

UNIVERZITA KARLOVA V PRAZE, PŘÍRODOVĚDECKÁ FAKULTA

Katedra zoologie



**Evoluce genovou duplikací: kvantifikace kopií
retrovirových elementů a behaviorálně
významných genů**

BAKALÁŘSKÁ PRÁCE

David Vinkler

Vedoucí práce: Doc. RNDr. Pavel Stopka, PhD.

Praha 2009

Poděkování

Děkuji mému školiteli, Doc. RNDr. Pavlu Stopkovi, PhD., za jeho ochotné a trpělivé vedení této práce. Velký dík patří rovněž mým rodičům za jejich obětavou podporu během celého mého studia.

Abstrakt

Před 40 lety přišel Susumo Ohno s tvrzením, že bez duplikace genů by byl vývoj mnohobuněčných organismů, obratlovců a savců z jednobuněčných organismů zcela nemožný, protože takové velké skoky v evoluci vyžadují vytvoření nových genových lokusů s funkcemi, jaké doposud neexistovaly. Cílem této práce je dokázat, že měl z velké části pravdu.

Klíčová slova: genová duplikace, evoluce, gen, genom, genová funkce

Abstract

40 years ago, Susumo Ohno proposed that without gene duplication the creation of metazoans, vertebrates and mammals from unicellular organisms would have been impossible, because such big leaps in evolution required the creation of new gene loci with previously nonexistent function. The aim of this paper is to prove that he was largely right.

Key Words: gene duplication, evolution, gene, genome, gene function

Obsah

1. Úvod.....	5
2. Krátký pohled do historie.....	7
3. Genové rodiny.....	9
4. Mechanismy a typy genové duplikace.....	11
4.1 Úsekové duplikace.....	11
4.2 Retrotranspozice.....	13
4.3 Polyploidie: Duplikace celého genomu.....	14
4.4 Polysomie.....	17
5. Evoluční osud duplikovaných genů.....	19
5.1 Ztráta funkce genu (nonfunkcionalizace).....	20
5.2 Zachování původní funkce genu.....	22
5.3 Vznik nové funkce genu (neofunkcionalizace).....	24
5.4 Rozdělení (pod)funkcí mezi vzniklé duplikáty (subfunkcionalizace).....	26
6. Závěr.....	28
7. Seznam použité literatury.....	29

1. Úvod

Jedním ze základních rysů života na Zemi je jeho velká rozmanitost. Různorodost organismů, jejich rozličných znaků a vlastností, je možná díky změnám v DNA, které se akumulují během evoluce. Vzhledem k tomu, že vlastnosti organismů jsou převážně určeny geneticky, evoluce organismů jde nezbytně ruku v ruce s evolucí na úrovni genomu, s evolucí genů a jejich vlastností.

Gen je ústředním pojmem dnešní biologie a během doby jeho existence došlo ke značným posunům v jeho chápání. Klasické pojetí genu, v němž je gen považován za jednotku genetické analýzy (dědičnosti), vystřídala ve většině biologických oborů molekulární definice genu. Ta praví, že strukturní gen je sekvence nukleotidů nesoucí informaci, která po transkripci a translaci řídí jeden polypeptidový řetězec. Dnes už víme, že ne všechny geny jsou překládány do aminokyselinového řetězce proteinu. Konečným produktem genu může být například některý typ RNA (rRNA či tRNA). V současném pojetí tak gen představuje souvislý úsek nukleové kyseliny kódující některou funkční makromolekulu (Flegr 2005). Je-li tato definice pro potřeby všech biologů nejvhodnější, a zda je konečná, o tom můžeme diskutovat, nicméně nás budou geny zajímat především z hlediska jejich evoluce a mechanismů, které se na ní podílejí.

Základní otázka nejen molekulární evoluce zní: jak se rodí nové geny, z čeho vznikají a jaké procesy se toho účastní. Tak jako buňka vzniká z buňky, ani geny nevznikají *de novo*, ale vyvíjejí se postupně jeden z druhého. Každý nový gen má svůj původ v některém dříve existujícím genu. Podobně jako samotné organismy i geny jsou vystaveny evolučním procesům a podléhají silám zrodu a zániku. Důkazem nepřetržitého vzniku nových genů během biologické evoluce jsou druhové odlišnosti v počtu genů a v rozdílných biologických funkcích jimi řízených produktů. Například genom octomilky obecné (*Drosophila melanogaster*) obsahuje 87 genů kódujících kutikulární proteiny, zatímco háďátko obecné (*Caenorhabditis elegans*) nemá žádný takový gen (Rubin et al. 2000).

Nezbytným prvkem biologické evoluce je proměnlivost. Prvotní a hlavní příčinou proměnlivosti u organismů jsou dědičné změny v sekvenci nukleové kyseliny - mutace. Dnes jsou mutační procesy považovány za nutnou, i když ne postačující součást biologické evoluce. Mutace dodávají surový genetický materiál, který je k dispozici přírodnímu výběru a dalším evolučním silám. Bez mutací by nebyly rozdíly mezi geny a přírodní výběr by tak neměl

z čeho vybírat. Následkem toho by se evoluce zastavila, nedošlo by ke vzniku rozmanitých forem života, organismy by nemohly reagovat na změny prostředí a nejspíše by vyhynuly.

Jako výchozí materiál pro vznik nového genu slouží zpravidla některý starý gen. Nahromaděním mutací, a za vhodného selekčního tlaku, může být takový gen změněn na gen nový s funkcí do té doby neznámou. Zdálo by se tedy, s nadsázkou řečeno, že kdykoli organismus „pocítí potřebu“ nového genu, stačí mu pozměnit některý stávající gen a získat tak požadovanou novou vlastnost. Tato představa v sobě nicméně ukrývá značnou potíž. Rozvoj nové biologické funkce genu totiž nutně provázejí (vyvolávají) změny v nukleotidovém řetězci, které zároveň vedou k postupnému zhoršování funkce staré. Pozměněný gen tak už není schopen dále vykonávat původní funkci, což může mít za následek ztrátu či výrazné snížení biologické zdatnosti nositele takového genu. Je tedy zřejmé, že přírodní výběr nepřipustí trvalý výskyt mutací postihujících funkčně kritická místa genu, který řídí některý pro organismus životně důležitý produkt. Na druhou stranu, jenom akumulací takových mutací může gen změnit svůj základní charakter a stát se genem novým (Ohno 1970). Elegantní řešení tohoto problému přináší mechanismus genové duplikace. Zdvojením genu se vytvoří nadbytečná kopie, která tak unikne z dosahu tlaku přírodního výběru a protože je mimo kontrolu, může akumulovat předtím zakázané mutace a dát vzniknout novému genu s doposud neznámou biologickou funkcí (Ohno 1970). Velké evoluční skoky jako je vývoj mnohobuněčných organismů z jednobuněčných, zrod obratlovců, savců a konečně i člověka si lze jen stěží představit bez vytvoření nových genů, které tyto rozsáhlé evoluční změny vyvolávají. A jenom z takového genu, který se nadbytečně zmnoží, může vzejít gen nový.

V následujících kapitolách se pokusím shrnout naše dosavadní poznání tohoto vůdčího principu molekulární evoluce, přičemž se zaměřím hlavně na mechanismy genové duplikace a na evoluční osud duplikovaných genů.

2. Krátký pohled do historie

Práce zabývající se evolučním významem duplikace genů mají dlouhou historii. Vzhledem ke své všudypřítomnosti a základní roli v molekulární evoluci, genové duplikace byly a jsou přitažlivým objektem studia pro vědce celého světa. Následuje stručný přehled několika málo základních myšlenek zajímavých, mimo jiného, tím, že značně předběhly svou dobu.

Už ve 30. letech minulého století, v době nástupu teorie syntézy, rozpoznali někteří biologové evoluční důležitost zdvojování genů. Haldane (1932) si jako jeden z prvních uvědomil, že duplikace mohou být prospěšné, protože vytvářejí nadbytečné kopie, které mohou být pozměněny, aniž by to ohrožovalo biologickou zdatnost organismu. Haldane (1933) byl také mezi prvními, kdo upozornili na skutečnost, že mutace mohou časem eliminovat jednu kopii duplikovaného páru z genomu. Dále učinil důležité zjištění, že duplikace následovaná inaktivací původního (rodičovského) genu může změnit vazbu mezi aktivními geny.

O dva roky později, zaujat Haldaneovou myšlenkou, představil Fisher (1935) první populačně-genetický model zabývající se evolučním osudem duplikovaných genů. Model, který předpokládal nekonečnou velikost populace a bral v úvahu zpětné mutace, ukázal, že nefunkční kopie (letální v případě jejího výskytu v obou lokusech) může mít v populaci stále významnou frekvenci, protože její výskyt bude krytý přítomností druhé funkční kopie.

V roce 1938 přišel Serebrovsky (1938) s myšlenkou, že duplikované geny mohou uniknout z dosahu tlaku přírodního výběru. Serebrovsky také uvažoval o možnosti, kdy budou změněny obě kopie duplikovaného páru. Z pozorování genů na X chromosomu octomilky (*Drosophila*) Serebrovsky vyvodil, že jeden gen může ovlivňovat četné vlastnosti organismu a po duplikaci mohou být tyto mnohočetné funkce původního genu rozděleny mezi duplikované kopie.

Zhruba ve stejné době vzácná tandemová duplikace, která má za následek extrémní redukci očí u mutantů octomilky obecné (*Drosophila melanogaster*), ukázala na spojitost mezi genovou duplikací a změnami morfologie (Bridges 1936).

Ze 40. let minulého století pocházejí první hlasy dávající dohromady zdvojování genů a vývoj komplexity organismů. Evoluce často vyžaduje přidání dalšího genetického materiálu: více komplexní organismy potřebují větší genomy. To si mezi prvními uvědomili Gulick (1944) a Metz (1947) a poukázali tak na nezastupitelnou úlohu, kterou sehrálo zmnožování genů v evoluci složitosti, komplexity živých systémů.

Duplikace genu (nebo vyšší genetické jednotky) a následná prospěšná mutace v jedné takto vzniklé kopii, to je ve stručnosti schéma zrodu nových genů, které navrhl Lewis (1951) v polovině 20. století.

Ve stejném roce došel Stephen (1951) k závěru, že pouhé alelické mutace nemohou vyvolat rozsáhlé evoluční změny a jediná cesta jak jich dosáhnout, je vytvoření nových genových lokusů zdvojením a funkčním rozrůzněním už existujících genů.

S ucelenou teorií, kterou představil širší veřejnosti ve své knize „Evoluce genovou duplikací“, přišel v roce 1970 genetik japonského původu Susumo Ohno (Ohno 1970). Ohno uvedl genovou duplikaci jako nejdůležitější samostatný činitel evoluce a prohlásil duplikaci genů za její hlavní hnací sílu. Teorie založená na evoluci obratlovců, u kterých Ohno předpokládal dvě kola tetraploidizací celého genomu, ve své době vzbudila značný ohlas, ale zdaleka ne všechny reakce byly příznivé. Chvillemi se dokonce zdálo, že upadne v zapomnění, tak jako mnohé jiné před ní. Snad definitivně dala Ohnovi za pravdu až záplava genomických dat všeho druhu, která přišla na přelomu tisíciletí. Dobře to odráží počet citací Ohnových prací z poslední doby, který je mnohonásobně vyšší než v letech předešlých.

3. Genové rodiny

Se znalostí sekvencí celých genomů některých organismů (Lander et al. 2001, Waterston et al. 2002) je dnes více než kdy jindy zřejmé, že mezi geny existují příbuzenské vztahy, které podobně jako u organismů odrážejí jejich evoluci ze společného předka. Na základě těchto vztahů je můžeme hierarchicky uspořádat do přirozených skupin vzájemně si příbuzných genů, které nazýváme genové rodiny či nadrodiny. Genovou rodinu můžeme tedy například definovat jako skupinu genů pocházejících ze společného předka, jejíž příslušníci si udržují podobné sekvence a mají podobné funkce (Dayhoff 1976). Členy určité genové rodiny nacházející se spolu v jednom genomu v důsledku duplikace a divergence jednoho původního genu označujeme někdy jako paralogní geny (paralogy), zatímco jako ortologní geny (ortology) nazýváme geny v různých genomech, které vznikly v důsledku rozpadu jednoho mateřského druhu na dva druhy dceřiné (Dayhoff 1976). Zjednodušeně řečeno, paralogy jsou výsledkem duplikace a ortology speciace.

Genomy organismů nejsou statické, neměnné, ba právě naopak jsou vysoce dynamické. Velikost genových rodin se v průběhu evoluce značně mění, přičemž počet funkčních genů v jednotlivých liniích vyplývá z dlouhodobé interakce mezi duplikacemi genů na jedné straně a jejich ztrátami na straně druhé. Některé rodiny tvoří jenom pár genů, zatímco další jich obsahují stovky jako například jedna z rodin pro chemoreceptory u háďátka obecného (*Caenorhabditis elegans*) mající 242 členů nebo rodina genů kódujících proteolytické enzymy zvané trypsiny u octomily obecné (*Drosophila melanogaster*), která obsahuje 111 genů (Gu et al. 2002). V genomu savců je tou vůbec největší nadrodina genů pro čichové receptory, která má okolo 1000 členů, nicméně zde existují výrazné mezidruhové rozdíly v jejich počtu. Přes 1000 genů pro čichové receptory mají takové druhy jako myš, potkan či vačice, kdežto na opačném konci spektra stojí primáti včetně člověka nebo zástupce ptakořitných ptakopysk, u kterých nalezneme méně než 400 těchto genů (Niimura a Nei 2007).

Obdobou genových rodin jsou rodiny či nadrodiny proteinů. Jako příklad nám poslouží starobylá nadrodina proteinů zvaných lipokaliny, které můžeme najít u celé řady organismů od bakterií až po obratlovce. Lipokaliny jsou nevelké mimobuněčné proteiny složené přibližně z 200 aminokyselinových zbytků, které vážou a transportují malé hydrofóbní molekuly (Flower 1996). Charakteristickým znakem těchto proteinů jsou v rámci rodiny velké rozdíly v aminokyselinovém řetězci (identita sekvencí může být dokonce menší než 20 %) a

navzdory tomu vysoce konzervovaná terciární struktura (Flower 1995). Dalším typickým rysem lipokalinů je jejich schopnost tvořit oligomery, což jsou makromolekuly ležící svou velikostí a počtem monomerních jednotek na rozhraní mezi nízkomolekulární sloučeninou a polymerem. Právě takovým oligomerem, konkrétně dimerem, jsou odorant vázající proteiny (odorant-binding proteins, dále jen „OBP“), u kterých se na malou chvíli zastavíme, neboť se jimi budu zabývat ve svém budoucím magisterském studiu. OBP jsou malé rozpustitelné proteiny, které jsou syntetizovány ve velkém množství v blízkosti čichových neuronů, a jejichž hlavním úkolem je vázat a přepravovat lipofilní odoranty ve vodném prostředí nosní sliznice savců a jiných obratlovců (Pelosi 1994). Právě schopnost vázat různé sloučeniny stála za jejich objevením. První OBP byl izolován z nosní sliznice hovězího skotu pomocí silného odorantu 2-isobutyl-3-methoxypyrazinu (Pelosi et al. 1982). OBP se vyskytují také u hmyzu, kde sice plní podobnou funkci, nicméně strukturně se velmi liší od svých protějšků u obratlovců. Přestože OBP známe přes 20 let a víme toho poměrně dost o jejich struktuře a molekulární podstatě, máme velmi málo informací o jejich expresi a ještě méně o funkcích, které v organismu vykonávají. Zatím to jediné, čím si můžeme být opravdu jisti, je jejich schopnost vázat rozličné látky jako jsou feromony, ale dá se očekávat, že tyto proteiny plní v organismu celou škálu dalších úkolů. Mohou například rozlišovat mezi jednotlivými odoranty a působit tak jako periferní filtr nebo odstraňovat nežádoucí a toxické sloučeniny (Steinbrecht 1998). S funkcí OBP rovněž úzce souvisí jejich exprese, která je podle všeho tkáňově specifická, liší se mezi jednotlivými členy této rodiny a odvíjí se od pohlaví a věku konkrétního zvířete. Dřívější práce zkoumající expresi OBP byly do jisté míry limitovány méně přesnými laboratorními nástroji. V dnešní době máme k dispozici daleko vnímavější metody jako jsou DNA čipy nebo qRT-PCR (real-time reverse-transkription PCR), což je vysoce citlivá technika mRNA detekce, která umožňuje namnožit RNA dokonce z jediné buňky. Právě tuto metodu spolu se systémem prob využívá naše laboratoř a díky ní získaná první data jsou velmi zajímavá, neboť ukazují na expresi OBP v takových tkáních jako je kůže, prostata, uterus nebo čichový lalok v mozku. Naším laboratorním objektem je myš (*Mus musculus*), u které v současnosti známe osm OBP a dá se předpokládat, že zanedlouho budou objeveny další. Studium exprese, funkcí a evoluce OBP představuje dosud málo probádanou a nesmírně zajímavou oblast výzkumu, kde odpovědi vyvolávají pouze další otázky.

4. Mechanismy a typy genové duplikace

Mechanismy genové duplikace můžeme rozdělit podle dvou kritérií. Za prvé podle velikosti duplikovaného úseku a za druhé podle toho, jestli vyžadují prostředníka v podobě RNA. V zásadě rozlišujeme následující typy duplikací: úsekové duplikace, retrotranspozice a duplikace celého genomu nebo jednotlivých chromosomů.

4.1 Úsekové duplikace

Úsekové duplikace, které například v genomu člověka tvoří nejméně 5 % všech sekvencí DNA (Lander et al. 2001), můžeme opět rozdělit na dva typy. Zmnožený úsek DNA buď může zůstat v místě svého vzniku nebo může být přemístěn do zcela odlišné oblasti genomu. V prvním případě na sebe jednotlivé duplikované úseky bezprostředně navazují a říkáme o nich, že jsou tandemově duplikovány. Tento druh duplikací vzniká většinou v důsledku nerovnoměrného crossing-overu mezi dvěma homologními chromosomy během meiózy. Výsledkem takové nereciproké rekombinace je duplikace určitého úseku DNA na jednom chromosomu a jeho delece na chromosomu druhém (Flegr 2005). V závislosti na místě překřížení, může duplikovaná oblast zahrnovat části genů, celé geny nebo několik genů. K duplikacím může rovněž vést nerovnoměrná výměna mezi dvěma chromatidami téhož chromosomu. Právě tandemové duplikace daly vzniknout jedněm z nejznámějších genových rodin včetně globinové nadrodiny (Shen et al. 1981) nebo rodiny Hox-genů (Holland a Takahashi 2005).

Na evoluci nových genů se též významnou měrou podílejí duplikace krátkých úseků DNA, čemuž výrazně napomáhá modulární charakter dnešních genů. Kódující sekvence jsou u eukaryontních organismů uspořádány do sérií relativně krátkých exonů, které jsou odděleny dlouhými nekódujícími introny. K tandemovým duplikacím exonů dochází nerovnoměrným překřížením kdekoli v oblasti intronů. Protože jsou exony jenom částí genu, nejenže nemohou po duplikaci vykonávat žádnou nezávislou funkci, navíc jejich přítomnost může mít negativní vliv na vlastnosti genového produktu. Mechanismus, který duplikovaným exonům umožní získat potřebnou funkční nadbytečnost, a zároveň zamezí případným škodlivým účinkům na strukturu transkriptu/proteinu, je zvláštní typ alternativního sestřihu (Letunic et al. 2002). Alternativní sestřih pre-mRNA je ústředním mechanismem genetické regulace u vyšších eukaryot a hlavním zdrojem různorodosti proteinů. Primární transkript může být totiž

sestřížen různými způsoby. V některém sestřihu určitý exon zůstane, v jiném je naopak vystřížen, díky čemuž může jeden gen syntetizovat několik různých, funkčně odlišných proteinů (Black 2003). Působením alternativního sestřihu na dva tandemově duplikované exony dojde k obdobné situaci, jakou známe u zdvojování celých genů: jeden exon zajišťuje původní funkci, zatímco druhý se vymaní z tlaku přírodního výběru a může vcelku volně hromadit mutace. Duplikace exonů představují pro geny důležitou možnost rozšíření jejich stávající funkce (Babushok et al. 2007, Moore et al. 2008) a navíc se zdá, že jsou až nečekaně běžným jevem molekulární evoluce (Letunic et al. 2002).

Ne vždy spolu zmnožené úseky DNA těsně sousedí, někdy mohou být kopie genů nebo jejich fragmentů rozptýleny po celém genomu. Duplikované lokusy mohou být tudíž od sebe značně vzdáleny a díky jejich větší fyzické separaci může častěji dojít k vytvoření nového vzoru exprese. Výkonnými přenašeči genů nebo jejich částí jsou transponovatelné elementy (transpozony). Transpozony jsou úseky DNA, které se mohou v rámci genomu přemisťovat z jednoho místa na druhé. Většina DNA transpozonů, o retrotranspozonech bude řeč později, se v rámci hostitelské DNA pohybuje mechanismem „vystřížení a vložení“ (cut-and-paste), což jim umožňuje speciální enzymy zvané transponázy, které si sami kódují (Miskey et al. 2004). Tyto mobilní genetické elementy jsou široce rozšířené ve všech organismech od bakterií až po člověka a tvoří značnou část genomu eukaryot (Kidwell a Lisch 2001, Lander et al. 2001, Waterston et al. 2002).

Téměř polovina genomu savců je tvořena různými typy repetitivních sekvencí DNA (Lander et al. 2001, Waterston et al. 2002), jejichž přítomnost má dalekosáhlé důsledky na strukturu celého genomu. Většinu z nich tvoří právě transponovatelné elementy. Členové rodiny Alu-sequencí patřící mezi SINE (Short Interspersed Nucleotide Elements) jsou 300 nukleotidů dlouhé mobilní genetické elementy, které nacházíme pouze v genomu primátů (Houck et al 1979). Tyto repetice se zhusta vyskytují na okrajích zmnožených úseků DNA, což poukazuje na jejich důležitou úlohu, kterou hrají při vzniku a šíření úsekových duplikací, a to jak uvnitř chromosomů tak mezi nimi (Bailey et al. 2003). Po svém zrodu před 60-40 milióny let se rodina Alu-sequencí rychle množila (Shen et al. 1991) a během této doby došlo rovněž k obrovskému nárůstu počtu duplikovaných úseků DNA, které dnes nacházíme v genomu lidí a jiných primátů (Bailey et al. 2003).

4.2 Retrotranspozice

Zvláštním typem transponovatelných elementů jsou retrotranspozony, které se po genomu pohybují pomocí RNA intermediátů. Retrotranspozony byly dosud popsány pouze u eukaryot, přičemž je obsahuje hlavně genom savců. Vůbec největší počet těchto sekvencí nacházíme v DNA lidí a ostatních primátů. Asi nejznámějším příkladem retrotransponu je právě lidský L1-transponovatelný element neboli LINE-1, který tvoří přibližně 17 % celkové DNA genomu člověka (Ostertag a Kazazian 2001). Ačkoliv většina dnešních L1-retrotransponů je neaktivních, některé jsou ještě schopné retrotranspozice (Sassaman et al 1997). DNA LINE-1 je nejprve transkribována do RNA, potom se pomocí enzymu reverzní transkriptázy opět vytvoří kopie DNA, která se následně může integrovat do dalšího místa v genomu (Ostertag a Kazazian 2001). Reverzní transkriptáza je kódována samotným LINE-1, stejně jako další nástroje potřebné k retrotranspozici, což ho činí soběstačným a na hostitelské DNA téměř nezávislým. Takové retrotranspozony označujeme jako autonomní, LINE-1 patří konkrétně mezi tzv. autonomní non-LTR (long terminal repeat) retrotranspozony (Ostertag a Kazazian 2001). Z předchozího vyplývá, že retrotranspozice má vždy duplikativní charakter. Protože primární transkript podléhá posttranskripčním úpravám jako je vystřížení intronů nebo napojení ocasu – polyadenylace, kopie DNA vzešlé z reverzní transkripce neobsahují introny, postrádají původní promotor a následkem přemístění na jiné místo v genomu se také mohou velmi vzdálit od své regulátorové oblasti. Po svém vzniku tudíž nejsou ve většině případů transkribovány, v důsledku čehož začnou hromadit mutace a stanou se z nich pseudogeny, které se označují jako „procesované“ (processed pseudogenes) nebo retropseudogeny (Zhang et al. 2003, Zhang et al. 2004). Ne vždy ztratí retropseudogeny svou funkci, ve skutečnosti až třetina z nich zůstává funkční a je přepisována do RNA, většina z nich výlučně ve varlatech (McCarrey a Thomas 1987, Harrison et al. 2005). V zásadě existují dvě cesty jak toho mohou dosáhnout: transkript může vzácně nést původní promotor nebo se po reverzní transkripci náhodou vloží do místa, které už regulační sekvenci obsahuje (McCarrey a Thomas 1987, Mighell et al. 2000). Takové funkční pseudogeny se nazývají retrogeny.

L-1 retrotranspozony jsou zdrojem genetické rozmanitosti a zásadním způsobem ovlivňují strukturu lidského genomu (Ostertag a Kazazian 2001). V první řadě značně přispívají k samotnému rozšíření genomu, ať už vlastní retrotranspozicí nebo poskytnutím nezbytných nástrojů pro retrotranspozici ostatním mobilním genetickým elementům jako jsou například Alu-sekvence. Dále přesouvají ostatní sekvence DNA z jedné oblasti genomu do druhé a

rovněž ovlivňují řadou mechanismů genovou expresi. Na rozdíl od úsekových duplikací popsaných v předchozí kapitole, které se většinou shlukují v určitých oblastech genomu, retrotranspozice jsou nahodilé, a co se týče místa vložení nerespektují žádná pravidla, což může mít velmi negativní dopady na hostitelskou DNA. Začlenění retrotranspozonu přímo do kódující sekvence genu zapříčiní zpravidla ztrátu jeho funkce. Inzerce LINE-1 způsobují u člověka řadu vážných onemocnění jako je například hemofilie A (Kazazian et al. 1988). Jinou příčinou lidských nemocí mohou být naopak delece určitých úseků DNA vyvolané nerovnoměrnou rekombinací mezi LINE-1 (Burwinkel a Kilimann 1998). Na druhou stranu transponovatelné elementy významnou měrou přispívají k evoluci nových genů a jejich funkcí. Některé z nich obsahují a přenášejí sekvence hostitelské DNA jako jsou exony nebo promotory, a když se přemístí do blízkosti genu nebo se vloží přímo do něj, mohou způsobit odlišnou úroveň jeho exprese (Kidwell a Lisch 2001). Patrně evolučně nejdůležitější jsou inzerce transponovatelných elementů do regulačních oblastí. Ty mohou způsobit poškození takové oblasti nebo naopak přidáním krátkých regulačních sekvencí změnu její genové exprese (Ostertag a Kazazian 2001).

Mezi lidské retrotranspozony patří také už zmíněná rodina Alu-sekvencí, která na rozdíl od LINE-1 patří mezi neautonomní retrotranspozony. Ačkoliv aparát potřebný k retrotranspozici Alu-sekvencím mohou poskytnout právě L-1 retrotranspozony, dochází k ní zřídka, protože pouze některé z Alu-sekvencí mohou být ještě i dnes přepisovány do RNA (Shaikh et al. 1997). Předpokládá se, že všichni členové rodiny Alu-sekvencí vzešly z jednoho nebo několika prapůvodních genů (master genes), které daly vzniknout jednotlivým podrodinám (Shen et al. 1991). V současnosti existuje více jak 1 000 000 kopií těchto sekvencí na haploidní genom, což představuje přibližně 10 % lidské DNA (Smit a Riggs 1996). Tyto retrotranspozony jsou patrně odpovědné za velký nárůst počtu duplikovaných úseků DNA, které se objevily v genomu primátů v jejich nedávné historii (Bailey et al. 2003).

4.3 Polyploidie: Duplikace celého genomu

Duplikace celého genomu představuje účinný způsob jak najednou vytvořit ohromné množství nadbytečného genetického materiálu, který je poté volně k dispozici molekulární evoluci. Současné zmnožení všech genů je jedinečnou příležitostí pro vznik tak zásadních evolučních novinek jakou je například kostra obratlovců (Zhang a Cohn 2008). Další výhodou takové duplikace je, že zdvojení každého strukturního genu provází i duplikace regulační oblasti, která jeho aktivitu řídí. Také odpadá potíž s poměrem dávek funkčně spřízněných

genů (Ohno 1970). I přes tyto značné výhody, které skýtá, je polyploidní evoluce v současné době mezi organismy poměrně vzácná. Polyploidní druhy nacházíme ve větším množství hlavně u kvetoucích rostlin včetně některých zemědělsky významných plodin (Masterson 1994, Leitch a Bennett 1997). V porovnání s rostlinnou říší, existuje u dnešních živočichů, v minulosti tomu mohlo být jinak, nepoměrně méně polyploidních druhů, ale i zde je najdeme, jak u bezobratlých (Viktorov 1997) tak obratlovců (Becak a Becak 1998, Jaillon et al. 2004,). Zarážející je zejména viditelná absence polyploidů u vyšších obratlovců jako jsou ptáci nebo savci. Obě tyto skupiny mají velmi diferencované pohlavní chromosomy a pevně ustálený mechanismus chromosomální determinace pohlaví. Je vysoce pravděpodobné, že změna poměru autosomů a pohlavních chromosomů v důsledku polyploidizace narušuje právě tento mechanismus určení pohlaví (Müller 1925). Proto bylo velkým překvapením nalezení pouštního hlodavce *Tympanoctomys barrerae*, jehož genom vykazoval známky tetraploidizace (Gallardo et al. 1999). Osmák *T. barrerae* patřící do čeledi osmákovití (Octodontidae) má nejvyšší počet chromosomů ($2n = 102$) (Contreras et al. 1990) a největší známý genom ze všech savců (Gallardo et al. 1999). Ostatní členové této endemické čeledi Jižní Ameriky mají diploidní počet chromosomů od $2n = 38$ do $2n = 78$ (Honeycutt et al 2003). Nicméně se zdá, že sláva tetraploidního druhu savce netrvala příliš dlouho. Svartman et al. (2005) porovnali genom *T. barrerae* s genomem příbuzného druhu *Octodon degus* ($2n = 58$), který je pokládán za předka všech ostatních členů čeledi Octodontidae a zjistili, že *T. barrerae* má jenom dva chromosomy od každého typu a je tudíž nepochybně diploidní. Podle stejných autorů neobvyklou velikost genomu tohoto druhu hlodavce nejlépe vysvětluje velký nárůst počtu repetitivních sekvencí a jejich rozšíření do různých oblastí genomu.

Pokud v průběhu mezidruhového křížení dojde ke splynutí dvou diploidních zárodečných buněk nebo k přímé hybridizaci dvou tetraploidních druhů vznikají polyploidi, kteří se označují jako alopolyploidi. Ke vzniku alopolyploida může dojít rovněž zkřížením dvou diploidních druhů, po kterém následuje zdvojení chromosomů u F1 hybridu (Chen a Ni 2006). Buňky takového alotetraploida obsahují dva různé diploidní soubory genetického materiálu, každý od jednoho druhu. Během redukčního dělení se budou chromosomy přednostně párovat se svými homology v rámci svého souboru a budou tak tvořit dvě nezávislé sady bivalentů (Ohno 1970). Mechanismy vzniku autotetraploida jsou stejné jako u výše zmíněného alotetraploida, jenom s tím rozdílem, že u autotetraploida zdvojené chromosomové sady pocházejí pouze od jednoho druhu, a jsou tudíž identické. Autopolyploidii mohou způsobit četné buněčné anomálie jako jsou například chromosomové nondisjunkce, endoreplikace nebo narušení buněčného cyklu (Storchova a Pellman 2004). Genom nově vzniklého

autotetraploida obsahuje čtyři homologní chromosomy každého typu a ty v meióze vytvářejí jeden kvadrivalent. Podle tohoto znaku tedy můžeme snadno rozlišit čerstvého autotetraploida od nového allotetraploida (Ohno 1970). Mnoho autopolyloidních druhů nacházíme u dnešních kvetoucích rostlin.

Během evoluce projde mnoho polyploidů procesem zvaným diploidisace. Aby stav polyploidie mohl přispět k vytvoření nových genových lokusů, musí se původně čtyři (v případě tetraploida) homologní chromosomy funkčně rozrůznit, což povede k obnovení diploidního stavu (Ohno 1970). Tento proces provázejí velké ztráty genů nebo dokonce celých chromosomů, v důsledku čehož se mění celková struktura genomu. Například analýza genomů *Saccharomyces cerevisiae* a příbuzných druhů zcela jasně ukázala, že tato kvasinka vznikla kompletní duplikací osmi chromosomů, které obsahoval genom jejího předka, po níž následovala ztráta téměř 90 % všech duplikovaných genů a obnovení normální ploidie (Kellis et al. 2004).

I když je většina duplikací menšího rozsahu, duplikace celého genomu byla hlášena u řady živočichů a rostlin, včetně *Arabidopsis thaliana* (Blanc a Wolfe 2004) nebo čtverzubce druhu *Tetraodon nigroviridis* (Jaillon et al. 2004). Je vysoce pravděpodobné, že duplikace celého genomu stála také u zrodu obratlovců. Množství důkazů, které nám poskytují fylogenetické a genomické analýzy, utvrzuje většinu biologů v přesvědčení, že v rané evoluci obratlovců proběhla dvě kola duplikací celého genomu (Dehal a Boore 2005). Myšlenka, kterou poprvé formuloval Ohno (1970), je dnes známá jako 2R-hypotéza (2R hypothesis). Tato hypotéza vzešla z pozorování skutečnosti, že některé rodiny genů jako například rodina Hox-genů jsou u obratlovců zastoupeny čtyřikrát, zatímco u jiných živočichů jen jednou (Lemons a McGinnis 2006). Silnou podporu této hypotéze přinesla sekvenace genomu kopinatce a jeho porovnání s genomem obratlovců (Putnam et al. 2008). U kopinatce si Hox-geny zachovaly archaickou strukturu jediného shluku genů, zatímco většina obratlovců má čtyři shluky těchto genů, jak odpovídá dvěma kolům tetraploidizace genomu (Putnam et al. 2008). V některých liniích kostnatých ryb skupiny Teleostei následovalo ještě další kolo duplikace celého genomu, takže u některých druhů nalézáme 6 až 7 shluků Hox-genů (Amores et al. 1998). Ačkoliv je stále jasnější, že genom dnešních obratlovců byl skutečně nejméně dvakrát celý zmnožen (nebo alespoň jeho podstatná část), přesný mechanismus těchto změn je neznámý. Nevíme zda k tomu došlo alo- nebo auto-tetraploidizací, případně jejich vzájemnou kombinací, ani v jakém časovém sledu po sobě tyto události následovaly. Navíc jak ukázal například Hughes (1999), ne všechny důkazy hovoří ve prospěch 2-R hypotézy. Čtyřnásobné

zastoupení některých genových rodin v genomu obratlovců může být rovněž výsledkem duplikcí menšího rozsahu zahrnujících jednotlivé geny nebo skupiny genů.

Další nevyřešenou otázkou zůstává, jestli těsně po duplikaci celého genomu následuje jeho rozsáhlá přestavba či nikoliv. Předpokládá se, že zmnožení všech chromosomových sad vyvolá nestabilitu genomu, v důsledku čehož dojde k jeho přeuspořádání. (Song et al. 1995, Pontes et al 2004, Semon a Wolfe 2007). K poněkud jinému závěru dospěli Hufton et al. (2008), kteří nenašli důkazy toho, že by dvě po sobě jdoucí kola tetraploidizace genomu, která se odehrála v počátcích vývoje obratlovců, provázelo významné přeskupení genů.

Jestliže jsou u diploidního organismu některé buňky/tkáně polyploidní, mluvíme o tzv. endopolyploidii (Edgar a Weaver 2001). Tento jev se často vyskytuje v buňkách, na které jsou kladeny vysoké nároky z hlediska syntézy určitých genových produktů. Proto není příliš překvapivé, že polyploidní buňky nacházíme například v játrech savců a jiných obratlovců. Zmnožení genetického materiálu dosahují buňky opakovanými koly replikace DNA bez následného rozdělení jádra a buňky (endomitóza). Důležitou úlohu sehrály v genetice a celé biologii polytenní chromosomy (Urata et al. 1995). Tyto ohromné chromosomy se běžně vyskytují u mnoha zástupců hmyzu z řádu dvoukřídlých (Diptera), nalézáme je zejména ve slinných žlázách larev, kde byly také u pakomára rodu *Chironomus* poprvé objeveny (Balbiani 1881). Polytenní chromosomy vznikají mnohonásobnou endomitózou, při níž nedochází k separaci sesterských chromatid, které zůstávají spojeny v oblasti centromery. Tyto chromosomy tak tvoří mnohočetné svazky vláken DNA, což larvám umožňuje produkovat velké množství proteinů zapojených do metamorfózy. Endopolyploidie není výjimkou ani u rostlin, u kterých můžeme nalézt polyploidní buňky například v endospermu či ve stonku (Galbraith et al. 1991).

4.4 Polysomie

Aneuploidie je změna počtu jednotlivých chromosomů, jejíž příčinou je meiotická nebo mitotická nondisjunkce. V prvním případě dochází k poruše oddělení páru chromosomů během jednoho ze dvou meiotických dělení, zatímco mitotická nondisjunkce je porucha separace sesterských chromatid jednoho z páru homologních chromosomů vedoucí k tomu, že obě chromatidy putují k témuž pólu vřetenka a vstupují tak do jedné dceřiné buňky. U diploidních organismů pak jedna dceřiná buňka získá tři homologní chromosomy (trisomie) a druhá pouze jeden (monosomie). Dojde-li k tomu v zárodečných buňkách, vznikají trisomická a monosomická individua.

Stav polysomie je u mnohých organismů, zvláště vyšších živočichů, vysoce škodlivý a proto se dá předpokládat, že tento způsob vytváření nadbytečných genových lokusů hrál v biologické evoluci, ve srovnání s ostatními typy duplikací, podružnou roli. Například u obratlovců jsou odchylky v počtu jednotlivých chromosomů letální nebo přinejmenším výrazně snižují biologickou zdatnost organismu. U člověka jsou, vedle trisomie pohlavních chromosomů, postnatálně slučitelné se životem pouze tři trisomie malých autosomů (13, 18, 21). Navíc postižení jedinci mají vážné malformace a často umírají záhy po narození (Lejeune et al. 1959, Edwards et al. 1960, Patau et al. 1960, Smith et al. 1960). Proto tento typ duplikací prakticky nemohl přispět k vývoji savců a jiných obratlovců.

Na rozdíl od diploidních druhů mohou být změny počtu jednotlivých chromosomů tolerovány u tetraploidních druhů. Monosomie může být užitečná při stabilizaci genomu tetraploidních druhů, které procházejí procesem diploidisace (Ohno 1970). Zkřížením dvou tetraploidních jedinců, z nichž každý má o jeden chromosom méně, může dojít k eliminaci nepotřebných chromosomů, které neuspěly v dosažení funkční diversifikace.

5. Evoluční osud duplikovaných genů

Mutace, genetický drift a přírodní výběr jsou tři základní síly, které determinují osud duplikovaných genů. Interakce těchto sil rozhodují jak často budou oba geny po duplikaci fixovány v populaci a také určují jejich případný budoucí osud. Klasický model předpokládá, že po duplikaci je jedna kopie genu nadbytečná, tím unikne před tlakem přírodního výběru a může akumulovat dříve zákázané mutace (Ohno 1970). Ve většině případů se z jedné kopie stává pseudogen (Haldane 1933, Fisher 1935, Nei a Roychoudhury 1973). Druhou možností je zachování obou duplikovaných kopií v genomu. K tomu může dojít několika způsoby. Přítomnost mnohočetných kopií jednoho genu někdy poskytuje organismu výhodu vyšší úrovně jeho exprese. Jinými slovy, obě duplikované kopie dále vykonávají stejnou funkci (Kondrashov a Kondrashov 2006). Nebo dojde v jedné kopii k fixaci prospěšné mutace a vzniku nové dříve neexistující funkce, zatímco druhá kopie vykonává dál funkci původní (Ohno 1970, Walsh 1995, Bergthorsson et al. 2007). Alternativní možností udržení obou kopií v populaci je rozdělení dvou (nebo více) odlišných funkcí původního genu mezi duplikované kopie. Každá kopie poté vykonává jednu funkci a vzájemně se tak doplňují (Force et al. 1999, Lynch a Force 2000, Lynch et al. 2001).

Základní otázka genové duplikace zní, jak rychle se v průběhu evoluce takové geny objevují. Lynch a Conery (2000) určili tuto rychlost s použitím sekvencí osmi eukaryotních organismů (*Homo sapiens*, *Mus musculus*, *Gallus gallus*, *Caenorhabditis elegans*, *Drosophila melanogaster*, *Arabidopsis thaliana*, *Oryza sativa* a *Saccharomyces cerevisiae*) zhruba na jednu duplikaci za gen za 100 miliónů let. Je nutné poznamenat, že Lynch a Conery brali v potaz (nemohli ostatně jinak) pouze takové genové páry, které byly po duplikaci fixovány v populaci a udržely se v genomu až do současnosti. S nadsázkou řečeno, aby se gen mohl objevit ve výpočtu, musel být zdvojen, fixován v populaci a následně přežít v genomu až do dnešní doby. Fixace nového genu v populaci je poměrně vzácná událost a to i v případě, že takový gen přináší organismu okamžitou selekční výhodu, navíc nejčastější osud, který gen po fixaci potká, je jeho inaktivace a přeměna v nefunkční pseudogen. Z těchto důvodů je rychlost odvozená Lynchem a Conerym až překvapivě vysoká. Jestliže by genom nějakého hypotetického organismu například obsahoval 20 000 genů, potom by každých 5000 let došlo k duplikaci některého z nich a k jeho rozšíření v genomu nebo-li každý milión let by takto vzniklo 200 genů (1 duplikace/gen/100 miliónů x 20 000 genů x 1 milión let). Polovina všech genů v genomu by téhož dosáhla během 50 miliónů let.

5.1 Ztráta funkce genu (nonfunkcionalizace)

Protože je jedna kopie genu po duplikaci nadbytečná, negativní mutace v jednom z duplikátů nebude mít ve většině případů vliv na biologickou zdatnost organismu a nebude odstraněna selekcí. A protože většina mutací postihujících organismus je škodlivých, obvyklý osud duplikovaného páru je inaktivace jedné kopie (Walsh 1995, Lynch a Conery 2000, Lynch et al. 2001). Z té se stává pseudogen, nefunkční sekvence DNA, která neřídí žádný genový produkt a nepodléhá evolučním omezením jako běžný funkční gen. Takový pseudogen poté setrvává v genomu až do té doby, než z něho bude zcela vymazán anebo zpětnou mutací nedojde k jeho aktivaci. Jinou možnost jeho aktivace nabízí genová konverze. Předpokládaná doba přeměny jedné kopie v nefunkční pseudogen je, vyjma populací s obrovskou efektivní velikostí, relativně krátká, dochází k ní během pár milionů let (Lynch a Force 2000).

Většina pseudogenů vzniká retrotranspozicí nebo genovou duplikací. Podle toho můžeme pseudogeny rozdělit do dvou typů. První představují „procesované pseudogeny“ (processed pseudogenes), které jsou výsledkem reverzní transkripce z RNA, po níž následuje začlenění takto vytvořené kopie DNA do některé oblasti genomu (Vanin 1985). Protože zralá funkční mRNA neobsahuje promotor, sekvence vzniklé reverzní transkripcí nejsou obvykle exprimovány a rychle hromadí mutace, v důsledku čehož se z nich ve většině případů stávají pseudogeny (Brosius 1999). Procesované pseudogeny jsou charakteristické absencí intronů, mají polyadeninový ocas, dále vykazují četné posunové mutace a obsahují předčasné stop kodony. Kvůli jejich původu se o nich někdy hovoří jako o retropseudogenech a uvažuje se o nich jako o zvláštním typu retrotranspozonů jako jsou Alu-sekvence nebo LINE (Long Interspersed Nuclear Elements). Patří sem většina známých savčích pseudogenů. Druhým typem jsou „duplikované pseudogeny“ (nonprocessed pseudogenes), které vznikají počáteční fixací škodlivé mutace v jedné kopii duplikovaného páru genů (Mighell et al. 2000). Tyto pseudogeny mají často zachovanou původní strukturu genu tvořenou exony a introny jakou nalézáme u jejich funkčních protějšků, byť někdy neúplnou, což je ostře odlišuje od retropseudogenů.

Pseudogenizace je široce rozšířený, nicméně zatím málo probádaný jev molekulární evoluce. Podle odhadů vzešlých ze sekvenace lidského genomu tvoří pseudogeny až 22 % všech genů v něm obsažených (Lander et al 2001). Tyto údaje mohou být ale značně podhodnocené, protože odlišit pseudogeny od funkčních genů je mnohdy obtížné nebo naopak jejich sekvence natolik divergovaly od sekvence rodičovského genu, že je v genomu už nedokážeme rozpoznat. Například podrobná analýza lidských chromosomů 21 a 22

odhalila, mimo 264 už známých, dalších 190 nových pseudogenů - na každé dva funkční geny připadal v této studii zhruba jeden pseudogen (Harrison et al. 2002). Na základě těchto výsledků odhadli autoři zmíněné práce celkový počet procesovaných pseudogenů v lidském genomu zhruba na 9000, přičemž dalších 10 000 připadlo na ostatní druhy pseudogenů. K podobným číslům dospěli Zhang et al. (2003), kteří našli více jak 8000 procesovaných pseudogenů a když zahrnuli do výpočtu i ostatní typy pseudogenů došli k číslu blízcímu se 20 000.

I přes zatím omezená data, která máme k dispozici, je zřejmé, že vysoký počet pseudogenů v genomu člověka není v živočišné říši ojedinělý. Analýza genomu háďátka obecného (*Caenorhabditis elegans*) identifikovala něco mezi jedním až dvěma tisíci pseudogeny – to je přibližně jeden pseudogen na každých osm funkčních genů (Harrison et al. 2001). Některé důkazy naznačují, že dokonce každý pátý popsáný gen háďátka může být pseudogenem (Mounsey et al 2002). Srovnávací studie myššího a lidského genomu odhalila v genomu myši okolo 5000 pseudogenů, z nichž odhadem 60 % vzniklo až poté, co se myšší a lidská linie rozdělily (Zhang et al. 2004).

U pseudogenů můžeme najít podobné rodiny jako u funkčních genů. Největší rodinu pseudogenů tvoří retropseudogeny odvozené od genů kódujících ribosomální proteiny. Srovnávací analýza provedená Balasubramaniamem et al. (2009) odhalila v genomu člověka, šimpanze, myši a potkana v tomto pořadí 1822, 1462, 2029 a 2848 těchto pseudogenů. Předchozí práce naznačovaly, že příčinou vysokého počtu těchto pseudogenů může být vyšší úroveň exprese jejich funkčních protějšků (Goncalves et al. 2000). Zdálo by se, že tato hypotéza nalezne podporu v práci Balasubramaniana a jeho kolegů, protože geny pro ribosomální proteiny bezesporu patří k těm nejvíce exprimovaným v celém genomu. Nicméně Balasubramaniam a jeho spolupracovníci takovou korelaci nenašli, z čehož lze usuzovat, že úroveň exprese nebude tím jediným hlavním faktorem, který určuje počet retropseudogenů.

Na pseudogeny nelze pohlížet jenom jako na vedlejší produkt evoluce. Význam pseudogenů zdaleka nespočívá jenom v jejich počtu, ale zdá se, že sloužily jako motor evolučních změn, zvláště během vývoje naší vlastní linie. Ačkoliv ještě není zcela jasné, zda proces pseudogenizace sám o sobě hraje v evoluci aktivní roli, některé důkazy tomu nasvědčují. Během své evoluce, poté co se linie lidí a šimpanzů před 5-7 milióny let rozdělila až po současnost, přišel člověk o celou řadu genů (Winter et al. 2001), přičemž se v některých případech mohlo jednat o aktivní reakci na měnící se podmínky prostředí. To vedlo Olsona (1999) k formulaci hypotézy známé jako „méně je více“ („less-is-more“ hypothesis). Podle Olsona jsou ztráty genů běžnou a častou evoluční odpovědí na změny prostředí a selekčního

tlaku. Protože mutace vedoucí ke ztrátě funkce genu jsou velmi početné, inaktivace genu může představovat nejrychlejší adaptivní odpověď na změnu selekce. Takový pseudogen může poté vyčkávat v genomu na další změny prostředí, které mohou vyvolat jeho zpětnou aktivaci. Podobný vzorec rozpínání a smršťování skutečně nacházíme u některých genových rodin včetně té u savců vůbec největší (Grus et al. 2005, Go 2006, Niimura a Nei 2007). Nadrodina genů pro čichové receptory prodělala u placentálních savců po jejich oddělení od ptakořitných a vačnatců pokaždé bouřlivý rozvoj, po kterém v jednotlivých liniích, podle funkčních požadavků každého druhu, následovaly ztráty těchto genů nebo naopak nárůst jejich počtu (Niimura a Nei 2007).

Pseudogeny mohou hrát důležitou roli i při vzniku nových druhů. Náhodná inaktivace jedné kopie duplikovaného páru ve dvou sesterkých taxonech může přispět k vytvoření reprodukčně-izolační bariéry (Lynch a Conery 2000). A v neposlední řadě podávají biologům méně zkreslený obraz o mutační rychlosti než sekvence pod vlivem selekce.

5.2 Zachování původní funkce genu

Potřebuje-li organismus větší množství určitého genového produktu, proteinu nebo RNA, zmnožení příslušného genu může tyto zvýšené nároky pokrýt. Ohno (1970) v takovém případě hovoří o genové duplikaci s cílem vytváření většího množství téhož produktu. Zesílení exprese určitého genu může organismus docílit několika způsoby, jednak trvalou přítomností duplikovaných kopií daného genu v genomu, dále jeho zvýšenou transkripcí a translací anebo jeho zmnožením v určitých etapách ontogenetického vývoje.

Některé úseky ontogenetického vývoje vyžadují značné kvantum určitého genového produktu, přičemž samotná zvýšená transkripce daného genu není schopna požadované množství produktu zajistit. Jedním mechanismem, kterým lze dosáhnout dostatečné úrovně genové exprese, je zmnožení potřebných genů vedoucí k nárůstu počtu templátů dostupných pro transkripci. Dnes už klasickým příkladem takové, vývojově regulované, genové duplikace jsou geny pro ribosomální RNA (rRNA) v oocytech obojživelníků.

Cytoplasma typické buňky obsahuje milióny ribosomů. Ribosomy jsou molekulové stroje, které vyrábějí proteiny. Jádra ribosomů jsou tvořena několika málo typy rRNA, která tvoří až 85 % celkové RNA přítomné v buňce. Geny pro rRNA se nacházejí v jadérku. U drápatky vodní (*Xenopus laevis*) vzniká v časných fázích oogenese přes 1000 extrachromosomálních jadérek a jádro oocyty obsahuje okolo 30 pg extrachromosomální rDNA (MacGregor 1968, Thiébaud 1979). Thiébaud (1979) spočítal, že jedno jadérko podle své velikosti nese od 500

do 11 000 tandemově duplikovaných kopií genů pro rRNA. Jediné jádro oocyty tak může obsahovat milióny kopií genů pro rRNA. Takové množství kopií umožňuje oocyty syntetizovat přibližně 300 000 ribosomů za vteřinu (Scheer 1973).

Amplifikace rDNA v brzkých fázích ontogenetického vývoje byla zjištěna i u hvězdice *Henricia hayashi* (Gaginskaya et al. 1988) nebo u nálevníka rodu *Tetrahymena* (Blomberg et al. 1997).

Důležitou adaptivní úlohu mají možná duplikace genů zapojených do imunitního systému (Bailey et al. 2002). Jak ukázala nedávná studie lidského genu CCL3L1 a jeho vztahu k HIV infekci, existuje souvislost mezi úsekovými duplikacemi vedoucími ke změně dóze genů odpovědných za imunitu a proměnlivostí fenotypové odpovědi na infekční nemoci (Gonzalez et al. 2005). CCL3L1 je nejúčinnější známý ligand pro CC chemokinový receptor 5 (CCR5), což je hlavní koreceptor pro virus HIV. Gonzalez et al. (2005) prokázali, že jedinci nesoucí více kopií genu CCL3L1 mají sníženou, oproti průměru v jejich populaci, vnímavost vůči HIV infekci a dochází u nich později k propuknutí nemoci AIDS.

Inkorporace mnohočetných kopií téhož genu do genomu s sebou nese otázku, jak si takové geny udrží stejnou funkci. Jedna možnost zachování vysoké identity sekvencí je prostřednictvím synchronizované evoluce (concerted evolution), kdy se členové genové rodiny v evoluci vyvíjejí jako jedna jednotka, která si vyměňuje genetickou informaci (Dover 1982). Mechanismem synchronizované evoluce je genová konverze, tedy nereciproká výměna genetické informace, při které sekvence jednoho úseku DNA slouží jako templát, podle kterého se přepíší sekvence jiných úseků. Druhou možností je silná negativní (= purifikující) selekce proti mutacím měnícím základní charakter genu (Piontkivska et al. 2002). Ačkoliv oba procesy vedou ke stejnému výsledku, jsou tu určité rozdíly, které nám je mohou pomoci rozlišit. Předpokládá se, že synchronizovaná evoluce má vliv na podobnost sekvencí mnohočetných kopií téhož genu hlavně uvnitř jednoho druhu, kdežto silná negativní selekce dokáže zachovat identitu sekvencí na vysoké úrovni i mezi vzdáleně příbuznými organismy jako jsou rostliny a živočichové (Piontkivska et al. 2002). Pokud na členy jedné genové rodiny působí pouze negativní selekce, nalezneme u nich vysoký počet synonymních mutací (Nei et al. 1997). Zatímco totiž působením genové konverze dojde k jejich odstranění, vůči selekci jsou víceméně imunní. Nicméně oba mechanismy se nezřídka vzájemně doplňují a rozpoznat je od sebe je obtížné, zvláště v případě pomalé rychlosti synchronizované evoluce.

Jiný problém existence mnohočetných kopií téhož genu představuje nerovnoměrný crossing-over, v jehož důsledku vznikají delece nebo přídavné duplikace, které postihují oblasti chromosomů tvořené tandemově duplikovanými geny (Ohno 1970). Jelikož přesné

spárování homologických úseků DNA ležících na dvou chromosomech je v tandemově duplikované oblasti obtížné, dochází k nelegitímnímu párování a výsledkem je přídatná duplikace na jednom chromosomu a naopak delece na chromosomu druhém. Výrazné delece snižují biologickou zdatnost organismu nebo jsou dokonce letální, takže jejich nositelé jsou odstraňováni přírodním výběrem. Nicméně i na první pohled škodlivé delece mohou být prospěšné, delece totiž může z genomu odstranit degenerované duplikáty, které působením mutací ztratily svou funkci. Následný nerovnoměrný crossing-over může obnovit původní rozsah duplikovaného segmentu, tentokrát ovšem tvořeného už vesměs funkčními kopiemi (Ohno 1970). Naproti tomu v duplikovaných úsecích dochází mnohem častěji k dalším duplikacím, což má za následek, že se v genomu organismu a následně i v genofondu druhu hromadí čím dál více duplikovaných úseků (Flegr 2005).

5.3 Vznik nové funkce genu (neofunkcionalizace)

Jednou z evolučně nejdůležitějších rolí genové duplikace je tvorba nových genů s dosud neznámou biologickou funkcí. Jak už bylo řečeno v úvodu, každý nový gen musel vzniknout z už existujícího genu. Zdvojením starého genu se vytvoří nadbytečná kopie, kterou přírodní výběr často ignoruje, v důsledku čehož má šanci hromadit předtím zakázané mutace a dát tak vzniknout novému genu (Ohno 1970). Protože druhá kopie stále vykonává původní, mnohdy pro organismus životně důležitou funkci, zrodem nového genu nedojde ke ztrátě či snížení biologické zdatnosti jeho nositele.

Prvním předpokladem vzniku nového genu z nadbytečné kopie genu starého je počáteční fixace několika neutrálních mutací, což genu umožní získat novou, byť ještě značně nedokonalou a slabou funkci. Druhým je přírodní výběr, jenž posléze podstatně urychlí fixaci prospěšných mutací, které povedou k rozvoji této nové funkce. Celý proces vzniku nové alely z jedné kopie duplikovaného páru genů je nicméně složitější a v cestě mu stojí několik závažných překážek. Předně musí zdvojený gen dosáhnout dostatečně vysoké frekvence v populaci a následně musí zůstat nedotčený do té doby, než získá prospěšnou mutaci a s ní novou biologickou funkci. Problémem je, že nadbytečná kopie je právě díky své nadbytečnosti selekčně neutrální a proto může snadno dojít k jejímu vymizení z populace působením genetického driftu nebo vlivem mutací k její inaktivaci a přeměně v nefunkční pseudogen. Geny ve většině případů pomáhá v populaci udržet přírodní výběr, ale ten jim na druhou stranu v důsledku tlaku na zachování původní funkce obvykle brání získat funkci novou. Nastává zde tedy rozpor označovaný jako Ohnovo dilema (Bergthorsson et al. 2007).

Tento problém má několik možných řešení, přičemž ve většině z nich hraje důležitou roli pozitivní selekce, která pomáhá udržet duplikované geny v populaci tak dlouho, dokud nezískají prospěšnou mutaci a s ní novou funkci. Přítomnost nadbytečných kopií v genomu může být žádoucí, protože chrání organismus před účinky škodlivých mutací (Clark 1994). Jak jsme si ukázali v předchozí kapitole, metabolické nároky organismu někdy vyžadují velké množství určitého genového produktu a zmnožení genu řídicího daný produkt může tyto požadavky pokrýt. Nicméně v tomto případě bude přírodní výběr opět bránit funkčnímu rozrůznění těchto mnohočetných kopií jednoho genu. Ohnovo dilema přivedlo některé vědce k myšlence, že novou funkci obdrží gen ještě před samotnou duplikací. Mnohočetné funkce takového genu jsou potom během doby uvolněné selekce rozděleny mezi vzniklé kopie (Piatigorsky a Wistow 1991). Zajímavým modelem řešícím Ohnovo dilema, jehož autory jsou Bergthorsson et al. (2007), je model známý pod zkratkou IAD (innovation-amplification-divergence). Model má tři kroky a v každém z nich hraje důležitou roli přírodní výběr. První krok vychází z předpokladu, že rodičovský gen kóduje protein mající v organismu na starost kromě své hlavní funkce ještě několik dalších funkcí, které ale v tu chvíli nejsou pro organismus životně důležité. Nicméně to se může z mnoha příčin rychle změnit, a jak se tato vedlejší funkce stává pro organismus čím dál cennější, přírodní výběr bude preferovat každé zvýšení její aktivity. K tomu může nejnázce dojít zmnožením původního genu, který bude poté selekcí ve prospěch postranní funkce udržován v populaci ve vysoké frekvenci. Druhý krok tedy představuje zmnožení genu a tím posledním je vznik nového genu z některé kopie genu původního. Dá se předpokládat, že stejný tlak přírodního výběru, který podporoval zmnožení původního genu, bude také zvýhodňovat další rozvoj postranní funkce v některé z mnoha přítomných kopií. Právě velký počet kopií značně zvyšuje šanci, že některá z nich získá prospěšné mutace, které povedou ke zdokonalení této požadované funkce. Jakmile se tak stane, vymaní se ostatní kopie z tlaku přírodního výběru, v důsledku čehož může dojít k jejich odstranění z populace genetickým driftem nebo působením mutací k jejich inaktivaci. Nakonec zůstane jedna kopie, která se zdokonalila natolik, že je schopna vykonávat novou funkci sama. Proces vylepšování nové funkce sice často doprovází současné zhoršování funkce původní, nicméně přírodní výběr starou funkci zachová alespoň v jedné z mnoha kopií rodičovského genu, takže v důsledku vzniku nové funkce nedojde ke ztrátě biologické zdatnosti organismu.

Ačkoliv se tedy může vznik nového genu jevit jako vysoce nepravděpodobná událost, známe celou řadu případů, kdy se tak skutečně stalo. ECP (eosinophil cationic protein) a EDN (eosinophil-derived neurotoxin) jsou dva toxické proteiny vyskytující se ve velkém množství

v granulích eosinofilů lidoopů a opic Starého světa, které podle všeho vznikly tandemovou duplikací rodičovského genu připomínajícího dnešní EDN po rozdělení linií opic Starého a Nového světa (Rosenberg et al. 1995). ECP má silný antibakteriální účinek, přičemž tuto schopnost postrádá jak EDN opic Starého světa tak příbuzný protein téhož jména u opic Nového světa, což jasně svědčí o tom, že ECP získal tuto funkci až po duplikaci. Jiným dobrým příkladem je evoluce trichromatického barevného vidění u primátů Starého a Nového světa (Kanwaljit et al. 1999) nebo evoluční osud receptoru NR3C4 pro androgeny u ryb skupiny Actinopterygii, který hraje klíčovou roli v determinaci pohlaví (Douard et al. 2008).

5.4 Rozdělení (pod)funkcí mezi vzniklé duplikáty (subfunkcionalizace)

Jak jsme si ukázali v minulých kapitolách, obvyklý osud, který potká duplikovaný pár genů je poměrně brzká inaktivace jednoho z nich a jeho přeměna v nefunkční pseudogen, zatímco druhý gen dál vykonává původní funkci. Vzácně může jedna kopie genu získat novou prospěšnou funkci, což povede k zachování obou genů z páru, protože pro organismus bude v tomto případě výhodné ponechat si je oba dva. Tomuto tradičnímu pohledu ale odporují empirická pozorování, která nám říkají, že v populaci se udržuje mnohem větší procento duplikovaných genů, než bychom měli na základě klasického modelu očekávat. To přimělo Force et al. (1999) k tomu, aby navrhli alternativní model, který by tento rozpor dokázal vysvětlit. Force a jeho spolupracovníci zaměřili svou pozornost na složité vztahy regulace eukaryotních genů a ukázali jak mohou škodlivé mutace v regulátorových oblastech přispět k zachování obou kopií duplikovaného páru genů v populaci. Současné strukturní geny mají často několik různých funkcí, přičemž každou z nich může kontrolovat jiný regulátorový gen (Piatigorsky a Wistow 1991). Jestliže duplikované geny následkem mutací v regulátorových oblastech přijdou o některou z těchto (pod)funkcí, musí se vzájemně doplnit tak, aby organismu zůstala zachována celá sada funkcí přítomná u rodičovského genu. Hlavním rozdílem tohoto DDC (duplication-degeneration-complementation) modelu oproti modelu klasickému je to, že v tomto případě škodlivé mutace spíše napomáhají než brání udržení duplikovaných genů v populaci. Protože negativní mutace jsou mnohem častější než mutace prospěšné, DDC model přináší hodnověrné vysvětlení existence velkého množství duplikovaných genů, které nacházíme v genomech současných organismů.

Příkladem takové „subfunkcionalizace“ mohou být dva geny kukuřice označované jako ZAG1 a ZMM2, které vznikly v důsledku alotetraploidizace zhruba před 11 milióny let. ZAG1 a ZMM2 jsou ortology jediného genu u houseníčku rolního (*Arabidopsis thaliana*) zvaného AGAMUS, který je exprimován v pestících a tyčinkách této rostliny během jejich vývoje. Naproti tomu ZAG1 je exprimován ve vysokém množství pouze v pestících, zatímco ZMM2 můžeme najít hlavně v tyčinkách, což svědčí o rozdělení původně jedné funkce rodičovského genu mezi tyto dva geny (Force et al. 1999).

Subfunkcionalizace nabízí významnou možnost jak odstranit evoluční omezení způsobená přítomností více funkcí u jednoho genu. Rozdělením mnohočetných funkcí rodičovského genu mezi duplikované kopie může eliminovat jeho pleiotropní účinek a tím otevřít evoluci dříve nepřístupnou cestu.

6. Závěr

Otázka původu nových genů je předmětem zájmu biologů každé doby, nehledě na to, jaká biologická disciplína zrovna udává směr a tempo dalšího vývoje. V procesu zrodu nových genů, a rovněž v celé historii života na Zemi, hraje důležitou roli mechanismus genové duplikace, který evoluci poskytuje tolik potřebný nový genetický materiál a bez něhož by byla schopnost organismů přizpůsobit se změnám prostředí vážně ohrožena. Je těžké si například představit, jak by mohl fungovat imunitní systém člověka a ostatních obratlovců bez přítomnosti stovek zmnožených genů. V současné době zároveň začíná vycházet najevo, že duplikace genů jsou až neobvykle častým jevem molekulární evoluce, přičemž rychlost duplikací se liší jak mezi geny v různých genomech tak mezi geny uvnitř genomu jednoho. Další evoluční osud duplikovaného genu se odvíjí od jeho funkce, exprese a v neposlední řadě od toho, v jakém organismu se nachází. Ve většině případů se z něho stává nefunkční pseudogen, zatímco vzácně může dát vzniknout novému genu s doposud neexistující funkcí. Duplikované kopie si také mohou mezi sebe rozdělit mnohočetné funkce rodičovského genu a vzájemně se tak doplňovat, díky čemuž se oba udrží v populaci. Pseudogeny mohou být aktivní odpovědí na měnící se podmínky prostředí a rovněž se mohou podílet na vzniku nových druhů vytvořením reprodukčně-izolační bariéry. Jednotlivé evoluční osudy jsou vzájemně propojeny, opakovaná kola duplikací na jedné straně a ztrát genů na straně druhé mohou uvolnit evoluční omezení a umožnit tak vznik nových genů. Nicméně stále zůstává spousta nezodpovězených otázek jako je například role přírodního výběru při fixaci duplikovaných genů v populaci a velkou záhadou rovněž zůstávají pseudogeny a jejich aktivní role v evoluci. Dnešní biologie se znalostí sekvencí celých genomů některých organismů představuje jedinečnou příležitost jak alespoň některým z těchto otázek porozumět

7. Seznam použité literatury

- Amores, A., Force, A., Yan, Y. L., Joly, L., Amemiya, C., Fritz, A., Ho, R. K., Langeland, J., Prince, V., Wang, Y. L. et al. (1998):** Zebrafish hox clusters and vertebrate genome evolution. *Science* 282: 1711-1714.
- Babushok, D. V., Ostertag, E. M. a Kazazian Jr., H. H. (2007):** Current topics in genome evolution: molecular mechanisms of new gene formation. *Cell. Mol. Life Sci.* 64: 542-554.
- Bailey, J. A., Gu, Z., Clark, R. A., Reinert, K., Samonte, R. V., Schwartz, S., Adams, M. D., Myers, E., W., Li, P. W. a Eichler, E. E. (2002):** Recent segmental duplications in the human genome. *Science* 297: 1003-1007.
- Bailey, J. A., Liu, G. a Eichler, E. E. (2003):** An Alu Transposition Model for the Origin and Expansion of Human Segmental Duplications. *Am. J. Hum. Genet.* 73: 823-834.
- Balasubramanian, S., Zheng, D., Liu, Y. J., Fang, G., Frankish, A., Carriero, N., Robiolotto, R., Cayting P. a Gerstein, M. (2009):** Comparative analysis of processed ribosomal protein pseudogenes in four mammalian genomes. *Genome Biology* 10: R2.
- Balbani, E. G. (1881):** Sur la structure de noyau des cellules salivaires chez les larves de *Chironomus*. *Zool. Anz.* 4: 637-641. 662–666.
- Becak, M. L. a Becak, W. (1998):** Evolution by polyploidy in Amphibia: new insights. *Cytogenet. Cell Genet.* 80:28-33.
- Bergthorsson, U., Andersson, D. I. a Roth, J. R. (2007):** Ohno's dilemma: Evolution of new genes under continuous selection. *PNAS* 104(43): 17004-17009.
- Black, D. L. (2003):** Mechanisms of Alternative Pre-Messenger RNA Splicing. *Annu. Rev. Biochem.* 27: 27-48.
- Blanc a Wolfe (2004):** Functional Divergence of Duplicated Genes Formed by Polyploidy during Arabidopsis Evolution. *The Plant Cell* 16: 1679-1691.
- Blomberg, P., Randolph, C., Yao, C. H., Yao, M. C. (1997):** Regulatory sequences for the amplification and replication of the ribosomal DNA minichromosome in *Tetrahymena thermophila*. *Mol. Cell. Biol.* 17: 7237-7247.
- Bridges, C. B. (1936):** The bar „gene“ a duplication. *Science* 83: 210-11.
- Brosius, J. (1999):** RNAs from all categories generate retrosequences that may be exapted as novel genes or regulatory elements. *Gene* 238: 115-134.
- Burwinkel B.a Kilimann M. W. (1998):** Unequal homologous recombination between LINE-1 elements as a mutational mechanism in human genetic disease. *Journal of Molecular Biology* 277 (3): 513-517.
- Clark, A. G. (1994):** Invasion and maintenance of a gene duplication. *Proc. Natl. Acad. Sci. USA* 91: 2950–2954.
- Contreras, L., Torres-Murra, J. a Spotorno, A. (1990):** The largest known chromosome number for a mammal, in a South American desert rodent, *Experientia* 46: 506-508.

- Dayhoff, M. O.** (1976): Origin and evolution of protein superfamilies. *Fed. Proc.* 35: 2132–2138.
- Dehal, P. a Boore, J. L.** (2005): Two rounds of whole genome duplication in the ancestral vertebrate. *Plos Biology* 3: e314.
- Dover, G.** (1982): Molecular drive: a cohesive mode of species evolution. *Nature* 299(5879): 111-117.
- Edgar, B. A. a Orr-Weaver, T. L.** (2001): Endoreplication cell cycles: more for less. *Cell* 105: 297–306.
- Edwards, J. H., Harnden, G. G., Cameron, A. H. et al.** (1960): A new trisomic syndrome. *Lancet* 1: 787-790.
- Fisher, R. A.** (1935): The sheltering of lethals. *The American Naturalist* 69: 446-455.
- Flegr, J.** (2005): *Evoluční biologie*. Academia, Praha.
- Flower, D. R.** (1995): Multiple molecular recognition properties of the lipocalin protein family. *J. Mol. Recogn.* 8: 185-195.
- Flower, D. R.** (1996): The lipocalin protein family: structure and function. *Biochem. J.* 318: 1-14.
- Force, A., Lynch, M., Pickett, F. B., Amores, A., Yan, Y. L. a Postlethwait, J.** (1999): Preservation of duplicate genes by complementary, degenerative mutations. *Genetics* 151: 1531-1545.
- Gaginskaya, E. R., Kasyanov, V. L., Kogan, G. L.** (1988): Amplification of ribosomal genes and formation of extrachromosomal nucleoli in oocytes of starfish *Henricia hayashi* (Asteroidea: Echinasteridae). *Cell Differ.* 23: 53-60.
- Galbraith, D. W., Karkins, K. R. a Knapp, S.** (1991): Systematic endopolyploidy in *Arabidopsis thaliana*. *Plant Physiology* 96: 985-989.
- Gallardo, M. H., Bickham, J. W., Honeycutt, R. L., Ojeda, R. A. a Kfhler, N.** (1999): Discovery o tetraploidy in a mammal. *Nature* 401: 341.
- Go, Y.** (2006): Lineage-specific expansions and contractions of the bitter taste receptor gene repertoire in vertebrates. *Mol. Biol. Evol.* 23: 964-972.
- Goncalves, I., Duret, L. a Mouchiroud, D.** (2000): Nature and structure of human genes that generate retropseudogenes. *Genome Res.* 10: 672-678.
- Gonzalez, E., Kulkarni, H., Bolivar, H., Mangano, A., Sanchez, R., Catano, G., Nibbs, R. J., Freedman, B. I., Quinones, M. P., Bamshad, M. J., Murthy, K. K., Rovin, B. H., Bradley, W., Clark, R. A., Anderson, S. A., O'Connell, R. J., Agan, B. K., Ahuja, S. S., Bologna, R., Sen, L., Dolan, M. J. a Ahuja, S. K.** (2005): The Influence of CCL3L1 Gene-Containing Segmental Duplications on HIV-1/AIDS Susceptibility. [Published online]: 10.1126/science. 1101160.
- Grus, W. E., Shi, P., Zhang, Y. P. a Zhang, J.** (2005): Dramatic variation of the vomeronasal pheromone receptor gene repertoire among five orders of placental and marsupial mammals. *Proc. Natl. Acad. Sci. U S A* 102: 5767-5772.
- Gu, Z., Cavalcanti, A., Chen, F., Bouman, P. a Li, W.** (2002): Extent of gene duplication in the genomes of *Drosophila*, nematode, and yeast. *Mol. Biol. Evol.* 19: 256–262.

- Gulick, A.** (1944): The chemical formulation of gene structure and gene action. *Adv. Enzymol.* 4: 1-39.
- Haldane, J. B. S.** (1933): The part played by recurrent mutation in evolution. *The American Naturalist* 67(708): 5-19.
- Haldane, J. B. S.** (1932): *The Causes of Evolution.* Cornell Univ. Press., Ithaca, NY.
- Harrison, P., Echols, N. a Gerstein, M.** (2001): Digging for Dead Genes: An Analysis of the Characteristics of the Pseudogene Population in the *C. elegans* Genome. *Nuc. Acids. Res.* 29: 818-830.
- Harrison, P. M., Hegyi, H. a Balasubramanian, S.** (2002): Molecular fossils in the human genome: identification and analysis of processed and non-processed pseudogenes in chromosomes 21 and 22. *Genome Res.* 12: 272-280.
- Harrison, P. M., Zheng, D., Zhang, Z., Carriero, N. a Gerstein, M.** (2005): Transcribed processed pseudogenes in the human genome: an intermediate form of expressed retrosequence lacking protein-coding ability. *Nucleic Acids Res.* 33: 2374 -2383.
- Honeycutt, R. L., Rowe, D. L. a Gallardo, M. H.** (2003): Molecular systematics of the South American caviomorph rodents: relationships among species and genera in the family Octodontidae. *Mol. Phylogenet. Evol.* 26: 476- 489.
- Holland, P. W. H. a Takahashi, T.** (2005): The evolution of homeobox genes: Implications for the study of brain development. *Brain Research Bulletin* 66 (4-6): 484-490.
- Houck, C. M., Rinehart, F. P. a Schmid, C. W.** (1979): Ubiquitous family of repeated DNA sequences in the human genome. *J. Mol. Biol.* 132: 289-306.
- Hufton, A. L., Groth, D., Vingron M. et al.** (2008): Early vertebrate whole genome duplications were predated by a period of intense genome rearrangement. *Genome Res.* 18: 1582-1591.
- Hughes, A. L.** (1999): Phylogenies of Developmentally Important Proteins Do Not Support the Hypothesis of Two Rounds of Genome Duplication Early in Vertebrate History. *J. Mol. Evol.* 48: 565-576
- Chen, Z. J. a Ni, Z.** (2006): Mechanisms of genomic rearrangements and gene expression changes in plant polyploids. *BioEssays* 28: 240-252.
- Jaillon, O., Aury, J. M., Brunet, F., Petit, J. L., Stange-Thomann, N. et al.** (2004): Genome duplication in the teleost fish *Tetraodon nigroviridis* reveals the early vertebrate proto-karyotype. *Nature* 431: 946-957.
- Kanwaljit, S. D., Dornum, M., Mollon, J. D. et al.** (1999): The Evolution of Trichromatic Color Vision by Opsin Gene Duplication in New World and Old World Primates. *Genome Res.* 9: 629-638
- Kazazian Jr., H. H., Wong, C., Youssoufian, H., Scott, A. F., Phillips, D. G. a Antonarakis, S. E.** (1988): Haemophilia A resulting from de novo insertion of L1 sequences represents a novel mechanism for mutation in man. *Nature* 332:164-166.
- Kellis, M., Birren, B. W. a Lander, E. S.** (2004): Proof and evolutionary analysis of ancient genome duplication in the yeast *Saccharomyces cerevisiae*. *Nature* 428: 617-624.
- Kidwell, M. G. a Lisch, D. R.** (2001): Perspective: Transposable elements, parasitic DNA, and genome evolution. *Evolution*, 55(1): 1-24.

- Kondrashov, F. A. a Kondrashov, A. S.** (2006): Role of selection in fixation of gene duplications. *Journal of Theoretical Biology* 239: 141-151.
- 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. International Human Genome Sequencing Consortium. *Nature* 409: 860–921.
- Leitch, I. L. a Bennett, M. D.** (1997): Polyploidy in angiosperms. *Trends Plant. Sci.* 2:470–476.
- Lejeune, J., Gautier, M. a Turpin, R.** (1959): Etude des chromosomes somatiques de neuf mongoliens. *Compte Rendus Hebdomadaires Des Seances De L Academie Des Sciences* 248(11): 1721-1722
- Lemons, D. a McGinnis, W.** (2006): Genomic evolution of Hox gene clusters. *Science* 313: 1918-1922.
- Letunic, I., Copley, R. R. a Bork, P.** (2002): Common exon duplication in animals and its role in alternative splicing. *Hum. Mol. Genet.* 11: 1561–1567.
- Lewis, E. B.** (1951): Pseudoallelism and gene evolution. *Cold Spring Harbor Symp. Q. Biol.* 16: 159-74.
- Lynch, M. a Conery, J. S.** (2000): The Evolutionary Fate and Consequences of Duplicate Genes. *Science* 290: 1151-1155.
- Lynch, M., a Force, A.** (2000): The probability of duplicate gene preservation by subfunctionalization. *Genetics* 154:459–473.
- Lynch, M., O’Hely, M., Walsh, B. a Force, A.** (2001): The Probability of Preservation of a Newly Arisen Gene Duplicate. *Genetics* 159: 1789-1804.
- Masterson, J.** (1994): Stomatal size in fossil plants: evidence for polyploidy in majority of angiosperms. *Science* 264: 421–424.
- McCarrey, J. R. a Thomas, K.** (1987): Human testis-specific PGK gene lacks introns and possesses characteristics of a processed gene. *Nature* 326: 501-505.
- Macgregor, H. C.** (1968): Nucleolar DNA in oocytes of *Xenopus laevis*. *J. Cell Sci.* 3: 437-444.
- Metz, C. W.** (1947): Duplication of chromosome parts as a factor in evolution. *The American Naturalist* 81: 81-103.
- Mighell, A. J., Smith, N. R., Robinson, P. A. a Markham, A. F.** (2000): Vertebrate pseudogenes. *FEBS Lett.* 468: 109-114.
- Miskey, C., Izsvák, Z., Kawakami, K. a Iwics, Z.** (2004): DNA transposons in vertebrate functional genomics. *Cell. Mol. Life Sci.* 62: 629–641
- Moore, A. D., Bjorklund, A. K., Ekman, D., Bornberg-Bauer, E. a Elofsson, A.** (2008): Arrangements in the modular evolution of proteins. *Trends Biochem. Sci.* 33: 444–451.
- Mounsey, A., Bauer, P. a Hope, I. A.** (2002): Evidence Suggesting That a Fifth of Annotated *Caenorhabditis elegans* Genes May Be Pseudogenes. *Genome Res.* 12: 770-775.
- Muller, H. J.** (1925): Why polyploidy is rarer in animals than in plants. *Amer. Nat.* 59: 346–353.

- Nei, M., a Roychoudhury, A. K.** (1973): Probability of fixation of nonfunctional genes at duplicate loci. *Am. Nat.* 107: 362-372.
- Nei, M., Gu, X. a Sitnikova, T.** (1997): Evolution by the birth-and-death process in multigene families of the vertebrate immune system. *Proc. Natl. Acad. Sci. USA* 94: 7799-7806.
- Niimura, Y., a Nei, M.** (2007): Extensive Gains and Losses of Olfactory Receptor Genes in Mammalian Evolution. *PloS ONE* 2: e708
- Ohno, S.** (1970): *Evolution by Gene Duplication*. Springer-Verlag, Heidelberg, Germany.
- Olson, M. V.** (1999): When less is more: gene loss as an engine of evolutionary change. *Am. J. Hum. Genet.* 64:18-23.
- Ostertag, E. M. a Kazazian Jr., H. H.** (2001): Biology of Mammalian L1 Retrotransposons. *Annu. Rev. Genet.* 35: 501-38.
- Patau, K., Smith, D. W., Therman, E., Inhorn, S. L. a Wagner, H. P.** (1960): Multiple Congenital Anomaly Caused By An Extra Autosome. *The Lancet* 275 (7128): 790-793.
- Pelosi, P., Baldaccini, N. E. a Pisanelli, A. M.** (1982): Identification of a specific olfactory receptor for 2-isobutyl-3-methoxypyrazine. *Biochem. J.* 201: 245-248.
- Pelosi, P.** (1994): Odorant-binding proteins. *Crit. Rev. Biochem. Mol. Biol.* 29: 199-228.
- Piatigorsky, J. a Wistow, G.** (1991). The recruitment of crystallins: New functions precede gene duplication. *Science* 252: 1078-1079.
- Piontkivska, H., Rooney, A. P. a Nei, M.** (2002): Purifying Selection and Birth-and-death Evolution in the Histone H4 Gene Family *Mol. Biol. Evol.* 19(5): 689-697.
- Pontes, O., Neves, N., Silva, M., Lewis, M. S., Madlung, A., Comai, L., Viegas, W. a Pikaard, C. S.** (2004): Chromosomal locus rearrangements are a rapid response to formation of the allotetraploid *Arabidopsis suecica* genome. *Proc. Natl. Acad. Sci.* 101: 18240-18245.
- Putnam, N. H., Butts, T., Ferrier, D. E. K., Furlong, R. F., Hellsten, U., Kawashima, T., Robinson-rechavi, M., Shoguchi, E., Terry, A. a Yu, J-K. et al.** (2008): The amphioxus genome and the evolution of the chordate karyotype. *Nature* 453: 1064-1072.
- Rosenberg, H. F., Dyer, K. D., Tiffany, H. L. a Gonzalez, M.** (1995): Rapid evolution of a unique family of primate ribonuclease genes. *Nat Genet.* 10(2):219-223.
- Rubin, G. M., Yandell, M. D., Wortman, J. R.** (2000): Comparative genomics of the eukaryotes. *Science* 287: 2204-2215.
- Sassaman, D. M., Dombroski, B. A., Moran, J. V., Kimberland, M. L., Naas, T. P., DeBerardinis, R. J., Gabriel, A., Swergold, G. D. a Kazazian Jr. H. H.** (1997): Many human L1 elements are capable of retrotransposition. *Nat. Genet.* 16: 37-43.
- Semon, M. a Wolfe, K. H.** (2007): Rearrangement rate following the whole-genome duplication in teleosts. *Mol. Biol. Evol.* 24: 860-867.
- Serebrovsky, A. S.** (1938): „Genes scute and achaete in *Drosophila melanogaster* and a hypothesis of gene divergency“. *C.R. Acad. Sci. URSS* 19: 77-81.

- Shaikh, T. H., Roy, A. M., Kim, J., Batzer, M. A. a Deininger, P. L.** (1997): cDNAs derived from primary and small cytoplasmic Alu (scAlu) transcripts. *J. Mol. Biol.* 271: 222–234.
- Shen, S. H., Slightom, J. L. a Smithies, O.** (1981): A history of the human fetal globin gene duplication. *Cell* 26: 191-203.
- Shen, M. R., Batzer, M. a Deininger, P.** (1991) Evolution of the master Alu gene(s). *J. Mol. Evol.* 33: 311–320.
- Scheer, U.** (1973). Nuclear pore flow rate of ribosomal RNA and chain growth rate of its precursor during oogenesis of *Xenopus laevis*. *Dev. Biol.* 30:13-28.
- Smit, A. F. a Riggs, A. D.** (1996): Tiggers and DNA transposon fossils in the human genome. *Proc. Natl. Acad. Sci. USA* 93(4):1443-1448.
- Smith, D. W., Patau, K., Therman, E., Inhorn, S. L.** (1960): A new autosomal trisomy syndrome: multiple congenital anomalies caused by an extra chromosome. *The Journal of Pediatrics* 57 (3): 338-345.
- Song, K., Lu, P., Tang, K. a Osborn, T. C.** (1995): Rapid genome change in synthetic polyploids of Brassica and its implications for polyploid evolution. *Proc. Natl. Acad. Sci.* 92: 7719-7723.
- Steinbrecht, R. A.** (1998): Odorant-Binding Proteins: Expression and Function. *Annals of the New York Academy of Sciences* 855: 323-32.
- Stephen, S. G.** (1951): Possible significance of duplication in evolution. *Adv. Genet.* 4: 247-65.
- Storchova, Z. a Pellman, D.** (2004): From polyploidy to aneuploidy, genome instability and cancer. *Nat. Rev. Mol. Cell Biol.* 5: 45–54.
- Svartman, M., Stone, G. a Stanyon, R.** (2005): Molecular cytogenetics discards polyploidy in mammals. *Genomics* 85: 425-430.
- Thiébaud, Ch. H.** (1979): Quantitative Determination of Amplified rDNA and Its Distribution During Oogenesis in *Xenopus laevis*. *Chromosoma* 73: 37-44.
- Urata, Y., Parmelee, S. J., Agard, D. A., a Sedat, J. W.** (1995): A three-dimensional structural dissection of *Drosophila* polytene chromosomes. *J. Cell Biol.* 131: 279–295.
- Vanin, E. F.** (1985): Processed pseudogenes: characteristic and evolution. *Annu. Rev. Genet.* 19: 253-72.
- Viktorov, A. G.** (1997): Diversity of polyploid races in the family Lumbricidae. *Soil Biol. Biochem.* 29: 217-21.
- Walsh, J. B.** (1995): How often do duplicated genes evolve new functions? *Genetics* 110: 345-364.
- Waterston, R. H., Lindblad-Toh, K., Birney, E., Rogers, J., Abril, J. F., Agarwal, P., Agarwala, R., Ainscough, R., Alexandersson, M., An, P., Antonarakis, S. E., Attwood, J. et al.** (2002): Initial sequencing and comparative analysis of the mouse genome. *Nature* 420: 520-562.
- Winter, H., Langbein, L., Krawczak, M., Cooper, D. N., Jave-Suarez, L. F., Rogers, M. A., Praetzel, S., Heidt, P. J. a Schweizer, J.** (2001): Human type I hair keratin pseudogene ϕ H_aA has functional orthologs in the chimpanzee and gorilla: evidence for

recent inactivation of the human gene after the Pan-Homo divergence. *Hum. Genet.* 108 : 37–42

Zhang, Z., Carriero, N. a Gerstein, M. (2004): Comparative analysis of processed pseudogenes in the mouse and human genomes. *Trends Genet.* 20: 62-67.

Zhang, G. J. a Cohn, M. J. (2008): Genome duplication and the origin of the vertebrate skeleton. *Current Opinion in Genetics & Development* 18:387-393.

Zhang, Z., Harrison, P. M., Liu, Y. a Gerstein, M. (2003): Millions of years of evolution preserved: a comprehensive catalog of the processed pseudogenes in the human genome. *Genome Res.* 13: 2541-2558.