

## REŠITVE S3-07

### 1. Del

Dano nukleotidno zaporedje damo v blastn. Če pogledamo zadetke, ki se nahajajo malo bolj spodaj dobimo tak rezultat:

<input checked="" type="checkbox"/>	<a href="#">Norovirus Hu/GI.1/8K/1979/USA_complete genome</a>	<a href="#">Norovirus Hu/GI...</a>	2001	2001	100%	0.0	100.00%	7617	<a href="#">KF429783.1</a>
<input checked="" type="checkbox"/>	<a href="#">Norovirus Hu/GI.1/8McIII/1973/USA_complete genome</a>	<a href="#">Norovirus Hu/GI...</a>	2001	2001	100%	0.0	100.00%	7626	<a href="#">KF429770.1</a>
<input checked="" type="checkbox"/>	<a href="#">Norovirus Hu/GI.1/8W/1968/USA_complete sequence</a>	<a href="#">Norovirus Hu/GI...</a>	2001	2001	100%	0.0	100.00%	7617	<a href="#">KF429765.1</a>
<input checked="" type="checkbox"/>	<a href="#">Norovirus Hu/GI.1/8MoIII/1972/USA_complete genome</a>	<a href="#">Norovirus Hu/GI...</a>	2001	2001	100%	0.0	100.00%	7623	<a href="#">KF429761.1</a>
<input checked="" type="checkbox"/>	<a href="#">Norovirus Hu/GI.1/CHA6A003_20091104/2009/USA_complete genome</a>	<a href="#">Norovirus Hu/GI...</a>	2001	2001	100%	0.0	100.00%	7600	<a href="#">KF039737.1</a>
<input checked="" type="checkbox"/>	<a href="#">Norovirus Hu/GI.1/CHA2A014/2008/USA_complete genome</a>	<a href="#">Norovirus Hu/GI...</a>	2001	2001	100%	0.0	100.00%	7600	<a href="#">KF039728.1</a>
<input checked="" type="checkbox"/>	<a href="#">Norovirus Hu/GI.1/CHA6A014/2009/USA_complete genome</a>	<a href="#">Norovirus Hu/GI...</a>	2001	2001	100%	0.0	100.00%	7600	<a href="#">KF039734.1</a>
<input checked="" type="checkbox"/>	<a href="#">Norovirus Hu/GI.1/CHA9A004_20110426/2011/USA_complete genome</a>	<a href="#">Norovirus Hu/GI...</a>	2001	2001	100%	0.0	100.00%	7600	<a href="#">KF039733.1</a>
<input checked="" type="checkbox"/>	<a href="#">Norovirus Hu/GI.1/CHA5A010/2009/USA_complete genome</a>	<a href="#">Norovirus Hu/GI...</a>	2001	2001	100%	0.0	100.00%	7600	<a href="#">KF039732.1</a>
<input checked="" type="checkbox"/>	<a href="#">Norovirus Hu/GI.1/CHA XO533/2009/USA_complete genome</a>	<a href="#">Norovirus Hu/GI...</a>	2001	2001	100%	0.0	100.00%	7600	<a href="#">KF039731.1</a>
<input checked="" type="checkbox"/>	<a href="#">Norovirus Hu/GI.1/CHA9A004_20110419/2011/USA_complete genome</a>	<a href="#">Norovirus Hu/GI...</a>	2001	2001	100%	0.0	100.00%	7600	<a href="#">KF039730.1</a>
<input checked="" type="checkbox"/>	<a href="#">Norovirus Hu/GI.1/CHA6A007/2010/USA_complete genome</a>	<a href="#">Norovirus Hu/GI...</a>	2001	2001	100%	0.0	100.00%	7600	<a href="#">KF039729.1</a>

Slika 1

Iz zadetkov ugotovimo, da je nukleotidno zaporedje iz človeškega norovirusa GI.

### 2. Del

- Norovirusi pri ljudeh povzročajo predvsem akutni gastroenteritis.
- Prvi članek, ki ima v naslovu norovirus GI, je bil objavljen leta 2003. V članku so virus določili z reverzno transkripcijo in uporabo PCR-ja.

[Detection of norovirus \(GI, GII\), Sapovirus and astrovirus in fecal samples using reverse transcription single-round multiplex PCR.](#)

Cite Yan H, Yagyu F, Okitsu S, Nishio O, Ushijima H.

J Virol Methods. 2003 Dec;114(1):37-44. doi: 10.1016/j.jviromet.2003.08.009.

PMID: 14599677

Slika 2

- Norovirus sodi v družino Caliciviridae (advanced search pod title/abstract napišeš famij).

### 3. Del

- Iz blasta gremo na genebank in znotraj zaporedja (gledamo zaporedje celotnega genoma) najdemo, da Norovirus GI ima proteazo. Ta proteaza 3C-like proteaza imenovana tudi calicivirin.

```
mat_peptide      /note= vrg
                  3302..3844
                  /gene="POL"
                  /locus_tag="H649_43482gpPOL"
                  /product="3C-like protease"
                  /note="3CLpro; calicivirin"
```

Slika 3

- Na uniprotu poiščemo Norovirus GI. Dobimo tri anotirane zadetke in kliknemo na tistega, kjer je anotiran poliprotein (koda: **Q83883**).

- Pod PTM/Processing najdemo 3C-like proteazo, vidimo da se proteaza se v poliproteinu norovirusa GI nahaja med aminokislinskimi ostanki 1101-1281.

+ Chain	PRO_0000460162	124-341	NS2	<a href="#">Tools</a> <a href="#">Add</a>
+ Chain	PRO_0000341619	399-761	NTPase	<a href="#">Tools</a> <a href="#">Add</a>
+ Chain	PRO_0000341620	762-962	NS4	<a href="#">Tools</a> <a href="#">Add</a>
+ Chain	PRO_0000341621	963-1100	Viral genome-linked protein	<a href="#">Tools</a> <a href="#">Add</a>
+ Modified residue		992	O-(5'-phospho-RNA)-tyrosine <a href="#">By Similarity</a>	
- Chain	PRO_0000341622	1101-1281	3C-like protease	<a href="#">Tools</a> <a href="#">Add</a>
<p><b>Sequence:</b>            APPTLWSRVTKFGSGWGFVVSPTVFIITTHVVPTGVKEFFGEPLSSIAIHQAGEFTQFRFSKMKRPDLTGMVLEEGCPEGTVCSVLIK            RDSGELLPLAVRMGAIASMRIQGRVLHGQSGMLLTGANAKGMDLGTIPGDCCGAPYVHHRGNDWVVCVHAAATKSGNTVVCAV            QAGEGETALE</p>				
+ Chain	PRO_0000341623	1282-1789	RNA-directed RNA polymerase	<a href="#">Tools</a> <a href="#">Add</a>

Slika 4

- V UniProtu pod features najdemo aktivno mesto proteaze in to so aminokislinski ostanki H, C in E.

Sequence: H				
Active site	1154	For 3CLpro activity	<a href="#">PROSITE-ProRule Annotation</a>	<a href="#">2 Publications</a>
Sequence: E				
Active site	1239	For 3CLpro activity	<a href="#">PROSITE-ProRule Annotation</a>	<a href="#">2 Publications</a>
Sequence: C				

Slika 5

- Pod fenotipi in variantami je napisano, da če E v aktivnem mestu spremenimo v G ali pa C v aktivnem mestu spremenimo v A je proteaza ni več aktivna.
- Proteaza se aktivira z avtokatalitično cepitvijo.
- RNA-polimeraza, ki je tudi del poliproteina potrebuje dva kofaktroja in sicer  $Mg^{2+}$  in  $Mn^{2+}$ .

#### 4. Del

- Človeška proteaza ima 3 aminokislinske ostanke več kot mišja.
- Za poravnavo mišje in človeške norovirusne proteaze je smiselno uporabiti lokalno poravnavo, ki je pokazana na sliki 6, saj gre za proteina s podobno funkcijo. Zato aminokislinski zaporedji proteaz poravnamo z water.
- Poravnava je identična v približno 60%.
- Poravnava se začne na prvem aminokislinskem ostanku.
- Aktivno mesto je poravnano. V proteazi mišjega norovirusa imamo v aktivnem mestu namesto aminokislina E (v proteazi človeškega norovirusa) aminokislino D.



- Na podlagi teh podatkov sklepamo, da je mišji norovirus smiselni raziskovalni model človeškega norovirusa.

## 5. Del

- Z water naredimo poravnavo proteaze Norovirusa GI s celotnim aminokislinskim zaporedjem poliproteina Norovirusa GII.4. Rezultat je prikazan na sliki 8.

```

# Identity:      118/181 (65.2%)
# Similarity:   145/181 (80.1%)
# Gaps:        0/181 ( 0.0%)
# Score: 673.0
#
#
#=====

neanotiran      998 APPSIWSRIVNFGSGWGFVSPSLFITSTHVIPQGAQEFFGVPIKQIQIH    1047
      |||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
anotiran        1  APPTLWSRVTKFGSGWGFVSPVTFITTHVVPTGVKEFFGGEPLSSIAIH     50

neanotiran     1048 KSGEFCRLRFKPIRTDVGMILEEAGEPTVATLLIKRPTGELMPLAAR    1097
      ::|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
anotiran       51  QAGEFTQFRFSKMRPDLTGMVLEEGCEGTVCSVLIKRDSGELLPLAVR    100

neanotiran     1098 MGTHATMKIQGRTVGGQMGMLLTGSNAKSMDLGTTPGDCGCPYIYKREND    1147
      ||..|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
anotiran      101  MGAIASMRIQGRLVHGQSGMLLTGANAKGMDLGTIPGDCGAPYVHVRGND    150

neanotiran     1148 YVIVGVHTAAARGGNTVICATQGESEATLE      1178
      :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
anotiran      151  WVVCVGHAAATKSGNTVCAVQAGEGETALE      181

```

Slika 8

Iz tega predvidevamo, da ima tudi genogrupa GII.4 proteazo, ki se nahaja med 998 in 1178 mestom. To je podobno mesto kot za proteazo Norovirusa GI. V poravnavi ni nobenih vrzeli, poravnani pa sta celotni zaporedji proteaz iz obeh virusov.

- Aktivno mesto so aminokislinski ostanki 1026, 1050, 1135 iz poliproteina.

## 6. Del

- Z blastp iščemo po zbirki PDB in na dnu seznama najdemo Tulane virus (višji niso ustrezni, saj so norovirusi – Southampton virus in Chiba virus pa sta seva norovirusa). Ta se ujema v 40%.
- Tulane virus naj bi bil potencialno ustrezen model za preučevanje inhibitorjev človeškega norovirusa, kljub majhnemu ujemanju v aminokislinskem zaporedju.

<input checked="" type="checkbox"/>	Chain A, Peptidase C37 [Norovirus]	Norovirus	265	265	100%	2e-91	66.30%	182	<a href="#">9DF5_A</a>
<input checked="" type="checkbox"/>	Chain A, Peptidase C37 [Norovirus]	Norovirus	265	265	100%	3e-91	65.75%	181	<a href="#">9DA0_A</a>
<input checked="" type="checkbox"/>	Chain A, Peptidase C37 [Norovirus]	Norovirus	263	263	100%	2e-90	65.75%	183	<a href="#">9D9T_A</a>
<input checked="" type="checkbox"/>	Chain A, Peptidase C37 [Norovirus]	Norovirus	262	262	100%	4e-90	65.75%	183	<a href="#">9DAI_A</a>
<input checked="" type="checkbox"/>	Chain A, Genome polyprotein [Norovirus Sydney 2212]	Norovirus Sydney...	256	256	95%	1e-87	66.28%	172	<a href="#">9T3I_A</a>
<input checked="" type="checkbox"/>	Chain A, RNA-dependent RNA polymerase domain of Protease-Polymerase precursor [Norovirus Hu/GII.4/Sydney/N...	Norovirus Hu/GII.4...	265	265	99%	5e-85	66.11%	707	<a href="#">9BI9_A</a>
<input checked="" type="checkbox"/>	Chain A, NS6 Protease [Murine norovirus_1]	Murine norovirus_1	215	215	100%	2e-71	57.46%	185	<a href="#">4X2V_A</a>
<input checked="" type="checkbox"/>	Chain A, NS6 PROTEASE [Murine norovirus_1]	Murine norovirus_1	215	215	100%	2e-71	57.46%	185	<a href="#">4ASH_A</a>
<input checked="" type="checkbox"/>	Chain A, NS6 protease [Murine norovirus_1]	Murine norovirus_1	212	212	98%	2e-70	57.30%	176	<a href="#">4X2X_A</a>
<input checked="" type="checkbox"/>	Chain A, NS6 Protease [Murine norovirus_1]	Murine norovirus_1	212	212	98%	2e-70	57.30%	179	<a href="#">4X2Y_A</a>
<input checked="" type="checkbox"/>	Chain A, NS6 Protease NS6 Protease [Murine norovirus_1]	Murine norovirus_1	212	212	98%	3e-70	57.30%	179	<a href="#">4X2Z_A</a>
<input checked="" type="checkbox"/>	Chain A, 3C-like protease [Tulane virus]	Tulane virus	40.4	40.4	40%	6e-04	31.51%	177	<a href="#">9ZFQ_A</a>
<input checked="" type="checkbox"/>	Chain A, 3C-like protease [Tulane virus]	Tulane virus	40.4	40.4	40%	7e-04	31.51%	178	<a href="#">9ZFN_A</a>

Slika 9

## 7. Del

- Na PDB najdemo kodi za Norovirusno GI proteazo (2FYQ) in proteazo Tulane virusa (9ZFN). V Chimerix naredimo superpozicijo. Primer superpozicije je prikazan na sliki 10.



Slika 10

- Strukturi sta si precej podobni. Zaporedje sekundarnih struktur se precej dobro ujema. Strukturi se večinoma razlikujeta v nestrukturiranih regijah.
- Glede na strukturo izgleda Tulane virus kot potencialno ustrezeni nadomestni model.
- Rezultati iz PDBeFolda prikazani na sliki 11 kažejo na dobro strukturno ujemanje proteaze Tulane virusa s protezami iz norovirusov.

Examined 198545 entries, (581769 chains). Displaying Matches 1-20 of 147.

Back to query next last page Sort by Q-score arrange by SCOP family match 1 jump

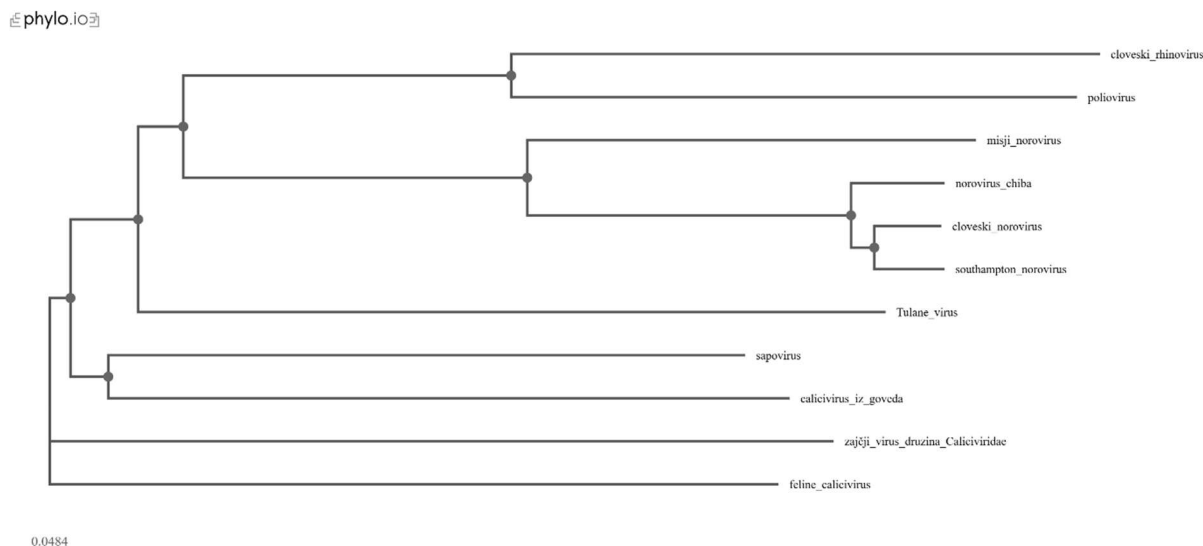
#	Scoring			RMSD	N <sub>align</sub>	N <sub>g</sub>	N <sub>seq</sub>	Query	Target (PDB entry)				Title
	Q	P	Z						% <sub>seq</sub>	Match	% <sub>seq</sub>	N <sub>res</sub>	
1	0.51	11.7	10.2	1.69	141	7	26	85	6w5j:A	92	170	<input type="checkbox"/>	1.85 A RESOLUTION STRUCTURE OF NOROVIRUS 3CL PROTEASE IN COMPLEX WITH INHIBITOR 7D
2	0.51	10.2	10.1	1.79	144	8	25	85	3ur6:B	85	173	<input type="checkbox"/>	1.5A RESOLUTION STRUCTURE OF APO NORWALK VIRUS PROTEASE
3	0.50	10.6	10.3	1.74	139	7	26	77	6b1b:A	77	165	<input type="checkbox"/>	1.95 A RESOLUTION STRUCTURE OF NOROVIRUS 3CL PROTEASE IN COMPLEX WITH A TRIAZOLE-BASED MACROCYCLIC INHIBITOR
4	0.50	11.3	10.0	1.96	143	6	25	85	5t6g:B	100	164	<input type="checkbox"/>	2.45 A RESOLUTION STRUCTURE OF NOROVIRUS 3CL PROTEASE IN COMPLEX WITH THE DIPEPTIDYL INHIBITOR 7M (HEXAGONAL FORM)
5	0.50	11.5	10.1	1.89	140	6	26	77	5t6d:B	91	161	<input type="checkbox"/>	2.10 A RESOLUTION STRUCTURE OF NOROVIRUS 3CL PROTEASE IN COMPLEX WITH THE DIPEPTIDYL INHIBITOR 7L (HEXAGONAL FORM)
6	0.50	10.3	10.2	1.79	141	7	26	77	4inh:H	77	168	<input type="checkbox"/>	STRUCTURAL BASIS OF SUBSTRATE SPECIFICITY AND PROTEASE INHIBITION IN NORWALK VIRUS
7	0.50	10.5	9.7	1.87	140	8	26	77	6b1c:A	77	162	<input type="checkbox"/>	2.25 A RESOLUTION STRUCTURE OF NOROVIRUS 3CL PROTEASE IN COMPLEX WITH A TRIAZOLE-BASED MACROCYCLIC INHIBITOR
8	0.50	9.6	9.9	1.82	140	8	25	77	5d6s:A	77	165	<input type="checkbox"/>	2.35A RESOLUTION STRUCTURE OF NOROVIRUS 3CL PROTEASE IN COMPLEX AN OXADIAZOLE-BASED, CELL PERMEABLE MACROCYCLIC (21-MER) INHIBITOR

Slika 11

- Eksperimentalno je boljše določena struktura Norovirusa GI, saj ima boljše ločljivost in R.

## 8. Naloga

- Večina izbranih 3C-like proteaz iz katerih sem naredila filogenetsko drevo, ki je prikazano na sliki 12, spada v družino Caliciviridae, kamor spadajo tudi norovirusi. Ti vključujejo človeške in mišje noroviruse, Tulane virus, sapovirus, feline calicivirus, rabbit hemorrhagic disease virus in bovine nebovirus. Za bolj oddaljeno primerjavo so bili dodani še predstavniki drugih virusnih družin, kot so human rhinovirus, poliovirus.
- Filogenetsko drevo narejeno na phylo.io:



Slika 12

- Filogenetsko drevo kaže, da se človeški norovirusi združujejo v jasno skupino. Najbližje sta si človeški norovirus in Southampton norovirus, medtem ko je Chiba norovirus nekoliko bolj oddaljen, vendar še vedno znotraj iste skupine. Mišji norovirus se nahaja blizu človeških norovirusov, vendar se od njih odcepi prej. To

potrjuje, da je soroden in primeren kot modelni virus, ni pa enak človeškim norovirusom. Tulane virus je od norovirusne skupine bolj oddaljen, vendar je še vedno uporaben za širšo primerjavo proteaz sorodnih kalicivirusov. Sapovirus, goveji kalicivirus, zajčji kalicivirus in feline kalicivirus so bolj oddaljeni predstavniki družine Caliciviridae. Rhinovirus in poliovirus tvorita ločeno skupino, kar je pričakovano, saj nista del družine Caliciviridae, ampak družine Picornaviridae.