo do GenBank, si pomagamo z »External links«

The screenshot shows the UniProt entry page for J7RUA5 - CAS9_STAAU. The 'External links' tab is highlighted with a red box. The page displays protein information such as 'CRISPR-associated endonuclease Cas9', 'cas9' gene, and 'Staphylococcus aureus' organism. It also shows 'Amino acids: 1053 (go to sequence)', 'Protein existence: Evidence at protein level', and 'Annotation score: 5/5'.

J7RUA5 - CAS9_STAAU

Proteinⁱ | CRISPR-associated endonuclease Cas9 | Amino acids | 1053 (go to sequence)

Geneⁱ | cas9 | Protein existenceⁱ | Evidence at protein level

Statusⁱ | UniProtKB reviewed (Swiss-Prot) | Annotation scoreⁱ | 5/5

Organismⁱ | Staphylococcus aureus

Entry | Variant viewer | Feature viewer | Genomic coordinates | Publications | **External links** | History

Tools | Download | Add | Add a publication | Entry feedback

S klin

S klikom na GenBank med »Sequence databases dobimo našo želeno povezavo do GenBank, kjer bomo našli informacijo o številu baznih parov

The screenshot shows the UniProt database interface. At the top, there are navigation tabs: ENZYME, Family and domain databases, and Sequence databases. The 'Family and domain databases' section lists various databases like Gene3D, HAMAP, InterPro, and NCBIfam. The 'Sequence databases' section is highlighted with a red box and contains a table with columns: NUCLEOTIDE SEQUENCE, PROTEIN SEQUENCE, MOLECULE TYPE, and STATUS. The table lists HE980450 with protein sequence CCK74173.1 and molecule type Genomic DNA.

Rezultat:

56 115 baznih parov

The screenshot shows the GenBank entry for Staphylococcus aureus subsp. aureus ORFX gene and pseudo SCCmec-SCC-SCCCRISPR element, strain M06/0171. The entry includes the accession number HE980450.1 and the length of the sequence, 56115 bp, which is highlighted with a red box. The entry also includes the definition, accession number, version, keywords, source, and organism information.