

Univerzita Karlova v Praze
Přírodovědecká fakulta

Speciální chemicko-biologické obory
Molekulární biologie a biochemie organismů



Dalibor Miklík

Integrační preference retrovirů a retrovirových vektorů
Genomic preferences in integration and expression of retroviruses and retrovirus-derived
vectors

Bakalářská práce

Školitel: RNDr. Jiří Hejnar, CSc.

Praha, 2010/2011

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, 28.04.2011

Podpis

Abstrakt

Integrace a vytvoření proviru jsou klíčovými kroky životního cyklu retrovirů. Genomové studie ukazují, že integrace není náhodným procesem a že skupiny příbuzných retrovirů prokazují různé integrační preference. Nejčastěji hodnocenými genomickými charakteristikami určujícími integrační preference retrovirů jsou transkripční jednotky, oblasti transkripčních startů a CpG ostrovy. S rozvíjejícími se poznatky o struktuře genomu je možná i definice nových, převážně epigenetických značek hrajících roli v integračních preferencích retrovirů. Integrace do různých oblastí genomu má přímý vliv na provirovou expresi, a tedy na produkci virového potomstva. Zatímco integrace do některých oblastí genomu má za následek umlčení proviru, které je zprostředkováno a udržováno množstvím nejrůznějších faktorů, jiné oblasti genomu jsou naopak schopny udržovat provirus v aktivním stavu. Tyto poznatky lze aplikovat ve vývoji bezpečných a účinných retrovirových vektorů, stejně jako ve využití modifikovaných retrovirů jako ukazatelů epigenetického či expresního stavu genomových úseků.

Abstract

Integration and provirus establishment are the key steps of retroviral life cycle. Genome-wide studies show that the integration is not a random process and that groups of related retroviruses display distinguishable patterns of integration preferences. The most rated genomic features forming the integration preferences of retroviruses are transcription units, transcription start sites and CpG islands. Whereas extending knowledge of genome structure, new, mainly epigenetic marks, which have a relationship to retroviral preferences are being defined. The integration into a specific genome region has a straight influence on the provirus expression and therefore on the production of virus progeny. While integration into some regions results in provirus silencing which is managed and maintained by variety of factors, some other genomic regions are *vice-versa* capable of stable provirus expression maintenance. These findings have implications for construction of safe and efficient retroviral vectors as well as for use of modified retroviruses as markers for an epigenetic and expression profile determination of genome regions.

Klíčová slova: integrace retrovirů, integrační preference, exprese proviru, retrovirové vektory, umlčování

Key words: retroviral integration, integration preferences, proviral expression, retroviral vectors, silencing

Obsah

SEZNAM POUŽITÝCH ZKRATEK.....	5
1. ÚVOD.....	7
2. INTEGRACE RETROVIROVÉ DNA.....	9
2.1. REVERZNĚ TRANSKRIPČNÍ KOMPLEX	9
2.2. PREINTEGRAČNÍ KOMPLEX	9
2.3. VSTUP PIC DO JÁDRA	10
2.4. INTEGRÁZA – STRUKTURA A FUNKCE.....	10
2.5. CHEMIE INTEGRAČNÍ REAKCE	11
2.6. POSTINTEGRAČNÍ ÚPRAVY	11
3. INTEGRACNÍ PREFERENCE.....	12
3.1. PRIMÁRNÍ SEKVENCE	12
3.2. NUKLEOZÓMY	13
3.3. CpG OSTROVY	13
3.4. HISTONOVÉ MODIFIKACE.....	14
3.5. OBLASTI TRANSKRIPČNÍCH STARTŮ.....	15
3.6. TRANSKRIPČNÍ JEDNOTKY	16
3.7. PROTEINOVÉ INTERAKCE	17
3.8. CHROMATINOVÁ STRUKTURA.....	17
3.9. BUNĚČNÝ TYP.....	18
3.10. SOUHRN INTEGRACNÍCH PREFERENCÍ RETROVIRŮ	18
4. EXPRESE A UMLČOVÁNÍ PROVIRU	20
4.1. HLAVNÍ FAKTORY V UMLČENÍ PROVIROVÉ EXPRESE.....	21
4.2. INTERAKČNÍ SÍŤ REPRESIVNÍCH EPIGENETICKÝCH MODIFIKACÍ.....	22
4.3. INICIACE UMLČOVÁNÍ PROVIRU	23
4.4. VLIV INTEGRACNÍHO MÍSTA NA EXPRESI PROVIRU	23
4.5. UMLČOVÁNÍ JAKO AKTIVNÍ OBRANA HOSTITELE	24
4.6. SOUHRN FAKTORŮ OVLIVŇUJÍCÍCH PROVIROVOU EXPRESI.....	25
5. PREFERENCE A EXPRESE RETROVIROVÝCH VEKTORŮ	27
5.3. ZMĚNA INTEGRACNÍCH PREFERENCÍ RETROVIROVÝCH VEKTORŮ	27
5.4. CÍLENÍ INTEGRACE MODIFIKACÍ VIROVÝCH PROTEINŮ.....	27
5.5. CÍLENÍ INTEGRACE MODIFIKACÍ BUNĚČNÝCH PROTEINŮ.....	28
5.6. UDRŽENÍ STABILNÍ EXPRESE VEKTORU	28
5.7. VYUŽITÍ A MOŽNOSTI RETROVIROVÝCH VEKTORŮ.....	29
6. ZÁVĚR.....	31
SEZNAM POUŽITÉ LITERATURY	32

Seznam použitých zkratek

5'-azaC	5'-azacytidine	5'-azacytidin
AIDS	acquired imunodeficiency syndrome	syndrom získané imunodeficiencie
ASLV	avian sarcoma leukosis virus	virus způsobující sarkomy a leukózy u ptáků
ASV	avian sarcoma virus	virus způsobující sarkomy u ptáků
BAF	barrier-to-autointegration factor	autointegraci zabráňující faktor
BLV	bovine leukemia virus	virus způsobující leukemii u krav
CA	capsid protein	kapsidový protein
CBF-1	CCAAT binding factor 1	CCAAT vazebný faktor 1
cDNA	complementary DNA	DNA kopie
DAXX	death-domain associated protein	protein obsahující doménu asociovanou se smrtí
GLP	G9a-like protein	G9a podobný protein
H3K20me	histone H3 lysine 20 methylation	methylace lysinu 20 hisonu H3
H3K27me	histone H3 lysine 27 methylation	methylace lysinu 27 hisonu H3
H3K4ac	histone H3 lysine 4 acetylation	acetylance lysinu 4 hisonu H3
H3K4me1/2/3	histone H3 lysine 4 mono-/di-/trimethylation	mono-/di-/trimethylace lysinu 4 hisonu H3
H3K9ac	histone H3 lysine 9 acetylation	acetylance lysinu 9 hisonu H3
H3K9me3	histone H3 lysine 9 trimethylation	trimethylace lysinu 9 hisonu H3
H4ac	histone 4 acetylation	Acetylance histonu H4
HDAC	histone deacetylase	histon deacetyláza
HERV	human endogenous retroviruses	lidské endogenní retroviry
HIV	human immunodeficiency virus	virus způsobující selhání imunity u lidí
HMT	histone methyltransferase	histonová methytransferáza
HP1	heterochromatin protein 1	heterochromatinový protein 1
HTLV-1	human T-lymphotropic virus type 1	lidský leukemický virus typu 1
CHAF1A	chromatin assembly factor 1, subunit A	podjednotka A chromatin sestavujícího faktoru 1
IN	integrase	integráza
kb	kilobase	kilobáze
LEDGF/p75	lens epithelium-derived growth factor/protein 75	růstový faktor odvozený z epitelu čočky/protein 75
LTR	long terminal repeat	dlouhé koncové repetice
MAR	matrix-attachnment region	oblasti kontaktu s jadernou matrix
MBD	methyl-CpG binding domain	methyl-CpG vazebná doména

MeC-P	methyl-CpG binding protein	methyl-CpG vazebný protein
MLV	murine leukemia virus	virus způsobující leukemii u myši
MMTV	mouse mammary tumor virus	virus způsobující rakovinu mléčné žlázy u myši
MoMLV	Moloney murine leukemia virus	Moloneyho myši leukemický virus
NCR CpG	noncoding region CpG	CpG nekódující oblasti
NFR	nucleosome-free regions	oblasti postrádající nukleozómy
NF-κB	nuclear factor κB	jaderný faktor κB
NLS	nuclear localization signal	jaderný lokalizační signál
PIC	preintegration complex	preintegrační komplex
PR	proteinase	proteínáza
RSV	Rous sarcoma virus	Rousův virus způsobující sarkomy
RT	reverse transcriptase	reverzní transkriptáza
RTC	reverse transcription complex	reverzně transkripční komplex
SIV	simian immunodeficiency virus	virus způsobující selhání imunity u opic
SCID	severe combined immunodeficiency	těžká kombinovaná imunodeficience
SETDB1	SET domain, bifurcated 1	SET doménu obsahující protein 1
Suv39H1	suppressor of variegation 3-9 homolog 1	homolog 1 supresoru variabilit 3-9
tDNA	target DNA	cílová DNA
TFBS	transcription factor-binding sites	vazebná místa transkripčních faktorů
TSA	trichostatin A	trichostatin A
TSD	target site duplication	duplikace cílových sekvencí
TSS	transcription start sites	oblasti transkripčních startů
TU	transcription units	transkripční jednotky
YY1	Yin Yang 1	jin jang 1
ZFD	zinc finger domain	doména zinkového prstu

1. Úvod

Retroviry jsou obalené viry z čeledi *Retroviridae* nesoucí v kapsidě spolu s množstvím virových proteinů i dvě molekuly RNA v kladném smyslu. Od ostatních RNA virů se liší především způsobem produkce virového potomstva – replikačním cyklem. Po vstupu viru do buňky je RNA pomocí reverzní transkripce, tedy přepisu z RNA do DNA, přepsána do formy lineární dvouřetězcové DNA označované jako cDNA. Tato cDNA je v jádře působením speciálního virového enzymu integrázy vložena do chromozomální DNA hostitele a tvoří provirus. Provirus je buněčnou mašinerií přepisován do mRNA, která slouží buďto jako předloha pro expresi virových proteinů, nebo je společně s dalšími složkami retrovirové kapsidy zabalena do virové částice. Viriony složené z lipidového obalu obsahujícího povrchové glykoproteiny (SU) a nukleoproteinového jádra (anglicky „core“) jsou sestavovány pod buněčnou membránou a způsobem připomínajícím pučení unikají do mezibuněčného prostoru. Takto vzniklé viriony jsou ještě nezralé a k jejich úplnému dozrání dochází až mimo buňky. Zralé nukleoproteinové jádro je složeno z kapsidového proteinu (CA), matrixového proteinu (MA), nukleokapsidového proteinu (NC), proteázy (PR), reverzní transkriptázy (RT), integrázy (IN), 2 molekul virové mRNA a 2 molekul buněčné tRNA. Takovéto zralé viriony již jsou schopny úplné infekce.

Genom retrovirů je tvořen RNA s 7-methylguanosiinovou čepičkou na 5'konci a polyA řetězcem na 3'konci. Oba konce RNA obsahují krátké netranslatované sekvence označované jako U5 (na 5'konci) a U3 (na 3'konci). Poblíž obou sekvencí se také nachází krátká repetice. Mezi těmito sekvencemi jsou pak umístěny strukturální geny viru a sekvence nutné pro zabalení virové mRNA do virionu nebo zahájení reverzní transkripce. Během reverzní transkripce dochází na obou koncích cDNA k vytvoření dlouhých koncových repetit (LTR).

Reverzní transkripce retrovirové RNA, integrace cDNA a exprese proviru byly a stále jsou žhavými tématy výzkumu na poli molekulární biologie a retrovirologie. Zájem o detailnější poznání základů těchto dějů byl zpočátku způsoben onkogenním potenciálem některých retrovirů. Další vlnu zájmu o replikační cyklus retrovirů rozhodně přinesl objev viru způsobujícího selhání imunity u lidí (HIV) jakožto původce syndromu získané imunodeficience (AIDS). Podrobné poznávání mechanismů reverzní transkripce a integrace je důležité právě pro vývoj nových účinných virostatik v boji proti AIDS. Také zjištění, že retroviry působí jako přirozené vektory pro přenos genů mezi organismy, zvýšilo zájem o výzkum retrovirového replikačního cyklu. Metodami genového inženýrství upravené retroviry jsou pak v různých formách využívány jako vektory pro přenos a expresi genů jak v laboratořích základního výzkumu, tak v klinice jako vektory pro genovou terapii. Nejpoužívanějšími se v této oblasti staly myší leukemické viry (MLV) z rodu gammaretrovirů, které byly mimo jiné použity pro genovou terapii těžké kombinované imunodeficience (SCID). Vývoj lymfoblastické leukemie u několika pacientů léčených pomocí retrovirové genové terapie (Hacein-Bey-Abina *et al.*,

2003) vedl kromě pozastavení klinických zkoušek také ke zvýšení zájmu o integrační mechanismy a preference retrovirů a k hledání a vývoji nových retrovirových vektorů.

Tato práce má za cíl shrnout poznatky o integračních preferencích retrovirů a vlivu těchto preferencí na expresi retrovirových genů s ohledem na rozdíly mezi jednotlivými retroviry a využití těchto poznatků při vývoji efektivních vektorů. V úvodních kapitolách jsou shrnuty hlavně základní informace o mechanismu integrace virové cDNA a vytvoření proviru. V kapitole o integračních preferencích retrovirů, jsou představeny genomové charakteristiky, které mohou jakkoliv definovat preferovaná integrační místa retrovirů, a to včetně jejich funkce v hostitelském genomu. Kapitola o expresi proviru je zaměřena hlavně na problematiku ztráty exprese proviru, přičemž jsou popsány možné příčiny tohoto jevu. V poslední části se pak nachází popis retrovirových vektorů z hlediska možnosti ovlivnění jejich integračních preferencí a exprese.

2. Integrace retrovirové DNA

Integrace, a tedy vytvoření proviru, je děj nezbytný pro správnou infekci buněk a produkci virového potomstva. Na druhou stranu bylo pozorováno, že v nově infikovaných buňkách dochází k hromadění neintegrované DNA a že i tato neintegrovaná DNA může mít určitou, ač velmi omezenou expresní aktivitu (přehledně uvedeno v Wu, 2008). Neintegrovaná DNA se však vyznačuje poměrně krátkou životností a taková infekce má spíše latentní charakter. K účinné a stabilní infekci je tedy třeba vytvoření proviru.

2.1. Reverzně transkripční komplex

Poté, co virus vstoupí do buňky, je v cytoplazmě hostitelské buňky díky částečnému rozvolnění kapsidy a přístupu volných nukleotidů započata reverzní transkripce virové RNA. Ze dvou virových RNA tak vzniká jediná DNA kopie (cDNA) virového genomu, jejíž struktura je až na malý počet terminálních nukleotidů shodná s provirovou DNA. Nukleoproteinový komplex, ve kterém reverzní transkripce probíhá je označován jako reverzně transkripční komplex (RTC). I když byly pozorovány rozdíly ve složení RTC mezi některými retroviry, základní charakteristikou tohoto nukleoproteinového komplexu zůstává zprostředkování reverzní transkripce a postupná ztráta kapsidových proteinů spojená s rozvolňováním nukleoproteinového jádra viru a celkovou maturací komplexu (Fassati and Goff, 1999; Fassati and Goff, 2001). Výjimku v tomto schématu tvoří viry z rodu spumaretrovirů, u nichž dochází k reverzní transkripci již v buňkách produkujících virus a do nových hostitelských buněk se tak dostávají i částice obsahující cDNA (přehledně o replikačním cyklu spumavirů v Delelis *et al.*, 2004).

2.2. Preintegrační komplex

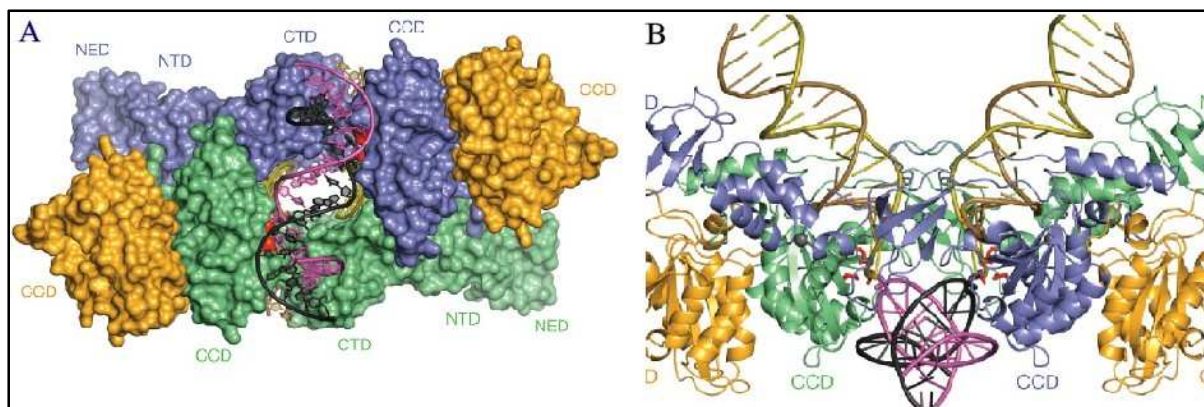
Po dokončení reverzní transkripce dozrává RTC při svém transportu podél mikrotubulů k jádru v preintegrační komplex (PIC). PIC je charakterizován jako nukleoproteinový komplex složený z cDNA a proteinů jak virového, tak i buněčného původu, schopný zprostředkovávat integrační reakci. Nejdůležitější složkou PIC je bezesporu IN, která zodpovídá za hlavní funkce PIC – zpracování 3' konců cDNA a samotnou integraci cDNA do genomu hostitelské buňky. U jiných virů navíc může hrát roli v cíleném transportu cDNA přes jadernou membránu. V PIC však bývají k nalezení také nejrůznější proteiny buněčného původu. Patří mezi ně např. faktory zabraňující autointegraci (BAF) (Lee and Craigie, 1998) nebo proteiny ovlivňující jaderný transport PIC a hrající roli v integračních preferencích retrovirů. Takovým proteinem je např. růstový faktor odvozený z epitelu čičky/protein 75 (LEDGF/p75). PIC je dobře izolovatelný z akutně infikovaných buněk a jeho schopnost samostatně zprostředkovávat integraci virové cDNA umožnila první *in vitro* experimenty, které položily základy výzkumu integrace a integračních preferencí.

2.3. Vstup PIC do jádra

Další překážku na cestě k úspěšné integraci virové cDNA do hostitelského genomu představuje jaderná membrána. Některé retroviry, jako Moloneyho myší leukemický virus (MoMLV), jsou schopny integrace pouze v dělících se buňkách, kdy se během mitózy jádro rozpadá a mitotické chromozómy jsou tak pro PIC volně přístupné (Roe *et al.*, 1993). Jiné viry, zastupované např. virem HIV-1, do jádra aktivně vstupují přes jaderné póry a jsou tak schopny integrace i v nedělících se buňkách (Weinberg *et al.*, 1991). Svou roli v aktivním transportu přes jaderné póry hraje nejspíše jak samotná cDNA, tak i proteinové složky PIC nesoucí jaderné lokalizační signály (NLS). Více informací o formování PIC a vstupu tohoto komplexu do jádra lze dohledat v přehledném článku autorů Suzuki and Craigie (2007).

2.4. Integráza – struktura a funkce

Integráza je nukleotidyltransferáza (EC 2.7.7) z enzymové superrodiny DDE(D) rekombináz, kódovaná 3' terminální částí Pol genu. Z polyproteinového prekurzoru je pomocí PR vystřižena právě při maturaci virionu, tedy až po uvolnění virových částic z infikovaných buněk. Monomer integrázy je tvořen 3 hlavními doménami: C-terminální (CTD), katalytickou (CCD) a N-terminální (NTD) doménou. Integráza je nejdůležitější složkou PIC, kde se ve formě tetramerů váže na konce cDNA (Cherepanov *et al.*, 2003). Nukleoproteinový komplex integrázy a případných buněčných faktorů s terminálními úseky cDNA se nazývá intazóm (Wei *et al.*, 1997). Celý tetramer je vlastně dimerem dimerů. Všechny interakce mezi DNA (cDNA i tDNA) a integrázou jsou zprostředkovávány vnitřními podjednotkami tetramerů. Samotná integráza pokrývá oblast asi 17 nukleotidů z každého konce cDNA. V prostoru mezi oběma dimery (ve velkém žlábků; viz obr.1) dochází k vazbě a výraznému ohybu tDNA, což nejspíš napomáhá samotné integraci. Celý intazóm prochází během integrační reakce několika fázemi charakterizovanými přítomností intermediátů. Všechny tyto struktury včetně přesných míst interakce integrázy s cDNA a tDNA byly v roce 2010 potvrzeny pomocí krystalografických struktur intazómu spumaretroviru a intermediátů integrační reakce (Hare *et al.*, 2010; Maertens *et al.*, 2010). Poprvé tak byl publikován detailní mechanistický popis integrace, který mimo jiné poskytuje silný nástroj pro homologní modelování intazómů různých retrovirů a vývoj efektivnějších inhibitorů integrázy v boji proti viru HIV.



Obr.1. Struktura intazómu. (A) Tetramer integrázy s tDNA, která je navázána do výrazného žlábků mezi dimery integrázy. (B) Tentýž model z jiného pohledu. Zde je vidět, kde dochází k vytvoření nových fosfoesterových vazeb. tDNA je značena fialovým a černým řetězcem, cDNA je žlutá. CCC – katalytická doména, CTD – C-terminální doména, NTD – N-terminální doména. Převzato a upraveno z Maertens *et al.* (2010).

2.5. Chemie integrační reakce

Po navázání na cDNA katalyzuje integráza vystřížení několika terminálních nukleotidů (nejčastěji 2 či 3) na 3' koncích obou řetězců cDNA (v anglické literatuře je reakce označována jako 3' processing), čímž je odkryta reaktivní hydroxylová (OH) skupina adenosinu z invariantního CA dinukleotidu. Tato OH skupina na ustupujících 3' koncích cDNA funguje jako nukleofil při transesterifikační reakci, tedy integraci cDNA do hostitelského genomu. Obě tyto reakce, jak úprava 3' konce, tak integrace, jsou v podstatě shodné svým mechanismem. V obou případech se jedná o nukleofilní substituci S_N2 typu, kdy výchozí látkou je alespoň jeden polynukleotidový řetězec a nukleofilem je hydroxylová skupina. Zatímco při úpravě 3' konců se jedná spíše o hydrolýzu (OH skupinu poskytuje voda) a produkty jsou vystřížený oligonukleotid a upravená cDNA, při integraci poskytuje OH skupinu 3' koncový adenosin a produkty reakce jsou nové fosfodiesterové vazby mezi virovou a cílovou DNA.

2.6. Postintegrační úpravy

Po integrační reakci dochází ke vzniku jednořetězcových mezer v tDNA a nekomplementárních přesahů 5' koncových nukleotidů integrované cDNA. Tyto přesahy tvoří nukleotidy komplementární k původním (později odstraněným) 3' koncům cDNA. Jednořetězcové mezery i přesahy jsou opraveny buněčnými reparačními mechanismy. Charakteristické také je to, že cílové fosfodiesterové vazby nejsou v tDNA umístěny přímo proti sobě, ale jsou od sebe vzdáleny několik nukleotidů. Okolo proviru, který se tak stává stabilní součástí genetické informace hostitelské buňky, tím vznikají krátké přímé repetice (TSD). Délka repetice se pohybuje mezi 4-6 nukleotidy a bývá charakteristická pro různé skupiny retrovirů.

3. Integrační preference

Integrace retrovirů, jak ukazuje stále více výsledků, není zcela náhodná. Na tuto skutečnost poukazovaly již první *in vitro* experimenty s integračními systémy obsahujícími PIC nebo pouhou integrázu. Dále bylo dokázáno, že různé skupiny retrovirů se v preferencích pro určitá místa v genomu méně či více liší. Mluvíme zde tedy o integračních preferencích retrovirů. Zájem o charakterizaci integračních preferencí vzrostl hlavně poté, co došlo k vývoji leukemických onemocnění a při léčbě genovou terapií, kde jsou využívány právě retrovirové vektory. Získání sekvence lidského genomu (Venter *et al.*, 2001; Lander *et al.*, 2001) spolu s rozvojem metod izolace integračních míst umožnilo efektivnější určování genomových charakteristik, které se mohou podílet na výběru integračního místa. Výběr místa integrace je navíc i důležitý z hlediska exprese proviru. Nejprozkoumanější jsou v této oblasti viry HIV, MLV a viru způsobujícího sarkomy a leukózy u ptáků (ASLV). Studie integračních preferencí však byly prováděny i u dalších retrovirů. Přesto, že bylo identifikováno množství genomových charakteristik, které nejspíše hrají roli v integraci retrovirů, přesný mechanismus výběru integračního místa zatím znám není.

3.1. Primární sekvence

První charakteristikou, která byla určována již při *in vitro* studiích s bezbuněčnými systémy, byla primární sekvence tDNA v místě integrace. Jak však bylo zjištěno, retrovirová integrace není sekvenčně specifická, tj. nebyla dosud objevena žádná sekvence specifická pro vazbu virové integrázy. Co však bylo nalezeno při rozsáhlejších studiích (ať už *in vivo* nebo při analýze dříve vyprodukovaných dat), byla slabá konsensus sekvence, která má navíc charakter palindromu (Holman and Coffin, 2005; Wu *et al.*, 2005), což odpovídá symetrické vazbě tetrameru IN na tDNA. Slabá konsenzuální sekvence znamená, že daná sekvence je charakterizována více či méně pravděpodobnou přítomností daného nukleotidu na daném místě sekvence v porovnání s náhodným výskytem nukleotidů. Žádný nukleotid přitom není úplně „zakázán“. K vyhledání konsenzuální sekvence je tedy zapotřebí velkého množství dat a určitě není překvapením, že takto určená sekvence se ve skutečnosti v místech integrace vyskytuje jen výjimečně. Ve studii autorů Wu *et al.* (2005) se takováto slabá konsenzuální sekvence vyskytovala v 10 z 334 integračních míst HIV-1. Mutační analýzy aminokyselinových zbytků IN interagujících přímo s tDNA navíc vykazovaly pouze mírné změny v nukleotidovém složení integračních míst (Maertens *et al.*, 2010).

Z primární sekvence však lze určit některé fyzikální vlastnosti cílové DNA. Jde například o určení ohebnosti dané oblasti či její náchylnost k různým deformacím, které může způsobit vazba proteinů, jako je např. virová integráza či buněčné DNA vazebné proteiny. Ačkoliv se slabé konsenzuální sekvence mohou pro cílová místa integrace jednotlivých retrovirů lišit, fyzikální rysy sekvencí popsané výše, jsou do určité míry retrovirům společné (Wu *et al.*, 2005). V této studii pak bylo poukázáno, že největší tendenci k ohybu měly cílové sekvence HIV-1 spolu s virem opičích

imunodeficiencie (SIV). Tendenci k integraci do lehce ohybatelných úseků DNA však vykazovaly i MLV či ASLV.

Primární sekvence, vzhledem k závěrům mnoha studií, rozhodně není faktorem, který by vysvětloval integrační preference retrovirů. Zdá se spíše, že nukleotidové složení integračního místa může poskytovat úseku DNA vlastnosti, které jsou pro integraci příznivé a konkrétní místo je IN vybráno až po navedení PIC k tDNA jinými faktory. Vzhledem k výraznému ohybu tDNA v žlábků mezi dimery IN a konkrétními interakcemi IN a tDNA (Maertens *et al.*, 2010), bude nejspíše právě ohebnost daného úseku nejdůležitější charakteristikou určenou primární sekvencí. Různá stavba slabých konsensus sekvencí zároveň poukazuje na rozdílné interakce PIC různých retrovirů a tDNA.

3.2. Nukleozómy

Nukleozómy jsou útvary tvořící 11 nm chromatinové vlákno připomínající korálky na niti. Tyto útvary jsou tvořeny oktamerem histonů a 146-147 nukleotidy DNA, která je kolem tohoto oktameru omotána v necelých 2 otočkách. DNA tak tvoří krátký levotočivý superhelix, kde velký žlábek DNA je otočen buďto směrem k histonům nebo od něj a to s periodickým opakováním. Pozice nukleozómů v DNA jsou důležitým faktorem tvorby chromatinového prostředí, přičemž jejich umístění může být kódováno v sekvenci DNA (Kaplan *et al.*, 2009; Segal *et al.*, 2006). Znalost tzv. „nukleozomálního kódu“ umožňuje tvorbu modelů pro predikci umístění nukleozómů. Nynější modely však správně odhadují umístění asi 50% nukleozómů. Navíc je stále vedena diskuze ohledně samotné existence „nukleozomového kódu“ *in vivo* (Zhang *et al.*, 2009; Zhang *et al.*, 2010).

Na vliv nukleozómů na integraci retrovirů bylo poukázáno již v *in vitro* experimentech s obnaženou DNA a DNA s nukleozómy (Pruss *et al.*, 1994). V těchto experimentech se ukázalo, že přítomnost nukleozómů v tDNA zvyšuje pravděpodobnost integrace do určitých míst na tDNA, která se nacházejí právě v centru nukleozómů. Integrace je navíc cílená do velkého žlábků nukleozomální DNA, který je obrácen směrem od histonového oktameru. Tyto závěry navíc potvrzuje studie autorů Wang *et al.* (2007), jejichž výsledky poukazují na preferenční integraci HIV-1 do míst, kde byla s využitím studie autorů Segal *et al.* (2006) predikována přítomnost nukleozómů. Preference integrace do nukleozomální DNA může být vysvětlena výrazným ohybem této DNA, který je potřebný pro vazbu PIC na tDNA, čímž může být usnadněn ohyb tDNA pro integrázu. Takovéto vysvětlení je v souladu s analýzami konsenzuálních sekvencí integračních míst stejně jako s krystalografickými strukturami intazómu a tDNA.

3.3. CpG ostrovy

Dinukleotidy CpG jsou v DNA eukaryot přítomny s velmi nízkou frekvencí, nižší, než by odpovídalo náhodné distribuci dinukleotidů. Cytosin vyskytující se v těchto vzácných dinukleotidech navíc často bývá methylován. Následnou deaminací může docházet k přeměně C na T, což je substituční mutace, která nemůže být opravena. Existují však oblasti genomu, kde se dinukleotidy CpG naopak vyskytují s

vysokou frekvencí (tedy blíží se matematicky očekávané frekvenci tohoto dinukleotidu). Tyto klastry dinukleotidů CpG se nazývají CpG ostrovy. CpG ostrovy, dosahující délky asi 0,2-5 kilobází (kb), se nejčastěji nacházejí v promotorových oblastech udržovacích (tzv. housekeeping, tj. ve všech tkáních exprimovaných) genů, ale lze je nalézt i v promotorech genů s tkáňově specifickou expresí. CpG ostrovy navíc zůstávají většinou nemetylované. Stupeň methylace takového promotoru souvisí s výší exprese genu a úplná methylace CpG ostrovu vede k jeho umlčení. Chromatin v oblastech CpG ostrovů také vykazuje rozvolněnější strukturu, neboť zde chybí histon H1 – protein propojující jednotlivé nukleozómy a umožňující tvorbu kondenzovanějšího chromatinu. Byly také nalezeny oblasti postrádající nukleozómy (NFR), které byly mapovány právě do oblastí CpG ostrovů (Choi, 2010). CpG ostrovy jsou navíc cílem pro vazbu některých regulačních proteinů. Nejznámější z nich je nejspíše specifický protein 1 (Sp1). (Baumann *et al.*, 2010; Cross and Bird, 1995)

CpG ostrovy se staly jedním s kriterií při hodnocení integračních preferencí, neboť právě rozvolněná forma chromatinu v blízkosti CpG ostrovů by mohla usnadňovat přístup PIC k tDNA a zároveň by dovozovala stabilní expresi proviru. Výrazná preference pro integraci do blízkosti CpG ostrovů byla však prokázána pouze u MLV (Mitchell *et al.*, 2004; Wu *et al.*, 2003). Podobné chování bylo zaznamenáno i u spumavirů (Trobridge *et al.*, 2006), jehož preference pro CpG ostrovy byla daleko slabší, než v případě MLV. Studie autorů Nowrouzi *et al.* (2006) a Santoni *et al.* (2010) však tento fakt (tedy preferenci FV pro CpG ostrovy) nepotvrdily. Ostatní retroviry vykazují pouze velmi slabé, pokud vůbec nějaké, preference pro tyto oblasti (více v kapitole 3.5 Oblasti transkripčních startů).

3.4. Histonové modifikace

Histonové modifikace patří do skupiny tzv. epigenetických značek. Jsou to tedy modifikace, které jsou schopny měnit fenotyp bez přímé změny v genetické informaci (tedy v pořadí nukleotidů). Bylo prokázáno, že konkrétní histonové modifikace jsou charakteristické pro určité genomové úseky, jako jsou aktivní promotory a umlčené geny či mohou rozlišovat mezi heterochromatinem a euchromatinem. Nejdůležitějšími histonovými modifikacemi vztahujícími se k expresi jsou methylace a acetylce aminokyseliny lysinu v N-koncové doméně histonů. Vysoký výskyt acetylce histonů je značkou aktivních promotorů, CpG ostrovů a regulačních segmentů, jako jsou enhancery a inzulátory (Roh *et al.*, 2005). Dají se tedy považovat za jakési značky aktivních genů a tedy i „aktivního“ euchromatinu, i když ne každá aktivní oblast genomu se vyznačuje vysokou acetylací histonů. Oproti tomu různé methylace různých histonů mohou mít různý vliv na genomové oblasti – jsou charakteristické jak pro oblasti aktivních promotorů, tak pro oblasti umlčených genů a heterochromatinu (Barski *et al.*, 2007). Mají tedy ambivalentní charakter.

Díky rostoucímu množství dat v genomických databázích a množství identifikovaných integračních míst retrovirů, je možno hledat preference i mezi histonovými modifikacemi. Byly nalezeny preferované histonové modifikace pro integraci HIV-1 a MLV, kterými jsou hlavně acetylce

histonů H3 a H4 (H3ac, H4ac) a mono-, di- a trimethylace histonu H3 na lysinu 4 (H3K4me1/2/3) pro HIV-1 i MLV. MLV navíc preferuje i místa bohatá na acetylaci lysinu 9 histonu H3 (H3K9ac) a metylaci histonu H2A.Z (H2A.Zme) (Santoni *et al.*, 2010; Wang *et al.*, 2007). Všechny tyto modifikace jsou spojovány s oblastmi aktivních genů. Zajímavý je pohled na H3K4me3. V práci autorů Wang *et al.* (2007) je na tuto modifikaci poukazováno, jako na jednu z preferenční epigenetických značek pro integraci HIV-1, zatímco práce autorů Santoni *et al.* (2010) poukazuje spíše na vliv H3K4me1, přičemž H3K4me3 v této studii vliv nemá. Vysvětlení je možno nalézt v rozložení jednotlivých značek, kdy stupně methylace H3K4 vzrůstají směrem k promotorové oblasti a H3K4me3 je tedy zastoupeno nejvíce přímo v oblastech promotorů (Barski *et al.*, 2007). Vzhledem k preferencím HIV-1 pro transkripčně aktivní oblasti (viz kapitola 3.6. Transkripční jednotky) závisí interpretace asociace methylovaných forem H3K4 s místy integrace HIV-1 na velikosti prohledávané oblasti (neboli okna) v okolí integračního místa. Problematika velikosti oken pro hledání asociací genomických charakteristik s integračními místy je širší a bylo na ni upozorněno již ve studii autorů Berry *et al.* (2006). V případě HIV-1 tedy při zvolení většího prohledávacího okna bude mít vliv H3K4me3, zatímco při zmenšení okna tento vliv poklesne a bude převažovat vliv H3K4me1 (Santoni *et al.*, 2010), který je daleko více zastoupen v oblastech transkripčně aktivních jednotek (Barski *et al.*, 2007). Naproti preferenčním histonovým modifikacím existují i modifikace označující velmi nepreferované oblasti pro integraci retroviru. Takovou značkou je pro HIV-1 trimethylace lysinu 27 histonu H3 (H3K27me3) (Wang *et al.*, 2007), která je známkou přítomnosti heterochromatinu a umlčených genů (Barski *et al.*, 2007).

Je otázkou, do jaké míry histonové modifikace samy o sobě ovlivňují cílení integrace retrovirů. Dosud totiž nebyly publikovány žádné důkazy o interakci integrázy ani žádné jiné složky PIC přímo s modifikovanými aminokyselinami histonů. PIC však může interagovat s modifikačními faktory či proteiny, které mají afinitu k modifikovaným histonům. Zatím se ale zdá, že všechny výše uvedené histonové modifikace podněčují a udržují ty podmínky, které tvoří ideální prostředí pro integraci retroviru (Wang *et al.*, 2007). Ukázalo se však, že histonové modifikace mohou poskytovat nástroj pro predikci integračních míst retrovirů, který je nezávislý na dalších genomických charakteristikách. Takovouto „superznačku“ vytvořili pro MLV Santoni *et al.* (2010), kteří kombinací jednotlivých značek H3K4me3, H3K4me1 a H3K9ac dokázali v okně 2 kb predikovat 75% ze známých integračních míst MLV.

3.5. Oblasti transkripčních startů

Jako oblasti transkripčních startů (TSS) jsou označovány oblasti v okolí místa začátku transkripce (tedy okolo +1 nukleotidu). To zahrnuje i oblasti promotorů. TSS jsou oblastmi s nízkým výskytem nukleozómů – tzv. oblasti bez nukleozómů (NFR). Ukazuje se však, že NFR jsou spíše oblasti obsahující velké množství nestabilních nukleozómů skládajících se z histonových variant H2A.Z a H3.3 (Jin *et al.*, 2009). Některé TSS jsou charakteristické přítomností CpG ostrovů (asi 40%

promotorů lidského genomu; Kim *et al.*, 2005) či již dříve zmiňovanými histonovými modifikacemi jako jsou H3ac, H4ac a H3K4me3.

TSS, jakožto genomová charakteristika pro hodnocení integračních preferencí, se dostaly do popředí díky schopnosti retrovirů ovlivnit expresi genu právě integrací do TSS. Taková integrace nese nebezpečí aktivace onkogenů, která je hlavním zdrojem genotoxicity retrovirů a retrovirových vektorů. Proto je hlavně míra integrace retrovirových vektorů do TSS také měřítkem pro bezpečnost jejich použití pro genovou terapii. Rozvolněná forma chromatinu v TSS by mohla být ideálním místem pro přístup PIC k tDNA. Výrazná preference pro TSS byla pozorována hlavě u gammaretrovirů, z nichž nejvýraznější v pohledu na integrace do TSS je MLV (Wu *et al.*, 2003). Tento fakt komplikuje používání vektorů odvozených od MLV v genové terapii.

3.6. Transkripční jednotky

Transkripční jednotky (TU) jsou nejspíše nejsledovanější genomovou charakteristikou při určování integračních preferencí retrovirů. Integrace do TU aktivního genu může totiž zajistit stabilní expresi proviru a jeho ochranu proti umlčení. Na druhou stranu není zcela žádoucí integrace do TU pro účely genové terapie, neboť taková integrace může narušit normální expresi genu.

Do aktivních TU a obecně do genově bohatých oblastí se integruje hlavně HIV-1 (Elleder *et al.*, 2002; Mitchell *et al.*, 2004; Schröder *et al.*, 2002; Wu *et al.*, 2003). HIV-1 navíc nepreferuje žádnou oblast TU a integruje se stejně pravděpodobně po celé délce TU. Stejná preference byla pozorována u SIV (Hematti *et al.*, 2004), což naznačuje, že integrace do TU může být společnou vlastností lentivirů. Slabá preference pro integraci do aktivních TU byla prokázána u dalších retrovirů, jako jsou MLV, ASLV (Mitchell *et al.*, 2004; Narezkina *et al.*, 2004), lidský leukemický virus typu 1 (HTLV-1) (Derse *et al.*, 2007; Ozawa *et al.*, 2004) nebo bovinní leukemický virus (BLV) (Murakami *et al.*, 2011). Integrace MLV jsou převážně nalézány v 5' částech TU, tedy poblíž TSS. Ostatní retroviry většinou prokazují velmi slabé, avšak statisticky významné preference pro TU. Preferenci pro aktivní TU také překvapivě vykazují *de novo* integrace lidských endogenních retrovirů (HERV), což poukazuje na vliv selekce integrovaných HERV a zvyšuje význam této selekce na úkor cílené integrace HERV mimo TU (Brady *et al.*, 2009).

Zdá se, že vliv na integraci může mít i stupeň exprese TU. Weidhaas *et al.* (2000) ve své studii poukázali na negativní vliv vysoké exprese genu pro integraci retroviru, což bylo potvrzeno i některými „většími“ studiemi (Narezkina *et al.*, 2004). Toto se zdá být obecným pravidlem integrace do aktivních TU. Byla však pozorována i integrace do transkripčně aktivnějších genů (Schröder *et al.*, 2002). Negativní vliv vysoké transkripční aktivity TU by mohl být důsledkem aktivity transkripčního komplexu, který by mohl bránit vazbě PIC na tDNA.

3.7. Proteinové interakce

Jak již bylo zmíněno v úvodních kapitolách, v PIC jsou k nalezení i proteiny buněčného původu, které mohou interagovat buďto s virovými proteiny nebo s cDNA. Interakce virové IN a buněčných proteinů se zdají být nejvýznamnějším faktorem v cílení virové integrace, a tak jsou tyto interakce hojně vyhledávány. Takovéto cílení integrace bylo již popsáno u kvasinkových retrotraspozónů, jako je Tf1 (Majumdar *et al.*, 2011). Nebyly zatím nalezeny žádné proteiny, které by svou vazbou na virovou DNA ovlivňovaly cílení integrace.

Nejprozkoumanější je interakce lentivirové IN a LEDGF/p75, což je buněčný transkripční koaktivátor vyskytující se v infikovaných buňkách v komplexu s virovou IN (Cherepanov *et al.*, 2003), který by mohl směřovat integraci do míst aktivních TU. Maertens *et al.* (2003) navíc prokázali, že LEDGF/p75 hraje roli v transportu HIV-1 IN do jádra buněk. Ciuffi *et al.* (2005) poukázali na fakt, že vypnutí exprese LEDGF/p75 v infikovaných buňkách sice kvantitativně snižuje četnost integrací do aktivních TU (jde většinou o geny, v jejichž regulaci LEDGF/p75 hraje roli), avšak HIV-1 si preference pro aktivní TU stále zachovává. Avšak další studie (shrnuté v práci autorů Llano *et al.*, 2006) ukázaly, že úplné vyřazení exprese LEDGF/p75 má téměř fatální dopad na integraci lentivirů (ve smyslu velmi snížené schopnosti integrace) a výsledky dříve zmíněné studie Ciuffi *et al.* (2005) mohou být způsobeny zbytkovou koncentrací LEDGF/p75, která je schopná podporovat integraci lentivirů. I v nepřítomnosti LEDGF/p75 stále existuje malé množství případů integrace, které mohou vykazovat na LEDGF/p75 nezávislou vazbu PIC na chromatin (Zheng *et al.*, 2010). Všechny tyto funkce LEDGF/p75 byly potvrzeny studií Gijssbers *et al.* (2010). Autoři navíc prokázali, že modifikací tohoto proteinu je možno změnit integrační preference lentivirů.

Kromě LEDGF/p75 byly pro HIV-1 identifikovány i další buněčné proteiny s možnou rolí v cílení integrace, jako je např. chromatinový remodelační komplex SWI/SNF umožňující rozvolnění chromatinu (Lesbats *et al.*, 2011).

3.8. Chromatinová struktura

Stav chromatinu je důležitým prvkem v integračních preferencích retrovirů, neboť jde o ukazatel přístupnosti potenciálního cíle integrace. Ačkoliv je očekávané, že retroviry se budou preferenčně integrovat pouze do rozvolněného chromatinu, nemusí tomu tak být. Taganov *et al.* (2004) na *in vitro* integracích dokonce poukázali na možné rozdíly mezi retroviry ve vztahu ke stavu chromatinu. Konkrétně se jedná ASLV, u něž se projevovala zvýšená míra integrace do kondenzovanějšího chromatinu (např. vazbou H1 histonu), zatímco u HIV-1 byl účinek chromatinové kondenzace opačný a integrace tak byla účinnější v rozvolněném chromatinu.

Vzhledem k chromatinové struktuře byla definována „nová“ genomová charakteristika mající vliv na integrační preference. Jedná se o oblasti chromatinu, kde je DNA v kontaktu s jadernou matrix (MAR). MAR tak představují exponovanou část chromatinu dobře přístupnou pro buněčné i virové faktory. Integrace do blízkosti MAR (do vzdálenosti 2 kb) je preferována myšími leukemickými viry,

HTLV-1 i HIV-1 (Johnson and Levy, 2005). Integrace do takových míst může mít pozitivní vliv na expresi proviru, ale může mít také dopad na expresi blízkých genů a případné aktivace onkogenů.

3.9. Buněčný typ

Většina výše uvedených analýz integračních preferencí byla prováděna na lidských buněčných kulturách (např. HeLa, SupT1 atd.), přičemž se integrační preference v jednotlivých buněčných typech nelišily. Studie autorů Barr *et al.* (2005) navíc prokázala, že rozdíly v integračních preferencích pro HIV-1 a ASLV nejsou ani mezi buňkami různých živočišných druhů (zde lidské a kuřecí buňky), na což poukázala i studie Faschinger *et al.* (2008) v případě HTLV-1 u myších a lidských buněk. V případě HIV-1 zůstaly jeho integrační preference shodné v dělicích i nedělicích se buňkách (Ciuffi *et al.*, 2006b). Lze tedy očekávat, že faktory ovlivňující cílení integrace budou společné pro více buněčných typů i živočišných druhů.

3.10. Souhrn integračních preferencí retrovirů

O tom, že integrace retrovirů není náhodná, je přinášeno stále více důkazů. Retroviry disponují různými preferencemi, což umožňuje jejich rozdělení do skupin s rozlišnými preferenčními charakteristikami. Toto rozdělení se shoduje s rozdělením podle sekvencí IN (Derse *et al.*, 2007). Nejvíce dat charakterizujících integrační preference, bylo získáno hlavně pro HIV-1, MLV a ASLV.

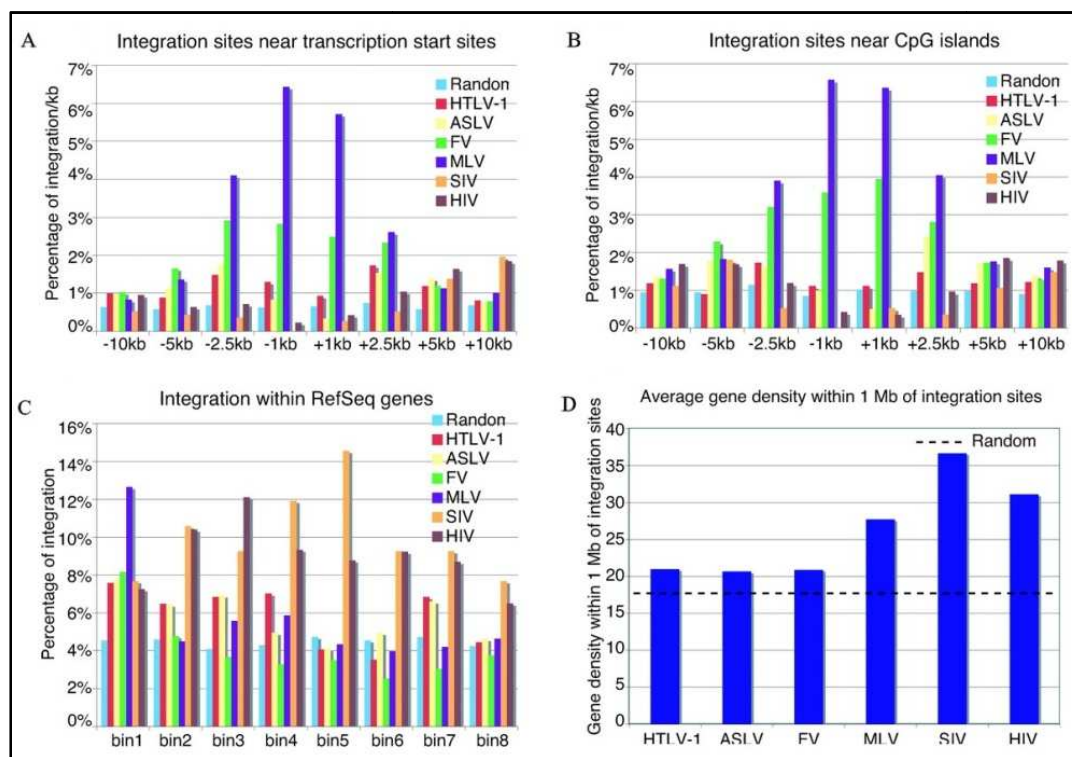
Lentiviry preferují integraci do aktivních TU a do genově bohatých oblastí genomu. Předpokládá se, že z velké části je za toto jejich chování zodpovědný LEDGF/p75, který hraje důležitou roli v jejich integraci. Přestože byla pro HIV-1 dokázána preferenční integrace do velkého žlábků nukleozomální DNA, kondenzovanější forma chromatinu integraci HIV-1 nepodporuje. Právě rozvolňování chromatinu je funkcí chromatin remodelačního komplexu SWI/SNF, který může být jedním z interakčních partnerů IN HIV-1. Integrační místa HIV-1 jsou navíc bohatá na methylace histonů charakteristické pro aktivní chromatin, a to především na H3K4me1, což je značka pro vzdálenější okolí TSS. To nepřímo podporuje preference HIV-1 pro TU.

MLV, zástupce gammaretrovirů, vykazuje preferenční integraci do oblastí TSS a do blízkosti CpG ostrovů. Tyto oblasti představují rozvolněný chromatin obsahující nejspíše malé množství velmi nestabilních nukleozómů. Do těchto oblastí se navíc váže množství buněčných faktorů, které by teoreticky mohly být součástí PIC či s ním interagovat a cílit tak integraci MLV do oblastí CpG ostrovů a TSS. Epigenetické značky asociované s integračními místy MLV jsou charakteristické právě pro oblasti TSS. Vytvořená epigenetická „superznačka“ zahrnující H3K4me3, H3K4me1 a H3K9ac je pak zcela nezávislou charakteristikou integrace MLV.

ASLV, představující alfaretroviry, disponují daleko náhodnější integrační strategií, než bylo popsáno pro předchozí případy, a to jak v lidských, tak i ptačích buňkách. I když u nich byla prokázána preferenční integrace do TU s velmi nízkou preferencí pro TSS, je jejich afinita k aktivním TU daleko nižší, než je tomu v případě HIV-1 či MLV. Tato slabá asociace ASLV s TU je zjevná i na

slabé preferenci ASLV pro místa obsahující H3K4me3 a H3K4me1, přičemž preference pro H3K4me1 je o něco silnější (Santoni *et al.*, 2010). RSV navíc, oproti HIV-1, prokazuje *in vitro* zvýšenou míru integrace do kondenzovanějších oblastí chromatinu.

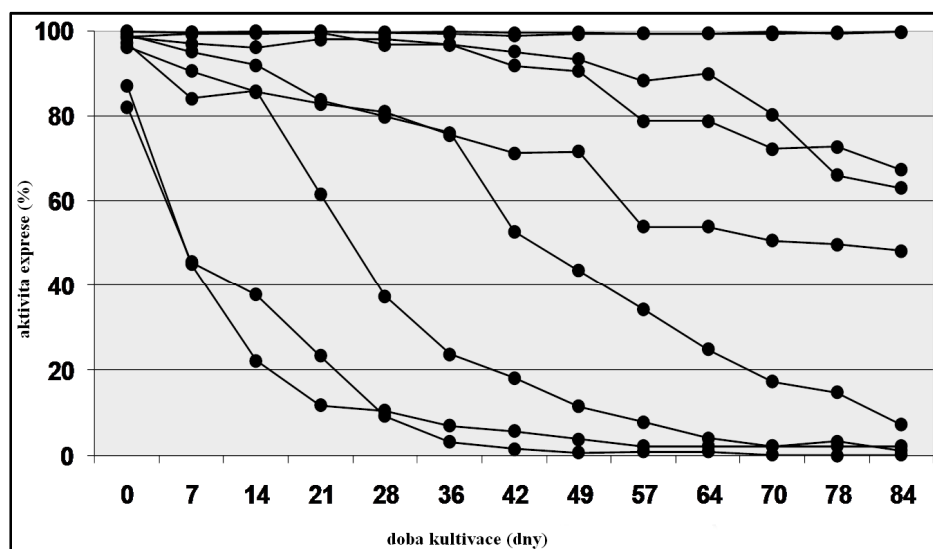
Pro většinu ostatních retrovirů nebylo dosud získáno takové množství dat, aby mohla být provedena podrobná analýza integračních preferencí. Každý retrovirus, u kterého byla identifikována integrační místa, je však možno přiřadit k některému z dříve zmíněných případů. HTLV-1 se svou slabou preferencí pro aktivní TU podobá spíše ASLV. FV nevykazuje preferenci pro TU, ale byla identifikována (avšak nepotvrzena) preference pro oblasti okolo CpG ostrovů, čímž by se FV podobal MLV. Nejnáhodnější integraci vykazuje myší virus způsobující nádory prsu (MMTV), u nějž nebyla pozorována preference pro TU ani pro TSS a CpG ostrovy (Faschinger *et al.*, 2008). Svými náhodnými „preferencemi“ se tak blíží ASLV a HTLV-1.



Obr.2. Integrační preference různých retrovirů. Znázornění různých integračních preferencí pomocí sloupcových grafů pro (A) TSS, (B) CpG ostrovy, (C) geny, (D) na geny bohaté oblasti. Na y ose jsou znázorněno množství integrací v % (A, B, C), případně množství genů v okolí 1 Mb od místa integrace (D). (A) a (B) ukazují množství integrací v okolí TSS, případně CpG ostrovů. (C) Znázornění integrací podél genu, přičemž bin1 odpovídá začátku genu a bin8 jeho konci. (D) Preference integrací do oblastí bohatých na geny. Random – četnost odpovídající náhodné integraci. Převzato a upraveno z Derse *et al.* (2007).

4. Exprese a umlčování proviru

Expresie integrovaného proviru je řízena z LTR, přesněji z 5' LTR. U3 oblast LTR obsahuje ve své sekvenci jak virový promotor, tak i enhancer. Do oblastí LTR se váže množství buněčných faktorů, které mohou hlavně pozitivně, ale i negativně ovlivňovat expresi proviru. Provirus je však součástí hostitelské DNA a podléhá chromatinově regulačním vlivům stejně tak jako každý jiný lokus genomové DNA. LTR tak není zcela dominantní transkripčně regulační oblastí pro provirus (Hoeben *et al.*, 1991). O tom, zda bude z provirového promotoru zahájena transkripce, rozhodují podmínky v hostitelské buňce, zvláště pak vlastnosti oblastí, kam se provirus integroval. Takové oblasti genomové DNA, ve kterých je umožněno zahájit a udržovat provirovou expresi, se nazývají permisivní. Vzhledem k pozorovanému umlčení exprese proviru není každé místo integrace permisivní a výběr integračního místa tak může ovlivnit další osud proviru (Obrázek 3).



Obr. 3. Stabilita exprese proviru. Body znázorňují výši exprese proviru v daném čase. Rozdíly mezi jednotlivými proviry lze přičíst rozdílným podmínkám v různých místech integrace. Převzato a upraveno z Šenigl *et al.* (2008).

Umlčování provirové exprese je pozorovatelné jak v buňkách původního, tedy permisivního, hostitele, tak, a to především, při tzv. heterotransmisi retrovirů do buněk nepermisivního hostitele. Schopnost stabilní exprese je ovšem různá i pro různé buněčné typy stejného hostitele. Takto bylo v minulosti pozorováno umlčování exprese ptačích (RSV, ASLV) či myších (MLV) retrovirů v lidských buňkách. Ptačí retroviry přitom jsou umlčováním postiženy nejvíce. Náchylnost k umlčování je nejspíše určena citlivostí retroviru k umlčovacími mechanismům hostitele. Tato citlivost může být zakódována v sekvencích LTR retrovirů a může poukazovat na společný vývoj retrovirů a jejich hostitelů. HIV-1 je v lidských buňkách umlčován také, avšak s daleko nižší účinností, než právě ptačí retroviry. Latentní stav HIV-1 v lidských buňkách je i klinicky velmi zajímavým úkazem, neboť je zodpovědný za perzistenci viru a vytváření virového rezervoáru u pacientů podstupujících léčbu proti AIDS.

Způsob umlčování retrovirů, charakteristiky permisivních a nepermisivních míst či vlastní vliv integračního místa na míru exprese jsou stále předměty výzkumu. Je však jisté, že na umlčování proviru mají vliv epigenetické modifikace jako je methylace DNA nebo histonové modifikace.

4.1. Hlavní faktory v umlčení provirové exprese

Hlavním znakem umlčeného proviru je methylace DNA (Hoeben *et al.*, 1991; Challita and Kohn, 1994). Na cytosin navázaná methylová skupina se objevuje především v CpG sekvencích LTR v oblastech okolo TSS a transkripčně regulačních sekvencí. Vektory nesoucí *in vitro* methylované LTR jsou po transfekci do cílových buněk rychle umlčovány (Bednarik *et al.*, 1990). Tatáž modifikace navíc podmiňuje daleko rychlejší a účinnější umlčování, jsou-li vektory přeneseny do nepermisivních buněk (Hejnar *et al.*, 1999). Methylace může postihovat selektivně pouze 5'LTR, jak bylo prokázáno u latentní infekce HTLV-1 (Koiwa *et al.*, 2002). Role methylace DNA v udržování represe provirové exprese, byla potvrzena reaktivací umlčených klonů pomocí 5'-azacytidinu (5'-azaC), což je inhibitor DNA methyltransferáz (DNMT).

Hlavní komponentou udržující metylaci proviru je dle současných výsledků *de novo* DNA methyltransferáza 3a (DNMT3a) (Poleshko *et al.*, 2010). DNMT3a je jedinou methyltransferázou, jejíž vyřazení pomocí RNA interference má stejný vliv na reaktivaci jako použití 5'-azaC, což je hlavní důkaz pro její roli při umlčování provirové exprese. Methylová skupina bývá vystavena ve velkém žlábků DNA, kam se váže množství proteinů. Může tak blokovat vazbu důležitých faktorů. Jsou ale i enzymy, které se přímo váží na methylované cytosiny (methyl-CpG binding domain, MBD; methyl CpG binding protein, MeC-P). Vazba těchto proteinů na methylované CpG a jejich další interakce mohou hrát důležitou roli pevných pojistek proti reaktivaci exprese proviru (Blažková *et al.*, 2009; Kauder *et al.*, 2009).

Další mechanismus zabraňující expresi proviru je deacetylace histonů, jak dokazuje reaktivace transkripce umlčených provirů pomocí trichostatinu A (TSA), inhibitoru histon deacetyláz (Katz *et al.*, 2007). Deacetylované N konce histonů jsou přístupné dalším modifikačním enzymům, jako jsou nejrůznější lysinové histon methyltransferázy (HMT), které mohou vytvářet podmínky pro inaktivní chromatinové prostředí. Hlavní histon deacetylázou je zde histon deacetyláza 1 (HDAC1) (Poleshko *et al.*, 2008). HDAC1 se k proviru neváže sama, ale je k němu přitahována pomocí DNA vazebných proteinů, jako jsou CCAAT vazebný faktor 1 (CBF-1) (Tyagi and Karn, 2007), transkripční faktor jin jang 1 (YY1) (He and Margolis, 2002) a jaderný faktor NF-κB (Williams *et al.*, 2006) v případě HIV-1 či protein nesoucí doménu asociovanou se smrtí (DAXX) (Greger *et al.*, 2005) interagující s IN a cDNA ASV. Umlčení spočívající v činnosti HDAC1 má však slabý charakter a takto umlčená exprese může být jednoduše obnovena (Blažková *et al.*, 2009; Pion *et al.*, 2003).

V umlčování exprese proviru je důležitá také methylace histonů a byly i identifikovány HMT sehrávající své role při umlčování provirů. Hlavními histonovými modifikacemi s represivním účinkem na expresi jsou zde H3K9me3 a H3K27me3. Bylo identifikováno hned několik HMT, které

mají co dočinění s umlčováním exprese. Jsou jimi homolog 1 supresoru variací 3-9 (Suv39H1), SET doménu obsahující protein 1 (SETDB1) a G9a spolu s G9a podobným proteinem (GLP; v G9a/GLP komplexu). Kromě své role v methylaci histonů, interagují tyto HMT s dalšími, v represi exprese vystupujícími faktory. Represivní účinek těchto modifikací (hlavně H3K9me) je dán také vazbou heterochromatinových proteinů, jako jsou CBX-5 a HP1, zde hlavně HP1 γ . Heterochromatinové proteiny jsou schopny rovněž interagovat s množstvím dalších proteinů, jimiž jsou i DNMT.

4.2. Interakční síť represivních epigenetických modifikací

Jak již bylo zmíněno v předešlé kapitole, umístění epigenetické modifikace přitahuje další proteiny a faktory působící na udržení a prohlubování umlčení exprese proviru. Takové represivní faktory mohou dokonce být přitahovány i samotnými modifikačními enzymy. Je tedy možné, že objevení se jedné represivní epigenetické značky způsobí (pokud to podmínky dovolí) spuštění modifikačních drah vyúsťujících v přítomnost dalších epigenetických značek na daném místě. Takováto komunikace je také důležitá v udržení umlčujících podmínek chromatinu. Obecné interakce epigenetických značek v umlčování exprese jsou uvedeny v souhrnných člancích autorů Fischle (2008) a Cedar a Bergman (2009).

Možnost ovlivnit funkci modifikačních enzymů se naskytuje již při jejich nasedání na histony, kdy různé histonové modifikace přímo zabraňují nasednutí některých modifikačních komplexů. To je případ H3K4me (značky aktivního chromatinu), jejíž přítomnost zabraňuje nasednutí HMT vytvářejících H3K9me (značky umlčeného chromatinu). Na místech výskytu H3K9me tak zároveň nelze nalézt H3K4me. Jiná represivní methylační značka H3K20me zabraňuje vytvoření H3K4ac. Acetylace histonů jsou obecně značkami pro aktivní chromatin a jako takové také brání ve vytváření represivních methylových značek. O odstranění histonových acetylací se v případě provirů stará nejspíše HDAC1 (Poleshko *et al.*, 2008). Ta disponuje širokým spektrem interakčních partnerů, mezi které patří např. MeCP-2, MBD proteiny, SETDB1 nebo některé DNA vazebné proteiny přitahující HDAC1 k proviru, jako jsou YY1, NF- κ B nebo DAXX. HDAC1, vyskytující se často ve větších nukleozómy remodelujících komplexech, tak mohou kromě specifických míst pro DNA vazebné proteiny přitahovat i místa methylace histonů i DNA. Velké interakční možnosti mají hlavně proteiny rozeznávající methylované CpG (MeCP-2, MBD proteiny), které tak mohou působit jako jakési adaptéry. Interagují jak s histon deacetylázovými komplexy (Kauder *et al.*, 2009), tak s SETDB1 či podjednotkou A chromatin sestavujícího faktoru 1 (CHAF1A). Právě interakce s CHAF1A může hrát důležitou roli udržování histonových modifikací v místech methylované DNA po její replikaci. Jako adaptér ale působí i G9a, který i bez své katalytické aktivity dokáže podněcovat DNA methylaci retrotranspozónů (Dong *et al.*, 2008; Tachibana *et al.*, 2008). Výsledky autorů Poleshko *et al.* (2010) ukazují, že centrálními hráči v umlčování exprese (nejen) provirů ASV mohou být SETDB1, MBD proteiny, CHAF1A, DNMT3a a HDAC1, přičemž právě trojice CHAF1A, MBD1 a SETDB1 je nejspíše zodpovědná za udržování umlčeného chromatinu v místech methylace DNA.

4.3. Iniciační umlčování proviru

Vzhledem k součinnosti tří epigenetických modifikací při umlčování exprese proviru – tedy methylace DNA, methylace a deacetylaci histonů – se nabízí otázka, která z nich je prvním krokem k umlčení proviru a zda jsou za tento postintegrační blok zodpovědné spíše histonové modifikace nebo methylace DNA.

Iniciační umlčovací modifikací by mohla být methylace DNA. DNMT jsou totiž schopny *de novo* methylovat CpG ostrovy (např. s histony bez H3K4me, jak je tomu při methylaci CpG v rané embryogenezi, viz Cedar and Bergman, 2009). Methylování CpG pak může skrze MBD a další 5'-methylcytosin vazebné proteiny spustit kaskádu modifikačních reakcí vedoucí k umlčení transkripce proviru. Možnost reaktivace exprese proviru pomocí TSA však poukazuje na funkci deacetylaci histonů jako možného udržovatele, ale i iniciátora umlčování, které je na CpG methylaci nezávislé (Katz *et al.*, 2007). Deacetylaci histonů (HDAC1) umožňuje nasedání takových faktorů, jakými jsou HMT katalyzující methylaci histonů, na které se tak mohou vázat další, represivní podporující faktory, jako je HP1 γ . Takovýto postup umlčování by odpovídal jednoduše reaktivovatelnému stavu umlčeného proviru, shodně s modelem autorů Blažková *et al.* (2009) navrženém pro HIV-1 (Obrázek 4). DNA methylování a ustanovení „uzamknutého“ stavu proviru by mohly následovat díky „nalákání“ DNMT k již slabě umlčenému proviru a upevnit tak tento stav. Není však vyloučeno, že iniciačním krokem v umlčení provirové exprese by mohla být třeba i methylování histonů.

4.4. Vliv integračního místa na expresi proviru

Proviróvá DNA je jako epigenetická *tabula rasa* vkládána do genomu hostitele, tedy do prostředí epigeneticky již poznamenaného, které však není statické, ale může vykazovat jistou proměnlivost v závislosti na širším kontextu okolního prostředí a na přítomnosti či nepřítomnosti různých faktorů. A právě to, do jakého prostředí je provirus integračním aparátem vložen, může mít vliv na jeho epigenetické obsazení a následně na jeho expresi.

Je možné povšimnout si jisté souvislosti mezi integračními preferencemi retrovirů a jejich schopností udržovat stabilní expresi. Čím vyšší je preference retrovirů pro TU a transkripčně aktivní oblasti, tím méně podléhají umlčování. Zatímco lentiviry významně preferují TU často vykazují velmi stabilní expresi, jiné retroviry jako ASLV, ale i MLV, jsou účinně umlčovány. Nejvýraznější je umlčování ptačích retrovirů v savčích buňkách. Ptačí retroviry přitom vykazují nejnáhodnější charakter integrace a nejslabší preference pro aktivní TU v porovnání s MLV a HIV-1. V souladu s těmito tvrzeními bylo navíc autory Plachý *et al.* (2010) prokázáno, že stabilně se exprimující se proviry RSV jsou nalézány v aktivních genech exprimovaných napříč nejrůznějšími tkáněmi. Tzv. housekeeping geny tedy nejspíše budou podporovat stabilní expresi účinněji než geny tkáňově

specifické. Podrobnější charakterizace míst podporujících aktivní expresi proviru však dosud nebyla publikována.

Otázkou zůstává, jak dokáže okolí proviru ovlivnit jeho expresi. Svou roli mohou hrát buněčné regulační sekvence a na ně vázané buněčné faktory (Yang and Dudley, 1992), ale také již tolikrát zmiňované epigenetické značky, respektive efektory vážící se na ně. Je-li v místě integrace prostředí podporující aktivní transkripci (tzn. H3K4me, H3ac, H4ac, atd), je pravděpodobné, že toto prostředí bude podporovat rozšíření aktivních chromatinových a histonových znaků i na provirus a zároveň zabraňovat umlčování. Přítomnost represivních znaků (CpG methylace, deacetylace histonů, H3K9me, atd.) bude naopak podporovat umlčení proviru. Na rozdíl od aktivních provirů nejsou ty umlčené vázány na konkrétní místa, jako je např. heterochromatin. Vzhledem k represivní funkci HP1 γ , který působí i v euchromatinu, může docházet k umlčení proviru i zde. Umlčené proviry se tak mohou vyskytovat na místech, která se navzájem svou charakteristikou liší a umlčení tak není místně specifické (Poleshko *et al.*, 2008). Proviry tedy bývají umlčeny mechanizmy specifickými pro danou oblast a společným rysem těchto míst může být neschopnost takovému umlčení zabránit.

Ovšem i v případě integrace proviru do místa s hojnou CpG methylací může být takový provirus po určité době exprimován, jak to bylo popsáno pro RSV (Hejnar *et al.*, 2003). Po integraci dochází k demethylaci DNA v okolí integračního místa a provirus je aktivní. Tento stav však není stabilní. Později dochází k opětné methylaci integračního místa spolu s provirem, čímž je jeho exprese umlčena. Takový provirus pak není možno reaktivovat pomocí 5'-azaC ani TSA. Ačkoliv není jasné, zda jde o obecné chování po integraci do oblastí bohatých na CpG methylace, neboť taková integrace byla zaznamenána pouze u ptačích retrovirů (Taganov *et al.*, 2004), jde o pozorování podporující model, kde je umlčování řízeno podmínkami panujícími v oblasti integračního místa. Místa s bohatou DNA methylací tedy nejsou výhodná pro dlouhodobou a stabilní expresi proviru.

4.5. Umlčování jako aktivní obrana hostitele

K umlčování proviru nemusí docházet působením okolního chromatinu rozšířením jeho represivních znaků na provirus. Může jít i o jakousi vnitřní imunitní odpověď buňky cílenou proti cizorodým segmentům DNA a tím pádem o formu cíleného zablokování exprese proviru. Takové umlčovací kroky bývají mířeny např. proti přímým repetitivním (Selker, 1999), kterými LTR jsou.

Svým způsobem může být cílená i methylace provirových CpG ostrovů. CpG dinukleotidy bývají v genomech vyšších organismů redukovány a buněčné CpG ostrovy jsou často proti methylaci chráněny. Přítomnost nechráněných CpG ostrovů, kterými retroviry bezesporu disponují, může přitahovat DNMT způsobující jejich methylaci. Předpokládá se, že zásadní, z hlediska DNA methylace, je methylace LTR. Pozorování latentní infekce HTLV-1 ukázalo, že methylace 5'LTR sice odpovídá umlčenému proviru, ale tato methylace se k 5'LTR rozšiřuje z vnitřních provirových sekvencí (Taniguchi *et al.*, 2005). Rozeznány a methylovány jsou tedy nejtíže sekvence z oblastí kódujících virové proteiny. Takové proviry, které jsou methylovány pouze na vnitřních sekvencích,

ale ne na LTR, již vykazují latentní infekci. DNA methylace CpG ostrovů mimo LTR hraje roli i v umlčování HIV-1. Jde o CpG ostrov nacházející se v nekódující oblasti proviru mezi 5'LTR a Gag genem (NCR CpG), který se navíc zdá být mezi různými typy HIV dobře konzervován (Chávez *et al.*, 2011). Kauder *et al.* (2009) navíc popsali, že methylovaný NCR CpG ostrov spolu s methylovaným CpG ostrovem nacházejícím se v 5'LTR jsou přítomné v umlčených provirech.

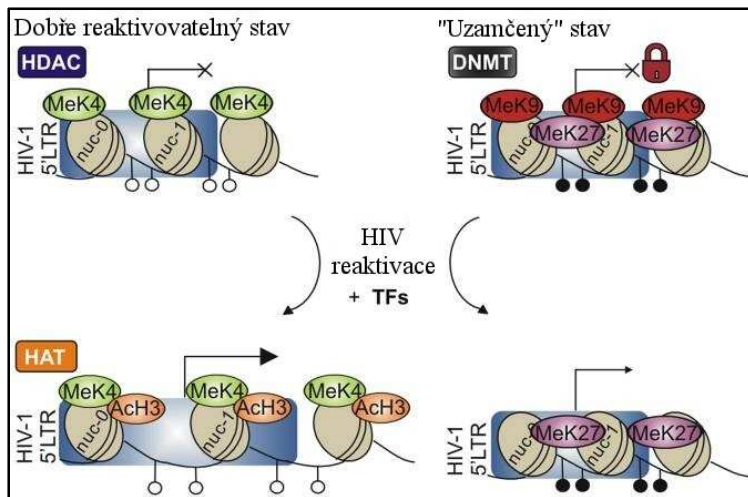
Greger *et al.* (2005) ukázali, že již na PIC ASV se váže DAXX protein, který interaguje s HDAC1 či HDAC2. Funkci DAXX a HDAC1 v umlčování některých provirů poté potvrdila studie autorů Katz *et al.* (2007). HDAC1 tedy může být specificky přitahována k proviru a iniciovat tak umlčení jeho exprese, jak zde již bylo popsáno dříve (kapitoly 4.2. Interakční síť represivních epigenetických modifikací a 4.3. Iniciale umlčování proviru).

4.6. Souhrn faktorů ovlivňujících provirovou expresi

Stabilní exprese a umlčování proviru jsou děje, na kterých se podílí množství buněčných faktorů, z nichž některé již byly identifikovány. Toho bylo dosaženo jejich inhibicí, případně cíleným vypnutím exprese daných faktorů. Jejich množství poukazuje na složitou síť, v níž hlavní roli hrají epigenetické modifikace.

Ačkoliv methylovaná DNA proviru odpovídá jeho umlčení, zdaleka ne vždy je inaktivní virus methylován. Methylace DNA nejspíše zodpovídá za udržování tohoto stavu, ale hlavně je zodpovědná za „uzamčení“ proviru v umlčeném stavu (Obrázek 4.). Iniciačním krokem k umlčení proviru však může být i deacetylace provirových histonů. Pokud však nebyly histony v oblasti proviru acetylovány (což může být případ nově nasednuvších histonů), může jeho inaktivaci zahájit i umístění represivních methylových značek, jako je H3K9me. Není tedy zcela jasné, která z modifikací je tou první, způsobující umlčení proviru. Faktorem určujícím, jakým způsobem bude provirus umlčen, mohou být podmínky panující v místě integrace.

Umlčení může být i důsledkem jakési buněčné obrany proti expresi cizorodých elementů (Greger *et al.*, 2005; Katz *et al.*, 2007; Poleshko *et al.*, 2008). E nukleační místo umlčování by se tak nacházelo na proviru, a to nejen na LTR, ale také na vnitřních sekvencích proviru (Taniguchi *et al.*, 2005). Vzhledem k faktu, že ne všechny proviry jsou umlčovány, i zde musí existovat jakási ochrana proviru proti těmto obranným mechanismům. Ochrannou ruku tak mohou nad proviry držet právě místně specifické podmínky.



Obr.4. Model umlčování a reaktivace HIV. Model dobře reaktivovatelného a „uzamčeného“ umlčení provirové exprese u HIV. Za „uzamčení“ proviru v umlčeném stavu může podle tohoto modelu hlavně přítomnost methylace DNA (znázorněno černými značkami). Modrým obdelníkem je znázorněna oblast 5'LTR. Šipky ukazují míru transkripce z promotoru HIV. Ke každému chromatinovému stavu jsou přiřazeny faktory hrající roli v ustanovení daného stavu. MeK4 – methylace lysinu 4, AcH3 – ecetylace histonu 3, MeK9 – methylace lysinu 9, MeK27 – methylace lysinu 27, HDAC – histon deacetylázy, HAT – histon acetyl transferázy, DNMT – DNA methyl transferázy. Převzato a upraveno z Blažková *et al.* (2009).

Ačkoliv se integrační místo zdá být nejdůležitějším faktorem určující osud proviru, detailní popis takových vztahů a interakcí chybí. Známé je pouze to, že stabilně exprimované proviry jsou integrovány do genů s rozsáhlým expresním profilem (Plachý *et al.*, 2010) a že integrace do míst bohatých na CpG methylaci vede k methylaci provirové DNA a umlčení jeho exprese. Přenos epigenetických znaků z okolí na provirus tak může být důležitým faktorem v regulaci genové exprese proviru. Podrobnější popis funkce epigenetických značek a regulačních sekvencí nacházejících se v okolí proviru ve vztahu k provirové expresi může být cílem dalšího zkoumání. Tyto poznatky mohou vést nejen k vytvoření modelu pro ochranu či represi exprese proviru, ale také k poznatku, jak tyto sekvence a epigenetické značky ovlivňují expresi vlastního genomu organismů.

5. Preference a exprese retrovirových vektorů

Retrovirové vektory využívané k transmissi genů sdílejí všechny vlastnosti s mateřskými retroviry. Mají tytéž integrační preference a vykazují i tutéž míru exprese s tendencí k umlčování. Právě tyto vlastnosti „divokých“ virů jsou překážkami k vytvoření dokonalých retrovirových vektorů využitelných např. v genové terapii. Stabilní expresi lze totiž nalézat hlavně u retrovirů integrujících se do aktivních, na geny bohatých oblastí, kde hrozí narušení exprese buněčných genů. Je-li naopak provirus často integrován do genově chudých oblastí, dochází často k jeho umlčování. Metodami genového inženýrství však lze tyto nedostatky upravovat. Základní snahou tak je (1) upravit integrační charakter retrovirů tak, aby nedocházelo k narušení expresního profilu genomu, a (2) docílit stabilní exprese transgenů a ochránit je tak před často pozorovaným umlčováním.

5.3. Změna integračních preferencí retrovirových vektorů

Integrační preference retrovirů lze měnit již záměnou retrovirových proteinů za proteiny retroviru s jinými preferencemi (Felice *et al.*, 2009; Lewinski *et al.*, 2006). Byla-li u lentivirového vektoru zaměněna sekvence kódující IN za IN pocházející z gammaretrovirového vektoru, došlo ke změně jeho preferencí od aktivních TU k oblastem TSS a CpG ostrovů (Lewinski *et al.*, 2006) či k oblastem bohatým na vazebná místa transkripčních faktorů (TFBS) (Felice *et al.*, 2009). Tento posun byl nejmarkantnější v případě záměny IN i Gag oblasti. Samotná záměna IN však také vedla k výrazné změně v preferencích vektoru, zatímco pouhá záměna Gag oblasti dávala za vznik vektoru s intermediálním chováním kombinujícím oba rodičovské vektory. V tomto směru je tedy dominantním virovým proteinem určujícím integrační preference IN, přičemž určitou doplňující roli hrají i proteiny kódované Gag oblastí virového genomu. Podobný efekt pak může mít i záměna U3 oblasti LTR obsahující regulační oblasti a TFBS, což může naznačovat i vliv buněčných DNA vazebných proteinů interagujících s cDNA na cílení integrace (Felice *et al.*, 2009). Změny v integračních preferencích vektorů lze dosáhnout i cílenými modifikacemi retrovirových a buněčných proteinů, které tyto preference ovlivňují.

5.4. Cílení integrace modifikací virových proteinů

První pokusy o řízenou integraci byly zaměřeny na modifikace IN, která je hlavním virovým faktorem integrace a jejího cílení. Jelikož IN neobsahuje sekvencně specifickou DNA vazebnou doménu, byly vytvořeny fúzní proteiny IN s DNA vazebnými doménami proteinů, jako je např. Sp1 (Peng *et al.*, 2002). Nejčastěji používané jsou právě domény tzv. zinkových prstů (ZFD), které jsou zodpovědné za vazbu široké škály regulačních faktorů k jejich cílovým sekvencím a mají schopnost zaměřovat integraci vektoru do určitých oblastí hostitelské DNA (Peng *et al.*, 2002; Tan *et al.*, 2006). Fúze IN a ZFD však může způsobovat narušení funkce IN (Bushman and Miller, 1997). Lim *et al.* (2010) publikovali výsledky určující místa v Gag-Pol kódující oblasti vektoru, kam je možno vložit sekvenci ZFD tak, aby nedocházelo k narušení permisivity vektoru. Tyto oblasti se nacházejí v proteinu 12

(p12) kódovaném Gag, RT a IN, oba produkty genu Pol. Vektory nesoucí takovéto fúzní proteiny, hlavně pokud jde o RT nebo IN, navíc vykazují změnu v integračních preferencích, kdy zhruba 40% pozorovaných integračních událostí je mířeno k očekávaným sekvencím. Přestože jsou definována místa permissivní pro modifikace virových proteinů, čímž je umožněna změna v cílení integrace, zatím dosažený posun integrační preference od TU není natolik markantní, aby jej bylo možno označit za cílenou integraci.

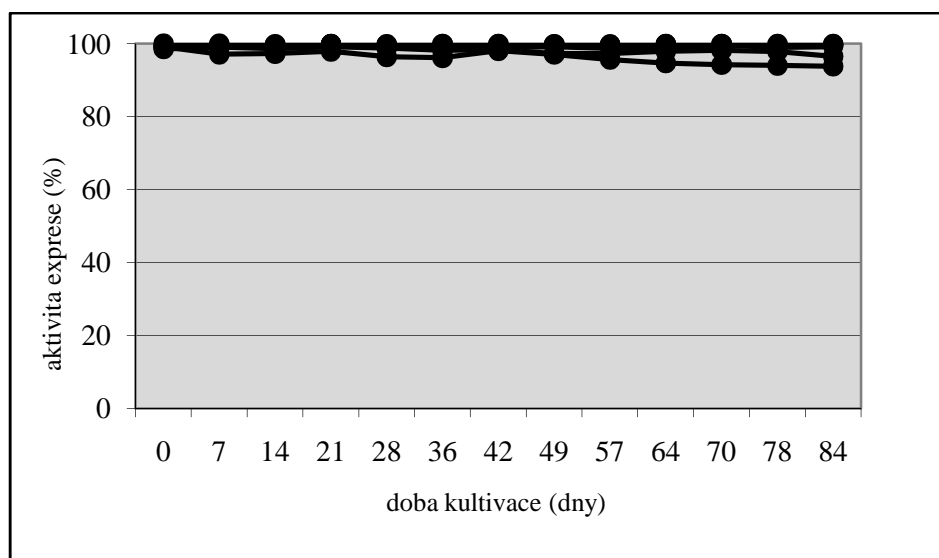
5.5. Cílení integrace modifikací buněčných proteinů

Možnou cestou k cílení integrace se ukazují být modifikace buněčných proteinů, u nichž byla prokázána interakce s IN a které tak mohou modulovat integrační preference. Jediným takto definovaným faktorem je LEDGF/p75 interagující s IN HIV-1. Úpravou LEDGF/p75, jenž byl fúzován se specifickými DNA vazebnými doménami, bylo dosaženo zacílení integrace v *in vitro* systému (Ciuffi *et al.*, 2006a), ale také v buněčné kultuře (Gijsbers *et al.*, 2010; Meehan *et al.*, 2009). Právě Gijsbers *et al.* (2010) dokázali pomocí fúze LEDGF/p75 s chromodoménou heterochromatin vazebného proteinu CBX1 (jinak také HP1 β) významně posunout integrační preference k oblastem s vysokým výskytem H3K9me, tedy do oblastí umlčených genů a heterochromatinu. Ačkoliv se jednalo o buňky, ve kterých byla snížena exprese LEDGF/p75, tyto výsledky poukazují na fakt, že integrační preference lze ovlivňovat i skrze buněčné interakční partnery IN. Hare *et al.* (2009) navíc dokázali vytvořit takovou od HIV odvozenou IN, která neinteraguje s buněčným LEDGF/p75, ale vytváří vazbu s LEDGF/p75 nesoucím supresorovou mutací. Spojením těchto postupů se tak otevírá možnost ovlivňovat cílení integrace pomocí rekombinantních proteinů obsahujících mutovanou IN vazebnou doménu a specifickou DNA vazebnou doménu.

5.6. Udržení stabilní exprese vektoru

Expresí retrovirových vektorů je variabilní v závislosti na typu cílových buněk, integračním místě a typu použitého vektoru. Stabilně exprimované proviry se často nacházejí v blízkosti CpG ostrovů či jiných regulačních sekvencích působících jako ochranné prvky buněčné exprese. Přítomnost takových sekvencí uvnitř proviru by tedy mohla snižovat frekvence umlčování exprese proviru a vést k vytvoření vektorů se stabilní expresí.

Skvencemi vkládanými do retrovirového vektoru jsou nejrůznější buněčné izolátory, jako je kuřecí cHS4 izolátor (Emery *et al.*, 2000; Rivella *et al.*, 2000), enhancery (Groth and Emery, 2010), CpG ostrovy (Hejnar *et al.*, 2001) či jejich části (Šenigl *et al.*, 2008) (obr. 5), ale i takové sekvence, které působí jako MAR (Buceta *et al.*, 2011). Vložení těchto sekvencí do oblasti proviru má pozitivní vliv na jejich expresi často spojovaný se zabráněním methylace provirových sekvencí či zabráněním vlivu methylace DNA, kdy i *in vitro* methylovaný provirus obsahující takovou regulační sekvenci vykazoval jistou míru exprese.



Obr. 5. Expresa modifikovaného retrovirového vektoru. Znázornění exprese vektoru odvozeného od ASLV s vloženým elementem z CpG ostrovu ze studie Šenigl *et al.* (2008). Znázorněné proviry vykazují stabilní expresi v kontrastu k expresi neupraveného vektoru z obr. 3.

Zvýšené míry exprese navíc bývá dosaženo pouze v některých buněčných typech. Toto je problém hlavně elementů, na něž se vážou buněčně specifické regulační faktory. Univerzální využití by v tomto směru mohly poskytovat právě sekvence odvozené od CpG ostrovů (Šenigl *et al.*, 2008), které fungují jako expresně regulační prvky napříč buněčnými typy. Dalším problémem vkládání takovýchto elementů je jejich pozice (často nejlépe fungují při vložení do určité oblasti proviru), orientace a velikost, neboť by měly co nejméně omezovat kapacitu vektoru. Pozitivně regulační elementy vkládané do vektorů také mohou způsobovat aktivaci okolního chromatinu a aktivovat tak umlčené geny, a to i na poměrně velké vzdálenosti (Singhal *et al.*, 2011). Tomuto efektu lze předejít např. použitím izolátorových sekvencí, které mají schopnost eliminovat vliv okolního chromatinu, ale vlastní aktivační funkci postrádají.

5.7. Využití a možnosti retrovirových vektorů

Cílem modifikací retrovirových vektorů je hlavně zajistit jejich vysokou, stabilní expresi a zamezit integraci do oblastí aktivních genů a TSS. V takových oblastech může totiž docházet k narušení exprese genů či aktivaci „spících“ protoonkogenů. Ideálem jsou tedy expresně velmi stabilní vektory se specificky cílenou integrací. U původních retrovirů se tyto jevy - mimogenová integrace a stabilní exprese - navzájem vylučují, a tak je potřeba důmyslných modifikací vektorů. Nejnadějnějšími se zdají být vektory odvozené od lentivirů, neboť ty bývají přirozeně expresně velmi stabilní. Lentiviry jsou navíc díky HIV způsobujícímu AIDS nejprozkoumanější skupinou retrovirů. Ačkoliv nebyla nikdy pozorována onkogeneze způsobená lentivirovými vektory, je jejich nevýhodou vysoká míra integrace do aktivních TU. Navíc bylo pozorováno, že lentivirový vektor, jehož integrační preference byla posunuta směrem k neaktivním oblastem chromatinu, si zachovává expresi velmi podobnou vektoru původnímu (Gijssbers *et al.*, 2010). Je však otázkou do jaké míry jsou preference takto upraveného

lentivirového vektoru pro TU nižší, či alespoň srovnatelné s jinými retroviry. Přesto by se dalo říci, že lentivirové vektory jsou i díky možnosti syntetizovat rekombinantní, integraci cílicí proteiny, nejbližší vektorovému ideálu. Využití však mohou nalézt např. i vektory odvozené od ptačích retrovirů, které již přirozeně vykazují velmi nízkou preferenci pro integrace do TU. Jejich největším záporem tedy zůstává sklon pro časté umlčování exprese, které je však možno (prozatím alespoň částečně) eliminovat pomocí inserce regulačních elementů.

Navzdory jejich nedokonalosti, byly retrovirové vektory úspěšně použity při genové terapii několika dědičných chorob, mezi které se řadí např. SCID, β -talasemie (Cavazzana-Calvo *et al.*, 2010) či Wiskottův Aldrichův syndrom (Boztug *et al.*, 2010). V počátcích byly ke genové terapii využívány gammaretrovirové vektory. Rozvoj klonálních expanzí či u některých pacientů až leukemických onemocnění po vystavení takové genové terapii, však poukazuje na nebezpečnost používání původních vektorů. Gammaretrovirové vektory jsou tak pomalu nahrazovány jinými, nejčastěji lentivirovými vektory. Ačkoliv je největší přínos retrovirových vektorů v genové terapii, přenosu a expresi transgenů, mohou sloužit i k charakterizaci exprese konkrétních oblastí genomu či nejruznějších expresně regulačních elementů, které bývají využívány např. k nastolení stabilní exprese transgenů.

6. Závěr

Integrační preference a exprese retroviru jsou děje úzce provázané a klíčové pro produkci jeho virového potomstva. Má-li dojít k efektivní expresi provirových genů, musí být provirus integrován do míst, která budou pro jeho expresi permissivní a poskytnou proviru ochranu před umlčovací mašinérií buňky. Taková místa představují hlavně aktivní TU, TFBS a oblasti CpG ostrovů. Jakákoliv integrace mimo tyto oblasti může vyústit v umlčení proviru. Ačkoliv bylo identifikováno množství faktorů činných v umlčování exprese proviru, hlavní roli v tom, zda bude provirus aktivní či nikoliv, hrají podmínky panující v místě integrace. Kombinace epigenetických značek, hlavně histonových modifikací a methylace DNA, je nejspíše rozhodující. Hlavní funkcí epigenetických značek je totiž převážně přitahování efektorů, které mohou rozšiřovat daný stav chromatinu na okolí, a tedy i na provirus. Jiné epigenetické značky mohou samy nasedání takových efektorů blokovat.

V epigeneticky poznamenaném prostředí genomu je tedy pro retrovirus důležité integrovat se do takového prostředí, které umožní rozšíření aktivních epigenetických značek, kterými jsou např. H3K4me, H3ac či H4ac, a naopak zabrání rozšíření represivních podmínek značených hlavně deacetylací histonů, H3K9me a methylocí DNA. Tento trend je ve vlivu integračních preferencí na míru umlčování exprese u retrovirů skutečně pozorovatelný, neboť čím více retrovirus preferuje aktivní TU, případně CpG ostrovy a TSS či místa bohaté na H3K4me, tím méně bývá umlčován. Tato závislost však není úplná, neboť svou roli hraje i sekvenční složení proviru a jeho citlivost k umlčování (např. množství methylovatelných CpG dinukleotidů). Není vyloučen ani vliv integrace proviru na remodelování okolního chromatinu, čímž by opět částečně poklesl vliv oblastí retroviru nejbližších.

Ucelený pohled na mozaiku integrace charakterizací nových epigenetických superznaček a faktorů hrajících roli v cílení integrace nalézá využití v konstrukci účinných a bezpečných retrovirových vektorů. Aplikací takových superznaček bude totiž možno predikovat integrační místa retrovirů, či naopak charakterizovat podmínky panující v oblasti integrovaného proviru. Společně s pozorováním míry exprese proviru a její proměnlivosti lze takto za pomoci retrovirů popsat stav, funkci či expresní profil a stabilitu dané genomové oblasti. Poznatky o možnostech cílené integrace spolu s dosažením stabilní exprese transgenů jsou pak reálnými cíly, ke kterým vývoj nových retrovirových vektorů jistě směřuje.

Seznam použité literatury

- 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.
- Baumann, M., Pontiller, J., and Ernst, W. (2010). Structure and basal transcription complex of RNA Polymerase II core promoters in the mammalian genome: An overview. *Molecular Biotechnology* 45, 241-247.
- 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.
- Berry, C., Hannenhalli, S., Leipzig, J., and Bushman, F. D. (2006). Selection of target sites for mobile DNA integration in the human genome. *PLoS Comput Biol* 2, e157.
- 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.
- 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.
- Brady, T., Lee, Y. N., Ronen, K., Malani, N., Berry, C. C., Bieniasz, P. D., and Bushman, F. D. (2009). Integration target site selection by a resurrected human endogenous retrovirus. *Genes & Development* 23, 633-642.
- Buceta, M., Galbete, J. L., Kostic, C., Arsenijevic, Y., and Mermod, N. (2011). Use of human MAR elements to improve retroviral vector production. *Gene Therapy* 18, 7-13.
- Bushman, F. D., and Miller, M. D. (1997). Tethering human immunodeficiency virus type 1 preintegration complexes to target DNA promotes integration at nearby sites. *Journal of Virology* 71, 458-464.
- 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.
- Cedar, H., and Bergman, Y. (2009). Linking DNA methylation and histone modification: Patterns and paradigms. *Nature Reviews Genetics* 10, 295-304.
- Ciuffi, A., Diamond, T. L., Hwang, Y., Marshall, H. M., and Bushman, F. D. (2006a). Modulating target site selection during human immunodeficiency virus DNA integration in vitro with an engineered tethering factor. *Human Gene Therapy* 17, 960-967.
- Ciuffi, A., Llano, M., Poeschla, E., Hoffmann, C., Leipzig, J., Shinn, P., Ecker, J. R., and Bushman, F. (2005). A role for LEDGF/p75 in targeting HIV DNA integration. *Nature Medicine* 11, 1287-1289.
- Ciuffi, A., Mitchell, R. S., Hoffmann, C., Leipzig, J., Shinn, P., Ecker, J. R., and Bushman, F. D. (2006b). Integration site selection by HIV-based vectors in dividing and growth-arrested IMR-90 lung fibroblasts. *Molecular Therapy* 13, 366-373.
- Cross, S. H., and Bird, A. P. (1995). CpG islands and genes. *Current Opinion in Genetics and Development* 5, 309-314.
- Delelis, O., Lehmann-Che, J., and Saïb, A. (2004). Foamy viruses -- a world apart. *Current Opinion in Microbiology* 7, 400-406.
- 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.
- Dong, K. B., Maksakova, I. A., Mohn, F., Leung, D., Appanah, R., Lee, S., Yang, H. W., Lam, L. L., Mager, D. L., Schübeler, D., *et al.* (2008). DNA methylation in ES cells requires the lysine methyltransferase G9a but not its catalytic activity. *EMBO Journal* 27, 2691-2701.

- 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.
- Fassati, A., and Goff, S. P. (1999). Characterization of intracellular reverse transcription complexes of moloney murine leukemia virus. *J Virol* 73, 8919-8925.
- Fassati, A., and Goff, S. P. (2001). Characterization of intracellular reverse transcription complexes of human immunodeficiency virus type 1. *J Virol* 75, 3626-3635.
- 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.
- Fischle, W. (2008). Talk is cheap - Cross-talk in establishment, maintenance, and readout of chromatin modifications. *Genes and Development* 22, 3375-3382.
- Gijsbers, R., Ronen, K., Vets, S., Malani, N., De Rijck, J., McNeely, M., Bushman, F. D., and Debysers, 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. *J Virol* 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., McCormack, M. P., Wulffraat, N., Leboulch, P., Lim, A., Osborne, C. S., Pawliuk, R., Morillon, E., *et al.* (2003). LMO2-associated clonal T cell proliferation in two patients after gene therapy for SCID-X1. *Science* 302, 415-419.
- 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.
- Hare, S., Shun, M. C., Gupta, S. S., Valkov, E., Engelman, A., and Cherepanov, P. (2009). A novel co-crystal structure affords the design of gain-of-function lentiviral integrase mutants in the presence of modified PSIP1/LEDGF/p75. *PLoS Pathogens* 5.
- 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.
- Hejnar, J., Elleder, D., Hájková, P., Walter, J., Blažková, J., and Svoboda, J. (2003). Demethylation of host-cell DNA at the site of avian retrovirus integration. *Biochemical and Biophysical Research Communications* 311, 641-648.
- 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.
- Hejnar, J., Plachý, J., Geryk, J., Machoň, O., Trejbalová, K., Guntaka, R. V., and Svoboda, J. (1999). Inhibition of the Rous sarcoma virus long terminal repeat-driven transcription by in vitro methylation: Different sensitivity in permissive chicken cells versus mammalian cells. *Virology* 255, 171-181.
- Hematti, P., Hong, B. K., Ferguson, C., Adler, R., Hanawa, H., Sellers, S., Holt, I. E., Eckfeldt, C. E., Sharma, Y., Schmidt, M., *et al.* (2004). Distinct genomic integration of MLV and SIV vectors in primate hematopoietic stem and progenitor cells. *PLoS Biology* 2.

- Hoeben, R. C., Migchielsen, A. A., van der Jagt, R. C., van Ormondt, H., and van der Eb, A. J. (1991). Inactivation of the Moloney murine leukemia virus long terminal repeat in murine fibroblast cell lines is associated with methylation and dependent on its chromosomal position. *J Virol* *65*, 904-912.
- 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.
- Challita, P. M., and Kohn, D. B. (1994). Lack of expression from a retroviral vector after transduction of murine hematopoietic stem cells is associated with methylation in vivo. *Proceedings of the National Academy of Sciences of the United States of America* *91*, 2567-2571.
- Chávez, L., Kauder, S., and Verdin, E. (2011). In vivo, in vitro, and in silico analysis of methylation of the HIV-1 provirus. *Methods* *53*, 47-53.
- Cherepanov, P., Maertens, G., Proost, P., Devreese, B., Van Beeumen, J., Engelborghs, Y., De Clercq, E., and Debysse, Z. (2003). HIV-1 integrase forms stable tetramers and associates with LEDGF/p75 protein in human cells. *Journal of Biological Chemistry* *278*, 372-381.
- Choi, J. K. (2010). Contrasting chromatin organization of CpG islands and exons in the human genome. *Genome Biology*.
- Jin, C., Zang, C., Wei, G., Cui, K., Peng, W., Zhao, K., and Felsenfeld, G. (2009). H3.3/H2A.Z double variant-containing nucleosomes mark 'nucleosome-free regions' of active promoters and other regulatory regions. *Nat Genet* *41*, 941-945.
- Johnson, C. N., and Levy, L. S. (2005). Matrix attachment regions as targets for retroviral integration. *Virology Journal* *2*.
- Kaplan, N., Moore, I. K., Fondufe-Mittendorf, Y., Gossett, A. J., Tillo, D., Field, Y., LeProust, E. M., Hughes, T. R., Lieb, J. D., Widom, J., and Segal, E. (2009). The DNA-encoded nucleosome organization of a eukaryotic genome. *Nature* *458*, 362-366.
- Katz, R. A., Jack-Scott, E., Narezkina, A., Palagin, I., Boimel, P., Kulkosky, J., Nicolas, E., Greger, J. G., and Skalka, A. M. (2007). High-frequency epigenetic repression and silencing of retroviruses can be antagonized by histone deacetylase inhibitors and transcriptional activators, but uniform reactivation in cell clones is restricted by additional mechanisms. *Journal of Virology* *81*, 2592-2604.
- Kauder, S. E., Bosque, A., Lindqvist, A., Planelles, V., and Verdin, E. (2009). Epigenetic regulation of HIV-1 latency by cytosine methylation. *PLoS Pathogens* *5*.
- Kim, T. H., Barrera, L. O., Zheng, M., Qu, C., Singer, M. A., Richmond, T. A., Wu, Y., Green, R. D., and Ren, B. (2005). A high-resolution map of active promoters in the human genome. *Nature* *436*, 876-880.
- Koiwa, T., Hamano-Usami, A., Ishida, T., Okayama, A., Yamaguchi, K., Kamihira, S., and Watanabe, T. (2002). 5'-long terminal repeat-selective CpG methylation of latent human T-cell leukemia virus type 1 provirus in vitro and in vivo. *J Virol* *76*, 9389-9397.
- 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.
- Lee, M. S., and Craigie, R. (1998). A previously unidentified host protein protects retroviral DNA from autointegration. *Proceedings of the National Academy of Sciences of the United States of America* *95*, 1528-1533.
- 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., Yamashita, M., Emerman, M., Ciuffi, A., Marshall, H., Crawford, G., Collins, F., Shinn, P., Leipzig, J., Hannenhall, S., *et al.* (2006). Retroviral DNA integration: Viral and cellular determinants of target-site selection. *PLoS Pathogens* *2*, 0611-0622.
- Lim, K. I., Klimczak, R., Yu, J. H., and Schaffer, D. V. (2010). Specific insertions of zinc finger domains into Gag-Pol yield engineered retroviral vectors with selective integration properties. *Proceedings of the National Academy of Sciences of the United States of America* *107*, 12475-12480.

- 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., Debysers, 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.
- Majumdar, A., Chatterjee, A. G., Ripmaster, T. L., and Levin, H. L. (2011). Determinants that specify the integration pattern of retrotransposon Tf1 in the *fbp1* promoter of *Schizosaccharomyces pombe*. *Journal of Virology* 85, 519-529.
- 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.
- 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.
- Murakami, H., Yamada, T., Suzuki, M., Nakahara, Y., Suzuki, K., and Sentsui, H. (2011). Bovine leukemia virus integration site selection in cattle that develop leukemia. *Virus Research* 156, 107-112.
- 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.
- Nowrouzi, A., Dittrich, M., Klanke, C., Heinkelein, M., Rammling, M., Dandekar, T., von Kalle, C., and Rethwilm, A. (2006). Genome-wide mapping of foamy virus vector integrations into a human cell line. *Journal of General Virology* 87, 1339-1347.
- Ozawa, T., Itoyama, T., Sadamori, N., Yamada, Y., Hata, T., Tomonaga, M., and Isobe, M. (2004). Rapid isolation of viral integration site reveals frequent integration of HTLV-1 into expressed loci. *Journal of Human Genetics* 49, 154-165.
- Peng, W. J., Chang, C. M., and Lin, T. H. (2002). Target integration by a chimeric Sp1 zinc finger domain-Moloney murine leukemia virus integrase in vivo. *Journal of Biomedical Science* 9, 171-184.
- Pion, M., Jordan, A., Biancotto, A., Dequiedt, F., Gondois-Rey, F., Rondeau, S., Vigne, R., Hejnar, J., Verdin, E., and Hirsch, I. (2003). Transcriptional suppression of in vitro-integrated human immunodeficiency virus type 1 does not correlate with proviral DNA methylation. *Journal of Virology* 77, 4025-4032.
- 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., Shalginiskikh, 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.
- Pruss, D., Bushman, F. D., and Wolffe, A. P. (1994). Human immunodeficiency virus integrase directs integration to sites of severe DNA distortion within the nucleosome core. *Proceedings of the National Academy of Sciences* 91, 5913-5917.
- Rivella, S., Callegari, J. A., May, C., Tan, C. W., and Sadelain, M. (2000). The cHS4 insulator increases the probability of retroviral expression at random chromosomal integration sites. *Journal of Virology* 74, 4679-4687.
- 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.
- Roh, T. Y., Cuddapah, S., and Zhao, K. (2005). Active chromatin domains are defined by acetylation islands revealed by genome-wide mapping. *Genes and Development* 19, 542-552.

- Santoni, F. A., Hartley, O., and Luban, J. (2010). Deciphering the code for retroviral integration target site selection. *PLoS Comput Biol* 6, e1001008.
- Segal, E., Fondufe-Mittendorf, Y., Chen, L., Thåström, A., Field, Y., Moore, I. K., Wang, J. P. Z., and Widom, J. (2006). A genomic code for nucleosome positioning. *Nature* 442, 772-778.
- Selker, E. U. (1999). Gene silencing: repeats that count. *Cell* 97, 157-160.
- 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.
- Singhal, R., Deng, X., Chenchik, A. A., and Kandel, E. S. (2011). Long-distance effects of insertional mutagenesis. *PLoS ONE* 6, e15832.
- Suzuki, Y., and Craigie, R. (2007). The road to chromatin - Nuclear entry of retroviruses. *Nat Rev Microbiol* 5, 187-196.
- Šenigl, F., Plachý, J., and Hejnar, J. (2008). The core element of a CpG island protects avian sarcoma and leukemia 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.
- Tachibana, M., Matsumura, Y., Fukuda, M., Kimura, H., and Shinkai, Y. (2008). G9a/GLP complexes independently mediate H3K9 and DNA methylation to silence transcription. *EMBO Journal* 27, 2681-2690.
- Tan, W., Dong, Z., Wilkinson, T. A., Barbas Iii, C. F., and Chow, S. A. (2006). Human immunodeficiency virus type 1 incorporated with fusion proteins consisting of integrase and the designed polydactyl zinc finger protein E2C can bias integration of viral DNA into a predetermined chromosomal region in human cells. *Journal of Virology* 80, 1939-1948.
- Taniguchi, Y., Nosaka, K., Yasunaga, J.-i., Maeda, M., Mueller, N., Okayama, A., and Matsuoka, M. (2005). Silencing of human T-cell leukemia virus type I gene transcription by epigenetic mechanisms. *Retrovirology* 2, 64.
- Trobridge, G. D., Miller, D. G., Jacobs, M. A., Allen, J. M., Kiem, H.-P., Kaul, R., and Russell, D. W. (2006). Foamy virus vector integration sites in normal human cells. *Proceedings of the National Academy of Sciences of the United States of America* 103, 1498-1503.
- Tyagi, M., and Karn, J. (2007). CBF-1 promotes transcriptional silencing during the establishment of HIV-1 latency. *EMBO Journal* 26, 4985-4995.
- 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.
- 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.
- 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.
- Weidhaas, B. J., Angelichio, E. L., Fenner, S., and Coffin, J. M. (2000). Relationship between retroviral DNA integration and gene expression. *Journal of Virology* 74, 8382-8389.
- Weinberg, J. B., Matthews, T. J., Cullen, B. R., and Malim, M. H. (1991). Productive human immunodeficiency virus type 1 (HIV-1) infection of nonproliferating human monocytes. *Journal of Experimental Medicine* 174, 1477-1482.
- 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.
- 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.
- Wu, Y. T. (2008). The second chance story of HIV-1 DNA: Unintegrated? Not a problem! *Retrovirology* 5.
- Yang, J. N., and Dudley, J. (1992). Endogenous Mtv-8 or a closely linked sequence stimulates rearrangement of the downstream V κ 9 gene. *Journal of Immunology* 149, 1242-1251.
- Zhang, Y., Moqtaderi, Z., Rattner, B. P., Euskirchen, G., Snyder, M., Kadonaga, J. T., Liu, X. S., and Struhl, K. (2009). Intrinsic histone-DNA interactions are not the major determinant of nucleosome positions in vivo. *Nat Struct Mol Biol* 16, 847-852.
- Zhang, Y., Moqtaderi, Z., Rattner, B. P., Euskirchen, G., Snyder, M., Kadonaga, J. T., Liu, X. S., and Struhl, K. (2010). Reply to Evidence against a genomic code for nucleosome positioning. *Nature Structural and Molecular Biology* 17, 920-923.
- Zheng, Y., Ao, Z., Jayappa, K. D., and Yao, X. (2010). Characterization of the HIV-1 integrase chromatin- and LEDGF/p75-binding abilities by mutagenic analysis within the catalytic core domain of integrase. *Virology Journal* 7.