

# Mogućnost ugradnje virusa SARS-CoV-2 u ljudski genom

---

Lesinger, Sara

Undergraduate thesis / Završni rad

2022

Degree Grantor / Ustanova koja je dodijelila akademski / stručni stupanj: **University of Zagreb, Faculty of Science / Sveučilište u Zagrebu, Prirodoslovno-matematički fakultet**

Permanent link / Trajna poveznica: <https://um.nsk.hr/um:nbn:hr:217:077953>

Rights / Prava: [In copyright](#)/[Zaštićeno autorskim pravom.](#)

Download date / Datum preuzimanja: **2024-11-06**



Repository / Repozitorij:

[Repository of the Faculty of Science - University of Zagreb](#)



Sveučilište u Zagrebu  
Prirodoslovno-matematički fakultet  
Biološki odsjek

Sara Lesinger

**Mogućnost ugradnje virusa SARS-CoV-2 u  
ljudski genom**

Završni rad

Zagreb, 2022.

University of Zagreb  
Faculty of Science  
Department of Biology

Sara Lesinger

**The possibility of integration of SARS-CoV-2  
virus into the human genome**

Bachelor thesis

Zagreb, 2022.

Ovaj završni rad je izrađen u sklopu studijskog programa Biologije na Zavodu za molekularnu biologiju Biološkog odsjeka Prirodoslovno-matematičkog fakulteta u Zagrebu, pod mentorstvom izv. prof. dr. sc. Ivane Ivančić Baće.

# TEMELJNA DOKUMENTACIJSKA KARTICA

---

Sveučilište u Zagrebu  
Prirodoslovno-matematički fakultet  
Biološki odsjek

Završni rad

## Mogućnost ugradnje virusa SARS-CoV-2 u ljudski genom

Sara Lesinger

Rooseveltov trg 6, 10000 Zagreb, Hrvatska

*Izješća o produljenoj detekciji RNA teškog akutnog respiratornog sindroma koronavirusa 2 (SARS-CoV-2) i o kontinuiranim pojavama PCR-pozitivnih testova kod pacijenata mnogo tjedana nakon oporavka od bolesti COVID-19, dovela su u pitanje mogućnost da se RNA virusa SARS-CoV-2 može prepisati reverznom transkripcijom i ugraditi u DNA ljudskih stanica. Zhang i sur. (2021) istražili su mogućnost bi li transkripcija ugrađenih sekvenci mogla objasniti neke od pozitivnih PCR testova viđenih kod pacijenata. Otkrili su da se kopije sekvenci DNA SARS-CoV-2 mogu ugraditi u genom zaraženih ljudskih stanica i dokaze koji upućuju na to da se veliki dio virusnih sekvenci prepisuje iz ugrađenih sekvenci DNA, stvarajući ljudsko-virusne kimerne transkripte. Parry i sur. (2021) i Smits i sur. (2021.) smatrali su potrebnim ponnije razmotriti predstavljene dokaze. Pokazali su da su podaci neprikladno interpretirani i ponovno su analizirali objavljene podatke. Nisu našli dokaze za genomsku ugradnju virusa, unatoč dokazanoj prisutnosti predloženog mehanizma. Činjenica da se retrotranskripcija i ugradnja SARS-CoV-2 u genom kod pacijenata događa u značajnoj učestalosti ili čak uopće, ostaje malo vjerojatna.*

**Ključne riječi:** reverzna transkripcija | ugradnja | LINE1 | sekvenciranje | SARS-CoV-2  
(16 stranica, 4 slike, 1 tablica, 27 literaturnih navoda, jezik izvornika: hrvatski)

Rad je pohranjen u Središnjoj biološkoj knjižnici

Mentor: izv. prof. dr. sc. Ivana Ivančić Baće

# BASIC DOCUMENTATION CARD

---

University of Zagreb  
Faculty of Science  
Department of Biology

Bachelor thesis

## The possibility of integration of SARS-CoV-2 virus into the human genome

Sara Lesinger

Rooseveltova trg 6, 10000 Zagreb, Croatia

*There have been reports about prolonged detection of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) RNA and recurrence of PCR-positive tests in patients many weeks after recovery from COVID-19. This raised the possibility that SARS-CoV-2 RNAs can be reverse-transcribed and integrated into the DNA of human cells. Zhang et al. (2021) investigated possibility whether transcription of the integrated sequences might account for some of the positive PCR tests seen in patients. They found that DNA copies of SARS-CoV-2 sequences can be integrated into the genome of infected human cells and evidence suggesting that a large fraction of the viral sequences are transcribed from integrated DNA copies, generating viral–host chimeric transcripts. Parry et al. (2021) and Smits et al. (2021) felt necessary to scrutinize the evidence presented. They showed that the data has been inappropriately interpreted and they reanalyzed published data. They found no evidence for genomic integration of the virus, despite demonstrated availability of the proposed machinery. The fact that retrotranscription and integration of the SARS-CoV-2 genome happens in patients at any notable frequency, or even at all, remains unlikely.*

**Keywords:** reverse transcription | integration | LINE1 | sequencing | SARS-CoV-2

(16 pages, 4 figures, 1 table, 27 references, original in: Croatian)

Thesis is deposited in Central Biological Library.

Mentor: izv. prof. dr. sc. Ivana Ivančić Baće

## SADRŽAJ

1.	<i>UVOD</i> .....	1
2.	<i>UGRADNJA SARS-COV-2 U LJUDSKI GENOM</i> .....	3
2.1.	Ugradnja sekvenci SARS-CoV-2 u DNA stanice domaćina u kulturi .....	3
2.1.1.	Nanopore sekvenciranje dugih sekvenci .....	3
2.1.2.	Illumina upareno sekvenciranje krajeva cijelog genoma .....	4
2.1.3.	Sekvenciranje ugrađenih DNA obilježenih Tn5 transpozonom .....	5
2.2.	Ekspresija virusno-staničnih kimernih transkripata u inficiranim stanicama i tkivima u kulturi iz pacijenata .....	5
2.3.	Kritički osvrt .....	7
3.	<i>UGRADNJA SARS-COV-2 U LJUDSKI GENOM NIJE MOGUĆA</i> .....	9
3.1.	Nema dokaza o reverznoj transkripciji i ugradnji SARS-CoV-2 u tkivu pacijenta .....	9
3.2.	DNA sekvenciranje dugih sekvenci ne otkriva dokaze ugradnje SARS-CoV-2 u ljudski genom .....	10
4.	<i>ZAKLJUČAK</i> .....	13
5.	<i>LITERATURA</i> .....	14
6.	<i>O SEBI</i> .....	16

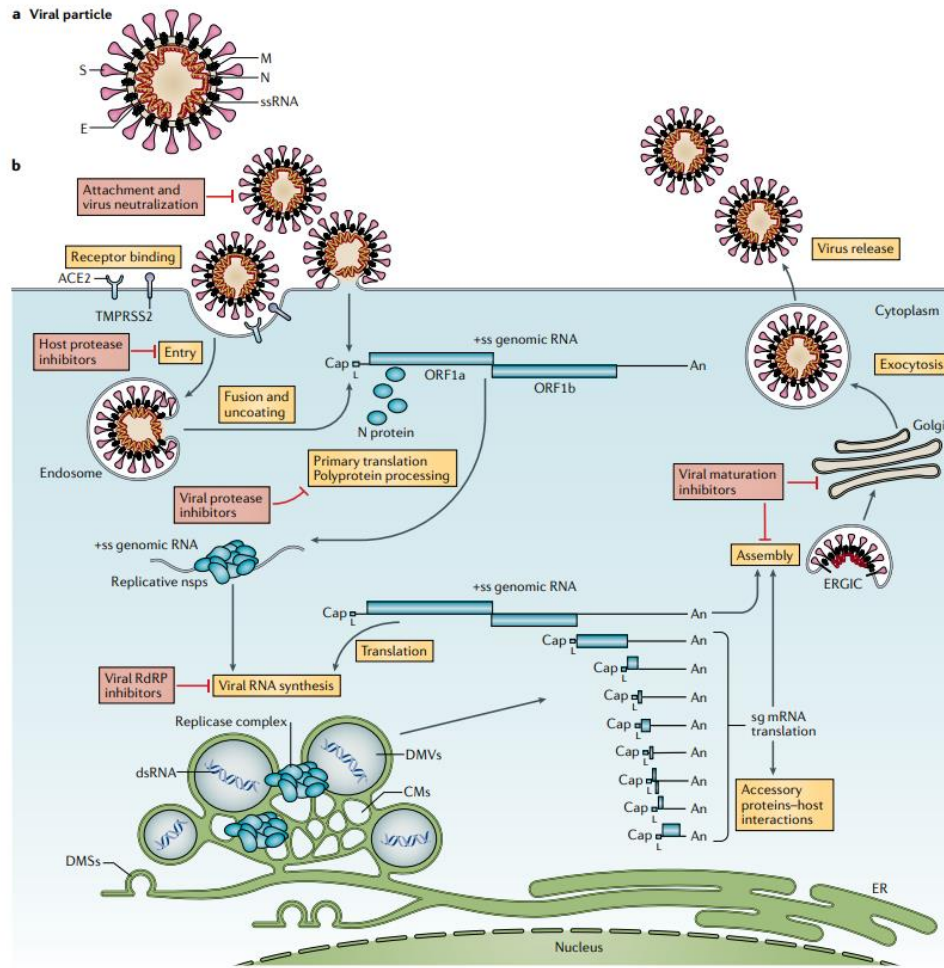
# 1. UVOD

Teški akutni respiratorni sindrom koronavirus 2 (engl. *Severe acute respiratory syndrome coronavirus 2*), skraćeno SARS-CoV-2, vrlo je prenosivi patogeni koronavirus koji se pojavio 2019. godine i uzrokovao pandemiju akutne respiratorne bolesti “koronavirusna bolest 2019” (COVID-19). Uz SARS-CoV-2, postoji sedam ljudskih koronavirusa od kojih neki uzrokuju blage bolesti gornjega dišnog te probavnog sustava čovjeka. Zoonotski koronavirusi potječu od životinja i prijenosom mogu uzrokovati zarazne bolesti kod ljudi (Ge i sur. 2013). Vjerovalo se da SARS-CoV-2 potječe od šišmiša, no kineski znanstvenici iznijeli su hipotezu o pangolinima kao životinjskom izvoru ovog virusa (Zhang i sur. 2020).

SARS-CoV-2 pripada skupini virusa iz porodice *Coronaviridae*. Koronavirusi su kuglastog oblika promjera 120 do 160 nm i obavijeni ovojnicom na kojoj se nalaze šiljasti transmembranski glikoproteinski izdanci duljine 12 do 24 nm koji joj daju izgled krune (Hrvatska enciklopedija). Genom korona virusa se sastoji od jednolančane pozitivne RNA sastavljene od 26 do 32 tisuće nukleotida, što ga čini najvećim među RNA virusima (Ravi i sur. 2022). Virusna čestica sadrži četiri glavna strukturna proteina: protein šiljka (engl. *spike protein*, S-protein), protein membrane (M), protein ovojnice (E) i nukleokapsidni protein (N) (Slika 1a). Virus se na stanicu domaćina veže pomoću S-proteina na angiotenzin-konvertirajući enzim 2 (ACE2) receptor (Cevik i sur. 2020). Nakon vezanja na receptor, virus ulazi u citoplazmu stanice domaćina i dolazi do fuzije virusne ovojnice i citoplazmatske membrane nakon čega slijede translacija i replikacija virusne genomske RNA te formiranje novih virusnih čestica. Virioni se zatim transportiraju na površinu stanične membrane pomoću vezikula i otpuštaju se egzocitozom (Yadav i sur. 2021) (Slika 1b).

SARS-CoV-2 pripada rodu betakoronavirusa kao i MERS-CoV (engl. *Middle East respiratory syndrome-related coronavirus*) i SARS-CoV (engl. *severe acute respiratory syndrome coronavirus*) koji koriste RNA-ovisnu RNA polimerazu (RdRP) za sintezu RNA kod replikacije genomske RNA i transkripcije subgenomskih RNA (V’Kovski i sur. 2021, Alanagreh i sur. 2020). Genetskom rekombinacijom unutar istih ili različitih skupina virusa povećava se raznolikost genoma što može dovesti do pojave novih varijanti virusa (Loomba i sur. 2020). Tijekom pandemije bolesti COVID-19, pojavile su se nove varijante SARS-CoV-2 virusa koje su podijeljene na dvije skupine: varijante od značaja (engl. *Variants of Concern*, VOCs) i varijante od interesa (engl. *Variants of Interest*, VOIs) (WHO 2022).





**Slika 1.** Virion koronavirusa i životni ciklus (Izvor: V’Kovski i sur. 2021)

- a) Virion koronavirusa sastoji se od strukturnih proteina: protein šiljka (S), protein membrane (M), protein ovojnice (E), nukleokapsidni protein (N). N zatvara pozitivni jednolančani RNA genom (+ssRNA), dok M i E osiguravaju njegovu ugradnju u virusnu česticu tijekom nastajanja novog virusa. S izdanci pružaju specifičnost za receptore pri ulasku u stanicu.
- b) Čestice koronavirusa vežu se za staničnu membranu pomoću staničnih receptora kao što je angiotenzin-konvertirajući enzim 2 (ACE2), a stanični faktori kao što je transmembranska serinska proteaza tipa 2 (TMPRSS2) pomažu kod ulaska virusa i njegove fuzije sa staničnom membranom. Nakon ulaska u stanicu, virusna RNA se oslobađa i počinje translacija dva velika otvorena okvira čitanja, ORF1a i ORF1b. Obradom dobivenih poliproteina pp1a i pp1ab nastaju pojedinačni nestrukturni proteini (nsps). Njihovom ekspresijom nastaje pogodni okoliš za replikaciju RNA i transkripciju subgenomskih mRNA (sg mRNAs). Translatirani strukturni proteini translociraju se u membrane endoplazmatskog retikuluma (ER) i prolaze kroz intermedijer ER-Golgi (ERGIC). Formirani virioni izlaze iz inficirane stanice egzocitozom. 3’ poliadenilirani rep (An), 5’ kapica (Cap), dvolančana RNA (dsRNA), vodeći slijed (L), RNA-ovisna RNA polimeraza (RdRP).

## **2. UGRADNJA SARS-COV-2 U LJUDSKI GENOM**

Zbog detekcije RNA virusa SARS-CoV-2 tjednima nakon inicijalne infekcije i kontinuiranog pojavljivanja pozitivnih rezultata PCR (engl. *polymerase chain reaction*) testova kod pacijenata koji su se oporavili od bolesti COVID-19, u ovom istraživanju Zhang i sur. (2021) istražili su vjerojatnost reverzne transkripcije i ugradnje virusne DNA u DNA ljudskih stanica u kulturi te mogućnost transkripcije ugrađenih sekvenci kao uzrok pozitivnih PCR testova kod pacijenata.

### **2.1. Ugradnja sekvenci SARS-CoV-2 u DNA stanice domaćina u kulturi**

Za istraživanje mogućnosti ugradnje sekvenci SARS-CoV-2 u DNA stanice domaćina u kulturi korištene su tri metode za detekciju genomske virusne sekvence unutar genoma zaraženih stanica: Nanopore sekvenciranje dugih sekvenci (engl. *Nanopore long-read sequencing*), Illumina upareno sekvenciranje krajeva cijelog genoma (engl. *Illumina paired-end whole genomic sequencing*) i sekvenciranje ugrađenih DNA obilježenih Tn5 transpozonom (engl. *Tn5 tagmentation-based DNA integration site enrichment sequencing*). Kako bi se povećala vjerojatnost prepoznavanja rijetkih ugradnji, potaknuta je prekomjerna ekspresija LINE1 (engl. *long interspersed nuclear element, L1*) u stanicama HEK293T transfekcijom pomoću ekspresijskog plazmida L1. Stanice su inficirane virusom SARS-CoV-2 i nakon dva dana iz njih je izolirana DNA i umnožena metodom PCR. Identificirana je kopija sekvenci DNA nukleokapside i predstavljena kao dokaz reverzne transkripcije i ugradnje RNA SARS-CoV-2 u genom stanice domaćina.

#### **2.1.1. Nanopore sekvenciranje dugih sekvenci**

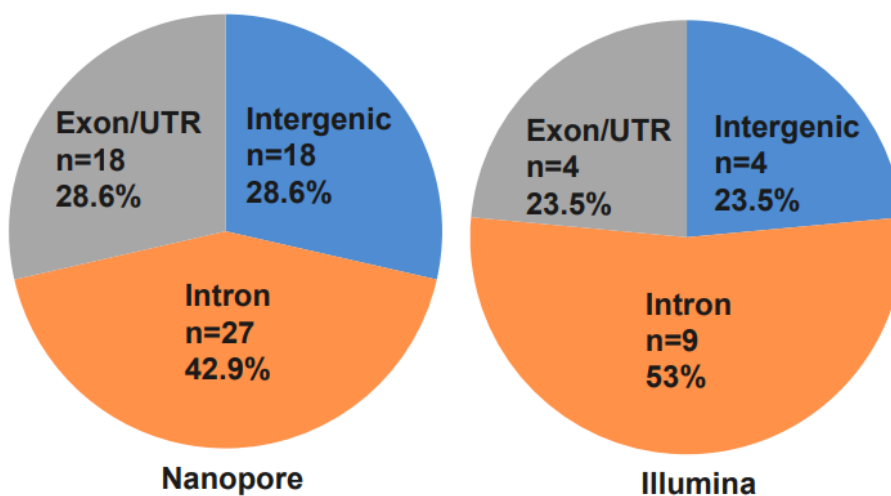
Kako bi direktno dokazali ugradnju sekvenci SARS-CoV-2 u genom inficiranih stanica HEK293T, njihova DNA podvrgnuta je Nanopore sekvenciranju. Otkrivene su ugrađene virusne sekvence DNA okružene duplikacijama ciljnog mjesta (engl. *target site duplication, TSD*) koje su sadržavale sekvence prepoznavanja endonukleaze L1. Ovi rezultati potiču na zaključak da sekvence virusa SARS-CoV-2 mogu biti ugrađene u genom ljudskih stanica u kulturi mehanizmom retropozicije pomoću retrotranspozona L1 (Lanciano i Cristofari 2020). Otprilike 71% navedenih ugrađenih virusnih sekvenci bilo je okruženo intronima ili intergenskim sekvencama, a 29%

egzonima (Slika 2). Kako se ljudski genom sastoji od 1,1% egzona, 24% introna i 75% intergenske DNA (Venter i sur. 2001), očekivalo se da nasumičnom ugradnjom nastaje puno manja povezanost s egzonima, a ovako visoki udio ugrađenih sekvenci pronađen unutar egzona implicira na preferencijalnu retropoziciju L1 u egzone.

### 2.1.2. Illumina upareno sekvenciranje krajeva cijelog genoma

Da bi se dodatno potvrdila ugradnja sekvenci SARS-CoV-2 u DNA ljudskog genoma, DNA iz transfeciranih i inficiranih stanica HEK293T korištena je u Illumina sekvenciranju. Sekvence virusne DNA bile su koncentrirane na 3' kraju SARS-CoV-2 genoma. Utvrđeno je 17 ugrađenih virusnih sekvenci mapiranjem ljudsko-virusnih kimernih DNA. 7 kimernih sekvenci (41%) sadržavalo je varijantu sekvence prepoznavanja endonukleaze L1, a slično kao i u prethodnom eksperimentu, približno 76% ugrađenih virusnih sekvenci bilo je okruženo intronima ili intergenskom DNA, a 24% egzonima (Slika 2).

U ukupno oko 32% ugrađenih sekvenci SARS-CoV-2, za 6 od 21 iz podataka Nanopore sekvenciranja i 4 od 10 iz Illumina podataka, nije pronađena sekvenca prepoznavanja endonukleaze L1 što upućuje na to da mehanizam retropozicije pomoću retrotranspozona L1 nije jedini mehanizam ugradnje, već da postoje i alternativni mehanizmi.



**Slika 2.** Raspodjela ugradnji DNA virusa SARS-CoV-2 u ljudskom genomu (Izvor: Zhang i sur. 2021)

Kružni dijagrami prikazuju postotak ljudsko-virusnih spojeva pronađenih unutar egzona, introna i intergenske DNA; Nanopore sekvenciranje (lijevo), Illumina sekvenciranje (desno).

### **2.1.3. Sekvenciranje ugrađenih DNA obilježenih Tn5 transpozonom**

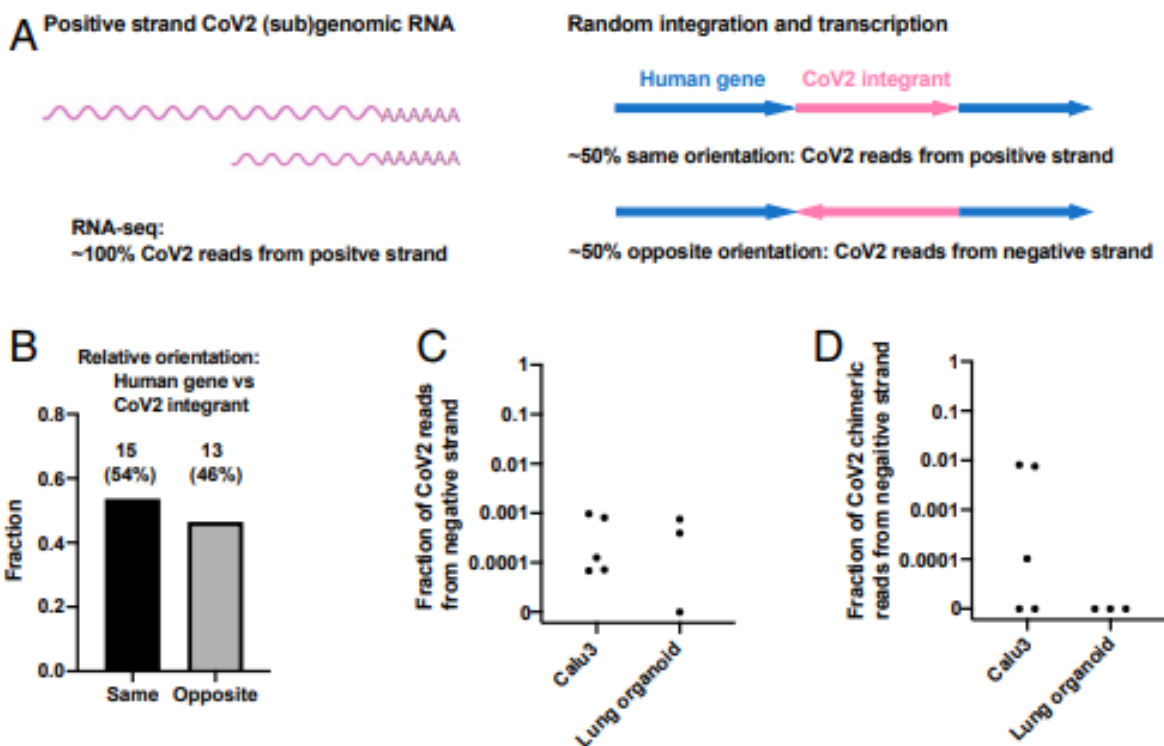
Za procjenu mogućnosti ugradnje sekvenci SARS-CoV-2 u genom inficiranih stanica koje nemaju prekomjernu ekspresiju retrotranspozona L1, inficirane stanice HEK293T obilježene su Tn5 transpozonom i sekvencirane. Detektirano je 7 ljudsko-virusnih kimernih sekvenci, a sve su sadržavale sekvencu prepoznavanja endonukleaze L1.

## **2.2. Ekspresija virusno-staničnih kimernih transkripata u inficiranim stanicama i tkivima u kulturi iz pacijenata**

Kako bi provjerili dolazi li do prepisivanja ugrađenih sekvenci SARS-CoV-2, čime bi nastali kimerni transkripti, analizirani su objavljeni podaci sekvenciranja RNA. Kimerne sekvence čine 0,004-0,14% svih sekvenciranih sekvenci SARS-CoV-2 i uglavnom se poklapaju sa sekvencom gena nukleokapside, što je vjerojatno povezano s time da njene RNA ima najviše pa je češća meta za reverznu transkripciju i ugradnju. Međutim, 1% kimernih sekvenci može biti posljedica grešaka nastalih prebacivanjem reverzne transkriptaze prilikom sinteze komplementarne DNA (engl. *complementary DNA*, cDNA) u pripremi biblioteke za sekvenciranje RNA. S obzirom na to da se u inficiranim stanicama nalaze mRNA (engl. *messenger RNA*) domaćina i pozitivnih lanaca virusne mRNA, identifikacija pravih kimernih transkripata je ugrožena zbog nastajanja lažnih kimera.

Znanstvenici su pretpostavili da bi nasumičnom ugradnjom polovica ugrađenih DNA kopija od SARS-CoV-2 trebala biti u suprotnoj orijentaciji od smjera transkripcije gena domaćina, odnosno da bi polovica nastalih kimernih sekvenci trebala sadržavati negativni lanac virusne sekvence povezan s pozitivnim lancem ljudske sekvence RNA (Slika 3A i 3B).

Replikacija virusne RNA podrazumijeva sintezu negativnog lanca RNA koji služi kao kalup za replikaciju genomske RNA i transkripciju subgenomskih pozitivnih lanaca RNA. U akutno inficiranim stanicama Calu3 i plućnim organoidima, između 0 i 0,1% virusnih pročitanih sekvenci dolazilo je od negativnog lanca RNA (Slika 3C i 3D). To znači da je udio pozitivnih lanaca RNA u tim stanicama barem 1000 puta viši, dijelom zahvaljujući velikoj količini stvorenih subgenomskih pozitivnih lanaca RNA tijekom replikacije virusa. Taj podatak znanstvenici su interpretirali kao dokaz da je vjerojatnost nastajanja lažnih kimera niska.



**Slika 3.** Pretpostavka orijentacije ugrađenih sekvenci i rezultati Nanopore sekvenciranja DNA stanica HEK293T s prekomjernom ekspresijom enzima retrotranspozona L1 (Izvor: Zhang i sur. 2021)

- Shema prikazuje pozitivne ili negativne lance RNA SARS-CoV-2 nastale od subgenomske RNA ili ugrađenih virusnih transkripata. Strelice predstavljaju smjer transkripcije gena domaćina (plavo) i orijentaciju ugrađenih sekvenci SARS-CoV-2 (ružičasto).
- Graf prikazuje postotak istosmjerno i suprotno orijentiranih ugrađenih sekvenci SARS-CoV-2.
- Graf prikazuje ukupni postotak virusnih sekvenci nastalih od negativnog lanca virusne RNA u stanicama Calu3 i plućnim organoidima.
- Kao za C), ali podaci su za kimerne sekvence.

Za razliku od akutno inficiranih stanica Calu3 i plućnih organoida, u tkivima iz pacijenata do 51% ukupnih sekvenciranih virusnih sekvenci i do 42,5% kimernih sekvenci potjecalo je od negativnog lanca RNA SARS-CoV-2. Analiza pojedinih stanica bronhoalveolarnog lavata (engl. *bronchoalveolar lavage fluid*, BALF) pacijenata teško oboljelih bolešću COVID-19 pokazala je da 40% svih virusnih sekvenci potječe od negativnih lanaca virusne RNA, no takvi podaci dobiveni su u 4 od 6 pacijenata i u 4 od 14 uzoraka autopsije. U tim fiksiranim uzorcima gotovo da nije bilo dokaza reprodukcije virusa, za razliku od akutno inficiranih stanica Calu3. Općenito, uzorci uzeti

iz pacijenata sadržavali su viši udio negativnih lanaca virusnih sekvenci, nego akutno inficirane stanice. Zhang i sur. (2021) smatraju da njihove analize nisu uspjele dokazati prisutnost velikog broja kimernih sekvenci iz BALF uzoraka zbog ograničene sposobnosti sekvenciranja kratkih sekvenci (engl. *short-read sequencing*). Oni smatraju da njihovi podaci pokazuju da bi velika količina virusnih transkripata, prisutna u tkivima pacijenata gdje i nema puno stanica pozitivnih na SARS-CoV-2, mogla dolaziti od transkripata virusa SARS-CoV-2 ugrađenih u genom domaćina.

### **2.3. Kritički osvrt**

Objavljeni rad Zhang i sur. (2021) izazvao je popriličnu reakciju javnosti, naročito zbog istovremenog predstavljanja mRNA cjepiva protiv bolesti COVID-19. Mnogi znanstvenici izrazili su svoje mišljenje te ih je velika većina zahtjevala ispravke i dodatne analize te istraživanja na temelju objavljenih podataka. Predstavljeni materijali su upitni i nisu dovoljni za dokaz tvrdnje da se virus SARS-CoV-2 prepisuje reverznom transkripcijom i ugrađuje u ljudski genom.

SARS-CoV-2 je koronavirus i ne sadrži enzim reverznu transkriptazu, pomoću koje retrovirusi prepisuju RNA u DNA. Mnogi retrovirusi su uzročnici tumora i povezani s bolestima poput sindroma stečene imunodeficijencije (engl. *Acquired Immunodeficiency Syndrome, AIDS*) uzrokovanog virusom humane imunodeficijencije (engl. *Human Immunodeficiency Virus, HIV*) koji koristeći vlastitu reverznu transkriptazu stvara komplementarnu DNA koju retrovirusni enzim integraza zatim ugrađuje u genom stanice domaćina (Hrvatska enciklopedija).

No ljudska DNA sadrži retrotranspozone LINE1 (L1) koji, ako se aktiviraju, mogu uzrokovati insercije, delecije i duplikacije u genomu, a tako bi potencijalno mogli prepisati RNA virusa u DNA koja bi se mogla ugraditi u genom stanice domaćina. Ugradnja pomoću retrotranspozona L1 odvija se unutar mjesta cijepanja endonukleaze L1 (TTTT/AA), stvara duplikacije TSD i poliadenilirani rep na 3' kraju. Zhang i sur. (2021) tvrde da se ugradnja odvila uz pomoć retrotranspozona L1, a ne navode nijedan dokaz o obilježjima spomenute ugradnje.

Korištene stanice HEK293T prekomjerno stvaraju enzime retrotranspozona L1 što povećava mogućnost prepoznavanja rijetkih ugradnji. Pokazalo se da virusna infekcija također stimulira ekspresiju pokretnih elemenata poput L1, no to ne znači da će doći do ugradnje virusa u genom stanice domaćina. Ovakav mehanizam nikada nije otkriven kod infekcije HIV-om, retrovirusom koji je vrlo dobro proučen.

Analiza kimernih sekvenci pokazuje da su one iznimno rijetke i mogu nastati kao posljedica pogreške tijekom pripreme biblioteke za sekvenciranje RNA ili zbog prebacivanja reverzne transkriptaze u sintezi komplementarne DNA. Koronavirusi se repliciraju pomoću diskontinuirane transkripcije, koja uključuje prebacivanje RNA-ovisne RNA polimeraze (Slika 1b). U oba slučaja kimerne sekvence bi nastale spajanjem virusne RNA s ljudskom mRNA koja se sastoji od egzona. Kada bi kimerne sekvence nastajale prepisivanjem ugrađenih virusnih sekvenci, one bi bile povezane većinom s intronima ili intergenskom DNA jer oni čine ~ 99% ljudskog genoma. Detekcija kimernih RNA ne predstavlja dokaz za reverznu transkripciju i ugradnju SARS-CoV-2 u ljudski genom.

### **3. UGRADNJA SARS-COV-2 U LJUDSKI GENOM NIJE MOGUĆA**

#### **3.1. Nema dokaza o reverznoj transkripciji i ugradnji SARS-CoV-2 u tkivu pacijenta**

Parry i sur. (2021) također smatraju da je istraživanje i rezultate koje su predstavili Zhang i sur. (2021) potrebno ispitati zbog mogućih posljedica na dugoročne učinke COVID-19 zaraze.

Skupili su podatke i iznova ih interpretirali. Smatraju da niska učestalost prepoznavanja kimernih sekvenci (Tablica 1) ukazuje na to da ugradnja SARS-CoV-2 u ljudski genom nije vjerojatna. Stanice HEK293T transfecirane s ekspresijskim plazmidom L1 povećavaju mogućnost prepoznavanja rijetkih ugradnji, a L1 mogu retrotransponirati bilo koju poliadeniliranu staničnu RNA. Kako su ugrađene virusne sekvence većinom nađene u egzonima, a takva preferencijalna aktivnost endonukleaze L1 nije poznata (Sultana i sur. 2019), spomenute tvrdnje nisu vjerodostojne. Zhang i sur. (2021) tvrde da je puno veći udio kimernih sekvenci dobivenih iz negativnog lanca RNA u tkivima pacijenata nego onima u uvjetima *in vitro*, dokaz ugradnje i transkripcije SARS-CoV-2. Parry i sur. (2021) međutim smatraju prezentirane podatke upitnima za dokaz navedenih tvrdnji iz više razloga. Pojava kimernih sekvenci je česta kod sekvenciranja RNA, pa tako i kod SARS-CoV-2, zbog fuzije komplementarnih DNA tijekom reverzne transkripcije u pripremi biblioteke RNA (Tablica 1). Sekvence negativnih lanaca RNA u stanicama pacijenta su očekivana pojava jer se oni stvaraju tijekom replikacije SARS-CoV-2 i kalup su za mRNA. U prirodi su retrotranskripcija i ugradnja normalne pojave, a koronavirusi se nikad nisu ugradili u embrionalne stanice nekog domaćina, temeljeno na sustavnom probiru u više od 750 životinjskih vrsta (Katzourakis i Gifford 2010).



**Tablica 1.** Pročitane sekvence genomske DNA i kimernih sekvenci pomoću različitih metoda sekvenciranja (Izvor: Parry i sur. 2021)

Accession	Cell line	Sequencing strategy	Chimeric reads/total reads	Percent of library
SRR14289057	HEK293T	Illumina Enrichment	2/71,263,270	0.000003
SRR14216062	Calu3	Illumina Enrichment	3/94,274,616	0.000003
SRR14216061	Calu3	Illumina Enrichment	2/103,608,699	0.000002
SRR14163829	HEK293T-L1	Nanopore	61/12,053,919	0.0005
SRR14136237	HEK293T-L1	Illumina	8/109,004,386	0.000007
SRR14136236	HEK293T-L1	Illumina	9/178,953,858	0.000005

### 3.2. DNA sekvenciranje dugih sekvenci ne otkriva dokaze ugradnje

#### SARS-CoV-2 u ljudski genom

Zbog mogućih značajnih kliničkih implikacija ugradnje SARS-CoV-2 u DNA inficirane stanice, Smits i sur. (2021) napravili su istraživanje koristeći Oxford Nanopore Technologies (ONT) sekvenciranje dugih sekvenci (engl. *long-read sequencing*) kako bi ispitali vjerodostojnost podataka u radu Zhang i sur. (2021).

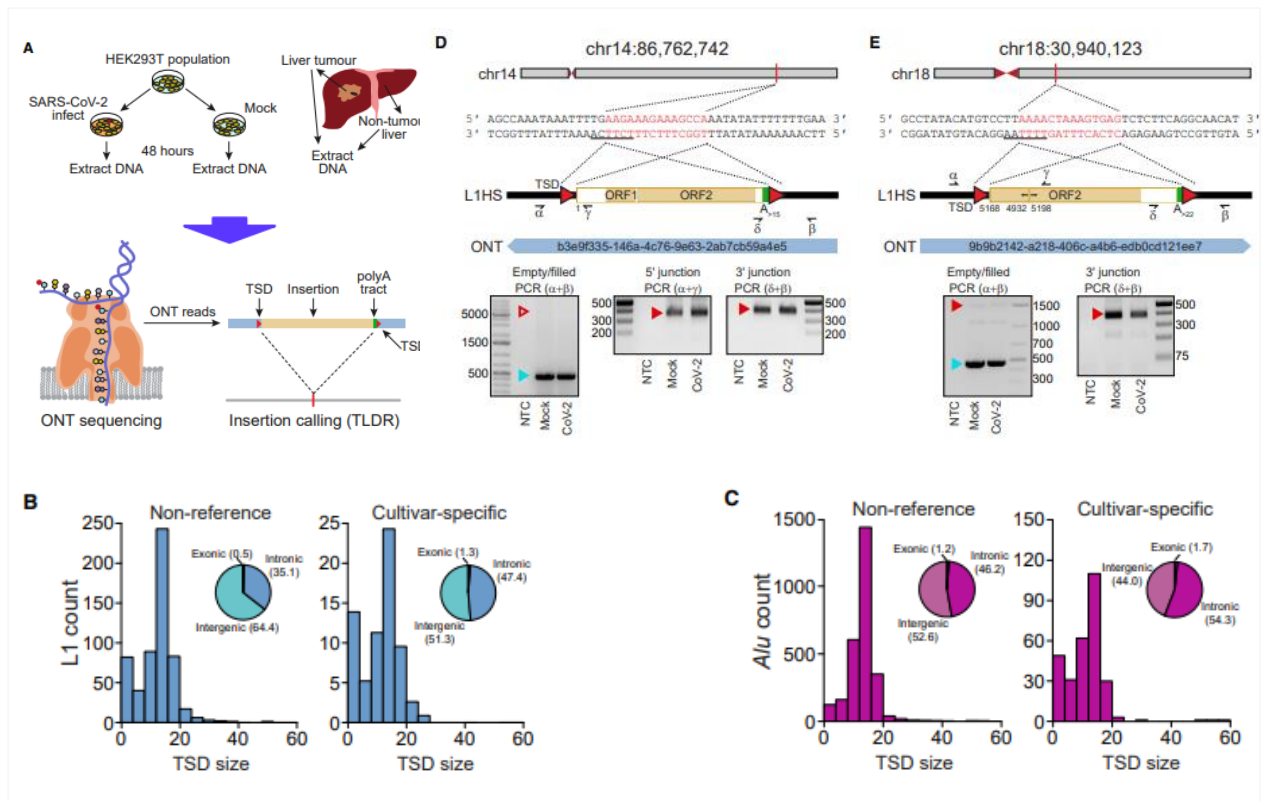
Sekvencirali su genomsku DNA iz stanica HEK293T inficiranih virusom SARS-CoV-2 i iz neinficiranih stanica HEK293T (Slika 4A). Tumorsko i normalno tkivo jetre pacijenta sa hepatocelularnim karcinomom (HCC), kojeg je uzrokovao virus hepatitisa B (HBV) (Prejac 2022), sekvencirano je kao pozitivna kontrola jer se HBV ugrađuje u mjesta oštećenja DNA. Kao negativna kontrola, korišteni su uzorci jetre i HCC-a pacijenta pozitivnog na virus hepatitisa C (HCV) koji se ne ugrađuje u genom domaćina. HCV je, kao i SARS-CoV-2, virus s jednolančanom pozitivnom RNA, ali ona ne sadrži 3' poliadenilirani rep koji je potreban za retropoziciju pomoću retrotranspozona L1. Svojim podacima dodali su i one od Zhang i sur. (2021) te su koristili softver TLDR (engl. *transposons from long DNA reads*) (Ewing i sur. 2020) za prepoznavanje ugrađenih sekvenci SARS-CoV-2, HBV i HCV kao i za ugradnje retrotranspozona koje sadrže barem jedno specifično poravnanje. Softverom TLDR nije identificirana ugradnja sekvenci SARS-CoV-2-a, HBV-a ni HCV-a (Smits i sur. 2021).

U stanicama inficiranima s virusom SARS-CoV-2 i neinficiranim stanicama, identificirano je 78 ugradnji L1 sa duplikacijama TSD veličine 14 bp (Slika 4B), od kojih je 69 detektirano

pomoću premošćujućeg sekvenciranja (engl. *single spanning read*), a 13 ih je sadržavalo 3' kraj nastao transdukcijom. Pokazalo se da stanice inficirane sa SARS-CoV-2 imaju puno više potencijalnih ugradnji L1, nego neinficirane stanice HEK293T. Zatim je odabrano 6 od 69 detektiranih ugradnji L1, koje su ručno provjerene: sve su sadržavale duplikacije TSD i 3' poliadenilirani rep te su smještene u motivu endonukleaze L1 (Slika 4D i 4E). Smits i sur. (2021) su zaključili da njihovi i prethodni (Ewing i sur. 2020) eksperimenti potvrđuju da se pomoću ONT sekvenciranja može vjerodostojno dokazati retropozicija i aktivnosti endonukleaze L1 u stanicama HEK293T bez prekomjerne ekspresije L1.

Pomoću direktnih poravnanja ONT sekvenci s genomom SARS-CoV-2, HBV, HCV i s mobilnim sekvencama L1HS (engl. *human-specific L1*) ispitali su mogućnost da su parametri računalne analize isključili prave ugradnje navedenih virusa. Ugradnje za genome SARS-CoV-2 i HCV nisu pronađene, a za HBV je pronađena samo jedna ugradnja u  $10^1$ - $10^4$  inficiranih hepatocita što dokazuje preciznost ONT sekvenciranja.

Ponovnom analizom podataka ONT sekvenciranja koje su proveli Zhang i sur. (2021), pronađeno je 555 sekvenci (od ukupno ~ 12 milijuna) koje su odgovarale sekvencama SARS-CoV-2 genoma prosječne veličine 924 bp. Osim što je ovo vrlo mali postotak, te sekvence bile su i neobično kraće od ostatka ukupnog seta podataka (2,686 kbp). Prosječni udio nukleotida SARS-CoV-2 unutar tih sekvenci bio je 52,3%. Za usporedbu, pročitane sekvence L1HS sadržavale su 17,1% nukleotida L1HS. Od svih 555 pročitanih sekvenci, 79 ih je sadržavalo i sekvence odgovarajuće ljudskom genomu. Pokazalo se da je 36,2% ugradnji SARS-CoV-2 u ljudski genom smješteno unutar egzona, dok je za ostale ugradnje L1 taj postotak puno manji (Slika 4B i 4C). Naposljetku se razmotrilo zašto softver TLDR nije pokazao dvije ugradnje virusa SARS-CoV-2 koje su u svom radu naveli Zhang i sur. (2021) i utvrdilo da su eliminirane zbog moguće dvosmislene interpretacije navedenih pročitanih sekvenci. Ugradnja na kromosomu X nije uzeta u obzir zbog kratkog poravnanja na 5' kraju, dok je sekvenca s kromosoma 22 bila komplementarna sekvenci s kromosoma 1. Analizom se potvrdila prisutnost sekvenci SARS-CoV-2 iz seta podataka (Zhang i sur. 2021), no identificirane sekvence su izuzetno kratke i mogle bi biti pogreške koje su Zhang i sur. (2021) interpretirali kao ugradnje virusa SARS-CoV-2.



**Slika 4.** Detekcija endogene retrotranspozicije pomoću retrotranspozona L1 u ljudskim stanicama (Izvor: Smits i sur. 2021, preuzeto i prilagođeno)

- E) Eksperimentalni dizajn. Stanice HEK293T podijeljene su na dva uzorka, od kojih je jedan inficiran virusom SARS-CoV-2, a drugi nije. Iz njih, kao i iz uzoraka pacijenata sa hepatocelularnim karcinomom (HCC), je izolirana DNA i podvrgnuta sekvenciranju ONT (Oxford Nanopore Technologies). Pomoću softvera TLDR (engl. *transposons from long DNA reads*), pronađene su nespecifične L1 i virusne ugradnje te duplikacije ciljnog mjesta (engl. *target site duplications*, TSD). TSD (crveni trokuti), poliadenilirani rep (zeleni pravokutnik), ONT sekvenca (plavi pravokutnik).
- F) Raspodjela duplikacija TSD za nespecifične i specifične ugradnje L1 pronađene u oba uzorka stanica HEK293T. Kružni dijagrami prikazuju postotak ugradnji pronađenih unutar egzona, introna i intergenske DNA.
- G) Kao za B), ali podaci su za ugradnje Alu.
- H) Detaljna karakterizacija ugradnje L1 u inficiranim stanicama HEK293T u poravnanju s kromosomom 14. Nukleotidi označeni crvenom bojom odgovaraju mjestu ugradnje. Podcrteni nukleotidi odgovaraju motivu endonukleaze L1. Crtana ilustracija prikazuje čitavu ugradnju L1 specifičnih za čovjeka (engl. *human-specific L1*, L1HS) okruženu sa duplikacijama TSD i 3' poliadeniliranim repom. Ispod ilustracije nalazi se prikaz odgovarajuće pročitane sekvence s identifikatorom. Simboli ( $\alpha$ ,  $\beta$ ,  $\delta$  i  $\gamma$ ) predstavljaju približan položaj početnica korištenih u metodi PCR. Veličine markera su navedene; kontrola bez uzorka (engl. *non-template control*, NTC). Crveni trokuti ukazuju na očekivanu veličinu ampikona L1; prazni trokut (nema produkta), popunjeni trokut (kapilarno sekvencirani produkt). Plavi trokuti označavaju očekivanu veličinu praznih mjesta.
- I) Kao za D), ali za ugradnju L1HS sa invertiranim ili deletiranim 5' krajem na kromosomu 18.

## 4. ZAKLJUČAK

Ugradnja DNA virusa SARS-CoV-2 u ljudski genom pomoću retrotranspozona LINE1, koju su objavili Zhang i sur. (2021) nije dokazana u istraživanjima Parry i sur. (2021) i Smits i sur. (2021), unatoč potvrđenoj prisutnosti retrotranspozona LINE1. Uočene su češće retrotranspozicije elemenata L1, Alu i SVA u inficiranim stanicama HEK293T vjerojatno zbog virusnog djelovanja koje stimulira aktivnost L1. Pristup Smits i sur. (2021) razlikuje se u nekoliko bitnih značajki od onog Zhang i sur. (2021). U svakom od istraživanja koristili su se različiti izolati virusa SARS-CoV-2, kod Smits i sur. multiplicitet infekcije (engl. *multiplicity of infection*, MOI) bio je duplo veći (1.0) nego kod Zhang i sur. (2021) (0.5), a metode ekstrakcije DNA, pripreme biblioteke za sekvenciranje, kao i preciznost i kvaliteta sekvenciranja stanica HEK293T, nisu bile iste. Zhang i sur. (2021) sekvencirali su stanice HEK293T transfecirane s ekspresijskim plazmidom L1, kakve se ne pojavljuju *in vivo* u pacijentima. Same stanice HEK293T predstavljaju puno povoljnije uvjete za aktivnost retrotranspozona L1, nego normalne stanice u kojima su retrotranspozoni L1 utišani. Također, visoka smrtnost inficiranih stanica smanjuje vjerojatnost opstanka ugrađenih sekvenci SARS-CoV-2 u ljudskom tijelu. Veliki postotak ugradnje virusa SARS-CoV-2 pronađen unutar egzona, Zhang i sur. (2021) su protumačili kao dokaz ugradnje virusa u ljudski genom, no ne poklapa se s aktivnostima retrotranspozona L1 dokazanim u prijašnjim istraživanjima. Zaključno, Smits i sur. (2021) tvrde da je retrotranspozicija SARS-CoV-2 teoretski moguća, ali vrlo malo vjerojatna, kao i za sve poliadenilirane RNA koje ne kodiraju reverznu transkriptazu.

## 5. LITERATURA

- Alanagreh, L. A., Alzoughool, F., & Atoum, M. (2020). The human coronavirus disease COVID-19: its origin, characteristics, and insights into potential drugs and its mechanisms. *Pathogens*, 9(5), 331.
- Cevik, M., Kuppalli, K., Kindrachuk, J., & Peiris, M. (2020). Virology, transmission, and pathogenesis of SARS-CoV-2. *bmj*, 371.
- Ewing, A. D., Smits, N., Sanchez-Luque, F. J., Faivre, J., Brennan, P. M., Richardson, S. R., ... & Faulkner, G. J. (2020). Nanopore sequencing enables comprehensive transposable element epigenomic profiling. *Molecular Cell*, 80(5), 915-928.
- Ge, X. Y., Li, J. L., Yang, X. L., Chmura, A. A., Zhu, G., Epstein, J. H., ... & Shi, Z. L. (2013). Isolation and characterization of a bat SARS-like coronavirus that uses the ACE2 receptor. *Nature*, 503(7477), 535-538.
- Katzourakis, A., & Gifford, R. J. (2010). Endogenous viral elements in animal genomes. *PLoS genetics*, 6(11), e1001191.
- koronavirusi. *Hrvatska enciklopedija, mrežno izdanje*. Leksikografski zavod Miroslav Krleža, 2021. Pristupljeno 15. 8. 2022. <<http://www.enciklopedija.hr/Natuknica.aspx?ID=70911>>.
- Lanciano, S., & Cristofari, G. (2020). Measuring and interpreting transposable element expression. *Nature Reviews Genetics*, 21(12), 721-736.
- Loomba, P., Wattal, C., Chakravarti, A., Dutta, S., Chhabra, M., Kale, P., ... & Gupta, E. (2020). Can the march of COVID-19 be halted. *Indian journal of medical microbiology*, 38(1), 128.
- Novi koronavirus i COVID-19 (2020) <https://www.plivazdravlje.hr/tekst/clanak/33401/Novi-koronavirus-i-COVID-19.html> (pristupljeno 15.8.2022.).
- Parry, R., Gifford, R. J., Lytras, S., Ray, S. C., & Coin, L. J. (2021). No evidence of SARS-CoV-2 reverse transcription and integration as the origin of chimeric transcripts in patient tissues. *Proceedings of the National Academy of Sciences*, 118(33), e2109066118.
- Prejac, J. (2022) *Hepatocelularni karcinom*. PLIVAZdravlje. <https://www.plivazdravlje.hr/aktualno/clanak/36372/Hepatocelularni-karcinom.html> (pristupljeno 25.8.2022.).

- Ravi, V., Saxena, S., & Panda, P. S. (2022). Basic virology of SARS-CoV 2. *Indian journal of medical microbiology*.
- retrovirusi. *Hrvatska enciklopedija, mrežno izdanje*. Leksikografski zavod Miroslav Krleža, 2021. Pristupljeno 8. 9. 2022. <<http://www.enciklopedija.hr/Natuknica.aspx?ID=68135>>.
- Smits, N., Rasmussen, J., Bodea, G. O., Amarilla, A. A., Gerdes, P., Sanchez-Luque, F. J., ... & Faulkner, G. J. (2021). No evidence of human genome integration of SARS-CoV-2 found by long-read DNA sequencing. *Cell Reports*, 36(7), 109530.
- Sultana, T., van Essen, D., Siol, O., Bailly-Bechet, M., Philippe, C., El Aabidine, A. Z., ... & Cristofari, G. (2019). The landscape of L1 retrotransposons in the human genome is shaped by pre-insertion sequence biases and post-insertion selection. *Molecular cell*, 74(3), 555-570.
- Venter, J. C., Adams, M. D., Myers, E. W., Li, P. W., Mural, R. J., Sutton, G. G., ... & Kalush, F. (2001). The sequence of the human genome. *science*, 291(5507), 1304-1351.
- V'kovski, P., Kratzel, A., Steiner, S., Stalder, H., & Thiel, V. (2021). Coronavirus biology and replication: implications for SARS-CoV-2. *Nature Reviews Microbiology*, 19(3), 155-170.
- WHO COVID-19 weekly epidemiological update vol. 74. World Health organization; January 2022.
- Yadav, R., Chaudhary, J. K., Jain, N., Chaudhary, P. K., Khanra, S., Dhamija, P., ... & Handu, S. (2021). Role of structural and non-structural proteins and therapeutic targets of SARS-CoV-2 for COVID-19. *Cells*, 10(4), 821.
- Zhang, L., Richards, A., Barrasa, M. I., Hughes, S. H., Young, R. A., & Jaenisch, R. (2021). Reverse-transcribed SARS-CoV-2 RNA can integrate into the genome of cultured human cells and can be expressed in patient-derived tissues. *Proceedings of the National Academy of Sciences*, 118(21), e2105968118.
- Zhang, L., Richards, A., Khalil, A., Wogram, E., Ma, H., Young, R. A., & Jaenisch, R. SARS-CoV-2 RNA reverse-transcribed and integrated into the human genome (preprint). <https://www.biorxiv.org/content/10.1101/2020.12.12.422516v1> (pristupljeno 18.8.2022.)
- Zhang, T., Wu, Q., & Zhang, Z. (2020). Probable pangolin origin of SARS-CoV-2 associated with the COVID-19 outbreak. *Current biology*, 30(7), 1346-1351.

## **6. O SEBI**

Rođena sam 2000. godine u Čakovcu. Završila sam program opće gimnazije u Gimnaziji Josipa Slavenskog Čakovec. Zbog zanimanja za biologiju, 2019. godine sam upisala preddiplomski studij Biologije Prirodoslovno-matematičkog fakulteta u Zagrebu. Slobodno vrijeme najviše volim iskoristiti vožnjom na biciklu, planinarenjem i uključivanjem u razne volonterske akcije. Članica sam Udruge studenata biologije – BIUS. Također, 2021. godine priključila sam se organizaciji WISE-a, Dana karijera na PMF-u.