

Svoluji k zapůjčení své diplomové práce ke studijním účelům a prosím, aby byla vedena přesná evidence vypůjčovateli. Převzaté údaje je vypůjčovateli povinen řádně ocitovat.

UNIVERZITA KARLOVA V PRAZE

PŘÍRODOVĚDECKÁ FAKULTA

Biologie

Genetika, molekulární biologie, virologie

Zaměření: Virologie



Bc. Dalibor Miklík

FUNKČNÍ ANALÝZA GENOMU POMOCÍ MAPOVÁNÍ INTEGRAČNÍCH MÍST PODPORUJÍCÍCH EXPRESI

RETROVIRŮ V LIDSKÝCH BUŇKÁCH

FUNCTIONAL GENOME ANALYSIS USING THE RETROVIRAL INTEGRATION SITES PERMISSIVE FOR PROVIRUS

EXPRESSION IN HUMAN CELLS

Diplomová práce

Vedoucí práce: RNDr. Jiří Hejnar, Csc

Praha, 2013

Prohlášení:

Prohlašuji, že jsem závěrečnou práci zpracoval samostatně a že jsem uvedl všechny použité informační zdroje a literaturu. Tato práce ani její podstatná část nebyla předložena k získání jiného nebo stejného akademického titulu.

V Praze, 01.05.2013

Podpis

PODĚKOVÁNÍ

Na tomto místě bych rád poděkoval RNDr. Jiřímu Hejnarovi, Csc za vedení této diplomové práce a vůbec za umožnění vypracovávat diplomovou práci v rámci výzkumného projektu Laboratoře virové a buněčné genetiky ÚMG AV ČR. Poděkování patří také Mgr. Filipu Šeniglovi, PhD za praktické vedení mé práce v laboratoři a inspirativní diskuze, které často posouvaly práci blíže vytyčeným cílům a pomáhaly vytvářet cíle nové. Díky však patří celému kolektivu Laboratoře virové a buněčné genetiky ÚMG AV ČR, a to nejen za shovívavost, trpělivost a podporu při mém praktickém působení v laboratoři, ale hlavně za udržování stále příjemné a přátelské atmosféry často posilující mou pracovní morálku a motivaci. V neposlední řadě bych chtěl zvláště poděkovat své rodině a nejbližším přátelům za plnou podporu při studiu i mimo něj.

Tato práce vznikla za finanční podpory grantu Grantové agentury České republiky GACR, GAP502/11/2207.

ABSTRAKT

Exprese genů retrovirů je závislá na vytvoření proviru – DNA kopii retrovirového genomu integrované do genomu hostitele. Stav transkripce proviru je pak závislý na podmínkách v místě integrace proviru. Existence umlčování exprese proviru je problémem při využívání retrovirových vektorů a překážkou pro eradikaci viru lidské imunodeficiency typu 1 (HIV-1) u infikovaných pacientů. S využitím buněčných klonů nesoucích jediný integrovaný provirus je v této diplomové práci zkoumána distribuce (epi)genomových charakteristik v místech integrace stabilně exprimovaných provirů. Celkově bylo získáno 245 a 255 expresních profilů klonů nesoucích proviry ptačího sarkomového a leukózního viru (ASLV) a HIV-1. Na databázích založená analýza 42 integračních míst ASLV a 3 integračních míst HIV-1 ukazuje, že udržení stabilní exprese provirů je závislé na integraci proviru do blízkosti transkripčních startů (TSS). Přítomnost histonových značek charakteristických pro okolí aktivních TSS a regulačních elementů v místech integrace stabilně exprimovaných provirů toto pozorování potvrzuje. Výsledky této práce mohou soužit jako základ dalších analýz vztahu integračních míst retrovirů a jejich exprese a měly by být brány na vědomí při dalších analýzách exprese retrovirů.

Klíčová slova: retrovirus, retrovirové vektory, ASLV, HIV-1, integrace, umlčování proviru, lidský genom, epigenetika, histonové modifikace

ABSTRACT

The expression of retroviral genes depends on the establishment of the provirus – the DNA copy of retroviral genome integrated into the host genome. The transcriptional state of provirus is then influenced by the environment at the site of integration. The phenomenon of proviral silencing is an obstacle to the usage of retroviral vectors and a barrier to the eradication of human immunodeficiency virus type 1 (HIV-1) from infected individuals. Taking advantage of single cell clones bearing one provirus, this diploma thesis investigates the distribution of (epi)genomic features at the sites occupied by stably expressed proviruses. In total, long-term expression profiles of 245 and 255 clones carrying avian sarcoma-leucosis virus (ASLV) and HIV-1, respectively, were obtained. The database-based analysis of 42 integration sites of ASLV and three integration sites of HIV-1 proviruses shows that proviral stable expression highly correlates with the transcriptional start sites (TSS) at the sites of integration. Histone marks characteristic for the proximity of active TSSs and regulatory elements at the sites of integration of stably expressed proviruses confirm this finding. The results presented in this thesis could inspire other analyses investigating the relationship between the integration site and the expression of retroviruses and should be noted by future analyses of retroviral expression.

Key words: retrovirus, retroviral vectors, ASLV, HIV-1, integration, proviral silencing, human genome, epigenetics, histone modifications

OBSAH

Seznam zkratk	11
1. Úvod	13
2. Cíle práce	15
3. Přehled literatury	16
3.1. Retroviry a jejich replikační cyklus	16
3.1.1. Systém a popis retrovirů	16
3.1.2. Replikační cyklus retrovirů	17
3.2. Integrace retrovirů	18
3.2.1. Mechanismus integrace	18
3.2.2. Integrace v kontextu hostitelského genomu	18
3.2.3. Integrační preference retrovirů	20
3.2.4. Buněčné faktory a cílení integrace	21
3.3. Exprese retrovirových genů	24
3.3.1. Buněčné faktory a umlčování retrovirové transkripce	25
3.3.2. Vliv místa integrace na expresi retrovirů	30
3.4. Využití retrovirů – retrovirové vektory	33
3.4.1. Modifikace retrovirových vektorů	34
3.4.2. Distribuce histonových modifikací a transkripční aktivita genomu	35
4. Materiál a metody	38
4.1. Pokusný materiál	38
4.1.1. Buněčné linie – specifika a kultivace	38
4.1.2. Retrovirové vektory	39
4.1.3. Plazmidy	40
4.1.4. Roztoky	40
4.2. Strategie a organizace pokusů	41

4.3. Produkce retrovirových vektorů.....	42
4.3.1. Produkce vektoru AG3.....	42
4.3.2. Produkce vektoru LTR-Tat-GFP.....	43
4.3.3. Sběr virových partikulí	43
4.3.4. Koncentrování viru.....	43
4.3.5. Titrace retrovirových vektorů	44
4.4. Infekce HEK293 a K562 retrovirovými vektory.....	44
4.5. Selektce a ustanovení buněčných klonů.....	45
4.5.1. Příprava buněk pro FACS	45
4.5.2. FACS	45
4.6. Stanovení expresních profilů buněčných klonů	45
4.6.1. Preselektce pomocí fluorescenční mikroskopie	45
4.6.2. Příprava buněk pro průtokovou cytometrii.....	46
4.6.3. Průtoková cytometrie	46
4.6.4. Vyhodnocení exprese klonů	46
4.7. Analýza GFP ⁻ subpopulací.....	46
4.8. Zamražení buněčných klonů.....	47
4.9. Izolace genomové DNA z buněk v tkáňové kultuře.....	47
4.9.1. Sběr buněk	47
4.9.2. Izolace genomové DNA.....	47
4.10. Splinkerette PCR	48
4.10.1. Štěpení genomové DNA.....	48
4.10.2. Příprava splinkerů	49
4.10.3. Ligace splinkerů a štěpené genomové DNA	49
4.10.4. Odstranění 3'LTR oblasti.....	49
4.10.5. Purifikace DNA před PCR reakcí.....	50

4.10.6. Nested PCR	50
4.11. Purifikace a sekvenace produktů splinkerette PCR.....	51
4.11.1. Agarózová elektroforéza.....	51
4.11.2. Purifikace DNA před sekvenační reakcí	51
4.11.3. Extrakce DNA z gelu	52
4.11.4. Sekvenování místa integrace	52
4.11.5. Purifikace produktu sekvenace.....	52
4.12. Aktivní a umlčená integrační místa	52
4.12.1. Izolace a purifikace DNA	52
4.12.2 Klonování do vektoru.....	53
4.12.3. Transformace bakterií.....	53
4.12.4. Výběr a sekvenace bakteriálních kolonií	53
4.12. Mapování sekvencí integračních míst do lidského genomu	53
4.12.1. Zpracování sekvence.....	53
4.12.2 Mapování sekvencí do lidského genomu	54
4.12.3 Lokalizace místa integrace.....	54
4.12.4. Integrační místa preference ASLV	54
4.13. Bioinformatická analýza integračních míst	54
4.13.1. Použitý software	54
4.13.2. Původ dat.....	55
4.13.3. Charakterizace místa integrace	56
4.14. RT-qPCR analýza cílených TU.....	57
4.14.1. Izolace RNA	57
4.14.2. Odstranění DNA.....	58
4.14.3. Vytvoření cDNA – RT PCR	58
4.14.4. qPCR.....	58

5. Výsledky	60
5.1. Exprese a integrace provirů ASLV v buňkách HEK293.....	60
5.2. Exprese provirů ASLV a HIV-1 v linii K562	63
5.3. Nestabilní umlčení provirů ASLV v lidských buňkách.....	68
5.4. Integrační místa stabilně exprimovaných provirů ASLV.....	70
5.5. Integrační místa nestabilně umlčovaných klonů ASLV.....	82
5.6. Integrační místa stabilně exprimovaných provirů HIV-1.....	84
6. Diskuze	87
7. Souhrn	94
8. Seznam použité literatury	95

SEZNAM ZKRATEK

AIDS	acquired immunodeficiency syndrome	syndrom získané imunodeficiencie
AP-1	activator protein 1	aktivátorový protein 1
ASLV	avian sarcoma leucosis virus	ptačí sarkomový a leukóзовý virus
ASV	avian sarcoma virus	virus způsobující sarkomy u ptáků
BAF	BRG1-associated factors	faktory asociované s BRG-1
BRG-1	Brahma-related gene 1	gen příbuzný Brahma
c-Myc	v-myc avian myelocytomatosis viral oncogene homolog	homolog ptačího virového onkogenu myelocytomatózy v-myc
CS	calf serum	telecí sérum
Daxx	death associated domain protein	protein obsahující doménu asociovanou se smrtí
DMEM	Dulbecco's modified Eagle's médium	Dulbeccovo modifikované Eaglovo médium
DNMT	de novo methyl transferase	de novo metyltransferáza
dNTP	deoxynucleosidetriphosphate	deoxynukleosidtrifosfáty
dpi	days post infection	dny po infekci
EGFP	enhanced green fluorescence protein	stabilizovaný zelený fluorescenční protein
FACS	fluorescence activated cell sorting	fluorescencí aktivovaná separace buněk
FCS	fetal calf serum	fetální telecí sérum
GFP	green fluorescence protein	zelený fluorescenční protein
H3K27ac	acetylation of lysine 27 at histon H3	acetylace lysinu 27 histonu H3
H3K27me3	trimethylation of lysine 27 at histon H3	trimetylace lysinu 27 histonu H3
H3K36me3	trimethylation of lysine 36 at histon H3	trimetylace lysinu 36 histonu H3
H3K4me1/2/3	mono-, di-, trimethylation of lysine 4 at histon H3	mono-, di-, trimetylace lysinu 4 histonu H3
H3K79me2	dimethylation of lysine 79 at histon H3	dimetylace lysinu 79 histonu H3
H3K9ac	acetylation of lysine 9 at histon H3	acetylace lysinu 9 histonu H3
H3K9me3	trimethylation of lysine 9 at histon H3	trimetylace lysinu 9 histonu H3
H4K20me1	monomethylation of lysine 20 at histon H4	monometylace lysinu 20 histonu H4
HDAC	histone deacetylase	histon deacetyláza
HIV-1	human immunodeficiency virus type 1	virus lidské získané imunodeficiencie typu 1
HRP2	histidine-rich protein 2	histidinem bohatý protein 2
HTLV-1	human T-cell leukemia virus	virus způsobující leukemii T-buněk

	1	u lidí
IN	integrase	integráza
interTU		mimo transkripční jednotku
intraTU		uvnitř transkripční jednotky
IRES	internal ribosome entry site	vnitřní místo pro nasedání ribozómů
kbp	kilobase pairs	tisíc párů bází
KS	chicken serum	kuřecí sérum
LEDGF/p75	lens epithelial-derived growth factor/protein 75	růstový faktor odvozený z epitelu čočky/protein 75
LMO2	LIM domain only 2	LIM doména 2
LTR	long terminal repetition	dlouhé koncové repetice
MFI	mean fluorescence intensity	průměrná intenzita fluorescence
MLV	murine leukemia virus	virus způsobující leukemii u myši
MMTV	mouse mammary tumor virus	virus způsobující tumor prsu u myši
MOI	multiplicity of infection	multiplicita infekce
NF-κB	nuclear factor κB	jaderný faktor κB
PBAF	Polybromo- and BRG1-associated factors	faktory asociované s BRG-1 a Polybromo
PBS	primer-binding site	místo pro vazbu primeru/tRNA
PIC	preintegration complex	preintegrační komplex
RSV	Rous sarcoma virus	Rousův virus způsobující sarkomy
RT	reverse transcriptase	reverzní transkriptáza
SIN	self-inactivated vectors	vektory s inaktivovaným TSS
Sp1	specificity protein 1	specifický protein 1
STAT1	signal transducers and activators 1	signální faktory a aktivátory 1
SWI/SNF	switch/sucrose nonfermentable	
Tat	trans-activator of transcription	transaktivátor transkripce
TBP	TATA box binding protein	TATA box vazebný protein
tps	weeks post sorting	týdny po separaci
TRIM28	tripartite motif containing 28	
TSS	transcription start sites	místo transkripčního startu
TU	transcription units	transkripční jednotka
VSV-G	G-protein of vesicular stomatitis virus	protein G viru vezikulární stomatitidy
X-SCID	X-bound severe combined immunodeficiency	X-vázaná kombinovaná imunodeficiencie
YY1	Yin Yang 1	jin jang 1
ZFP809	zinc finger protein 809	protein 28 se zinkovým prstem

1. ÚVOD

Retroviry představují jednu z nejvýznamnějších skupin živočišných virů jak v lidské, tak ve veterinární medicíně. Výjimečnost retrovirů spočívá v jejich replikačním cyklu. Retroviry nejdříve přepisují svůj genom tvořený RNA do kopie ve formě DNA, kterou následně integrují do genomu hostitelské buňky. Integrovaná DNA retroviru, tzv. provirus, je pak předlohou pro přepis retrovirových mRNA. Závislost replikačního cyklu na integraci do hostitelského genomu je vlastnost zcela unikátní a pro retroviry charakteristická. Exprese retrovirového proviru je pak zcela závislá na buněčném aparátu. Integrace však má důležitý důsledek i pro hostitelskou buňku, neboť ta je již perzistentně infikována. Perzistentní infekce buněk germinální linie tak např. dává vzniknout endogenním retrovirům, které jsou zachovány u dalších generací potomků infikovaného jedince.

Jelikož integrovaný provirus nelze odstranit, mají buňky obranné mechanismy působící na úrovni transkripce retrovirových genů, ale i na dalších úrovních genové exprese provirů. Dominantní vliv na expresi proviru mají nejspíše podmínky v místě integrace proviru, které určují, zda bude konkrétní provirus schopen aktivní transkripce či bude jeho exprese umlčena již v kroku transkripce. Přítomnost transkripčně umlčených provirů má velký význam pro léčbu retrovirových onemocnění. Buňky nesoucí neexprimující provirus totiž nevykazují znaky infekce retrovirem a nemohou být z organismu odstraněny. Tento fenomén je v současnosti ve velkém zájmu medicíny, neboť přítomnost latentního rezervoáru vytváří překážku pro úplné vyléčení pacientů infikovaných virem lidské získané imunodeficiency typu 1 (HIV-1) způsobující onemocnění AIDS. Umlčování retrovirové exprese je také problémem pro retrovirové vektory, které jsou využívány např. pro vytváření geneticky modifikovaných organismů či pro genovou terapii.

Expresní stav genomu je pod vlivem epigenetických faktorů. Mezi epigenetické znaky ovlivňující stav chromatinu jsou počítány hlavně histonové modifikace a metylace DNA. Jak se ukazuje, právě tyto epigenetické faktory a přítomnost epigenetických značek v oblasti proviru hrají roli v nastavení jejich expresního stavu. V současnosti je pozornost

zaměřena hlavně na epigenetické mechanismy umlčování exprese retrovirů a předpokládá se, že inhibitory enzymů umísťujících represivní epigenetické značky do oblasti proviru mohou být účinnými nástroji pro kompletní léčbu pacientů infikovaných HIV-1. Velmi málo je však známo o tom, jaké charakteristiky mají místa integrace různě transkripčně aktivních provirů a jaká místa naopak stabilně podporují aktivní expresní stav proviru. Tyto poznatky mohou být důležité právě pro vytváření expresně stabilních retrovirových vektorů či mohou nastínit, jak striktní jsou požadavky pro aktivní stav proviru a jak časté je tak umlčování provirové exprese. Díky své závislosti na charakteru integračního místa mohou retroviry také fungovat jako funkčně genomické sondy stavu chromatinu v místě integrace.

Tématem závislosti stabilní exprese retrovirů na funkčně genomických a epigenetických vlastnostech místa integrace v lidském genomu se zabývá tato diplomová práce. V laboratoři buněčné a virové genetiky ÚMG AV ČR, kde byla tato práce vypracována, již byl vytvořen model předpovídající závislost stabilně exprimovaných provirů ptačího sarkomového a leukózoového viru (ASLV) na integraci do funkčně a epigeneticky definovaných oblastí (Šenigl et al., 2012). Tato diplomová práce pak tento model rozšiřuje. Jak je v této práci popsáno, stabilní exprese ptačího retroviru ASLV v lidských buňkách je silně spjata s integrací proviru do míst, která po epigenetické stránce definují lokusy s aktivní transkripcí, hlavně pak místa, kde bývá transkripce lokusu zahajována. Další výsledky pak naznačují, že by tato vlastnost mohla být společná pro další skupiny retrovirů.

2. CÍLE PRÁCE

Hlavním cílem této diplomové práce je definovat charakteristiky integračních míst stabilně exprimovaných provirů ASLV v lidských buňkách. Tento cíl pak může být rozdělen do několika dílčích bodů:

- Příprava a expresní charakterizace lidských buněk exprimujících ASLV
- Určení místa integrace provirů stabilně exprimujících buněčných klonů
- Popis místa integrace identifikovaných provirů z hlediska genomiky a epigenomiky

Výsledky této práce budou využity jako předběžná data pro další analýzu většího a reprezentativnějšího souboru integračních míst.

3. PŘEHLED LITERATURY

3.1. Retroviry a jejich replikační cyklus

3.1.1. Systém a popis retrovirů

Retroviry (*Retroviridae*) jsou obalené viry s genomem ve formě RNA v pozitivním smyslu. Retroviry se dále dělí na orthoretroviry (*Orthoretrovirinae*), kam je zařazena většina dnes známých retrovirů, a spumaviry (*Spumavirinae*). Mezi orthoretroviry dále řadíme alfa-, beta-, gama-, delta-, epsilon- a lentiviry. Retroviry jsou infekčním agens mnoha živočichů včetně člověka, u nichž vyvolávají nejrůznější onemocnění. Lidskými retroviry jsou lidský leukemický virus typu 1 (HTLV-1) patřící mezi deltaviry a virus lidské získané imunodeficience typu 1 (HIV-1), patřící mezi lentiviry, který způsobuje onemocnění AIDS. První popsáný retrovirus, virus Rousova sarkomu (RSV) vyvolávající nádorové bujení u ptáků patří do skupiny ptačích alfaretrovirů. Jiným známým retrovirem je pak myší leukemický virus (MLV) řadící se mezi gamaretroviry který je znám hlavně díky jeho dřívějšímu využívání jakožto retrovirového vektoru pro účely genové terapie.

Uvnitř proteinové kapsidy obalené fosfolipidovou membránou původem z buněčné plazmatické membrány nesou retroviry dvě celogenomové RNA v kladném smyslu (jsou tedy diploidní). Tyto RNA mají vzhled buněčné mRNA – na 5' konci obsahují 7-metylguanovinovou čepičku, na 3' konci pak polyA sekvenci. Tyto molekuly slouží pro tvorbu dvouvláknové DNA kopie genomu retroviru. Kromě genomové RNA nesou retroviry ve své kapsidě také molekuly tRNA a enzymy kódované v genomu retrovirů, které hrají důležité role hlavně v prvních fázích replikačního cyklu retrovirů. DNA retrovirů dosahuje velikostí okolo 7 – 11 tisíc párů bází (kbp). Na obou koncích retrovirové DNA jsou oblasti dlouhých koncových repetitiv (LTR), kdy 5'LTR slouží retroviru jako vlastní promotorová oblast. Oblast mezi oběma LTR pak tvoří regulační sekvence (polypurinový trakt, enkapsidační signál, místo vazby tRNA - PBS) a retrovirové geny kódující virové proteiny a polyproteiny (geny *gag*, *pro*, *pol*, *env*, případně další geny u komplexních retrovirů či akutně transformujících retrovirů).

3.1.2. Replikační cyklus retrovirů

Na počátku replikačního cyklu retrovirů stojí vstup do hostitelské buňky. Retroviry pronikají do buněk pomocí receptorem zprostředkované fúze retrovirového obalu a buněčné plazmatické membrány. Retroviry využívají širokou škálu různých receptorů, které určují specifitu retroviru pro určitý buněčný typ. U viru HIV-1 je pro vstup do buněk nutný protein (receptor) CD4, přítomný na membráně specifických lymfocytů, a koreceptor, kterým může být protein CXCR4 či CCR5.

Ihned po vstupu viru do buňky je rozvolněna proteinová kapsida retroviru a retrovirový nukleoproteinový komplex putuje podél mikrotubulů k jádru. Rozvolnění kapsidy umožňuje zahájit jeden z unikátních a pro retroviry specifických dějů – reverzní transkripci. Při reverzní transkripci je virovým enzymem přítomným v kapsidě retroviru – reverzní transkriptázou (RT) – podle dvou celogenomových molekul RNA vytvořena dvouvláknová DNA forma retrovirového genomu, která je velmi podobná retrovirové RNA. Právě reverzní transkripce je odpovědná za vysokou rychlost vzniku mutací a rekombinací u retrovirů a je jedním z dějů cílených antiretrovirovou terapií.

Po dokončení reverzní transkripce disociuje RT spolu s kapsidovými proteiny z komplexu s DNA a dochází k maturaci v tzv. preintegrační komplex (PIC). Ten obsahuje řadu jak retrovirových, tak i buněčných faktorů, které mohou sehrávat důležité role v dalších fázích replikačního cyklu retrovirů. Nejdůležitější složkou PIC je však další pro retroviry unikátní enzym – integráza (IN). Hlavní funkcí integrázy je odštěpení nukleotidů z 3' konce a integrace retrovirové DNA do genomu hostitele. Aby však mohlo dojít k samotné integraci, musí PIC nejdříve překonat překážku ve formě jaderného obalu. Bylo postulováno, že integrace většiny retrovirů, s výjimkou lentivirů, je závislá na průchodu buňky mitózou, kdy dochází k rozpadu jaderné membrány a zpřístupnění jaderné DNA (Lewis and Emerman, 1994). Toto platí pro MLV, patřící mezi gamaretroviry (Roe et al., 1993). Pro ptačí virus vyvolávající sarkomy (ASV) byl prokázán alespoň v malé míře transport PIC do jádra, za který je zodpovědná právě IN a dosud neznámé buněčné faktory (Andrake et al., 2008). Vliv buněčného cyklu na rané fáze replikačního cyklu retrovirů je pak shrnut v souhrnném článku autorů Katz et al. (2005).

Integrovaná forma retrovirové DNA se nazývá provirus. Ačkoliv si některé proviry kódují faktory regulující provirovou expresi, stává se replikační cyklus retroviru vytvořením proviru závislým na buňce. Provirus se stává novým buněčným genem s vlastním promotorem, je buněčnými faktory exprimován a replikuje se společně s buněčnou DNA.

Jakmile dojde k syntéze dostatečného množství virových proteinů, dochází ke skládání komponent virových partikulí, balení celogenomových RNA a nakonec k tzv. pučení virionů z plazmatické membrány buňky a následné maturaci virionů mimo buňku.

3.2. Integrace retrovirů

3.2.1. Mechanismus integrace

Integrace retrovirové DNA a vytvoření proviru je zcela závislé na retrovirové integráze (IN). Tento enzym patří mezi nukleotidyltransferázy (EC 2.7.7) ze superrodiny DDE(D) rekombináz. Po dokončení reverzní transkripce tvoří IN hlavní složku PIC (Wei et al., 1997). Součástí tohoto komplexu bývá také množství buněčných faktorů. IN se váže na konce retrovirové DNA ve formě tetramerů. Tyto tetramery jsou vlastně dimery dimerů, kdy se na každý konec DNA vážou dvě molekuly IN a cílová DNA je navázána do prostoru mezi dimery. Struktury jednotlivých intermediátů integrace byly zobrazeny a detailně popsány pomocí krystalografických struktur, čímž byl umožněn mechanistický popis integrace retrovirové DNA do hostitelského genomu (Hare et al., 2010; Maertens et al., 2010). Chemicky je integrace transesterifikační reakcí, kdy dochází k nukleofilní substituci S_N2 typu. Nukleofilem je zde OH skupina na 3' konci retrovirové DNA a výchozí látkou je polynukleotidový řetězec hostitelské DNA. Po proběhnutí transesterifikace dochází k úpravě cílového místa integrace (odstranění přesahujících 5' konců retrovirové DNA a doplnění jednovláknových mezer okolo budoucího proviru) a okolo proviru vznikají krátké přímé repetice. Přesný průběh peri- a postintegračních událostí včetně zodpovědných buněčných faktorů však stále není znám.

3.2.2. Integrace v kontextu hostitelského genomu

Integrace retrovirové DNA není omezena na jeden genomový lokus, ale pokrývá celý hostitelský genom (**Obr. 1**). Dalo by se tedy říct, že je sekvenčně náhodná. Při detailnější analýze, kterou umožnilo hlavně získání sekvence lidského genomu (Lander et

al., 2001; Venter et al., 2001) si však lze všimnout, že různé skupiny retrovirů se liší frekvencí integrací do oblastí s různými charakteristikami (Derse et al., 2007; Elleder et al., 2002; Mitchell et al., 2004; Narezkina et al., 2004; Schröder et al., 2002; Wu et al., 2003). Retroviry tedy vykazují jisté integrační preference. Vzhledem k tomu, že pro integraci nebyla nalezena žádná specifika na úrovni sekvence DNA, ale pouze slabé, reálně se málo vyskytující palindromatické sekvence (Holman and Coffin, 2005; Wu et al., 2005), bude výběr místa integrace jednotlivých provirů pravděpodobně závislý převážně na buněčných faktorech interagujících s IN a stavu chromatinu hostitelské DNA.



Obrázek 1. Distribuce integračních míst retrovirů. Integrační místa tří různých retrovirů jsou zobrazena v lidském karyotypu. Modře jsou znázorněna integrační místa HIV-1, fialově místa MLV, zeleně místa ASLV. Červeně je pak na jednotlivých chromozómech zobrazena intenzita transkripce daných lokusů. Z obrázku je znatelné, že retroviry se integrují po celé délce genomu. Převzato z Mitchell et al. (2004).

3.2.3. Integrační preference retrovirů

Výskyt integrací retrovirů byl zpočátku hodnocen hlavně vzhledem k charakteristikám genomovým (**Obr. 2**). Zde můžeme zařadit např. integrace dovnitř či vně transkripčních jednotek (TU) a vzdálenost integrace od míst transkripčních startů (TSS), transkripční aktivitu TU či CpG ostrovy, což jsou oblasti bohaté na CG dinukleotidy, které bývají často asociovány s TSS většiny lidských TU (Bird, 1986). S postupným odhalováním funkčních prvků genomu však bylo možné hodnotit i frekvence integrací poblíž oblastí obohacených o nejrůznější epigenetické značky či vazebných míst pro proteinové faktory. Pro takové typy studií bylo třeba popsat velké množství míst integrací. Integrační preference byly do současnosti dobře definovány hlavně pro HIV-1, MLV, ASLV, jejichž výběr integračních míst se statisticky liší (Mitchell et al., 2004; Wu et al., 2003). Hlavním nositelem těchto odlišných preferencí je virová IN. Dojde-li totiž k výměně genu pro IN mezi retroviry, získává retrovirus preferenci podle původu IN (Lewinski et al., 2006). Roli integrázy v integračních preferencích retrovirů potvrzuje i zjištění, že retroviry příbuzné podle sekvence IN mají podobné preference integrace (Derse et al., 2007).

HIV-1 pro svou integraci preferuje oblasti aktivního, rozvolněného chromatinu (Taganov et al., 2004) a preferenčně se integruje do oblastí aktivních TU a to po celé jejich délce (Elleder et al., 2002; Mitchell et al., 2004; Schröder et al., 2002) a nepreferuje tedy oblasti TSS ani CpG ostrovy. V TU je nalézáno okolo 70% provirů HIV-1. Fakt, že stejnou preferenci vykazoval HIV-1 i v ptačích buňkách poukazuje na konzervovanost faktorů ovlivňujících výběr místa integrace. Doposud nebyla popsána žádná výrazná epigenetická determinanta preferovaná integrací HIV-1. Epigenetické modifikace identifikované ve vyšší míře v místě integrace zahrnují monometylacii lysinu 4 histonu H3 (H3K4me1) a acetylaci histonů H3 a H4 (H3ac, H4ac), což jsou značky aktivního chromatinu (Santoni et al., 2010; Wang et al., 2007). Autoři Wang et al. (2007) pak ve své studii poukazují na preferenční integraci HIV-1 do míst, která jsou okupována nukleozómy, a to zvláště do velkých žlábků nukleozomální DNA. Mikroskopické techniky pak určily, že k integraci HIV-1 by mělo docházet častěji na periferii jádra (Albanese et al., 2008; Di Primio et al., 2013).

Pro MLV je charakteristická integrace do blízkosti TSS aktivních TU (Mitchell et al., 2004; Wu et al., 2003). Do TU se integruje asi 40% provirů a 20 - 30% provirů MLV je nalézáno v těsné blízkosti TSS a CpG ostrovů. Častá je také integrace do blízkosti vazebných míst pro transkripční faktory (Felice et al., 2009). S tímto faktem pak souvisí i charakterizované epigenetické znaky nalézané v blízkosti provirů MLV. Těmi jsou trimetylace histonu H3 na lysinu 4 (H3K4me3), modifikace charakteristická právě pro oblasti TSS a H3K4me1 (Santoni et al., 2010). Znakem pro preferenci byla i blízkost místa vazby proteinu STAT1, či přítomnost RNA polymerázy II. Ve své práci pak autoři Santoni et al. (2010) pro MLV definovali integrační „superznačku“ složenou z modifikací H3K4me3, H3K4me1 a acetylovaného histonu H3 na lysinu K9 (H3K9ac). V okolí této „superznačky“, tedy alespoň jedné ze zmíněných histonových modifikací, se pak nacházelo 75% provirů MLV. Okolí samotné H3K4me3 zahrnovalo něco přes 60% provirů.

Oproti HIV-1 a MLV nevykazuje ASLV žádnou výraznou preferenci pro dosud popisované charakteristiky a jeho integrace se tedy nejvíce blíží integraci náhodné (Mitchell et al., 2004; Narezkina et al., 2004; Santoni et al., 2010). To platí jak v lidských buňkách, tak pro přirozeného hostitele ASLV, ptáky (Barr et al., 2005; Reinisova et al., 2008). Proviry ASLV byly nalézány v TU asi ve 40 - 57% integrací, což je jen o málo vyšší pravděpodobnost, než pravděpodobnost náhodné integrace. Na rozdíl od HIV-1 byla u ASLV pozorována účinnější integrace do kompaktního chromatinu (Taganov et al., 2004).

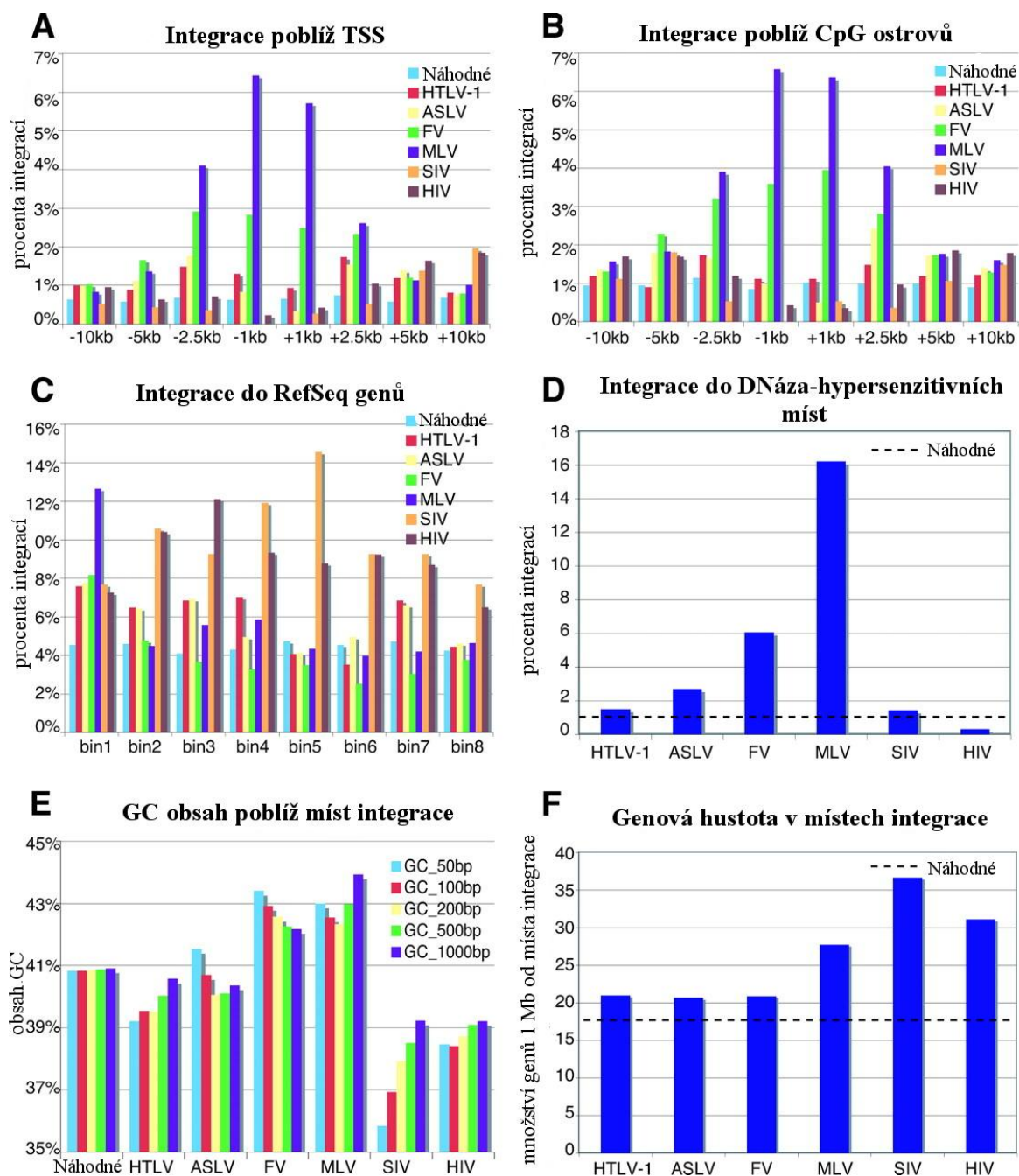
Integrační preference byly zjišťovány i u dalších retrovirů. Lidský retrovirus HTLV-1 vykazuje podobný výběr místa integrace jako ASLV s asi 46% provirů integrovaných v TU (Derse et al., 2007) s rovnoměrnou distribucí v TU bez vyšší preference k oblastem TSS a CpG ostrovům. Virus vyvolávající tumory mléčné žlázy u myši (MMTV) pak vykazuje integraci velmi blízkou integraci náhodné (Faschinger et al., 2008).

3.2.4. Buněčné faktory a cílení integrace

Mechanismus nenáhodného výběru místa integrace retroviru není jasný. Předpokládá se však vliv buněčných faktorů přítomných v PIC a interagujících s IN.

U HIV-1 byl identifikován s IN interagující buněčný růstový faktor odvozený z epitelu čočky/protein 75 (LEDGF/p75), který hraje roli ve vstupu PIC do jádra (Cherepanov et al., 2003; Maertens et al., 2003), integraci (Llano et al., 2006) a ovlivňuje integrační preferenci směrem k aktivním TU, TU koregulovaných faktorem LEDGF/p75 a místům bohatým na nukleotidy AT (Ciuffi et al., 2006). LEDGF/p75 ovlivňuje integrační preferenci nejspíše pomocí ne plně specifického přitahování („tethering“) právě k místům aktivních TU (Meehan et al., 2009) a vzorec tohoto přitahování lze měnit modifikací LEDGF/p75. Strukturální studie autorů Eidahl et al. (2013) pak ukázala, že LEDGF/p75 vykazuje vysokou vazebnou afinitu pro mononukleozómy, která je zprostředkována hlavně interakcí s trimetylovaným lysinem 36 histonu H3 (H3K36me3). Tato značka je charakteristická pro vnitřní sekvence aktivních TU. Zajímavostí však je, že v žádné ze studií zabývajících se integračními preferencemi HIV-1 nebyla H3K36me3 identifikována jako značka preferovaných míst integrace. Dalším faktorem ovlivňujícím preferenci HIV-1 může být buněčný histidinem bohatý protein 2 (HRP2), jehož funkce je znatelná hlavně v buňkách s absencí LEDGF/p75 (Schrijvers et al., 2012; Wang et al., 2012a).

Kromě případu HIV-1 a LEDGF/p75 či HRP2 nebyly identifikovány žádné další faktory hrající roli ve výběru místa integrace retrovirů. Přesný mechanismus nenáhodných integrací retrovirů je tak stále neznámý.



Obrázek 2 Integrační preference retrovirů. Grafy představující různé preference pro výběr místa integrace u vybraných retrovirů – HTLV-1, ASLV, „foamy“ virus (FV), MLV, virus opičí získané imunodeficiencie (SIV), HIV-1. Integrace jsou vyobrazeny proti náhodně vygenerovaným místům v genomu. **A**, procento provirů integrovaných v různých vzdálenostech od TSS genů do vzdálenosti 10 kbp od TSS. Záporná hodnota značí vzdálenost intergenových integrací, pozitivní intragenových. **B**, Integrace do 10 kbp od CpG ostrovů. **C**, intragenové integrace podél cílených genů z databáze RefSeq. Délky genů jsou normalizovány a rozděleny na 8 částí (bin1-bin8), přičemž bin1 značí začátek a bin8 konec genu. **D**, Integrace do míst senzitivních na štěpení DNázou I, která odpovídají místům rozvolněného chromatinu bez nukleozómů. **E**, Procentuální obsah dinukleotidů GC v okolí integrací. **F**, Množství genů identifikovaných ve vzdálenosti 1 Mbp od místa integrace. Převzato a upraveno podle Derse et al. (2007).

3.3. Exprese retrovirových genů

Transkripce genů proviru je řízena z oblasti 5'LTR. Tato oblast, stejně jako 3'LTR, je rozdělena na segmenty U3, R a U5, přičemž pro normální transkripci proviru je nejdůležitější oblast U3 obsahující promotor proviru, TSS a enhancery. Nachází se zde také četná místa vazby různých transkripčních faktorů, které mohou fungovat jako aktivátory, ale i represory retrovirové exprese. Některé komplexní retroviry jako HIV-1 či HTLV-1 kódují vlastní transaktivátory transkripce. Ty musí být v první fázi exprimovány z proviru. Vazbou na regulační sekvence proviru pak v druhé fázi provirové exprese zesilují transkripci z provirového promotoru.

Retroviry jsou často cílem transkripční represe. A právě represe neboli umlčování retrovirové transkripce je způsob, jakým se buňky mohou bránit množení již integrovaného proviru. Takový způsob restrikce je charakteristický např. pro umlčování endogenních retrovirů při embryonálním vývoji (přehledně v článku Rowe a Trono, 2011). Umlčování retrovirové exprese je příčinou přítomnosti latentního rezervoáru u pacientů infikovaných HIV-1 podstupujících antiretrovirovou terapii. Po vysazení terapie může dojít k opětovnému šíření viru z latentního rezervoáru. Latence HIV-1 je tak překážkou pro úplné vyléčení pacientů a způsoby reaktivace umlčených provirů a možné eradikace infikovaných buněk jsou v současnosti intenzivně studovány. Souhrn faktorů zodpovědných za latenci HIV-1 a možná řešení toho problému lze nalézt v souhrnném článku autorů Shirakawa et al. (2013). Umlčování proviru je také problémem pro retrovirové vektory, které bývají využívány hlavně za účelem dlouhodobé exprese transgenů.

Kromě latence HIV-1 jsou dalšími modely pro studium umlčování provirové exprese MLV v embryonálních kmenových buňkách či ASLV v buňkách lidských. V obou případech dochází k účinnému umlčování exprese provirů. Zatímco MLV působí jako model pro restrikci exprese retrovirů v kmenových buňkách, je umlčování ASLV v savčích buňkách výsledkem obrany již diferencovaných, pro ASLV nepermissivních buněk.

3.3.1. Buněčné faktory a umlčování retrovirové transkripce

Umlčování exprese retrovirů je děj, který může probíhat na úrovni peri- či postintegrační. Je otázkou, jak jsou faktory odpovědné za umlčení exprese retrovirových genů k proviru přitahovány. Jednou z možností je restrikce pomocí transkripčních faktorů, ať už se jedná o absenci aktivátorů či přítomnost represorů. Samotná oblast 5'LTR obsahuje sekvence pro řadu takových faktorů (**Obr. 3A**). U HIV-1 byly identifikovány faktory vážící se do oblasti 5'LTR, které přispívají k nastolení latence HIV-1 skrze přitahování histonových deacetyláz (HDAC, **Obr. 3B**) (Coull et al., 2000; He and Margolis, 2002; Jiang et al., 2007; Williams et al., 2006). Mezi tyto faktory patří proteiny YY1, NF-κB p50, Sp1 či c-Myc. Nově identifikovaným faktorem pro ustanovení latentní infekce je vazebné místo pro transkripční faktor AP-1 přítomné v oblasti 5'LTR (Duverger et al., 2013). Stav tohoto vazebného místa pak koreluje se schopností viru nastolit latentní infekci. Jako důležité transkripční aktivátory HIV-1 byly identifikovány NF-κB a NF-AT. Tyto faktory se ve své aktivní formě vyskytují v aktivovaných T-buňkách a naopak chybí v klidových T-buňkách, které představují rezervoár latentních provirů HIV-1. Nově však bylo ukázáno, že latentní proviry se vyskytují i v aktivovaných T-buňkách a že tyto proviry (nebo alespoň část z nich) mohou být aktivovány kontaktem T-buněk s dendritickými buňkami (van der Sluis et al., 2013). Lze tedy předpokládat, že k aktivaci umlčeného proviru HIV-1 může být potřeba víc, než doposud známé faktory.

Důležitou roli v ustanovení latence HIV-1 hraje umístění nukleozómů v okolí TSS proviru (**Obr. 3A**). Hlavní je pak umístění nukleozómu nuc-1, který je u neaktivních provirů umístěn těsně za TSS proviru ve směru transkripce (van Lint et al., 1996; Verdin et al., 1993). Odstranění tohoto nukleozómu je spojeno s acetylací okolí TSS (inhibicí HDAC) a hlavně s činností chromatin remodelačního komplexu BAF, který udržuje nuc-1 v pozici bránící transkripci proviru (Rafati et al., 2011). Odstranění tohoto represivního SWI/SNF chromatin remodelačního komplexu způsobuje retrovirový transaktivátor Tat, který do oblasti TSS přináší řadu aktivačních faktorů a komplexů včetně histon acetylázy p300 a aktivačního chromatin remodelačního SWI/SNF komplexu PBAF, jejichž činností dochází k posunu nuc-1 a uvolnění transkripčního startu. Tyto poznatky poukazují na důležitou roli acetylace, respektive deacetylace histonů v transkripčním umlčování HIV-1. Jinou epigenetickou modifikací, která udržuje HIV-1 v umlčeném stavu je také

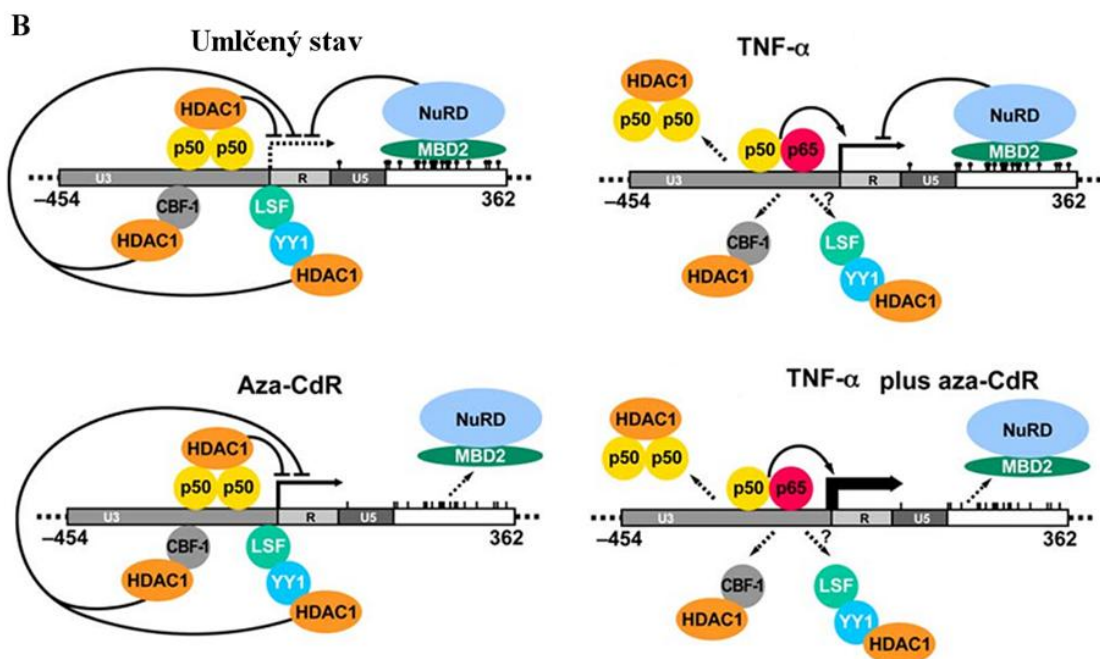
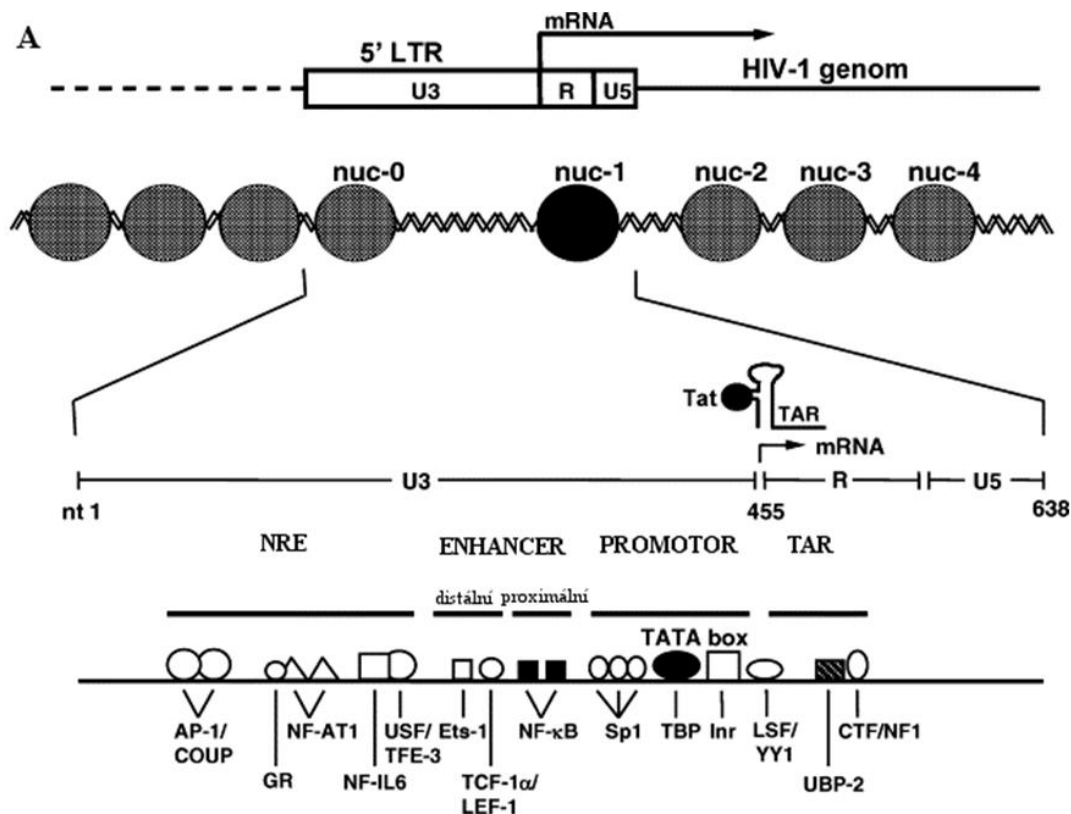
metylace DNA (Bednarik et al., 1990). Předpokládá se, že tato modifikace hraje roli epigenetického zámku, který provirus uzavře v umlčeném stavu a znesnadňuje tak jeho reaktivaci (Blažková et al., 2009).

Jiným příkladem specifického rozeznávání retrovirové sekvence aparátem buňky je umlčování exprese MLV v myších embryonálních kmenových buňkách. Pro toto velmi účinné umlčování je důležitá oblast vazby prolinové tRNA, tedy pro primer vazebné místo (PBS). Tato sekvence je specificky rozeznávaná DNA vazebným, sekvenčně specifickým proteinem ZFP809 (Wolf and Goff, 2009), který k provirům přináší transkripční korepresor TRIM28 (Wolf and Goff, 2007). Zamezením exprese obou těchto faktorů, stejně jako záměnou prolinové PBS za PBS pro jinou tRNA, dochází k narušení umlčování exprese MLV. Tyto faktory pak umlčují expresi také HTLV-1, který obsahuje PBS pro prolinovou tRNA. Tento velmi specifický aparát má buňky nejspíše ochraňovat před nechtěnou integrací exogenních či endogenních retrovirů v časných vývojových stádiích embrya. Pro umlčování retrovirů v kmenových buňkách ovšem existují i jiné mechanismy, které jsou společné pro ostatní systémy čítající různé typy buněk a retrovirů. Tyto mechanismy také velmi účinně umlčují proviry v kmenových buňkách a udržují proviry ve stavu represe i po diferenciaci kmenových buněk. Pro tuto represivní „paměť“ je nejvhodnějším kandidátem metylace DNA. Bylo však zjištěno, že k umlčování a „zapamatování“ umlčení dochází i v nepřítomnosti *de novo* metytransferáz DNMT3a a DNMT3b (Yao et al., 2004). Yao et al. (2004) navíc poukázali na fakt, že umlčení exprese provirů, které dávají vznik populacím aktivních i umlčených provirů, nemusí být v kmenových buňkách stálé, ale že provirus může oscilovat mezi aktivním a umlčeným stavem.

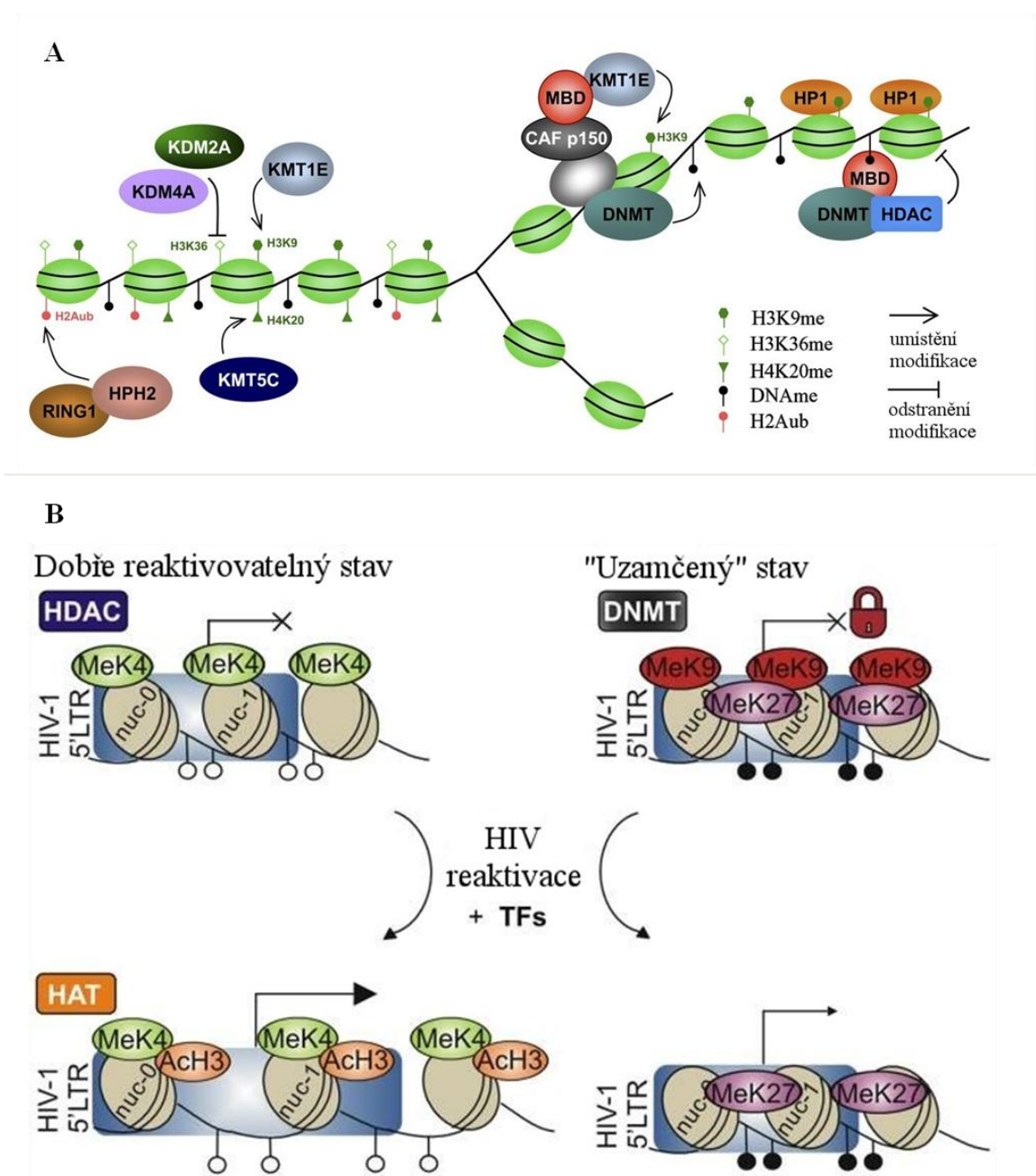
Další možností, jak zacílit represivní faktory na integrovaný provirus, je vazba buněčných faktorů na PIC a jejich zanášení do místa integrace. Vazba represivních faktorů na PIC by tak dovoľovala označení proviru ještě před jeho integrací do hostitelského genomu. Takovýto mechanismus je navržen pro umlčování ASLV v savčích buňkách. Greger et al. (2005) identifikovali Daxx jako buněčný protein interagující s IN ASLV ihned po infekci. Role Daxx proteinu pak nejspíše spočívá v přitahování HDAC a DNMT k proviru, čímž jako adaptér zajišťuje umlčování ASLV (Greger et al., 2005; Poleshko et al., 2008; Shalginskikh et al., 2013). Roli DNMT, zvláště pak DNMT3a a

DNMT3b v umlčování exprese ASLV v lidských buňkách potvrzuje i studie autorů Šenigl et al. (2012). Z výsledků této práce ale zároveň vyplývá fakt, že DNMT3a a DNMT3b jsou jen jedním z článků umlčovacích mechanismů buňky a umlčování exprese ASLV je zprostředkovávána i jinými, na DNMT nezávislými mechanismy. Bylo také pozorováno zvýšení exprese Daxx při infekci ASLV, což potvrzuje antivirovou aktivitu tohoto proteinu. Podobně jako u ASLV byla popsána interakce s IN a možná role v umlčování exprese proviru u HIV-1 (Huang et al., 2008). Jedním z prvních popsaných buněčných komplexů interagujících s IN HIV-1 však byl chromatin remodelační komplex SWI/SNF (Kalpana et al., 1994; Lesbats et al., 2011). Předpokládaná funkce této interakce je v usnadnění integrace do chromatinu. Avšak role těchto komplexů v umlčování provirové exprese a udržování provirů v latentním stavu (Boese et al., 2009; Rafati et al., 2011) poukazuje na možné cílení SWI/SNF komplexů do místa integrace pomocí interakce s IN s cílem blokovat expresi HIV-1.

Jak je vidět, za umlčování retrovirů stojí řada faktorů a mechanismů a umlčování retrovirové exprese ještě stále není vyřešenou otázkou. Společným jmenovatelem výše zmíněných mechanismů umlčování retrovirů však bývá epigenetická modifikace proviru (**Obr. 4**). Výrazné je zvláště přitahování HDAC k integrovanému proviru. Přítomnost HDAC by pak měla vést k deacetylaci histonů a umožňovat tak epigenetické uzavření proviru v heterochromatinu, na jehož pomyslném vrcholu je metylace DNA uzamykající provirus v inaktivním stavu. Fakt, že DNA metylace nebývá vždy nutná pro umlčení proviru, napovídá, že tato epigenetická modifikace nejspíše nebude prvním krokem k umlčování proviru. Samotný děj umlčování provirové exprese a to, která modifikace jej bude iniciovat, pak bude pravděpodobně záviset na souhře více faktorů.



Obrázek 3. Buněčné faktory a transkripce proviru. A, stavba 5'LTR oblasti HIV-1. Je označeno TSS proviru, umístění nukleozómů okolo TSS a místa vazby pro transkripční faktory uvnitř LTR. Při zobrazení nukleozómů je zvýrazněn nuc-1, který v této pozici brání transkripci proviru. NRE - negativní regulační element, TAR - Tat vazebné místo. Převzato a upraveno podle Demonté et al. (2004). B, funkce některých buněčných faktorů a komplexů při umlčování a reaktivaci exprese HIV-1. Zdůrazněny jsou zde role HDAC a DNA metylace (černá lízátka). Aktivační látky jsou TNF- α - tumor nekrotizující faktor α , Aza-CdR - 5-aza-2-deoxycytidine, inhibitor DNA metyl transferáz. Převzato a upraveno podle Kauder et al. (2009).



Obrázek 4. Epigenetické faktory a umlčování exprese retrovirů. Modely zobrazující souhru historonových modifikací a metylace DNA při umlčování exprese integrovaného proviru u ASLV a HIV-1. **A**, model souhry faktorů, které nejspíše hrají roli při umlčování ASLV v lidských buňkách. Převzato a upraveno podle Poleshko et al. (2010). **B**, model uzamčení HIV-1 v umlčeném stavu pomocí metylace DNA. Modrý obdélník znázorňuje oblast 5'LTR, černá lízátka zobrazují metylovanou DNA, prázdná lízátka zobrazují nemetylované CpG dinukleotidy. Zvýrazněn je nuc-1, který bývá při umlčeném stavu umístěn těsně za transkripčním startem HIV-1 a brání tak expresi proviru. DNMT – *de novo* metyltransferázy, HDAC – histon deacetylázy, HAT – histon acetyltransferáza. Převzato a upraveno podle Blažková et al. (2009).

3.3.2. Vliv místa integrace na expresi retrovirů

Jednotlivé proviry vykazují z pohledu exprese velmi heterogenní populaci. Za tuto variabilitu je zodpovědný charakter místa integrace proviru. Ačkoliv je tento fakt obecně přijímán, není doposud úplně jasné, jakým způsobem určuje místo integrace expresní stav proviru a jaká místa jsou charakteristická pro určitou hladinu či stabilitu provirové exprese.

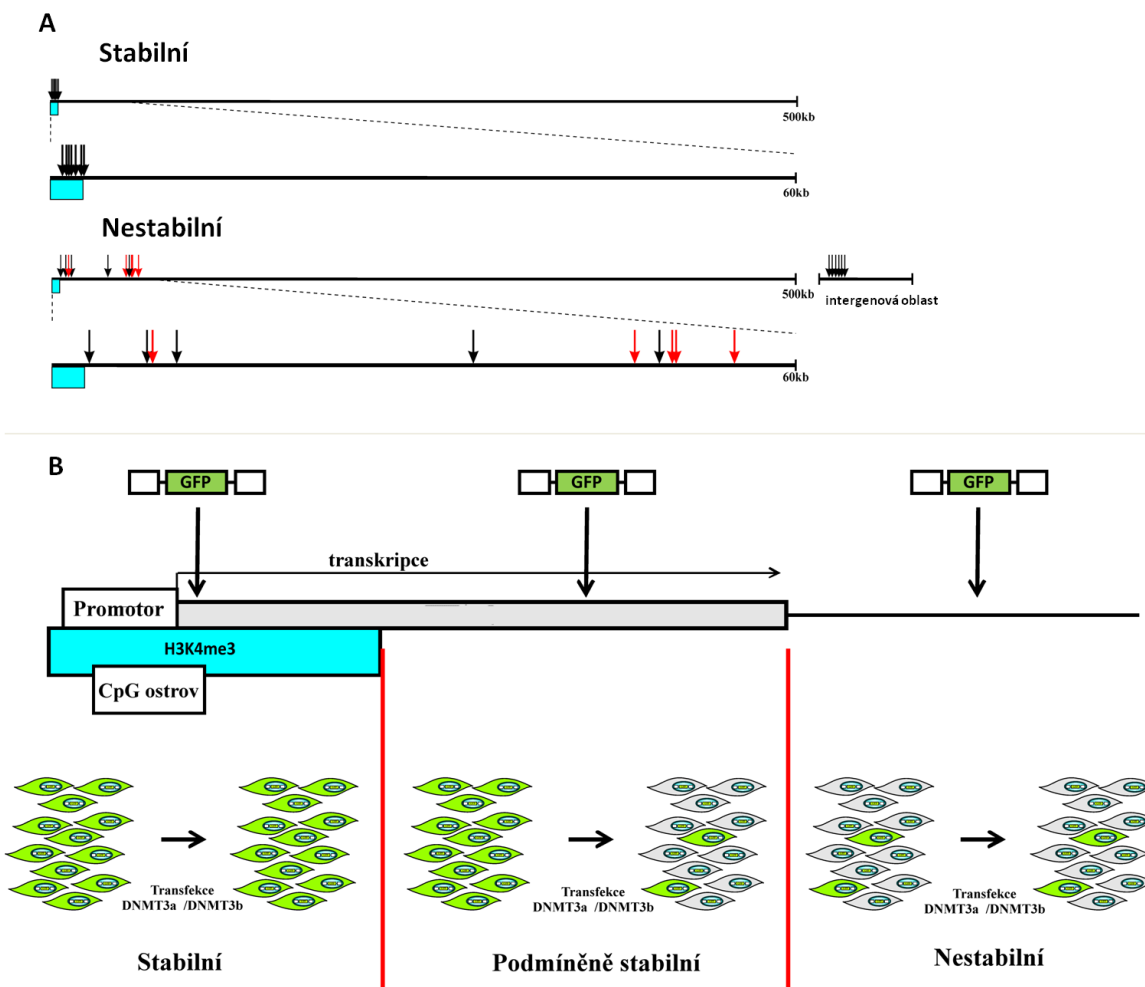
Na heterogenost bazální exprese HIV-1 poukazovala studie autorů Jordan et al. (2001). Tito autoři ukázali, že jednotlivé proviry vykazují různou míru exprese a že variabilita exprese je stírána expresí virového Tat proteinu, který funguje jako transaktivátor. Umlčené proviry HIV-1, které bylo možné *in vitro* reaktivovat, se ve zvýšené míře vyskytovaly v oblastech repetitivních alfa charakteristických např. pro centromerické oblasti, v oblastech chudých na geny a v oblastech vysoce aktivních TU, kde byly často nalézány proviry se shodnou orientací transkripce s transkripcí TU (Jordan et al., 2003; Lewinski et al., 2005; Shan et al., 2011). Poslední ze jmenovaných charakteristik je poměrně překvapivá a poukazuje na významný vliv transkripční interference, kdy vysoká aktivita transkripce TU brání účinné transkripci proviru. Tyto výsledky byly získány převážně na klidových CD4⁺ T-buňkách, které se zdají být významným rezervoárem latentních provirů. van der Sluis et al. (2013) však recentně ukazují, že latentní proviry se nacházejí také v aktivovaných T-buňkách. Tyto proviry tedy neodpovídají na aktivovaný stav T-buněk a umlčení jejich exprese je nejspíše ustanoveno ihned po integraci do míst podporujících umlčený stav provirů. Expresně stabilní proviry HIV-1 jsou pak podle autorů Lewinski et al. (2005) integrovány do středně aktivních TU a do menších intergenových oblastí, které jsou však bohaté na CpG ostrovy.

Vztah transkripčního stavu proviru a místa integrace byl studován také na modelu ASLV. Tento ptačí retrovirus totiž vykazuje náhodnější distribuci integrace a méně stabilní expresi hlavně při infekci savčích, pro ASLV nepermissivních buněk. Proviry indukující nádory u kuřat a disponující tedy dlouhodobou expresí onkogenu *v-src* vykazovaly nenáhodné rozložení integrací do oblastí TU exprimovaných ve středních hladinách v širokém souboru tkání (Plachý et al., 2010). Detailnější pohled na vliv integračního místa na expresi proviru pak přinesla práce autorů Šenigl et al. (2012) využívající právě model integrace a exprese ASLV v lidských buňkách. Tato práce

definovala model, kdy proviry integrované do TU jsou umlčovány mechanismem závislým na DNMT a tedy metylací DNA, zatímco proviry mimo TU jsou umlčovány i v nepřítomnosti DNMT. Stabilně exprimující proviry necitlivé na přítomnost DNMT pak byly nalezeny výhradně uvnitř TU, v blízkosti TSS, CpG ostrovů a v oblastech bohatých na H3K4me3 (**Obr. 5**). Tento model tedy předpokládá, že proviry integrované mimo tyto oblasti budou velmi náchylné k umlčování jejich exprese. Je však otázkou, zda takovýto model bude platit i pro jiné retroviry.

Recentně byla publikována data o závislosti expresní aktivity HTLV-1 na místě integrace v lidském genomu (Melamed et al., 2013). Výsledky této práce naznačují možný podíl transkripční interference na umlčování exprese provirů integrovaných uvnitř TU a závislost aktivity provirů integrovaných mimo TU na orientaci blízkých TU. Práce popisuje i vztah exprese a výskytu míst pro vazbu některých transkripčních faktorů. Přítomnost STAT1 poblíž 5'LTR proviru tak byla asociována s vyšší expresí proviru, zatímco přítomnost faktoru BRG-1 asociuje spíše s umlčenými proviry.

Funkční charakteristiky místa integrace mají dominantní vliv na expresní aktivitu jednotlivých provirů, což potvrzují dosud publikované výsledky. Z těch zatím vyplývá, že existuje mnoho míst, kde je integrovaný virus umlčen, zatímco pro udržení transkripční aktivity proviru budou potřeba více specifické podmínky v místě integrace. Sklon k umlčení či stabilní aktivitě proviru pak bude záležet na typu retroviru a infikovaných buněk. Integrace do aktivních TU pak nemusí znamenat stabilní expresi proviru, ale naopak může znamenat potíže ve formě transkripční interference, jak bylo ukázáno pro HIV-1 a HTLV-1. Naopak oblasti TSS aktivních TU, které bývají preferovány např. MLV, mohou nejspíše poskytovat ochranu před umlčením transkripce i při infekci nepermissivních hostitelů (ASLV v lidských buňkách).



Obrázek 5. Model závislosti exprese proviru na místě integrace. **A**, poloha integrovaných provirů ASLV v genomu lidských buněk. Integrace jsou znázorněny vzhledem k poloze v TU (dlouhé úsečky) a v intergenových oblastech (krátká úsečka napravo). Modře je znázorněna oblast výskytu H3K4me3. Stabilně exprimované proviry se nalézají výhradně poblíž transkripčních startů a v oblastech bohatých na H3K4me3. Nestabilní proviry jsou pak rozmístěny v různých vzdálenostech od transkripčního startu a v intergenových oblastech. Červeně jsou znázorněny proviry, jejichž transkripce probíhá proti transkripci TU. **B**, Model závislosti integrace a exprese ASLV v lidských buňkách. Stabilní proviry jsou integrovány uvnitř TU, v oblastech H3K4me3. Podmíněně stabilní proviry umlčované působením DNMT3a či DNMT3b jsou integrovány do TU ale mimo oblasti H3K4me3. Nestabilní proviry jsou integrovány mimo TU a jsou umlčovány i v nepřítomnosti DNMT3a a DNMT3b. Převzato a upraveno z Šenigl et al. (2012). DNMT – *de novo* metyltransferázy, H3K4me3 – trimetylace lysinu 4 histonu H3, GFP – zelený fluorescenční protein.

3.4. Využití retrovirů – retrovirové vektory

Retroviry díky své vlastnosti integrovat svou genetickou informaci do genomu hostitele a zajistit tak její přenos do dalších generací buněk nacházejí využití ve formě retrovirových vektorů. Ty bývají využívány hlavně při transdukcii cizorodých genů do buněk za cílem dosáhnout stabilní exprese transgenů ať už jde o buněčné kultury, tvorbu transgenních zvířat či genovou terapii geneticky podmíněných chorob. První pokusy o genovou terapii pomocí retrovirových vektorů využívající gamaretrovirové vektory však poukázaly na nebezpečí používání těchto vektorů. Při jinak úspěšné genové terapii 10 pacientů trpících X-vázanou kombinovanou imunodeficiencí (X-SCID) totiž došlo u dvou pacientů k vývoji leukemie T buněk (Hacein-Bey-Abina et al., 2003a). Klonální proliferace T buněk byla způsobena integrací gamaretrovirového vektoru do blízkosti TSS protoonkogenu LMO2 a k jeho aktivaci (Hacein-Bey-Abina et al., 2003b). Dalším z nedostatků retrovirových vektorů je také umlčování jejich exprese po integraci do genomu. Cílem vývoje retrovirových vektorů je tak odstranění těchto nedostatků a zajištění bezpečné integrace a zároveň stabilizace exprese retrovirů. Ačkoliv se stále nepodařilo dosáhnout jejich dokonalosti, byly již několikrát úspěšně využity k vyléčení pacientů s dědičnými chorobami, jako jsou dědičné formy imunodeficiency, β -thalasemie (Cavazzana-Calvo et al., 2010) či Wiskottova Aldrichova syndromu (Boztug et al., 2010), a jsou využívány pro genetickou manipulaci imunitních buněk pro imunoterapii. Práce na vylepšení vlastností těchto vektorů tedy mají potenciál pro vývoj vektorů pro přípravu transgenních živočichů a genovou terapii.

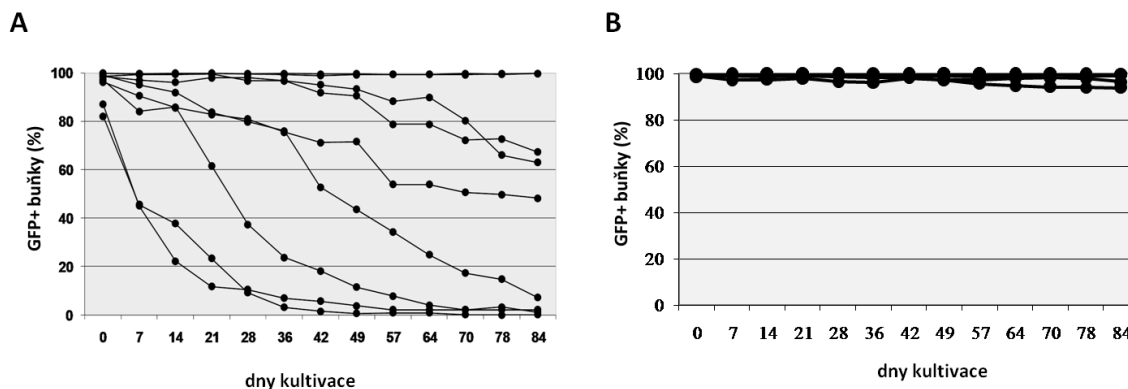
Předpoklad, že retrovirová exprese je závislá na místě integrace, navozuje myšlenku využít retrovirové vektory jako sondy pro funkční genomiku. Gierman et al. (2007) takto využili lentivirový vektor jako expresní sondu. Integrovaný provirus pak vykazoval hladinu exprese podle expresní aktivity TU vykazujících určitou podobnou hladinu transkripce. Autoři tak mohli vytvořit závěry o „doménové“ regulaci exprese jednotlivých TU na transkripční úrovni, která je nejspíše společná pro TU dané oblasti.

3.4.1. Modifikace retrovirových vektorů

Vývoj v konstrukci retrovirových vektorů je v současnosti směřován hlavně ke stabilizaci exprese. Stabilizace provirové exprese je prováděna vkládáním regulačních elementů do sekvence proviru. Často využívanými elementy bývají endogenní enhancery a izolátory, které by měly expresi proviru zesilovat či izolovat od vlivu okolního chromatinu (Emery et al., 2000; Groth and Emery, 2010; Miccio et al., 2011). Retrovirovou expresi také stabilizovalo vložení CpG ostrovů či jejich částí (**Obr. 6**) jak u lentivirových vektorů (Brendel et al., 2012), tak i u ASLV při infekci nepermissivních savčích buněk (Hejnar et al., 2001; Machoň et al., 1998; Šenigl et al., 2008). Brendel et al. (2012) ukázali, že tzv. element AU2COE odvozený od CpG ostrova nejen zesiluje expresi vnášeného genu, ale navíc zachovává expresní specifitu tkáňově specifického promotoru. Nevýhodou toho elementu však je, že obsahuje i dva opačně orientované transkripční starty endogenních TU, jejichž exprese je CpG ostrovem řízena. Toto může vést ke vzniku transkriptů začínajících z TSS uvnitř CpG ostrova a obsahujících buď vložený gen, nebo LTR spolu se sekvencí přiléhající k proviru. Šenigl et al. (2008) naopak využívají pouze malou část CpG ostrova, který obsahuje vazebná místa pro Sp1, faktoru, který je nejspíše zodpovědný za stabilizaci exprese proviru (**Obr. 6**).

Retrovirové vektory jsou také často využívány ve formě vektorů s inaktivovaným vlastním TSS (SIN). Tyto vektory nesou delecii v U3 oblasti LTR, ve které se vyskytuje promotor, a provirus tak bývá transkripčně defektní. Do vektoru pak bývá vložen jiný promotor, který může být i specifický pro určitou tkáň. Takto lze tedy konstruovat vektory se silnější, či tkáňově specifickou expresí. Nevýhodou toho postupu může být snížení kapacity vektoru vloženým promotorem.

Na rozdíl od stabilizace transkripce, nebylo zatím dosaženo úplného cílení inegrace vektorů. Bylo zjištěno, že integrační preference retroviru lze modifikovat buďto záměnou IN za IN jiného retroviru, či modifikací buněčných interakčních partnerů IN (Gijssbers et al., 2010). Posun v otázce místně cílené integrace retrovirů by však mohl znamenat objev definovatelných, místně specifických endonukleáz (endonukleázy se zinkovými prsty či nově technologie TALEN).



Obrázek 6. Stabilizace exprese retrovirového vektoru. Exprese retrovirového vektoru odvozeného od ASLV je stabilizována pomocí vložení krátkého úseku DNA z CpG ostrova. **A)** záznam exprese vektoru po infekci savčích buněk. Různá stabilita exprese je způsobena vlivem míst integrace. **B)** exprese modifikovaného vektoru po infekci savčích buněk. Vložený CpG ostrov blokuje vliv místa integrace proviru a stabilizuje tak expresi vektoru. Převzato a upraveno z Šenigl et al. (2008).

3.4.2. Distribuce histonových modifikací a transkripční aktivita genomu

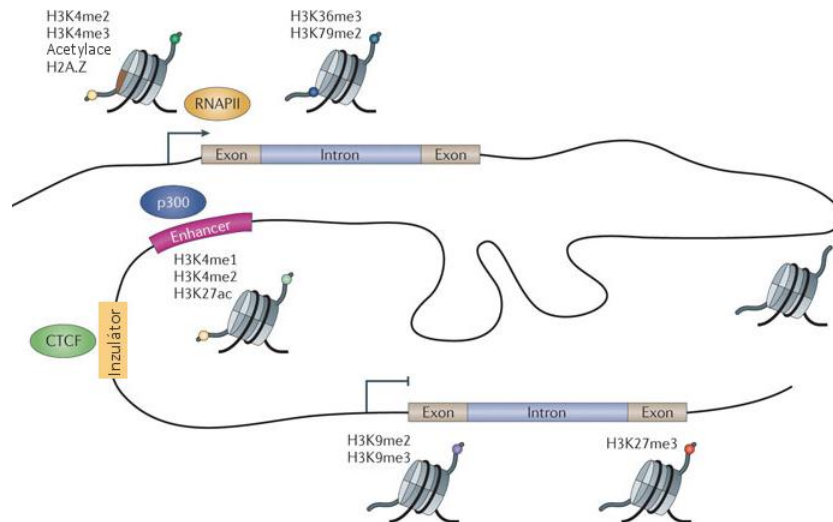
V předešlých kapitolách byly epigenetické modifikace často zmiňovány v souvislosti s transkripčním stavem provirů. Epigenetické modifikace jsou obecně důležitou součástí mnoha regulačních kaskád ovlivňující děje na úrovni genomu – od reparace DNA až po definici různých transkripčních stavů chromatinu. Na úrovni DNA je důležitou epigenetickou modifikací metylace DNA, která má na transkripci represivní účinky. Na úrovni histonů, bazických DNA vazebných proteinů, jejichž N-konec bývá často cílem posttranslačních modifikací, má jasnou roli v regulaci transkripce acetylace histonů, která umožňuje rozvolnění chromatinu a podporuje tak aktivní stav chromatinu. Methylace histonů má na rozdíl od acetylace různé funkce a záleží na typu histonu a na aminokyselině, která je metylována (trimethylace lysinu 4 histonu H3 se označuje jako H3K4me3). Jak se však ukazuje, distribuce a kombinace přítomnosti různě metylovaných histonů definuje funkční stav chromatinu (Zhou et al., 2011).

Aktivní stav oblasti TSS a regulačních oblastí genomu označují různé metylační stavy H3K4. H3K4me2 a H3K4me3 jsou přítomny v promotorových a 5' transkribovaných oblastech aktivních TU (Schneider et al., 2004). Zatímco hladina H3K4me2 klesá s metylací DNA (Okitsu and Hsieh, 2007), hladina H3K4me3 pozitivně koreluje s transkripční aktivitou TU. Přítomnost H3K4me2 v okolí aktivních TSS má nejspíše represivní funkci, kdy přitahováním HDAC zabraňuje protisměrné transkripci nekódujících RNA v okolí TSS aktivních TU (Kim et al., 2012). H3K4me1 je histonovou

modifikací charakteristickou pro oblasti enhancerů (Barski et al., 2007; Heintzman et al., 2007). H3K4me1 se také objevuje v okolí TSS dále za těmito oblastmi, zatímco H3K4me3 se ve zvýšené míře v oblastech enhancerů nevyskytuje. V 5' oblastech TU ve směru transkripce tedy metylační stav H3K4 postupně klesá.

Pro vnitřní oblasti transkripčně aktivních TU je charakteristická přítomnost vysoké hladiny H3K36me3, H3K79me2 a H4K20me1 (Barski et al., 2007; Li et al., 2007). Hladina H3K36me3 ve směru transkripce podél TU stoupá, zatímco hladiny H3K79me2 a H4K20me1 se zvyšují za TSS uvnitř TU a po dosažení maxima opět ve směru transkripce klesají. H3K36me3 má represivní roli, neboť přitahuje HDAC, jejichž působením dochází ke zpomalení výměny nukleozómů v transkribované oblasti a zamezování vzniku kryptické transkripce (Venkatesh et al., 2012). H3K36me3 je také substrátem pro vazbu DNMT3a, čímž způsobuje metylaci DNA uvnitř transkripčně aktivních TU. H3K79me2 a H4K20me1 pravděpodobně hrají roli při elongaci transkripce (Nguyen and Zhang, 2011; Vavouri and Lehner, 2012), jejich přesná funkce však není známá. Vysoká hladina H4K20me1 byla např. asociována s vysoce aktivními TU, jejichž TSS jsou řízeny CpG ostrovy. Další metylační stavy H4K20 bývají nacházeny v heterochromatinu, v aktivních TU však bývá nacházena pouze H4K20me1.

Metylace histonů může mít i represivní vliv na transkripci. H3K9me3 a H3K27me3 jsou histonové značky přítomné v oblastech heterochromatinu. Jejich přítomnost v oblasti TSS značí umlčení transkripce TU (Barski et al., 2007).



Obrázek 7. Distribuce histonových modifikací a transkripce. Zobrazeny jsou funkční genomové elementy a přítomnost nejčastěji se vyskytujících histonových modifikací, které definují stav těchto elementů. Zobrazeny jsou TU (Exon, Intron, Exon), enhancer a inzulátor. U enhanceru a inzulátoru jsou zobrazeny proteiny, které se do těchto oblastí často vážou (p300, CTCF). RNAPII – RNA polymeráza II. Převzato a upraveno z Heintzman et al. (2007).

4. MATERIÁL A METODY

4.1. Pokusný materiál

4.1.1. Buněčné linie – specifika a kultivace

AviPack

AviPack je ptačí buněčná linie vytvořená pro tvorbu infekčních částic retrovirových vektorů. Jde o upravenou buněčnou linii DF-1 stabilně exprimující geny *gag*, *pol* a *env* ASLV (Plachý et al., 2010).

Buňky byly kultivovány v médiu DMEM (Sigma) s 4% telecího séra (CS), 3% fetálního telecího séra (FCS), 1% kuřecího séra (KS) a antibiotiky (penicilin/streptomycin, 100 mg/ml každého) v atmosféře s obsahem 3% CO₂ při teplotě 37°C. Pasáž buněk byla prováděna každý 2. - 3. den. Při pasáži byly buňky opláchnuty fosfátovým puftrem (PBS) a enzymaticky rozvolněny (trypsin + EDTA, Invitrogen). Působení trypsinu bylo zastaveno přidáním kultivačního media.

HEK293

HEK293 je nenádorová lidská buněčná linie odvozená z buněk ledvin embrya (Human Embryonal Kidney). Jde o adherentní lidské buňky. Tyto buňky byly zprvu vybrány vzhledem k jejich nenádorovému původu. Tato buněčná linie byla první volbou pro experimentální práci, později však byla nahrazena linií K562.

Buňky byly kultivovány na miskách potažených poly-L-lysinem (1 mg/ml v H₂O). Kultivační médium se skládá z DMEM, 5% CS, 5% FCS a směsi antibiotik penicilin/streptomycin, 100 mg/ml každého). Kultivace probíhala v atmosféře s obsahem 3% CO₂ při teplotě 37°C. Pasáž buněk byla prováděna každý 2. - 3. den (postup stejný, jako u AviPack).

K562

K562 je lidská buněčná linie myeloidního typu. Jde o buněčnou kulturu odvozenou od krevních buněk pacientky s chronickou myeloidní leukémií v blastické krizi. Buňky linie K562 jsou kulaté a suspenzní. Tato buněčná linie byla vybrána z důvodu snadné kultivace při velkém množství buněčných klonů a dostupnosti velkého množství

informací v genomových databázích (UCSC Genome Bioinformatics, <http://genome.ucsc.edu/>). Na linii K562 byla provedena většina z prezentovaných experimentů.

Buňky byly kultivovány v médiu RPMI 1640 (Sigma) s 5% FCS, 5% CS a antibiotiky penicilin/streptomycin, 100 mg/ml každého) v atmosféře s obsahem 5 % CO₂ při teplotě 37°C. Pasáž buněk byla prováděna každý 2. - 3. den.

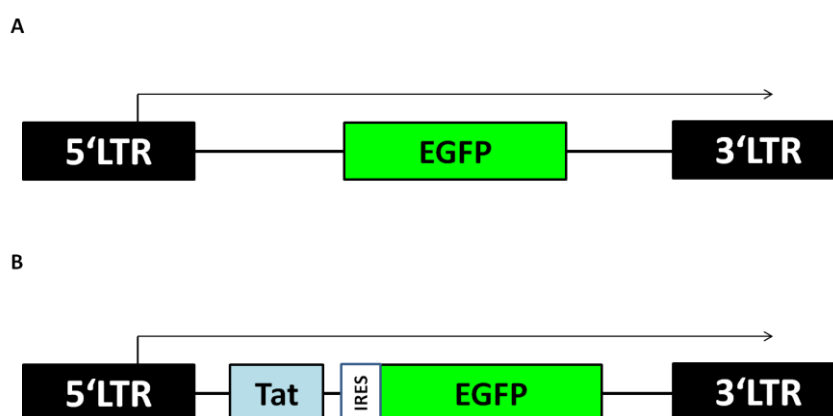
4.1.2. Retrovirové vektory

AG3

Retrovirový vektor AG3 je replikačně defektní vektor odvozený od ptačího retroviru ASLV (Šenigl et al., 2012). Vektor kromě obou LTR a enkapsidačního signálu (ψ) ASLV kóduje stabilizovaný zelený fluorescenční protein (EGFP), jehož exprese je řízena z provirového promotoru v 5' LTR (**Obr. 8A**).

LTR-Tat-GFP

Vektor LTR-Tat-GFP je replikačně defektní lentivirový vektor odvozený od lidského viru HIV-1. Vektor obsahuje obě oblasti LTR, oblast vazby Tat proteinu (TAR), virový transkripční faktor Tat a EGFP (**Obr. 8B**). Transkripce obou genů je pod kontrolou promotoru v 5' LTR HIV-1. Translace GFP je řízena z vnitřního místa pro nasedání ribozómů (IRES). Vektor byl vyvinut v laboratoři Dr. Verdina (Jordan et al., 2001).



Obrázek 8. Použité retrovirové vektory. A) retrovirový vektor AG3 odvozený od ASLV. B) retrovirový vektor LTR-Tat-GFP odvozený od HIV-1. Vektory jsou znázorněny ve formě provirů. Šipka ukazuje směr transkripce, která je iniciována v oblasti 5'LT. LTR – dlouhá koncová repetice, EGFP – stabilizovaný zelený fluorescenční protein, Tat – transaktivátor transkripce, IRES – vnitřní místo pro nasedání ribozómu.

4.1.3. Plazmidy

pAG3

pAG3 je plazmid použitý pro produkci AG3 vektoru. Příprava a popis celého plazmidu je popsána v práci Šenigl et al., 2012.

pHIV-1-TG

pHIV-1-TG je 7 kbp dlouhý plazmid použitý pro produkci LTR-Tat-GFP vektoru. Plazmid byl vytvořen F. Šeniglem v Laboratoři virové a buněčné genetiky ÚMG AV ČR. Genom vektoru byl amplifikován z integrovaného proviru buněčné linie H12 poskytnuté laboratoří Dr. Verdina a následně vložen do vektoru pGEM[®]-T Easy (Promega).

pVSV-G

pVSV-G je 6,5 kbp dlouhý plazmid použitý pro pseudotypizaci retrovirových vektorů. Plazmid nese obalový protein G viru vezikulární stomatitidy (VSV-G) pod kontrolou promotoru cytomegalovu.

pMD2.G, psPAX2

pMD2.G (Addgene plasmid 12259) a psPAX2 (Addgene plasmid 12260) jsou plazmidy vytvořené v laboratoři prof. Trono a využívané pro produkci lentivirových vektorů druhé generace. Plazmid pMD2.G nese obalový glykoprotein VSV-G, zatímco plazmid psPAX2 nese lentivirové geny (gag, pol, rev) potřebné pro produkci infekčních partikulí.

4.1.4. Roztoky

50x TAE

121, 1 g Tris Base (2 M), 50 ml 0,5 M EDTA pH 8,0 (50 mM), 28 ml kyseliny octové (5,6%).

T₁₀ E_{0,1}

1 ml 1 M Tris pH 7,6 (10 mM), 20 ml 0,5 M EDTA pH 8 (0,1 mM)

Lyzační pufr

0,5 M EDTA pH 8,0, 10% SDS, H₂O v poměru 5:1:4.

2x HBS

NaCl 820 mg, Na₂HPO₄ 11 mg, HEPES 595 mg, H₂O ad 50 ml; pH 7,1.

4.2. Strategie a organizace pokusů

Příprava retrovirových vektorů a infekce

Transfekce plazmidů do kuřecí buněčné linie AviPack, sběr, koncentrování a zamražení virové suspenze.

Infekce lidské buněčné linie retrovirovými vektory o nízké multiplicitě infekce (MOI). Tato infekce by měla zvýšit pravděpodobnost integrace maximálně jednoho proviru na buňku.

Selekce buněčných klonů a stanovení expresních profilů (**Obr. 9A**)

Selekce a vytvoření klonů buněk pozitivních na expresi GFP (GFP⁺ buňky) pomocí fluorescenčně aktivované buněčné separace (FACS). Separace populací GFP⁺ buněk či buněk negativních na expresi GFP (GFP⁻ buňky) krátce po infekci.

Stanovení dlouhodobých expresních profilů jednotlivých buněčných klonů pomocí průtokové cytometrie a identifikace klonů vykazujících stabilní expresi GFP.

Selekce GFP⁻ buněk u vybraných klonů se známým expresním profilem a analýza výskytu GFP⁺ buněk pomocí průtokové cytometrie.

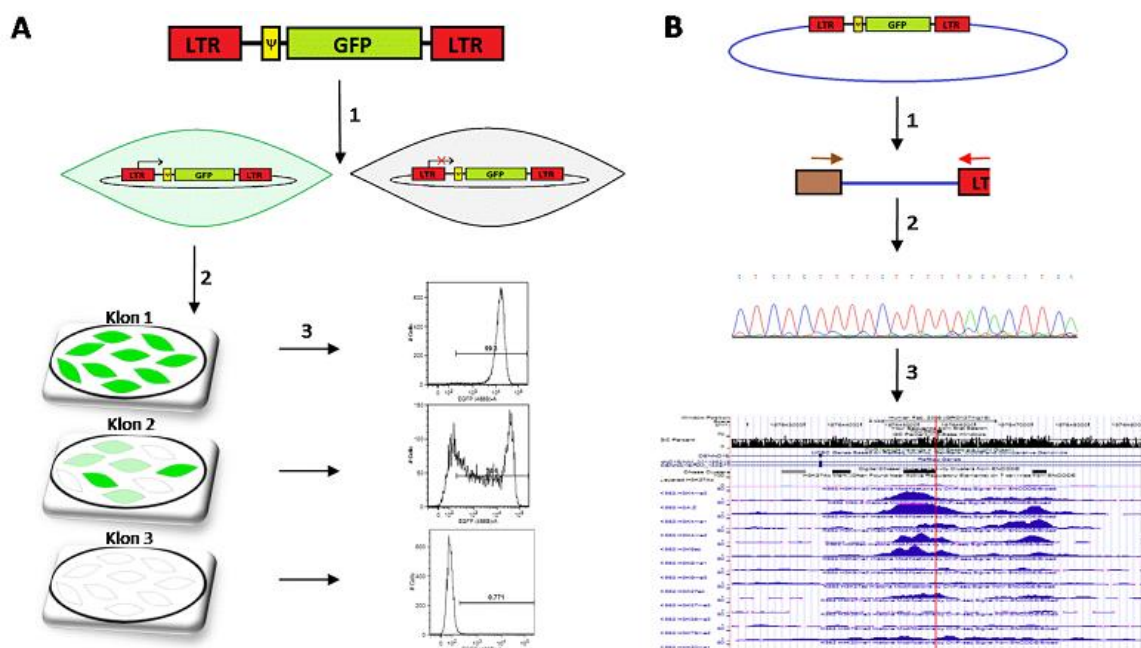
Identifikace a analýza integračních míst (**Obr. 9B**)

Izolace genomové DNA vybraných klonů. Amplifikace integračního místa pomocí splinkerette PCR a určení sekvence přiléhající k 5'LTR proviru.

Získání kontrolních souborů integračních míst odpovídajících neselektovaným provirům integrační preference ASLV, a provirům aktivním či umlčeným krátce po infekci.

Mapování získaných sekvencí do lidského genomu. Bioinformatická analýza míst integrací pomocí dostupných dat obsažených v genomových databázích.

Další přímá charakterizace míst integrace stabilně exprimujících provirů.



Obrázek 9. Základní přehled a uspořádání experimentů. A) Expresní profilování. 1) Lidská buněčná linie je infikována retrovirovým vektorem nesoucím GFP jako ukazatel exprese proviru. 2) 3 dny po infekci jsou pomocí FACS rozklonovány buňky GFP⁺ buňky a jsou tak ustanoveny buněčné klony. 3) U jednotlivých klonů jsou sledovány dlouhodobé expresní profily. **B)** Charakterizace místa integrace proviru. Místo integrace proviru je identifikováno s využitím splinkerette PCR, kdy je genomová DNA 1) štěpena restriční endonukleázou. K fragmentům genomové DNA je ligován splinker a DNA přiléhající k 5'LTR je amplifikována pomocí nested PCR. Primery nasedají na oblast splinkeru a LTR. 2) Amplifikovaná oblast je sekvenována. 3) Získaná sekvence je mapována do lidského genomu a oblast integrace je charakterizována pomocí informací v databázích. LRT – dlouhá koncová repetice, GFP – zelený fluorescenční protein.

4.3. Produkce retrovirových vektorů

4.3.1. Produkce vektoru AG3

Buněčná linie AviPack byla kotransfekována plazmidy pAG3 a pVSV-G metodou kalcium fosfátové precipitace. Buňky byly vysety v počtu 10^7 buněk na 150mm Petriho misku. Po 24 h bylo buňkám vyměněno kultivační médium, kdy bylo k buňkám přidáno 24 ml čerstvého media. Kalcium fosfátová precipitace byla provedena 2 h po výměně média. Složení transfekčních roztoků je uvedeno v **tabulce 1**. Roztok B obsahující transfekovanou DNA byl opatrně přikapáván k Roztoku A, který musel být stále protřepáván. Po smíchání obou roztoků došlo k vytvoření jemně zakalené suspenze, která byla opatrně přidána ke kultivovaným buňkám. Po 4,5 h byly buňky jednou

opláchnuty 15% roztokem glycerolu v HBS a jednou PBS. Po promytí bylo buňkám opět přidáno 15 ml čerstvého kultivačního media.

Tabulka 1. Složení transfekčních roztoků

Roztok A		Roztok B	
Látka	Objem	Látka	Objem, množství
2x HBS	2205 μ l	2M CaCl	270 μ l
HPO ₄ + H ₂ PO ₄	45 μ l	pVSV-G	10 μ g
		pAG3	50 μ g
		H ₂ O	ad 2250 μ l

Látky jsou v tabulce uvedeny v zásobních koncentracích a ředěních.

4.3.2. Produkce vektoru LTR-Tat-GFP

Vektor HIV-1 byl produkován kotransfekcí plazmidů pMD2.G, psPAX2 a pHIV-1-TG do buněk 293T-17. Postup produkce vektoru HIV-1 je stejný, jako při produkci vektoru AG3. Buňky byly vysety na 100mm Petriho misku a každý z plazmidů byl přidán v množství 10 μ g.

4.3.3. Sběr virových partikulí

Virové partikule byly sbírány z kultivačního media transfekovaných buněk 48 a 72 h po transfekci. Odebrán byl vždy celý objem kultivačního media (15 ml). Takto odebrané virové suspenze byly zamraženy v -80°C.

4.3.4. Koncentrování viru

Virová suspenze byla nejdříve podrobena centrifugaci (200 x g, 20 min, pokojová teplota). Výsledný supernatant byl přenesen do ultracentrifugačních zkumavek a podroben ultracentrifugaci (23 000 rpm, 150 min, 4°C, rotor SW28, Beckman Optima100). Po odstranění supernatantu byl výsledný pelet resuspendován ve 400 μ l vychlazeného DMEM média. Takto vzniklá virová suspenze byla po protřepání (Orbit™ P2, Labnet, 100 min, 4°C) ponechána odstát (přes noc, 4°C). Další den byla suspenze opět stejným způsobem protřepána a rozdělena po 200 μ l do zamrazovacích ampulí. Tyto výsledné virové suspenze pak byly uloženy do -80°C.

4.3.5. Titrace retrovirových vektorů

Fluorescenční titr virových vektorů byl odhadován 3 dny po infekci (dpi) na lidských buněčných liniích. Titr viru AG3 byl odečítán počítáním GFP⁺ HEK 293 buněk pod fluorescenčním mikroskopem (Olympus IX51). Titr AG3 byl odhadnut na $3,52 \times 10^6$ FU/ml (fluorescenčních jednotek na ml). Titr vektoru LTR-Tat-GFP byl odhadován pomocí ředící řady virové suspenze na buňkách K562 s využitím průtokové cytometrie. Pro další infekce byl vybrán poměr virová suspenze/počet buněk, při kterém bylo v infikované populaci 3 dpi přítomno 0,3 % GFP⁺ buněk.

4.4. Infekce HEK293 a K562 retrovirovými vektory

Lidské buněčné linie HEK293 a K562 byly infikovány při MOI okolo 0,2 tak, aby 3 dpi bylo v kultuře přítomno maximálně 0,5 % GFP⁺ buněk. Virová suspenze byla vždy rozmrazena v 37°C (vodní lázeň) a naředěna kultivačním médiem.

HEK293

Požadované množství buněk bylo vyseto na Petriho misku potaženou poly-L-lysinem. Po 3 - 4 h bylo odsáto médium a buňky byly infikovány v celkovém objemu 200 μ l virového inokula v kultivačním médiu se sérem. Buňky byly poté ponechány 30 min ve tmě (během této doby byla miska protřepána) a následně bylo přidáno kultivační médium v objemu 3 ml.

K562

Buňky byly nejdříve centrifugovány (200 x g, 5 min, 22°C). Buněčný pelet byl poté resuspendován ve virové suspenzi, která dle potřeby naředěna kultivačním médiem v poměru 1:9. Buňky byly v této suspenzi ponechány po dobu 30 min ve tmě, přičemž byla suspenze po 15 min protřepána. Infikované buňky byly poté vysety na 60mm Petriho misku a suspenze byla doplněna 3 ml čerstvého kultivačního media. V případě buněk určených pro separaci smíšené populace GFP⁻ buněk se MOI pohybovala okolo 1 – důvodem byla snaha zajistit vysokou proinfikovanost populace a záchyt co největšího počtu provirů umlčených již 3 dpi.

4.5. Selekce a ustanovení buněčných klonů

4.5.1. Příprava buněk pro FACS

Infikované buňky byly sklizeny a stočeny (200 x g, 5 min, 22°C). Buněčný pelet byl opět resuspendován v 1 ml media bez séra. Takto byly zpracovány i buňky neinfikované, které sloužily jako negativní kontrola pro rozlišení GFP⁺ a GFP⁻ buněk.

4.5.2. FACS

K roztřídění buněk podle fluorescence byl využit přístroj BD InfluxTM Cell Sorter (BD Biosciences) ze Servisní laboratoře cytometrie a světelné mikroskopie Ústavu molekulární genetiky AVČR. Buňky byly tříděny 3 dpi pouze tehdy, nepřesahovalo-li množství GFP⁺ buněk 3 dpi 0,5 % z celkové buněčné populace.

HEK 293

V případě HEK293 byly buňky separovány dvakrát. GFP⁺ buňky byly nejdříve separovány do jedné zkumavky a vysety na Petriho misku v 1 ml kultivačního media. 44 dpi byly GFP⁺ buňky znovu separovány, ale tentokrát po jedné buňce na jamku se 100 µl kultivačního media do 96 jamkové kultivační destičky. Dno jamek bylo před separací potaženo poly-L-lysinem.

K562

GFP⁺ buňky byly 3 dpi separovány po jedné buňce do jamek 96 jamkové kultivační destičky. Jamky byly naplněny 100 µl kultivačního media. Frakce buněk určená pro analýzu aktivních integrací ve smíšené buněčné populaci AG3 vektoru byla 3 dpi separací rozdělena na populaci GFP⁺ a GFP⁻ buněk. Pro tyto účely byly buňky separovány do zkumavek s 10 ml media, následně centrifugovány (200 x g, 5 min, 22°C) a vysety na 6 jamkové kultivační destičky v objemu 2 ml kultivačního media (každá populace do jedné jamky).

4.6. Stanovení expresních profilů buněčných klonů

4.6.1. Preselekce pomocí fluorescenční mikroskopie

Buněčné klony byly v kultivačních 96 jamkových deskách bez pasáže kultivovány po dobu 2 - 3 týdnů. Po této době byly jednotlivé klony mikroskopicky (Olympus IX51) kontrolovány na expresi GFP. V této fázi byly vyřazeny klony, u nichž nebyla pozorována

fluorescence, nebo klony u nichž byla pozitivní pouze minoritní část klonu. Vybrané klony byly přeneseny na 24 jamkové kultivační destičky a kultivovány v 1 ml kultivačního media.

4.6.2. Příprava buněk pro průtokovou cytometrii

Den před měřením na průtokovém cytometru bylo klonům odsáto médium a přidáno 600 μ l čerstvého kultivačního media. V den měření bylo 300 μ l buněčné suspenze měřených klonů přeneseno do 96 jamkové kultivační destičky. Dále kultivovaným klonům bylo médium doplněno na 1 ml. Kultivační destičky byly centrifugovány (540x g, 3 min, 4°C) a buněčné pelety byly resuspendovány ve 100 μ l roztoku Hoechst 33258 v PBS (1 μ g/ml).

4.6.3. Průtoková cytometrie

Cytofluorescence klonů byla stanovována každý týden od 30. do 60. dpi (v případě K562) a 80 dpi (v případě HEK293) na cytometru BD LSRII v Servisní laboratoři cytometrie a světelné mikroskopie Ústavu molekulární genetiky AVČR. Měření bylo prováděno v „High Throughput“ módu. Měření bylo nastaveno na maximální počet 10 000 Hoechst negativních buněk (živé buňky). Byla měřena intenzita fluorescence a procento GFP⁺, Hoechst negativních buněk.

4.6.4. Vyhodnocení exprese klonů

Data naměřená na cytometru byla zpracována v programu FlowJo 9.5. Za stabilně exprimující klony byly považovány ty, u nichž i) procento GFP⁺ buněk nekleslo ani po 60 dpi pod 90 %, ii) procento GFP⁺ buněk bylo menší, než 90 %, ale hladina GFP⁺ buněk byla v čase stálá. Kromě procenta GFP⁺ buněk byla zpracována i intenzita fluorescence jednotlivých klonů. Data o expresi klonů byla dále graficky zpracována v programu R 2.15.1.

4.7. Analýza GFP⁻ subpopulací

U vybraných klonů linie K562 s proviry AG3, u nichž byl stanoven expresní profil, byly separovány GFP⁻ buňky (postup viz **Selekce a ustanovení buněčných klonů**). Buňky byly separovány do 15ml zkumavek s 10 ml kultivačního media a následně centrifugovány (200 x g, 5 min, 22°C) a přeneseny na 24 jamkové kultivační destičky v 1 ml kultivačního media. Po 1 - 2 týdnech po separaci (tps) byla měřena exprese

negativních subpopulací pomocí průtokové cytometrie (viz. **Stanovení expresních profilů buněčných klonů**).

4.8. Zamražení buněčných klonů

Buňky byly sklizeny a centrifugovány (200 x g, 5 min, 22°C). Buněčný pelet byl resuspendován v kultivačním médiu s přidavkem FCS (10%) a DMSO (10%) v celkovém objemu 1 ml. V zamrazovací ampuli byla suspenze přenesena do zamrazovací krabičky plněné isopropanolem (Nalgene), která byla vložena do -80°C. Po 24 h byly ampule vyjmuty a uloženy do -80°C.

4.9. Izolace genomové DNA z buněk v tkáňové kultuře

4.9.1. Sběr buněk

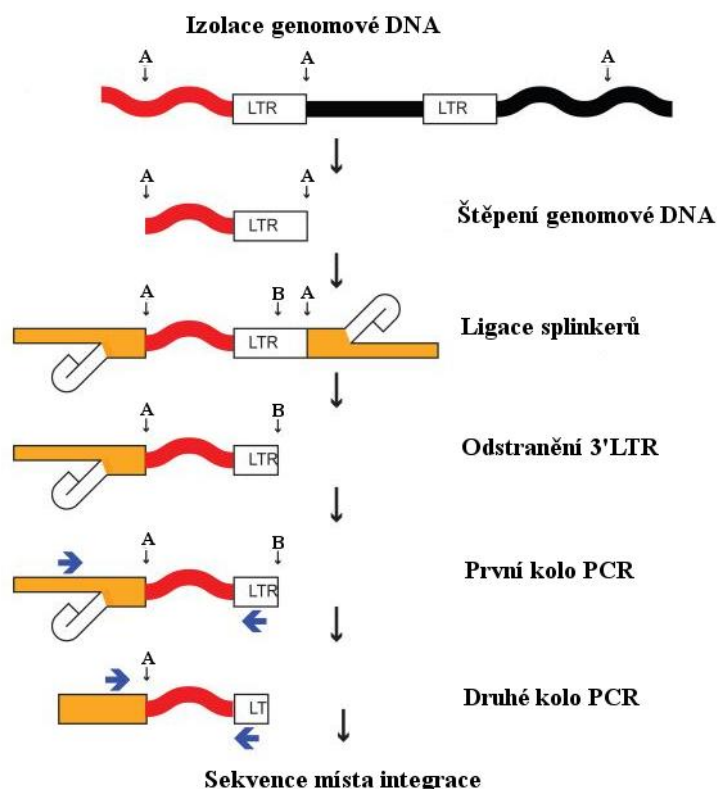
Buňky byly vysety na 6 jamkové kultivační destičky, odkud byly po dosažení dostatečné hustoty sklizeny a centrifugovány (200 x g, 5 min, 4°C). Po odstranění supernatantu byl takto získaný buněčný pelet buď přímo zpracován, nebo uložen do -20°C.

4.9.2. Izolace genomové DNA

Na buněčný pelet bylo aplikováno 600 µl lyzačního roztoku. Pelet byl ustříženou špičkou opatrně resuspendován. Poté bylo přidáno 200 µl lyzačního roztoku. K 800 µl suspenze bylo přidáno 40 µl proteinázy K (20 mg/ml). Suspenze poté rotovala přes noc při 55°C. Další den byla k suspenzi přidána RNáza A (100 mg/ml) v množství 2,5 µl a suspenze rotovala další hodinu při 37°C. Po vyjmutí bylo k buněčnému lyzátu přidáno 800 µl směsi fenol:chloroform (1:1, ph 7,9). Po opatrném protřepání byla směs centrifugována (16 100 x g, 3 min, pokojová teplota). Odebraná horní vodná fáze obsahující DNA byla smíchána se stejným objemem vychlazeného (-20°C) 96% etanolu. Po protřepání se směs ponechala stát při pokojové teplotě po dobu 5 min. Následovala centrifugace (16 100 x g, 15 min, 12°C) a odstranění supernatantu. Pelet vysrážené DNA byl opláchnut vychlazeným (-20°C) 80% etanolem a znovu centrifugován (16 100 x g, 15 min, 12°C). Po odstranění supernatantu byl pelet vysušen při 37°C a rozpuštěn v 50 µl T₁₀E_{0,1} pufru. Výsledná koncentrace genomové DNA byla změřena pomocí nanofotometru (IMPLEN). Vzorky byly uloženy ke skladování při 4°C.

4.10. Splinkerette PCR

Pro amplifikaci sekvence přilehlé k provirusu byla zvolena metoda splinkerette-PCR (Uren et al., 2009). Tento postup zahrnuje štěpení genomové DNA pomocí restriktivního enzymu, ligaci splinkerů, po kterém následuje druhé štěpení restriktivním enzymem pro odstranění 3'LTR oblasti a následně nested PCR s primery specifickými pro oblasti LTR a splinkerů (Obr. 10).



Obrázek 10. Splinkerette PCR pro amplifikaci integračních míst provirů. Genomová DNA obsahující integrovaný provirus je nejprve štěpena restriktivním enzymem A. Následuje ligace splinkerů a odstranění 3'LTR restriktivním enzymem B. Sekvence genomové DNA přiléhající k 5'LTR (červená) je amplifikována s využitím nested PCR s primery nasadajícími na oblast LTR a splinkeru. LTR – dlouhá koncová repetice.

4.10.1. Štěpení genomové DNA

Celková genomová DNA v množství 2 µg byla štěpena ve 20µl směsi s restriktivní endonukleázou. Složení reakcí se lišilo podle typu použitého retrovirového vektoru. Pro štěpení DNA obsahující provirus vektoru AG3 byla použita restriktivní endonukleáza DpnII či MseI (New England BioLabs) ve výsledném množství 10 U endonukleázy na reakci. Pro štěpení DNA s provirem vektoru HIV-1 byla vytvořena směs restriktivních endonukleáz NheI, SpeI a XbaI (New England BioLabs), kdy směs obsahovala 7,5 U jednotlivých

endonukleáz na jednu reakci, nebo byla použita samotná MseI (New England Biolabs). Reakce byly doplněny o 10x koncentrovaný reakční pufr dle doporučení výrobce (DpnII, MseI – Pufr 2, NSX směs – Pufr 4, New England BioLabs) a 100 x koncentrovaný telecí sérový albumin (BSA, New England BioLabs), obojí ve výsledné koncentraci 1x. Směs byla doplněna vodou do konečného objemu 20 μ l. Štěpící reakce probíhala přes noc při 37°C. Poté byly reakce tepelně inaktivovány (65°C, 20 min).

4.10.2. Příprava splinkerů

Pro přípravu splinkerů byly využity oligonukleotidy HMspAa (5'-CGAAGAGTAACCGTTGCTAGGAGAGACCGTGGCTGAATGAGACTGGTGTCGACACTAGTGG-3'), HMspBb-DpnII (5'-GATCCCACTAGTGTGCGACACCAGTCTCATTTTTTTTTTTTCAAAAAA-3'), HMspBb-MseI (5'-TACCACTAGTGTGCGACACCAGTCTCATTTTTTTTTTTTCAAAAAA-3') a HMspBb-ASX (5'-CTAGCCACTAGTGTGCGACACCAGTCTCATTTTTTTTTTTTCAAAAAA-3'). Splinkery byly připraveny smícháním oligonukleotidu HMspAa se specifickým oligonukleotidem HMspBb, který obsahuje na 5' konci sekvenci komplementární k sekvenci rozeznávané použitými restričními endonukleázami. Sekvence 3' konce oligonukleotidu HMspBb pak tvoří vlásenku, zvyšující specifitu splinkerette-PCR.

Hybridizační směs obsahovala oligonukleotidy HMspAa a HMspBb (každý ve výsledné koncentraci 4 μ M), 10 μ l 10 x koncentrovaného SuRE/Cut pufr M (Roche) a vodu (ad 100 μ l). Hybridizační reakce probíhala při programu: 94°C (3min), ochlazení na 21°C rychlostí 1°C/15 s. Po provedení hybridizační reakce byla směs naředěna 100 μ l vody.

4.10.3. Ligace splinkerů a štěpené genomové DNA

Splinkery a štěpená genomová DNA byly ligovány ve směsi obsahující 3 μ l štěpené DNA (300 ng DNA), 5 μ l směsi splinkerů, 400 U T4 DNA ligázy (New England BioLabs), T4 DNA ligázový pufr (výsledná koncentrace 1x) a byla doplněna vodou do objemu 20 μ l. Ligační reakce probíhala přes noc při teplotě 15°C. Následující den byla reakce inaktivována při teplotě 65°C (20 min).

4.10.4. Odstranění 3'LTR oblasti

Pro zamezení amplifikace vnitřních provirových sekvencí bylo potřeba odstranit spojení mezi těmito vnitřními sekvencemi a 3'LTR proviru. To bylo provedeno pomocí

restrikčního štěpení. Pro tyto účely byly využity enzymy Bsu36I (New England BioLabs) při použití vektoru AG3 a endonukleázy PvuII či DraI (New England BioLabs) při použití vektoru LTR-Tat-GFP. Reakce se skládala z celé ligační reakce (20 μ l), 10 x koncentrovaného restrikčního pufru dle doporučení výrobce (New England BioLabs), 100 x koncentrovaného BSA (New England BioLabs) a 10 U restrikční endonukleázy. Směs byla doplněna vodou do konečného objemu 50 μ l. Reakce probíhala při teplotě 37°C (přes noc) a následující den byla tepelně inaktivována (podle návodu výrobce).

4.10.5. Purifikace DNA před PCR reakcí

DNA byla dále přečištěna pro účely amplifikace sekvencí pomocí nested PCR. Pro purifikaci DNA byl využit komerční kit (QIAquick PCR Purification Kit, Qiagen). DNA byla z kolonek eluována ve 22 μ l vody.

4.10.6. Nested PCR

Amplifikace neznámé genomové sekvence přiléhající k 5'LTR proviru byla provedena pomocí metody nested PCR (dvě po sobě následující PCR reakce) s primery hybridizujícími s oblastmi LTR a splinkeru. Pro první kolo reakce byly použity primery Splink1 5'-CGAAGAGTAACCGTTGCTAGGAGAGACC-3' a spPCR-AG3-R 5'-CCACCTTACTTCCACCAATCGGCATG-3' při použití vektoru AG3, nebo Splink1 a spPCR-HIV-1-R 5'-TATCTGATCCCTGGCCCTGGTGTGTAG-3' při použití vektoru HIV-1. Pro druhé kolo amplifikace byly použity primery Splink2 5'-GTGGCTGAATGAGACTGGTGTGCGAC-3' a spinPCR-AG3-R 5'-CGGTGCTTTTTCTCTCCTTGTAAGGCA-3' při použití vektoru AG3, nebo Splink2 a spinPCR-HIV-1-R 5'-CTGCCAATCAGGGAAGTAGCCTTGTGTG-3' při použití vektoru HIV-1.

Složení reakcí a programy pro nested PCR:

Splink1 PCR

DNA	90 ng
Pufr k DNA polymeráze (10x)	5,0 μ l
dNTP (2,5 mM každý nukleotid)	6,0 μ l
Splink1 (10 μ M)	1,5 μ l
spAG3-R nebo spHIV-1-R (10 μ M)	1,5 μ l
DNA polymeráza*	1 U
H ₂ O	ad 50 μ l

	2 cykly	32 cyklů	
94°C 3 min	94°C 5 sec 68°C 30 sec 72°C 2 min	94°C 15 sec 61°C 30 sec 72°C 2 min	72°C 5 min

Splink2 PCR

Splink1 PCR reakce	2,0 µl
Pufř k DNA polymeráze (10x)	5,0 µl
dNTP (2,5 mM každý nukleotid)	6,0 µl
Splink2(10 µM)	1,5 µl
spinAG3-R nebo spinHIV-R (10 µM)	1,0 µl
DNA polymeráza*	1 U
H ₂ O	ad 50 µl

	32 cyklů	
94°C 3 min	94°C 15 sec 60°C 30 sec 72°C 2 min	72°C 5 min

* jako DNA polymerázy byly použity polymerázy rTaq DNA polymeráza (TaKaRa) a BioTherm DNA polymeráza (GeneCraft).

4.11. Purifikace a sekvenace produktů splinkerette PCR

4.11.1. Agarózová elektroforéza

Agarózová elektroforéza byla využita k ověření a identifikaci produktů splinkerette PCR. Vzorby byly nanášeny a následně migrovány v prostředí 1 - 2% (w/v) agarózového gelu (Agarose SERVA Premium rozpuštěná v 1x TAE pufřu) s obsahem ethidium bromidu (0,5 µg/ml). Proužky DNA byly foceny na prohlížečce vybavené UV lampou a fotoaparátem (Olympus SP-500 UZ).

4.11.2. Purifikace DNA před sekvenační reakcí

Produkty splinkerette PCR byly přečištěny pomocí komerčního kitu (USB® ExoSAP-IT® PCR Product Cleanup) dle návodu výrobce, kdy jsou k 5 µl PCR reakce přidány 2 µl směsi enzymů (exonukleáza I a alkalická fosfatáza) v pufřu. Tato směs slouží

k odstranění přebytečných volných deoxynukleosidtrifosfátů (dNTP) a primerů. Reakce pak byla inkubována 15 min při 37°C a následně 15 min při 80°C.

4.11.3. Extrakce DNA z gelu

V případě přítomnosti více proužků DNA či „smírů“ byla DNA pro účely sekvenace purifikována extrakcí z agarózového gelu. Agarózový gel obsahující separační agarózu (UltraPure™ Agarose, Invitrogen v 1x TAE pufru) byl připraven stejným způsobem, jaký je popsán v kapitole **Agarózová elektroforéza**. Příslušné proužky DNA byly z gelu vyřezány a DNA byla extrahována z gelu pomocí komerčního kitu dle návodu výrobce (QIAEX II Gel Extraction Kit, Qiagen). DNA byla ze silikonových kuliček eluována v 10 µl H₂O, z čehož bylo 7 µl odebráno pro sekvenaci.

4.11.4. Sekvenování místa integrace

Purifikovaná DNA byla sekvenována Sangerovou dideoxy metodou s využitím komerčního kitu (BigDye® Terminator v3.1 Cycle Sequencing Kit, Applied Biosystems®). Reakce se skládala ze 7 µl DNA, 2 µl sekvenačního kitu a 1 µl spinPCR-AG3-R nebo spinPCR-HIV-1-R primeru (10µM). Program sekvenační reakce: 96°C po dobu 20 s, 35 cyklů 96°C 20 min, 58°C 10 min, 60°C 4 min.

4.11.5. Purifikace produktu sekvenace

Po sekvenační reakci byla DNA opět purifikována pro analýzu sekvence pomocí vysrážení DNA ethanolem. K 10µl sekvenační reakci bylo přidáno 80 µl vody, 10 µl octanu sodného, 320 µl 96% ethanolu a 5 mg glykogenu pro zviditelnění výsledného peletu. Reakce se po promíchání nechala 5 min odstát. Následovala centrifugace (16 100 x g, 15 min 15°C), odsátí supernatantu a znovu centrifugace v 70% ethanolu. Po odsátí supernatantu byl pelet vysušen ve 37°C.

4.12. Aktivní a umlčená integrační místa

4.12.1. Izolace a purifikace DNA

Buňky separované na GFP⁺ a GFP⁻ populace byly zpracovány podle kapitol **Štěpení genomové DNA až Agarózová elektroforéza**. Ze separačního agarózového gelu byly z každého vzorku (GFP⁺ a GFP⁻) vyřezány proužky obsahující DNA fragmenty (produkty Splink2 PCR) o velikostech 250 - 800 bp. DNA byla z gelu extrahována dle

postupu popsaného v kapitole **Extrakce DNA z gelu**. DNA byla eluována 13 μl vody, z čehož bylo 11 μl odebráno pro další zpracování.

4.12.2 Klonování do vektoru

DNA izolována v kroku **Izolace a purifikace DNA** byla klonována do vektoru pGEM pomocí klonovacího kitu pGEM[®]-T Easy (Promega). PCR produkty byly do vektoru ligovány v 15 μl reakci dle návodu výrobce. Reakce obsahovala 5,5 μl Splink2 PCR produktu. Ligace probíhala přes noc při 4°C.

4.12.3. Transformace bakterií

Bakterie Epicurian Coli[®] XL1-Blue (50 μl) byly transfekovány metodou tepelného šoku. Ligační reakce (15 μl) byla smíchána s bakteriemi a inkubována 30 min na ledě. Poté byla bakteriální směs inkubována 1 min při 42°C ve vodní lázni. Následně byla bakteriální směs opět přemístěna na led (4°C), kde byla udržována po dobu 1 min, po které byla každá transformační reakce rozdělena do dvou zkumavek obsahujících 600 μl LB media bez AMP. V tomto médiu byly bakterie kultivovány 45 min při 37°C. Nakonec byly bakterie rozetřeny na misky (100 mm) s agarem obsahující 100 $\mu\text{g/ml}$ ampicilinu, X-gal (50 μl) a IPTG (4 μl). Z každé zkumavky byly naočkovány dvě misky - se 100 a 200 μl bakteriálního inokula.

4.12.4. Výběr a sekvenace bakteriálních kolonií

Kolonie pozitivní na přítomnost vektoru s vloženým Splink2 PCR fragmentem (bílé kolonie) byly vybrány a odeslány na sekvenaci do GATC Biotech (<http://www.gatc-biotech.com/en/home>). Takto bylo ke sekvenaci odesláno 45 klonů s fragmenty z GFP⁺ buněk a stejný počet klonů s fragmenty z GFP⁻ buněk.

4.12. Mapování sekvencí integračních míst do lidského genomu

4.12.1. Zpracování sekvence

Získané sekvence by měly v ideálním případě obsahovat část 5'LTR, genomovou sekvenci a sekvenci Splink2 primeru. Pro to, aby byla sekvence dále mapována, bylo podmínkou, aby genomová sekvence přímo přiléhala k poslednímu nukleotidu 5'LTR (tedy k sekvenci 5'-GACTACA-3' u AG3). Délka příslušné genomové sekvence pak byla

určena nejbližším místem rozeznávaným použitou restriční endonukleázou (5'-GATC-3' u AG3) či maximální čitelnou délkou produktu sekvenace.

4.12.2 Mapování sekvencí do lidského genomu

K mapování sekvencí integračních míst retrovirů do lidského genomu byl využit BLAT (<http://genome.ucsc.edu/cgi-bin/hgBlat?command=start>) na webových stránkách UCSC Genome Bioinformatics. Pro mapování byla využita nejnovější verze genomu z února 2009 (GRCh37/hg19). Minimální délka vložené sekvence byla 21 nukleotidů (nt).

4.12.3 Lokalizace místa integrace

Za integrační místo byla považována taková sekvence, která vytvořila jedinečnou shodu s referenčním genomovým lokusem. Sekvence musela vykazovat vysokou identitu a *START* shody musel odpovídat pozici 1 mapované sekvence. Velmi důležitým ukazatelem pak bylo skóre, které muselo být mezi ostatními výsledky jedinečné. Místu integrace (iPos) pak byla přidělena souřadnice odpovídající chromozomální a nukleotidové pozici prvního nukleotidu přiléhajícího k 5'LTR (tedy pokud byl *STRAND* +, byla pozice určena polohou nukleotidu *START*, v případě, že byl *STRAND* -, odpovídalo iPos pozici nukleotidu *END*).

4.12.4. Integrační místa preference ASLV

Integrační místa (226) odpovídající preferenci integrace ASLV (Narezkina et al., 2004) byla získána z databáze GenBank (AY653309 až AY653534). Z databáze byly získány přímo souřadnice míst integrace. Ty byly převedeny na souřadnice aktuální verze genomu (z verze hg16 na verzi hg19) pomocí nástroje LiftOver na webových stránkách UCSC Genome Bioinformatics (<http://genome.ucsc.edu/cgi-bin/hgLiftOver?hgid=316086979>).

4.13. Bioinformatická analýza integračních míst

4.13.1. Použitý software

Pro bioinformatickou analýzu byl využíván Table Browser (<http://genome.ucsc.edu/cgi-bin/hgTables?command=start>) na webových stránkách UCSC Genome Bioinformatics a program MySQL 5.5 Command Line Client. Pro grafickou

interpretaci dat a statistické testování pak byl využíván program R 2.15.1. Všechny tyto programy běžely na operačním systému Windows 7.

4.13.2. Původ dat

Data pro analýzu integračních míst byla získána z databáze UCSC Genome Bioinformatics (<ftp://hgdownload.cse.ucsc.edu/goldenPath/hg19/database/>). Názvy konkrétních souborů jsou uvedeny v **tabulce 2**. Vždy byly získány dva soubory, jeden ve formátu .txt.gz (data), a druhý ve formátu .sql (struktura datové tabulky pro MySQL).

Tabulka 2. Seznam tabulek využitých pro analýzu integrčních míst.

Zdroj dat pro	Záznam („track“)	název tabulky (.txt)
Transkripční jednotky	UCSC Genes	knownGene
Transkripční starty	SwitchGear TSS	switchDbTss
CpG ostrovy	CpG Islands	cpGIslandExt
G-pruhování	Chromosome Band	cytoBand
Repetitivní elementy	RepeatMasker	rmsk
H2A.Z	Broad Histone	wgEncodeBroadHistoneK562H2azStdPk
H3K4me1	Broad Histone	wgEncodeBroadHistoneK562H3k4me1StdPk
H3K4me2	Broad Histone	wgEncodeBroadHistoneK562H3k4me2StdPk
H3K4me3	Broad Histone	wgEncodeBroadHistoneK562H3k4me3StdPk
H3K9ac	Broad Histone	wgEncodeBroadHistoneK562H3k9acStdPk
H3K9me3	Broad Histone	wgEncodeBroadHistoneK562H3k9me3StdPk
H3K79me2	Broad Histone	wgEncodeBroadHistoneK562H3k79me2StdPk
H4K20me1	Broad Histone	wgEncodeBroadHistoneK562H4k20me1StdPk
H3K27me3	Broad Histone	wgEncodeBroadHistoneK562H4k27me3StdPk
H3K27ac	Broad Histone	wgEncodeBroadHistoneK562H4k27acStdPk
H3K36me3	Broad Histone	wgEncodeBroadHistoneK562H3k36me3StdPk

Tabulky jsou k dispozici na adrese <ftp://hgdownload.cse.ucsc.edu/goldenPath/hg19/database/>. Tabulky byly staženy a data z nich byla zpracována pomocí programu MySQL 5.5 Comand Line Client.

4.13.3. Charakterizace místa integrace

Genomové a epigenomové prvky přítomné v místě integrace a jeho okolí byly nejprve zkontrolovány vizuálně pomocí grafického rozhraní Genome Browser na webových stránkách UCSC Genome Bioinformatics (<http://genome.ucsc.edu/cgi-bin/hgGateway>). Další charakterizace byla prováděna pomocí MySQL tabulek a aplikace Table Browser na webových stránkách UCSC Genome Bioinformatics (viz **Původ dat**).

Chromozomální pozice a pozice vzhledem ke G pruhování chromozómů

Chromozomální pozice byla vizualizována na pozadí lidského karyotypu pomocí webových stránek databáze Ensembl (<http://www.ensembl.org/index.html>). Pozice vzhledem ke G pruhování byla získána z dat Chromosome Band přístupných na webových stránkách UCSC Genome Bioinformatics. Byla-li pozice integrace (iPos) přítomna mezi pozicemi *chromStart* a *chromEnd* G negativního (*gneg*) pruhu, byla tato integrace považována za přítomnou mimo oblasti G pozitivní a označena jako integrace do pruhu R.

Transkripční jednotky (TU) a transkripční starty (TSS)

Za TU je zde považována jakákoliv transkribovaná oblast zaznamenaná v datech UCSC Genes (jde o data získaná složením databází RefSeq, GenBank, CCDS, UniProt, Rfam a tRNA Genes). Tato definice by měla postihovat místa transkripce jak kódujících, tak nekódujících RNA. Integrační místo bylo považováno za přítomné v intraTU oblasti tehdy, nacházelo-li se iPos mezi pozicemi *txStart* a *txEnd* jakékoliv TU zaznamenané v záznamu UCSC Genes. Místa TSS pak byla získána ze záznamu SwitchGear TSS. TSS jsou v tomto záznamu určovány dle tzv. genových modelů. Ke každému genovému modelu je přiřazeno několik TSS a tyto TSS jsou seřazeny dle jejich validity. Vzdálenost od TSS pak byla vypočítána jako vzdálenost od *chromStart* nejbližšího TSS, jenž byl nejpravděpodobnějším TSS pro daný genový model (R1 TSS). Složením informací získaných ze SwitchGear a UCSC Genes pak byla určena pozice iPos vzhledem k TU.

CpG ostrovy

Data pro CpG ostrovy byla získána ze záznamu CpG Islands. Vzdálenost od CpG ostrova byla spočítána jako vzdálenost od nejbližší *chromStart* či *chromEnd* genomové

pozice CpG ostrova od iPos. Pokud se iPos vyskytovala mezi pozicemi *chromStart* a *chromEnd* jednoho CpG ostrova, byla vzdálenost od CpG ostrova 0.

Repetitivní elementy

Data pro přítomnost repetitivních elementů byla získána ze záznamu RepeatMasker. Integrace byla považována za přítomnou uvnitř repetitivního elementu, byla-li iPos mezi pozicemi *genoStart* a *genoEnd* repetitivního elementu. Do souboru provirů integrovaných do repetitivních elementů byly zahrnuty i ty proviry, jejichž přesnou pozici nebylo možno díky integraci do oblasti repetitivního elementu přesně určit.

Histony, histonové modifikace

Data pro proteinové faktory histonové modifikace byla získána ze záznamů Broad Histone. Jedná se o data, která byla získána pomocí metody Chip-Seq. Data byla zpracovávána ve formě ohraničených oblastí s navýšeným signálem dané modifikace, tj. oblastí obohacených o danou histonovou modifikaci. Pokud byla iPos mezi *chromStart* a *chromEnd* pozicemi oblasti faktoru, byla integrace považována za přítomnou uvnitř oblasti a její vzdálenost od histonové modifikace byla 0.

4.14. RT-qPCR analýza cílených TU

U vybraných TU, do kterých byla cílená integrace, byla provedena analýza transkripční aktivity pomocí RT-qPCR. Analýza byla provedena na RNA izolované z neinfikovaných buněk linie K562.

4.14.1. Izolace RNA

Buňkám bylo odsáto médium a byl na ně přidán RNazol RT (Biotech) v objemu 1 ml. Buňky byly seškrábnuty a přeneseny do 2ml mikrozkuřavky. Dále bylo k buňkám přidáno 400 μ l vody. Buněčná směs byla řádně promíchána a poté ponechána stát při pokojové teplotě po dobu 15 min. Po uplynutí této doby byl 1 ml supernatantu přenesen do čisté mikrozkuřavky a RNA byla vysrážena pomocí přidání 400 μ l 75% ethanolu. Směs byla řádně promíchána a ponechána stát 10 min při pokojové teplotě. Následovala centrifugace (12000 x g, 10 min, 15°C). Vzniklý pelet byl ještě dvakrát opláchnut 400 μ l 75% ethanolem a centrifugován (8000 x g, 5 min, 15°C). Precipitát byl poté vysušen při pokojové teplotě a nakonec rozpuštěn v 50 μ l vody.

Koncentrace RNA byla změřena na nanofotometru (IMPLEN). Pro kontrolu byly 2 µg RNA naneseny na 1% agarózový gel. Takto izolovaná RNA byla skladována při -80°C.

4.14.2. Odstranění DNA

Před vytvořením cDNA byla směs izolované RNA podrobena působení DNázy. Směs obsahovala 1 µg RNA, 2 U DNázy (New England Biolabs), 10 µl 5 x koncentrovaného M-MLV reakčního pufru (Promega) a vodu (ad 50 µl). Po celou dobu přípravy směsi byla RNA udržována na ledě. Reakce pak byla 15 min při 37°C a následně 5 min při 70°C.

4.14.3. Vytvoření cDNA – RT PCR

Pro tvorbu cDNA byl použit 1 µg RNA a bylo přidáno 5 µl směsi dNTP (dATP, dTTP, dGTP, dCTP, 10 µM každý), 2 µl směsi náhodných hexanukleotidů (pdN6, Promega) sloužících jako primery, 10 U M-MLV reverzní transkriptázy (Promega) a voda (ad 50 µl). Reverzní transkripce byla prováděna při programu: 37°C 60 min, 70°C 10 min. Výsledná směs cDNA byla naředěna 50 µl vody a skladována při -20°C.

4.14.4. qPCR

Pro qPCR analýzu byl využit CFX96™ Real-Time PCR Detection System (Bio-Rad). Každý vzorek byl analyzován v triplikátu. Každá reakce obsahovala 2 µl cDNA, 10 µl 2x MESA GREEN qPCR MasterMix Plus for SYBR® Assay kitu (Eurogentec), po 0,5 µl každého z primerů (10 µM) a vodu (ad 20 µl). Reakce běžela při programu: 95°C po 5 min, 40 cyklů 95°C 15 s, 63°C 20 s, 72°C 25 s, poté 72°C 10 min, 63°C 5 s. Následně byla měřena denaturační křivka, kdy byla teplota zvyšována po 0,5°C na 95°C, přičemž byla při každém zvýšení měřena fluorescence (5 s). Program byl zakončen polymerací při 72°C po dobu 8 min.

Použité primery a cílové TU jsou vypsány v **tabulce 3**. Primery pro qPCR analýzu byly vybrány z databáze PrimerBank (Wang et al., 2012b). K analyzovaným TU byl přidán pro normalizaci transkripce také konstitutivně transkribovaný (housekeeping) gen pro TATA box vazebný protein (TBP). TBP je v Laboratoři virové a buněčné genetiky AV ČR, běžně využívaným referenčním genem pro relativní kvantifikaci transkripce analyzovaných genů. Kvantifikace exprese TU pak byla vztahována k množství TBP.

Analyzovaný soubor také obsahoval vzorky bez přidané cDNA (blank) jako kontrolu kontaminace.

Tabulka 3. Názvy cílených TU a PrimerBank ID primerů použitých pro RT-qPCR.

TU (GeneSymbol)	PrimerBank ID		
ARRB1	320461704c3	DGCR8	298358603c1
CEP128	194018736c1	FEM1B	52851431c1
CFLAR	187608576b1	KLF16	221219061c1
COQ4	166795284c3	NACA	333033786c1
		SUV39H1	112382253c1
		TRAPPC10	183227692c1

Názvy TU jsou uvedeny ve formátu geneSymbol. Konkrétní sekvence primerů spolu s jejich charakteristikami je možné nalézt podle PrimerBank ID na webové adrese <http://pga.mgh.harvard.edu/primerbank/index.html>.

5. VÝSLEDKY

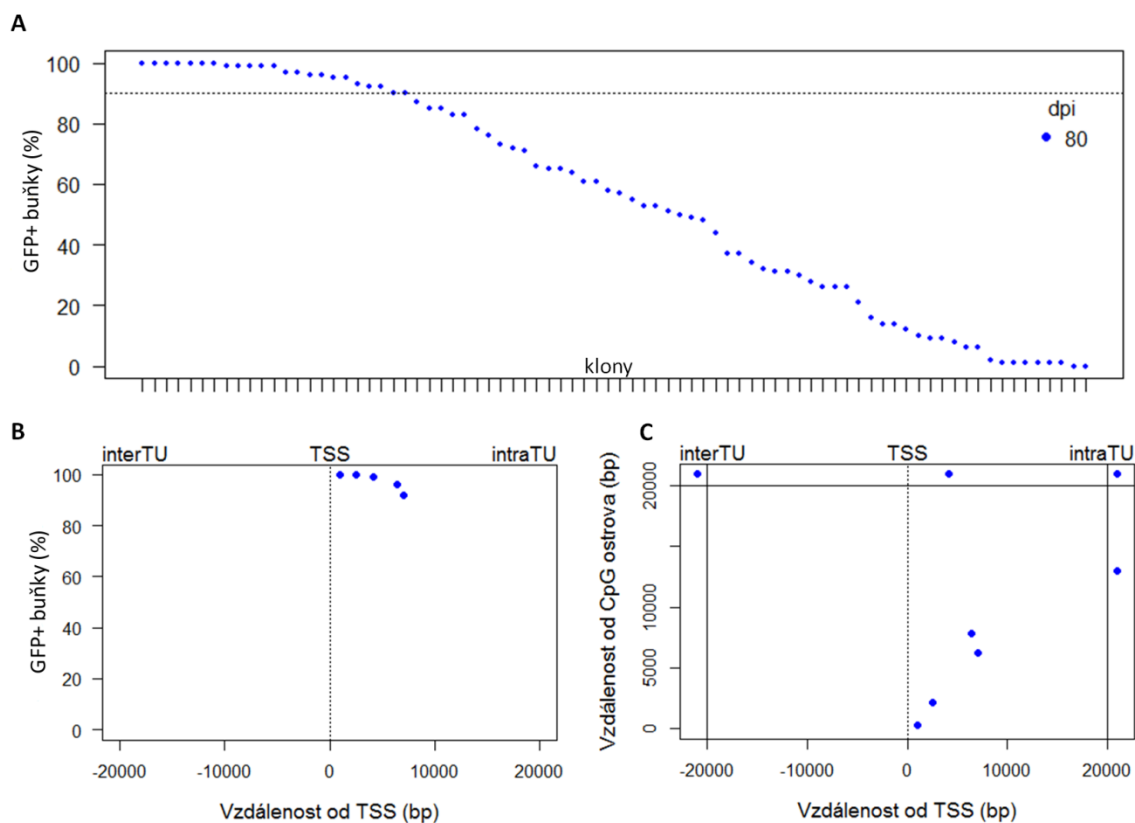
5.1. Exprese a integrace provirů ASLV v buňkách HEK293

Lidská buněčná linie HEK293 byla infikována při MOI 0,2 vektorem AG3 a 3 dpi byly separovány GFP⁺ buňky. Po hromadné kultivaci byly 44 dpi pomocí FACS GFP⁺ buňky separovány po jedné a jejich kultivací v oddělených jamkách byly ustanoveny buněčné klony. Procento GFP⁺ buněk bylo stanoveno 80 dpi u 80 buněčných klonů (**Obr. 11A**). U 23 klonů bylo pozorováno alespoň 90 % GFP⁺ buněk v klonální populaci. Tyto klony byly považovány za stabilní a byly vybrány pro analýzu místa integrace provirů. Z 23 stabilně exprimujících klonů se metodou splinkerette PCR podařilo identifikovat celkem 10 unikátních integračních míst provirů ASLV. U ostatních provirů se buď nezdařila amplifikace sekvence místa integrace, bylo u klonů nalezeno více než jeden provirus, nebo různé klony obsahovaly provirus integrovaný v totožném místě. Klony s více proviry či s proviry v již identifikovaném místě byly z analýzy vyřazeny. Identifikované sekvence byly mapovány pomocí nástroje BLAT na webových stránkách UCSC Genome Bioinformatics (<http://genome.ucsc.edu/>). Místa integrace pak byla analyzována vzhledem k přítomnosti TU, TSS a CpG ostrovů (**Obr. 11B, 11C**). Vzhledem k absenci většího souboru epigenetických dat specifických pro linii HEK239, nebyla místa integrace detailně analyzována pro přítomnost epigenetických značek.

Z 10 provirů bylo 9 integrováno do TU. Celkem 5 provirů bylo nalezeno v okolí 20 kbp od nejbližšího TSS. Jejich vzdálenost od TSS se pohybovala od 1 do 7 kbp. Tytéž proviry byly nalezeny poblíž oblastí CpG ostrovů (vzdálenost od 300 bp po 13 kbp). U ostatních 5 integračních míst nebyly identifikovány žádné TSS ani CpG ostrovy v blízkém okolí (20 kbp). I přes malé množství izolovaných integračních míst, mohly tyto výsledky naznačovat neúplnou závislost stabilní exprese proviru na integraci do TU a poblíž TSS a CpG ostrovů. Byly identifikovány stabilně exprimující proviry (97 % a 99 % GFP⁺ buněk) integrované v TU, ale ve vzdálenosti více než 20 kbp od TSS či CpG ostrova. U jednoho proviru (exprese 92 %) pak nebyla ve vzdálenosti 20 kbp identifikována ani jedna ze zmíněných charakteristik.

Ačkoliv tato data vykazují odchylku od již publikovaných dat, není kvůli malému počtu identifikovaných integračních míst možné určit, zda jde o zcela výjimečná

integrační místa či pravidlo. Nízký počet identifikovaných míst integrace tak znemožňuje detailnější statistickou charakteristiku. Jedním z důvodů nízkého výtěžku integračních míst byl nejspíše i způsob kultivace klonů, kdy byly buňky po dlouhou dobu (asi 40 dní) kultivovány ve směsné populaci. Na tento fakt poukazuje častý výskyt shodných integračních míst u různých klonů. Tyto výsledky také poukázaly na vzácný výskyt stabilně exprimujících provirů ASLV v lidských buňkách a nutnost analýzy většího množství klonů. Pro detailní analýzu míst integrace ASLV pak chyběla epigenetická data specifická pro tuto linii. Tato fakta podpořila výběr jiné lidské buněčné linie vhodnější pro analýzu integračních míst, kterou se stala linie K562. Vzhledem k malému počtu míst a k možné roli faktorů specifických pro danou linii, nebyla data získána na linii HEK293 použita v další analýze.



Obrázek 11. Expres a integrační místa ASLV v HEK293. **A)** Expres provirů ASLV byla měřena u 80 buněčných klonů lidské linie HEK293. Procento buněk exprimujících GFP bylo stanoveno 80 dpi. Znáznorněny jsou všechny tyto klony. Klony jsou ve sloupcích seřazeny sestupně dle množství buněk exprimujících GFP. Přerušovaná čára znázorňuje hranici 90 %. **B)** Graf znázorňuje klony nesoucí proviry integrované do vzdálenosti 20 kbp od TSS. Přerušovaná čára znázorňuje TSS. Napravo je znázorněna oblast TU (intraTU), nalevo pak oblast mimo TU (interTU). Osa y znázorňuje množství buněk exprimujících GFP pro dané klony. **C)** Vzdálenost od TSS a CpG ostrovů jednotlivých provirů. Oblasti za hranicemi 20000 bp znázorňují proviry integrované ve vzdálenosti vyšší než 20 kbp od TSS či CpG ostrova. Provirus znázorněný v levém horním rohu je tedy vzdálený od nejbližšího CpG ostrova i TSS více než 20 kbp. bp – páry bazí, dpi – dny po infekci, TSS – místo transkripčního startu, TU – transkripční jednotka.

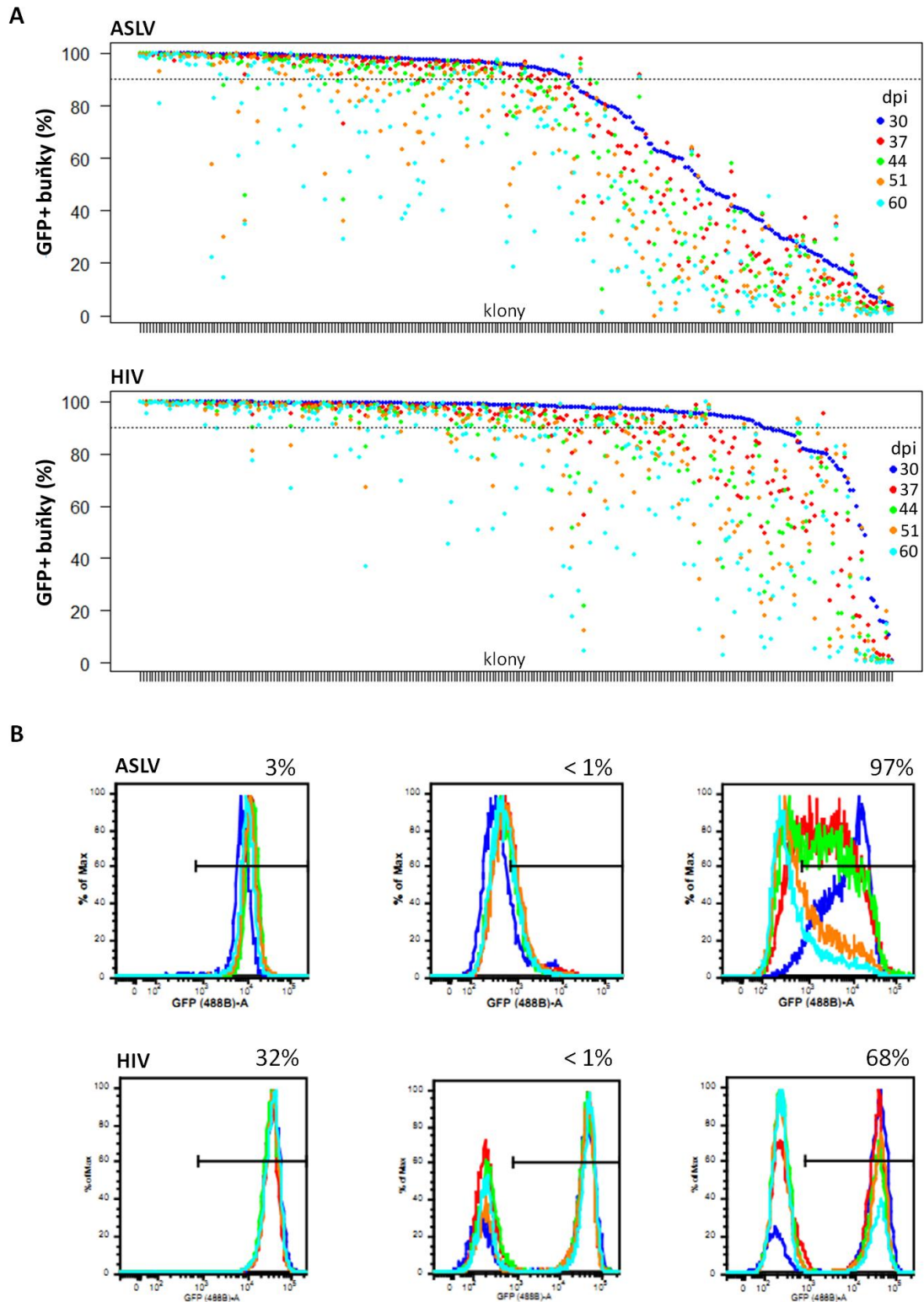
5.2. Exprese provirů ASLV a HIV-1 v linii K562

Lidská buněčná linie K562 byla pro charakterizaci integračních míst stabilně exprimujících provirů ASLV vybrána z důvodů dobré charakterizace z pohledu funkční genomiky a epigenetických modifikací a pro možnost snadné kultivace mnoha paralelních klonů (jedná se o buňky rostoucí v suspenzní kultuře). Buňky linie K562 byly infikovány retrovirovými vektory AG3 a LTR-Tat-GFP při MOI přibližně 0,2. Nízká MOI zvyšuje pravděpodobnost integrace pouze jednoho proviru na buňku. LTR-Tat-GFP je retrovirový vektor odvozený od HIV-1 (Jordan et al., 2001). Kromě GFP nese LTR-Tat-GFP i virový transaktivátor Tat. Tento vektor byl použit za účelem porovnání s ASLV a analýzy závislosti exprese na místě integrace u HIV-1 s přirozenou regulací exprese (přítomnost Tat proteinu). Vektor LTR-Tat-GFP by měl oproti vektoru AG3 vykazovat preferenční integraci do aktivních TU.

Z infikované buněčné kultury byly 3 dpi separovány GFP⁺ buňky pomocí FACS a následně kultivovány až do 60 dpi. Exprese GFP u jednotlivých klonů byla počínaje 30 dpi každý týden měřena pomocí průtokové cytometrie. Tím byla získána informace o dynamice umlčování exprese provirů v jednotlivých klonech a také informace o průměrné intenzitě exprese (MFI) klonů. Soubor charakteristik popisujících expresi GFP konkrétního klonu v době 30 – 60 dpi je zde označován jako expresní profil klonu. Klony byly po separaci i v době měření selektovány na expresi GFP. Klony vykazující expresní profily umlčování exprese GFP byly odstraňovány. Pro získání celkového přehledu o expresní aktivitě provirové populace však byly stanoveny expresní profily také u některých klonů s umlčovanými proviry (**Obr. 12A**). Celkově byly takto získány expresní profily 245 klonů s ASLV a 255 klonů s proviry HIV-1. Z 245 expresně charakterizovaných klonů ASLV bylo identifikováno 57 klonů obsahujících 60 dpi alespoň 90 % GFP⁺ buněk. Tento počet odpovídá zhruba 3 % z asi 1900 klonů získaných separací GFP⁺ buněk 3 dpi. To podporuje vzácnost stabilně exprimovaných provirů ASLV při infekci savčích buněk. Dále byly nalezeny klony s méně než 90 % GFP⁺ buněk 60 dpi, které však tento poměr udržovaly po celou dobu sledování expresních profilů klonů. Výskyt těchto klonů, jejichž počet GFP⁺ buněk často nepřekračoval 50 %, byl ještě vzácnější, a pohyboval se pod hodnotou 1 % separované populace. Většina klonů (97 % z klonů separovaných 3 dpi) s proviry ASLV pak vykazovala klesající počty GFP⁺ buněk, které v rámci klonu klesly 60

dpi pod 90 % a vykazovaly tedy znaky výrazného umlčování provirové exprese. Stabilně exprimující proviry byly selektovány také u klonů s proviry vektoru LTR-Tat-GFP. Již z **obr. 12A** je patrné, že vektor LTR-Tat-GFP vykazoval oproti vektoru AG3 vyšší stabilitu exprese. Vysoce stabilní expresi, tedy přítomnost více než 90 % GFP⁺ buněk, vykazovalo celkem 129 klonů, což odpovídá asi 32 % populace GFP⁺ buněk separovaných 3 dpi. Byly identifikovány i klony obsahující GFP⁺ buňky, jejichž počet byl nižší než 90 % avšak stabilní. Stejně jako u ASLV, byl výskyt těchto klonů velice vzácný a odpovídal méně než 1 % klonů separovaných 3 dpi. Na rozdíl od ASLV byly u těchto klonů patrné dvě zcela oddělené populace GFP⁺ a GFP⁻ buněk (**Obr. 12B**). Zcela oddělené populace GFP⁺ a GFP⁻ buněk byly pozorovatelné také u klonů nesoucích umlčované proviry.

Po identifikaci stabilně exprimujících provirů byl analyzován vztah mezi stabilitou exprese a intenzitou exprese u jednotlivých klonů. MFI GFP⁺ buněk se mezi jednotlivými klony HIV-1 mohla lišit až 60 x, u klonů ASLV pak až 70 x. **Obr. 13A** znázorňuje vztah stability exprese vyjádřené procenty exprimujících buněk klonu a MFI celkové populace či pouze GFP⁺ buněk klonu v době 60 dpi. ASLV vykazuje postupný pokles MFI s poklesem výskytu exprimujících buněk jak u celkové populace, což je ovlivněno výskytem GFP⁻ buněk, tak u populace GFP⁺ buněk (srovnání klonů nad a pod 50 % exprimujících buněk, Studentův t-test, $p < 2,2 \times 10^{-16}$). Klony vykazující více než 50 % GFP⁻ buněk tak obsahují proviry, jejichž exprese je slabší než exprese provirů u klonů s více než 50 % GFP⁺ buněk. Tento fakt poukazuje na postupné umlčování (tedy postupné vyhasínání GFP) provirů ASLV v savčích buňkách. U HIV-1 je však situace jiná. K umlčování provirů zde se snižující se MFI nedochází postupně jako u ASLV, ale skokově. Pokud jsou mezi jednotlivými klony porovnány pouze GFP⁺ buňky, je průměrná exprese podobná u klonů s méně či více než 50 % exprimujících buněk (Studentův t-test, $p = 0,6459$, průměrná MFI byla 27721 a 28748).

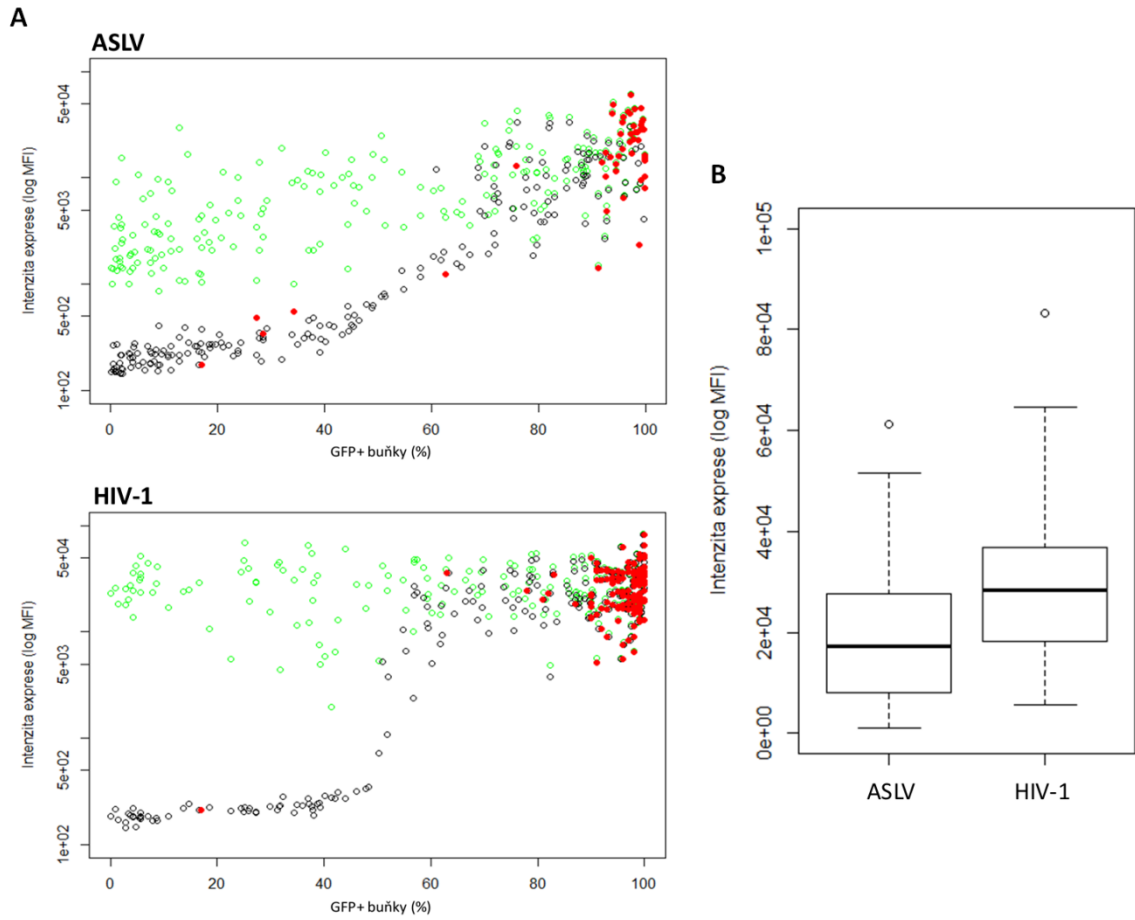


Obrázek 12. Expres ASLV a HIV-1 v buňkách K562. A) Znáznornění exprese buněčných klonů nesoucích proviry ASLV (245) a HIV-1 (255) v době 30 – 60 dpi. Jednotlivé klony jsou seřazeny ve sloupcích podél vodorovné osy. Jednotlivé časy, kdy byl stanovován poměr buněk exprimujících GFP, jsou odlišeny barevně: 30 dpi (tmavě modrá), 37 dpi (červená), 44 dpi (zelená), 51 dpi (žlutá), 60 dpi (světle modrá). Přerušovanou čarou je znázorněna hranice 90 %. **B)** Znáznornění různých skupin klonů dle jejich expresních profilů. Histogramy znázorňují relativní počet buněk na ose y a intenzitu fluorescence na ose x. Je znázorněna i oblast, ve které byly buňky považovány za pozitivní pro expresi GFP. Skupiny podle

expresních profilů jsou seřazeny zleva následujícím způsobem: vysoce stabilně exprimující, nízké stabilně exprimující, umlčované. Čísla nad histogramy vyjadřují, jakou část buněk tyto klony představují z populace separované 3 dpi (tedy z buněk, které byly 3 dpi pozitivní na expresi GFP). Barevné rozlišení histogramů odpovídá časovému rozlišení jednotlivých měření v grafu **A**). dpi – dny po infekci, ASLV – ptačí virus způsobující sarkomy a leukózy, HIV-1 – virus lidské získané imunodeficiency.

Porovnána byla také intenzita exprese ASLV a HIV-1. K tomuto srovnání byly vybrány stabilně exprimující klony. Jak ukazuje **obr. 13B**, stabilně exprimované proviry HIV-1 vykazují vyšší intenzitu exprese než proviry ASLV (Studentův t-test, $p = 3,78 \times 10^{-5}$).

Výsledky expresní analýzy buněčných klonů lidských buněk s proviry ASLV a HIV-1 ukazují na rozdílnou četnost výskytu stabilně exprimovaných provirů u obou retrovirů. U HIV-1 je výskyt takových provirů asi 10 x častější než u ASLV v lidských buňkách. Heterogenita mezi jednotlivými proviry téhož retroviru v rámci jednoho buněčného systému poukazuje na vliv místa integrace na expresi proviru. Tato heterogenita se projevuje nejen různou stabilitou exprese proviru, ale také intenzitou exprese různě integrovaných provirů. Analýza intenzity exprese také poukazuje na různou dynamiku umlčování provirové exprese u ASLV a HIV-1. Variabilní intenzita exprese je však pozorovatelná nejen mezi různě integrovanými proviry, ale také v rámci jednoho klonu mezi postupně umlčovanými proviry ASLV. Tento jev však není pozorovatelný u HIV-1. Buňky klonů s proviry vektoru LTR-Tat-GFP vykazovaly buď stav plné exprese GFP omezené pravděpodobně jen vlivem místa integrace, nebo stav úplného umlčení bez výskytu přechodných intenzit exprese GFP. V celkové klonální populaci tak byl na rozdíl od ASLV, u umlčovaných provirů pozorován skokový přechod od maximální k minimální MFI.



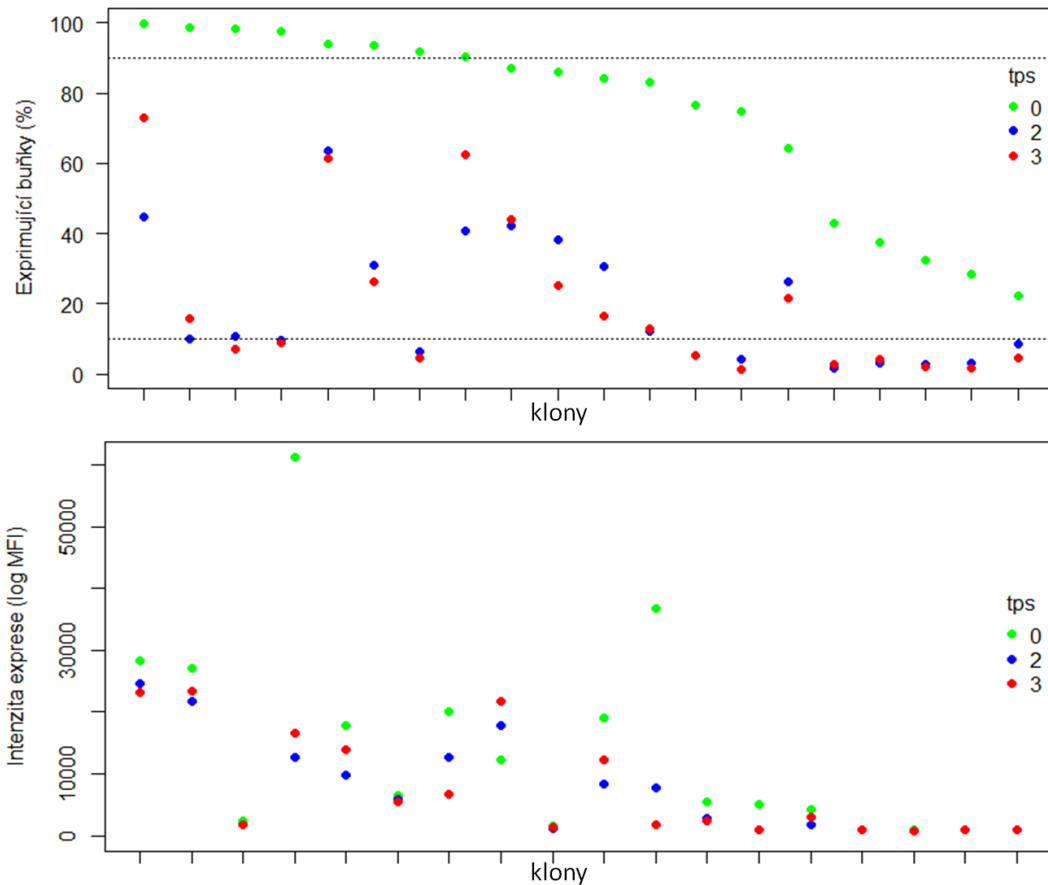
Obrázek 13. Stabilita a intenzita exprese ASLV a HIV-1 v buňkách K562. A) Na grafech jsou znázorněny všechny klony nesoucí proviry ASLV či HIV-1, u nichž byla sledována expresní aktivita provirů. Znázorněné výsledky pocházejí z měření 60 dpi. Černě jsou znázorněny intenzity exprese celých populací buněčných klonů (pozitivní i negativní buňky), zatímco zeleně jsou znázorněny intenzity exprese pouze pozitivních buněk klonu. Červenými body jsou pak zvýrazněny ty klony, které byly považovány za stabilně exprimující. Intenzity exprese jsou znázorněny v MFI (logaritmičká škála). **B)** Krabicové diagramy intenzity exprese v MFI (logaritmičká škála) stabilně exprimujících klonů K562 nesoucích ASLV či HIV-1. MFI odpovídá intenzitě klonů v 60 dpi. MFI – průměrná intenzita fluorescence. ASLV – ptačí virus způsobující sarkomy a leukózy, HIV-1 – virus lidské získané imunodeficiency typu 1.

5.3. Nestabilní umlčení provirů ASLV v lidských buňkách

Yao et al. (2004) ve své práci poukázali na možnost spontánní reaktivity provirů MLV integrovaných v myších embryonálních kmenových buňkách. K zodpovězení otázky, zda je spontánní reaktivity exprese již jednou umlčených provirů možná i v diferencovaných lidských buňkách, byl využit systém buněk K562 s integrovanými proviry ASLV.

Po dokončení analýzy exprese klonů, byly z 18 vybraných klonů separovány pomocí FACS GFP⁻ buňky. Takto ustanovené negativní subpopulace byly posléze analyzovány na přítomnost reaktivovaných GFP⁺ buněk. **Obr. 14** ukazuje expresi těchto subpopulací 2 a 3 týdny po separaci (tps) v porovnání s expresí původních klonů před separací. Jelikož se účinnost separace pohybovala v rozmezí 90 – 100 %, byly subpopulace pokládány za pozitivní, pokud bylo v subpopulaci přítomno alespoň 10 % GFP⁺ buněk. Z 18 negativních subpopulací byla u 12 pozorována přítomnost více než 10 % GFP⁺ buněk. U většiny z těchto nestabilně umlčených provirů však docházelo během následné kultivace k poklesu procenta exprimujících buněk. Ke zvyšování exprese směrem ke stavu v původním klonu docházelo u dvou subpopulací. Ani v jednom z případů však nebyl pozorován návrat negativní subpopulace na hladinu exprese původního klonu. Analýza MFI ukázala, že GFP⁺ buňky objevující se v negativních subpopulacích klonů vykazují MFI podobnou MFI GFP⁺ buněk původního klonu. Expresní profily původních klonálních populací pak ukazují, že klony s nestabilně umlčenými proviry vykazovaly po dobu 30 - 60 dpi spíše stabilní hladinu exprese. Přítomnost GFP⁺ buněk u negativních subpopulací však nebyla podmíněná vysokým podílem exprimujících buněk u původních klonů, neboť byla reaktivity pozorována i u klonů s 60 % exprimujících buněk.

Umlčené proviry ASLV jsou tedy v lidských buňkách schopné spontánně obnovit svou expresi. Schopnost vymanit se z umlčeného stavu však nemají všechny proviry ASLV. Spontánní reaktivity byla pozorována u klonů vykazujících spíše stabilní poměry GFP⁺ buněk v klonálních populacích pohybujících se na hladinách alespoň 60 % GFP⁺ buněk. Tyto výsledky opět poukazují na vliv místa integrace na expresi proviru.



Obrázek 14. Exprese negativně separovaných buněčných subpopulací ASLV. Na grafech jsou znázorněny počty (%) a intenzita exprese v MFI (logaritmičká škála) buněk exprimujících GFP u subpopulací vytvořených separací negativních buněk z původních buněčných klonů K562 nesoucích proviry ASLV. Exprese byla stanovována 2 a 3 tps (modré a zelené body). Časem 0 tps je znázorněna exprese původních klonů před separací. Jednotlivé subpopulace jsou seřazeny podle % buněk exprimujících GFP v čase 0 tps. Přerušované čáry znázorňují 10 % a 90 %. MFI – průměrná intenzita fluorescence, tps – týdny po separaci.

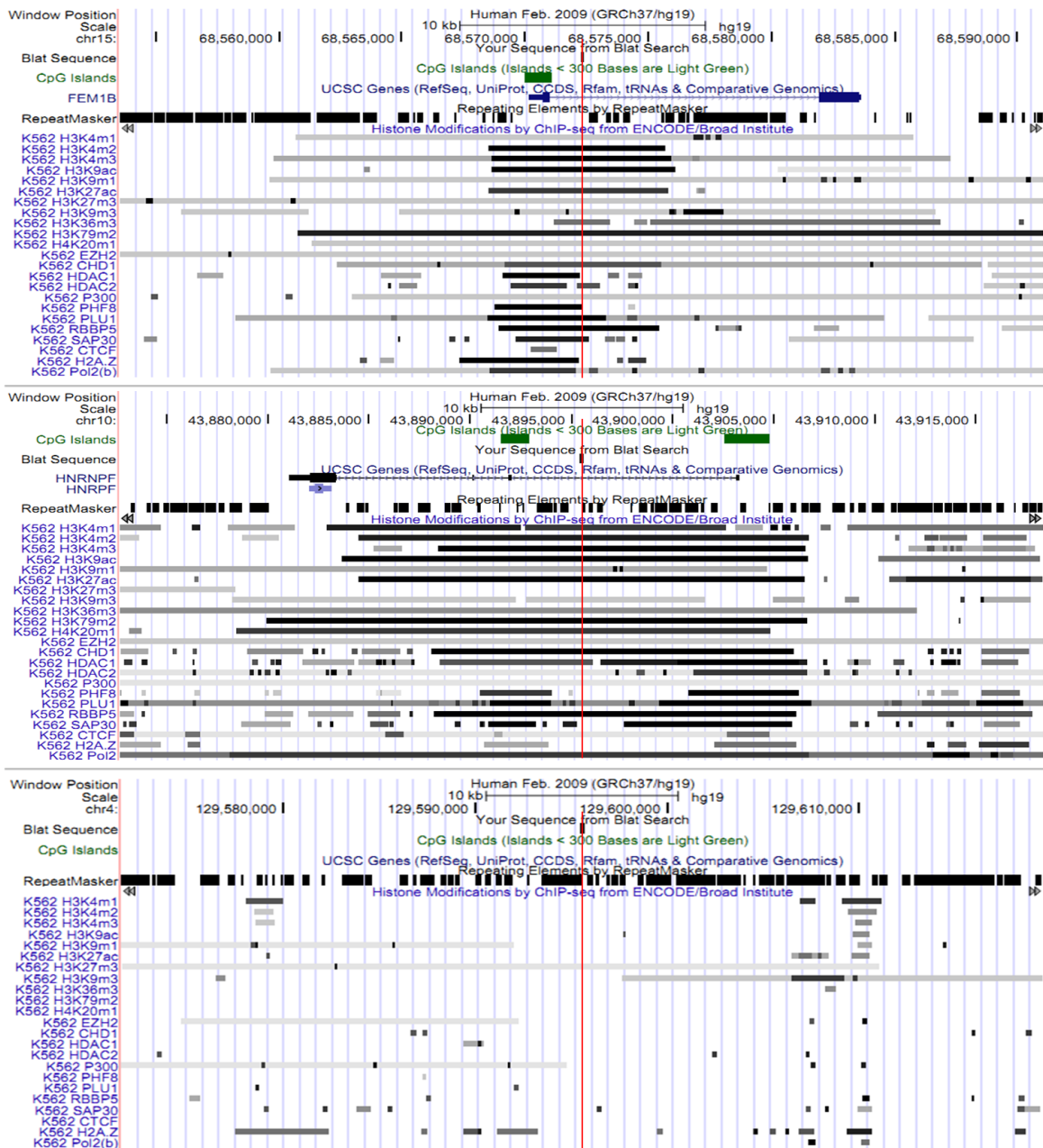
5.4. Integrační místa stabilně exprimovaných provirů ASLV

Místa integrace provirů byla izolována pomocí metody splinkerette PCR (Uren et al., 2009). Úskalí této metody může spočívat v její závislosti na štěpení genomové restriktivní endonukleázou, která může DNA štěpit příliš blízko či daleko od proviru a znemožňovat tak mapování či amplifikaci sekvence místa integrace. V této práci byly pro izolaci integračních míst provirů ASLV využity enzymy DpnII či MseI. Pomocí této metody pak bylo izolováno, mapováno a charakterizováno celkem 42 míst integrace stabilně exprimujících provirů integrovaných v genomu linie K562. Grafická ukázka výstupu mapování je demonstrována na **obr. 15**. Pro srovnání bylo také izolováno 16 integračních míst z buněk získaných ihned po separaci exprimujících buněk 3 dpi. Do analýzy bylo zahrnuto také 226 integračních míst ASLV z práce autorů Narezkina et al. (2004) jako kontrola neselektované integrační preference ASLV (dále jako preference). Byla provedena také separace negativních buněk 3 dpi, ovšem z těchto vzorků se nezdařilo získat žádné integrační místo. Získané sekvence se podobaly spíše defektním LTR provirů. Příčinou neúspěchu v získání sekvencí integračních míst z GFP⁻ buněk mohl být také nízký počet buněk nesoucích provirus. Zatímco 226 integračních míst preference by mělo simulovat neselektovanou integrační preferenci ASLV, 16 integračních míst klonů s proviry aktivními 3 dpi (dále jako aktivní) představuje populaci provirů, z níž byly stabilní proviry selektovány. Z předchozích výsledků exprese GFP lze také odhadovat, že asi 4 % aktivních provirů by měla být schopna stabilní exprese. Tato místa by měla vykazovat posun od integrace podle neselektované integrační preference k místům permisivním pro aktivní expresi ASLV.

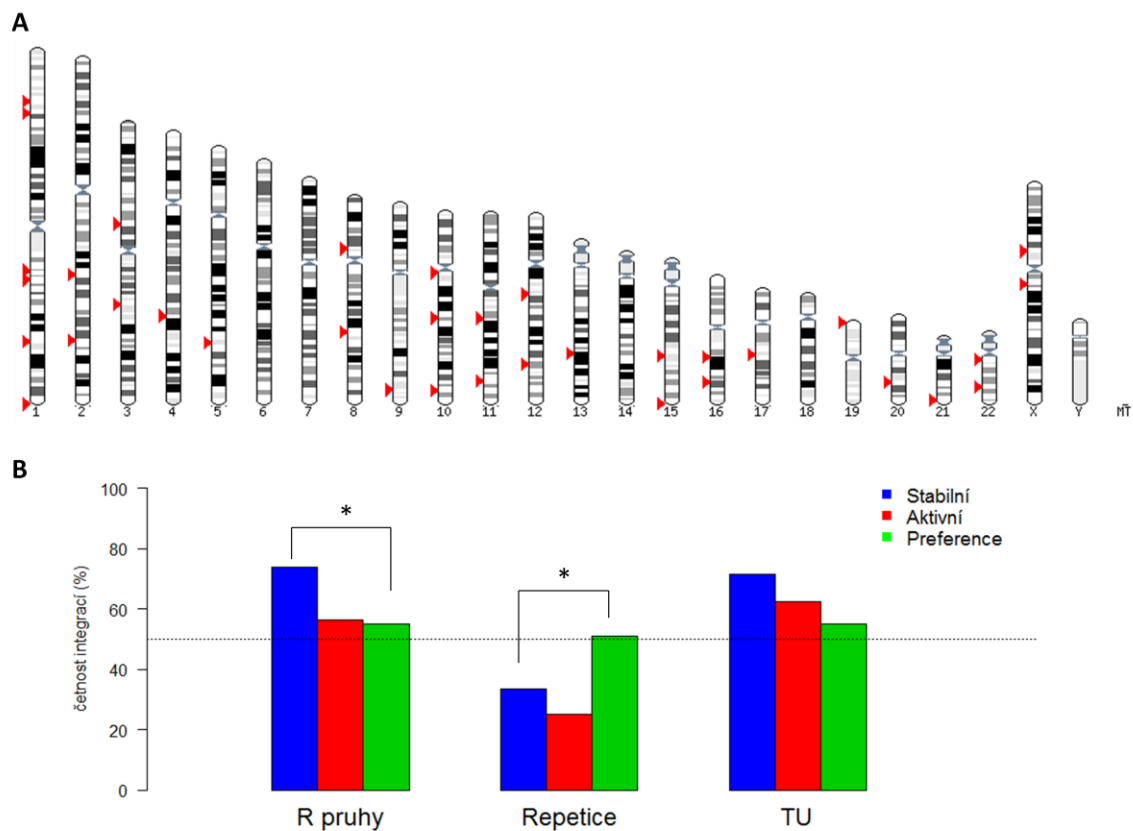
V první řadě byla místa integrace srovnávána s výskytem genomových charakteristik, jako je pozice v rámci chromozómů, TU, TSS či CpG ostrovů. Integrační místa jsou rozprostřena po celém genomu, jak je patrné ze zobrazení integračních míst na pozadí lidského karyotypu (**Obr. 16A**). Stabilně exprimované proviry byly oproti integrační preferenci ASLV častěji přítomny v oblastech chromozómových pruhů R, které by měly odpovídat oblastem aktivního euchromatinu. Naopak méně často byly stabilně aktivní proviry oproti místům integrace odpovídajícím integrační preferenci ASLV integrovány do repetitivních elementů (**Obr. 16B**). Ačkoliv byl oproti neselektované integrační preferenci pozorován malý nárůst integrací do oblastí TU, a to jak u stabilně

aktivních, tak u aktivních provirů, nebyl tento nárůst statisticky prokazatelný (Fisherův exaktní test, $p = 0,06$, rozdíl mezi stabilními proviry a neselektovanou preferencí ASLV).

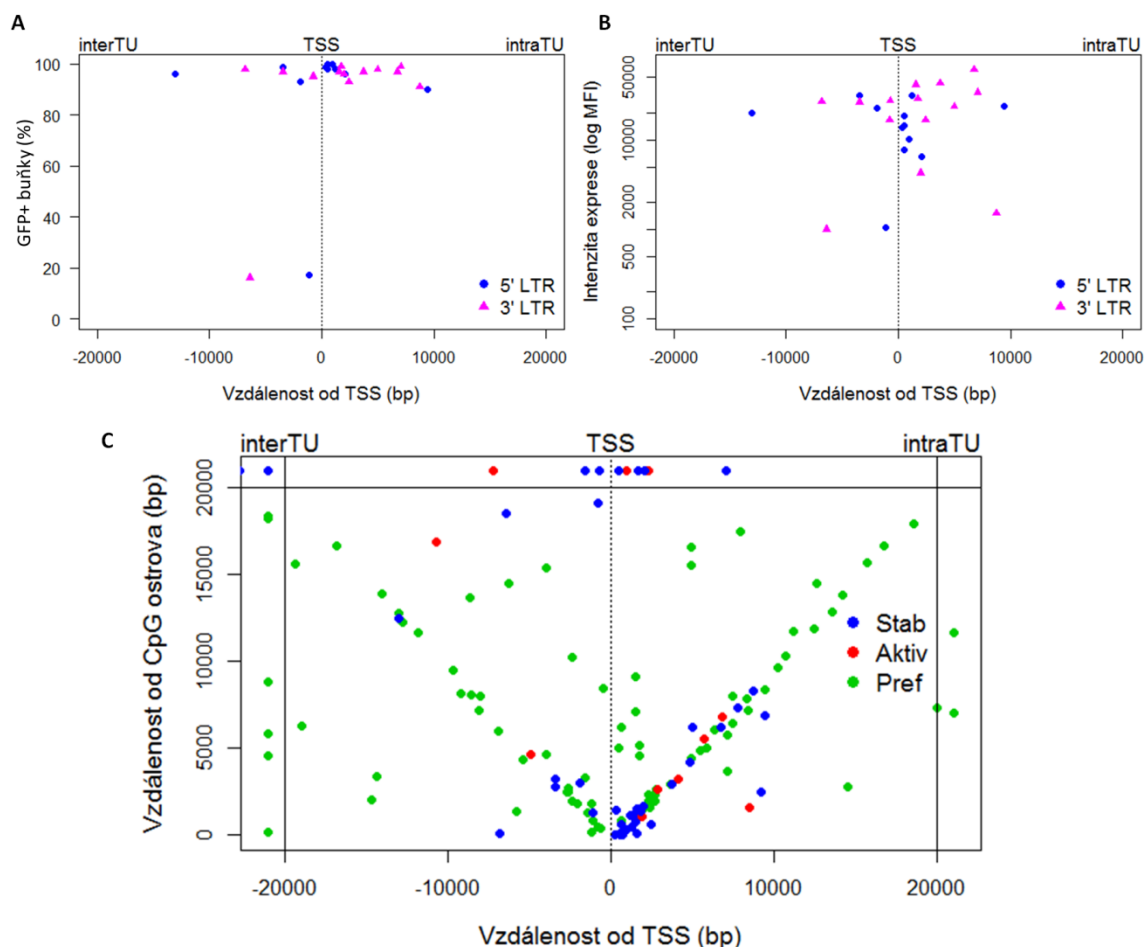
Integrace aktivních provirů byla poté charakterizována vzhledem k vzdálenostem od TSS a CpG ostrovů. Přítomnost TSS a CpG ostrovů byla analyzována v okolí 20 kbp od místa integrace. V této vzdálenosti od TSS se nacházelo 95 % stabilně exprimovaných provirů, přičemž vzdálenost těchto provirů od TSS se pohybovala v rozmezí 203 - 13000 bp. Procento GFP⁺ buněk v klonu, stejně jako intenzita exprese, byly nezávislé na vzdálenosti od TSS a orientaci stabilních provirů k TSS (**Obr. 17A, 17B**). Nebylo tedy pozorováno, že by stabilní proviry integrované ve větší vzdálenosti od TSS vykazovaly nižší MFI než proviry integrované blíže TSS. Závislost na orientaci proviru byla patrná pouze u těch stabilně aktivních provirů, které byly integrovány do TU (intraTU) do vzdálenosti 1 kbp od TSS (24 % provirů). Všechny stabilně exprimující proviry integrované v této oblasti byly orientovány shodně se směrem transkripce TU. Nejbližší provirus orientovaný proti směru transkripce se nacházel ve vzdálenosti asi 1500 kbp od TSS. V okolí 20 kb od CpG ostrovů se nacházelo 80 % stabilně exprimujících provirů. Jak lze vypozorovat z **obr. 17C**, jednalo se o CpG ostrovy, s nimiž byly asociovány blízké TSS (přítomnost bodů na diagonálách, tj. vzdálenost TSS je podobná vzdálenosti proviru od CpG ostrova). U 19 % provirů se pak integrační místo nacházelo ve vzdálenosti menší než 10 kbp od TSS, přičemž jejich vzdálenost od CpG ostrovů byla podstatně větší (**Obr. 17C**).



Obrázek 15. Ukázky výstupu mapování integračních míst ASLV. Na obrázku jsou znázorněny grafické výstupy 3 míst integrace (svíslá červená čára) ASLV a jejich okolí. Ve výstupech jsou znázorněny CpG ostrovy (zelené obdélníky), TU podle UCSC Genes, repetitivní elementy podle RepeatMasker a oblasti obohacené o přítomnost histonových modifikací či proteinových faktorů podle ENCODE/Broad Institute. Tmavost obdélníku u histonových modifikací a proteinových faktorů udává intenzitu signálu daného faktoru v dané oblasti. Grafická podoba místa integrace byla získána z aplikace Genome Browser na webových stránkách UCSC Genome Bioinformatics (<http://genome.ucsc.edu/>).

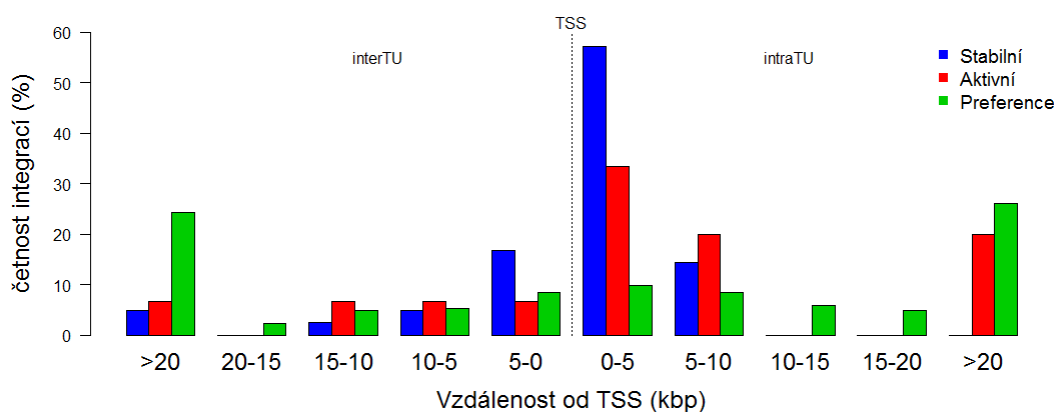


Obrázek 16. Integrační místa ASLV a genomové charakteristiky. A) Integrace stabilně exprimovaných provirů v rámci lidského karyotypu. Místa integrace jsou znázorněna červenou šipkou. **B)** Četnost integrací různých skupin provirů ASLV do oblastí chromozomálních R pruhů, repetitivních elementů a TU. Znázorněny jsou stabilně aktivní proviry (modrá), proviry aktivní 3 dpi (červená) a neselektované proviry integrační preference ASLV (zelená). Signifikance rozdílů ve frekvenci integrace mezi stabilně exprimovanými proviry a proviry znázorňujícími neselektovanou integrační preferenci ASLV byly testovány pomocí Fisherova exaktního testu. Signifikantní rozdíly jsou znázorněny hvězdičkou (*, $p < 0,05$). Přerušovaná čára znázorňuje frekvenci 50%. TU – transkripční jednotka.



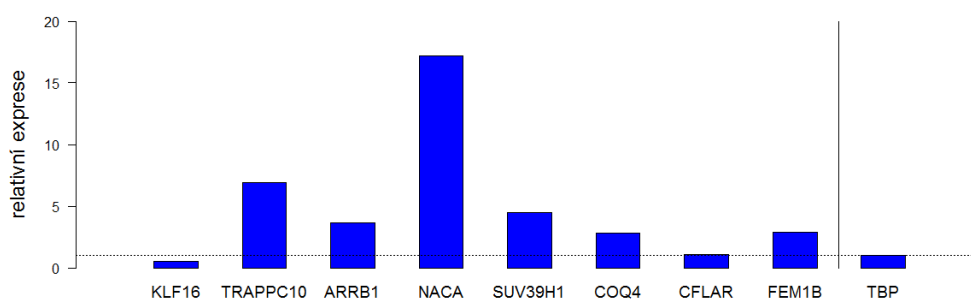
Obrázek 17. TU, TSS a CpG ostrovy v místě integrace stabilně exprimovaných provirů ASLV. A) Integrace provirů ASLV stabilně exprimujících klonů do oblastí TU (intraTU) či mimo TU (interTU). Znázorněna je vzdálenost od TSS a množství exprimujících buněk klonu 60 dpi. Negativní hodnoty vzdáleností od TSS znázorňují vzdálenosti provirů v interTU oblasti. Proviry orientované k TSS svým 5'LTR (místo provirového TSS) jsou znázorněny modrými body. Opačně orientované proviry (3'LTR k TSS) jsou znázorněny růžovými trojúhelníky. Znázorněny jsou proviry integrované do vzdálenosti 20 kbp od nejbližší oblasti TSS. Přeručovaná čára znázorňuje TSS **B)** Závislost intenzity exprese stabilně exprimovaných provirů (MFI, logaritmičká škála) na integraci do intraTU a interTU oblastí, vzdálenosti od TSS a orientaci proviru. **C)** Znázornění vzdálenosti integrovaných provirů od TSS a CpG ostrovů. Znázorněny jsou integrační místa stabilně exprimovaných provirů (modré body), provirů aktivních 3 dpi (červené body) a provirů znázorňujících neselektovanou integrační preferenci ASLV (zelené body). Jsou znázorněny proviry integrované do vzdálenosti 20 kbp od dané charakteristiky. Proviry integrované dále než 20 kbp jsou znázorněny za hranicemi 20 kbp. TSS – místo transkripčního startu, intraTU – oblast uvnitř transkripční jednotky, interTU – oblast mimo transkripční jednotku, MFI – průměrná intenzita fluorescence, bp – páry bazí.

Z výše uvedených výsledků (**Obr. 17**) je možné soudit, že distribuce integrovaných stabilně exprimujících provirů bude výrazně posunuta k oblastem TSS. Naprostá většina provirů je integrována do okolí 10 kbp od TSS (93 %). Proviry integrované do intraTU oblasti navíc tuto vzdálenost ani nepřekračují. Průměrná délka TU, do kterých jsou stabilně exprimované proviry integrovány, dosahuje téměř 40 kbp, přičemž většina přesahuje délku 10 kbp. Integrace do maximální vzdálenosti od TSS 10 kbp tedy není dána omezenou délkou TU. Distribuce zobrazená na **obr. 18** navíc ukazuje, že téměř 60 % stabilně exprimovaných provirů je integrováno do TU maximálně 5 kbp od TSS. Do oblasti 5 kbp od TSS je integrována také podstatná část provirů nacházejících se mimo TU (interTU oblasti). Proviry integrované v interTU oblasti lze, na rozdíl od těch v intraTU oblasti, nalézt i dále než 20 kbp od nejbližšího TSS (5 % provirů).



Obrázek 18. Frekvence integrací ASLV poblíž TSS. Graf znázorňuje četnosti integrací stabilně exprimovaných provirů (modrý sloupec), provirů aktivně 3 dpi (červený sloupec) a provirů odpovídajících neselektované preferenci integrací ASLV (zelený sloupec) v různých vzdálenostech od TSS. Vzdálenosti jsou rozděleny do oken, které jsou tvořeny 5kbp úseky až do vzdálenosti 20 kbp od nejbližšího TSS. Graf je dále rozdělen na úseky uvnitř TU (intraTU) a mimo TU (interTU). Vyhodnoceno bylo celkem 42 stabilních provirů, 16 aktivních a 226 provirů preference ASLV.

Integrace do TU nabízí otázku ohledně transkripční aktivity cílených TU. Bylo tedy vybráno 8 TU, do nichž byly integrovány stabilně exprimované proviry. Transkripční aktivita těchto TU byla měřena pomocí metody RT-qPCR v neinfikovaných buňkách linie K562. Relativní exprese TU byla získána porovnáním s endogenním housekeeping genem TBP. Analyzované TU většinou vykazovaly poměrně vysokou transkripční aktivitu (**Obr. 19**). Výskyt H3K4me3 v oblasti TSS, či do 1 kbp od 98 % TSS asociovaných se stabilními proviry ukazuje, že jsou tyto proviry integrovány do blízkostí, či přímo do transkripčně aktivních TU. Mezi TU, do kterých byly stabilně se exprimující proviry integrovány, se nacházely 4 (NACA, OAZ1, RPS4X, STAU1), které ve své práci She et al. (2009) identifikovali jako konstitutivně exprimované housekeeping geny. To odpovídá asi 13 % všech integrací do TU. Tato frekvence integrací do housekeeping genů však nepředstavuje žádný rozdíl oproti neselektované integrační preferenci ASLV.

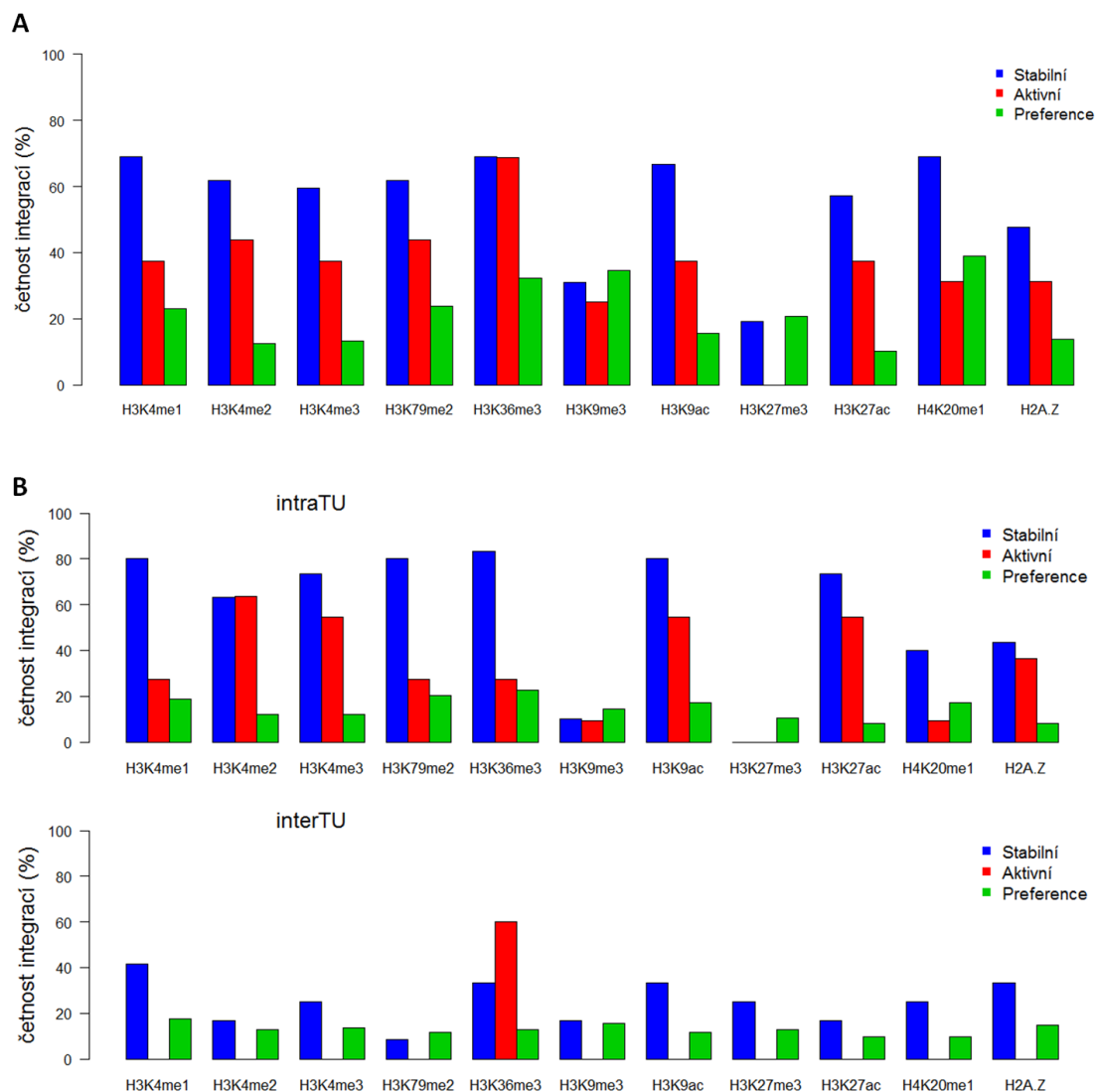


Obrázek 19. Relativní transkripční aktivita TU cílených stabilně aktivními proviry. Graf zobrazuje relativní transkripci 8 endogenních TU pomocí metody RT-qPCR. Do těchto TU byly integrovány stabilně exprimované proviry ASLV v lidských buňkách. Množství mRNA bylo stanovováno v neinfikovaných buňkách K562. Výsledky transkripce cílených TU byly vztaženy k množství mRNA referenčního housekeeping genu TBP. Hladina transkripce TBP (relativní exprese = 1) je znázorněna přerušovanou čarou. Názvy TU jsou uvedeny podle symbolů genů (Gene Symbols) dle <http://www.genenames.org>.

Kromě genomových charakteristik byly vyhodnocovány i histonové modifikace vyskytující se v místě integrace provirů. Nejprve byla vyhodnocována četnost integrací stabilně exprimovaných, aktivních a neselektovaných provirů do oblastí obohacených o výskyt jedenácti histonových modifikací (**Obr. 20**). Stabilně aktivní proviry se vyskytovaly častěji téměř ve všech z těchto modifikací (p hodnoty jsou uvedeny v **tabulce 4**). Zvýšená frekvence integrací stabilně aktivních provirů oproti integrační preferenci ASLV nebyla

pozorována u histonových modifikací H3K9me3 a H3K27me3, což jsou represivní histonové modifikace. Nejvyšší rozdíl ve frekvenci integrace pak byl pozorován u acetylovaných histonů (H3K9ac, H3K27ac) a u di- a trimetylace H3K4, což jsou naopak modifikace asociované s oblastmi aktivních TSS. Aktivní proviry, které nebyly selektovány na stabilitu exprese, vykazovaly ve většině případů četnosti integrací do oblastí modifikací podobající se frekvenci mezistupně mezi neselektovanou preferencí ASLV a stabilně aktivními proviry (**Obr. 20A**). Toto platilo hlavně pro modifikace H3K4me1 a H4K20me1. V oblastech bohatých na tyto modifikace byly stabilně exprimované proviry nacházeny častěji než proviry aktivní 3 dpi, ale i na expresi neselektované proviry ASLV, zatímco aktivní a neselektované proviry vykazovaly stejnou frekvenci integrací do oblastí H3K4me1 a H4K20me1. Nápadně shodná pak byla frekvence, se kterou se stabilně exprimované a aktivní proviry nacházely v oblastech obohacených o histonovou modifikaci H3K36me3. Tato histonová modifikace byla v místech integrace stabilně exprimovaných a aktivních provirů přítomná stejně často, přičemž se obě skupiny výrazně lišily od neselektovaných integrací.

Histonové modifikace byly dále vyhodnoceny v závislosti na tom, zda byly v předešlé analýze nalezeny v intraTU či interTU oblastech. V intraTU se nacházelo 30 stabilně exprimovaných, 11 aktivních a 124 neselektovaných provirů integrační preference ASLV. Naproti tomu v interTU oblastech bylo identifikováno 12 stabilně exprimovaných, 5 aktivních a 102 neselektovaných provirů odpovídajících integrační preferenci ASLV. Stabilně exprimované proviry v intraTU oblastech byly ve velké míře integrovány do oblastí všech histonových modifikací spojených s aktivním chromatinem (**Obr. 20B**). Výrazných rozdílů v porovnání s aktivními proviry dosahovaly frekvence integrací stabilně exprimujících provirů do oblastí modifikací H3K4me1, H3K79me2, H3K36me3 a H4K20me1. U provirů v interTU oblastech pak byla překvapivě výrazná frekvence výskytu v oblastech bohatých na H3K36me3. Tyto proviry byly nalezeny v oblastech s transkripční aktivitou, v nichž se nacházely predikované TU (Ensembl Genes). Některé interTU proviry jsou tak nejspíše integrovány do TU, které však nejsou zavedeny v záznamu UCSC Genes. Výskyt 3 z 5 aktivních provirů z interTU oblastech v místech bohatých na H3K36me3 pak poukazuje na výraznou frekvenci výskytu aktivních provirů v transkripčně aktivních TU.



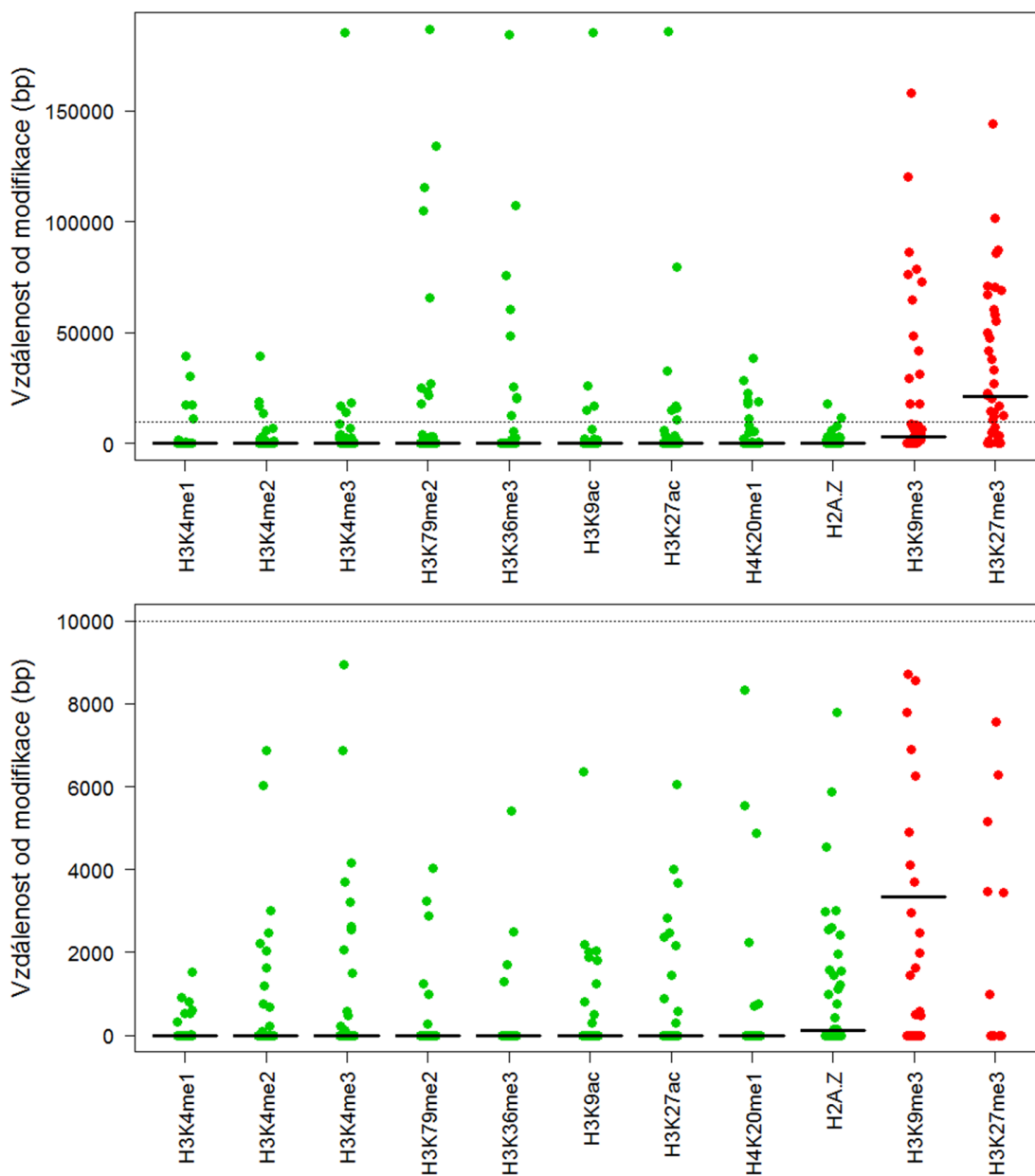
Obrázek 20. Frekvence integrací ASLV do oblastí histonových modifikací. Grafy znázorňují četnosti integrací stabilně exprimovaných provirů (modrý sloupec), provirů aktivních 3 dpi (červený sloupec) a provirů odpovídajících neselektované preferenci integrací ASLV (zelený sloupec) do oblastí obohacených o různé histonové modifikace. **A)** Frekvence integrací všech provirů v daných kategoriích. **B)** Proviry rozdělené podle toho, zda byly integrovány uvnitř známé TU (intraTU) či mimo TU (interTU). Bylo analyzováno celkem 42 integračních míst stabilních provirů, 16 integračních míst aktivních provirů a 226 míst integrace provirů odpovídajících preferenci ASLV. TU – transkripční jednotka.

Tabulka 4. Statistická signifikance rozdílů v integraci do oblastí histonových modifikací.

Histonové modifikace	Porovnávané skupiny provirů (p hodnoty)		
	Aktivní x Preference	Stabilní x Preference	Stabilní x Aktivní
H3K4me1	0,2246	$1,454 \times 10^{-8}$	0,0380
H3K4me2	0,0032	$4,642 \times 10^{-11}$	0,2467
H3K4me3	0,0187	$8,324 \times 10^{-10}$	0,1533
H3K79me2	0,1309	$4,17 \times 10^{-06}$	0,2467
H3K36me3	0,0052	$1,08 \times 10^{-05}$	1,0000
H3K9me3	0,5878	0,8582	0,7555
H3K9ac	0,0352	$5,067 \times 10^{-11}$	0,0723
H3K27me3	0,0469	1,0000	0,0916
H3K27ac	0,0060	$9,269 \times 10^{-11}$	0,0916
H4K20me1	0,6057	$3,53 \times 10^{-04}$	0,0159
H2A.Z	0,0699	$3,01 \times 10^{-06}$	0,3753

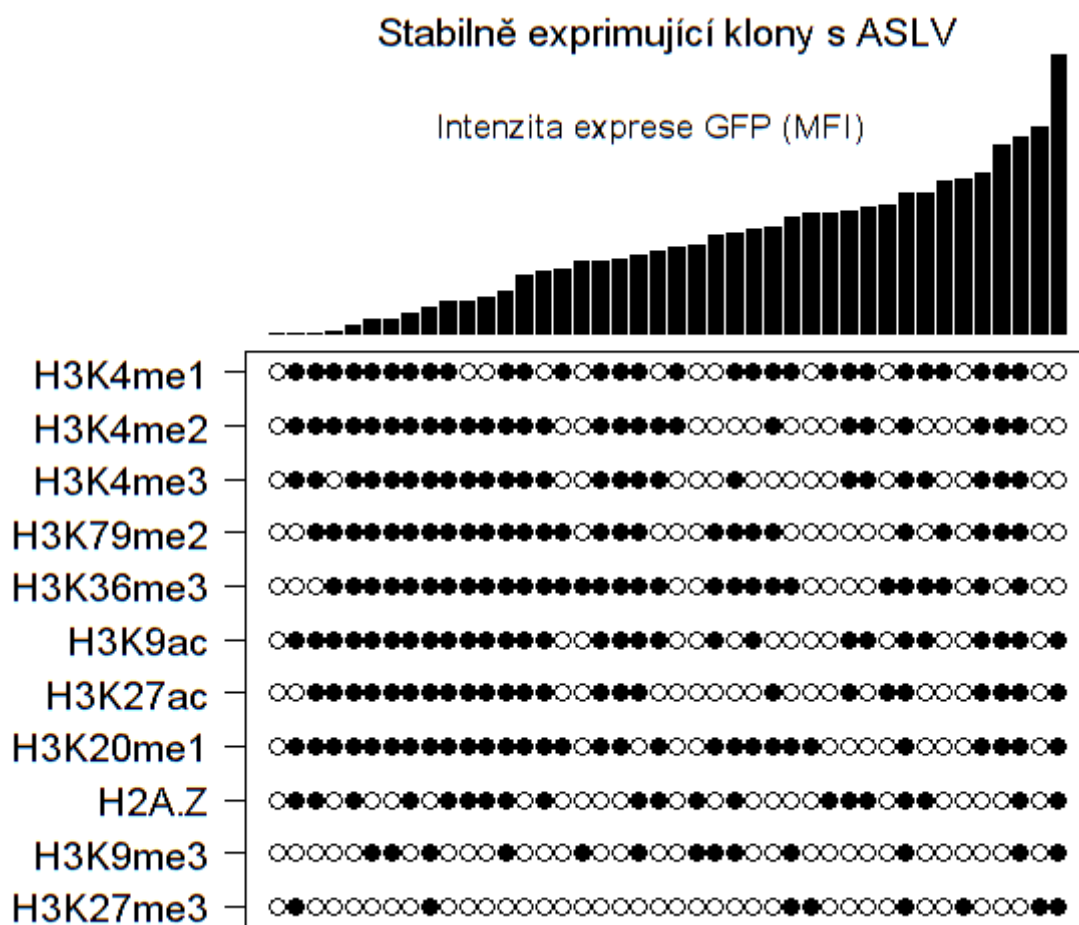
V tabulce jsou uvedeny p-hodnoty Fisherova exaktního testu frekvencí integrací stabilně aktivních provirů (Stabilní), provirů aktivních 3 dpi (Aktivní) a provirů odpovídajících preferenci ASLV (Preference). Rozdíl je signifikantní, je-li $p < 0,05$.

Vyhodnocována byla i vzdálenost stabilně exprimovaných provirů od oblastí bohatých na různé histonové modifikace (**Obr. 21**). Mediány vzdáleností míst integrace od oblastí s modifikacemi asociovanými s aktivním chromatinem byly rovny 0. Výjimkou byla histonová varianta H2A.Z, jejíž medián vzdáleností provirů byl 120 bp. To znamená, že většina stabilně exprimovaných provirů byla přímo asociovaná s oblastmi histonových modifikací aktivního chromatinu, především s oblastmi obohacenými o metylace H3K4, acetylaci histonů a H4K20me1. Výrazně pak byly posunuty vzdálenosti míst integrace od oblastí bohatých na H3K9me3 a H3K27me3. Distribuce vzdáleností provirů např. ukazuje, že 90 % stabilně aktivních provirů je integrováno do vzdálenosti 10 kbp od míst obohacených o H3K4me3. To je asi dvojnásobný nárůst integrace do této oblasti oproti neselektovaným provirům ASLV (43 %, Fisherův exaktní test, $p = 1,274 \times 10^{-9}$). Stejně navýšení stabilně aktivních provirů je pak pozorovatelné i pro oblasti výskytu H3K9ac.



Obrázek 21. Vzdálenost stabilně exprimovaných provirů od oblastí histonových modifikací. Na grafu jsou zobrazeny vzdálenosti v jednotkách bp stabilně exprimovaných provirů ASLV od oblastí zvýšeného výskytu různých histonových modifikací. Zobrazeny jsou proviry ve všech vzdálenostech (horní graf) a detail vzdáleností do 10 kbp od histonových modifikací (spodní graf). Vodorovné čáry v každém sloupci symbolizují mediány vzdáleností. Zeleně jsou pak zobrazeny aktivační modifikace, červeně pak modifikace, u nichž se předpokládá spíše represivní role. Přerušovaná čára vyznačuje vzdálenost 10 kbp – páry bazí.

Na **obr. 22** jsou znázorněny všechny stabilně exprimující klony a histonové modifikace, jejichž výskyt je zvýšen právě v místě integrace identifikovaného proviru. Pokud jsou klony seřazeny podle intenzity exprese GFP, nelze vypořádat žádnou souvislost mezi přítomností určité epigenetické modifikace přímo v místě integrace proviru a intenzitou exprese daného klonu.



Obrázek 22. Integrace do oblastí histonových modifikací jednotlivých stabilních klonů ASLV. Jednotlivé proviry stabilně exprimujících klonů jsou seřazeny ve sloupcích ve směru vodorovné osy podle intenzity exprese (MFI) jednotlivých klonů. MFI klonů se pohybovala od 1000 (úplně vlevo) po 61162 (úplně vpravo). V řádcích jsou pomocí plných či prázdných kruhů vyobrazeny jednotlivé histonové modifikace. Pokud byl provirus daného klonu nalezen integrován v oblasti obohacené o danou histonovou modifikaci, obsahuje pak jeho sloupec v daném řádku plný kruh.

Proviry identifikované ve stabilně exprimujících kloněch vykazovaly integrace preferenčně do oblastí aktivního chromatinu, do TU a blízkosti TSS. Na častou integraci do TU a do blízkosti TSS ukazují i histonové modifikace přítomné v místě integrace. Stabilně aktivní proviry vykazovaly oproti neselektovaným provirům ASLV preferenci pro všechny histonové modifikace asociované s aktivním chromatinem, TSS a transkribovanými TU. Proviry selektované na expresi brzy po infekci vykazovaly frekvence integrací odlišné od těch neselektovaných a byly tak mezistupněm mezi stabilně exprimovanými a neselektovanými proviry. Stejnou frekvenci v integracích jako stabilně exprimované proviry pak vykazovaly aktivní proviry hlavně v oblastech bohatých na H3K36me3. Ve vztahu k TU pak byly stabilně aktivní proviry integrovány v transkripčně vysoce aktivních TU. Nebyla pozorována žádná výrazná preference pro orientaci v TU. Výjimku tvořily proviry integrované v úplné blízkosti TSS v intraTU oblasti (do 1 kbp), kdy směr transkripce proviru odpovídal směru transkripce TU. Taktéž nebyla zaznamenána žádná závislost intenzity exprese klonu na výskytu některé histonové modifikace v místě integrace proviru.

5.5. Integrační místa nestabilně umlčovaných klonů ASLV

Bylo identifikováno 5 integračních míst provirů, jejichž klony vykazovaly existenci exprimujících buněk po separaci buněk negativních na expresi GFP. Tato místa vykazovala značnou heterogenitu. 3 proviry byly nalezeny v oblasti TU a vykazovaly různou vzdálenost od TSS s minimem okolo 0,2 kbp a maximem okolo 23 kbp (**tabulka 5**). Přítomnost H3K4me3 v rozmezí 0 – 9 kbp od místa integrace však naznačuje, že aktivní TSS by mohly být vzdáleny do 10 kbp. U dvou provirů je známa transkripční aktivita TU, do kterých byly proviry integrovány. Jejich transkripční aktivita dosahovala dvou a čtyřnásobku aktivity referenčního genu TBP, přičemž se jejich vzdálenost od TSS lišila asi o 2 kbp, kdy dále od TSS byl integrován provirus přítomen v aktivnější TU. Integrační místa jsou známa také u 3 provirů izolovaných z klonů, jejichž negativní subpopulace neobsahovaly 2 a 3 tps více než 10% GFP⁺ buněk. Tyto klony tedy obsahovaly proviry, jejichž umlčení bylo stabilní. Proviry z těchto klonů byly integrovány do blízkosti TSS, na což poukazuje i asociace se všemi metylačními stavy H3K4, H3K79me2 a H2A.Z (**tabulka 6**). Provirus klonu, který po separaci vykazoval nejvyšší podíl exprimujících buněk (64 % 2 tps) byl naopak integrován do oblasti bohaté na H3K4me1 a

H2A.Z. Tato specifika poukazují na pravděpodobnou integraci do oblasti enhanceru. H3K4me1 je až na klon vykazující expresi 11% exprimujících buněk 2 tps, specifickou modifikací asociovanou s reaktivovanými proviry.

Tabulka 5. Exprese a genomové charakteristiky integračních míst provirů nestabilně umlčovaných klonů.

klon	Exprese klonu			Genomové charakteristiky			
	60 dpi (%)	MFI 60dpi	2 tps (%)	TSS (kbp)	TU*	Orientace k TU**	CpG (kbp)
K2-1D5	99	27073	64	>20	-	NA	>20
K3-1H3	100	28330	45	0,2	+	+	8,3
K3-2B3	91	1510	41	8,7	+	-	6,2
K2-1F2	96	20011	38	13,0	-	-	12,4
K3-2A10	97	61162	11	6,7	+	-	0,023
K3-1D4	17	1051	9	1,1	-	-	1,3
K2-1A2	96	6573	6	2,1	+	+	0
K3-2D4	34	1010	3	0,6	+	+	>20

Klony v tabulce jsou seřazeny podle % exprimujících buněk 2 tps negativních subkonů. Klony vykazující méně než 10% pozitivních buněk byly považovány za negativní (poslední 3 klony). * proviry identifikované v TU jsou označeny jako +. ** relativní orientace k TU, kdy + znamená, že směr transkripce proviru je shodný s TU, do které je integrován či vzdálena do 20 kb, NA znamená, že nebyla v okolí 20 kbp nalezena žádná TU. dpi – dny po infekci, tps – týdny po separaci, MFI – průměrná intenzita exprese, TSS – místo transkripčního startu, TU – transkripční jednotka, CpG – CpG ostrov.

Tabulka 6. Vzdálenost od histonových modifikací provirů ASLV nestabilně umlčovaných klonů.

klon	Vzdálenost od oblasti bohaté na histonovou modifikaci (kbp)										
	H3K4 me1	H3K4 me2	H3K4 me3	H3K79 me2	H3K36 me3	H3K9 ac	H3K27 ac	H3K20 me2	H2 A.Z	H3K9 me3	H3K27 me3
K2-1D5	0	6,9	8,9	22	20,6	6,3	6,1	22,5	0	7,0	41,6
K3-1H3	0	0	0	2,9	1,3	0	0,3	0,7	0	1,6	22,7
K3-2B3	0	0	2,6	0	0	0	0	0	7,8	120	47,5
K2-1F2	0,6	2,2	4,1	3,2	5,4	2,2	2,2	5,5	0	0	55,0
K3-2A10	30,5	0,1	3,7	0,3	48,6	0	0	0	0	0	0
K3-1D4	0	0	0	0	24,9	0	0	0	0	6,2	5,1
K2-1A2	0	0	0	0	0	0	0	0	0,1	0	0
K3-2D4	0	0	0	105	107	0	32,8	0	0	64,7	0

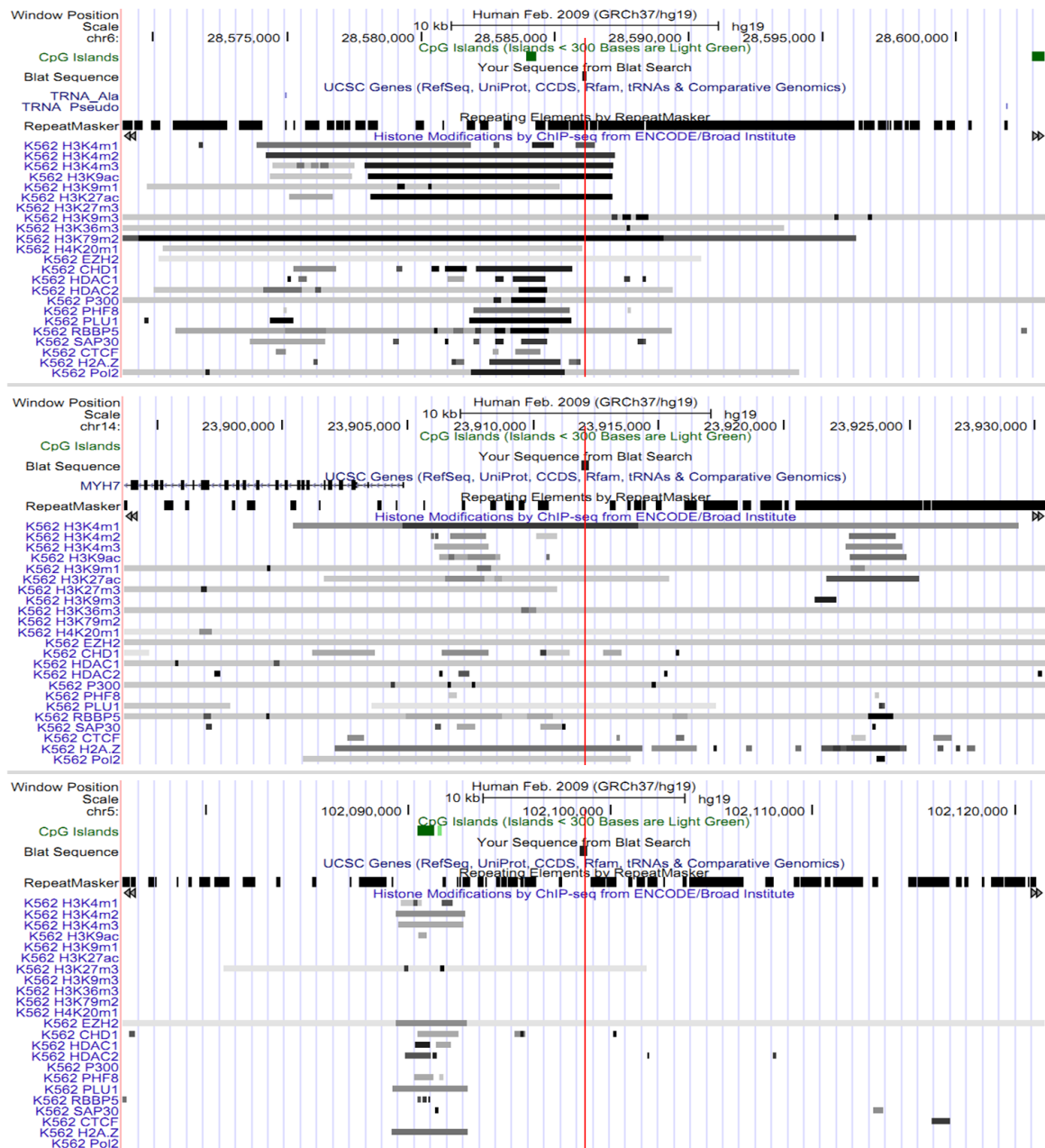
Klony v tabulce jsou seřazeny podle procenta exprimujících buněk 2 tps negativních subpopulací. Klony vykazující méně než 10% pozitivních buněk byly považovány za negativní (poslední 3 klony). Pokud byl provirus integrován v oblasti bohaté na histonovou modifikaci, je jeho vzdálenost od této modifikace 0.

5.6. Integrační místa stabilně exprimovaných provirů HIV-1

V rámci této práce byla izolována 3 integrační místa provirů lentivirového vektoru LTR-Tat-GFP (**Obr. 23**). Tyto klony vykazovaly v době 60 dpi expresní stabilitu alespoň 90 % (100 %, 91 % a 90 %). Žádný z provirů identifikovaných ve stabilně exprimujících kloněch nebyl nalezen v intraTU oblasti. Nejbližší TSS bylo identifikováno u proviru s expresní stabilitou 91 % exprimujících buněk 60 dpi, kde byla jeho vzdálenost od TSS asi 7 kbp. Od ostatních provirů však byly TSS vzdáleny 30 a 20 kbp. Ve vzdálenosti 1,7 a 6,8 kbp však byly u těchto klonů nalezeny CpG ostrovy. Podle genomových charakteristik tak byly tyto proviry integrovány mimo TU, avšak do blízkosti TSS či CpG ostrovů neasociovaných s žádnými známými TSS.

Lepší přehled o místě integrace provirů HIV-1 ukázala analýza histonových modifikací. Nejstabilněji exprimovaný provirus (100 % GFP⁺ buněk 60 dpi) byl identifikován uvnitř oblasti bohaté na všechny histonové modifikace kromě H3K27me3. Takový profil bývá charakteristický pro blízkosti TSS intraTU oblastí. Ostatní proviry pak vykazovaly různé profily pro přítomnost histonových modifikací. Jeden byl integrován do blízkosti necelých 6 kbp od H3K4me2/3 bohaté oblasti a do H3K27me3 bohaté oblasti. Druhý provirus byl integrován v oblasti bohaté na H3K4me1, H3K36me3, H3K27ac, H3K20me1 a H2A.Z. Tyto dva proviry se srovnatelnou stabilitou exprese (90 % a 91 %) tak vykazovaly znaky integrace do vzdálenosti 5,8 a 4 kbp od oblasti bohaté na H3K4me3. Provirus integrováný do oblasti H3K36me3 a H3K20me1 vykazoval nejvyšší intenzitu exprese ze všech tří provirů.

Ačkoliv nebyla v místě integrace analyzovaných stabilně exprimujících provirů HIV-1 nalezena žádná TU, vykazovala integrační místa charakteristiky pro místa s aktivní transkripcí v blízkosti aktivních TSS. Na tento fakt poukazují histonové modifikace (hlavně metylované varianty H3K4, H3K36me3 a H4K20me1) přítomné v oblasti integrace, ale také přítomnost CpG ostrovů, které však dosud nebyly asociovány s žádnými známými TU.



Obrázek 23. Grafická podoba integračních míst HIV-1. Na obrázku jsou znázorněny grafické výstupy 3 míst integrace (svíslá červená čára) stabilně exprimovaných klonů HIV-1 a jejich okolí. Počet GFP⁺ buněk klonů 60 dpi byl (seshora) 100 %, 91 %, 9 0%. Ve výstupech jsou znázorněny CpG ostrovy (zelené obdélníky), TU podle UCSC Genes, repetitivní elementy podle RepeatMasker a oblasti obohacené o přítomnost histonových modifikací či proteinových faktorů podle ENCODE/Broad Institute. Tmavost obdélníku u histonových modifikací a proteinových faktorů udává intenzitu signálu daného faktoru v dané oblasti. Grafická podoba místa integrace byla získána z aplikace Genome Browser na webových stránkách UCSC Genome Bioinformatics (<http://genome.ucsc.edu/>).

Tabulka 7. Exprese a genomové charakteristiky stabilně exprimovaných provirů HIV-1.

klon	Exprese 60 dpi		Genomové charakteristiky		
	%	MFI	TU	TSS (kbp)	CpG (kbp)
KH-1A12	100	20112	-	30,9	1,7
KH-1A8	90	19911	-	20	6,8
KH-1A9	91	30943	-	7,3	22,8

dpi – dny po infekci, MFI – průměrná intenzita fluorescence, TU – transkripční jednotka, TSS – místo transkripčního startu, CpG – CpG ostrov.

Tabulka 8. Integrace stabilně exprimovaných provirů HIV-1 do oblastí histonových modifikací.

klon	Vzdálenost od oblasti bohaté na histonovou modifikaci (kbp)										
	H3K4 me1	H3K4 me2	H3K4 me3	H3K79 me2	H3K36 me3	H3K9 ac	H3K27 ac	H4K20 me1	H2A. Z	H3K9 me3	H3K27 me3
KH-1A12	0	0	0	0	0	0	0	0,03	0,1	0	92,1
KH-1A8	6,2	5,6	5,8	91,7	154,4	7,6	91,5	92,0	5,6	219,3	0
KH-1A9	0	1,2	4	25,7	0	1,6	0	0	0	9,0	1,2

V tabulce jsou shrnuty vzdálenosti od oblastí bohatých na histonové modifikace. Pokud byl provirus identifikován v oblasti obohacené o danou histonovou modifikaci, byla jeho vzdálenost od modifikace 0. kbp – tisíc páru bazí.

6. DISKUZE

Retrovirová exprese je závislá na integraci retrovirové DNA do genomu a vytvoření proviru. Transkripční aktivita proviru pak silně závisí na podmínkách panujících v místě integrace proviru, které mohou jeho transkripci směřovat k umlčení či udržovat transkripci proviru. Jak ukazují výsledky této diplomové práce, stabilní exprese proviru je podmíněna integrací do oblastí aktivního chromatinu a zvláště pak do okolí TSS v intraTU oblastech.

Analýza buněčných klonů nesoucích jediný integrovaný provirus je dobrým nástrojem pro analýzu exprese provirů. Jak ukázala expresní analýza provirů ASLV a HIV-1 v lidských buňkách K562, stabilně exprimované proviry jsou minoritní složkou provirů integrovaných do hostitelského genomu, zatímco exprese většiny provirů je umlčována. Stabilně exprimované proviry HIV-1 byly asi 10 x častěji přítomné, než stabilně exprimované proviry ASLV. Systém stanovování expresních profilů provirů prezentovaný v této práci však neumožňuje úplný odhad frekvence stabilně exprimovaných provirů ze všech integrovaných provirů. Frekvence popsané v této práci totiž vycházejí z populace provirů aktivních 3 dpi. Existují spekulace, že většina provirů je umlčena brzy po integraci a proviry aktivní 3 dpi tak mohou představovat pouhý zlomek celkové populace integrovaných provirů. Vzácný výskyt stabilně exprimovaných provirů, spolu s téměř náhodnou integrací do genomu však poukazuje na vhodnost využití ASLV v hledání integračních míst podporujících stabilní expresi retrovirů. Exprese HIV-1 se od ASLV liší také dynamikou umlčování exprese a průměrnou intenzitou exprese stabilních klonů. Za tyto rozdíly je pravděpodobně zodpovědný retrovirový transaktivátor Tat. Ten je velmi účinným aktivátorem transkripce HIV-1 u různě integrovaných provirů (Jordan et al., 2001). Bazální transkripce velké části provirů HIV-1 však nepostačuje na expresi Tat proteinu a udržení stabilní exprese HIV-1. Stabilní exprese HIV-1 tak bude nejspíše také omezena na místa integrace zamezujících umlčování exprese HIV-1.

Klonální systém kultivace buněk s integrovanými proviry také dovoluje připisovat expresní profily klonů konkrétním místům integrace provirů. K identifikaci integračních míst provirů byla využita metoda splinkerette PCR. Tato metoda je závislá na štěpení genomové DNA restrikcími endonukleázami (Uren et al., 2009). Pokud je provirus integrován příliš blízko či daleko od místa štěpení, nemusí být integrační místo proviru

nalezeno. Ačkoliv byly pro tuto práci vybrány enzymy s častou frekvencí štěpení v lidském genomu, integrační místa provirů některých stabilně exprimujících klonů se identifikovat nepodařilo. V některých případech byly naopak identifikovány dva proviry v jednom buněčném klonu, což vedlo k vyřazení takového klonu z analýzy. Důvodem pro přítomnost více provirů může být nepřesnost separace buněk (umístění více buněk do jedné jamky) či infekce více virovými partikulami. I přes použitou nízkou MOI je pravděpodobnější druhá možnost. Výskyt takových klonů byl však spíše sporadický a expresní profily většiny klonů by měly odpovídat identifikovaným provirům. Teoreticky se však může stát, že i při výskytu více provirů může být identifikován pouze jeden, který neodpovídá za expresní profil daného klonu. Ačkoliv je pravděpodobnost takového jevu velmi nízká, je třeba jej brát v potaz při analýze izolovaných integračních míst.

V této práci byla identifikována místa integrace 42 stabilně exprimovaných provirů ASLV a 3 stabilně exprimovaných provirů HIV-1. Stabilně exprimované proviry ASLV vykazovaly častou integraci do míst aktivního chromatinu, na což poukazovala zvýšená integrace do chromozómových R pruhů a hlavně histonové modifikace přítomné v místě integrace. Na rozdíl od modelu vycházejícího z práce autorů Šenigl et al. (2012), však nebyl pozorován striktní výskyt provirů v blízkosti CpG ostrovů, TSS a oblastí bohatých na H3K4me3. Většina provirů sice byla integrována do blízkosti TSS, tyto TSS však nebyly vždy asociovány s CpG ostrovy. CpG ostrovy jsou přítomny v promotorových oblastech přibližně 70% (ale možná i více) známých TU (Saxonov et al., 2006). Pozorovaná asociace exprimovaných provirů s CpG ostrovy tak může představovat efekt doprovázející jejich asociaci s TSS.

Podstatná část stabilně exprimovaných provirů byla nalezena integrovaná do TU. Častou integraci do TU podporovala i zvýšená přítomnost stabilně exprimovaných provirů v oblastech se zvýšeným výskytem H3K36me3, který je charakteristický právě pro intraTU oblasti aktivních TU. Stejně zvýšená frekvence výskytu provirů v oblastech H3K36me3 však byla pozorována téměř u všech provirů, které vykazovaly expresní aktivitu 3 dpi. Aktivní proviry ASLV jsou na rozdíl od těch stabilně exprimovaných nalézány po celé délce TU v různých vzdálenostech od TSS, zatímco stabilně exprimované se nalézají v oblastech blízkých k TSS. Funkce H3K36me3 je pravděpodobně v zabraňování iniciaci transkripce z kryptických TSS pomocí přitahování HDAC do intraTU

oblastí (Venkatesh et al., 2012). Tato histonová modifikace by tak mohla bránit stabilní expresi integrovaných provirů. Integrace do aktivních TU tak může udržet po krátkou dobu proviry expresně aktivní, pro udržení stabilní exprese však bude pravděpodobně zapotřebí přítomnost dalších faktorů.

IntraTU oblasti aktivních TU by podle některých studií měly zabraňovat transkripci proviru pomocí transkripční interference (Han et al., 2008; Lenasi et al., 2008; Shan et al., 2011). Stabilně exprimované proviry se však nalézají často v transkripčně aktivních TU, jejichž transkripční aktivita může být i poměrně vysoká. Ačkoliv se v tom všechny práce zcela neshodují, na expresní aktivitu by mohla mít vliv i orientace proviru vůči směru transkripce TU. Nebyla však pozorována žádná korelace mezi procenty GFP⁺ buněk či intenzitou exprese GFP stabilně exprimujících klonů a orientací provirů v intraTU oblastech. Ve vzdálenosti do 1 kbp však byly pozorovány pouze proviry orientované shodně se směrem transkripce TU. Proviry orientované proti směru transkripce TU, pokud se nachází v blízkosti TSS, tak mohou být nejspíše umlčeny mechanismem transkripční interference, zatímco transkripce shodně orientovaných provirů je zvýšená. Ostatní stabilně exprimované proviry integrované až do vzdálenosti 10 kbp však vykazovaly různou orientaci, převážně však orientaci proti směru transkripce TU. Je tedy otázkou, zda je požadavek na orientaci proviru v intraTU oblasti závislý také na vzdálenosti od TSS, nebo jsou ve hře další faktory zabraňující transkripční interferenci. Proviry HIV-1 vykazující transkripční interferenci ve studiích autorů Han et al. (2008) a Lenasi et al. (2008) byly od TSS vzdáleny vždy více než 10 kbp. Možnou roli zde tak mohla hrát právě i vzdálenost provirů od TSS a absence faktorů, které by bránily umlčení proviru způsobenému ať už transkripční interferencí nebo např. přítomností H3K36me3.

Stabilně exprimované proviry integrované v intraTU oblastech byly kromě H3K36me3 asociovány s řadou histonových modifikací charakteristických pro aktivní chromatin a zvláště pro oblasti blízké k TSS. Jednou z těchto modifikací je i H3K4me2. Tato modifikace má podobný represivní účinek na intragenovou iniciaci transkripce, jako H3K36me3 (Kim et al., 2012). IntraTU oblasti, ve kterých byly nalézány stabilně exprimované proviry, byly navíc obohaceny o modifikace s aktivačním potenciálem. Výrazný byl výskyt stabilně exprimovaných provirů v oblastech bohatých na H3K4me1 a H4K20me1, kde se stabilní proviry vyskytovaly častěji než proviry aktivní. Zatímco

modifikace H3K4me1 je charakteristická hlavně pro místa vazby transkripčních faktorů, H4K20me1 a H3K79me2 značí hlavně intraTU oblasti blízké aktivním TSS. H4K20me1 se pak ve zvýšené míře objevuje u takových TU, jejichž exprese je řízena z CpG ostrovů (Karlič et al., 2010). V některých případech lze vyzorovat, že jsou stabilně exprimované proviry integrovány do oblastí, kde se nacházejí histonové modifikace, které by se měly za normálních podmínek vylučovat (H3K9ac a H3K9me3). Toto pozorování může být způsobeno buď přechodem mezi oběma modifikacemi v daném místě, nebo falešně pozitivním označením dané oblasti za obohacenou o danou histonovou modifikaci. Průměrné signály H3K9me3 oblastí s integrovanými stabilně exprimovanými proviry vykazovaly nízký průměrný signál H3K9me3, což poukazuje na možnost pouze nevýznamného nárůstu této modifikace v analyzovaných oblastech.

Integrační místa některých stabilně exprimovaných provirů se však nacházela v interTU oblastech. Tato místa se však ve většině případů nacházela poblíž aktivních TSS a pouze ve výjimečných případech byla nalezena ve větších vzdálenostech od TSS. Proviry integrované do míst vzdálených od TSS však ve svém okolí často obsahovaly oblasti vykazující charakteristiky oblastí vazby transkripčních faktorů (H3K4me1, H2A.Z). V těchto místech také byla pozorována vazba některých transkripčních faktorů (např. CTCF či MAF proteiny). Vliv vazby transkripčních faktorů však nebyl v této práci analyzován. Je tedy otázkou, zda místa vzdálená známým TSS představují lokusy s výjimečnými vlastnostmi podporujícími stabilní expresi provirů či jde o artefakt. V případě potvrzení výjimečných vlastností takových míst by byla dobrými cíly pro specifické cílení integrace retrovirových vektorů (tzv. „safe harbors“).

Nebyla pozorována souvislost mezi intenzitou exprese proviru a přítomností některé z histonových modifikací. Intenzita exprese proviru, tak může záviset na jiných faktorech. Možný vliv může mít transkripční aktivita větší oblasti genomu. Regulace exprese genů bývá kromě genově specifických faktorů ovlivněna intenzitou exprese okolních genů (80 – 90 TU), které tak tvoří jednu regulační doménu a intenzita exprese integrovaného proviru do těchto domén může být řízena právě intenzitou exprese v doméně (Gierman et al., 2007). Lidský genom je v interfázním jádře skutečně uspořádán do intrachromozomálních topologických domén (Bickmore and van Steensel, 2013; Dixon et al., 2012). Expese proviru tak může být řízena nejen charakteristikami

nejbližšího okolí proviru, ale i povahou interakcí s oblastmi uvnitř topologické domény a expresní aktivitou uvnitř této domény. Jak však ukazují výsledky autorů Šenigl et al. (2012) a této diplomové práce, na stabilitu exprese provirů budou mít pravděpodobně vliv podmínky v místě integrace a jeho nejbližším okolí.

Zvláštní skupinu tvořily proviry ASLV, které sice podléhaly umlčování, ale vykazovaly spontánní reaktivaci exprese. Umlčování těchto provirů ASLV je tedy nestabilní. Takové chování provirů bylo již pozorováno u MLV v myších embryonálních kmenových buňkách (Yao et al., 2004). Identifikace nestabilně umlčených provirů v diferencovaných lidských buňkách v této práci ukazuje, že nemusí jít o vlastnost specifickou pouze pro kmenové buňky. Buňky K562 sice nepředstavují terminálně diferencovaný buněčný typ, vývojově jsou však již vzdálené buňkám kmenovým. Otázkou zůstává mechanismus nestabilního umlčení a spontánní reaktive. Volnost umlčení by mohla poukazovat na umlčení pomocí modifikace histonů bez DNA metylace. K ověření této teorie by mohly vést experimenty s různými inhibitory enzymů (např. HDAC, DNMT atd.) ovlivňujících přítomnost epigenetických značek. Oscilace mezi aktivním a umlčeným stavem může záviset také např. na buněčném cyklu. U většiny klonů, které vykazovaly známky spontánní reaktive, bylo často pozorovatelné opětovné umlčování (tedy pokles množství GFP⁺ buněk v separované subpopulaci). Toto napovídá, že proviry by mohly s různou pravděpodobností dynamicky oscilovat mezi aktivním a umlčeným stavem. Fakt, že většina klonů vykazovala stabilní expresi, vede k myšlence, že nestabilní umlčování a opětovná spontánní reaktive, pokud je jejich pravděpodobnost vyvážená, mohou udržovat navenek stabilní expresi proviru. Integrační místa takových provirů vykazují různý charakter. Místa podporující dobrou reaktivaci byla poměrně vzdálená od TSS či CpG ostrovů a byla asociována s H3K4me1 a H2A.Z, což ukazuje na blízkost regulačních sekvencí. Tato místa byla také vzdálena oblastem obohacených o H3K27me3. Vzdálenost od této represivní modifikace a asociace s místy vazby transkripčních faktorů tak může být nedostatečná pro udržení stálé exprese, avšak může umožňovat únik z umlčeného stavu.

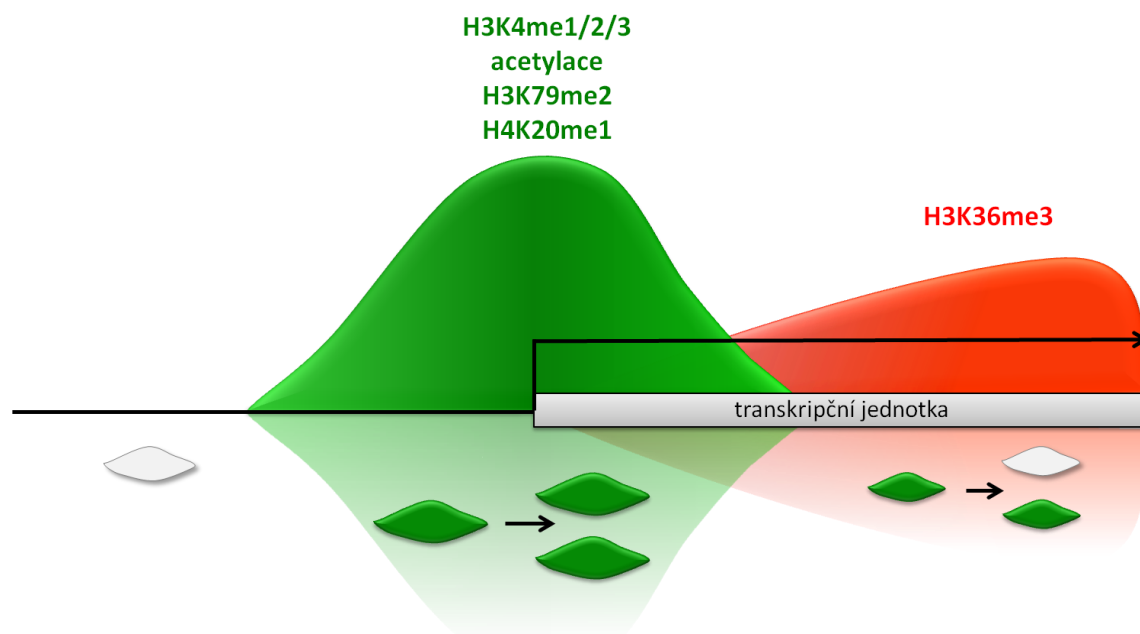
Integrační místa stabilně exprimovaných provirů HIV-1 vykazovala charakteristiky podobné těm, které byly popsány pro proviry ASLV. Proviry se nacházely poblíž TSS či CpG ostrovů, které sice nebyly asociovány s žádnými známými TU, ale mohou teoreticky

TSS obsahovat. Přítomné histonové modifikace tuto úvahu podporují. Nejstabilnější z izolovaných provirů byl integrován do oblasti obohacené o modifikace vyskytující se v nejbližším okolí promotorových oblastí uvnitř TU (např. H3K4me3, H3K79me2, H4K20me1 a H3K36me3). Ačkoliv jsou data v této práci omezená pouze na 3 známá integrační místa HIV-1, naznačují, že nastolení stabilní exprese HIV-1 se může řídit podobnými, nebo i stejnými pravidly závislosti na místech integrace jako ASLV. Výskyt aktivních provirů HIV-1 a ASLV v místech bez popsaných TU, avšak v oblastech bohatých na H3K36me3 naznačuje, že aktivní proviry by mohly být využívány pro identifikaci zatím nepotvrzených TU.

Ačkoliv jsou data prezentovaná v této diplomové práci výsledkem předběžné analýzy, ukazují na důležitost charakteristik integračního místa na expresi a hlavně stabilní expresi retrovirů (**Obr. 24**). ASLV a HIV-1 vykazují odlišnou schopnost nastolit aktivní expresi a liší se i v intenzitě exprese a dynamice jejího umlčování. Tento rozdíl je pravděpodobně výsledkem adaptace HIV-1 na lidského hostitele, zatímco lidské buňky jsou pro ASLV nepermissivní. Odlišná frekvence výskytu stabilně exprimovaných provirů však může být také důsledkem rozdílné integrační preference ASLV a HIV-1. HIV-1 se totiž preferenčně integruje do aktivních TU. Oblast aktivních TU je oblast, kde byly preferenčně pozorovány expresně aktivní proviry. Pokud je navíc provirus integrován v blízkosti TSS, bude pravděpodobně expresně stabilní. Z tohoto pohledu by mohla být pro časté nastavení stabilní exprese výhodná taková integrační preference, kdy by se retrovirus často integroval do oblastí TSS. Takovou integrační preferenci lze pozorovat u gamaretrovirů (MLV). Je však otázkou, zda a jak se liší frekvence stabilně exprimovaných provirů MLV, HIV-1 a ASLV a zda má posun integrační preference k TSS znatelný vliv na expresi proviru. K zodpovězení této otázky by bylo zapotřebí provést další studie podobné této u dalších retrovirů.

Z výsledků této práce plynou také důsledky pro retrovirové vektory. Ty se využívají právě kvůli schopnosti udržet vkládaný gen stabilně aktivní i v populaci dceřiných buněk. Pokud nejsou tyto vektory nějakým způsobem modifikovány, budou buňky vykazující stabilní expresi vloženého genu obsahovat provirus integrovaný pravděpodobně poblíž TSS některé z aktivních TU a teoreticky tak mohou ovlivňovat

přirozenou transkripční aktivitu TU. Řešením je úprava retrovirových vektorů vedoucí ke stabilizaci exprese a změně integračních preferencí či přímo cílení integrace retrovirů.



Obrázek 24. Model závislosti exprese provirů ASLV na místě integrace. Zelená oblast označuje místa, ve kterých byla nalezena většina provirů vykazujících vysokou expresní aktivitu, a jejich charakteristiky. Tyto oblasti byly často obohaceny o metylace H3K4, acetylace histonů, H3K79me2 a H4K20me1. Proviry, které vykazovaly aktivní expresi 3 dpi byly preferenčně integrovány v oblastech obohacených o H3K36me3. Jen malé procento z těchto provirů je schopno udržet stabilní expresi. Proviry, jejichž exprese je rychle umlčována, jsou pak nejspíše integrovány mimo tyto oblasti. Integrační místa u HIV-1 ukazují, že by tento model mohl být platný pro retroviry obecně.

7. SOUHRN

V této diplomové práci byly získány expresní profily celkem 245 buněčných klonů s integrovanými proviry ASLV, z čehož bylo asi 60 klonů stabilních. Navíc bylo identifikováno 130 stabilně exprimujících klonů s proviry HIV-1. Expresní analýzy buněčných klonů pak ukazují na 10 x častější výskyt stabilně exprimovaných provirů HIV-1 oproti ASLV (32% a 3% z expresně aktivních klonů 3 dpi). Ke stabilně exprimujícím klonům byly identifikovány také klony vykazující spontánní reaktivaci (nestabilní umlčení) exprese u buněk nesoucích expresně umlčené proviry.

Z buněčných klonů se stabilním expresním profilem bylo izolováno a charakterizováno celkem 42 integračních míst provirů ASLV a 3 integrační místa provirů HIV-1. Integrační místa ASLV se vyskytovala převážně v TU a poblíž TSS. Tyto charakteristiky byly podpořeny analýzou výskytu histonových modifikací v místech integrace. Častá přítomnost histonových modifikací spjatých s aktivní transkripcí ukazovala na přítomnost provirů v aktivním chromatinu (H3K4me3, H3K9ac, H3K36me3). Jako histonové modifikace charakteristické pro stabilně aktivní proviry byly identifikovány hlavně modifikace H4K20me1 a H3K4me1. Tyto dvě modifikace potvrzují preferenční výskyt stabilně exprimovaných provirů ASLV ve vnitřních oblastech transkripčně aktivních TU, v blízkosti TSS a poblíž regulačních oblastí obsahujících vazebná místa pro transkripční faktory. Integrační místa stabilně exprimovaných provirů HIV-1 vykazovala stejné charakteristiky jako ASLV.

8. SEZNAM POUŽITÉ LITERATURY

- Albanese, A., Arosio, D., Terreni, M., and Cereseto, A. (2008). HIV-1 pre-integration complexes selectively target decondensed chromatin in the nuclear periphery. *PLoS ONE* 3.
- Andrake, M. D., Sauter, M. M., Boland, K., Goldstein, A. D., Hussein, M., and Skalka, A. M. (2008). Nuclear import of Avian Sarcoma Virus integrase is facilitated by host cell factors. *Retrovirology* 5.
- Barr, S. D., Leipzig, J., Shinn, P., Ecker, J. R., and Bushman, F. D. (2005). Integration targeting by avian sarcoma-leukosis virus and human immunodeficiency virus in the chicken genome. *Journal of Virology* 79, 12035-12044.
- Barski, A., Cuddapah, S., Cui, K., Roh, T. Y., Schones, D. E., Wang, Z., Wei, G., Chepelev, I., and Zhao, K. (2007). High-resolution profiling of histone methylations in the human genome. *Cell* 129, 823-837.
- Bednarik, D. P., Cook, J. A., and Pitha, P. M. (1990). Inactivation of the HIV LTR by DNA CpG methylation: Evidence for a role in latency. *EMBO Journal* 9, 1157-1164.
- Bickmore, Wendy A., and van Steensel, B. (2013). Genome Architecture: Domain Organization of Interphase Chromosomes. *Cell* 152, 1270-1284.
- Bird, A. P. (1986). CpG-rich islands and the function of DNA methylation. *Nature* 321, 209-213.
- Blažková, J., Trejbalová, K., Gondois-Rey, F., Halfon, P., Philibert, P., Guiguen, A., Verdin, E., Olive, D., Van Lint, C., Hejnar, J., and Hirsch, I. (2009). CpG methylation controls reactivation of HIV from latency. *PLoS Pathogens* 5.
- Boese, A., Sommer, P., Holzer, D., Maier, R., and Nehrbass, U. (2009). Integrase interactor 1 (Ini1/hSNF5) is a repressor of basal human immunodeficiency virus type 1 promoter activity. *Journal of General Virology* 90, 2503-2512.
- Boztug, K., Schmidt, M., Schwarzer, A., Banerjee, P. P., Diez, I. A., Dewey, R. A., Bohm, M., Nowrouzi, A., Ball, C. R., Glimm, H., *et al.* (2010). Stem-cell gene therapy for the Wiskott-Aldrich syndrome. *N Engl J Med* 363, 1918-1927.
- Brendel, C., Muller-Kuller, U., Schultze-Strasser, S., Stein, S., Chen-Wichmann, L., Krattenmacher, A., Kunkel, H., Dillmann, A., Antoniou, M. N., and Grez, M. (2012). Physiological regulation of transgene expression by a lentiviral vector containing the A2UCOE linked to a myeloid promoter. *Gene Ther* 19, 1018-1029.
- Cavazzana-Calvo, M., Payen, E., Negre, O., Wang, G., Hehir, K., Fusil, F., Down, J., Denaro, M., Brady, T., Westerman, K., *et al.* (2010). Transfusion independence and HMGA2 activation after gene therapy of human β -thalassaemia. *Nature* 467, 318-322.
- Ciuffi, A., Diamond, T. L., Hwang, Y., Marshall, H. M., and Bushman, F. D. (2006). Modulating target site selection during human immunodeficiency virus DNA integration in vitro with an engineered tethering factor. *Human Gene Therapy* 17, 960-967.
- Coull, J. J., Romerio, F., Sun, J.-M., Volker, J. L., Galvin, K. M., Davie, J. R., Shi, Y., Hansen, U., and Margolis, D. M. (2000). The Human Factors YY1 and LSF Repress the Human Immunodeficiency Virus Type 1 Long Terminal Repeat via Recruitment of Histone Deacetylase 1. *Journal of Virology* 74, 6790-6799.
- Demonté, D., Quivy, V., Colette, Y., and Van Lint, C. (2004). Administration of HDAC inhibitors to reactivate HIV-1 expression in latent cellular reservoirs: Implications for the development of therapeutic strategies. *Biochemical Pharmacology* 68, 1231-1238.

Derse, D., Crise, B., Li, Y., Princler, G., Lum, N., Stewart, C., McGrath, C. F., Hughes, S. H., Munroe, D. J., and Wu, X. (2007). Human T-cell leukemia virus type 1 integration target sites in the human genome: Comparison with those of other retroviruses. *Journal of Virology* *81*, 6731-6741.

Di Primio, C., Quercioli, V., Allouch, A., Gijbsers, R., Christ, F., Debyser, Z., Arosio, D., and Cereseto, A. (2013). Single-Cell Imaging of HIV-1 Provirus (SCIP). *Proceedings of the National Academy of Sciences*.

Dixon, J. R., Selvaraj, S., Yue, F., Kim, A., Li, Y., Shen, Y., Hu, M., Liu, J. S., and Ren, B. (2012). Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature advance online publication*.

Duverger, A., Wolschendorf, F., Zhang, M., Wagner, F., Hatcher, B., Jones, J., Cron, R. Q., van der Sluis, R. M., Jeeninga, R. E., Berkhout, B., and Kutsch, O. (2013). An AP-1 Binding Site in the Enhancer/Core Element of the HIV-1 Promoter Controls the Ability of HIV-1 To Establish Latent Infection. *Journal of Virology* *87*, 2264-2277.

Eidahl, J. O., Crowe, B. L., North, J. A., McKee, C. J., Shkriabai, N., Feng, L., Plumb, M., Graham, R. L., Gorelick, R. J., Hess, S., *et al.* (2013). Structural basis for high-affinity binding of LEDGF PWWP to mononucleosomes. *Nucleic Acids Research*.

Elleder, D., Pavlicek, A., Paces, J., and Hejnar, J. (2002). Preferential integration of human immunodeficiency virus type 1 into genes, cytogenetic R bands and GC-rich DNA regions: insight from the human genome sequence. *Febs Letters* *517*, 285-286.

Emery, D. W., Yannaki, E., Tubb, J., and Stamatoyannopoulos, G. (2000). A chromatin insulator protects retrovirus vectors from chromosomal position effects. *Proceedings of the National Academy of Sciences of the United States of America* *97*, 9150-9155.

Faschinger, A., Rouault, F., Sollner, J., Lukas, A., Salmons, B., Gunzburg, W. H., and Indik, S. (2008). Mouse mammary tumor virus integration site selection in human and mouse genomes. *J Virol* *82*, 1360-1367.

Felice, B., Cattoglio, C., Cittaro, D., Testa, A., Miccios, A., Ferrari, G., Luzi, L., Recchia, A., and Mavilio, F. (2009). Transcription factor binding sites are genetic determinants of retroviral integration in the human genome. *PLoS ONE* *4*.

Gierman, H. J., Indemans, M. H. G., Koster, J., Goetze, S., Seppen, J., Geerts, D., Van Driel, R., and Versteeg, R. (2007). Domain-wide regulation of gene expression in the human genome. *Genome Research* *17*, 1286-1295.

Gijbsers, R., Ronen, K., Vets, S., Malani, N., De Rijck, J., McNeely, M., Bushman, F. D., and Debyser, Z. (2010). LEDGF hybrids efficiently retarget lentiviral integration into heterochromatin. *Molecular Therapy* *18*, 552-560.

Greger, J. G., Katz, R. A., Ishov, A. M., Maul, G. G., and Skalka, A. M. (2005). The Cellular Protein Daxx Interacts with Avian Sarcoma Virus Integrase and Viral DNA To Repress Viral Transcription. *Journal of Virology* *79*, 4610-4618.

Groth, A. C., and Emery, D. W. (2010). A functional screen for regulatory elements that improve retroviral vector gene expression. *Blood Cells, Molecules, and Diseases* *45*, 343-350.

Hacein-Bey-Abina, S., Von Kalle, C., Schmidt, M., Le Deist, F., Wulffraat, N., McIntyre, E., Radford, I., Villeval, J. L., Fraser, C. C., Cavazzana-Calvo, M., and Fischer, A. (2003a). A serious adverse event after successful gene therapy for X-linked severe combined immunodeficiency [1]. *N Engl J Med* *348*, 255-256.

- Hacein-Bey-Abina, S., Von Kalle, C., Schmidt, M., McCormack, M. P., Wulffraat, N., Leboulch, P., Lim, A., Osborne, C. S., Pawliuk, R., Morillon, E., *et al.* (2003b). LMO2-associated clonal T cell proliferation in two patients after gene therapy for SCID-X1. *Science* 302, 415-419.
- Han, Y., Lin, Y. B., An, W., Xu, J., Yang, H.-C., O'Connell, K., Dordai, D., Boeke, J. D., Siliciano, J. D., and Siliciano, R. F. (2008). Orientation-Dependent Regulation of Integrated HIV-1 Expression by Host Gene Transcriptional Readthrough. *Cell Host & Microbe* 4, 134-146.
- Hare, S., Gupta, S. S., Valkov, E., Engelman, A., and Cherepanov, P. (2010). Retroviral intasome assembly and inhibition of DNA strand transfer. *Nature* 464, 232-236.
- He, G., and Margolis, D. M. (2002). Counterregulation of chromatin deacetylation and histone deacetylase occupancy at the integrated promoter of human immunodeficiency virus type 1 (HIV-1) by the HIV-1 repressor YY1 and HIV-1 activator Tat. *Molecular and Cellular Biology* 22, 2965-2973.
- Heintzman, N. D., Stuart, R. K., Hon, G., Fu, Y., Ching, C. W., Hawkins, R. D., Barrera, L. O., Van Calcar, S., Qu, C., Ching, K. A., *et al.* (2007). Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. *Nat Genet* 39, 311-318.
- Hejnar, J., Hájková, P., Plachý, J., Elleder, D., Stepanets, V., and Svoboda, J. (2001). CpG island protects Rous sarcoma virus-derived vectors integrated into nonpermissive cells from DNA methylation and transcriptional suppression. *Proceedings of the National Academy of Sciences of the United States of America* 98, 565-569.
- Holman, A. G., and Coffin, J. M. (2005). Symmetrical base preferences surrounding HIV-1, avian sarcoma/leukosis virus, and murine leukemia virus integration sites. *Proceedings of the National Academy of Sciences of the United States of America* 102, 6103-6107.
- Huang, L., Xu, G.-l., Zhang, J.-q., Tian, L., Xue, J.-l., Chen, J.-z., and Jia, W. (2008). Daxx interacts with HIV-1 integrase and inhibits lentiviral gene expression. *Biochemical and Biophysical Research Communications* 373, 241-245.
- Cherepanov, P., Maertens, G., Proost, P., Devreese, B., Van Beeumen, J., Engelborghs, Y., De Clercq, E., and Debyser, Z. (2003). HIV-1 integrase forms stable tetramers and associates with LEDGF/p75 protein in human cells. *Journal of Biological Chemistry* 278, 372-381.
- Jiang, G., Espeseth, A., Hazuda, D. J., and Margolis, D. M. (2007). c-Myc and Sp1 Contribute to Proviral Latency by Recruiting Histone Deacetylase 1 to the Human Immunodeficiency Virus Type 1 Promoter. *Journal of Virology* 81, 10914-10923.
- Jordan, A., Bisgrove, D., and Verdin, E. (2003). HIV reproducibly establishes a latent infection after acute infection of T cells in vitro. *EMBO Journal* 22, 1868-1877.
- Jordan, A., Defechereux, P., and Verdin, E. (2001). The site of HIV-1 integration in the human genome determines basal transcriptional activity and response to Tat transactivation. *EMBO Journal* 20, 1726-1738.
- Kalpana, G., Marmon, S., Wang, W., Crabtree, G., and Goff, S. (1994). Binding and stimulation of HIV-1 integrase by a human homolog of yeast transcription factor SNF5. *Science* 266, 2002-2006.
- Karlič, R., Chung, H.-R., Lasserre, J., Vlahoviček, K., and Vingron, M. (2010). Histone modification levels are predictive for gene expression. *Proceedings of the National Academy of Sciences* 107, 2926-2931.
- Katz, R. A., Greger, J. G., and Skalka, A. M. (2005). Effects of cell cycle status on early events in retroviral replication. *Journal of Cellular Biochemistry* 94, 880-889.
- Kauder, S. E., Bosque, A., Lindqvist, A., Planelles, V., and Verdin, E. (2009). Epigenetic Regulation of HIV-1 Latency by Cytosine Methylation. *PLoS Pathog* 5, e1000495.

Kim, T., Xu, Z., Clauder-Münster, S., Steinmetz, Lars M., and Buratowski, S. (2012). Set3 HDAC Mediates Effects of Overlapping Noncoding Transcription on Gene Induction Kinetics. *Cell* 150, 1158-1169.

Lander, E. S., Linton, L. M., Birren, B., Nusbaum, C., Zody, M. C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., Fitzhugh, W., *et al.* (2001). Initial sequencing and analysis of the human genome. *Nature* 409, 860-921.

Lenasi, T., Contreras, X., and Peterlin, B. M. (2008). Transcriptional Interference Antagonizes Proviral Gene Expression to Promote HIV Latency. *Cell Host & Microbe* 4, 123-133.

Lesbats, P., Botbol, Y., Chevereau, G., Vaillant, C., Calmels, C., Arneodo, A., Andreola, M. L., Lavigne, M., and Parissi, V. (2011). Functional coupling between HIV-1 integrase and the SWI/SNF chromatin remodeling complex for efficient in vitro integration into stable nucleosomes. *PLoS Pathogens* 7.

Lewinski, M. K., Bisgrove, D., Shinn, P., Chen, H., Hoffmann, C., Hannenhalli, S., Verdin, E., Berry, C. C., Ecker, J. R., and Bushman, F. D. (2005). Genome-wide analysis of chromosomal features repressing human immunodeficiency virus transcription. *Journal of Virology* 79, 6610-6619.

Lewinski, M. K., Yamashita, M., Emerman, M., Ciuffi, A., Marshall, H., Crawford, G., Collins, F., Shinn, P., Leipzig, J., Hannenhalli, S., *et al.* (2006). Retroviral DNA integration: Viral and cellular determinants of target-site selection. *PLoS Pathogens* 2, 0611-0622.

Lewis, P. F., and Emerman, M. (1994). Passage through mitosis is required for oncoretroviruses but not for the human immunodeficiency virus. *Journal of Virology* 68, 510-516.

Li, B., Carey, M., and Workman, J. L. (2007). The Role of Chromatin during Transcription. *Cell* 128, 707-719.

Llano, M., Saenz, D. T., Meehan, A., Wongthida, P., Peretz, M., Walker, W. H., Teo, W., and Poeschla, E. M. (2006). An essential role for LEDGF/p75 in HIV integration. *Science* 314, 461-464.

Maertens, G., Cherepanov, P., Pluymers, W., Busschots, K., De Clercq, E., Debyser, Z., and Engelborghs, Y. (2003). LEDGF/p75 is essential for nuclear and chromosomal targeting of HIV-1 integrase in human cells. *Journal of Biological Chemistry* 278, 33528-33539.

Maertens, G. N., Hare, S., and Cherepanov, P. (2010). The mechanism of retroviral integration from X-ray structures of its key intermediates. *Nature* 468, 326-329.

Machoň, O., Strmen, V., Hejnar, J., Geryk, J., and Svoboda, J. (1998). Sp1 binding sites inserted into the Rous sarcoma virus long terminal repeat enhance LTR-driven gene expression. *Gene* 208, 73-82.

Meehan, A. M., Saenz, D. T., Morrison, J. H., Garcia-Rivera, J. A., Peretz, M., Llano, M., and Poeschla, E. M. (2009). LEDGF/p75 proteins with alternative chromatin tethers are functional HIV-1 cofactors. *PLoS Pathogens* 5.

Melamed, A., Laydon, D. J., Gillet, N. A., Tanaka, Y., Taylor, G. P., and Bangham, C. R. M. (2013). Genome-wide Determinants of Proviral Targeting, Clonal Abundance and Expression in Natural HTLV-1 Infection. *PLoS Pathog* 9, e1003271.

Miccio, A., Poletti, V., Tiboni, F., Rossi, C., Antonelli, A., Mavilio, F., and Ferrari, G. (2011). The GATA1-HS2 Enhancer Allows Persistent and Position-Independent Expression of a β -globin Transgene. *PLoS ONE* 6, e27955.

Mitchell, R. S., Beitzel, B. F., Schroder, A. R. W., Shinn, P., Chen, H., Berry, C. C., Ecker, J. R., and Bushman, F. D. (2004). Retroviral DNA integration: ASLV, HIV, and MLV show distinct target site preferences. *PLoS Biology* 2.

- Narezkina, A., Taganov, K. D., Litwin, S., Stoyanova, R., Hayashi, J., Seeger, C., Skalka, A. M., and Katz, R. A. (2004). Genome-wide analyses of avian sarcoma virus integration sites. *Journal of Virology* 78, 11656-11663.
- Nguyen, A. T., and Zhang, Y. (2011). The diverse functions of Dot1 and H3K79 methylation. *Genes & Development* 25, 1345-1358.
- Okitsu, C. Y., and Hsieh, C.-L. (2007). DNA Methylation Dictates Histone H3K4 Methylation. *Molecular and Cellular Biology* 27, 2746-2757.
- Plachý, J., Kotáb, J., Divina, P., Reinišova, M., Šenigl, F., and Hejnar, J. (2010). Proviruses selected for high and stable expression of transduced genes accumulate in broadly transcribed genome areas. *J Virol* 84, 4204-4211.
- Poleshko, A., Einarson, M. B., Shalginskikh, N., Zhang, R., Adams, P. D., Skalka, A. M., and Katz, R. A. (2010). Identification of a functional network of human epigenetic silencing factors. *Journal of Biological Chemistry* 285, 422-433.
- Poleshko, A., Palagin, I., Zhang, R., Boimel, P., Castagna, C., Adams, P. D., Skalka, A. M., and Katz, R. A. (2008). Identification of cellular proteins that maintain retroviral epigenetic silencing: Evidence for an antiviral response. *Journal of Virology* 82, 2313-2323.
- Rafati, H., Parra, M., Hakre, S., Moshkin, Y., Verdin, E., and Mahmoudi, T. (2011). Repressive LTR Nucleosome Positioning by the BAF Complex Is Required for HIV Latency. *PLoS Biol* 9, e1001206.
- Reinisova, M., Pavlicek, A., Divina, P., Geryk, J., Plachy, J., and Hejnar, J. (2008). Target site preference of subgroup C Rous sarcoma virus integration into the chicken DNA. *Open Genomics J* 1, 6-12.
- Roe, T., Reynolds, T. C., Yu, G., and Brown, P. O. (1993). Integration of murine leukemia virus DNA depends on mitosis. *EMBO Journal* 12, 2099-2108.
- Rowe, H. M., and Trono, D. (2011). Dynamic control of endogenous retroviruses during development. *Virology* 411, 273-287.
- Santoni, F. A., Hartley, O., and Luban, J. (2010). Deciphering the code for retroviral integration target site selection. *PLoS Comput Biol* 6, e1001008.
- Saxonov, S., Berg, P., and Brutlag, D. L. (2006). A genome-wide analysis of CpG dinucleotides in the human genome distinguishes two distinct classes of promoters. *Proceedings of the National Academy of Sciences of the United States of America* 103, 1412-1417.
- Shalginskikh, N., Poleshko, A., Skalka, A. M., and Katz, R. A. (2013). Retroviral DNA Methylation and Epigenetic Repression Are Mediated by the Antiviral Host Protein Daxx. *Journal of Virology* 87, 2137-2150.
- Shan, L., Yang, H. C., Rabi, S. A., Bravo, H. C., Shroff, N. S., Irizarry, R. A., Zhang, H., Margolick, J. B., Siliciano, J. D., and Siliciano, R. F. (2011). Influence of host gene transcription level and orientation on HIV-1 latency in a primary-cell model. *Journal of Virology* 85, 5384-5393.
- She, X., Rohl, C., Castle, J., Kulkarni, A., Johnson, J., and Chen, R. (2009). Definition, conservation and epigenetics of housekeeping and tissue-enriched genes. *BMC Genomics* 10, 269.
- Shirakawa, K., Chavez, L., Hakre, S., Calvanese, V., and Verdin, E. (2013). Reactivation of latent HIV by histone deacetylase inhibitors. *Trends in Microbiology*.
- Schneider, R., Bannister, A. J., Myers, F. A., Thorne, A. W., Crane-Robinson, C., and Kouzarides, T. (2004). Histone H3 lysine 4 methylation patterns in higher eukaryotic genes. *Nat Cell Biol* 6, 73-77.

- Schrijvers, R., Vets, S., De Rijck, J., Malani, N., Bushman, F., Debyser, Z., and Gijsbers, R. (2012). HRP-2 determines HIV-1 integration site selection in LEDGF/p75 depleted cells. *Retrovirology* *9*, 84.
- Schröder, A. R. W., Shinn, P., Chen, H., Berry, C., Ecker, J. R., and Bushman, F. (2002). HIV-1 integration in the human genome favors active genes and local hotspots. *Cell* *110*, 521-529.
- Šenigl, F., Auxt, M., and Hejnar, J. (2012). Transcriptional provirus silencing as a crosstalk of de novo DNA methylation and epigenomic features at the integration site. *Nucleic Acids Research* *40*, 5298-5312.
- Šenigl, F., Plachý, J., and Hejnar, J. (2008). The core element of a CpG island protects avian sarcoma and leukosis virus-derived vectors from transcriptional silencing. *Journal of Virology* *82*, 7818-7827.
- Taganov, K. D., Cuesta, I., Daniel, R., Cirillo, L. A., Katz, R. A., Zaret, K. S., and Skalka, A. M. (2004). Integrase-specific enhancement and suppression of retroviral DNA integration by compacted chromatin structure in vitro. *Journal of Virology* *78*, 5848-5855.
- Uren, A. G., Mikkers, H., Kool, J., van der Weyden, L., Lund, A. H., Wilson, C. H., Rance, R., Jonkers, J., van Lohuizen, M., Berns, A., and Adams, D. J. (2009). A high-throughput splinkerette-PCR method for the isolation and sequencing of retroviral insertion sites. *Nature protocols* *4*, 789-798.
- van der Sluis, R. M., van Montfort, T., Pollakis, G., Sanders, R. W., Speijer, D., Berkhout, B., and Jeeninga, R. E. (2013). Dendritic Cell-induced Activation of Latent HIV-1 Provirus in Actively Proliferating Primary T Lymphocytes. *PLoS Pathog* *9*, e1003259.
- van Lint, C., Emiliani, S., Ott, M., and Verdin, E. (1996). Transcriptional activation and chromatin remodeling of the HIV-1 promoter in response to histone acetylation. *EMBO Journal* *15*, 1112-1120.
- Vavouri, T., and Lehner, B. (2012). Human genes with CpG island promoters have a distinct transcription-associated chromatin organization. *Genome Biology* *13*, R110.
- Venkatesh, S., Smolle, M., Li, H., Gogol, M. M., Saint, M., Kumar, S., Natarajan, K., and Workman, J. L. (2012). Set2 methylation of histone H3 lysine 36 suppresses histone exchange on transcribed genes. *Nature* *489*, 452-455.
- Venter, J. C., Adams, M. D., Myers, E. W., Li, P. W., Mural, R. J., Sutton, G. G., Smith, H. O., Yandell, M., Evans, C. A., Holt, R. A., *et al.* (2001). The sequence of the human genome. *Science* *291*, 1304-1351.
- Verdin, E., Paras Jr, P., and Van Lint, C. (1993). Chromatin disruption in the promoter of human immunodeficiency virus type 1 during transcriptional activation. *EMBO Journal* *12*, 3249-3259.
- Wang, G. P., Ciuffi, A., Leipzig, J., Berry, C. C., and Bushman, F. D. (2007). HIV integration site selection: Analysis by massively parallel pyrosequencing reveals association with epigenetic modifications. *Genome Research* *17*, 1186-1194.
- Wang, H., Jurado, K. A., Wu, X., Shun, M.-C., Li, X., Ferris, A. L., Smith, S. J., Patel, P. A., Fuchs, J. R., Cherepanov, P., *et al.* (2012a). HRP2 determines the efficiency and specificity of HIV-1 integration in LEDGF/p75 knockout cells but does not contribute to the antiviral activity of a potent LEDGF/p75-binding site integrase inhibitor. *Nucleic Acids Research* *40*, 11518-11530.
- Wang, X., Spandidos, A., Wang, H., and Seed, B. (2012b). PrimerBank: a PCR primer database for quantitative gene expression analysis, 2012 update. *Nucleic Acids Research* *40*, D1144-D1149.
- Wei, S. Q., Mizuuchi, K., and Craigie, R. (1997). A large nucleoprotein assembly at the ends of the viral DNA mediates retroviral DNA integration. *EMBO Journal* *16*, 7511-7520.

Williams, S. A., Chen, L. F., Kwon, H., Ruiz-Jarabo, C. M., Verdin, E., and Greene, W. C. (2006). NF- κ B p50 promotes HIV latency through HDAC recruitment and repression of transcriptional initiation. *EMBO Journal* 25, 139-149.

Wolf, D., and Goff, S. P. (2007). TRIM28 Mediates Primer Binding Site-Targeted Silencing of Murine Leukemia Virus in Embryonic Cells. *Cell* 131, 46-57.

Wolf, D., and Goff, S. P. (2009). Embryonic stem cells use ZFP809 to silence retroviral DNAs. *Nature* 458, 1201-1204.

Wu, X., Li, Y., Crise, B., and Burgess, S. M. (2003). Transcription start regions in the human genome are favored targets for MLV integration. *Science* 300, 1749-1751.

Wu, X. L., Li, Y., Crise, B., Burgess, S. M., and Munroe, D. J. (2005). Weak palindromic consensus sequences are a common feature found at the integration target sites of many retroviruses. *Journal of Virology* 79, 5211-5214.

Yao, S., Sukonnik, T., Kean, T., Bharadwaj, R. R., Pasceri, P., and Ellis, J. (2004). Retrovirus silencing, variegation, extinction, and memory are controlled by a dynamic interplay of multiple epigenetic modifications. *Molecular Therapy* 10, 27-36.

Zhou, V. W., Goren, A., and Bernstein, B. E. (2011). Charting histone modifications and the functional organization of mammalian genomes. *Nat Rev Genet* 12, 7-18.