

Přírodovědecká fakulta Univerzity Karlovy v Praze
Katedra buněčné a vývojové biologie



Charakterizace genového obsahu chromosomu Z u kura domácího

DIPLOMOVÁ PRÁCE

autor: **Bc. Libor Mořkovský**

vedoucí práce: **RNDr. Radka Storchová, Ph.D.**

Praha 2010

Prohlašuji, že diplomovou práci jsem vypracoval samostatně s použitím uvedené literatury.

Libor Mořkovský

Matematik potřebuje ke své práci tužku, papír a odpadkový koš.

Evoluční biolog si vystačí s tužkou a papírem.

Chtěl bych poděkovat své školitelce, Radce Storchové, za vedení mojí práce a také za komentáře a návrhy ke vznikající diplomové práci. Dále bych chtěl poděkovat Petru Divinovi (Informační technologie ÚMG AV ČR) za poskytnutí účtu na serveru vhodném pro provádění rozsáhlejších výpočtů.

Abstrakt

Podle teoretického modelu by měly pohlavně antagonistické mutace být více nebo naopak méně zastoupeny na chromosomech X a Z ve srovnání s autosomy. Zda budou zastoupeny více či méně závisí na průměrné míře dominance těchto mutací. Protože průměrná míra dominance nově vznikajících mutací není známá, je obtížné předpovědět důsledky působení pohlavně antagonistické selekce. Abychom lépe porozuměli vlivu pohlavně antagonistické selekce na genový obsah chromosomu Z u kura domácího, podrobili jsme analýze veřejně dostupná expresní data z několika somatických tkání a z germinálních i somatických buněk vaječníku.

Zjistili jsme, že na chromosomu Z jsou ve větší míře zastoupeny geny s preferenční expresí v somatických buňkách vaječníku (kumulární buňky). Geny s preferenční expresí v germinálních buňkách ani geny s preferenční expresí v somatických tkáních nevykazují odlišné rozmístění na chromosomech od rozmístění náhodného. Tyto výsledky napovídají, že na chromosomu Z se hromadí alely výhodné pro samice. Nepřítomnost signálu u germinálních buněk může být vysvětlena buď jako důsledek nižší intenzity pohlavně antagonistické selekce v těchto buňkách, nebo jako důsledek meiotické inaktivace chromosomu Z. Naše výsledky jsou v souladu se závěry podobných výzkumů u savců a naznačují, že na utváření genového obsahu chromosomu Z se výrazněji podílí recesivní pohlavně antagonistické mutace.

Klíčová slova

kur domácí, pohlavní antagonizmus, kompenzace genové dávky, meiotická inaktivace pohlavních chromosomů, chromosom Z, genový obsah, genové čipy
Affymetrix

Abstract

Theory predicts that sexually antagonistic mutations will be over- or under-represented on the X and Z chromosomes, depending on the average dominance coefficient of the mutations. However, as little is known about the dominance coefficients for new mutations, the effect of sexually antagonistic selection is difficult to predict. To elucidate the role of sexually antagonistic selection in the evolution of Z chromosome gene content in chicken, we analyzed publicly available microarray data from several somatic tissues as well as somatic and germ cells of the ovary.

We found that the Z chromosome is enriched for genes showing preferential expression in ovarian somatic cells, but not for genes with preferential expression in primary oocytes or non-sex-specific somatic tissues. Our results suggest that sexual antagonism leads to higher abundance of female-benefit alleles on the Z chromosome. No bias towards Z-linkage of oocyte-enriched genes can be explained by lower intensity of sexually antagonistic selection in ovarian germ cells compared to ovarian somatic cells. An alternative explanation would be that meiotic Z chromosome inactivation hinders accumulation of oocyte-expressed genes on the Z chromosome. Our results are consistent with findings in mammals and indicate that recessive rather than dominant sexually antagonistic mutations shape the gene content of the X and Z chromosomes.

Keywords

chicken, sexual antagonism, dosage compensation, meiotic sex chromosome inactivation, Z chromosome, gene content, Affymetrix gene chips

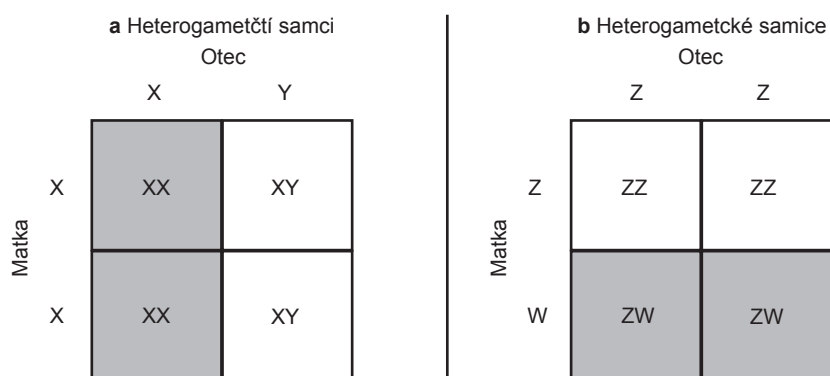
Obsah

1. Úvod	7
1.1 Pohlavní chromosomy	8
1.2 Pohlavně antagonistická selekce	13
1.3 Kompenzace genové dávky	15
1.4 Meiotická inaktivace pohlavních chromosomů	21
2. Cíle práce	24
3. Materiál a metody	26
3.1 Data z genových expresních čipů	26
3.2 Anotace, normalizace a sumarizace expresních dat	27
3.3 Hodnocení diferenciální exprese a statistická analýza	27
4. Výsledky	30
5. Diskuse	34
5.1 Pohlavně antagonistická selekce a míra dominance pohlavně antagonistických mutací	34
5.2 Meiotická inaktivace pohlavních chromosomů a defeminizace chromosomu Z	36
6. Závěr	38
7. Použitá literatura	39
8. Přílohy	43
8.1 Článek publikovaný v <i>Journal of Molecular Evolution</i>	43
8.2 Skripty jazyka R využité pro analýzu	52
8.3 Výsledky pro datové sady 1 a 2	57
8.4 Seznam genů s preferenční expresí v kumulárních buňkách.....	60

1. Úvod

Existence pohlavních chromosomů je zajímavým fenoménem. Mají je všichni savci a ptáci. Pohlavní chromosomy nalezneme i u některých ryb, obojživelníků a plazů či některých řádů hmyzu, zatímco jiní zástupci těchto skupin se obejdou bez nich. Pokud dojde ke vzniku pohlavních chromosomů, hrají významnou roli v evolučních procesech, jako je vznik pohlavního dimorfismu, pohlavní výběr a samotný vznik nových druhů (speciace). Velká časová škála, na které takové evoluční procesy probíhají, znemožňuje jejich přímé studium. Aktuální stav genomu a transkriptomu, které zkoumat díky moderním molekulárním metodám lze, můžeme považovat za otisk probíhajících evolučních procesů. Studium genů lokalizovaných na pohlavních chromosomech nám pomůže nahlédnout do mechanismů, které tyto procesy řídí.

Tradičními – a tedy nejprobádanějšími – genetickými modely jsou myš domácí a octomilka. Oba tyto organismy mají pohlaví určené systémem XY. To znamená, že homogametickým pohlavím (tj. pohlavím se dvěma stejnými pohlavními chromosomy, a tudíž jedním typem gamet) je samice, heterogametickým pohlavím (tj. pohlavím se dvěma odlišnými pohlavními chromosomy, a tudíž dvěma typy gamet) samec (obr. 1).



Obr. 1: Různé systémy určení pohlaví. Na rozdíl od savců a octomilek (a) mají ptáci heterogametické samice (b). Díky této odlišnosti jsou zajímavým modelem pro testování evolučních hypotéz.

Kur domácí (*Gallus gallus*) je pro studium pohlavních chromosomů zajímavým modelem, protože jeho systém určení pohlaví je oproti myši a octomilce opačný – samice jsou heterogametické, samci homogametičtí. Pro odlišení se pohlavní chromosomy u tohoto systému nazývají Z a W (obr. 1). Kur je nejvíce probádaným gene-

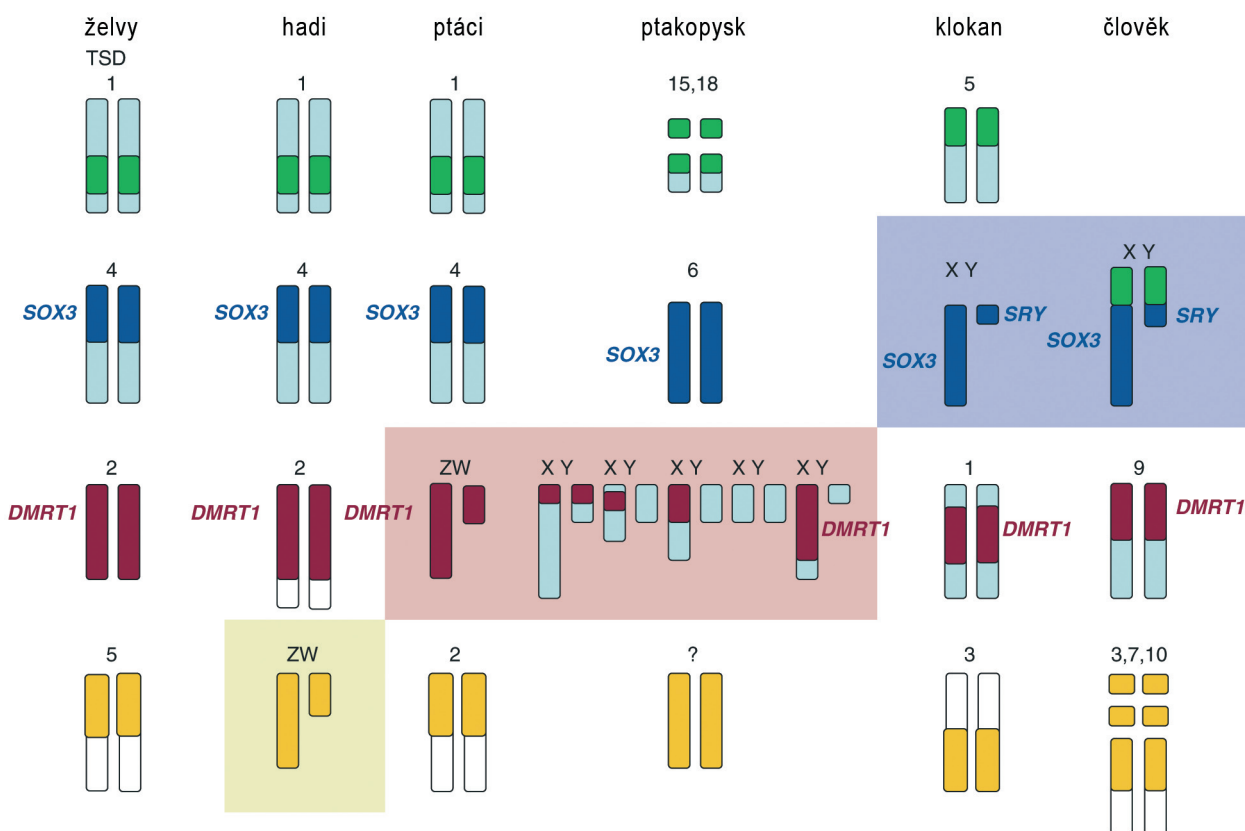
tickým modelem se systémem určení pohlaví ZW, kompletní genom byl publikován v roce 2004 (International Chicken Genome Sequencing Consortium, 2004). Využití modelu se systémem určení pohlaví ZW, ve srovnání se studii provedenými na modelech s XY určením pohlaví, by mohlo umožnit odhalit obecnější mechanismy, které ovlivňují charakter genového složení pohlavních chromosomů.

V úvodu práce nejprve čtenáře blíže seznámím s jevy, které se podle dosavadních znalostí mohou podílet na utváření genového obsahu pohlavních chromosomů, zasaženými v kontextu modelu zvoleného pro mou práci – kura domácího. Je to pohlavně antagonistická selekce, kompenzace genové dávky v jednom z pohlaví a meiotická inaktivace pohlavních chromosomů.

1.1 Pohlavní chromosomy

Pohlavní chromosomy byly objeveny koncem 19. století, nejprve u ruměnice pospolné (*Pyrrhocoris apterus*; Henking, 1891), posléze u savců (Painter, 1912) a u rostlin (Blackburn, 1923). Aktuální poznatky o přítomnosti pohlavních chromosomů v různých organismech jsou poněkud roztržité. Je to dáno relativní komplikovaností provedení karyologického vyšetření, obtížnou rozpoznatelností homomorfních (viz dále) pohlavních chromosomů a také samotnou situací v přírodě a nerovnoměrným zaměřením vědeckého výzkumu.

U obratlovců je situace prozkoumána asi nejlépe (přehled v Graves a Peichel, 2010). Placentální savci sdílejí homologní genově bohatý chromosom X a většina jich má malý degenerovaný chromosom Y obsahující *SRY*, gen určující samčí pohlaví nositele. Výjimkou jsou slepušky, krysa japonská a hraboš mandarínský, u kterých nebyl nalezen gen *SRY* a samci mají genotyp X0. Vačnatci mají menší chromosom X, který je homologní pouze s částí chromosomu X placentálů, a drobný chromosom Y. Vejcorodí mají zvláštní komplex pěti X a pěti Y chromosomů. Ptáci mají chromosom Z, který není homologní k savčímu X. Chromosom W u ptáků vykazuje různý stupeň degenerace. U plazů již situace není tak jednoduchá. Hadi mají podobně jako ptáci vzájemně homologní chromosom Z, který však není homologní ani s ptačím chromosomem Z, ani se savčím chromosomem X (viz také obr. 2). Různé stupně



Obr. 2: Ortologie pohlavních chromosomů a autosomů u amniot. Čtyři konzervované regiony obsahující ortologní geny jsou vyznačeny zelenou, červenou, modrou a žlutou. Pokud leží na pohlavních chromosomech, jsou podbarveny. *Upraveno podle Graves a Peichel (2010).*

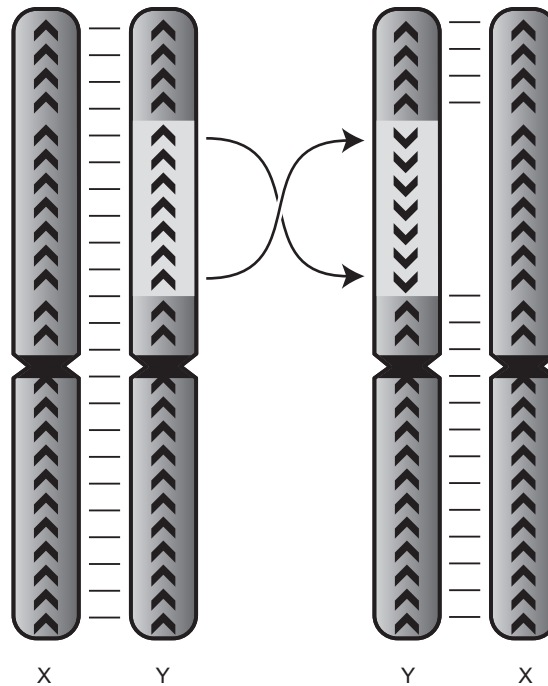
degenerace chromosomu W pozorované u hadů vedly k navržení teorie o vzniku pohlavních chromosomů z páru autosomů (Ohno, 1967). Ostatní linie plazů mají různé systémy určení pohlaví – XY, ZW i teplotní určení pohlaví (*temperature-dependent sex determination*, TSD). U obojživelníků i ryb nalezneme systémy XY, ZW i TSD. U skokana vrásčitého (*Rana rugosa*) žijícího na japonských ostrovech jsou známy populace s určením pohlaví XY, populace s určením pohlaví ZW a existuje i hybridní zóna, kde se jedinci z různých populací kříží a vznikají životaschopní hybridy (Ogata et al., 2007).

V rostlinné říši je asi 5 % druhů dvoudomých (mají oddělené samčí a samičí jedince – nutný, ale ne postačující předpoklad pro přítomnost pohlavních chromosomů). Nějakého dvoudomého zástupce má přibližně 75 % čeledí. Ze známějších jmenujme například silenku širolistou, šťovík kyselý, kopřivu dvoudomou, konopí seté, chmel otáčivý, jinan dvoulaločný a špenát setý (přehled ve Vyskot a Hobza, 2004).

Mnoho řádů hmyzu má pohlaví určeno pohlavními chromosomy. Obvyklý je systém XY a odvozené varianty (XX/X0, více pohlavních chromosomů), pouze motýli (*Lepidoptera*) a chrostíci (*Trichoptera*) mají systém určení pohlaví ZW. U motýlů nalezneme různé odvozené varianty, chrostíci mají systém ZZ/Z0 (přehled v Traut, 1999). Blanokřídlí mají pohlaví určené haplodiploidně, diploidní jedinci jsou samice, haploidní samci.

U hub je situace poněkud odlišná – dochází zde k pohlavnímu rozmnožování, ale gamety nejsou morfologicky rozlišitelné (isogamie). Nerozlišujeme proto pohlaví, ale párovací typy. Párovací typy nemusí být na rozdíl od pohlaví pouze dva, například sněť kukuřičná (*Ustilago maydis*) má párovací typy kódované dvěma nezávislými lokusy. Pro jeden z těchto lokusů existuje více alel, takže ve výsledku existují desítky párovacích typů gamet. Regiony chromosomů, které nesou geny kódující párovací typy gamet, sdílejí některé vlastnosti s pohlavními chromosomy ostatních organismů – dochází v nich ke ztrátě rekombinace, hromadění sobeckých genetických elementů a meiotické inaktivaci (přehled ve Fraser a Heitman, 2004).

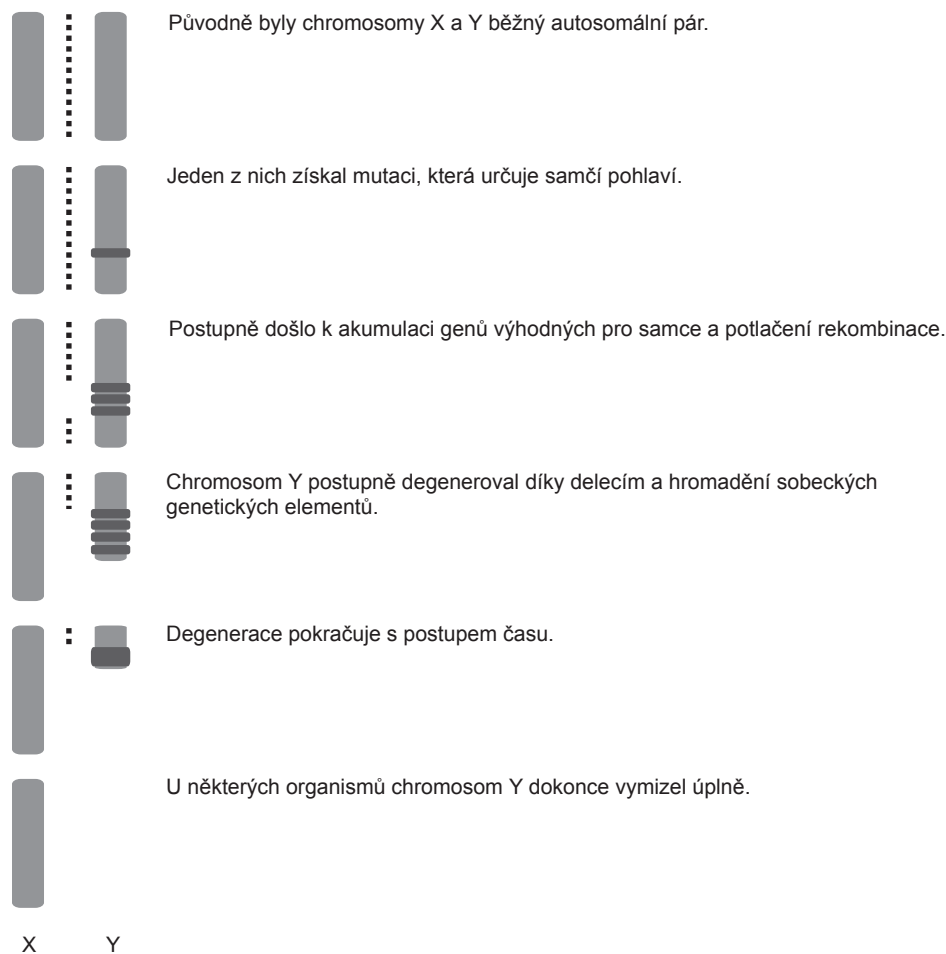
Pohlavní chromosomy se vyvinuly pravděpodobně několikrát nezávisle na sobě z páru běžných autosomů (Bull, 1983). Postup evoluce pohlavních chromosomů lze rozdělit do několika logicky navazujících kroků. V jednom z kaskády genů podílejících se na určení pohlaví jedince v ontogenezi vznikne mutace určující pohlaví svého nositele. U savců například zmutoval gen *SOX3* na variantu nazvanou *SRY* (Waters et al., 2007). Tím se z páru autosomů stane pár pohlavních chromosomů. Tyto nascentní pohlavní chromosomy se liší pouze na molekulární úrovni, nejsou rozlišitelné v mikroskopu. Nazýváme je proto homomorfní. Následně dochází k postupnému omezování rekombinace mezi chromosomy, převážně díky fixaci inverzí (obr. 3). Ztrátu rekombinace si lze představit jako evolučně výhodnou, alespoň v kratším časovém měřítku. V okolí genu určujícího pohlaví se začnou kumulovat další mutace výhodné právě pro pohlaví určené původně zmutovaným genem. Vazba s tímto genem jim zaručí, že se budou vždy nacházet ve správném pohlaví, a rekombinace je jediný způsob, jak mohou o tuto vazbu přijít (Waters et al., 2007).



Obr. 3: Chromosomální inverze. Při vzniku inverze dochází ke ztrátě homologního párování, a tudíž i možnosti rekombinace mezi chromosomy.

Následně začíná nepárový chromosom kvůli nepřítomnosti rekombinace genetiky degenerovat. Dochází postupně k jeho heterochromatinizaci a obvykle díky deletcím i k jeho zmenšování. Díky tomu je později možné od sebe pohlavní chromosomy odlišit (heteromorfní pohlavní chromosomy). Tento proces zjevně probíhá u savců, ptáků i hadů (obr. 4). Přesto tato cesta není jedinou, po které se může evoluce páru pohlavních chromosomů vydat. Silenka širolistá – jeden z významných modelů pro studium pohlavních chromosomů u rostlin – disponuje chromosomem Y, který je nejen větší než chromosom X, ale také větší než všechny její ostatní chromosomy a sestává převážně z euchromatinu (Lardon et al., 1999). Některé z výše popsaných mechanismů ale pracují i u silenky. Existují totiž náznaky genetické degenerace chromosomu Y, Guttman (1998) publikuje nález nefunkčního homologu genu lokalizovaného na chromosomu X.

Nestejná provázanost pohlavních chromosomů s pohlavími dává prostor k působení různých evolučních tlaků. Ty postupně vedou ke zřetelné odlišnosti páru pohlavních chromosomů od autosomů. Pohlavní chromosomy vykazují odlišnosti jak v epigenetických modifikacích a s nimi související transkripční aktivitě (Graves



Obr. 4: Postupný vývoj pohlavních chromosomů. Pro jednoduchost je v popiskách použito označení chromosomů X a Y, ale podobný proces probíhá i u Z a W.

et al., 2006), tak při podrobnějším prozkoumání ve vlastnostech genů na nich lokalizovaných (Vicoso a Charlesworth, 2006). Mezi epigenetické modifikace pohlavních chromosomů patří kompenzace dávky genů a meiotická inaktivace u heterogametickeho pohlaví. Geny pohlavních chromosomů jsou častěji specificky exprimovány v jednom pohlaví. Některé z těchto vlastností mohou být důsledkem působení přírodního výběru na protichůdné zájmy samců a samic (pohlavně antagonistické selekce), jiné naopak pouze následky diferenciacce pohlavních chromosomů (Graves et al., 2006). V tomto komplikovaném systému se sice rozsáhlou, ale také velmi nejednoznačnou evidencí není lehké rozlišit, co je příčina a co důsledek.

1.2 Pohlavně antagonistická selekce

Reprodukční strategie samců a samic se liší. Vezměme například lokus kódující velikost plodu v embryonálním vývoji. Zájmem samce je, aby jeho potomek byl co největší a nejživotoschopnější. Zájmem samice je ve zdraví přežít vrh a případně donosit i potomky dalších samců. Takto protichůdné zájmy označujeme jako pohlavně antagonistické. Pohlavní antagonismus představuje problém pro evoluci lokusu ovlivňujícího procesy, v nichž se liší strategie samců a samic. Pokud je takový lokus na autosomech, je vystaven stejné expozici v obou pohlavích. Ve výsledku bude tedy ovlivňován součtem obou evolučních tlaků. Nevznikne-li žádný další mechanismus, který by odděloval selekční tlaky samčí a samičí, bude lokus udržován ve stavu, který není optimální ani pro jedno pohlaví.

Zastoupení pohlavních chromosomů v každém pohlaví a také jejich dědičnost jsou pohlavně specifické, čehož využil Rice (1984) k navržení matematického modelu pohlavně antagonistické selekce. Model charakterizuje pohlavně antagonistickou alelu třemi parametry: relativním zvýšením fitness jednoho pohlaví S , relativním snížením fitness pohlaví opačného T a relativní dominancí alely h . Všechny parametry modelu jsou z intervalu $<0, 1>$. Pokud je zkoumaná alela autosomální, bude pozitivně selektována, pouze pokud je S větší než T , bez ohledu na míru její dominance h . Alely na chromosomu Y (chromosomu W) se nacházejí stále ve stejném pohlaví, jsou proto pod konstantní selekcí na výhodnost pro toto pohlaví, čili budou selektovány pro všechna $S > 0$.

Pokud je lokalizována na chromosomu X (nebo na chromosomu Z u systému ZW), může být pohlavně antagonistická alela pozitivně selektována i přesto, že snížení fitness, které představuje pro jedno pohlaví je větší, než výhoda, kterou přináší pohlaví druhému. Homogametické pohlaví může těžit z dominantní alely lokalizované na chromosomu X (chromosomu Z), protože tento chromosom v něm tráví 2/3 (evolučního) času. Selekce bude upřednostňovat takovou alelu, pokud:

$$h > 1/(2 - T) \quad (1)$$

Heterogametickému pohlaví naopak přinášejí výhody alely dostatečně recesivní, protože jsou v tomto pohlaví v hemizygotní expozici (bez druhé alely), zatímco

v homogameticím pohlaví je nevýhodná alela s velkou pravděpodobností skryta za jinou (dominantnější) alelou. Selektce bude upřednostňovat takovou alelu, pokud:

$$h < 1/(2 + T) \quad (2)$$

Pohlavně antagonistická alela se tedy mnohem snáze rozšíří v populaci, vznikne-li na pohlavním chromosomu. Jakmile se taková alela více rozšíří v populaci, vznikne selekční tlak na úpravu exprese jejího lokusu nebo konkrétní alely jako takové. Gen bude exprimovaný v tom pohlaví, kterému přináší výhodu, a nebude exprimovaný v pohlaví opačném, pro které je nevýhodný. Lze předpokládat, že nadprůměrná proporce pohlavně dimorfních znaků bude kódována – nebo alespoň regulována – lokusy na pohlavních chromosomech (obr. 5).

mutace		míra dominance	
		D	r
přínosná pro	♀	X, W	Z, W
	♂	Z, Y	X, Y

Obr. 5: Chromosomy, na kterých je preferováno hromadění pohlavně antagonistických mutací, tak jak vyplývá z teoretického modelu, který navrhl Rice (1984). Mutace jsou rozděleny podle dominance – D znamená dostatečně dominantní podle rovnice (1), r znamená dostatečně recesivní podle rovnice (2) – a podle pohlaví, pro které přináší výhodu ($S > 0$). Dominantní mutace se hromadí na chromosomech, které tráví více evolučního času v pohlaví, pro které je mutace výhodná. Recesivní mutace se hromadí na chromosomech, které jsou v hemizygotní expozici v pohlaví, pro které je mutace výhodná.

Mnohá srovnání genové exprese na úrovni celých transkriptomů potvrzují, že pohlaví je jedním z hlavních faktorů podílejících se na vnitrodruhové variabilitě expresních úrovní, a to nejen v gonádách (Parisi et al., 2003; Ranz et al., 2003; Mank et al., 2008), ale i v somatických tkáních (Parisi et al., 2003; Yang et al., 2006; Mank et al., 2008). Pohlavně specifickou úroveň exprese daného lokusu můžeme považovat za důkaz pohlavního antagonismu působícího na tento lokus. Ve spojení s globálním pohledem na expresní úrovně lze pohlavní specifitu exprese využít k testování platnosti matematického modelu pohlavně antagonistické selekce. Je zajímavé, že výsledky dosavadních studií nejsou konzistentní mezi jednotlivými druhy

(přehled v Vicoso a Charlesworth, 2006; Gurbich a Bachtrög, 2008). Chromosom X u savců je obohacen o geny s přednostní expresí v somatických tkáních samců, a naopak ochuzen o geny exprimované ve spermatogenních tkáních (Khil et al., 2004). U octomilky mají na chromosomu X prokazatelně nižší zastoupení geny specifické pro samce, bez ohledu na konkrétní tkáň (Parisi et al., 2003; Sturgill et al., 2007). U kura domácího je na chromosomu Z výrazně nižší zastoupení genů více exprimovaných v samicích (Kaiser a Ellegren, 2006; Storchova a Divina, 2006).

Pokud bychom se rigidně drželi teorie pohlavně antagonistické selekce, jedinou možností, jak vysvětlit tyto nesrovnalosti, by byl závěr, že různé skupiny organismů mají své specifické míry dominance náhodně vzniklých mutací. V rámci dosavadních poznatků molekulární biologie je těžké si představit konkrétní molekulární mechanismus, který by toto zajišťoval. Prozaičtějším vysvětlením mohou být epigenetické mechanismy, které v průběhu ontogeneze ovlivňují expresi genů na úrovni jednotlivých chromosomů, obvykle tím, že umlčují nebo zesilují expresi právě na pohlavních chromosomech. Nezohlednění těchto jevů může snadno vést k dezinterpretaci výsledků studií zabývajících se testováním predikcí matematického modelu pohlavního antagonismu, protože ten s nimi nepočítá. Přesto existují recentní studie – například Naurin et al. (2010), snažící se syntetizovat moderní poznatky s výše zmíněným modelem. Vznik epigenetických mechanismů z něj naopak vyvozují.

1.3 Kompenzace genové dávky

Díky postupné degeneraci nepárového pohlavního chromosomu – chromosom Y, případně chromosom W – jsou buňky odlišných pohlaví vystaveny působení různého počtu kopií genů lokalizovaných na nerekombinující části párového pohlavního chromosomu (různé genové dávky). Homogametické pohlaví má pro takový lokus dvě kopie, heterogametické pohlaví pouze jednu. Za mechanistického předpokladu, že míra transkripce odpovídá počtu kopií genu v buňce, je bez přítomnosti dalších korekčních mechanismů výsledkem vyšší transkripce pro daný lokus u homogametického pohlaví. Změna genové dávky pro jednotlivé geny má prokazatelný efekt na biochemické pochody v buňce (Pollack et al., 2002). Změna genové dávky pro celý

chromosom je obvykle letální, jak můžeme soudit například z dobře prozkoumaných vrozených aneuploidií u člověka. Pouze aneuploidie malých chromosomů (chromosom Y – 104 genů, chromosom 21 – 311 genů, chromosom 18 – 451 genů a chromosom 13 – 508 genů) a poměrně velkého chromosomu X (1203 genů) (počty genů podle Ensembl release 59, bez pseudogenů; Birney et al., 2004) jsou slučitelné se životem. Aneuploidie autosomů jsou obvykle provázeny výraznějším fenotypem než aneuploidie chromosomu X (X0 – Turnerův syndrom, XXY – Klinefelterův syndrom). Důvodem je pravděpodobně existence mechanismů, které jsou schopny transkripčně utiřit (inaktivovat) téměř celý chromosom X.

Výzkum inaktivace pohlavních chromosomů začíná popisem „Barrova tělíska“ u kočičích neuronů (Barr a Bertram, 1949). Již na konci 50. let 20. století bylo známo, že i v buňkách samic jiných savců (myš, krysa, vačice, člověk) se nalézá podobný chromatin, a analýzou dědičnosti znaků – různých chorob a barvy srsti – bylo prokázáno, že se jedná o inaktivovaný chromosom X. Bylo také zjištěno, že je inaktivován náhodně, nezávisle na tom, zda byl zděděn od otce, nebo od matky, takže samičí organismy jsou mosaikou buněk s různými chromosomy X (přehled v Lyon, 1962; recentně v Payer a Lee, 2008). Tento fenomén nabízí vysvětlení pro menší vliv aneuploidií chromosomu X na výsledný fenotyp. Samičí buňky po inaktivaci využívají převážně jeden chromosom X, proto snáze kompenzují genotyp X0, při genotypu XXY je naopak využít inaktivační mechanismus i u samců (mužů).

Další výzkum ukázal, že ve tkáních vačnatců (Cooper et al., 1993) a v placentách myši (Takagi a Sasaki, 1975; Okamoto et al., 2005) je inaktivován specificky chromosom X zděděný od otce. Molekulární mechanismy zajišťující inaktivaci specifického chromosomu X se u myši a vačnatců shodují a jsou odlišné od mechanismu náhodně inaktivujícího chromosom X u placentálů. Sdílení tohoto znaku mezi vačnatci a placentálními savci naznačuje, že tento mechanismus je evolučně starší než náhodná inaktivace. Další, recentně objevený, mechanismus navíc vyrovnává úroveň exprese genů z aktivního chromosomu X s úrovněmi autosomální exprese (Nguyen a Disteche, 2006).

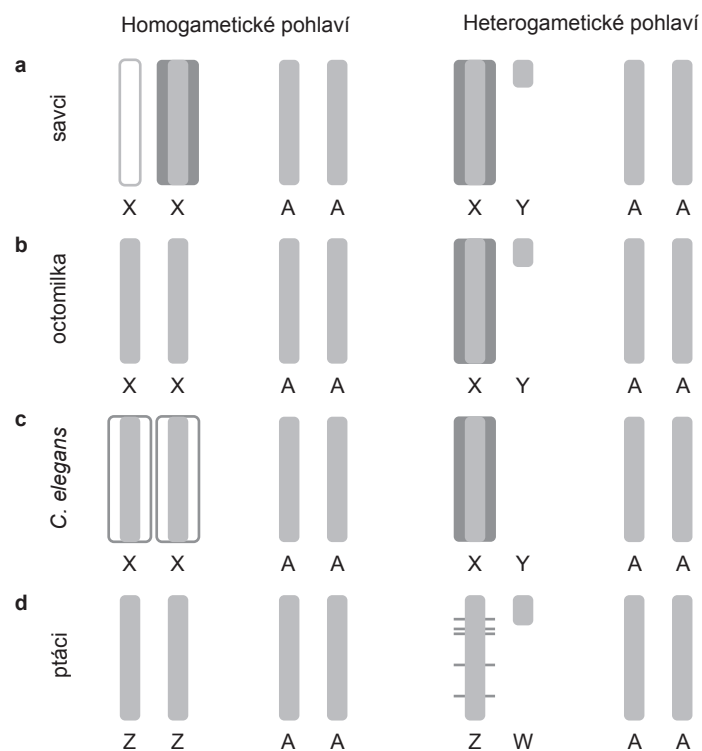
U bezobratlých nacházíme sice poněkud odlišné mechanismy kompenzující genovou dávku pro chromosom X, ale výsledek je vcelku podobný. Samci octomilky (XY) díky zvýšené transkripci jejich jediného chromosomu X dosahují zároveň vyrovnání expresních úrovní mezi pohlavími i mezi pohlavními chromosomy a autosomy (Fagegaltier a Baker, 2004; Rodriguez et al., 2007). Samci (X0) i hermafroditi (XX) háďátka (*Caenorhabditis elegans*) mají zvýšenou úroveň transkripce z chromosomu X (Gupta et al., 2006), což je dostačující pro vyrovnání expresních úrovní mezi chromosomem X a autosomy v samcích. V hermafroditech by tento mechanismus sám o sobě vedl k přílišné expresi z chromosomů X, dochází u nich proto zároveň k celkovému snížení exprese na obou chromosomech X (McDonel et al., 2006; Ercan et al., 2007), které vyrovnává expresi genů z chromosomu X a genů lokalizovaných na autosomech a zároveň vede ke srovnání hladiny exprese genů vázaných na X mezi hermafrodity a samci.

Tato škála mechanismů (obr. 6) lišících se na molekulární úrovni, ale vedoucích k velmi podobnému výsledku naznačuje, že samčí heterogamie vznikla v těchto evolučně vzdálených organismech vícekrát nezávisle na sobě. Složitě molekulární mašinerie potřebné k dosažení celochromosomové kompenzace se tedy pravděpodobně vyvinuly konvergentně jako doplněk postupné degenerace chromosomu Y. Mohlo by se zdát, že existence mechanismu kompenzujícího genovou dávku na úrovni celého chromosomu je nezbytná. To bylo ostatně dlouho považováno za primární důvod existence tohoto jevu.

Nový vhled přinesl výzkum prováděný u ptáků, konkrétně u kura domácího. Dvě nezávislé studie (Ellegren et al., 2007; Itoh et al., 2007) potvrdily, že geny lokalizované na chromosomu Z mají v buňkách samců (homogametické pohlaví, ZZ) vyšší expresní úroveň než v buňkách samic. Na chromosomu Z existují oblasti, které jsou kompenzovány lokálně, pomocí zvýšení transkripce u samic, mechanismem podobným jako u octomilky – acetylovaný lysin 16 na histonu H4 (Teranishi et al., 2001; Bioni et al., 2005). Většina genů je však bez kompenzace. Podobný nepoměr v množství transkriptů z chromosomu Z mezi samci a samicemi byl prokázán i u zebříčky (Itoh et al., 2007) a u bource morušového (Zha et al., 2009).

Protože výjimky nejsou v evoluci (a hlavně v evolučních teoriích) nic neobvyklého, nabízelo se, že ptáci budou jednou z těchto výjimek v otázce kompenzace genové dávky. Data z čínského výzkumu bource morušového a další náznaky z výzkumu motýlů (Traut et al., 2007) hovoří spíše pro obecnou absenci globální kompenzace genové dávky u organismů s homogametickými samci. Je též zřejmé, že globální (celochromosomový) mechanismus není k úspěšnému přežívání druhu zapotřebí. Je dostačující, když jsou kompenzovány pouze jednotlivé geny citlivé na změnu genové dávky (Mank, 2009).

Jaké se pro tyto rozdíly nabízí vysvětlení? Zatím neexistuje žádná teorie, která by uspokojivě vysvětlovala všechny výše zmíněné jevy zároveň. Možnost, že by se míra kompenzace genové dávky vyvíjela postupně od vzniku pohlavních chromosomů

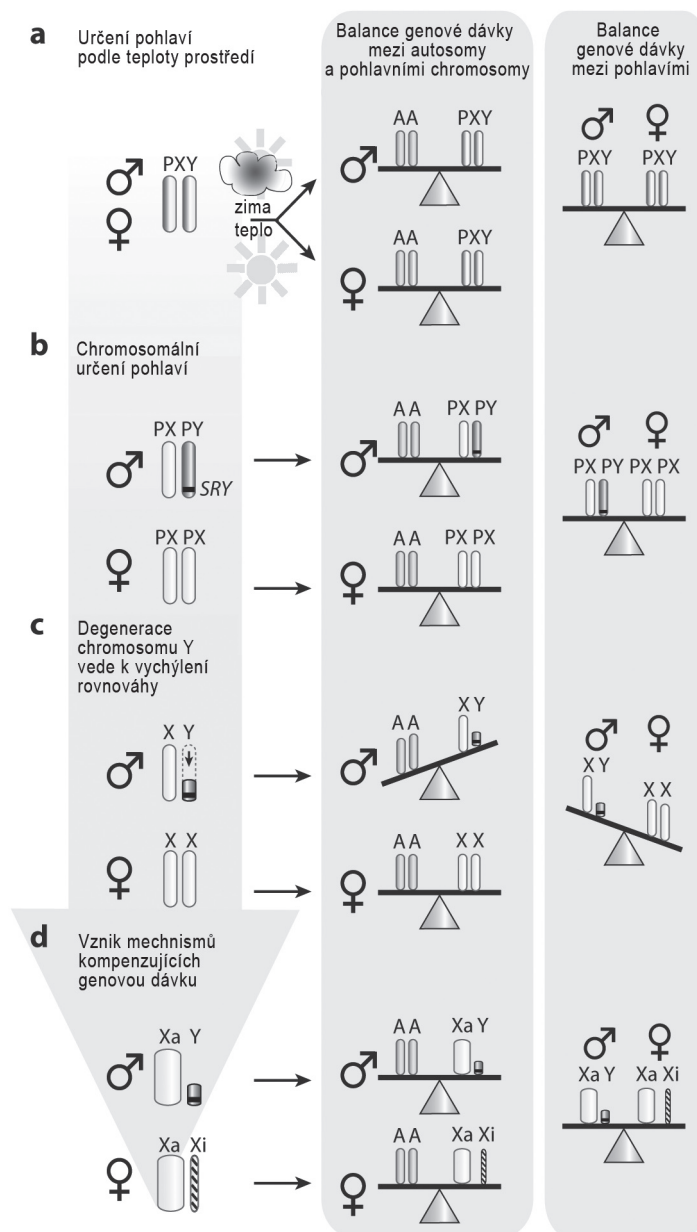


Obr. 6: Schématický přehled mechanismů kompenzujících genovou dávku u různých skupin organismů. Běžná průměrná expresní úroveň genů na chromosomu je naznačena středním odstínem šedi. Prázdný chromosom naznačuje inaktivaci, tmavý odstín šedi a rozšíření chromosomu hypertranskripci (prázdný rozšířený chromosom naznačuje druhotně regulovanou hypertranskripci). (a) u savců je jeden chromosom X u samic inaktivován, aktivní kopie chromosomu X je v obou pohlavích hypertranskribována, (b) u octomilky je hypertranskribován chromosom X u samců, (c) u hermafroditů háďátka *C. elegans* je na obou chromosomech X snížena expresní úroveň všech genů přibližně na polovinu, zároveň však pravděpodobně dochází ke zvýšení transkripce z chromosomu X jak u hermafroditů, tak u samců, (d) u ptáků dochází pouze k lokální kompenzaci genové dávky zvýšením exprese jednotlivých genů. *Upraveno podle Mank, 2009.*

(obecně akceptovaná teorie u savců, obr. 7), je vyloučena jak podobným stářím pohlavních chromosomů savčích (160 milionů let; Potrzebowski et al., 2008) a ptačích (150 milionů let; Stiglec et al., 2007), tak nesplněným předpokladem, že se kompenzace genové dávky vyvíjí zároveň s degenerací nepárového pohlavního chromosomu. Chromosom W u ptáků je degenerován do stavu velmi podobného savčímu chromosomu Y, zatímco míra kompenzace transkripčních úrovní je u těchto skupin diametrálně odlišná.

Potřeba inaktivace chromosomu X zděděného od otce v placentách savců by mohla vzniknout díky působení pohlavního antagonismu, kde zájmem otce je co největší vzrůst plodu, zájmem matky přežít ve zdraví výživu plodu a případně donosit i potomstvo dalších partnerů. Tato inaktivace pravděpodobně navazuje na meiotickou inaktivaci pohlavních chromosomů (viz další kapitolu), která probíhá při spermatogenezi (přehled v Payer a Lee, 2008). Zatímco u ptakořitných a vačnatců setrvává v inaktivovaném stavu vždy chromosom zděděný po otci, placentálové disponují evoluční novinkou – inaktivací náhodného chromosomu X. Ta se pravděpodobně vyvinula díky selekčnímu tlaku na funkční diploidii pro chromosom X u samic. Ty jsou u vačnatců a ptakořitných díky inaktivaci vždy stejného chromosomu X v podstatě hemizygotní, a tedy vystavené účinkům škodlivých recesivních mutací stejně jako samci.

Vzhledem k tomu, že většina zmíněných teorií počítá s nezbytností vyváženosti genové dávky pro samotnou životaschopnost organismu, je vhodné zmínit, že existují skupiny organismů, kde tyto teorie pravděpodobně platí velmi omezeně. U velké většiny rostlin nové druhy vznikají často polyploidizací jak ze stejných, tak z různých otcovských druhů. Z živočišné říše jsou významným příkladem blanokřídlí, u kterých je pohlaví určeno haplodiploidně, což znamená rozdílný počet ne jednoho, ale všech chromosomů. Jednou z možností je, že roli nehraje absolutní počet chromosomů, ale pouze vzájemné poměry jejich počtů.



Obr. 7: Schématický nákras vývoje potřeby kompenzace genové dávky spřažené s vývojem pohlavních chromosomů u savců. (a) pohlaví jedince je určováno podmínkami prostředí, samci se karyotypem neliší od samic (PXY – proto-XY, pár chromosomů nesoucí gen z kaskády určující pohlaví, např. *SOX3*), (b) vznikem mutace určující pohlaví nositele (např. *SRY*) se z autosomů stanou pohlavní chromosomy (PX – proto-X, PY – proto-Y, homomorfní), (c) s přibývajícím mírou degenerace chromosomu Y přibývá na chromosomu X genů, které jsou v samcích přítomné pouze v jedné kopii, což vytváří disbalanci mezi geny na chromosomu X a autosomy, a zároveň vzniká disbalance mezi geny lokalizovanými na chromosomu X v samcích a samicích, (d) pro kompenzaci těchto disbalancí se u savců vyvinuly dva mechanismy: Xa (aktivní kopie chromosomu X) je přepisována přibližně dvojnásobnou měrou, což vyrovnává transkripční úroveň mezi Xa a autosomy, druhá kopie chromosomu X je v samicích inaktivována (Xi), pravděpodobně aby hypertranskripční mechanismus nevedl k disbalanci v samicích. Který z těchto dvou mechanismů se vyvinul dříve, či zda se oba vyvinuly současně, zatím není známo. *Upraveno podle Payer a Lee, 2008.*

1.4 Meiotická inaktivace pohlavních chromosomů

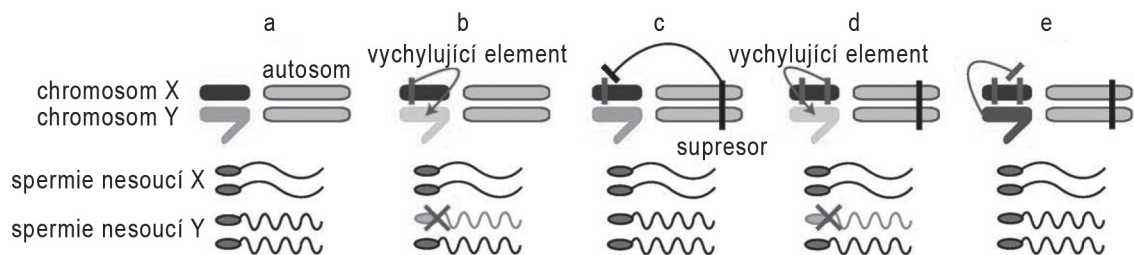
Meiotická inaktivace přináší do škály molekulárních mechanismů umlčujících pohlavní chromosomy několik dalších. Bývá běžně označována zkratkou MSCI (*meiotic sex chromosome inactivation*). MSCI začíná v průběhu pachytene, při párování homologních chromosomů. Na myším modelu bylo prokázáno, že v této fázi jsou transkripčně umlčeny nespárované oblasti jak na autosomech, tak na pohlavních chromosomech. MSCI se proto zdá být případem obecnějšího mechanismu – meiotické inaktivace nespárovaného chromatinu (Baarends et al., 2005; Turner et al., 2005). Pro organismy s degenerovaným nepárovým pohlavním chromosomem meiotická inaktivace nespárovaného chromatinu znamená transkripční umlčení téměř celého párového pohlavního chromosomu (X/Z) v heterogametickém pohlaví v průběhu meiosis.

Průběh MSCI u myši a háďátka *C. elegans*, kde byla důkladně studována, je spojen s různými epigenetickými modifikacemi histonů – deacetylace, methylace, ubiquitinizace – a také s nahrazením některých histonů za jejich méně běžné varianty. To způsobí vymizení transkripční aktivity u většiny genů a cytologicky pozorovatelnou heterochromatinizaci chromosomu. Některé změny přetrvávají až do samotné spermatogeneze (Turner, 2007). V průběhu meiosis nebo těsně po ní je transkripčně aktivních asi 20 genových rodin mikroRNA a pouze několik jednotlivých genů lokalizovaných na chromosomu X (Namekawa et al., 2006; Turner et al., 2006; Mueller et al., 2008; Song et al., 2009). Meiotická inaktivace byla prokázána též u vřeckovité houby *Neurospora crassa*, kde hlavním mechanismem inaktivace pohlavních chromosomů je RNA interference (RNAi). Protože RNAi je posttranskripční mechanismus, který funguje na základě homologního párování vláken RNA, umlčuje nejen geny z chromosomu X, ale i jejich homology lokalizované na autosomech (Shiu et al., 2001). Není zatím jasné, zda RNAi je jediným mechanismem podílejícím se na MSCI u *N. crassa*, ani zda se RNAi podílí významnějším způsobem na MSCI u savců nebo u *C. elegans*. U octomilek je důkaz existence MSCI podložen poměrně komplikovanými experimenty. U tohoto rodu totiž nedochází k heterochromatinizaci pohlavních chromosomů. Nejprve byla přítomnost MSCI navržena jako mecha-

nismus zodpovědný za pohlavně specifickou sterilitu hybridů s chromosomálními translokacemi mezi X a autosomy (Lifschytz a Lindsley, 1972). Recentní důkaz přítomnosti MSCI u octomilek využívá transgenní octomilky s vneseným reporterovým genem opatřeným promotorem specifickým pro spermatogenezi. U takto modifikovaných mušek byl prokázán rozdíl v expresi reporterového genu závislý na tom, zda byl konstrukt lokalizován na autosomech nebo pohlavních chromosomech (Hense et al., 2007).

Recentní výzkum (Schoenmakers et al., 2009) potvrdil přítomnost meiotické inaktivace také u organismů s heterogametickými samicemi – v průběhu oogeneze u kura domácího. Způsob, kterým je inaktivace dosaženo, se odlišuje od těch již známých z myši nebo *N. crassa*. Během meiosis dochází u kura k heterolognímu spárování chromosomů Z a W, což vylučuje využití mechanismu umlčujícího nespárovaný chromatin (viz výše). Během pachytene nebyly detekovány ani běžné histonové modifikace známé ze savčích modelů (fosforylace serinu 139 na histonu *H2AX*, asociovaná s výskytem dvouvláknových zlomů – Rogakou et al., 1998; Mahadevaiah et al., 2001). Tyto modifikace se objevují až v pozdějších fázích meiosis. To by mohlo naznačovat existenci selekčního tlaku směřujícího k inaktivaci pohlavních chromosomů, odlišného od pouhého důsledku nespecifické meiotické inaktivace nespárovaného chromatinu.

Nabízí se několik vysvětlení, proč k MSCI dochází. Podobně jako u fenoménu kompenzace genové dávky můžeme také u MSCI pozorovat rozmanitou škálu mechanismů vedoucích k velmi podobnému výsledku, z čehož můžeme usuzovat, že MSCI je výsledkem určitého selekčního tlaku a ne jenom průvodní jev evoluce. Důvodem vzniku obecnějšího mechanismu inaktivujícího nespárovaný chromatin může být kontrola průběhu meiosis s cílem zamezení vzniku chromosomálních abnormalit a aneuploidií. Může také zároveň sloužit jako obrana před cizorodou DNA, jako jsou transpozony a retroviry (přehled v Kelly a Aramayo, 2007; Turner, 2007). Evoluční silou vedoucí ke specifické inaktivaci pohlavních chromosomů by mohla být například selekce mechanismu potlačujícího sobecké genetické elementy, které ovlivněním meiosis vychylují poměr pohlaví potomstva (obr. 8). Pohlavní chromo-



Obr. 8: Schéma postupné (ko-)evoluce sobeckých genetických elementů vychylujících poměr pohlaví a jejich supresorů. (a) schematicky znázorněný genotyp heterogametického samce a dva typy spermií, které produkuje, (b) nově vzniklý sobecký element schopný poškozovat spermie nesoucí chromosom Y se rozšíří v populaci a tím vychýlí poměr pohlaví, (c) selekce zvýhodní nositele mutace suprimující efekt sobeckého elementu – bude mít víc synů, tím pádem i více vnuků a supresor se rozšíří v populaci, (d, e) proces se opakuje. *Upraveno podle Meiklejohn a Tao, 2010.*

somy jsou ideálním útočištěm pro tento typ sobeckých elementů a paušální umlčení celých chromosomů během inkriminovaného buněčného děje se nabízí jako vcelku účinný mechanismus (Meiklejohn a Tao, 2010).

Konkrétní zjištění podkládající tuto teorii nalezneme například u octomilek. Na modelovém organismu *D. simulans* byly popsány tři systémy elementů vychylujících poměr pohlaví potomstva a jejich supresorů: *Paris*, *Winters* a *Durham* (přehled v Tao et al., 2007). Všechny tři elementy lokalizované na chromosomu X způsobují vývojové chyby u spermií obsahujících chromosom Y, takže vychylují poměr pohlaví potomstva ve prospěch samic. Tím zvyšují své poměrné zastoupení v populaci. Systém *Paris* je dvojice genů na chromosomu X objevených při křížení geograficky oddělených populací *D. simulans* se supresory lokalizovanými na chromosomu Y a na autosomech (zatím se neví na kterých konkrétně). Podobný je systém *Winters*, sestávající z dvojice genů na chromosomu X a supresoru *Nmy* (*not much yang*) na chromosomu 3. Byl objeven v inbrední linii vzniklé z *D. simulans* a *D. sechellia*. Všechny tři geny zahrnuté v systému *Winters* jsou pravděpodobně paralogy, účinek supresoru zajišťuje RNA interference. Systém *Durham* je zatím nejméně probádaným systémem. Byl objeven při introgresi druhů *D. simulans* a *D. mauritiana*. Dynamicky se vyvíjející sobecké elementy a jejich supresory by mohly mít mnoho společného s geny hybridní sterility a v důsledku i se vznikem nových druhů. Proto by bylo zajímavé se jejich výzkumu věnovat hlouběji.

2. Cíle práce

Cílem této diplomové práce je objasnění role pohlavního antagonismu, kompenzace dávky genů a meiotické inaktivace pohlavních chromosomů při formování genového obsahu chromosomu Z u kura domácího. Předchozí studie ukázaly, že ptačí chromosom Z je obohacen o geny s preferenční expresí v samcích a ochuzen o geny vykazující preferenční expresi v samicích (Kaiser a Ellegren, 2006; Storchová a Divina, 2006). Protože se v té době předpokládalo, že u ptáků existuje kompenzace genové dávky stejně jako u ostatních probádaných organismů (i když přesný mechanismus nebyl znám), výsledky zmíněných studií byly interpretovány jako možný důsledek působení pohlavně antagonistické selekce. Později se ukázalo, že u ptáků chybí mechanismus kompenzující genovou dávku pro celé pohlavní chromosomy (Ellegren et al., 2007; Itoh et al., 2007). Proto je potřeba navrhnout jiné experimenty objasňující roli pohlavního antagonismu při formování genového obsahu chromosomu Z.

První pokus rozlišit vliv pohlavně antagonistické selekce a nekompletní kompenzace genové dávky vycházel z předpokladu, že většina pohlavního antagonismu se projeví až v dospělém věku (Mank a Ellegren, 2009). Tato studie proto využívá genové expresní čipy připravené z kuřecích embryí i dospělých jedinců, a všechna srovnání expresních úrovní jsou prováděna v rámci stejného pohlaví. Autoři studovali fyzické pozice genů vykazujících rozdílnou expresi mezi gonádami dospělců a embryí. Geny s vyšší expresí ve vaječnicích dospělých samic (čili považované za výhodné pro samice) jsou na chromosomu Z zastoupeny méně než by odpovídalo náhodnému rozdělení. Geny s nižší expresí ve vaječnicích dospělých samic (považované za nevýhodné pro samice) jsou na chromosomu Z zastoupeny naopak více. Geny exprimované ve varlatech nevykazují signifikantní rozdíly v lokalizaci mezi autosomy a chromosomem Z, jsou-li vybírány pomocí podobných kritérií. To naznačuje, že selekční tlaky na regulaci exprese jsou u samic významnější než u samců. Autoři tato data interpretují jako evidenci dominantního modelu dědičnosti pohlavně antagonistických mutací (Mank a Ellegren, 2009).

Alternativní vysvětlení pro pozorovaný odliv „samičích“ genů z chromosomu Z nabízí recentně publikovaný důkaz transkripční inaktivace chromosomu Z v prů-

běhu meiosis (Schoenmakers et al., 2009). Meiotická inaktivace chromosomu Z začíná v pachytene (ta probíhá ve všech oocytech těsně po líhnutí kuřat) a trvá několik dní až do časně diplotene (Schoenmakers et al., 2009). Přestože byla prokázána reaktivace chromosomu Z v pozdějších stádiích oogeneze, je možné, že některé epigenetické modifikace přetrvávají déle a ovlivňují expresi v oocytech i u dospělých jedinců. Lokalizace na inaktivním chromosomu není pro geny s důležitou rolí v oogenezi příliš výhodná, lze proto očekávat, že jejich zastoupení na chromosomu Z nebude nijak výrazné.

Cílem mojí práce bylo pokusit se s využitím veřejně dostupných expresních dat rozlišit důsledky působení pohlavně antagonistické selekce a meiotické inaktivace chromosomu Z na charakter genů lokalizovaných na chromosomu Z. To bylo umožněno využitím expresních dat získaných z oddělených buněčných populací oocytů (germinální buňky) a kumulárních buněk (somatické buňky vytvářející obal kolem oocytu), které dohromady tvoří jeden funkční celek – ovariální folikul. Předpokladem bylo, že pohlavně antagonistická selekce se projeví v obou typech buněk stejně. Podle míry dominance „průměrné“ pohlavně antagonistické mutace budou pak geny s preferenční expresí v daném typu buněk zastoupeny na chromosomu Z více nebo méně, než by odpovídalo náhodě. Pokud je hlavní silou formující obsah chromosomu Z jeho meiotická inaktivace, lze předpokládat, že ve svém zastoupení na chromosomu Z budou ovlivněny pouze geny s preferenční expresí v germinálních buňkách (tj. oocytech), zatímco geny s preferenční expresí v somatických tkáních nebudou ovlivněny.

3. Materiál a metody

3.1 Data z genových expresních čipů

Data pro naši analýzu byla získána z veřejně dostupné databáze Gene Expression Omnibus (<http://www.ncbi.nlm.nih.gov/geo/>, Edgar et al., 2002). Využili jsme jedno-kanálové expresní čipy firmy Affymetrix – platforma GPL3213, *Affymetrix Chicken Genome Array*. Expresní data pro somatické tkáně – mozek, srdce a játra – byla získána z experimentů, které provedl Ellegren et al. (2007) a Itoh et al. (2007). Data z vaječníků – roztríděné populace kumulárních buněk a primárních oocytů – pocházejí z experimentu, který provedl Elis et al. (2008). Všechny somatické tkáně pocházejí z kuřecích embryí, 14 nebo 18 dnů starých. Kumulární buňky pocházejí z F1 folikulů dospělých slepic, použité primární oocyty jsou sebrané záhy po ovulaci. Pro analýzu byla použita pouze data ze samičích tkání, aby byl vyloučen vliv chybějící kompenzace genové dávky u chromosomu Z na výběr tkáňově specifických genů. Seznam jednotlivých použitých hybridizačních experimentů je v tabulce 1, data jsou také na příloženém CD.

Tabulka 1: Seznam použitých expresních čipů využitých pro analýzu.

Tkáň	Identifikátory vzorků v GEO	Publikováno v
mozek	GSM157823, GSM157824, GSM157825, GSM157826, GSM157827	Itoh et al. 2007
játra	GSM157879, GSM157881, GSM157882, GSM157884, GSM157886	Itoh et al. 2007
srdce	GSM157809, GSM157811, GSM157813, GSM157815, GSM157817	Itoh et al. 2007
mozek	GSM215459, GSM215460, GSM215461, GSM215462	Ellegren et al. 2007
srdce	GSM215473, GSM215474, GSM215475, GSM215476	Ellegren et al. 2007
primární oocyty	GSM189423, GSM189424	Elis et al. 2008
kumulární buňky	GSM189428, GSM189427	Elis et al. 2008

3.2 Anotace, normalizace a sumarizace expresních dat

Ke zpracování dat byl vyžit programový balík Bioconductor verze 2.1 (<http://www.bioconductor.org>, Gentleman et al., 2004), který je rozšířením statistického systému R Project for Statistical Computing (použita verze 2.6.0, <http://www.r-project.org>). Oligonukleotidové sondy na čipech byly přiřazeny ke genovým identifikátorům databáze Ensembl (Ensembl vydání 46) za pomoci `cdf` souboru (*custom chip description file*) „`ggchickenggensg`“ verze 10.0.0 (Dai et al., 2005). Takto bylo přiřazeno 11 998 sad oligonukleotidových sond ke genům se známým umístěním na chromosomech.

Expresní úrovně pro jednotlivé geny byly vypočteny pomocí algoritmu gcRMA (*GC-content aware Robust Multiarray Average*, Wu et al., 2004). Ten využívá sekvenci každé oligonukleotidové sondy k výpočtu jejího koeficientu afinity, který vyjadřuje pravděpodobný podíl nespecificky vázané cDNA na celkovém signálu. Dále je v algoritmu gcRMA využita kvantilová normalizace a lineární aditivní model pro každou sadu sond. Použitý lineární model sestává z koeficientu afinity, logaritmované hodnoty skutečného signálu a nezávislé chyby (Irizarry et al., 2003).

3.3 Hodnocení diferenciální exprese a statistická analýza

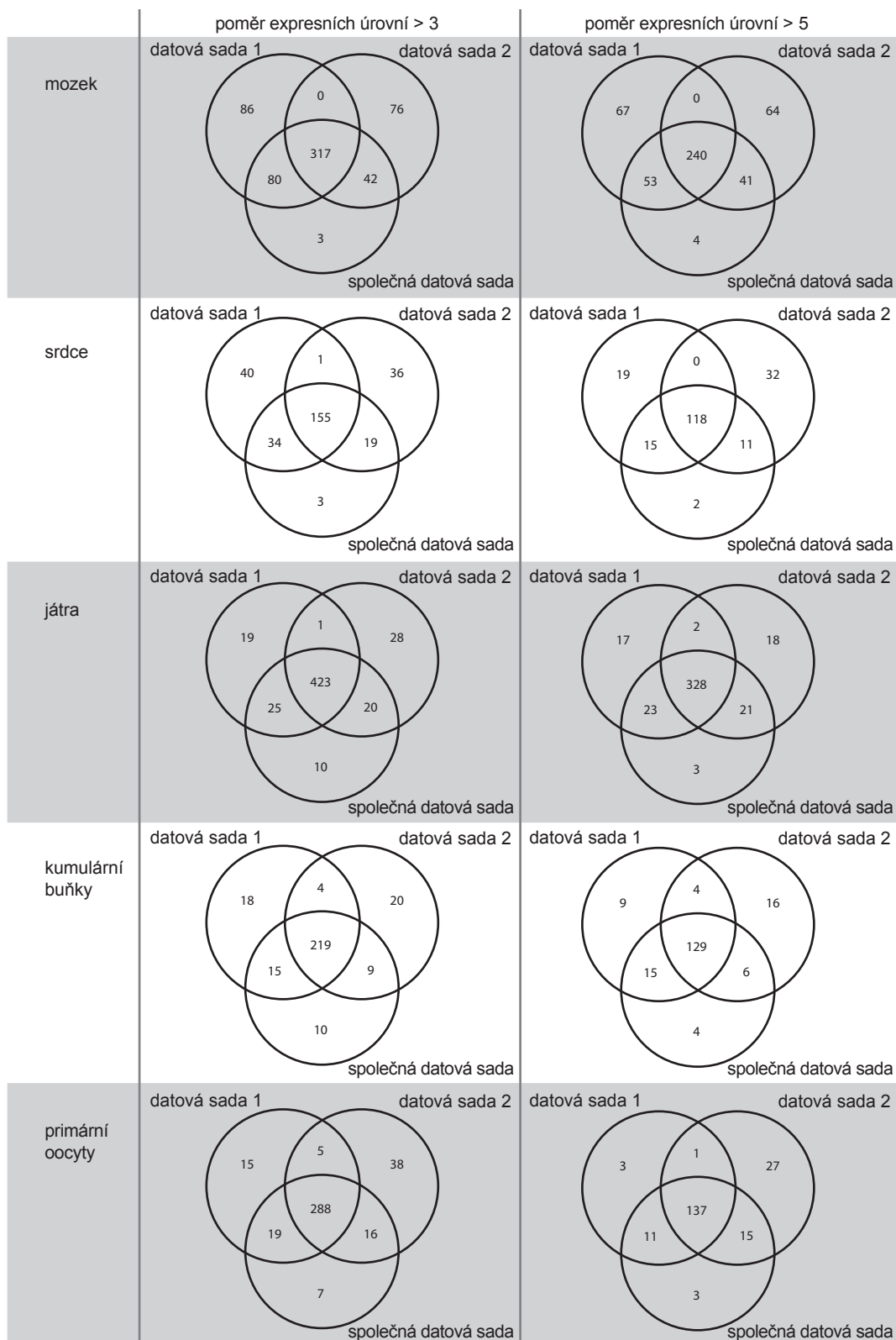
Pro rozhodování, zda je ten který gen signifikantně více exprimován v jedné tkáni oproti tkáni druhé, byl použit modul z balíku Bioconductor – `limma` (*Linear Models for Microarray Data*, verze 2.12.0, Smyth, 2004). Pro zhodnocení statistické významnosti je využit lineární model, tentokrát pro výsledné expresní úrovně genů – na rozdíl od gcRMA, kde je použit pro jednotlivé sondy. Následně je pro zvýšení síly testů použita Bayesovská statistika, využívající informace o expresi všech genů zároveň. Jako korekce na mnohonásobná porovnání mezi dvěma tkáněmi (na úrovni jednotlivých sad sond) je použita metoda *false discovery rate* (Benjamini a Hochberg, 1995). Korekce mnohonásobného porovnávání mezi více tkáněmi (na úrovni celých čipů) byla provedena pomocí metody vnořených F-testů (parametr `method="nestedF"` funkce `decideTests()`).

Geny přednostně exprimované v každé tkáni byly vybrány podle následujících kritérií: (1) signifikantně vyšší exprese v dané tkáni ve srovnání s každou další tkání, (2) poměr expresních úrovní genu v dané tkáni oproti každé další tkáni větší než tři (větší než pět), (3) hodnota expresní úrovně v dané tkáni je větší než 100 (prahová hodnota běžně užívaná u čipů firmy Affymetrix pro odlišení signálu od pozadí).

Očekávané počty genů přednostně exprimovaných v jednotlivých tkáních byly vypočítány na základě počtu anotovaných genů v použitém *cdf* souboru. Nulová hypotéza je, že umístění na chromosomu neovlivní pravděpodobnost, se kterou bude gen přednostně exprimován v některé tkáni. Očekávané a pozorované počty byly porovnány pomocí testu dobré shody (χ^2 test dobré shody). Viz také skripty v příloze 8.2 nebo na přiloženém CD.

Přestože projekt MAQC-I (*MicorArray Quality Control*) prokázal velkou míru konzistence mezi surovými daty vyprodukovanými v různých laboratořích, v procesu přípravy čipu existuje mnoho zdrojů možné variability (MAQC Consortium, 2006). Data využitá v naší analýze pocházejí z různých laboratoří, proto byl proces výběru genů s preferenční expresí v jednotlivých tkáních proveden třikrát, s částečně odlišnými vstupními datovými sadami. Ve všech analýzách byla použita data připravená z jater (Itoh et al., 2007) a z vaječnicků (kumulární buňky a ovulované oocyty, Elis et al., 2008). V první analýze byla dále zahrnuta data z mozku a srdce připravená v laboratoři A. P. Arnolda (Itoh et al., 2007). V druhé analýze byla použita data z mozku a srdce připravená v laboratoři H. Ellegrena (Ellegren et al., 2007). Ve třetí analýze byla využita všechna dostupná data pro vybrané tkáně (mozek, srdce, játra, kumulární buňky, oocyty). Překryv mezi geny určenými jako tkáňově specifické v první a druhé analýze byl 52–82 %, viz také obrázek 9.

Statistické testy rozdělení vybraných genů mezi jednotlivými chromosomy vykazovaly shodné výsledky pro všechny tři datové sady. V další práci jsou proto zmiňovány pouze výsledky pro třetí (kombinovanou) datovou sadu. Výsledky pro 1. a 2. datovou sadu jsou uvedeny v příloze 8.3 (tabulky P1, P2 a P3, obr. P1). Shodné závěry všech tří analýz naznačují, že je nepravděpodobné, aby pozorované rozdíly byly způsobeny odlišným zacházením s materiálem v různých laboratořích.



Obr. 9: Vennovy diagramy zobrazující překryvy sad genů s preferenční expresí v jednotlivých tkáních. Každý kruh reprezentuje analýzu s použitím jedné datové sady, součet čísel v každém kruhu je tedy celkový počet genů s preferenční expresí v dané tkáni pro danou datovou sadu. Číslo uprostřed vyjadřuje počet genů s preferenční expresí pro všechny datové sady.

4. Výsledky

Abychom rozlišili vlivy pohlavně antagonistické selekce a meiotické inaktivace chromosomu Z na genový obsah chromosomu Z, podrobili jsme analýze rozmístění genů preferenčně exprimovaných v různých tkáních. Analyzované tkáně lze rozdělit na pohlavně nesespecifické – mozek, srdce a játra, a specifické pro samice – kumulární buňky a ovulované oocyty. Předpokládáme, že geny s preferenční expresí ve tkáních specifických pro samice jsou geny výhodné pro samice. Geny s preferenční expresí v pohlavně nesespecifických tkáních představují kontrolní skupinu, která by neměla být ovlivněna pohlavně antagonistickou selekcí.

Tabulka 2: Pozorované a očekávané počty genů s preferenční expresí v jednotlivých tkáních na autosomech (A) a chromosomu Z (Z).

Tkáň	Poměr ^a	Pozorované			Očekávané		p ^b
		A	Z	% Z	A	Z	
mozek	3	423	19	4,30	421	21	0,8714
	5	324	14	4,14	322	16	0,8518
srdce	3	205	6	2,84	201	10	0,4445
	5	141	5	3,42	139	7	0,7681
játra	3	455	23	4,81	455	23	0,8799
	5	364	11	2,93	357	18	0,2558
kumulární buňky	3	212	41	16,21	241	12	0,0000
	5	125	29	18,83	147	7	0,0002
primární oocyty	3	316	14	4,24	314	16	0,8518
	5	158	8	4,82	158	8	0,7978

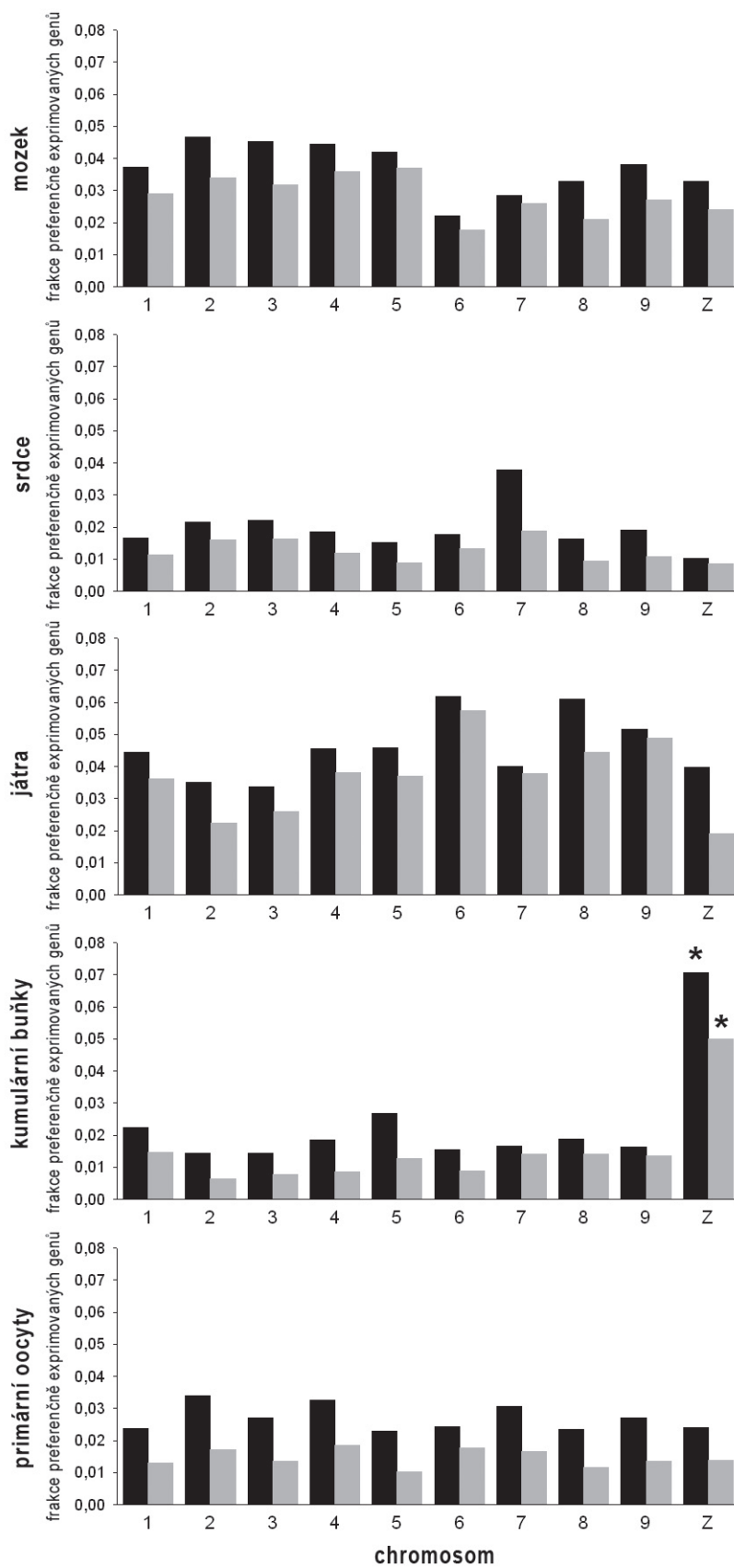
^a poměr expresních úrovní

^b test dobré shody, df=1, statisticky významné hodnoty vyznačeny tučně

Geny s preferenční expresí byly vybrány pro dvě různé limitní hodnoty poměru expresních úrovní – geny s poměrem větším než tři (volnější kritérium) a geny s poměrem větším než pět (přísnější kritérium). Celkové počty genů vybraných pro jednotlivé tkáně jsou uvedeny v tabulce 2. Geny s preferenční expresí v kumulárních buňkách jsou ve srovnání s náhodným rozdělením zastoupeny na chromosomu Z 3,4 krát více pro poměr expresních úrovní větší než tři a 4,1 krát více pro poměr expresních úrovní větší než pět. Tento nepoměr je signifikantní (test dobré shody, $p < 0,001$, viz tabulku 2). Geny s preferenční expresí v primárních oocytech a ani geny s preferenční expresí v testovaných somatických tkáních (mozek, srdce a játra)

nevykazují odlišnosti proti očekávaným počtům genů (viz tabulka 2). Rozdíl v počtu pozorovaných a očekávaných genů s preferenční expresí na chromosomu Z je signifikantní i po aplikaci korekce na mnohonásobná porovnání (Bonferroniho korekce, $p < 0,01$).

V předchozí analýze jsme se zabývali rozložením genů mezi chromosomem Z na jedné straně a autosomy na straně druhé. Taková analýza nám však neumožní odhalit jednotlivé autosomy, které by také mohly nést nenáhodný počet námi studovaných genů. V další analýze jsme proto studovali distribuci genů s preferenční expresí v jednotlivých tkáních na konkrétních chromosomech. V této analýze byly zahrnuty pouze chromosomy s více než 350 anotovanými geny, tj. chromosom Z a chromosomy 1–9. K porovnání pozorovaných a očekávaných hodnot byl opět použit test dobré shody. Jediný vzorek, ve kterém byly nalezeny počty genů signifikantně odlišné od očekávaných hodnot po provedení korekce na mnohonásobná porovnání (Bonferroniho korekce), jsou kumulární buňky (test dobré shody, $p < 0,00001$; obr. 10). Výrazně největší příspěvek k signifikanci testu měl počet genů na chromosomu Z (tabulka 3). Při vyloučení tohoto chromosomu z analýzy nebyly pozorované počty vybraných genů signifikantně odlišné od očekávaných (test dobré shody, $p > 0,05$). To napovídá, že nenáhodná distribuce genů s preferenční expresí v kumulárních buňkách je specifická pro chromosom Z. Na jednotlivých autosomech jsou tyto geny rozmístěny rovnoměrně.



Obr. 10: Zastoupení genů preferenčně exprimovaných v jednotlivých tkáních na jednotlivých chromosomech.

Zahrnuty jsou pouze chromosomy s více než 350 anotovanými geny. Černé sloupce reprezentují zastoupení genů s poměrem expresních úrovní (oproti všem ostatním tkáním) větším než tři, šedé sloupce zastoupení genů s poměrem expresních úrovní větším než pět. * značí signifikantní rozdíl oproti očekávanému počtu.

Tabulka 3: Pozorované (P) a očekávané (O) počty genů preferenčně exprimovaných v kumulárních buňkách podle jednotlivých chromosomů.

Chr	Poměr exp. úrovní 3			Poměr exp. úrovní 5		
	P	O	$(P - O)^2/O$	P	O	$(P - O)^2/O$
1	38	36	0,15	25	22	0,50
2	16	24	2,41	7	14	3,75
3	15	22	2,16	8	13	2,12
4	17	19	0,30	8	12	1,24
5	21	17	1,21	10	10	0,00
6	7	10	0,68	4	6	0,57
7	7	9	0,42	6	5	0,06
8	8	9	0,11	6	5	0,05
9	6	8	0,40	5	5	0,02
Z	41	12	67,68	29	7	62,41
Sum	176	164	75,52 *	108	100	70,70 *

* test dobré shody, $df=9$, $p < 0.000001$

Vyšší zastoupení genů s konkrétním profilem genové exprese by mohlo být způsobeno také častější genovou duplikací na chromosomu Z bez následné úpravy expresního profilu nově vzniklého genu. Podrobili jsme proto analýze všechny geny vybrané pro poměr expresních úrovní větší než tři (41 genů; tabulky P4 a P5 v příloze 8.4). Tato sada díky použitému kritériu zahrnuje i druhý výběr. Zkoumanými aspekty byla fyzická pozice na chromosomu Z a homologie k ostatním genům ve výběru. Pomocí nástroje BioMart (Haider et al., 2009) nad daty z databáze Ensembl (Birney et al., 2004) jsme zjistili, že pouze dva z 41 genů (ENSGALG00000012575 a ENSGALG00000015184) jsou paralogy. Tyto dva geny leží vedle sebe na chromosomu Z, což naznačuje, že mohly vzniknout tandemovou duplikací. V porovnání s celkovým počtem vybraných genů (41) nemohou dva paralogní geny vysvětlit nepoměrně vyšší zastoupení genů s preferenční expresí v kumulárních buňkách na chromosomu Z.

5. Diskuse

Abychom lépe porozuměli mechanismům, které utvářejí genový obsah chromosomu Z, analyzovali jsme rozmístění vybraných genů v genomu kura domácího. Byly vybrány geny preferenčně exprimované v různých pohlavně nesespecifických tkáních, v somatických buňkách vaječníku a v primárních oocytech. Zjistili jsme, že chromosom Z je obohacen o geny s preferenční expresí v kumulárních buňkách (somatické pohlavně specifické buňky). Geny preferenčně exprimované ve všech ostatních vzorcích nevykazovaly žádné statisticky signifikantní odchylky od náhodného rozmístění na chromosomech. Vyšší zastoupení vybraných genů na chromosomu Z není způsobeno genovými duplikacemi. Navrhujeme tedy pohlavně antagonistickou selekci jako hlavní mechanismus zodpovědný za obohacení chromosomu Z o geny s preferenční expresí v kumulárních buňkách. V dalším textu srovnáváme naše výsledky s výsledky předchozích studií zabývajících se genovým obsahem chromosomů X a Z.

5.1 Pohlavně antagonistická selekce a míra dominance pohlavně antagonistických mutací

Matematický model pohlavně antagonistické selekce předpovídá hromadění mutací výhodných pro samce na chromosomu X v případě, že nově vznikající mutace jsou v průměru recesivní, a naopak hromadění mutací výhodných pro samice v případě, že nově vznikající mutace jsou v průměru dominantní (Rice, 1984; viz kapitolu 1.2). Analogicky chromosom Z by měl být obohacen o mutace výhodné pro samice, pokud jsou nově vznikající mutace v průměru recesivní, nebo obohacen o mutace výhodné pro samce, pokud jsou nově vznikající mutace v průměru dominantní (v podstatě opak stavu pro chromosom X). Protože průměrná míra dominance nově vzniklé mutace není právě jednoduše zjistitelný parametr, jsou výsledky pohlavně antagonistické selekce obtížně předpověditelné (Vicoso a Charlesworth, 2006; van Doorn, 2009).

Doposud provedené experimentální studie poskytují nejednoznačné výsledky. Pro v průměru dominantní pohlavně antagonistické mutace hovoří studie Mank a Ellegren (2009). Porovnává genové expresní úrovně v celých vaječnicích z dospěl-

ců a z embryí. Geny s vyšší expresí ve vaječnicích dospělců – považovány za geny výhodné pro samice – jsou na chromosomu Z zastoupeny méně, než předpovídá náhodné rozmístění. Naopak geny s nižší expresní úrovní v dospělcích oproti embryím – považované za nevýhodné pro samice – jsou na chromosomu Z zastoupeny více, než by předpovídalo náhodné rozmístění. U octomilek je popsáno nižší zastoupení genů s vyššími expresními úrovněmi v samcích na chromosomu X (Parisi et al., 2003; Ranz et al., 2003; Sturgill et al., 2007), které je konzistentní s dominantním modelem dědičnosti pohlavně antagonistických mutací. U savců jsou ale výsledky právě opačné. Geny s vyššími expresními úrovněmi v somatických nebo pre-meiotických tkáních u samců jsou na chromosomu X zastoupeny více (Wang et al., 2001; Lercher et al., 2003; Khil et al., 2004; Divina et al., 2005). To hovoří pro v průměru recesivní nově vzniklé pohlavně antagonistické mutace (obr. 11).

	zastoupení genů na X/Z
ptáci	nižší
octomilka	nižší
savci	vyšší

Obr. 11: Zastoupení genů považovaných za výhodné pro heterogametické pohlaví na chromosomu X/Z podle dosavadních studií.

Nejednoznačnost těchto výsledků může být vysvětlena působením epigenetických mechanismů, které se mezi různými organismy mohou podstatně lišit. Nedostatek genů s vyšší expresí v samcích na chromosomu X u octomilky může být důsledkem celochromosomového zvýšení expresních úrovní, ke kterému dochází kvůli kompenzaci genové dávky u samců (Baker et al., 1994). To, že je kompenzace dosaženo pomocí globálního zvýšení transkripce, může omezit možnosti dalšího zvyšování transkripčních úrovní u jednotlivých genů v důsledku pohlavně antagonistické selekce. Chromosom X se tak stává nevhodným pro geny s vyšší expresí v samcích (Rogers et al., 2003; Vicoso a Charlesworth, 2009). U kura domácího může být nižší zastoupení genů exprimovaných pouze ve vaječnicích dospělců (oproti embryonálním vaječnicům) i vyšší zastoupení genů exprimovaných pouze v embryonálních vaječnicích (oproti vaječnicům dospělců) spíše důsledkem meiotické inaktivace chromosomu Z (Schoenmakers et al., 2009) než působení pohlavně antagonistické

selekce. Navíc zjištění, že nenáhodné rozmístění na chromosomech vykazují geny s rozdílnou expresí mezi vaječníky z dospělců a z embryí, ale nevykazují geny s rozdílnou expresí mezi varlaty z dospělců a z embryí (Mank a Ellegren, 2009), hovoří spíše pro efekt meiotické inaktivace chromosomu Z než pro důsledek působení pohlavně antagonistické selekce. Ta by měla ovlivnit obě pohlaví stejnou měrou (Ellegren a Parsch, 2007).

V této práci jsme ukázali, že geny preferenčně exprimované v kumulárních buňkách, u kterých se předpokládá výhodnost pro samice, se nacházejí na chromosomu Z častěji, než by odpovídalo náhodnému rozmístění. Toto zjištění koresponduje s výsledky získanými u savců (Wang et al., 2001; Lercher et al., 2003; Khil et al., 2004; Divina et al., 2005) a odpovídá spíše recesivnímu způsobu dědičnosti pohlavně antagonistických mutací. Zůstává otázkou, proč se podobně nenáhodné rozmístění na chromosomech netýká genů s preferenční expresí v primárních oocytech. Protože primární oocyty tvoří s kumulárními buňkami jeden funkční celek, dalo by se očekávat, že míra pohlavně antagonistické selekce bude podobná. Jako možný důvod navrhuje existenci meiotické inaktivace chromosomu Z. Ta ovlivňuje právě oocyty, ve kterých na rozdíl od kumulárních buněk probíhá meiosa.

5.2 Meiotická inaktivace pohlavních chromosomů a defeminizace chromosomu Z

Meiotická inaktivace pohlavních chromosomů (MSCI) začíná u savců během pachytene 1. meiotického dělení (McKee a Handel, 1993; Turner 2007). Po profázi I nastává částečná reaktivace, avšak 87 % genů zůstává neaktivní i po dokončení meiosis (Namekawa et al., 2006). MSCI je pravděpodobně důvodem nižšího zastoupení genů exprimovaných v savčích spermatocytech a spermatidách na chromosomu X (Emerson et al. 2004; Khil et al. 2004; Mueller et al. 2008; Potrzebowski et al. 2008). MSCI probíhá také u octomilky (Hense et al., 2007) a háďátka *C. elegans* (Kelly et al., 2002). Může přispívat k pozorovanému nižšímu zastoupení genů, které jsou více exprimované v samcích, na chromosomu X u obou druhů (Betran et al. 2002; Parisi et al. 2003; Ranz et al. 2003; Sturgill et al. 2007; Vibranovski et al. 2009 a Reinke et al. 2004).

U kura, podobně jako u ostatních organismů, MSCI nastupuje během pachyete, ale trvá pouze do časně diplotene (Schoenmakers et al., 2009). Možným vysvětlením pro tento rozdíl je srovnání velikosti a specializovanosti spermie a oocyty. Velká a vcelku generická buňka oocyty bude s větší pravděpodobností potřebovat některý z přibližně 750 genů lokalizovaných na chromosomu Z (Namekawa a Lee, 2009). Je přesto možné, že stejně jako u savců nedochází ani u ptáků k úplné reaktivaci celého chromosomu Z. I kdyby reaktivace již v diplotene byla úplná, selekční tlak způsobený periodou inaktivace by mohl být dostatečný na to, aby vyrovnal účinek pohlavně antagonistické selekce.

Rozdíl v zastoupení genů preferenčně exprimovaných v primárních oocytech a v kumulárních buňkách na chromosomu Z by mohl být způsoben tím, že efekt pohlavně antagonistické selekce, která přispívá k hromadění genů výhodných pro samice na chromosomu Z, je v primárních oocytech kompenzován efektem MSCI, která brání hromadění genů účastnících se meiosis a následného vývoje oocyty.

6. Závěr

V této práci jsme se za pomoci analýzy rozmístění genů preferenčně exprimovaných v několika somatických tkáních, kumulárních buňkách a primárních oocytech kura domácího pokusili objasnit vlivy pohlavně antagonistické selekce a meiotické inaktivace pohlavních chromosomů na utváření genového obsahu chromosomu Z. Jedinou odchylku od náhodného rozmístění na chromosomech vykazují geny přednostně exprimované v kumulárních buňkách. Geny přednostně exprimované v primárních oocytech ani geny přednostně exprimované v somatických tkáních odchylky od náhodného rozmístění na chromosomech nevykazují. Tyto výsledky jsme interpretovali jako důsledek pohlavně antagonistické selekce s průměrně recesivním modelem dědičnosti nově vznikajících pohlavně antagonistických mutací, způsobující hromadění mutací vhodných pro samice na chromosomu Z. Působení pohlavně antagonistické selekce je v primárních oocytech, kde by se dal očekávat stejný selekční tlak jako v kumulárních buňkách, zbrzděno meiotickou inaktivací pohlavních chromosomů.

Výsledky této práce byly publikovány v odborném časopise *Journal of Molecular Evolution*. Článek je přetištěn v příloze k této práci (příloha 8.1).

MOŘKOVSKÝ L, STORCHOVÁ R, PLACHÝ J, IVÁNEK R, DIVINA P, HEJNAR J. The Chicken Z Chromosome Is Enriched for Genes with Preferential Expression in Ovarian Somatic Cells. *J. Mol. Evol.* 2010 ;**70(2)**:129–136.

K získání přesnější představy o vlivu meiotické inaktivace pohlavních chromosomů na genový obsah chromosomu Z by bylo vhodné podrobněji prozkoumat změny genových expresních úrovní v průběhu embryonálního vývoje. Dále by bylo vhodné ověřit kvalitu genových expresních dat generovaných čipy Affymetrix, například pomocí RNA-Seq (přehled ve Wang et al., 2009).

Tato práce byla podpořena juniorským grantem Grantové agentury Akademie věd ČR (KJB501110701) řešeným Radkou Storchovou.

7. Použitá literatura

- BAARENDS WM, WASSENAAR E, VAN DER LAAN R, HOGERBRUGGE J, SLEDDENS-LINKELS E, HOEIJMAKERS JHJ, DE BOER P, GROOTEGOED JA.** Silencing of unpaired chromatin and histone H2A ubiquitination in mammalian meiosis. *Mol. Cell. Biol.* 2005;**25**(3):1041–1053.
- BAKER BS, GORMAN M, MARIN I.** Dosage compensation in *Drosophila*. *Annu. Rev. Genet.* 1994;**28**(1):491–521.
- BARR ML, BERTRAM EG.** A morphological distinction between neurones of the male and female, and the behaviour of the nucleolar satellite during accelerated nucleoprotein synthesis. *Nature* 1949;**163**(4148):676–677.
- BENJAMINI Y, HOCHBERG Y.** Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. Ser. B-Stat. Methodol.* 1995;**57**(1):289–300.
- BIRNEY E, ANDREWS TD, BEVAN P, CACCAMO M, CHEN Y, CLARKE L, COATES G, CUFF J, CURWEN V, CUTTS T, DOWN T, EYRAS E, FERNANDEZ-SUAREZ XM, GANE P, GIBBINS B, GILBERT J, HAMMOND M, HOTZ H, IYER V, JEKOSCH K, KAHARI A, KASPRZYK A, KEEFE D, KEENAN S, LEHVASLAIHO H, McVICKER G, MELSOPP C, MEIDL P, MONGIN E, PETTETT R, POTTER S, PROCTOR G, RAE M, SEARLE S, SLATER G, SMEDLEY D, SMITH J, SPOONER W, STABENAU A, STALKER J, STOREY R, URETA-VIDAL A, WOODWARK KC, CAMERON G, DURBIN R, COX A, HUBBARD T, CLAMP M.** An overview of Ensembl. *Genome Res.* 2004;**14**(5):925–928.
- BISONI L, BATTLE-MORERA L, BIRD AP, SUZUKI M, McQUEEN HA.** Female-specific hyperacetylation of histone H4 in the chicken Z chromosome. *Chromosome Res.* 2005;**13**(2):205–214.
- BLACKBURN KB.** The cytological aspects of the determination of sex in the dioecious forms of *Lychnis*. *J. Exp. Biol.* 1924;**1**(3):413.
- BULL JJ.** Evolution of sex determining mechanisms. *California: Menlo Park* 1983.
- COOPER DW, JOHNSTON PG, WATSON JM, GRAVES JAM.** X-inactivation in marsupials and monotremes. In: *Seminars in Developmental Biology*. 1993 p. 117–128.
- DAI M, WANG P, BOYD AD, KOSTOV G, ATHEY B, JONES EG, BUNNEY WE, MYERS RM, SPEED TP, AKIL H, WATSON SJ, MENG F.** Evolving gene/transcript definitions significantly alter the interpretation of GeneChip data. *Nucleic Acids Res.* 2005;**33**(20):e175.
- DIVINA P, VLCEK C, STRNAD P, PACES V, FOREJT J.** Global transcriptome analysis of the C57BL/6J mouse testis by SAGE: evidence for nonrandom gene order. *BMC Genomics* 2005;**6**(1):29.
- EDGAR R, DOMRACHEV M, LASH AE.** Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Res.* 2002;**30**(1):207–210.
- ELIS S, BATELLIER F, COUTY I, BALZERGUE S, MARTIN-MAGNIETTE M, MONGET P, BLESBOIS E, GOVOROUN M.** Search for the genes involved in oocyte maturation and early embryo development in the hen. *BMC Genomics* 2008;**9**(1):110.
- ELLEGREN H, PARSCH J.** The evolution of sex-biased genes and sex-biased gene expression. *Nat. Rev. Genet.* 2007;**8**(9):689–698.
- ELLEGREN H, HULTIN-ROSENBERG L, BRUNSTROM B, DENCKER L, KULTIMA K, SCHOLZ B.** Faced with inequality: chicken do not have a general dosage compensation of sex-linked genes. *BMC Biology* 2007;**5**(1):40.
- EMERSON JJ, KAESSMANN H, BETRAN E, LONG M.** Extensive gene traffic on the mammalian X chromosome. *Science* 2004;**303**(5657):537–540.
- ERCAN S, GIRESI PG, WHITTLE CM, ZHANG X, GREEN RD, LIEB JD.** X chromosome repression by localization of the *C.elegans* dosage compensation machinery to sites of transcription initiation. *Nature Genet.* 2007;**39**(3):403–408.
- FAGEGALTIER D, BAKER BS.** X chromosome sites autonomously recruit the dosage compensation complex in *Drosophila* males. *PLoS. Biol.* 2004;**2**(11):e341.
- FRASER JA, HEITMAN J.** Evolution of fungal sex chromosomes. *Mol. Microbiol.* 2004;**51**(2):299–306.

- GENTLEMAN RC, CAREY VJ, BATES DM, BOLSTAD B, DETTLING M, DUDDOIT S, ELLIS B, GAUTIER L, GE Y, GENTRY J, HORNIK K, HOTHORN T, HUBER W, IACUS S, IRIZARRY R, LEISCH F, LI C, MAECHLER M, ROSSINI AJ, SAWITZKI G, SMITH C, SMYTH G, TIERNEY L, YANG JY, ZHANG J.** Bioconductor: open software development for computational biology and bioinformatics. *Genome Biol.* 2004;**5(10)**:R80–R80.
- GRAVES JAM, PEICHEL CL.** Are homologies in vertebrate sex determination due to shared ancestry or to limited options? *Genome Biol.* 2010;**11(4)**:205.
- GRAVES JA, KOINA E, SANKOVIC N.** How the gene content of human sex chromosomes evolved. *Curr. Opin. Genet. Dev.* 2006;**16(3)**:219–224.
- GUPTA V, PARISI M, STURGILL D, NUTTALL R, DOCTOLERO M, DUDKO O, MALLEY J, EASTMAN PS, OLIVER B.** Global analysis of X-chromosome dosage compensation. *J. Biol.* 2006;**5(1)**:3.
- GURBICH TA, BACHTROG D.** Gene content evolution on the X chromosome. *Curr. Opin. Genet. Dev.* 2008;**18(6)**:493–498.
- GUTTMAN DS, CHARLESWORTH D.** An X-linked gene with a degenerate Y-linked homologue in a dioecious plant. *Nature* 1998;**393(6682)**:263–266.
- HAIDER S, BALLESTER B, SMEDLEY D, ZHANG J, RICE P, KASPRZYK A.** BioMart Central Portal—unified access to biological data. *Nucleic Acids Res.* 2009;**37(suppl_2)**:W23–27.
- HENSE W, BAINES JF, PARSCH J.** X Chromosome Inactivation during *Drosophila* Spermatogenesis. *PLoS. Biol.* 2007 Oct;**5(10)**:e273.
- INTERNATIONAL CHICKEN GENOME SEQUENCING CONSORTIUM.** Sequence and comparative analysis of the chicken genome provide unique perspectives on vertebrate evolution. *Nature* 2004;**432(7018)**:695–716.
- IRIZARRY RA, HOBBS B, COLLIN F, BEAZER-BARCLAY YD, ANTONELLIS KJ, SCHERF U, SPEED TP.** Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics* 2003;**4(2)**:249–264.
- ITOH Y, MELAMED E, YANG X, KAMPF K, WANG S, YEHYA N, VAN NAS A, REPLOGLE K, BAND M, CLAYTON D, SCHADT E, LUSIS A, ARNOLD A.** Dosage compensation is less effective in birds than in mammals. *J. Biol.* 2007;**6(1)**:2.
- KAISER VB, ELLEGREN H.** Nonrandom distribution of genes with sex-biased expression in the chicken genome. *Evolution* 2006;**60(9)**:1945–1951.
- KHIL PP, SMIRNOVA NA, ROMANIENKO PJ, CAMERINI-OTERO RD.** The mouse X chromosome is enriched for sex-biased genes not subject to selection by meiotic sex chromosome inactivation. *Nature Genet.* 2004;**36(6)**:642–646.
- LARDON A, GEORGIEV S, AGHMIR A, LE MERRER G, NEGRUTIU I.** Sexual dimorphism in white campion: complex control of carpel number is revealed by Y chromosome deletions. *Genetics* 1999;**151(3)**:1173–1185.
- LERCHER MJ, URRUTIA AO, HURST LD.** Evidence that the human X chromosome is enriched for male-specific but not female-specific genes. *Mol. Biol. Evol.* 2003;**20(7)**:1113–1116.
- LIFSCHYTZ E, LINDSLEY DL.** The role of X-chromosome inactivation during spermatogenesis. *Proc. Natl. Acad. Sci. U. S. A.* 1972;**69(1)**:182–186.
- LYON MF.** Sex chromatin and gene action in the mammalian X-chromosome. *Am. J. Hum. Genet.* 1962;**14(2)**:135.
- MAHADEVAIAH SK, TURNER JM, BAUDAT F, ROGAKOU EP, DE BOER P, BLANCO-RODRIGUEZ J, JASIN M, KEENEY S, BONNER WM, BURGOYNE PS.** Recombinational DNA double-strand breaks in mice precede synapsis. *Nature Genet.* 2001;**27(3)**:271–276.
- MANK JE.** The W, X, Y and Z of sex-chromosome dosage compensation. *Trends Genet.* 2009;**25(5)**:226–233.
- MANK JE, ELLEGREN H.** Sex-linkage of sexually antagonistic genes is predicted by female, but not male, effects in birds. *Evolution* 2009;**63(6)**:1464–1472.
- MANK JE, HULTIN-ROSENBERG L, WEBSTER MT, ELLEGREN H.** The unique genomic properties of sex-biased genes: insights from avian microarray data. *BMC Genomics* 2008;**9(1)**:148.
- MAQC CONSORTIUM.** The MicroArray Quality Control (MAQC) project shows inter- and intraplatform reproducibility of gene expression measurements. *Nat. Biotechnol.* 2006;**24(9)**:1151–1161.
- MCDONEL P, JANS J, PETERSON BK, MEYER BJ.** Clustered DNA motifs mark X chromosomes for repression by a dosage compensation complex. *Nature* 2006;**444(7119)**:614–618.

- McKEE BD, HANDEL MA.** Sex chromosomes, recombination, and chromatin conformation. *Chromosoma* 1993;**102(2)**:71–80.
- MEIKLEJOHN CD, TAO Y.** Genetic conflict and sex chromosome evolution. *Trends Ecol. Evol.* 2010;**25(4)**:215–223.
- MUELLER JL, MAHADEVAIAH SK, PARK PJ, WARBURTON PE, PAGE DC, TURNER JMA.** The mouse X chromosome is enriched for multicopy testis genes showing postmeiotic expression. *Nature Genet.* 2008;**40(6)**:794–799.
- NAMEKAWA SH, LEE JT.** XY and ZW: Is Meiotic Sex chromosome inactivation the rule in evolution? *PLoS Genet.* 2009;**5(5)**
- NAMEKAWA SH, PARK PJ, ZHANG L, SHIMA JE, McCARREY JR, GRISWOLD MD, LEE JT.** Postmeiotic sex chromatin in the male germline of mice. *Curr. Biol.* 2006;**16(7)**:660–667.
- NAURIN S, HANSSON B, BENSCH S, HASSELQUIST D.** Why does dosage compensation differ between XY and ZW taxa? *Trends Genet.* 2010;**26(1)**:15–20.
- NGUYEN DK, DISTECHE CM.** Dosage compensation of the active X chromosome in mammals. *Nature Genet.* 2006;**38(1)**:47–53.
- OGATA M, HASEGAWA Y, OHTANI H, MINEYAMA M, MIURA I.** The ZZ/ZW sex-determining mechanism originated twice and independently during evolution of the frog, *Rana rugosa*. *Heredity* 2007;**100(1)**:92–99.
- OHNO S.** Sex chromosomes and sex-linked genes. *Springer Berlin* 1967.
- OKAMOTO I, OTTE AP, ALLIS CD, REINBERG D, HEARD E.** Epigenetic dynamics of imprinted X inactivation during early mouse development. *Science* 2004;**303(5658)**:644.
- PAINTER T.** The Y-chromosome in mammals. *Science* 1912;503–504.
- PARISI M, NUTTALL R, NAIMAN D, BOUFFARD G, MALLEY J, ANDREWS J, EASTMAN S, OLIVER B.** Paucity of genes on the drosophila X chromosome showing male-biased expression. *Science* 2003;**299(5607)**:697–700.
- PAYER B, LEE JT.** X chromosome dosage compensation: how mammals keep the balance. *Annu. Rev. Genet.* 2008;**42**:733–772.
- POLLACK JR, SØRLIE T, PEROU CM, REES CA, JEFFREY SS, LONNING PE, TIBSHIRANI R, BOTSTEIN D, BØRRESEN-DALE A, BROWN PO.** Microarray analysis reveals a major direct role of DNA copy number alteration in the transcriptional program of human breast tumors. *Proc. Natl. Acad. Sci. U. S. A.* 2002;**99(20)**:12963–12968.
- POTRZEBOWSKI L, VINCKENBOSCH N, MARQUES AC, CHALMEL F, JÉGOU B, KAESSMANN H.** Chromosomal gene movements reflect the recent origin and biology of therian sex chromosomes. *PLoS. Biol.* 2008;**6(4)**:e80.
- RANZ JM, CASTILLO-DAVIS CI, MEIKLEJOHN CD, HARTL DL.** Sex-dependent gene expression and evolution of the drosophila transcriptome. *Science* 2003;**300(5626)**:1742–1745.
- RICE WR.** Sex chromosomes and the evolution of sexual dimorphism. *Evolution* 1984;**38(4)**: pp. 735–742.
- ROGAKOU EP, PILCH DR, ORR AH, IVANOVA VS, BONNER WM.** DNA double-stranded breaks induce histone H2AX phosphorylation on serine 139. *J. Biol. Chem.* 1998;**273(10)**:5858–5868.
- ROGERS DW, CARR M, POMIANKOWSKI A.** Male genes: X-pelled or X-cluded? *Bioessays* 2003;**25(8)**:739–741.
- SHIU PK, RAJU NB, ZICKLER D, METZENBERG RL.** Meiotic silencing by unpaired DNA. *Cell* 2001 Dec;**107(7)**:905–916.
- SCHOENMAKERS S, WASSENAAR E, HOOGERBRUGGE JW, LAVEN JSE, GROOTEGOED JA, BAARENS WM.** Female meiotic sex chromosome inactivation in chicken. *PLoS Genet.* 2009;**5(5)**: e1000466.
- SMYTH GK.** Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Stat. Appl. Genet. Mol. Biol.* 2004;**3**:Article3.
- SONG R, RO S, MICHAELS JD, PARK C, McCARREY JR, YAN W.** Many X-linked microRNAs escape meiotic sex chromosome inactivation. *Nature Genet.* 2009;**41(4)**:488–493.
- STIGLEC R, EZAZ T, GRAVES J.** A new look at the evolution of avian sex chromosomes. *Cytogenet. Genome Res.* 2007;**117(1–4)**:103–109.

- STORCHOVA R, DIVINA P.** Nonrandom representation of sex-biased genes on chicken Z chromosome. *J. Mol. Evol.* 2006;**63(5)**:676–681.
- STURGILL D, ZHANG Y, PARISI M, OLIVER B.** Demasculinization of X chromosomes in the *Drosophila* genus. *Nature* 2007;**450(7167)**:238–241.
- TAKAGI N, SASAKI M.** Preferential inactivation of the paternally derived X chromosome in the extraembryonic membranes of the mouse. 1975.
- TAO Y, ARARIPE L, KINGAN SB, KE Y, XIAO H, HARTL DL.** A sex-ratio meiotic drive system in *Drosophila simulans*. II: An X-linked distorter. *PLoS Biol.* 2007;**5(11)**:e293.
- TERANISHI M, SHIMADA Y, HORI T, NAKABAYASHI O, KIKUCHI T, MACLEOD T, PYM R, SHELDON B, SOLOVEI I, MACGREGOR H.** Transcripts of the MHM region on the chicken Z chromosome accumulate as non-coding RNA in the nucleus of female cells adjacent to the DMRT1 locus. *Chromosome Res.* 2001;**9(2)**:147–165.
- TRAUT W.** The evolution of sex chromosomes in insects: differentiation of sex chromosomes in flies and moths. *Eur. J. Entomol.* 1999;**96**:227–236.
- TRAUT W, SAHARA K, MAREC F.** Sex Chromosomes and Sex Determination in Lepidoptera. *Sex. Dev.* 2007;**1(6)**:332–346.
- TURNER JMA, MAHADEVAIAH SK, FERNANDEZ-CAPETILLO O, NUSSENZWEIG A, XU X, DENG C, BURGOYNE PS.** Silencing of unsynapsed meiotic chromosomes in the mouse. *Nature Genet.* 2005;**37(1)**:41–47.
- TURNER JM, MAHADEVAIAH SK, ELLIS PJ, MITCHELL MJ, BURGOYNE PS.** Pachytene asynapsis drives meiotic sex chromosome inactivation and leads to substantial postmeiotic repression in spermatids. *Dev. Cell* 2006;**10(4)**:521–529.
- VAN DOORN GS.** Intralocus sexual conflict. *Ann. NY Acad. Sci.* 2009;**1168(The Year in Evolutionary Biology 2009)**:52–71.
- VICOSO B, CHARLESWORTH B.** Evolution on the X chromosome: unusual patterns and processes. *Nat. Rev. Genet.* 2006;**7(8)**:645–653.
- VICOSO B, CHARLESWORTH B.** The deficit of male-biased genes on the *D. melanogaster* X chromosome is expression-dependent: a consequence of dosage compensation? *J. Mol. Evol.* 2009;**68(5)**:576–583.
- VON HENKING H.** Untersuchungen über die ersten Entwicklungsvorgänge in den Eiern der Insekten. II. Ueber Spermatogenese und deren Beziehung zur Entwicklungen bei *Pyrrhocoris apterus*. *Zeit. wiss. Zool.* 1891;**51**
- VYSKOT B, HOBZA R.** Gender in plants: sex chromosomes are emerging from the fog. *Trends Genet.* 2004;**20(9)**:432–438.
- WATERS PD, WALLIS MC, MARSHALL GRAVES JA.** Mammalian sex—origin and evolution of the Y chromosome and SRY. *Semin. Cell Dev. Biol.* 2007; **18**:389–400
- WANG PJ, MCCARREY JR, YANG F, PAGE DC.** An abundance of X-linked genes expressed in spermatogonia. *Nature Genet.* 2001;**27(4)**:422–426.
- WANG Z, GERSTEIN M, SNYDER M.** RNA-Seq: a revolutionary tool for transcriptomics. *Nat. Rev. Genet.* 2009;**10(1)**:57–63.
- WU Z, IRIZARRY RA, GENTLEMAN R, MARTINEZ-MURILLO F, SPENCER F.** A model-based background adjustment for oligonucleotide expression arrays. *J. Am. Stat. Assoc.* 2004;**99(468)**:909–917.
- YANG X, SCHADT EE, WANG S, WANG H, ARNOLD AP, INGRAM-DRAKE L, DRAKE TA, LUSIS AJ.** Tissue-specific expression and regulation of sexually dimorphic genes in mice. *Genome Res.* 2006;**16(8)**:995.
- ZHA X, XIA Q, DUAN J, WANG C, HE N, XIANG Z.** Dosage analysis of Z chromosome genes using microarray in silkworm, *Bombyx mori*. *Insect Biochem. Mol. Biol.* 2009;**39(5–6)**:315–321.

8. Přílohy

8.1 Článek publikovaný v *Journal of Molecular Evolution*

Článek je přetištěn na následujících osmi stranách.

MOŘKOVSKÝ L, STORCHOVÁ R, PLACHÝ J, IVÁNEK R, DIVINA P, HEJNAR J. The Chicken Z Chromosome Is Enriched for Genes with Preferential Expression in Ovarian Somatic Cells. *J. Mol. Evol.* 2010;**70(2)**:129–136.

The Chicken Z Chromosome Is Enriched for Genes with Preferential Expression in Ovarian Somatic Cells

Libor Mořkovský · Radka Storchová · Jiří Plachý · Robert Ivánek · Petr Divina · Jiří Hejnar

Received: 25 September 2009 / Accepted: 15 December 2009 / Published online: 27 December 2009
© Springer Science+Business Media, LLC 2009

Abstract Theory predicts that sexually antagonistic mutations will be over- or under-represented on the X and Z chromosomes, depending on their average dominance coefficients. However, as little is known about the dominance coefficients for new mutations, the effect of sexually antagonistic selection is difficult to predict. To elucidate the role of sexually antagonistic selection in the evolution of Z chromosome gene content in chicken, we analyzed publicly available microarray data from several somatic tissues as well as somatic and germ cells of the ovary. We found that the Z chromosome is enriched for genes showing preferential expression in ovarian somatic cells, but not for genes with preferential expression in primary oocytes or non-sex-specific somatic tissues. Our results suggest that sexual antagonism leads to a higher abundance of female-benefit alleles on the Z chromosome. No bias toward Z-linkage for oocyte-enriched genes can be explained by lower intensity of sexually antagonistic

selection in ovarian germ cells compared to ovarian somatic cells. An alternative explanation would be that meiotic Z chromosome inactivation hinders accumulation of oocyte-expressed genes on the Z chromosome. Our results are consistent with findings in mammals and indicate that recessive rather than dominant sexually antagonistic mutations shape the gene content of the X and Z chromosomes.

Keywords Z chromosome · Sexual antagonism · Meiotic sex chromosome inactivation · Birds · Oogenesis · Gene expression

Introduction

Homogametic sex chromosomes, X and Z, are characterized by nonrandom content of genes with preferential or specific expression in one sex (hereafter sex-biased genes). This phenomenon has been described in various organisms, including human (Hurst and Randerson 1999; Lercher et al. 2003; Saifi and Chandra 1999), mouse (Divina et al. 2005; Khil et al. 2004; Mueller et al. 2008; Wang et al. 2001), *Drosophila* (Parisi et al. 2003; Ranz et al. 2003; Sturgill et al. 2007), *C. elegans* (Reinke et al. 2004), chicken (Kaiser and Ellegren 2006; Storchova and Divina 2006) and silkworm (Arunkumar et al. 2009). Interestingly, the proportion of sex-biased genes on the sex chromosomes is not consistent across various tissues and/or taxa (reviewed in Gurbich and Bachtrog 2008; Vicoso and Charlesworth 2006). For instance, the mammalian X chromosome is enriched for male-biased genes expressed in somatic tissues, while genes expressed in spermatogenic cells during meiosis are less abundant on this chromosome (Khil et al. 2004). In contrast, the *Drosophila* X chromosome is

Libor Mořkovský and Radka Storchová have contributed equally to this work.

Electronic supplementary material The online version of this article (doi:10.1007/s00239-009-9315-5) contains supplementary material, which is available to authorized users.

L. Mořkovský · R. Storchová (✉)
Department of Zoology, Faculty of Science, Charles University in Prague, Viničná 7, 128 44 Prague 2, Czech Republic
e-mail: radkas@natur.cuni.cz

L. Mořkovský
Department of Cell Biology, Faculty of Science,
Charles University in Prague, Prague, Czech Republic

J. Plachý · R. Ivánek · P. Divina · J. Hejnar
Institute of Molecular Genetics, Academy of Sciences of the
Czech Republic, Prague, Czech Republic

depleted of male-biased genes regardless of their tissue expression patterns (Parisi et al. 2003; Sturgill et al. 2007). In chicken, the Z chromosome is depleted of female-biased genes in somatic as well as germinal tissues (Kaiser and Ellegren 2006; Storchova and Divina 2006).

The observed inconsistencies indicate that the sex chromosomes are shaped by several evolutionary forces whose impact differs among tissues and/or taxa. One of these forces is sexually antagonistic selection that leads to nonrandom genomic distribution of mutations with sex-biased fitness effects (Rice 1984). The accumulation of sexually antagonistic mutations in a gene will eventually lead to selection for modifier mutations that alter expression in a sex-specific manner, thus leading to genes with sex-biased expression (Connallon and Knowles 2005; Mank 2009). Sexually antagonistic selection can lead either to over-representation or under-representation of sex-biased genes on the X/Z chromosome, depending on the average dominance of sexually antagonistic mutations (Ellegren and Parsch 2007). The other factors that can affect the genomic distribution of sex-biased genes include meiotic sex chromosome inactivation in the heterogametic sex (Khil et al. 2004; Mueller et al. 2008), epigenetic modifications of the X chromosome associated with dosage compensation (Rogers et al. 2003; Vicoso and Charlesworth 2009) and incomplete dosage compensation in ZW species (Arnold et al. 2008; Ellegren et al. 2007; Itoh et al. 2007; Zha et al. 2008).

Disentangling the relative importance of various selective forces shaping the X/Z chromosome gene content might be a difficult task. Recently, Mank and Ellegren (2009b) introduced the first attempt to decouple the effect of incomplete dosage compensation and sexually antagonistic selection on the Z chromosome gene content in chicken. They studied genomic distribution of genes with differential expression between adult and embryonic gonads, which are expected to differ in the intensity of sexually antagonistic selection. Their results revealed that genes that are up-regulated in the adult ovary (i.e., female-benefit genes) are less abundant on the Z chromosome, while genes down-regulated in the adult ovary (i.e., female-detriment genes) are enriched on Z (Mank and Ellegren 2009b). These results are consistent with a dominant mode of inheritance for sexually antagonistic mutations. Under such scenario, female-benefit genes are expected to be under-represented on the Z chromosome, since this chromosome spends less time in females than in males. However, there is an alternative explanation that meiotic Z chromosome inactivation, i.e. transcriptional silencing of the Z chromosome during oogenesis (Schoenmakers et al. 2009), results in paucity of genes expressed in adult ovary on the Z chromosome. Meiotic Z inactivation begins in pachytene that occurs just after hatching and lasts a few

days until early diplotene (Schoenmakers et al. 2009). Although reactivation of the Z chromosome has been demonstrated in the later stages of oogenesis, it is possible that some epigenetic modifications persist and influence gene expression also in adult ovaries (Schoenmakers et al. 2009).

In this study, we analyzed publicly available microarray data from several somatic tissues as well as somatic and germ cells of the adult ovary. These data allowed us to distinguish between the effects of sexually antagonistic selection and meiotic sex chromosome inactivation on the Z chromosome gene content. We assumed that sexually antagonistic selection will result in nonrandom genomic distribution of genes with preferential expression in somatic as well as germ cells of the ovary, as both groups of genes are expected to have female-benefit effect. These genes can be either less or more abundant on the Z chromosome depending on the average dominance of sexually antagonistic mutations. On the other hand, meiotic Z chromosome inactivation is expected to result in lower abundance of genes expressed in ovarian germ cells but should not affect genomic distribution of genes expressed in ovarian somatic cells or other somatic tissues.

Methods

Chicken Microarray Data

We used chicken microarray data publicly available in Gene Expression Omnibus repository (<http://www.ncbi.nlm.nih.gov/ GEO/>—Affymetrix Chicken Genome Array platform GPL3213) from several somatic tissues including brain, heart, and liver (Ellegren et al. 2007; Itoh et al. 2007) and sorted populations of primary oocytes and granulosa cells (Elis et al. 2008). All somatic tissues were obtained from 14 or 18 days old embryos. Granulosa cells were collected from mature preovulatory (F1) follicles of adult hens. Primary oocytes came from F1 follicles and just ovulated oocytes. Since populations of F1 oocytes were contaminated with granulosa cells (Elis et al. 2008), we used only data from ovulated primary oocytes that were contamination-free in our analysis. Only microarray data from female individuals were included in the analysis to eliminate the effect of differential expression of Z-linked genes in males and females due to incomplete dosage compensation. The extent of local gene-specific dosage compensation on the Z chromosome is similar across the somatic tissues and gonads (Mank and Ellegren 2009a; Melamed and Arnold 2007) and should not influence the selection of tissue-enriched genes. The GEO samples used in this study are listed in Table 1.

Table 1 A list of chicken microarray data used in the analysis

Tissue	GEO Sample IDs	Publication
Brain	GSM157823, GSM157824, GSM157825, GSM157826, GSM157827	Itoh et al. (2007)
Liver	GSM157879, GSM157881, GSM157882, GSM157884, GSM157886	Itoh et al. (2007)
Heart	GSM157809, GSM157811, GSM157813, GSM157815, GSM157817	Itoh et al. (2007)
Brain	GSM215459, GSM215460, GSM215461, GSM215462	Ellegren et al. (2007)
Heart	GSM215473, GSM215474, GSM215475, GSM215476	Ellegren et al. (2007)
Primary oocytes	GSM189423, GSM189424	Elis et al. (2008)
Granulosa cells	GSM189428, GSM189427	Elis et al. (2008)

Normalization and Annotation of Microarray Data

Data analysis was performed using the software Bioconductor 2.1 (<http://www.bioconductor.org/>) (Gentleman et al. 2004) and R project for statistical computing (version 2.6.0, <http://www.r-project.org/>). The probes were annotated to Ensembl gene identifiers (Ensembl release 46) using the custom chip description file (cdf) “ggchickengensg” version 10.0.0 (Dai et al. 2005) from the Bioconductor repository. This file contains 11,998 probe sets assigned to Ensembl genes with known chromosomal location. Expression levels were calculated using GC robust multi-array average (gcrMA) with the help of the probe sequence (Irizarry et al. 2003). For the comparison of microarray data from different experiments, the gcrMA preprocessing step was performed for all the data in one batch.

Statistical Analysis

We used Linear Models for Microarray Data Package, “limma” version 2.12.0 (Smyth 2005), for statistical evaluations of expression differences. A linear model was fitted for each gene in a given series of arrays using the lmFit function. To rank the differential expression of genes, we applied the eBayes function of the empirical Bayes method. A correction for multiple testing was performed using the Benjamini and Hochberg false-discovery-rate (FDR) method. Adjustment for multiple testing between tissues was done by the nested *F*-test approach (decideTests parameter method set to “nestedF” in limma).

Tissue-enriched genes (also referred to as genes with preferential/predominant expression in a given tissue) were defined as genes fulfilling the following criteria: (1) Significantly higher expression in a given tissue compared to any other tissue in the analysis, (2) fold-change value higher than three or five in a given tissue than in any other tissue in the analysis, and (3) average expression level higher than 100. This level is often used to discriminate between true expression and signal noise in microarray data. Because of a limited number of tissues used in our analysis and embryonic origin of non-sex-specific tissues,

our selection of tissue-enriched genes should be considered as approximate. The observed and expected numbers of tissue-enriched genes on individual chromosomes were compared using the χ^2 test. Expected gene counts were calculated using total numbers of annotated genes on GPL3213 GeneChips using “ggchickengensg” annotation.

The MicroArray Quality Control (MAQC) project has demonstrated high consistency among data from Affymetrix GeneChips prepared in different laboratories, which justifies intraplatform comparisons (MAQC Consortium 2006). Nevertheless, to independently assess whether differences among microarray data prepared in different laboratories had not introduced any bias into the analysis, we performed normalization and selection of tissue-enriched genes in two analyses, each time using microarray data partially prepared in different laboratories. In the first analysis, we included expression data from brain and heart prepared in the laboratory of A. P. Arnold (Itoh et al. 2007). In the second analysis, we included the data from brain and heart prepared in the laboratory of H. Ellegren (Ellegren et al. 2007). The data from liver, oocytes and granulosa cells were the same in both analyses (Itoh et al. 2007; Elis et al. 2008). Our results show that 52–82% of tissue-enriched genes were identical in both analyses (Supplementary Fig. 1). We then performed the third analysis with the combined data set. All statistical analyses were performed separately for all three data sets. Because we obtained the same results in terms of genomic distribution of tissue-enriched genes for all three data sets, only results from the combined data set are presented. Results for the data set 1 and data set 2 are shown in Supplementary material. Consistent results across analyses suggest that the observed patterns in genomic distribution of tissue-enriched genes are unlikely to be caused by differences in microarray data prepared in different laboratories.

Results

To clarify the effects of sexually antagonistic selection and meiotic Z chromosome inactivation on the Z

chromosome gene content, we analyzed genomic distribution of genes with preferential expression in several non-sex-specific tissues as well as somatic and germ cells of the ovary (hereafter tissue-enriched genes). Tissue-enriched genes were selected using two different fold-change thresholds (fold-change 3 and fold-change 5) (see [Methods](#)). Our results show that genes with preferential expression in granulosa cells, which represent ovarian somatic cells, are 3.4 times (for fold-change 3) or 4.1 times (for fold-change 5) more common on the Z chromosome than expected assuming random genomic distribution. This represents significant enrichment (χ^2 test, $p < 0.001$, [Table 2](#)). Interestingly, such enrichment was not found for genes with preferential expression in primary oocytes (χ^2 test, $p > 0.05$, [Table 2](#)). Genes with predominant expression in non-sex-specific tissues, including brain, heart and liver, also did not show any significant deviations from expected distribution (χ^2 test, $p > 0.05$, [Table 2](#)). Enrichment of the Z chromosome for genes with preferential expression in granulosa cells was also significant after applying Bonferroni correction for multiple comparisons ($p < 0.01$; in total 20 tests were performed).

To explore whether the nonrandom genomic distribution of genes with preferential expression in granulosa cells is specific for the Z chromosome, we compared proportions of tissue-enriched genes on individual chromosomes ([Fig. 1](#)). Only chromosomes with more than 350 annotated genes were included into the analysis. These chromosomes involved the Z chromosome and autosomes 1–9. χ^2 test was then used to compare the observed and expected numbers of tissue-enriched genes on individual chromosomes. The only tissue that showed significant deviations from random genomic expectations when Bonferroni

correction for multiple comparisons was applied were granulosa cells (χ^2 test, $df = 9$, $p < 0.00001$). The significant deviation from expected gene counts was mostly due to over-representation of genes with preferential expression in granulosa cells on the Z chromosome ([Fig. 1](#) and [Table 3](#)). When the Z chromosome was excluded from the analysis, no significant deviations from expected gene counts were observed (χ^2 test, $df = 8$, $p > 0.05$). Our results thus suggest that nonrandom distribution of genes with predominant expression in granulosa cells is specifically related to the Z chromosome.

The observed enrichment of the Z chromosome for genes with preferential expression in granulosa cells could in principle be caused by a higher duplication rate of these genes on the Z chromosome rather than more frequent fixation of new female-benefit mutations on this chromosome. To exclude this possibility, we analyzed chromosomal location and homology of 41 Z-linked genes showing preferential expression in granulosa cells (fold-change 3) ([Supplementary Table 3](#)). Using Ensembl genome browser, we found that only two of these genes, ENSGALG00000012575 and ENSGALG00000015184, were paralogous. Both genes are located next to each other on the chromosome, suggesting that they might have originated by tandem duplication. Nevertheless, the presence of these two paralogues cannot explain the marked over-representation of granulosa cell-enriched genes on the Z chromosome.

Discussion

To elucidate the selective forces that shape the Z chromosome gene content in chicken, we analyzed genomic distribution of genes with preferential expression in several non-sex-specific tissues as well as somatic and germ cells of the ovary. We found that the Z chromosome is enriched for genes with preferential expression in granulosa cells—the ovarian somatic cells—but not for genes with predominant expression in non-sex-specific tissues or genes with preferential expression in primary oocytes. The observed enrichment of the Z chromosome for genes with preferential expression in granulosa cells cannot be attributed to recent duplications of these genes on the Z chromosome. Instead, we suggest that sexually antagonistic selection resulting in higher abundance of female-benefit alleles on the Z chromosome is responsible for the observed pattern. Below, we compare our results with previous studies on the X and Z chromosome gene content and discuss our findings in the light of theories on the evolution of sex chromosome gene content.

Table 2 Observed and expected numbers of tissue-enriched genes on the autosomes (A) and the Z chromosome (Z)

Tissue	Fold-change	Observed			Expected		p^*
		A	Z	%Z	A	Z	
Brain	3	423	19	4.30	421	21	0.8714
	5	324	14	4.14	322	16	0.8518
Heart	3	205	6	2.84	201	10	0.4445
	5	141	5	3.42	139	7	0.7681
Liver	3	455	23	4.81	455	23	0.8799
	5	364	11	2.93	357	18	0.2558
Granulosa cells	3	212	41	16.21	241	12	0.0000
	5	125	29	18.83	147	7	0.0002
Primary oocytes	3	316	14	4.24	314	16	0.8518
	5	158	8	4.82	158	8	0.7978

* χ^2 test, $df = 1$, significant values are indicated in bold

Fig. 1 The proportions of tissue-enriched genes on individual chromosomes. Only chromosomes with more than 350 annotated genes are shown. Black and gray bars represent genes with fold-change three and five, respectively. Granulosa cells were the only tissue showing significant deviations from random genomic expectations when Bonferroni correction for multiple testing was applied (χ^2 test, $df = 9, p < 0.00001$). This deviation was mostly due to a higher abundance of tissue-enriched genes on the Z chromosome

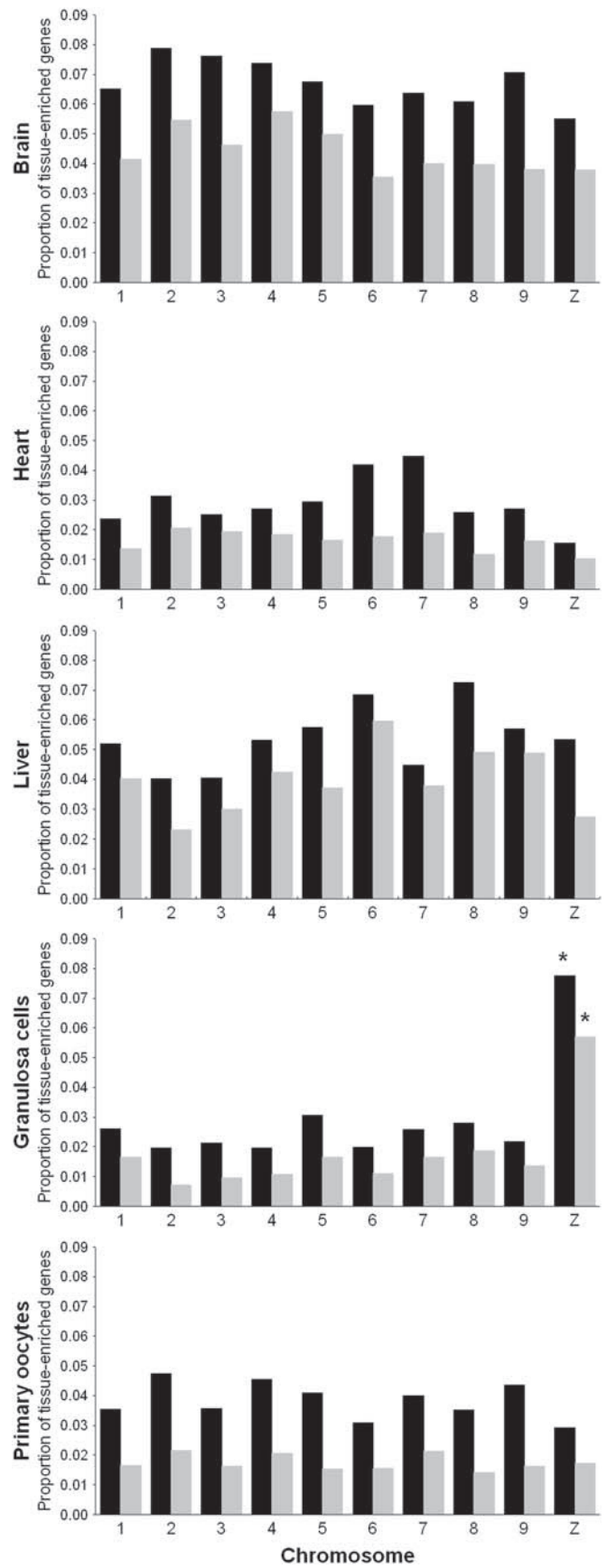


Table 3 Observed (O) and expected (E) numbers of genes with preferential expression in granulosa cells on individual chromosomes

Chr	Fold-change 3			Fold-change 5		
	O	E	$(O - E)^2/E$	O	E	$(O - E)^2/E$
1	38	36	0.15	25	22	0.50
2	16	24	2.41	7	14	3.75
3	15	22	2.16	8	13	2.12
4	17	19	0.30	8	12	1.24
5	21	17	1.21	10	10	0.00
6	7	10	0.68	4	6	0.57
7	7	9	0.42	6	5	0.06
8	8	9	0.11	6	5	0.05
9	6	8	0.40	5	5	0.02
Z	41	12	67.68	29	7	62.41
Sum	176	164	75.52	108	100	70.70

* χ^2 test, $df = 9$, $p < 0.000001$

Only chromosomes with more than 350 annotated genes are shown

Sexually Antagonistic Selection and the Dominance of Sexually Antagonistic Mutations

Theoretical models predict that sexually antagonistic selection should result in a higher abundance of male-advantageous alleles on the X chromosome if new beneficial mutations are on average recessive (Rice 1984). The reason is that male-advantageous recessive mutations are more likely to become fixed on the X chromosome than on autosomes, because they are immediately exposed to selection in hemizygous males. On the other hand, if mutations are on average dominant, the X chromosome is expected to become enriched for female-advantageous alleles and depleted of male-advantageous alleles because this chromosome spends two-thirds of its time in males, but only one-third in females (Rice 1984). By analogy, the Z chromosome should become enriched for female-benefit alleles if new mutations are mostly recessive, but depleted of female-benefit alleles if new mutations are mostly dominant. Since the level of dominance of new sexually antagonistic mutations is not known, the effect of sexually antagonistic selection is difficult to predict (van Doorn 2009; Vicoso and Charlesworth 2006).

Empirical results provide mixed evidence on the role of sexually antagonistic selection in the evolution of sex chromosome gene content. Consistently with a dominant mode of inheritance for sexually antagonistic genes, Mank and Ellegren (2009b) found that genes that are up-regulated in the adult ovary relative to the embryonic ovary (i.e., female-benefit genes) are under-represented on the avian Z chromosome, while genes down-regulated in the adult ovary compared to the embryonic ovary (i.e., female-

detriment genes) are over-represented on the Z chromosome. The paucity of male-biased genes on the *Drosophila* X chromosome is also consistent with the dominant nature of sexually antagonistic mutations (Parisi et al. 2003; Ranz et al. 2003; Sturgill et al. 2007). The data from mammals, however, show just the opposite pattern. Male-biased genes expressed in somatic tissues or pre-meiotic germ cells are more common on the X chromosome, indicating that sexually antagonistic mutations are on average recessive (Divina et al. 2005; Khil et al. 2004; Lercher et al. 2003; Wang et al. 2001).

The inconsistency of these results could be explained by various epigenetic effect that shape the X/Z chromosome gene content alongside with sexually antagonistic selection. Thus, under-representation of male-biased genes on the *Drosophila* X chromosome might be caused by up-regulation of the single X chromosome in males owing to dosage compensation (Baker et al. 1994). This up-regulation might limit further increase in X-linked gene expression in males, resulting in the paucity of male-biased genes on the X chromosome (Rogers et al. 2003; Vicoso and Charlesworth 2009). In chicken, the paucity of genes up-regulated in the adult ovary relative to the embryonic ovary as well as enrichment of genes down-regulated in the adult ovary compared to the embryonic ovary on the Z chromosome could be related to meiotic inactivation of the Z chromosome (Schoenmakers et al. 2009) rather than sexually antagonistic selection. Indeed, the observation that genes with differential expression between embryonic and adult gonads show nonrandom genomic distribution only in females but not in males (Mank and Ellegren 2009b) is more compatible with the action of meiotic Z chromosome inactivation than with sexually antagonistic selection that is expected to affect both female-effect and male-effect genes (Ellegren and Parsch 2007).

In this study, we show that genes with preferential expression in granulosa cells, that are expected to have female-benefit effect, are more abundant on the Z chromosome. This is consistent with results from mammals (Divina et al. 2005; Khil et al. 2004; Lercher et al. 2003; Wang et al. 2001) and indicates that recessive rather than dominant sexually antagonistic mutations drive the evolution of the X and Z chromosomes. The question still remains why the Z chromosome is not enriched also for genes with preferential expression in primary oocytes. Because primary oocytes and granulosa cells together form one functional unit, the ovary follicle, the level of sexual antagonism in both cell types is expected to be similar. We thus suggest that meiotic sex chromosome inactivation might hinder accumulation of oocyte-expressed genes on the Z chromosome.

Meiotic Sex Chromosome Inactivation and Defeminization of the Z Chromosome

In organisms with heterogametic males, the X chromosome is subject to meiotic sex chromosome inactivation in the male germline (McKee and Handel 1993; Turner 2007). In mammals, meiotic X inactivation begins at pachytene and although partial reactivation occurs after the prophase I, about 87% of X-linked genes remain suppressed post-meiotically (Namekawa et al. 2006). Transcriptional silencing of the X chromosome during male meiosis has also been observed in *Drosophila* (Hense et al. 2007) and *C. elegans* (Kelly et al. 2002). It has been suggested that meiotic X chromosome inactivation is the most likely explanation for the paucity of spermatocyte- and spermatid-expressed genes on the mammalian X chromosome (Emerson et al. 2004; Khil et al. 2004; Mueller et al. 2008; Potrzebowski et al. 2008) and might contribute to underrepresentation of male-biased genes on the X in *Drosophila* (Betran et al. 2002; Parisi et al. 2003; Ranz et al. 2003; Sturgill et al. 2007; Vibrationovski et al. 2009) and *C. elegans* (Reinke et al. 2004).

In birds, the Z chromosome becomes transcriptionally silenced during pachytene and remains inactive until early diplotene (Schoenmakers et al. 2009). This period occurs just after hatching and lasts approximately 5–6 days. It has been suggested that reactivation of the Z chromosome at early diplotene may be necessary to allow expression of genes that are required for further maintenance and growth of the large and long-living oocytes (Namekawa and Lee 2009). However, it is possible that, similarly as in mammals, the Z chromosome is not reactivated completely and that some Z-linked genes might be transcriptionally suppressed even in later stages of oogenesis. This could result in the relative paucity of genes with preferential expression in adult oocytes on the Z chromosome. Alternatively, even a short period of Z chromosome inactivation might create selective pressure against Z-linkage of genes expressed in the female germline.

Our results thus suggest that genes expressed in primary oocytes can be exposed to two evolutionary forces with opposite effects: Sexual antagonism that increases Z-linkage and meiotic Z inactivation that decreases it. Balance between these two forces might explain why we did not observe any bias in Z-linkage for oocyte-expressed genes. Further studies on gene expression changes during the course of oogenesis should shed more light on the evolutionary forces shaping the Z chromosome gene content.

Acknowledgments We thank Šárka Takáčová for linguistic revision of the manuscript. This work was supported by junior grant GAAV (KJB501110701) to RS and by the Ministry of Education, Youth and Sports of the Czech Republic (MSM0021620828).

References

- Arnold AP, Itoh Y, Melamed E (2008) A bird's-eye view of sex chromosome dosage compensation. *Annu Rev Genom Hum Genet* 9:109–127
- Arunkumar KP, Mita K, Nagaraju J (2009) Silkworm Z chromosome is enriched in testis-specific genes. *Genetics* 182:493–501
- Baker BS, Gorman M, Marin I (1994) Dosage compensation in *Drosophila*. *Annu Rev Genet* 28:491–521
- Betran E, Thornton K, Long M (2002) Retroposed new genes out of the X in *Drosophila*. *Genome Res* 12:1854–1859
- Connallon T, Knowles LL (2005) Intergenomic conflict revealed by patterns of sex-biased gene expression. *Trends Genet* 21:495–499
- Dai M, Wang P, Boyd AD, Kostov G, Athey B, Jones EG, Bunney WE, Myers RM, Speed TP, Akil H, Watson SJ, Meng F (2005) Evolving gene/transcript definitions significantly alter the interpretation of GeneChip data. *Nucleic Acids Res* 33:e175
- Divina P, Vlcek C, Strnad P, Paces V, Forejt J (2005) Global transcriptome analysis of the C57BL/6 J mouse testis by SAGE: Evidence for nonrandom gene order. *BMC Genom* 6:29
- Elis S, Batellier F, Couty I, Balzergue S, Martin-Magniette ML, Monget P, Blesbois E, Govoroun MS (2008) Search for the genes involved in oocyte maturation and early embryo development in the hen. *BMC Genom* 9:110
- Ellegren H, Parsch J (2007) The evolution of sex-biased genes and sex-biased gene expression. *Nat Rev Genet* 8:689–698
- Ellegren H, Hultin-Rosenberg L, Brunstrom B, Dencker L, Kultima K, Scholz B (2007) Faced with inequality: Chicken do not have a general dosage compensation of sex-linked genes. *BMC Biol* 5:40
- Emerson JJ, Kaessmann H, Betran E, Long M (2004) Extensive gene traffic on the mammalian X chromosome. *Science* 303:537–540
- Gentleman RC, Carey VJ, Bates DM, Bolstad B, Dettling M, Dudoit S, Ellis B, Gautier L, Ge Y, Gentry J, Hornik K, Hothorn T, Huber W, Iacus S, Irizarry R, Leisch F, Li C, Maechler M, Rossini AJ, Sawitzki G, Smith C, Smyth G, Tierney L, Yang JY, Zhang J (2004) Bioconductor: Open software development for computational biology and bioinformatics. *Genome Biol* 5:R80
- Gurbich TA, Bachtrog D (2008) Gene content evolution on the X chromosome. *Curr Opin Genet Dev* 18:493–498
- Hense W, Baines JF, Parsch J (2007) X chromosome inactivation during *Drosophila* spermatogenesis. *PLoS Biol* 5:e273
- Hurst LD, Randerson JP (1999) An eXceptional chromosome. *Trends Genet* 15:383–385
- Irizarry RA, Hobbs B, Collin F, Beazer-Barclay YD, Antonellis KJ, Scherf U, Speed TP (2003) Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics* 4:249–264
- Itoh Y, Melamed E, Yang X, Kampf K, Wang S, Yehya N, Van Nas A, Replogle K, Band MR, Clayton DF, Schadt EE, Lusk AJ, Arnold AP (2007) Dosage compensation is less effective in birds than in mammals. *J Biol* 6:2
- Kaiser VB, Ellegren H (2006) Nonrandom distribution of genes with sex-biased expression in the chicken genome. *Evolution* 60:1945–1951
- Kelly WG, Schaner CE, Dernburg AF, Lee MH, Kim SK, Villeneuve AM, Reinke V (2002) X-chromosome silencing in the germline of *C. elegans*. *Development* 129:479–492
- Khil PP, Smirnova NA, Romanienko PJ, Camerini-Otero RD (2004) The mouse X chromosome is enriched for sex-biased genes not subject to selection by meiotic sex chromosome inactivation. *Nat Genet* 36:642–646
- Lercher MJ, Urrutia AO, Hurst LD (2003) Evidence that the human X chromosome is enriched for male-specific but not female-specific genes. *Mol Biol Evol* 20:1113–1116

- Mank JE (2009) Sex chromosomes and the evolution of sexual dimorphism: Lessons from the genome. *Am Nat* 173:141–150
- Mank JE, Ellegren H (2009a) All dosage compensation is local: Gene-by-gene regulation of sex-biased expression on the chicken Z chromosome. *Heredity* 102:312–320
- Mank JE, Ellegren H (2009b) Sex-linkage of sexually antagonistic genes is predicted by female, but not male, effects in birds. *Evolution* 63:1464–1472
- MAQC Consortium (2006) The MicroArray Quality Control (MAQC) project shows inter- and intraplatform reproducibility of gene expression measurements. *Nat Biotechnol* 24:1151–1161
- McKee BD, Handel MA (1993) Sex chromosomes, recombination, and chromatin conformation. *Chromosoma* 102:71–80
- Melamed E, Arnold AP (2007) Regional differences in dosage compensation on the chicken Z chromosome. *Genome Biol* 8:R202
- Mueller JL, Mahadevaiah SK, Park PJ, Warburton PE, Page DC, Turner JMA (2008) The mouse X chromosome is enriched for multicopy testis genes showing postmeiotic expression. *Nat Genet* 40:794–799
- Namekawa SH, Lee JT (2009) XY and ZW: Is meiotic sex chromosome inactivation the rule in evolution? *PLoS Genet* 5:e1000493
- Namekawa SH, Park PJ, Zhang LF, Shima JE, McCarrey JR, Griswold MD, Lee JT (2006) Postmeiotic sex chromatin in the male germline of mice. *Curr Biol* 16:660–667
- Parisi M, Nuttall R, Naiman D, Bouffard G, Malley J, Andrews J, Eastman S, Oliver B (2003) Paucity of genes on the *Drosophila* X chromosome showing male-biased expression. *Science* 299:697–700
- Potrzebowski L, Vinckenbosch N, Marques AC, Chalmel F, Jegou B, Kaessmann H (2008) Chromosomal gene movements reflect the recent origin and biology of therian sex chromosomes. *PLoS Biol* 6:e80
- Ranz JM, Castillo-Davis CI, Meiklejohn CD, Hartl DL (2003) Sex-dependent gene expression and evolution of the *Drosophila* transcriptome. *Science* 300:1742–1745
- Reinke V, Gil IS, Ward S, Kazmer K (2004) Genome-wide germline-enriched and sex-biased expression profiles in *Caenorhabditis elegans*. *Development* 131:311–323
- Rice WR (1984) Sex chromosomes and the evolution of sexual dimorphism. *Evolution* 38:735–742
- Rogers DW, Carr M, Pomiankowski A (2003) Male genes: X-pelled or X-cluded? *Bioessays* 25:739–741
- Saifi GM, Chandra HS (1999) An apparent excess of sex- and reproduction-related genes on the human X chromosome. *Proc Biol Sci* 266:203–209
- Schoenmakers S, Wassenaar E, Hoogerbrugge JW, Laven JS, Grootegoed JA, Baarends WM (2009) Female meiotic sex chromosome inactivation in chicken. *PLoS Genet* 5:e1000466
- Smyth G (2005) Limma: Linear models for microarray data. In: Gentleman R, Carey VJ, Huber W, Dudoit S, Irizarry RA (eds) *Bioinformatics and computational biology solutions using R and bioconductor*. Springer, New York, pp 397–420
- Storchova R, Divina P (2006) Nonrandom representation of sex-biased genes on chicken Z chromosome. *J Mol Evol* 63:676–681
- Sturgill D, Zhang Y, Parisi M, Oliver B (2007) Demasculinization of X chromosomes in the *Drosophila* genus. *Nature* 450:238–241
- Turner JM (2007) Meiotic sex chromosome inactivation. *Development* 134:1823–1831
- van Doorn GS (2009) Intralocus sexual conflict. *Ann N Y Acad Sci* 1168:52–71
- Vibrantovski MD, Zhang Y, Long M (2009) Out of the X chromosomal gene movement in the *Drosophila* genus. *Genome Res* 19:897–903
- Vicoso B, Charlesworth B (2006) Evolution on the X chromosome: Unusual patterns and processes. *Nat Rev Genet* 7:645–653
- Vicoso B, Charlesworth B (2009) The deficit of male-biased genes on the *D. melanogaster* X chromosome is expression-dependent: A consequence of dosage compensation? *J Mol Evol* 68:576–583
- Wang PJ, McCarrey JR, Yang F, Page DC (2001) An abundance of X-linked genes expressed in spermatogonia. *Nat Genet* 27:422–426
- Zha X, Xia Q, Duan J, Wang C, He N, Xiang Z (2008) Dosage analysis of Z chromosome genes using microarray in silkworm, *Bombyx mori*. *Insect Biochem Mol Biol* 39:315–321

8.2 Skripty jazyka R využité pro analýzu

Některé příložené skripty nejsou samostatně spustitelné, jsou spíše bloky používané pro interaktivní práci s programem R, vkládané do příkazového řádku pomocí schránky. Naleznete je i na příloženém CD.

Vytvoření anotačního datového objektu (init.R)

```
#
# retrieves gene annotations for genes identified by custom cdf file
# -creates data.frame "g" with annotations having the same indices as genes
#   in cdf file
# -creates list "set" with slots representing chromosomes, containing indices
#   of genes localized on these chromosomes
#

library(affy);
library(gcrma);
library(biomaRt);

#
# load some chip to get the cdf rownames
#

# read in experiment data
pd <- read.AnnotatedDataFrame("pData_leto.txt", header=TRUE, row.names="name");

# load cel files according to paths in data
celData <- ReadAffy(filenamees=pd$filename, phenoData=pd, cdfname="ggchickenggensg",
                    verbose=TRUE);

#
# extract the names
#

# get annotation for all genes in table
geneNames <- substr(featureNames(celData), 1, 18);

#
# query biomaRt for gene annotations
#

ens_mart <- useMart("ensembl", dataset="ggallus_gene_ensembl");
genes <- getGene(id=geneNames, type="ensembl_gene_id", mart=ens_mart);

#
# map the unordered table to the cdf table
#

# create a mapping to the genes list
geneMap <- as(sapply(geneNames, grep, genes$ensembl_gene_id), "integer");

# use the mapping to create table with same order as cdf
g <- genes[geneMap,]
```

```

#
# create sets of genes by chromosomes, dumping gene counts in each set
#

# for all chromosomes in annotation in a loop
set <- NULL
for (chr in levels(as.factor(genes$chromosome_name))) {
  tmp_name <- paste("chr", chr, sep="")
  cat("chromosome", chr, "-> ")

  set[[tmp_name]] <- which(g[, "chromosome_name"] == chr)
  set[[tmp_name]] <- set[[tmp_name]][order(g[set[[tmp_name]], "start_position"])]

  cat(length(set[[tmp_name]]), "genes\n")
  rm(tmp_name)
}

# save the objects
save(g, set, file="common2")

```

Výběr genů s preferenční expresí v konkrétní tkáni (sel_spec.R)

```

#
# select tissue specific genes
#

library(gcrma);
library(limma);
load("common2");

celDir <- "/home/liborm/data/celfiles/"
prefix <- "run4"
lfcThreshold <- log2(5)
outDir <- "lfc5"

# create global variables needed in both functions
# - otherwise they get a local scope in read_data()
fit <- NULL;
exData <- NULL;
pd <- NULL;

#
# create cont[[tissue]] matrices
#
# run1, run3 and run4 contrasts
cont <- NULL
cont$brain <- cbind(gra=c(1, -1, 0, 0,0), liver=c(1,0,0,-1,0),
                  heart=c(1,0,-1,0,0), ov=c(1, 0, 0, 0, -1))
cont$gra <- cbind(brain=c(-1, 1, 0, 0,0), liver=c(0,1,0,-1,0),
                 heart=c(0,1,-1,0,0), ov=c(0, 1, 0, 0, -1))
cont$heart <- cbind(brain=c(-1, 0, 1, 0,0), liver=c(0,0,1,-1,0),
                  gra=c(0,-1,1,0,0), ov=c(0, 0, 1, 0, -1))
cont$liver <- cbind(brain=c(-1, 0, 0, 1,0), gra=c(0,-1,0,1,0),
                  heart=c(0,0,-1,1,0), ov=c(0, 0, 0, 1, -1))
cont$ov <- cbind(brain=c(-1, 0, 0, 0,1), gra=c(0,-1,0,0,1),
                heart=c(0,0,-1,0,1), liver=c(0, 0, 0, -1, 1))

# remove chrW set, because there are no genes in it
set$chrW <- NULL

#
# read_data reads data and fits the linear model
# for current prefix
#
# using "write global" operator <<-
read_data <- function() {
  pd <<- read.AnnotatedDataFrame(paste(prefix, "_pdata.txt", sep=""),
                                header=TRUE, row.names="sample_id");

```

```

celData <- ReadAffy(filename=paste(celDir, pd$series_id, "/", sampleNames(pd),
                                ".CEL.gz", sep=""),
                  phenoData=pd, cdfname="ggchickenggensg", verbose=TRUE);

exData <- gcrma(celData);

design <- model.matrix(~0 + pd$tissue)
fit <- lmFit(exData, design)
}

#
# dump_results fits tissue specific contrasts for all given in cont object
# applies empirical bayes correction, does the thresholding and outputs the
# results to a disk file for further use
#
dump_results <- function() {
  results <- data.frame()

  # for each tissue
  for(tissue in names(cont)) {
    fit2 <- contrasts.fit(fit, cont[[tissue]])
    fit2 <- eBayes(fit2);

    # matrix of expression values just for current tissue
    tiss_exp <- exprs(exData)[,which(pd$tissue == tissue)]
    tiss_means <- rowMeans(tiss_exp)

    # calculate statistic decisions
    dt <- decideTests(fit2, method="nestedF", lfc=lfcThreshold)

    # dump the table with test decisions and means
    write.table(file=paste(outDir, "/", prefix, "_", tissue, ".txt", sep=""),
              cbind(dt, means=tiss_means))

    # calculate sums for each chromosome
    for(chr in names(set)) {
      # create a subset with expression means for current chromosome
      exp_chr_spec <- tiss_means[set[[chr]][rowSums(dt[set[[chr]],]) == dim(dt)[2]]

      results[chr, paste(tissue, "_spec0", sep="")] <- length(exp_chr_spec)
      results[chr, paste(tissue, "_spec50", sep="")] <- sum(exp_chr_spec > log2(50))
      results[chr, paste(tissue, "_spec100", sep="")] <- sum(exp_chr_spec > log2(100))
    }

    # for each tissue sum up
    set_spec <- tiss_exp[rowSums(dt) == dim(dt)[2],]
    results["all_minExp", paste(tissue, "_spec0", sep="")] <- min(set_spec);
    results["all_100", paste(tissue, "_spec0", sep="")] <-
      sum(rowMeans(set_spec) > log2(100));
    results["all_50", paste(tissue, "_spec0", sep="")] <-
      sum(rowMeans(set_spec) > log2(50));
    results["all", paste(tissue, "_spec0", sep="")] <- dim(set_spec)[1];
  }

  write.table(file=paste(outDir, "/", prefix, "_results.txt", sep=""), results)
}

#
# batch process all data (_male are calculated using also chips from males)
#
prefixes <- c("set1", "set2", "set_all", "set1_male", "set_all_male")
for (prefix in prefixes) {
  read_data();

  lfcThreshold <- log2(3)
  outDir <- "lfc3"
  dump_results();
}

```

```

lfcThreshold <- log2(5)
outDir <- "lfc5"
dump_results();
}

# wrap up the results and send via e-mail
system("tar -cjf res.tar.bz2 lfc3/*_results.txt lfc5/*_results.txt");
system("echo Otevres WinRarem|
      mutt -s \"Dalsi vysledky\" -a res.tar.bz2 -c radkas@natur.cuni.cz");

```

Výpočet překryvů mezi analýzami jednotlivých datových sad (venn.R)

```

#
# calculate overlaps of given gene sets
#

library(limma)

venn <- function(postfix="", limit=0) {

  # write all output to a file
  sink(paste("venn_", dataDir, postfix, ".txt", sep=""));

  # alternative tissue lists
  #for(tiss in c("brain", "gra", "heart", "ov")) {
  #for(tiss in levels(as.factor(pd$tissue))) {
  for(tiss in c("brain", "gra", "heart", "liver", "ov")) {

    data<-NULL
    flags <- NULL
    res_spec <- NULL

    # for each data set
    for(r in vennpfx) {
      # read the data from file
      # and remove spike-in probes rows by subsetting
      data[[r]] <-
        read.table(paste(dataDir, "/", r, "_", tiss, ".txt", sep=""))[-c(1:96),]

      # exclude means from table for venn counting
      flags[[r]] <- data[[r]][,which(colnames(data[[r]]) != "means")]

      # count cases where gene was overx against each other
      # creating logical vectors
      res_spec[[r]] <- (rowSums(flags[[r]]) == dim(flags[[r]])[2]) &
        (data[[r]][,"means"] > limit);
    }

    res_spec <- as.data.frame(res_spec);

    # output using vennCounts function from limma package
    cat("\n", tiss, "\n");
    print(vennCounts(res_spec))
  }
  sink();
}

#
# sample input for interactive session
#

dataDir <- "lfc3"
vennpfx <- c("set1", "set2", "set_all")

venn(postfix="lfc_100_fix", limit=log2(100))

```

Výpočet testu dobré shody (chisq.R)

```
#
# summary table for chisq -- observed and expected gene counts
#
# for each tissue
dataDir <- "lfc3"
workPfx <- c("set1", "set2", "set_all", "set1_male", "set_all_male")

res <- NULL

annotA <- sum(sapply(set, length)) - length(set$chrZ) - length(set$chrMT)
annotZ <- length(set$chrZ)

for(prefix in workPfx) {
  # load the results
  t <- read.table(paste(dataDir, "/", prefix, "_results.txt", sep=""))
  sumA <- colSums(t[1:28,])
  sumZ <- t[30,]

  # for each tissue calc chisq for exp > 0
  for(tiss in c("brain", "gra", "heart", "liver", "ov")) {

    # build column name to use
    col_name <- paste(tiss, "_spec0", sep="")
    observSum <- as.numeric(sumA[col_name] + sumZ[col_name])
    expectA <- as.numeric(round((observSum * annotA) / (annotA + annotZ)))
    expectZ <- as.numeric(round((observSum * annotZ) / (annotA + annotZ)))

    pval <- chisq.test(matrix(
      c(as.numeric(sumA[col_name]), as.numeric(sumZ[col_name]),
        expectA, expectZ), nrow=2)
      )$p.value

    obsPercZ <- (sumZ[col_name] * 100) / (sumA[col_name] + sumZ[col_name])

    res <- rbind(res, c(prefix, col_name, sumA[col_name],
      sumZ[col_name], obsPercZ, expectA, expectZ, pval)
      )

    # repeat for 100 expression limit
    col_name <- paste(tiss, "_spec100", sep="")
    observSum <- as.numeric(sumA[col_name] + sumZ[col_name])
    expectA <- as.numeric(round((observSum * annotA) / (annotA + annotZ)))
    expectZ <- as.numeric(round((observSum * annotZ) / (annotA + annotZ)))

    pval <- chisq.test(matrix(
      c( as.numeric(sumA[col_name]), as.numeric(sumZ[col_name]),
        expectA, expectZ), nrow=2)
      )$p.value

    obsPercZ <- (sumZ[col_name] * 100) / (sumA[col_name] + sumZ[col_name])

    res <- rbind(res, c(prefix, col_name, sumA[col_name],
      sumZ[col_name], obsPercZ, expectA, expectZ, pval)
      )

  }
}
colnames(res) <-
  c("pfx", "tissue", "obsA", "obsZ", "percZ", "expA", "expZ", "pval")
write.table(file=paste(dataDir, "/", "chi_stats.txt", sep=""), res)
```


8.3 Výsledky pro datové sady 1 a 2

Tabulka P1: Pozorované a očekávané počty genů s preferenční expresí v jednotlivých tkáních na autosomech (A) a chromosomu Z (Z) pro datové sady 1 a 2.

Tkáň	Poměr ^a	Datová sada	Pozorované			Očekávané		p ^b
			A	Z	% Z	A	Z	
mozek	3	1	463	20	4,14	460	23	0,7550
	3	2	416	19	4,37	414	21	0,8714
	5	1	345	15	4,17	343	17	0,8565
	5	2	329	16	4,64	328	17	1,0000
srdce	3	1	224	6	2,61	219	11	0,3229
	3	2	206	5	2,37	201	10	0,2930
	5	1	148	4	2,63	145	7	0,5391
	5	2	156	5	3,11	153	8	0,5712
játra	3	1	445	23	4,91	445	23	0,8798
	3	2	450	22	4,66	449	23	1,0000
	5	1	359	11	2,97	352	18	0,2557
	5	2	356	13	3,52	351	18	0,4629
kumulární buňky	3	1	220	36	14,06	244	12	0,0005
	3	2	213	39	15,48	240	12	0,0001
	5	1	129	28	17,83	149	8	0,0008
	5	2	126	29	18,71	148	7	0,0002
primární oocyty	3	1	315	12	3,67	311	16	0,5623
	3	2	334	13	3,75	330	17	0,5755
	5	1	145	7	4,61	145	7	0,7844
	5	2	173	7	3,89	171	9	0,7981

^a poměr expresních úrovní

^b test dobré shody, df=1, statisticky významné hodnoty vyznačeny tučně

Tabulka P2: Pozorované (P) a očekávané (O) počty genů preferenčně exprimovaných v kumulárních buňkách podle jednotlivých chromosomů pro datovou sadu 1.

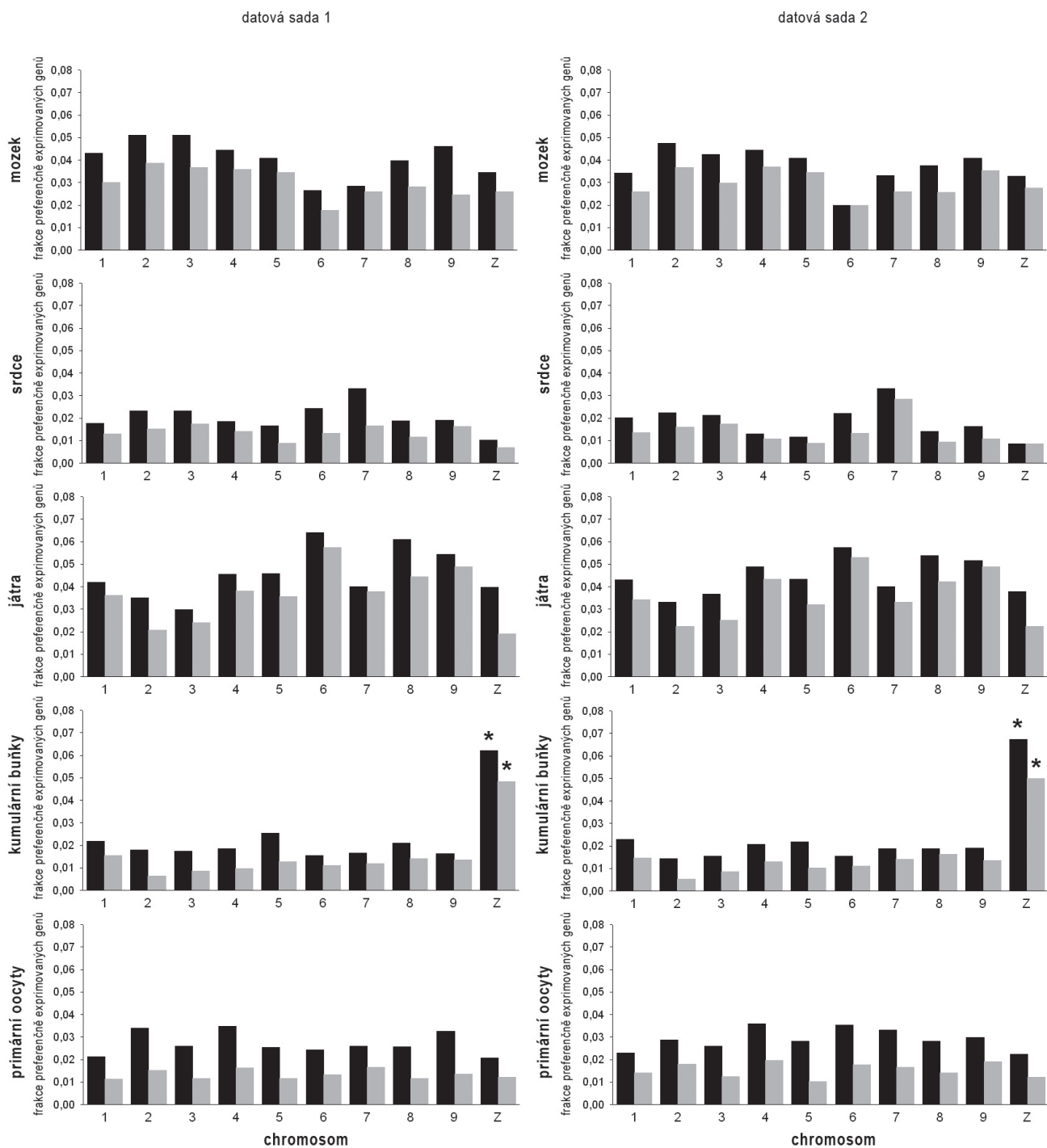
Chr	Poměr exp. úrovní 3			Poměr exp. úrovní 5		
	P	O	$(P - O)^2/O$	P	O	$(P - O)^2/O$
1	37	36	0,02	26	22	0,68
2	20	24	0,61	7	15	3,96
3	18	22	0,77	9	14	1,54
4	17	20	0,36	9	12	0,77
5	20	17	0,64	10	10	0,01
6	7	10	0,74	5	6	0,15
7	7	9	0,46	5	6	0,05
8	9	9	0,00	6	6	0,03
9	6	8	0,44	5	5	0,01
Z	36	12	45,10	28	8	54,89
Sum	177	166	49,14	110	102	62,08

* test dobré shody, df=9, $p < 0.000001$

Tabulka P3: Pozorované (P) a očekávané (O) počty genů preferenčně exprimovaných v kumulárních buňkách podle jednotlivých chromosomů pro datovou sadu 2.

Chr	Poměr exp. úrovní 3			Poměr exp. úrovní 5		
	P	O	$(P - O)^2/O$	P	O	$(P - O)^2/O$
1	39	36	0,34	25	22	0,46
2	16	23	2,36	6	14	4,91
3	16	22	1,53	9	13	1,44
4	19	19	0,01	12	12	0,00
5	17	16	0,02	8	10	0,45
6	7	10	0,66	5	6	0,12
7	8	9	0,09	6	5	0,05
8	8	9	0,10	7	6	0,40
9	7	8	0,07	5	5	0,01
Z	39	12	59,04	29	7	61,73
Sum	176	164	64,23	112	101	69,58

* test dobré shody, df=9, $p < 0.000001$



Obr. P1: Zastoupení genů preferenčně exprimovaných v jednotlivých tkáních na jednotlivých chromosomech pro datové sady 1 a 2. Zahrnuty jsou pouze chromosomy s více než 350 anotovanými geny. Černé sloupce reprezentují zastoupení genů s poměrem expresních úrovní (oproti všem ostatním tkáním) větším než tři, šedé sloupce zastoupení genů s poměrem expresních úrovní větším než pět. * značí signifikantní rozdíl od očekávaného počtu.

8.4 Seznam genů s preferenční expresí v kumulárních buňkách

Tabulka P4: Seznam genů s preferenční expresí v kumulárních buňkách pro poměr expresních úrovní větší než tři. Identifikátor genu odpovídá databázi Ensembl, vydání 46. Začátek a konec jsou pozice na chromosomu Z.

Ensembl ID genu	Začátek (bp)	Konec (bp)	Jméno genu
ENSGALG00000000215	45111055	45125036	IPI00586567.2
ENSGALG00000000241	45677856	45686316	Q5ZK67_CHICK
ENSGALG00000000255	46340505	46451863	XR_027032.1
ENSGALG00000002505	71207726	71209821	
ENSGALG00000002580	8597143	8599753	IPI00589938.2
ENSGALG00000003547	10327865	10338425	NP_001007983.1
ENSGALG00000003792	11668247	11709850	NP_001026622.1
ENSGALG00000003803	11746041	11756873	IPI00818969.1
ENSGALG00000004390	22222797	22387909	XR_027204.1
ENSGALG00000005437	31359105	31391130	NP_001026609.1
ENSGALG00000008237	72329216	72363594	IPI00582067.3
ENSGALG00000010164	27123323	27225427	SMARCA2
ENSGALG00000010166	27352711	27367737	VLDLR_CHICK
ENSGALG00000010180	27630692	27746246	IPI00601698.3
ENSGALG00000012575	38667131	38741615	XR_027180.1
ENSGALG00000013377	10286523	10327766	LMBD2_CHICK
ENSGALG00000014668	56458036	56684292	IPI00591535.2
ENSGALG00000014678	56157463	56212769	IPI00602590.2
ENSGALG00000014690	55570989	55625946	IPI00600831.3
ENSGALG00000014697	1290952	1330993	NP_989892.1
ENSGALG00000014719	16675929	16827865	C5orf35
ENSGALG00000014724	16876559	16914709	IPI00571929.2
ENSGALG00000014727	17606535	17657651	XR_026650.1
ENSGALG00000014730	18241276	18252802	IPI00584088.3
ENSGALG00000014863	13109506	13121624	IPI00574406.1
ENSGALG00000014880	14212108	14321882	IPI00588854.2
ENSGALG00000014903	15198520	15242625	IPI00594966.2
ENSGALG00000014978	22920786	23023263	IPI00586050.2
ENSGALG00000015028	28156680	28159130	B1AC67_CHICK
ENSGALG00000015105	31672108	31708616	PSIP1_CHICK
ENSGALG00000015110	34354371	34367915	C9orf61
ENSGALG00000015115	34554572	34621088	XR_027016.1
ENSGALG00000015184	38048991	38146662	NP_989568.1

ENSGALG00000015205	30549535	30561335	TYRP1_CHICK
ENSGALG00000015284	50219811	50229875	IPI00575417.2
ENSGALG00000015301	50494188	50523340	IPI00593582.3
ENSGALG00000015315	50910991	50919001	NP_001026602.1
ENSGALG00000015677	64690161	64727867	SUSD1
ENSGALG00000015684	64942669	64950821	XR_026950.1
ENSGALG00000023157	45143711	45162808	IPI00813637.1
ENSGALG00000023301	30614316	30614687	IPI00594920.2

Tabulka P5: Seznam genů s preferenční expresí v kumulárních buňkách pro poměr expresních úrovní větší než tři. Uvedeny jsou pouze geny s přiřazeným popisem.

Ensembl ID genu	Popis
ENSGALG00000000241	Putative uncharacterized protein [Source:UniProtKB/TrEMBL;Acc:Q5ZK67]
ENSGALG00000000255	PREDICTED: Gallus gallus similar to mannosidase, alpha, class 2A, member 1 (LOC415611), mRNA. [Source:RefSeq DNA;Acc:XR_027032]
ENSGALG000000003547	S-phase kinase-associated protein 2 (p45) [Source:RefSeq peptide;Acc:NP_001007983]
ENSGALG000000003792	FYN binding protein (FYB-120/130) [Source:RefSeq peptide;Acc:NP_001026622]
ENSGALG000000004390	PREDICTED: Gallus gallus similar to AP-3 complex beta3A subunit (LOC427646), mRNA. [Source:RefSeq DNA;Acc:XR_027204]
ENSGALG000000005437	zinc finger, DHHC domain containing 21 [Source:RefSeq peptide;Acc:NP_001026609]
ENSGALG00000010164	Probable global transcription activator SNF2L2 (EC 3.6.1.-)(ATP-dependent helicase SMARCA2)(SNF2-alpha)(SWI/SNF-related matrix-associated actin-dependent regulator of chromatin subfamily A member 2)(hBRM) [Source:UniProtKB/Swiss-Prot;Acc:P51531]
ENSGALG00000010166	Very low-density lipoprotein receptor Precursor (VLDL receptor)(VLDL-R)(Vitellogenin receptor)(VTG receptor) [Source:UniProtKB/Swiss-Prot;Acc:P98165]
ENSGALG00000012575	PREDICTED: Gallus gallus similar to Transducin-like enhancer of split 1 (E(sp1) homolog, Drosophila) (LOC427450), mRNA. [Source:RefSeq DNA;Acc:XR_027180]
ENSGALG00000013377	LMBR1 domain-containing protein 2 [Source:UniProtKB/Swiss-Prot;Acc:Q5F3F5]
ENSGALG00000014697	Sma- and Mad-related protein 2 [Source:RefSeq peptide;Acc:NP_989892]
ENSGALG00000014719	Uncharacterized protein C5orf35 [Source:UniProtKB/Swiss-Prot;Acc:Q8NE22]

ENSGALG00000014727	PREDICTED: Gallus gallus hypothetical protein LOC769469 (LOC769469), mRNA. [Source:RefSeq DNA;Acc:XR_026650]
ENSGALG00000015028	Relaxin-like peptide locus B [Source:UniProtKB/TrEMBL;Acc:B1AC67]
ENSGALG00000015105	Lens epithelium-derived growth factor [Source:UniProtKB/Swiss-Prot;Acc:Q5XXA9]
ENSGALG00000015110	Uncharacterized protein C9orf61 Precursor (Protein X123) [Source:UniProtKB/Swiss-Prot;Acc:Q15884]
ENSGALG00000015115	PREDICTED: Gallus gallus similar to MAM domain containing 2 (LOC427247), mRNA. [Source:RefSeq DNA;Acc:XR_027016]
ENSGALG00000015184	transducin-like enhancer of split 4 (E(sp1) homolog, Drosophila) [Source:RefSeq peptide;Acc:NP_989568]
ENSGALG00000015205	5,6-dihydroxyindole-2-carboxylic acid oxidase Precursor (DHICA oxidase)(EC 1.14.18.-)(Tyrosinase-related protein 1)(TRP-1)(TRP1) [Source:UniProtKB/Swiss-Prot;Acc:O57405]
ENSGALG00000015315	tripartite motif protein TRIM14 [Source:RefSeq peptide;Acc:NP_001026602]
ENSGALG00000015677	Sushi domain-containing protein 1 Precursor [Source:UniProtKB/Swiss-Prot;Acc:Q6UWL2]
ENSGALG00000015684	PREDICTED: Gallus gallus hypothetical LOC431639 (LOC431639), mRNA. [Source:RefSeq DNA;Acc:XR_026950]
