



VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

BRNO UNIVERSITY OF TECHNOLOGY

FAKULTA CHEMICKÁ

FACULTY OF CHEMISTRY

ÚSTAV FYZIKÁLNÍ A SPOTŘEBNÍ CHEMIE

INSTITUTE OF PHYSICAL AND APPLIED CHEMISTRY

**ANALÝZA INVERZNÍCH REPETIC V GENOMU
LIDSKÝCH PATOGENŮ**

ANALYSES OF INVERTED REPEATS IN HUMAN PATOGEN GENOMES

BAKALÁŘSKÁ PRÁCE

BACHELOR'S THESIS

AUTOR PRÁCE

AUTHOR

Anna Hanzlíková

VEDOUCÍ PRÁCE

SUPERVISOR

doc. Mgr. Václav Brázda, Ph.D.

BRNO 2020

Zadání bakalářské práce

Číslo práce: FCH-BAK1509/2019 Akademický rok: 2019/20
Ústav: Ústav fyzikální a spotřební chemie
Studentka: **Anna Hanzlíková**
Studijní program: Chemie a chemické technologie
Studijní obor: Chemie pro medicínské aplikace
Vedoucí práce: **doc. Mgr. Václav Brázda, Ph.D.**

Název bakalářské práce:

Analýza inverzních repetit v genomu lidských patogenů

Zadání bakalářské práce:

Literární rešerše lokálních struktur DNA se zaměřením na inverzní repetice, využití nástroje Palindrome finder pro charakterizaci přítomnosti a lokalizaci inverzních repetit v genomech organismů patogenních pro člověka. Analýza a zpracování dat.

Termín odevzdání bakalářské práce: 31. 7. 2020:

Bakalářská práce se odevzdává v děkanem stanoveném počtu exemplářů na sekretariát ústavu. Toto zadání je součástí bakalářské práce.

Anna Hanzlíková

student(ka)

doc. Mgr. Václav Brázda, Ph.D.

vedoucí práce

prof. Ing. Miloslav Pekař, CSc.

vedoucí ústavu

prof. Ing. Martin Weiter, Ph.D.

děkan

V Brně dne 31.1.2020

ABSTRAKT

Patogeny jsou organismy, které jsou příčinou různých onemocnění hostitele. Patří zde: priony, viry, bakterie, houby, prvoci a živočichové. Tato bakalářská práce je zaměřená konkrétně na viry, způsobující lidská onemocnění jako jsou závažné respirační syndromy, onemocnění jater či rakovina děložního čípku.

Cílem této bakalářské práce bylo pomocí webové aplikace Palindrome analyser charakterizovat přítomnost a umístění inverzních repetit v genomech organismů, které jsou patogenní pro lidský organismus. Byly vybrány 4 viry, z nichž dva jsou ze skupiny DNA virů a dva ze skupiny RNA virů. S ohledem na propuknutí pandemie na začátku roku 2020, kterou způsobil virus SARS-CoV-2, je v této bakalářské práci zařazen. Z RNA virů byly tedy vybrány: SARS-CoV a SARS-CoV-2. Z DNA virů byly vybrány: virus Hepatitis B a lidský papilomavirus.

Sekvence virových genomů byly získány z databáze NCBI (National Center for Biotechnology). Poté byly pomocí programu Palindrome analyser, který je dostupný online, analyzovány všechny čtyři viry na přítomnost inverzních repetit, jejich polohu a velikost. Největším genomem byl SARS-CoV-2 s velikostí 29 903 bp, který měl zároveň nejvíce inverzních repetit.

ABSTRACT

Pathogens are organisms that cause various host diseases. These include prions, viruses, bacteria, fungi, protozoa and animals. This bachelor thesis is focused specifically on viruses causing human diseases such as severe respiratory syndromes, liver diseases or cervical cancer.

The aim of this bachelor thesis was to characterize the presence and location of inverted repeats in the genomes of organism using the web application Palindrome analyzer. Four viruses were selected, two of them are from the group of DNA viruses and two from the group of RNA viruses. In view of the outbreak of a pandemic in early 2020 caused by virus SARS-CoV-2, is included in this bachelor thesis. Thus, SARS-CoV and SARS-CoV-2 were selected from RNA viruses and hepatitis B virus and human papillomavirus were selected from the DNA viruses.

The sequences of the viral genomes were obtained from the NCBI (National Center for Biotechnology) database. Then, all four viruses were analyzed for the presence of inverse repeats, their location and size using the Palindrome analyzer, which is available online. The largest genome was SARS-CoV-2 of 29 903 bp, which also had the most inverse repeats.

KLÍČOVÁ SLOVA

Inverzní repetice, křížové struktury, patogeny, viry

KEYWORDS

Inverted repeats, cruciforms, pathogens, viruses

HANZLÍKOVÁ, Anna. *Analýza inverzních repetit v genomu lidských patogenů*. Brno, 2020. Dostupné také z: <https://www.vutbr.cz/studenti/zav-prace/detail/124252>. Bakalářská práce. Vysoké učení technické v Brně, Fakulta chemická, Ústav fyzikální a spotřební chemie. Vedoucí práce Václav Brázda.

PROHLÁŠENÍ

Prohlašuji, že jsem bakalářskou práci vypracovala samostatně a že všechny použité literární zdroje jsem správně a úplně citovala. Bakalářská práce je z hlediska obsahu majetkem Fakulty chemické VUT v Brně a může být využita ke komerčním účelům jen se souhlasem vedoucího bakalářské práce a děkana FCH VUT.

.....
podpis studentky

PODĚKOVÁNÍ

Ráda bych poděkovala svému vedoucímu práce, doc. Mgr. Václavovi Brázdovi Ph.D., za odborné vedení, cenné rady, ochotu mít se mnou občas trpělivost a čas, který mi věnoval na konzultacích. Taky bych chtěla poděkovat Ing. Otílii Porubiakové za pomoc při zpracovávání bakalářské práce.

OBSAH

Obsah.....	5
1 Úvod.....	6
2 Teoretická část.....	7
2.1 Patogeny.....	7
2.1.1 Fakultativní patogeny	7
2.1.2 Obligátní patogeny	7
2.2 Priony.....	7
2.3 Viry	8
2.3.1 Historie zkoumání virů	9
2.3.2 Baltimorova klasifikace	9
2.3.3 Hepatitis B virus	11
2.3.4 Lidský papilomavirus	12
2.3.5 SARS-CoV	13
2.3.6 SARS-CoV-2	15
2.4 Bakterie	17
2.4.1 Historie poznávání bakterie	17
2.4.2 Struktura bakterie	17
2.4.3 Patogenita bakterií	18
2.5 Křížové struktury	18
2.6 Palindrome analyser	19
3 Experimentální část	21
3.1 Metody	21
3.1.1 Palindrome analyser	21
4 Výsledky a diskuze.....	23
4.1 SARS-CoV.....	23
4.2 SARS-CoV-2	24
4.3 Hepatitis B virus	25
4.4 Lidský papilomavirus typu 16	27
5 Závěr.....	29
6 Seznam použitých zdrojů	30
7 Seznam použitých zkratk a symbolů	36

1 ÚVOD

Počet virů a bakterií na Zemi je ohromující, zabírají v podstatě každé prostředí. Drtivá většina virů a bakterií nemá jakýkoliv negativní vliv na lidské zdraví, jelikož lidské tělo je složitý a prosperující ekosystém [1]. Lidské tělo obsahuje přibližně 10^{13} lidských buněk a asi 10^{14} bakteriálních, fungálních a protozoálních buněk, které představují tisíce mikrobiálních druhů. Tyto mikroby nazýváme „normální mikroflóra“ a jsou obvykle omezeny na konkrétních částech těla, včetně kůže a úst. Konkrétně asi jeden z miliardy mikrobiálních druhů je lidským patogenem. V současné době již bylo popsáno přibližně 1400 lidských patogenů, přičemž odhadovaný počet existujících mikrobiálních druhů patogenů na Zemi je bilion [2].

Vzhledem k tomu, že na začátku roku 2020 probíhala „koronavirová krize“ v nejméně 188 zemích po celém světě, tak v této práci byla věnována analýze a porovnání genomů dvou koronavirů s dalšími dvěma DNA viry. Prvním z koronavirů je SARS-CoV, virus který způsobuje onemocnění zvané SARS (z angl. *Severe Acute Respiratory Syndrome*). Toto onemocnění se projevilo v provincii Kuang-tung v Číně na konci roku 2002. Jedná se o onemocnění postihující zejména dolní cesty dýchací, které vede k závažnému onemocnění plic [3].

Druhým je SARS-CoV-2, který byl poprvé pozorován na konci roku 2019 v čínském městě Wu-chan. Na jaře roku 2020 bylo onemocnění způsobené tímto virem označeno za pandemii. Tak jako SARS-CoV postihuje zejména dolní cesty dýchací [4].

Dalším virem je virus Hepatitidy B, který se od ostatním virů hepatitidy liší tím, že se jedná o DNA vir. Je běžnou příčinou širokého spektra onemocnění jater, od akutní po chronickou hepatitidu, cirhózu či rakovinu jater [5].

Posledním analyzovaným virem je lidský papilomavirus, který způsobuje infekci HPV. Existuje více než 150 typů lidských papilomavirů, avšak jen část je pro lidský organismus nebezpečná. Například genotypy HPV 16 a 18 jsou zodpovědné za cca 70 % rakoviny děložního čípku u žen [6].

2 TEORETICKÁ ČÁST

2.1 Patogeny

Patogen je organismus, který způsobuje onemocnění hostiteli, přičemž závažnost symptomů onemocnění se označuje jako virulence. Stupeň virulence souvisí přímo se schopností organismu způsobovat onemocnění navzdory mechanismům odolnosti hostitele. Je ovlivňován řadou faktorů, jako je množství patogenních organismů, cesta vstupu do těla hostitele či obranné mechanismy hostitele [1]. Termín patogen se začal používat už v roce 1880. Patogeny mohou být priony, viry, bakterie, prvoci, houby a živočichové. Studium, které se zabývá nemocemi a může zahrnovat tyto patogeny, se nazývá patologie [2]. Každý živý organismus je ovlivněn patogeny, včetně bakterií, na které se zaměřují specializované viry nazývané fágy [7]. Patogeny se rozdělují na fakultativní (oportunní) a obligátní (primární) [1].

2.1.1 Fakultativní patogeny

Fakultativní patogeny jsou organismy, u nichž je hostitel pouze jedním z míst, kde se může reprodukovat. Fakultativní patogeny jsou především bakterie a houby, které příležitostně způsobují infekce. Infikují pouze oslabené nebo imunokompromitované hostitele. Oslabení hostitele může být způsobeno poraněním kůže, sliznice, či po lékařských zákrocích. Mezi fakultativní patogeny patří např. *Escherichia coli*, či *Neisseria meningitidis* [8].

2.1.2 Obligátní patogeny

Obligátní patogeny vyžadují více hostitelů k dokončení jejich životního cyklu. Definitivním hostitelem, který podporuje dospělou formu patogenu, je nejčastěji obratlovec. Mezihostitelem bývá většinou členovec či měkkýš. Tohle střídání hostitelů bývá zejména u virů (virus Zika), bakterií (např. *Borrelia* způsobující Lymeovu chorobu) a prvoků (např. *Plasmodii*, která způsobuje malárii) [8]. Obligátní patogeny mají schopnost vyvolat onemocnění i u zdravých jedinců [7].

2.2 Priony

Jedná se o nejjednodušší infekční částice tvořené pouze molekulou bílkoviny, jež způsobují neurodegenerativní onemocnění savců, tzv. transmisivní spongiformní encefalopatie (TSE). Prusinerův objev prionového proteinu (PrP) v 80. letech 20. století významně urychlil poznávání biologie a patogeneze TSE [9]. Slovo jako takové pochází z ‘proteinaceous infectious particle’ a tento termín odkazuje na původně kacírskou hypotézu, že infekční částice, které způsobují tato onemocnění, sestávají pouze z proteinu. Všechny předtím známé patogeny (viry, bakterie) obsahují nukleové kyseliny, které jim umožňují reprodukci [10].

TSE od ostatních neurodegenerativních onemocnění, jakými jsou např. Parkinsonova či Alzheimerova choroba, odděluje jedna vlastnost, a tou je přenosnost (transmisivnost) [9]. Přenos z jednoho organismu na druhý je velice snadný v rámci stejného druhu, ale je možný i mezi druhy různými. TSE jsou těžké, bez výjimky smrtelné choroby. Jsou charakterizovány asymptomatickou inkubační dobou, která v případě člověka může dosahovat až několik desítek let. Onemocnění většinou probíhají pomalu a jsou nevyléčitelná. Některé orgány, především

mozek, mícha a sítnice jsou těžce infikovány již dříve, než se objeví klinické projevy dané nemoci. Mezi klinické příznaky patří poruchy či ztráta paměti, projevy motorických dysfunkcí (poruchy chůze nebo rovnováhy), změny v chování či demence [11].

PrP je glykoprotein, jehož izoforma nazývaná buněčný prionový protein (PrP^C) je normální součástí živočišných těl. PrP však může existovat také ve formě abnormálního konformačního izomeru (PrP^{Sc}), který je infekční a u savců patrně způsobuje TSE. Mezi TSE patří tzv. nemoc šílených krav, scrapie ovcí a Creutzfeldt-Jakobova choroba u lidí [12].

Stejně jako u savců, byly i v houbách nalezeny proteiny se samovolně se rozmnožujícími konformacemi. Tyto fungální priony sdílejí mnoho podobností se savčími priony. Díky snadné genetické manipulaci a rychlému růstu plísni mohl postupovat plísňový prionový výzkum rychlým tempem [10]. Mezi kvasinkovými a savčími priony existuje plno rozdílů, například kvasinkové priony nemusí způsobovat onemocnění a ani žádné nemusí přenášet [12].

2.3 Viry

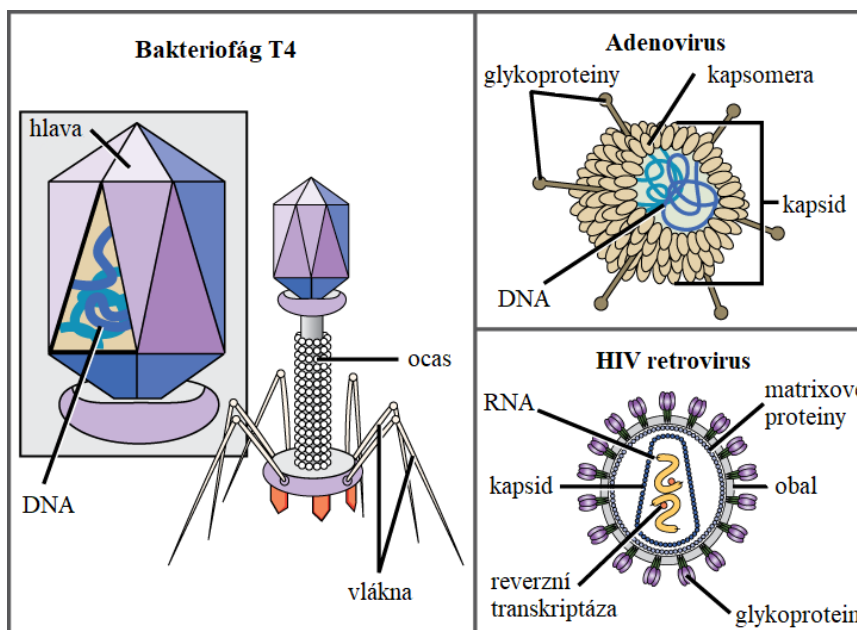
Viry jsou malé částice, které se samy nemohou reprodukovat [13]. Většina virů má jako svůj genetický materiál buď RNA nebo DNA, jiné mají genomy, které mají více nukleových kyselin a tvoří segmenty [14]. Celá částice infekčního viru, nazývaná virion, se skládá z nukleové kyseliny a vnějšího obalu proteinu [13]. Na Obrázku 1 lze vidět uspořádání jednotlivých virů, jako je bakteriofág T4, adenovirus, který patří mezi neobalené dsDNA viry, a HIV retrovirus [14].

Viry musí vstoupit i vystoupit ze své hostitelské buňky, aby dokončily cyklus replikace [13]. Jelikož plazmatická membrána působí jako bariéra mezi buňkami a jejich prostředím, viry byly nuceny vyvinout strategii, která by jim umožnila překonat tuto lipidovou dvojrivrstvu. V případě mnoha obalených virů, vstup do hostitelské buňky na začátku replikace zahrnuje membránové fúzní či štěpné reakce mezi obalem viru a plazmatickou membránou hostitelské buňky [15].

V lytickém cyklu se bakteriofág připojí na receptor bakteriální buňky, který je specifický. Buněčná stěna se natráví, proděraví a DNA je z hlavy fágu vypuzena do buňky. Uvnitř hostitelské buňky se virová DNA přepíše do specifické virové mRNA (transkripce) a na ribozomech se začne tvořit bílkovina k tvorbě kapsidu (translace). Následně proběhne replikace virové DNA, která kondenzuje a nastane kompletace virionů neboli maturace. Hostitelská buňka je zničena a uvolněné viriony napadají další buňky [13].

Při lyzogenním cyklu se virová DNA nereplikuje, ale začlení se do chromozomu hostitelské buňky. Takto se pak přenáší do dceřiných buněk. Hostitelská buňka obsahující tzv. profága je imunní vůči infekci stejným fágem a označuje se jako lyzogenní. Kdykoliv je možná aktivace proviru, např. infekcí či UV zářením, spustí se lytický cyklus. Profág se vyčlení z chromozomu, replikuje se a hostitelská buňka se rozpadá [13].

Mezi choroby způsobené DNA viry patří plané i pravé neštovice, hepatitida B, opary na rtech a některé pohlavní choroby: genitální opar a genitální bradavice. Na druhou stranu mezi onemocnění způsobené RNA viry patří např.: hepatitida A i C, spalničky, vzteklna, klíšťová encefalitida a chřipka [14].



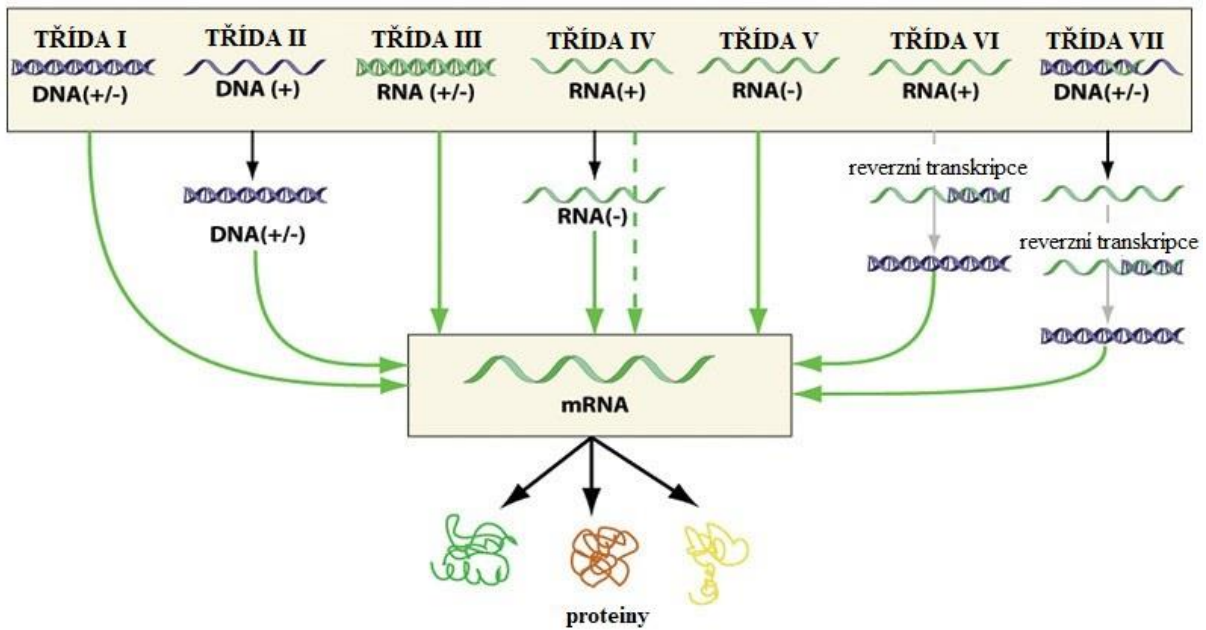
Obrázek 1: Složité viriony - bakteriofág T4, adenovirus a HIV retrovirus [14]

2.3.1 Historie zkoumání virů

Viry byly poprvé objeveny po vývoji porcelánového filtru, nazývaného Chamberlandův-Pasteurův filtr, který mohl odstranit všechny viditelné bakterie v mikroskopu v jakémkoliv kapalném vzorku. V roce 1886 Adolph Meyer prokázal, že onemocnění tabákové mozaiky může být přeneseno z nemocné rostliny na zdravou rostlinu pomocí tekutých rostlinných extraktů. O 6 let později, tedy v roce 1892 Dmitri Ivanowski prokázal, že toto onemocnění lze přenášet i potom, co Chamberland-Pasteurův filtr odstraní z extraktu všechny životaschopné bakterie [16]. Stále to bylo mnoho předtím, než bylo prokázáno, že tyto „nepropustné“ infekční agens nebyly jednoduše velmi malé bakterie, ale byly novým typem velmi malých částic způsobujících onemocnění [13].

2.3.2 Baltimorova klasifikace

Baltimorova klasifikace, viz Obrázek 2, je systém klasifikace virů podle typu genetického materiálu, který je obsažen ve virových částicích. Byla poprvé navržena v roce 1971 Davidem Baltimorem, pozdějším nositelem Nobelovy ceny za fyziologii a medicínu [14]. Jedná se o jednoduchou a běžně používanou alternativu k systému ICTV (Mezinárodní výbor pro klasifikaci virů, z angl. International Committee on Taxonomy of Viruses), která člení viry podle více hledisek. V původním návrhu Davida Baltimorea bylo šest skupin virů (I. až VI.) – dnes se do sedmé skupiny vyčlenila později objevená skupina hepadnavirů s unikátně stavěným genomem. Systém umožňuje zjednodušit si komplikované a rozmanité typy životních cyklů, jimiž viry oplývají [17].



Obrázek 2: Baltimorova klasifikace virů podle genetického materiálu [18]

2.3.2.1 Třídy I-II

Různé typy virů DNA se běžně používají ve studiích o replikaci DNA, tvorbě mRNA a transformaci onkogenních buněk [15]. Viry třídy I obsahují jednu molekulu dvouvláknové DNA (dsDNA). V případě nejběžnějšího typu zvířecího viru I. třídy vstupuje virová DNA do buněčného jádra, kde buněčné enzymy přepisují DNA a výslednou RNA zpracovávají na virovou mRNA [13]. Mezi příklady těchto virů patří: adenoviry, SV40, herpesviry, papillomaviry a mnoho dalších [14]. Viry II. skupiny mají jako genom jednovláknovou DNA (ssDNA). Předtím, než může dojít k transkripci na mRNA, převádějí své jednovláknové genomy na meziproduct dsDNA [13]. Mezi viry této skupiny patří např. parvoviry [18].

2.3.2.2 Třídy III-V

Všechny živočišné viry, patřící do třídy III–VI mají RNA genomy. Celá řada zvířat, od hmyzu po člověka, je těmito viry infikována. Tyto viry byly velice užitečné při studiích syntézy a translace mRNA, při tvorbě membrány a intracelulárního transportu [13]. Viry skupiny IV mají jako genom ssRNA s pozitivní polaritou. Pozitivní polarita znamená, že genomická RNA může sloužit přímo jako mRNA. Naopak viry skupiny V obsahují genomy ssRNA se zápornou polaritou, což znamená, že jejich sekvence je komplementární k mRNA [14]. Mezi zástupce III. třídy patří reoviry nebo birnaviry [18], v IV. třídě se nachází flaviviry, coronaviry a picornaviry. Do V. třídě můžeme zařadit orthomyxoviry, arenaviry a plno dalších [14].

2.3.2.3 Třída VI

Viry třídy VI jsou obalené viry, jejichž genom se skládá ze dvou identických plus řetězců RNA. Tyto viry jsou známé jako retroviry, protože jejich RNA genom řídí tvorbu molekuly DNA. Molekula DNA pak nakonec funguje jako templát pro syntézu virové mRNA [13]. Některé retroviry obsahují geny způsobující rakovinu (nazývané onkogeny). Buňky infikované

takovými retroviry jsou onkogeně transformovány do nádorových buněk. Mezi známými lidskými retroviry jsou lidský T-buněčný lymfotrofický tanus, který způsobuje leukémii a lidský imunodeficienci virus, který způsobuje nedostatečnou imunitu [14].

2.3.2.4 Třída VII

Viry skupiny VII mají parciální dsDNA genomy a vytvářejí meziprodukty ssRNA, které působí jako mRNA, ale jsou převedeny zpět do genomů dsDNA pomocí reverzní transkriptázy, což je nezbytné pro replikaci genomu [18]. Nejznámějším zástupcem hepadnavirů je virus hepatitidy B (HBV) [14].

2.3.3 Hepatitis B virus

Virus hepatitidy B (HBV) je malý virus DNA s neobvyklými rysy, které jsou podobné retrovirům. Jedná se o virus z rodu Orthohepadnavirus, který je součástí čeledi Hepadnaviridae. Podobné viry se vyskytují u savců, např. u svišťů, veverek, Pekingských kachen a jiných. Na základě porovnání je HBV klasifikován do osmi genotypů A až H. Každý genotyp má odlišné geografické rozdělení [19]. HBV infikuje více než 300 milionů lidí na celém světě a je běžnou příčinou onemocnění jater a rakoviny jater. Infekce HBV vede k širokému spektru onemocnění jater od akutní (včetně selhání jater) po chronickou hepatitidu, cirhózu či hepatocelulární karcinom. Inkubační doba hepatitidy B je variabilní v rozmezí 1 měsíce až půl roku [5].

Objev australského antigenu Baruchem Blumbergem a jeho kolegů, který byl později identifikován jako povrchový antigen viru hepatitidy B, byl hlavním průlomem ke zlepšení globálního zdraví. Po celá desetiletí před Blumbergovým objevem, byl neznámý virus v krvi a plazmě podezřelý z potransfuzní hepatitidy. Uznání, že australský antigen byl markerem virové hepatitidy, usnadnilo vytvoření protokolu o screeningu krve, který vedl ke dvojnásobnému až trojnásobnému snížení incidence hepatitidy po transfuzi, přičemž zbývající případy pravděpodobně byly způsobeny virem hepatitidy C (HCV), který nebyl identifikován dalších 23 let [20]. Objev australského antigenu také usnadnil vývoj vakcíny, která výrazně snížila globální břemeno infekce HBV a Baruch Blumberg získal za objev HBV Nobelovu cenu za medicínu v roce 1976 [21].

2.3.3.1 Struktura částice HBV

V infekčním séru se pomocí elektronové mikroskopie vizualizují tři typy virových částic. Dvě z virových částic jsou menší sférické struktury o průměru 20 nm a vlákna různých délek se šířkou přibližně 22 nm [19]. Obálka obsahuje malé množství lipidů buněčného původu a tři povrchové proteiny hepatitidy B: velký, střední a malý, které tvoří s disulfidem vázané homodimery a heterodimery [22]. Sféry a vlákna jsou složeny z povrchového antigenu hepatitidy B a lipidů odvozených od hostitele bez virových nukleových kyselin, a proto jsou neinfekční. Infekční virion HBV (Daneova částice) má kulovitou strukturu s dvojitou skořepinou o průměru 42-47 nm, sestávající z lipidové obálky obsahující HBsAg, která obklopuje vnitřní ikoseadrický (dvacetistěnný) nukleokapsid složený z jádrového antigenu hepatitidy B (HBcAg) komplexovaného s virově kódovanou polymerázou a genomem virové DNA. Genom HBV je částečně dvouvláknová kruhová DNA o párech asi 3,2 kb [19].

2.3.3.2 Čtecí rámce

Virový genom kóduje čtyři překrývající se otevřené čtecí rámce (ORF: S, C, P a X). S ORF kóduje virové povrchové obalové proteiny (HBsAg) a lze je strukturálně a funkčně dělit do oblastí pre-S1, pre-S2 a S. Mnoho iniciačních kodonů translace je znakem genů S a C, které vedou k příbuzným, ale funkčně odlišným proteinům. C ORF kóduje buďto virový nukleokapsid HBcAg nebo antigen hepatitidy B (HBeAg) v závislosti na tom, zda je translace iniciována z jádra nebo precore oblasti. Jádrový protein má tu vlastnost, že se sám sestaví do kapsidovité struktury a na svém C-konci má vysoce bazické aminokyseliny s vazebnou aktivitou k RNA [23].

Polymeráza je velký protein (přibližně 800 AMK), který je kódovaný P ORF a je funkčně rozdělen do tří domén: terminální proteinová doména, doména reverzní transkriptázy a doména ribonukleázy H. X ORF kóduje 16,5 kDa protein (HBxAg) s mnoha funkcemi, včetně přenosu signálu, transkripční aktivace, opravy DNA a inhibice degradace proteinu. Mechanismus této aktivity a biologické funkce je do značné míry pořád neznámý. Je však prokázáno, že HBxAg je nezbytný pro produkci infekce HBV [24].

2.3.4 Lidský papilomavirus

S prevalencí 70 milionů případů a každoročním výskytem 14 milionů nových přenosů jsou infekce HPV anogenitálního traktu nejčastějšími pohlavně přenosnými chorobami v USA. Přibližně 80 % žen ve věku do 50 let tuto infekci již získalo. Celosvětově se odhaduje, že se každoročně vyskytne 500 000 případů rakoviny děložního čípku, což má za následek i více než 250 000 úmrtí [6]. V roce 2018 byla rakovina čípku 4. nejčastější rakovinou u žen a celkově 8. nejčastější rakovinou vůbec [25]. Břemeno rakoviny děložního čípku je v zemích s nízkými příjmy neúměrně vysoké kvůli nedostatku zdrojů na očkování a léčby. I když jsou k dispozici bezpečné a účinné vakcíny pro prevenci infekce HPV, nechrání však ty, kteří již byly infikováni, a nechrání před všemi typy HPV [26].

Rizikové faktory: Kouření, sexuální život (první pohlavní styk, počet sexuálních partnerů), užívání perorálních kontraceptiv a UV záření [27].

2.3.4.1 Struktura HPV

Lidský papilomavirus (HPV) je neobalený, dvouřetězcový, cirkulární DNA virus z rodiny Papillomaviridae [27]. Existuje více než 150 typů, přičemž genotypy HPV 16 a 18 jsou zodpovědné za přibližně 70 % rakovin děložního čípku. Zbývajících 30 % je způsobeno jinými karcinogenními typy HPV. HPV16 je nejvíce karcinogenní typ z hlediska počtu případů cervikální intraepiteliální neoplasie (CIN) nebo rakoviny děložního čípku. Nízkorizikové genotypy 6 a 11 způsobují většinou genitálních bradavic [6]. Virus vstupuje do epitelu narušením kůže či sliznice, a infikuje bazální kmenové buňky. Virová DNA může zůstat po určitou dobu nezávislým epizomem, než se začlení do genomu hostitele. HPV se přednostně integruje na křehkých místech v lidské DNA, kde je řetězec náchylný ke zlomení [27].

Navzdory různým asociacím nemocí mají částice papilomaviru společnou neobalenou ikosedráltní strukturu o průměru 50 až 60 nm. Jejich genomy obsahují dvouřetězcové kruhy (epizomy) o přibližně 8 000 párů bází, které obsahují osm nebo devět ORF. Ačkoliv je počet genů omezen malou velikostí genomu, počet kódovaných proteinů je mnohem větší, protože

genová exprese zahrnuje použití více promotorů a složitých vzorců pro sestřih. Mapování jemných struktur ukazuje, že virový plášť obsahuje 360 molekul proteinu L1 uspořádaných do 72 kapsomerů, z nichž každá je tvořena 5 molekulami L1 [28]. Částice lidského papilomaviru také obsahují variabilní počet molekul L2, které nejsou úplně exponovány na povrchu virionu, kromě jejich N-terminálních 120 aminokyselin. Během infekce je L2 rozštěpen furinem během infekce. Virový genom také kóduje regulační proteiny, které stimulují vstup do buněčného cyklu a proliferaci buněk, a také proteiny, které zprostředkovávají replikaci genomu viru, sestavení viru a pravděpodobně také účinné uvolňování a přenos viru [29].

2.3.4.2 Rozdělení lidských papilomavirů

Lidské typy jsou rozděleny do pěti rodů na základě rozdílů v jejich DNA sekvenci, přičemž jednotlivé typy mají nukleotidovou sekvenci, která je nejméně z 10 % odlišná od ostatních sekvencí papilomavirů. V posledních letech umožňují citlivé detekční metody identifikovat velký počet nových typů HPV (především typů beta a gama) z odebraných výtěrů kožního epitelu, nebo z vytrhaných vlasů. Počet typů beta se téměř zdvojnásobil (z 25 na 45) a počet typů gama se za poslední dekádu zvýšil téměř osmkrát (ze 7 na 54) [28].

2.3.5 SARS-CoV

SARS-CoV způsobuje závažný akutní respirační syndrom (z ang. Severe Acute Respiratory Syndrome). SARS-CoV je relativně nové infekční virové onemocnění, které je doprovázeno závažnými klinickými projevy dolních cest dýchacích. Patogeneze SARS-CoV je velice komplexní, s více faktory, které vedou k těžkému poškození plic. V průběhu nemoci může být infikováno více buněk či orgánů, včetně epitelových buněk ledvin, neuronů v mozku a několika typů imunitních buněk. Na Obrázku 3 lze vidět hlavní mechanismy, které vedou k poškození plic [4].

Genom SARS-CoV je jednovláknová RNA s pozitivním nábojem sestávající z 29 751 bází. Fylogenetická analýza odhalila, že tento druh CoV je jen mírně příbuzný jiným CoV, včetně dvou lidských: HCoV-OC43 a HCoV-229E. U těchto lidských CoV je známo, že způsobují infekce dolních cest dýchacích, běžné nachlazení a průjem [30]. SARS je přenášen na člověka a mezi lidmi buďto přímým kontaktem (kůží), kapičkami nebo vzdušnými cestami. Virus byl taky izolován z několika zvířat, včetně psů mývalovitých, avšak nejsou považovány za skutečný zdroj [4].

2.3.5.1 Historie

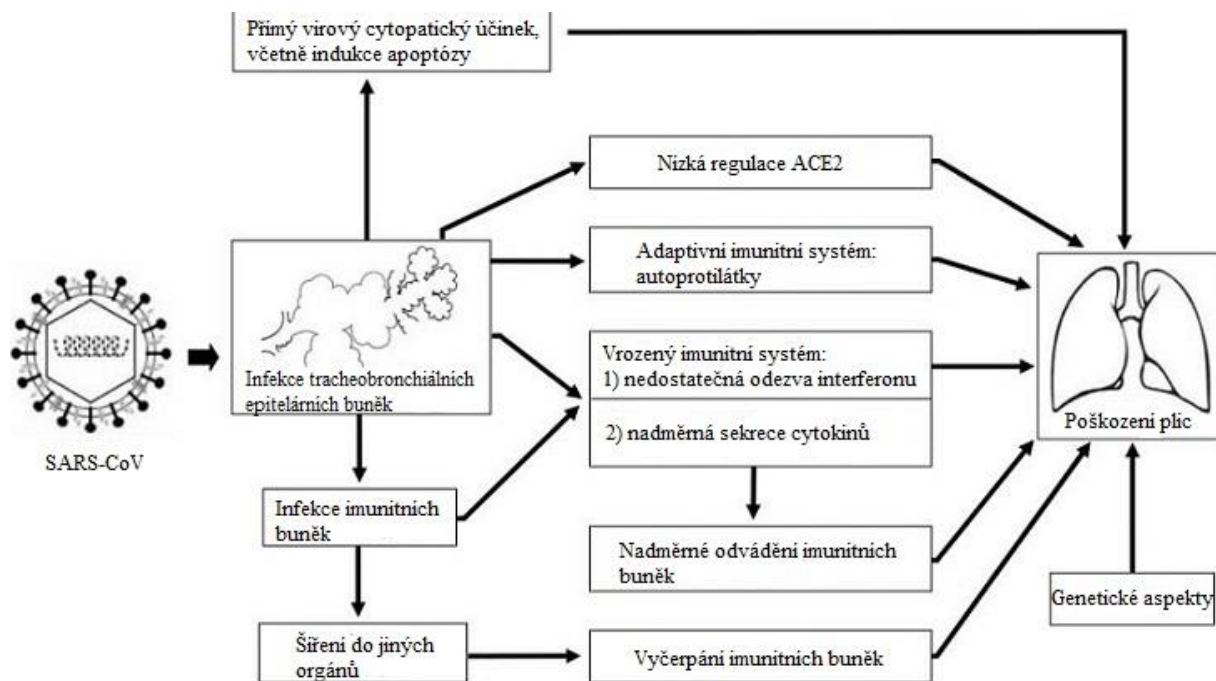
SARS-CoV se objevil v provincii Kuang-tung v Číně v listopadu roku 2002. Do března následujícího roku se toto onemocnění rozšířilo po celém světě. Do července roku 2003 bylo nahlášeno přes 8 000 případů s tímto onemocněním, včetně přibližně 900 úmrtí z 32 zemí po celém světě [31].

2.3.5.2 Detekce infekce

SARS-CoV může být detekován reverzní transkriptázovou polymerázovou řetězovou reakcí (RT-PCR) ve stolici, moči a také v krvi. Infekce tudíž není omezena pouze na dýchací cesty. Na rozdíl od většiny ostatních respiračních virů je zátěž SARS-CoV ve stolici v prvních dnech

nemoci nízká, avšak kolem 11. dne nemoci vrcholí. Takle informace odpovídá nízké citlivosti diagnostických testů v prvních několika dnech nemoci [32]. Exhalační vzorky a stolice jsou užitečné klinické vzorky pro diagnostiku RT-PCR, ale stolice jsou méně uspokojivým vzorkem během prvních 5 dnů nemoci. Nasopharyngeální aspiráty jsou vzorky dýchacích cest, ale lze použít i výtěry z krku, i když s nižším diagnostickým výnosem. Když pacienti produkují sputum (výměšek dýchacího ústrojí a vykašlaná hmota), je to také vynikající klinický vzorek, u většiny pacientů je však kašel neproduktivní. Testy RT-PCR druhé generace jsou schopné detekovat SARS-CoV u nasopharyngeálních aspirátů u přibližně 80 % pacientů s SARS-CoV během prvních 3 dnů nemoci [33].

Virus může také způsobit buněčnou fúzi vedoucí k tvorbě syncytií (mnohojaderný buněčný útvar, který vznikne rozrušením buněčných membrán několika sousedních buněk). Tyto cytopatické účinky jsou způsobeny kroky virové replikace, jako je mobilizace vezikul za vzniku virového replikačního komplexu, což vede k narušení Golgiho komplexů. Souběžně s výsledky na jiné koronaviry bylo prokázáno, že SARS-CoV způsobuje tvorbu syncytií v plicních tkáních. Další podobností s jinými koronaviry je potenciál SARS-CoV způsobit fibrózu tkáně. Jako molekulární mechanismus pro tuto fibrózu bylo prokázáno, že N protein indukuje promotorovou aktivitu genu protrombinázy, který koreluje s ukládáním fibrinu [34].



Obrázek 3: Hlavní mechanismy přispívající k patogenezi SARS [4]

Některé orgány, jako jsou plíce a střevo, byly studovány a patologické léze v těchto orgánech jsou velice dobře známy. V dýchacích cestách se jedná o difúzní alveolární poškození s různými stupni akutních exsudativních rysů, včetně otoků a hyalinních membrán. Makrofágová či smíšená buněčná infiltrace, vícejaderné obří buňky, atypické reaktivní pneumocyty a cévní poškození. Pozitivní in situ hybridizační signály v pneumocytech, lymfocytech a makrofázích [35].

Ve střevech byla vyčerpána lymfoidní tkáň (specializovaná tkáň, kde se tvoří, zrají a množí

lymfocyty). Pozitivní in situ hybridizační signály v mukózních epiteliálních buňkách. V játrech byla pozorována v některých případech nekroza a našly se důkazy apoptózy [36]. Dále byla pozorována deplece lymfocytů ve slezině a lymfatických uzlinách s architektonickým narušením. Splenicá atrofie bílé buničiny a opět pozitivní in situ hybridizační signály v imunitních buňkách [35].

2.3.6 SARS-CoV-2

2.3.6.1 Dřívější β -koronaviry

V posledních dvou desetiletích došlo k dvěma událostem, kdy zkřížení živočišných β -koronavirů na člověka mělo za následek vážné onemocnění. První doložený případ byl v letech 2002–2003, kdy nový β -koronavirus s původem v netopýrech přešel na člověka prostřednictvím zprostředkujícího hostitele, a to koček z palmy civet v čínské provincii Guangdong. Tento virus označený jako závažný akutní respirační syndrom koronavirus postihl 8 422 lidí převážně v Číně a Hongkongu a způsobil 916 úmrtí (úmrtnost 11 %) [37]. O deset let později, tedy v roce 2012 se v Saudské Arábii objevil respirační syndrom na Středním východě (MERS-CoV), který byl také netopýřího původu. Mezihostitelem tohoto respiračního syndromu se stali velbloudi a zasáhl 2 494 lidí, přičemž způsobil 858 úmrtí (úmrtnost 34 %) [38].

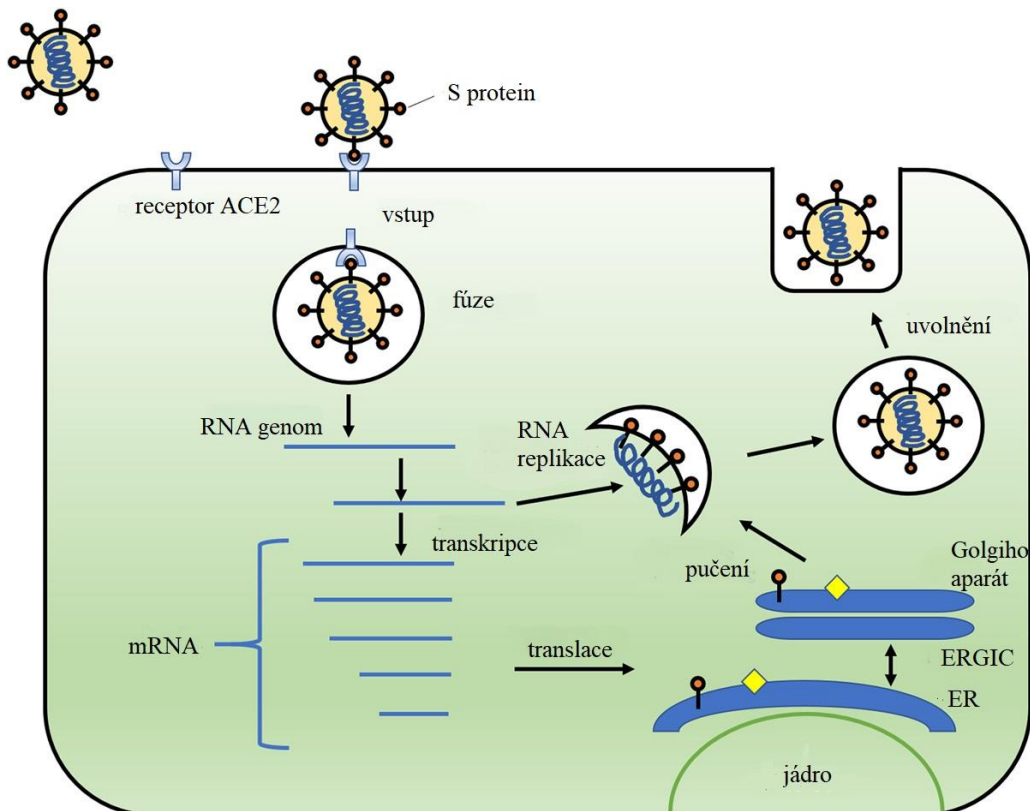
2.3.6.2 Rozdělení koronavirů

Koronaviry jsou rozděleny do čtyř skupin: alfa, beta, gama a delta. Alfa a beta koronaviry jsou schopny infikovat savce, zatímco gama a delta koronaviry mají tendenci infikovat ptáky. SARS-CoV-2 patří do kategorie betaCoV. Dříve bylo identifikováno šest CoV, které mohou infikovat člověka, mezi nimiž je α -CoVs: HCoV-229E, HCoV-NL63 a β -CoVs: HCoV-HKU1 a HCoV-OC43. Tyto koronaviry s nízkou patogenitou způsobují mírné respirační symptomy podobné běžnému nachlazení. Další β -koronaviry, SARS-CoV a MERS-CoV, vedou k závažným a potenciálně fatálním infekcím dýchacích cest [3].

2.3.6.3 Struktura a mechanismy

SARS-CoV-2 má kulatý nebo eliptický, často pleomorfní (mnohotvarný) tvar a průměr přibližně 60–140 nm. Jeho jednořetězcový RNA genom obsahuje 29 891 nukleotidů kódujících 9 860 aminokyselin. Stejně jako ostatní, CoV je citlivý na ultrafialové záření a teplo. Kromě toho mohou být tyto viry účinně inaktivovány lipidovými rozpouštědly obsahujícími ether (75%), ethanol, dezinfekčními prostředky obsahující chlor, kyselinu peroxyoctovou a chloroform [39]. Z genetického hlediska Chan a kol. prokázali, že genom nového lidského koronaviru, izolovaného z pacienta s atypickou pneumonií po návštěvě Wu-chanu, měl 89% nukleotidovou identitu s genomem netopýra SARS-CoVZXC21 a 82% identitu s genomem lidského SARS-CoV [40]. SARS-CoV-2 by mohl být přenášen z netopýrů prostřednictvím neznámých zprostředkujících hostitelů na lidi. Jeho původ není prozatím zcela pochopen. Nyní je ale zřejmé, že by tento vir mohl k infikování lidí použít stejný enzym konvertující angiotensin 2 (ACE2), což je stejný receptor jako u SARS-CoV [41]. Na Obrázku 4 je vidět životní cyklus SARS-CoV-2, v němž mu při vstupu pomáhá receptor ACE2, dále musí projít transkripcí, translací, replikací RNA a pučením.

Koronaviry jsou obalené RNA viry s pozitivním řetězcem s nukleokapsidem. Pro řešení patogenetických mechanismů SARS-CoV-2 je nutné vzít v úvahu jeho virovou strukturu a genom. V CoV je genomická struktura organizována v + ssRNA o délce přibližně 30 kb. Transkripce funguje prostřednictvím replikačně-transkripčního komplexu (RCT) organizovaného ve dvojmembránových vezikulách a syntézou sekvencí subgenomických RNA (sgRNA). K transkripčnímu ukončení dochází u regulačních sekvencí pro transkripci, umístěných mezi takzvanými otevřenými čtecími rámci (ORF), které fungují jako templáty pro produkci subgenomických mRNA. V atypickém CoV genomu musí být přítomno alespoň šest ORF [39]. Protein S je důležitý pro SARS-CoV-2. S protein zprostředkovává vazbu na receptor a membránovou fúzi a je rozhodující pro stanovení hostitelského tropismu a přenosové kapacity [42]. Protein S je funkčně rozdělen do domény S1, která je zodpovědná za vazbu receptoru. S2 doména je zodpovědná za fúzi buněčné membrány [43]. Analýza struktury naznačila, že doména vázající receptor byla složena z jádra a externí subdomény [42].



Obrázek 4: Schématický model životního cyklu COVID-19 [44]

2.3.6.4 Infekce

Infekce je přenášena velkými kapičkami, které jsou generovány během kašle a kýchlání symptomatickými pacienty, ale může se vyskytnout také u asymptomatických lidí nebo před nástupem příznaků. Pacienti mohou být infekční tak dlouho, dokud příznaky přetrvávají, a dokonce i po klinickém zotavení. Infikované kapičky se mohou šířit 1–2 metry daleko a mohou se ukládat na povrchy. Virus může zůstat životaschopný na površích několik dní, pokud k tomu má příznivé podmínky, avšak běžnými dezinfekčními prostředky, jako

je chlornan sodný, peroxid vodíku atd., je lze zcela zničit během minuty [45]. Infekce je získána buďto inhalací těchto kapiček nebo dotykem s kontaminovanými povrchy, a poté dotykem nosu, úst a očí. Podle současných informací nebyl transplacentální přenos z těhotných žen na jejich plod popsán [44]. Je však popsáno novorozenecké onemocnění způsobené postnatálním přenosem. Inkubační doba se pohybuje od 2 do 14 dní [45].

2.4 Bakterie

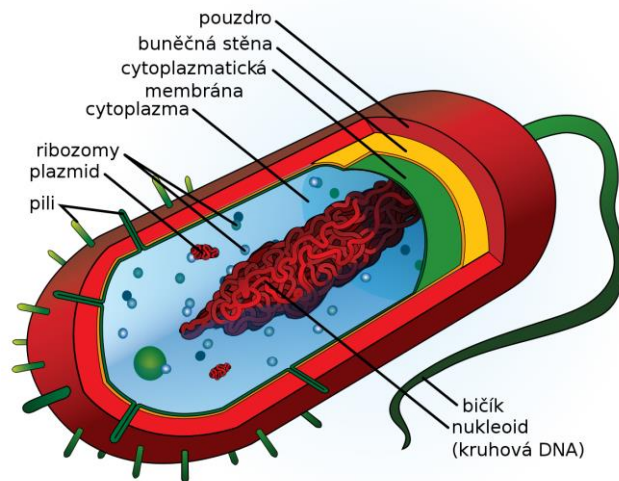
Bakterie jsou jednobuněčné prokaryotické mikroorganismy, které postrádají jadernou membránu. Jsou metabolicky aktivní a dělí se štěpením. Mají zejména kokovitý či tyčinkovitý tvar a zpravidla dosahují velikosti v řádu několika mikrometrů. Studium bakterií se zabývá bakteriologie. Bakterie jsou všude kolem nás, lze je najít od vrcholů hor v Antarktidě po hlubinné podloží v oceánu [46].

2.4.1 Historie poznávání bakterie

Prvním doloženým člověkem, který pozoroval bakterie, byl nizozemský přírodovědec Antoni van Leeuwenhoek žijící v 17. století. Bakterie poprvé pozoroval v roce 1676, díky jednočočkovému mikroskopu vlastní výroby. Během 19. století německý biolog Ferdinand Cohn jako vůbec první přišel se systémem klasifikace bakterií, navíc se mu podařilo objevit bakteriální spory u rodu *Bacillus* [47]. Jiným velkým vědcem té doby byl Louis Pasteur. Ten je znám jako otec pasterace, techniky likvidace bakterií teplem. Pasteur také dokázal, že bakterie nevznikají spontánně z neživé hmoty a prosazoval názor, že jsou odpovědné za nemoci [48]. Vědecké bádání jiného vědce potvrdilo, že se Pasteur nemýlil. Šlo o německého vědce a mikrobiologa Roberta Kocha pozdějšího nositele Nobelovy ceny. Ten mimo jiné přišel s postuláty, které pomáhají identifikovat patogen odpovědný za danou nemoc [49].

2.4.2 Struktura bakterie

Běžné formy bakterií se skládají z několika základních částí, jejichž schéma je vyobrazeno na Obrázku 5. Cytoplazmatická membrána je velmi tenká polopropustná obalová část, která je tvořena ze specifického buněčného stavebního materiálu tzv. fosfolipidů. Tato biomembrána slouží jako ochranné pásmo, které zprostředkovává transport látek do a z buňky. Tento základní biochemický cyklus se uskutečňuje za pomoci transportních proteinů. Vnější část bakterie je tvořena pevnou buněčnou vrstvou, jejíž hlavní složkou je peptidoglyken [50]. Jedná se o chemickou látku běžně se řadící do kategorie glykosidů. Právě tato organická sloučenina dokáže perfektně zpevnit buňku. Zabraňuje jejímu poškození a propouští většinu důležitých látek. Vevnitř jednobuněčného organismus se nachází jádro prokaryotické buňky, jenž nemá stálý tvar a standardně jej vytváří jediná molekula DNA, která obsahuje genetickou informaci. Tento základní buněčný komponent zaujímá zhruba patnáct procent z celkového objemu buňky [2]. Uvnitř samotné buňky se nalézá základní hmota zvaná cytoplazma, která obsahuje kromě vody, také bílkovinnou složku i další rozpuštěné látky. V cytoplazmě se nachází obrovské množství cytoplazmatických útvarů, jako jsou například ribozomy. Právě v těchto speciálních částicích probíhá produkce vysokomolekulárních látek přírodní povahy takzvaných bílkovin [51].



Obrázek 5: Schéma prokaryotické buňky - bakterie [52]

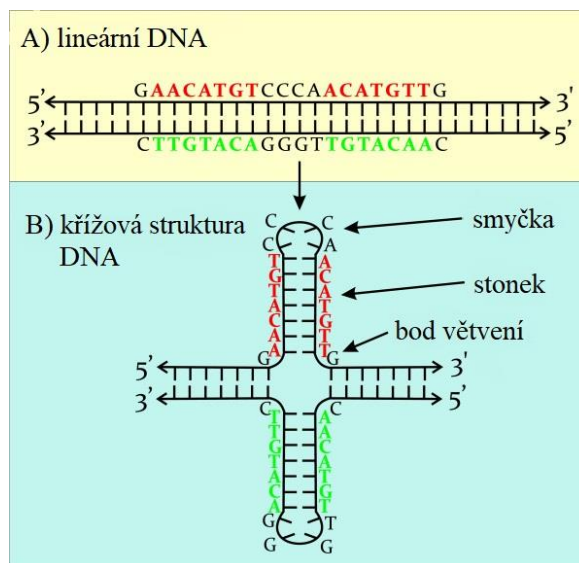
2.4.3 Patogenita bakterií

Bakterie způsobují řadu různých infekcí, které se mohou pohybovat v závažnosti od neznámých po fulminující. Schopnost bakterie způsobovat onemocnění odráží její relativní patogenitu, díky tomuto mohou být bakterie rozděleny do tří skupin. Oportunistické patogeny jsou izolované patogeny od pacientů, jejichž obranné mechanismy hostitele byly narušeny [50]. Poslední skupinou jsou nepatogeny, protože zřídka či vůbec nezpůsobují lidské onemocnění. Jejich kategorizace se však může změnit z důvodu přizpůsobivosti bakterií a škodlivého účinku moderních terapií, chemoterapií či imunoterapií [1].

Citlivost bakteriální infekce závisí na fyziologickém a imunologickém stavu hostitele a na virulenci jednotlivých bakterií. Než se vytvoří zvýšené množství specifických protilátek, musí „nespecifické“ mechanismy rezistence hostitele (např. polymorfonukleární neutrofil) bránit hostitele proti mikrobům [1]. Vývoj účinné specifické imunity může vyžadovat čas až několik týdnů [50]. U většiny zdravých jedinců se bakterie z normální flóry, které příležitostně pronikají do těla (například při čištění zubů), odstraňují pomocí buněčných mechanismů u hostitele. Naproti tomu, jedinci s oslabeným imunitním systémem jsou náchylní k častým, opakujícím se infekcím, a to i u méně virulentních bakterií [1]. Kojenci jsou například velmi citliví na určité patogeny. Mezi ně se řadí např. streptokoky skupiny B. Důvodem je to, že jejich imunitní buňky ještě nejsou zcela vyvinuty, a tím pádem nemohou navázat ochrannou imunitní odpověď na důležité bakteriální antigeny [53].

2.5 Křížové struktury

Inverzní repetice (IR) jsou úseky DNA, které jsou přítomny v genomech a jejich sekvence nukleotidů je stejná zepředu i zezadu. Inverzní repetice, která neobsahuje mezeru mezi jednotlivými inverzními repeticemi, se nazývá palindrom. Křížové struktury, neboli kruciformy, vznikají z obrácených (inverzních) repetic o délce šesti a více nukleotidů [54]. Pro vznik křížových struktur je důležitá přítomnost C-G párů. Díky jejich vyšší stabilitě jsou výhodnější pro tvorbu křížové struktury A-T páry [55]. Dalším parametrem je negativní, ale i pozitivní nadšroubovicové vinutí. Křížové struktury jsou obvykle složeny ze smyčky, stonku a bodu větvení [56], což je vyobrazeno na Obrázku 6.



Obrázek 6: Tvorba křížové struktury [57]

Rozpoznání křížové struktury v DNA se zdá být velice důležité nejen pro stabilitu genomu, ale také pro mnoho základních biologických procesů. Bylo prokázáno, že mnoho proteinů vykazuje vazebné vlastnosti specifické pro křížovou strukturu [56]. Vytváření křížové struktury z nepřímých inverzí repetit, které obsahují mezery, závisí nejen na délce mezery, ale také na její posloupnosti. Z přímých inverzních repetit (bez obsahu mezery) se vytváří křížové struktury s minimálně jednovláknovou smyčkou [55].

Vytváření křížové struktury *in vivo* bylo prokázáno u prokaryot i eukaryot pomocí několika metodických přístupů. Přítomnost struktury kříže byla poprvé popsána v cirkulární plazmidové DNA, kde negativní hustota superhelixu může stabilizovat tvorbu křížové struktury [58]. Křížové struktury jsou cílovými strukturami pro protoonkogenní protein DEK a mnoho dalších regulačních a strukturních proteinů [59].

Nadměrný výskyt inverzních repetit, které se vyskytují nepravidelně v DNA všech organismů, je zaznamenán v blízkosti oblasti promotoru či v místech zahájení replikace [55]. Křížové struktury mohou ovlivnit umístění nukleosomů, stupeň nadšroubovicovitého vinutí DNA nebo i tvorbu dalších sekundárních struktur [59].

2.6 Palindrome analyser

V roce 2016 byla vyvinuta nová platforma pod názvem Palindrome analyser. Palindrome analyser byl vyvinut ve spolupráci Biofyzikálního ústavu akademie věd v ČR a Mendelovy univerzity v Brně, a je volně dostupný na webové adrese: <http://bioinformatics.ibp.cz> [60].

Tato aplikace dokáže zobrazit a analyzovat inverzní repetice v různých sekvencích genomů, podle třech klíčových parametrů: délka hledaných palindromů, mezerník (tvorí smyčku a odděluje repetice) a počet neshod, tudíž počet nespárovaných bází. Také integruje výpočet energie sekvenace DNA v lineární a křížové struktuře, založené na termodynamice nejbližšího souseda [60].

Program je napsán v programovacím jazyce Java, poskytuje úložnou kapacitu pro importované sekvence a umožňuje ukládat výsledky analýz. Analýza je použitelná pro

lineární a křížové sekvence. Palindromový analyzátor lze použít třemi způsoby. První je klasický režim sever/klient, kdy aplikace běží na centrálním webovém serveru a návštěvníci provádí analýzy pomocí serverových prostředků dostupných online. V tomto režimu stačí znát platnou URL adresu aplikace. Druhý je režim aplikace, kdy je zapotřebí podporovaná platforma Javy. Výhodou tohoto režimu je rychlejší analýza, jelikož uživatel nemusí data nahrávat na sdílený server. Posledním režimem, ve kterém může být aplikace použita je tzv. „backend over“ API (rozhraní aplikace) pro vlastní výpočty bez použití uživatelského rozhraní [60].

3 EXPERIMENTÁLNÍ ČÁST

3.1 Metody

3.1.1 Palindrome analyser

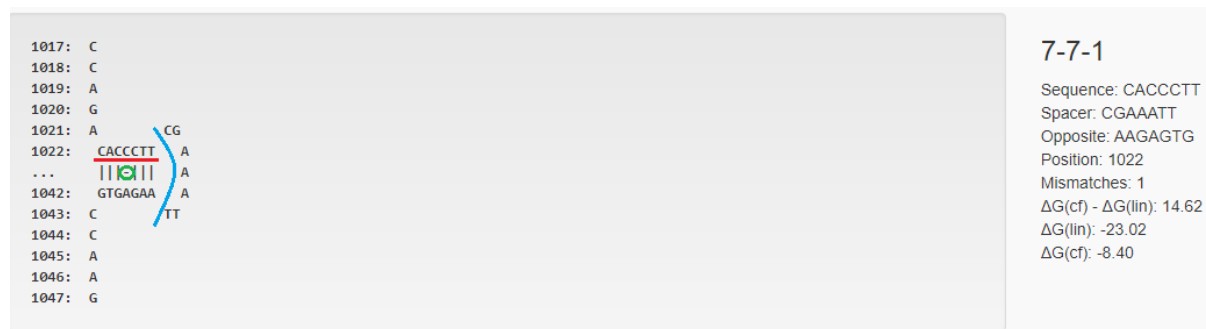
Byly importovány kódy čtyř virových genomů z NCBI (Nation Center for Biotechnology Information) do volně dostupné webové aplikace Palindrome analyser, dostupné na [61]. Konkrétně se jednalo o dva RNA viry: SARS-CoV a SARS-CoV-2 a dva DNA viry: Hepatitis B virus a Lidský papilomavirus typu 16. Pro analýzu inverzních repetic ve všech genomech byly nastaveny následující parametry: délka IR 6–30, délka smyčky 0–10 a počet chyb 0 a 1.

Výstup webové aplikace Palindrome analyser, viz Obrázek 7, zobrazuje na levé straně délku sekvence, poté počet bází, které tvoří smyčku a poslední číslo udává počet nepárování bází. Dále následuje pozice inverzní repetice a energie, která je potřebná pro její vytvoření. Úplně napravo se nacházejí báze sekvence tvořící křížovou strukturu.

Length - Spacer - Mismatches	↑ Position	$\Delta G(\text{cf}) - \Delta G(\text{lin})$	Palindrome
6-7-1	Q 3151	15.08	TGAGAA ACACTCA TCCTCA
6-3-1	Q 3141	11.63	TCTCCA CCT TTGAGA

Obrázek 7: Výstup z analýzy viru Hepatitis B pomocí Palindrome analyser

Po rozkliknutí detailů lze vidět i znázornění křížové struktury. Na Obrázku 8 je vidět vlásenku o délce sedm bází, která je znázorněna červeně a smyčka, rovněž tvořena sedmi bázemi, která je znázorněna modře. Nachází se zde i jedna chyba neboli nespárování, což je označeno zeleným kolečkem.



Obrázek 8: Schéma křížové struktura viru SARS-CoV z webové aplikace Palindrome Analyser

Na konci každé analýzy se nachází přehled informací o celém genomu. Tento přehled lze vidět na Obrázku 9 a zahrnuje informace jako je velikost sekvence, celkový počet křížových struktur, minimální a maximální energie, aby mohla být vytvořena křížová struktura a tři tabulky uvádějící kolikrát se daná délka inverzních repetic, délka smyčky či chyba opakuje v celé sekvenci virového genomu.

Size of sequence	Found # of palindromes	$\Delta G(\text{cf}) - \Delta G(\text{lin})$	Request duration
3182 bp	96	min: 2.65 max: 25.06	0.338s

Amounts by length		Amounts by spacer		Amounts by mismatches	
6	63x	0	7x	0	7x
7	23x	1	11x	1	89x
8	9x	2	13x		
9	1x	3	8x		
		4	11x		
		5	9x		
		6	6x		
		7	5x		
		8	7x		
		9	10x		
		10	9x		

Obrázek 9: Seřazení dle délky IR, smyčky a chybovosti u viru Hepatitis B z Palindrome analyser

4 VÝSLEDKY A DISKUZE

4.1 SARS-CoV

Tento virový genom má velikost 29 751 bp, což z něj dělá druhý největší genom ze čtyř zkoumaných. Křížových struktur se v SARS-CoV nachází 1 088, přičemž nejčastěji se jedná o inverzní repetice o délce 6 bází. Početnost klesá s narůstající délkou inverzních repetic, což je vyobrazeno v Tabulce 1. Nejdelší IR má 12 bází, ale vyskytuje se v genomu pouze třikrát, zatímco IR o velikosti 6 bází je v genomu přítomný 672krát. Většina křížových struktur obsahuje jednu chybu, opravdu jen malá část (necelých 7 %) je bez chyby. Nejčtenější délka smyčky je o velikosti 1 báze, viz Tabulka 2, přičemž tato hodnota se v celém genomu SARS-CoV vyskytuje celkem 134krát.

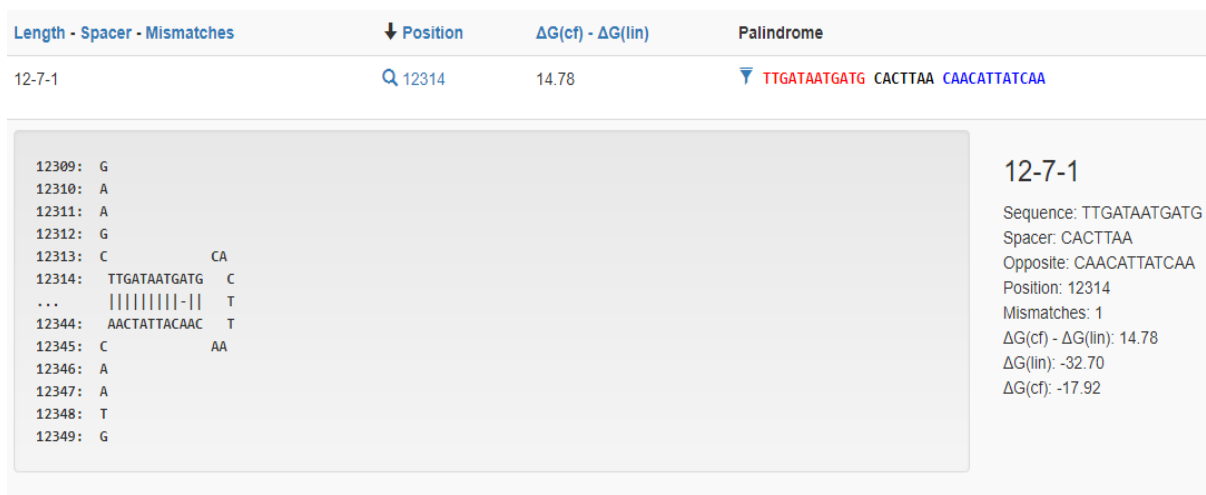
Tabulka 1: Velikost, počet a frekvence IRs v genomu SARS-CoV dle délky IR

Velikost IR [bp]	Počet IR	Frekvence [počet IR/velikost]
6	672	0,0226
7	241	0,0081
8	99	0,0033
9	46	0,0015
10	21	0,0007
11	6	0,0002
12	3	0,0001

Tabulka 2: Velikost, počet a frekvence IRs v genomu SARS-CoV dle délky smyčky

Velikost smyčky [bp]	Počet IR	Frekvence [počet IR/velikost]
0	110	0,0037
1	134	0,0045
2	124	0,0042
3	125	0,0042
4	110	0,0037
5	82	0,0028
6	93	0,0031
7	85	0,0029
8	81	0,0027
9	75	0,0025
10	69	0,0023

Na Obrázku 10 je vyobrazena nejdelší inverzní repetice genomu SARS-CoV, která má délku 12 bází, přičemž délku smyčky tvoří 7 bází. Tato inverzní repetice se nachází na 12 314. místě a lze tu taky vidět nespárování, a to mezi dvěma adeniny.



Obrázek 10: Největší IR v genomu SARS-CoV

4.2 SARS-CoV-2

S velikostí 29 903 bp se jedná o největší zkoumaný genom. V Tabulce 3 lze vidět, že SARS-CoV-2 oproti SARS-CoV obsahuje nejdelší IR o délce 13 bází. Celkově se zde nachází 1 203 křížových struktur, které mají nejčastější délku šest bází, a to celkem 737krát. Necelých 8 % křížových struktur neobsahuje jedinou chybu, zatímco zbylých 92 % obsahuje jednu chybu. SARS-CoV-2 sdílí se SARS-CoV to, že nejčastěji se vyskytovaná délka smyčky má velikost 1 bázi, což je vyobrazeno v Tabulka 4. I v tomhle genomu se potvrzuje pravidlo, že s narůstající délkou inverzních repetíc klesá jejich početnost. Přibližně třetina inverzních repetíc zcela postrádala smyčku.

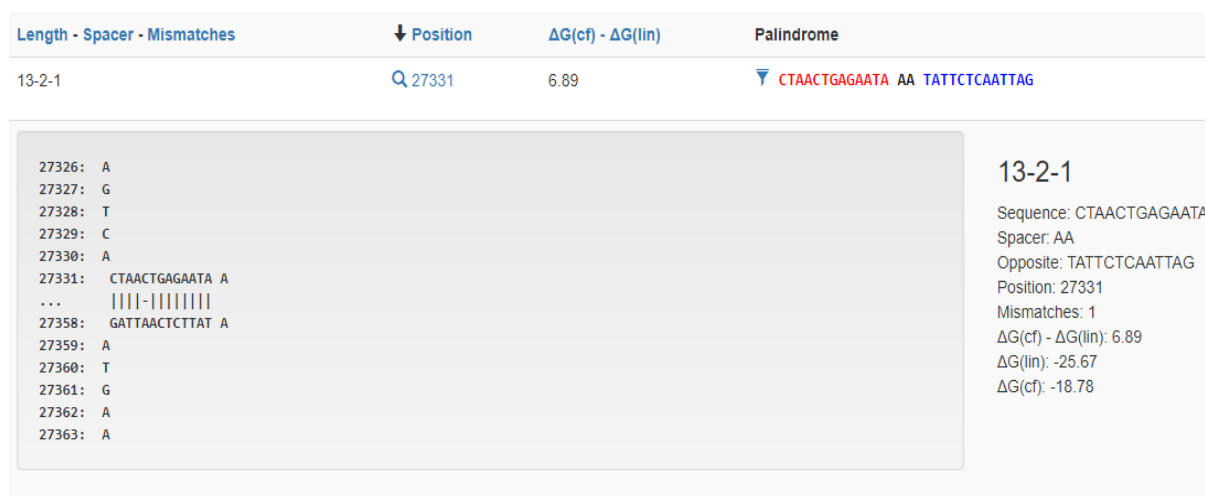
Tabulka 3: Velikost, počet a frekvence IRs v genomu SARS-CoV-2 dle délky IR

Velikost IR [bp]	Počet IR	Frekvence [počet IR/velikost]
6	737	0,0246
7	263	0,0088
8	127	0,0042
9	39	0,0013
10	28	0,0009
11	4	0,0001
12	3	0,0001
13	2	0,0001

Tabulka 4: Velikost, počet a frekvence IRs v genomu SARS-CoV-2 dle délky smyčky

Velikost smyčky [bp]	Počet IR	Frekvence [počet IR/velikost]
0	124	0,0041
1	142	0,0047
2	124	0,0041
3	120	0,0040
4	113	0,0038
5	99	0,0033
6	115	0,0038
7	100	0,0033
8	99	0,0033
9	67	0,0022
10	100	0,0033

V genomu SARS-CoV-2, na Obrázku 11, lze pozorovat na 27 331. pozici nejdelší inverzní repetici, která obsahuje 13 bází a 2 báze tvoří smyčku. Opět je zde jedno nespárování a to konkrétně mezi cytosinem a adeninem.



Obrázek 11: Největší IR v genomu SARS-CoV-2

4.3 Hepatitis B virus

DNA virový genom, který má velikost pouze 3 182 bp, obsahuje 98 křížových struktur. Délky inverzních repetice se pohybují mezi šesti až devíti bázemi, viz Tabulka 5, přičemž inverzní repetice o délce devět bází zde najdeme pouze jednou. Nejčastěji se zde vyskytuje smyčka o délce dvou bází, viz Tabulka 6. Ani v tomto genomu není výjimkou, že většina křížových struktur je s chybovostí jedna, a to skoro až 93 %.

Tabulka 5: Velikost, počet a frekvence IRs v genomu Hepatitis B dle délky IR

Velikost IR [bp]	Počet IR	Frekvence [počet IR/velikost]
6	63	0,0198
7	23	0,0072
8	9	0,0028
9	1	0,0003

Tabulka 6: Velikost, počet a frekvence IRs v genomu Hepatitis B dle délky smyčky

Velikost smyčky [bp]	Počet IR	Frekvence [počet IR/velikost]
0	7	0,0022
1	11	0,0035
2	13	0,0041
3	8	0,0025
4	11	0,0035
5	9	0,0028
6	6	0,0019
7	5	0,0016
8	7	0,0022
9	10	0,0031
10	9	0,0028

Obrázek 11 zobrazuje nejdelší inverzní repetici v genomu Hepatitis B. Tato inverzní repetice je složena z 9 bází, dalších 5 bází tvoří smyčku a je umístěna na 1 195. pozici. Na obrázku lze vidět i jednu chybu, a to mezi dvěma cytosiny.

Length - Spacer - Mismatches	↓ Position	ΔG(cf) - ΔG(lin)	Palindrome
9-5-1	Q 1195	16.35	CAACCCCA CTGGC TGGGGCTTG


```

1190: T
1191: G
1192: A
1193: C
1194: G      C
1195: CAACCCCA T
...   ||| - |||| G
1217: GTTCGGGGT G
1218: G      C
1219: T
1220: C
1221: A
1222: T
                    
```

9-5-1

Sequence: CAACCCCA
Spacer: CTGGC
Opposite: TGGGGCTTG
Position: 1195
Mismatches: 1
ΔG(cf) - ΔG(lin): 16.35
ΔG(lin): -33.23
ΔG(cf): -16.88

Obrázek 12: Největší IR v genomu Hepatitis B

4.4 Lidský papilomavirus typu 16

Druhý DNA virový genom má délku o velikosti 7 906 bp. Křížových struktur v tomto typu lidského papilomaviru je celkem 333, z toho se 227krát jedná o inverzní repetice o délce šesti bází. Tak jako u všech ostatních zkoumaných genomů se s narůstající délkou IR snižuje početnost, avšak v tomto genomu nejsou přítomny inverzní repetice o délce 11 a 12 bází, což je zobrazeno v Tabulce 7. V Tabulce 8 je nejčastější délkou smyčky v tomto případě nula bází, tudíž se jedná převážně o tzv. palindromy. Většina křížových struktur vykazuje jednu chybu.

Tabulka 7: Velikost, počet a frekvence IRs v genomu Human papillomavirus type 16 dle délky IR

Velikost IR [bp]	Počet IR	Frekvence [počet IR/velikost]
6	227	0,0287
7	65	0,0082
8	25	0,0032
9	9	0,0011
10	5	0,0006
13	1	0,0001
14	1	0,0001

Tabulka 8: Velikost, počet a frekvence IRs v genomu Human papillomavirus type 16 dle délky smyčky

Velikost smyčky [bp]	Počet IR	Frekvence [počet IR/velikost]
0	41	0,0052
1	36	0,0046
2	29	0,0037
3	21	0,0027
4	37	0,0047
5	31	0,0039
6	32	0,0040
7	20	0,0025
8	27	0,0034
9	35	0,0044
10	24	0,0030

Na Obrázku 13 lze pozorovat nejdelší inverzní repetici z tohoto, i ze všech zkoumaných genomů. Skládá se ze 14 bází a 4 báze utváří smyčku. Opět se zde nachází jedno nespárování, a to mezi adeninem a adeninem. Tato inverzní repetice se nachází na pozici 3 228, což je lehce před polovinou genomu.

Length - Spacer - Mismatches	↓ Position	ΔG(cf) - ΔG(lin)	Palindrome
14-4-1	3228	8.39	AATTACATAATGTA TATG TACATAATGTAATT

<pre> 3223: T 3224: T 3225: T 3226: T 3227: T T 3228: AATTACATAATGTA A ... ----- 3259: TTAATGTAATACAT T 3260: G G 3261: T 3262: T 3263: A 3264: C </pre>	<p>14-4-1</p> <p>Sequence: AATTACATAATGTA Spacer: TATG Opposite: TACATAATGTAATT Position: 3228 Mismatches: 1 ΔG(cf) - ΔG(lin): 8.39 ΔG(lin): -29.51 ΔG(cf): -21.12</p>
--	--

Obrázek 13: Největší IR v genomu Lidský papilomavirus typu 16

5 ZÁVĚR

V teoretické části byly popsány různé druhy patogenů, a to priony, viry a bakterie. Konkrétně jsem se zaměřila na viry: SARS-CoV, SARS-CoV-2, virus Hepatitis B a lidský papilomavirus, které byly následně podrobené bioinformatické analýze.

Genomy těchto čtyř virů byly staženy z databáze NCBI, který poskytuje celá jména genomů, fylogenetické skupiny a přesná přístupová čísla. Palindrome analyser je webová aplikace, která je upravená pro čtení NCBI identifikátorů. Tato aplikace byla použita v experimentální části bakalářské práce, kde byla zkoumána přítomnost, pozice a délka inverzních repetit v genomech výše zmíněných virů.

Již dříve bylo prokázáno, že inverzní repetice u virů hrají důležitou roli v procesech jako je transkripce, replikace, genová exprese či genová rekombinace [60].

V genomu lidského papilomaviru byla celková frekvence nejvyšší, a to 0,0421 IR/bp. Naopak nejnižší celková frekvence výskytu IR byla nalezena v genomu viru Hepatitis B, a to 0,0308 IR/bp. U všech genomů se vzrůstající velikosti inverzních repetit klesal počet inverzních repetit. Zhruba jedna desetina z celkového množství nalezených inverzních repetit ze všech genomů neobsahovala chybu. Oproti inverzním repetitím, které jsou v SARS-CoV-2 časté, se sekvence formující G-kvadruplexy v tomto viru prakticky nevyskytují [62].

6 SEZNAM POUŽITÝCH ZDROJŮ

- [1] PETERSON, Johnny W. Bacterial Pathogenesis. In: Samuel BARON, ed. *Medical Microbiology* [online]. 4th vyd. Galveston (TX): University of Texas Medical Branch at Galveston, 1996 [vid. 2020-03-17]. ISBN 978-0-9631172-1-2. Dostupné z: <http://www.ncbi.nlm.nih.gov/books/NBK8526/>
- [2] Microbiology by numbers. *Nature Reviews Microbiology* [online]. 2011, **9**(9), 628–628. ISSN 1740-1526, 1740-1534. Dostupné z: doi:10.1038/nrmicro2644
- [3] YIN, Yudong a Richard G. WUNDERINK. MERS, SARS and other coronaviruses as causes of pneumonia: MERS, SARS and coronaviruses. *Respirology* [online]. 2018, **23**(2), 130–137. ISSN 13237799. Dostupné z: doi:10.1111/resp.13196
- [4] GU, Jiang a Christine KORTEWEG. Pathology and Pathogenesis of Severe Acute Respiratory Syndrome. *The American Journal of Pathology* [online]. 2007, **170**(4), 1136–1147. ISSN 0002-9440. Dostupné z: doi:10.2353/ajpath.2007.061088
- [5] LIANG, T. Jake. Hepatitis B: The Virus and Disease. *Hepatology (Baltimore, Md.)* [online]. 2009, **49**(5 Suppl), S13–S21. ISSN 0270-9139. Dostupné z: doi:10.1002/hep.22881
- [6] GEREIN, Valentin, Eugen RASTORGUEV, Johann GEREIN, Wolfgang DRAF a Joachim SCHIRREN. Incidence, age at onset, and potential reasons of malignant transformation in recurrent respiratory papillomatosis patients: 20 years experience. *Otolaryngology–Head and Neck Surgery* [online]. 2005, **132**(3), 392–394. ISSN 0194-5998, 1097-6817. Dostupné z: doi:10.1016/j.otohns.2004.09.035
- [7] ALBERTS, Bruce, Alexander JOHNSON, Julian LEWIS, Martin RAFF, Keith ROBERTS a Peter WALTER. Introduction to Pathogens. *Molecular Biology of the Cell. 4th edition* [online]. 2002 [vid. 2020-03-17]. Dostupné z: <https://www.ncbi.nlm.nih.gov/books/NBK26917/>
- [8] BOLEK, Matthew G., Heather R. TRACY a John JANOVY. The Role of Damsel flies (Odonata: Zygoptera) as Paratenic Hosts in the Transmission of *HaliPEGUS eccentricus* (Digenea: Hemiuridae) to Anurans. *Journal of Parasitology* [online]. 2010, **96**(4), 724–735. ISSN 0022-3395, 1937-2345. Dostupné z: doi:10.1645/GE-2365.1
- [9] CLARKE, Anthony R., Graham S. JACKSON a John COLLINGE. The molecular biology of prion propagation. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences* [online]. 2001, **356**(1406), 185–195. ISSN 0962-8436, 1471-2970. Dostupné z: doi:10.1098/rstb.2000.0764
- [10] *What Is a Prion?* - *Scientific American* [online]. [vid. 2020-02-06]. Dostupné z: <https://www.scientificamerican.com/article/what-is-a-prion-specifica/>
- [11] DORMONT, Dominique. Prion diseases: pathogenesis and public health concerns. *FEBS Letters* [online]. 2002, **529**(1), 17–21. ISSN 00145793. Dostupné z: doi:10.1016/S0014-5793(02)03268-4

- [12] COLBY, David W. a Stanley B. PRUSINER. Prions. *Cold Spring Harbor Perspectives in Biology* [online]. 2011, **3**(1) [vid. 2020-01-30]. ISSN 1943-0264. Dostupné z: doi:10.1101/cshperspect.a006833
- [13] LODISH, Harvey, Arnold BERK, S. Lawrence ZIPURSKY, Paul MATSUDAIRA, David BALTIMORE a James DARNELL. Viruses: Structure, Function, and Uses. *Molecular Cell Biology. 4th edition* [online]. 2000 [vid. 2020-01-30]. Dostupné z: <https://www.ncbi.nlm.nih.gov/books/NBK21523/>
- [14] *OpenStax: Biology | CH21: VIRUSES | Top Hat* [online]. [vid. 2020-03-17]. Dostupné z: <https://tophat.com/marketplace/science-&-math/biology/textbooks/oer-openstax-biology-openstax-content/79/4124/>
- [15] MARAMOROSCH, Karl, Aaron J SHATKIN a Polly ROY. *Advances in virus research. Vol. 64, Vol. 64*, [online]. Amsterdam; New York: Elsevier Academic Press, 2005 [vid. 2020-07-18]. ISBN 978-0-08-045508-2. Dostupné z: <http://search.ebscohost.com/login.aspx?direct=true&scope=site&db=nlebk&db=nlabk&AN=166228>
- [16] FIELDS, Bernard N, David M KNIPE a Peter M HOWLEY. *Fields virology*. Philadelphia: Wolters Kluwer Health/Lippincott Williams & Wilkins, 2007. ISBN 978-0-7817-6060-7.
- [17] D, Baltimore. Expression of animal virus genomes. *Bacteriological reviews*. 1971, **35**(3), 235–41. ISSN 0005-3678.
- [18] Classification of virus | Virology. *Microbe Notes* [online]. 4. únor 2019 [vid. 2020-07-20]. Dostupné z: <https://microbenotes.com/classification-of-virus/>
- [19] GAVILANES, F., J. M. GONZALEZ-ROS a D. L. PETERSON. Structure of hepatitis B surface antigen. Characterization of the lipid components and their association with the viral proteins. *The Journal of Biological Chemistry*. 1982, **257**(13), 7770–7777. ISSN 0021-9258.
- [20] ALTER, H. J., P. V. HOLLAND, R. H. PURCELL, J. J. LANDER, S. M. FEINSTONE, A. G. MORROW a P. J. SCHMIDT. Posttransfusion hepatitis after exclusion of commercial and hepatitis-B antigen-positive donors. *Annals of Internal Medicine* [online]. 1972, **77**(5), 691–699. ISSN 0003-4819. Dostupné z: doi:10.7326/0003-4819-77-5-691
- [21] PURCELL, R. H. a J. L. GERIN. Hepatitis B subunit vaccine: a preliminary report of safety and efficacy tests in chimpanzees. *The American Journal of the Medical Sciences*. 1975, **270**(2), 395–399. ISSN 0002-9629.
- [22] HUMANS, IARC Working Group on the Evaluation of Carcinogenic Risk to. *HEPATITIS B VIRUS* [online]. B.m.: International Agency for Research on Cancer, 2012 [vid. 2020-03-29]. Dostupné z: <https://www.ncbi.nlm.nih.gov/books/NBK304352/>
- [23] HATTON, T., S. ZHOU a D. N. STANDRING. RNA- and DNA-binding activities in hepatitis B virus capsid protein: a model for their roles in viral replication. *Journal of Virology*. 1992, **66**(9), 5232–5241. ISSN 0022-538X.

- [24] BOUCHARD, Michael J. a Robert J. SCHNEIDER. The Enigmatic X Gene of Hepatitis B Virus. *Journal of Virology* [online]. 2004, **78**(23), 12725–12734. ISSN 0022-538X, 1098-5514. Dostupné z: doi:10.1128/JVI.78.23.12725-12734.2004
- [25] Worldwide cancer data. *World Cancer Research Fund* [online]. 6. srpen 2018 [vid. 2020-07-20]. Dostupné z: <https://www.wcrf.org/dietandcancer/cancer-trends/worldwide-cancer-data>
- [26] SCHIFFMAN, Mark, Philip E CASTLE, Jose JERONIMO, Ana C RODRIGUEZ a Sholom WACHOLDER. Human papillomavirus and cervical cancer. *The Lancet* [online]. 2007, **370**(9590), 890–907. ISSN 01406736. Dostupné z: doi:10.1016/S0140-6736(07)61416-0
- [27] ARALDI, Rodrigo Pinheiro, Thalita Araujo SANT'ANA, Diego Grando MÓDOLO, Thatiana Correa DE MELO, Diva Denelle SPADACCI-MORENA, Rita DE CASSIA STOCCO, Janete Maria CERUTTI a Edislane Barreiros DE SOUZA. The human papillomavirus (HPV)-related cancer biology: An overview. *Biomedicine & Pharmacotherapy = Biomedecine & Pharmacotherapie* [online]. 2018, **106**, 1537–1556. ISSN 1950-6007. Dostupné z: doi:10.1016/j.biopha.2018.06.149
- [28] DE VILLIERS, Ethel-Michele. Cross-roads in the classification of papillomaviruses. *Virology* [online]. 2013, **445**(1–2), 2–10. ISSN 00426822. Dostupné z: doi:10.1016/j.virol.2013.04.023
- [29] CHEN, X. S., R. L. GARCEA, I. GOLDBERG, G. CASINI a S. C. HARRISON. Structure of small virus-like particles assembled from the L1 protein of human papillomavirus 16. *Molecular Cell* [online]. 2000, **5**(3), 557–567. ISSN 1097-2765. Dostupné z: doi:10.1016/s1097-2765(00)80449-9
- [30] SATIJA, Namita a Sunil K. LAL. The Molecular Biology of SARS Coronavirus. *Annals of the New York Academy of Sciences* [online]. 2007, **1102**(1), 26–38. ISSN 0077-8923. Dostupné z: doi:10.1196/annals.1408.002
- [31] GRONEBERG, David A, Rolf HILGENFELD a Peter ZABEL. Molecular mechanisms of severe acute respiratory syndrome (SARS). *Respiratory Research* [online]. 2005, **6**(1), 8. ISSN 1465-9921. Dostupné z: doi:10.1186/1465-9921-6-8
- [32] PEIRIS, J. S. M., C. M. CHU, V. C. C. CHENG, K. S. CHAN, I. F. N. HUNG, L. L. M. POON, K. I. LAW, B. S. F. TANG, T. Y. W. HON, C. S. CHAN, K. H. CHAN, J. S. C. NG, B. J. ZHENG, W. L. NG, R. W. M. LAI, Y. GUAN a K. Y. YUEN. Clinical progression and viral load in a community outbreak of coronavirus-associated SARS pneumonia: a prospective study. *The Lancet* [online]. 2003, **361**(9371), 1767–1772. ISSN 0140-6736, 1474-547X. Dostupné z: doi:10.1016/S0140-6736(03)13412-5
- [33] POON, Leo L.M., Kwok Hung CHAN, On Kei WONG, Wing Cheong YAM, Kwok Yung YUEN, Yi GUAN, Y.M.Dennis LO a Joseph S.M. PEIRIS. Early diagnosis of SARS Coronavirus infection by real time RT-PCR. *Journal of Clinical Virology* [online]. 2003, **28**(3), 233–238. ISSN 13866532. Dostupné z: doi:10.1016/j.jcv.2003.08.004
- [34] ANTONIO, Gregory E., K. T. WONG, David S. C. HUI, Alan WU, Nelson LEE, Edmund H. Y. YUEN, C. B. LEUNG, T. H. RAINER, Peter CAMERON, Sydney S. C. CHUNG, Joseph J. Y. SUNG a Anil T. AHUJA. Thin-section CT in patients with severe

- acute respiratory syndrome following hospital discharge: preliminary experience. *Radiology* [online]. 2003, **228**(3), 810–815. ISSN 0033-8419. Dostupné z: doi:10.1148/radiol.2283030726
- [35] LANG, Zhen-Wei, Li-Jie ZHANG, Shi-Jie ZHANG, Xin MENG, Jun-Qiang LI, Chen-Zhao SONG, Ling SUN, Yu-Sen ZHOU a Dominic E. DWYER. A clinicopathological study of three cases of severe acute respiratory syndrome (SARS). *Pathology* [online]. 2003, **35**(6), 526–531. ISSN 0031-3025. Dostupné z: doi:10.1080/00313020310001619118
- [36] SHI, Xueying, Encong GONG, Dongxia GAO, Bo ZHANG, Jie ZHENG, Zifen GAO, Yanfeng ZHONG, Wanzhong ZOU, Bingquan WU, Weigang FANG, Songlin LIAO, Shenglan WANG, Zhigang XIE, Min LU, Lin HOU, Haohao ZHONG, Hongquan SHAO, Ning LI, Congrong LIU, Fei PEI, Jingping YANG, Yuping WANG, Zhihui HAN, Xiaohong SHI, Qianying ZHANG, Jiangfeng YOU, Xiang ZHU a Jiang GU. Severe acute respiratory syndrome associated coronavirus is detected in intestinal tissues of fatal cases. *The American Journal of Gastroenterology* [online]. 2005, **100**(1), 169–176. ISSN 0002-9270. Dostupné z: doi:10.1111/j.1572-0241.2005.40377.x
- [37] CHAN-YEUNG, Moira a Rui-Heng XU. SARS: epidemiology. *Respirology* [online]. 2003, **8**(s1), S9–S14. ISSN 1323-7799, 1440-1843. Dostupné z: doi:10.1046/j.1440-1843.2003.00518.x
- [38] WHO | Middle East respiratory syndrome coronavirus (MERS-CoV). *WHO* [online]. [vid. 2020-07-20]. Dostupné z: <http://www.who.int/emergencies/mers-cov/en/>
- [39] PERLMAN, Stanley a Jason NETLAND. Coronaviruses post-SARS: update on replication and pathogenesis. *Nature Reviews. Microbiology* [online]. 2009, **7**(6), 439–450. ISSN 1740-1534. Dostupné z: doi:10.1038/nrmicro2147
- [40] CHAN, Jasper Fuk-Woo, Kin-Hang KOK, Zheng ZHU, Hin CHU, Kelvin Kai-Wang TO, Shuofeng YUAN a Kwok-Yung YUEN. Genomic characterization of the 2019 novel human-pathogenic coronavirus isolated from a patient with atypical pneumonia after visiting Wuhan. *Emerging Microbes & Infections* [online]. 2020, **9**(1), 221–236. ISSN 2222-1751. Dostupné z: doi:10.1080/22221751.2020.1719902
- [41] ZHOU, Peng, Xing-Lou YANG, Xian-Guang WANG, Ben HU, Lei ZHANG, Wei ZHANG, Hao-Rui SI, Yan ZHU, Bei LI, Chao-Lin HUANG, Hui-Dong CHEN, Jing CHEN, Yun LUO, Hua GUO, Ren-Di JIANG, Mei-Qin LIU, Ying CHEN, Xu-Rui SHEN, Xi WANG, Xiao-Shuang ZHENG, Kai ZHAO, Quan-Jiao CHEN, Fei DENG, Lin-Lin LIU, Bing YAN, Fa-Xian ZHAN, Yan-Yi WANG, Geng-Fu XIAO a Zheng-Li SHI. A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature* [online]. 2020, **579**(7798), 270–273. ISSN 0028-0836, 1476-4687. Dostupné z: doi:10.1038/s41586-020-2012-7
- [42] LI, Fang. Structure, Function, and Evolution of Coronavirus Spike Proteins. *Annual Review of Virology* [online]. 2016, **3**(1), 237–261. ISSN 2327-056X, 2327-0578. Dostupné z: doi:10.1146/annurev-virology-110615-042301
- [43] HE, Yuxian, Yusen ZHOU, Shuwen LIU, Zhihua KOU, Wenhui LI, Michael FARZAN a Shibo JIANG. Receptor-binding domain of SARS-CoV spike protein induces highly

- potent neutralizing antibodies: implication for developing subunit vaccine. *Biochemical and Biophysical Research Communications* [online]. 2004, **324**(2), 773–781. ISSN 0006291X. Dostupné z: doi:10.1016/j.bbrc.2004.09.106
- [44] HE, Feng, Yu DENG a Weina LI. Coronavirus disease 2019: What we know? *Journal of Medical Virology* [online]. nedatováno, **n/a**(n/a) [vid. 2020-04-14]. ISSN 1096-9071. Dostupné z: doi:10.1002/jmv.25766
- [45] KAMPF, G., D. TODT, S. PFAENDER a E. STEINMANN. Persistence of coronaviruses on inanimate surfaces and their inactivation with biocidal agents. *The Journal of Hospital Infection* [online]. 2020, **104**(3), 246–251. ISSN 1532-2939. Dostupné z: doi:10.1016/j.jhin.2020.01.022
- [46] DYKHUIZEN, Daniel. Species Numbers in Bacteria. *Proceedings. California Academy of Sciences*. 2005, **56**(6 Suppl 1), 62–71. ISSN 0068-547X.
- [47] *bacteria* | Origin and meaning of bacteria by Online Etymology Dictionary [online]. [vid. 2020-07-20]. Dostupné z: <https://www.etymonline.com/word/bacteria>
- [48] *Pasteur's Papers on the Germ Theory* [online]. [vid. 2020-07-20]. Dostupné z: <https://biotech.law.lsu.edu/cphl/history/articles/pasteur.htm>
- [49] The Nobel Prize in Physiology or Medicine 1905. *NobelPrize.org* [online]. [vid. 2020-07-20]. Dostupné z: <https://www.nobelprize.org/prizes/medicine/1905/summary/>
- [50] VAN HEYNINGEN, S. Molecular basis of bacterial pathogenesis (volume xi of 'the bacteria: a treatise on structure and function': Edited by Barbara H. Iglewski and Virginia L. Clark; Academic Press; San Diego, 1990; xi + 473 pages; £85.00. *FEBS Letters* [online]. 1991, **287**(1–2), 227–228. ISSN 00145793. Dostupné z: doi:10.1016/0014-5793(91)80060-G
- [51] DAUBIN, Vincent, Nancy A. MORAN a Howard OCHMAN. Phylogenetics and the cohesion of bacterial genomes. *Science (New York, N.Y.)* [online]. 2003, **301**(5634), 829–832. ISSN 1095-9203. Dostupné z: doi:10.1126/science.1086568
- [52] *File:Average prokaryote cell- en.svg* [online]. 2020 [vid. 2020-07-20]. Dostupné z: https://en.wikipedia.org/w/index.php?title=File:Average_prokaryote_cell-_en.svg&oldid=948192438
- [53] FINLAY, B. B. a S. FALKOW. Common themes in microbial pathogenicity. *Microbiological Reviews*. 1989, **53**(2), 210–230. ISSN 0146-0749.
- [54] MURCHIE, Alastair I.H. a David M.J. LILLEY. The mechanism of cruciform formation in supercoiled DNA: initial opening of central basepairs in salt-dependent extrusion. *Nucleic Acids Research* [online]. 1987, **15**(23), 9641–9654. ISSN 0305-1048, 1362-4962. Dostupné z: doi:10.1093/nar/15.23.9641
- [55] VOLOGODSKII, A V a N R COZZARELLI. Conformational and Thermodynamic Properties of Supercoiled DNA. *Annual Review of Biophysics and Biomolecular Structure* [online]. 1994, **23**(1), 609–643. ISSN 1056-8700, 1545-4266. Dostupné z: doi:10.1146/annurev.bb.23.060194.003141

- [56] BRÁZDA, Václav, Rob C LAISTER, Eva B JAGELSKÁ a Cheryl ARROWSMITH. Cruciform structures are a common DNA feature important for regulating biological processes. *BMC Molecular Biology* [online]. 2011, **12**(1), 33. ISSN 1471-2199. Dostupné z: doi:10.1186/1471-2199-12-33
- [57] Changes associated with transition from the linear to cruciform state... *ResearchGate* [online]. [vid. 2020-07-20]. Dostupné z: https://www.researchgate.net/figure/Changes-associated-with-transition-from-the-linear-to-cruciform-state-in-the-p53-target_fig1_51546715
- [58] PANAYOTATOS, N. a A. FONTAINE. A native cruciform DNA structure probed in bacteria by recombinant T7 endonuclease. *The Journal of Biological Chemistry*. 1987, **262**(23), 11364–11368. ISSN 0021-9258.
- [59] JAGELSKÁ, Eva B., Hana PIVONKOVÁ, Miroslav FOJTA a Václav BRÁZDA. The potential of the cruciform structure formation as an important factor influencing p53 sequence-specific binding to natural DNA targets. *Biochemical and Biophysical Research Communications* [online]. 2010, **391**(3), 1409–1414. ISSN 0006291X. Dostupné z: doi:10.1016/j.bbrc.2009.12.076
- [60] BRÁZDA, Václav, Jan KOLOMAZNÍK, Jiří LÝSEK, Lucia HÁRONÍKOVÁ, Jan COUFAL a Jiří ŠT'ASTNÝ. Palindrome analyser – A new web-based server for predicting and evaluating inverted repeats in nucleotide sequences. *Biochemical and Biophysical Research Communications* [online]. 2016, **478**(4), 1739–1745. ISSN 0006291X. Dostupné z: doi:10.1016/j.bbrc.2016.09.015
- [61] SAYERS, Eric W, Richa AGARWALA, Evan E BOLTON, J Rodney BRISTER, Kathi CANESE, Karen CLARK, Ryan CONNOR, Nicolas FIORINI, Kathryn FUNK, Timothy HEFFERON, J Bradley HOLMES, Sunghwan KIM, Avi KIMCHI, Paul A KITTS, Stacy LATHROP, Zhiyong LU, Thomas L MADDEN, Aron MARCHLER-BAUER, Lon PHAN, Valerie A SCHNEIDER, Conrad L SCHOCH, Kim D PRUITT a James OSTELL. Database resources of the National Center for Biotechnology Information. *Nucleic Acids Research* [online]. 2019, **47**(D1), D23–D28. ISSN 0305-1048, 1362-4962. Dostupné z: doi:10.1093/nar/gky1069
- [62] BARTAS, Martin, Václav BRÁZDA, Natália BOHÁLOVÁ, Alessio CANTARA, Adriana VOLNÁ, Tereza STACHUROVÁ, Kateřina MALACHOVÁ, Eva B. JAGELSKÁ, Otília PORUBIAKOVÁ, Jiří ČERVENĚ a Petr PEČINKA. In-Depth Bioinformatic Analyses of Nidovirales Including Human SARS-CoV-2, SARS-CoV, MERS-CoV Viruses Suggest Important Roles of Non-canonical Nucleic Acid Structures in Their Lifecycles. *Frontiers in Microbiology* [online]. 2020, **11** [vid. 2020-07-26]. ISSN 1664-302X. Dostupné z: doi:10.3389/fmicb.2020.01583

7 SEZNAM POUŽITÝCH ZKRATEK A SYMBOLŮ

- A – adenin
- ACE2 – enzym konvertující angiotensin 2
- AMK – aminokyselina
- bp – pár bází
- C – cytosin
- CIN – cervikální intraepiteliární neoplazie
- CoV – koronavirus
- DNA – deoxyribonukleová kyselina
- dsDNA – dvouřetězcová DNA
- ERGIC – organela v eukaryotických buňkách
- G – guanin
- HBcAg – vnitřní jádro (základní antigen) HBV
- HBeAg – část HBcAg, vylučována během replikace
- HBsAg – povrchový antigen HBV
- HBV – virus hepatitidy B
- HCV – virus hepatitidy C
- HIV – virus lidské imunitní nedostatečnosti (human immunodeficiency virus)
- HPV – lidský papilomavirus (human papilloma virus)
- ICTV – mezinárodní výbor pro klasifikaci virů
- IR – inverzní repetice
- kDa – kilodalton
- MERS-CoV – respirační syndrom na Středním východě
- mRNA – mediátorová RNA
- NCBI – national center for biotechnology information
- ORF – otevřený čtecí rámec
- RCT – replikačně-transkripční komplex
- RNA – ribonukleová kyselina
- RT-PCR – polymerázová řetězová reakce v reálném čase
- SARS-CoV – vir způsobující těžký akutní respirační syndrom
- sgRNA – subgenomická mRNA
- ssDNA – jednořetězcová DNA
- ssRNA – jednořetězcová RNA
- SV40 – opičí virus 40
- T – thymin
- TSE – transmisivní spongiformní encefalopatie
- UV – ultrafialová oblast