

Jihočeská univerzita v Českých Budějovicích

Přírodovědecká fakulta



Změny v genomu viru klíšové encefalitidy
u variant s různou historií pasáží a odlišnými
biologickými vlastnostmi

Diplomová práce

Bc. Renata Strouhalová

Školitel: Prof. RNDr. Libor Grubhoffer, CSc. (PřF JU, PAÚ BC AV ČR)

Školitel specialista: Mgr. Václav Hönig (PřF JU, PAÚ BC AV ČR)

České Budějovice

2011

Strouhalová, R., 2011: Změny v genomu viru klíšťové encefalitidy u variant s různou historií pasáží a odlišnými biologickými vlastnostmi. [Genome changes of tick-borne encephalitis virus in variants with different passage history and biological properties. Mgr. Thesis, in Czech] – 96 p., Faculty of Science, The University of South Bohemia, České Budějovice, Czech Republic.

Anotace:

Tick-borne encephalitis virus (strain Hypr) was serially subcultured in PS cells and tick cell line IRE/CTVM19, producing four different viral variants. Biological properties of these new variants were investigated in mouse model. Possible determinants of virulence were found by full-genome sequencing. The role of glycosylation for tick-borne encephalitis virus was evaluated.

Tato práce byla financována granty:

Ministerstva školství, mládeže a tělovýchovy České republiky LC 06009 a Grantové agentury ČR 524/08/1509.

Prohlašuji, že svoji diplomovou práci jsem vypracovala samostatně pouze s použitím pramenů a literatury uvedených v seznamu citované literatury.

Prohlašuji, že v souladu s § 47b zákona č. 111/1998 Sb. v platném znění souhlasím se zveřejněním své diplomové práce, a to v nezkrácené podobě – v úpravě vzniklé vypuštěním vyznačených částí archivovaných Přírodovědeckou fakultou - 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 kvalifikační práce. Rovněž souhlasím s porovnáním textu mé kvalifikační práce s databází kvalifikačních prací Theses.cz provozovanou Národním registrem vysokoškolských kvalifikačních prací a systémem na odhalování plagiátů.

3. 1. 2011

.....

Poděkování:

Na tomto místě bych ráda poděkovala svému školiteli Prof. RNDr. Liboru Grubhofferovi, CSc. za možnost pracovat na tomto tématu, za poskytnutí zázemí a všestranné podpory.

Největší dík patří mému školiteli-specialistovi Vaškovi Hönigovi, který má hlavní zásluhu na vzniku této práce, který mě spoustu věcí naučil, poradil a pomohl, kdykoliv jsem potřebovala a v nejhorších chvílích mě vždy dokázal povzbudit.

Velký dík patří Martinovi za to, že mi pomáhal v mých laboratorních začátcích, naučil mě základní virologické metody a vštípil mi dokonale pravidla bezpečnosti práce s viry. Díky za to, že byl se mnou ve chvílích největšího zoufalství a přesvědčil mě to nevzdat.

Dík patří také Hance, Jankovi a Jirkovi i všem ostatním, kteří mi kdy při práci poradili nebo pomohli.

Zejména bych ale chtěla poděkovat svým rodičům za podporu, lásku a za možnost studovat.

Obsah:

1. Úvod	1
1.1 Virus klíšťové encefalitidy	1
1.1.1 Ekologické a epidemiologické charakteristiky infekce VKE.....	2
1.1.2 Patogeneze klíšťové encefalitidy.....	3
1.1.3 Molekulárně biologické charakteristiky viru	4
1.1.4 Cyklus viru v buňce.....	10
1.1.5 Glykosylace	12
1.2 Variabilita VKE	12
1.2.1 Variabilita RNA virů ve srovnání s DNA viry.....	12
1.2.2 Evoluce VKE a stabilita jeho genomu.....	13
1.2.3 Pasážování viru	14
1.2.4 „Quasispecies“	16
1.2.5 Molekulární determinanty virulence	17
2. Cíle práce.....	20
3. Materiál a metody	21
3.1 Virus	21
3.2 Buněčné linie	21
3.3 Pasážování viru.....	21
3.4 Plaková titrace viru	22
3.5 Izolace virové RNA	23
3.6 Reverzní transkripce	23
3.7 Polymerázová řetězová reakce (PCR).....	23
3.8 Separace PCR produktů na agarózovém gelu.....	25
3.9 Sekvenční analýza	25
3.10 3D modely proteinu E.....	25
3.11 Myši – testování virulence.....	26
3.12 Myši – průběh infekce.....	26
3.13 Real-time PCR	27
3.14 Statistická analýza	28
3.15 Metoda konformačního polymorfismu jednořetězců – SSCP	28
3.16 Klonování.....	30
3.16.1 Příprava ligační směsi.....	30
3.16.2 Transformace	31

3.16.3 PCR z kolonií	31
3.16.4 Pomnožení bakterií v tekutém médiu	31
3.16.5 Izolace plazmidu	31
4. Výsledky.....	33
4.1 Pasážování viru a hledání determinant virulence	33
4.1.1 Změna titru viru v průběhu pasážování a rozdíly ve velikosti plaků.....	33
4.1.2 Virulence variant VKE v myším modelu.....	35
4.1.3 Sledování dynamiky množení viru v orgánech.....	37
4.1.4 Sekvenční analýza kompletního genomu všech čtyř variant VKE	40
4.1.5 Potenciální determinanty virulence.....	43
4.2 Variabilita v rámci jednotlivých izolátů (přítomnost „quasispecies“).....	46
4.2.1 Metoda konformačního polymorfismu jednořetězců (SSCP)	46
4.2.2 Klonování	47
4.3 Inhibice glykosylace tunikamycinem	50
5. Diskuze.....	51
5.1 Biologické vlastnosti nově získaných variant viru	51
5.2 Dynamika množení viru v orgánech	53
5.3 Potenciální determinanty virulence.....	54
5.4 Variabilita v rámci virových variant.....	58
5.5 Význam glykosylace VKE pro biologické vlastnosti viru	60
6. Závěr.....	62
7. Seznam zkratk	63
8. Literatura.....	64
9. Příloha.....	76

1. Úvod

1.1 Virus klíšťové encefalidity

Virus klíšťové encefalidity (VKE) patří do rodu *Flavivirus*, čeledi *Flaviviridae*. Je původcem klíšťové encefalidity, jednoho z nejnebezpečnějších onemocnění centrálního nervového systému, vyskytujícího se v Evropě a v Asii (Heinz *et al.*, 2000). Incidence tohoto onemocnění neustále roste, každoročně je celosvětově zaznamenáno 10-12 tisíc případů (Süss, 2003).

Onemocnění klíšťovou encefalidou bylo poprvé popsáno roku 1931 rakouským lékařem Schneiderem, který si povšiml jeho pravidelného sezónního výskytu (Schneider, 1931). Samotný VKE, jako původce tohoto onemocnění, byl objeven roku 1937 ruským vědcem Zilberem ve východním Rusku, který zároveň prokázal přenos VKE na člověka pomocí klíšťe *Ixodes persulcatus* (Zilber *et al.*, 1939).

VKE je přenášen klíšťaty a patří tedy mezi arboviry - viry přenášené na obratlovce prostřednictvím krevsajících členovců. Rod *Flavivirus* zahrnuje kromě VKE i další viry, které lze v rámci tohoto rodu obecně rozdělit do tří skupin: flaviviry přenášené klíšťaty (např. virus Langat, virus Powassan), flaviviry přenášené komáry (např. virus žluté zimnice, virus západního Nilu, virus dengue) a flaviviry, u kterých nebyl dosud žádný vektor prokázán (např. virus Modoc, virus Apoi, virus Rio Bravo) (Varelas-Wesley *et al.*, 1982; Gaunt *et al.*, 2001).

Blízce příbuznými viry VKE ze skupiny klíšťaty přenášených flavivirů jsou virus vrtivky (Louping ill virus), virus Langat a virus Powassan, které také mohou způsobovat encefalidity u lidí, ale jen zřídka na epidemické úrovni. Další tři příbuzné viry ze stejné skupiny (virus Omské hemoragické horečky, virus horečky Kyasanurského lesa a virus Alkhurma) způsobují častěji než encefalidity ještě mnohem nebezpečnější hemoragické horečky (Charrel *et al.*, 2001).

Samotný VKE se ještě dále dělí do tří subtypů: subtyp evropský (prototypový kmen Neudoerfl), sibiřský (prototypový kmen Aina) a dálně-východní (prototypový kmen Sofjin).

Tyto odlišné genetické linie byly definovány na základě srovnávací analýzy sekvencí kódujících protein E různých kmenů VKE (Ecker *et al.*, 1999).

1.1.1 Ekologické a epidemiologické charakteristiky infekce VKE

V přirozeném prostředí je virus klíšťové encefalidity přenášen v cyklu zahrnujícím klíšťata jako vektory viru a obratlovce jako jeho hostitele. Přirozeným prostředím výskytu viru jsou lesní ekosystémy Evropy a Asie, ve kterých se udržuje dostatečná vlhkost nutná pro život klíšťat (Nuttall *et al.*, 1994).

Klíště *I. ricinus* je v Evropě dominantním druhem klíšťe a zároveň hlavním vektorem evropského subtypu VKE. Jeho sezónní aktivita obecně vrcholí dvakrát do roka, mezi květnem a červnem a zářím a říjnem. Klíště *I. persulcatus* se vyskytuje především v oblasti Uralu, Sibiře a Dálného východu a je hlavním přenašečem dalších dvou subtypů VKE – sibiřského a dálně-východního. K vrcholu jeho aktivity dochází pouze jednou ročně a to v průběhu května a června (Korenberg *et Kovalevskii*, 1999). V Rusku byl VKE sporadicky izolován také z 18 dalších druhů klíšťat, jako například *Dermacentor* spp. a *Hyalomma* spp. (Zlobin *et Gorin*, 1996). V České republice byl VKE izolován například z klíšťe *I. hexagonus* (Křivanec *et al.*, 1988) nebo *Haemaphysalis inermis* (Grešíková *et Nosek*, 1966). Samotná detekce viru však nepotvrzuje zapojení těchto klíšťat do cirkulace viru v přírodě, natož jejich význam pro přenos viru na člověka.

Prevalence VKE v klíšťatech v endemických oblastech Evropy se pohybuje v rozmezí 0,5-5 % (Süss *et al.*, 1999; Süss *et al.*, 2002), zatímco v určitých oblastech Ruska byla zaznamenána prevalence až 40 % (Gritsun *et al.*, 2003).

Virus je na fyziologii i chování klíšťat velmi dobře adaptován (Nuttall *et al.*, 1994). VKE v klíšťeti perzistuje přes metamorfózu, a tak jednou infikované klíště zůstává nakaženo po celý zbytek života. Pokud se tedy klíště nakazí jako larva, zůstane infikované i jako nymfa a dospělec (tzv. transstadiální přenos). Vzácně může dojít i k přenosu transovariálnímu, kdy se virus přenáší z matky na potomstvo. Klíšťata se VKE infikují nejčastěji při sání na infikovaném hostiteli, u kterého se právě vyskytuje virus v krvi – viremický přenos. Kromě viremického přenosu existuje ještě jiný, velmi důležitý mechanismus, napomáhající cirkulaci viru mezi jedinci klíšťat, označovaný jako „co-feeding“. Pro přenos viru z jednoho klíšťe na druhé při současném sání (tzv. sousání neboli „co-feeding“) na témže hostiteli není totiž u tohoto hostitele nutné dosažení virémie (Labuda *et al.*, 1993). Dochází pouze k lokálnímu přenosu v podkoží v místě sání infikovaného klíšťe na klíště dosud neinfikované (Jones *et*

al., 1987). K přenosu viru mezi klíštětem infikovaným a neinfikovaným může dojít dokonce i jejich sousáním na imunních jedincích (Labuda *et al.*, 1997).

Z důvodu celoživotní perzistence VKE v klíšťatech jsou klíšata nejen významným rezervoárem tohoto viru, ale také na virus působí dlouhodobým selekčním tlakem, který může ovlivňovat genotyp a fenotyp viru (Nuttall *et al.*, 1991; Labuda *et al.*, 1994). Proto se adaptace vzniklé na podkladě interakce s různými druhy vektorů mohou projevit různou mírou patogenity pro obratlovčí hostitele (Leonova, 1997; Romanova *et al.*, 2007).

Primárními obratlovčími hostiteli VKE jsou drobní hlodavci (Nuttall *et al.*, 1994). Další obratlovci jako ptáci, vysoká zvěř a koně slouží také jako hostitelé viru, nehrají ale pravděpodobně při přenosu viru mezi klíšaty tak důležitou roli (Nuttall *et al.*, 1994). Ve srovnání s klíšaty mají savčí hostitelé funkci pouze krátkodobých hostitelů s omezeným trváním infekce. I tak je přenos přes obratlovce velice důležitý pro perzistenci viru v přírodě (Gritsun *et al.*, 2003). Zásadní roli v udržení cirkulace viru v přírodě hraje zřejmě výše zmíněný mechanismus neviremického přenosu viru (Labuda *et al.*, 1997).

1.1.2 Patogeneze klíšťové encefalitidy

K nákaze člověka dochází ve většině případů po přisátí infikovaného klíštěte. Přibližně 70-95 % infekcí v endemických oblastech mívá však buď subklinický průběh nebo jsou zcela bezpříznakové (Gritsun *et al.*, 2003). Druhý možný způsob nákazy je požitím kontaminovaného a nepasterizovaného mléka, zejména kozího (Gritsun *et al.*, 2003). Třetím způsobem nákazy je vdechnutí kontaminovaného aerosolu, což se týká zejména laboratorních nálezů při práci s vysokými koncentracemi viru (Scherer *et al.*, 1980).

Po přisátí infikovaného klíštěte a uvolnění viru do hostitele dochází nejprve k jeho pomnožení v kůži a podkožním vazivu (Labuda *et al.*, 1996). Poté se virus šíří lymfatickým systémem do spádových lymfatických uzlin, kde se dále množí. Z lymfatických uzlin proniká virus do krevního řečiště a iniciuje se tak viremické stadium infekce (Málková *et al.*, 1959; McMinn *et al.*, 1996a). Šíří se krví a napadá různé orgány, zejména retikulo-endoteliálního systému (slezinu, játra a kostní dřev). Během viremické fáze dochází také k průniku viru hematoencefalickou bariérou do mozku (McMinn *et al.*, 1996a). Způsob, jakým encefalitogenní flaviviry bariéru překonávají a napadají mozek, není dosud zcela znám. Mezi možnými mechanismy jsou mimo jiné zmiňovány pasivní difúze (s ovlivněním permeability membrány TNF- α), transcytóza nebo šíření po čichovém nervu (Haglund *et al.*, 2003; Samuel *et al.*, 2006). Schopnost viru replikovat se v periferních tkáních, indukovat

virémii a proniknout do centrálního nervového systému je označována jako „neuroinvasivita“. Naopak schopnost viru vyvolat cytopatický efekt v mozku a způsobit encefalitidu se nazývá „neurovirulence“ (McMinn, 1997).

Jak již bylo zmíněno výše, patogenita jednotlivých subtypů VKE je, navzdory jejich antigenní a biologické podobnosti, rozdílná. Onemocnění evropským subtypem má mírnější charakter a také nižší úmrtnost (asi 1-2 %) ve srovnání s ostatními dvěma subtypy. Probíhá ve dvou fázích. Po přísátí nakaženého klíštěte je inkubační doba dlouhá asi 3-7 dní. Poté nastává první fáze onemocnění, doprovázená příznaky podobnými chřipce, jako je horečka, bolest hlavy, malátnost nebo únava. Mezi první a druhou fází je přibližně sedmidenní bezpříznakový interval. Druhá fáze je charakteristická neurologickými poruchami různé závažnosti, ve 20-30 % případů se objevují středně těžké meningitidy až těžké encefalitidy doprovázené páteřním ochrnutím. U 10-20 % pacientů s vážným průběhem onemocnění se vyskytují dlouhotrvající nebo permanentní neuropsychiatrické následky infekce (Burke *et* Monath, 2001; Gritsun *et al.*, 2003).

Infekce dálně-východním subtypem VKE vyvolává u člověka velmi vážné poruchy centrální nervové soustavy, jako jsou meningoencefalitidy nebo polyencefalitidy, doprovázené ztrátou vědomí a dlouhotrvajícími pocity vyčerpání. Nejzávažnější formy onemocnění zahrnují poškození neuronů v různých částech mozku a také poškození míchy. Mortalita se pohybuje mezi 20-60 %. Charakteristická je také absence chronické formy onemocnění (Gritsun *et al.*, 2003).

Infekce způsobená sibiřským subtypem vyvolává méně vážná onemocnění než infekce subtypem dálně-východním, je často doprovázena neparalytickými horečnatými formami encefalitidy. Úmrtnost se většinou pohybuje v rozsahu 6-8 % a chronické formy onemocnění jsou poměrně časté (Burke *et* Monath, 2001).

1.1.3 Molekulárně biologické charakteristiky viru

Viriony viru klíšťové encefalitidy mají průměr přibližně 50 nm a skládají se z elektronově hustého jádra, obklopeného lipidovou dvojvrstvou, která obsahuje dva obalové glykoproteiny – protein E („envelope“) a protein M („membrane“). Intracelulární (dosud nezralé) viriony obsahují prekurzor prM, z něhož se protein M vyštěpuje při průchodu virionu ven z buňky (Murphy, 1980). Jádro je tvořeno kapsidovým proteinem C a jednořetězcovou RNA pozitivní polarity o celkové délce 11 kb. Genomová RNA obsahuje jeden otevřený čtecí rámec („Open Reading Frame“ - ORF) a kóduje polyprotein tvořený asi

3400 aminokyselinami, který je ko-translačně a post-translačně štěpen buněčnými a virovými proteázami na tři strukturální proteiny (C, M a E) a sedm nestrukturních proteinů (NS1, NS2A, NS2B, NS3, NS4A, NS4B a NS5). Otevřený čtecí rámec všech flavivirů je na 5' a 3' konci ohraničen netranslatovanou oblastí („untranslated region“ – UTR) (Gritsun *et al.*, 2003).

Koncové oblasti

5' koncová oblast genomu flavivirů nese methylguanositovou čepičku (m⁷GpppAmp), je dlouhá 95-132 bazí a začíná konzervovanou dvojicí nukleotidů AG (Chambers *et al.*, 1990a). Kromě zmíněné dvojice nukleotidů obsahuje tato sekvence ještě několik dalších konzervovaných oblastí, které se pravděpodobně podílejí na formování sekundárních struktur RNA (Rice, 1996).

3' koncová oblast vykazuje výraznou heterogenitu jak v délce (~ 300 - 700 nt), tak i v nukleotidové sekvenci, a to nejen mezi různými druhy flavivirů, ale dokonce i mezi různými kmeny téhož viru. Rozdíly se vyskytují převážně v proximální části 3' oblasti, za STOP kodonem, kde byly pozorovány dlouhé delece, inserce nebo opakující se sekvence (Wallner *et al.*, 1995). Narozdíl od proximální části, distální část (samotný 3' konec), dlouhá přibližně 340 nukleotidů, je vysoce konzervovaná (Gritsun *et al.*, 1997).

3' oblast některých kmenů VKE je dlouhá pouze 350 nt, což znamená, že proximální část téměř chybí (Wallner *et al.*, 1995). Vzhledem k vysoké variabilitě nemá pravděpodobně tato část žádnou funkci při replikaci viru (Mandl *et al.*, 1998). Je možné, že tato hypervariabilní oblast slouží jako „mezerník“, který odděluje distální část od zbytku virového genomu, což je nutné pro navázání RNA polymerázy a buněčných faktorů, podílejících se na transkripci (Gritsun *et al.*, 1997).

Narozdíl od většiny dalších buněčných a virových molekul RNA s pozitivním řetězcem, RNA flavivirů postrádá na svém 3' konci poly(A) oblast. Některé kmeny VKE (Neudoerfl nebo 263) mají ale homoadenosinový úsek, vyskytující se v proximální části 3' koncové oblasti (Wallner *et al.*, 1995).

Distální část 3' konce je rozdělena do tří strukturálních regionů (I-III), obsahujících nezávislé sekundární struktury, které spolu pravděpodobně navzájem nijak neinteragují. Region I, navzdory podobnosti sekundárních struktur mezi jednotlivými flaviviry, vykazuje poměrně vysokou sekvenční variabilitu, a to i u blízce příbuzných kmenů. Oproti tomu region II se v sekundární struktuře u klíšťaty přenášených flavivirů výrazně liší od flavivirů přenášených komáry. Na samém konci regionu III vytváří posledních asi 90 bazí dlouhou stabilní vlásenku („long stable hairpin“ - 3' LSH), jejíž struktura je mezi flaviviry, navzdory

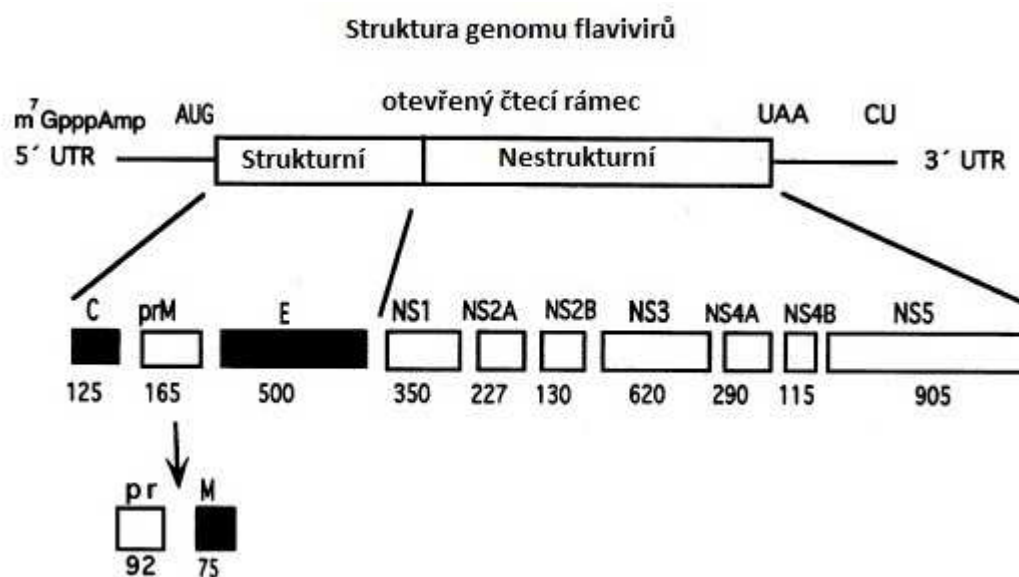
rozdílům v sekvenci, také konzervovaná. Předpokládá se, že tato vlásenka slouží ke specifické interakci mezi molekulou RNA a některými hostitelskými proteiny, které jsou součástí replikačního komplexu (Gritsun *et al.*, 1997; Proutski *et al.*, 1997).

RNA v obou koncových nekódujících oblastech tvoří sekundární struktury, které jsou většinou konzervované a pravděpodobně slouží pro regulaci amplifikace, translace a balení viru (Gritsun *et al.*, 1997; Proutski *et al.*, 1997).

V genomu flavivirů byly nalezeny dvě potenciální cyklizační sekvence – CS“A“ na 5' konci a CS“B“ na 3' konci. Tyto sekvence umožňují interakce mezi konci virového genomu a pravděpodobně mají určitou funkci při iniciaci replikace RNA, kdy je nutné, aby RNA-polymeráza nasedla současně na oba konce virového genomu (Khromykh *et al.*, 2001).

Strukturní proteiny

Tři geny kódující strukturní komponenty virionu – protein C, protein prM/M a protein E zabírají asi 25 % ORF na 5' konci genomu. Zbytek ORF (asi 75 %) kóduje nestrukturní proteiny v pořadí NS1, NS2A, NS2B, NS3, NS4A, NS4B a NS5 (Rice, 1996) (Obr. 1).



Obr. 1. Struktura genomu flavivirů (McMinn, 1997, upraveno).

Protein C

Tento vysoce bazický protein o velikosti 11 kDa vykazuje v porovnání s proteiny E a M mezi různými flaviviry méně sekvenčních homologií. Během translace je ze všech virových proteinů syntetizován jako první. O jeho struktuře a funkci je zatím známo poměrně málo (Mandl *et al.*, 1988). Velké množství zásaditých aminokyselinových zbytků na

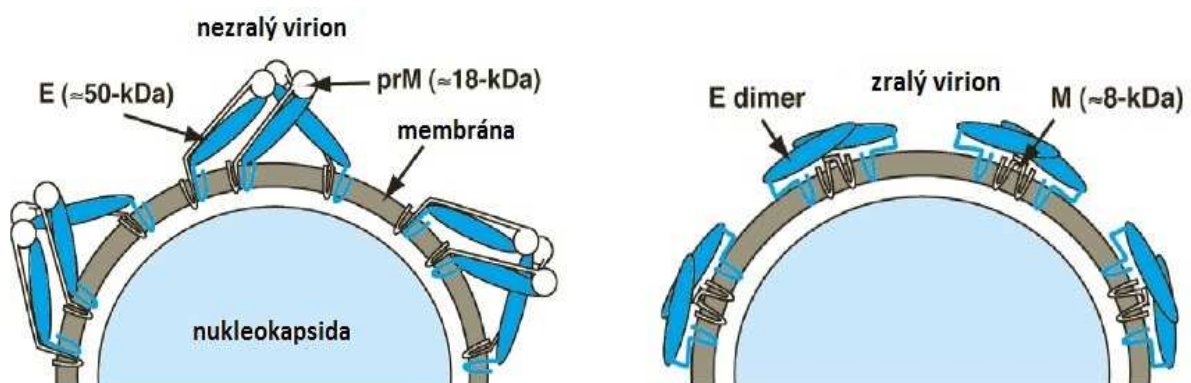
jeho koncích napovídá jeho možné funkci při balení virové RNA. Mezi těmito koncovými oblastmi se nachází vnitřní hydrofobní doména, zprostředkávající pravděpodobně interakci s membránou. Nascentní protein C obsahuje také C-terminální hydrofobní sekvenci, která slouží jako signální peptid a zahajuje translokaci proteinu prM do lumen endoplazmatického retikula. Později během syntézy je tato signální sekvence odštěpena virovou proteázou NS2B-NS3. Toto štěpení je nutné k následnému uvolnění N-konce proteinu prM (Kofler *et al.*, 2002).

Proteiny prM a M

Jak již bylo zmíněno výše, glykoprotein prM (26 kDa) je translokován do lumen endoplazmatického retikula prostřednictvím signální sekvence na C-konci proteinu C. N-konec proteinu je převážně hydrofilní a obsahuje 6 konzervovaných cysteinových zbytků, tvořících disulfidické můstky (Nowak *et Wengler*, 1987). Tato část proteinu obsahuje také jedno potenciální N-glykosylační místo. Protein prM je prekurzorem poměrně malého membránového proteinu M (8 kDa), který je složkou pouze zralých virionů a nachází se na vnějším povrchu virové částice (Chambers *et al.*, 1990a).

Kromě funkce prekurzoru má protein prM také úlohu chaperoninu, potřebného pro správné sbalení proteinu E. V časně fázi sekreční dráhy procházejí totiž viriony přes kyselé kompartmenty a vezikuly Golgiho aparátu a nízké pH by za nepřítomnosti prM způsobilo nevratnou konformační změnu proteinu E a tím vznik nefunkčního virionu. Protein prM má proto zásadní roli při maturaci virionu (Guirakhoo *et al.*, 1991).

Krátce předtím, než je virion uvolněn z buňky, je prM štěpen buněčnou proteázou furinem, vzniká protein M a dochází tím k maturaci virionu (Stadler *et al.*, 1997) (Obr. 2).



Obr. 2. Obalové proteiny nezralého a zralého virionu (Lindenbach *et al.*, 2007, upraveno).

Protein E

Protein E (53 kDa) je hlavním povrchovým proteinem virové částice. Interaguje s buněčnými receptory a zprostředkovává fúzi viru s buněčnou membránou. Také indukuje tvorbu virus neutralizujících protilátek, které hrají důležitou roli v ustavení ochranné imunitní odpovědi (Heinz, 1986). Při procesu vstupu viru do buňky se tento protein uplatňuje v interakci s receptorem při klatrinem zprostředkované endocytóze. V endosomu pak pokles pH způsobí změnu konformace proteinu E a umožní fúzi virové membrány s membránou endosomu, vedoucí k uvolnění virových částic do cytoplazmy (Heinz *et Allison*, 2000, 2001).

Trojrozměrná struktura proteinu E byla studována pomocí rentgenové krystalografie (Rey *et al.*, 1995). Protein se skládá ze dvou monomerních podjednotek zakotvených do obalové membrány. Každý z monomerů je tvořen třemi doménami označovanými I-III.

Doména I je centrální doménou. Obsahuje N-konec proteinu tvořený asi 120 aminokyselinami, uspořádanými do osmi antiparalelních beta listů, které dohromady tvoří strukturu beta soudku. Dále obsahuje dva disulfidické můstky a potenciální glykosylační místo v poloze 154, kde se na asparagin váže N-glykosidickou vazbou antenálně rozvětvený glykan (Grubhoffer *et al.*, 1990).

Doménu II je možné považovat za subdoménu domény I, protože je tvořena ze dvou nezávislých dlouhých smyček, které z domény I vyčnívají a tvoří tak prstovitou strukturu. Doména II je z většiny tvořena beta listy a jedna ze dvou smyček je stabilizována třemi disulfidickými můstky (Rey *et al.*, 1995). Část domény II je zodpovědná za pH-dependentní fúzi virové membrány s membránou endosomu (aminokyseliny v polohách 98 – 113 na cd smyčce) (Heinz *et Allison*, 2003; Hurrelbrink *et McMinn*, 2003). Při vystavení kyselému pH podléhá doména II největším konformačním změnám (Rey *et al.*, 1995).

Doména III připomíná svou strukturou molekulu imunoglobulinu (Heinz, 2003). Obsahuje C-konec proteinu, jehož asi 100 posledních aminokyselin vytváří beta soudek tvořený sedmi antiparalelními beta listy. Dále obsahuje jeden disulfidický můstek a je spojena s doménou I prostřednictvím 15 aminokyselin dlouhého řetězce. Laterální část domény III se pravděpodobně váže k dosud neznámému buněčnému receptoru (Rey *et al.*, 1995).

Nejvíce konzervovanou částí proteinu E je sekvence 12-ti cysteinových zbytků, které vytvářejí šest výše zmíněných disulfidických můstků v přesně definovaných oblastech. Jejich lokalizace je shodná u všech známých flavivirů (Nowak *et Wengler*, 1987).

Antigenní struktura proteinu E byla studována pomocí monoklonálních protilátek a je rovněž u všech flavivirů stejná. Protein E obsahuje tři antigenní domény: A, B a C (Heinz *et al.*, 1983; Roehrig *et al.*, 1983; McMinn *et al.*, 1995). Antigenní doména A je lokalizována na strukturní doméně II a nese epitopy pro zkříženou reaktivitu mezi různými flaviviry. Doména B se nachází na strukturní doméně III a nese epitopy odpovídající za subtypově specifickou reaktivitu, stejně jako doména C, ležící na doméně I (Rey *et al.*, 1995).

Nestrukturní proteiny

NS1 je s membránou asociovaný glykoprotein o velikosti 46 kDa, který je během syntézy translokován do endoplazmatického retikula a odštěpen od proteinu E hostitelskou signální proteázou (Falgout *et Markoff*, 1995). NS1 obsahuje tři potenciální glykosylační místa a 12 konzervovaných cysteinů, které tvoří disulfidické můstky (Blitvich *et al.*, 2001; Wallis *et al.*, 2004).

NS1 má významnou, ale dosud neznámou roli v replikaci virové RNA. Cílené mutace v glykosylačních místech proteinu vedly k dramatickým defektům RNA a tím také k omezené produkci viru (Muylaert *et al.*, 1996).

NS2A je malý hydrofobní protein o velikosti 22 kDa. Na svém N-konci je od proteinu NS1 odštěpen dosud neznámým hostitelským enzymem endoplazmatického retikula (Falgout *et Markoff*, 1995). Plní pravděpodobně funkci při formování virové částice. Mutace v tomto proteinu blokuje produkci viru (Kümmerer *et Rice*, 2002).

NS2B je také malý (14 kDa), s membránou asociovaný protein (Clum *et al.*, 1997). Vytváří stabilní komplex s proteinem NS3 a slouží jako kofaktor NS2B-NS3 serinové proteázy (Falgout *et al.*, 1991).

NS3 je druhým největším proteinem (70 kDa). Je to multifunkční protein, vysoce konzervovaný mezi všemi flaviviry. Jeho aktivita je nutná pro zpracování virového polyproteinu a pro replikaci RNA. Na jeho N-konci se nachází katalytická doména NS2B-NS3 serinové proteázy (Bazan *et Fletterick*, 1989; Chambers *et al.*, 1990b). C-koncová oblast kóduje helikázu, sloužící k replikaci virové RNA (Gorbalenya *et al.*, 1989). Posledním enzymem, kódovaným proteinem NS3, je RNA trifosfatáza, která slouží k defosforylaci 5' konce genomové RNA před přidáním čepičky (Wengler *et Wengler*, 1993).

NS4A (16 kDa) a **NS4B** (27 kDa) jsou malé hydrofobní proteiny (Lindenbach *et Rice*, 1999). NS4A se podílí na tvorbě replikačního komplexu (Mackenzie *et al.*, 1998). NS4B je transmembránový protein se zatím ne zcela známou funkcí (Lin *et al.*, 1993; Miller *et al.*, 2006).

NS5 je největší (103 kDa) a nejvíce konzervovaný protein VKE. Tento bazický, multifunkční protein má methyltransferázovou aktivitu a také funguje jako RNA-dependentní RNA polymeráza (Koonin, 1993). C-konec proteinu zahrnuje oblasti vysoce homologní k RNA-dependentním RNA polymerázám ostatních RNA virů s pozitivní polaritou genomu (Koonin *et Dolja*, 1993).

1.1.4 Cyklus viru v buňce

Jak již bylo zmíněno výše, na vstupu viru do hostitelské buňky se podílí strukturní protein E, který interaguje s receptorovými molekulami na povrchu buněk. Rozpoznání a vazba na receptor je proces, kterého se pravděpodobně účastní hned několik povrchových receptorových molekul současně. Domnělými primárními receptory flavivirů jsou proteiny DC-SIGN a DC-SIGNR. Oba tyto lektiny vážou N-glykany bohaté na manosu. DC-SIGN se vyskytuje na povrchu dendritických buněk, DC-SIGNR na povrchu endoteliálních buněk kapilár. Jejich podíl ve vazbě viru na hostitelskou buňku byl prokázán u viru dengue a viru západního Nilu, i když ke vstupu virionu do buňky je pravděpodobně ještě potřeba přítomnosti dalších receptorových molekul (Tassaneeritthep *et al.*, 2003; Davis *et al.*, 2006).

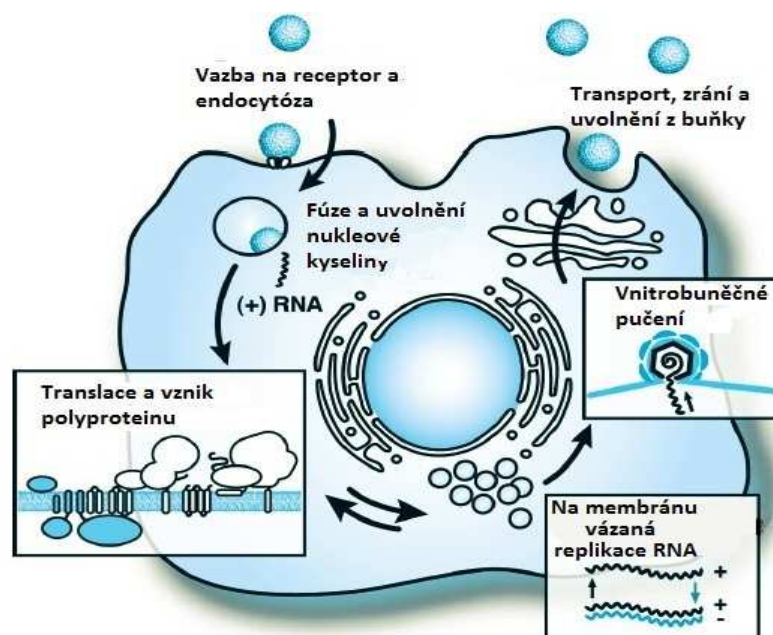
Významnými receptory jsou také glykosaminoglykany. Nacházejí se v hojném počtu na povrchu buněk a liší se navzájem svojí stavbou a množstvím mezi různými typy buněk a tkání (Bernfield *et al.*, 1999; Tumova *et al.*, 2000). Heparan sulfát je nejčastější glykosaminoglykan, se kterým viry interagují (Bruett *et al.*, 2000; Chen *et al.*, 1997). Předpokládá se, že afinita virové částice k heparan sulfátu by mohla být důležitou determinantou patogenity a tkáňového tropismu viru a mohla by hrát roli v jeho adaptaci na různé hostitele (Mandl *et al.*, 2001; Romanova *et al.*, 2007, Kozlovskaya *et al.*, 2010).

Po navázání viru na buněčný povrch dojde k receptorem zprostředkované, na klatrinu závislé endocytóze. Virus pučí v klatrinových váčcích do cytosolu a putuje k endosomům. Nízké pH uvnitř endosomů způsobí reorganizaci proteinu E z dimerů na trimery. Poté se protein E vztyčí a zasáhne do membrány endosomu, při čemž dochází k fúzi virového obalu s membránou endosomálního váčku a uvolnění nukleokapsidy do cytoplazmy hostitelské buňky (Heinz *et Allison*, 2001, 2003).

V cytoplazmě dojde k uvolnění RNA z nukleokapsidy a začíná translace virového polyproteinu, který je již během vzniku nebo ihned po něm glykosylován buněčnými glykosyltransferázami. Poté jsou z něj odštěpovány jednotlivé strukturní a nestrukturní proteiny činností buněčných a virových proteáz (Lindenbach *et al.*, 2007).

Současně probíhá v membránových strukturách poblíž endoplazmatického retikula (ER) také virová replikace. Replikace začíná syntézou řetězce RNA s negativní polaritou, který pak slouží jako templát pro tvorbu dalších řetězců s polaritou pozitivní. Syntéza virové RNA je asymetrická, kdy plus řetězců vzniká desetkrát více než minus řetězců (Muylaert *et al.*, 1996). Plus řetězce RNA slouží zároveň i jako mRNA, tedy jako templát pro vznik polyproteinu.

Po odštěpení jednotlivých proteinů z virového polyproteinu a replikaci RNA dochází ke vzniku nezralých a neinfekčních virionů v lumen ER. Na cytoplazmatické straně ER dochází ke vzniku nukleokapsidy interakcí proteinu C s virovou RNA. Obalové proteiny E a prM jsou naproti tomu po svém odštěpení od polyproteinu translokovány do lumen ER. Nezralé viriony vznikají pučením nukleokapsidy do lumen ER. Poté jsou sekrečními dráhami transportovány až do Golgiho aparátu. Během tohoto transportu dochází k dodatečným úpravám virového obalu (Lindenbach *et al.*, 2007). Proteiny E a prM jsou v nezralých virionech seskupeny ve formě heterodimeru. Při průchodu přes kyselé kompartmenty a vezikuly Golgiho aparátu by totiž nízké pH za nepřítomnosti prM způsobilo nevratnou konformační změnu proteinu E a tím vznik nefunkčního virionu. V Golgiho aparátu je protein prM štěpen hostitelskou proteázou furinem. Uvolňuje se část pr a vzniká protein M. Současně dochází k reorganizaci proteinu E do formy homodimeru, čímž se z nezralého virionu stává virion zralý. Zralé a infekční viriony se fúzí s cytoplazmatickou membránou uvolňují z buňky (Stadler *et al.*, 1997).



Obr. 3. Životní cyklus viru v hostitelské buňce (Lindenbach *et al.*, 2007, upraveno).

1.1.5 Glykosylace

Glykosylace je jednou z ko- a post-translačních modifikací proteinů. Je to proces, při kterém se na molekulu proteinu naváže oligosacharid za vzniku glykoproteinu. Jednou z nejčastějších modifikací proteinů je N-glykosylace, při které dochází k připojení manosylové vidličky na aminoskupinu (-NH₂) asparaginového postranního řetězce proteinu. Tento asparagin musí být součástí konzervovaného motivu Asn-X-Ser/Thr, kde X může být jakákoli aminokyselina s výjimkou prolinu (Chambers *et al.*, 1990a). K N-glykosylaci dochází již během syntézy virových obalových proteinů na ribozomech drsného ER. Při následném průchodu proteinů přes ER a Golgiho aparát jsou navázané oligosacharidy dále upravovány hostitelskými enzymy, čímž získávají svou konečnou strukturu. Různé enzymy upravují oligosacharidové řetězce různým způsobem. Výsledkem je glykoprotein s lišícími se oligosacharidovými strukturami. Viry využívají této služby hostitelských buněk k modifikaci svých proteinů, vyskytujících se na povrchu virionu, což naznačuje důležitost virových glykoproteinů v udržení stability virionu, v antigenicitě a invazi do hostitelských buněk (Vigerust *et Shepherd*, 2007).

Pouze tři proteiny v genomu flavivirů obsahují ve svých sekvencích potenciální glykosylační místa. Prvním z nich je protein prM, který má jedno až tři N-glykosylační místa. Druhým proteinem je obalový protein E, který buď neobsahuje žádné nebo maximálně dvě potenciální glykosylační místa. Třetím proteinem je nestrukturní protein NS1, který obsahuje ve svojí sekvenci dvě až tři potenciální glykosylační místa, z nichž dvě jsou unikátní pro VKE a třetí je konzervované u všech flavivirů (Chambers *et al.*, 1990a).

N-glykosylaci lze zablokovat antibiotikem tunikamycinem, obsahujícím glukosamin. Tunikamycin specificky inhibuje N-glykosylaci, zprostředkovanou dolichol pyrofosfátem. Protože je tato reakce nutná pro tvorbu hlavní sekvence N-glykosydicke vázaných oligosacharidů, zabraňuje tunikamycin N-glykosylaci proteinů. Antibiotikum tunikamycin lze proto využít ke studiu významu glykosylace u virů (Olden *et al.*, 1979).

1.2 Variabilita VKE

1.2.1 Variabilita RNA virů ve srovnání s DNA viry

Viry užívají nejrůznější funkční a adaptivní strategie, které jim pomáhají vyhnout se obranným mechanismům svých hostitelů. Tyto strategie mají odlišný charakter mezi RNA a DNA viry.

RNA viry využívají častých mutací k tomu, aby rozšířily spektrum hostitelů, aby unikly z dosahu protilátek nebo cytotoxických lymfocytů (McKnight *et al.*, 1995; McMichael *et Phillips*, 1997). Díky jejich limitované genetické komplexitě, dané spíše menší velikostí jejich genomů (3-30 Kb), jsou RNA viry tolerantní k většímu počtu mutací a tudíž velmi variabilní (Eigen *et Schuster*, 1979; Domingo *et Holland*, 1997). Naproti tomu DNA viry s velkým genomem (např. herpesviry, poxviry, adenoviry nebo iridoviry) mají komplexní genetickou informaci, kde tlak na její zachování snižuje jejich celkovou toleranci k mutacím a jejich genom je proto stabilní (Domingo *et al.*, 1995; Domingo *et Holland*, 1997). Vyvinuly si ale jiný mechanismus, jak se bránit imunitní odpovědi hostitelského organismu, který by se dal označit jako tzv. „proteinové mimikry“. Jsou schopné potlačit svými proteiny některé povrchové molekuly buňky nutné pro její rozpoznání imunitním systémem nebo naopak napodobit extracelulární receptory a domény pro cytokiny, makrofágy a další buňky imunitního systému (Murphy, 1993; Alcamí *et al.*, 1998).

To, že jsou viry s jednodušším genomem tolerantnější k mutagenезi než viry komplexnější, dokazuje vyšší počet pozorovaných mutací v jednoduchých replikonech a také evoluce „proofreadingových“ a postreplikačních mechanismů, nutných během replikace DNA u některých komplexních DNA virů (Holland *et al.*, 1982; Domingo *et al.*, 1995; Domingo *et Holland*, 1997).

1.2.2 Evoluce VKE a stabilita jeho genomu

Evoluce a epidemiologie flavivirů je do značné míry dána ekologickými potřebami jejich hostitelů, krevsajícími členovci a obratlovci.

Molekulárně biologické studie poukazují na to, že se flaviviry přenášené klíšťaty vyvíjely pozvolna během několika posledních tisíciletí, kdy se šířily severním a západním směrem přes lesy Asie a Evropy (Zanotto *et al.*, 1995; Gould *et al.*, 2003).

K tomu, že se VKE vyvíjel velice pozvolna, přispělo několik faktorů. Prvním z nich je dlouhý životní cyklus klíšťat (*Ixodes spp.*), který může trvat až pět let. Druhým faktorem je fakt, že toto klíště saje krev během svého životního cyklu pouze třikrát, jako larva, nymfa a dospělec. Třetím faktorem je závislost klíšťat na určitém životním prostředí, jako jsou lesy s vlhkými a hustými podrosty nebo mokřiny. Posledním faktorem přispívajícím k postupné evoluci VKE je fakt, že se virus může přenášet přímo mezi klíšťaty, a to buď transovariálně nebo prostřednictvím „co-feedingu“ (Labuda *et al.*, 1997; Gould *et al.*, 2003).

Významným rysem evoluce VKE je diverzita jeho subtypů. Ta je mnohem větší u subtypu sibiřského a dálně-východního, které jsou přenášeny klíštětem *I. persulcatus*, než u evropských kmenů viru přenášených klíštětem *I. ricinus*. Zatímco kmeny sibiřského a dálně-východního subtypu VKE se vyvíjely pravděpodobně tisíce let, kmeny evropského subtypu se rozšířily během několika posledních století (Kuno *et al.*, 1998; Grard *et al.*, 2007).

Variabilita aminokyselinových sekvencí proteinu E v rámci téhož subtypu VKE je poměrně nízká, dosahuje pouze 2,2 %. Variabilita mezi jednotlivými subtypy je 5,6 % (Ecker *et al.*, 1999). Na rozdíl od komáry přenášených flavivirů nebyly také dosud u klíštěcích flavivirů pozorovány žádné genetické rekombinace (Grard *et al.*, 2007). Z toho vyplývá, že VKE, navzdory všeobecně vysoké mutační rychlosti RNA virů, je geneticky velmi stabilní. Jeho stabilita je pravděpodobně způsobena vysokým stupněm adaptace viru na ekosystém a cyklem přenosu, který zahrnuje dva velmi odlišné hostitele (Ecker *et al.*, 1999).

1.2.3 Pasážování viru

I přesto, že je genom VKE považován za geneticky velmi stabilní, byly v minulosti identifikovány varianty viru podstatně se lišící biologickými vlastnostmi, jako jsou míra virulence pro laboratorní myši, velikost plaků nebo citlivost vůči teplotě (Mayer *et al.* 1969; Kopecký *et al.*, 1991). Takové varianty je možné vyselektovat i uměle pomocí dlouhodobého pasážování viru v určitém typu prostředí. Varianty viru lépe přizpůsobené jednomu prostředí jsou pak znevýhodněny v prostředí druhém. Hostitel působí na virus určitým selekčním tlakem a mění jeho genotyp, díky tomu může při pasážování docházet i ke změně virové neuroinvazivity nebo neurovirulence (Labuda *et al.*, 1994; Romanova *et al.*, 2007; Mitzel *et al.*, 2008).

Dlouhodobým pasážováním viru v klíšťatech byly vyselektovány varianty méně virulentní pro laboratorní myši, tvořící menší plaky v kultuře savčích buněk a rychleji se množící v klíšťatech (Chunikhin *et al.*, 1986; Dzhivanian *et al.*, 1986, 1988, 1991). Pasážování VKE ve slinných žlázách klíštěte *I. ricinus* mělo rovněž za následek sníženou virulenci pro obratlovce. Přestože po sedmé pasáži viru na klíštěti přežilo 95 % infikovaných myší, které vykazovaly velmi slabou nebo dokonce žádnou virémii, bylo možné z jejich mozků virus izolovat. Z této skutečnosti se dá usuzovat, že i virus pasážovaný na klíšťatech způsobuje u myší perzistentní infekci (Labuda *et al.*, 1994).

Kromě takto zaznamenaných změn byl však zároveň také pozorován i spontánní návrat k původnímu fenotypu. Pasážováním VKE v klíštěti *Hyalomma plumbeum* došlo

k výraznému snížení virulence pro laboratorní myš, ale následné pasážování viru v myších mozcích jeho původní fenotypové vlastnosti obnovilo (Dzhivanian *et al.*, 1988). Obdobně i ve studii Kaluzové *et al.* (1994) došlo k podobnému zvratu, od atenuovaného fenotypu po pasážování v klíšťatech, k původnímu fenotypu po následném pasážování v myších mozcích. Po obnovení původního virulentního fenotypu byl virus dále pasážován v myších mozcích a byla tak vyselektována dokonce ještě virulentnější varianta, než původní (Kaluzová *et al.*, 1994).

Podobně jako lze dlouhodobým pasážováním v klíšťatech získat méně virulentní varianty viru, je možné dlouhodobým pasážováním viru v mozcích sajících myšek zvýšit jeho virulenci. Atenuovaný termosenzitivní kmen VKE (263), izolovaný z klíštěte *I. ricinus*, byl sériově pasážován dvojím způsobem: po pěti pasážích v myších vznikl kmen viru 263-m5; jednou pasáží původního kmene v buněčné kultuře odvozené z prasečích ledvin za zvýšené teploty (40°C) byl vyselektován kmen viru označený jako 263-TR. Oba tyto kmene byly termorezistentní, vysoce virulentní pro laboratorní myši, a vytvářely větší plaký v kultuře savčích buněk v porovnání s původním kmenem 263. Navíc měly oba nově vzniklé kmene identické genomy. Sekvenováním byly zjištěny dvě unikátní aminokyselinové záměny oproti původní variantě, v nestrukturních proteinech NS2B a NS3 a také téměř identická delece v 3' UTR oblasti (Růžek *et al.*, 2008).

Mikroevolucí VKE v závislosti na změně hostitele viru se zabývá i další studie. Sibiřský kmen VKE izolovaný z klíštěte *I. persulcatus* byl pasážován 5x v myších mozcích (varianta EK-328) a 17x v klíštěti *Hyalomma marginatum* pomocí umělé inokulace a poté ještě znovu 5x v myších (varianta M). Poté byly srovnávány vlastnosti obou nově vyselektovaných variant. Varianta M, adaptovaná na klíšťata, vykazovala pomalejší replikaci v kultuře savčích buněk, byla méně virulentní pro laboratorní myši po subkutánní inokulaci (snížená neuroinvasivita) a naopak lépe se množila v klíšťatech. Co se týče velikosti plaků, při plakové titraci vytvářela varianta M většinou malé plaký, pouze výjimečně i plaký velké, srovnatelné s velikostí plaků u varianty EK-328 (1-2 velké plaký na 100 malých plaků). Při srovnání genomu bylo detekováno celkem 15 nukleotidových substitucí u varianty M ve srovnání s variantou EK-328, z čehož 6 záměn mělo za následek změnu v aminokyselinové sekvenci. Tyto aminokyselinové záměny se vyskytovaly ve strukturních proteinech E a prM a v nestrukturních proteinech NS2A a NS4A. Pro zjištění, která z těchto aminokyselinových substitucí se podílí na změně fenotypu u variant adaptovaných na různá prostředí, byly izolovány jeden velký a několik malých plaků varianty M a pasážovány v kultuře savčích buněk. Klon izolovaný z velkého plaku si udržel tento fenotypový znak již po jedné pasáži.

Klony odvozené z malých plaků tvořily po první pasáži buď přímo velké plaky nebo plaky smíšené. Po dalších dvou pasážích v savcích buňkách navíc byly malé plaky zcela nahrazeny plaky velkými. Tyto klony, tvořící velké plaky, si zachovaly aminokyselinové substituce v nestrukturních proteinech, ale v proteinu E došlo k opětovným změnám a předpokládá se, že tyto změny zodpovídají za rozdíly ve fenotypových vlastnostech (Romanova *et al.*, 2007).

Výše zmíněné studie poukazují na potenciální možnost výskytu tzv. „quasispecies“ VKE, jejichž přítomnost by mohla být vysvětlením velmi rychlých změn probíhajících v genomu viru během sériových pasáží (Labuda *et al.*, 1994). Výsledky naznačují, že se virus vyskytuje ve formě heterogenních populací, které obsahují jak varianty viru adaptované na reprodukci v klíšťatech (atenuované varianty), tak varianty adaptované na reprodukci v savcích (virulentnější varianty). Změna hostitele má za následek změnu v poměru těchto variant v populaci (Romanova *et al.*, 2007). Rychlá změna neurovirulence během adaptace viru na klíšťata/savce je tedy způsobená spíše selekcí příslušných „quasispecies“, než náhodnými mutacemi vzniklými během pasážování viru (Růžek *et al.*, 2008).

1.2.4 „Quasispecies“

V 70. letech minulého století navrhl Manfred Eigen se svými kolegy nový model molekulární evoluce, vysvětlující adaptabilitu a rychlý vývoj jednoduchých replikonů, které pravděpodobně sídlily na Zemi v počátečních stádiích vývoje života. Tento evoluční model kladl důraz na vývoj pomocí mutací a vyvrátil tak koncept divokých („wild-type“) genomů jako pevně definovaných sekvencí nukleotidů. Genetickou organizaci jednoduchých replikonů na populační úrovni, poprvé navrženou Eigenem, Schustrem a jejich kolegy, charakterizují „quasispecies“ (Eigen, 1971; Eigen *et Schuster*, 1979).

Termínem „quasispecies“ je ve virologii míněna heterogenní a kontinuálně se vyvíjející populace virionů RNA virů v rámci téhož hostitele, mající společný původ. Nukleotidové záměny, kterými se od sebe jednotlivé viriony liší, mohou mít vliv na jejich virulenci. Tyto blízké příbuzné, mutantní a často také rekombinantní virové genomy vznikají jako důsledek mutací, genetického driftu, kompetice a selekce (Domingo *et al.*, 1978; Eigen *et Schuster*, 1979; Domingo *et al.*, 1995; Domingo *et Holland*, 1997; Smith *et al.*, 1997).

Bakteriofág Q β , replikující se v *E. coli*, byl prvním, u jehož populace byly popsány znaky „quasispecies“. Od té doby je tento termín používán v souvislosti s heterogenními populacemi i u mnoha dalších virů, zejména viru lidské imunitní nedostatečnosti (HIV) a viru hepatitidy C, ale také u některých picornavirů, rhabdovirů, hepadnavirů, arenavirů a dalších (Holland *et al.*, 1982; Domingo *et al.*, 1997).

„Quasispecies“ mají mnoho biologických důsledků, zahrnujících rychlou antigenní proměnlivost, změny v tkáňovém tropismu, ve virulenci, rozsahu hostitelů a expresi viru (Domingo, 1989, 1996; Domingo *et al.*, 1997; Sa-Carvalho *et al.*, 1997). Některé z těchto důsledků mají přímý vliv i na perzistenci viru v přírodě.

1.2.5 Molekulární determinanty virulence

Jak již bylo zmíněno výše, perzistence VKE v přírodě vyžaduje schopnost viru replikovat se ve dvou velmi odlišných hostitelích. Během replikace viru v klíšťatech dochází k selekci, díky které jsou potom virem, množící se v různých druzích klíšťat, různě patogenní pro člověka (Leonova, 1997). I kmeny přenášené stejným druhem klíštěte se ale mohou navzájem lišit. Při srovnávání genomů evropského prototypového kmene VKE Neudoerfl s virulentnějším kmenem Hypr (oba přenášené klíštětem *I. ricinus*) bylo identifikováno velké množství nukleotidových záměn v rámci celého genomu, vedoucích ke změnám v sekvenci aminokyselin, zodpovědných za rozdíly v neuroinvasivitě a neurovirulenci (Wallner *et al.*, 1996).

Jednou ze strategií, vedoucích k pochopení, jak hostitel ovlivňuje replikaci a patogenitu viru, je identifikace molekulárních determinant virulence. V případě klíšťaty přenášených flavivirů nejsou determinanty virulence, specificky se uplatňující při replikaci v klíštěcích nebo savčích hostitelích, zatím přesně známy, stejně jako mechanismus, jakým tyto determinanty ovlivňují patogenitu viru (Mitzel *et al.*, 2008).

V případě VKE je za hlavní determinantu virulence všeobecně považován obalový protein E a to zejména díky jeho zásadním funkcím během životního cyklu VKE. Mezi tyto funkce patří interakce s buněčnými receptory, fúze viru s membránou endosomu a také indukce tvorby virus neutralizujících protilátek. Mutace v jakékoli ze tří domén proteinu E mohou ovlivnit virulenci (Holzmann *et al.*, 1997). Souhrnně pro všechny flaviviry platí, že mutace, mající největší vliv na atenuaci viru, se nacházejí na laterální straně domény III (tato část se pravděpodobně podílí na vazbě na receptor), dále na rozhraní mezi doménami I a II a nakonec na vrcholu domény II a v místě jejího kontaktu s druhou podjednotkou dimeru (poslední dvě zmíněné oblasti proteinu E se pravděpodobně účastní fúze s buněčnou membránou) (Rey *et al.*, 1995; McMinn, 1997).

Několik studií potvrdilo, že stačí jediná aminokyselinová záměna v těchto kritických oblastech proteinu E k tomu, aby došlo ke ztrátě neuroinvasivity, aniž by byla zároveň také ovlivněna neurovirulence nebo snížena schopnost viru replikovat se *in vitro*. Atenuace těchto

variant je pravděpodobně způsobena sníženou schopností proteinu E zprostředkovat vstup viru do příslušných buněk periferních tkání hostitele pomocí receptorem zprostředkované endocytózy (Holzmann *et al.*, 1990; Cecilia *et al.*, 1991; McMinn *et al.*, 1995; McMinn *et al.*, 1996b; Holzmann *et al.*, 1997).

Ve výše zmíněné studii Labudy *et al.* (1994) byla pasážováním viru v klíšťatech vyselektována méně virulentní varianta. Rozdíl v sekvenci původního viru a atenuované varianty v rámci proteinu E vedl k jediné unikátní změně aminokyseliny (Glu (84) → Lys). Tato změna pravděpodobně způsobila zvýšení pozitivního náboje proteinu E a ovlivnila místo na tomto obalovém proteinu, kterým se váže na receptory erytrocytů. Tato varianta viru následně zcela postrádala hemaglutinační aktivitu.

Protein E se jako významná determinanta virulence ukázal i ve studii Romanové *et al.* (2007). Sibiřský kmen VKE izolovaný z klíštěte *I. persulcatus* zde byl pasážován ve dvojnásobném prostředí: v myších mozcích (varianta EK-328) a v klíštěti *Hyalomma marginatum* a poté ještě znovu v myších (varianta M). Mezi těmito dvěma nově vyselektovanými variantami bylo identifikováno šest aminokyselinových záměn, z nichž dvě byly v proteinu E. První z těchto substitucí, Thr (426) → Ile, se nacházela v konzervované sekvenci v místě dvou předpokládaných alfa-helixů (Allison *et al.*, 1999). Tato oblast patří v rámci flavivirů k nejkonzervovanějším (Stiasny *et al.*, 1996). U VKE se však kmeny přenášené různými druhy klíšťat liší právě v aminokyselině v pozici 426. Kmeny evropského subtypu, jejichž hlavním přenašečem je klíště *I. ricinus*, mají v této poloze alanin, zatímco kmeny sibiřského a dálněvýchodního subtypu, s hlavním přenašečem *I. persulcatus*, mají v poloze 426 threonin. V této studii, kdy byl virus izolován z klíštěte *I. persulcatus* a adaptován na klíště *H. marginatum*, došlo ke změně threoninu na izoleucin. Druhá aminokyselinová záměna, Glu (122) → Gly, již byla identifikována v předchozích studiích (Mandl *et al.*, 2001; Kroschewski *et al.*, 2003). Tato substituce má za následek zvýšení pozitivního náboje na proteinu E a zvýšenou afinitu ke glykosaminoglykanům, což vede ke dvěma zásadním fenotypovým změnám – k tvorbě malých plaků v kultuře savčích buněk a ke snížené neuroinvasivitě u myší (Mandl *et al.*, 2001).

Na klíšťata adaptovaná varianta M (Romanova *et al.*, 2007) měla vlastnosti podobné mutantním variantám, vyselektovaným adaptací VKE na nádorovou linii buněk, získaných z ledviny mláďete křečka („baby hamster kidney cells“ – BHK-21) (Mandl *et al.*, 2001). V této studii bylo provedeno celkem 16 nezávislých pokusů, které vedly k získání 12 kombinací různých substitucí v rámci celého proteinu E. Substituce Glu (122) → Gly se opakovala celkem třikrát. Avšak nejen tato substituce, ale i všechny ostatní, měly za následek zvýšení pozitivního náboje na proteinu E. Bylo vytvořeno několik infekčních klonů, nesoucích různé

kombinace těchto mutací. Všechny tyto klony projevovaly zvýšenou infektivitu pro BHK buňky, způsobenou zvýšenou vazbou na heparan sulfát. Zároveň tvořily malé plaky v kultuře buněk PS (buňky odvozené z prasečí ledviny) a vykazovaly sníženou neuroin vazivitu pro laboratorní myši.

Nejen změny v proteinu E vedou ke změnám fenotypu viru, jak bylo potvrzeno v dalších studiích. Pro zjištění determinant virulence, spojených s replikací v různých hostitelích, byl virus Langat (varianta TP21 pp) adaptován na růst v klíštěcích buněčných liniích (varianta ISEp20). Varianta ISEp20 se od původní lišila třemi aminokyselinovými substitucemi, v proteinech M, NS3 a NS4A. Tato klíštěcí varianta vykazovala atenuovaný fenotyp pro laboratorní myši. V této studii byl potvrzen význam nejen strukturních, ale také nestrukturních proteinů v adaptaci na hostitele (Mitzel *et al.*, 2008).

Obdobně ve studii Růžka *et al.* (2008) bylo prokázáno, že se nestrukturní proteiny při adaptaci viru na hostitele uplatňují. Konkrétně došlo k mutacím v proteinech NS2B a NS3, které ovlivnily neuroin vazivitu pro laboratorní myši a měly rovněž za následek rezistenci viru vůči vyšším teplotám.

Avšak i změny v nekódujících oblastech virového genomu ovlivňují fenotypové vlastnosti viru. Mutanty VKE s delecí ve variabilní (proximální) části 3' UTR oblasti nemají žádný vliv na fenotyp viru. Naopak delece ve vysoce konzervované (distální) části 3' UTR oblasti způsobují, že je virus značně atenuovaný pro laboratorní myši, ačkoliv v tkáňových kulturách nevykazuje žádné změny v růstu (Mandl *et al.*, 1998). Tyto změny fenotypu nejsou způsobeny pouze ztrátou určitých sekvenčních motivů 3' UTR oblasti, ale spíše přerušáním nebo přeorganizováním jinak konformačně konzervovaných sekundárních struktur RNA (Proutski *et al.*, 1999).

Souhrnně z výše jmenovaných studií vyplývá, že virulence je multifaktoriální vlastnost a že každá mutace, ovlivňující jakýmkoliv způsobem replikaci viru v těle hostitele, má pravděpodobně také vliv na jeho virulenci (Heinz, 2003).

2. Cíle práce

- u Příprava variant vybraného kmene klíšťové encefalitidy mnohonásobným pasážováním v kulturách klíštěcích a savčích buněk.
- u Stanovení biologických vlastností nově získaných variant viru.
- u Získání celogenomových sekvencí, porovnání sekvencí s biologickými vlastnostmi, pokus o identifikaci možných determinant biologických vlastností na úrovni genomu.
- u Pokus o průkaz variability vybraných částí genomu v rámci jednotlivých izolátů.

3. Materiál a metody

3.1 Virus

Kmen viru klíšťové encefalitidy Hypr byl izolován na území Moravy v roce 1953 z krve nemocného desetiletého dítěte (Pospíšil *et al.*, 1954). Virus byl poté pětkrát pasážován v mozcích sajících myšek. Ve formě 20% mozkové suspenze byl uchováván při $-70\text{ }^{\circ}\text{C}$.

3.2 Buněčné linie

Virus byl pasážován *in vitro* ve dvou odlišných prostředích – na savčí buněčné linii PS a na klíštěcích buňkách IRE/CTVM19. Kromě pasážování viru byly buňky PS používány i pro plakovou titraci.

Buněčná linie PS je odvozená z prasečí ledviny („porcine kidney stable“). Kultivace těchto buněk probíhala v médiu L15 (Leibowitz) s 3 % prekolostrálního telecího séra (PTS), 1 % směsi antibiotik a antimykotik (penicilin, streptomycin, amphotericin B) (Sigma) a 1 % glutaminu při $37\text{ }^{\circ}\text{C}$. Pasáž buněk byla prováděna metodou suché trypsinizace.

Buněčná linie IRE/CTVM19 je odvozena z buněk klíštěte *I. ricinus* (Bell-Sakyi, 2004). Tyto buňky byly kultivovány v médiu L15 (Leibowitz) s 10 % tryptosofosfátového bujónu (TPB), 20 % bovinního fetálního séra (BOFES), 1 % směsi antibiotik a antimykotik a 1 % glutaminu při $28\text{ }^{\circ}\text{C}$. Pasáže buněk byly prováděny naředěním do čerstvého kultivačního média.

3.3 Pasážování viru

Počáteční koncentrace viru před první pasáží byla 3×10^4 PFU/ml. Pasážování probíhalo v jednotlivých jamkách 24-jamkového panelu, v odpovídajícím médiu pro buňky PS a IRE/CTVM19. K 500 μl média s buňkami (počet buněk PS přibližně 10^6 , počet buněk IRE/CTVM19 přibližně 10^5) bylo přidáno 100 μl virové suspenze. Buňky s virem byly inkubovány při odpovídajících teplotách ($37\text{ }^{\circ}\text{C}$ pro PS a $28\text{ }^{\circ}\text{C}$ pro IRE/CTVM19) v atmosféře 0,5% CO_2 po dobu čtyř dní. Poté byla veškerá buněčná suspenze s virem přenesena z jamky do 1,5 ml mikrozkuřavky a zmrazena při $-70\text{ }^{\circ}\text{C}$. Při opětovném rozmrazení došlo k popraskání buněk a vylití jejich obsahu, včetně dosud neuvolněných virových částic. Suspenze byla dále centrifugována při 2650 g 5 min. a supernatant přenesen do nové 1,5 ml mikrozkuřavky. Na další pasáž viru bylo použito 100 μl tohoto supernatantu, zbytek byl uchováván při teplotě $-70\text{ }^{\circ}\text{C}$. Opakováním stejného postupu bylo

provedeno celkem 50 pasáží viru v každém z typů buněk a získány dvě odlišné varianty viru.

Kromě krátkodobých pasáží byla na buňkách IRE/CTVM19 provedena i dlouhodobá pasáž. 100 μ l původní virové suspenze (koncentrace 3×10^4 PFU/ml) bylo inkubováno v 500 μ l kultivačního média pro klíštěcí buňky při 28 °C v atmosféře 0,5% CO₂ po dobu jednoho měsíce. Po prvních 14 dnech bylo ke vzorku přidáno dalších 500 μ l čerstvého média. Po měsíci byl odebrán celý objem suspenze do 1,5 ml mikrozkuřavky, zmražen a poté centrifugován při 2650 g 5 min. Následně byl vzniklý supernatant přenesen do nové mikrozkuřavky a uchován při -70 °C.

Třetím odlišným způsobem pasážování byla získána další varianta viru. Při tomto pasážování bylo k médiu s buňkami přidáváno antibiotikum tunikamycin, který inhibuje N-glykosylaci. Pasáže probíhaly na buňkách PS. K 500 μ l média s buňkami byl přidán nejprve tunikamycin (koncentrace v médiu byla 0,5 μ g/ml) a poté 100 μ l původní virové suspenze. Vše bylo inkubováno při 37 °C v atmosféře 0,5% CO₂ po dobu čtyř dní. Po odebrání buněčné suspenze s virem, jejím zmražení a následné centrifugaci, byl získaný supernatant použit na další pasáž, při které se opakoval stejný postup. Pasážování viru s tunikamycinem bylo ukončeno u třetí pasáže.

3.4 Plaková titrace viru

Plaková titrace je metoda sloužící ke stanovování množství virových částic ve vzorku. Výsledná koncentrace virových částic je vyjádřena počtem plakotvorných jednotek (PFU – „plaque forming units“) na 1 ml virové suspenze. Plaková titrace byla prováděna na savcích buňkách PS (De Madrid *et* Porterfield, 1969) ve 24-jamkových panelech. Virová suspenze byla naředěna desítkovou řadou v kultivačním médiu pro buňky PS. Poté bylo k naředěným vzorkům přidáno po 300 μ l buněčné suspenze o koncentraci $4-5 \times 10^5$ buněk/ml a obsah panelu byl řádně promíchán. Po 4 hodinách inkubace při 37 °C v atmosféře 0,5% CO₂ byly jamky převrstveny 400 μ l přelivu (roztok 1,5% karboxymethylcelulózy a 2x koncentrovaného kultivačního média v poměru 1:1). Vše bylo znovu inkubováno při 37 °C v atmosféře 0,5% CO₂ po dobu 5 dní. Po této době byly vzniklé plaky vizualizovány barvením. Panel byl nejprve promyt ve fyziologickém roztoku a poté byly živé buňky obarveny v roztoku naftalenové černě (1 g naftalenové černě, 60 ml ledové kyseliny octové, 13,6 g octanu sodného, doplněno do 1 l destilované vody), čímž došlo ke zviditelnění ložisek virem zlyzovaných buněk, tzv. plaků.

3.5 Izolace virové RNA

Izolace virové RNA byla prováděna pomocí kitu *QIAamp Viral RNA Mini Kit* (QIAGEN) dle návodu výrobce. V posledním kroku byla purifikovaná RNA eluována z kolonky 60 µl vymývacího pufru do 1,5 ml mikroskopické kumavky a uchovávána při -70 °C.

3.6 Reverzní transkripce

Vyizolovaná virová RNA byla prostřednictvím reverzní transkripce přepsána do cDNA. Reverzní transkripce byla prováděna pomocí kitu *First Strand cDNA Synthesis Kit* (Fermentas) podle instrukcí výrobce. Pro syntézu cDNA byl použit reverse primer 1F (0,01 mM) (Tab. 1). cDNA byla dále používána jako templát pro PCR reakci a uchovávána při -70 °C.

3.7 Polymerázová řetězová reakce (PCR)

Amplifikace cDNA probíhala pomocí polymerázové řetězové reakce. Amplifikační směs pro jednu reakci obsahovala 12,5 µl Combi PPP Master Mixu (2x koncentrovaný: 150 mM Tris-HCl, pH 8,8; 40 mM (NH₄)₂SO₄; 0,02% Tween 20; 5 mM MgCl₂; 400 µM dNTPs; 100 U/ml Taq Purple DNA polymeráza; protilátka anti-Taq DNA polymeráza (38nM); stabilizátory; aditiva) (Top-Bio), 4 µl templátové cDNA a 1 µl od každého z příslušných „forward“ a „reverse“ primerů (0,01 mM; přehled všech použitých primerů je uveden v Tab. 1). Reakční směs byla doplněna do celkového objemu 25 µl deionizovanou vodou (Top-Bio). PCR reakce probíhala v přístroji *Master cycler* nebo *Master cycler Gradient* (Eppendorf) za následujících reakčních podmínek:

1. predenaturace 95 °C, 5 min.
2. 30 cyklů: a) denaturace 94 °C, 30 s
b) nasedání primerů 42-55 °C *, 30 s
c) elongace 72 °C, 1 min.
3. postelongace 72 °C, 5 min.
4. udržovací teplota 14 °C

* Teplota nasedání primerů byla závislá na použité kombinaci primerů (Tab. 1).

Tab. 1 Použité primery.

Název primeru	Polarita*	Sekvence primeru 5'→3'	Teplota nasedání [°C]	Poloha v genomu**
1C	F	AGA TTT TCT TGC ACG T	48	1-16
1D	R	TTC CTT CAG CTC TGA TC	48	532-516
2C	F	GGT CAT TAC TCT GTT GGG GA	45	450-469
2D	R	ACA CAT CCA TTG AAG GCT TC	45	1102-1083
1A	F	TGT CTA CGC TTC GCG TTG CAC ACA	55	963-986
1B	R	GGA ACC CCA GCG AGA GCC TTC AGT A	55	1787-1763
2A	F	CGT GTT GAA TTT GGG GCT CCT CAC G	54	1693-1717
2B	R	TTC GTT CCG TGT CCA CAG CGC A	54	2491-2470
1M	F	GCT TTC AAC AGC ATC TTC GG	49	2302-2321
1N	R	TGT TAT ACC AGG ATC TGG GTC C	49	3229-3208
2M	F	ACT TTT GGT CAC TGA CCT GAG G	49	3105-3126
2N	R	TCC ATG ATG CAC AGT CAC CA	49	4032-4013
1O	F	GCT GGT CCT GGA ATT GGG GC	52,8	3855-3874
1P	R	GGA CCG GCC ACA GCA TCA TC	52,8	4808-4789
2O	F	GGT CGA GAG CGT GGT GAT AG	50,8	4627-4646
2P	R	CAG GAG GTG TCG CTG TCA TC	50,8	5575-5556
1R	F	CCT ATG TCA ACA GAC GGC TAC T	50	5408-5429
1S	R	GGC TCC TAA AGG TAA CCA AGT C	50	6358-6337
2R	F	GCC CGT GGC CAC CTT CTA TG	52,8	6114-6133
2S	R	CCC CCC GTT CAC TCC ACA AC	52,8	6976-6957
1K	F	CCG AGG CGG GAA AAC AGA GA	50,8	6821-6840
1L	R	AGG ATG CCA GTG CGC CTG TA	50,8	7763-7744
TBE-7547 (K2)***	F	CTG ACA CGT TGT GGA CGA TG	50	7547-7566
TBE-c8732 (K7)***	R	AAC ACT CTC TGC TGT CCG AAA G	50	8732-8711
TBE-8641 (K3)***	F	GTG AAA CTT CTC AGC TGG CC	42	8641-8660
TBE-c9685 (K8)***	R	TTG CCA AAT CTG TCA TCC	42	9695-9678
TBE-9244 (K4)***	F	AAT GGA GGA CTC TTC TAT GCA G	49	9244-9265
TBE-c10140 (K6)***	R	CAT GAA AGG GTT GTC CAG AAT C	49	10140-10119
TBE-2E hypr	F	TTG TTC ACA CCA CTT CCA CG	47	9792-9811
TBE-2T hypr	F	CGT TGG GAG GAA AGA AAG AG	47	10218-10237
1F	R	AGC GGG TGT TTT TCC GAG TC	47	10835-10816

* F – „forward“ primer, R – „reverse“ primer

** číslování odpovídá genomu kmene Hypr (U39292)

*** upraveno podle Kupča *et al.*, 2010

3.8 Separace PCR produktů na agarózovém gelu

Produkty PCR reakce byly separovány a vizualizovány pomocí gelové elektroforézy na 1% agarózovém gelu (1 % agarózy v 1x TAE pufru; 50x TAE pufr: 242 g TRIS, 57,1 ml kyseliny octové, 100 ml 0,5 M EDTA, 1000 ml H₂O, pH = 8,0). Před nanesením bylo ke vzorkům přidáno 5 µl nanášecího pufru (6x Orange DNA Loading Dye, Fermentas) s 5x koncentrovaným barvivem *SYBR Green* (*SYBR Green I*, Amresco). Separace probíhala po dobu 30-45 minut při napětí 100-130 V. Vizualizace byla prováděna na transiluminátoru (*Visi-blue™ Transilluminator*). Pro odečtení velikosti produktu byl použit 1 kb velikostní standard (*O'GeneRuler™ 1 kb DNA Ladder*, Fermentas) obsahující 5x koncentrované barvivo *SYBR Green*.

Proužky odpovídající velikosti byly z gelu vyříznuty skalpelem a DNA z nich byla eluována pomocí kitu *illustra™ GFX™ PCR DNA and Gel band purification Kit* (GE Healthcare) podle návodu výrobce. V posledním kroku byla purifikovaná DNA eluována z kolonky 30 µl elučního pufru do 1,5 ml mikrozkušavky a uskladněna při -20 °C. Výsledek purifikace byl elektroforeticky ověřen.

3.9 Sekvenční analýza

Přímou sekvenční analýzou obou vláken PCR produktů eluovaných z agarózového gelu byly zjištěny jejich nukleotidové sekvence. Sekvenční analýza byla prováděna na přístroji *ABI PRISM 3130xl* (Applied Biosystems) v Laboratoři genomiky, která je společnou laboratoří Přírodovědecké fakulty Jihočeské univerzity a Biologického centra Akademie věd České republiky.

Pro spojení komplementárních sekvencí, jejich úpravu a vytvoření konsenzuálních sekvencí byl použit program *MEGA* verze 4 (Tamura *et al.*, 2007). Aligmenty byly vytvořeny metodou *ClustalW* také v programu *MEGA* verze 4. Kontrolní identifikace sekvencí byla provedena pomocí databáze *NCBI BlastSearch* (<http://www.ncbi.nlm.nih.gov>).

3.10 3D modely proteinu E

Na základě získaných sekvencí byl pomocí software *GENO3D* (<http://geno3d-pbil.ibcp.fr>) a *SwissPdbViewer* (Peitsch 1995; Arnold *et al.*, 2006; Kiefer *et al.*, 2009) vytvořen 3D model proteinu E u varianty Hypr 40. P. PS. Model byl následně porovnán se stejně vytvořeným modelem pro původní variantu Hypr 0. P. Pro homologní modelování byla jako templát použita struktura proteinu E (PDB: 1SVB) modelovaná na základě krystalografické analýzy (Rey *et al.*, 1995). Různé varianty modelů byly porovnány

pomocí Swiss-Model, Structure Assessment (<http://swissmodel.expasy.org>) a nejlepší varianta byla použita pro vizualizaci aminokyselinových záměn. Aminokyselinová záměna u varianty Hypr DL. P. IRE nebyla modelována, protože se nachází v koncové části sekvence, která nebyla zahrnuta do referenčního modelu. Aminokyselinové záměny byly v modelech vizualizovány pomocí programů SwissPdbViewer a YASARA (<http://www.yasara.org>).

3.11 Myši – testování virulence

Rozdíly v biologických vlastnostech jednotlivých variant viru získaných různými způsoby pasážování byly testovány na myších. Pro stanovení virulence bylo použito celkem 50 myší kmene CD1 (AnLab s.r.o., Praha). Testováno bylo 5 různých variant viru (9 myší na každou variantu a 5 myší kontrolních). Všechny myši byly samice, vážící 15-20g, jejich stáří v době zahájení pokusu bylo 25-35 dní. Myši byly subkutánně (pod kůží za krkem) infikovány dávkou 100 PFU viru v 200 µl média (L15 s 3 % PTS, 1 % ATB a 1 % glutaminu). Po infekci byly myši denně kontrolovány po dobu následujících 30 dní a byly zaznamenávány počty uhynulých jedinců.

3.12 Myši – průběh infekce

Rozdíly v průběhu infekce mezi jednotlivými variantami viru s různou historií pasáží byly zjišťovány pomocí stanovování množství viru v orgánech myší v různých intervalech po infekci. Průběh infekce byl zkoumán pouze u 3 variant viru (0. P., 40. P. PS, 40. P. IRE). Celkem bylo na tento pokus použito 46 myší (15 myší pro každou variantu a 1 myš kontrolní). Myši kmene CD1 (AnLab s.r.o., Praha) byly samice, vážící 15-20 g, jejich stáří bylo 25-35 dní. Myši byly infikovány subkutánně (pod kůží za krkem) dávkou 100 PFU viru v 200 µl média (L15 s 3 % PTS, 1 % ATB a 1 % glutaminu).

Odběr vzorků byl prováděn po následujících osm dní. Každý den byly náhodně vybrány 2 myši z každé ze tří skupin (od každé varianty viru; celkem tedy 6 myší denně s výjimkou 7. dne, kdy byla usmrcena jen 1 myš z každé skupiny). Myši byly anestetizovány a humánně usmrceny cervikální dislokací. Byla jim odebrána krev, slezina a mozek. Krev byla ponechána 30 min. při pokojové teplotě za účelem sedimentace koagula a následně centrifugována 5 min. při 1030 g. Supernatant (krevní sérum) byl poté odebrán a zmrazen při -70 °C pro pozdější analýzu. Orgány byly zváženy a homogenizovány ve vodě pomocí

ocelové kuličky v homogenizátoru TissueLyser II (Retsch) při maximální rychlosti (frekvence 30/s) po dobu 2 min. Ze slezin byla přidáním vody vytvořena 20% suspenze, z mozků 33,3% suspenze. Po homogenizaci byly vzorky 10 min. centrifugovány při 16000 g a 4 °C, supernatant byl odebrán a zmražen při -70 °C pro pozdější analýzy.

3.13 Real-time PCR

Zjišťování přítomnosti viru a stanovení jeho množství v krvi a orgánech bylo prováděno pomocí metody real-time PCR. Tato metoda je založena na principu konvenční polymerázové řetězové reakce. Real-time PCR však umožňuje kromě amplifikace také kvantifikaci templátu. Detekce amplifikovaného produktu je prováděna na základě měření fluorescence. Fluorescenční substrát (v našem případě specifická sonda), se váže na syntetizovanou DNA a úroveň detekované fluorescence je pak úměrná množství nasyntetizované nukleové kyseliny. Kvantifikovat množství DNA je možné díky tomu, že existuje přímá úměra mezi počátečním množstvím templátu a cyklem, ve kterém začne fluorescence exponenciálně růst. Hodnota CT („threshold cycle“) je bod (počet cyklů), ve kterém fluorescence dosáhne námi zvolené prahové hodnoty.

Kvantitativní analýza byla prováděna pomocí absolutní kvantifikace, kdy se neznámá koncentrace cílové DNA vypočítá s použitím kalibrační křivky, která je tvořena ze vzorků se známou koncentrací DNA. Za účelem vytvoření kalibrační křivky byla připravena ředící řada pěti standardů cDNA s množstvím viru 10^4 , 10^3 , 10^2 , 10^1 a 10^0 PFU/4 μ l (množství cDNA potřebné pro 1 reakci). Všechny vzorky ředící řady byly analyzovány v triplikacích. Z kalibrační křivky bylo pak odečítáno množství viru v jednotlivých vzorcích, odpovídající počtu PFU. Všechny vzorky byly, stejně jako u ředící řady, analyzovány v triplikacích.

Amplifikační směs pro 1 reakci obsahovala 12,5 μ l Probe qPCR Master Mixu 2x (2x koncentrovaný: Probe qPCR pufr – $(\text{NH}_4)_2\text{SO}_4$, MgCl_2 , KCl ; dNTPs, Hot Start Taq DNA polymeráza), 4 μ l templátové cDNA, 1,5 μ l od každého z příslušných „forward“ a „reverse“ primerů (0,01 mM, Tab. 2) a 0,5 μ l hydrolyzační sondy typu TaqMan[®] (BHQ1-FAM, 200 nM, sekvence sondy: 5'- ACG CCA CTA GCG ACC CTG CAC AAC A-3'). Reakční směs byla doplněna do celkového objemu 25 μ l deionizovanou vodou (Fermentas).

Tab. 2 Primery pro real-time PCR.

Název primeru	Polarita*	Sekvence primeru 5'→3'	Teplota nasedání [°C]	Poloha v genomu**
E2(F) (TBE)	F	ACA CGG GAG ACT ATG TTG CCG CA	60	1409-1431
E(R) (TBE)	R	CCG TTG GAA GGT GTT CCA CT	60	1606-1587

* F – „forward“ primer, R – „reverse“ primer

** číslování odpovídá genomu kmene Hypr (U39292)

PCR reakce byla provedena v termálním cycleru *Rotor Gene 3000* (Corbett Research). Amplifikační program sestával z následujících kroků:

1. počáteční denaturace (holding) 95 °C, 10 min.
2. 45 cyklů: a) denaturace 95 °C, 15 s
b) nasedání primerů + elongace 60 °C, 60 s

Odečet fluorescence byl prováděn na konci každého cyklu. Výsledky byly zpracovány pomocí programu Rotor Gene 5.0 (Corbett Research).

3.14 Statistická analýza

Statistická významnost výsledků titru viru v orgánech a stanovení střední doby přežití byla vyhodnocena programem Statistica (StatSoft CR, Praha, verze 9.). Pro vyhodnocení byla použita metoda analýzy variance (Fischer LSD Post Hoc test). Data, která nevykazovala normální rozdělení, byla zlogaritmována. Výsledky, jejichž hladina významnosti p byla menší než 5%, byly považovány za signifikantní.

3.15 Metoda konformačního polymorfismu jednořetězců – SSCP

(„Single-Strand Conformation Polymorphism“)

Metoda konformačního polymorfismu jednořetězců (SSCP) je elektroforetická metoda, založená na principu různé rychlosti migrace jednovláknových molekul DNA, které se liší svou sekundární strukturou (konformací), v polyakrylamidovém gelu. Unikátní konformace jednovláknového fragmentu DNA je dána intramolekulárními interakcemi, které odpovídají nukleotidové sekvenci. Pomocí SSCP je možné ve vzorku rozdělit sekvenční varianty, lišící se i v pouhém jednom nukleotidu, pokud záměna způsobí změnu konformace. Proto byla tato metoda vybrána pro studium „quasispecies“ VKE.

Efektivita, s jakou je SSCP schopna rozdělit různé sekvenční varianty, je ovlivňována jak samotnou sekvencí nukleotidů, tak délkou fragmentů a podmínkami separace, jako jsou

pH, teplota, napětí, koncentrace polyakrylamidového gelu, koncentrace a složení nanášecího i elektroforetického pufru apod.

Analýza SSCP byla prováděna za různých podmínek a kombinací faktorů. Kromě různě koncentrovaných polyakrylamidových gelů (jejich složení je uvedeno v Tab. 3) byly vyzkoušeny dva nanášecí pufrы – formamidový (20 mM EDTA, 95 % formamid, 0,05 % xylen cyanol, 0,05 % bromfenolová modř) a s hydroxidem sodným (20 mM EDTA, 10 mM NaOH, 15 % ficoll, 0,05 % xylen cyanol, 0,05 % bromfenolová modř) v různých poměrech nanášecího pufru a DNA (3:1, 1:1). Separace probíhala při různých teplotách (5 °C, 9 °C, laboratorní teplota) a různých časových intervalech (2, 3, 4, 6, 12, 14 hod.). Kromě pufru 1x TBE byl vyzkoušen i pufr 1x TAE (jeho složení je uvedeno v kapitole 3.8). V několika pokusech bylo do gelu přidáno i denaturační činidlo močovina (5% močovina v 1x TBE nebo 1x TAE). Separace probíhala při konstantním napětí 240 V.

Tab. 3 Složení polyakrylamidových gelů.

<u>6% PA gel:</u>	1 ml 30% Akrylamid mix (akrylamid: bis 37,5 : 1) (Merck)
	1,25 ml 1xTBE (10xTBE: 54 g Tris, 27,5 g kyselina boritá, 20 ml 0,5 M EDTA, 1000 ml H ₂ O, pH=8)
	2,7 ml MQ H ₂ O
	0,05 ml 10% APS (persíran amonný)
	0,004 ml TEMED (SERVA)
<u>8% PA gel:</u>	1,3 ml 30% Akrylamid mix (akrylamid: bis 37,5 : 1)
	1,25 ml 1xTBE
	2,4 ml MQ H ₂ O
	0,05 ml 10% APS
	0,003 ml TEMED
<u>10% PA gel:</u>	1,7 ml 30% Akrylamid mix (akrylamid: bis 37,5 : 1)
	1,25 ml 1xTBE
	2,0 ml MQ H ₂ O
	0,05 ml 10% APS
	0,002 ml TEMED
<u>12% PA gel:</u>	2,0 ml 30% Akrylamid mix (akrylamid: bis 37,5 : 1)
	1,25 ml 1xTBE
	1,7 ml MQ H ₂ O
	0,05 ml 10% APS
	0,002 ml TEMED

PCR produkty byly před nanesením na gel smíchány s nanášecím pufrům. Jako vhodnější se ukázal nanášecí pufr s formamidem v poměru objemů 1:1 (10 µl PCR produktu smícháno s 10 µl pufru). DNA byla poté zdenaturována v termálním cycleru (94 °C, 10-15 min.). Po vytažení ze cycleru byly vzorky okamžitě uloženy na led, aby nedošlo k renaturaci DNA a následně naneseny na gel.

Po ukončení elektroforézy byl gel barven ve 100 ml 0,5x pufru TBE (nebo 0,5x TAE), do kterého byl přidán 1 ml 100x koncentrovaného barviva SYBR Green (SYBR Green I, Amresco). Barvení probíhalo po dobu 30-45 min. Vizualizace byla prováděna na transiluminátoru (Visi-blue™ Transilluminator).

Proužky byly z gelu vyříznuity, přeneseny do mikroskopu a zality 100 µl elučního pufru (Elution Buffer, Fermentas). Eluce DNA z gelu probíhala při 37 °C po dobu 12-15 hod. Poté byly vzorky vortexovány, centrifugovány a eluční pufr s uvolněnou DNA byl přenesen do nové mikroskopu. Pro další analýzy byly vzorky uchovávány při -20 °C.

3.16 Klonování

Další metodou, použitou pro studium „quasispecies“, je klonování. Klonování bylo prováděno pomocí kitu *CloneJET PCR Cloning Kit* (Fermentas). Tento kit využívá pozitivní selekce založené na sekvenci vektoru. Klonovací vektor (pJET1.2/blunt) obsahuje sebevražedný gen (kódující pro bakterie letální restrikční enzym), v jehož sekvenci se nachází klonovací místo. Pokud je ligace úspěšná a do klonovacího místa je vnesen cílový fragment DNA, přeruší se tím sekvence sebevražedného genu a bakterie jsou tudíž životaschopné. Pokud však ligace neproběhne úspěšně, vektor bez inzertu bakterii zabije. Na plotnách proto nakonec vyrostou pouze kolonie bakterií obsahující rekombinantní vektor s inzertem.

3.16.1 Příprava ligační směsi

DNA určená ke klonování byla nejprve amplifikována pomocí PCR a fragmenty odpovídající velikosti byly izolovány z gelu. Koncentrace eluované DNA byla zjištěna spektrofotometricky (*Nano-drop ND-1000*, Thermo Scientific). Ligační směs byla připravena podle pokynů výrobce. Množství DNA v ligační směsi bylo přibližně 30 ng. Ligační směs byla inkubována při laboratorní teplotě po dobu 20 min.

3.16.2 Transformace

K 25 μ l kompetentních bakterií *Escherichia coli* (TOP10, Invitrogen) byly přidány 4 μ l ligační směsi. Výsledná směs byla promíchána kroužením a ponechána stát 20 min. na ledu. Poté byl proveden tepelný šok ponořením na 30 s do 42 °C horké vodní lázně a vrácením vzorků na led. Ke směsi bylo přidáno 200 μ l SOC média (LB médium obohacené o glukózu a hořečnaté soli). Vše bylo inkubováno 60 min. při 37 °C v třepačce. Po inkubaci byla bakteriální směs rovnoměrně rozetřena na předem připravené misky, obsahující tuhé LB médium (10 g trypton, 5 g kvasnicový extrakt, 10 g NaCl, 1 l voda) s ampicilinem (konečná koncentrace 100 μ g/ml). Bakterie rostly na miskách obrácených dnem vzhůru následujících 12 – 15 hod. při teplotě 37 °C.

3.16.3 PCR z kolonií

Pro ověření zaklonování fragmentu o správné délce bylo provedeno PCR z bakteriálních kolonií a délka produktů byla zjištěna pomocí gelové elektroforézy. Jako templát sloužily přímo části bakteriálních kolonií. Před přenesením do PCR směsi byly kolonie nejprve přeočkovány na nové plotny a znovu inkubovány při 37 °C po dobu nejméně 12 hod. Pro PCR z kolonií byly použity primery příslušné k zaklonovanému fragmentu DNA (Tab. 1).

3.16.4 Pomnožení bakterií v tekutém médiu

Z přeočkovaných kolonií bylo odebráno maximální množství bakteriální masy a přeneseno do 1,8 ml tekutého LB média s antibiotikem (ampicilin nebo carbenicillin, koncentrace 50 μ g/ml). Vzorky byly umístěny na třepačku a inkubovány po dobu nejméně 12 hod., při teplotě 37 °C. Poté byly centrifugovány při 8100 g po dobu 1 min. Supernatant byl odebrán a bakteriální pelet uchováván při -20 °C pro další analýzy.

3.16.5 Izolace plazmidu

Z bakteriálního peletu byly izolovány plazmidy pomocí kitu *GeneJET™ Plasmid Miniprep Kit* (Fermentas) podle návodu výrobce. V posledním kroku byla purifikovaná plazmidová DNA eluována z kolonky 50 μ l vymývacího pufru do 1,5 ml mikrozskumavky a uskladněna při -20 °C.

Před použitím purifikované vektorové DNA pro sekvenační reakci byla spektrofotometricky (*Nano-drop ND-1000*, Thermo Scientific) změřena její koncentrace. Do sekvenační reakce bylo použito 100 ng DNA. Jako sekvenační primery byly použity primery ohraničující klonovací místo na vektoru (Tab. 4).

Tab. 4 Primery vektoru pJET1.2.

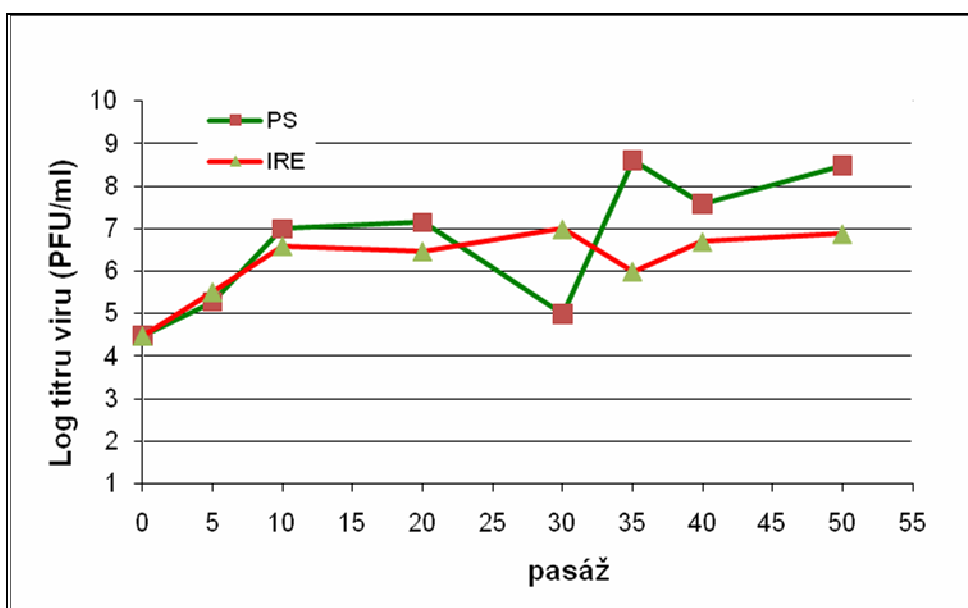
Název primeru	Sekvence primeru 5'→3'	Teplota nasedání [°C]	Poloha v genomu
pJET1.2 forward sequencing primer	CGA CTC ACT ATA GGG AGA GCG GC	55	310-332
pJET1.2 reverse sequencing primer	AAG AAC ATC GAT TTT CCA TGG CAG	55	428-405

4. Výsledky

4.1 Pasážování viru a hledání determinant virulence

4.1.1 Změna titru viru v průběhu pasážování a rozdíly ve velikosti plaků

Virus VKE kmene Hypr byl pasážován souběžně na savčích buňkách PS a na klíštěcích buňkách IRE/CTVM19. Plakovou titrací byl stanoven počáteční titr původní varianty viru (dále jen 0. P.) 3×10^4 PFU/ml. Bylo provedeno celkem 50 pasáží v obou prostředích. Pro kontrolu titru viru a velikosti plaků byla provedena po 5., 10., 20., 30., 35., 40. a 50. pasáži plaková titrace (Obr. 4, 5 a 6).

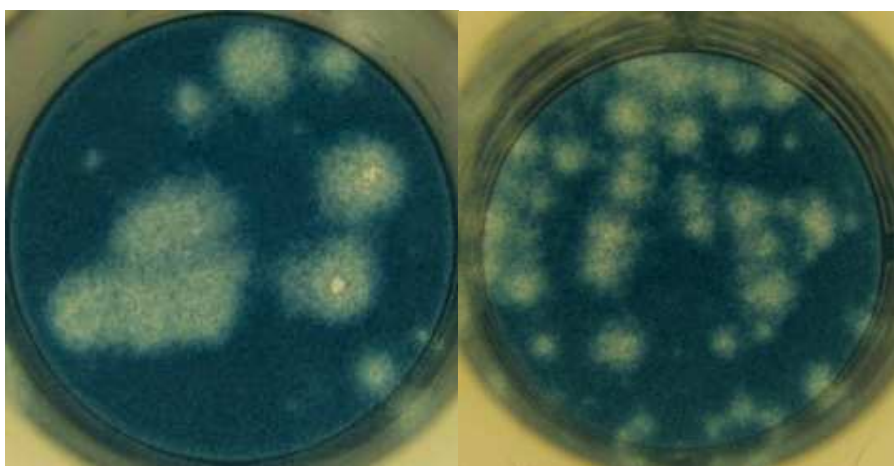


Obr. 4 Graf závislosti titru viru na počtu pasáží znázorňuje změny titru viru během pasážování.

Titry varianty viru pasážované přes klíštěcí buňky se od 10. pasáže pohybovaly víceméně stabilně v rozmezí $10^6 - 10^7$ PFU/ml. Titry varianty viru pasážované přes savčí buňky na počátku stoupany rychleji, ale později došlo k mnohem větším výkyvům hodnot než u klíštěcí varianty. Titr se u savčí varianty pohyboval od 10. pasáže v rozmezí $10^5 - 10^8$ PFU/ml. Příčinou výrazného poklesu titru viru mezi 20. a 30. pasáží byla stále se zvyšující rychlost množení viru při stále stejném množství buněk PS. Virus během čtyř dnů trvající inkubace zabíjel buňky rychleji, než se samy stačily množit, čímž si likvidoval prostředí nutné pro svou replikaci. Proto byla od 30. pasáže zkrácena doba inkubace viru s PS buňkami na dva dny. To způsobilo výrazný nárůst titru viru v následujících pěti pasážích o více než tři řády. Následně titr viru opět poklesl, pravděpodobně ze stejného důvodu, jako

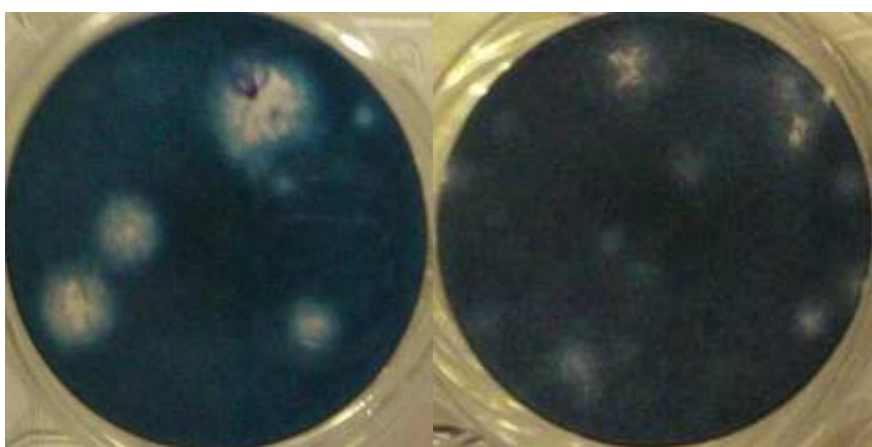
poprvé. I při zkrácené inkubaci mezi pasážemi bylo takto vysoké množství viru v buněčné kultuře limitní.

První změny ve velikosti plaků byly zachyceny ve 30. pasáži. U varianty viru pasážovaného přes savčí buňky převažovaly plaky velké, charakterově shodné s plaky původního viru. Kromě těchto plaků se objevilo i několik plaků menších. Virus pasážovaný přes klíštěcí buňky vytvářel v kultuře savčích buněk plaky přibližně poloviční velikosti, tedy výrazně menší, než původní virus (Obr. 5).



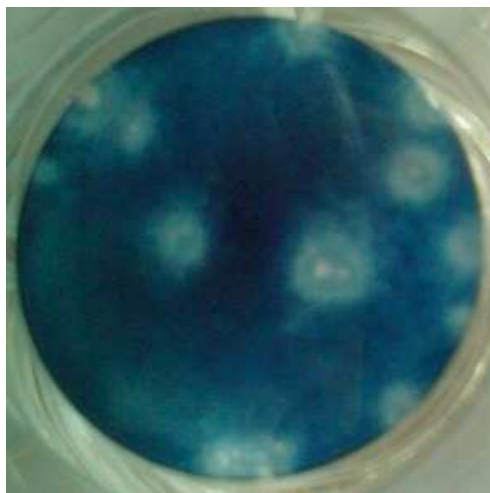
Obr. 5 Charakter plaků kmene Hypr po 30. pasáži viru přes savčí (vlevo) a klíštěcí (vpravo) buňky.

Při pokračování pasáží až do 40. pasáže (a kontrolně do 50. pasáže) nedošlo k žádné změně ve velikosti plaků oproti 30. pasáži. Vzorky ze 40. pasáže obou variant (dále jen 40. P. PS a 40. P. IRE) byly vybrány pro další analýzy (Obr. 6).



Obr. 6 Charakter plaků kmene Hypr po 40. pasáži viru přes savčí (vlevo) a klíštěcí (vpravo) buňky.

Další variantou viru, jejíž biologické vlastnosti byly dále studovány, byl virus pasážený dlouhodobě na klíštěcích buňkách (dále jen DL. P. IRE). Po měsíční inkubaci stoupl titr viru jen nepatrně oproti původní variantě, na hodnotu 9×10^3 PFU/ml. V kultuře buněk PS se objevila směs velkých a menších plaků (Obr. 7). Velké plaky se velikostně shodovaly s plaky 0. P., zatímco menší plaky odpovídaly svou velikostí spíše 30. a 40. P. IRE.

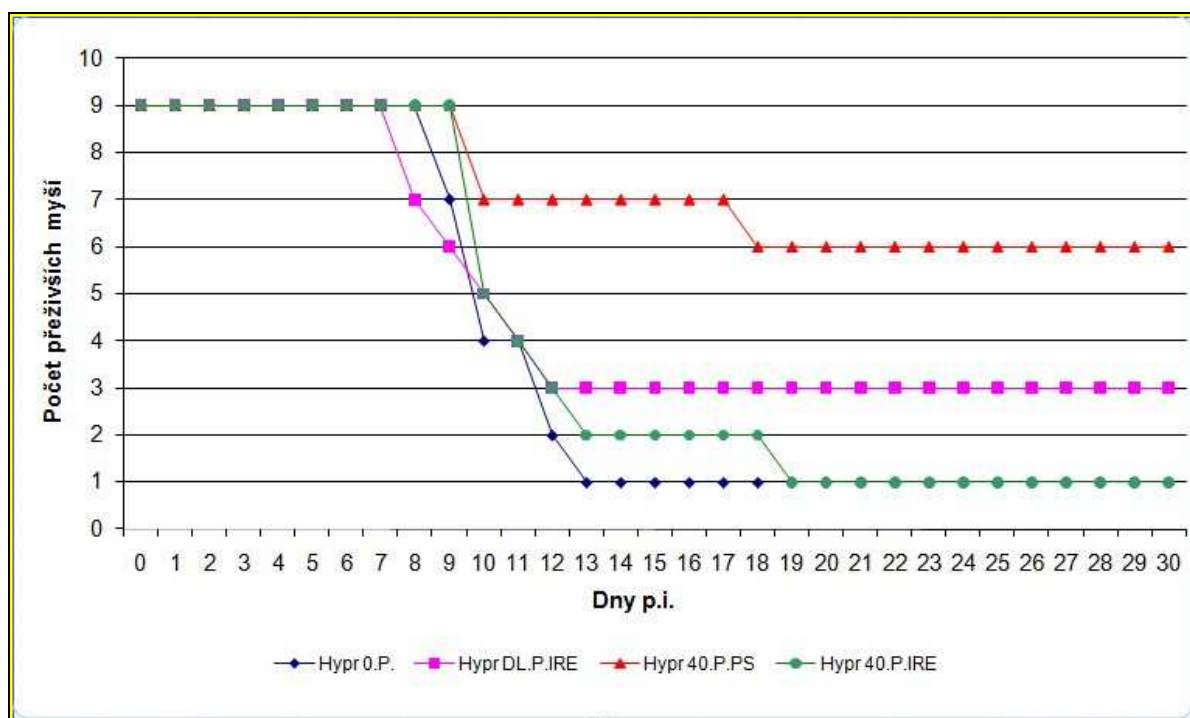


Obr. 7 Charakter plaků kmene Hypr po dlouhodobé pasáži na klíštěcích buňkách.

Pasážováním viru v různém prostředí a za různých podmínek byly připraveny tři varianty VKE kmene Hypr (DL. P. IRE, 40. P. PS a 40. P. IRE). Jejich biologické vlastnosti a míra patogenity ve srovnání s původní variantou (0. P.) byly dále studovány na myším modelu.

4.1.2 Virulence variant VKE v myším modelu

Testování, zda dlouhodobé pasážování vedlo k vyselektování variant s různými biologickými vlastnostmi, jak naznačovaly výsledky plakové titrace, bylo prováděno na myším modelu. Po subkutánní inokulaci dospělých myši kmene CD1 dávkou 100 PFU byly myši pozorovány po dobu 30 dní. Na následujícím grafu (Obr. 8) je znázorněn průběh pokusu pro infekci jednotlivými variantami viru.



Obr. 8 Graf závislosti počtu přeživších jedinců na čase po inokulaci.

Střední doby přežití, jejich směrodatné odchylky a počty přeživších myší jsou pro jednotlivé varianty uvedeny v Tab. 5. Statisticky průkazně kratší střední dobu přežití vykazovaly myši infikované variantami 0. P. a 40. P. IRE oproti 40. P. PS. (ANOVA, $p < 0,05$; Fisher LSD, $p < 0,05$).

Tab. 5 Střední doba přežití a počet přeživších myší CD1 po subkutánní inokulaci různými variantami viru kmene Hypr (doba trvání pokusu 30 dní).

Varianta viru	Střední doba přežití (dny)	Směrodatná odchylka (dny)	Počet přeživších myší (%)
0. P.	11,7	7	1 (11 %)
DL. P. IRE	15,8	10,7	3 (33 %)
40. P. PS	23,8	9,7	7 (67 %)
40. P. IRE	12,8	7	1 (11 %)

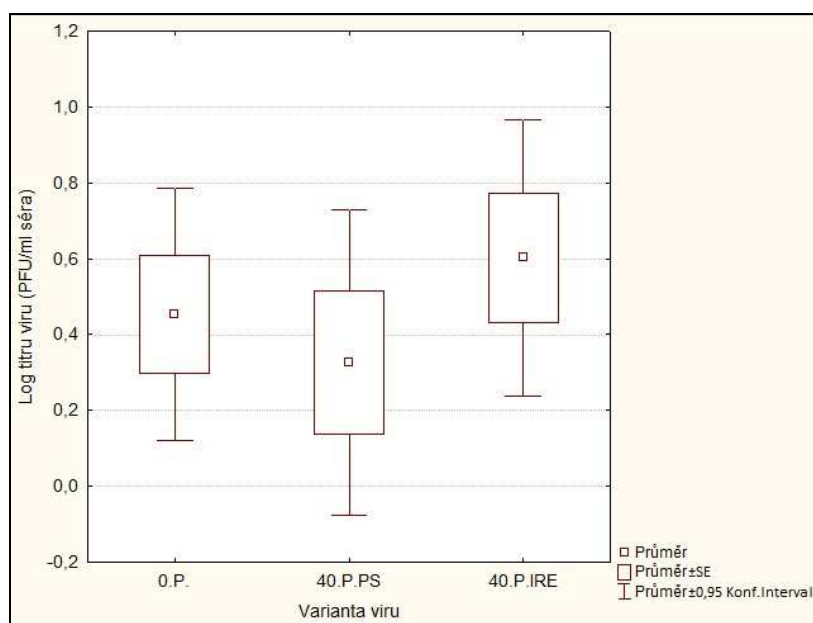
Z výsledků tedy vyplývá, že nejrychlejší průběh a nejvyšší míru virulence měla původní varianta viru, nejnižší patogenitu naopak vykazovala 40. P. PS.

4.1.3 Sledování dynamiky množení viru v orgánech

V předchozím pokusu bylo zjištěno, že varianty viru adaptované na odlišné prostředí, jsou různě virulentní pro laboratorní myši.

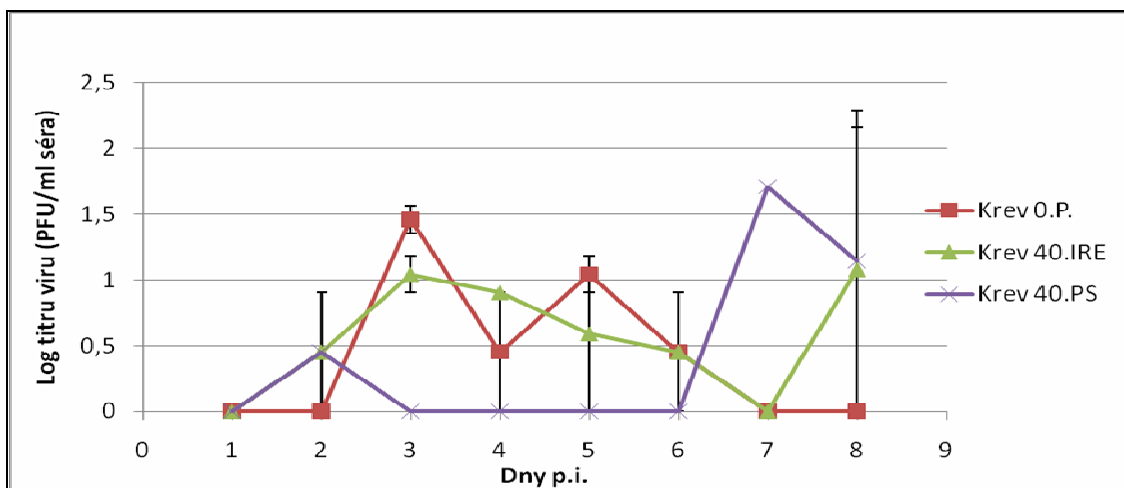
Pro bližší osvětlení pravděpodobných mechanismů byla sledována dynamika množení viru v krvi a vybraných orgánech. Do tohoto pokusu byly zahrnuty pouze původní varianta viru a obě 40. pasáže. Každou z těchto variant bylo infikováno celkem 15 myši. Po dobu následujících 8 dní byly každý den náhodně vybrány 2 myši z každé varianty (7. den pouze jedna myš) a humánně usmrceny. Byla jim odebrána krev, slezina a mozek. Pomocí real-time PCR bylo zjištěno množství viru v těchto orgánech.

Na prvním grafu (Obr. 9) je znázorněno porovnání množství viru v krvi (log PFU/ml) u jednotlivých variant viru. Žádný z rozdílů nebyl statisticky průkazný.



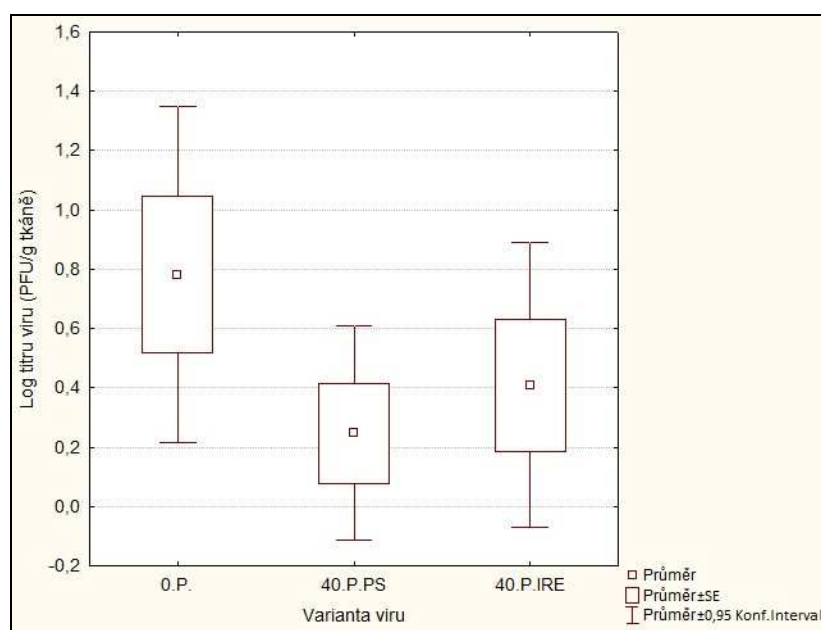
Obr. 9 Graf množství viru v krvi u jednotlivých variant.

Na dalším grafu (Obr. 10) je patrný rozdílný charakter průběhu virémie u myši infikovaných 40. P. PS oproti oběma zbývajícím variantám. Nástup virémie byl pro všechny varianty viru zachycen shodně již 2. - 3. den po inokulaci, ovšem u varianty 40. P. PS nebyl nalezen žádný další pozitivní vzorek až do 7. dne. Vzorek zachycený 2. den po infekci byl podle real-time analýzy těsně nad hranicí detekce. U 40. P. PS se tedy virus 3. - 6. den, vzhledem k jednomu pozitivnímu vzorku ze 2. dne, pravděpodobně v krvi vyskytoval, ovšem v množství pod hranicí detekce metody. U 0. P. a 40. P. IRE byly charakter dynamiky výskytu viru i dosahovaný titr obdobný.



Obr. 10 Graf závislosti množství viru v krvi na čase po inokulaci (každý bod grafu je průměrem dvou hodnot získaných ze dvou myší; chybové úsečky představují směrodatnou odchylku od průměru).

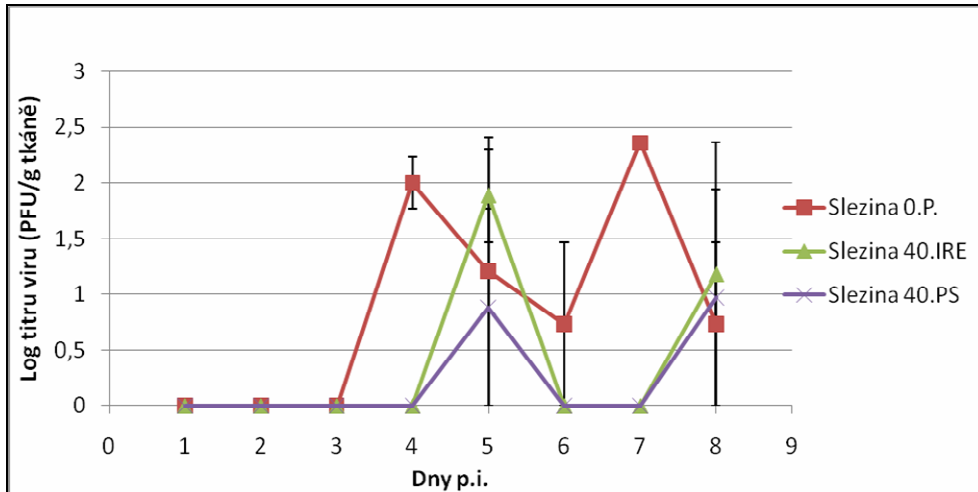
Na následujícím grafu (Obr. 11) je znázorněno množství viru ve slezině (log PFU/g) u jednotlivých virových variant. Statistickou analýzou nebyl žádný z rozdílů vyhodnocen jako signifikantní.



Obr. 11 Graf množství viru ve slezině u jednotlivých variant viru.

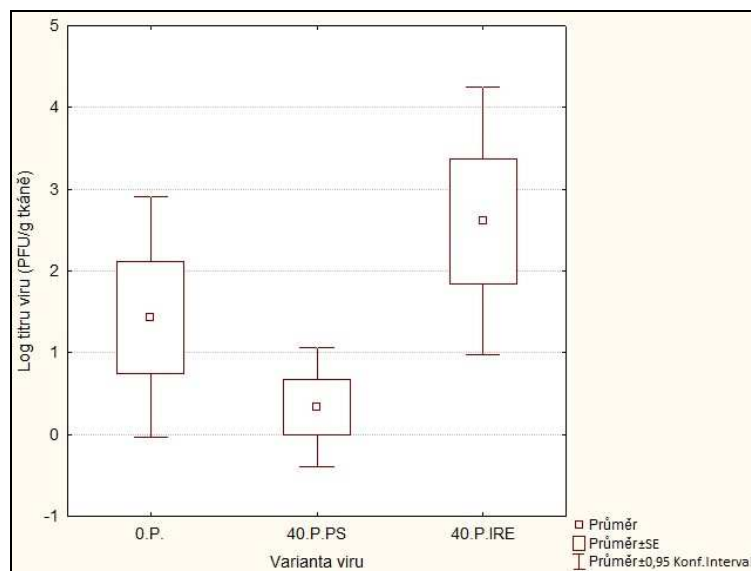
Na Obr. 12 je vynesena závislost množství viru ve slezině na čase po inokulaci. Titr varianty 40. P. PS dosahuje v korespondujících částech křivky přibližně poloviční hodnoty než zbývající dvě varianty, ale objevuje se ve slezině ve stejný den jako 40. P. IRE a jen o den

později než 0. P. Zatímco v krvi se nejrychleji rozšířila 40. P. IRE a až o den později byla přítomnost viru v krvi objevena také u 0. P., ve slezině se virus objevil dříve u 0. P. a o den později u 40. P. IRE. Z grafu je také patrné, že se u všech tří variant projevil dvoufázový nárůst titru viru ve slezině.



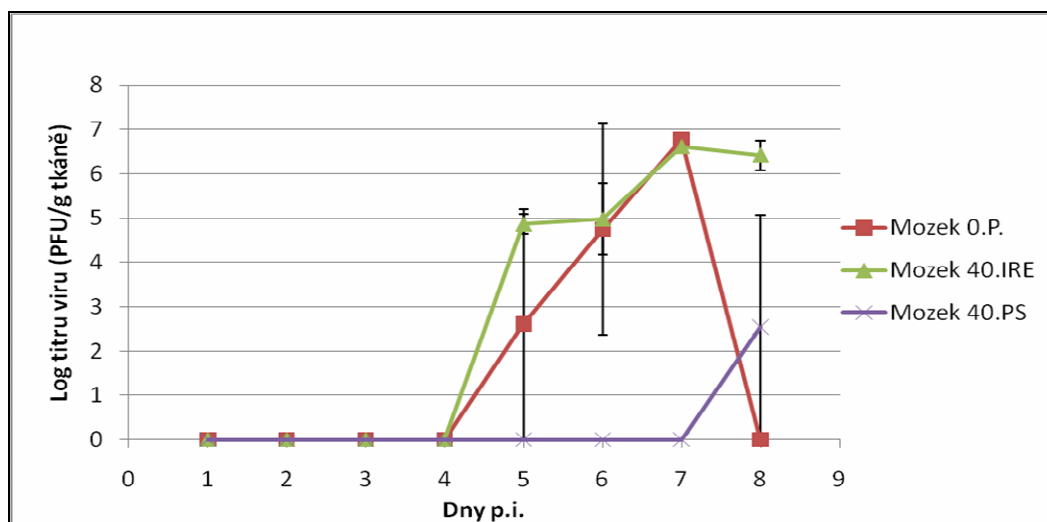
Obr. 12 Graf závislosti množství viru ve slezině na čase po inokulaci (každý bod grafu je průměrem dvou hodnot získaných ze dvou myší; chybové úsečky představují směrodatnou odchylku od průměru).

Na dalším grafu (Obr. 13) je znázorněno množství viru v mozku (log PFU/g) u jednotlivých variant viru. Statistickým vyhodnocením získaných výsledků byl zjištěn signifikantní rozdíl mezi variantami viru 40. P. PS a 40. P. IRE.



Obr. 13 Graf množství viru v mozku u jednotlivých variant.

Rozdíl mezi variantami je zřejmý i z dalšího grafu (Obr. 14), znázorňujícího závislost množství viru v mozku na čase po inokulaci. U myší infikovaných variantami 0. P. a 40. P. IRE byl virus přítomen v mozku již 5. den po inokulaci a během dalších dvou dní jeho množství v tomto orgánu strmě rostlo. U myší infikovaných variantou 40. P. PS se virus v mozku objevil poprvé až 8. den po inokulaci a to pouze u jednoho jedince.



Obr. 14 Graf závislosti množství viru v mozku na čase po inokulaci (každý bod grafu je průměrem dvou hodnot získaných ze dvou myší; chybové úsečky představují směrodatnou odchylku od průměru).

Výsledky průběhu infekce u myší infikovaných jednotlivými variantami viru odpovídají výsledkům stanovení virulence. Infekce původní variantou a 40. pasáží viru adaptovaného na klíčtější buňky měla až na malé odchylky velmi podobný průběh. U myší infikovaných 40. P. PS byl oproti dvěma zbývajícím virovým variantám v krvi i v orgánech dosažen výrazně nižší titr. Varianta viru adaptovaná mnohonásobným pasážováním na savčí buňky vykazovala výrazně nižší neuroinvasivitu než zbývajících dvě varianty.

4.1.4 Sekvenční analýza kompletního genomu všech čtyř variant VKE

Za účelem identifikace konkrétních úseků genomu VKE, zodpovědných za rozdílné biologické vlastnosti jednotlivých variant viru, byla provedena kompletní sekvenční genomová analýza původní varianty 0. P. i všech tří nově vyselektovaných variant – DL. P. IRE, 40. P. PS a 40. P. IRE (Příloha 9.1).

Celková délka genomu všech čtyř variant byla shodná a to 10 835 nukleotidů. V koncových nekódujících oblastech ani v sekvencích jednotlivých proteinů nebyly tedy pozorovány žádné delece ani inserce. Za 5' netranslatovanou oblastí, dlouhou 132 bazí, následoval START kodon ATG, od kterého začíná proteosyntéza rozsáhlého virového polyproteinu. Sekvence polyproteinu, o celkové délce 10 239 bazí, byla zakončena STOP kodonem TAA. Za ním následovala 3' netranslatovaná oblast, dlouhá 458 nukleotidů. Základní charakteristika všech proteinů včetně nekódujících oblastí je uvedena v Tab. 6.

Tab. 6 Základní charakteristika jednotlivých částí genomu VKE.

Oblast genomu	Délka v nt	Počet AK	Poloha v genomu
5' UTR	132	---	1 - 132
Protein C	333	111	136 – 468
Protein prM	279	93	469 – 747
Protein M	225	75	748 – 972
Protein E	1488	496	973 – 2460
Protein NS1	1056	352	2461 – 3516
Protein NS2A	690	230	3517 – 4206
Protein NS2B	393	131	4207 – 4599
Protein NS3	1863	621	4600 – 6462
Protein NS4A	447	149	6463 – 6909
Protein NS4B	756	252	6910 – 7665
Protein NS5	2709	903	7666 – 10374
3' UTR	458	---	10378 - 10835

*číslování podle genomové sekvence kmene Hypr (U39292)

Na základě porovnání celogenomových sekvencí nově vyselektovaných variant viru a varianty původní bylo identifikováno celkem 20 jednonukleotidových záměn. Dvě z nich se vyskytovaly v koncových nekódujících oblastech. Ze zbylých 18 záměn jich 12 způsobilo i změnu v aminokyselinové sekvenci. Nejvíce se od původní varianty odlišovala 40. P. PS, a to 11 nukleotidovými záměnami, z nichž 6 vedlo ke změně aminokyseliny. Je tedy pravděpodobné, že jedna nebo i více z těchto záměn jsou zodpovědné za sníženou virulenci 40. P. PS. Následovala 40. P. IRE, která se od 0. P. lišila v 9 nukleotidech a 4 aminokyselinách. Nejpodobnější sekvenci s původní variantou měla DL. P. IRE, lišící se ve 3

nukleotidech a 2 aminokyselinách. Konkrétní údaje týkající se jednotlivých záměn jsou uvedeny v Tab. 7, frekvence záměn pak v Tab. 8.

Tab. 7 Rozdíly v nukleotidových a aminokyselinových sekvencích mezi jednotlivými virovými variantami (před šipkou je původní báze (aminokyselina) a v závorce její poloha v rámci proteinu, za šipkou pak změněná báze (aminokyselina) a v závorce varianta, u které k této změně došlo).

Oblast genomu	Rozdíly v nukleotidové sekvenci	Rozdíly v aminokyselinové sekvenci
5' UTR	G (52) → A (40. P. IRE)	
Protein C	A (315) → G (40. P. IRE)	
Protein prM	C (101) → T (40. P. IRE)	Thr (34) → Ile (40. P. IRE)
Protein E	A (913) → G (40. P. PS)	Thr (305) → Ala (40. P. PS)
	C (1078) → T (40. P. PS)	Pro (360) → Ser (40. P. PS)
	A (1411) → T (DL. P. IRE)	Met (471) → Leu (DL. P. IRE)
Protein NS1	G (169) → A (40. P. IRE)	Val (57) → Ile (40. P. IRE)
	G (237) → A (40. P. PS)	
Protein NS2A	T (605) → C (40. P. PS)	Val (202) → Ala (40. P. PS)
Protein NS2B	G (97) → A (40. P. IRE)	Val (33) → Met (40. P. IRE)
Protein NS3	T (978) → C (40. P. PS, 40. P. IRE)	
	A (1314) → G (DL. P. IRE)	
Protein NS4B	C (240) → T (40. P. PS)	
	G (253) → T (40. P. IRE)	Ala (85) → Ser (40. P. IRE)
	T (262) → A (40. P. PS)	Phe (88) → Ile (40. P. PS)
Protein NS5	G (333) → A (40. P. PS, 40. P. IRE)	
	A (529) → C (DL. P. IRE)	Thr (177) → Pro (DL. P. IRE)
	C (2332) → T (40. P. PS)	Leu (778) → Phe (40. P. PS)
	A (2588) → G (40. P. PS)	Asn (863) → Ser (40. P. PS)
3' UTR	T (282) → C (40. P. PS, 40. P. IRE)	

Tab. 8 Frekvence nukleotidových záměn pro jednotlivé varianty viru a oblasti genomu.

Oblast genomu	Procento nukleotidových záměn oproti 0. P.			
	40. P. PS	40. P. IRE	DL. P. IRE	Celkem
5'-UTR		0,75 % (1/132)		0,25 % (1/396)
Protein C		0,3 % (1/333)		0,1 % (1/999)
Protein prM		0,35 % (1/279)		0,12 % (1/837)
Protein E	0,13 % (2/1488)		0,07 % (1/1488)	0,07 % (3/4464)
Protein NS1	0,09 % (1/1056)	0,09 % (1/1056)		0,06 % (3/3168)
Protein NS2A	0,14 % (1/690)			0,05 % (1/2070)
Protein NS2B		0,25 % (1/393)		0,08 % (1/1179)
Protein NS3	0,05 % (1/1863)	0,05 % (1/1863)	0,05 % (1/1863)	0,05 % (3/5589)
Protein NS4B	0,27 % (2/756)	0,13 % (1/756)		0,13 % (3/2268)
Protein NS5	0,11 % (3/2709)	0,04 % (1/2709)	0,04 % (1/2709)	0,06 % (5/8127)
3' - UTR	0,22 % (1/458)	0,22 % (1/458)		0,15 % (2/1374)
Celkové procento záměn v genomu	0,1 % (11/10 835)	0,08 % (9/10 835)	0,03 % (3/10 835)	

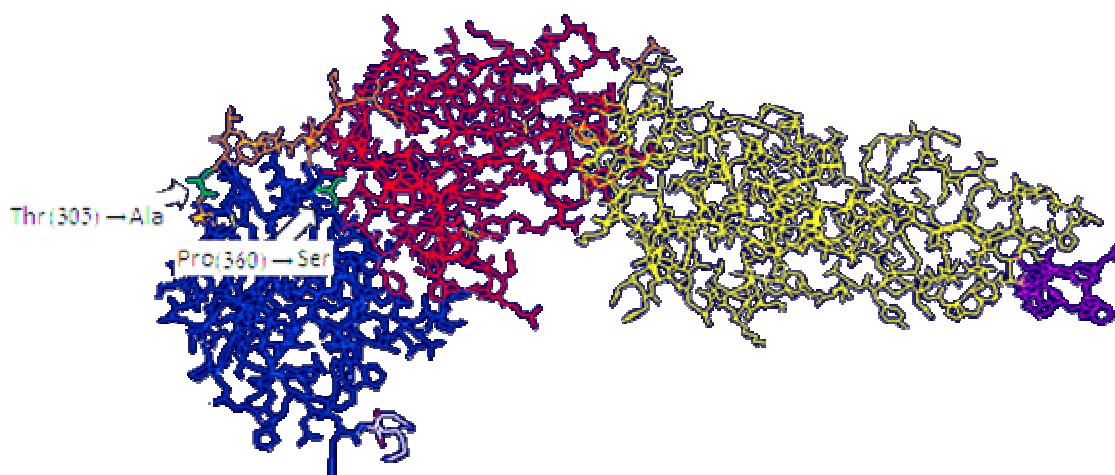
Nejvyšší frekvenci nukleotidových záměn vykazovaly podle očekávání oba koncové netranslatované úseky. Z kódujících sekvencí se jako nejvariabilnější jevíly úseky kódující strukturální protein prM a nestrukturní protein NS4B.

4.1.5 Potenciální determinanty virulence

Vzhledem k možnosti identifikace úseků genomu ovlivňujících virulenci viru je nejzajímavější srovnání varianty Hypr 40. P. PS se zbývajícími.

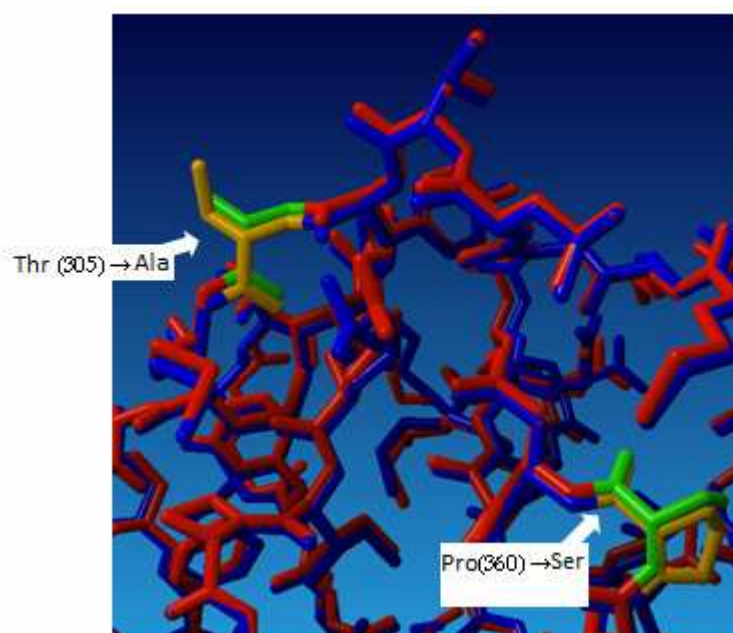
Protein E

V rámci proteinu E byly u varianty 40. P. PS identifikovány 2 nukleotidové záměny, které měly obě za následek změnu AK. Záměna Thr (305) → Ala je umístěna v oblasti spojující doménu I a doménu III („D I – D III linker“). Druhá záměna Pro (360) → Ser je umístěna v blízkosti oblasti, u které se předpokládá zapojení do vazby na buněčné receptory (Obr. 15).



Obr. 15. Model proteinu E varianty 40. P. PS (doména I – červená, doména II – žlutá, doména III – modrá, fúzní peptid – fialová, „DIII-DI linker“ – oranžová; záměny vyznačeny zeleně).

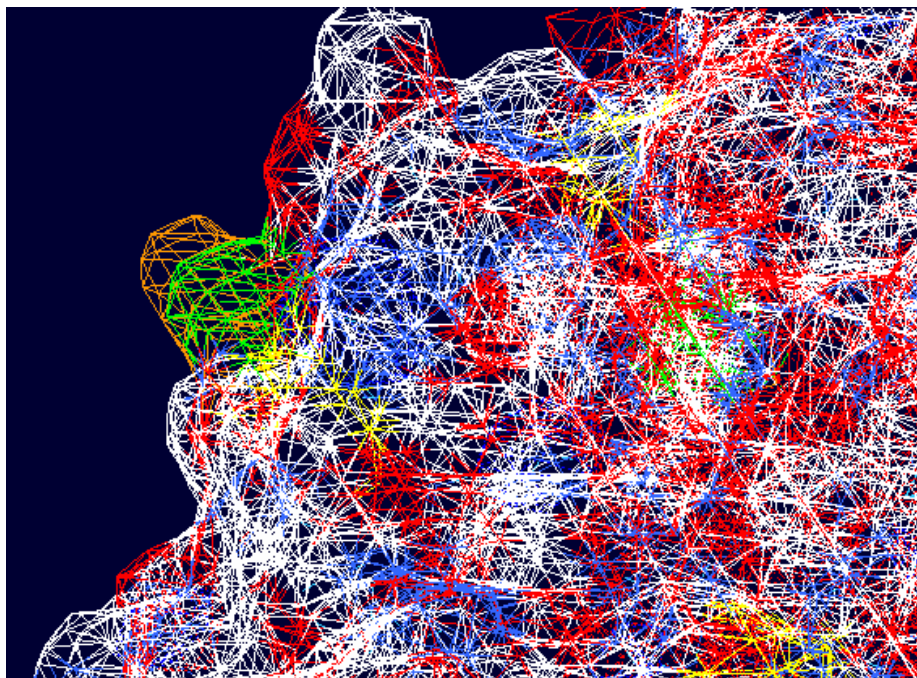
Detail změny prostorového uspořádání postranních řetězců AK v proteinu E u 40. P. PS je uveden na Obr. 16.



Obr. 16 Detail aminokyselinových záměn v proteinu E (0. P. – oranžová, 40. P. PS – zelená).

V případě záměny u 40. P. PS v pozici 305 Thr → Ala se jednalo o záměnu polárního threoninu za nepolární alanin. V druhém případě, Pro → Ser v pozici 360, byl zaměněn

nepolární prolin za polární serin. Změna povrchové struktury proteinu je vyznačena na Obr. 17. Z pohledu náboje byly obě změny elektroneutrální.



Obr. 17 Změna molekulárního povrchu proteinu E (0. P. – oranžová, 40. P. PS – zelená).

Protein NS1

V proteinu NS1 došlo u 40. P. PS k záměně G (237) → A. Tato nukleotidová záměna ale nezpůsobila záměnu AK.

Protein NS2A

V sekvenci kódující protein NS2A byla identifikována záměna v pozici T (605) → C. Tato mutace měla za následek záměnu nepolární AK (valinu) za jinou nepolární AK (alanin).

Protein NS3

V proteinu NS3 došlo k identické nukleotidové záměně u variant 40. P. PS a 40. P. IRE. Varianta 40. P. IRE se však virulencí ani dynamikou množení prakticky nelišila od původního kmene, tato záměna tedy pravděpodobně nemá vliv na tyto vlastnosti.

Protein NS4B

V proteinu NS4B byly identifikovány dvě unikátní záměny u varianty 40. P. PS a to v pozicích 240 C → T a 262 T → A. V prvním případě se jednalo pouze o záměnu v sekvenci

nukleotidů, v druhém případě měla záměna nukleotidu za důsledek záměnu nepolárního fenylalaninu za nepolární izoleucin.

Protein NS5

V proteinu NS5 došlo celkem ke 3 záměnám. Záměna v pozici 333 G → A byla společná pro 40. P. PS a 40. P. IRE. Tato záměna se týkala pouze nukleotidové sekvence. Obě další záměny, C → T v pozici 2332 a A → G v pozici 2588, byly unikátní pro 40. P. PS a měly za následek změnu AK. V prvním případě se jednalo o záměnu dvou nepolárních AK, leucinu za fenylalanin, ve druhém případě o záměnu dvou polárních AK, asparaginu za serin.

3' nekódující úsek

V 3' UTR byla nalezena jediná nukleotidová záměna v pozici 282 T → C. Tato záměna byla opět společná pro 40. P. PS a 40. P. IRE.

4.2 Variabilita v rámci jednotlivých izolátů (přítomnost „quasispecies“)

Jelikož existují hypotézy, že adaptace viru na různá prostředí probíhá mechanismem selekce z ko-existujících sekvenčních variant („quasispecies“), pokusili jsme se stanovit sekvenční variabilitu v rámci našich variant viru. Přítomnost „quasispecies“, byla testována pomocí dvou metod – metodou konformačního polymorfismu jednořetězců (SSCP) a klonováním.

4.2.1 Metoda konformačního polymorfismu jednořetězců (SSCP)

Jak je již zmíněno v kapitole 3.15, tato metoda je schopná za jistých podmínek a při správné kombinaci faktorů, ovlivňujících její průběh, oddělit od sebe v polyakrylamidovém gelu různé sekvenční varianty téhož vzorku, lišící se i v pouhém jednom nukleotidu. V této práci se bohužel nepodařilo tyto ideální podmínky najít. I přesto, že se vzorky na gelu rozdělovaly na 2 – 6 proužků různé délky, po osekvenování byla vždy zjištěna u všech těchto variant naprosto shodná sekvence.

4.2.2 Klonování

Při klonování je, stejně jako při SSCP, možné oddělit jednotlivé sekvenční varianty v rámci téhož vzorku. Pokud se totiž u vzorku vyskytuje více sekvenčních variant s určitou frekvencí, je pravděpodobné, že se různé varianty zaklonují zvlášť a po procesu klonování vyrostou na jedné plotně kolonie bakterií s inzerty, jejichž sekvence se liší jedním nebo více nukleotidy.

Části genomu, které byly podrobeny klonování, jsou část genu kódujícího protein E (796 nt), gen pro protein C (333 nt), 5' UTR (132 nt) a 3' UTR (458 nt). Nukleotidové záměny byly objeveny pouze v sekvencích obou proteinů a v 5' koncové oblasti a byla spočtena jejich frekvence (Tab. 9). Celkem bylo analyzováno 1261 nt.

Tab. 9 Frekvence nukleotidových záměn v proteinech a 5' UTR v rámci jednotlivých variant (procento zaměněných nukleotidů).

Část genomu	Varianta viru	Procento záměn v rámci varianty
5' konec	0. P.	0,5 %
	DL. P. IRE	0,9 %
	40. P. PS	0,4 %
	40. P. IRE	0,9 %
Protein C	0. P.	0,2 %
	DL. P. IRE	0,1 %
	40. P. PS	0,2 %
	40. P. IRE	0,4 %
Protein E	0. P.	0,4 %
	DL. P. IRE	0,3 %
	40. P. PS	0,4 %
	40. P. IRE	0,4 %

V průměru byla nejvyšší nukleotidová diverzita pozorována v nekódující 5' oblasti. Jednotlivé virové varianty se až na 40. P. IRE lišily velmi málo.

Většina jednonukleotidových záměn byla nalezena vždy jen u jediného klonu v rámci virové varianty, několik záměn bylo však detekováno v několika koloniích nezávisle. První z těchto záměn byla nalezena v 5' nekódující oblasti varianty 40. P. IRE v pozici 52. Po klonování byly získány sekvence z celkem deseti kolonií. Osm z nich má v poloze 52 adenin, dvě guanin. Zajímavé je, že guanin se v této pozici vyskytl u všech kolonií získaných

z ostatních variant viru (Obr. 18). Nukleotidová záměna převažující v koloniích se shoduje s celogenomovou sekvencí.

	46	112
40.P.IREcelogenom.sekvence		
	GCGGTTAGTTTGAAAGAAATATTCTTTTGTTTTACCAGTCGTGAACGTGTTGAGAAAAAGACAGCT	
40.P.IRE kolonie č. 1	GCGGTTGGTTTGAAAGAAATATTCTTTTGTTTTACCAGTCGTGAACGTGTTGAGAAAAAGACAGCT	
kolonie č. 2	ACGGTTGGTTTGAAAGAAATATTCTTTTGTTTTACCAGTCGTGAACGTGTTGAGAAAAAGACAGCT	
kolonie č. 3	GCGGTTAGTTTGAAAGAAATATTCTTTTGTTTTACCAGTCGTGAACGTGTTGAGAAAAAGACAGCT	
kolonie č. 4	GCGGTTAGTTTGAAAGAAATATTCTTTTGTTTTACCAGTCGTGAACGTGTTGAGAAAAAGACAGCT	
kolonie č. 5	GCGGTTAGTTTGAAAGAAATATTCTTTTGTTTTACCAGTCGTGAACGTGTTGAGAAAAAGGACAGCT	
kolonie č. 6	GCGGTTAGTTTGAAAGAAATATTCTTTTGTTTTACCAGTCGTGAACGTGTTGAGAAAAAGACAGCT	
kolonie č. 7	GCGGTTAGTTTGAAAGAAATATTCTTTTGTTTTACCAGTCGTGAACGTGTTGAGAAAAAGACAGCT	
kolonie č. 8	GCGGTTAGTTTGAAAGAAATATTCTTTTGTTTTACCAGTCGTGAACGTGTTGAGAAAAAGACAGCT	
kolonie č. 9	GCGGTTAGTTTGAAAGAAATATTCTTTTGTTTTACCAGTCGTGAACGTGTTGAGAAAAAGACAGCT	
kolonie č. 10	GCGGTTAGTTTGAAAGAAATATTCTTTTGTTTTACCAGTCGTGAACGTGTTGAGAAAAAGACAGCT	
0.P.celogenom. sekvence	GCGGTTGGTTTGAAAGAAATATTCTTTTGTTTTACCAGTCGTGAACGTGTTGAGAAAAAGACAGCT	
0.P. všechny kolonie (10)	GCGGTTGGTTTGAAAGAAATATTCTTTTGTTTTACCAGTCGTGAACGTGTTGAGAAAAAGACAGCT	
DL.P.IREcelogenom. sekvence	GCGGTTGGTTTGAAAGAAATATTCTTTTGTTTTACCAGTCGTGAACGTGTTGAGAAAAAGACAGCT	
DL.P.IRE všechny kolonie (11)	GCGGTTGGTTTGAAAGAAATATTCTTTTGTTTTACCAGTCGTGAACGTGTTGAGAAAAAGACAGCT	
40.P.PScelogenom. sekvence	GCGGTTGGTTTGAAAGAAATATTCTTTTGTTTTACCAGTCGTGAACGTGTTGAGAAAAAGACAGCT	
40.P.PS všechny kolonie (6)	GCGGTTGGTTTGAAAGAAATATTCTTTTGTTTTACCAGTCGTGAACGTGTTGAGAAAAAGACAGCT	

Obr. 18 Porovnání nukleotidových sekvencí 5' oblasti mezi jednotlivými koloniemi (záměny vyznačeny barevně).

K podobné záměně došlo u 40. P. IRE i v sekvenci proteinu C v pozici 315. Tato záměna byla ale synonymní, nezpůsobila tedy změnu AK. Po klonování mělo sedm kolonií v této pozici guanin, stejně jako celogenomová sekvence. Ve zbylých čtyřech koloniích byl detekován adenin, čímž se tyto sekvence shodovaly s celogenomovými sekvencemi všech zbylých variant a téměř všech jejich kolonií. Výjimku tvoří jedna kolonie 40. P. PS, která má v poloze 315, stejně jako většina kolonií 40. P. IRE, guanin (Obr. 19).

	301	327
40.P.IRE celogenom.sekvence		
	ATGAGCTGGTTGCTG	GTCATTACTCTG
40.P.IRE kolonie č. 1	ATGAGCTGGTTGCTA	AGTCATTACTCTG
kolonie č. 2	ATGAGCTGGTTGCTA	AGTCATTACTCTG
kolonie č. 3	ATGAGCTGGTTGCTA	AGTCATTACTCTG
kolonie č. 4	ATGAGCTGGTTGCTG	GTCATTACTCTG
kolonie č. 5	ATGAGCTGGTTGCTG	GTCATTACTCTG
kolonie č. 6	ATGAGCTGGTTGCTG	GTCATTACTCTG
kolonie č. 7	ATGAGCTGGTTGCTG	GTCATTACTCTG
kolonie č. 8	ATGAGCTGGTTGCTG	GTCATTACTCTG
kolonie č. 9	ATGAGCTGGTTGCTG	GTCATTACTCTG
kolonie č. 10	ATGAGCTGGTTGCTG	GTCATTACTCTG
kolonie č. 11	ATGAGCTGGTTGCTA	AGTCATTACTCTG
40.P.PS celogenom.sekvence	ATGAGCTGGTTGCTA	AGTCATTACTCTG
40.P.PS většina kolonií (8)	ATGAGCTGGTTGCTA	AGTCATTACTCTG
kolonie č. 9	ATGAGCTGGTTGCTG	GTCATTACTCTG
0.P. celogenom.sekvence	ATGAGCTGGTTGCTA	AGTCATTACTCTG
0.P. všechny kolonie (7)	ATGAGCTGGTTGCTA	AGTCATTACTCTG
DL.P.IRE všechny kolonie (13)	ATGAGCTGGTTGCTA	AGTCATTACTCTG
DL.P.IRE celogenom.sekvence	ATGAGCTGGTTGCTA	AGTCATTACTCTG

Obr. 19 Porovnání jednotlivých nukleotidových sekvencí mezi koloniemi v rámci proteinu C (záměny vyznačeny barevně).

Další podobná záměna se objevila v sekvenci proteinu E v poloze 913 u 40. P. PS. Tato varianta měla v celogenomové sekvenci a v sekvenci dvou kolonií guanin, zatímco u sekvence třetí kolonie se vyskytoval adenin, přítomný u převažujících variant a všech kolonií ostatních vzorků (Obr. 20).

	899	925
40.P.PS celogenom.sekvence		
	AAGGTCTTACGTACG	CAATGTGTGACA
40.P.PS kolonie č. 1	AAGGTCTTACGTACA	CAATGTGTGACA
kolonie č. 2	AAGGTCTTACGTACG	CAATGTGTGACA
kolonie č. 3	AAGGTCTTACGTACG	CAATGTGTGACA
0.P. celogenom.sekvence	AAGGTCTTACGTACA	CAATGTGTGACA
0.P. všechny kolonie (5)	AAGGTCTTACGTACA	CAATGTGTGACA
DL.P.IRE celogenom.sekvence	AAGGTCTTACGTACA	CAATGTGTGACA
DL.P.IRE všechny kolonie (5)	AAGGTCTTACGTACA	CAATGTGTGACA
40.P.IRE celogenom.sekvence	AAGGTCTTACGTACA	CAATGTGTGACA
40.P.IRE všechny kolonie (8)	AAGGCCTTACGTACA	CAATGTGTGACA

Obr. 20 Porovnání nukleotidových sekvencí proteinu E mezi koloniemi (záměny vyznačeny barevně).

4.3 Inhibice glykosylace tunikamycinem

Různé literární zdroje poukazují na význam glykosylace virových proteinů pro interakci s hostitelskou buňkou. Vliv glykosylace na patogenitu viru byl testován v následujícím pokusu.

Trojím pasážováním původní varianty VKE kmene Hypr na buňkách PS za přítomnosti antibiotika tunikamycinu byla vyselektována další variantaviru – 3. P. TM. Titr této varianty vzrostl během tří pasáží z původních 3×10^4 PFU/ml na $5,5 \times 10^7$ PFU/ml. Charakter plaků odpovídal původní variantě viru.

Vliv inhibice glykosylace na virulenci VKE byl zjišťován na myším modelu. Celkem 9 myši bylo infikováno dávkou 100 PFU a sledováno během dalších 30 dní. Všechny myši přežily a ani u jediné se neprojevyly žádné příznaky onemocnění.

Pro kontrolu toho, že za sníženou virulencí nestojí nějaká náhodná nezávislá mutace, byla zjištěna nukleotidová sekvence proteinu E, jako hlavního proteinu, odpovědného za interakci s buněčnými receptory a fúzi s membránou hostitelské buňky. Tato sekvence byla ale naprosto shodná se sekvencí proteinu E původní varianty. Výsledky tedy ukazují, že glykosylace proteinů VKE je nutným procesem, bez něhož je výrazně narušena virulence viru.

5. Diskuze

5.1 Biologické vlastnosti nově získaných variant viru

Dlouhodobým pasážováním viru ve dvou různých prostředích byly vyselektovány rozdílné varianty viru (40. P. PS a 40. P. IRE), lišící se biologickými vlastnostmi. První rozdíl byl patrný z charakteru plaků v kultuře buněk PS, kdy varianta viru adaptovaná na savčí buňky vytvářela ve většině případů velké plaky, zatímco varianta viru adaptovaná na buňky klíštěcí produkovala plaky přibližně poloviční velikosti.

Předpoklad, že varianta 40. P. PS, tvořící větší plaky v kultuře buněk PS, bude virulentnější, se během pokusu na laboratorních myších nepotvrdil. Výsledky stanovení střední doby přežití vykazovaly velkou variabilitu, z rychlosti nástupu onemocnění a počtů přeživších jedinců je však možné usuzovat míru virulence jednotlivých virových variant. Po subkutánní inokulaci vykazovala varianta 40. P. PS u myší ve srovnání s původní variantou a dokonce i s variantou adaptovanou na klíštěcí buňky, atenuovaný fenotyp. Velikost plaků v kultuře savčích buněk tedy nemusí vždy souviset s virulencí viru *in vivo* (Mayer, 1964). Obdobné výsledky byly získány v případě mutanty kmene VKE Oshima 5-10 (Goto *et al.*, 2003) i viru Langat (Mitzel *et al.*, 2008). Tvorba malých plaků v kultuře savčích buněk bývá spojována s mutacemi způsobujícími zvýšenou afinitu viru k heparan sulfátu – tato možnost bude diskutována v rámci dalšího odstavce.

Výsledky testování virulence na myším modelu se také lišily od očekávání, že varianta 40. P. PS, adaptovaná dlouhodobým pasážováním na savčí buňky, bude mít díky této adaptaci určitou selekční výhodu pro replikaci v myších a bude tedy pro myši virulentnější než varianta původní. Stejně tak byla vyvrácena i domněnka, že varianta 40. P. IRE, adaptovaná naopak na buňky klíštěcí, bude v myším modelu atenuovaná. Velkou roli v tomto případě zřejmě sehrál fakt, že virus byl pasážován *in vitro* na buněčných liniích a ne přímo v živých organismech (v myších a klíšťatech). Je totiž poměrně častým jevem, že virus pasážovaný *in vitro* na savčích buněčných liniích vykazuje sníženou virulenci *in vivo*. Mechanismus, který tuto atenuaci v živých organismech způsobuje, není ale zatím zcela znám. Nejčastěji popisovanými změnami v souvislosti s atenuací VKE *in vivo* jsou mutace v proteinu E, které vedou ke zvýšení jeho pozitivního náboje, čímž se zvyšuje afinita k heparan sulfátu (glykosaminoglykan). Ačkoli je afinita k heparan sulfátu pravděpodobně pro virus v buněčné kultuře (případně v klíšťatech) prospěšná, v myších způsobuje tato nově získaná vlastnost jeho atenuovaný fenotyp (Mandl *et al.*, 2001; Goto *et al.*, 2003; Romanova *et al.*, 2007; Kozlovskaya *et al.*, 2010). Mutanty se zvýšenou afinitou ke glykosaminoglykanům

navíc často tvoří drobnější plaky v kultuře savčích buněk, podobně jako 40. P. IRE. Zvýšení pozitivního náboje na proteinu E je nutnou, ne však jedinou podmínkou pro zvýšení afinity k heparan sulfátu (Kozlovskaya *et al.*, 2010). Aminokyselinové záměny v proteinu E, nalezené v této práci u 40. P. PS (v sekvenci 40. P. IRE nebyla v tomto proteinu nalezena žádná záměna), byly ovšem elektroneutrální, proto příčina atenuace této virové varianty pravděpodobně nesouvisela se změnou afinity ke glykosaminoglykanům. Pravděpodobnějším vysvětlením může být fakt, že varianta 40. P. PS měla být sice po dlouhodobém pasážování dokonale adaptovaná na savčí buňky, ve skutečnosti byla ale tato adaptace omezena pouze na jeden typ buněk a to buňky prasečí ledviny. Organismus hostitele je prostředím mnohem komplexnějším. Nejen, že se virus musí vypořádat s imunitním systémem, ale vnitřní prostředí hostitele mu poskytuje i daleko větší variabilitu co do typu buněk a pravděpodobně i buněčných receptorů (Kaufmann *et al.*, 2010).

Stejně jako pasážování *in vitro* na buněčné linii PS zřejmě zodpovídá za atenuaci virové varianty 40. P. PS, je pravděpodobné, že i pasážování viru na klíštěcích buňkách IRE/CTVM19 ovlivnilo fenotypové vlastnosti 40. P. IRE. Z výsledků stanovení střední doby přežití je zřejmé, že varianta viru 40. P. IRE byla pro myši, navzdory původním předpokladům, stejně patogenní jako původní virus. Dlouhodobým pasážováním viru přímo v klíštětech byly ale naproti tomu v minulosti mnohokrát vyselektovány varianty méně virulentní pro laboratorní myši, tvořící menší plaky v kultuře savčích buněk (Chunikhin *et al.*, 1986; Dzhivianian *et al.*, 1986, 1988, 1991; Romanova *et al.*, 2007). V práci Labudy *et al.* (1994) byly pasážovány dva kmeny VKE simultánně přes klíštěata a zatímco u jednoho došlo k výraznému snížení virulence pro myš již v páté pasáži, u druhého zůstala virulence téměř nezměněna i po pasáži sedmé. Snížení virulence viru po pasážování v systému klíštěcích buněk je tedy možným, nikoli však nutným důsledkem. Přestože se ale virulence 40. P. IRE adaptací na klíštěcí buňky nezměnila, varianta vytvářela menší plaky v kultuře buněk PS. Tímto se, stejně jako u 40. P. PS potvrdilo, že velikost plaků nemusí souviset s virulencí viru.

Dlouhodobá pasáž na klíštěcích buňkách (DL. P. IRE) vytvářela velké plaky v kultuře buněk PS a byla jen částečně atenuována ve srovnání s původním virem. Pokud uvažujeme, že velikost plaků nemusí souviset s virulencí, důvodem atenuovaného fenotypu u DL. P. IRE může být právě delší období inkubace viru s klíštěcími buňkami. Během jeden měsíc trvající inkubace nebyl vyvíjen selekční tlak na rychlost množení viru, protože nemusel napadat nové buňky každé čtyři dny, jak se tomu dělo u 40. P. IRE. Tím pádem se virus nestihl adaptovat tak dobře na měnící se podmínky (různé typy buněk a receptorů), nebyly

selektovány rychle se množící varianty, a proto také vykazoval nižší virulenci u myši než 40. P. IRE.

5.2 Dynamika množení viru v orgánech

Po subkutánní inokulaci dochází nejprve k pomnožení viru v kůži a podkožním vazivu (Labuda *et al.*, 1996), pravděpodobně v dendritických buňkách kůže (Lindenbach *et al.*, 2007). Infikované dendritické buňky putují do spádových lymfatických uzlin, kde dochází k další replikaci viru. Poté je virus uvolněn do krevního oběhu a nastává tak viremická fáze infekce (Málková *et Fraňková*, 1959). Krví se dále šíří a napadá různé orgány. Jedním z nejvíce postižených orgánů je slezina, která hraje významnou roli v imunitní obraně organismu (McMinn *et al.*, 1996a).

Průběh infekce u jednotlivých variant viru poměrně dobře odpovídal průběhu pokusu stanovení virulence pro laboratorní myši. V krvi se všechny varianty viru objevily druhý až třetí den po infekci. Virémie u 0. P. a 40. P. IRE dosahovala obdobných hodnot a trvala přibližně do 7. dne, 40. P. PS však vykazovala až do 7. dne velmi nízkou virémii (na hranici detekce). Opožděný nástup virémie a nižší titr viru v krvi u 40. P. PS jsou pravděpodobně způsobeny sníženou schopností 40. P. PS pomnožit se v podkoží a napadat dendritické buňky.

Ve slezině se virus všech tří variant objevil třetí až pátý den, u všech variant byla pozorována dvoufázová dynamika množení. Titr 40. P. PS dosahoval oproti oběma zbývajícím variantám přibližně polovičních hodnot.

Po pomnožení v krvi a ve vnitřních orgánech prochází virus přes hematoencefalickou bariéru a napadá mozek. Pokud je virus dostatečně neurovirulentní, způsobí v mozku cytopatický efekt a vyvolá vznik encefalitidy (McMinn *et al.*, 1996a). Obě virulentnější varianty, 0. P. i 40. P. IRE, byly detekovány v mozcích myši již 5. den po infekci, tedy poměrně záhy poté, co byly poprvé detekovány v krvi. Titr těchto dvou variant v mozku velmi strmě rostl až do 7. dne. Varianta 40. P. PS byla v mozku detekována poprvé až 8. den, kdy byl tento pokus ukončen vzhledem k virulenci obou zbývajících variant. Je pravděpodobné, že maximálních hodnot by dosáhl titr této varianty až po ukončení pokusu. Tato varianta tedy opět vykazovala atenuovaný fenotyp.

Poté, co se virus dostane do mozku, dochází většinou k poklesu jeho množství v krvi a orgánech. Po úspěšném napadení mozku a vyvolání cytopatického efektu, na kterém se zřejmě podílejí i buňky imunitního systému (Růžek *et al.*, 2009), nastává druhá fáze virémie,

kdy virus zaplavuje celý organismus. Tento mechanismus vysvětluje i dvoufázový nárůst titru viru ve slezině u myši infikovaných všemi variantami viru.

Porovnáním výsledků se stanovením virulence zjistíme, že myši začínaly hynout v době, kdy byl virus v mozku již pomnožen do vysokých titrů (přibližně 3-4 dny poté, co byl virus poprvé v mozku detekován). Rozdíly ve virulenci mezi 0. P, 40. P. IRE a atenuovanou variantou 40. P. PS mohou být přisuzovány nižší efektivitě množení viru před vstupem do mozku (tedy snížení neuroinvasivity). Vzhledem k načasování pokusu stanovení dynamiky množení bohužel není možné posoudit, zda se na snížení virulence podílela i snížená neurovirulence. Definitivní odpověď bychom získali porovnáním titrů variant viru v mozku po intracerebrální inokulaci. Snížená neuroinvasivita byla jako příčina snížené virulence viru popsána například ve studii Mitzelové *et al.* (2008).

5.3 Potenciální determinanty virulence

Variety viru s různou historií pasáží vykazovaly různé fenotypové vlastnosti co do morfologie plaků, míry patogenity pro laboratorní myši i dynamiky množení viru v myším modelu. Porovnáním celogenomových sekvencí jednotlivých variant byly identifikovány nukleotidové a aminokyselinové záměny unikátní pro jednotlivé varianty lišící se biologickými vlastnostmi. V následujících odstavcích bude těmto potenciálním determinantám virulence věnována pozornost.

Biologickými vlastnostmi se od původní varianty viru 0. P. nejvíce odlišovala 40. P. PS. V sekvenci této varianty bylo identifikováno také nejvíce nukleotidových záměn (celkem 11), z nichž 6 bylo nesynonymních a způsobilo tedy i změnu aminokyseliny. Dvě z těchto aminokyselinových záměn se vyskytovaly v sekvenci proteinu E a byla jim věnována zvláštní pozornost, vzhledem k několika zásadním funkcím, které tento protein během životního cyklu viru zastává.

Z 3D modelů proteinu E bylo zjištěno, že záměna Thr (305) → Ala leží v oblasti spojující doménu I a doménu III, která je označována jako „D I – D III linker“. Tento 15 aminokyselin dlouhý řetězec leží v polohách 296 – 310 AK. Během vystavení proteinu E kyselému pH dochází k výrazným konformačním změnám, které umožňují zaujmutí správné pozice domény III v trimeru proteinu E (Bressanelli *et al.*, 2004). Předpokládáme tedy, že mutace, přesto že nemá vliv na rozložení náboje, může stericky ovlivňovat interakce mezi monomery proteinu E a tedy i celý proces fúze virionů s membránou endosomu.

Teoreticky může virulenci ovlivňovat i druhá záměna v proteinu E Pro (360) → Ser, která je umístěna v doméně III, v blízkosti oblasti, u které se předpokládá zapojení do vazby

na buněčné receptory. Receptorová doména dosud nebyla přesně vymezena, a proto není možné vyloučit z podílu na vazbě k receptoru veškeré aminokyseliny v blízkosti této oblasti. Dosud bylo v předpokládané receptorové oblasti domény III identifikováno několik záměn, které měly vliv na virulenci viru. Přímo u VKE byly takové záměny nalezeny v polohách 384 (Holzmann *et al.*, 1990), 310 (Mandl *et al.*, 2000) nebo 368 (Holzmann *et al.*, 1997). Všechny tyto záměny vedly ke snížené virulenci na myším modelu. Z formování 3D struktury proteinu E viru Langat vyplývá, že podíl na vazbě na buněčný receptor mohou mít i aminokyseliny v poloze 347, 348 a řetězec β -listu 6 (388 – 394 AK) (Mukherjee *et al.*, 2006).

Další aminokyselinovou záměnou unikátní u 40. P. PS byla záměna Val (202) → Ala v proteinu NS2A. Tato záměna leží v N-koncové části proteinu NS2A. Tento protein je asociovaný s membránou a je součástí replikačního komplexu flavivirů (Mackenzie *et al.*, 1998). Podílí se na skládání virionu a uvolňování infekčních částic z hostitelských buněk. Ovlivnění skládání virionů bylo prokázáno pro mutace ve štěpicím místě (Kümmerer *et al.*, 2002) a pro změny vedoucí k narušení hydrofobních domén N (Leung *et al.*, 2008). Naše mutace Val (202) → Ala byla záměnou hydrofobní AK za jinou hydrofobní AK a je lokalizována v N-koncové části proteinu, takže pravděpodobně nebude mít přímý vliv na virulenci viru. Protein NS2 je obecně jedním z nejméně konzervovaných proteinů VKE (Chambers *et al.*, 1990a).

Další aminokyselinová záměna v rámci 40. P. PS byla nalezena v proteinu NS4B, v pozici Phe (88) → Ile. NS4B je transmembránový protein se zatím ne zcela známou funkcí (Miller *et al.*, 2006). Bylo ale zjištěno, že záměny v tomto proteinu mohou souviset s adaptací flavivirů na specifického hostitele (Mitzel *et al.*, 2008). Je proto možné, že tato záměna souvisela s přizpůsobením varianty 40. P. PS na buňky PS.

Poslední dvě aminokyselinové záměny v sekvenci 40. P. PS byly nalezeny v oblasti kódující protein NS5, a to v polohách Leu (778) → Phe a Asn (863) → Ser. Tento vysoce konzervovaný, bifunkční protein funguje jako methyltransferáza a také jako RNA-dependentní RNA polymeráza (Koonin, 1993). S-adenosyl methionin methyltransferáza je zodpovědná za metylaci guaninu v pozici N-7 a ribózy v pozici 2'-O při formování RNA čepičky, jejíž přítomnost je nezbytná pro iniciaci translace. Funkční doména S-adenosyl methionin methyltransferázy se nachází na N-konci proteinu v polohách 1 – 296. Další významnou částí AK sekvence tohoto proteinu jsou oblasti mezi aminokyselinami 320 – 405, kde leží dvě sekvence pro směřování proteinu do jádra buňky („NLS - nuclear localization sequences“). Katalytická doména RNA-dependentní RNA polymerázy se nachází v poloze AK 270 – 900 a obsahuje 6 vysoce konzervovaných oblastí (Yap *et al.*, 2007). Obě

aminokyselinové záměny nalezené u 40. P. PS leží v oblasti katalytické domény virové RNA polymerázy, ovšem mimo konzervované úseky, a pravděpodobně tedy neovlivňují její funkci.

Druhou variantou, vykazující alespoň částečně atenuovaný fenotyp ve srovnání s variantou původní, byla DL. P. IRE. Tato varianta se od původního viru lišila 3 nukleotidovými záměnami, z nichž 2 vedly k záměně aminokyseliny. Jedna z těchto AK záměn, Met (471) → Leu, se nacházela v proteinu E, v oblasti ležící mimo hlavní ektodoménu proteinu. Tato oblast se podílí na zakotvení proteinu do membrány, na interakci s proteinem prM a na pH závislých konformačních změnách (Hurrelbrink *et al.*, 2003). Poloha 471 spadá konkrétně do oblasti sloužící k zakotvení proteinu do membrány („anchor“). I v jiné studii byla u VKE po adaptaci na klíštěcí buňky nalezena mutace v této oblasti mimo ektodoménu proteinu E, konkrétně v poloze 426 (oblast stopky – „stem“). I tato záměna se podílela na snížení virulence pro laboratorní myši (Romanova *et al.*, 2007). Také v další studii měla mutace, ležící mimo hlavní ektodoménu proteinu E v poloze 496, vliv na neuroinvazivitu viru (Gritsun *et al.*, 2001). Je tedy možné, že záměna v poloze 471 v sekvenci proteinu E u DL. P. IRE způsobila částečnou atenuaci viru.

Další aminokyselinová záměna DL. P. IRE se nacházela v proteinu NS5, konkrétně v pozici Thr (177) → Pro. Spadala tedy do funkční domény pro methyltransferázu. Z dřívějších studií vyplývá, že jednonukleotidové záměny v proteinu NS5 mají sice vliv na atenuaci viru, ale podílejí se spíše na kumulativním efektu jednobodových mutací, kde největší vliv na snížení virulence mají mutace v proteinu E a mutace v NS5 k atenuaci pouze přispívají (Gritsun *et al.*, 2001; Hayasaka *et al.*, 2004).

Varianta 40. P. IRE vykazovala sice téměř shodnou virulenci pro laboratorní myši jako původní varianta 0. P., přesto se od ní svojí sekvencí lišila v 9 nukleotidech, z nichž 4 způsobily i záměnu AK. Navíc se tato varianta od všech zbývajících lišila morfologií plaků.

Jedna z mutací, vedoucích ke změně AK, se nacházela v pozici Thr (34) → Ile v proteinu prM. Touto záměnou došlo ke zrušení jediného potenciálního glykosylačního místa proteinu prM Asn -X-Thr, kde threonin byl nahrazen izoleucinem. Protein prM sehrává důležitou úlohu jako chaperonin proteinu E (Guirakhoo *et al.*, 1991). Goto *et al.* (2005) studovali význam glykosylací proteinů prM a E pomocí VLP („virus like particles“), které mají povrchové vlastnosti shodné s infekčními viriony. Bylo zjištěno, že mutace v glykosylačním místě prM způsobí významný pokles sekrece VLP oproti glykosylované

variantě (Goto *et al.*, 2005). Je tedy možné, že mutace v proteinu prM měla za následek hromadění virionů v buňkách, tedy i pomalejší šíření viru z buňky do buňky a tedy i plaky menší velikosti v kultuře buněk PS.

Druhou mutací, jež způsobila změnu AK, byla záměna v pozici Val (57) → Ile v proteinu NS1. Tento s membránou asociovaný glykoprotein má důležitou, ale dosud ne zcela známou funkci v replikaci virové RNA. Mutace v glykosylačních místech proteinu měly za následek defekty RNA, které vedly k omezené produkci viru (Muylaert *et al.*, 1996). Záměna ve 40. P. IRE ale neležela ani v jednom ze tří potenciálních glykosylačních míst, ani v žádném z 12 konzervovaných cysteinů, tvořících disulfidické můstky.

Další aminokyselinová záměna se nacházela v pozici Val (33) → Met v proteinu NS2B. Tento protein vytváří stabilní komplex s proteinem NS3 a slouží jako kofaktor NS2B-NS3 serinové proteázy (Falgout *et al.*, 1991). Na kofaktorové aktivitě se podílí 47 AK zbytků ležících v centrální části proteinu, kterou se NS2B vmezeřuje do ohybu domény serinové proteázy (Erbel *et al.*, 2006). Vzhledem k tomu, že mutace v poloze 33 leží mimo kofaktorovou doménu, pravděpodobně přímo neovlivňuje proteázovou aktivitu komplexu.

Poslední mutace se nacházela v poloze Ala (85) → Ser v proteinu NS4B. Jak je již uvedeno výše, záměny v tomto proteinu se pravděpodobně podílejí na adaptaci na specifického hostitele (Mitzel *et al.*, 2008). V tomto případě by tedy mohla tato mutace souviset s adaptací 40. P. IRE na klíštěcí buňky.

Jak je již zmíněno výše, nejen mutace v proteinech, ale i v koncových nekódujících oblastech, mohou ovlivňovat patogenitu viru. Právě oba koncové netranslatované úseky vykazovaly v naší práci oproti kódujícím úsekům vyšší frekvenci nukleotidových záměn. U 40. P. IRE se vyskytovala mutace v 5' oblasti (pozice G (52) → A) i v 3' oblasti (pozice T (282) → C). K naprosto shodné mutaci došlo na 3' konci i u varianty 40. P. PS.

Obě koncové nekódující oblasti se podílejí svými sekundárními strukturami na regulaci replikace, translace a balení viru. Proto zejména vícebazové delece ovlivňují virulenci a životaschopnost viru (Cahour *et al.*, 1995; Proutski *et al.*, 1999). Rozsah atenuace záleží na tom, v jaké konzervované oblasti k deleci došlo (Proutski *et al.*, 1999).

V minulosti byly popsány i jednobazové mutace v 5' oblasti v souvislosti s produkcí menších plaků (Butrapet *et al.*, 2000; Romanova *et al.*, 2007). Je proto možné, že záměna G (52) → A u 40. P. IRE měla vliv na jedinou fenotypovou vlastnost, kterou se tato varianta lišila od ostatních, a tou je tvorba malých plaků.

Mutace v 3' oblasti T (282) → C byla shodná u 40. P. PS a u 40. P. IRE. Tato mutace sice ležela v konzervovanější distální části 3' konce, ale vzhledem k tomu, že se obě varianty lišily jak velikostí plaků, tak mírou virulence *in vivo*, neměla pravděpodobně tato záměna vliv na nově získané fenotypové vlastnosti ani u jedné z variant.

Tvorba malých plaků v kultuře buněk PS bývá nejčastěji spojována s mutacemi zvyšujícími afinitu k heparan sulfátu (Mandl *et al.*, 2001, Romanova *et al.*, 2007). Změny afinity jsou v těchto případech vyvolány lokálním zvýšením pozitivního náboje na povrchu proteinu E. V našem případě nebyla u varianty 40. P. IRE mutace v proteinu E zaznamenána, tvorba malých plaků musí tedy být vyvolána jiným mechanismem. Mutanty s vysokou afinitou ke glykosaminoglykanům bývají méně virulentní pro myši.

Souhrnně lze tedy říci, že za sníženou virulenci *in vivo* u 40. P. PS pravděpodobně zodpovídají obě záměny v proteinu E, zatímco mutace v proteinu NS4B vznikla nejspíše jako důsledek adaptace viru na buňky PS. Za pouze částečně atenuovaný fenotyp DL. P. IRE pravděpodobně také zodpovídá zejména mutace v proteinu E, ovšem za přispění mutace v proteinu NS5. Tvorba malých plaků u 40. P. IRE je pravděpodobně důsledkem mutací v proteinu prM a v 5' koncové oblasti, zatímco mutace v proteinu NS4B je důsledkem adaptace viru na klíčecí buňky.

5.4 Variabilita v rámci virových variant

Výsledky několika prací zaměřených na studium atenuovaných kmenů naznačují, že adaptace viru na určité prostředí neprobíhá na základě aktuálních náhodných mutací, ale že je adaptovaná varianta spíše selektována z již existujících sekvenčních variant (Labuda *et al.*, 1994; Růžek *et al.*, 2008; Romanova *et al.*, 2007). Při plakové titraci varianty 40. P. PS byly navíc kromě převažujících plaků, velikostí shodných s původní variantou 0. P., zachyceny i plaky velikostně podobné variantě 40. P. IRE. Z těchto důvodů jsme se rozhodli blíže prozkoumat sekvenční variabilitu v rámci našich virových variant.

Metoda konformačního polymorfismu jednořetězců je široce užívanou technikou pro detekci jednonukleotidových mutací ve směsi DNA fragmentů o stejné délce. Proto je vhodná také ke studiu „quasispecies“ - heterogenní směsi geneticky rozdílných, ale blízce příbuzných variant v rámci téhož izolátu (Enomoto *et al.*, 1994; Maggi *et al.*, 1997). V našem případě se nepodařilo jednotlivé sekvenční varianty, později potvrzené klonováním, pomocí této metody oddělit. Jak již bylo zmíněno výše, existuje mnoho faktorů, které mohou průběh SSCP analýzy ovlivnit. Kromě nejrůznějších fyzikálních a chemických podmínek vlastní separace je to především samotná nukleotidová sekvence a délka fragmentů DNA. Všeobecně se předpokládá, že záměny jedné báze se projeví spíše u fragmentů kratších než

300 bp. Vzhledem k tomu, že v této práci byly pro SSCP analýzu použity fragmenty o velikosti 500-700 bp, je možné, že právě velikost fragmentů byla příčinou neúspěchu při použití této metody.

Jako alternativní metody bylo použito klonování. Před klonováním byly vybrané úseky amplifikovány pomocí PCR. Použitá Taq polymeráza nemá korekturní aktivitu, a proto se předpokládá, že některé ze záměn byly zavedeny uměle v amplifikačním kroku. Celkovou nukleotidovou variabilitu u jednotlivých variant však porovnáváme relativně, takže z tohoto pohledu by neměla chybovost polymerázy výsledky významněji ovlivnit.

Klonovány a následně sekvenovány byly úseky 5' nekódující oblasti, proteinu E a proteinu C. Hlavní pozornost byla soustředěna na nukleotidovou variabilitu v pozicích, ve kterých se vzájemně lišily celogenomové sekvence jednotlivých virových variant. Úseky zahrnující tyto pozice byly osekvenovány v několika nezávislých klonech, aby tak byla minimalizována možnost vnášení mutací chybovostí Taq DNA polymerázy při PCR. Konkrétně se jednalo o záměny v 5' nekódující oblasti (G (52) → A), v proteinu C (A (315) → G) a v proteinu E (A (913) → G).

V případě záměny v proteinu C se jednalo o synonymní mutaci, která neměla za následek změnu v proteinu. Pravděpodobně tedy nebyla vystavena žádnému selekčnímu tlaku. To potvrzuje i skutečnost, že obě možnosti zároveň (A i G) se vyskytly nezávisle u 40. P. IRE a 40. P. PS. V tomto ohledu je poněkud překvapující poměr ve frekvenci výskytu obou bazí u 40. P. PS, kde silně převažuje A (8:1). Vysvětlením by mohl být nedávný vznik této mutace.

Vzhledem k adaptaci viru na prostředí jsou mnohem zajímavější obě zbývající záměny – v 5' nekódující oblasti a v sekvenci proteinu E (nekonzervativní záměna - Thr (305) → Ala). V obou případech byly pozorovány stejné okolnosti: u jedné varianty viru byla zjištěna nukleotidová variabilita v rámci varianty, přičemž minoritně zastoupená база se shodovala s celogenomovou sekvencí zbývajících variant i všech jejich kolonií. Prakticky shodný charakter výsledků byl pozorován ve studii Mitzelové *et al.* (2008) na viru Langat a Jerzakové *et al.* (2005) na viru západního Nilu. Je tedy možné, že po vzniku mutace, která se následně zdá být pro dané prostředí výhodnou, je určitým mechanismem stále zachovávána i původní verze, i když s nižší frekvencí. Takový mechanismus odpovídá i výsledkům Romanové *et al.* (2007), kde byl po mnohonásobné pasáži přes klíště získán kmen VKE v kultuře savčích buněk tvořící převážně malé plaky. Kromě malých plaků se s nízkou frekvencí vyskytovaly i plaky velké. Po plakové purifikaci a sekvenování byly zjištěny nukleotidové záměny unikátní pro virus tvořící malé plaky a pro fenotypové revertanty.

Fenotypová variabilita po mnohonásobném pasážování VKE přes klíšťata byla pozorována i v práci Labudy *et al.* (1994). Dále ve studii Růžka *et al.* (2008) byly z přirozeně atenuovaného kmene selektovány 2 virulentní varianty s identickými záměnami v genomu, jedna během 5 pasáží v myších, druhá během jedné pasáže na buňkách PS při 40 °C. Bohužel nemůžeme s jistotou říci, zda námi analyzované mutace opravdu poskytují v daném prostředí selekční výhodu a tedy zda se nejedná pouze o ekvivalentní alternativy bez vlivu na vlastnosti viru. Alespoň záměna v proteinu E, mající za následek změnu aminokyseliny v propojení domény I s doménou III, však funkci proteinu velmi pravděpodobně ovlivňuje, jak bylo diskutováno výše. V případě mutace v 5' oblasti jejímu významu napovídá zastoupení bazí mezi klony, kde silně převažuje A nad G (8:2). Přímý vztah jednonukleotidové záměny v 5' NTR ke změně fenotypových vlastností, konkrétně velikosti plaků, je spíše méně pravděpodobný. Jiní autoři poukazují na kumulativní efekt jednotlivých jednonukleotidových záměn a jejich kombinací na vlastnosti viru (Davis *et al.*, 2007; Mitzel *et al.*, 2008). Pro sekundární strukturu RNA a tedy i replikaci viru mohou mít význam i synonymní mutace, jejich hromadění a kombinace.

5.5 Význam glykosylace VKE pro biologické vlastnosti viru

Glykosylace je důležitou ko- a post-translační modifikací virových proteinů. V genomu flavivirů se nachází 1 potenciální glykosylační místo v proteinu prM (pozice 32-34 AK), 2 v proteinu E (první v pozicích 154-156, další na C-konci proteinu) a 3 v proteinu NS1 (v pozicích 85-87, 207-209, 223-225). Glykoproteiny, vzniklé při glykosylaci, mají význam pro kompletaci virionu a jeho stabilitu, ovlivňují vstup viru do hostitelských buněk a jeho následné uvolnění i antigenicitu viru (Chambers *et al.*, 1990a; Vigerust *et al.* Shepherd, 2007).

Pomocí mutací v potenciálních glykosylačních místech bylo zjištěno, že ztráta glykosylace má různý (často i zcela opačný) efekt v závislosti na umístění glykosylace, konkrétním flaviviru, typu prostředí, ve kterém se virus množí (Beasley *et al.*, 2005; Bryant *et al.*, 2007; Moudy *et al.*, 2009), a fázi replikačního cyklu viru (Hanna *et al.*, 2005; Goto *et al.*, 2005).

V této práci byla zachycena jediná mutace mající za následek zrušení potenciálního glykosylačního místa, a to mutace v pozici Thr (34) → Ile proteinu prM (tato záměna byla diskutována výše). V případě viru klíšťové encefalitidy je informací k možné roli jednotlivých glykosylací v interakci viru s různými typy buněk poměrně málo. Proto jsme se rozhodli studovat vliv glykosylace pomocí kompletní inhibice glykosylace tunikamycinem.

Virová varianta 3. P. TM vykazovala v kultuře buněk PS obdobnou morfologii plaků jako původní varianta. Při testování virulence na myším modelu všechny myši přežily po celou dobu sledování (30 dní), aniž by se u nich projevil jakékoli příznaky onemocnění. Sekvenací proteinu E, jakožto hlavního strukturního proteinu, byly vyloučeny náhodné mutace, ke kterým mohlo během pasážování dojít. Výsledky našeho pokusu, stejně jako ostatních studií, tedy naznačují, že glykosylace virových proteinů je významným procesem, ovlivňujícím zásadním způsobem virulenci viru. Význam jednotlivých glykosylací je pro jednotlivé viry výrazně specifický a informací ke glykosylacím VKE je poměrně málo. Můžeme tedy pouze odhadovat, že absence glykosylace proteinu E může významně ovlivňovat vstup viru do buňky. Podle výše zmiňované práce Gota *et al.* (2005) glykosylace proteinů prM a E ovlivňují efektivitu sekrece virionů z buňky.

6. Závěr

Virus klíšťové encefalidity kmene Hypr byl pasážován na dvou odlišných typech buněk a bylo sledováno, jakým způsobem adaptace viru na různá prostředí ovlivní jeho biologické vlastnosti. Varianta viru adaptovaná na savčí buňky vykazovala opožděný nástup virémie v myším modelu, který byl způsoben sníženou neuroinvazivitou. Za atenuovaný fenotyp pravděpodobně zodpovídaly dvě aminokyselinové záměny v proteinu E. Virulence varianty viru pasážované na klíšťecích buňkách se v porovnání s původní variantou nezměnila, jediným rozdílem byla produkce malých plaků v kultuře buněk PS. Ta byla pravděpodobně způsobena mutací v 5' koncové oblasti a v glykosylačním místě proteinu prM. Variabilita klonů různým způsobem pasážovaného viru naznačuje přítomnost více sekvenčních variant viru v rámci téhož izolátu. Ze studia glykosylace vyplývá, že tato úprava naprosto zásadním způsobem ovlivňuje virulenci viru.

7. Seznam zkratek

Ala	alanin
Asn	asparagin
ATB	antibiotikum
BOFES	bovinní fetální sérum
DNA	deoxyribonukleová kyselina
ER	endoplazmatické retikulum
Glu	kyselina glutamová
Gly	glycin
Ile	izoleucin
Leu	leucin
Lys	lysin
Met	methionin
nt	nukleotid
ORF	otevřený čtecí rámec
Phe	fenylalanin
Pro	prolin
PTS	prekolostrální telecí sérum
RNA	ribonukleová kyselina
Ser	serin
Thr	threonin
UTR	netranslatovaná oblast
Val	valin
VKE	virus klíšřové encefalidity

8. Literatura

- Alcamí A., Symons J. A., Khanna A., Smith G. L. (1998). Poxviruses: capturing cytokines and chemokines. *Seminars in Virology*. 5: 419-427.
- Allison S. L., Stiasny K., Standler K., Mandl C. W., Heinz F. X. (1999). Mapping of functional elements in the stem-anchor region of tick-borne encephalitis virus envelope protein E. *Journal of Virology*. 73: 5605-5612.
- Arnold K., Bordoli L., Kopp J., Schwede T. (2006). The SWISS-MODEL Workspace: A web-based environment for protein structure homology modelling. *Bioinformatics*. 22: 195-201.
- Bazan J. F., Fletterick R. J. (1989). Detection of a trypsin-like serine protease domain in flaviviruses and pestiviruses. *Virology*. 171: 637-639.
- Beasley D. W., Whiteman M. C., Zhang S., Huang C. Y., Schneider B. S., Smith D. R., Gromowski G. D., Higgs S., Kinney R. M., Barrett A. D. (2005). Envelope protein glycosylation status influences mouse neuroinvasion phenotype of genetic lineage 1 West Nile virus strains. *Journal of Virology*. 79: 8339-8347.
- Bell-Sakyi L. (2004). *Ehrlichia ruminantium* grows in cell lines from four ixodid tick genera. *Journal of comparative pathology*. 130: 285-293.
- Bernfield M., Götte M., Park P. W., Reizes O., Fitzgerald M. L., Lincecum J., Zako M. (1999). Functions of cell surface heparan sulfate proteoglycans. *Annual review of biochemistry*. 68: 729-777.
- Blitvich B. J., Scanlon D., Shiell B. J., Mackenzie J. S., Pham K., Hall R. A. (2001). Determination of the intramolecular disulfide bond arrangement and biochemical identification of the glycosylation sites of the nonstructural protein NS1 of Murray Valley encephalitis virus. *The Journal of General Virology*. 82: 2251-2256.
- Bressanelli S., Stiasny K., Allison S. L., Stura E. A., Duquerroy S., Lescar J., Heinz F. X., Rey F. A. (2004). Structure of a flavivirus envelope glycoprotein in its low-pH-induced membrane fusion conformation. *The EMBO Journal*. 23: 728-738.
- Bruett L., Barber S. A., Clements J. E. (2000). Characterization of a membrane-associated protein implicated in visna virus binding and infection. *Virology*. 271: 132-141.
- Bryant J. E., Calvert A. E., Mesesan K., Crabtree M. B., Volpe K. E., Silengo S., Kinney R. M., Huang C. Y.-H., Miller B. R., Roehrig J. T. (2007). Glycosylation of the dengue 2 virus E protein at N67 is critical for virus growth *in vitro* but not for growth in intrathoracically inoculated *Aedes aegypti* mosquitoes. *Virology*. 366: 415-423.
- Burke D. S., Monach T. P. (2001). Flaviviruses. In: Knipe D. M., Howley P. M. (Eds.), *Fields Virology*, vol. 1. Lippincott Williams & Wilkins, London, New York, Tokyo, pp. 1043-1125.

- Butrapet S., Huang C. Y.-H., Pierro D. J., Bhamarapravati N., Gubler D. J., Kinney R. M. (2000).** Attenuation markers of a candidate dengue type 2 vaccine virus, strain 16681 (PDK-53), are defined by mutations in the 5' noncoding region and nonstructural proteins 1 and 3. *Journal of Virology*. 7: 3011-3019.
- Cahour A., Pletnev A., Vazielle-Falcoz M., Rosen L., Lai C. J. (1995).** Growth-restricted dengue virus mutants containing deletions on the 5' noncoding region of the RNA genome. *Virology*. 207: 68-76.
- Cecilia D., Gould E. A. (1991).** Nucleotide changes responsible for loss of neuroinvasiveness in Japanese encephalitis virus neutralization-resistant mutants. *Virology*. 181: 70-77.
- Chambers T. J., Hahn C. S., Galler R., Rice C. M. (1990a).** Flavivirus genome organization, expression, and replication. *Annual review of microbiology*. 44: 649-688.
- Chambers T. J., Weir R. C., Grakoui A., McCourt D. W., Bazan J. F., Fletterick R. J., Rice C. M. (1990b).** Evidence that the N-terminal domain of yellow fever virus NS3 is a serine protease responsible for site-specific cleavages in the viral polyprotein. *Proceedings of the National Academy of Sciences of the United States of America*. 87: 8898-8902.
- Charrel R. N., Zaki A. M., Attoui H., Fakeeh M., Billoir F., Yousef A. I., de Chesse R., De Micco P., Gould E. A., de Lamballerie X. (2001).** Complete cosiny sequence of the Alkhurma virus, a tick-borne flavivirus causing severe hemorrhagic fever in humans in Saudi Arabia. *Biochemical and Biophysical Research Communications* 287: 445-461.
- Chen Y., Maguire T., Hileman R. E., Fromm J. R., Esko J. D., Linhardt R. J., Marks R. M. (1997).** Dengue virus infectivity depends on envelope protein binding to target cell heparan sulfate. *Nature Medicine*. 3: 866-871.
- Chunikhin S. P., Reshetnikov I. N., Liapustin V. N. (1986).** Variability of tick-borne encephalitis virus during passage through *Ixodes* ticks and small mammals. *Meditinskaiia parazitologii i parazitarnye bolezni*. 6: 58-61.
- Clum S., Ebner K. E., Padmanabhan R. (1997).** Cotranslational membrane insertion of the serine proteinase precursor NS2B-NS3 (Pro) of dengue virus type 2 is required for efficient in vitro processing and is mediated through the hydrophobic regions of NS2B. *The journal of biological chemistry*. 272: 30715-30723.
- Davis C. T., Galbraith S. E., Zhang S., Whiteman M. C., Li L., Kinney R. M., Barrett A. D. T. (2007).** A combination of naturally occurring mutations in north american west Nile virus nonstructural protein genes and in the 3' untranslated region alters virus phenotype. *Journal of Virology*. 11: 6111-6116.
- Davis C. W., Nguyen H. Y., Hanna S. L., Sanchez M. D., Doms R. W., Pierson T. C. (2006).** West Nile virus discriminates between DC-SIGN and DC-SIGNR for cellular attachment and infection. *Journal of Virology*. 80: 1290-1301.

- De Madrid A. T., Porterfield J. S. (1969).** A simple microculture method for the study of group B arboviruses. *Bulletin of WHO.* 40: 113 - 121.
- Domingo E. (1989).** RNA virus evolution and the control of viral disease. In: Junker E. (editor). *Progress in drug research.* Basel: Birkhauser. 33: 93-133.
- Domingo E. (1996).** Biological significance of viral quasispecies. *Viral Hepatitis Review.* 2: 247-261.
- Domingo E., Holland J. J. (1997).** RNA virus mutations and fitness for survival. *Annual Review of Microbiology.* 51: 151-178.
- Domingo E., Holland J. J., Biebricher C., Eigen M. (1995).** Quasispecies: the concept and the word. In: Gibbs A., Calisher C., García-Arenal F. (Eds.), *Molecular basis of virus evolution.* Cambridge: Cambridge University Press. p. 171-180.
- Domingo E., Menéndez-Arias L., Quiñones-Mateu M. E., Holguín A., Gutierrez-Rivas M., Martínez M. A., Quer J., Novella I. S., Holland J. J. (1997).** Viral quasispecies and the problem of vaccine-escape and drug-resistant mutants. *Progress in drug research.* 48: 99-128.
- Domingo E., Sabo D. L., Taniguchi T., Weissmann C. (1978).** Nucleotide sequence heterogeneity of an RNA phage population. *Cell.* 13: 735-744.
- Dzhivaniyan T. I., Chunikhin S. P., Lisak V. M., Kashtanova G. M., Korolev M. B. (1986).** Immunochemical characteristics of the antigens of a variant tick-borne encephalitis virus adapted to *Hyalomma plumbeum* ticks. *Voprosy virusologii.* 1: 92-96.
- Dzhivaniyan T. I., Karganova G. G., Sobolev S. G., Korolev M. B., Kashtanova G. M., Chuprinskaia M. V., Lashkevich V. A. (1991).** The properties of the particles formed in the reproduction of an acute in vitro infection with the tick-borne encephalitis virus adapted to *H. plumbeum* ticks. *Voprosy virusologii.* 36: 297-300.
- Dzhivaniyan T. I., Korolev M. B., Karganova G. G., Lisak V. M., Kashtanova G. M., Chuprinskaya M. V. (1988).** Changes in the host-dependent characteristics of the tick-borne encephalitis virus during its adaptation to ticks and its readaptation to white mice. *Voprosy virusologii.* 33: 589-595.
- Ecker M., Allison S. L., Meixner T., Heinz F. X. (1999).** Sequence analysis and genetic classification of tick-borne encephalitis viruses from Europe and Asia. *The Journal of General Virology.* 80: 179-185.
- Eigen M. (1971).** Self-organization of matter and the evolution of biological macromolecules. *Naturwissenschaften.* 58: 465-523.
- Eigen M., Schuster P. (1979).** The hypercycle. A principle of natural self-organization. Berlin: Springer.
- Enomoto N., Kurosaki M., Tanaka Y., Marumo F., Sato C. (1994).** Fluctuation of hepatitis C virus quasispecies in persistent infection and interferon treatment revealed by single-strand conformation polymorphism analysis. *Journal of General Virology.* 95: 1361-1369.

- Erbel P., Schiering N., D'Arcy A., Rénatus M., Kroemer M., Lim S. P., Yin Z., Keller T. H., Vasudevan S. G., Hommel U. (2006).** Structural basis for the activation of flaviviral NS3 proteases from dengue and West Nile virus. *Nature structural & molecular biology* 13: 372-373.
- Falgout B., Markoff L. (1995).** Evidence that flavivirus NS1-NS2A cleavage is mediated by a membrane-bound host protease in the endoplasmic reticulum. *Journal of Virology*. 69: 7232-7243.
- Falgout B., Pethel M., Zhang Y. M., Lai C. J. (1991).** Both nonstructural proteins NS2B and NS3 are required for the proteolytic processing of dengue virus nonstructural proteins. *Journal of Virology*. 65: 2467-2475.
- Gaunt M. W., Sall A. A., de Lamballerie X., Falconar A. K. I., Dzhivanian T. I., Gould E. A. (2001).** Phylogenetic relationships of flaviviruses correlate with their epidemiology, disease association and biogeography. *The Journal of General Virology*. 82: 1867-1876.
- Gorbalenya A. E., Koonin E. V., Donchenko A. P., Blinov V. M. (1989).** Two related superfamilies of putative helicases involved in replication, recombination, repair and expression of DNA and RNA genomes. *Nucleic Acids Research*. 17 (12): 4713-4730.
- Goto A., Hayasaka D., Yoshii K., Mizutani T., Kariwa H., Takashima I. (2003).** A BHK-21 cell culture-adapted tick-borne encephalitis virus mutant is attenuated for neuroinvasiveness. *Vaccine*. 21: 4043-4051.
- Goto A., Yoshii K., Obara M., Ueki T., Mizutani T., Kariwa H., Takashima I. (2005).** Role of the N-linked glycans of the prM and E envelope proteins in tick-borne encephalitis virus particle secretion. *Vaccine*. 23: 3043-3052.
- Gould E. A., de Lamballerie X., Zanotto P. M., Holmes E. C. (2003).** Origins, evolution, and vector/host coadaptations within the genus *Flavivirus*. *Advances in Virus Research*. 59: 277-314.
- Grard G., Moureau G., Charrel R. N., Lemasson J. J., Gonzalez J. P., Gallian P., Gritsun T. S., Holmes E. C., Gould E. A., de Lamballerie X. (2007).** Genetic characterization of tick-borne flaviviruses: new insights into evolution, pathogenetic determinants and taxonomy. *Virology*. 36: 80-92.
- Grešíková M., Nosek J. (1966).** Isolation of tick-borne encephalitis virus from *Haemaphysalis inermis* tick. *Acta Virologica*. 10: 359.
- Gritsun T. S., Desai A., Gould E. A. (2001).** The degree of attenuation of tick-borne encephalitis virus depends on the cumulative effects of point mutations. *The Journal of General Virology*. 82: 1667-1675.
- Gritsun T. S., Lashkevich V. A., Gould E. A. (2003).** Tick-borne encephalitis. *Antiviral Research*. 57: 129-147.

- Gritsun T. S., Venugopal K., Zanutto P. M. de A., Mikhailov M. V., Sall A. A., Holmes E. C., Polkinghorne I., Frolova T. V., Pogodina V. V., Lashkevich V. A., Gould E. A. (1997).** Complete sequence of two flaviviruses isolates from Siberia and the UK: analysis and significance of the 5' and 3'-UTRs. *Virus Research*. 49: 27-39.
- Grubhoffer L., Guirakhoo F., Heinz F. X., Kunz Ch. (1990).** Interaction of tick-borne encephalitis virus protein E with labelled lectins. In: Kocourek J., Feed D. L. J. (Eds.), *LECTINS, Biol., Biochem., Clin. Bichem.*, Vol. 7, Sigma Chem. Comp., St. Louis, pp. 313-319.
- Guirakhoo F., Heinz F. X., Mandl C. W., Holzman H., Kunz C. (1991).** Fusion activity of flaviviruses: comparison of mature and immature (prM-containing) tick-borne encephalitis virions. *The Journal of General Virology*. 57: 263-274.
- Haglund M., Günther G. (2003).** Tick-borne encephalitis-pathogenesis, clinical course and long-term follow-up. *Vaccine*. 21: S1/11-S1/18.
- Hanna S. L., Pierson T. C., Sanchez M. D., Ahmed A. A., Murtadha M. M., Doms R. W. (2005).** N-linked glycosylation of West Nile virus envelope proteins influences particle assembly and infectivity. *Journal of Virology*. 79: 13262-13274.
- Hayasaka D., Gritsun T. S., Kentarou Y., Ueki T., Goto A., Mizutani T., Kariwa H., Iwasaki T., Gould E. A., Takashima I. (2004).** Amino acid changes responsible for attenuation of virus neurovirulence in an infectious cDNA clone of the Oshima strain of tick-borne encephalitis virus. *The Journal of General Virology*. 85: 1007-1018.
- Heinz F. X. (1986).** Epitope mapping of flavivirus glycoproteins. *Advances in Virus Research*. 31: 103-168.
- Heinz F. X. (2003).** Molecular aspects of TBE virus research. *Vaccine*. 21: 3-10.
- Heinz F. X., Allison S. L. (2000).** Structures and mechanisms in flavivirus fusion. *Advances in Virus Research*. 55: 231-269.
- Heinz F. X., Allison S. L. (2001).** The machinery for flavivirus fusion with host cell membranes. *Current Opinion in Microbiology*. 4: 450-455.
- Heinz F. X., Allison S. L. (2003).** Flavivirus structure and membrane fusion. *Advances in Virus Research*. 59: 63-97.
- Heinz F. X., Berger R., Tuma W., Kunz C. (1983).** A topological and functional model of epitopes on the structural glycoprotein of tick-borne encephalitis virus defined by monoclonal antibodies. *Virology*. 126: 525-537.
- Heinz F. X., Collet M. S., Purcell R. H., Gould E. A., Howard C. R., Houghton M., Moormann R. J. M., Rice C. M., Thiel H. J. (2000).** Family Flaviviridae. In: van Regenmortel M. H. V., Fauquet C. M., Bishop D. H. L., Carstens E., Estes M. K., Lemon S., Maniloff J., Mayo M. A., McGeogch D., Pringle C. R., Wickner R. B. eds. *Virus Taxonomy*. In: *Proceedings of the 7th*

- International Committee for the Taxonomy of Viruses. Academic Press, San Diego, pp. 859-878.
- Holland J. J., Spindler K., Horodyski F., Grabau E., Nichol S., VandePol S. (1982).** Rapid evolution of RNA genomes. *Science*. 215: 1577-1585.
- Holzmann H., Heinz F. X., Mandl C., Guirakhoo F., Kunz C. (1990).** A single amino acid substitution in envelope protein of tick-borne encephalitis virus leads to attenuation in the mouse model. *Journal of Virology*. 64: 5156-5159.
- Holzmann H., Stiasny K., Ecker M., Kunz C., Heinz F. X. (1997).** Characterization of monoclonal antipody-escape mutants of tick-borne encephalitis virus with reduced neuroinvasiveness in mice. *The Journal of General Virology*. 78:31-37.
- Hurrelbrink R. J., McMinn P. (2003).** Molecular determinants of virulence: the structural and functional basis for flavivirus attenuation. *Advances in Virus Research*. 60: 1-42.
- Jerzak G., Bernard K. A., Kramer L. D., Ebel G. D. (2005).** Genetic variation in *West Nile Virus* from naturally infected mosquitoes and birds suggests quasispecies structure and strong purifying selection. *The Journal of General Virology*. 86: 2175-2183.
- Jones L. D., Davies C. R., Steele G. M., Nuttall P. A. (1987).** A novel mode of arbovirus transmission involving a nonviremic host. *Science*. 237: 775-777.
- Kaluzová M., Elečková E., Žuffová E., Pastorek J., Kaluz Š., Kouch O., Labuda M. (1994).** Reverted virulence of attenuated tick-borne encephalitis virus mutant is not accompanied with the changes in deduced viral envelope protein amino acid sequence. *Acta virologica*. 38: 133-140.
- Kaufmann B., Rossmann M. G. (2010).** Molecular mechanisms involved in the early step of flavivirus cell entry. *Microbes and Infection*. 1-9.
- Khromykh A. A., Meka H., Guyatt K. J., Westaway E. G. (2001).** Essential role of cyclization sequences in flavivirus RNA replication. *Journal of Virology*. 75: 6719-6728.
- Kiefer F., Arnold K., Künzli M., Bordoli L., Schwede T. (2009).** The SWISS-MODEL Repository and associated resources. *Nucleic Acids Research*. 37: D387-D392.
- Kofler R. M., Heinz F. X., Mandl C. W. (2002).** Capsid protein C of tick-borne encephalitis virus tolerates large internal deletion and is a favorable target for attenuation of virulence. *Journal of Virology*. 76: 3534-3543.
- Koonin E. V. (1993).** Computer-assisted identification of a putative methyltransferase domain in NS5 protein of flaviviruses and lambda 2 protein of reovirus. *The Journal of General Virology*. 74: 733-740.
- Koonin E. V., Dolja V. V. (1993).** Evolution and taxonomy of positive-strand RNA viruses: implication of comparative analysis of amino acid sequences. *Critical reviews in biochemistry and molecular biology*. 28: 375-430.

- Kopecký J., Křivanec K., Tomková E. (1991).** Attenuated temperature-sensitive mutants of tick-borne encephalitis (TBE) virus isolated from natural focus. In: Dusbábek F., Bukva V. (Eds.), *Modern Acarology*. Academia, Prague and SPB Academic Publishing bv, The Hague. Vol. 2: 11-19.
- Korenberg E. I., Kovalevskii Y. V. (1999).** Main features of tick-borne encephalitis eco-epidemiology in Russia. *Zentralblatt für Bakteriologie: international journal of medical microbiology*. 289: 525-539.
- Kozlovskaya L. I., Osolodkin D. I., Shevtsova A. S., Romanova L. I., Rogova Y. V., Dzhivaniyan T. I., Lyapustin V. N., Pivanova G. P., Gmyl A. P., Palyulin V. A., Karganova G. G. (2010).** GAG-binding variants of tick-borne encephalitis virus. *Virology*. 398: 262-272.
- Kroschewski H., Allison S. L., Heinz F. X., Mandl C. W. (2003).** Role of heparan sulfate for attachment and entry of tick-borne encephalitis virus. *Virology*. 308: 92-100.
- Křivanec K., Kopecký J., Tomková E., Grubhoffer L. (1988).** Isolation of TBE virus from the tick *Ixodes hexagonus*. *Folia Parasitologica*. 35: 273-276.
- Kümmerer B. M., Rice C. M. (2002).** Mutations in the yellow fever virus nonstructural protein NS2A selectively block production of infectious particles. *Journal of Virology*. 76: 4773-4784.
- Kuno G., Chang G. J., Tsuchiya K. R., Karabatsos N., Cropp C. B. (1998).** Phylogeny of the genus Flavivirus. *Journal of virology*. 72: 73-83.
- Kupča A. M., Essbauer S., Zoeller G., deMendonca P. G., Brey R., Rinder M., Pfister K., Spiegel M., Doerrbecker B., Pfeffer M., Dobler G. (2010).** Isolation and molecular characterization of a tick-borne encephalitis virus strain from a new tick-borne encephalitis focus with severe cases in Bavaria, Germany. *Ticks and tick-borne diseases*. 1: 44-51.
- Labuda M., Austyn J. M., Zuffova E., Kozuch O., Fuchsberger N., Lysy J., Nuttall P. A. (1996).** Importance of localized skin infection in tick-borne encephalitis virus transmission. *Virology*. 219: 357-366.
- Labuda M., Jiang W. R., Kaluzova M., Kozuch O., Nuttall P. A., Weismann P., Eleckova E., Zuffova E., Gould E. A. (1994).** Change in phenotype of tick-borne encephalitis virus following passage in *Ixodes ricinus* ticks and associated amino acid substitution in the envelope protein. *Virus Research*. 31: 305-315.
- Labuda M., Jones L. D., Williams T., Danielova V., Nuttall P. A. (1993).** Efficient transmission of tick-borne encephalitis virus between co-feeding ticks. *Journal of Medical Entomology*. 30: 295-299.
- Labuda M., Kozuch O., Zuffova E., Eleckova E., Hails R. S., Nuttall P. A. (1997).** Tick-borne encephalitis virus transmission between ticks co-feeding on specific immune natural rodent hosts. *Virology*. 235: 138-143.

- Leonova G. N. (1997).** Tick-borne Encephalitis in Primorskii Region: Virological, Ecological, and Epidemiological Aspects. Dal'nauka Press, Vladivostok.
- Leung J. Y., Pijlman G. P., Kondratieva N., Hyde J., Mackenzie J. M., Khromykh A. A. (2008).** Role of nonstructural protein NS2A in flavivirus assembly. *Journal of Virology*. 10: 4731-4741.
- Lin C., Amberg S. M., Chambers T. J., Rice C. M. (1993).** Cleavage at a novel site in the NS4A region by the yellow fever virus NS2B-3 proteinase is a prerequisite for processing at the downstream 4A/4B signalase site. *Journal of virology*. 67: 2327-2335.
- Lindenbach B. D., Rice C. M. (1999).** Genetic interaction of flavivirus nonstructural proteins NS1 and NS4A as a determinant of replicase function. *Journal of Virology*. 73: 4611-4621.
- Lindenbach B. D., Thiel H. J., Rice C. M. (2007).** *Flaviviridae: the viruses and their replication*. pp. 1101-1152. In: Knipe D. M., Howley P. M. (eds.), *Fields Virology*, 5th ed. Lippincott-Raven, Philadelphia.
- Mackenzie J. M., Khromykh A. A., Jones M. K., Westaway E. G. (1998).** Subcellular localization and some biochemical properties of the flavivirus Kunjin nonstructural proteins NS2A and NS4A. *Virology*. 245: 203-215.
- Maggi F., Fornai C., Vatteroni M. L., Giorgi M., Morrica A., Pistello M., Cammarota G., Marchi S., Ciccorossi P., Bionda A., Bendinelli M. (1997).** Differences in hepatitis C virus quasispecies composition between liver, peripheral blood mononuclear cells and plasma. *The Journal of General Virology*. 78: 1521-1525.
- Málková D., Fraňková V. (1959).** The lymphatic system in the development of experimental tick-borne encephalitis in mice. *Acta Virologica*. 3: 210-214.
- Mandl C. W., Allison S. L., Holzmann H., Meixner T., Heinz F. X. (2000).** Attenuation of tick-borne encephalitis virus by structure-based site-specific mutagenesis of a putative flavivirus receptor binding site. *Journal of Virology*. 20: 9601-9609.
- Mandl C. W., Heinz F. X., Kunz C. (1988).** Sequence of the structural proteins of tick-borne encephalitis virus (Western subtype) and comparative analysis with other flaviviruses. *Virology*. 166: 197-205.
- Mandl C. W., Holzmann H., Meixner T., Rauscher S., Stadler P. F., Allison S. L., Heinz F. X. (1998).** Spontaneous and engineered deletions in the 3' noncoding region of tick-borne encephalitis virus: construction of highly attenuated mutants of a flavivirus. *Journal of Virology*. 72: 2132-2140.
- Mandl C. W., Kroschewski H., Allison S. L., Kofler R., Holzmann H., Meixner T., Heinz F. X. (2001).** Adaptation of tick-borne encephalitis virus to BHK-21 cells results in the formation of multiple heparan sulfate binding sites in the envelope protein and attenuation in vivo. *Journal of Virology*. 75: 5627-5637.

- Mayer V. (1964).** Study of the virulence of tick-borne encephalitis virus. III. Biological evaluation of large-plaque and small-plaque variants of viruses of the tick-borne encephalitis complex. *Acta virologica*. 8: 507-520.
- Mayer V., Kožuch O. (1969).** Study of the virulence of tick-borne encephalitis virus. XI. Genetic heterogeneity of the virus from naturally infectious *Ixodes ricinus* ticks. *Acta virologica* 13: 469-482.
- McKnight A., Weiss R. A., Shotton C., Takenchi Y., Hoshino H., Chapham P. R. (1995).** Change in tropism upon immune escape by human immunodeficiency virus. *Journal of virology*. 69: 3167-3170.
- McMichael A. J., Phillips R. E. (1997).** Escape of human immunodeficiency virus from immune control. *Annual Review of Immunology*. 15: 271-296.
- McMinn P. C. (1997).** The molecular basis of virulence of the encephalitogenic flaviviruses. *The Journal of General Virology*. 78: 2711-2722.
- McMinn P. C., Dalgarno L., Weir R. C. (1996a).** A comparison of the spread of Murray Valley encephalitis viruses of high or low neuroinvasiveness in the tissues of Swiss mice after peripheral inoculation. *Virology*. 220: 414-423.
- McMinn P. C., Lee E., Hartley S., Roehrig J. T., Dalgarno L., Weir R. C. (1995).** Murray Valley encephalitis virus envelope protein antigenic variants with altered haemagglutination properties and reduced neuroinvasiveness in mice. *Virology*. 211:10-20.
- McMinn P. C., Weir R. C., Dalgarno L. (1996b).** A mouse-attenuated envelope protein variant of Murray Valley encephalitis virus with altered fusion activity. *The Journal of General Virology*. 77: 2085-2088.
- Miller S., Sparacio S., Bartenschlager R. (2006).** Subcellular localization and membrane topology of the dengue virus type 2 nonstructural protein 4B. *The Journal of biological chemistry*. 281: 8854-8863.
- Mitzel D. N., Best S. M., Masnick M. F., Porcella S. F., Wolfenbarger J. B., Bloom M. E. (2008).** Identification of genetic determinants of a tick-borne flavivirus associated with host-specific adaptation and pathogenicity. *Virology*. 381: 268-276.
- Moudy R. M., Zhang B., Shi P.-Y., Kramer L. D. (2009).** West Nile virus envelope protein glycosylation is required for efficient viral transmission by *Culex* vector. *Virology*. 387: 222-228.
- Mukherjee M., Dutta K., White M. A., Cowburn D., Fox R. O. (2006).** NMR solution structure and backbone dynamics of domain III of the E protein of tick-borne Langat flavivirus suggest a potential site for molecular recognition. *Protein Science*. 15: 1342-1355.
- Murphy F. (1980).** Togavirus morphology and morphogenesis. In: Schlesinger, R. W. (Eds.), *The Togaviruses. Biology, Structure, Replication*. Academic Press, New York, p. 241-316.

- Murphy P. M. (1993).** Molecular mimicry and the generation of host defense protein diversity. *Cell*. 72: 823-826.
- Muylaert I. R., Galler R. G., Rice C. M. (1996).** Mutagenesis of the N-linked glykosylation sites of the yellow fever virus NS1 protein: effects on virus replication and mouse neurovirulence. *Virology*. 222: 159-168.
- Nowak T., Wengler G. (1987).** Analysis of disulphides present in the membrane proteins of West Nile flavivirus. *Virology*. 156: 127-137.
- Nuttall P. A., Jones L. D., Davies C. R. (1991).** The role of arthropod vectors in arbovirus evolution. *Advances in Disease Vector Research*. 15-45.
- Nuttall P. A., Jones L. D., Labuda M., Kaufman W. R. (1994).** Adaptations of arboviruses to tick. *Journal of Medical Entomology*. 31: 1-9.
- Nuttall P. A., Labuda M. (1994).** Tick-borne Encephalitis Subgroup. In: Sonenshine D. E., Mather T. M. (Eds.), *Ecological Dynamics of Tick-borne Zoonoses*. Oxford Press, New York, p. 351-391.
- Olden K., Pratt R. M., Jaworski C., Yamada K. M. (1979).** Evidence for role of glycoprotein carbohydrates in membrane transport: Specific inhibition by tunicamycin. *Cell Biology*. 2: 791-795.
- Peitsch M. C. (1995).** Protein modeling by E-mail. *Nature Biotechnology*. 13: 658-660.
- Pospíšil L., Jandásek L., Pešek J. (1954).** Izolace nových kmenů meningoencefalitidy v brněnském kraji za letní období 1953. *Lékařské listy*. 9: 5-3.
- Proutski V., Gould E. A., Holmes E. C. (1997).** Secondary structure of the 3' untranslated region of flaviviruses: similarities and differences. *Nucleic Acids Research*. 25: 1194-1202.
- Proutski V., Gritsun T. S., Gould E. A., Holmes E. C. (1999).** Biological consequences of deletions within the 3'-untranslated region of flavivirus may be due to rearrangements of RNA secondary structure. *Virus Research*. 64: 107-123.
- Rey F. A., Heinz F. X., Mandl C., Kunz C., Harrison S. C. (1995).** The envelope glycoprotein from tick-borne encephalitis virus at 2 Å resolution. *Nature*. 375: 291-298.
- Rice C. M. (1996).** *Flaviviridae: the viruses and their replication*. In *Fields Virology*, 3rd edn, pp. 931-959. Edited by Fields, B. N., Knipe, D. M., Howley, P. M. Philadelphia: Lippincott-Raven.
- Roehrig J. T., Mathews J. H., Trent D. W. (1983).** Identification of epitopes on the E glycoprotein of Saint Louis encephalitis virus using monoclonal antibodies. *Virology*. 128: 118-126.
- Romanova L. I., Gmyl A. P., Dzhivanian T. I., Bakhmutov D. V., Lukashev A. N., Gmyl L. V., Rumyantsev A. A., Burenkova L. A., Lashkevich V. A., Karganova G. G. (2007).** Microevolution of tick-borne encephalitis virus in course of host alteration. *Virology*. 326: 75-84.

- Růžek D., Gritsun T. S., Forrester N. L., Gould E. A., Kopecký J., Golovchenko M., Rudenko N., Grubhoffer L. (2008).** Mutations in the NS2B and NS3 genes affect mouse neuroinvasiveness of a Western European field strain of tick-borne encephalitis virus. *Virology*. 374: 249-255.
- Růžek D., Salát J., Palus M., Gritsun T. S., Gould E. A., Dyková I., Skallová A., Jelínek J., Kopecký J., Grubhoffer L. (2009).** CD8⁺ T-cells mediate immunopathology in tick-borne encephalitis. *Virology*. 384: 1-6.
- Sa-Carvalho D., Rieder E., Baxt B., Rodarte R., Tanuri A., Mason P. (1997).** Tissue culture adaptation of foot-and-mouth disease virus selects viruses that bind to heparin and are attenuated in cattle. *Journal of virology*. 71: 5115-5123.
- Samuel M. A., Diamond M. S. (2006).** Pathogenesis of West Nile Virus infection: a balance between virulence, innate and adaptive immunity, and viral evasion. *Journal of Virology*. 19: 9349-9360.
- Scherer W. F., Eddy G. A., Monath T. P., Walton T. E. (1980).** Laboratory safety for arboviruses and certain other viruses of vertebrates. *The American journal of tropical medicine and hygiene*. 29: 1359-1381.
- Schneider H. (1931).** Über epidemische akute „Meningitis serosa“. *Wiener klinische Wochenschrift*. 44: 350-352.
- Smith D. B., McAllister J., Casino C., Simmonds P. (1997).** Virus ‚quasispecies‘: making a mountain out of a molehill? *The Journal of General Virology*. 78: 1511-1519.
- Stadler K., Allison S. L., Schalich J., Heinz F. X. (1997).** Proteolytic activation of tick-borne encephalitis virus by furin. *Journal of Virology*. 71: 8475-8481.
- Stiasny K., Allison S. L., Marchler-Bauer A., Kunz C., Heinz F. X. (1996).** Structural requirements for low-pH-induced rearrangements in the envelope glycoprotein of tick-borne encephalitis virus. *Journal of Virology*. 70: 8142-8147.
- Süss J. (2003).** Epidemiology and ecology of TBE relevant to the production of effective vaccines. *Vaccine*. 21: 19-35.
- Süss J., Schrader C., Abel U., Bormane A., Duks A., Kalnina V. (2002).** Characterization of tick-borne encephalitis (TBE) foci in Germany and Latvia (1997-2000). *International Journal of Medical Microbiology*. 291 (suppl 33): 34-42.
- Süss J., Schrader C., Abel U., Voigt W. P., Schosser R. (1999).** Annual and seasonal variation of tick-borne encephalitis virus (TBEV) prevalence in ticks in selected hot spot areas in Germany using an RT-PCR: results from 1997 and 1998. *Zentralblatt für Bakteriologie : international journal of medical microbiology*. 289: 564-578.
- Tamura K., Dudley J., Nei M., Kumar S. (2007).** MEGA 4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Molecular Biology and Evolution* 24: 1596-1599.

- Tassaneeritthep B., Burgess T. H., Granelli-Piperno A., Trumpfheller C., Finke J., Sun W., Eller M. A., Pattanapanyasat K., Sarasombath S., Birx D. L., Steinman R. M., Schlesinger S., Marovich M. A. (2003).** DC-SIGN (CD209) mediates dengue virus infection of human dendritic cells. *The Journal of Experimental Medicine*. 197: 823-829.
- Tumova S., Woods A., Couchman J. R. (2000).** Heparan sulfate proteoglycans on the cell surface: versatile coordinators of cellular functions. *The international journal of biochemistry & cell biology*. 32: 269-288.
- Varelas-Wesley I., Calisher C. H. (1982).** Antigenic Relationship of Flaviviruses with Undetermined Arthropod-Borne Status. *The American Journal of Tropical Medicine and Hygiene*. 31: 1273-1284.
- Vigerust D. J., Shepherd V. L. (2007).** Virus glycosylation: role in virulence and immune interactions. *Trends in Microbiology*. 15: 211-218.
- Wallis T. P., Huang C.-Y., Nimkar S. B., Young P. R., Gorman J. J. (2004).** Determination of the disulfide bond arrangement of dengue virus NS1 protein. *The Journal of Biological Chemistry*. 279: 20729-20741.
- Wallner G., Mandl C., Ecker M., Holzmann H., Stiasny K., Kunz C., Heinz F. X. (1996).** Characterization and complete genome sequences of high- and low-virulence variants of tick-borne encephalitis viruses. *The Journal of General Virology*. 77: 1035-1042.
- Wallner G., Mandl C., Kunz C., Heinz F. X. (1995).** The flavivirus 3'-noncoding region: extensive size heterogeneity independent of evolutionary relationships among strains of tick-borne encephalitis virus. *Virology*. 213: 169-178.
- Wengler G., Wengler G. (1993).** The NS3 nonstructural protein of flaviviruses contains an RNA triphosphatase activity. *Virology*. 197: 265-273.
- Zanotto P. M., Gao G. F., Gritsun T, Marin M. S., Jiang W. R., Venugopal K., Reid H. W., Gould E. A. (1995).** An arbovirus cline across the northern hemisphere. *Virology*. 210: 152-159.
- Yap T. L., Xu T., Chen Y.-L., Malet H., Egloff M.-P., Canard B., Vasudevan S. G., Lescar J. (2007).** Crystal structure of the dengue virus RNA-dependent RNA polymerase catalytic domain at 1.85-Angstrom resolution. *Journal of Virology*. 9: 4753-4765.
- Zilber L. A. (1939).** Vesennij (vesenne-letnij) epidemičeskij kleščevoj encefalit. *Arkhiv Biologičeskikh Nauk*. 56: 2.
- Zlobin V. I., Gorin O. Z. (1996).** Tick-borne encephalitis: etiology, epidemiology and prophylactics in Siberia. In: Philippova, L. B., ed. *Kleshchevoi entsefalit: Etiologija. Epidemiologija i profilaktika v sibirii*. Novosibirsk: Nauka, 177.

Internetové zdroje: <http://www.ncbi.nlm.nih.gov> <http://swissmodel.expasy.org>
 <http://geno3d-pbil.ibcp.fr> <http://www.yasara.org>

9. Příloha

Příloha 9.1 Kompletní genomové sekvence všech čtyř variant VKE studovaných v této práci; nukleotidové záměny jsou vyznačeny barevně a podtržením (zeleně záměny v nekódujících oblastech; modře záměny v nukleotidové sekvenci, červeně záměny v nukleotidové sekvenci vedoucí ke změně AK).

5' koncová oblast:

	10	20	30	40	50	60	70	
0.P.	AGATTTTCTT	GCACGTGCGT	GCGTTTGCTT	CGGACAGCAT	TAGCAGCGGT	TGGTTTGAAA	GAAATATTCT	
DL.P.IRE	AGATTTTCTT	GCACGTGCGT	GCGTTTGCTT	CGGACAGCAT	TAGCAGCGGT	TGGTTTGAAA	GAAATATTCT	
40.P.PS	AGATTTTCTT	GCACGTGCGT	GCGTTTGCTT	CGGACAGCAT	TAGCAGCGGT	TGGTTTGAAA	GAAATATTCT	
40.P.IRE	AGATTTTCTT	GCACGTGCGT	GCGTTTGCTT	CGGACAGCAT	TAGCAGCGGT	<u>TA</u> GTTTGAAA	GAAATATTCT	
	80	90	100	110	120	130	START kodon	
0.P.	TTTGTITTTA	CCAGTCGTGA	ACGTGTTGAG	AAAAAGACAG	CTTAGGAGAA	CAAGAGCTGG	GG	ATG
DL.P.IRE	TTTGTITTTA	CCAGTCGTGA	ACGTGTTGAG	AAAAAGACAG	CTTAGGAGAA	CAAGAGCTGG	GG	ATG
40.P.PS	TTTGTITTTA	CCAGTCGTGA	ACGTGTTGAG	AAAAAGACAG	CTTAGGAGAA	CAAGAGCTGG	GG	ATG
40.P.IRE	TTTGTITTTA	CCAGTCGTGA	ACGTGTTGAG	AAAAAGACAG	CTTAGGAGAA	CAAGAGCTGG	GG	ATG

Protein C:

	145	155	165	175	185	195	205	
0.P.	GTCAAGAAGG	CCATCCTGAA	AGGTAAGGGG	GGCGGTCCCC	CTCGACGAGT	GTCGAAAGAG	ACCGCAACGA	
DL.P.IRE	GTCAAGAAGG	CCATCCTGAA	AGGTAAGGGG	GGCGGTCCCC	CTCGACGAGT	GTCGAAAGAG	ACCGCAACGA	
40.P.PS	GTCAAGAAGG	CCATCCTGAA	AGGTAAGGGG	GGCGGTCCCC	CTCGACGAGT	GTCGAAAGAG	ACCGCAACGA	
40.P.IRE	GTCAAGAAGG	CCATCCTGAA	AGGTAAGGGG	GGCGGTCCCC	CTCGACGAGT	GTCGAAAGAG	ACCGCAACGA	
	215	225	235	245	255	265	275	
0.P.	AGACGCGTCA	ACCCAGAGTC	CAAATGCCAA	ATGGGCTCGT	GTTGATGCGC	ATGATGGGGA	TCTTGTGGCA	
DL.P.IRE	AGACGCGTCA	ACCCAGAGTC	CAAATGCCAA	ATGGGCTCGT	GTTGATGCGC	ATGATGGGGA	TCTTGTGGCA	
40.P.PS	AGACGCGTCA	ACCCAGAGTC	CAAATGCCAA	ATGGGCTCGT	GTTGATGCGC	ATGATGGGGA	TCTTGTGGCA	
40.P.IRE	AGACGCGTCA	ACCCAGAGTC	CAAATGCCAA	ATGGGCTCGT	GTTGATGCGC	ATGATGGGGA	TCTTGTGGCA	
	285	295	305	315	325	335	345	
0.P.	TGCCGTAGCT	GGCACCGCGA	GAAACCCCGT	ATTGAAGGCG	TTCTGGAACT	CAGTCCCTCT	GAAACAGGCC	
DL.P.IRE	TGCCGTAGCT	GGCACCGCGA	GAAACCCCGT	ATTGAAGGCG	TTCTGGAACT	CAGTCCCTCT	GAAACAGGCC	
40.P.PS	TGCCGTAGCT	GGCACCGCGA	GAAACCCCGT	ATTGAAGGCG	TTCTGGAACT	CAGTCCCTCT	GAAACAGGCC	
40.P.IRE	TGCCGTAGCT	GGCACCGCGA	GAAACCCCGT	ATTGAAGGCG	TTCTGGAACT	CAGTCCCTCT	GAAACAGGCC	
	355	365	375	385	395	405	415	
0.P.	ACAGCAGCAC	TGCGGAAGAT	CAAAAGGACA	GTGAGTGCCC	TAATGGTTGG	CTTGCAAAAA	CGCGGGAAAA	
DL.P.IRE	ACAGCAGCAC	TGCGGAAGAT	CAAAAGGACA	GTGAGTGCCC	TAATGGTTGG	CTTGCAAAAA	CGCGGGAAAA	
40.P.PS	ACAGCAGCAC	TGCGGAAGAT	CAAAAGGACA	GTGAGTGCCC	TAATGGTTGG	CTTGCAAAAA	CGCGGGAAAA	
40.P.IRE	ACAGCAGCAC	TGCGGAAGAT	CAAAAGGACA	GTGAGTGCCC	TAATGGTTGG	CTTGCAAAAA	CGCGGGAAAA	
	425	435	445	455	465			
0.P.	GGAGGTCAGC	GACGGACTGG	ATGAGCTGGT	TGCTAGTCAT	TACTCTGTTG	GGG		
DL.P.IRE	GGAGGTCAGC	GACGGACTGG	ATGAGCTGGT	TGCTAGTCAT	TACTCTGTTG	GGG		
40.P.PS	GGAGGTCAGC	GACGGACTGG	ATGAGCTGGT	TGCTAGTCAT	TACTCTGTTG	GGG		
40.P.IRE	GGAGGTCAGC	GACGGACTGG	ATGAGCTGGT	TGCTAGTCAT	TACTCTGTTG	GGG		

Protein prM:

	478	488	498	508	518	528	538
0.P.	ATGACGATTG	CTGCAACCGT	GAGGAAAAGAA	AGGGACGGCT	CAACTGTGAT	CAGAGCTGAA	GGAAAAGGACG
DL.P.IRE	ATGACGATTG	CTGCAACCGT	GAGGAAAAGAA	AGGGACGGCT	CAACTGTGAT	CAGAGCTGAA	GGAAAAGGACG
40.P.PS	ATGACGATTG	CTGCAACCGT	GAGGAAAAGAA	AGGGACGGCT	CAACTGTGAT	CAGAGCTGAA	GGAAAAGGACG
40.P.IRE	ATGACGATTG	CTGCAACCGT	GAGGAAAAGAA	AGGGACGGCT	CAACTGTGAT	CAGAGCTGAA	GGAAAAGGACG

	548	558	568	578	588	598	608
0.P.	CAGCAACTCA	GGTGCCTGTG	GAGAATGGCA	CCTGTGTGAT	CCTGGCTACT	GACATGGGGT	CATGGTGTGA
DL.P.IRE	CAGCAACTCA	GGTGCCTGTG	GAGAATGGCA	CCTGTGTGAT	CCTGGCTACT	GACATGGGGT	CATGGTGTGA
40.P.PS	CAGCAACTCA	GGTGCCTGTG	GAGAATGGCA	CCTGTGTGAT	CCTGGCTACT	GACATGGGGT	CATGGTGTGA
40.P.IRE	CAGCAACTCA	GGTGCCTGTG	GAGAATGGCA	I CCTGTGTGAT	CCTGGCTACT	GACATGGGGT	CATGGTGTGA

	618	628	638	648	658	668	678
0.P.	TGATTCACTG	TCCTATGAGT	GTGTGACCAT	AGATCAAGGA	GAAGAGCCTG	TTGACGTGGA	TTGTTTTTGC
DL.P.IRE	TGATTCACTG	TCCTATGAGT	GTGTGACCAT	AGATCAAGGA	GAAGAGCCTG	TTGACGTGGA	TTGTTTTTGC
40.P.PS	TGATTCACTG	TCCTATGAGT	GTGTGACCAT	AGATCAAGGA	GAAGAGCCTG	TTGACGTGGA	TTGTTTTTGC
40.P.IRE	TGATTCACTG	TCCTATGAGT	GTGTGACCAT	AGATCAAGGA	GAAGAGCCTG	TTGACGTGGA	TTGTTTTTGC

	688	698	708	718	728	738	
0.P.	CGGAATGTTG	ATGGAGTCTA	TCTGGAGTAT	GGACGCTGTG	GGAAACAGGA	AGGCTCACGG	ACAAGGCCG
DL.P.IRE	CGGAATGTTG	ATGGAGTCTA	TCTGGAGTAT	GGACGCTGTG	GGAAACAGGA	AGGCTCACGG	ACAAGGCCG
40.P.PS	CGGAATGTTG	ATGGAGTCTA	TCTGGAGTAT	GGACGCTGTG	GGAAACAGGA	AGGCTCACGG	ACAAGGCCG
40.P.IRE	CGGAATGTTG	ATGGAGTCTA	TCTGGAGTAT	GGACGCTGTG	GGAAACAGGA	AGGCTCACGG	ACAAGGCCG

Protein M:

	757	767	777	787	797	807	817
0.P.	TCAGTGCTGA	TCCCATCCCA	TGCCCAGGGA	GAGCTGACGG	GGAGGGGACA	CAAATGGCTA	GAAGGAGACT
DL.P.IRE	TCAGTGCTGA	TCCCATCCCA	TGCCCAGGGA	GAGCTGACGG	GGAGGGGACA	CAAATGGCTA	GAAGGAGACT
40.P.PS	TCAGTGCTGA	TCCCATCCCA	TGCCCAGGGA	GAGCTGACGG	GGAGGGGACA	CAAATGGCTA	GAAGGAGACT
40.P.IRE	TCAGTGCTGA	TCCCATCCCA	TGCCCAGGGA	GAGCTGACGG	GGAGGGGACA	CAAATGGCTA	GAAGGAGACT

	827	837	847	857	867	877	887
0.P.	CGCTGCGAAC	ACATCTCACC	AGAGTTGAGG	GATGGGTTTG	GAAGAACAGG	CTACTTGCCC	TGGCGATGGT
DL.P.IRE	CGCTGCGAAC	ACATCTCACC	AGAGTTGAGG	GATGGGTTTG	GAAGAACAGG	CTACTTGCCC	TGGCGATGGT
40.P.PS	CGCTGCGAAC	ACATCTCACC	AGAGTTGAGG	GATGGGTTTG	GAAGAACAGG	CTACTTGCCC	TGGCGATGGT
40.P.IRE	CGCTGCGAAC	ACATCTCACC	AGAGTTGAGG	GATGGGTTTG	GAAGAACAGG	CTACTTGCCC	TGGCGATGGT

	897	907	917	927	937	947	957
0.P.	CACCGTTGTG	TGGTTGACCC	TGGAGAGTGT	GGTGACCAGG	GTCGCCGTTT	TGGTTGTGCT	CCTGTGTTTG
DL.P.IRE	CACCGTTGTG	TGGTTGACCC	TGGAGAGTGT	GGTGACCAGG	GTCGCCGTTT	TGGTTGTGCT	CCTGTGTTTG
40.P.PS	CACCGTTGTG	TGGTTGACCC	TGGAGAGTGT	GGTGACCAGG	GTCGCCGTTT	TGGTTGTGCT	CCTGTGTTTG
40.P.IRE	CACCGTTGTG	TGGTTGACCC	TGGAGAGTGT	GGTGACCAGG	GTCGCCGTTT	TGGTTGTGCT	CCTGTGTTTG

	967	
0.P.	GCGCCGGTCT	ACGCT
DL.P.IRE	GCGCCGGTCT	ACGCT
40.P.PS	GCGCCGGTCT	ACGCT
40.P.IRE	GCGCCGGTCT	ACGCT

Protein E:

	982	992	1002	1012	1022	1032	1042
0.P.	TCGCGTTGCA	CACATTTGGA	AAACAGGGAC	TTTGTGACTG	GTACTCAGGG	GACTACGAGG	GTCACCTTGG
DL.P.IRE	TCGCGTTGCA	CACATTTGGA	AAACAGGGAC	TTTGTGACTG	GTACTCAGGG	GACTACGAGG	GTCACCTTGG
40.P.PS	TCGCGTTGCA	CACATTTGGA	AAACAGGGAC	TTTGTGACTG	GTACTCAGGG	GACTACGAGG	GTCACCTTGG
40.P.IRE	TCGCGTTGCA	CACATTTGGA	AAACAGGGAC	TTTGTGACTG	GTACTCAGGG	GACTACGAGG	GTCACCTTGG
	1052	1062	1072	1082	1092	1102	1112
0.P.	TGCTGGAAct	GGGTGGATGT	GTTACCATAA	CAGCTGAGGG	GAAGCCTTCA	ATGGATGTGT	GGCTTGACGC
DL.P.IRE	TGCTGGAAct	GGGTGGATGT	GTTACCATAA	CAGCTGAGGG	GAAGCCTTCA	ATGGATGTGT	GGCTTGACGC
40.P.PS	TGCTGGAAct	GGGTGGATGT	GTTACCATAA	CAGCTGAGGG	GAAGCCTTCA	ATGGATGTGT	GGCTTGACGC
40.P.IRE	TGCTGGAAct	GGGTGGATGT	GTTACCATAA	CAGCTGAGGG	GAAGCCTTCA	ATGGATGTGT	GGCTTGACGC
	1122	1132	1142	1152	1162	1172	1182
0.P.	CATTACCAG	GAGAACCCTG	CTAAGACACG	TGAGTACTGT	TTACACGCCA	AGTTGTCGGA	CACTAAGGTT
DL.P.IRE	CATTACCAG	GAGAACCCTG	CTAAGACACG	TGAGTACTGT	TTACACGCCA	AGTTGTCGGA	CACTAAGGTT
40.P.PS	CATTACCAG	GAGAACCCTG	CTAAGACACG	TGAGTACTGT	TTACACGCCA	AGTTGTCGGA	CACTAAGGTT
40.P.IRE	CATTACCAG	GAGAACCCTG	CTAAGACACG	TGAGTACTGT	TTACACGCCA	AGTTGTCGGA	CACTAAGGTT
	1192	1202	1212	1222	1232	1242	1252
0.P.	GCAGCTAGAT	GCCCAACAAT	GGGACCAGCC	ACTTTGGCTG	AAGAACACCA	GGGTGGCACA	GTGTGCAAGA
DL.P.IRE	GCAGCTAGAT	GCCCAACAAT	GGGACCAGCC	ACTTTGGCTG	AAGAACACCA	GGGTGGCACA	GTGTGCAAGA
40.P.PS	GCAGCTAGAT	GCCCAACAAT	GGGACCAGCC	ACTTTGGCTG	AAGAACACCA	GGGTGGCACA	GTGTGCAAGA
40.P.IRE	GCAGCTAGAT	GCCCAACAAT	GGGACCAGCC	ACTTTGGCTG	AAGAACACCA	GGGTGGCACA	GTGTGCAAGA
	1262	1272	1282	1292	1302	1312	1322
0.P.	GAGATCAGAG	TGATCGAGGC	TGGGGCAACC	ACTGTGGACT	TTTTGGAAAG	GGTAGCATTG	TGGCCTGTGT
DL.P.IRE	GAGATCAGAG	TGATCGAGGC	TGGGGCAACC	ACTGTGGACT	TTTTGGAAAG	GGTAGCATTG	TGGCCTGTGT
40.P.PS	GAGATCAGAG	TGATCGAGGC	TGGGGCAACC	ACTGTGGACT	TTTTGGAAAG	GGTAGCATTG	TGGCCTGTGT
40.P.IRE	GAGATCAGAG	TGATCGAGGC	TGGGGCAACC	ACTGTGGACT	TTTTGGAAAG	GGTAGCATTG	TGGCCTGTGT
	1332	1342	1352	1362	1372	1382	1392
0.P.	CAAGGCGGCT	TGTGAGGCAA	AAAAGAAAGC	CACAGGACAT	GTGTACGACG	CCAACAAAAT	AGTGTACACG
DL.P.IRE	CAAGGCGGCT	TGTGAGGCAA	AAAAGAAAGC	CACAGGACAT	GTGTACGACG	CCAACAAAAT	AGTGTACACG
40.P.PS	CAAGGCGGCT	TGTGAGGCAA	AAAAGAAAGC	CACAGGACAT	GTGTACGACG	CCAACAAAAT	AGTGTACACG
40.P.IRE	CAAGGCGGCT	TGTGAGGCAA	AAAAGAAAGC	CACAGGACAT	GTGTACGACG	CCAACAAAAT	AGTGTACACG
	1402	1412	1422	1432	1442	1452	1462
0.P.	GTCAAAGTCG	AACCACACAC	GGGAGACTAT	GTTGCCGCAA	ACGAGACACA	TAGTGGGAGG	AAGACGGCAT
DL.P.IRE	GTCAAAGTCG	AACCACACAC	GGGAGACTAT	GTTGCCGCAA	ACGAGACACA	TAGTGGGAGG	AAGACGGCAT
40.P.PS	GTCAAAGTCG	AACCACACAC	GGGAGACTAT	GTTGCCGCAA	ACGAGACACA	TAGTGGGAGG	AAGACGGCAT
40.P.IRE	GTCAAAGTCG	AACCACACAC	GGGAGACTAT	GTTGCCGCAA	ACGAGACACA	TAGTGGGAGG	AAGACGGCAT
	1472	1482	1492	1502	1512	1522	1532
0.P.	CCTTCACAGT	TTCTTCAGAG	AAAACCATTC	TGACTATGGG	TGAGTATGGA	GATGTGTCTC	TGTTGTGTAG
DL.P.IRE	CCTTCACAGT	TTCTTCAGAG	AAAACCATTC	TGACTATGGG	TGAGTATGGA	GATGTGTCTC	TGTTGTGTAG
40.P.PS	CCTTCACAGT	TTCTTCAGAG	AAAACCATTC	TGACTATGGG	TGAGTATGGA	GATGTGTCTC	TGTTGTGTAG
40.P.IRE	CCTTCACAGT	TTCTTCAGAG	AAAACCATTC	TGACTATGGG	TGAGTATGGA	GATGTGTCTC	TGTTGTGTAG

		1542	1552	1562	1572	1582	1592	1602
0.P.		GGTCGCTAGT	GGCGTTGACT	TGGCCCAGAC	CGTCATCCTT	GAGCTTGACA	AGACAGTGGGA	ACACCTTCCA
DL.P.IRE		GGTCGCTAGT	GGCGTTGACT	TGGCCCAGAC	CGTCATCCTT	GAGCTTGACA	AGACAGTGGGA	ACACCTTCCA
40.P.PS		GGTCGCTAGT	GGCGTTGACT	TGGCCCAGAC	CGTCATCCTT	GAGCTTGACA	AGACAGTGGGA	ACACCTTCCA
40.P.IRE		GGTCGCTAGT	GGCGTTGACT	TGGCCCAGAC	CGTCATCCTT	GAGCTTGACA	AGACAGTGGGA	ACACCTTCCA
		1612	1622	1632	1642	1652	1662	1672
0.P.		ACGGCTTGGC	AGGTCCACAG	GGACTGGTTT	AATGATCTGG	CTCTGCCATG	GAAACATGAG	GGAGCGCAAA
DL.P.IRE		ACGGCTTGGC	AGGTCCACAG	GGACTGGTTT	AATGATCTGG	CTCTGCCATG	GAAACATGAG	GGAGCGCAAA
40.P.PS		ACGGCTTGGC	AGGTCCACAG	GGACTGGTTT	AATGATCTGG	CTCTGCCATG	GAAACATGAG	GGAGCGCAAA
40.P.IRE		ACGGCTTGGC	AGGTCCACAG	GGACTGGTTT	AATGATCTGG	CTCTGCCATG	GAAACATGAG	GGAGCGCAAA
		1682	1692	1702	1712	1722	1732	1742
0.P.		ACTGGAATAA	CGCAGAAAGA	TTGGTTGAAT	TTGGGGCTCC	TCATGCTGTC	AAGATGGATG	TGTACAACCT
DL.P.IRE		ACTGGAATAA	CGCAGAAAGA	TTGGTTGAAT	TTGGGGCTCC	TCATGCTGTC	AAGATGGATG	TGTACAACCT
40.P.PS		ACTGGAATAA	CGCAGAAAGA	TTGGTTGAAT	TTGGGGCTCC	TCATGCTGTC	AAGATGGATG	TGTACAACCT
40.P.IRE		ACTGGAATAA	CGCAGAAAGA	TTGGTTGAAT	TTGGGGCTCC	TCATGCTGTC	AAGATGGATG	TGTACAACCT
		1752	1762	1772	1782	1792	1802	1812
0.P.		CGGAGACCAG	ACTGGAGTGT	TACTGAAGGC	TCTCGCTGGG	GTTCTGTGG	CACACATTGA	GGGAACCAAG
DL.P.IRE		CGGAGACCAG	ACTGGAGTGT	TACTGAAGGC	TCTCGCTGGG	GTTCTGTGG	CACACATTGA	GGGAACCAAG
40.P.PS		CGGAGACCAG	ACTGGAGTGT	TACTGAAGGC	TCTCGCTGGG	GTTCTGTGG	CACACATTGA	GGGAACCAAG
40.P.IRE		CGGAGACCAG	ACTGGAGTGT	TACTGAAGGC	TCTCGCTGGG	GTTCTGTGG	CACACATTGA	GGGAACCAAG
		1822	1832	1842	1852	1862	1872	1882
0.P.		TACCACCTGA	AGAGTGGCCA	TGTGACCTGC	GAAGTGGGAC	TGGAAAAACT	GAAGATGAAA	GGTCTTACGT
DL.P.IRE		TACCACCTGA	AGAGTGGCCA	TGTGACCTGC	GAAGTGGGAC	TGGAAAAACT	GAAGATGAAA	GGTCTTACGT
40.P.PS		TACCACCTGA	AGAGTGGCCA	TGTGACCTGC	GAAGTGGGAC	TGGAAAAACT	GAAGATGAAA	GGTCTTACGT
40.P.IRE		TACCACCTGA	AGAGTGGCCA	TGTGACCTGC	GAAGTGGGAC	TGGAAAAACT	GAAGATGAAA	GGTCTTACGT
		1892	1902	1912	1922	1932	1942	1952
0.P.		ACACAATGTG	TGACAAAACA	AAGTTCACAT	GGAAGAGAGC	TCCAACAGAC	AGTGGGCATG	ATACAGTGGT
DL.P.IRE		ACACAATGTG	TGACAAAACA	AAGTTCACAT	GGAAGAGAGC	TCCAACAGAC	AGTGGGCATG	ATACAGTGGT
40.P.PS		ACACAATGTG	TGACAAAACA	AAGTTCACAT	GGAAGAGAGC	TCCAACAGAC	AGTGGGCATG	ATACAGTGGT
40.P.IRE		ACACAATGTG	TGACAAAACA	AAGTTCACAT	GGAAGAGAGC	TCCAACAGAC	AGTGGGCATG	ATACAGTGGT
		1962	1972	1982	1992	2002	2012	2022
0.P.		CATGGAAGTC	ACATTCTCTG	GAACAAAGCC	CTGTAGGATC	CCAGTCAGGG	CAGTGGCACA	TGGATCTCCA
DL.P.IRE		CATGGAAGTC	ACATTCTCTG	GAACAAAGCC	CTGTAGGATC	CCAGTCAGGG	CAGTGGCACA	TGGATCTCCA
40.P.PS		CATGGAAGTC	ACATTCTCTG	GAACAAAGCC	CTGTAGGATC	CCAGTCAGGG	CAGTGGCACA	TGGATCTCCA
40.P.IRE		CATGGAAGTC	ACATTCTCTG	GAACAAAGCC	CTGTAGGATC	CCAGTCAGGG	CAGTGGCACA	TGGATCTCCA
		2032	2042	2052	2062	2072	2082	2092
0.P.		GATGTGAACG	TGGCCATGCT	GATAACGCCA	AACCCAACAA	TTGAAAACAA	TGGAGGTGGC	TTCATAGAGA
DL.P.IRE		GATGTGAACG	TGGCCATGCT	GATAACGCCA	AACCCAACAA	TTGAAAACAA	TGGAGGTGGC	TTCATAGAGA
40.P.PS		GATGTGAACG	TGGCCATGCT	GATAACGCCA	AACCCAACAA	TTGAAAACAA	TGGAGGTGGC	TTCATAGAGA
40.P.IRE		GATGTGAACG	TGGCCATGCT	GATAACGCCA	AACCCAACAA	TTGAAAACAA	TGGAGGTGGC	TTCATAGAGA

	2102	2112	2122	2132	2142	2152	2162
0.P.	TGCAGCTGCC	CCCAGGGGAC	AACATCATCT	ATGTTGGGGA	ACTGAGTCAT	CAATGGTTCC	AAAAAAGGGAG
DL.P.IRE	TGCAGCTGCC	CCCAGGGGAC	AACATCATCT	ATGTTGGGGA	ACTGAGTCAT	CAATGGTTCC	AAAAAAGGGAG
40.P.PS	TGCAGCTGCC	CCCAGGGGAC	AACATCATCT	ATGTTGGGGA	ACTGAGTCAT	CAATGGTTCC	AAAAAAGGGAG
40.P.IRE	TGCAGCTGCC	CCCAGGGGAC	AACATCATCT	ATGTTGGGGA	ACTGAGTCAT	CAATGGTTCC	AAAAAAGGGAG

	2172	2182	2192	2202	2212	2222	2232
0.P.	TAGCATCGGA	AGAGTTTTCC	AAAAGACCAA	GAAAGGCATA	GAAAGACTGA	CAGTGATAGG	GGAGCACGCC
DL.P.IRE	TAGCATCGGA	AGAGTTTTCC	AAAAGACCAA	GAAAGGCATA	GAAAGACTGA	CAGTGATAGG	GGAGCACGCC
40.P.PS	TAGCATCGGA	AGAGTTTTCC	AAAAGACCAA	GAAAGGCATA	GAAAGACTGA	CAGTGATAGG	GGAGCACGCC
40.P.IRE	TAGCATCGGA	AGAGTTTTCC	AAAAGACCAA	GAAAGGCATA	GAAAGACTGA	CAGTGATAGG	GGAGCACGCC

	2242	2252	2262	2272	2282	2292	2302
0.P.	TGGGACTTCG	GTTCTGCTGG	AGGCTTCTCG	AGTTCAATTG	GGAAGGCGGT	GCACACGGTC	CTTGGTGGTG
DL.P.IRE	TGGGACTTCG	GTTCTGCTGG	AGGCTTCTCG	AGTTCAATTG	GGAAGGCGGT	GCACACGGTC	CTTGGTGGTG
40.P.PS	TGGGACTTCG	GTTCTGCTGG	AGGCTTCTCG	AGTTCAATTG	GGAAGGCGGT	GCACACGGTC	CTTGGTGGTG
40.P.IRE	TGGGACTTCG	GTTCTGCTGG	AGGCTTCTCG	AGTTCAATTG	GGAAGGCGGT	GCACACGGTC	CTTGGTGGTG

	2312	2322	2332	2342	2352	2362	2372
0.P.	CTTTCAACAG	CATCTTCGGG	GGAGTGGGGT	TTCTACCAAA	GCTTCTATTA	GGAGTGGCAT	TGGCTTGGTT
DL.P.IRE	CTTTCAACAG	CATCTTCGGG	GGAGTGGGGT	TTCTACCAAA	GCTTCTATTA	GGAGTGGCAT	TGGCTTGGTT
40.P.PS	CTTTCAACAG	CATCTTCGGG	GGAGTGGGGT	TTCTACCAAA	GCTTCTATTA	GGAGTGGCAT	TGGCTTGGTT
40.P.IRE	CTTTCAACAG	CATCTTCGGG	GGAGTGGGGT	TTCTACCAAA	GCTTCTATTA	GGAGTGGCAT	TGGCTTGGTT

	2382	2392	2402	2412	2422	2432	2442
0.P.	GGGCCTGAAC	ATGAGAAACC	CTACAATGTC	CATGAGCTTT	CTCTTGGCTG	GAGGTCTGGT	TTTGGCCATG
DL.P.IRE	GGGCCTGAAC	ATGAGAAACC	CTACAATGTC	CATGAGCTTT	CTCTTGGCTG	GAGGTCTGGT	TTTGGCCATG
40.P.PS	GGGCCTGAAC	ATGAGAAACC	CTACAATGTC	CATGAGCTTT	CTCTTGGCTG	GAGGTCTGGT	TTTGGCCATG
40.P.IRE	GGGCCTGAAC	ATGAGAAACC	CTACAATGTC	CATGAGCTTT	CTCTTGGCTG	GAGGTCTGGT	TTTGGCCATG

	2452
0.P.	ACCCTTGGAG TGGGGGCG
DL.P.IRE	ACCCTTGGAG TGGGGGCG
40.P.PS	ACCCTTGGAG TGGGGGCG
40.P.IRE	ACCCTTGGAG TGGGGGCG

Protein NS1:

	2470	2480	2490	2500	2510	2520	2530
0.P.	GATGTTGGCT	GCGCTGTGGA	CACGGAACGA	ATGGAGCTCC	GCTGTGGCGA	AGGCCTGGTC	GTGTGGAGAG
DL.P.IRE	GATGTTGGCT	GCGCTGTGGA	CACGGAACGA	ATGGAGCTCC	GCTGTGGCGA	AGGCCTGGTC	GTGTGGAGAG
40.P.PS	GATGTTGGCT	GCGCTGTGGA	CACGGAACGA	ATGGAGCTCC	GCTGTGGCGA	AGGCCTGGTC	GTGTGGAGAG
40.P.IRE	GATGTTGGCT	GCGCTGTGGA	CACGGAACGA	ATGGAGCTCC	GCTGTGGCGA	AGGCCTGGTC	GTGTGGAGAG

	2540	2550	2560	2570	2580	2590	2600
0.P.	AGGTCTCAGA	ATGGTATGAC	AACTATGCCT	ACTACCCGGA	GACACCGGGG	GCCCTTGCAT	CAGCCATAAA
DL.P.IRE	AGGTCTCAGA	ATGGTATGAC	AACTATGCCT	ACTACCCGGA	GACACCGGGG	GCCCTTGCAT	CAGCCATAAA
40.P.PS	AGGTCTCAGA	ATGGTATGAC	AACTATGCCT	ACTACCCGGA	GACACCGGGG	GCCCTTGCAT	CAGCCATAAA
40.P.IRE	AGGTCTCAGA	ATGGTATGAC	AACTATGCCT	ACTACCCGGA	GACACCGGGG	GCCCTTGCAT	CAGCCATAAA

		2610	2620	2630	2640	2650	2660	2670
0.P.		GGAGACATTT	GAGGAGGGAA	GCTGTGGCGT	AGTCCCCCAA	AACAGGCTCG	AGATGGCCAT	GTGGAGAAGC
DL.P.IRE		GGAGACATTT	GAGGAGGGAA	GCTGTGGCGT	AGTCCCCCAA	AACAGGCTCG	AGATGGCCAT	GTGGAGAAGC
40.P.PS		GGAGACATTT	GAGGAGGGAA	GCTGTGGCGT	AGTCCCCCAA	AACAGGCTCG	AGATGGCCAT	GTGGAGAAGC
40.P.IRE		GGAGACATTT	GAGGAGGGAA	GCTGTGGCA ^T	AGTCCCCCAA	AACAGGCTCG	AGATGGCCAT	GTGGAGAAGC
		2680	2690	2700	2710	2720	2730	2740
0.P.		TCGGTCACAG	AGCTGAATCT	GGCTCTGGCG	GAAGGTGAGG	CAAATCTTAC	AGTGATGGTG	GACAAGTTTG
DL.P.IRE		TCGGTCACAG	AGCTGAATCT	GGCTCTGGCG	GAAGGTGAGG	CAAATCTTAC	AGTGATGGTG	GACAAGTTTG
40.P.PS		TCGGTCACAG	AGCTGAATCT	GGCTCTAGCG	GAAGGTGAGG	CAAATCTTAC	AGTGATGGTG	GACAAGTTTG
40.P.IRE		TCGGTCACAG	AGCTGAATCT	GGCTCTGGCG	GAAGGTGAGG	CAAATCTTAC	AGTGATGGTG	GACAAGTTTG
		2750	2760	2770	2780	2790	2800	2810
0.P.		ACCCCACTGA	CTACCGAGGT	GGTGTCCCTG	GTTTACTGAA	AAAAGGAAAG	GACATAAAAG	TCTCCTGGAA
DL.P.IRE		ACCCCACTGA	CTACCGAGGT	GGTGTCCCTG	GTTTACTGAA	AAAAGGAAAG	GACATAAAAG	TCTCCTGGAA
40.P.PS		ACCCCACTGA	CTACCGAGGT	GGTGTCCCTG	GTTTACTGAA	AAAAGGAAAG	GACATAAAAG	TCTCCTGGAA
40.P.IRE		ACCCCACTGA	CTACCGAGGT	GGTGTCCCTG	GTTTACTGAA	AAAAGGAAAG	GACATAAAAG	TCTCCTGGAA
		2820	2830	2840	2850	2860	2870	2880
0.P.		AAGCTGGGGC	CATTCAATGA	TCTGGAGCAT	TCCTGAGGCC	CCCCGTCGCT	TCATGGTGGG	CACGGAAGGA
DL.P.IRE		AAGCTGGGGC	CATTCAATGA	TCTGGAGCAT	TCCTGAGGCC	CCCCGTCGCT	TCATGGTGGG	CACGGAAGGA
40.P.PS		AAGCTGGGGC	CATTCAATGA	TCTGGAGCAT	TCCTGAGGCC	CCCCGTCGCT	TCATGGTGGG	CACGGAAGGA
40.P.IRE		AAGCTGGGGC	CATTCAATGA	TCTGGAGCAT	TCCTGAGGCC	CCCCGTCGCT	TCATGGTGGG	CACGGAAGGA
		2890	2900	2910	2920	2930	2940	2950
0.P.		CAAAGTGAGT	GTCCCCTGGA	GAGACGGAAG	ACAGGTGTTT	TCACGGTGGC	AGAATTCGGG	ATTGGCCTGA
DL.P.IRE		CAAAGTGAGT	GTCCCCTGGA	GAGACGGAAG	ACAGGTGTTT	TCACGGTGGC	AGAATTCGGG	ATTGGCCTGA
40.P.PS		CAAAGTGAGT	GTCCCCTGGA	GAGACGGAAG	ACAGGTGTTT	TCACGGTGGC	AGAATTCGGG	ATTGGCCTGA
40.P.IRE		CAAAGTGAGT	GTCCCCTGGA	GAGACGGAAG	ACAGGTGTTT	TCACGGTGGC	AGAATTCGGG	ATTGGCCTGA
		2960	2970	2980	2990	3000	3010	3020
0.P.		GAACAAAGGT	CTTCTTGAT	TTCAGACAGG	AACCAACACA	TGAGTGTGAC	ACAGGAGTGA	TGGGAGCTGC
DL.P.IRE		GAACAAAGGT	CTTCTTGAT	TTCAGACAGG	AACCAACACA	TGAGTGTGAC	ACAGGAGTGA	TGGGAGCTGC
40.P.PS		GAACAAAGGT	CTTCTTGAT	TTCAGACAGG	AACCAACACA	TGAGTGTGAC	ACAGGAGTGA	TGGGAGCTGC
40.P.IRE		GAACAAAGGT	CTTCTTGAT	TTCAGACAGG	AACCAACACA	TGAGTGTGAC	ACAGGAGTGA	TGGGAGCTGC
		3030	3040	3050	3060	3070	3080	3090
0.P.		CGTCAAGAAC	GGCATGGCAA	TCCACACAGA	TCAAAGTCTC	TGGATGAGAT	CAATGAAAAA	TGACACAGGC
DL.P.IRE		CGTCAAGAAC	GGCATGGCAA	TCCACACAGA	TCAAAGTCTC	TGGATGAGAT	CAATGAAAAA	TGACACAGGC
40.P.PS		CGTCAAGAAC	GGCATGGCAA	TCCACACAGA	TCAAAGTCTC	TGGATGAGAT	CAATGAAAAA	TGACACAGGC
40.P.IRE		CGTCAAGAAC	GGCATGGCAA	TCCACACAGA	TCAAAGTCTC	TGGATGAGAT	CAATGAAAAA	TGACACAGGC
		3100	3110	3120	3130	3140	3150	3160
0.P.		ACTTACATAG	TTGAACTTTT	GGTCACTGAC	CTGAGGAACT	GCTCATGGCC	TGCTAGCCAC	ACTATCGATA
DL.P.IRE		ACTTACATAG	TTGAACTTTT	GGTCACTGAC	CTGAGGAACT	GCTCATGGCC	TGCTAGCCAC	ACTATCGATA
40.P.PS		ACTTACATAG	TTGAACTTTT	GGTCACTGAC	CTGAGGAACT	GCTCATGGCC	TGCTAGCCAC	ACTATCGATA
40.P.IRE		ACTTACATAG	TTGAACTTTT	GGTCACTGAC	CTGAGGAACT	GCTCATGGCC	TGCTAGCCAC	ACTATCGATA

	3170	3180	3190	3200	3210	3220	3230
0.P.	ATGCTGACGT	GGTGGACTCG	GAGTTATTC	TTCCGGCGAG	CTTGGCAGGA	CCTAGATCC T	GGTACAACAG
DL.P.IRE	ATGCTGACGT	GGTGGACTCG	GAGTTATTC	TTCCGGCGAG	CTTGGCAGGA	CCTAGATCCT	GGTACAACAG
40.P.PS	ATGCTGACGT	GGTGGACTCG	GAGTTATTC	TTCCGGCGAG	CTTGGCAGGA	CCTAGATCCT	GGTACAACAG
40.P.IRE	ATGCTGACGT	GGTGGACTCG	GAGTTATTC	TTCCGGCGAG	CTTGGCAGGA	CCTAGATCCT	GGTACAACAG

	3240	3250	3260	3270	3280	3290	3300
0.P.	GATACCCGGT	TATTCAGAAC	AGGTGAAGGG	GCCATGGAAG	CACACGCCTA	TCCGAGTCAT	CAGAGAGGAG
DL.P.IRE	GATACCCGGT	TATTCAGAAC	AGGTGAAGGG	GCCATGGAAG	CACACGCCTA	TCCGAGTCAT	CAGAGAGGAG
40.P.PS	GATACCCGGT	TATTCAGAAC	AGGTGAAGGG	GCCATGGAAG	CACACGCCTA	TCCGAGTCAT	CAGAGAGGAG
40.P.IRE	GATACCCGGT	TATTCAGAAC	AGGTGAAGGG	GCCATGGAAG	CACACGCCTA	TCCGAGTCAT	CAGAGAGGAG

	3310	3320	3330	3340	3350	3360	3370
0.P.	TGTCCCGGCA	CGACCGTTAC	CATCAACGCC	AAGTGTGACA	AAAGAGGAGC	ATCTGTGAGG	AGTACCACAG
DL.P.IRE	TGTCCCGGCA	CGACCGTTAC	CATCAACGCC	AAGTGTGACA	AAAGAGGAGC	ATCTGTGAGG	AGTACCACAG
40.P.PS	TGTCCCGGCA	CGACCGTTAC	CATCAACGCC	AAGTGTGACA	AAAGAGGAGC	ATCTGTGAGG	AGTACCACAG
40.P.IRE	TGTCCCGGCA	CGACCGTTAC	CATCAACGCC	AAGTGTGACA	AAAGAGGAGC	ATCTGTGAGG	AGTACCACAG

	3380	3390	3400	3410	3420	3430	3440
0.P.	AGAGTGCAA	GGTTATCCCA	GAATGGTGCT	GCCGAGCGTG	CACAATGCCA	CCAGTGACGT	TCCGGACTGG
DL.P.IRE	AGAGTGCAA	GGTTATCCCA	GAATGGTGCT	GCCGAGCGTG	CACAATGCCA	CCAGTGACGT	TCCGGACTGG
40.P.PS	AGAGTGCAA	GGTTATCCCA	GAATGGTGCT	GCCGAGCGTG	CACAATGCCA	CCAGTGACGT	TCCGGACTGG
40.P.IRE	AGAGTGCAA	GGTTATCCCA	GAATGGTGCT	GCCGAGCGTG	CACAATGCCA	CCAGTGACGT	TCCGGACTGG

	3450	3460	3470	3480	3490	3500	3510
0.P.	AACTGATTGC	TGGTATGCCA	TGGAATACG	GCCAGTCCAT	GACCAGGGGG	GGCTTGTTTCG	CTCAATGGTG
DL.P.IRE	AACTGATTGC	TGGTATGCCA	TGGAATACG	GCCAGTCCAT	GACCAGGGGG	GGCTTGTTTCG	CTCAATGGTG
40.P.PS	AACTGATTGC	TGGTATGCCA	TGGAATACG	GCCAGTCCAT	GACCAGGGGG	GGCTTGTTTCG	C TCAATGGTG
40.P.IRE	AACTGATTGC	TGGTATGCCA	TGGAATACG	GCCAGTCCAT	GACCAGGGGG	GGCTTGTTTCG	CTCAATGGTG

0.P. GTTGCG
DL.P.IRE GTTGCG
40.P.PS GTTGCG
40.P.IRE GTTGCG

Protein NS2A:

	3526	3536	3546	3556	3566	3576	3586
0.P.	GACAACGGTG	AATTACTCAG	TGAGGGAGGG	GTCCCCGGAA	TAGTGGCATT	GTTTGTGGTC	CTTGAATACA
DL.P.IRE	GACAACGGTG	AATTACTCAG	TGAGGGAGGG	GTCCCCGGAA	TAGTGGCATT	GTTTGTGGTC	CTTGAATACA
40.P.PS	GACAACGGTG	AATTACTCAG	TGAGGGAGGG	GTCCCCGGAA	TAGTGGCATT	GTTTGTGGTC	CTTGAATACA
40.P.IRE	GACAACGGTG	AATTACTCAG	TGAGGGAGGG	GTCCCCGGAA	TAGTGGCATT	GTTTGTGGTC	CTTGAATACA

	3596	3606	3616	3626	3636	3646	3656
0.P.	TCATCCGCAG	GAGACCCTCC	ACGGGATCAA	CGGTTGTGTG	GGGGGGCATC	GTCGTTCTCG	C TTGCTTGT
DL.P.IRE	TCATCCGCAG	GAGACCCTCC	ACGGGATCAA	CGGTTGTGTG	GGGGGGCATC	GTCGTTCTCG	C TTGCTTGT
40.P.PS	TCATCCGCAG	GAGACCCTCC	ACGGGATCAA	CGGTTGTGTG	GGGGGGCATC	GTCGTTCTCG	C TTGCTTGT
40.P.IRE	TCATCCGCAG	GAGACCCTCC	ACGGGATCAA	CGGTTGTGTG	GGGGGGCATC	GTCGTTCTCG	C TTGCTTGT

		3666	3676	3686	3696	3706	3716	3726
0.P.	CACCGGGATG	GTCAGGATGG	AGAGCCTGGT	GCGCTATGTG	GTGGCAGTGG	GGATCACATT	CCACCTTGAG	
DL.P.IRE	CACCGGGATG	GTCAGGATGG	AGAGCCTGGT	GCGCTATGTG	GTGGCAGTGG	GGATCACATT	CCACCTTGAG	
40.P.PS	CACCGGGATG	GTCAGGATGG	AGAGCCTGGT	GCGCTATGTG	GTGGCAGTGG	GGATCACATT	CCACCTTGAG	
40.P.IRE	CACCGGGATG	GTCAGGATGG	AGAGCCTGGT	GCGCTATGTG	GTGGCAGTGG	GGATCACATT	CCACCTTGAG	
		3736	3746	3756	3766	3776	3786	3796
0.P.	CTAGGGCCAG	AGATCGTGGC	CTTGATGCTA	CTCCAGGCTG	TGTTTGAGCT	GAGGGTGGGT	TTGCTCAGCG	
DL.P.IRE	CTAGGGCCAG	AGATCGTGGC	CTTGATGCTA	CTCCAGGCTG	TGTTTGAGCT	GAGGGTGGGT	TTGCTCAGCG	
40.P.PS	CTAGGGCCAG	AGATCGTGGC	CTTGATGCTA	CTCCAGGCTG	TGTTTGAGCT	GAGGGTGGGT	TTGCTCAGCG	
40.P.IRE	CTAGGGCCAG	AGATCGTGGC	CTTGATGCTA	CTCCAGGCTG	TGTTTGAGCT	GAGGGTGGGT	TTGCTCAGCG	
		3806	3816	3826	3836	3846	3856	3866
0.P.	CATTTGCGCT	GCGCAGAAGC	CTCACCGTCC	GAGAGATGGT	GACCACCTAC	TTTCTTTTGC	TGGTCTGGA	
DL.P.IRE	CATTTGCGCT	GCGCAGAAGC	CTCACCGTCC	GAGAGATGGT	GACCACCTAC	TTTCTTTTGC	TGGTCTGGA	
40.P.PS	CATTTGCGCT	GCGCAGAAGC	CTCACCGTCC	GAGAGATGGT	GACCACCTAC	TTTCTTTTGC	TGGTCTGGA	
40.P.IRE	CATTTGCGCT	GCGCAGAAGC	CTCACCGTCC	GAGAGATGGT	GACCACCTAC	TTTCTTTTGC	TGGTCTGGA	
		3876	3886	3896	3906	3916	3926	3936
0.P.	ATTGGGGCTG	CCGAGTGCGA	ACCTTGAGGA	TTTCTGGAAA	TGGGGTGATG	CACTGGCCAT	GGGGGCGCTG	
DL.P.IRE	ATTGGGGCTG	CCGAGTGCGA	ACCTTGAGGA	TTTCTGGAAA	TGGGGTGATG	CACTGGCCAT	GGGGGCGCTG	
40.P.PS	ATTGGGGCTG	CCGAGTGCGA	ACCTTGAGGA	TTTCTGGAAA	TGGGGTGATG	CACTGGCCAT	GGGGGCGCTG	
40.P.IRE	ATTGGGGCTG	CCGAGTGCGA	ACCTTGAGGA	TTTCTGGAAA	TGGGGTGATG	CACTGGCCAT	GGGGGCGCTG	
		3946	3956	3966	3976	3986	3996	4006
0.P.	ATATTCAGGG	CTTGCACGGC	AGAAGGAAAG	ACTGGAGCGG	GGCTCCTGCT	CATGGCTCTC	ATGACACAGC	
DL.P.IRE	ATATTCAGGG	CTTGCACGGC	AGAAGGAAAG	ACTGGAGCGG	GGCTCCTGCT	CATGGCTCTC	ATGACACAGC	
40.P.PS	ATATTCAGGG	CTTGCACGGC	AGAAGGAAAG	ACTGGAGCGG	GGCTCCTGCT	CATGGCTCTC	ATGACACAGC	
40.P.IRE	ATATTCAGGG	CTTGCACGGC	AGAAGGAAAG	ACTGGAGCGG	GGCTCCTGCT	CATGGCTCTC	ATGACACAGC	
		4016	4026	4036	4046	4056	4066	4076
0.P.	AGGATGTGGT	GACTGTGCAT	CATGGACTGG	TGTGCTTCTT	GAGTGCAGCT	TCGGCTTGCT	CGGTCTGGAG	
DL.P.IRE	AGGATGTGGT	GACTGTGCAT	CATGGACTGG	TGTGCTTCTT	GAGTGCAGCT	TCGGCTTGCT	CGGTCTGGAG	
40.P.PS	AGGATGTGGT	GACTGTGCAT	CATGGACTGG	TGTGCTTCTT	GAGTGCAGCT	TCGGCTTGCT	CGGTCTGGAG	
40.P.IRE	AGGATGTGGT	GACTGTGCAT	CATGGACTGG	TGTGCTTCTT	GAGTGCAGCT	TCGGCTTGCT	CGGTCTGGAG	
		4086	4096	4106	4116	4126	4136	4146
0.P.	GCTGCTCAGG	GGACACAGAG	AGCAGAAGGG	ATTGACCTGG	ATTGTCCCC	TGGCTAGATT	GCTTGGGGGA	
DL.P.IRE	GCTGCTCAGG	GGACACAGAG	AGCAGAAGGG	ATTGACCTGG	ATTGTCCCC	TGGCTAGATT	GCTTGGGGGA	
40.P.PS	GCTGCTCAGG	GGACACAGAG	AGCAGAAGGG	ATTGACCTGG	ATTGCCCCC	TGGCTAGATT	GCTTGGGGGA	
40.P.IRE	GCTGCTCAGG	GGACACAGAG	AGCAGAAGGG	ATTGACCTGG	ATTGTCCCC	TGGCTAGATT	GCTTGGGGGA	
		4156	4166	4176	4186	4196	4206	
0.P.	GAGGGCTCTG	GAATCAGACT	GCTGGCGTTT	TGGGAGCTGT	CAGCTCACAG	AGGAAGACGA		
DL.P.IRE	GAGGGCTCTG	GAATCAGACT	GCTGGCGTTT	TGGGAGCTGT	CAGCTCACAG	AGGAAGACGA		
40.P.PS	GAGGGCTCTG	GAATCAGACT	GCTGGCGTTT	TGGGAGCTGT	CAGCTCACAG	AGGAAGACGA		
40.P.IRE	GAGGGCTCTG	GAATCAGACT	GCTGGCGTTT	TGGGAGCTGT	CAGCTCACAG	AGGAAGACGA		

Protein NS2B:

		4216		4226		4236		4246		4256		4266		4276
0.P.	TCTTTCAGTG	AACCACTAAC	TGTGGTAGGG	GTCATGCTGA	CATTGGCTAG	CGGCATGATG	CGACACACCT							
DL.P.IRE	TCTTTCAGTG	AACCACTAAC	TGTGGTAGGG	GTCATGCTGA	CATTGGCTAG	CGGCATGATG	CGACACACCT							
40.P.PS	TCTTTCAGTG	AACCACTAAC	TGTGGTAGGG	GTCATGCTGA	CATTGGCTAG	CGGCATGATG	CGACACACCT							
40.P.IRE	TCTTTCAGTG	AACCACTAAC	TGTGGTAGGG	GTCATGCTGA	CATTGGCTAG	CGGCATGATG	CGACACACCT							
		4286		4296		4306		4316		4326		4336		4346
0.P.	CCCAGGAGGC	TCTCTGTGCA	CTCGCAGTGG	CCTCGTTTCT	CTTGTTGATG	CTGGTGCTGG	GGACAAGAAA							
DL.P.IRE	CCCAGGAGGC	TCTCTGTGCA	CTCGCAGTGG	CCTCGTTTCT	CTTGTTGATG	CTGGTGCTGG	GGACAAGAAA							
40.P.PS	CCCAGGAGGC	TCTCTGTGCA	CTCGCAGTGG	CCTCGTTTCT	CTTGTTGATG	CTGGTGCTGG	GGACAAGAAA							
40.P.IRE	CCCAGGAGGC	TCTCTGTGCA	CTCGCA A TGG	CCTCGTTTCT	CTTGTTGATG	CTGGTGCTGG	GGACAAGAAA							
		4356		4366		4376		4386		4396		4406		4416
0.P.	GATGCAGCTG	GTTGCCGAAT	GGAGTGGCTG	TGTGGAATGG	CATCCGGAAC	TAGTGAATGA	GGGTGGAGAG							
DL.P.IRE	GATGCAGCTG	GTTGCCGAAT	GGAGTGGCTG	TGTGGAATGG	CATCCGGAAC	TAGTGAATGA	GGGTGGAGAG							
40.P.PS	GATGCAGCTG	GTTGCCGAAT	GGAGTGGCTG	TGTGGAATGG	CATCCGGAAC	TAGTGAATGA	GGGTGGAGAG							
40.P.IRE	GATGCAGCTG	GTTGCCGAAT	GGAGTGGCTG	TGTGGAATGG	CATCCGGAAC	TAGTGAATGA	GGGTGGAGAG							
		4426		4436		4446		4456		4466		4476		4486
0.P.	GTTAGCCTGC	GGGTCCGTCA	GGACGCGATG	GGAAACTTTC	ACTTGACTGA	GCTCGAGAAA	GAAGAGAGAA							
DL.P.IRE	GTTAGCCTGC	GGGTCCGTCA	GGACGCGATG	GGAAACTTTC	ACTTGACTGA	GCTCGAGAAA	GAAGAGAGAA							
40.P.PS	GTTAGCCTGC	GGGTCCGTCA	GGACGCGATG	GGAAACTTTC	ACTTGACTGA	GCTCGAGAAA	GAAGAGAGAA							
40.P.IRE	GTTAGCCTGC	GGGTCCGTCA	GGACGCGATG	GGAAACTTTC	ACTTGACTGA	GCTCGAGAAA	GAAGAGAGAA							
		4496		4506		4516		4526		4536		4546		4556
0.P.	TGATGGCTTT	TTGGTTGATT	GCCGGCTTGG	CAGCTTCAGC	CATTCATTGG	TCAGGCATTC	TTGGTGTGAT							
DL.P.IRE	TGATGGCTTT	TTGGTTGATT	GCCGGCTTGG	CAGCTTCAGC	CATTCATTGG	TCAGGCATTC	TTGGTGTGAT							
40.P.PS	TGATGGCTTT	TTGGTTGATT	GCCGGCTTGG	CAGCTTCAGC	CATTCATTGG	TCAGGCATTC	TTGGTGTGAT							
40.P.IRE	TGATGGCTTT	TTGGTTGATT	GCCGGCTTGG	CAGCTTCAGC	CATTCATTGG	TCAGGCATTC	TTGGTGTGAT							
		4566		4576		4586		4596						
0.P.	GGGACTGTGG	ACGCTCACGG	AAATGCTGAG	GTCATCCCGA	AGG									
DL.P.IRE	GGGACTGTGG	ACGCTCACGG	AAATGCTGAG	GTCATCCCGA	AGG									
40.P.PS	GGGACTGTGG	ACGCTCACGG	AAATGCTGAG	GTCATCCCGA	AGG									
40.P.IRE	GGGACTGTGG	ACGCTCACGG	AAATGCTGAG	GTCATCCCGA	AGG									

Protein NS3:

		4609		4619		4629		4639		4649		4659		4669
0.P.	TCTGACCTGG	TTTTCTCTGG	ACAGGGGGGT	CGAGAGCGTG	GTGACAGACC	TTTCGAGGTT	AAGGACGGTG							
DL.P.IRE	TCTGACCTGG	TTTTCTCTGG	ACAGGGGGGT	CGAGAGCGTG	GTGACAGACC	TTTCGAGGTT	AAGGACGGTG							
40.P.PS	TCTGACCTGG	TTTTCTCTGG	ACAGGGGGGT	CGAGAGCGTG	GTGACAGACC	TTTCGAGGTT	AAGGACGGTG							
40.P.IRE	TCTGACCTGG	TTTTCTCTGG	ACAGGGGGGT	CGAGAGCGTG	GTGACAGACC	TTTCGAGGTT	AAGGACGGTG							
		4679		4689		4699		4709		4719		4729		4739
0.P.	TCTACAGGAT	TTTCAGCCCC	GGCTTGTTCT	GGGGTCAGAA	CCAGGTGGGA	GTTGGCTACG	GTTCCAAAGG							
DL.P.IRE	TCTACAGGAT	TTTCAGCCCC	GGCTTGTTCT	GGGGTCAGAA	CCAGGTGGGA	GTTGGCTACG	GTTCCAAAGG							
40.P.PS	TCTACAGGAT	TTTCAGCCCC	GGCTTGTTCT	GGGGTCAGAA	CCAGGTGGGA	GTTGGCTACG	GTTCCAAAGG							
40.P.IRE	TCTACAGGAT	TTTCAGCCCC	GGCTTGTTCT	GGGGTCAGAA	CCAGGTGGGA	GTTGGCTACG	GTTCCAAAGG							

	4749	4759	4769	4779	4789	4799	4809
0.P.	TGTCTTGAC	ACGATGTGGC	ATGTGACGAG	AGGAGCGGCG	CTGTCTATTG	ATGATGCTGT	GGCCGGTCCC
DL.P.IRE	TGTCTTGAC	ACGATGTGGC	ATGTGACGAG	AGGAGCGGCG	CTGTCTATTG	ATGATGCTGT	GGCCGGTCCC
40.P.PS	TGTCTTGAC	ACGATGTGGC	ATGTGACGAG	AGGAGCGGCG	CTGTCTATTG	ATGATGCTGT	GGCCGGTCCC
40.P.IRE	TGTCTTGAC	ACGATGTGGC	ATGTGACGAG	AGGAGCGGCG	CTGTCTATTG	ATGATGCTGT	GGCCGGTCCC
	4819	4829	4839	4849	4859	4869	4879
0.P.	TACTGGGCTG	ATGTGAGGGA	GGATGTTGTG	TGCTACGGAG	GAGCCTGGAG	TCTGGAGGAA	AAATGGAAAG
DL.P.IRE	TACTGGGCTG	ATGTGAGGGA	GGATGTTGTG	TGCTACGGAG	GAGCCTGGAG	TCTGGAGGAA	AAATGGAAAG
40.P.PS	TACTGGGCTG	ATGTGAGGGA	GGATGTTGTG	TGCTACGGAG	GAGCCTGGAG	TCTGGAGGAA	AAATGGAAAG
40.P.IRE	TACTGGGCTG	ATGTGAGGGA	GGATGTTGTG	TGCTACGGAG	GAGCCTGGAG	TCTGGAGGAA	AAATGGAAAG
	4889	4899	4909	4919	4929	4939	4949
0.P.	GTGAAACAGT	ACAGGTTTCAT	GCCTTCCCAC	CAGGGAAGGC	CCATGAGGTG	CATCAGTGCC	AGCCTGGGGA
DL.P.IRE	GTGAAACAGT	ACAGGTTTCAT	GCCTTCCCAC	CAGGGAAGGC	CCATGAGGTG	CATCAGTGCC	AGCCTGGGGA
40.P.PS	GTGAAACAGT	ACAGGTTTCAT	GCCTTCCCAC	CAGGGAAGGC	CCATGAGGTG	CATCAGTGCC	AGCCTGGGGA
40.P.IRE	GTGAAACAGT	ACAGGTTTCAT	GCCTTCCCAC	CAGGGAAGGC	CCATGAGGTG	CATCAGTGCC	AGCCTGGGGA
	4959	4969	4979	4989	4999	5009	5019
0.P.	GTTGATCCTT	GACACCGGAA	GGAAGCTTGG	GGCAATACCA	ATTGATTTGG	TAAAAGGAAC	TCAGGCAGC
DL.P.IRE	GTTGATCCTT	GACACCGGAA	GGAAGCTTGG	GGCAATACCA	ATTGATTTGG	TAAAAGGAAC	TCAGGCAGC
40.P.PS	GTTGATCCTT	GACACCGGAA	GGAAGCTTGG	GGCAATACCA	ATTGATTTGG	TAAAAGGAAC	TCAGGCAGC
40.P.IRE	GTTGATCCTT	GACACCGGAA	GGAAGCTTGG	GGCAATACCA	ATTGATTTGG	TAAAAGGAAC	TCAGGCAGC
	5029	5039	5049	5059	5069	5079	5089
0.P.	CCCATTCTTA	ACGCCCAGGG	AGTGGTCGTG	GGGCTATACG	GAAATGGCCT	AAAGACCAAT	GAGACCTACG
DL.P.IRE	CCCATTCTTA	ACGCCCAGGG	AGTGGTCGTG	GGGCTATACG	GAAATGGCCT	AAAGACCAAT	GAGACCTACG
40.P.PS	CCCATTCTTA	ACGCCCAGGG	AGTGGTCGTG	GGGCTATACG	GAAATGGCCT	AAAGACCAAT	GAGACCTACG
40.P.IRE	CCCATTCTTA	ACGCCCAGGG	AGTGGTCGTG	GGGCTATACG	GAAATGGCCT	AAAGACCAAT	GAGACCTACG
	5099	5109	5119	5129	5139	5149	5159
0.P.	TCAGCAGCAT	TGCTCAAGGG	GAAGCGGAGA	AGAGTCGACC	CAACCTCCCA	CAGGCTGTTG	TGGGCACGGG
DL.P.IRE	TCAGCAGCAT	TGCTCAAGGG	GAAGCGGAGA	AGAGTCGACC	CAACCTCCCA	CAGGCTGTTG	TGGGCACGGG
40.P.PS	TCAGCAGCAT	TGCTCAAGGG	GAAGCGGAGA	AGAGTCGACC	CAACCTCCCA	CAGGCTGTTG	TGGGCACGGG
40.P.IRE	TCAGCAGCAT	TGCTCAAGGG	GAAGCGGAGA	AGAGTCGACC	CAACCTCCCA	CAGGCTGTTG	TGGGCACGGG
	5169	5179	5189	5199	5209	5219	5229
0.P.	CTGGACATCA	AAGGGTCAGA	TCACAGTGCT	GGACATGCAC	CCAGGCTCAG	GGAAGACCCA	CAGAGTCCTC
DL.P.IRE	CTGGACATCA	AAGGGTCAGA	TCACAGTGCT	GGACATGCAC	CCAGGCTCAG	GGAAGACCCA	CAGAGTCCTC
40.P.PS	CTGGACATCA	AAGGGTCAGA	TCACAGTGCT	GGACATGCAC	CCAGGCTCAG	GGAAGACCCA	CAGAGTCCTC
40.P.IRE	CTGGACATCA	AAGGGTCAGA	TCACAGTGCT	GGACATGCAC	CCAGGCTCAG	GGAAGACCCA	CAGAGTCCTC
	5239	5249	5259	5269	5279	5289	5299
0.P.	CCGGAGCTCA	TTCGCCAATG	CATTGACAGG	CGCCTGAGAA	CGTTGGTGTT	GGCTCCAACT	CGTGTGGTAC
DL.P.IRE	CCGGAGCTCA	TTCGCCAATG	CATTGACAGG	CGCCTGAGAA	CGTTGGTGTT	GGCTCCAACT	CGTGTGGTAC
40.P.PS	CCGGAGCTCA	TTCGCCAATG	CATTGACAGG	CGCCTGAGAA	CGTTGGTGTT	GGCTCCAACT	CGTGTGGTAC
40.P.IRE	CCGGAGCTCA	TTCGCCAATG	CATTGACAGG	CGCCTGAGAA	CGTTGGTGTT	GGCTCCAACT	CGTGTGGTAC

		5309		5319		5329		5339		5349		5359		5369
0.P.	TCAAAGAAAT	GGAGCGCGCC	TTGAATGGGA	AACGGGTCAG	GTTCCACTCA	CCAGCAGTCA	GTGACCAACA							
DL.P.IRE	TCAAAGAAAT	GGAGCGCGCC	TTGAATGGGA	AACGGGTCAG	GTTCCACTCA	CCAGCAGTCA	GTGACCAACA							
40.P.PS	TCAAAGAAAT	GGAGCGCGCC	TTGAATGGGA	AACGGGTCAG	GTTCCACTCA	CCAGCAGTCA	GTGACCAACA							
40.P.IRE	TCAAAGAAAT	GGAGCGCGCC	TTGAATGGGA	AACGGGTCAG	GTTCCACTCA	CCAGCAGTCA	GTGACCAACA							
		5379		5389		5399		5409		5419		5429		5439
0.P.	GGCTGGAGGG	GCAATTGTCG	ATGTGATGTG	TCACGCAACC	TATGTCAACA	GACGGCTACT	ACCACAGGGG							
DL.P.IRE	GGCTGGAGGG	GCAATTGTCG	ATGTGATGTG	TCACGCAACC	TATGTCAACA	GACGGCTACT	ACCACAGGGG							
40.P.PS	GGCTGGAGGG	GCAATTGTCG	ATGTGATGTG	TCACGCAACC	TATGTCAACA	GACGGCTACT	ACCACAGGGG							
40.P.IRE	GGCTGGAGGG	GCAATTGTCG	ATGTGATGTG	TCACGCAACC	TATGTCAACA	GACGGCTACT	ACCACAGGGG							
		5449		5459		5469		5479		5489		5499		5509
0.P.	AGACAAAATT	GGGAGGTGGC	AATCATGGAT	GAGGCCCACT	GGACTGACCC	TCACAGCATA	GCTGCCAGAG							
DL.P.IRE	AGACAAAATT	GGGAGGTGGC	AATCATGGAT	GAGGCCCACT	GGACTGACCC	TCACAGCATA	GCTGCCAGAG							
40.P.PS	AGACAAAATT	GGGAGGTGGC	AATCATGGAT	GAGGCCCACT	GGACTGACCC	TCACAGCATA	GCTGCCAGAG							
40.P.IRE	AGACAAAATT	GGGAGGTGGC	AATCATGGAT	GAGGCCCACT	GGACTGACCC	TCACAGCATA	GCTGCCAGAG							
		5519		5529		5539		5549		5559		5569		5579
0.P.	GTCATTGTA	CACTCTGGCA	AAAGAAAACA	AGTGTGCACT	GGTCTTGATG	ACAGCGACAC	CTCCTGGTAA							
DL.P.IRE	GTCATTGTA	CACTCTGGCA	AAAGAAAACA	AGTGTGCACT	GGTCTTGATG	ACAGCGACAC	CTCCTGGTAA							
40.P.PS	GTCATTGTA	CACTCTGGCA	AAAGAAAACA	AGTGTGCACT	GGTCTTGATG	ACAGCGACAC	CTCCTGGCAA							
40.P.IRE	GTCATTGTA	CACTCTGGCA	AAAGAAAACA	AGTGTGCACT	GGTCTTGATG	ACAGCGACAC	CTCCTGGCAA							
		5589		5599		5609		5619		5629		5639		5649
0.P.	GAGTGAACCT	TTTCCGGAGT	CTAACGGAGC	CATCACTAGT	GAGGAAAGAC	AGATTCTGA	TGGAGAGTGG							
DL.P.IRE	GAGTGAACCT	TTTCCGGAGT	CTAACGGAGC	CATCACTAGT	GAGGAAAGAC	AGATTCTGA	TGGAGAGTGG							
40.P.PS	GAGTGAACCT	TTTCCGGAGT	CTAACGGAGC	CATCACTAGT	GAGGAAAGAC	AGATTCTGA	TGGAGAGTGG							
40.P.IRE	GAGTGAACCT	TTTCCGGAGT	CTAACGGAGC	CATCACTAGT	GAGGAAAGAC	AGATTCTGA	TGGAGAGTGG							
		5659		5669		5679		5689		5699		5709		5719
0.P.	CGTGACGGGT	TTGACTGGAT	CACTGAGTAT	GAAGGACGCA	CAGCCTGGTT	TGTCCCTTCG	ATTGCAAAAAG							
DL.P.IRE	CGTGACGGGT	TTGACTGGAT	CACTGAGTAT	GAAGGACGCA	CAGCCTGGTT	TGTCCCTTCG	ATTGCAAAAAG							
40.P.PS	CGTGACGGGT	TTGACTGGAT	CACTGAGTAT	GAAGGACGCA	CAGCCTGGTT	TGTCCCTTCG	ATTGCAAAAAG							
40.P.IRE	CGTGACGGGT	TTGACTGGAT	CACTGAGTAT	GAAGGACGCA	CAGCCTGGTT	TGTCCCTTCG	ATTGCAAAAAG							
		5729		5739		5749		5759		5769		5779		5789
0.P.	GTGGTGCTAT	AGCTCGCACC	TTGAGACAGA	AAGGGAAAAG	TGTGATCTGT	TTGAACAGCA	AAACCTTTGA							
DL.P.IRE	GTGGTGCTAT	AGCTCGCACC	TTGAGACAGA	AAGGGAAAAG	TGTGATCTGT	TTGAACAGCA	AAACCTTTGA							
40.P.PS	GTGGTGCTAT	AGCTCGCACC	TTGAGACAGA	AAGGGAAAAG	TGTGATCTGT	TTGAACAGCA	AAACCTTTGA							
40.P.IRE	GTGGTGCTAT	AGCTCGCACC	TTGAGACAGA	AAGGGAAAAG	TGTGATCTGT	TTGAACAGCA	AAACCTTTGA							
		5799		5809		5819		5829		5839		5849		5859
0.P.	AAAGGACTAC	TCAAGAGTGA	GGGATGAGAA	GCCTGACTTT	GTGGTGACGA	CTGATATCTC	GGAGATGGGA							
DL.P.IRE	AAAGGACTAC	TCAAGAGTGA	GGGATGAGAA	GCCTGACTTT	GTGGTGACGA	CTGATATCTC	GGAGATGGGA							
40.P.PS	AAAGGACTAC	TCAAGAGTGA	GGGATGAGAA	GCCTGACTTT	GTGGTGACGA	CTGATATCTC	GGAGATGGGA							
40.P.IRE	AAAGGACTAC	TCAAGAGTGA	GGGATGAGAA	GCCTGACTTT	GTGGTGACGA	CTGATATCTC	GGAGATGGGA							

		5869	5879	5889	5899	5909	5919	5929
0.P.		GCCAACCTTG	ACGTGAGCCG	CGTCATAGAT	GGGAGGACAA	ACATCAAGCC	TGAAGAGGTT	GATGGGAAGG
DL.P.IRE		GCCAACCTTG	ACGTGAGCCG	CGTCATAGAT	GGGAGGACAA	ACATCAAGCC	TGAAGAGGTT	GATGGGAAGG
40.P.PS		GCCAACCTTG	ACGTGAGCCG	CGTCATAGAT	GGGAGGACAA	ACATCAAGCC	TGAAGAGGTT	GATGGGAAGG
40.P.IRE		GCCAACCTTG	ACGTGAGCCG	CGTCATAGAT	GGGAGGACAA	ACATCAAGCC	TGAAGAGGTT	GATGGGAAGG
		5939	5949	5959	5969	5979	5989	5999
0.P.		TCGAGCTCAC	CGGGACCAGG	CGAGTGACCA	CGGCTTCCGC	TGCCCAACGG	CGCGGAAGAG	TTGGTCGGCA
DL.P.IRE		TCGAGCTCAC	CGGGACCAGG	CGAGTGACCA	CGGCTTCCGC	TGCCCAACGG	CGCGGAAGAG	TTGGTCGGCA
40.P.PS		TCGAGCTCAC	CGGGACCAGG	CGAGTGACCA	CGGCTTCCGC	TGCCCAACGG	CGCGGAAGAG	TTGGTCGGCA
40.P.IRE		TCGAGCTCAC	CGGGACCAGG	CGAGTGACCA	CGGCTTCCGC	TGCCCAACGG	CGCGGAAGAG	TTGGTCGGCA
		6009	6019	6029	6039	6049	6059	6069
0.P.		AGACGGACGA	ACAGATGAAT	ACATATACTC	TGGACAGTGT	GATGATGATG	ACAGTGGATT	AGTGAATGG
DL.P.IRE		AGACGGACGA	ACAGATGAAT	ACATATACTC	TGGACAGTGT	GATGATGATG	ACAGTGGATT	AGTGAATGG
40.P.PS		AGACGGACGA	ACAGATGAAT	ACATATACTC	TGGACAGTGT	GATGATGATG	ACAGTGGATT	AGTGAATGG
40.P.IRE		AGACGGACGA	ACAGATGAAT	ACATATACTC	TGGACAGTGT	GATGATGATG	ACAGTGGATT	AGTGAATGG
		6079	6089	6099	6109	6119	6129	6139
0.P.		AAGGAGGCGC	AAATACTTCT	TGACAACATA	ACAACCTTGC	GGGGGCCCGT	GGCCACCTTT	TATGGACCAG
DL.P.IRE		AAGGAGGCGC	AAATACTTCT	TGACAACATA	ACAACCTTGC	GGGGGCCCGT	GGCCACCTTT	TATGGACCAG
40.P.PS		AAGGAGGCGC	AAATACTTCT	TGACAACATA	ACAACCTTGC	GGGGGCCCGT	GGCCACCTTT	TATGGACCAG
40.P.IRE		AAGGAGGCGC	AAATACTTCT	TGACAACATA	ACAACCTTGC	GGGGGCCCGT	GGCCACCTTT	TATGGACCAG
		6149	6159	6169	6179	6189	6199	6209
0.P.		AACAGGACAA	GATGCCGGAG	GTGGCCGGTC	ACTTTCGACT	CACTGAAGAG	AAAAGAAAGC	ACTTCCGACA
DL.P.IRE		AACAGGACAA	GATGCCGGAG	GTGGCCGGTC	ACTTTCGACT	CACTGAAGAG	AAAAGAAAGC	ACTTCCGACA
40.P.PS		AACAGGACAA	GATGCCGGAG	GTGGCCGGTC	ACTTTCGACT	CACTGAAGAG	AAAAGAAAGC	ACTTCCGACA
40.P.IRE		AACAGGACAA	GATGCCGGAG	GTGGCCGGTC	ACTTTCGACT	CACTGAAGAG	AAAAGAAAGC	ACTTCCGACA
		6219	6229	6239	6249	6259	6269	6279
0.P.		TCTTCTCACT	CATTGTGACT	TTACACCGTG	GCTGGCGTGG	CACGTCCGAG	CGAATGTGTC	CAGCGTCACG
DL.P.IRE		TCTTCTCACT	CATTGTGACT	TTACACCGTG	GCTGGCGTGG	CACGTCCGAG	CGAATGTGTC	CAGCGTCACG
40.P.PS		TCTTCTCACT	CATTGTGACT	TTACACCGTG	GCTGGCGTGG	CACGTCCGAG	CGAATGTGTC	CAGCGTCACG
40.P.IRE		TCTTCTCACT	CATTGTGACT	TTACACCGTG	GCTGGCGTGG	CACGTCCGAG	CGAATGTGTC	CAGCGTCACG
		6289	6299	6309	6319	6329	6339	6349
0.P.		GATCGAAGCT	GGACATGGGA	AGGGCCGGAG	GCAAATGCCG	TGGATGAGGC	CAGTGGTGAC	TTGGTCACCT
DL.P.IRE		GATCGAAGCT	GGACATGGGA	AGGGCCGGAG	GCAAATGCCG	TGGATGAGGC	CAGTGGTGAC	TTGGTCACCT
40.P.PS		GATCGAAGCT	GGACATGGGA	AGGGCCGGAG	GCAAATGCCG	TGGATGAGGC	CAGTGGTGAC	TTGGTCACCT
40.P.IRE		GATCGAAGCT	GGACATGGGA	AGGGCCGGAG	GCAAATGCCG	TGGATGAGGC	CAGTGGTGAC	TTGGTCACCT
		6359	6369	6379	6389	6399	6409	6419
0.P.		TCAGGAGCCC	GAATGGGGCG	GAGAGAACTC	TCAGGCCGGT	GTGGAAGGAC	GCACGCATGT	TCAAAGAGGG
DL.P.IRE		TCAGGAGCCC	GAATGGGGCG	GAGAGAACTC	TCAGGCCGGT	GTGGAAGGAC	GCACGCATGT	TCAAAGAGGG
40.P.PS		TCAGGAGCCC	GAATGGGGCG	GAGAGAACTC	TCAGGCCGGT	GTGGAAGGAC	GCACGCATGT	TCAAAGAGGG
40.P.IRE		TCAGGAGCCC	GAATGGGGCG	GAGAGAACTC	TCAGGCCGGT	GTGGAAGGAC	GCACGCATGT	TCAAAGAGGG

		6429	6439	6449	6459	
0.P.	ACGTGACATC	AAAGAGTTCG	TGGCGTACGC	ATCTGGGCGT	CGC	
DL.P.IRE	ACGTGACATC	AAAGAGTTCG	TGGCGTACGC	ATCTGGGCGT	CGC	
40.P.PS	ACGTGACATC	AAAGAGTTCG	TGGCGTACGC	ATCTGGGCGT	CGC	
40.P.IRE	ACGTGACATC	AAAGAGTTCG	TGGCGTACGC	ATCTGGGCGT	CGC	

Protein NS4A:

		6472	6482	6492	6502	6512	6522	6532
0.P.	AGTTTCGGAG	ATGTTCTGAC	AGGAATGTGC	GGAGTTCGG	AGCTTTTGCG	GCACAGATGC	GTCAGTGCCC	
DL.P.IRE	AGTTTCGGAG	ATGTTCTGAC	AGGAATGTGC	GGAGTTCGG	AGCTTTTGCG	GCACAGATGC	GTCAGTGCCC	
40.P.PS	AGTTTCGGAG	ATGTTCTGAC	AGGAATGTGC	GGAGTTCGG	AGCTTTTGCG	GCACAGATGC	GTCAGTGCCC	
40.P.IRE	AGTTTCGGAG	ATGTTCTGAC	AGGAATGTGC	GGAGTTCGG	AGCTTTTGCG	GCACAGATGC	GTCAGTGCCC	

		6542	6552	6562	6572	6582	6592	6602
0.P.	TGGACGTCTT	CTACACACTC	ATGCATGAGA	AACCTGACAG	CAGGGCAATG	AGAATGGCGG	AGAGAGACGC	
DL.P.IRE	TGGACGTCTT	CTACACACTC	ATGCATGAGA	AACCTGACAG	CAGGGCAATG	AGAATGGCGG	AGAGAGACGC	
40.P.PS	TGGACGTCTT	CTACACACTC	ATGCATGAGA	AACCTGACAG	CAGGGCAATG	AGAATGGCGG	AGAGAGACGC	
40.P.IRE	TGGACGTCTT	CTACACACTC	ATGCATGAGA	AACCTGACAG	CAGGGCAATG	AGAATGGCGG	AGAGAGACGC	

		6612	6622	6632	6642	6652	6662	6672
0.P.	CCCAGAGGCC	TTTCTGACCA	TGGTTGAGAT	GATGGTGTG	GGTTTGCAA	CCCTGGGTGT	CATCTGGTGC	
DL.P.IRE	CCCAGAGGCC	TTTCTGACCA	TGGTTGAGAT	GATGGTGTG	GGTTTGCAA	CCCTGGGTGT	CATCTGGTGC	
40.P.PS	CCCAGAGGCC	TTTCTGACCA	TGGTTGAGAT	GATGGTGTG	GGTTTGCAA	CCCTGGGTGT	CATCTGGTGC	
40.P.IRE	CCCAGAGGCC	TTTCTGACCA	TGGTTGAGAT	GATGGTGTG	GGTTTGCAA	CCCTGGGTGT	CATCTGGTGC	

		6682	6692	6702	6712	6722	6732	6742
0.P.	TTCGTCGTCC	GGACTTCAAT	CAGCCGTATG	ATGCTGGGCA	CGCTGGTCCT	GCTGGCCTCT	TTGCTACTCT	
DL.P.IRE	TTCGTCGTCC	GGACTTCAAT	CAGCCGTATG	ATGCTGGGCA	CGCTGGTCCT	GCTGGCCTCT	TTGCTACTCT	
40.P.PS	TTCGTCGTCC	GGACTTCAAT	CAGCCGTATG	ATGCTGGGCA	CGCTGGTCCT	GCTGGCCTCT	TTGCTACTCT	
40.P.IRE	TTCGTCGTCC	GGACTTCAAT	CAGCCGTATG	ATGCTGGGCA	CGCTGGTCCT	GCTGGCCTCT	TTGCTACTCT	

		6752	6762	6772	6782	6792	6802	6812
0.P.	TGTGGGCAGG	TGGCGTCGGC	TATGGGAATA	TGGCCGAGT	GGCCCTCATC	TTCTACACGT	TGCTGACGGT	
DL.P.IRE	TGTGGGCAGG	TGGCGTCGGC	TATGGGAATA	TGGCCGAGT	GGCCCTCATC	TTCTACACGT	TGCTGACGGT	
40.P.PS	TGTGGGCAGG	TGGCGTCGGC	TATGGGAATA	TGGCCGAGT	GGCCCTCATC	TTCTACACGT	TGCTGACGGT	
40.P.IRE	TGTGGGCAGG	TGGCGTCGGC	TATGGGAATA	TGGCCGAGT	GGCCCTCATC	TTCTACACGT	TGCTGACGGT	

		6822	6832	6842	6852	6862	6872	6882
0.P.	GCTGCAGCCC	GAGGCGGAA	AACAGAGAAG	CAGTGATGAC	AACAAACTGG	CATATTTCTT	GCTGACGCTC	
DL.P.IRE	GCTGCAGCCC	GAGGCGGAA	AACAGAGAAG	CAGTGATGAC	AACAAACTGG	CATATTTCTT	GCTGACGCTC	
40.P.PS	GCTGCAGCCC	GAGGCGGAA	AACAGAGAAG	CAGTGATGAC	AACAAACTGG	CATATTTCTT	GCTGACGCTC	
40.P.IRE	GCTGCAGCCC	GAGGCGGAA	AACAGAGAAG	CAGTGATGAC	AACAAACTGG	CATATTTCTT	GCTGACGCTC	

		6892	6902	
0.P.	TGCAGCCTTG	CTGGACTGGT	TGCAGCC	
DL.P.IRE	TGCAGCCTTG	CTGGACTGGT	TGCAGCC	
40.P.PS	TGCAGCCTTG	CTGGACTGGT	TGCAGCC	
40.P.IRE	TGCAGCCTTG	CTGGACTGGT	TGCAGCC	

Protein NS4B:

		6919	6929	6939	6949	6959	6969	6979
0.P.	AATGAGATGG	GTTTTCTGGA	GAAGACCAAG	GCAGACTTGT	CCACGGTGCT	GTGGAGTGAA	CGGGAGGAAC	
DL.P.IRE	AATGAGATGG	GTTTTCTGGA	GAAGACCAAG	GCAGACTTGT	CCACGGTGCT	GTGGAGTGAA	CGGGAGGAAC	
40.P.PS	AATGAGATGG	GTTTTCTGGA	GAAGACCAAG	GCAGACTTGT	CCACGGTGCT	GTGGAGTGAA	CGGGAGGAAC	
40.P.IRE	AATGAGATGG	GTTTTCTGGA	GAAGACCAAG	GCAGACTTGT	CCACGGTGCT	GTGGAGTGAA	CGGGAGGAAC	
		6989	6999	7009	7019	7029	7039	7049
0.P.	CCCGGCCATG	GAGTGAATGG	ACGAATGTGG	ACATCCAGCC	AGCGAGGTCC	TGGGGGACCT	ATGTGCTGGT	
DL.P.IRE	CCCGGCCATG	GAGTGAATGG	ACGAATGTGG	ACATCCAGCC	AGCGAGGTCC	TGGGGGACCT	ATGTGCTGGT	
40.P.PS	CCCGGCCATG	GAGTGAATGG	ACGAATGTGG	ACATCCAGCC	AGCGAGGTCC	TGGGGGACCT	ATGTGCTGGT	
40.P.IRE	CCCGGCCATG	GAGTGAATGG	ACGAATGTGG	ACATCCAGCC	AGCGAGGTCC	TGGGGGACCT	ATGTGCTGGT	
		7059	7069	7079	7089	7099	7109	7119
0.P.	GGTGCTCTCG	TTACACCTT	ACATCATCCA	CCAAGTGCAG	ACCAAAATCC	AACAACCTGT	CAACAGTGCC	
DL.P.IRE	GGTGCTCTCG	TTACACCTT	ACATCATCCA	CCAAGTGCAG	ACCAAAATCC	AACAACCTGT	CAACAGTGCC	
40.P.PS	GGTGCTCTCG	TTACACCTT	ACATCATCCA	CCAAGTGCAG	ACCAAAATCC	AACAACCTGT	CAACAGTGCC	
40.P.IRE	GGTGCTCTCG	TTACACCTT	ACATCATCCA	CCAAGTGCAG	ACCAAAATCC	AACAACCTGT	CAACAGTGCC	
		7129	7139	7149	7159	7169	7179	7189
0.P.	GTGGCATCTG	GTGCACAGGC	CATGAGAGAC	CTTGGGGGAG	GTGCCCCCTT	CTTTGGTGTG	GCGGGACATG	
DL.P.IRE	GTGGCATCTG	GTGCACAGGC	CATGAGAGAC	CTTGGGGGAG	GTGCCCCCTT	CTTTGGTGTG	GCGGGACATG	
40.P.PS	GTGGCATCTG	GTGCACAGGC	CATGAGAGAT	CTTGGGGGAG	GTGCCCCCTT	CATTGGTGTG	GCGGGACATG	
40.P.IRE	GTGGCATCTG	GTGCACAGGC	CATGAGAGAC	CTTGGGGGAG	GTGCCCCCTT	CTTTGGTGTG	GCGGGACATG	
		7199	7209	7219	7229	7239	7249	7259
0.P.	TCATGACCCT	CGGGGTGGTG	TCACTGATTG	GAGCTACTCC	CACCTCATTG	ATGGTGGGCG	TTGGCCTGGC	
DL.P.IRE	TCATGACCCT	CGGGGTGGTG	TCACTGATTG	GAGCTACTCC	CACCTCATTG	ATGGTGGGCG	TTGGCCTGGC	
40.P.PS	TCATGACCCT	CGGGGTGGTG	TCACTGATTG	GAGCTACTCC	CACCTCATTG	ATGGTGGGCG	TTGGCCTGGC	
40.P.IRE	TCATGACCCT	CGGGGTGGTG	TCACTGATTG	GAGCTACTCC	CACCTCATTG	ATGGTGGGCG	TTGGCCTGGC	
		7269	7279	7289	7299	7309	7319	7329
0.P.	GGCACTCCAT	CTGGCCATTG	TGGTGTCTGG	TCTGGAGGCT	GAATTGACAC	AGAGAGCTCA	TAAGTCTTT	
DL.P.IRE	GGCACTCCAT	CTGGCCATTG	TGGTGTCTGG	TCTGGAGGCT	GAATTGACAC	AGAGAGCTCA	TAAGTCTTT	
40.P.PS	GGCACTCCAT	CTGGCCATTG	TGGTGTCTGG	TCTGGAGGCT	GAATTGACAC	AGAGAGCTCA	TAAGTCTTT	
40.P.IRE	GGCACTCCAT	CTGGCCATTG	TGGTGTCTGG	TCTGGAGGCT	GAATTGACAC	AGAGAGCTCA	TAAGTCTTT	
		7339	7349	7359	7369	7379	7389	7399
0.P.	TTCTCTGCAA	TGGTGCGCAA	CCCCATGGTG	GATGGGGATG	TCATCAACCC	ATTCGGGGAG	GGGGAGGCAA	
DL.P.IRE	TTCTCTGCAA	TGGTGCGCAA	CCCCATGGTG	GATGGGGATG	TCATCAACCC	ATTCGGGGAG	GGGGAGGCAA	
40.P.PS	TTCTCTGCAA	TGGTGCGCAA	CCCCATGGTG	GATGGGGATG	TCATCAACCC	ATTCGGGGAG	GGGGAGGCAA	
40.P.IRE	TTCTCTGCAA	TGGTGCGCAA	CCCCATGGTG	GATGGGGATG	TCATCAACCC	ATTCGGGGAG	GGGGAGGCAA	
		7409	7419	7429	7439	7449	7459	7469
0.P.	AACCTGCTCT	ATATGAAAGG	AGAATGAGTC	TGGTGCTGGC	CATAGTGTG	TGCCTCATGT	CGGTGGTCAT	
DL.P.IRE	AACCTGCTCT	ATATGAAAGG	AGAATGAGTC	TGGTGCTGGC	CATAGTGTG	TGCCTCATGT	CGGTGGTCAT	
40.P.PS	AACCTGCTCT	ATATGAAAGG	AGAATGAGTC	TGGTGCTGGC	CATAGTGTG	TGCCTCATGT	CGGTGGTCAT	
40.P.IRE	AACCTGCTCT	ATATGAAAGG	AGAATGAGTC	TGGTGCTGGC	CATAGTGTG	TGCCTCATGT	CGGTGGTCAT	

	7479	7489	7499	7509	7519	7529	7539
0.P.	GAACCGAACG	GTGGCCTCCA	TAACAGAGGC	TTCAGCTGTG	GGACTGGCAG	CAGCGGGACA	GCTGCTCAGG
DL.P.IRE	GAACCGAACG	GTGGCCTCCA	TAACAGAGGC	TTCAGCTGTG	GGACTGGCAG	CAGCGGGACA	GCTGCTCAGG
40.P.PS	GAACCGAACG	GTGGCCTCCA	TAACAGAGGC	TTCAGCTGTG	GGACTGGCAG	CAGCGGGACA	GCTGCTCAGG
40.P.IRE	GAACCGAACG	GTGGCCTCCA	TAACAGAGGC	TTCAGCTGTG	GGACTGGCAG	CAGCGGGACA	GCTGCTCAGG

	7549	7559	7569	7579	7589	7599	7609
0.P.	CCGGAGGCTG	ACACACTGTG	GACGATGCCG	GTTGCTTGTG	GCATGAGTGG	TGTGGTCAGG	GGTAGCCTGT
DL.P.IRE	CCGGAGGCTG	ACACACTGTG	GACGATGCCG	GTTGCTTGTG	GCATGAGTGG	TGTGGTCAGG	GGTAGCCTGT
40.P.PS	CCGGAGGCTG	ACACACTGTG	GACGATGCCG	GTTGCTTGTG	GCATGAGTGG	TGTGGTCAGG	GGTAGCCTGT
40.P.IRE	CCGGAGGCTG	ACACACTGTG	GACGATGCCG	GTTGCTTGTG	GCATGAGTGG	TGTGGTCAGG	GGTAGCCTGT

	7619	7629	7639	7649	7659	
0.P.	GGGGGTTTCT	GCCTCTTGGG	CATAGACTCT	GGCTTCGAGC	TTCTGGGGGT	AGGCGT
DL.P.IRE	GGGGGTTTCT	GCCTCTTGGG	CATAGACTCT	GGCTTCGAGC	TTCTGGGGGT	AGGCGT
40.P.PS	GGGGGTTTCT	GCCTCTTGGG	CATAGACTCT	GGCTTCGAGC	TTCTGGGGGT	AGGCGT
40.P.IRE	GGGGGTTTCT	GCCTCTTGGG	CATAGACTCT	GGCTTCGAGC	TTCTGGGGGT	AGGCGT

Protein NS5:

	7675	7685	7695	7705	7715	7725	7735
0.P.	GGTGGTTCG	AGGGAGACAC	GCTTGGAGAT	CTCTGAAAC	GGAGGCTGAA	CAACTGCACC	AGGGAGGAAT
DL.P.IRE	GGTGGTTCG	AGGGAGACAC	GCTTGGAGAT	CTCTGAAAC	GGAGGCTGAA	CAACTGCACC	AGGGAGGAAT
40.P.PS	GGTGGTTCG	AGGGAGACAC	GCTTGGAGAT	CTCTGAAAC	GGAGGCTGAA	CAACTGCACC	AGGGAGGAAT
40.P.IRE	GGTGGTTCG	AGGGAGACAC	GCTTGGAGAT	CTCTGAAAC	GGAGGCTGAA	CAACTGCACC	AGGGAGGAAT

	7745	7755	7765	7775	7785	7795	7805
0.P.	TCTTTGTGTA	CAGGCGCACT	GGCATCCTGG	AGACGGAACG	TGACAAGGCT	AGAGAGTTGC	TCAGAAGAGG
DL.P.IRE	TCTTTGTGTA	CAGGCGCACT	GGCATCCTGG	AGACGGAACG	TGACAAGGCT	AGAGAGTTGC	TCAGAAGAGG
40.P.PS	TCTTTGTGTA	CAGGCGCACT	GGCATCCTGG	AGACGGAACG	TGACAAGGCT	AGAGAGTTGC	TCAGAAGAGG
40.P.IRE	TCTTTGTGTA	CAGGCGCACT	GGCATCCTGG	AGACGGAACG	TGACAAGGCT	AGAGAGTTGC	TCAGAAGAGG

	7815	7825	7835	7845	7855	7865	7875
0.P.	AGAGACCAAT	ATGGGATTGG	CTGTCTCTCG	GGGCACGGCA	AAGCTTGCCT	GGCTCGAGGA	ACGCGGATAT
DL.P.IRE	AGAGACCAAT	ATGGGATTGG	CTGTCTCTCG	GGGCACGGCA	AAGCTTGCCT	GGCTCGAGGA	ACGCGGATAT
40.P.PS	AGAGACCAAT	ATGGGATTGG	CTGTCTCTCG	GGGCACGGCA	AAGCTTGCCT	GGCTCGAGGA	ACGCGGATAT
40.P.IRE	AGAGACCAAT	ATGGGATTGG	CTGTCTCTCG	GGGCACGGCA	AAGCTTGCCT	GGCTCGAGGA	ACGCGGATAT

	7885	7895	7905	7915	7925	7935	7945
0.P.	GCCACCCTCA	AGGGAGAGGT	GGTAGATCTT	GGATGTGGAA	GGGGCGGCTG	GTCCTATTAT	GCGGCATCCC
DL.P.IRE	GCCACCCTCA	AGGGAGAGGT	GGTAGATCTT	GGATGTGGAA	GGGGCGGCTG	GTCCTATTAT	GCGGCATCCC
40.P.PS	GCCACCCTCA	AGGGAGAGGT	GGTAGATCTT	GGATGTGGAA	GGGGCGGCTG	GTCCTATTAT	GCGGCATCCC
40.P.IRE	GCCACCCTCA	AGGGAGAGGT	GGTAGATCTT	GGATGTGGAA	GGGGCGGCTG	GTCCTATTAT	GCGGCATCCC

	7955	7965	7975	7985	7995	8005	8015
0.P.	GACCGGCAGT	CATGAGTGTC	AGGGCATAACA	CCATTGGTGG	AAGAGGGCAT	GAGGCTCCAA	AGATGGTAAC
DL.P.IRE	GACCGGCAGT	CATGAGTGTC	AGGGCATAACA	CCATTGGTGG	AAGAGGGCAT	GAGGCTCCAA	AGATGGTAAC
40.P.PS	GACCGGCAGT	CATGAGTGTC	AGGGCATAACA	CCATTGGTGG	AAGAGGGCAT	GAAGCTCCAA	AGATGGTAAC
40.P.IRE	GACCGGCAGT	CATGAGTGTC	AGGGCATAACA	CCATTGGTGG	AAGAGGGCAT	GAAGCTCCAA	AGATGGTAAC

		8025	8035	8045	8055	8065	8075	8085
0.P.	AAGCCTGGGT	TGGAACCTGA	TCAAATTTAG	ATCAGGAATG	GACGTGTTCA	GCATGCAGCC	ACACCGGGCT	
DL.P.IRE	AAGCCTGGGT	TGGAACCTGA	TCAAATTTAG	ATCAGGAATG	GACGTGTTCA	GCATGCAGCC	ACACCGGGCT	
40.P.PS	AAGCCTGGGT	TGGAACCTGA	TCAAATTTAG	ATCAGGAATG	GACGTGTTCA	GCATGCAGCC	ACACCGGGCT	
40.P.IRE	AAGCCTGGGT	TGGAACCTGA	TCAAATTTAG	ATCAGGAATG	GACGTGTTCA	GCATGCAGCC	ACACCGGGCT	
		8095	8105	8115	8125	8135	8145	8155
0.P.	GACACTGTCA	TGTGTGACAT	CGGAGAGAGC	AGCCCAGATG	CCGCTGTGGA	GGGTGAGAGG	ACAAGGAAAG	
DL.P.IRE	GACACTGTCA	TGTGTGACAT	CGGAGAGAGC	AGCCCAGATG	CCGCTGTGGA	GGGTGAGAGG	ACAAGGAAAG	
40.P.PS	GACACTGTCA	TGTGTGACAT	CGGAGAGAGC	AGCCCAGATG	CCGCTGTGGA	GGGTGAGAGG	ACAAGGAAAG	
40.P.IRE	GACACTGTCA	TGTGTGACAT	CGGAGAGAGC	AGCCCAGATG	CCGCTGTGGA	GGGTGAGAGG	ACAAGGAAAG	
		8165	8175	8185	8195	8205	8215	8225
0.P.	TGATATTGCT	CATGGAGCAA	TGGAAAAACA	GGAACCCAC	GGCTGCCTGT	GTGTTCAAGG	TGCTGGCCCC	
DL.P.IRE	TGATATTGCT	CATGGAGCAA	TGGAAAAACA	GGAACCCACC	GGCTGCCTGT	GTGTTCAAGG	TGCTGGCCCC	
40.P.PS	TGATATTGCT	CATGGAGCAA	TGGAAAAACA	GGAACCCAC	GGCTGCCTGT	GTGTTCAAGG	TGCTGGCCCC	
40.P.IRE	TGATATTGCT	CATGGAGCAA	TGGAAAAACA	GGAACCCAC	GGCTGCCTGT	GTGTTCAAGG	TGCTGGCCCC	
		8235	8245	8255	8265	8275	8285	8295
0.P.	ATACCGCCCA	GAAGTGATAG	AAGCACTGCA	CAGATTCCAA	CTGCAATGGG	GGGGGGTCT	GGTGAGGACC	
DL.P.IRE	ATACCGCCCA	GAAGTGATAG	AAGCACTGCA	CAGATTCCAA	CTGCAATGGG	GGGGGGTCT	GGTGAGGACC	
40.P.PS	ATACCGCCCA	GAAGTGATAG	AAGCACTGCA	CAGATTCCAA	CTGCAATGGG	GGGGGGTCT	GGTGAGGACC	
40.P.IRE	ATACCGCCCA	GAAGTGATAG	AAGCACTGCA	CAGATTCCAA	CTGCAATGGG	GGGGGGTCT	GGTGAGGACC	
		8305	8315	8325	8335	8345	8355	8365
0.P.	CCTTTTCAA	GGAACTCCAC	CCATGAGATG	TATTACTCAA	CAGCTGTCAC	TGGGAACATA	GTGAACTCCG	
DL.P.IRE	CCTTTTCAA	GGAACTCCAC	CCATGAGATG	TATTACTCAA	CAGCTGTCAC	TGGGAACATA	GTGAACTCCG	
40.P.PS	CCTTTTCAA	GGAACTCCAC	CCATGAGATG	TATTACTCAA	CAGCTGTCAC	TGGGAACATA	GTGAACTCCG	
40.P.IRE	CCTTTTCAA	GGAACTCCAC	CCATGAGATG	TATTACTCAA	CAGCTGTCAC	TGGGAACATA	GTGAACTCCG	
		8375	8385	8395	8405	8415	8425	8435
0.P.	TCAATGTACA	GTCGAGGAAA	CTTTTGGCTC	GGTTTGGAGA	CCAGAGAGGG	CCAACCAGGG	TGCCTGAACT	
DL.P.IRE	TCAATGTACA	GTCGAGGAAA	CTTTTGGCTC	GGTTTGGAGA	CCAGAGAGGG	CCAACCAGGG	TGCCTGAACT	
40.P.PS	TCAATGTACA	GTCGAGGAAA	CTTTTGGCTC	GGTTTGGAGA	CCAGAGAGGG	CCAACCAGGG	TGCCTGAACT	
40.P.IRE	TCAATGTACA	GTCGAGGAAA	CTTTTGGCTC	GGTTTGGAGA	CCAGAGAGGG	CCAACCAGGG	TGCCTGAACT	
		8445	8455	8465	8475	8485	8495	8505
0.P.	TGACCTGGGA	GTTGGAACGA	GGTGTGTGGT	CTTAGCTGAG	GACAAGGTGA	AAGAACAAGA	TGTACAAGAG	
DL.P.IRE	TGACCTGGGA	GTTGGAACGA	GGTGTGTGGT	CTTAGCTGAG	GACAAGGTGA	AAGAACAAGA	TGTACAAGAG	
40.P.PS	TGACCTGGGA	GTTGGAACGA	GGTGTGTGGT	CTTAGCTGAG	GACAAGGTGA	AAGAACAAGA	TGTACAAGAG	
40.P.IRE	TGACCTGGGA	GTTGGAACGA	GGTGTGTGGT	CTTAGCTGAG	GACAAGGTGA	AAGAACAAGA	TGTACAAGAG	
		8515	8525	8535	8545	8555	8565	8575
0.P.	AGGATCAAAG	CGTTGCGGGA	GCAATACAGC	GAAACCTGGC	ACATGGACGA	GGAACATCCG	TACCGGACAT	
DL.P.IRE	AGGATCAAAG	CGTTGCGGGA	GCAATACAGC	GAAACCTGGC	ACATGGACGA	GGAACATCCG	TACCGGACAT	
40.P.PS	AGGATCAAAG	CGTTGCGGGA	GCAATACAGC	GAAACCTGGC	ACATGGACGA	GGAACATCCG	TACCGGACAT	
40.P.IRE	AGGATCAAAG	CGTTGCGGGA	GCAATACAGC	GAAACCTGGC	ACATGGACGA	GGAACATCCG	TACCGGACAT	

		8585	8595	8605	8615	8625	8635	8645
0.P.		GGCAGTATTG	GGGCAGCTAC	CGCACGGCAC	CAACCGGCTC	GGCGGCGTCA	CTGATCAATG	GGGTTGTGAA
DL.P.IRE		GGCAGTATTG	GGGCAGCTAC	CGCACGGCAC	CAACCGGCTC	GGCGGCGTCA	CTGATCAATG	GGGTTGTGAA
40.P.PS		GGCAGTATTG	GGGCAGCTAC	CGCACGGCAC	CAACCGGCTC	GGCGGCGTCA	CTGATCAATG	GGGTTGTGAA
40.P.IRE		GGCAGTATTG	GGGCAGCTAC	CGCACGGCAC	CAACCGGCTC	GGCGGCGTCA	CTGATCAATG	GGGTTGTGAA
		8655	8665	8675	8685	8695	8705	8715
0.P.		ACTTCTCAGC	TGGCCATGGA	ACGCACGGGA	AGATGTGGTG	CGCATGGCCA	TGACTGACAC	AACGGCTTTC
DL.P.IRE		ACTTCTCAGC	TGGCCATGGA	ACGCACGGGA	AGATGTGGTG	CGCATGGCCA	TGACTGACAC	AACGGCTTTC
40.P.PS		ACTTCTCAGC	TGGCCATGGA	ACGCACGGGA	AGATGTGGTG	CGCATGGCCA	TGACTGACAC	AACGGCTTTC
40.P.IRE		ACTTCTCAGC	TGGCCATGGA	ACGCACGGGA	AGATGTGGTG	CGCATGGCCA	TGACTGACAC	AACGGCTTTC
		8725	8735	8745	8755	8765	8775	8785
0.P.		GGACAGCAGA	GAGTGTTCAA	AGATAAAGTT	GACACAAAGG	CACAGGAACC	TCAGCCCGGC	ACAAGAGTCA
DL.P.IRE		GGACAGCAGA	GAGTGTTCAA	AGATAAAGTT	GACACAAAGG	CACAGGAACC	TCAGCCCGGC	ACAAGAGTCA
40.P.PS		GGACAGCAGA	GAGTGTTCAA	AGATAAAGTT	GACACAAAGG	CACAGGAACC	TCAGCCCGGC	ACAAGAGTCA
40.P.IRE		GGACAGCAGA	GAGTGTTCAA	AGATAAAGTT	GACACAAAGG	CACAGGAACC	TCAGCCCGGC	ACAAGAGTCA
		8795	8805	8815	8825	8835	8845	8855
0.P.		TCATGAGAGC	AGTAAATGAT	TGGATTTTGG	AACGACTGGC	GCAGAAAAGC	AAACCGCGTA	TGTGCAGCAG
DL.P.IRE		TCATGAGAGC	AGTAAATGAT	TGGATTTTGG	AACGACTGGC	GCAGAAAAGC	AAACCGCGTA	TGTGCAGCAG
40.P.PS		TCATGAGAGC	AGTAAATGAT	TGGATTTTGG	AACGACTGGC	GCAGAAAAGC	AAACCGCGTA	TGTGCAGCAG
40.P.IRE		TCATGAGAGC	AGTAAATGAT	TGGATTTTGG	AACGACTGGC	GCAGAAAAGC	AAACCGCGTA	TGTGCAGCAG
		8865	8875	8885	8895	8905	8915	8925
0.P.		AGAGGAATTC	ATAGCAAAAAG	TGAAATCAAA	CGCAGCCTTG	GGAGCTTGGT	CAGATGAGCA	AAACAGATGG
DL.P.IRE		AGAGGAATTC	ATAGCAAAAAG	TGAAATCAAA	CGCAGCCTTG	GGAGCTTGGT	CAGATGAGCA	AAACAGATGG
40.P.PS		AGAGGAATTC	ATAGCAAAAAG	TGAAATCAAA	CGCAGCCTTG	GGAGCTTGGT	CAGATGAGCA	AAACAGATGG
40.P.IRE		AGAGGAATTC	ATAGCAAAAAG	TGAAATCAAA	CGCAGCCTTG	GGAGCTTGGT	CAGATGAGCA	AAACAGATGG
		8935	8945	8955	8965	8975	8985	8995
0.P.		GCGAGTGCAA	GAGAGGCTGT	AGAGGATCCT	GCATTCTGGC	ACCTCGTGGA	TGAAGAGAGA	GAAAGGCACC
DL.P.IRE		GCGAGTGCAA	GAGAGGCTGT	AGAGGATCCT	GCATTCTGGC	ACCTCGTGGA	TGAAGAGAGA	GAAAGGCACC
40.P.PS		GCGAGTGCAA	GAGAGGCTGT	AGAGGATCCT	GCATTCTGGC	ACCTCGTGGA	TGAAGAGAGA	GAAAGGCACC
40.P.IRE		GCGAGTGCAA	GAGAGGCTGT	AGAGGATCCT	GCATTCTGGC	ACCTCGTGGA	TGAAGAGAGA	GAAAGGCACC
		9005	9015	9025	9035	9045	9055	9065
0.P.		TCATGGGGAG	ATGCGCGCAC	TGCGTGTACA	ACATGATGGG	CAAGAGAGAG	AAGAAACTGG	GAGAGTTCGG
DL.P.IRE		TCATGGGGAG	ATGCGCGCAC	TGCGTGTACA	ACATGATGGG	CAAGAGAGAG	AAGAAACTGG	GAGAGTTCGG
40.P.PS		TCATGGGGAG	ATGCGCGCAC	TGCGTGTACA	ACATGATGGG	CAAGAGAGAG	AAGAAACTGG	GAGAGTTCGG
40.P.IRE		TCATGGGGAG	ATGCGCGCAC	TGCGTGTACA	ACATGATGGG	CAAGAGAGAG	AAGAAACTGG	GAGAGTTCGG
		9075	9085	9095	9105	9115	9125	9135
0.P.		AGTGGCGAAG	GGAAGTCGGG	CCATTTGGTA	CATGTGGCTG	GGGAGTCGCT	TTCTGGAGTT	CGAGGCTCTT
DL.P.IRE		AGTGGCGAAG	GGAAGTCGGG	CCATTTGGTA	CATGTGGCTG	GGGAGTCGCT	TTCTGGAGTT	CGAGGCTCTT
40.P.PS		AGTGGCGAAG	GGAAGTCGGG	CCATTTGGTA	CATGTGGCTG	GGGAGTCGCT	TTCTGGAGTT	CGAGGCTCTT
40.P.IRE		AGTGGCGAAG	GGAAGTCGGG	CCATTTGGTA	CATGTGGCTG	GGGAGTCGCT	TTCTGGAGTT	CGAGGCTCTT

	9145	9155	9165	9175	9185	9195	9205
0.P.	GGATTCTTGA	ATGAAGACCA	TTGGGCCTCT	AGAGAGTCCA	GTGGAGCTGG	AGTTGAGGGA	ATAAGCTTGA
DL.P.IRE	GGATTCTTGA	ATGAAGACCA	TTGGGCCTCT	AGAGAGTCCA	GTGGAGCTGG	AGTTGAGGGA	ATAAGCTTGA
40.P.PS	GGATTCTTGA	ATGAAGACCA	TTGGGCCTCT	AGAGAGTCCA	GTGGAGCTGG	AGTTGAGGGA	ATAAGCTTGA
40.P.IRE	GGATTCTTGA	ATGAAGACCA	TTGGGCCTCT	AGAGAGTCCA	GTGGAGCTGG	AGTTGAGGGA	ATAAGCTTGA
	9215	9225	9235	9245	9255	9265	9275
0.P.	ACTACCTGGG	CTGGCACCTC	AAGAAGTTGT	CAACCCTGAA	TGGAGGACTC	TTCTATGCAG	ATGACACAGC
DL.P.IRE	ACTACCTGGG	CTGGCACCTC	AAGAAGTTGT	CAACCCTGAA	TGGAGGACTC	TTCTATGCAG	ATGACACAGC
40.P.PS	ACTACCTGGG	CTGGCACCTC	AAGAAGTTGT	CAACCCTGAA	TGGAGGACTC	TTCTATGCAG	ATGACACAGC
40.P.IRE	ACTACCTGGG	CTGGCACCTC	AAGAAGTTGT	CAACCCTGAA	TGGAGGACTC	TTCTATGCAG	ATGACACAGC
	9285	9295	9305	9315	9325	9335	9345
0.P.	TGGCTGGGAC	ACGAAAGTTA	CCAATGCAGA	CTTAGAGGAT	GAAGAACAGA	TCCTACGGTA	CATGGAGGGT
DL.P.IRE	TGGCTGGGAC	ACGAAAGTTA	CCAATGCAGA	CTTAGAGGAT	GAAGAACAGA	TCCTACGGTA	CATGGAGGGT
40.P.PS	TGGCTGGGAC	ACGAAAGTTA	CCAATGCAGA	CTTAGAGGAT	GAAGAACAGA	TCCTACGGTA	CATGGAGGGT
40.P.IRE	TGGCTGGGAC	ACGAAAGTTA	CCAATGCAGA	CTTAGAGGAT	GAAGAACAGA	TCCTACGGTA	CATGGAGGGT
	9355	9365	9375	9385	9395	9405	9415
0.P.	GAGCACAAAC	AACTGGCAAC	CACAATAATG	CAAAAAGCAT	ATCATGCCAA	AGTCGTGAAG	GTCGCGAGGC
DL.P.IRE	GAGCACAAAC	AACTGGCAAC	CACAATAATG	CAAAAAGCAT	ATCATGCCAA	AGTCGTGAAG	GTCGCGAGGC
40.P.PS	GAGCACAAAC	AACTGGCAAC	CACAATAATG	CAAAAAGCAT	ATCATGCCAA	AGTCGTGAAG	GTCGCGAGGC
40.P.IRE	GAGCACAAAC	AACTGGCAAC	CACAATAATG	CAAAAAGCAT	ATCATGCCAA	AGTCGTGAAG	GTCGCGAGGC
	9425	9435	9445	9455	9465	9475	9485
0.P.	CCTCCCGTGA	TGGAGGCTGC	ATCATGGATG	TCATCACAAG	AAGAGATCAA	AGAGGTTCCG	GCCAGGTTGT
DL.P.IRE	CCTCCCGTGA	TGGAGGCTGC	ATCATGGATG	TCATCACAAG	AAGAGATCAA	AGAGGTTCCG	GCCAGGTTGT
40.P.PS	CCTCCCGTGA	TGGAGGCTGC	ATCATGGATG	TCATCACAAG	AAGAGATCAA	AGAGGTTCCG	GCCAGGTTGT
40.P.IRE	CCTCCCGTGA	TGGAGGCTGC	ATCATGGATG	TCATCACAAG	AAGAGATCAA	AGAGGTTCCG	GCCAGGTTGT
	9495	9505	9515	9525	9535	9545	9555
0.P.	GACCTATGCC	CTTAACACCC	TCACCAACAT	AAAGGTGCAA	CTAATCCGAA	TGATGGAAGG	GGAAGGGGTC
DL.P.IRE	GACCTATGCC	CTTAACACCC	TCACCAACAT	AAAGGTGCAA	CTAATCCGAA	TGATGGAAGG	GGAAGGGGTC
40.P.PS	GACCTATGCC	CTTAACACCC	TCACCAACAT	AAAGGTGCAA	CTAATCCGAA	TGATGGAAGG	GGAAGGGGTC
40.P.IRE	GACCTATGCC	CTTAACACCC	TCACCAACAT	AAAGGTGCAA	CTAATCCGAA	TGATGGAAGG	GGAAGGGGTC
	9565	9575	9585	9595	9605	9615	9625
0.P.	ATAGAGGCAG	CGGATGCACA	CAACCCGAGA	CTGCTTCGAG	TGGAGCGCTG	GCTGAAAGAA	CATGGAGAAG
DL.P.IRE	ATAGAGGCAG	CGGATGCACA	CAACCCGAGA	CTGCTTCGAG	TGGAGCGCTG	GCTGAAAGAA	CATGGAGAAG
40.P.PS	ATAGAGGCAG	CGGATGCACA	CAACCCGAGA	CTGCTTCGAG	TGGAGCGCTG	GCTGAAAGAA	CATGGAGAAG
40.P.IRE	ATAGAGGCAG	CGGATGCACA	CAACCCGAGA	CTGCTTCGAG	TGGAGCGCTG	GCTGAAAGAA	CATGGAGAAG
	9635	9645	9655	9665	9675	9685	9695
0.P.	AGCGTCTTGG	AAGAATGCTC	GTCAGTGGTG	ACGATTGTGT	GGTGAGGCC	TTGGATGACA	GATTTGGCAA
DL.P.IRE	AGCGTCTTGG	AAGAATGCTC	GTCAGTGGTG	ACGATTGTGT	GGTGAGGCC	TTGGATGACA	GATTTGGCAA
40.P.PS	AGCGTCTTGG	AAGAATGCTC	GTCAGTGGTG	ACGATTGTGT	GGTGAGGCC	TTGGATGACA	GATTTGGCAA
40.P.IRE	AGCGTCTTGG	AAGAATGCTC	GTCAGTGGTG	ACGATTGTGT	GGTGAGGCC	TTGGATGACA	GATTTGGCAA

	9705	9715	9725	9735	9745	9755	9765
0.P.	AGCACTTTAC	TTTCTGAATG	ACATGGCCAA	GACCAGGAAG	GACATTGGGG	AATGGGAGCA	CTCGGCCGGC
DL.P.IRE	AGCACTTTAC	TTTCTGAATG	ACATGGCCAA	GACCAGGAAG	GACATTGGGG	AATGGGAGCA	CTCGGCCGGC
40.P.PS	AGCACTTTAC	TTTCTGAATG	ACATGGCCAA	GACCAGGAAG	GACATTGGGG	AATGGGAGCA	CTCGGCCGGC
40.P.IRE	AGCACTTTAC	TTTCTGAATG	ACATGGCCAA	GACCAGGAAG	GACATTGGGG	AATGGGAGCA	CTCGGCCGGC
	9775	9785	9795	9805	9815	9825	9835
0.P.	TTATCAAGCT	GGGAGGAGGT	CCCCTTTTGT	TCACACCACT	TCCACGAGCT	AGTGATGAAG	GACGGACGCA
DL.P.IRE	TTATCAAGCT	GGGAGGAGGT	CCCCTTTTGT	TCACACCACT	TCCACGAGCT	AGTGATGAAG	GACGGACGCA
40.P.PS	TTATCAAGCT	GGGAGGAGGT	CCCCTTTTGT	TCACACCACT	TCCACGAGCT	AGTGATGAAG	GACGGACGCA
40.P.IRE	TTATCAAGCT	GGGAGGAGGT	CCCCTTTTGT	TCACACCACT	TCCACGAGCT	AGTGATGAAG	GACGGACGCA
	9845	9855	9865	9875	9885	9895	9905
0.P.	CCTTGGTGGT	GCCGTGCAGA	GACCAAGATG	AACTCGTTGG	GAGGGCGCGC	ATCTCACCGG	GATGCGGCTG
DL.P.IRE	CCTTGGTGGT	GCCGTGCAGA	GACCAAGATG	AACTCGTTGG	GAGGGCGCGC	ATCTCACCGG	GATGCGGCTG
40.P.PS	CCTTGGTGGT	GCCGTGCAGA	GACCAAGATG	AACTCGTTGG	GAGGGCGCGC	ATCTCACCGG	GATGCGGCTG
40.P.IRE	CCTTGGTGGT	GCCGTGCAGA	GACCAAGATG	AACTCGTTGG	GAGGGCGCGC	ATCTCACCGG	GATGCGGCTG
	9915	9925	9935	9945	9955	9965	9975
0.P.	GAGTGTCGC	GAGACGGCCT	GCCTTTCAAA	AGCCTACGGG	CAGATGTGGC	TGCTGAGCTA	CTTCCATCGG
DL.P.IRE	GAGTGTCGC	GAGACGGCCT	GCCTTTCAAA	AGCCTACGGG	CAGATGTGGC	TGCTGAGCTA	CTTCCATCGG
40.P.PS	GAGTGTCGC	GAGACGGCCT	GCCTTTCAAA	AGCCTACGGG	CAGATGTGGC	TGCTGAGCTA	CTTCCATCGG
40.P.IRE	GAGTGTCGC	GAGACGGCCT	GCCTTTCAAA	AGCCTACGGG	CAGATGTGGC	TGCTGAGCTA	CTTCCATCGG
	9985	9995	10005	10015	10025	10035	10045
0.P.	CGAGACCTGA	GGACGCTCGG	GCTTGCCATC	AACTCAGCAG	TGCCTGTCGA	TTGGGTTCTT	ACCGGCCGCA
DL.P.IRE	CGAGACCTGA	GGACGCTCGG	GCTTGCCATC	AACTCAGCAG	TGCCTGTCGA	TTGGGTTCTT	ACCGGCCGCA
40.P.PS	CGAGACCTGA	GGACGCTCGG	GCTTGCCATC	AACTCAGCAG	TGCCTGTCGA	TTGGGTTCTT	ACCGGCCGCA
40.P.IRE	CGAGACCTGA	GGACGCTCGG	GCTTGCCATC	AACTCAGCAG	TGCCTGTCGA	TTGGGTTCTT	ACCGGCCGCA
	10055	10065	10075	10085	10095	10105	10115
0.P.	CGACATGGAG	CATCCATGCC	AGTGGGGCCT	GGATGACCAC	AGAAGACATG	CTGGACGTCT	GGAACCGGGT
DL.P.IRE	CGACATGGAG	CATCCATGCC	AGTGGGGCCT	GGATGACCAC	AGAAGACATG	CTGGACGTCT	GGAACCGGGT
40.P.PS	CGACATGGAG	CATCCATGCC	AGTGGGGCCT	GGATGACCAC	AGAAGACATG	CTGGACGTCT	GGAACCGGGT
40.P.IRE	CGACATGGAG	CATCCATGCC	AGTGGGGCCT	GGATGACCAC	AGAAGACATG	CTGGACGTCT	GGAACCGGGT
	10125	10135	10145	10155	10165	10175	10185
0.P.	GTGGATTTTG	GACAACCCTT	TCATGCAGAA	CAAGGGAAAAG	GTCATGGAGT	GGAGGGATGT	TCCGTACCTC
DL.P.IRE	GTGGATTTTG	GACAACCCTT	TCATGCAGAA	CAAGGGAAAAG	GTCATGGAGT	GGAGGGATGT	TCCGTACCTC
40.P.PS	GTGGATTTTG	GACAACCCTT	TCATGCAGAA	CAAGGGAAAAG	GTCATGGAGT	GGAGGGATGT	TCCGTACCTC
40.P.IRE	GTGGATTTTG	GACAACCCTT	TCATGCAGAA	CAAGGGAAAAG	GTCATGGAGT	GGAGGGATGT	TCCGTACCTC
	10195	10205	10215	10225	10235	10245	10255
0.P.	CCTAAAGCTC	AGGACATGTT	ATGTTCTCC	CTCGTTGGGA	GGAAAGAAAAG	AGCAGAATGG	GCCAAGAACA
DL.P.IRE	CCTAAAGCTC	AGGACATGTT	ATGTTCTCC	CTCGTTGGGA	GGAAAGAAAAG	AGCAGAATGG	GCCAAGAACA
40.P.PS	CCTAAAGCTC	AGGACATGTT	ATGTTCTCC	CTCGTTGGGA	GGAAAGAAAAG	AGCAGAATGG	GCCAAGAACA
40.P.IRE	CCTAAAGCTC	AGGACATGTT	ATGTTCTCC	CTCGTTGGGA	GGAAAGAAAAG	AGCAGAATGG	GCCAAGAACA

	10265	10275	10285	10295	10305	10315	10325
0.P.	TCTGGGGAGC	AGTGGAAAAG	GTGAGGAAGA	TGATAGGGCC	TGAAAAGTTC	AAGGACTATC	TCTCTGTAT
DL.P.IRE	TCTGGGGAGC	AGTGGAAAAG	GTGAGGAAGA	TGATAGGGCC	TGAAAAGTTC	AAGGACTATC	TCTCTGTAT
40.P.PS	TCTGGGGAGC	AGTGGAAAAG	GTGAGGAAGA	TGATAGGGCC	TGAAAAGTTC	AAGGACTATC	TCTCTGTAT
40.P.IRE	TCTGGGGAGC	AGTGGAAAAG	GTGAGGAAGA	TGATAGGGCC	TGAAAAGTTC	AAGGACTATC	TCTCTGTAT

	10335	10345	10355	10365	STOP kodon	
0.P.	GGACCGCCAT	GACCTGCACT	GGGAGCTCAG	ACTGGAGAGC	TCAATAATC	TAA
DL.P.IRE	GGACCGCCAT	GACCTGCACT	GGGAGCTCAG	ACTGGAGAGC	TCAATAATC	TAA
40.P.PS	GGACCGCCAT	GACCTGCACT	GGGAGCTCAG	ACTGGAGAGC	TCAATAATC	TAA
40.P.IRE	GGACCGCCAT	GACCTGCACT	GGGAGCTCAG	ACTGGAGAGC	TCAATAATC	TAA

3' koncová oblast:

	10387	10397	10407	10417	10427	10437	10447
0.P.	ACCCAGACTG	TGACAGAGCA	AAACCCGGAG	GGCTCGTAAA	AGATTGTCCG	GAACCAAAAAG	AAAAGCAAGC
DL.P.IRE	ACCCAGACTG	TGACAGAGCA	AAACCCGGAG	GGCTCGTAAA	AGATTGTCCG	GAACCAAAAAG	AAAAGCAAGC
40.P.PS	ACCCAGACTG	TGACAGAGCA	AAACCCGGAG	GGCTCGTAAA	AGATTGTCCG	GAACCAAAAAG	AAAAGCAAGC
40.P.IRE	ACCCAGACTG	TGACAGAGCA	AAACCCGGAG	GGCTCGTAAA	AGATTGTCCG	GAACCAAAAAG	AAAAGCAAGC

	10457	10467	10477	10487	10497	10507	10517
0.P.	AACTCACAGA	GTGTGAGACC	CCCCTGACCA	GCAAAGGGGC	AGATCGGTCA	GGGGTGAGGG	ATGCCCCCAG
DL.P.IRE	AACTCACAGA	GTGTGAGACC	CCCCTGACCA	GCAAAGGGGC	AGATCGGTCA	GGGGTGAGGG	ATGCCCCCAG
40.P.PS	AACTCACAGA	GTGTGAGACC	CCCCTGACCA	GCAAAGGGGC	AGATCGGTCA	GGGGTGAGGG	ATGCCCCCAG
40.P.IRE	AACTCACAGA	GTGTGAGACC	CCCCTGACCA	GCAAAGGGGC	AGATCGGTCA	GGGGTGAGGG	ATGCCCCCAG

	10527	10537	10547	10557	10567	10577	10587
0.P.	AGTGCATTAC	GGCAGCACGC	CAGTGAGAGT	GGCGACGGGA	AAATGGTCGA	TCCCGACGTA	GGGCACTCTG
DL.P.IRE	AGTGCATTAC	GGCAGCACGC	CAGTGAGAGT	GGCGACGGGA	AAATGGTCGA	TCCCGACGTA	GGGCACTCTG
40.P.PS	AGTGCATTAC	GGCAGCACGC	CAGTGAGAGT	GGCGACGGGA	AAATGGTCGA	TCCCGACGTA	GGGCACTCTG
40.P.IRE	AGTGCATTAC	GGCAGCACGC	CAGTGAGAGT	GGCGACGGGA	AAATGGTCGA	TCCCGACGTA	GGGCACTCTG

	10597	10607	10617	10627	10637	10647	10657
0.P.	AAAAACTTTG	TGAGACCCCC	TGCATCATGA	TAAGGCCGAA	CATGGTGCAT	GAAAAGGGAG	GCCCCCGGAA
DL.P.IRE	AAAAACTTTG	TGAGACCCCC	TGCATCATGA	TAAGGCCGAA	CATGGTGCAT	GAAAAGGGAG	GCCCCCGGAA
40.P.PS	AAAAACTTTG	TGAGACCCCC	TGCATCATGA	TAAGGCCGAA	CATGGTGCAT	GAAAAGGGAG	GCCCCCGGAA
40.P.IRE	AAAAACTTTG	TGAGACCCCC	TGCATCATGA	TAAGGCCGAA	CATGGTGCAT	GAAAAGGGAG	GCCCCCGGAA

	10667	10677	10687	10697	10707	10717	10727
0.P.	GTACGCTTCC	GGGAGGAGGG	AAGAGAGAAA	TTGGCAGCTC	TCTTCAGGAT	TTTTCCTCCT	CCTATACAAA
DL.P.IRE	GTACGCTTCC	GGGAGGAGGG	AAGAGAGAAA	TTGGCAGCTC	TCTTCAGGAT	TTTTCCTCCT	CCTATACAAA
40.P.PS	G <u>C</u> ACGCTTCC	GGGAGGAGGG	AAGAGAGAAA	TTGGCAGCTC	TCTTCAGGAT	TTTTCCTCCT	CCTATACAAA
40.P.IRE	G <u>C</u> ACGCTTCC	GGGAGGAGGG	AAGAGAGAAA	TTGGCAGCTC	TCTTCAGGAT	TTTTCCTCCT	CCTATACAAA

	10737	10747	10757	10767	10777	10787	10797
0.P.	ATCCCCCTC	GGTAGAGGGG	GGGCGGTICT	TGTTCTCCCT	GAGCCACCAT	CACCCAGACA	CAGATAGTCT
DL.P.IRE	ATCCCCCTC	GGTAGAGGGG	GGGCGGTICT	TGTTCTCCCT	GAGCCACCAT	CACCCAGACA	CAGATAGTCT
40.P.PS	ATCCCCCTC	GGTAGAGGGG	GGGCGGTICT	TGTTCTCCCT	GAGCCACCAT	CACCCAGACA	CAGATAGTCT
40.P.IRE	ATCCCCCTC	GGTAGAGGGG	GGGCGGTICT	TGTTCTCCCT	GAGCCACCAT	CACCCAGACA	CAGATAGTCT

	10807	10817	10827	
0.P.	GACAAGGAGG	TGATGTGTGA	CTCGGAAAAA	CACCCGCT
DL.P.IRE	GACAAGGAGG	TGATGTGTGA	CTCGGAAAAA	CACCCGCT
40.P.PS	GACAAGGAGG	TGATGTGTGA	CTCGGAAAAA	CACCCGCT
40.P.IRE	GACAAGGAGG	TGATGTGTGA	CTCGGAAAAA	CACCCGCT