Doktorské studijní programy v biomedicíně

Univerzita Karlova v Praze

a Akademie věd České republiky

Program: Buněčná a vývojová biologie

Předseda oborové rady: Doc. RNDr. Jan Černý, Ph.D.

Školicí pracoviště: Univerzita Karlova, Přírodovědecká fakulta, Katedra buněčné biologie

Autor: Mgr. Eva Seifertová

Školitel: RNDr. Ing. Vladimír Krylov, Ph.D.

S disertací je možno se seznámit v příslušných knihovnách Přírodovědecké fakulty Univerzity Karlovy v Praze.

Abstrakt

Význam modelového organizmu *Xenopus tropicalis* v posledním desetiletí mnohonásobně stoupl. Jeho genom je oproti většině obojživelníků pouze diploidní a byl osekvenován, ovšem snaha o jeho kompletaci byla zatím neúspěšná. Příčinou je velké množství repetic znemožňující propojení jednotlivých úseků genomu do větších celků. I přes tento nedostatek zájem výzkumníků o tento druh neopadá- byly vynalezeny techniky jako gynogeneze či injikace morfolino oligonukleotidů do oplodněných oocytů, existuje knihovna BAC klonů s mnohonásobným pokrytím genomu. Pro následný výzkum je nicméně nezbytná co nejlepší znalost genomu tohoto druhu.

Tato dizertační práce je zaměřena především na vyvinutí nových technik použitelných pro genetické mapování. V prvním projektu byla sestavena vazebná a fyzická mapa druhu *X. tropicalis*. Protože tato obsahovala nezmapované oblasti, byla v rámci dalšího výzkumu vyvinuta zcela unikátní technika genetického mapování využívající mikrodisekci chromozomů a chromozomálních oblastí s následnou sekvenací takto získané DNA. Tímto způsobem jsme sekvenací pouhých 15 kopií krátkého raménka chromozómu 7 přiřadili k vyřezané oblasti velké množství genů a scaffoldů, jejichž poloha byla doposud neznámá, a rovněž se nám podařilo poukázat na nedostatky v současných verzích genomu tohoto druhu. Vyřezané chromozómy byly navíc použity i pro studii příbuznosti druhů *X. tropicalis* a *X. laevi*s metodou Zoo-FISH a data získaná ze sekvenace umožnila studii rozdílů chromozómu 7 mezi pohlavími.

Nově vyvinutý přístup sekvenace mikrodisektovaných oblastí je možné použít i v klinické medicíně, onkologii, případně pro genetické mapování druhů s neznámým genomem a s morfologicky odlišnými chromozómy.

Abstract

The diploid amphibian *Xenopus tropicalis* represents a significant model organism for studies of early development, genes function and evolution. Such techniques as gynogenesis, injection of morpholino antisense oligonucleotide into fertilized eggs or transgenesis were established. In the recent ten years, many efforts have been made to complete the sequence information. *X. tropicalis* genome has been sequenced but the completion of its assembly only on the basis of sequence data has been impossible. Therefore, our first work was focused on one of approaches for a genome completing- genetic mapping. First of all, the genetic map of *Xenopus tropicalis* was established pursuant linkage and physical positions of markers. Since the map contained gaps, we developed a new method for genetic mapping based on the next generation sequencing of laser microdissected arm. Using Illumina next generation sequencing of fifteen copies of a short arm of chromosome 7, we obtained new insights into its genome by localizing previously unmapped genes and scaffolds as well as recognizing mislocalized portions of the genome assembly. This was the first time laser microdissection and sequencing of specific chromosomal regions has been used for the purpose of genome mapping. These data were also used in the evolution study of the sex determining area placed on the q arm of chromosome 7, which showed that *Xenopus tropicalis* sex chromosomes contain large pseaudoautosomal areas. Moreover, we made Zoo-FISH analysis using *X. tropicalis* microdissected chromosomes as probes for labeling *Xenopus laevis* chromosomes, which revealed similarity of meiotic quartets even after 65 million years of separate evolution.

Our novel approach for next generation sequencing of microdissected chromosomal area is also applicable to species without sequenced genomes or for clinical applications in medical cytogenetics and oncology where tissue availability may be limiting. This method is likely to be of widespread use in species where individual chromosomes are distinguishable by cytological methods.

Obsah

1. Česká část

1.1. Literární přehled

Xenopus tropicalis má deset párů chromozómů, které je možné rozlišit pod mikroskopem podle jejich délky a poměru ramen. V této práci je použita nomenklatura chromozómů dle (Khokha et al., 2009). Jeho genom je velký 1,7 Gb a byl osekvenován (Hellsten et al., 2010). Kompletace genomu nebyla zatím dokončena. V současné době existují dvě verze (v.4.1 a v.7.1) obsahující přibližně 19.000, respektive 12.000 scaffoldů, které se doposud nepodařilo propojit a poloha většiny z nich je neznámá. Kromě toho je až 25% sekvencí samotných scaffoldů neznámá (představuje je „N"). Hlavní příčinou těchto nedostatků je velké množství repetic, které znemožňují bioinformatickým softwarům scaffoldy propojit. V nové verzi genomu *X. tropicalis* (v.7.1) se sice podařilo vytvořit 10 superscaffoldů představujících 10 chromozómů, ovšem pouze na základě syntenií s kuřetem. Vzhledem k tomu, že jsou tato data neověřená, je verze v.7.1 považována za velmi kontroverzní. Navíc je po porovnání fyzické mapy s touto verzí genomu zřejmé, že nesouhlasí pořadí genů na některých superscaffoldech a tedy že tato assembly obsahuje velké chyby.

Pohlavní chromozómy jsou u druhu *X. tropicalis* nerozlišitelné (homomorfní) a nebyly u něj doposud nalezeny žádné známé geny určující pohlaví. Oblast určující pohlaví nebyla nalezena ani po sekvenaci genomu a velmi dlouho nebylo jisté, které z pohlaví je u tohoto druhu heterogametické. (Hellsten et al., 2010) nalezli 22 markerů se silnou vazbou na samičí pohlaví. Tyto markery se podle naší genetické mapy (Wells et al., 2011) nacházejí na krátkém raménku chromozómu 7.

Genom druhu *X. tropicalis* byl osekvenován tradičním sekvenačním přístupem- Sanger sekvenováním. V současné době jsou pro účely sekvenování čím dál častěji využívány techniky tzv. sekvenování nové generace (Next Generation Sequencing, NGS). Oproti Sanger sekvenování je možné získat mnohonásobně levněji a rychleji velmi velké množství čtení, ta jsou ovšem velmi krátká (přibližně 30-100 bází) a některá rovněž obsahují velké množství chyb. Pro celogenomové sekvenování s následnou kompletací genomu je sice možné NGS použít, ale je nutné zhruba třicetinásobné pokrytí genomu. Tento problém je možné zjednodušit selektivní sekvenací pouze určité části genomu, jako jsou jednotlivé chromozómy či ramena.

Chromozómy je možné separovat dvěma odlišnými způsoby- průtokovou cytometrií nebo mikrodisekcí. První způsob má velké výtěžky, ale chromozómy se musí velikostně velmi lišit. Dodnes existuje pouze několik protokolů pro oddělení chromozómů tímto způsobem. Druhý přístup nemá na velikostní rozdíly tak výrazné nároky, ovšem výtěžky z mikrodisekce jsou velmi nízké. Pro následnou analýzu je proto nutné získanou DNA namnožit.

Celogenomová amplifikace (Whole Genome Amplification- WGA) slouží k mnohonásobnému namnožení DNA. Nevýhodou těchto metod je jejich neschopnost amplifikovat neselektivně všechny genové sekvence. Příčinou jsou preferenční amplifikace některých sekvencí a rovněž nízká kvalita a malé množství amplifikovaného materiálu. Podle studie (Höckner et al., 2009) má největší podíl úspěšně amplifikovaných markerů WGA přístup OmniPlex využívající fragmentaci DNA s následným připojením adaptérových sekvencí, které slouží pro nasedání specifických primerů. Konkrétním studovaným kitem byl GenomePlex Single Cell Kit (Sigma-Aldrich, Vienna, Austria). Ostatní přístupy fungující na principu amplifikace z nespecifických primerů (PEP, DOP-PCR, MDA) vykazovaly mnohem menší podíl amplifikovaných markerů. Rovněž následná NGS sekvenace z mikrodisektovaných a následně amplifikovaných chromozómů metodou PEP není v porovnání s OmniPlex metodou příliš úspěšná. (Weise et al., 2010) získali po 454-pyrosekvenování pouze 1000 sekvencí, z nichž pouze 15-20% mapovalo do referenčního genomu, zatímco při použití GenomePlex kitu a sekvenováním na přístroji Illumina získali (Ma et al., 2010) z mikrodisektovaných lidských chromozómů více než 24.000 haplotypových markerů.

Ke kompletaci genomu je možné využít i fyzickou a vazebnou mapu. V dnešní době je nejvyužívanější technikou fyzického mapování Fluorescenční In Situ Hybridizace (FISH), kterou je možné zobrazit skutečnou polohu studovaného markeru přímo na chromozómu. Tento přístup je poměrně komplikovaný a proto je možné zobrazit polohu pouze několika genů najednou. Vazebné mapování umožňuje vytvoření mapy celého genomu tvořené velkým množstvím markerů najednou. Značnou nevýhodou tohoto typu mapování je fakt, že mapy vytvořené nezávisle na sobě není možné propojit. Poloha markerů je vyjádřena jako relativní pozice v závislosti na pravděpodobnosti rekombinace, kde 1% pravděpodobnost je vyjádřena jako 1 centimorgan (cM). Poměr cM/počet bází není lineární a závisí na poloze markeru- klesá v okolí centromery či v sex určující oblasti. Kromě toho nemusí být některé části genomu ve vazebné mapě vůbec zahrnuty. Z tohoto důvodu je velmi vhodné propojovat

výsledky z obou typů mapování a rovněž z výsledků kompletace sekvenovaného genomu (assembly).

1.2. Cíle práce

Hlavním cílem této dizertační práce bylo vyvinout nové přístupy pro genetické mapování použitelné u modelového organizmu *Xenopus tropicalis.*

Konkrétní cíle byly následující:

- Sestavení genetické mapy obojživelníka rodu *Xenopus tropicalis, a* to fyzické mapy genů z vybraných scaffoldů metodou FISH.

- Studie podobnosti chromozómů druhů X. tropicalis a *X. laevis* pomocí techniky Zoo-FISH a s tím spojená optimalizace technik mikrodisekce chromozómů a celogenomové amplifikace.

- Přiřazení scaffoldů a genů ke krátkému raménku chromozómu 7 druhu *X. tropicalis* za použití nově vyvinuté techniky genetického mapování pomocí sekvenování nové generace mikrodisektovaných chromozomálních oblastí.

- Charakterizace pohlavních chromozómů a jejich rozdílů u druhu *X. tropicalis.*

1.3. Metody

1.3.1 Příprava buněk

K přípravě metafázních preparátů byla použita primární buněčná kultura z testikulárních buněk druhu *X. tropicalis*. Buňky byly trypsinovány a hypotonizovány v 38mM KCl po dobu 5 minut, fixovány a uloženy přes noc v teplotě -20°C.

1.3.2 Laserová mikrodisekce

Fixované buňky byly nakapány na sklíčka potažená polytén- naftalénovou mambránou (P.A.L.M. Gmbh, Bernried, Německo) a po 10ti minutách schnutí nabarveny 3% roztokem Giemsy. Následně bylo přístrojem PALM MicroLaser (P.A.L.M. Gmbh, Bernried, Německo) spojeným s invertovaným mikrokopem (Olympus) mikrodisektováno přibližně 15 kopií kopií chromozómů nebo chromozomálních ramen. Vyřezané oblasti byly katapultovány laserovým paprskem do kloboučku zkumavky, který obsahoval 4ml PCR oleje. Jednotlivé chromozómy byly rozpoznány podle velikosti a poměru ramen dle (Tymowska, 1973), nomenklatura je použitá dle (Khokha et al., 2009). Aby byla zajištěna mikrodisekce celého krátkého raménka chromozómu 7, byl tento řez veden mezi centromerou a sekundární konstrikcí.

Veškeré mikromanipulační pokusy byly dělány ve spolupráci s Ústavem veterinárního lékařství v Brně.

1.3.3 Celogenomová amplifikace

Chromozómy/chromozomální ramena byly amplifikovány dvoustupňovou amplifikací komerčně dostupnými kity WGA4 GenomePlex Single Cell Kit a WGA3 GenomePlex WGA Reamplification Kit (obojí Sigma-Aldrich). V první fázi byla DNA amplifikována WGA4 kitem dle instrukcí výrobce a produkt byl přečištěn přes kolonku Quiquick Gel Extraction Kitu (Quiagen). 20ng získaného produktu bylo následně reamplifikováno dle instrukcí výrobce kitem WGA4, produkt z tohoto kroku byl přečištěn srážením v etanolu. Celkový výtěžek z celogenomové amplifikace a reampifikace bylo přibližně 20μg DNA fragmentů o délce 500-600 bp.

1.3.4 Tvorba knihovny a sekvenování nové generace

DNA byla sekvenována metodou Illumina v MRC National Insitute for Medical Research, Londýn, Velká Británie.

Po dvoustupňové amplifikaci obsahovaly fragmenty DNA na svých koncích adaptorové sekvence. Pro eliminaci jejich sekvenování byl použit Nextera DNA sample kit (Illumina) dle instrukcí výrobce. Při použití tohoto kitu primery nenasedají na fragmentovanou DNA do koncových oblastí, a tudíž nedochází k následné sekvenaci adaptorů. 50ng produktu bylo smícháno s Nextera Illumina- Compatible Enzym-Mix a nízkomolekulárním pufrem. Vytvořená knihovna obsahovala fragmenty DNA o délce přibližně 200-400 bází, z čehož 135 bází byla adaptorová sekvence pro nasedání NGS primerů. Kvalita DNA knihovny byla určena kity KAPA library qualification kit a QuBit na přístroji Agient Bioanalyser. DNA byla osekvenována na přístroji Illumina, délka jednotlivých čtení byla 80 bazí.

1.3.5 Analýza dat získaných ze sekvenace

Z Illumina Sekvenace jsme získali celkem 35 miliónů čtení o délce 80 bází. Tato čtení byla bioinformatickým softwarem Bowtie (http://bowtie-bio.sourceforge.net) namapována do dvou verzí genomu *X. tropicalis* (v.4.1 a v.7.1). Do následné analýzy byla zahrnuta jen čtení s maximálně dvěma odlišnými bázemi oproti genové sekvenci (sekvence s 97,5% a větší shodou), k čemuž byly použity Bowtie indexy m=1 a v=2. Pro shrnutí výsledků byly použity PERL skripty, Microsoft Excel a vizualizace zásahů do jednotlivých scaffoldů pomocí programu GBrowse (http://gbrowse.org).

1.3.6 FISH-TSA

Fyzické mapování genů z vybraných scaffoldů bylo provedeno metodou Fluorescenční In Situ Hybridizace spojené s tyramidovou amplifikací dle (Krylov et al., 2007). Chromozómy pro mapování pocházely z výše uvedené primární testikulární kultury. Geny mapované při analýze krátkého raménka chromozómu 7 pocházely ze scaffoldů nerovnoměrně pokrytých zásahy.

1.3.7 Příprava malovacích sond

Pro přípravu malovacích sond byly použity dva přístupy- dvoustupňová DOP-PCR s Dig-11-dUTP dle (Kubickova et al., 2002) a výše uvedená metoda využívající WGA4 a WGA3 kity (Sigma-Aldrich). Při použití WGA 3 kitu byl oproti instrukcím výrobce nahrazen zásobní roztok 1,5µl 10mM dATP,dCTP,DTP a 6,5 mMd dTTP s 2 µl 1µM Dig-11-UTP.

1.3.8 Zoo-FISH

Sondy z mikrodisektovaných chromozómů *X. tropicalis* byly hybridizovány nejprve s chromozómy téhož druhu pro potvrzení jejich kvality a následně na chromozómy příbuzného druhu *X. laevis*. Hybridizační roztok obsahoval 50% formamid, 2xSSC, 10% dextran sulfát, 5µg kompetitorové DNA *X. tropicalis* (Applied Genetics Laboratories, Melbourne, USA) a 2ml sondy. Tento roztok byl denaturován při 72°C po dobu 5 minut a renaturován 80 minut při 37°C. Sklíčka s chromozómy byla 2 minuty denaturována v 70% formamidu s 2xSSC, dehydrována v rostoucí řadě metanolu a nechána oschnout. K takto ošetřeným sklíčkům byl přidán hybridizační roztok, preparáty byly přikryty krycím sklíčkem a hybridizace probíhala přes noc ve vlhké komůrce při 37°C. Poté byla sklíčka dvakrát omyta 50% formamidem, 2x v SSC a 1x v TNT, vždy 5 minut, blokována 30 minut TNB pufrem a sonda byla detekována systémem antidigoxigenin-tretrametylrhodamin (Roche diagnostics, GmbH, Německo). Nakonec byla sklíčka omyta TNT pufrem a zamontována v montovacím médiu s DAPI (Intimex, Česká republika). Sondy byly zobrazeny pod mikroskopem (Olympus BX40), obrázky foceny kamerou SONY a upraveny programem ACC (SOFO, Česká republika).

Výše uvedený postup byl použit pro hybridazi sond s chromozómy obou druhů (*X. tropicalis* i *X. laevis*), hybridizační krok byl u preparátů z buněk druhu *X. laevis* prodloužen na 72 hodin.

1.4. Výsledky a diskuze

1.4.1. Genetická mapa druhu *Xenopus tropicalis*

Prvním projektem této dizertační práce bylo sestavení vazebné a fyzické mapy druhu *Xenopus tropicalis*. Vazebné mapování bylo provedeno ve spolupracujících laboratořích Laboratory of Biology and Biochemistry, Houston, USA a v Human Genome Sequencing center, Houston, USA. Na základě výsledků vyplývajících z vazebné mapy byly vybrány geny pro fyzické mapování, které proběhlo v naší laboratoři.

Za účelem vazebného mapování byli kříženi jedinci inbredních kmenů Nigerian a Ivory coast. Mezi 190 jedinci F2 generace bylo nalezeno celkem 2886 SSLP markerů (Simple Sequence Length Polymorphism), které byly programem JoinMap3.0 sestaveny do vazebné mapy. Ta měla celkem deset velkých vazebných skupin o délce 77-138 cM a 103-506 markerů. Zbylé 4 malé vazebné skupiny obsahovaly dohromady 54 markerů a byly 6-40 cM dlouhé. Vzniklá mapa obsahovala 758 z celkových 1300 scaffoldů delších než 100kbp a zahrnovala přibližně 62% z genomu *X. tropicalis* v.4.1.

Poloha malých vazebných skupin a rovněž přiřazení jednotlivých vazebných skupin k chromozómům bylo zjišťováno pomocí Fluorescenční In Situ Hybridizace spojené s pyramidovou amplifikací (FISH-TSA) (Krylov et al., 2007). Celkem bylo lokalizováno 69 genů. Malé vazebné skupiny byly lokalizovány na chromozómech 3, 5 a dvě na chromozómu 8. Velké vazebné skupiny byly přiřazeny k příslušným chromozómům a rovněž byla určena jejich orientace (p-q, q-p). Markery z jedné vazebné skupiny vždy mapovaly na stejný chromozóm, čímž byla potvrzena správnost vazebného mapování i assembly v.4.1. Na druhou stranu nebyla shodná poloha některých markerů v assembly v.4.1 a ve vazebné mapě. Celkem 45 scaffoldů obsahovalo markery s pozicí ve dvou a více vazebných skupinách. Chimerismus těchto scaffoldů nicméně nebyl potvrzen jinými metodami.

Fyzická mapa rovněž poukázala na nedostatek vazebné mapy u chromozómu 2, který neměl vazebně zmapováno celé krátké raménko. Důvodem mohla být nepřítomnost zvolených markerů či změněná rekombinace v této oblasti. q rameno chromozómu 2 vykazuje velkou syntenii s lidským chromozómem 2 a dlouhou dobu existovalo podezření, že se zde nachází gen určující pohlaví (Mácha et al., 2012). Ten je ovšem pravděpodobně umístěn na druhé oblasti nezahrnuté ve vazebné mapě- na distálních 15 cM krátkého raménka chromozómu 7.

Zde bylo nalezeno 22 markerů se silnou vazbou na samičí pohlaví (Olmstead et al., 2010), což mohlo vést ke snížené rekombinaci této oblasti spojené se sníženým množstvím polymorfních markerů.

Na základě této genetické mapy byla sestavena nová verze genomu druhu *Xenopus tropicalis* v.7.1. a stala se důležitým zdrojem informací pro genetické studie.

1.4.2 Příprava malovacích sond druhu *X. tropicalis* z mikrodisektovaných chromozómů a rekonstrukce tetraploidního karyotypu *X. laevis* metodou Zoo-FISH

K oddělení druhů *Xenopus tropicalis* a *Xenopus laevis* došlo přibližně před 50-65 miliony let (Evans et al., 2004). *X. laevis* je alotetraploidní druh, který vznikl spojením dvou příbuzných diploidních druhů a v dnešní době je již funkčně diploidizován. I přes velmi dlouhou evoluci existuje u druhů *X. laevis* a X. *tropicalis* značná podobnost chromozomálních kvartetů, kterou jsme potvrdili technikou Zoo-FISH.

Z buněk primární testikulární kultury druhu *X. tropicalis* bylo vyřezáno přibližně 15 kopií každého chromozómu. Kontrolní sonda byla vyrobena z chromozómu 9 technikou DOP-PCR s Dig-11-UTP dle (Kubickova et al., 2002). Tento přístup se ukázal jako neúspěšný, protože sonda značila na preparátu pouze sekundární konstrikci nacházející se na příslušném chromozómu. Pro výrobu další sondy byla použita celogenomová amplifikace kitem GenomePlex Single Cell WGA kit (WGA4). 20 ng produktu bylo označeno Dig-11-dUTP kitem GenomePlex DNA reamplification kit (WGA3). Takto vyrobené sondy byly ve značení sice úspěšnější, ale ani v tomto případě nedošlo k označení celého chromozómu 9. Příčinou byl pravděpodobně fakt, že se v oblasti sekundární konsrikce nachází velké množství repetic a chromozóm 9 nebyl proto pro výrobu kontrolní sondy příliš vhodný.

Ostatní sody vyrobené pomocí WGA4 a WGA3 kitů barvily vždy pouze příslušné chromozómy. Výjimkou byly sondy z chromozómů 2 a 4, které značily nespecificky i jiné chromozomální oblasti. S největší pravděpodobností došlo k chybě při mikrodisekci, protože tyto dva chromozómy jsou pod mikroskopem obtížně rozlišitelné. Z tohoto důvody nebyly tyto sondy pro následnou Zoo-FISH studii použity.

Sondy z chromozómů druhu *Xenopus tropicalis* byly hybridizovány s chromozómy druhu *Xenopus laevis*. Výsledky jsou shrnuty v Tabulce 1.

Xenopus tropicalis	1	3	5	6	7	8	9	10
Xenopus laevis	chr1+ chr 2	qchr 12+ chr 16	chr 13+ chr 17	chr 6+ chr 9	chr 7+ chr 10+ satelit chr 12+ p13	chr 11+ chr 14	1/2q chr 15+ 1/2q chr 18	Pericentrická oblast chr.14+chr 18

Tabulka 1: Zoo-FISH analýza druhu *X. laevis* použitím sond pocházejících z druhu *X. tropicalis*

Sonda z chromozómu 9 druhu *X. tropicalis* (Xt) značila na chromozómech *X. laevis* (Xl) pouze malou oblast. Pravděpodobně se opět jednalo o následek velkého množství repetic na chromozómu Xt9, nebo toho, že *X. laevis* nemá na svých chromozómech žádnou sekundární konstrikci.

Zoo-FISH studie potvrdila výzkum rekonstrukce tetraploidního genomu druhu *X. laevis* (Schmid and Steinlein, 1991) a zároveň podpořila teorii o vzniku tohoto druhu ze dvou jiných diploidních druhů (2n=20), u kterých nejprve došlo k translokaci jednoho chromozómu jiný chromozóm. Následně se tyto druhy zkřížily, po oplození ovšem nedošlo k mitóze a vzniklý druh byl tetraploidní (2n=18, 4n=36). Pravděpodobně se jednalo o chromozóm odpovídající chromozómu Xt10, protože žádný jiný druh z rodu Xenopus tento chromozóm nemá. Protože sonda z chromozómu Xt10 značila chromozómy Xl14 a Xl18, které byly zároveň značeny sondami z chr. Xt8 a Xt9, předpokládáme, že chromozóm Xt10 translokoval na tyto dva chromozómy.

1.4.3 Efektivní vysokorychlostní sekvenace laserově mikrodisektovaného chromozomálního ramene

Jak již bylo několikrát řečeno, doposud se žádnými metodami nepodařilo vytvořit kompletní mapu genomu druhu *X. tropicalis*. Ve vazebné mapě se stále nachází velké nemapované oblasti (především celé krátké rameno chromozómu 2 a 0-15 cM na krátkém raménku chromozómu 7), a obě současné verze genomu obsahují více než 10 tisíc scaffoldů s neznámou polohou. Protože není možné všechny geny mapovat fyzicky, byla vynalezena nová metoda pro genetické mapování. Jejím základním principem je mikrodisekce chromozomální oblasti a její následná sekvenace přístupem vysokorychlostní sekvenace nové generace. Tímto způsobem je možné k vyřezané oblasti přiřadit velké množství genů a

scaffoldů a rovněž poukázat na špatně sestavené (chimérické) scaffoldy, které se v příslušné oblasti nachází.

V této studii jsme se zaměřili na krátké raménko chromozómu 7, které je díky sekundární konstrikci snadno rozlišitelné pod mikroskopem, obsahuje vazebně nemapovanou oblast a pravděpodobně i gen určující pohlaví (Olmstead et al., 2010). Nejprve bylo vyřezáno 15 kopií p ramen chromozómu 7, DNA byla amplifikována kity WGA4 a WGA 3 a takto získaný materiál byl sekvenován metodou Illumina. Celkově jsme získali více než 35 miliónů čtení o délce 80 bazí. Bioinformatickým softwarem Bowtie (Langmead et al., 2009) byla tato čtení namapována do obou současných verzí genomu a pozice jednotlivých zásahů ve scaffoldech byla zobrazena programem GBrowse.

Do starší verze assembly v.4.1 mapovalo s minimálně 97,5 % shodou 3.900.340 čtení. Všechny scaffoldy byly seřazeny podle počtu zásahů na 1 kb jejich délky. Scaffoldy mapované fyzicky na p rameno chromozómu 7 a i většina scaffoldů mapovaných vazebně měla koeficient počtu zásahů/kilobázi vyšší než 29. V oblasti vazebné mapy odpovídající centromeře tento poměr rapidně klesl. Scaffold 297, fyzicky namapovaný těsně pod centromerou byl s 17 zásahy na kb určen jako „hraniční scaffold“, tedy scaffold s nejmenším počtem zásahů na kb. Tento scaffold se pravděpodobně nacházel na hranici řezu a nemusel být přítomen ve všech vyřezaných raménkách. Všechny zásahy do scaffoldů s nižším poměrem zásah/kb byly vzaty jako falešně pozitivní. Celkem šlo o 7% zásahů.

Analýza zásahů do scaffoldů ukázala, že jich velké množství (67%) mapovalo do scaffoldů již dříve namapovaných do p raménka chromozómu 7. 22,5% pak mapovalo do scaffoldů bez známé polohy a 3,5% do scaffoldů nacházejících se v jiných oblastech. Protože všechny tyto scaffoldy měly vysoký poměr zásahů na kilobázi, předpokládáme, že se náchází ve studované oblasti.

Nejzajímavější výsledky přineslo zobrazení zásahů do scaffoldů. Ukázalo se, že se přibližně u 15% scaffoldů s velikostí přes 100kb vyskytují velké oblasti bez jediného zásahu. Tyto oblasti obsahovaly v porovnání s hustě pokrytými oblastmi přibližně stejné množství genů. FISH analýzou fyzické polohy několika vybraných genů ze scaffoldů 75, 266 a 270 jsme zjistili, že se geny z nepokryté oblasti nenacházejí na chromozómu 7, zatímco geny z oblastí s velkým a množstvím zásahů byly vždy lokalizovány do mikrodisektované oblasti. Tyto výsledky nasvědčují, že je až 15% scaffoldů z assembly 4.1 sestaveno chybně.

Analýza nové verze assembly (v.7.1) rovněž poukázala na některé její nedostatky. Většina zásahů (80%) mapovala do superscaffoldu 7 (odpovídajícímu chromozómu 7), ale bylo nalezeno i několik oblastí s velkým množstvím zásahů na jeho dlouhém raménku a rovněž v jiných supercaffoldech. V assembly 7.1 se nám podařilo do vyřezané oblasti přiřadit 14 scaffoldů.

Přestože se nejedná o první pokus sekvenace mikrodisektovaných chromozomálních ramen, je naše studie první, která tuto techniku použila pro účely genetického mapování. Navíc byl náš přístup v porovnání s dříve použitou metodikou výrazně úspěšnější. Oproti studii (Weise et al., 2010), ve které byla pro celogenomovou amplifikaci použita DOP-PCR a pro sekvenaci 454 pyrosekvenování, jsme obrželi více než 30.000x víc čtení a 10.000x víc zásahů do genomu.

1.4.4 Velké pseudoautosomální oblasti na pohlavních chromozómech žáby *Silurana* (*Xenopus) tropicalis*

Pohlavní chromozómy jsou u druhu *Xenopus tropicalis* morfologicky nerozlišitelné a nepodařilo se je určit ani po sekvenaci samičího genomu (Hellsten et al., 2010). Oblast určující pohlaví se dle studie nachází na krátkém raménku chromozómu 7 (Olmstead et al., 2010). To, že se dlouhou dobu nedařilo určit oblast určující pohlaví, je pravděpodobně zapříčiněno velmi malými rozdíly mezi pohlavními chromozómy, které jsou způsobeny mnohonásobnou (až 32x) změnou pohlavních chromozómů v průběhu evoluce (Schmid et al., 2012; Evans and Charlesworth, 2013). Další příčinou může být nezměněná rekombinace na pohlavních chromozómech.

Cílem této studie bylo nalézt rozdíly mezi samčím a samičím chromozómem 7. Metodou „RAD tags“ byl sekvenován genom čtyř samců a čtyř samic a získaná čtení byla mapována do genomu *X. tropicalis*. Předpokládali jsme, že do genomu bude mapovat více sekvencí získaných z heterogametického samičího pohlaví, protože samci by měli některé části genomu postrádat. To se ovšem nepotvrdilo. Absolutní počet mapovaných čtení byl sice u samic vyšší, ovšem do chromozómu 7 jich mapovalo v porovnání s ostatními chromozómy zhruba stejné množství.

Přesvědčivé výsledky nepřinesla ani analýza sekvenace krátkého raménka chromozómu 7. Chromozómy pocházející ze samců postrádaly pouze 2 z 22 scaffoldů s velmi

silnou vazbou na pohlaví ze studie (Olmstead et al., 2010), zatímco podle našich předpokladů by se na samčích chromozómech neměla nacházet většina či žádný z nich. Kromě toho nebyly nalezeny pohlavní rozdíly u markerů použité ve výše uvedené studii. Je ovšem možné, že není možné při použití současných verzí genomu druhu *Xenopus tropicalis* takové rozdíly nalézt, protože se v nich příslušné oblasti vůbec nenachází.

Naší studií jsme nenalezli žádné rozdíly mezi pohlavními chromozómy samic a samců, a proto předpokládáme, že je oblast určující pohlaví velmi malá. Tento závěr odpovídá i starší studii, kdy nebyly nalezeny rozdíly mezi bandovanými samčími a samičími chromozómy (Uno et al., 2008a). Kromě toho existuje možnost, že je pohlaví u druhu *X. tropicalis* určeno nějakým složitějším způsobem, jako je tomu například u ryby *Danio renio* (zebřička). U té je pohlaví určeno několika geny nacházejících se na různých chromozómech (Anderson et al., 2012).

1.5 Závěr

V minulém desetiletí byly naše poznatky o genetice druhu *Xenopus tropicalis* mnohonásobně prohloubeny. Přispěly k tomu i publikace presentované v této práci. V rámci prvního projektu byla sestavena první genetická mapa obsahující deset vazebných skupin, které odpovídají deseti chromozómům druhu *X. tropicalis*. Fyzickým namapováním celkem 69 genů se podařilo přiřadit sestavené vazebné skupiny k příslušným chromozómům. Genetická mapa zahrnující minimálně 62% genomu tohoto druhu (1.055Gb) sloužila jako důležitý zdroj dat pro sestavení nové verze genomu (v.7.1). Hlavním nedostatkem mapy jsou nezahrnuté krátké rameno chromozómu 2 a distální konec p ramene chromozómu 7. I přes tuto nedokonalost se jedná o velmi důležitý nástroj pro genetiku, genomiku, cytogenetiku i evoluční studie.

Ve druhé studii se nám podařilo vyvinout techniku pro výrobu sond z mikrodisektovaných chromozómů druhu *X. tropicalis*. Malovací sondy použité pro Zoo-FISH studii příbuznosti s druhem *Xenopus laevis* poukázaly na velkou podobnost chromozómů těchto druhů i po 50-65 miliónech let samostatného vývoje. Dále jsme zjistili, že v evoluci vznikly minimálně dva druhy k nereciproční translokací chromozómu 10 druhu *X. tropicalis* na chromozómy 8 a 9. Poté došlo k jejich křížení, což vedlo ke vzniku nových druhů, mimo jiné i *Xenopus laevis*.

V dalším článku jsme vyvinuli novou techniku pro genetické mapování. Mikrodisekcí krátkého raménka chromozómu 7 s následnou celogenomovou amplifikací a sekvenací vysocerychlostní metodou Illumina jsme získali dostatečné množství čtení pro přiřazení více než 5 Mb nových sekvencí ke krátkému raménku chromozómu 7. Navíc se nám podařilo poukázat na velké nedostatky obou dostupných verzí genomu, včetně špatně sestavených oblastí. Tuto techniku je možné použít pro genetické mapování i u živočišných druhů s neznámou sekvencí a rovněž v klinickém výzkumu či biomedicíně pro zjištění translokací genomu.

Poslední studie se zabývala pohlavními chromozómy u druhu *Xenopus tropicalis*. Protože se nám žádnou z použitých technik nepodařilo nalézt sekvenční rozdíly mezi pohlavími, předpokládáme, že u tohoto druhu je oblast určující pohlaví velmi malá a k rekombinaci může docházet téměř v rámci celého pohlavního chromozómu. Hlavní příčinou této skutečnosti je pravděpodobně opakovaná změna pohlavních chromozómů a častá polyploidizace, ke kterým u obojživelníků dochází.

Hlavním cílem této práce bylo vyvinutí nových přístupů a metod pro genetické mapování aplikovatelných i u jiných živočišných druhů. Použitý přístup využívající laserově mikrodisektované chromozomální rameno a jeho NGS sekvenaci nebyl za podobným účelem nikdy použit. Oproti podobným studiím byla naše metodika úspěšnější jak v počtu získaných sekvencí, tak v počtu sekvencí mapovaných do genomu. Tato technika má navíc velmi malé nároky na množství vstupního materiálu, což zcela jistě umožňuje její využití i v dalších oblastech výzkumu či klinické medicíny.

2. English part

2.1. Introduction

Genome of *Xenopus tropicalis* is 1,7Gb large and is divided into 10 chromosomes (1n). In this work, the classification in which chromosomes are numbered according to their decreasing size (Khokha et al., 2009) is used. Genome of *X. tropicalis* has been sequenced (Hellsten et al., 2010), there are currently two versions of its assembly. The older version (v.4.1) contains approximately 19,000 scaffolds and is 1.51 Gbp large. Although the assembly v.4.1 was published in 2005, it is still updated and used because of the controversies in the newer version (v.7.1). On the other hand, both versions of *Xenopus tropicalis* genome are extremely uncompleted and up to 25% of bases in scaffolds sequence are unknown (represented by "N"). The difficulties in assembly probably originate in large amount of repetition of *X. tropicalis* genome, resulting in incapability of bioinformatics software link/connect contigs to larger complexes. The newer assembly version (v.7.1), released in 2011, was based on the data from arising genetic map and synteny with other species. Approximately 70% of the genome is incorporated to 10 superscaffolds representing 10 chromosomes. Differences between these two versions of assemblies are huge- the same gene can even be placed in two different chromosomes. Moreover, the v.7.1 is considered very controversial due to the use of unverified synteny data.

For *X. tropicalis* genome sequencing the "traditional" Sanger sequencing method was used. However, there are also new techniques of Next Generation Sequencing (NGS), such as Illumina, 454-pyrosequencing, Solid etc. These approaches are also called "high- throughput" and produce millions of reads (approximately 30-100bp long) in a short time. This type of sequencing can be applied for genome assembly, but the shortness of reads requires at least 30x coverage. In clinical medicine, the NGS approaches are used in many contexts, including reproductive medicine, oncology and infectious disease. The mostly used source of DNA is whole-nuclear, but also individual chromosomes can serve for the purposes of NGS sequencing.

Chromosomes can be separated with two techniques: flow sorting or microdissection. Flow sorting produces high numbers of chromosomes, but until now only few protocols for sorting exist, and these are mostly focused on human chromosomes. However, any

distinguishable chromosome or even chromosomal part can be separated with laser microdissection. On the other hand, this approach yields only few chromosomes and for following research methods is usually required some type of whole- genome DNA amplification (WGA).

There are four different types for whole genome DNA amplification- PEP, DOP-PCR, MDA and OmniPlex. The basic principle of these methods is annealing of non-specific primers (PEP, DOP-PCR, MDA), or ligation of adapter sequences to fragmented DNA, which serve for annealing of specific primers (OmniPlex). The disadvantage of all the approaches is disability to amplify an entire genome. The reason is obvious- the amount of template DNA is usually small and it can also be damaged after some types of isolation. Although the primers are short, some fragments are replicated more frequently. According to (Höckner et al., 2009), the highest rate of successfully amplified markers was achieved with GenomePlex Single Cell Kit linker adapter PCR approach (Sigma-Aldrich, Vienna, Austria), which uses the OmniPlex WGA method. In this study, chromosomes were obtained by needle microdissection. So far, the following NGS sequencing of microdissected chromosomes was successful only while using GenomePlex kit for the whole genome amplification. (Weise et al., 2010), microdissected 10 arms of human chromosome 12, amplified DNA with DOP-PCR and sequenced with 454-pyrosequencing leading to only app. 1000 obtained reads from which 154 were mapped to the target. On the other hand, (Ma et al., 2010) determined more than 24,000 human haplotypes from microdissected chromosomes while using GenomePlex kit and Illumina sequencing.

The genome of a model organism can also be completed on the basis of a genetic map. The linkage map shows relative position of markers with the respect to recombination frequency (1% of recombination= 1centimorgan, cM). Many markers can be placed in one run, but the relationship between the distance in bp and cM is nonlinear. The recombination frequency depends on the position of markers - is decreased around a centromere, in the sex determining area and contrary decreased in hot spots. Moreover, even genes on different chromosomes can be in genetic linkage. The last disadvantage of linkage mapping arises from the type of chosen markers- some chromosomal parts can lack them and thus these are not included in the map. Physical mapping using FISH (Fluorescent In Situ Hybridization) technique lacks these disadvantages because of a direct visualization of markers to a chromosome. However, this approach is more complicated and enables placing only few

markers simultaneously. In *Xenopus tropicalis*, the protocol for physical mapping with FISH technique is available (Krylov et al., 2007).

Because this amphibian has homomorphic (undistinguishable) sex chromosomes and lacks all known sex genes, the sex chromosomes and even the type of sex determination were unknown until 2010. Even after the genome was sequenced, the sex determining area remained undiscovered. (Olmstead et al., 2010) found 22 AFLP markers with linkage to female sex. According to our linkage map (Wells et al., 2011) the position of these markers is on the distal part of the short arm of the chromosome 7.

2.2. Aims of the study

The main goal of this study was to find new approaches for genetic mapping applicable in model organism *Xenopus tropicalis.*

The specific aims were:

- Compilation of a genetic map of *Xenopus tropicalis*. Because this project was very complex, our laboratory focused on the physical mapping of genes from chosen scaffolds.

- Optimization of laser microdissection of *X. tropicalis* chromosomes, whole genome amplification and using labeled DNA as probes for Zoo-FISH study for painting *Xenopus laevis* chromosomes.

- Identification of scaffolds and genes on the short arm of *X. tropicalis* chromosome 7 with newly developer approach using next-generation sequencing of microdissected chromosomes

- Characterization of the sex chromosomes of *X. tropicalis*

2.3. Methods

2.3.1 Preparation of cells

Metaphase spreads were prepared from *X. tropicalis* euploid primary cell culture of dissected testes. Cells were trypsinized and hypotonized for 5 minutes in 38 mM KCl, fixated and stored overnight at -20°C.

2.3.2 Laser microdissection

Cells were dropped on a thin glass slide with a polythene naphthalene membrane (P.A.L.M Gmbh, Bernried, Germany), allowed to dry and stained for 10 minutes with 3% Giemsa in H_2O. Approximately 15 copies of chromosomes or chromosomal arms were harvested using PALM MicroLaser system (P.A.L.M Gmbh, Bernried, Germany), coupled with an inverted microscope (Olympus) with 100x magnification under an oil immersion objective. With a single laser pulse, chromosomes/chromosomal arms were catapulted into the cap of PCR tube with 4μL PCR oil. Chromosomes were distinguished according to their size and the presence p/q arm ratio according to (Tymowska, 1973), nomenclature follows (Khokha et al., 2009).

The laser cut on the short arm of the chr.7 was led under the centromere for ensuring dissection of a whole arm. Chromosomes were microdissected in the cooperation with Veterinary Research Institute, Brno.

2.3.3 Whole genome amplification

WGA was performed in two steps. In the first step, WGA4 GenomePlex Single Cell Kit (Sigma-Aldrich) was used according to manufacturer´s manual. PCR product from primary amplification was purified with Quiquick Gel Extraction Kit (Quiagen) and 20 ng was reamplified with GenomePlex WGA Reamplification Kit (Sigma-Aldrich). The product of reamplification was cleaned up by ethanol precipitation. The total yield from starting 15 chromosomes was approximately 20 μg of DNA with the length of fragments approximately 500-600 bp.

2.3.4 High throughput sequencing and library construction for next generation sequencing

The high throughput sequencing was performed in MRC National Institute for Medical Research, London, United Kingdom.

Due to two runs of WGA amplification, the fragments of DNA contained WGA adaptors at both ends. To reduce their sequencing, the fragmenting/adaptor ligation method was used. DNA was fragmented and simultaneously primers for sequencing were aligned, using Nextera DNA sample prep kit (Illumina, Inc.) according to manufacturer´s instructions. With this approach, primers were inserted away from DNA ends. 50 ng of product was mixed with Nextera Illumina-Compatible Enzym Mix and low molecular buffer, yielding libraries with the fragments size 200-400 bp. The length of fragment included 135 bp of adapter sequence. The quality of libraries was established with Agilent Bioanalyser, OT-PCR using KAPA library quantification kit and QuBit. 6.5 pM of DNA was loaded on a lane and sequenced on Illumina NGS sequenator. The lengths of reads were 80bp.

2.3.5 Data analysis

35 million of 80 bases long reads were obtained from Illumina sequencing. The reads were mapped to both versions of *X. tropicalis* assemblies (v.4.1 and v.7.1) using bioinformatics software Bowtie (http://bowtie-bio.sourceforge.net). As high number of repetitions and WGA-primers were expected, only unique hits with 97, 5% and higher identity (less than 2 mismatches per 80 bases) were selected (Bowtie indexes m=1 and v=2). The selected sequences were counted and analyzed with PERL scripts and Microsoft Excel. For the visualization of hits to scaffolds, we used G-Browse (GBrowse)(http://gbrowse.org).

2.3.6 FISH-TSA

Metaphase spreads for FISH-TSA were prepared from the same culture as chromosomes for laser microdissection. Genes for visualization were chosen from selected scaffolds and also from “hit rich” and “hit poor” regions from above G-Browse hit analysis. The protocol for FISH-TSA followed (Krylov et al., 2007) without changes.

2.3.7 Preparation of chromosome painting probes

Two techniques were used for a preparation of a whole chromosome painting probes: two rounds of DOP-PCR with Dig-11-dUTP according to (Kubickova et al., 2002) and

Genome Plex Single cell WGA4 followed with WGA3 reamplification kit (Sigma Aldrich). WGA4 kit was used according to manufacturer´s instructions. 20ng of amplification product was labeled with WGA3 Reamplification kit (Sigma Aldrich), with minor changes- the stock solution was replaced with 1,5µl of 10mM dATP, dCTP,dGTP and 6,5mM dTTP and 2 µl of 1µM Dig-11-UTP.

2.3.8 ZOO-FISH

The protocols for the chromosomes painting were very similar in both species (*X. tropicalis* and *X. laevis*), only in the case of *X. laevis* chromosomes, the hybridization step was prolonged up to 72 hours and the volume of the probe was 1,5x increased.

The hybridization mixture for painting *X. tropicalis* chromosomes contained 50% formamide, 2x SSC, 10% dextran sulphate, 5µg of *X. tropicalis* competitor DNA (Applied Genetics Laboratories, Melbourne, USA) and 2ml of probe. This mixture was denaturized for 10min in 72°C and reannealed for 80 min at 37°C. Slides with chromosomal spreads were denatured in 70% formamid, 2 x SSC for 2 min, dehydrated in 70%, 90% and 100% methanol (3 min for each step) and let dry out. The hybridization mixture was added to chromosomes and hybridized overnight under 22x22 mm coverslip in a wet chamber at 37°C. After that, the slides were washed twice in 50% formamide, than in 2x SSC at 42°C, 3x in 2xSSC and 1xTNT at RT, each for 5 min. Slides were blocked with TNB buffer for 30 min, the probe was detected with antidigoxigenin- tetramethylrhodamin (Roche diagnostics, GmbH, Germany). Finally, the slides were washed with TNT buffer and mounted in medium containing DAPI (Intimex, Czech Republic). The probes were visualized with Olympus BX40 microscope, pictures taken with SONY camera and images processed with ACC program (SOFO, Czech Republic).

2.4. Results and discussion

2.4.1 A genetic map of *Xenopus tropicalis*

In the first study, the genetic map of *Xenopus tropicalis* was compiled. This project was very complex and consisted of two separate parts- firstly; the linkage map was generated and on its basis a physical map was created.

The linkage map was assembled in the Laboratory of Biology and Biochemistry, Houston, USA and in Human Genome Sequencing center, Houston, USA. It included 2886 SSLP markers (Simple Sequence Length Polymorphism). Polymorphic markers were searched between inbreed Nigerian and Ivory Coast Strains individuals. The linkage analysis of markers in 190 siblings in F2 cross panel was performed with JoinMap 3.0. The map consisted of 10 major and 4 minor linkage groups (LG). Each major LG included 103-506 markers with length 77-138 centimorgans. The rest 54 markers yielded 4 minor LGs, which length was 6-40cM. The total length of the linkage map covered 1345, 9 cM. A comparison of the map and v.4.1 assembly revealed, that it included 758 from 1300 scaffolds larger than 100kbp and incorporated at least 62% of *Xenopus tropicalis* genome (1,055 Gbp).

The places of some markers were not in correspondence with v.4.1 assembly. Markers from 45 scaffolds revealed position in two and more different linkage groups, indicating possible mistakes and occurrence of chimeric scaffolds in assembly v.4.1, which, however, was not confirmed by other methods.

The assigning of LG to chromosomes, orientation and position of 4 minor linkage groups was determined by Fluorescence *In Situ* Hybridization coupled with a Tyramide Amplification Step (FISH-TSA) (Krylov et al., 2007). For each chromosome and minor LG, several cDNA markers were found in Ensembl or GenBank database. Completely, we located 69 genes. All markers from the same LG were detected on the same chromosome, supporting correctness of the linkage analysis.

With FISH-TSA, all major linkage groups were assigned to chromosomes and their orientation was specified. Minor linkage groups were placed to chromosomes 3, 5 and two to chromosome 8. The numbering of LG was changed to correspond with the number of chromosome according to (Khokha et al., 2009). All chromosomes, excluding chromosome 2

and 7, were covered by corresponding linkage group over their whole length. The marker with linkage position 0cM in LG 2 (chromosome 2) physically mapped to the centromere, concluded that p arm of the chr.2 is not represented in the genetic map. The reason could be either low or no content of polymorphic markers or altered recombination frequency in the area. According to (Mácha et al., 2012), short arm of chr.2 shows high synteny with human chromosome X, which could lead to impossibility to map this area with linkage analysis. LG 7 contains 15cM large gap at its beginning. This area probably contains sex gene (Olmstead et al., 2010), which can lead to changed recombination frequency of the area.

The genetic map of *Xenopus tropicalis* became significant tool for genetic and genomic studies and also served as an important source of information for new v.7.1 assembly.

2.4.2 Preparation of *X. tropicalis* whole chromosome painting probes using laser microdissection and reconstruction of *X. laevis* tetraploid karyotype by Zoo-FISH

This study focused on the technique for chromosome microdissection of *Xenopus tropicalis* chromosomes, selection the method for whole genome amplification and establishing the Zoo-FISH technique in the species *X. tropicalis* (Xt) and its relative *Xenopus laevis (*Xl*)*. Although long (50-65 million years) separate evolution of these two species (Evans et al., 2004), there exist high similarity of chromosomal quartets, enabling usage of Zoo-FISH probes.

Approximately 15 copies of each *X. tropicalis* chromosome were microdissected from primary testicular culture. In the first run, the painting probe for chromosome 9 was prepared by two rounds of DOP-PCR with Dig-11-dUTP (Kubickova et al., 2002). This approach was unsuccessful, because the probe labeled only secondary constriction and surrounding area on the short arm of the Xt chr.9. The second method for a probe preparation was based on usage of GenomePlex Single Cell WGA kit (WGA4) followed by labeling 20ng of PCR product by Dig-11-dUTP with GenomePlex DNA reamplification kit (WGA3). Even this approach labeled only a part of the Xt chromosome 9, but this was probably caused by a high amount of repetitions in its sequence, and this method was chosen for the preparation of remaining probes. All probes except Xt2 and Xt4 labeled relevant Xl chromosomes specifically. Probes from Xt2 and Xt4 labeled non-specifically other chromosomes. This could be caused by

wrongly distinguished chromosomes during microdissection, because these two chromosomes are very similar.

The Xt probes were afterwards hybridized with *X. laevis* chromosomes. Probes labeled chromosomes as showed in the Table 1.

Xenopus tropicalis	1	3	5	6	7	8	9	10
Xenopus laevis	chr1+ chr 2	qchr 12+ chr 16	chr 13+ chr 17	chr 6+ chr 9	chr 7+ chr 10+ satelit chr 12+ p13	chr 11+ chr 14	1/2q chr 15+ 1/2q chr 18	Pericentric area of chr.14+chr 18

Table **1**: Labeling of *X. laevis* chromosomes with probes from *X. tropicalis* chromosomes

The Xt9 probe labeled only small parts of Xl chromosomes, probably also because of the high amount of repetitions. Another reason can be that Xl chromosomes lack secondary constriction located on Xt9. Labeling of probes from Xt1 and Xt6 correspond with (Schmid and Steinlein, 1991) study, in which was reconstructed the tetraploid karyotype of *X. laevis*.

Our results also support the theory that *X. laevis* originated from two diploid species (2n=20). There the same chromosome translocated independently to one recipient in the first and to another one in the second species. These two species with 18 chromosomes in the diploid set joint together through the allotetraploidization event. As a result *X. laevis* with 2n = 36 chromosomes emerged.

2.4.3 Efficient high- throughput sequencing of a laser microdissected chromosome arm

The genetic map of *Xenopus tropicalis* contained gaps, which were impossible to fill up with recent methods. Two largest areas without any mapped marker were whole short arm of chromosome 2 and 15 cM at distal end of p arm of chromosome 7. Event thought the *X.tropicalis* genome was sequenced (Hellsten et al., 2010), neither one of two latest assembly versions does cover whole ten chromosomes, including two mentioned areas. The main reason is a high amount of repetitions in its genome, which makes it impossible for assembly software to connect scaffolds to whole-chromosome superscaffold. Until now, the position of

scaffolds was usually determined by physical or linkage analysis, because both versions of assembly comprised many scaffolds with unknown location.

This study describes a new approach for localizing previously unmapped scaffolds and genes and also for recognizing of wrongly localized and even assembled scaffolds. It is based on microdissection of a whole chromosome or a chromosomal area, followed by two rounds of whole genome amplification with WGA4 and WGA3 kits, high-throughput sequencing with Illumina technology, and bioinformatics analysis of obtained reads.

We focused on the short arm of chromosome 7. This area is easily distinguishable under a microscope due to the presence of secondary constriction and is also supposed to contain sex determining locus (Olmstead et al., 2010). With previously described methodology, we obtained approximately 35 million Illumina reads with length of 80 bases. These were mapped by bioinformatics software Bowtie (Langmead et al., 2009) to both versions of assembly and the position of hits in scaffolds was visualized using the GBrowse software.

To the v.4.1 assembly 3,900,340 Illumina reads were mapped. For each scaffold, the ratio of hits per one kb of its length (hit/kb ratio) was counted. All scaffolds physically mapped to the short arm of chr.7 had hit/ length ratio higher than 29hit/kb, same as most scaffolds from the linkage map of short arm of chr. (=LG) 7. This value decreased rapidly in the area around the centromere (69-71 cM). A physically mapped scaffold 279 with a position between the centromere and the secondary constriction (17hit/kb) was used as bordering scaffold. All scaffolds with lower hit/kb ratio were accepted as false positive. We expect that scaffold 279 is located on the cutting edge.

From all hits to scaffolds, 67% mapped to scaffolds included in the linkage map of short arm of chr.7, 3,5% to scaffolds from other areas and 22,5% hits were to scaffolds with unknown position. We expect that a physical position of these scaffolds corresponds with the microdissected area. The rest of the hits were to false positive scaffolds.

With visualization of hits to scaffolds, we found that approximately 15% of them were covered with hits unevenly with large (more than 100kb) gaps without any hit. These “hit-poor” areas contained genes and the amount of repetitions was comparable with other “hit-rich” areas. It is expected that these scaffolds were misassembled and their gapped parts are not situated on the microdissected area. Our assumption was confirmed by a FISH-TSA

analysis of genes from scaffolds 75, 266 and 270. Whereas all genes from hit-rich area located on short arm of chr.7, these from hit-poor areas located to chr.3 and two to chr.4.

Assembly v.7.1 contains 10 superscaffolds representing 10 chromosomes. 80% of hits from Illumina sequencing located to superscaffold 7, mostly to 0-60Mb area. The 0-15Mb area contained few smaller domains without any hit. 14 scaffolds without position were assigned to p arm of chr.7 by our analysis.

Our approach for sequencing of microdissected chromosomal area is innovative in usage of GenomePlex WGA kits and Illumina NGS sequencing. Compared to the previously used method in (Weise et al., 2010) study, where DOP-PCR was used for the whole genome amplification and 454 for NGS sequencing, we obtained more than 30.000x more reads and 10.000x more hits. This was also the first time the next generation sequencing was used for a purpose of a genetic mapping and verifying the genome assembly.

2.4.4 A Large Pseudoautosomal Region on the Sex Chromosomes of the Frog *Silurana tropicalis*

Xenopus (Silurana) tropicalis has homomorphic sex chromosomes. In (Olmstead et al., 2010) study, 22 markers with strong linkage to sex were found. Their position corresponds with the short arm of chromosome 7. Interestingly, the sex determining region was not characterized during genome sequencing (Hellsten et al., 2010), where sequenced genome was female. The difference between sexes is probably small, which can be caused by multiple (more than 32x) change of sex chromosomes (Evans and Charlesworth, 2013). Other reason may be a recombination along whole length of the sex chromosome.

The aim of the study was to characterize region with a strong linkage to sex. With Illumina technology for reduced genome sequencing approach ("RAD tags"), many small portions of *X. tropicalis* genome were sequenced in four female and four male individuals. It was expected that more females´ sequences would map to the genome, because this sex is heterogametic and males are expected to lack some parts of the genome. However, this was not confirmed, and the same proportions of reads were mapped in both sexes (87,8% -90,8% in females and 85-90,8% in males). The total number of mapped reads was higher in females, but there was no higher proportion of reads mapping to chromosome 7 compare to other chromosomes.

Neither Illumina sequencing of the microdissected male p arm of chr.7 produced significant results. This sex lacked only two out of 22 scaffolds with strong linkage to sex from (Olmstead et al., 2010) study, while our assumption was that male sex chromosome would not contain no or most of them.

We did not find significant sequence differences in markers used in (Olmstead et al., 2010), and any of them was fixed to female sex. However, current assembly (v.7.1) contains only 80% of whole genomic sequence and according to our previous study; genome of short arm of chr.7 contains misassembled regions. Thus, we were not able to find the sex differences on the bases of the assembly, because it lacks that portion of genome sequences.

In (Uno et al., 2008), no sex differences in C-banded heterochromatin of *X. tropicalis* were found. This corresponds with our results. We expect that the sex determining region is extremely small with minimal sex differences. The position of the sex gene still remains unknown. Moreover, the sex of *X. tropicalis* can be determined also by several genes on different chromosomes, same way as in zebrafish (Anderson et al., 2012).

2.5. Conclusion

All projects presented in this thesis are related to the genetics of the frog *Xenopus tropicalis*. In the first study, we compiled a genetic map of this specie. It covers all its 10 chromosomes and became a very important tool for genomics, genetics and also for cytogenetic and evolutionary studies. *X. tropicalis* genome (assembly v.7.1) was also newly assembled based on the genetic map. The genomic representation of the map is minimally 62% (1.055 Gb) of *X. tropicalis* genome and it incorporates 758 of the 1300 scaffolds larger than 100 kb. Physical map contains 69 genes mapped with FISH-TSA technique. The largest gap of the map includes whole short arm of chromosome 2 and distal end of p arm of chromosome 7.

Our other study addresses known shortcomings in both publicly- available assemblies of the *Xenopus tropicalis* genome as well as large gaps in the meiotic map of the chromosome 7, where sex determining of this species is thought to reside. Our new approach contributes more than 5 Mb of unmapped or incorrectly localized sequence; much of it is probably situated in the sex-determining area. We also identified poor sequence joins in both assemblies.

On the basis of data from the next generation sequencing of short arm of chromosome 7 and "RAD tags", we concluded that differences between sexes are minimal. The sex-divergent region was not identified in our research, not event in tightly linked region. Thus, the sex determining region is very small and nearly whole sex chromosome is autosomal. This fact can cause high incidence of sex chromosome turnover and also frequent polyploidization in amphibians.

The last study was focused on studying differences and relationship among chromosomes of two relative species X. *tropicalis* and X. *laevis*. The probes from microdissected *X. tropicalis* chromosomes were hybridized with *X. laevis* chromosomes. Some of them labeled relevant chromosomal quartets even after 50-65 million of separate evolution. Furthermore, we found out that *X. tropicalis* chromosome 10 underwent non-reciprocal translocation to at least two different ancient recipients before whole genome polyploidization.

Another aim of the studies was developing a new methods and approaches applicable even in other species. This was the first time the DNA from laser microdissected chromosomes of *X. tropicalis* was amplified and used as probe for Zoo-FISH or sequenced with high throughput technology. Moreover, the improved method for NGS sequencing of microdissected chromosomal arm produced 300x more reads compare to previously published methods. This approach is applicable also for species without sequenced genomes or for clinical application in medical cytogenetics and oncology, where tissue availability may be limiting. This method is likely to be of widespread use in species where individual chromosomes are distinguishable by cytological methods.

3. Publikace

Wells DE, Gutierrez L, Xu Z, Krylov V, Macha J, Blankenburg KP, Hitchens M, Bellot LJ, Spivey M, Stemple DL, Kowis A, Ye Y, Pasternak S, Owen J, Tran T, Slavikova R, Tumova L, Tlapakova T, Seifertova E, Scherer SE, Sater AK. **A genetic map of Xenopus tropicalis.** Dev Biol. 2011 Jun 1;354(1):1-8. doi: 10.1016/j.ydbio.2011.03.022. Epub 2011 Mar 31. PubMed PMID: 21458440; PubMed Central PMCID: PMC3098391.
IF = 3.868

Krylov V, Kubickova S, Rubes J, Macha J, Tlapakova T, Seifertova E, Sebkova N. **Preparation of Xenopus tropicalis whole chromosome painting probes using laser microdissection and reconstruction of X. laevis tetraploid karyotype by Zoo-FISH.** Chromosome Res. 2010 Jun;18(4):431-9. doi: 10.1007/s10577-010-9127-x. PubMed PMID: 20390340.
IF = 2,847

Seifertova E, Zimmerman LB, Gilchrist MJ, Macha J, Kubickova S, Cernohorska H,Zarsky V, Owens ND, Sesay AK, Tlapakova T, Krylov V**. Efficient high-throughput sequencing of a laser microdissected chromosome arm. BMC Genomics**. 2013 May28;14:357. doi: 10.1186/1471-2164-14-357. PubMed PMID: 23714049; PubMed Central PMCID: PMC3701504.
IF = 4.40

Bewick AJ, Chain FJ, Zimmerman LB, Sesay A, Gilchrist MJ, Owens ND, Seifertova E, Krylov V, Macha J, Tlapakova T, Kubickova S, Cernohorska H, Zarsky V, Evans BJ. **A large pseudoautosomal region on the sex chromosomes of the frog Silurana tropicalis**. Genome Biol Evol. 2013;5(6):1087-98. doi: 10.1093/gbe/evt073. PubMed PMID: 23666865; PubMed Central PMCID: MC3698919.
IF = 4.759

4. Použitá literatura

Bowtie: an ultrafast, memory-efficient short read aligner. http://bowtie-bio.sourceforge.net/.

Anderson JL, Rodríguez Marí A, Braasch I, Amores A, Hohenlohe P, Batzel P, Postlethwait JH. 2012. Multiple sex-associated regions and a putative sex chromosome in zebrafish revealed by RAD mapping and population genomics. PLoS One 7: e40701.

Evans BJ, Charlesworth B. 2013. The effect of nonindependent mate pairing on the effective population size. Genetics 193: 545-556.

Evans BJ, Kelley DB, Tinsley RC, Melnick DJ, Cannatella DC. 2004. A mitochondrial DNA phylogeny of African clawed frogs: phylogeography and implications for polyploid evolution. Mol Phylogenet Evol 33: 197-213.

GBrowse. GBrowse: Genome annotation viewer. [http://gmod.org].

Hellsten U, Harland RM, Gilchrist MJ, Hendrix D, Jurka J, Kapitonov V, Ovcharenko I, Putnam NH, Shu S, Taher L, Blitz IL, Blumberg B, Dichmann DS, Dubchak I, Amaya E, Detter JC, Fletcher R, Gerhard DS, Goodstein D, Graves T, Grigoriev IV, Grimwood J, Kawashima T, Lindquist E, Lucas SM, Mead PE, Mitros T, Ogino H, Ohta Y, Poliakov AV, Pollet N, Robert J, Salamov A, Sater AK, Schmutz J, Terry A, Vize PD, Warren WC, Wells D, Wills A, Wilson RK, Zimmerman LB, Zorn AM, Grainger R, Grammer T, Khokha MK, Richardson PM, Rokhsar DS. 2010. The genome of the Western clawed frog Xenopus tropicalis. Science 328: 633-636.

Höckner M, Erdel M, Spreiz A, Utermann G, Kotzot D. 2009. Whole genome amplification from microdissected chromosomes. Cytogenet Genome Res 125: 98-102.

Khokha MK, Krylov V, Reilly MJ, Gall JG, Bhattacharya D, Cheung CY, Kaufman S, Lam DK, Macha J, Ngo C, Prakash N, Schmidt P, Tlapakova T, Trivedi T, Tumova L, Abu-Daya A, Geach T, Vendrell E, Ironfield H, Sinzelle L, Sater AK, Wells DE, Harland RM, Zimmerman LB. 2009. Rapid gynogenetic mapping of Xenopus tropicalis mutations to chromosomes. Dev Dyn 238: 1398-1346.

Krylov V, Mácha J, Tlapáková T, Takác M, Jonák J. 2003. The c-SRC1 gene visualized by in situ hybridization on Xenopus laevis chromosomes. Cytogenet Genome Res 103: 169-172.

Krylov V, Tlapakova T, Macha J. 2007. Localization of the single copy gene Mdh2 on Xenopus tropicalis chromosomes by FISH-TSA. Cytogenet Genome Res 116: 110-112.

Kubickova S, Cernohorska H, Musilova P, Rubes J. 2002. The use of laser microdissection for the preparation of chromosome-specific painting probes in farm animals. Chromosome Res 10: 571-577.

Langmead B, Trapnell C, Pop M, Salzberg SL. 2009. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol 10: R25.

Ma L, Xiao Y, Huang H, Wang Q, Rao W, Feng Y, Zhang K, Song Q. 2010. Direct determination of molecular haplotypes by chromosome microdissection. Nat Methods 7: 299-301.

Mácha J, Teichmanová R, Sater AK, Wells DE, Tlapáková T, Zimmerman LB, Krylov V. 2012. Deep ancestry of mammalian X chromosome revealed by comparison with the basal tetrapod Xenopus tropicalis. BMC Genomics 13: 315.

Olmstead AW, Lindberg-Livingston A, Degitz SJ. 2010. Genotyping sex in the amphibian, Xenopus (Silurana) tropicalis, for endocrine disruptor bioassays. Aquat Toxicol 98: 60-66.

Schmid M, Steinlein C. 1991. Chromosome banding in Amphibia. XVI. High-resolution replication banding patterns in Xenopus laevis. Chromosoma 101: 123-132.

Schmid M, Steinlein C, Bogart JP, Feichtinger W, Haaf T, Nanda I, del Pino EM, Duellman WE, Hedges SB. 2012. The hemiphractid frogs. Phylogeny, embryology, life history, and cytogenetics. Cytogenet Genome Res 138: 69-384.

Schmid M, Steinlein C, Bogart JP, Feichtinger W, León P, La Marca E, Díaz LM, Sanz A, Chen SH, Hedges SB. 2010. The chromosomes of terraranan frogs. Insights into vertebrate cytogenetics. Cytogenet Genome Res 130-131: 1-568.
Tymowska J. 1973. Karyotype analysis of Xenopus tropicalis Gray, Pipidae. Cytogenet Cell Genet 12: 297-304.
Uno Y, Nishida C, Oshima Y, Yokoyama S, Miura I, Matsuda Y, Nakamura M. 2008a. Comparative chromosome mapping of sex-linked genes and identification of sex chromosomal rearrangements in the Japanese wrinkled frog (Rana rugosa, Ranidae) with ZW and XY sex chromosome systems. Chromosome Res 16: 637-647.
Uno Y, Nishida C, Yoshimoto S, Ito M, Oshima Y, Yokoyama S, Nakamura M, Matsuda Y. 2008b. Diversity in the origins of sex chromosomes in anurans inferred from comparative mapping of sexual differentiation genes for three species of the Raninae and Xenopodinae. Chromosome Res 16: 999-1011.
Weise A, Timmermann B, Grabherr M, Werber M, Heyn P, Kosyakova N, Liehr T, Neitzel H, Konrat K, Bommer C, Dietrich C, Rajab A, Reinhardt R, Mundlos S, Lindner TH, Hoffmann K. 2010. High-throughput sequencing of microdissected chromosomal regions. Eur J Hum Genet 18: 457-462.
Wells DE, Gutierrez L, Xu Z, Krylov V, Macha J, Blankenburg KP, Hitchens M, Bellot LJ, Spivey M, Stemple DL, Kowis A, Ye Y, Pasternak S, Owen J, Tran T, Slavikova R, Tumova L, Tlapakova T, Seifertova E, Scherer SE, Sater AK. 2011. A genetic map of Xenopus tropicalis. Dev Biol 354: 1-8.