



Zdravotně  
sociální fakulta  
Faculty of Health  
and Social Sciences

Jihočeská univerzita  
v Českých Budějovicích  
University of South Bohemia  
in České Budějovice

**Fylogenetická analýza a molekulární detekce  
koronavirů**

## **BAKALÁŘSKÁ PRÁCE**

Studijní program:

**SPECIALIZACE VE ZDRAVOTNICTVÍ/ ZDRAVOTNÍ  
LABORANT**

**Autor:** Lucie Marhounová

**Vedoucí práce:** doc. RNDr. Miroslav Šíp, DrSc.

České Budějovice 2022

## **Prohlášení**

Prohlašuji, že svoji bakalářskou práci s názvem „*Fylogenetická analýza a molekulární detekce koronavirů*“ jsem vypracoval/a samostatně pouze s použitím pramenů v seznamu citované literatury.

Prohlašuji, že v souladu s § 47b zákona č. 111/1998 Sb. v platném znění souhlasím se zveřejněním své bakalářské práce, a to v nezkrácené podobě elektronickou cestou ve veřejně přístupné části databáze STAG provozované Jihočeskou univerzitou v Českých Budějovicích na jejích internetových stránkách, a to se zachováním mého autorského práva k odevzdanému textu této kvalifikační práce. Souhlasím dále s tím, aby toutéž elektronickou cestou byly v souladu s uvedeným ustanovením zákona č. 111/1998 Sb. zveřejněny posudky školitele a oponentů práce i záznam o průběhu a výsledku obhajoby bakalářské práce. Rovněž souhlasím s porovnáním textu mé bakalářské práce s databází kvalifikačních prací Theses.cz provozovanou Národním registrem vysokoškolských kvalifikačních prací a systémem na odhalování plagiátů.

V Českých Budějovicích dne

.....

### **Poděkování**

Touto cestou bych ráda poděkovala mému vedoucímu bakalářské práce doc. RNDr. Miroslavu Šípovi, DrSc. za jeho odborné vedení a poskytnutí cenných rad při konzultacích v rámci bakalářské práce. Dále bych ráda poděkovala paní Mgr. Dagmar Riegert Bystřické, Ph.D. za praktické seznámení s detekční metodou PCR. V neposlední řadě náleží velký dík mé rodině, příteli a blízkým přátelům za trpělivost a podporu.

# Fylogenetická analýza a molekulární detekce koronavirů

## Abstrakt

Cílem této bakalářské práce bylo seznámení se s tvorbou fylogenetických stromů, které popisují vztahy mezi koronaviry, porovnání a posouzení vhodnosti používaných a navržených primerů pro detekci nového typu koronaviru SARS-CoV-2.

Dle fylogenetických stromů lze určit evoluční vztah zkoumaných organismů, které v průběhu evoluce podléhají změnám ve svých sekvencích tzv. mutacím. Tyto změny určují jejich vývoj.

Teoretická část bakalářské práce byla zaměřena na obecné seznámení se s virem a následně koronaviry, kam se nově řadí SARS-CoV-2. SARS-CoV-2 představoval a pořád představuje velkou hrozbu pro celý svět. Dále byla probraná problematika fylogenetiky a na závěr detekční metoda PCR.

V praktické části byly vytvořeny fylogenetické stromy pomocí počítačového programu MEGA-X za využití metody Neighbor-Joining, které zkoumaly fylogenetické vztahy mezi koronaviry zaměřené hlavně na SARS-CoV-2. Nejbližší koronaviry tohoto nového typu koronaviru SARS-CoV-2 byly dále porovnávány pomocí speciálního softwaru BLAST, který uváděl procentuální shodu studovaných sekvencí. Během těchto analýz se zjistilo, že SARS-CoV-2 je nejvíce podobný netopýřímu koronaviru RaTG13.

Primery, které se běžně používají při detekci SARS-CoV-2 nejsou zveřejňovány, proto byly použity primery z webové aplikace CoVrimer, kde jsou dostupné primery, které byly použity pro vědecké výzkumy. Ty byly porovnány s návrhem primerů. Návrh byl proveden v rámci řešení bakalářské práce pomocí speciální aplikace Pirmer-BLAST. Primery, které byly získány z aplikace CoVrimer, nebyly navrženy pro komplexní rozlišení všech studovaných variant SARS-CoV-2.

## Klíčová slova

koronavirus; SARS-CoV-2; mutace; fylogeneze; MEGA-X; ClustalW; Neighbor-Joining; PCR

## **Phylogenetic analysis and molecular detection of coronaviruses**

### **Abstract**

The aim of this bachelor thesis was to get acquainted with the creation of phylogenetic trees, which describe the relationships between coronaviruses and compare and assess the suitability of used and designed primers for the detection of a new type of coronavirus SARS-CoV-2.

According to phylogenetic trees, it is possible to determine the evolutionary relationship of the studied organisms, which during evolution are subject to changes in their sequences, so-called mutations. These changes determine their development.

The theoretical part of the bachelor thesis was focused on general acquaintance with viruses and subsequently coronaviruses and a new type of coronavirus SARS-CoV-2. SARS-CoV-2 has posed and continues to pose a major threat to the world. Furthermore, the issue of phylogenetics and finally the PCR detection method were discussed.

In the practical part, phylogenetic trees were created using the computer program MEGA-X using the Neighbor-Joining method, which examined the phylogenetic relationships between coronaviruses focused mainly on SARS-CoV-2. The nearest coronaviruses of this new type of SARS-CoV-2 coronavirus were further compared using special BLAST software, which reported the percent agreement of the studied sequences. During these analyzes, SARS-CoV-2 was found to be most like the bat coronavirus RaTG13.

Primers that are commonly used to detect SARS-CoV-2 are not disclosed, therefore primers from the CoVrimer web application have been used, where primers that have been used for scientific research are available. These were compared with the primers designed by the author. The design was performed as part of a bachelor's thesis using a special application Primer-BLAST. The primers obtained from the CoVrimer application were not designed to comprehensively distinguish all SARS-CoV-2 variants studied.

### **Key words**

coronavirus; SARS-CoV-2; mutation; phylogeny; MEGA-X; ClustalW; Neighbor-Joining; PCR

## Obsah

Úvod.....	8
1. Viry.....	9
1.1 Historie.....	9
1.2 Struktura virů .....	9
1.2.1 Virová částice .....	10
1.2.2 Kapsida .....	10
1.3 Klasifikace.....	10
1.4 Dělení dle přítomnosti obalu.....	11
1.4.1 Neobalené viry.....	11
1.4.2 Obalené viry.....	11
1.5 Replikace.....	12
1.6 Mutace.....	14
1.7 Patogeneze a rozvoj infekce.....	14
1.7.1 Brána vstupu .....	14
1.7.2 Primární pomnožení.....	15
1.7.3 Sekundární pomnožení .....	15
2. Koronaviry.....	16
2.1 Seznámení .....	16
2.2 Výskyt, šíření a klinické příznaky.....	17
2.3 SARS-CoV-2.....	18
2.4 Onemocnění .....	18
2.5 Varianty SARS-CoV-2.....	19
2.5.1 Varianta B.1.1.7.....	19
2.5.2 Varianta B.1.351 .....	20
2.5.3 Varianta P.1 .....	20
2.5.4 Varianta B.1.617.2.....	20
2.5.5 Varianta B.1.1.529 .....	20
3. Fylogeneze.....	21
3.1 Fylogenetické stromy .....	21
3.1.1 Zakořeněné stromy .....	22
3.1.2 Nezakoreněné stromy .....	23
3.2 Význam mutací pro fylogenetiku.....	24
4. Molekulární detekce .....	25

4.1	PCR .....	25
4.1.1	Historie metody.....	25
4.1.2	Princip.....	26
4.2	RT-PCR.....	27
4.2.1	Princip RT-PCR.....	27
5.	Cíl práce a výzkumné otázky .....	28
5.1	Cíl práce .....	28
6.	Metodika a metody .....	29
6.1	Fylogenetická analýza.....	29
6.2	Materiál .....	29
6.2.1	GenBank .....	31
6.3	Metodika rekonstrukce fylogenetických stromů .....	31
6.4	MEGA-X.....	31
6.5	Metody .....	32
6.5.1	ClustalW .....	32
6.5.2	Neighbor-Joining .....	32
6.6	BLAST .....	33
6.7	Primery .....	34
6.8	Běžně používané PCR primery pro detekci SARS-CoV-2 .....	34
6.9	Návrh primerů .....	37
7.	Výsledky.....	39
7.1	Fylogenetická analýza.....	39
7.2	Primery .....	47
8.	Diskuse .....	50
8.1	Fylogenetické souvislosti.....	50
8.2	Primery .....	53
9.	Závěr.....	54
10.	Literatura.....	55
11.	Seznam obrázků a tabulek .....	61
11.1	Seznam obrázků .....	61
11.2	Seznam tabulek .....	61
12.	Seznam zkratk .....	62
13.	Přílohy.....	64
13.1	Seznam příloh.....	64

## Úvod

Koronaviry, dříve známé jako „nevýznamná“ čeleď virů, která způsobovala běžné onemocnění, které se projevovalo rýmou a kašlem, jsou dnes známé po celém světě, a to hlavně ve spojení s novým typem koronaviru SARS-CoV-2. Tento nový typ koronaviru je 7. typem koronaviru, který způsobuje onemocnění u lidí. Jednotlivé typy se dále vyvíjejí, tím se mění i jejich schopnost infikovat člověka a dále se mění i obraz vznikajícího onemocnění. V současné době je SARS-CoV-2 velkým problémem. Způsobuje závažná respirační onemocnění, na která zemřelo již mnoho lidí.

Základem mapování a hledání nových mutací tohoto viru je zarovnání sekvencí a fylogenetické vztahy se zkoumají sestavováním fylogenetických stromů. Tyto stromy zkoumají podobnost jednotlivých typů SARS-CoV-2.

Nejběžnější metodou stanovení nového typu koronaviru je již od první vlny metoda PCR a přímo pro detekci SARS-CoV-2 a jeho mutačních variant je nejběžnější metoda RT-PCR. Mutační varianty jsou rozeznávány právě na základě jejich charakteristických mutací.

Součástí teoretické části práce je popis virů od jejich historie, struktury až po jejich replikaci a následný průběh infekce. Dále je představena rodina koronavirů, kde je blíže popsán SARS-CoV-2. Další část zahrnuje přiblížení tematiky fylogeneze a fylogenetických stromů. A na závěr se teoretická část věnuje molekulární analýze, kde je blíže představena PCR metoda.

Praktická část bakalářské práce je zaměřena na fylogenetickou analýzu koronavirů, kde se zkoumá jejich podobnost vůči novému typu koronaviru SARS-CoV-2, pro pochopení jeho původu, jsou použity i sekvence zvířecích koronavirů. Ty koronaviry, které se vyskytují ve fylogenetických stromech nejbliže sekvenci SARS-CoV-2, jsou zkoumány blíže a to přesně, tak že byla zkoumána jejich procentuální podobnost s touto sekvencí.

Dále byly zkoumány primery pro detekci SARS-CoV-2 a rozlišení jejich mutačních variant. Tyto primery byly porovnávány, na základě toho, na jaký úsek sekvencí cílí a jak jsou komplexní pro rozlišení studovaných variant SARS-CoV-2.



# 1. Viry

## 1.1 Historie

V latině slovo virus původně znamenalo „šťáva či jed“ a v lékařské terminologii dostalo význam „infekční činitel“. Ve druhé polovině 19. století se o virech začalo hovořit jako o submikroskopických infekčních agens, tedy jako o mikroorganismech, které nelze spatřit v optickém mikroskopu. [53]

Německý chemik Adolf Mayer (1843–1942) byl první komu se pokusně podařilo přenést virové onemocnění. Při studii v Holandsku se mu povedlo přenést mosaikovou chorobu tabáku na zdravou rostlinu šťávou z rozestřených listů. Potvrdil to Rus Dimitrij Ivanovskij (1864–1920). Ten v roce 1892 dokázal, že šťáva z nakažených listů je i po filtraci přes Chamberlandovu svíčku (porcelánový filtr, který obsahoval jemné póry zadržující bakterie) infekční. [53]

Martinus Willem Beijerinck byl Mayerův spolupracovník, kterému se povedlo prokázat, že šťáva z nakažených listů si i po filtraci zanechává infekčnost a že se tento mikroorganismus dokáže pomnožit. Přišel na to, že šťáva, která je přefiltrovaná se dá naředit a po pomnožení v živé rostlině získává svou původní „sílu“. Tohoto mikroba považoval za „contagium vivum fluidum“, což znamená přenosné živé tekuté agens. Poté tohoto mikroba označil jako virus. [53]

## 1.2 Struktura virů

Na rozdíl od ostatních mikroorganismů jsou viry poměrně malými (20–300 nm) nebuněčnými částicemi, které nelze rozeznat v optickém mikroskopu. Je velmi sporné, zda jsou tyto částice živé, nejsou totiž schopné samostatné reprodukce bez hostitelské buňky. Proto viry zařazujeme mezi obligátní intracelulární parazity, nelze je totiž běžně kultivovat v nebuněčném prostředí. [4]

S živými organismy mají přece jen něco společného a to, že mají vlastní genetický kód pro replikaci. Viry mají svojí genetickou informaci zapsanou vždy jen v jediném typu nukleové kyseliny, proto je rozdělujeme na DNA (dle deoxyribonukleotidové kyseliny) a RNA (dle ribonukleotidové kyseliny) viry. [4]

### **1.2.1 Virová částice**

Základní virová částice, která je schopná infikovat hostitelskou buňku, je označována jako virion. Vnitřek virionu se nazývá nukleoid. Nukleoid se skládá z nukleové kyseliny a obklopuje ho proteinová schránka, kterou nazýváme kapsida. [53]

Nukleová kyselina funguje jako genom, který nese dva druhy genů. Tyto geny jsou buď determinující strukturní bílkoviny, které vytvářejí část virionu a nestrukturní (funkční) bílkoviny neboli enzymy účastníci se stavby virionu. Příklad takových enzymů jsou DNA-polymerázy. [45]

### **1.2.2 Kapsida**

Kapsida chrání genom před okolními vlivy a slouží k adsorpci k povrchu hostitelské buňky. [45] Virová kapsida se skládá z proteinových podjednotek neboli protomer, které se opakují. Protomery jsou většinou jednoho nebo jen několika málo typů, pro nejlepší využití virového genomu. Jejich tvar jim napomáhá ke vzájemnému pravidelnému přikládání se k sobě, bez jakýchkoliv vlivů, jako jsou například enzymy. Tímto vytvářejí větší celek. Tento celek je znám jako nukleokapsida. [53]

S ohledem na symetrii, lze snadno rozlišit dva hlavní typy kapsidy, a to na kapsidy s kubickou symetrií a symetrií spirální (helikoidální). [53]

Kapsidy s kubickou (ikosaedrální) symetrií se vyskytují u většiny DNA virů a také u některých RNA virů. Tato kapsida je těleso o 12 vrcholech a 20 stěnách, které tvoří pravidelný dvacetistěn se třemi osami symetrie. Vrcholy kapsidy utvářejí stavební prvky pentony a stěny kapsidy hexony, přičemž pentony jsou tvořeny pěticí a hexony šesticí do kruhu spojených totožných molekul polypeptidů. Tyto subjednotky nazýváme kapsoméry. Jejich počet tvořící kapsidu, je pro každou čeleď virů charakteristický. [53]

U virů s helikoidální symetrií hovoříme jen o nukleokapsidě, protože se kapsida a nukleotid nedají dobře rozeznat. Tento typ kapsidy většinou rozeznáváme u řady RNA virů. Je tvořen jednotlivými polypeptidy (protoméry), které obkládají závitnici vlákna nukleové kyseliny a vzniká tak dutý válec. [53]

## **1.3 Klasifikace**

Dnes je již tisíce druhů virů, proto se používá řazení virů do hierarchického systému kategorií, z kterých jsou nejdůležitější kategorie:

Čeleď → podčeleď → rod → druh → (případně poddruh → kmen → varianta)

Názvy jsou tvořeny podobně jako u taxonomie bakterií. Můžeme tedy pozorovat označení čeledi Herpesviridae, do které spadá podčeleď Herpesvirinae a v té je jako člen rodu Herpesvirus. Názvy virů bývají často deskriptivní. Například označení Picornaviridae se skládá z pico jako malý a RNA, tudíž vyjadřuje velikost. Jako další příklad můžeme použít označení čeledi Orthomyxoviridae, které je složeno z ortho jako pravý a myxo jako hlen, což nám vyjadřuje afinitu k hleny v dýchacích cestách a podobně. Klasifikace virů probíhá podle komplexu jejich vlastností. [45]

Viry se od bakterií liší tím, že jejich genom nese pouze jeden typ nukleové kyseliny, jak bylo již zmíněno výše. Ty tedy můžeme rozdělit buď na RNA nebo DNA viry, což je základem orientačního dělení. Dle toho se viry rozdělují na skupinu DNA virů a RNA virů. Každá z těchto skupin se dále dělí na skupinu s dvouvláknovou a jednovláknovou nukleovou kyselinou a dále skupina RNA virů je rozdělena na RNA viry s pozitivním provazcem a negativním provazcem. Jako další kritéria jsou použita tvar, velikost, přítomnost či nepřítomnost obalu, symetrie kapsidy, místo replikace (v jádře či cytoplazmě) a počet kapsomer.[45]

## **1.4 Dělení dle přítomnosti obalu**

### ***1.4.1 Neobalené viry***

Viriony nejjednodušších virů jsou pouze holé nukleokapsidy (např. pikonaviry, adenoviry a další), tedy viry neobalené. Tento typ virů bývá odolný k zevním vlivům prostředí a většinou je rezistentní vůči tukovým rozpouštědlům a éteru. Obvykle nejsou inaktivovány kyselým pH neboli vysokým potenciálem vodíku v žaludku, a proto snadno pronikají do zažívacího traktu, mohou být vylučovány stolicí, a dokonce přežívají i v odpadních a povrchových vodách. [4]

### ***1.4.2 Obalené viry***

Viry obalené obsahují kromě kapsidy navíc ještě lipoproteinový obal. Ten je tvořen lipidovou dvojvrstvou a tou se při průniku přes membrány hostitelské buňky obalí kapsida. [4] Kapsida obsahuje i bílkoviny hostitele, glykoproteiny obalu jsou určeny virem, napomáhají určit hostitelskou specifitu (jakého hostitele či jeho buňky vir infikuje).[45]

Dále se tyto obalené viry rozdělují dle lokace, kde tyto nově vytvořené nukleokapsidy vznikají, na viry které získávají obal pučením přes membránu jadernou (viry z čeledi Herpesviridae) nebo pučením při prostupu cytoplazmatickým retikulem z cisteren (viry například z čeledi Arenaviridae, Flaviridae a Coronaviridae) a pučením na cytoplazmatické membráně (viry z čeledi Orthomyxoviridae, Paramyxoviridae a další). [4]

Některé obalené viry mají na vnitřní straně lipidové dvojvrstvy obsažený protein, který se označujeme jako membránový protein neboli M protein. [4] Lipidová dvojvrstva je prostoupena glykosidickou bílkovinou, která tvoří výběžky ven z lipidu a ty jsou označovány jako peplomery [45], které mohou být zakotveny v M proteinu nebo jsou usazené v lipidové membráně jejich hydrofobními konci a jsou složeny z biologicky aktivních glykoproteinů. Jejich hlavní účel spočívá v adsorpci virionů na povrch vnímavých buněk. Na rozdíl od virů, které postrádají obal, jsou tyto obalené viry obecně labilnější a citlivé vůči fyzikálním a chemickým vlivům a jsou citlivé vůči éteru a tukovým rozpouštědlům. Nepronikají do oblasti zažívacího traktu, protože jsou inaktivovány kyselým pH. Přenáší se přímým kontaktem se zdrojem nákazy, nejčastější přenos je kapénkovou infekcí či kontaktem sliznic. [4]

### **1.5 Replikace**

Jak bylo již zmíněno, viry jsou replikace schopné pouze v živých buňkách. K tomu, aby se vir mohl množit, potřebuje tzv. permissivní buňku. Tato buňka obsahuje receptory na své membráně, které vytváří vazbu mezi virem a buněčným prostorem a také obsahuje buněčné enzymy důležité při replikaci. [8] Zabezpečení podmínek pro reprodukci virových genomů a syntézy proteinů, které jsou součástí virionů, je zajištěno pomocí replikace virů v hostitelské buňce. [4]

Strukturální a nestrukturální virové proteiny jsou vytvářeny pomocí překlada virových mediátorových RNA na ribozomech, které jsou uloženy v cytoplazmě a odtud putují do místa určení. Viry, u kterých se kompletují nukleokapsidy uvnitř jádra se přemisťují po transkripci genomu mediátorové RNA do cytoplazmy, tam jsou přeloženy na ribozomech a nové polypeptidy jsou transportovány zpět do jádra.[8] Zásadou pro pomnožení viru je průnik virového genomu do hostitelské buňky. Virová nukleová kyselina putuje do místa replikace. U DNA viru většinou probíhá replikace v jádře na rozdíl od RNA virů, kde skoro vždy probíhá replikace v cytoplazmě. [4]

Začátek replikace je po přilnutí, adsorpci na hostitelskou buňku a tento proces replikace se ukončuje po pomnožení mnohočetných replik infekčních virových partikulí. Cyklus replikace u virů má několik fází, které jsou různé dle závislosti na charakteru genomu, zda se jedná o DNA či RNA, struktury nukleové kyseliny, kde záleží, zda je dvouvláknová či jednovláknová a zda má tato jednovláknová RNA pozitivní či negativní provazec. Je mnoho způsobů, jak replikace může probíhat, ale vždy probíhá v určitých fázích, které jsou společné pro všechny viry. Tyto fáze jsou adsorpce, penetrace, odstranění obalu, aktivace genomu, syntéza viru a uvolnění z buňky. [45]

Jako první dojde k reakci glykoproteinového a bílkovinného antireceptoru, který je na povrchu virionu s receptorem na povrchu buňky. Tato reakce může za to, že se virion naváže na povrch hostitelské buňky. Při nedostatku jedné nebo obou součástí, nemůže proběhnout adsorpce, a tudíž ani infekce. Specifitu viru určuje druh hostitele a buňky. Po procesu, kdy dojde k adsorpci, může nastat změna konfigurace na povrchu virionu a antireceptory a recepční místa, která jsou na buňce v přebytečné míře. [45]

Ihned po absorpci vir proniká do buňky pomocí endocytózy, pomocí přímého průniku přes cytoplazmatickou membránu nebo fúzí. Endocytóza je děj, při kterém se virion dovnitř buňky dostává vchlípením buněčné membrány. Tyto viriony putují do endozomu (místo, které obsahuje virové partikule), kde se hromadí. Tady se pronikání může lišit dle obaleného viru a neobaleného. Vir, který nemá obal proniká přímo cytoplazmatickou membránou (translokace). Naopak u obalených virů dochází ke splynutí (fúzi) obalu s buněčnou membránou. V takovém případě obal zůstává v membráně a do buňky vstupuje jen nukleokapsida. [45]

Odstranění obalu a kapsidy je důležitý krok pro to, aby se mohl genom exprimovat, a to kvůli tomu, že vytvářejí mechanickou překážku při přístupu k němu. Od nukleové kyseliny se obal oddělí buď spontánně, nebo za pomoci buněčných enzymů. Po odstranění obalu může genom exprimovat kódované funkce, které slouží k syntéze virů. [45]

## 1.6 Mutace

U virů v genomu nejčastěji dochází k mutacím bodovým. Tyto mutace jsou zapříčiněné nepřesnou činností virových polymeráz, které připojují chybné nukleotidy během replikace virového genomu. U RNA virů, jejichž polymerázám chybí korekční (proof-reading, opravná) aktivita, jsou takové mutace mnohem významnější a u takových virů je také častější výskyt bodových mutací. U DNA virů, které jsou stabilnější jsou tyto mutace méně časté. [8]

Bodové mutace mohou vzniknout i v důsledku účinku ionizujícího záření nebo pomocí chemických vlivů. V genomu virů nemusejí vznikat jen mutace bodové, ale mohou vznikat také delece nebo inserce různého rozsahu. [8]

U všech bodových mutacích nemusí nutně vést ke změnám aminokyselinového složení kódovaných proteinů a tím dále ke změně ve fenotypu virů. Pomocí degenerovaného genetického kódu, kdy jsou aminokyseliny zakódovány několika triplety, mohou nastat tzv. synonymní mutace. Takový triplet, který kóduje příslušnou aminokyselinu, je pomocí těchto mutací nahrazován jiným tripletem, který kóduje shodnou aminokyselinou. [8]

## 1.7 Patogeneze a rozvoj infekce

Na rozdíl od bakteriálních infekcí tyto virové probíhají složitěji. Rozdíl je takový, že proliferace viru, která zapříčiní poškození organismu neprobíhá vně tkání, ale je závislá na buňkách. Virová infekce probíhá v několika stádiích a v jakékoli z nich se může zastavit. [45]

### 1.7.1 Brána vstupu

Vir ke vstupu do organismu potřebuje nalézt místo, které je pro něj vlastním a nejvýhodnějším. Existují tři cesty vstupu do organismu a umístění v orgánech. Prvním způsobem vniknutí viru do těla je přes sliznici dýchacích cest nebo přes trávicí trakt s následnou infekcí, která vzniká v místě vstupu. Tato cesta je běžná brána vstupu u myxovirů, koronavirů, rotavirů a dalších. Druhá vstupní brána je na sliznicích, kde dochází k následnému šíření virů krevním oběhem (virémie) nebo neurony k cílovému orgánu. Touto cestou do těla pronikají například viry spalniček, viry příušnic a hepatitidy A. Jako poslední cestou, je vstup po vpichu injekční stříkačky přímo do krevního oběhu nebo kousnutím hmyzu. Tímto způsobem putuje virus hepatitidy B nebo flaviviry. [45]

### ***1.7.2 Primární pomnožení***

Infekce může být nalezena již v místě vstupu. Například u rinovirů, které se pomnožují již v horních cestách dýchacích, především v nosní sliznici. Jako další příklad jsou rotaviry, které se množí již v trávicím traktu a papilomaviry v epidermis. [45]

### ***1.7.3 Sekundární pomnožení***

Tímto místem je citlivá tkáň, kam se vir dostane. Na takovém místě nastává nejzávažnější poškození napadeného hostitele, v této fázi již bývají poznatelné charakteristické příznaky, které mohou již napomáhat ke stanovení diagnózy. Poškození se může projevovat odlišně v závislosti na viru, který se do tkáně dostal. [45]

## 2. Koronaviry

### 2.1 Seznámení

První zmínka o koronavirech pochází z 60. let 20. století. Byly popsány jako původci onemocnění, které se projevovalo dýchacími a zažívacími obtížemi jak u lidí, tak i u zvířat. [49] Do roku 2002 byly popsány pouze dva lidské koronaviry (HCoV). Byl to HCoV-229E, který byl izolovaný u lidí jako první a HCoV-OC43. V roce 2002 byl identifikován Koronavirus těžkého akutního respiračního syndromu (SARS-CoV). V roce 2012 se objevil Koronavirus blízkovýchodního respiračního syndromu (MERS-CoV). Dnes již nejznámější Koronavirus těžkého akutního respiračního syndromu 2 (SARS-CoV-2), byl rozpoznán u lidí v roce 2019. [31] [27]

Koronaviry jsou zařazeny do řádu Nidovirales, který je rozdělen na 2 čeledi, na čeleď Coronaviridae (známá i jako Orthocoronavirinae) a Arteriviridae. Tyto čeledi mají společný fylogenetický původ. Do čeledě Coronaviridae zahrnujeme 2 rody, a to rod Koronavirus a Torovirus. [8] Tato čeleď zahrnuje podčeleď Coronavirinae, která je složena z rodů: Alphacoronavirus, Betacoronavirus, Gammacoronavirus a Deltacoronavirus. Rody jsou členěny dle jejich fylogenetických vztahů a struktury jejich genomů. [37] Betakoronaviry dále rozdělujeme do pěti podrodů Embecovirus, Sarbekovirus, Hibecovirus, Merbecovirus a Nobekovirus. [50] Onemocnění u lidí jsou obvykle způsobena alfa-koronaviry a beta-koronaviry. [37] Tyto rody jsou schopné infikovat pouze savce. Představují velkou hrozbu pro hospodářská zvířata, zahrnují například virus střevního průjmu prasat, koronavirus syndromu akutního průjmu prasat a virus přenosné gastroenteritidy u prasat. [11] Gamakoronaviry a deltakoronaviry jsou známé primárně pro infekci ptačích hostitelů, mohou však infikovat i savce. Způsobují ptačí koronavirovou infekční bronchitidu, která se projevuje infekčním onemocněním horních dýchacích cest u drůbeže. [26]

Dosud je známo 7 HCoV. Čtyři běžné HCoV způsobují mírnější onemocnění: HCoV-229E (název odvozen od kódu vzorku), HCoV-OC43 (z angl. Organ Culture 43, dle tracheální orgánové kultury, ze které byl získán), HCoV-NL63 (NetherLand 63, poprvé identifikován v Nizozemsku) a HCoV-HKU1 (Hong Kong University 1, podle hongkongského pacienta). Další tři HCoV, jsou známé pro těžký průběh onemocnění a jsou to SARS-CoV, MERS-CoV a nový typ koronaviru SARS-CoV-2. [24][36] Tyto tři HCoV řadíme do betakoronavirů. [37]



Koronaviry patří mezi viry, které jsou poměrně velké s velikostí v průměru 100 nm. Koronavirus je obalený vir a jeho nukleokapsida má helikoidální symetrii, kterou jsme si již výše přiblížili. [45] Genom se skládá z jedné molekuly lineární, dále z jednovláknové RNA, která má pozitivní polaritu. [8] Ze všech RNA virů jsou genomy koronavirů nejdelší. [27] [8]

Koronaviry vlastní lipidový obal, který obsahuje výběžky paličkovitě zakončené. Jsou dva druhy takových výběžků. Delší výběžky jsou určené k adsorpci viru na buňku. Kratší plní funkci M proteinu a procházejí lipidovým obalem. Koronaviry jsou citlivé k éteru a rozpouštědlům, která rozpouštějí tuky. Takové druhy, které se vyskytují ve střevech, jsou rezistentní vůči kyselému pH. [4]

Koronaviry jsou kódovány čtyřmi hlavními strukturálními proteiny, konkrétně povrchový spike protein (S), membránový (M), obalový (E) a nukleokapsidový (N) protein. S protein má funkci při vstupu virové částice do buňky. Je to jeden z proteinů koronavirů, který je schopen vyvolat imunitní odpověď hostitele. M protein je nejčastější virový protein ve virové částici a dává definitivní tvar virovému obalu. E protein pomáhá při patogenezí, hromadění a uvolňování viru. N protein je víceúčelový, například zvyšuje účinnost transkripce viru a usnadňuje práci M proteinu. [14]

## **2.2 Výskyt, šíření a klinické příznaky**

Výskyt infekcí vyvolaných koronaviry s průběhem zánětů dýchacích cest je po celém světě. Infekce je šířena pomocí kapének. Virus je množen v epitelu. Nákaza může probíhat skrytě, jindy se dostavuje rýma a zánět nosohltanu, zhruba po 3–5denní inkubaci. Ve výjimečných případech vir napadá i nižší části dýchacího traktu a vyvolává pneumonie. V séru jsou přítomny humorální protilátky. Jako obranné protilátky se uplatňují hlavně specifické imunoglobuliny třídy A. Časté jsou i reinfekce. Jsou i případy, kdy se koronaviry diagnostikovaly ve stolici u osob s akutním nebakteriálním průjmem nebo u dětí, které prodělávají akutní nekrotizující enterokolitidu.[4]

### 2.3 SARS-CoV-2

Koncem prosince 2019 ohlásily čínské zdravotní úřady vypuknutí onemocnění v průběhu pneumonie neznámého původu ve Wu-chanu. O několik dnů později byl zveřejněn genom nového koronaviru. Dříve nazýván Nový koronavirus 2019, nyní již jako SARS-CoV-2, tento virus řadíme do čeledi Coronaviridae, rodu Betacoronavirus a podrodu Sarbecovirus. Od objevu tohoto viru se šíří po celém světě, způsobil několik tisíc úmrtí a má obrovský dopad na zdravotnické systémy a ekonomiku. [10] SARS-CoV-2 je považován za jednoho ze sedmi koronavirů, které způsobují lidské infekce a patří do stejné skupiny koronavirů, které způsobují těžký akutní respirační syndrom. [14]

Genom koronavirů zahrnuje různý počet otevřených čtecích rámců (ORF). První 2/3 genomu tvoří 5' ORF (ORF1a/b), který podléhá translaci v drsném endoplazmatickém retikulu hostitelské buňky na dva proteiny fosfatázy 1. Tyto proteiny jsou dále štěpeny pomocí proteáz, a tím vzniká 16 nestrukturálních proteinů. Poslední třetinu představuje 3' ORF, který je složen z genů kódujících strukturální a doplňkové proteiny. Mezi hlavní strukturální proteiny se řadí S protein, E protein, M protein a N protein, jak je již zmíněno výše. S protein dokáže rozpoznat receptor na hostitelské buňce, který se nazývá Angiotensin-konvertující enzym 2. N protein dokáže vázat RNA a je zásadní při sestavení virionů. Nový typ koronaviru SARS-CoV-2 obsahuje navíc šest doplňkových proteinů ORF3a, ORF6, ORF7a, ORF8 a ORF10. Jejich funkce nejsou zcela prozkoumané. ORF1a a ORF1ab představují důležitost pro replikaci viru. Dále byly u SARS-CoV-2 identifikovány: 5' nepřeložená oblast (UTR) a 3'-UTR.[20]

### 2.4 Onemocnění

Tento virus způsobuje koronavirové onemocnění, které nese název COVID-19 (z angl. Coronavirus disease). S porovnáním s HCoV, které způsobovaly podobné onemocnění, má SARS-CoV-2 vyšší přenosovou kompetenci. [14]

Nejběžnější příznaky, které lidé s onemocněním COVID-19 popisují jsou horečka, kašel, dušnost, bolest hlavy a únava. Naopak mezi méně časté příznaky se řadí průjem, ztráta čichu, bolesti břicha. Jsou i spousty případů, kdy se onemocnění jeví jako asymptomatické, což je vzhledem k možnému přenosu nemoci nebezpečné. Tento virus je přenášen prostřednictvím kapének. Jeho inkubační doba se pohybuje od 3 do 14 dnů. [41]

## 2.5 Varianty SARS-CoV-2

U SARS-CoV-2 bylo zjištěno, že je náchylný k chybám, které se objevují v jeho genetickém kódu během replikace. Každý měsíc nastávají většinou 1-2 nové změny nukleotidů (mutace). Většina takových mutací byla z hlediska klinického obrazu nemoci zanedbatelná, ale objevilo se několik mutací, které byly závažnější. [21] Dnes existuje několik desítek tisíc různých variant viru, které se navzájem liší alespoň jednou mutací. [34]

Jako první na začátku pandemie byla zpozorována mutace SARS-CoV-2, která dostala název D614G. Tato mutace je lokalizována v oblasti aminokyselin na pozici 614 v proteinu spike. Viry s touto mutací byly prokázány za infekčnějším, ale neměly dopad na závažnost onemocnění. První variantou SARS-CoV-2, která nesla tuto mutaci byla varianta B.1.1.7. Další varianta SARS-CoV-2 je označovaná jako B.1.351 nebo 501Y.V2, byla odhalena v Jižní Africe. Jedna z mutací, která ji způsobila se nazývá N501Y. Další významná varianta SARS-CoV-2 je označována jako B.1.1.7 nebo 501Y.V1.[21] Do dnes bylo popsáno několik dalších variant SARS-CoV-2. Proto Světová zdravotnická organizace (z angl. World Health Organization) a Centra pro kontrolu a prevenci nemocí (z angl. Centers for Disease Control and Prevention) navrhly klasifikační systém, kde rozdělují nově vznikající varianty SARS-CoV-2 na varianty zájmu (VOI, z angl. variants of interest) a nebezpečnější varianty vzbuzující obavy (VOC, z angl. variants of concern). VOC spojujeme s vyšší přenositelností a také mohou snížit ochranu vůči vakcinaci. [1]

Mezi VOC varianty řadíme varianty Alfa (B.1.1.7), Beta (B.1.351), Gamma (P.1), Delta (B.1.617.2), Omicron (B.1.1.529). A mezi VOI řadíme varianty Epsilon (B.1.427 a B.1.429), Zeta (P.2), Eta (B.1.525), Theta (P.3), Iota (B.1.526), Kappa (B.1.617.1), Lambda (C.37) a Mu (B.1.621). [1]

### 2.5.1 Varianta B.1.1.7

Tuto varianty nazýváme také jako Britská varianta SARS-CoV-2. Dle názvu je zřejmé, že byla poprvé detekována v jihovýchodní Anglii, a to v září roku 2020. Tato varianta byla přenositelnější než předchozí varianty, a proto se rychle rozšířila do několika stovek zemí po celém světě. Nejcharaktičtější mutace, kterou tato varianta nese je mutace N501Y. [19] [12] Tato mutace je lokalizovaná v S proteinu, který má schopnost vázat receptory na lidských buňkách, a tak je infikovat virem. [34]

Tato varianta obsahuje několik dalších mutací, a to v proteinech ORF1ab, S, ORF8 a N. V genu S se nachází delece na pozicích 21764-21770 (del69-70) v genomu, které zvyšují přenositelnost viru.[19] Mutace N501Y je substituce na pozici 501 v S genu asparaginu na tyrosin a del69-70 je delece u šesti bází, které kódují histidin a valin. [51]

### **2.5.2 Varianta B.1.351**

Varianta B.1.351 je označovaná jako Jihoafrická varianta SARS-CoV-2, která obsahuje opět několik mutací v proteinu S. Jedna z mutací, která se shodují s těmi, které byly zjištěny u britské varianty, je mutace N501Y. Znepokojivé je to, že tato varianta obsahuje tři mutace, které jsou lokalizované v receptorové doméně v proteinu S. [34]

### **2.5.3 Varianta P.1**

Varianta P.1 pochází z Brazílie, kde se objevila v listopadu roku 2020 ve městě Manaus. Tato varianta obsahuje 17 mutací, v S proteinu se nachází 10 mutací. Mezi ně patří tři klíčové mutace (N501Y, K417T a E484K), které se nachází v S proteinu přesně v doméně vázající receptor. [6][16]

### **2.5.4 Varianta B.1.617.2**

Varianta B.1.617.2 známější pod názvem delta varianta byla poprvé detekována v prosinci 2020 v Indii. Tato varianta je až o 60 % přenositelnější než varianta alfa. Obsahuje 23 mutací z nichž 12 mutací v S proteinu. [44]

### **2.5.5 Varianta B.1.1.529**

Variantu B.1.1.529 známe pod názvem Omicron. Poprvé ji detekovali v listopadu roku 2021 v Jižní Africe. [38] Omicron je zatím řazen mezi variantu SARS-CoV-2, která nese největší počet mutací ze všech již výše zmíněných variant. [28] Tato varianta nese 32 mutací jen v S proteinu. Je mnohem přenositelnější a má schopnost vysoké reinfekčnosti, což znamená, že snadno onemocní i osoby, které nedávno prodělaly COVID-19 způsobený jinou variantou SARS-CoV-2.[29]

Všechny výše zmíněné varianty nesou mutaci D614G, která se vyznačuje substitucí kyseliny asparagové za glycin, a to v proteinu S na pozici 614. Tento typ mutace se vyskytuje ve více zemích. [51]

### 3. Fylogeneze

Při studiu vývoje rychle mutujících virů se uplatňuje fylogenetika. Významnou roli při pokládání základů fylogentiky zaujmul Charles Darwin, který ve své publikaci *On the origin of Species* (O původu druhů) zobrazuje existující druhy a historické souvislosti a od toho pochází představa Stromu života. [48] Další významnou osobou ve fylogenetice je Ernst Haeckel, ten se při tvoření stromů inspiroval právě u Darwina, a proto se těmto stromům říká darwinovské větvicí diagramy. [13]

Jednoduše můžeme říci, že se fylogenetika zabývá evolucí života na Zemi. Podrobněji fylogenetika studuje fylogenezi, což je tzv. analýza vývojových linií. Tyto vývojové linie jsou označovány jako taxony. Průběh fylogeneze můžeme odvodit na základě informací o kladogenezi a anagenezi a následně můžeme vytvořit fylogenetický strom, tedy popis procesu jednotlivého větvení taxonů. Kladogeneze je proces postupného odvětřování linií. Výjimečně mohou vzniknout nové linie pomocí splynutí dvou linií. Anageneze jsou změny, které nastávají v průběhu jedné linie. Takové vlastnosti se jinak nazývají jako znaky. [17]

Kladogeneze a anageneze je spíše spojována pouze se sexuálními organismy. U takových organismů chápeme druh jako množinu organismů, které si mezi sebou jsou schopné vyměňovat genetickou informaci. Definovat druhy můžeme, ale i u organismů asexuálních, a to pomocí ekologické podobnosti nebo morfologické podobnosti. Každý jedinec u asexuálních organismů je druhem, a to z toho důvodu, že je izolovanou evoluční linií. Geny, které nesou potomci asexuálních mikroorganismů vznikají mikroevolučními procesy.[2]

Při studování fylogeneze předpokládáme, že každé dva různé druhy mají společného předka, který může být i velice vzdálený. [17] Fylogenetická analýza je důležitá pro pochopení evoluce virů. [10]

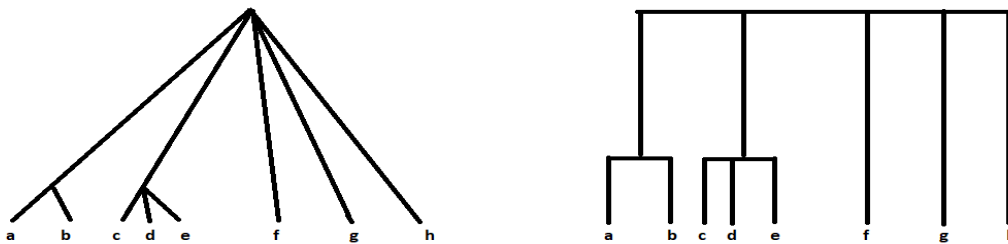
#### 3.1 Fylogenetické stromy

Stromy dělíme na zakořeněné a nezakořeněné. Zejména u zakořeněných stromů se používají k vyjádření evolučních dějin a nazýváme je jako stromy fylogenetické nebo také evoluční stromy. Většina statických modelů evoluce předpokládá, že modelové stromy jsou zakořeněné binární stromy, to znamená, že každý uzel, který není listem, má přesně dva potomky. [55]

Pokud, bychom uvažovali z pohledu matematického, je pro nás takový fylogenetický strom spíše graf (tj. jeden pár  $G = (V, E)$ . Kdy  $V$  nám udává sadu vrcholů a  $E$  je hranová množina), který je propojený nebo acyklický. Ekvivalentně je strom grafem, takže pro každou z dvojic vrcholů  $v, w$  v grafu existuje jedinečná cesta mezi nimi. Když je strom zakořeněný, označíme jeho kořen jako  $r(T)$ . Dále označujeme množinu vrcholů  $T$  jako  $V(T)$ , hrany pomocí  $E(T)$  a listy  $T$  pomocí psacího velkého  $L(T)$ . [55]

### 3.1.1 Zakořeněné stromy

V zakořeněném stromu můžeme orientovat jeho hrany ve směru kořene  $r = r(T)$ , takže všechny jiné vrcholy než  $r$ , mají jeden uzel. Pro všechny uzly platí:  $v$  není rovno  $r$ , tedy existuje jedinečný vrchol  $w$  takový, že  $v \rightarrow w$  je oblouk ve stromu (kde,  $w$  nazýváme předeek  $v$  a  $v$  nazýváme jako potomek  $w$ ). Dva nebo více vrcholů, které sdílí jednoho překda se nazývají spolu potomci. Vrchol bez listu nazýváme list nebo špička a všechny ostatní uzly nazýváme vnitřní uzly. Vrchol s více než dvěma potomky představuje polytomii. Zakořeněný strom, který nemá žádné polytomie považujeme za strom binární, bifunkční nebo plně vyřešený a pokud není binární hovoříme o stromu vícevrstvě. Okraje stromu se označují jako větve nebo interuzly. [55]



**Obrázek 1: Fylogenetické stromy 1: dva způsoby zakreslení jednoho totožného fylogenetického stromu**

Zdroj: [55]

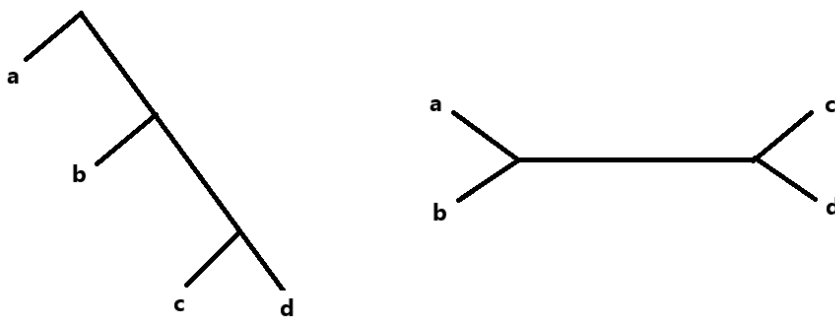
Listy na fylogenetickém stromu představují sledované taxony (obecně existující druhy nebo i různí jedinci stejného druhu) a vnitřní uzly představují předky taxonů u listů. Zakořeněný strom můžeme nakreslit s kořenem r nahoře, dole, vpravo nebo vlevo. Pokud nakreslíme strom s kořenem nahoře, tak někdy grafická znázornění stromů obsahují délky větví pro to, aby to pomohlo navrhnout relativní rychlosti změn a/nebo skutečné množství času, který uplynul. Topologii stromu znázorňuje grafický model bez větví. [55]

Existuje několik variant, jak takovou topologii zakořeněného stromu nakreslit a dva způsoby máme ukázané na obrázku 1. Tento obrázek ukazuje dva příklady reprezentace stejné evoluční historie. Strom nalevo je standartní v počítačové vědě a strom vpravo často najdeme v biologické systematice. Strom nalevo má kořen, který má 5 potomků, tudíž při interpretaci stromu vpravo bychom měli vědět, že vodorovné čáry neodpovídají hranám. [55]

### **3.1.2 Nezakořeněné stromy**

Použití nezakořeněných stromů je v počítačové vědě přednostní a má svůj význam i ve fylogenetice. Jedním z důvodů, proč je výstup nezakořeněných stromů přednostní, je ten, že většina metod pro stanovení stromů je založena na časově reverzibilních modelech evoluce sekvence, což znamená, že kořen stromu není identifikovatelný. [55]

Nezakořeněnou verzi stromu, který je zakořeněný  $T$  označujeme jako  $T_u$ . Ta je utvořena potlačením kořene a následným zpracováním stromu jako nezakořeněného. To znamená, že když má kořen  $T$  dva potomky, tak jsou přilehlé navzájem, anebo když má kořen více jak dva potomky, tak se grafický model nemění. Například pokud budeme chtít změnit zakořeněný strom na nezakořeněný  $T_u$  viz obrázek 2, tak budeme ignorovat rozložení kořenového stromu a vytvoříme nezakořeněný strom. Každý zakořeněný strom má totiž jedinečnou nezakořeněnou verzi stromu. [55]



**Obrázek 2: Fylogenetické stromy 2:** Tento obrázek znázorňuje na levé straně strom zakořeněný a na pravé strom nezakořeněný.

Zdroj: [55]

### 3.2 Význam mutací pro fylogenetiku

Tak jako fylogeneze souvisí s biologickou evolucí, tak pro ni mají zásadní význam i mutace. Mutace se projevují jako změny ve struktuře genetického materiálu a při takových změnách dochází ke změně smyslu genetické informace. Můžeme říct, že ne všechny mutace musejí být škodlivé, bez některých nových mutací by se biologická evoluce mohla zastavit. [17]



## 4. Molekulární detekce

Molekulární detekce u virů se používají zejména pro průkaz virového genomu. Molekulární detekce má tu výhodu, že lze uskutečnit diagnostiku i u inaktivovaného viru, jinak je vysoce citlivá a specifická. Řadíme sem metody, které prokazují nukleové kyseliny, a tudíž je potřeba znát nukleotidové sekvence příslušného virového genomu. [8]

Nejčastěji se používají metody amplifikační. To jsou metody, které fungují na principu enzymatického pomnožení části nukleové kyseliny viru a poté její zvýraznění. Jednou z nejvýznamnějších metod je polymerázová řetězová reakce (PCR, z angl. Polymerase Chain Reaction). [8]

### 4.1 PCR

Dnes jsou známy téměř kompletní nukleotidové sekvence několika genomů, a to i lidského. Sekvence jsou dostupné v mnoha databázích, jako je například GenBank (veřejně dostupná databáze nukleotidových sekvencí), proto je možné izolovat geny nebo vybrané sekvence DNA bez toho, aby byly použity klonovací vektory nebo hostitelské buňky. Amplifikace DNA sekvencí se výhradně provádí *in vitro*. Sekvence mohou být amplifikovány milionkrát, ale i vícekrát během několika málo hodin. K provedení tohoto postupu stačí znát krátké nukleotidové sekvence, které ohraničují danou zkoumanou sekvenci. Všechno, co jsme si zde popsali, lze provést pomocí PCR. K PCR reakci jsou nutné oligonukleotidy, které jsou komplementární ke známým sekvencím, které ohraničují určitou oblast k tomu, aby mohla ve zkumavce začít enzymatická amplifikace části DNA těchto sekvencí.[47]

#### 4.1.1 Historie metody

Mullis, Faloon a Saiki jsou zodpovědní za vyvinutí vysoce účinné enzymové metody v letech 1983–1985, pomocí které se *in vitro* amplifikují vybrané sekvence nukleových kyselin. Tato metoda byla nazvaná jako PCR. [56] Kary Mullis za svou práci o PCR v roce 1993 obdržel Nobelovu cenu za chemii. [47]

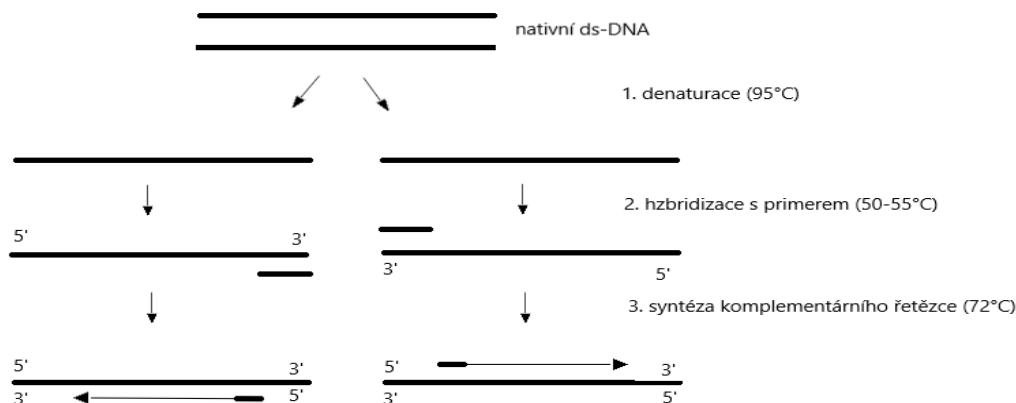
### 4.1.2 Princip

PCR funguje na principu schopnosti dvouvláknové DNA se při vysoké teplotě denaturovat a poté po snížení teploty se zpět renaturovat, a to všechno za zachování pravidla komplementarity bází. PCR reakce je zahájena denaturací vzorku DNA na dva jednoduché řetězce při vysoké teplotě 95°C. V další části reakce dochází k ochlazení vzorku na 50-60 °C, což vede k připojení primerů na komplementární 3' konce cílových DNA. Primery jsou v této reakci přítomny ve větších počtech než dlouhé řetězce analyzované DNA, které jsou v této části PCR odděleny. [3]

Základ pro syntézu nových vláken DNA jsou hybridizované primery. V PCR reakci je dostatečné množství deoxynukleotidtrifosfátů, kvůli dostatečnému množství substrátu pro syntézu nových vláken. Tato syntéza je katalyzována pomocí termostabilní DNA polymerázy. Často se využívá DNA polymeráza, která je izolovaná z bakterie *Thermus aquaticus*, ta se vyskytuje v horkých pramenech a označujeme ji jako Taq polymerázu. [3]

Tento enzym slouží k prodlužování vláken DNA směrem od obou primerů, a to ve směru od 5' konce po konec 3' při teplotě 72 °C. Po ukončení syntézy u obou vláken, je zkumavka opět zahřátá na 95 °C a opět dochází k denaturaci nově vytvořených DNA duplexů a celý proces se začne opakovat. Toto opakování vede k rychlému namnožení cílových sekvencí DNA. Po každém následujícím opakování se počty kopií úseků mezi připojenými primery zdvojnásobí. [3] Schéma reakce PCR je znázorněno na obrázku 3.

Počet kopií určené sekvence DNA roste logaritmicky (přibližně okolo 106 x po 20 cyklech), přičemž se ostatní úseky výchozí DNA již neamplifikují. Při PCR reakci je důležité pro její průběh zajistit vzorek s DNA obsahující zkoumaný úsek, primery neboli syntetické oligonukleotidy (má úlohu jako základ pro syntézu nových vláken), směs všech 4 deoxynukleotidtrifosfátů, hořčnaté ionty, Taq polymerázu a PCR pufr. Reakční objem bývá okolo 20 a 100 µl. [3]



**Obrázek 3:** Schéma reakce PCR: Zkratka ds-DNA znamená dvouvláknová DNA.

Zdroj: [3]

## 4.2 RT-PCR

Jedna z metod pro detekci koronaviřů a dnes hlavně SARS-CoV-2 je nejčastěji používaná metoda reverzní transkriptázová polymerázová řetězová reakce v reálném čase (RT-PCR). Tento test je nejčastěji prováděn z výtěrů z nosofaryngu. [43] Pomocí RT-PCR je tedy detekovatelný SARS-CoV-2 a to pomocí ORF1ab, který zahrnuje RNA-dependentní RNA polymerázu a proteinů E, N a S. RT-PCR je tedy metoda, která je založena na metodě PCR a doplňuje jí přesným monitorováním produktů v reálném čase.[52]

### 4.2.1 Princip RT-PCR

Metoda RT-PCR je založena na rozlišení a měření určitých sekvencí nukleových kyselin ve vzorku, a to i v malém množství. Při RT-PCR dochází k amplifikaci cílové sekvence, ta se následně monitoruje pomocí fluorescenčního značení. Konečný produkt je charakterizován pomocí zvyšování teplot, a to pro to, aby se dalo určit, kdy se tento produkt složen ze dvou vláken rozdvojí. [42]

## **5. Cíl práce a výzkumné otázky**

### **5.1 Cíl práce**

Cílem této práce bylo stanovení fylogenetických souvislostí na základě publikovaných sekvencí koronavirů, a to zejména mutačních variant nového typu koronaviru SARS-CoV-2 a na základě těchto fylogenetických souvislostí zkonstruovat fylogenetický strom. Výzkumnou otázkou je, které z koronavirů jsou nejvíce příbuzné SARS-CoV-2.

Dalším cílem bylo porovnání a posouzení používaných PCR primerů pro rozlišení jednotlivých variant SARS-CoV-2 a určit primery, které reagují pouze na nově vzniklý virus SARS-CoV-2, ale nereagují na ostatní typy koronavirů.

## **6. Metodika a metody**

Fylogenetické souvislosti byly popsány na základě sestavení fylogenetických stromů pomocí programu MEGA-X. Na základě fylogenetických analýz byly zkoumány procentuální podobnosti mezi koronaviry. Primery byly navrženy pomocí webového nástroje Primer-BLAST a byly porovnány s již používanými primery pro detekci SARS-CoV-2.

### **6.1 Fylogenetická analýza**

Cílem fylogenetické analýzy je rekonstrukce stromu, podle kterého lze popsat evoluční vztahy mezi organismy, které jsou v tu chvíli studovány. Genomy podléhají hromadění mutací, kvůli kterým se dále vyvíjejí. Tyto změny v nukleotidových sekvencích mezi dvěma genomy by měly značit, jak nedávno sdílely společného předka. Dá se očekávat, že dva genomy, které se rozcházejí v nedávné době, budou nést méně mutací než dva genomy, které mají staršího společného předka. Dle toho při porovnání tří a více genomů by mělo být možné odhalit evoluční vztahy mezi nimi. [7]

### **6.2 Materiál**

Ve fylogenetice se dnes již převážně využívají molekulární data (proteinové nebo nukleotidové sekvence). Nukleotidové sekvence DNA uvádí více informací pro fylogenetiku než sekvence proteinové, protože v proteinových sekvencích se ukazují pouze ty mutace, které změnilý smysl aminokyseliny, přičemž v nukleotidových sekvencích jsou ukázány všechny změny. Z toho důvodu byly pro tuto fylogenetickou analýzu koronavirů použity sekvence nukleotidové.[7]

Sekvence pro fylogenetickou analýzu byly získány z GenBank pomocí webového nástroje NCBI Virus na stránkách Národního centra pro biotechnologické informace (NCBI) na internetovém odkaze: <https://www.ncbi.nlm.nih.gov/labs/virus/vssi/#/>. Sekvence byly vybírány na základě 4 rodů koronavirů (alfa, beta, gamma a delta). Byly vybírány sekvence lidských i zvířecích koronavirů. Podmínkou pro výběr bylo, aby měly sekvence kompletní genom. Jako hlavní sekvence sloužila referenční sekvence nového typu koronaviru SARS-CoV-2 s přístupovým kódem NC\_045512. Sekvence byly staženy ve formátu FASTA (formát souboru sekvencí).

Pro rekonstrukci prvních dvou fylogenetických stromů byly vybrány sekvence pro rod alfa, beta, gamma a delta. Pro první strom byly použity sekvence celých genomů a pro druhý strom z těchto sekvencí byly použity úseky sekvencí proteinu S.

Pro rod alfa byly vybrány sekvence koronavirů pod těmito přístupovými čísly: NC\_005831.2, LC687294.1, NC\_032107.1, NC\_048216.1, NC\_028752.1, NC\_002645.1, MW202340.1, OL415262.1, OL410607.1, NC\_046964, MH687938.1, MW560716.1, MH726400.1, MT906864.1, NC\_048211.1, NC\_034972.1, NC\_032730.1, KF294380.2.

Jako zástupci rodu beta byly vybrány následující sekvence koronavirů pod těmito přístupovými čísly: AY597011.2, NC\_026011.1, NC\_017083.1, NC\_006213.1, KU886219.1, KT368891.1, NC\_009021.1, EF065514.1, NC\_025217.1, KY417143.1, KC881006.1, NC\_004718.3, MG772934.1, MT121216.1, MN996532.2, MW729043.1, NC\_045512.2, MT371574.1, MN611520.1, NC\_038294.1, NC\_019843.3, MN611519.1.

Dále byly vybrány sekvence pro rod gamma s přístupovými čísly: NC\_048213.1, MT984583.1, NC\_048214.1, MN690608.1, NC\_010648.1, EU111742.1.

Pro rod delta byly vybrány sekvence s přístupovými čísly: NC\_011547.1, NC\_039208.1, MG812375.1 a OM256446.1. Tato poslední sekvence s přístupovým číslem OM256446.1 byla použita pouze ve stromu, který srovnává podobnost koronavirů na základě proteinu S. Všechny sekvence byly staženy z toho internetového odkazu: [https://www.ncbi.nlm.nih.gov/labs/virus/vssi/#/virus?SeqType\\_s=Nucleotide](https://www.ncbi.nlm.nih.gov/labs/virus/vssi/#/virus?SeqType_s=Nucleotide).

Třetí fylogenetický strom se zaměřoval na hledání podobností mezi koronaviry rodu beta a jeho podrodů Embecovirus, Merbecovirus, Nobecovirus, Hibecovirus a Sarbecovirus.

Pro podrod Embecovirus byly vybrány tyto sekvence pod přístupovými čísly: KU886219.1, FJ938064.1, JX860640.1, NC\_006213.1, NC\_017083.1, JN874560.1, KT368891.1, KF906249.1, LC592689.1, LC061272.1, NC\_025011.1, KM349743.1, NC\_001846.1, JX169866.1, AC\_000192.1, AY597011.2, OM725430.1,

dále sekvence zastupující podrod Merbecovirus: MN611520.1, MN611519.1, MK564475.1, NC\_019843.3, NC\_038294.1, OL622035.1, MH029552.1,

sekvence zastupující podrod Hibecovirus: NC\_025217.1, KF636752.1.

a sekvence zastupující podrod Nobecovirus: EF065514.1, NC\_009021.1, MG762674.1.

Nakonec byly vybrány sekvence zastupující podrod Sarbekovirus: OM725430.1, NC\_014470.1, AY427439.1, AY274119.3, NC\_004718.3, KY417143.1, KJ473815.1, DQ071615.1, MW251308.1, MG772934.1, MZ376737.1, MW729043.1, OM846619.1, ON004292.1, MT534285.1, NC\_045512.2, MZ264787.1 a MT371574.1. Všechny sekvence byly staženy z toho internetového odkazu: [https://www.ncbi.nlm.nih.gov/labs/virus/vssi/#/virus?SeqType\\_s=Nucleotide](https://www.ncbi.nlm.nih.gov/labs/virus/vssi/#/virus?SeqType_s=Nucleotide).

### **6.2.1 GenBank**

GenBank je veřejně dostupná databáze nukleotidových sekvencí. Do databáze jsou sekvence získávány pomocí zadávání z jednotlivých laboratoří a rozsáhlých projektů sekvenování. Po obdržení dat zaměstnanci přidělují jednotlivým sekvencím přístupová čísla. Pokrytí po celém světě je zajištěno pomocí každodenního vyměňování dat s European Nucleotide Archive a DNA Data Bank of Japan. Sekvence GenBank jsou zveřejněna prostřednictvím databáze Národního centra pro biotechnologické informace (NCBI). V této databázi jsou odkazy na další informace (taxonomie, genomy, proteinové sekvence a jejich struktury a literatura z biomedicínských časopisů). [5]

### **6.3 Metodika rekonstrukce fylogenetických stromů**

Jak již bylo popsáno výše, tak cílem fylogenetické analýzy je rekonstrukce stromu, který znázorňuje evoluční vztahy mezi zkoumanými organismy. Základním krokem před rekonstrukcí fylogenetického stromu je zarovnání studovaných nukleotidových sekvencí a získání srovnávacích dat, která jsou dále převedena do rekonstruovaného fylogenetického stromu.[32]

### **6.4 MEGA-X**

Zvolený počítačový program pro rekonstrukci fylogenetických stromů byl program pro analýzu molekulární evoluční genetiky neboli MEGA (z aj. Molecular Evolutionary Genetics Analysis). Do tohoto programu byly vloženy studované sekvence ve formátu FASTA, nejprve se pracovalo v módu Align viz příloha 1.

MEGA-X nabízí několik analytických metod a nástrojů pro určení fylogenetiky. Lze jej bezplatně stáhnout na webových stránkách: [www.megasoftware.net](http://www.megasoftware.net). Tento program byl citován v evoluční analýze více než ve stovce studií zabývajících se fylogenetickou analýzou. [32] MEGA obsahuje dvě metody, které jsou určeny k zarovnání sekvencí, jedná se o metody ClustalW a Muscle. Dále obsahuje metody pro sestavení fylogenetických stromů, jedná se o Neighbor Joining, UPGMA, Maximum Parsimony, Bayesian Inference a Maximum Likelihood. [23]

## **6.5 Metody**

### **6.5.1 ClustalW**

Sekvence byly zarovnány pomocí metody ClustalW, která jak bylo již zmíněno je součástí programu MEGA-X. Je to metoda, která je určena k zarovnání více nukleotidových nebo proteinových sekvencí. Zarovnání probíhá třemi kroky, nejprve se jedná o párové zarovnání, následně se generuje vodící strom a na závěr jde o postupné zarovnání. [35] Metoda je vhodná pro srovnání hodně dlouhých sekvencí. [15] Zarovnané sekvence byly uloženy jako nový soubor, který mohl být dále použit při rekonstrukci fylogenetického stromu, tento proces byl opakován pro další dva stromy.

### **6.5.2 Neighbor-Joining**

Zarovnané a uložené sekvence byly opět vloženy do programu MEGA-X, kde se dále pokračovalo v módu Analyze viz příloha 1. Pro rekonstrukci fylogenetických stromů byla zvolena metoda Neighbor-Joining. Metoda byla nastavena na 1000 replikací bootstrapu (opakování). Dále byl použit model Kimura 2-parameter, tento model je založen na rozlišování tranzic a tranzverzí. [30] Tranzice jsou záměny purinového nukleotidu za purinový nebo pyrimidinového nukleotidu za pyrimidinový a u tranzverzí, dochází k záměně purinového nukleotidu za pyrimidinový a naopak. Mezi purinový nukleotid je zařazen thymin (T) a cytosin (C) a pyrimidinový nukleotid adenin (A) a guanin (G). [47] Další nastavení bylo zanecháno jako výchozí, viz příloha 2.

Neighbor-Joining metoda je založena na metodách distanční matice. Tyto metody jsou závislé na míře „genetické vzdálenosti“ mezi studovanými sekvencemi. Počet neshod mezi zarovnanými sekvencemi, definují tuto vzdálenost. Další metoda, která využívá metody distanční matice je metoda neváženého párování s aritmetickým průměrem neboli UPGMA (z angl. unweighted pair group method with arithmetic mean).[22]



Neighbor-Joining konstruuje fylogenetický strom na základě matice párových evolučních vzdáleností, které se týkají souboru studovaných taxonů. Vzdálenost mezi taxonomickým párem se získá ze sekvenčních dat například pomocí modelu Kimura 2-parametr.[18] Tímto vznikne hvězdicový strom, na kterém lze pozorovat jeden vnitřní vrchol, ze kterého vedou listy pomocí, kterých jsou reprezentovány studované taxonomické jednotky. Strom se následně rozkládá pomocí shlukování nejbližších taxonomických jednotek a postupně se v každém kroku celková délka stromu zmenšuje. [22]

Mezi další metody rekonstrukce fylogenetických stromů se řadí maximum parsimony, maximum likelihood, molekulární hodiny, metoda nejmenších čtverců a minimální evoluce.[22]

## **6.6 BLAST**

Pro vypočítání shody mezi sekvencemi celých genomů a následně úseků těchto sekvencí byl použit software BLAST (algoritmus pro porovnání sekvencí z angl. Basic Local Alignment Search Tool) skrze webové stránky NCBI na internetovém odkaze: <https://blast.ncbi.nlm.nih.gov/Blast.cgi>. Byla použita funkce: Globální zarovnání, která využívá Needlemanův-Wunschův algoritmus viz. příloha 3. Sekvence lze do programu zadat přístupovým číslem (při porovnání celých genomů), nebo ručně zkopírováním a vložením určitého úseku či celé sekvence ve formátu FASTA. Lze také vložit soubor sekvencí.

Pro vypočítání podobnosti sekvencí byla jako referenční sekvence použita sekvence nového typu koronaviru SARS-CoV-2, která byla porovnávána s nejbližšími sekvencemi koronavirů, na základě fylogenetických analýz.

Algoritmus Needlemanův-Wunschův se používá k zarovnání proteinových i nukleotidových sekvencí v bioinformatice. Vyvinuli ho a publikovali v roce 1970 Saul B. Needleman a Christian D. Wunsch. [39] Jak je již zmíněno výše, tento algoritmus provádí globální zarovnání sekvencí, kdy je brán ohled na celou délku sekvence. Principem tohoto algoritmu je tabulka, kde v horním řádku je zapsaná sekvence jednotlivými bázemi o délce  $m$  a v levém sloupci je zapsaná druhá sekvence, jejíž délka je značena jako  $n$ . Velikost tabulky je definována jako  $((m+1)(n+1))$ . Algoritmus každému zarovnání přiděluje skóre, přičemž účelem je vyhledat všechny uspořádání, která mají nejvyšší skóre neboli maximum. Maximum bude buď hodnota skóre, která přísluší poli vlevo + hodnota sankce za použití mezery (inzerce) nebo hodnota skóre, která přísluší poli nahoře + hodnota sankce za použití mezery (delece) nebo hodnota skóre, která přísluší poli vlevo nahoře + hodnota sankce shody/neshody znaků, které jsou na pozici dané sekvence, dle skórovací matice. Výsledná cesta průchodu, který je zapisován šipkami, reprezentuje optimální zarovnání v matici.[40]

## 6.7 Primery

Primery jsou krátké jednovláknové řetězce nukleové kyseliny neboli oligonukleotidy, které jsou obvykle složeny z 20 až 25 bází. Tyto úseky jsou komplementární k cílovému místu na nukleotidové sekvenci. Pro správný výsledek PCR metody je klíčové, aby byly primery navrhnuté správně. Primery by měly obsahovat srovnatelný počet GC:AT párů. [33]

## 6.8 Běžně používané PCR primery pro detekci SARS-CoV-2

Používané PCR primery pro detekci SARS-CoV-2 nejsou běžně zveřejňovány v laboratorních kitech, které se pro tuto detekci používají. Existuje však webová softwarová aplikace CoVrimer, která je volně dostupná na webové adrese: <http://konulabapps.bilkent.edu.tr:3838/CoVrimer/>. Pomocí tohoto webového serveru jsou uživatelé schopni vyhledávat a porovnávat existující zveřejněné sekvence párů primerů. Primery, které jsou dostupné na tomto webovém serveru, byly využívány spíše pro experimentální účely, které jsou popsány ve vědeckých článcích. [54]

Softwarová aplikace CoVrimer nabízí pod složkou zvanou Align Published Primer Sets 97 publikovaných párů primerů pro detekci SARS-CoV-2. Primery byly vybírány takové, aby jejich cílovým genem byl gen S. Těchto párů primerů aplikace nabízí 23.

V programu MEGA-X byl otevřen soubor, který byl sestaven pomocí použitých sekvencí HCoV pro tvorbu fylogenetických stromů, jednalo se o 22 sekvencí, které zastupovaly lidské koronaviry viz tabulka 1.

*Tabulka 1: Počet použitých sekvencí*

Název lidské koronaviru	Počet použitých sekvencí
SARS-CoV-2	10
MERS-CoV	4
SARS-CoV	3
HCoV OC43	1
HCoV NL63	1
HCoV HKU1	2
HCoV 229E	1

Sekvence primerů forward a reverse byly vkládány do políčka s lupou, které má úlohu pro vyhledávání úseku vložené sekvence ve všech zastoupených sekvencích v souboru. Na tomto základě bylo vybráno 10 párů primerů, které nasedaly na všech 10 sekvencích SARS-CoV-2, nenedaly na ostatní použité sekvence HCoV a ohraničovaly charakteristické mutace příslušných variant SARS-CoV-2. Tyto primery jsou zapsány v tabulce 2.

*Tabulka 2: Vybrané primery*

Pár primerů	Směr primeru	Sekvence primerů 5'→3'
1.	Forward	CCTACTAAATTAATGATCTCTGCTTTACT
	Revers	CAAGCTATAACGCAGCCTGTA
2.	Forward	TTGGCAAATTC AAGACTCACTTT
	Revers	TGTGGTTCATAAAAATTCCTTTGTG
3.	Forward	TCAAGACTCACTTTCTTCCAC
	Revers	ATTTGAAACAAAGACACCTTCAC
4.	Forward	AAGACTCACTTTCTTCCACAG
	Revers	CAAAGACACCTTCACGAGG
5.	Forward	CCTACTAAATTAATGATCTCTGCTTTACT
	Revers	CAAGCTATAACGCAGCCTGTA
6.	Forward	CAATGGTTTAACAGGCACAGG
	Revers	CTCAAGTGTCTGTGGATCACG
7.	Forward	GGTTTTAATTGTTACTTTCCTTTACAA
	Revers	GAAAGTACTACTACTCTGTATGGTTGG
8.	Forward	TCCTTTACAATCATATGGTTTCCA
	Revers	GCATGTAGAAGTTCAAAGAAAGT
9.	Forward	CTGAAATCTATCAGGCCGGTAG
	Revers	CTACTCTGTATGGTTGGTAACC
10.	Forward	GTTACTTTCCTTTACAATCATATG
	Revers	TTTAGGTCCACAAACAGTTGC

Dále byly primery vloženy do speciálního softwaru Primer-BLAST, který je volně přístupný na stránkách NCBI (<https://www.ncbi.nlm.nih.gov/tools/primer-blast>) a byly zkoumány jejich vlastnosti. Kde byla jako šablona PCR použita sekvence SARS-CoV-2 (NC\_045512) a do parametrů primerů byly postupně zadávány jednotlivé forward a revers primery a na závěr byla vybrána kontrola specifičnosti, viz příloha 4. Na základě specifičnosti a na základě cílové oblasti byly vybrány dva páry primerů pro srovnání s následným návrhem primerů.

## 6.9 Návrh primerů

Primery byly navrhovány, tak aby cílily na daný úsek sekvence pomocí speciálního softwaru Primer-BLAST, který byl použit při posouzení vhodnosti primerů, které jsou dostupné v softwarové aplikaci CoVprimer.

Cílová oblast byla vybrána na základě vytvoření alignmentu v programu BioEdit, který zarovnával pět celogenomových sekvencí pomocí funkce ClustalW. Jednalo se o referenční sekvenci SARS-CoV-2 (NC\_045512), dále o sekvenci VOC alfa (MZ376737), beta (OM725430), gamma (MZ264787) a delta (OM846619). U těchto VOC byly zkoumány změny v bázích v proteinu S, které jsou charakteristické pro studované VOC SARS-CoV-2. Pozice bází u referenční sekvence SARS-CoV-2, které podlely změně u studovaných sekvencí, byly vloženy do speciální funkce Mutace v archivu čtení sekvence dat (z angl. Mutations in Sequence Read Archive), kterou nabízí webová stránka NCBI Virus ([https://www.ncbi.nlm.nih.gov/labs/virus/vssi/#/scov2\\_snp](https://www.ncbi.nlm.nih.gov/labs/virus/vssi/#/scov2_snp)). Tato funkce, po zadání pozice, na které byla změna, vyhodnotí, o jakou mutaci se jedná (synonymní, nesynonymní), popíše změnu v tripletu a uvede název mutace. Na tomto základě byly vybrány tři cílové mutace L452R, E484K a A570D viz tabulka 3.

*Tabulka 3: Cílové mutace*

Název mutace	Změna v tripletu	Povaha mutace
L452R	CTG → CGG	NESYNONIMNÍ
E484K	GAA → AAA	NESYNONYMNÍ
A570D	GCT → GAT	NESYNONYMNÍ

Mutace L452R podléhá změně aminokyselin, kdy se mění leucin (CTG) za arginin (CGG), u mutace E484K dochází k záměně kyseliny glutamové (GAA) za lysin (AAA) a u mutace A570D se mění alanin (GCT) za asparagin (GAT).

Tyto tři mutace byly nalezeny v alignmentu, kde mutaci L452R (na pozici 22918) nese VOC delta, mutaci E484K (na pozici 23013) nesou VOC beta a gamma a mutaci A570D (na pozici 23272) nese VOC alfa. Jedná se o mutace bodové tzv. substituce, kde dochází ke záměně jedné či více po sobě jdoucích párů bází za jiné. V takovém případě může dojít ke změně aminokyseliny, čím se mohou měnit vlastnosti vzniklé bílkoviny, takové změny se nazývají nesynonymní.[25] Substituce můžeme rozdělit na dvě skupiny, jednu nazýváme jako tranzici a druhou jako tranzverzi.[47]

Navržený pár primerů je vložen do následující tabulky 4, celková délka produktu je 512 párů bází, forward i revers primer mají délku 20 párů bází.

*Tabulka 4: Navržený pár primerů*

Směr primeru	Sekvence primerů 5' → 3'
Forward	AAATCGCTCCAGGGCAAAC
Revers	TGGATCACGGACAGCATCAG

## 7. Výsledky

Praktická část byla jako první zaměřena na fylogenetickou analýzu koronavirů, která spočívala v rekonstrukci tří fylogenetických stromů pomocí počítačového softwaru MEGA-X, jak bylo již zmíněno výše. Tato práce zahrnovala také porovnání částí sekvencí koronavirů. Dále byla praktická část zaměřena na porovnání primerů, které byly již použity pro detekci SARS-CoV-2 s našimi návrhy primerů.

### 7.1 Fylogenetická analýza

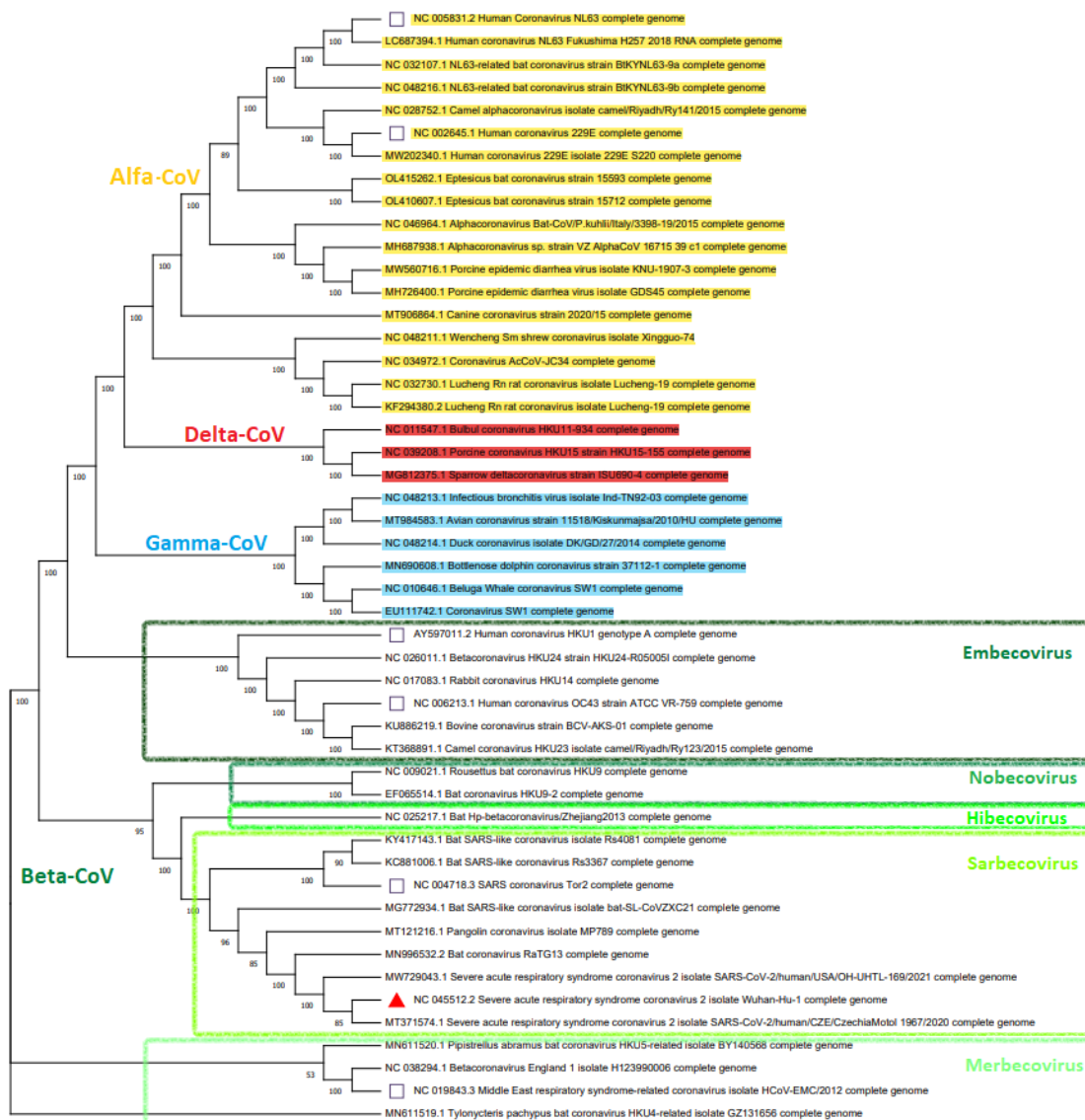
Fylogenetická analýza prvních dvou stromů byla provedena pro všechny čtyři rody koronavirů (alfa, beta, gamma a delta). Jako první byla provedena fylogenetická analýza zaměřena na celé genomy studovaných koronavirů.

Výsledný fylogenetický strom viz obrázek 4 znázorňuje podobnost mezi rody alfa, beta, gamma a delta koronavirů. Přičemž u rodu beta koronavirů je porovnávána podobnost mezi jeho podrody Embecovirus, Nobecovirus, Hibecovirus, Sarbecovirus a Merbecovirus.

Referenční sekvence nového typu koronaviru SARS-CoV-2 s přístupovým číslem NC\_045512 je v tomto fylogenetickém stromu zařazena do rodu betakoronavirů a jeho podrodu Sarbecovirus. Dva nejpodobnější genomy k NC\_045512 jsou izoláty MW729043 a MT371574, protože to jsou izoláty SARS-CoV-2 nejsou pro další analýzu brány v potaz. Dalšími nejbližšími izoláty jsou netopýří koronavirus RaTG13 (MN996532.2), luskouní koronavirus MP789 (MT121216.1) a netopýří koronavirus podobný SARS ZXC21 (MG772934).

Při porovnání podobnosti celých genomů se nejvíce se SARS-CoV-2 podobal netopýří koronavirus RaTG13, a to s procentuální shodou 96 %, 89% shoda byla u luskouního koronaviru MP789 a 88% shoda u netopýřího koronaviru podobnému SARS ZXC21. SARS-CoV (NC\_004718.3) nese pouze 80% shodu, viz tabulka 5.

Dále lze na fylogenetickém stromu pozorovat zařazení dalších lidských koronavirů. MERS (NC\_019843), je součástí rodu betakoronavirů podrodu Merbecovirus a jeho procentuální shoda se SARS-CoV-2 činí 57 % viz tabulka 8. Jako další lidské koronaviry, které jsou zařazeny do rodu beta jsou HCoV-OC43 (NC\_006213) a HCoV-HKU1 (AY597011), oba jsou zařazeny do podrodu Embocovirus a jejich procentuální shoda se SARS-CoV-2 je v obou případech 56 % viz tabulka 9.



**Obrázek 4: Fylogenetický strom pro čeled' Coronavirinae za použití celogenomových sekvencí. Čísla zapsaná u uzlů stromu, značí procentuální shodu bootstappingu. Taxony jsou značeny přístupovými čísly a názvem koronaviru. Žlutou barvou jsou vyznačeny alfa CoV (koronaviry), červenou barvou delta CoV, modrou gamma koronaviry a zelenou barvou beta koronaviry, které jsou rozděleny do zelených obdelníků, dle jejich podrodu Embecovirus, Nobecovirus, Hibecovirus, Sarbecovirus a Merbecovirus. Červeným trojúhelníkem je značena referenční sekvence SARS-CoV-2 (NC\_045512.2), fialovými čtverci vyplněnými bílou barvou je značeno dalších 6 lidských koronavirů.**

Zdroj: vlastní



Fylogenetická analýza ukazuje, že SARS-CoV-2, SARS-CoV, MERS-CoV a HCoV-NL63 jsou blízce související s koronaviry netopýřími, naopak HCoV-HKU1 je blízce příbuzný s králičím koronavirem, HCoV-OC43 s bovinním a velbloudím koronavirem a HCoV-229E také blízce souvisí velbloudím koronavirem viz obrázek 4.

Následující tabulky 5, 6, 7, 8, 9 ukazují totožnost celogenomových sekvencí studovaných koronavirů v rodě beta s referenční sekvencí SARS-CoV-2. V tabulkách je uváděn celonárodní název koronavirů.

**Tabulka 5: Procentuální shoda celých genomů podrodu Sarbecovirus**

Přístupové číslo	Název koronaviru	Totožnost s NC_045512.2 (%)
MW729043	SARS-CoV-2	99
MT371574	SARS-CoV-2	99
MN996532	Bat CoV RaTG13	96
MT121216	Pangolin CoV isolate MP789	89
MG772934	Bat SARS-like CoV isolate bat-SL-CoVZXC21	88
NC_004718	SARS CoV Tor2	80
KC881006	Bat-SARS-like CoV Rs3367	80
KY417143	Bat-SARS-like CoV Rs4081	80

**Tabulka 6: Procentuální shoda celých genomů s podrodem Hibecovirus**

Přístupové číslo	Název koronaviru	Totožnost s NC_045512.2 (%)
NC_025217	Bat Hp-betaCoV	58

**Tabulka 7: Procentuální shoda celých genomů s podrodem Nobecovirus**

Přístupové číslo	Název koronaviru	Totožnost s NC_045512.2 (%)
NC_009021	Rousettus bat CoV HKU9	57
EF055514	Bat CoV HKU9-2	57

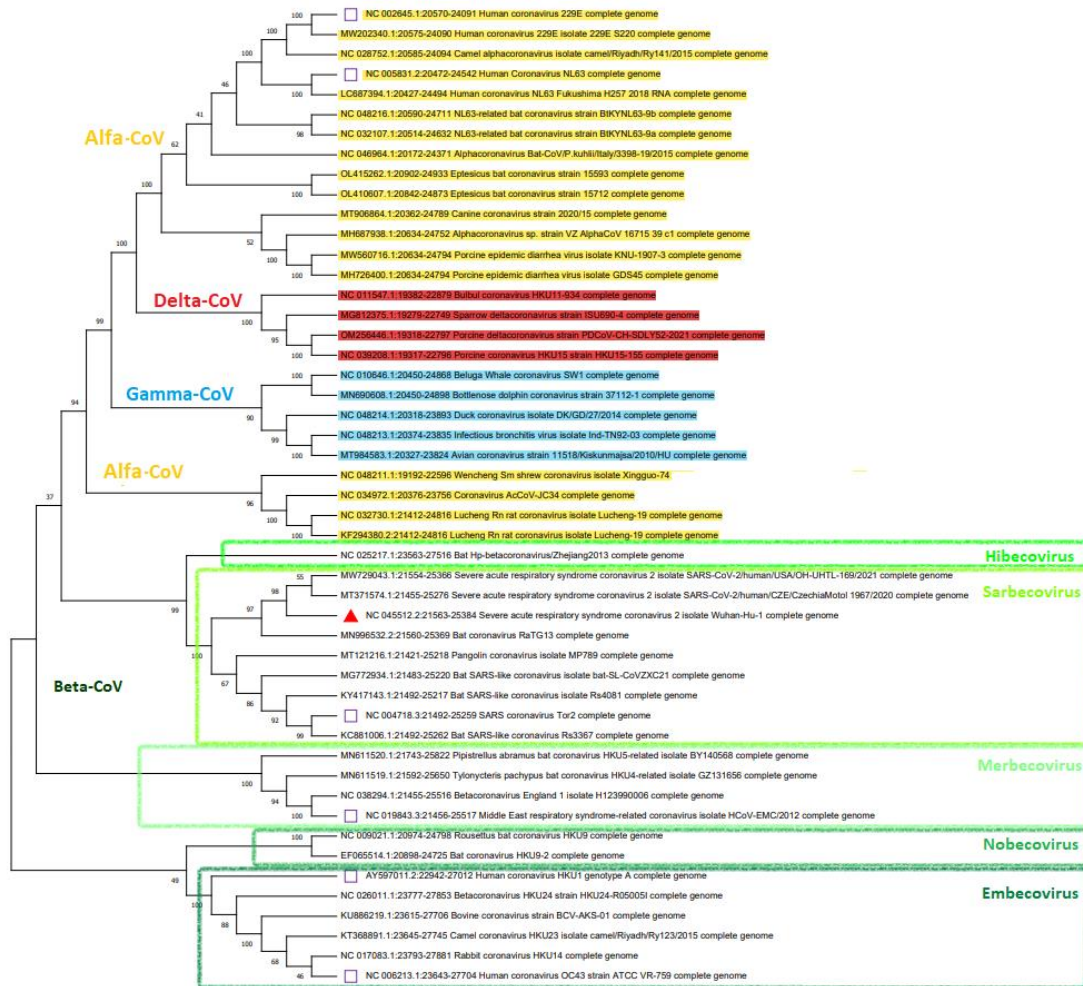
**Tabulka 8: Procentuální shoda celých genomů s podrodem Merbecovirus**

Přístupové číslo	Název koronaviru	Totožnost s NC_045512.2 (%)
MN611520	Pipistrellus bat CoV HKU5	57
NC_038294	BetaCoV England 1	57
NC_019843	MERS-CoV	57
MN611519	Bat CoV HKU4	57

**Tabulka 9: Procentuální shoda celých genomů s podrodem *Embecovirus***

Přístupové číslo	Název koronaviru	Totožnost s NC_045512.2 (%)
AY597011	HCoV HKU1	56
NC_026011	BetaCoV HKU24	55
NC_017083	Rabbit CoV HKU14	56
NC_006213	HCoV OC43	56
KU886219	Bovine CoV strain BCV-AKS-01	56
KT368891	Camel CoV HKU23	56

Fylogenetický strom uvedený níže (obrázek 5) je založený na fylogenetických souvislostech proteinu S. Znázorňuje, že rod alfa čeledi Coronavirinae již nejsou součástí jednotné linie. Uzel, který se odchýlil podobností, nese větve čtyř proteinů S, které patří hlodavčím koronaviřům, důvodem zřejmě bude vyšší odlišnost hlodavčích koronaviřových proteinů S od ostatních studovaných zvířecích a lidských S proteinů koronaviřů v rodě alfa.



**Obrázek 5:** Fylogenetický strom pro čeleď Coronavirinae za použití úseků sekvencí genu S. Čísla zapsaná u uzlů stromu, značí procentuální shodu bootstappingu. Taxony jsou značeny přístupovými čísly, následuje pozice proteinu S a název koronaviř. Žlutou barvou jsou vyznačeny alfa CoV, červenou barvou delta CoV, modrou gamma koronaviř a zelenou barvou beta koronaviř, které jsou rozděleny do zelených obdelníků, dle jejich podrodu Embecovirus, Nobecovirus, Hibecovirus, Sarbecovirus a Merbecovirus. Červeným trojúhelníkem je značena referenční sekvence SARS-CoV-2 (NC\_045512.2), fialovými čtverci vyplněnými bílou barvou je značeno dalších 6 lidských koronaviřů.

Zdroj: vlastní

Fylogenetická analýza sekvencí proteinu S u čeledi Coronavirinae viz obrázek 5 ukazuje, že SARS-CoV-2 (NC\_045512) je opět řazen do rodu beta a podrodu Sarbecovirus. Fylogeneticky je protein S SARS-CoV-2 nejvíce příbuzný proteinu S netopýřího koronaviru RaTG13 (MN996532). Další velice blízké jsou opět sekvence luskouního koronaviru MP789 (MT121216) a netopýřího koronaviru podobnému SARS ZXC21 (MG772934). Větev spolu s nimi sdílí SARS (NC\_004718) a dva izoláty netopýřích koronavirů podobných SARS (KY417143 a KC881006).

V následující tabulce 10 byla sledována procentuální podobnost jednotlivých proteinů u nejpodobnějších netopýřích koronavirů, které jsou zmíněny v předchozím odstavci s geny u nového typu koronaviru SARS-CoV-2. V levém sloupci jsou názvy proteinů a v druhém horním řádku jsou zapsána přístupová čísla koronavirů. U každého přístupového čísla, je tučně vyznačena hodnota, která udává nejvyšší procento shody. U přístupového čísla MG772934, které náleží, netopýřímu koronaviru podobnému SARS ZXC21, nemohly být proteiny ORF6, ORF7a a ORF7b porovnány, protože dle stránek NBCI je tento koronavirus neobsahuje.

*Tabulka 10: Procentuální podobnost jednotlivých proteinů*

Název proteinu	Podobnost s NC_045512.2 (%)					
	MN996532	MT121216	MG772934	KY417143	NC_004718	KC881006
ORF1ab	97	90	89	80	80	80
S	<b>93</b>	84	76	72	74	72
ORF3a	96	93	89	76	76	76
E	99	99	99	94	94	94
M	96	93	93	84	86	84
ORF6	98	96	Nelze porovnat	77	<b>41</b>	77
ORF7a	96	93	Nelze porovnat	84	82	84
ORF7b	99	92	Nelze porovnat	85	83	85
ORF8	97	<b>80</b>	<b>54</b>	<b>19</b>	55	<b>19</b>
N	97	96	91	87	88	87

Dále byly fylogenetické vztahy zkoumány blíže v rodě beta čeledi Coronavirinae viz obrázek 6. Tato fylogenetická analýza zkoumá podobnost koronavirových proteinů S mezi jednotlivými podrody betakoronavirů. Podrod Sarbecovirus ukazuje fylogenetické souvislosti mezi VOC SARS-CoV-2, které jsou na obrázku popsány.



**Obrázek 6: Fylogenetický strom pro rod beta čeledi Coronaviridae za použití úseků sekvencí genu S. Čísla zapsaná u uzlů stromu, značí procentuální shodu bootstappingu. Taxony jsou značeny přístupovými čísly, následuje pozice proteinu S a název koronavirovu. Beta koronavirové, jsou zde rozděleny do zelených obdelníků, dle jejich podrodu Embecovirus, Nobecovirus, Hibecovirus, Sarbecovirus a Merbecovirus. Červeným trojúhelníkem je značena referenční sekvence SARS-CoV-2 (NC\_045512.2), zelená kolečka označují jednotlivé varianty SARS-CoV-2, které jsou popsány na obrázku.**

Zdroj: vlastní

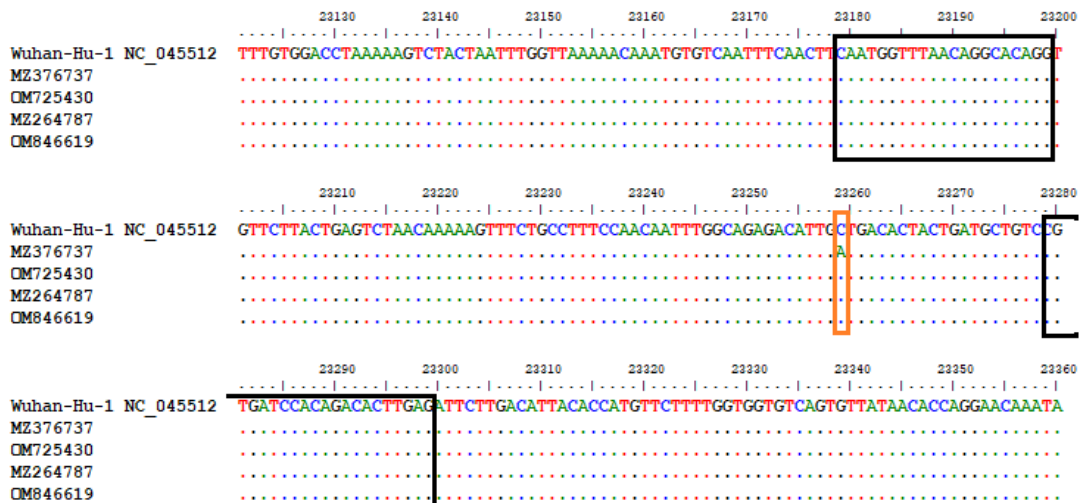
## 7.2 Primery

Součástí praktické části v této bakalářské práci bylo porovnání a posouzení vhodnosti běžně používaných primerů s našimi návrhy primerů. Cílovou sekvencí pro nasedání primerů byl protein S. Primery by měly být komplementární se všemi studovanými VOC SARS-CoV-2, tudíž by oblast, na kterou nasedají neměla podléhat mutacím a měly by ohraničovat takový úsek sekvence, kde došlo v průběhu vyvíjení SARS-CoV-2 ke změnám, které jsou charakteristické pro určité VOC SARS-CoV-2. Z tohoto důvodu byly vybrány tři cílové mutace L452R, E484K a A570D.

Z veřejně dostupných primerů bylo vybráno deset párů primerů, které ohraničovaly úsek sekvence, kde se nacházely některé mutace. Na základě porovnání v příloze 5 byly vybrány dva z nich pro porovnání s naším návrhem primerů, které jsou vloženy do tabulky 11, jedná se o pár primerů číslo 6 a číslo 9 z tabulky 2 a přílohy 5. První pár primerů značen černě cílil na sekvenci, která nese mutaci A570D, která je značena oranžově viz obrázek 7 a druhý pár primerů je značen černě a cílí na úsek sekvence, která nese mutaci E484K, ta je značena modře viz obrázek 8. Na tomto obrázku lze vidět ještě dvě mutace, jedná se o mutaci T478K, kde došlo u VOC delta ke změně v jedné bázi (C za A) a druhou mutací je N501Y, která není příliš vhodná pro odlišení jednotlivých VOC, z důvodu výskytu u třech (alfa, beta a gamma) ze čtyř zkoumaných VOC.

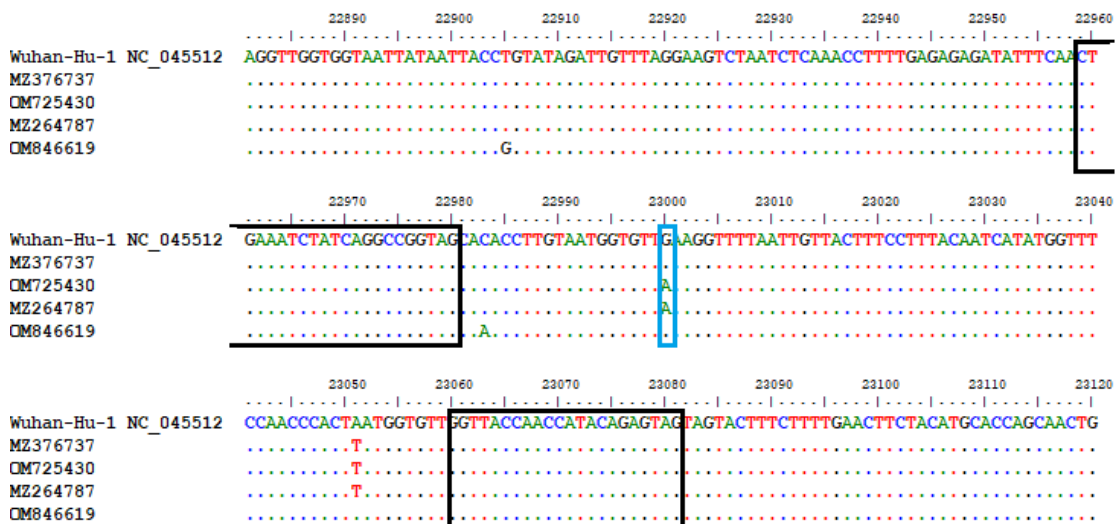
*Tabulka 11: Vybrané 2 páry primerů*

Pár primerů	Směr primeru	Sekvence primerů 5'→3'
1.	Forward	CAATGGTTTAACAGGCACAGG
	Revers	CTCAAGTGTCTGTGGATCACG
2.	Forward	CTGAAATCTATCAGGCCGGTAG
	Revers	CTACTCTGTATGGTTGGTAACC



Obrázek 7: Část sekvence S proteinu s vyznačenými primery 1.

Zdroj: vlastní

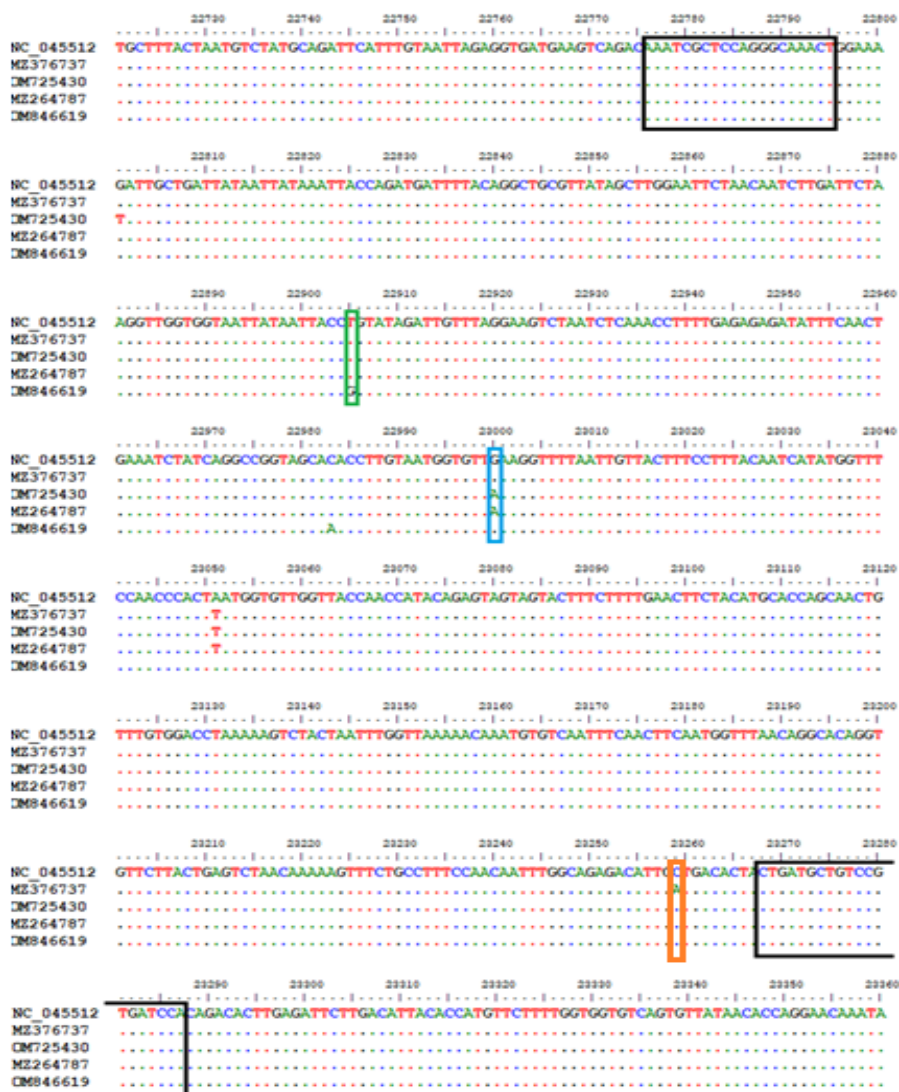


Obrázek 8: Část sekvence S proteinu s vyznačenými primery 2.

Zdroj: vlastní



Navrhovaný pár primerů (forward 5'-AAATCGCTCCAGGGCAAACCT-3' a revers 5'-TGGATCACGGACAGCATCAG-3') je značen černě a mutace, které se vyskytují v cílové sekvenci jsou L452R značena zeleně, E484K značena modře a A570D značena oranžově viz obrázek 9. Tento primer je specifický pro rozlišení všech čtyř zkoumaných VOC SARS-CoV-2. Ohraničuje také mutaci K417N, kde došlo ke změně (G za T) u VOC beta, což by se také mohlo využít při rozlišování variant. A dalšími mutacemi jsou T478K a N501Y, které jsou popsány výše.



Obrázek 9: Část sekvence S proteinu s vyznačenými primery 3.

Zdroj: vlastní

## 8. Diskuse

### 8.1 Fylogenetické souvislosti

Stanovení fylogenetických souvislostí koronavirů bylo provedeno na základě fylogenetické analýzy s použitím již publikovaných sekvencí koronavirů. Pomocí počítačové softwaru MEGA-X a metody Neighbor-Joining byly rekonstruovány tři fylogenetický stromy. První dva stromy zkoumaly fylogenetické souvislosti mezi všemi rody koronavirů (alfa, beta, gamma a delta) a třetí strom byl zaměřený na zkoumání fylogenetických souvislostí jen v rodě beta a jeho podrodů.

Ve všech třech případech fylogenetické analýzy byly studované sekvence SARS-CoV-2 zařazeny do rodu beta a podrodu Sarbecovirus společně se SARS. Rod beta dále zahrnuje tři lidské koronaviry HCoV-HKU1 a HCoV-OC43, které spadají do podrodu Embecovirus a MERS-CoV, který je řazen do podrodu Merbecovirus. Zbylé dva lidské koronaviry HCoV-NL63 a HCoV-229E jsou řazeny do rodu alfa. Fylogenetická analýza celogenomových sekvencí, viz obrázek 4 vykazuje největší podobnost SARS-CoV-2 s netopýřím koronavirem RaTG13, izolátem luskouního koronaviru MP789 a izolátem netopýřního koronaviru ZXC21 podobnému SARS. Vedlejší linie nejpříbuznějších celogenomových sekvencí je tvořena SARS-CoV a dvěma netopýřními koronaviry podobnými SARS.

Výsledky v tabulce 5 naznačují, že fylogenetická analýza vycházela z těchto podobností, tudíž čím blíže se sekvence nachází ve fylogenetickém stromu k SARS-CoV-2, tím je jeho podobnost k SARS-CoV-2 větší. Pokud budeme brát v potaz tabulky 5, 6, 7, 8 a 9, tak můžeme říct, že SARS-CoV-2 nevykazuje ani u jednoho podrodu, kromě podrodu Sarbecovirus podobnosti vyšší než 58 %. Nejpodobnější koronaviry k SARS-CoV-2 jsou netopýří koronaviry, které jsou řazeny do podrodu Sarbecovirus a SARS-CoV. V tabulce 5 vyšla podobnost SARS-CoV-2 s netopýřím koronavirem RaTG13 rovna 96 % a s netopýřím koronavirem podobným SARS rovna 88 %, zatímco podobnost se SARS je pouze 80 %. Tyto výsledky naznačují, že je pravděpodobné, že má SARS-CoV-2 původ spíše u netopýřích koronavirů než u lidského koronaviru SARS.

Výsledky fylogenetické analýzy, která zkoumala podobnost celých genomů koronavirů, byly porovnány se studií od Cheina et al. Tato studie vykazuje velmi podobné výsledky, odlišnosti spočívají ve zvolení studovaných sekvencí koronavirů, kdy analýza nezahrnuje netopýří koronavirus ZXC21 a netopýří koronaviry podobné SARS. Byly ale použity sekvence netopýřího koronaviru RaTG13 a luskouních koronavirů, které vykazují stejnou podobnost jako v naší analýze. [9]

Dle Singha a Yia jsou koronaviry SARS-CoV a MERS-CoV příkladem potenciálního přenosu virů ze zvířat na člověka, a proto je budoucí přenos koronavirů na lidskou populaci nevyhnutelný. Novému koronaviru SARS-CoV-2 přiřazují také zoonotický původ. Důvodem je častá rekombinace koronavirů, kdy při těsném kontaktu zvířat s koronaviry dojde k výměně virů čímž dochází k rekombinaci mezi různými kmeny. Je dost možné, že tyto události přispěly v evoluční historii SARS-CoV-2 k vývoji silného kmene, který je schopný lehce infikovat lidské buňky. [46]

Protein S je důležitý z hlediska přenosu viru na hostitelskou buňku. Tento protein je vázán na lidský angiotenzin-konvertující enzym 2, který se nachází na buněčném povrchu, tento proces umožňuje proniknutí viru na lidského těla. [46] Z tohoto důvodu byla provedena fylogenetická analýza proteinů S koronavirů. Tato analýza prokázala opět největší podobnost SARS-CoV-2 k netopýřímu koronaviru RaTG13, která vykazuje podobnost 93 % v tomto proteinu viz tabulka 10. Lze vidět, že na rozdíl od fylogenetické analýzy, kde se porovnávaly celé genomy, tak u fylogenetického stromu, viz obrázek 5 byly izoláty luskouního koronaviru a netopýří koronavirus podobný SARS řazeny do větve se SARS a dalšími netopýřími koronaviry podobným SARS. Další velká změna, kterou lze na tomto stromu pozorovat, je rozdělení rodu alfa koronavirů. K tomuto rozdělení nejspíše došlo z důvodu větších změn, které se objevují pouze v proteinu S. V oddělené větvi jsou čtyři koronaviry, které infikují hlodavce, zatímco ve větvi s ostatními alfa koronaviry se již takové koronaviry nevyskytují, což naznačuje, že hlodavčí koronaviry mají vysoké odlišnosti právě v proteinu S.

Z tabulky 10 je zřejmé, že nejpodobnější strukturu se SARS-CoV-2 ve všech studovaných proteinech má netopýří koronavirus RaTG13, což bylo zmíněno výše. Nejvyšší podobnost je u proteinu E, a to u všech studovaných koronavirů. SARS se nejméně podobá v proteinu ORF6. A zbylé studované koronaviry jsou nejvíce odlišné v proteinu ORF8, dva netopýří koronaviry podobné SARS vykazují pouhých 19 %. Za tento veliký rozdíl může jejich odlišná délka, kdy u SARS-CoV-2, je délka ORF8 80 nukleotidových bází a u netopýřího koronaviru podobnému SARS je délka 365 nukleotidových bází jedná se tedy o rozdíl 285 bází.

Dále byly fylogenetické souvislosti porovnávány za použití sekvencí betakoronavirů, které charakterizovaly jednotlivé VOC (alfa, beta, gamma a delta) SARS-CoV-2. Na obrázku 6 varianta beta není zařazena do linie SARS-CoV-2, tato anomálie je způsobena tím, že tento izolát obsahuje v sekvenci proteinu S na pozicích 21836-21864 místo jednotlivých bází zapsaná písmena N (může obsahovat jakoukoli bázi) viz obrázek 10. Další varianty jsou již velmi podobné referenční sekvenci SARS-CoV-2.

```
>OM725430.1:21551-25363 Severe acute respiratory syndrome coronavirus 2 isolate SARS-CoV-2/human/SouthAfrica/NHLS-UCT-GS-0807/2020, complete genome
ATGTTTGTGTTTTCTGTTTTATTGCCACTAGTCTCTAGTCAGTGTGTTAATCTTACAACCAGAAGTCAAT
TACCCCTGCATACACTAATTCTTTCACACGTGGTGTATTACCCTGACAAAAGTTTCAGATCCTCAGT
TTTACATTCAACTCAGGACTTGTCTTACCTTTCTTTTCCAATGTTACTTGGTTCCATGCTATACATGTC
TCTGGGACCAATGGTACTAAGAGGTTTGTAAACCTGCTCCTACCAATTAATGATGGTGTATTGCTT
CCACTGAGAACTTAACATAATAAGAGGCTGGATTTTGGTACTACTTTAGATTGAAAGACCCAGTCCCT
ACTTATTGTTAATAACGCTACTAATGTTGTTATTAAGTCTGTGAATTTCAATTTGTAATGATCCATT
TTGGGTGTTTATTACCACAAAAACAACAAAAGTTGGATGGAAAAGTGAGTTCAGAGTTTATTCTAGTGCGA
ATAATTGCACCTTTGAATATGTCCTCAGCCTTTCTTATGGACCTTGAAGGAAAACAGGGTAATTTCAA
AAATCTTAGGGAATTTGTGTTAAGAATATTGATGGTATTTTAAAAATATATTCTAAGCACACGCCTATT
AATTTAGTGCCTGCTCCTCAGGGTTTTTCGGCTTTAGAACCATTTGGTAGATTTGCCAATAGGTATTA
ACATCACTAGGTTTCAAACCTTACATATAAGTTATTTGACTCCTGGTGATTCTTCTCAGGTTGGACAGC
TNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
NNNNNNNNNNATGCTGTAGACTGTGCACCTCTCAGAAAACAAAGTGTACGTTGAAATCCTTCA
CTGTAGAAAAGGAATCTATCAAACCTTAACTTTAGAGTCCAACCAACAGAACTATTGTTAGATTTCC
TAATATTACAACCTTGTGCCCTTTTGGTGAAGTTTTAACGCCACCAGATTTGCATCTGTTTATGCTTGG
AACAGGAAGAGAATCAGCAACTGTGTTGCTGATTATTCTGCTCATATAATCCGCATCATTTTCCACTT
TTAAGTGTATGGAGTGTCTCTACTAAATTAATGATCTCTGCTTACTAATGCTATGCGAGATTCATT
TGTAATTAGAGGTGATGAAGTCAGACAAATCGCTCCAGGGCAAACCTGGAAATATTGCTGATTATAATTAT
AAATTACCAGATGATTTACAGGCTGCGTTATAGCTTGGAAATCTAACAACTTTGATTCTAAGGTTGGTG
GTAATTATAATTACCTGTATAGATTGTTTAGGAAGTCTAATCTCAAACCTTTGAGAGAGATATTTCAAC
TGAAATCTATCAGGCCGTTAGCACACCTTGTAATGGTGTAAAGGTTTTAATTGTTACTTTCTTTTACAA
TCATATGTTTCCAACCCACTATGGTGTGGTTACCAACCATACAGAGTAGTAGTACTTTCTTTTGAAC
```

**Obrázek 10:** Část sekvence proteinu S u jihoafrické varianty SARS-CoV-2.

Zdroj: vlastní

## 8.2 Primery

Pro porovnání a posouzení vhodnosti primerů pro rozlišení studovaných VOC SARS-CoV-2 při použití detekční metody PCR, byly vybrány dostupné primery, které nasedají na cílovou sekvenci. První pár primerů cílí na úsek sekvence, kde dochází k mutaci pouze u VOC alfa, tudíž je vhodný pro rozlišení jen jedné varianty SARS-CoV-2, proto není moc vhodný pro komplexní stanovení více mutačních variant SARS-CoV-2.

Druhý pár primerů je specifitější pro rozlišení studovaných VOC SARS-CoV-2. Nasedá na úsek sekvence, která nese tři mutace. Mutace E484K, kterou nesou dvě VOC SARS-CoV-2 (beta a gamma). Mutace T478K, která je charakteristická pro VOC delta a N501Y, která není příliš vhodná pro odlišení jednotlivých studovaných VOC, z důvodu výskytu u třech (alfa, beta a gamma) ze čtyř zkoumaných VOC, jak bylo popsáno již výše v kapitole „Výsledky“. Dá se říct, že tento primer je vhodný pro rozlišení všech studovaných VOC, ale pokud nebude brána v potaz mutace N501Y, tak je tento primer vhodný pro rozlišení beta, gamma a delta varianty.

Navrhovaný pár primerů na obrázku 3 cílí na úsek sekvence, kde se vyskytují charakteristické mutace, pro rozlišení všech studovaných VOC. Zahrnuje dvě mutace (L452R a E484K), na které cílí kity, vyráběné pro detekci SARS-CoV-2, které jsou schopny rozlišit jednotlivé zkoumané VOC tohoto nového typu koronaviru. Všechny tři páry primerů, jsou vhodné na detekci pouze SARS-CoV-2 protože nasedají pouze na sekvence SARS-CoV-2 a nenedávají na ostatní sekvence HCoV.

## 9. Závěr

Teoretická část této bakalářské práce byla zaměřena na bližší seznámení s virem, kdy byla blíže popsána skupina koronavirů, další částí bylo přiblížení fylogenetiky a následně detekční metody PCR.

Praktická část práce byla věnována především diagnostice koronavirů pomocí metody PCR. Autorka práce byla seznámena s touto metodou detailně na pracovišti klinické genetiky. Vzhledem k významu fylogenetiky pro detekci různých variant virů byly v rámci praktické části dále zkoumány fylogenetické souvislosti koronavirů, které byly hlavně zaměřeny na sledování podobnosti koronavirů s novým typem koronaviru SARS-CoV-2. SARS-CoV-2 vykazoval největší podobnost s netopýřím koronavirem RaTG13, luskouním koronavirem, dále s koronaviry podobnými SARS a SARS-CoV. Z našich výsledků vyplývá, že nový typ koronaviru má pravděpodobně zoonotický původ.

Úseky, které vykazovaly největší odlišnosti od SARS-CoV-2 byl protein S u netopýřeho koronaviru RaTG13, dále protein ORF6 u SARS a u ostatních netopýřích koronavirů podobnými SARS a pangolíního koronaviru ORF8.

Dále byly porovnávány páry primerů pro detekci SARS-CoV-2, které jsou schopny rozlišit zkoumané VOC tohoto koronaviru. Podařilo se nalézt úsek sekvence, kde se vyskytují mutace charakteristické pro jednotlivé studované VOC a navrhnout pár primerů, který tento úsek ohraničuje. Tímto způsobem je tedy principiálně možné jednotlivé VOC rozlišit.

## 10. Literatura

- [1] ALEEM, A., SAMAD, A. B., SLENKER, A. K. *Emerging Variants of SARS-CoV-2 And Novel Therapeutics Against Coronavirus (COVID-19)* [Updated 2022 Jan 5]. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2022 Jan-. Available from: <https://www.ncbi.nlm.nih.gov/books/NBK570580/>
- [2] BALÁŽOVÁ, A., ČERNÝ, J., DAMAŠKA, A., et al. *Život je jen náhoda aneb Evoluce života na Zemi*. Praha: Česká zemědělská univerzita v Praze, Ústřední komise Biologické olympiády, 2015, 143 s. ISBN 978-80-213-2565-4
- [3] BARTŮŇKOVÁ, J., PAULÍK, M. *Vyšetřovací metody v imunologii*. 2., přeprac. a dopl. vyd. Praha: Grada, 2011, IV, 168 s. ISBN 978-80-247-3533-7.
- [4] BEDNÁŘ, M. *Lékařská mikrobiologie: bakteriologie, virologie, parazitologie*. Praha: Triton, 2009, 560 s. ISBN 978-80-9028-966-5.
- [5] BENSON, D. A., CAVANAUGH, M., CLARK, K. KARSCH-MIZRACHI, I., OSTELL, J., PRUITT, K. D., a SAYERS, E. W. GenBank. *Nucleic Acids Research* [online]. 2018, **46**(D1), D41-D47 [cit. 2022-04-25]. ISSN 0305-1048. Dostupné z: doi:10.1093/nar/gkx1094
- [6] BOEHM, E., KRONIG, I., NEHER, R. A., ECKERLE, I., VETTER, P. a KAISER, L. *Novel SARS-CoV-2 variants: the pandemics within the pandemic*. *Clinical Microbiology and Infection* [online]. 2021, 27(8), 1109-1117 [cit. 2022-02-22]. ISSN 1198743X. Dostupné z: doi:10.1016/j.cmi.2021.05.022
- [7] BROWN, T. A. *Genomes: Chapter 16, Molecular Phylogenetics* [online]. 2nd edition. Oxford: Wiley-Liss, 2002 [cit. 2022-04-25]. ISBN 0-471-25046-5. Dostupné z: <https://www.ncbi.nlm.nih.gov/books/NBK21122/>
- [8] CELER, V. *Obecná virologie*. Hradec Králové: Nucleus HK, 2010, 143 s. ISBN 978-80-87009-70-3.
- [9] CHEN, Z., BOON, S. S., WANG, M. H., CHAN, R. W.Y. a CHAN, P. K.S. Genomic and evolutionary comparison between SARS-CoV-2 and other human coronaviruses. *Journal of Virological Methods* [online]. 2021, 289 [cit. 2022-04-27]. ISSN 01660934. Dostupné z: doi:10.1016/j.jviromet.2020.114032
- [10] CIOTTI, M., ANGELETTI, S., MINIERI, M., et al. COVID-19 Outbreak: An Overview. *Chemotherapy* [online]. 2020, **64**(5-6), 215-223 [cit. 2020-12-01]. ISSN 0009-3157. Dostupné z: doi:10.1159/000507423
- [11] CUI, J., LI, F. a SHI, Z. Origin and evolution of pathogenic coronaviruses. *Nature Reviews Microbiology* [online]. 2019, **17**(3), 181-192 [cit. 2022-01-26]. ISSN 1740-1526. Dostupné z: doi:10.1038/s41579-018-0118-9
- [12] DAVIES, N. G., ABBOTT, S., C. BARNARD, R. C. et al. Estimated transmissibility and impact of SARS-CoV-2 lineage B.1.1.7 in

England. *Science* [online]. 2021, **372**(6538) [cit. 2022-02-22]. ISSN 0036-8075. Dostupné z: doi:10.1126/science.abg3055

[13] DAYRAT, B., LINDER, P. The Roots of Phylogeny: How Did Haeckel Build His Trees? *Systematic Biology* [online]. 2003, **52**(4), 515-527 [cit. 2021-02-16]. ISSN 1076-836X. Dostupné z: doi:10.1080/10635150390218277

[14] DHAMA, K., KHAN, S., TIWARI, R., et al. Coronavirus Disease 2019–COVID-19. *Clinical Microbiology Reviews* [online]. 2020, **33**(4), e00028-20 [cit. 2021-01-05]. ISSN 0893-8512. Dostupné z: doi:10.1128/CMR.00028-20

[15] EDGAR, R. C., BATZOGLOU, S. Multiple sequence alignment. *Current Opinion in Structural Biology* [online]. 2006, **16**(3), 368-373 [cit. 2021-03-30]. ISSN 0959440X. Dostupné z: doi:10.1016/j.sbi.2006.04.004

[16] FARIA, N. R., MELLAN, T. A., WHITTAKER, Ch. et al. Genomics and epidemiology of the P.1 SARS-CoV-2 lineage in Manaus, Brazil. *Science* [online]. 2021, **372**(6544), 815-821 [cit. 2022-02-22]. ISSN 0036-8075. Dostupné z: doi:10.1126/science.abh2644

[17] FLEGR, J. *Evoluční biologie*. 2., opr. a rozš. vyd. Praha: Academia, 2009. ISBN 978-80-200-1767-3.

[18] GASCUEL, O. Neighbor-Joining Revealed. *Molecular Biology and Evolution* [online]. 2006, 23(11), 1997-2000 [cit. 2022-04-25]. ISSN 0737-4038. Dostupné z: doi:10.1093/molbev/msl072

[19] GALLOWAY, S. E., PAUL, P., MACCANNELL, D. R. et al. Emergence of SARS-CoV-2 B.1.1.7 Lineage — United States, December 29, 2020–January 12, 2021. *MMWR. Morbidity and Mortality Weekly Report* [online]. 2021, **70**(3), 95-99 [cit. 2021-03-01]. ISSN 0149-2195. Dostupné z: doi:10.15585/mmwr.mm7003e2

[20] GIOVANETTI, M., BENEDETTI, F., CAMPISI, G. et al. Evolution patterns of SARS-CoV-2: Snapshot on its genome variants. *Biochemical and Biophysical Research Communications* [online]. 2021, **538**, 88-91 [cit. 2022-04-25]. ISSN 0006291X. Dostupné z: doi:10.1016/j.bbrc.2020.10.102

[21] GRUBAUGH, N. D., HODCROFT, E. B., FAUVER, J. R., PHELAN, A. L., CEVIK, M. Public health actions to control new SARS-CoV-2 variants. *Cell* [online]. 2021 [cit. 2021-02-26]. ISSN 00928674. Dostupné z: doi:10.1016/j.cell.2021.01.044

[22] HAMPL, V. Molekulární taxonomie [online]. In: Praha: Hampl, 2016 [cit. 2022-04-25]. Dostupné z: <http://web.natur.cuni.cz/~vlada/moltax/>

[23] HALL, B. G. Building Phylogenetic Trees from Molecular Data with MEGA. *Molecular Biology and Evolution* [online]. 2013, **30**(5), 1229-1235 [cit. 2022-04-25]. ISSN 0737-4038. Dostupné z: doi:10.1093/molbev/mst012



- [24] HASÖKSÜZ, M., KILIÇ, S. a SARAÇ, F. Coronaviruses and SARS-COV-2. *TURKISH JOURNAL OF MEDICAL SCIENCES* [online]. 2020, **50**(SI-1), 549-556 [cit. 2021-12-14]. ISSN 13036165. Dostupné z: doi:10.3906/sag-2004-127
- [25] CHALUPOVÁ-KARLOVSKÁ, V. *Obecná biologie: středoškolská učebnice evoluce, biologie buňky, genetika: s 558 řešenými testovými otázkami*. 2., opr. vyd. Olomouc: Nakladatelství Olomouc, 2010. ISBN isbn978-80-7182-282-0.
- [26] JACKWOOD, M. W., HALL, D. a HANDEL, A. Molecular evolution and emergence of avian gammacoronaviruses. *Infection, Genetics and Evolution*[online]. 2012, **12**(6), 1305-1311 [cit. 2022-01-26]. ISSN 15671348. Dostupné z: doi:10.1016/j.meegid.2012.05.003
- [27] MACHADO, J. D., SCOTT, R., GUIRALES, S. a JANIES, D. A. Fundamental evolution of all Orthocoronavirinae including three deadly lineages descendent from Chiroptera-hosted coronaviruses: SARS-CoV, MERS-CoV and SARS-CoV-2. *Cladistics* [online]. 2021, **37**(5), 461-488 [cit. 2021-12-14]. ISSN 0748-3007. Dostupné z: doi:10.1111/cla.12454
- [28] KANDEEL, M., MOHAMED, M. E. M., EL-LATEEF, H. M. A., VENUGOPALA, K. N. a EL-BELTAGI, H. S. Omicron variant genome evolution and phylogenetics. *Journal of Medical Virology* [online]. 2022, 94(4), 1627-1632 [cit. 2022-02-22]. ISSN 0146-6615. Dostupné z: doi:10.1002/jmv.27515
- [29] KANNAN, S., ALI, P. S. S., a SHEEZA, A. Omicron (B.1.1.529) – variant of concern – molecular profile and epidemiology: a mini review. *European Review for Medical and Pharmacological Sciences*. Maldives, 2021, **24**(25), 8019-8022. Dostupné z: doi:10.26355/eurrev\_202112\_27653
- [30] KIMURA, M. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *Journal of Molecular Evolution* [online]. 1980, 16(2), 111-120 [cit. 2022-04-25]. ISSN 0022-2844. Dostupné z: doi:10.1007/BF01731581
- [31] KORSMAN, S. N.J., ZYL, G. U., NUTT, L., ANDERSSON, M. I., a PREISER, W. Human coronaviruses. *Virology* [online]. Elsevier, 2012, 2012, s. 94-95 [cit. 2021-12-14]. ISBN 9780443073670. Dostupné z: doi:10.1016/B978-0-443-07367-0.00040-9
- [32] KUMAR, S., STECHER, G., LI, M., KNYAZ, CH., TAMURA, K. a BATTISTUZZI, F. U. MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. *Molecular Biology and Evolution* [online]. 2018, **35**(6), 1547-1549 [cit. 2022-04-25]. ISSN 0737-4038. Dostupné z: doi:10.1093/molbev/msy096
- [33] LABGUIDE.CZ. Průvodce laboratoří. In: *Metody: PCR* [online]. c2014-2019 [cit. 2022-04-25]. Dostupné z: <http://labguide.cz>

- [34] LE PAGE, M. Threats from new variants. *New Scientist* [online]. 2021, **249**(3316), 8-9 [cit. 2021-03-01]. ISSN 02624079. Dostupné z: doi:10.1016/S0262-4079(21)00003-8
- [35] LI, K.-B. ClustalW-MPI: ClustalW analysis using distributed and parallel computing. *Bioinformatics* [online]. 2003, **19**(12), 1585-1586 [cit. 2022-04-25]. ISSN 1367-4803. Dostupné z: doi:10.1093/bioinformatics/btg192
- [36] LIU, D.X., LIANG, J.Q., FUNG, T.S. Human Coronavirus-229E, -OC43, -NL63, and -HKU1 (Coronaviridae). *Encyclopedia of Virology* [online]. Elsevier, 2021, 2021, s. 428-440 [cit. 2021-03-23]. ISBN 9780128145166. Dostupné z: doi:10.1016/B978-0-12-809633-8.21501-X
- [37] METTENLEITER, T. C. The First “Virus Hunters”. In *Loeffler’s Footsteps – Viral Genomics in the Era of High-Throughput Sequencing* [online]. Elsevier, 2017, 2017, s. 1-16 [cit. 2022-01-26]. Advances in Virus Research. ISBN 9780128125984. Dostupné z: doi:10.1016/bs.aivir.2017.07.005
- [38] National Center for Immunization and Respiratory Diseases (NCIRD), Division of Viral Diseases. CDC COVID-19 Science Briefs [Internet]. Atlanta (GA): Centers for Disease Control and Prevention (US); 2020-. Science Brief: Omicron (B.1.1.529) Variant. [Updated 2021 Dec 2]. Available from: <https://www.ncbi.nlm.nih.gov/books/NBK575856/>
- [39] NEEDLEMAN, S.B. a WUNSCH., CH.D. A general method applicable to the search for similarities in the amino acid sequence of two proteins. *Journal of Molecular Biology* [online]. 1970, **48**(3), 443-453 [cit. 2022-04-25]. ISSN 00222836. Dostupné z: doi:10.1016/0022-2836(70)90057-4
- [40] Needlemanův–Wunschův algoritmus. In: *Wikipedia: the free encyclopedia* [online]. San Francisco (CA): Wikimedia Foundation, 2001- [cit. 2022-04-25]. Dostupné z: <https://cs.wikipedia.org/wiki/Needleman%20v%20Wunsch%20algoritmus>
- [41] OLIVEIRA, B. E., OLIVEIRA, L. C., SABINO E. C., OKAY T. S. SARS-CoV-2 and the COVID-19 disease: a mini review on diagnostic methods. *Revista do Instituto de Medicina Tropical de São Paulo* [online]. 2020, **62** [cit. 2021-01-05]. ISSN 1678-9946. Dostupné z: doi:10.1590/s1678-9946202062044
- [42] SEIFI, M., GHASEMI, A., HEIDARZADEH, S., KHOSRAVI, M., NAMIPASHAKI, A., MEHRI, V., ALIZADEH, A., DANAEI, N. Overview of Real-Time PCR Principles. HERNANDEZ-RODRIGUEZ, Patricia, ed. *Polymerase Chain Reaction* [online]. InTech, 2012, 2012-05-30 [cit. 2021-03-03]. ISBN 978-953-51-0612-8. Dostupné z: doi:10.5772/39220
- [43] SETHURAMAN, N., JEREMIAH, S.S., RYO, A. Interpreting Diagnostic Tests for SARS-CoV-2. *JAMA* [online]. 2020, **323**(22) [cit. 2021-03-02]. ISSN 0098-7484. Dostupné z: doi:10.1001/jama.2020.8259

- [44] SHIEHZADEGAN, S., ALAGHEMAND, N., FOX, M. a VENKETARAMAN, V. Analysis of the Delta Variant B.1.617.2 COVID-19. *Clinics and Practice* [online]. 2021, **11**(4), 778-784 [cit. 2022-02-22]. ISSN 2039-7283. Dostupné z: doi:10.3390/clinpract11040093
- [45] SCHINDLER, J. *Mikrobiologie: pro studenty zdravotnických oborů. 2., dopl. a přeprac. vyd.* Praha: Grada, 2014, XXIV, 215 s. Sestra (Grada). ISBN 978-80-247-4771-2.
- [46] SINGH, D. a YI, S.V. On the origin and evolution of SARS-CoV-2. *Experimental & Molecular Medicine* [online]. 2021, **53**(4), 537-547 [cit. 2022-04-25]. ISSN 1226-3613. Dostupné z: doi:10.1038/s12276-021-00604-z
- [47] SNUSTAD, D. P., SIMMONS, J. M., RELICHOVÁ, J., et al. *Genetika*. Druhé, aktualizované vydání. Přeložil DOŠKAŘ, J., přeložil FAJKUS, J., přeložil HOŘÍN, P., přeložil KNOLL, A., přeložil KUGLÍK, P., přeložil ŠMARDA, J., přeložil ŠMARDOVÁ, J., přeložil VESELSKÁ, R., přeložil VYSKOT. B., Brno: Masarykova univerzita, 2017, XIX, 844 s. ISBN 978-80-210-8613-5.
- [48] SOLTIS, P.S., FOLK, R.A., SOLTIS, D.E. Darwin review: angiosperm phylogeny and evolutionary radiations. *Proceedings of the Royal Society B: Biological Sciences* [online]. 2019, **286**(1899) [cit. 2021-02-16]. ISSN 0962-8452. Dostupné z: doi:10.1098/rspb.2019.0099
- [49] SOUČEK, M., ŠPINAR J., VORLÍČEK, J. *Vnitřní lékařství*. Praha: Grada, 2011, 807–1577 s. ISBN 978-80-210-5418-9.
- [50] TABIBZADEH, A., ESGHAEI, M. SOLTANI, S. et al. Evolutionary study of COVID-19, severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) as an emerging coronavirus: Phylogenetic analysis and literature review. *Veterinary Medicine and Science* [online]. 2021, **7**(2), 559-571 [cit. 2022-04-25]. ISSN 2053-1095. Dostupné z: doi:10.1002/vms3.394
- [51] TANG, J.W., TAMBYAH, P.A., HUI, D.S. Emergence of a new SARS-CoV-2 variant in the UK. *Journal of Infection* [online]. 2020 [cit. 2021-03-01]. ISSN 01634453. Dostupné z: doi:10.1016/j.jinf.2020.12.024
- [52] VAN KASTEREN, P.B., VAN DER VEER, B., VAN DEN BRINK, S. et al. Comparison of seven commercial RT-PCR diagnostic kits for COVID-19. *Journal of Clinical Virology* [online]. 2020, **128** [cit. 2021-03-03]. ISSN 13866532. Dostupné z: doi:10.1016/j.jcv.2020.104412
- [53] VOTAVA, M. *Lékařská mikrobiologie speciální*. Brno: Neptun, 2003, 495 s. ISBN 80-902896-6-5.
- [54] VURAL-OZDENIZ, M., AKTURK, A., DEMIRDIZEN, M., LEKA, R., ACAR, R., a KONU, O. CoVrimer: A tool for aligning SARS-CoV-2 primer sequences and selection of conserved/degenerate primers. *Genomics* [online]. 2021, **113**(5), 3174-3184 [cit. 2022-04-25]. ISSN 08887543. Dostupné z: doi:10.1016/j.ygeno.2021.07.020

[55] WARNOW, T. *Computational phylogenetics: an introducing to desining methods for phylogeny estimatin*. Cambridge: Cambridge University Press, 2018. ISBN 978-1-107-18471-8.

[56] ZOUHAR, M., GAAR, V. *Základní metody diagnostiky a determinance karanténních háďátek z rodu Globodera (G. rostochiensis a G. pallida) pro potřebu praxe*. V Praze: Česká zemědělská univerzita, 2003, 71 s. ISBN 80-213-1087-1.

## 11. Seznam obrázků a tabulek

### 11.1 Seznam obrázků

Obrázek 1: Fylogenetické stromy 1 .....	22
Obrázek 2: Fylogenetické stromy 2.....	24
Obrázek 3: Schéma reakce PCR.....	27
Obrázek 4: Fylogenetický strom pro čeleď Coronavirinae za použití celogenomových sekvencí.....	40
Obrázek 5: Fylogenetický strom pro čeleď Coronavirinae za použití úseků sekvencí genu S.....	44
Obrázek 6: Fylogenetický strom pro rod beta čeledi Coronaviridae za použití úseků sekvencí genu S .....	46
Obrázek 7: Část sekvence S proteinu s vyznačenými primery 1.....	48
Obrázek 8: Část sekvence S proteinu s vyznačenými primery 2.....	48
Obrázek 9: Část sekvence S proteinu s vyznačenými primery 3.....	49
Obrázek 10: Část sekvence proteinu S u jihoafrické varianty SARS-CoV-2.....	52

### 11.2 Seznam tabulek

Tabulka 1: Počet použitých sekvencí.....	35
Tabulka 2: Vybrané primery.....	36
Tabulka 3: Cílové mutace .....	37
Tabulka 4: Navržený pár primerů.....	38
Tabulka 5: Procentuální shoda celých genomů podrodu Sarbecovirus .....	42
Tabulka 6: Procentuální shoda celých genomů s podrodem Hibecovirus .....	42
Tabulka 7: Procentuální shoda celých genomů s podrodem Nobecovirus .....	42
Tabulka 8: Procentuální shoda celých genomů s podrodem Merbecovirus .....	42
Tabulka 9: Procentuální shoda celých genomů s podrodem Embecovirus .....	43
Tabulka 10: Procentuální podobnost jednotlivých proteinů.....	45
Tabulka 11: Vybrané 2 páry primerů.....	47

## 12. Seznam zkratek

A	Adenin
AAA	Lysin
BLAST	Algoritmus pro porovnávání sekvencí (z angl. Basic Local Alignment Search Tool)
C	Cytosin
CGG	Arginin
COVID-19	Koronavirové onemocnění 2019 (z angl. Coronavirus disease)
CTG	Leucin
Del69-70	Delece
DNA	Deoxyribonukleotidová kyselina
E	Obalový protein
FASTA	Formát souboru sekvence
G	Guanin
GAA	Kyselina glutamová
GAT	Asparagin
GCT	Alanin
GenBank	Veřejně dostupná databáze nukleotidových sekvencí
HCoV	Lidský koronavirus
M	Membránový protein
MEGA-X	Program pro analýzu molekulární evoluční genetiky (z angl. Molecular Evolutionary Genetics Analysis).
MERS-CoV	Koronavirus blízkovýchodního respiračního syndromu
NCBI	Národního centra pro biotechnologické informace
N	Nukleokapsidový protein
ORF	Otevřený čtecí rámec
PRC	Polymerázová řetězová reakce (z angl. Polymerase Chain Reaction)
RNA	Ribonukleotidová kyselina
RT-PCR	Reverzní transkriptázová polymerázová řetězová reakce v reálném čase
S	Povrchový, spikový protein
SARS-CoV	Koronavirus těžkého akutního respiračního syndromu (z angl. Severe acute respiratory syndrome-related coronavirus)
SARS-CoV-2	Koronavirus těžkého akutního respiračního syndromu 2
T	Thymin
Taq polymeráza	DNA polymeráza izolovaná z bakterie <i>Thermus aquanticus</i>
UPGMA	Metoda neváženého párování s aritmetickým průměrem (z angl.

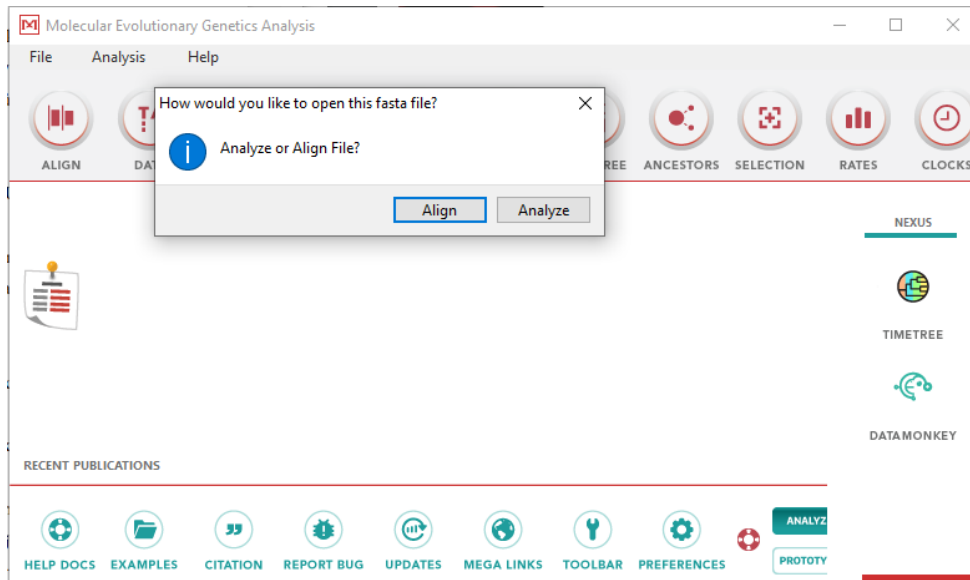
	Unweighted Pair Group Method with Arithmetic Mean)
UTR	Nepřeložená oblast
VOC	Varianta vzbuzující obavy (z angl. variants of concern)
VOI	Varianta zájmu (z angl. variants of interest)

## 13. Přílohy

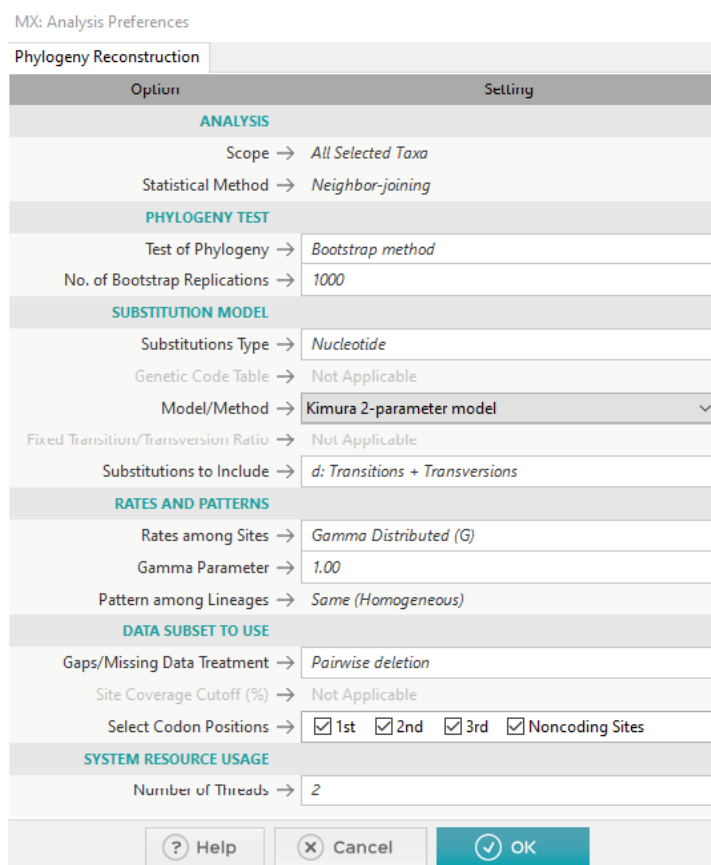
### 13.1 Seznam příloh

Příloha 1: Program MEGA-X: Výběr Align/Analyze. ....	65
Příloha 2: Nastavení metody Neighbor-Joining.....	65
Příloha 3: Nástroj BLAST: Globální zarovnání .....	66
Příloha 4: Nástroj Primer-BLAST .....	66
Příloha 5: Tabulka primerů .....	67





*Priloha 1: Program MEGA-X: Vyběr Align/Analyze.*



*Priloha 2: Nastavení metody Neighbor-Joining*

An official website of the United States government [Here's how you know](#)

**NIH** National Library of Medicine  
National Center for Biotechnology Information

**BLAST®** » Global Alignment Home Recent Results Saved Strategies

**Needleman-Wunsch Global Align Nucleotide Sequences**

**Nucleotide** Protein

Needleman-Wunsch alignment of two nucleotide sequences [?](#) [Reset page](#)  
[Bookmark](#)

**Enter Query Sequence**

Enter accession number, gi, or FASTA sequence [?](#) [Clear](#) Query subrange [?](#)

Input limited to 100,000 letters for either input sequence. The total length of both query and subject may not exceed 150,000 letters.

From   
To

Or, upload file  Soubor nevybrán [?](#)

Job Title   
Enter a descriptive title for your BLAST search [?](#)

**Enter Subject Sequence**

Enter accession number, gi, or FASTA sequence [?](#) [Clear](#) Subject subrange [?](#)

Input limited to 100,000 letters for either input sequence. The total length of both query and subject may not exceed 150,000 letters.

From   
To

Or, upload file  Soubor nevybrán [?](#)

**BLAST**  Show results in a new window

**+ Algorithm parameters**

*Priloha 3: Nástroj BLAST: Globální zarovnání*

[Retrieve recent results](#) [Publication](#) [Tips for finding specific primers](#)

**PCR Template**

Enter accession, gi, or FASTA sequence (A refseq record is preferred) [?](#) [Clear](#) Range [?](#) [Clear](#)

NC\_045512

Forward primer From  To   
Reverse primer

Or, upload FASTA file  Soubor nevybrán

**Primer Parameters**

Use my own forward primer (5'→3' on plus strand)  [?](#) [Clear](#)

Use my own reverse primer (5'→3' on minus strand)  [?](#) [Clear](#)

PCR product size Min  Max

# of primers to return

Primer melting temperatures (T<sub>m</sub>) Min  Opt  Max  Max T<sub>m</sub> difference  [?](#)

**Exon/intron selection**

A refseq mRNA sequence as PCR template input is required for options in the section [?](#)

Exon junction span  [?](#)

Exon junction match Min 5' match  Min 3' match  Max 3' match   
Minimal and maximal number of bases that must anneal to exons at the 5' or 3' side of the junction [?](#)

Intron inclusion  Primer pair must be separated by at least one intron on the corresponding genomic DNA [?](#)

Intron length range Min  Max  [?](#)

**Primer Pair Specificity Checking Parameters**

Specificity check  Enable search for primer pairs specific to the intended PCR template [?](#)

Search mode  [?](#)

Database  [?](#)

Exclusion  Exclude predicted Refseq transcripts (accession with XM, XR prefix)  Exclude uncultured/environmental sample sequences [?](#)

Organism  [Add organism](#)

Enter an organism name (or organism group name such as enterobacteriaceae, rodents), taxonomy id or select from the suggestion list as you type. [?](#)

Entrez query (optional)  [?](#)

Primer specificity stringency Primer must have at least  total mismatches to unintended targets, including

*Priloha 4: Nástroj Primer-BLAST*

Příloha 5: Tabulka primerů

Pár primerů	Směr primeru (F/R)	Sekvence primerů 5' → 3'	Teplota nasedání	Zastoupení GC (%)	Cílové mutace
1.	F	CCTACTAAATTAATGATCTCTGCTTTACT	58,15	30	K417N
	R	CAAGCTATAACGCAGCCTGTA	58,18	47	
2.	F	TTGGCAAAATTCAAGACTCACTTT	58,02	33,33	S982A
	R	TGTGGTTCATAAAAATTCCTTTGTG	56,98	32,00	T1027I
3.	F	TCAAGACTCACTTTCTTCCAC	55,48	42,86	S982A
	R	ATTTGAAACAAAGACACCTTCAC	56,13	34,78	T1027I
4.	F	AAGACTCACTTTCTTCCACAG	55,47	42,86	S982A
	R	CAAAGACACCTTCACGAGG	55,88	52,63	T1027I
5.	F	CCTACTAAATTAATGATCTCTGCTTTACT	58,15	30,00	K417N
	R	CAAGCTATAACGCAGCCTGTA	58,18	47,62	
6.	F	CAATGGTTTAACAGGCACAGG	57,70	47,62	A570D
	R	CTCAAGTGTCTGTGGATCACG	58,67	52,38	
7.	F	GGTTTTAATTGTTACTTTCCTTTACAA	55,58	25,93	N501Y
	R	GAAAGTACTACTACTCTGTATGGTTGG	58,81	40,74	
8.	F	TCCTTTACAATCATATGGTTTCCA	56,07	33,33	N501Y
	R	GCATGTAGAAGTTCAAAGAAAGT	56,14	33,33	
9.	F	CTGAAATCTATCAGGCCGGTAG	58,08	50,00	N501Y
	R	CTACTCTGTATGGTTGGTAACC	55,64	45,45	E484K
10.	F	GTTACTTTCCTTTACAATCATATG	52,00	29,17	N501Y
	R	TTTAGGTCCACAAACAGTTGC	56,93	42,86	