

Molekularna evolucija i epidemiologija virusa SARS-CoV-2

Schwerer, Mia

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:217:210557>

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

Download date / Datum preuzimanja: **2024-07-22**



Repository / Repozitorij:

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



Sveučilište u Zagrebu
Prirodoslovno-matematički fakultet
Biološki odsjek

Mia Schwerer

**Molekularna evolucija i epidemiologija
virusa SARS-CoV-2**

Završni rad

Zagreb, 2022.

University of Zagreb
Faculty of Science
Department of Biology

Mia Schwerer

**Molecular evolution and epidemiology of
SARS-CoV-2 virus**

Bachelor thesis

Zagreb, 2022.

Ovaj završni rad je izrađen u sklopu studijskog programa Preddiplomski sveučilišni studij Molekularna biologija na Zavodu za mikrobiologiju Biološkog odsjeka Prirodoslovno-matematičkog fakulteta u Zagrebu, pod mentorstvom izv. prof. dr. sc. Silvije Černi.

TEMELJNA DOKUMENTACIJSKA KARTICA

Sveučilište u Zagrebu
Prirodoslovno-matematički fakultet
Biološki odsjek

Završni rad

Molekularna evolucija i epidemiologija virusa SARS-CoV-2

Mia Schwerer

Rooseveltov trg 6, 10000 Zagreb, Hrvatska

Koronavirusi često uzrokuju bolesti u ljudi, a jedan od njih, virus SARS-CoV-2, nakon pojave u Kini 2019. postao je problem svjetskih razmjera. S obzirom na to da je njegov genom pozitivna jednolančana RNA molekula, ovaj virus ima relativno visoku stopu mutacija. Brojne mutacije kao glavna pokretačka sila uzrokuju brzu evoluciju što sa sobom često nosi promjene u biološkim karakteristikama virusa. Posebno su česte i problematične one u dijelu genoma koji kodira za protein šiljka jer promjena samo jedne aminokiseline može značajno utjecati na interakcije između proteina šiljka i receptora na stanici domaćina čime se mijenja infektivnost samog virusa. S obzirom na prisutnost određenih mutacija do sada je identificirano pet varijanti koje izazivaju zabrinutost zbog povećane sposobnosti rasprostranjivanja. Praćenjem virusne filodinamike, koja se bavi utjecajem epidemioloških, imunoloških i evolucijskih procesa na gensku raznolikost virusa moguće je dobiti uvid u geografsku rasprostranjenost i patogenezu spomenutih varijanti te procijeniti eventualnu opasnost od razvoja novih. S ciljem olakšavanja istraživanja i boljeg razumijevanja evolucije i rasprostranjivanja virusa gripe i koronavirusa pokrenuta je platforma GISAID u sklopu koje su javno dostupne brojne nukleotidne sekvence te epidemiološki podatci. Upravo na temelju tih podataka provode se mnoge analize koje doprinose razumijevanju evolucije i patogeneza virusa SARS-CoV-2.

Ključne riječi: filodinamika, protein šiljka, varijante koje izazivaju zabrinutost, GISAID
(25 stranica, 10 slika, 64 literaturnih navoda, jezik izvornika: hrvatski)
Rad je pohranjen u Središnjoj biološkoj knjižnici

Mentor: izv. prof. dr. sc. Silvija Černi

BASIC DOCUMENTATION CARD

University of Zagreb
Faculty of Science
Department of Biology

Bachelor thesis

Molecular evolution and epidemiology of SARS-CoV-2 virus

Mia Schwerer

Rooseveltova trg 6, 10000 Zagreb, Croatia

After its appearance SARS-CoV-2 virus shortly became a global problem. Given that its genome is a positive single-stranded RNA, this virus has a relatively high mutation rate. Numerous mutations as the main driving force cause rapid evolution, which often brings changes in the biological characteristics of virus. Mutations in the part of genome that codes for spike protein are especially common and problematic. The reason for this is that changing a single amino acid can significantly affect interactions between spike protein and receptor on the host cell, thus changing the infectivity. Due to the presence of certain mutations there are until now identified five variants of concern. By monitoring viral phylodynamics, which deals with the influence of epidemiological, immunological and evolutionary processes on genetic variability, it is possible to gain insight into the geographical distribution and pathogenesis of variants and assess the eventual risk of developing new variants. With the aim of facilitating better understanding of the evolution of viruses and spread of influenza and coronaviruses, GISAID platform was launched, within which nucleotide sequences and epidemiological data are publicly available. Based on these data, many analyses are carried out that contribute to the understanding of the evolution and pathogenesis of SARS-CoV-2.

Keywords: phylodynamics, spike protein, variants of concern, GISAID
(25 pages, 10 figures, 0 tables, 64 references, original in: croatian)
Thesis is deposited in Central Biological Library.

Mentor: assoc. prof. Silvija Černi

Popis kratica

ACE2	angiotenzin konvertirajući enzim 2 (engl. <i>angiotensin-converting enzyme 2</i>)
CD	domena za povezivanje (engl. <i>connecting domain</i>)
CH	središnja zavojnica (engl. <i>central helix</i>)
CoV	koronavirus (engl. <i>coronavirus</i>)
COVID-19	koronavirusna bolest (engl. <i>coronavirus disease</i>)
CT	citoplazmatski rep (engl. <i>cytoplasmic tail</i>)
DNA	deoksiribonukleinska kiselina (engl. <i>deoxyribonucleic acid</i>)
E	protein ovojnice (engl. <i>envelope protein</i>)
FP	fuzijski peptid (engl. <i>fusion peptide</i>)
GISAID	globalna inicijativa za dijeljenje svih podataka o gripi (engl. <i>The Global Initiative on Sharing All Influenza Data</i>)
HR1	sedmerostruko ponavljanje 1
HR2	sedmerostruko ponavljanje 2
ICTV	međunarodni odbor za taksonomiju virusa (engl. <i>The International Committee on Taxonomy of Viruses</i>)
M	protein membrane (engl. <i>membrane protein</i>)
MERS-CoV	<i>Middle east respiratory syndrome coronavirus</i>
mRNA	glasnička ribonukleinska kiselina (engl. <i>messenger ribonucleic acid</i>)
N	nukleokapsidni protein (engl. <i>nucleocapside protein</i>)
NCBI	Nacionalni centar za biotehnoške informacije (engl. <i>National Center for Biotechnology Information</i>)
NTD	N-terminalna domena (engl. <i>N-terminus domain</i>)
ORF	otvoreni okvir čitanja (engl. <i>open reading frame</i>)
RBD	receptor vezujuća domena (engl. <i>receptor-binding domain</i>)
RNA	ribonukleinska kiselina (engl. <i>ribonucleic acid</i>)
S	protein šiljka (engl. <i>spike protein</i>)
SSE	dogadjaji superširenja (engl. <i>superspread event</i>)

ssRNA	jednolančana ribonukleinska kiselina (engl. <i>single stranded ribonucleic acid</i>)
TM	transmembranska domena (engl. <i>transmembrane domain</i>)
TMPRSS2	transmembranska serinska proteaza 2 (engl. <i>transmembrane serine protease 2</i>)
UTR	netranslatirajuća regija (engl. <i>untranslated region</i>)
VOC	varijante koje izazivaju zabrinutost (engl. <i>variants of concern</i>)
VOI	varijante od interesa (engl. <i>variants of interest</i>)
WHO	Svjetska zdravstvena organizacija (engl. <i>World Health Organisation</i>)

Sadržaj

1. Uvod	1
2. Virus SARS-CoV-2.....	2
2. 1. Klasifikacija.....	2
2. 2. Građa	3
2. 3. Protein šiljka.....	4
3. Molekularna epidemiologija	7
4. Molekularna evolucija.....	7
4. 1. Varijante koje izazivaju zabrinutost	8
5. Filogenetika.....	10
5. 1. Viralna filodinamika.....	12
5. 2. Filodinamika virusa SARS-CoV-2.....	14
6. Globalna inicijativa za dijeljenje svih podataka o gripu	16
7. Zaključak.....	19
8. Literatura	20
9. Životopis.....	25

1. Uvod

Svjetske pandemije i epidemije tijekom povijesti nisu bile velika rijetkost te je kroz godine istraživanja otkriveno puno o mikrobima koji ih uzrokuju. Osim velikih svjetskih pandemija unazad 20-ak godina izbile su i brojne manje epidemije, ali niti jedna nije prouzročila toliko problema kao ona koju uzrokuje virus SARS-CoV-2 (Severe acute respiratory syndrome coronavirus 2). Novi koronavirus koji uzrokuje bolest COVID-19 pojavio se u Wuhanu (Kina) u prosincu 2019. godine (Singh i Yi, 2021). Iako u početku skoro u potpunosti nepoznat i neistražen ubrzo se počeo prenositi i izvan granica Azijskog kontinenta te je prerastao u globalnu pandemiju i postao jedan od glavnih tema brojnih istraživanja (Singh i Yi, 2021). Prema podacima Svjetske zdravstvene organizacije (engl. *World Health Organisation* – WHO) od samog početka pandemije do lipnja 2022. u svijetu je zabilježeno preko 528 000 000 slučajeva zaraze te je posljedično preminulo preko 6 290 000 osoba (<https://covid19.who.int/>).

Pandemija COVID-19 nije prva poznata pandemija uzrokovana koronavirusom. Relativno nedavno, svijet je bio suočen s dvije druge pandemije čiji je uzročnik također bio virus iz istog roda koji je upravo iz tog razloga dijelio brojne karakteristike s virusom SARS-CoV-2, kao što je spektar domaćina koje inficira te organizacija i veličina genoma (Abdelghany *i sur.*, 2021). Pandemija koju je uzrokovao virus SARS-CoV (Severe acute respiratory syndrome coronavirus) između 2002. i 2003. godine bila je puno manjih razmjera te geografski uže lokalizirana na područje Saudijske Arabije. Na istom je prostoru zabilježena i pandemija uzrokovana virusom MERS-CoV (Middle east respiratory syndrome coronavirus) koji je prvi puta detektiran 2011. godine, ali je ta pandemija bila nešto duljeg trajanja, a novi slučajevi su se pojavljivali sve do 2017. godine. Kao što se pretpostavlja i za prijašnje pandemije provedena istraživanja ukazuju na to da je podrijetlo virusa životinjsko, odnosno da je došlo do interspecijske transmisije što je dovelo do brze adaptacije virusa na novog domaćina, u ovom slučaju čovjeka (Bashor *i sur.*, 2021).

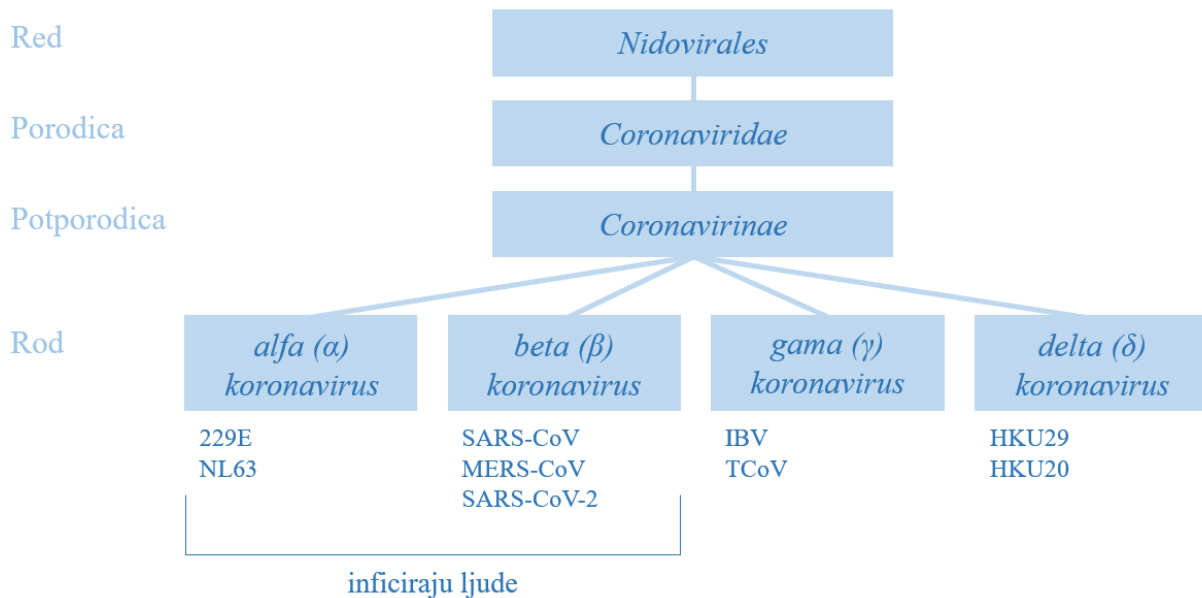
Ubrzo nakon pojave pandemije razvijeno je i odobreno nekoliko cjepiva protiv virusa SARS-CoV-2, a među njima i dva mRNA cjepiva, koja imaju prednost nad ostalima zbog mogućnosti brzog razvoja i prilagođavanja virusu koji brzo evoluiraju (Hogan i Pardi, 2022). No i dalje je aktualan problem kontinuirane pojave novih varijanti koje sve uspješnije zaobilaze imunost stečenu cijepljenjem ili preboljenjem što smanjuje efikasnost postojećih cjepiva (Mascellino *i sur.*,

2021). Stoga se danas teži boljem razumijevanju virusne evolucije te praćenju novih varijanti kako bi se buduće aktivnosti kontrole rasprostanjivanja virusa i razvoja bolesti mogle usmjeriti u pravom smjeru.

2. Virus SARS-CoV-2

2. 1. Klasifikacija

Koronavirusi su RNA virusi iz reda *Nidovirales* koji pripadaju u porodicu *Coronaviridae*, unutar koje su na temelju genetičke i serološke karakterizacije pripadnici potporodice *Coronavirinae* podijeljeni u četiri roda; alfa-, beta-, gama- i delta-koronavirusi (Rotondo *i sur.*, 2021). Alfa- i beta-koronavirusi inficiraju isključivo sisavce uključujući i ljude, a gama- i delta-koronavirusi ptice, iako postoje izuzetci koji inficiraju i sisavce kao što su kitovi (Singh i Yi, 2021). Klasifikacija koronavirusa prikazana je na Slici 1.

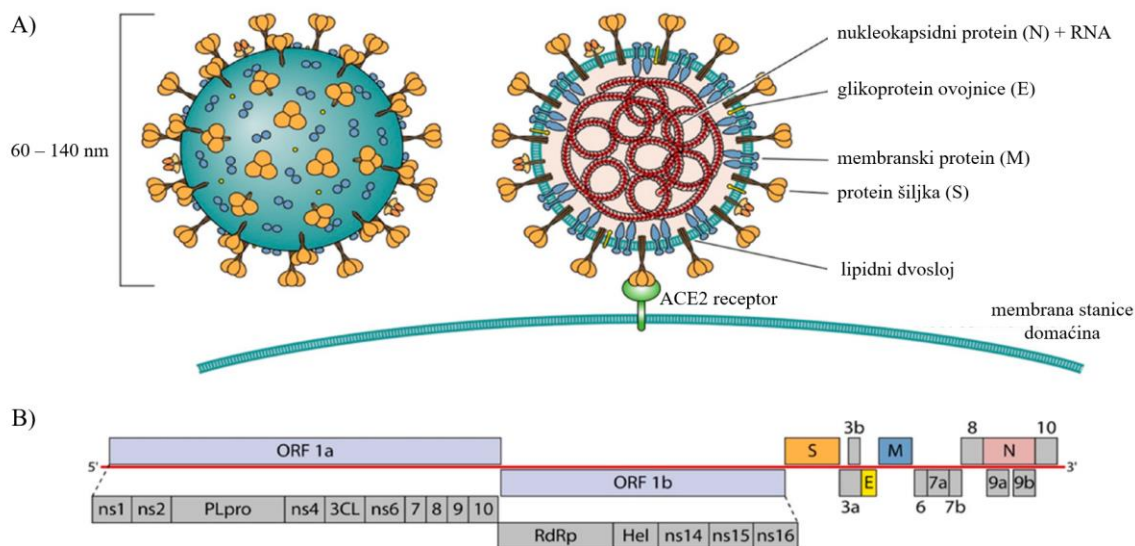


Slika 1. Klasifikacija koronavirusa. Podjela u rodove na temelju seroloških i filogenetskih istraživanja. Preuzeto i prilagođeno iz Zhou *i sur.*, 2021.

Prema virusnoj taksonomskoj klasifikaciji određenoj od strane Međunarodnog odbora za taksonomiju virusa (engl. *The International Committee on Taxonomy of Viruses – ICTV*) (<https://talk.ictvonline.org/>) identifikacija i klasifikacija koronavirusa temelji se na sličnosti aminokiselinskih sljedova sedam domena koje su kodirane otvorenim okvirom čitanja (engl. *open reading fram – ORF*) ORF1ab. Virusi SARS-CoV i SARS-CoV-2 dijele preko 90 % sličnosti te su svrstani u SARS klaster vrlo srodnih virusa (Yang *i sur.*, 2020).

2. 2. Građa

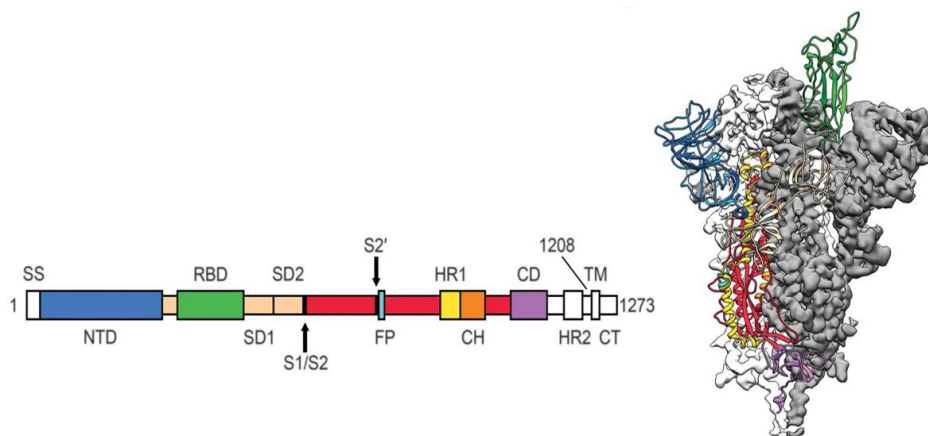
Čestica virusa SARS-CoV-2 je sferičnog oblika promjera između 60 i 140 nm (Gitman, Shaban i Paniz-Mondolfi, 2021). Sadrži jednolančanu pozitivnu RNA molekulu (+ssRNA) te mu je genom veličine oko 30 kb. Na 5' netranslatiranom kraju (engl. *untranslated region – UTR*) nalazi se dugi otvoreni okvir čitanja ORF1ab koji kodira 16 nestrukturnih proteina (Zhang *i sur.*, 2021). Zatim slijedi regija koja kodira četiri strukturna proteina; protein šiljka (S), protein ovojnice (E), protein membrane (M) i nukleokapsidni protein (N) (Phan, 2020). Proteini ovojnice zajedno s glikoproteinima membrane čine ovojnicu virusne čestice i važni su za ulazak u stanicu domaćina, replikaciju te kasnije sastavljanje i pupanje novih čestica (Singh i Yi, 2021). Nukleokapsida koju sačinjavaju nukleokapsidni proteini je helikalno simetrična, što je inače karakteristika -RNA virusa te je u ovom slučaju iznimka (Maier, Bickerton i Britton, 2015). Na Slici 2. shematski je prikazana struktura virusne čestice virusa SARS-CoV-2 i struktura njegovog genoma.



Slika 2. Struktura čestice virusa SARS-CoV-2. A) Shematski prikaz vanjske građe te poprečni presjek kroz česticu. B) Shematski prikaz strukture genoma. Preuzeto i prilagođeno iz Gitman, Shaban i Paniz-mondolfi, 2021.

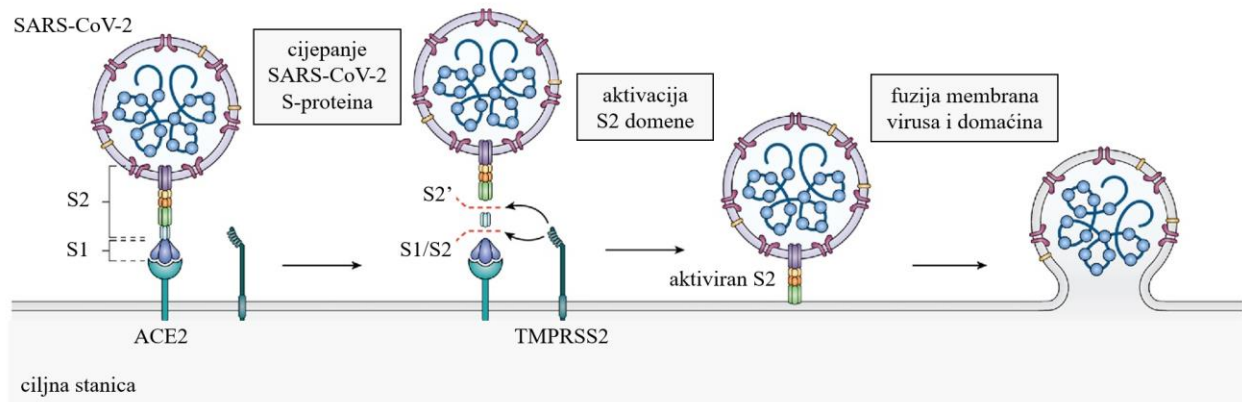
2. 3. Protein šiljka

Proteini šiljka nalazi se na vanjskoj strani membrane koja obavija virusnu česticu, a organizirani su u homotrimere veličine 3,5 angstrema (Wrapp *i sur.*, 2020). Virusu daju njegov karakterističan izgled te mu omogućuje interakciju sa staničnim receptorima i ulazak u ciljnu stanicu domaćina (Harvey *i sur.*, 2021). Upravo zbog najveće stope mutacija i promjena u genskoj sekvenci, protein koji može otkriti najviše o virusnoj evoluciji je protein šiljka. S protein kojeg sačinjavaju 1 273 aminokiseline veličine je između 180 i 200 kDa i sastoji se od ekstracelularnog N-kraja, transmembranske regije i kratkog intracelularnog C-kraja (Maier, Bickerton i Britton, 2015). Čine ga dvije podjedinice; S1 koja sadrži N-terminalnu domenu (engl. *N-terminus domain* – NTD) i receptor vezujuću domenu (engl. *receptor binding domain* – RBD) te S2 koja sadrži fuzijski peptid (engl. *fusion peptide* – FP), regiju sedmerostrukog ponavljanja 1 i 2 (engl. *heptad repeat 1/2* – HR1/2), središnju zavojnicu (engl. *central helix* – CH), domenu za povezivanje (engl. *connector domain* – CD), transmembransku domenu (engl. *transmembrane domain* – TM) i citoplazmatski rep (engl. *cytoplasmic tail* – CT) (Wrapp *i sur.*, 2020). Shematski prikaz organizacije regije genoma koji kodira za S protein s naznačenim domena te njegova 3D struktura prikazani su na Slici 3.



Slika 3. Shematski prikaz regije gena koja kodira za S protein virusa SARS-CoV-2 i njegove 3D strukture. Domene su prikazane različitim bojama; SS, signalna sekvenca; NTD, N-terminalna domena; RBD, receptor vezujuća domena; S1/S2 i S2', mjesta cijepanja; FP, fuzijski peptid; HR1, sedmerostruko ponavljanje 1; CH, središnja zavojnica; CD, domena za povezivanje; HR2, sedmerostruko ponavljanje 2; TM, transmembranska domena; CT; citoplazmatski rep. Preuzeto i prilagođeno prema (Wrapp *i sur.*, 2020).

Prilikom interakcije virusnih proteina šiljka i proteina na membrani stanice domaćina ključnu ulogu ima membranski receptor ACE2 (engl. *angiotensin-converting enzyme 2*). ACE2 je protein uključen u renin angiotenzinski sustav koji ima izuzetno bitnu ulogu u kardiovaskularnoj regulaciji i virusnim infekcijama. ACE2 je metalopeptidaza koja se sastoji od 805 aminokiselina te je kao transmembranski glikoprotein prisutan na većini stanica u obliku jedne ekstracelularne katalitičke domene na staničnoj membrani (Medina-Enríquez *i sur.*, 2020). Do vezanja S proteina za ACE2 receptor dolazi zbog elektrostatskih sila, a zbog veće gustoće električnog polja S proteina virusa SARS-CoV-2 u usporedbi s virusom SARS-CoV dolazi do jačih interakcijama što rezultira većom infektivnošću i bržim širenjem bolesti (Shirbhate *i sur.*, 2021). Receptor ACE2 nije jedini stanični protein potreban za omogućavanje ulaska virusa u stanicu i patogenezu. Uz njega u spomenutom procesu ključna je transmembranska serinska proteaza 2 (engl. *transmembrane serine protease 2* – TMPRSS2) (Hoffmann *i sur.*, 2020; Fraser *i sur.*, 2022). TMPRSS2 cijepa S protein na S1/S2 i S2' restrikcijskim mjestima čime nastaju dvije podjedinice što omogućava fuziju membrana virusa i stanice domaćina te ulazak virusne čestice u samu stanicu i nastavak infekcije (Peiffer *i sur.*, 2021). Shematski prikaz procesa ulaska virusa SARS-CoV-2 u stanicu domaćina prikazan je na Slici 4.



Slika 4. Mehanizam ulaska virusa SARS-CoV-2 u stanicu domaćina. Interakcija s ACE2 receptorom i cijepanje na S1/S2 te S2' restriksijskim mjestima od strane proteaze TMPRSS2 što uzrokuje aktivaciju S proteina i fuziju membrana te ulazak virusa u stanicu. Preuzeto i prilagođeno iz (Hartenian *i sur.*, 2020).

3. Molekularna epidemiologija

Epidemiologija je znanost koja istražuje zdravstvene probleme fokusirajući se na tri faktora, a to su domaćin, uzročnik bolesti i okoliš. Proučavanje spomenutih faktora i njihove povezanosti vodi ka otkrivanju uzroka i tijeka širenja bolesti unutar populacije (Mpolya *i sur.*, 2009). Prije 50-ak godina, točnije 1973. godine, Kilbourne je prvi puta upotrijebio pojam molekularne epidemiologije, grane epidemiologije koja kao spoj molekularne biologije i epidemiologije korištenjem molekularnih metoda pruža bolji uvid u patogenezu, prijenos, prevenciju i terapiju zaraznih bolesti (Kilbourne, 1973). Od tada pa sve do danas ne prestaje rasti broj znanstvenih radova u kojima se koriste upravo metode molekularne epidemiologije (Riapis, Filatov i Salova, 2003). Razvoj spomenute grane epidemiologije omogućava istraživanje uzročnika bolesti na molekularnoj razini te korištenjem novih tehnologija i metoda rješava probleme koje to tada nije bilo moguće premostiti metodama korištenim u klasičnoj epidemiologiji. Pojmom molekularna epidemiologija često se obuhvaćaju metode tipizacije sojeva i varijanti patogena koji uzrokuju zarazne bolesti. Upravo filogenetika koju čini širok spektar različitih metoda je vrlo često korišten alat u molekularno epidemiološkim istraživanjima (Hall i Barlow, 2006).

4. Molekularna evolucija

Praćenje virusne evolucije izuzetno je bitno kako bi se bolje razumjelo podrijetlo samog virusa, epidemiologija te potencijalno pretpostavilo moguće daljnje rasprostranjivanje virusa i nastanak novih varijanti (Lauring, 2020). Za njezino razumijevanje i praćenje nužno je uzeti u obzir i glavnu pokretačku silu evolucije – mutacije. Mutacije unose promijene koje omogućavaju da prirodna selekcija djeluje te da se zadržavaju one koje su za jedinku najpovoljnije (Engelhardt i Shakhnovich, 2019). Velika većina mutacija nema utjecaj na jedinku, dio utječe negativno, dok samo mali postotak doprinosi povoljnim značajkama (Duffy, 2018). Zbog svoje građe genoma i kemijske prirode RNA molekule, RNA virusi imaju visoku stopu mutacija. Za RNA viruse procijenjeno je da se ona kreće u intervalu između 10^{-4} i 10^{-6} mutacija po nukleotidu po infekcijistanice (engl. *substitutions per nucleotide per cell infection – s/n/c*), a za DNA viruse je nešto niža te se nalazi u intervalu od 10^{-8} do 10^{-6} s/n/c (Sanjuán *i sur.*, 2010).

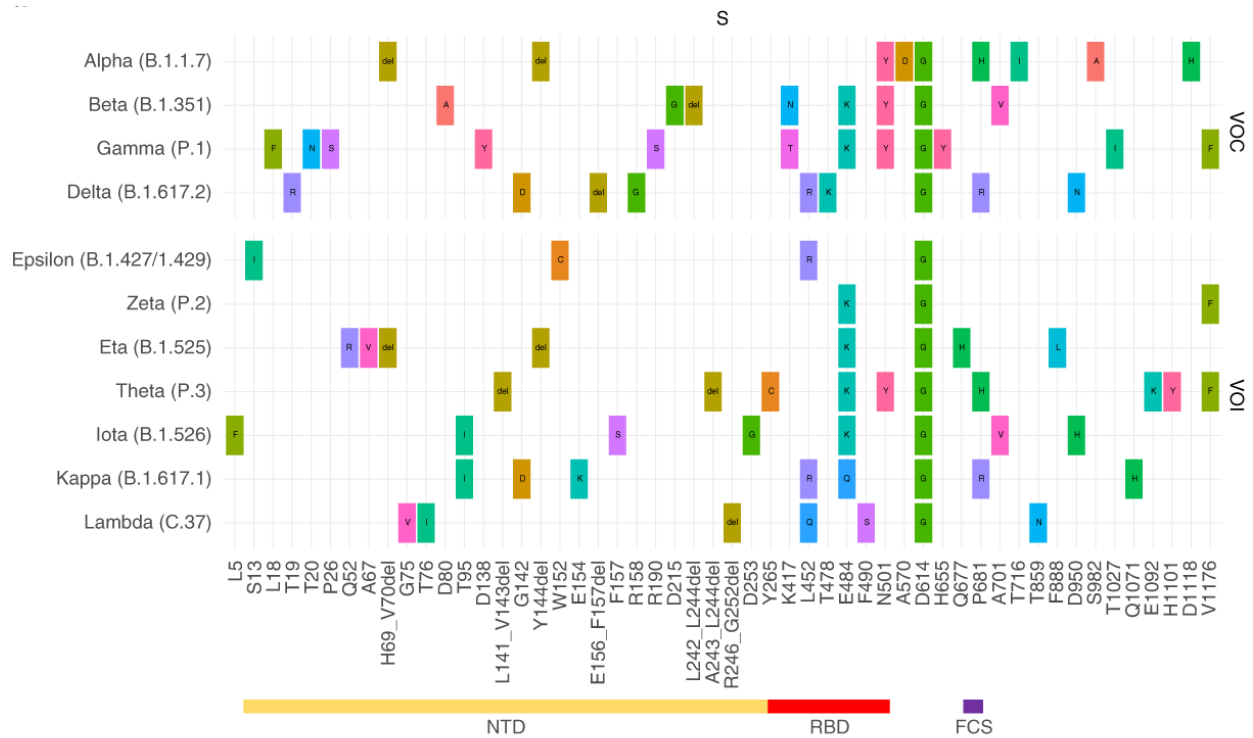
4. 1. Varijante koje izazivaju zabrinutost

Postoje brojne varijante virusa koje nose određene mutacije u svojem genomu što ih čini genetički različitim. U slučaju da nose mutacije za koje je otkriveno i dokazano da imaju značajan utjecaj na fenotip, kao što je uspješnije rasprostranjivanje i veća virulencija ili poboljšana sposobnost bijega od imunskog sustava, te varijante nazivaju se varijantama koje izazivaju zabrinutost (engl. *variants of concern* – VOCs) (Sanyaolu *i sur.*, 2021). Ukoliko se za mutacije koje posjeduje određena varijanta samo pretpostavlja takav utjecaj, bez postojećih dokaza, onda se one nazivaju varijantama od interesa (engl. *variants of interest* – VOIs) (<https://www.ecdc.europa.eu/en/covid-19/variants-concern>).

Od početka pandemije do danas od strane WHO proglašeno je pet varijanti koje izazivaju zabrinutost; alfa, beta, gama, delta i omikron. Prva detektirana varijanta koja izaziva zabrinutost bila je alfa varijanta poznata i kao B.1.1.7 koja se pojavila u rujnu 2020. godine u Ujedinjenom Kraljevstvu te je ubrzo postala dominantna zbog brzog širenja bolesti koju izaziva. Ono po čemu se razlikuje od prvotno detektirane varijante iz Wuhana je 17 ne-sinonimnih mutacija i osam dodatnih mutacija u S proteinu; Δ H69-V70, Δ Y144, N501Y, A570D, P681H, T716I, S982A i D1118H. Od navedenih mutacija tri su od posebnog interesa, a to su delecija dvije aminokiseline u NTD (Δ H69-V70), supstitucija 501. aminokiseline iz asparagina (N) u tirozin (Y) (N501Y) u RBD i supstitucija 681. aminokiseline iz prolina (P) u histidin (H) (P681H) koja se nalazi proksimalno od furinskog mjesta cijepanja podjedinica (Shen *i sur.*, 2021). Najznačajniji utjecaj na fenotip ima mutacija N501Y koja samostalno ima skoro jednake fenotipske karakteristike kao i kombinacija svih osam mutacije S proteina. To upućuje na izuzetno veliki utjecaj spomenute supstitucije na prijenos virusa među domaćinima, a podatci da je ona uočena i kod ostalih varijanti koje izazivaju zabrinutost zbog povećane prenosivosti i prijenosa dodatno podupire tu pretpostavku (Liu *i sur.*, 2021). Beta varijanta (B.1.351) detektirana je u svibnju 2020. godine u Južnoafričkoj Republici (D. Zhou *i sur.*, 2021), a karakterizira ju neobično velik broj mutacija (<https://gvn.org/covid-19/beta-b-1-351/>). Od 21 mutacije devet ih je u dijelu genoma koji kodira S protein. Uz N501Y mutaciju koja se pojavljuje i u alfa varijanti treba istaknuti najvažnije, a to su; E484K, K417N, ORF1B delecija u RBD te L18F, D80A, D215G, Δ 242-244, R264I, A701V u NTD (Lemieux *i sur.*, 2021). Iste godine u studenom u Brazilu detektirana je još jedna varijanta koja izaziva zabrinutost, gama varijanta (P. 1) koja se ubrzo proširila i izvan granica Brazila i Južnoameričkog kontinenta (Banho *i sur.*, 2022). Od 21 mutacije koje definiraju gama varijantu

njih deset nalazi se na S proteinu, od čega tri u domeni RBD, a to su supstitucije K417T, E484K i N501Y. Uočljivo je kako su te mutacije identične odnosno vrlo slične onima u beta varijanti što ukazuje na konvergenciju i nezavisnu pojavu istih mutacija kod više varijanti (Nonaka *i sur.*, 2021). Krajem 2020. godine u Indiji prepoznata je još jedna nova virusna varijanta koja izaziva zabrinutost, a to je delta (B.1.617.2). Ono što karakterizira tu varijantu i razlikuje ju od dosad spomenutih je supstitucija u S proteinu gdje je na 681. mjestu prolin zamijenjen argininom (P681R). Ta mutacija doprinosi boljem cijepanju S proteina te time pospješuje fuziju i ulazak virusa u stanicu (Saito *i sur.*, 2021). Također mutacije L452R i T478K u RBD te delecija Δ 156-157 u NTD imaju značajan utjecaj na imunski odgovor jer se nalaze u ključnim antigenim regijama proteina (<https://gvn.org/covid-19/delta-b-1-617-2/>).

Omikron (B.1.1.529) je najrecentnije proglašena varijanta koja izaziva zabrinutost, a prvi puta je detektirana u Južnoafričkoj Republici. Preko 60 % svih mutacija koje ovu varijantu odvajaju od originalnog, prvotno izoliranog, soja iz Wuhana iz prosinca 2019. godine nalazi se u dijelu genoma koji kodira za S protein (Jung *i sur.*, 2022). Unutar S1 domene nalazi se 31 mutacija točnije supstitucije i delecije, od čega njih 11 u domeni NTD (A67V, Δ H69, Δ V70, T95I, G142D, Δ V143, Δ Y144, Δ Y145, Δ N211, L212I i ins214EPE), 15 u domeni RBD (G339D, S371L, S373P, S375F, K417N, N440K, G446S, S477N, T478K, E484A, Q493R, G496S, Q498R, N501Y i Y505H) i pet u domeni CT (T547K, D614G, H655Y, N679K i P681H) (Sarkar *i sur.*, 2021). Ukratko pregled supstitucija i delecija aminokiselina u dijelu genoma koji kodira za S protein varijanta virusa SARS-CoV-2 koje izazivaju zabrinutost (alfa, beta, gama, delta i omikron) i varijantama od interesa (epsilon, zeta, eta, theta, jota, kapa i lambda) prikazan je na Slici 6. Može se uočiti kako su neke mutacije, primjerice D614G prisutne kod svih varijanti dok je velik dio mutacija usko vezan uz samo jednu ili dvije varijante, primjerice D215G kod beta varijante i D138Y kod gama varijante.



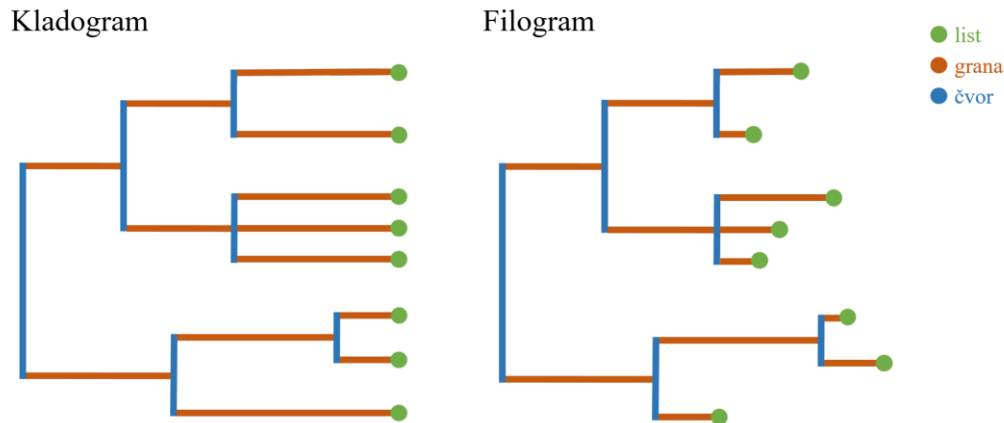
Slika 6. Pregled mutacija aminokiselina (supstitucija i delecija) u dijelu genoma koji kodira za S protein u varijantama virusa SARS-CoV-2 koje izazivaju zabrinutost (VOC) i varijantama od interesa (VOI). Preuzeto iz Oude Munnink *i sur.*, 2021.

5. Filogenetika

Filogenetika, područje biologije koje se bavi međusobnim odnosima između organizama, vrsta i/ili varijanti, postala je neophodna u praćenju epidemiologije (Crandall i Posada, 2002). Pomoću filogenetskih analiza moguće je otkriti podrijetlo uzročnika bolesti te obrasce rasprostranjivanja i pojave novih varijanti (Yang i Rannala, 2012) što se preklapa s ciljevima epidemioloških istraživanja i razlog je zašto je filogenetika toliko korištena i korisna u epidemiologiji.

Međusobni odnosi mogu se odrediti na temelju takozvanih filogenetskih biljeaga, a prikazuju se filogenetskim stablima. Filogenetski biljezi su karakteristike koje su unutar promatranih uzoraka varijabilne te se na temelju njihovih razlika uspostavljaju odnosi. Mogu biti klasični biljezi, poput morfoloških karakteristika, ili molekularni biljezi, poput sekvenci nukleotida ili aminokiselina (Ke

i sur., 2017). Ukoliko se koriste molekularni biljezi govorimo o molekularnoj filogeniji koja je danas sve češća. Upotrebnom molekularnih biljega dobivaju se precizniji rezultati jer su ti podatci nedvosmisleni i ima ih izuzetno puno što omogućava dobru i kvalitetnu statističku analizu (Amit Roy, 2014). Postupak izrade stabla molekularnom filogenijom sastoji se od nekoliko koraka. Prvo je potrebno sakupiti sve sekvence, sekvenciranjem vlastitih uzoraka ili pretraživanjem i sakupljanjem istih iz baza podataka kao što je Nacionalni centar za biotehnoške informacije (engl. *National Center for Biotechnology Information* – NCBI). Zatim se sve sekvence višestruko sravnjuju (engl. *multiple alignment*) te se na temelju višestrukog sravnjenja pomoću neke od metoda rekonstruira stablo (Hall, 2013). Najčešće korištene filogenetske metode za rekonstrukciju stabala su; metode maksimalne štedljivosti (engl. *maximum parsimony methods*), metode maksimalne vjerojatnosti (engl. *maximum likelihood methods*), metode udaljenosti (engl. *distance methods*) i Bayesijske metode (engl. *Bayesian methods*). Metode se međusobno razlikuju po matematičkim i statističkim modelima na kojima se temelje, a ovisno o krajnjem cilju bira se ona metoda koja je najpogodnija (Yang i Rannala, 2012). Analizom topologije dobivenog stabla donose se zaključci o odnosima promatranih taksona, a za razumijevanje i mogućnost interpretacije stabla bitno je znati što predstavljaju listovi, grane i čvorovi. Duljina grana može, ali i ne mora biti informativna, a ukoliko je, ona može predstavljati razlike odnosno mutacije (filogram) ili vrijeme (kronogram). S druge strane čvorovi, točnije njihov položaj je informativan te upućuje na odnose između taksi. Čvorovi mogu biti unutarnji ili vanjska, a vanjski se često nazivaju i listovima. Oni predstavljaju taksone koje proučavamo, u molekularnoj filogenetici to su najčešće jedinice iz kojih potječu sekvence na temelju kojih je izrađeno stablo (Gregory, 2008). Usporedba kladograma i filograma s naznačenim elementima filogenetskog stabla prikazana je na Slici 5.



Slika 5. Prikaz dva tipa filogenetskih stabala; kladograma (duljina grana prikazuje vrijeme) i filograma (duljina grana prikazuje razlike, odnosno mutacije). Različitim bojama prema legendi naznačeni su ključni elementi stabala (grane, čvorovi i listovi).

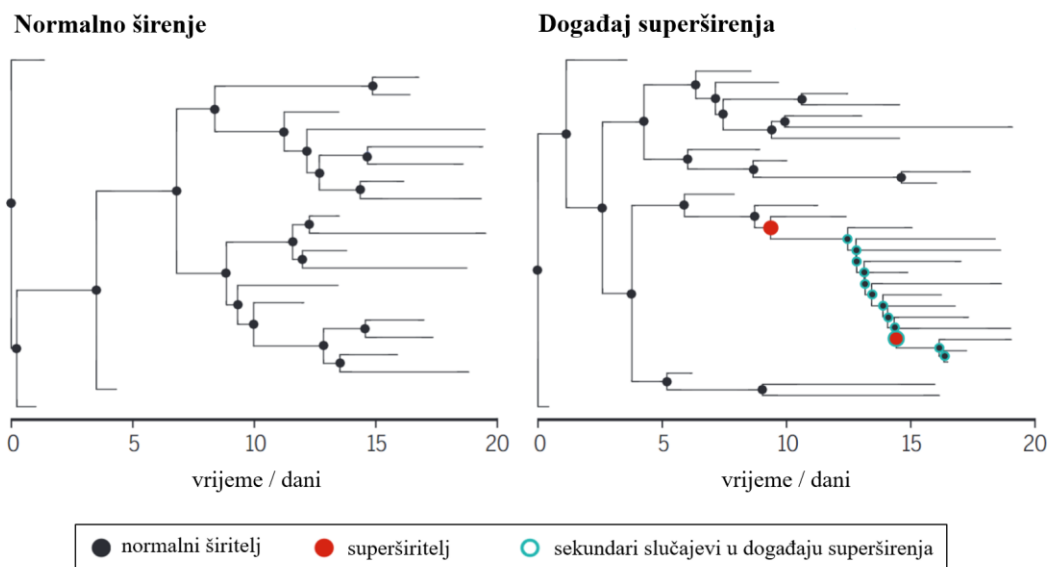
5. 1. Viralna filodinamika

Viralna filodinamika je područje koje se bavi utjecajem epidemioloških, imunoloških i evolucijskih procesa na viralnu filogeniju (Volz, Koelle i Bedford, 2013). Pojam filodinamike uveden je 2004. godine kada je Grenfell sa suradnicima uočio važnost poveznice između epidemioloških procesa i evolucije patogena. Za prva istraživanja, radi jednostavnosti, bilo je bitno pronaći model u kojem se epidemiološki i populacijski procesi odvijaju u sličnim vremenskim okvirima. Upravo RNA virusi zbog svoje visoke stope mutacija, velikih populacija i kratkog generacijskog vremena bili su idealni za takav tip istraživanja (Grenfell *i sur.*, 2004).

Filodinamičke analize temelje se na statističkim i matematičkim modelima kao što su supstitucijski model i model molekularnog sata (Attwood *i sur.*, 2022). Supstitucijski modeli su matematički modeli koji opisuju promijene odnosno mutacije u nukleotidnim sekvencama i čine osnovu u evolucijskim analizama genetičkih podataka na molekularnoj razini (Arenas, 2015). Pojavom naprednijih softvera za filogenetičke analize, kao što je metoda najveće vjerojatnosti (engl. *maximum likelihood method* – ML) (Postillone i Perez, 2017), uočena je važnost odabira što boljeg supstitucijskog modela jer on direktno utječe i na kvalitetu same filogenetske analize (Posada i Crandall, 2001). Danas su poznati brojni supstitucijski modeli između kojih je moguće odabrati onaj koji najbolje odgovara ciljnoj regiji genoma. Postojanje spomenute raznolikosti

modela nužno je s obzirom na to da je molekularna evolucija često heterogena te se ne može uvijek opisati istim modelom već su potrebne određene izmjene kako bi se ona što bolje opisala za određeni dio genoma (Arbiza *i sur.*, 2011). Drugi bitan model koji se koristi u filogenetičkim analizama je molekularni sat. To je hipoteza prema kojoj DNA molekula i proteini evoluiraju relativno konstantnom brzinom u vremenu što omogućuje utvrđivanje kada su se dvije vrste odvojile odnosno gdje se na vremenskoj skali nalazi njihov zadnji zajednički predak (Pascual-García, Arenas i Bastolla, 2019). Iako neočekivano, stopa brzine molekularne evolucije relativno je konstanta, a taj fenomen se objašnjava time da je većina mutacija neutralna, odnosno da nemaju pozitivan ili negativan utjecaj na fitnes i preživljavanje jedinki (Bromham i Penny, 2003).

Viralna filogenija, kao dio područja kojima se bavi filodinamika, često je vrlo dinamična i pruža efikasan uvid u brojne aspekte kako samog virusa uzročnika tako i bolesti te njezinog širenja. Ono što uvelike utječe na filogeniju su događaji supersirenja (engl. *superspread event* – SSE). Ovakvi događaji podrazumijevaju kada jedan slučaj zaraze uzrokuje više sekundarnih infekcija no što je očekivano s obzirom na osnovni reprodukcijski broj virusa (R_0). R_0 predstavlja prosječan broj sekundarnih infekcija u populaciji te su stoga SSE često tzv. katalizatori epidemija (Lemieux *i sur.*, 2021). Na Slici 7. prikazan je efekt superširitelja. Superširitelj je označen crvenom točkom, a svaki čvor na stablu prikazuje novi prijenos, odnosno transmisiju virusa. Po broju čvorova vidljivo kako je superširitelj izazvao brojne sekundarne infekcije.



Slika 7. Primjer efekta superširenja prikazan na filogenetskom stablu. Superširiteljima (označeni crvenim točkama) nazivamo one jedinice i slučajeve zaraze koji uzrokuju više sekundarnih infekcija (označeni plavo okruženim točkama) no što je očekivano s obzirom na osnovni reprodukcijski broj virusa (R_0). Preuzeto i prilagođeno iz Hillmayer *i sur.*, 2021.

5. 2. Filodinamika virusa SARS-CoV-2

U samom početku pandemije COVID-19 svaka informacija i novo saznanje o rasprostranjenosti virusa bilo je izuzetno korisno prilikom donošenja strategija suzbijanja širenja bolesti, a upravo filodinamika može doprinijeti tim saznanjima. Filodinamički modeli i analize mogu se koristiti u razne svrhe, kao što su; otkrivanje podrijetla virusa, praćenje epidemioloških parametara i prijenosu virusa i ispitivanje efektivnosti mjera protiv širenja zaraze (Attwood *i sur.*, 2022).

Upravo na temelju filodinamike pojašnjeni su evolucijski događaji i razvoj novih varijanti virusa SARS-CoV-2. Tako je potvrđeno da je mutacija N501Y uočena kod alfa varijante postojala već i ranije kod beta varijante. Ustanovljeno je kako neke od varijanti koje izazivaju zabrinutost imaju identične mutacije što je rezultat konvergentne evolucije, primjerice varijante koje nose N501Y i E484K mutacije (beta i gama) pojavile su se neovisno na više geografski udaljenih

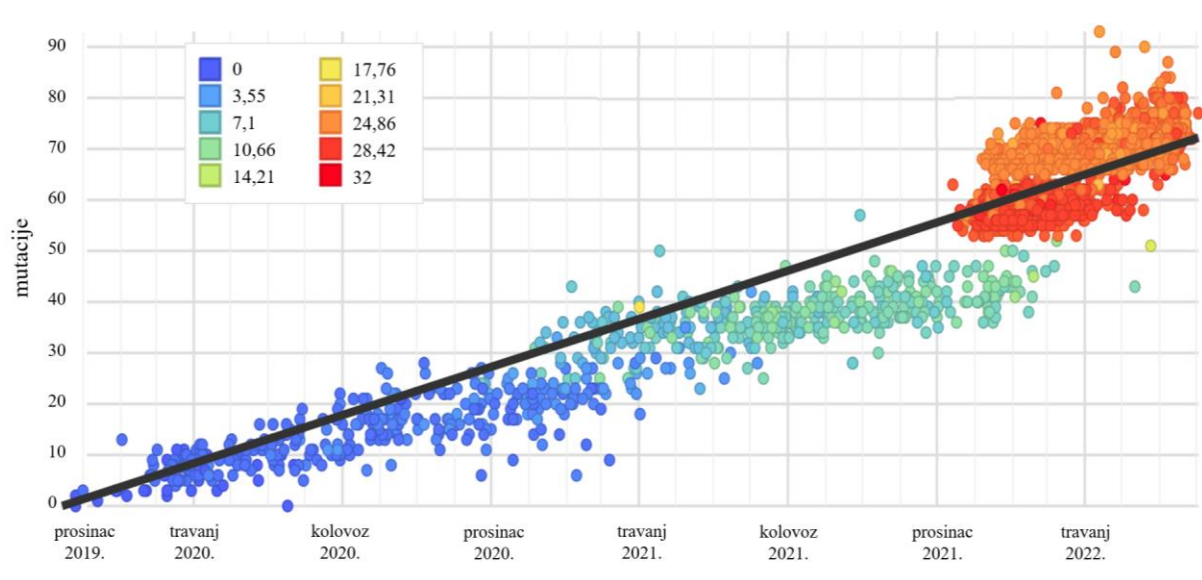
lokacija te ih stoga nazivamo homoplazijama (karakteristika koja je zajednička dvama organizmima, a nije postojala kod njihovog zadnjeg zajedničkog pretka) (Attwood *i sur.*, 2022). Upravo ta istovremena pojava mutacija implicira na postojanje istih selektivnih pritisaka koji djeluju na virus i uzrokuju zadržavanje te fiksiranje određenih mutacija unutar populacije (Gutierrez, Escalera-zamudio i Pybus, 2020). Seleksijski pritisci mogu biti širokog spektra, primjerice potreba za većom prenosivosti na nove domaćine ili zaobilaženjem imunosnog odgovora domaćina (Day *i sur.*, 2020). Osim tijeka evolucije moguće je odrediti i redosljed kojim su se mutacije pojavljivale, primjerice kod gama varijante do pojave mutacija specifičnih za tu varijantu došlo je u dvije faze u razmaku od nekoliko mjeseci. Također filodinamika može pomoći u otkrivanju podrijetla virusa ili virusnih varijanti. Nagla pojava varijanti s 30 ili više novih značajnih supstitucija i/ili delecija u odnosu na genom pretka je malo vjerojatna. Ukoliko dođe do pojave takvog događaja vjerojatno je do njega došlo zbog prolongirane infekcije u imunosuprimiranom domaćinu ili većeg broja infekcija koje nisu bile detektirane i uzorkovane te kasnije ponovnog vraćanja u populaciju i uzorkovanja. Primjer takve nagle pojave genetski značajno različitiije varijante je pojava omikron varijante. Osim spomenutog slučaja na promjene u frekvenciji mutacija tijekom epidemije mogu utjecati i brojni drugi čimbenici poput efekta osnivača i nejednolikog uzorkovanja stoga to također treba uzeti u obzir prilikom filodinamičkih istraživanja i donošenja zaključaka (Attwood *i sur.*, 2022).

6. Globalna inicijativa za dijeljenje svih podataka o gripi

Inicijativa GISAID (engl. *The Global Initiative on Sharing All Influenza Data*) je inicijativa u sklopu koje je pokrenuta internetska platforma u svrhu brzog i jednostavnije distribucije podataka o virusu gripe i koronavirusima. Uključuje sekvencionirane genetičke informacije i epidemiološke podatke, a ideja je tim načinom doprinijeti i olakšati istraživanja s ciljem boljeg razumijevanja kako virusi evoluiraju i rasprostranjuju se tijekom epidemija i pandemija (Yang *i sur.*, 2020). Tijekom prijašnjih epidemija kao što su ona koju uzrokuje virus svinjske gripe (H1N1) iz 2009. godine te ptičje gripe (H7N9) iz 2013. godine razvijene su metode za dijeljenje podataka te je stoga vrlo brzo nakon pojave virusa SARS-CoV-2 razvijena i EpiCoV baza podataka koja je uvelike doprinijela i ubrzala razvoju dijagnostičkih metoda i cjepiva (Khare *i sur.*, 2021). Razvijene su i brojne kolaboracije odnosno platforme sa sličnim ciljevima, a neki od primjera gdje su korišteni podatci preuzeti iz GISAID baze podataka su; *CoV-Spectrum*, *CoVerage*, *CovidTag*, *CoVGlobe*, *CoronaTrend*, *CoVariants*. Brojne grafičke prikaze, vizualizacije i analize podataka iz GISAID baze napravljene su u okviru *Next strain* projekta. To je projekt koji za cilj ima pružiti javnosti prikaze javno dostupnih podataka o evoluciji patogena i širenju epidemija. Putem otvorenih kodova žele omogućiti korištenje njihovih alata kako bi svatko mogao vizualizirati podatke na isti način te time doprinijeti razumijevanju epidemija i njihovu rješavanju (Hadfield *i sur.*, 2018).

Kao referentna sekvenca korištena u bazi odabrana je genska sekvenca izolirana 2019. iz uzorka iz Wuhana koji predstavlja jedan od prvih slučajeva infekcije koja je uzrokovala bolest COVID-19 te je izolat nazvan WIV04 (Okada *i sur.*, 2020). Na GISAID platformi moguće je pronaći vizualizacije genomske epidemiologije SARS-CoV-2 uz poduzorkovanje na globalnoj razini, a neki od grafičkih prikaza su filogenetsko stablo, geografska rasprostranjenost s naznačenim udjelima pojedinih varijanti, raznolikost genoma s naznačenim mjestima čestih mutacija te frekvencije određenih varijanti kroz vrijeme. S obzirom na to da u bazi postoji skoro 12 milijuna cjelovito sekvencioniranih genoma virusa SARS-CoV-2, a na grafičkim prikazima je moguće jasno prikazati tek oko 4 000 genoma, postojeći podatci su poduzorkovani, odnosno u svrhu vizualizacije korišten je samo dio podataka.

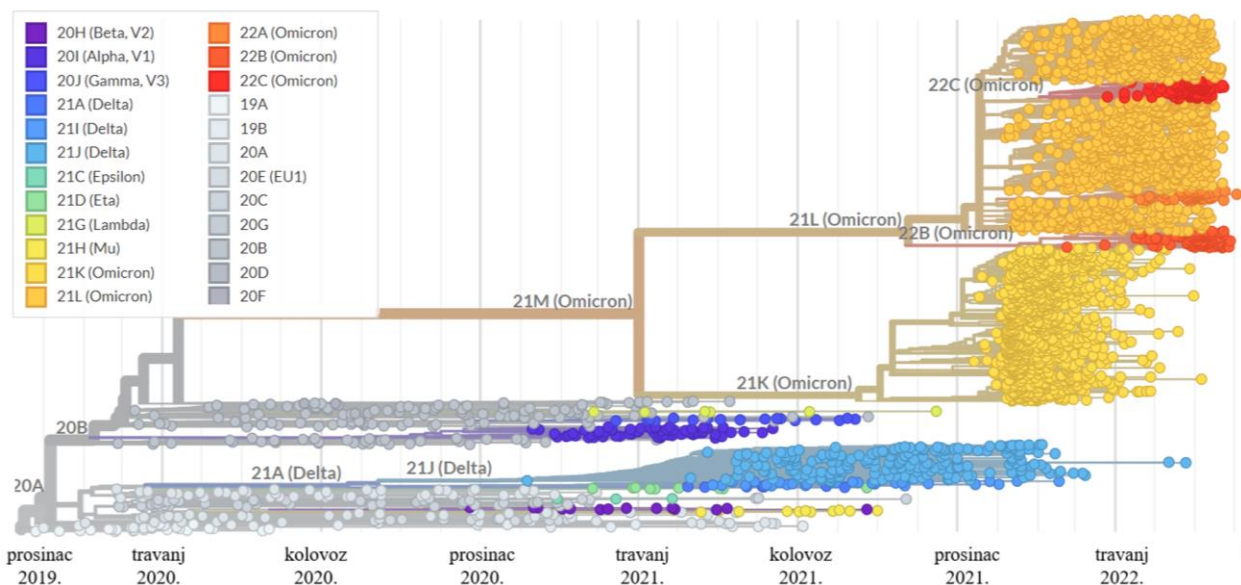
Filogenetski prikaz je moguće prilagoditi i prikazati na više načina, ovisno o tome što se njime želi vidjeti i istaknuti. Na Slici 8. prikazana je ovisnost broja mutacija u S1 podjedinici S proteina virusa SARS-CoV-2 o vremenu u periodu od prosinca 2019. godine do srpnja 2022. godine. Može se uočiti kako se s vremenom mutacije nakupljaju te da taj porast kroz vrijeme ima linearan trend.



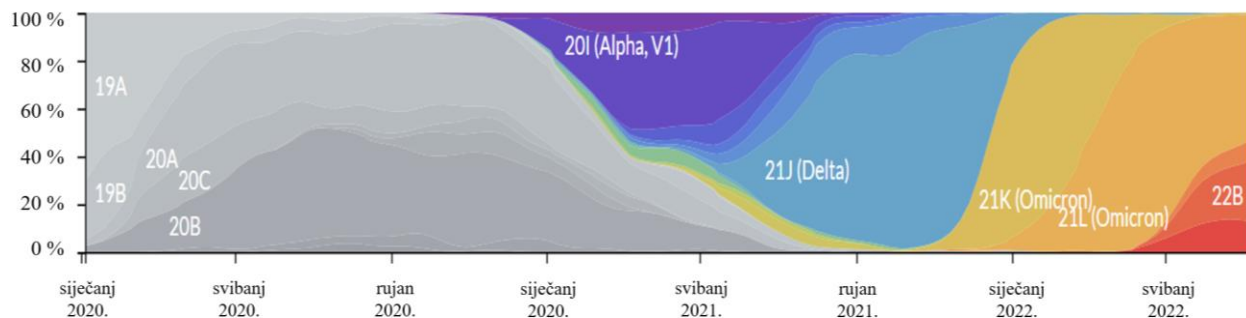
Slika 8. Ovisnost broja mutacija u S1 podjedinici S proteina virusa SARS-CoV-2 o vremenu u periodu od prosinca 2019. godine do srpnja 2022. godine. Različite boje prema legendi predstavljaju broj mutacija. Prikaz je izrađen analizom 2 928 uzoraka sekvencioniranih genoma koji se nalaze u bazi podataka GISAID. Preuzeto s <https://www.gisaid.org/phylogenetics/global/nextstrain/> (pristupljeno 8. srpnja 2022.).

Usporedbom filogenetskog stabla varijanti virusa SARS-CoV-2 (Slika 9.) i grafičkog prikaza frekvencija pojedinih virusnih varijanti (Slika 10.) jasno je vidljiva analogija. Početkom 2021. godine prvi puta se pojavljuju varijante koje izazivaju zabrinutost, odnosno one koje izazivaju povećanu stopu širenja zaraze. Iz kladograma je vidljivo kako se alfa, beta i gama varijante razvijaju i šire neovisno jedna o drugoj u sličnom vremenskom periodu. Nešto kasnije, početkom 2022. godine dolazi do pojave još jedne varijante koja izaziva zabrinutost, a to je omikron. Omikron

se od dotada poznatih varijanti razlikuje u dvije stvari; ima neočekivano velik broj mutacija i uzrokuje izuzetno velik broj infekcija.



Slika 9. Kronogram koji prikazuje filogeniju virusa SARS-CoV-2. Različite boje prema legendi predstavljaju različite virusne varijante. Prikaz je zrađen analizom 2 928 uzoraka sekvencioniranih genoma koji se nalaze u bazi podataka GISAID. Preuzeto s <https://www.gisaid.org/phylogenetics/global/nextstrain/> (pristupljeno 8. srpnja 2022.).



Slika 10. Grafički prikaz frekvencija virusnih varijanti. Prikaz je izrađen analizom 2 928 uzoraka sekvencioniranih genoma koji se nalaze u bazi podataka GISAID. Preuzeto s <https://www.gisaid.org/phylogenetics/global/nextstrain/> (pristupljeno 8. srpnja 2022.).

7. Zaključak

Pandemija koju je uzrokovao virus SARS-CoV-2, kao i sve pandemije, pojavila se neočekivano, ali u vrlo kratkom periodu od njezine pojave znanstvenici su se upoznali i razotkrili brojne bitne informacije o samom uzročniku pandemije, njegovoj evoluciji, širenju zaraze te utjecaju na ljudsko zdravlje. Kao i svaki organizam i virus evoluiraju pod utjecaj selekcijskih pritisaka koji uzrokuju fiksiranje određenih mutacija u njegovu genomu. Filogenetskim analizama i rekonstrukcijom filogenetskih stabala koja daju uvid u međusobne odnose omogućeno je praćenje pojave i širenja mutacija u prostoru čime se dobiva uvid u evoluciju virusa. Globalna internetska platforma GISAID predstavlja inicijativu koja je uvelike doprinijela brojnim znanstvenicima i istraživačkim timovima. U samo nekoliko klikova svima su javno dostupne izuzetno velike količine kvalitetnih informacija o genskim sekvencama i epidemiologiji čijom se obradom dobiva uvid u brojne aspekte koronavirusa i virusa gripe. Na temelju podataka iz GISAID baze svakim danom objavljeno je sve više znanstvenih radova koji prate filodinamiku virusa SARS-CoV-2, a te informacije omogućuju preciznije usmjerivanje budućih pokušaja kontrole i prevencije širenja bolesti. Filodinamika će nastaviti igrati jednu od glavnih uloga u daljnjem praćenju kako ove, trenutno aktualne, tako i svih budućih pandemija i epidemija uzrokovanih mikroorganizmima.

8. Literatura

- Abdelghany, T. M., Ganash, M., Bakri, M. M., Qanash, H., Al-Rajhi, A. M. H., Elhussieny, N. I. (2021): SARS-CoV-2, the other face to SARS-CoV and MERS-CoV: Future predictions. *Biomedical Journal. Elsevier Ltd*, 44(1): 86–93.
- Patwardhan, A., Ray, S., Roy, A. (2014): Molecular Markers in Phylogenetic Studies-A Review. *Journal of Phylogenetics & Evolutionary Biology*. 02(02).
- Arbiza, L., Ptericio, M., Dopazo, H., Posada, D. (2011). Genome-wide heterogeneity of nucleotide substitution model fit. *Genome Biology and Evolution*, 3(1): 896–908.
- Arenas, M. (2015): Trends in substitution models of molecular evolution. *Frontiers in Genetics*. 6:319.
- Attwood, S. W., Hill, S. C., Aanensen, D. M., Connor, T. R., Pybus, O. G. (2022): Phylogenetic and phylodynamic approaches to understanding and combating the early SARS-CoV-2 pandemic. *Nature Reviews Genetics*. Springer US, 012456789.
- Banho, C. A. *i sur*. (2022): Impact of SARS-CoV-2 Gamma lineage introduction and COVID-19 vaccination on the epidemiological landscape of a Brazilian city. *Communications Medicine*. Springer US, 2(1): 1–11.
- Bashor, L., Gagne, R. B., Bosco-Lauth, A. M., Bowen, R. A., Stenglein, M. i VandeWoude, S. (2021): SARS-CoV-2 evolution in animals suggests mechanisms for rapid variant selection. *Proceedings of the National Academy of Sciences of the United States of America*, 118(44): 1–10.
- Bromham, L. i Penny, D. (2003): The modern molecular clock. *Nature Reviews Genetics*, 4(3): 216–224.
- Crandall, K. A. i Posada, D. (2002): Phylogenetic Approaches to Molecular Epidemiology. *The Molecular Epidemiology of Human Viruses*: 25–40.
- Day, T., Gandon, S., Lion, S. i Otto, S. P. (2020): On the evolutionary epidemiology of SARS-CoV-2. *Current Biology Magazine*, 30: 849–857.
- Duffy, S. (2018): Why are RNA virus mutation rates so damn high?. *PLoS Biology*, 16(8): 1–6.
- Engelhardt, D. i Shakhnovich, E. I. (2019): Mutation rate variability as a driving force in adaptive evolution. *Physical Review E. American Physical Society*, 99(2): 1–7.
- European Centre for Disease Prevention and Control (ECDC). Situation updates on COVID-19. <https://www.ecdc.europa.eu/en/covid-19/variants-concern> (pristupljeno 27. lipnja 2022.).

- Fraser, B. J. *i sur.* (2022): Structure and activity of human TMPRSS2 protease implicated in SARS-CoV-2 activation. *Nature Chemical Biology*.
- Gitman, M. R., Shaban, M. V i Paniz-mondolfi, A. E. (2021): Laboratory Diagnosis of SARS-CoV-2 Pneumonia. *Diagnostics*, 11, 1270: 1–17.
- Global Virus Network. COVID-19 Variants, Beta (B.1.351) (2022). <https://gvn.org/covid-19/beta-b-1-351/> (pristupljeno 7. srpnja 2022.).
- Global Virus Network. COVID-19 Variants, Delta (B.1.617.2) (2022). <https://gvn.org/covid-19/delta-b-1-617-2/> (pristupljeno 7. srpnja 2022.).
- Gregory, T. R. (2008): Understanding Evolutionary Trees. *Evolution: Education and Outreach*, 1(2): 121–137.
- Grenfell, B. T., Pybus, O. G., Gog, J. R., Wood, J. L. N., Daly, J. M., Mumford, J. A., Holmes i E. C. (2004): Unifying the Epidemiological and Evolutionary Dynamics of Pathogens. *Science*, 303: 327–332.
- Gutierrez, B., Escalera-zamudio, M. i Pybus, O. G. (2020): Parallel molecular evolution and adaptation in viruses. *Current Opinion in Virology*, 34: 90–96.
- Hadfield, J., Megill, C., Bell, S. M., Huddleston, J., Potter, B., Callender, C., (2018): NextStrain: Real-time tracking of pathogen evolution“, *Bioinformatics*, 34(23): 4121–4123.
- Hall, B. G. (2013): Building phylogenetic trees from molecular data with MEGA“, *Molecular Biology and Evolution*, 30: 1229–1235.
- Hall, B. G. i Barlow, M. (2006): Phylogenetic analysis as a tool in molecular epidemiology of infectious diseases. *Annals of Epidemiology*, 16(3): 157–169.
- Hartenian, E., Nandakumar, D., Lari, A., Ly, M., Tucker, J. M. i Glaunsinger, B. A. (2020): The molecular virology of coronaviruses. *Journal of Biological Chemistry*, 295: 12910–12934.
- Harvey, W. T. *i sur.* (2021): SARS-CoV-2 variants, spike mutations and immune escape. *Nature Reviews Microbiology*. Springer US, 19: 409–424.
- Hoffmann, M. *i sur.* (2020): SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven Protease Inhibitor. *Cell*, 181: 271-280.
- Hogan, M. J. i Pardi, N. (2022): mRNA Vaccines in the COVID-19 Pandemic and Beyond. *Annual Review of Medicine*, 73: 17–39.
- International Committee on Taxonomy of Viruses (ICTV) (2022). <https://talk.ictvonline.org/> (pristupljeno 6. lipnja 2022.).

- Jung, C., Kmiec, D., Koepke, L., Zech, F., Jacob, T., Sparre, K. M. J. i Kirchhoff, F. (2022): Omicron: What Makes the Latest SARS-CoV-2 Variant of Concern So Concerning?. *Journal of Virology*. American Society for Microbiology, 96(6).
- Ke, H. M., Yu. C. P., Liu, Y. C., Tsai, I. J. (2017): Popmarker: Identifying Phylogenetic Markers at the Population Level. *Evolutionary Bioinformatics*, 13.
- Khare, S., Gurry, C., Freitas, L., Schultz, M. B., Bach, G., Diallo, A., Akite, N., (2021): GISAID's Role in Pandemic Response. *China CDC Weekly*, 3(49): 1049–1051.
- Kilbourne, E. D. (1973): The molecular epidemiology of influenza. *Journal of Infectious Diseases*, 127(4): 478–487.
- Lauring, A. S. (2020): Within-Host Viral Diversity: A Window into Viral Evolution. *Annual Review of Virology*, 7: 63–81.
- Lemieux, J. E. *i sur.* (2021): Phylogenetic analysis of SARS-CoV-2 in Boston highlights the impact of superspreading events. *Science*, 371(6529).
- Liu, J. *i sur.* (2021): The N501Y spike substitution enhances SARS-CoV-2 transmission. *bioRxiv*.
- Maier, H. J., Bickerton, E. i Britton, P. (2015): Coronaviruses: Methods and protocols. *Coronaviruses: Methods and Protocols*, 1282(1): 1–282.
- Mascellino, M. T., Timoteo, F. D., De Angelis, M., Oliva A. (2021): Overview of the Main Anti-SARS-CoV-2 Vaccines: Mechanism of Action, Efficacy and Safety. *Infection and Drug Resistance*, 14: 4501–4502.
- Medina-Enríquez, M. M., Lopez-Leon, S., Carlos-Escalante, J. A., Aponte-Torres, Z., Cuapio, A. i Wegman-Ostrosky, T. (2020): ACE2: the molecular doorway to SARS-CoV-2. *Cell and Bioscience*. *BioMed Central*, 10(1): 1–17.
- Mpolya, E. A., Furuse, Y., Nukiwa, N., Suzuki, A., Kamigaki, T. i Oshitani, H. (2009): Pandemic (H1N1) 2009 Virus Viewed from an Epidemiological Triangle Model. *Journal of Disaster Research*, 4(5): 356–364.
- Nonaka, C. K. V. *i sur.* (2021): SARS-CoV-2 variant of concern P.1 (Gamma) infection in young and middle-aged patients admitted to the intensive care units of a single hospital in Salvador, Northeast Brazil, February 2021. *International Journal of Infectious Diseases*. Elsevier Ltd, 111: 47–54.

- Okada, P. *i sur.* (2020): Early transmission patterns of coronavirus disease 2019 (COVID-19) in travellers from Wuhan to Thailand, January 2020. *Eurosurveillance*. European Centre for Disease Prevention and Control (ECDC), 25(8).
- Oude Munnink, B. B., Worp, N., Nieuwenhuijse, D. F., Sikkema, R. S., Haagman, B., Fouchier, R. A. M. i Koopmans, M. (2021): The next phase of SARS-CoV-2 surveillance: real-time molecular epidemiology. *Nature Medicine*. Springer US, 27(9): 1518–1524.
- Pascual-García, A., Arenas, M. i Bastolla, U. (2019): The Molecular Clock in the Evolution of Protein Structures. *Systematic Biology*, 68(6): 987–1002.
- Phan, T. (2020): Genetic diversity and evolution of SARS-CoV-2. *Infection, Genetics and Evolution*. Elsevier, 81(February): 104260.
- Posada, D. i Crandall, K. A. (2001): Selecting the best-fit model of nucleotide substitution. *Systematic Biology*, 50(4): 580–601.
- Postillone, M. B. i Perez, S. I. (2017): Mitochondrial-DNA phylogenetic information and the reconstruction of human population history: The south american case. *Human Biology*, 89(3): 229–250.
- Riapis, L. A., Filatov, N. N. i Salova, N. I. (2003): Molecular epidemiology of infectious diseases. *Zhurnal mikrobiologii, epidemiologii, i immunobiologii*, (5): 49–54.
- Rotondo, J. C. *i sur.* (2021): Sars-cov-2 infection: New molecular, phylogenetic, and pathogenetic insights. efficacy of current vaccines and the potential risk of variants, *Viruses*, 13, 1687.
- Saito, A. *i sur.* (2021): SARS-CoV-2 spike P681R mutation, a hallmark of the Delta variant, enhances viral fusogenicity and pathogenicity. *bioRxiv*.
- Sanyaolu, A., Okorie, C., Marinkovic, A., Haider, N., Abbasi, A. F., Jaferi, U., Prakash, S. i Balendra V. (2021): The emerging SARS-CoV-2 variants of concern. *Therapeutic Advances in Infectious Disease*, 8: 9–11.
- Sanjuán, R., Nebot, M. R., Chirico, N., Mansky, L. M. i Belshaw, R. (2010): Viral Mutation Rates. *Journal of Virology*, 84(19): 9733–9748.
- Sarkar, R., Lo, M., Saha, R., Dutta, S. i Chawla-Sarkar, M. (2021): S glycoprotein diversity of the Omicron variant. *medRxiv*.
- Shen, X. *i sur.* (2021): SARS-CoV-2 variant B.1.1.7 is susceptible to neutralizing antibodies elicited by ancestral Spike vaccines. *bioRxiv*.

- Shirbhate, E., Pandey, J., Patel, V., K., Kamal, M., Jawaid, T., Gorain, B., Kesharwani, P. i Rajak, H. (2021): Understanding the role of ACE-2 receptor in pathogenesis of COVID-19 disease: a potential approach for therapeutic intervention. *Pharmacological Reports*. Springer International Publishing, 73(6): 1539–1550.
- Singh, D. i Yi, S. V. (2021): On the origin and evolution of SARS-CoV-2. *Experimental and Molecular Medicine*. Springer US, 53(4): 537–547.
- The Global Initiative on Sharing All Influenza Data. Genomic epidemiology of SRAS-CoV-2 with subsampling focused globally over the past 6 months (2022). <https://www.gisaid.org/phylogenomics/global/nextstrain/> (pristupljeno 8. srpnja 2022.).
- Volz, E. M., Koelle, K. i Bedford, T. (2013): Viral Phylodynamics. *PLoS Computational Biology*, 9(3).
- WHO COVID-19 Dashboard. Geneva: World Health Organization. (2020). <https://covid19.who.int/> (pristupljeno 6. lipnja 2022.).
- Wrapp, D., Wang, N., Corbett, K. S., Goldsmith, J. A., Hsieh, C. L., Abiona, O., Graham, B. S. i McLellan, J. S. (2020): Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. *Science*, 367(6483): 1260–1263.
- Yang, Y. *i sur.* (2020): SARS-CoV-2: characteristics and current advances in research. *Virology Journal*, 17(1): 1–17.
- Yang, Z. i Rannala, B. (2012): Molecular phylogenetics: Principles and practice. *Nature Reviews Genetics*. Nature Publishing Group, 13(5): 303–314.
- Zhang, Q., Xiang, R., Huo, S., Zhou, Y., Jiang, S., Wang, Q. i Yu, F. (2021): Molecular mechanism of interaction between SARS-CoV-2 and host cells and interventional therapy. *Signal Transduction and Targeted Therapy*. Springer US, 6(1).
- Zhou, D. *i sur.* (2021): Evidence of escape of SARS-CoV-2 variant B.1.351 from natural and vaccine-induced sera. *Cell*. Elsevier Inc., 184(9): 2348-2361.e6.
- Zhou, H. *i sur.* (2021): A Review of SARS-CoV2: Compared With SARS-CoV and MERS-CoV. *Frontiers in Medicine*, 8(December): 1–14.

9. Životopis

Mia Schwerer rođena je 28. rujna 2000. godine u Zagrebu. Završila je dvojezični program XVIII. Gimnazije u Zagrebu na njemačkom jeziku te upisala preddiplomski studij Molekularne biologije na Prirodoslovno-matematičkom fakultetu u Zagrebu. Od 2016. godine volontira kao mentorica u Znanstveno edukacijskom centru Višnjan gdje u sklopu programa *Youth science camp* i *Summer school of science* osmišljava i provodi sedmodnevne projekte za djecu osnovnoškolske i srednjoškolske dobi. Tijekom prve godine studija bila je uključena u rad studentske udruge eSTUDENT te je kao suvoditeljica tima Popularizacija znanosti sudjelovala u organizaciji dvaju projekata; Mozak voli zdravo, u sklopu kojeg je organizirana jednodnevna konferencija i brojne radionice te Znanost za sve, koji obuhvaća organizaciju znanstvenih predavanja za studente i osmišljanje znanstvenih radionica za djecu. Također, od rujna 2020. godine aktivna je članica Udruge studenata biologije – BIUS, a u siječnju 2021. godine postala je glavna urednica studentskog časopisa *In Vivo*. Tijekom preddiplomskog studija sudjelovala je i u osmišljavanju i snimanju videa za manifestaciju Noć biologije. U akademskoj godini 2021./2022. bila je i studentska predstavica u Vijeću Biološkog odsjeka te dio organizacijskog odbora Simpozija studenata bioloških usmjerenja 2022. Dobitnica je i Rektorove nagrade za individualni znanstveni rad pod naslovom „Fenotipska i genotipska varijabilnost izolata roda *Saprolegnia* iz Hrvatskih pastrvskih ribnjaka“.