

Stopy retrovirů v lidském genomu

Retroviry mají svou dědičnou informaci uloženou v molekulách RNA. V jejich životním cyklu je zapojen enzym reverzní transkriptáza, která umožňuje přepis RNA do DNA, jež pak další enzym – integráza – začleňuje do genomu hostitelské buňky (viz také Živa 2006, 1: 6–8 a předchozí článek na str. 101–104 tohoto čísla). Jestliže dojde k integraci retroviru v zárodečné (germinální) linii, může se stát součástí genomu hostitelského organismu. V lidském genomu tvoří endogenní retroviróvé elementy (Human Endogenous Retroviruses, HERVs) zhruba 8 % DNA. Až donedávna se soudilo, že jednotlivé elementy nejsou až na výjimky v genomu aktivní, tedy, že nedochází k tvorbě virových partikulí. Během evoluce prošly mutacemi a ztratily schopnost dokončit životní cyklus a infikovat další buňky. Nedávno však bylo objeveno, že některým lidským endogenním retrovirům (nebo některým jejich genům) určitá schopnost genové exprese zůstala a dokonce mohou v lidském genomu plnit zásadní funkce. Příkladem může být gen *env* retroviru HERV-W (viz Živa 2007, 5: 236–237), kódující protein syncytin-1, který je zapojen do fúze buněk lidské placenty. Jak je to však s transkripcí a genovou expresí ostatních endogenních retrovirů člověka? Jsou skutečně jen tichými pozůstatky nákaz našich dávných předků?

Z nejnovějších výzkumů vyplývá, že přepis (transkripce) HERV, tedy vznik mediátorové RNA (mRNA) není žádnou vzácností, zatímco její translace do proteinů zůstává spíše výjimečná, protože většina HERVs je natolik mutována, že nemůže dát vzniknout funkčním proteinům. V některých případech se mohou tvořit krátké peptidy.

Struktura retroviru

Na obr. 1 je schematicky znázorněna obecná struktura infekčního retroviru. Na obou koncích retroviróvého genomu jsou dlouhé koncové repetice – Long Terminal Repeats (LTR). Skládají se jednak z přímých repetit a jednak z unikátních sekvencí, které obsahují regulační prvky důležité pro transkripci, jako jsou promotory (na nichž dochází k zahájení transkripce), enhancey (zesilovače transkripce) a terminátorové (ukončující) sekvence. Mezi sekvencemi LTR se nacházejí tři geny: *gag*, *pol* a *env*. *Gag* kóduje nukleokapsidu – strukturální protein virové kapsidy, *pol* kóduje integrázu a reverzní transkriptázu, *env* obalové

glykoproteiny. Endogenní retroviróvé elementy mají tuto strukturu analogickou, ale vzhledem k jejich stáří v lidském genomu je redukována a pozměněna, obsahuje mnoho mutací, delecí a inzercí. To zapříčiňuje porušení čtecího rámce (Open Reading Frame, ORF) genů HERV. Jde o část kódující protein, která začíná startovacím kodonem a končí některým ze stop kodonů. Např. z přibližně 4 000 míst v lidském genomu, kde se vyskytuje gen *env* některého endogenního retroviru, jen 16 obsahuje celý úsek, který potenciálně může kódovat protein Env. U genů *gag* a *pol* vychází rozdíl mezi počtem míst v lidském genomu obsahujících jejich ORF neporušený a mutovaný ještě větší. U genu *gag* z 9 500 výskytů se 17 ORF zachovalo kompletních a u genu *pol* je to jen 13 z 20 900. Extrémní případ představuje ztráta všech virových genů a výskyt pouze soliterních sekvencí LTR.

V lidském genomu bylo identifikováno kolem 200 různých rodin endogenních retroviróvého elementů. Byly rozděleny do tří hlavních skupin na základě podobnosti

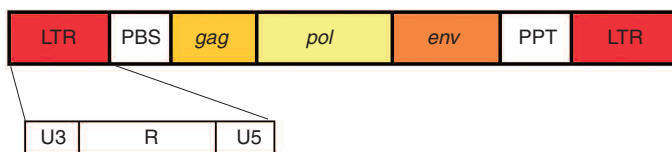
k infekčním retrovirům a každá skupina se dále dělí na mnoho podskupin. Přehled rodin lidských endogenních retrovirů ukazuje tab. 1. Z ní je mimo jiné patrná jejich různorodost. I když existuje základní schéma endogenních retrovirů, může se jejich konkrétní podoba značně lišit. HERV skupiny I jsou příbuzné gamaretrovirům (např. virus myši leukemie) a mezi nejdůležitější podskupiny patří HERV-W a HERV-H. Elementy skupiny II vykazují příbuznost k betaretrovirům, mezi které náležejí viry tumoru myši mléčné žlázy (HERV-K). Skupina III je vzdáleně příbuzná spumavirům (např. Foamy virus) a jako důležité podskupiny lze jmenovat HERV-L a HERV-S.

Aktivita HERV

Pokud chceme určit, které endogenní retroviróvé elementy a do jaké míry jsou v lidském genomu přepisovány, musíme je analyzovat v transkriptomech konkrétních jedinců. Transkriptom je veškerá mRNA přítomná v jednom okamžiku v buňce. Z této mRNA můžeme reverzní transkripcí získat komplementární – cDNA (complementary DNA), z níž lze dále připravit knihovny cDNA a tuto cDNA následně sekvenovat. Jednotlivé tkáně daného organismu obsahují odlišnou cDNA.

V našich experimentech jsme analyzovali transkripty z různých tkání (játra, žaludek, střeva, kožní, plicní a embryonální plicní fibroblasty, myoblasty příčné pruhovaných svalů, buňky epitelu a cévního epitelu, embryonální kmenové buňky) 15 osob. Jejich cDNA byla sekvenována novými vysokokapacitními metodami – pomocí Illumina sekvenační technologie, případně 454 sekvenační technologie. Každá z těchto metod má své výhody a nevýhody. Výstupem Illumina sekvenování je obrovské množství krátkých sekvencí, v našem případě o délce 100 nukleotidů. Oproti tomu 454 sekvenování poskytuje mnohem delší část jednotlivých úseků, ale ve významně menším množství. V desítkách milionů získaných sekvencí jsme pak identifikovali ty, které měly původ v některém lidském endogenním retroviru z databáze HERVd. V Ústavu molekulární genetiky AV ČR, v. v. i., byla vyvinuta databáze lidských endogenních retrovirů (HERVd) dostupná na internetové adrese <http://herv.img.cas.cz>. V ní je možné nalézt konkrétní retroviry, jejich umístění na chromozomech a zařazení do rodin. Databáze poskytuje velmi užitečný nástroj, ale její současná podoba je statická, nelze z ní bohužel zjistit informace o transkripci HERV.

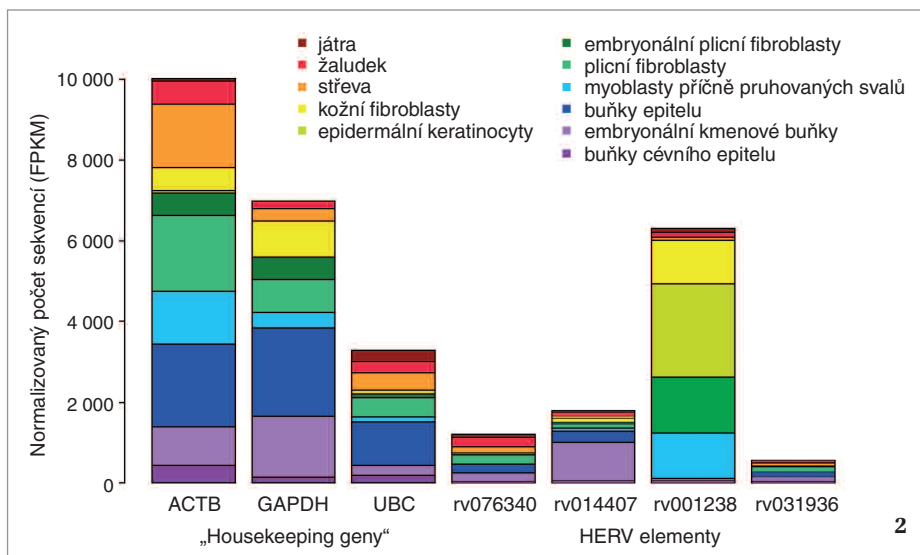
Podarilo se nám najít transkripty retroviróvého elementů v rozličných tkáních (příklad ukazuje obr. 2 na následující str.). Míra transkripce HERV byla srovnávána s přepisem tzv. „housekeeping genů“, které zajišťují základní funkce ve všech buňkách organismu; většinou se účastní jejich metabolismu.



Tab. 1 Rozdělení lidských endogenních retrovirů (HERVs)

HERV	Zastoupení v lidském genomu [%]	Příbuzné infekční retroviry	Stáří [milióny let]
Skupina I	2,5	Gamaretroviry	20–80
Skupina II	0,5	Betaretroviry	0–50
Skupina III	6,0	Spumaviry – lidský Foamy virus	20–150

1 Schéma stavby (endogenního) retroviru. LTR – Long Terminal Repeat (dlouhá koncová repetice), U3 – unikátní sekvence na 3' konci, U5 – unikátní sekvence na 5' konci; *gag*, *pol*, *env* – geny; PBS – místo vazby primeru (odkud začíná replikace), PPT – polypurinový úsek, R – repetice. Blíže v textu



2 Transkripce některých elementů HERV a „housekeeping genů“ v různých tkáních. Míra transkripce je udávána jako normalizovaný počet sekvencí (FPKM, fragments per kilobase of exon per million reads mapped). Část dat pochází z volně dostupných archivů SRA (Sequence Read Archives), šlo o vzorky tkání zdravých lidských jedinců zpracované metodami vysoce výkonného sekvenování

(High Throughput Sequencing). V ostatních případech jsme použili naše vlastní data, upravená stejným způsobem. Blíže v textu. Orig. M. Famfulíková

Z obr. 2 je vidět, že lidské endogenní retroviry jsou přepisovány ve všech sledovaných tkáních. Některé HERV byly preferenčně transkribovány v určitých tkáních, jiné byly aktivní ve všech zkoumaných

tkáních. Zajímavé poznatky přineslo srovnání transkripce HERV ve zdravé a nádorové tkáni. V nádorových tkáních můžeme pozorovat zvýšenou aktivitu HERV, což by mohlo být způsobeno především celkovou deregulací genové exprese v buňkách nádoru. Ačkoli se nezdá, že by tyto retroviry mohly vyvolat vznik nádorového bujení, jejich aktivita na průběh onemocnění pravděpodobně vliv má. Navíc z některých transkriptů vzniklých z HERV mohou být překládány krátké peptidy, jež buňku různým způsobem ovlivňují. Přestože je dnes již většina HERV lokusů neaktivních, kvůli množství mutací existují i výjimky, jejichž ORF je kompletní a dává vzniknout proteinům. Transkripce HERV je pod kontrolou jejich LTR oblastí. Transkripčně aktivní endogenní retroviry mohou potenciálně interagovat s proteiny hostitele a stát se příčinou nemoci, nebo přispívat k jejímu rozvoji. Elementy HERV jsou stále více spojovány např. s rakovinným bujením, progresí HIV, neurologickými a autoimunitními poruchami, schizofrenií a bipolární poruchou.

Endogenní retroviry nejsou jen tichými společníky, možná mají mnohem větší význam, než si myslíme.

Článek vznikl v rámci projektu Akademie věd ČR Otevřená věda.

Jan Votýpka, Kateřina Jirků Pomajbíková

Co je nového v biologii: Biomy našich těl – změna paradigmatu

Desetisekundový intenzivní milostný polibek zajistí nejen utvrzení vzájemné náklonnosti a uvolnění řady hormonů měnících fyziologický stav i chování milenců, ale též vzájemné předání přibližně 80 milionů bakterií zejména rodu *Lactobacillus* a *Bifidobacterium*. S tímto zjištěním přišli nizozemští vědci (Kort a kol. 2014) a dokládají tak, že studium mikrobiomu (mikrobioty, dříve mikrobioty, tedy souhrnu všech mikroorganismů v tělech hostitelů) je čím dál tím oblíbenějším tématem a každou chvíli se objeví článek popisující nový mikrobiom, nebo vliv mikrobiomu či některé jeho frakce na zdraví člověka.

Názor, že jsme chodící zoologické zahrady, by ještě před několika lety většina lidí považovala za poněkud přehnaný. Paradigma se však značně změnilo a fakt, že na každou lidskou buňku v našem těle připadá 10 buněk jiných organismů, začíná být součástí učebnic. (Ve vědě znamená paradigma zásadní pojetí určitého oboru, které je považováno za vzorové a sdílí ho většina vědců v určité historické etapě. Vymezuje, co má být zkoumáno, jakými přístupy a metodami, podle jakých pravidel. Paradigmata ve vědě jsou dočasná a jejich změna obvykle probíhá radikálním obratem.) Počítáme-li pouze bakterie a mikroskopické

houby, tvoří lidské buňky asi čtvrtinu z celkového počtu buněk v našem organismu. Jsme složeni přibližně z 37 trilionů buněk *Homo sapiens*, k tomu musíme přičíst zhruba 10^{14} mikrobů, kteří představují asi 1 kg z naší hmotnosti. V každém z nás přebývá více než jeden tisíc různých druhů bakterií a hub, v naprosté většině dosud nepojmenovaných. Ještě překvapivější však je, že mikroorganismy v našich tělech s dvěma miliony genů stonásobně převyšují počet genů lidských. A tak bychom se mohli zeptat: „Člověče, kdo jsi?“
Co tedy jsme, když většina buněk není našich? Zatímco jednodušší jednodruhový

pohled na naši individualitu zcela dostává pro převážnou část každodenních sociálních interakcí, skutečného biologického superjedince definujeme jako součet lidského organismu s jeho mikrobiomem. Je to právě tato vzájemná symbiotická asociace, která je schopna existence a přežívání v čase historickém i evolučním. Jak básník Walt Whitman výstižně uvádí v díle *Song of Myself*: „I am large, I contain multitudes“ (Jsem velký, nosím ti davu).

Ačkoli se pojem ekosystém stává čím dál tím více diskutovaným termínem, který byl vytvořen spíše z didaktických důvodů a ve skutečnosti možná ani neexistuje, zůstává přesto hojně používán. Je jím míněna ucelená a částečně uzavřená část přírody, např. louka, která však na všech úrovních komunikuje se svým okolím. Ekosystémy širšího až kontinentálního rozsahu označujeme za biomy, ať již vodní (biom volného oceánu či šelfových moří) nebo suchozemské (lesní biomy, savany a stepi či chladné tundry a polární pustiny). Pro ekosystém není stanovena velikost, a tak za něj lze považovat v extrémních případech celou biosféru i několik desítek mililitrů vody v listové růžici bromelie. Našemu pojetí ale často uniká jedno zcela specifické prostředí – všudypřítomné, na očích, a přesto skryté. Jde o prostředí těl organismů (říkejme jim hostitelé), které dokonale naplňuje v jiných případech jen těžko uchopitelnou definici ekosystému – splňuje totiž podmínku jasně prostorové vymezenosti a částečného (v některých případech i značného) uzavření vůči okolí.

Ekosystém hostitele tedy nejlépe naplňuje naši představu o ekosystému, ačkoli ho ekologové často přehlížejí. Vzhledem k nepatrné velikosti obyvatel nacházejících se v tělech svých hostitelů označujeme jejich