

**UNIVERZITA PALACKÉHO V OLOMOUCI**

**Přírodovědecká fakulta**

**Katedra buněčné biologie a genetiky**



**Hodnocení výskytu populačních frekvencí spádové oblasti  
Morava ve vybraných STR lokusech – využití pro paternitní  
a identifikační účely**

**Diplomová práce**

**Bc. Kateřina Čížková**

Studijní program: Biologie

Studijní obor: Molekulární a buněčná biologie

Forma studia: Prezenční

**Olomouc 2010**

**Vedoucí práce: Mgr. Radek Vodička, Ph.D.**

Prohlašuji, že jsem tuto diplomovou práci vypracovala samostatně za odborného vedení svého školitele Mgr. Radka Vodičky, Ph.D. v průběhu navazujícího magisterského studia, a že uvádím veškerou použitou literaturu.

Dne:

Podpis.....

Děkuji svému školiteli Mgr. Radku Vodičkovi, Ph.D za odborné vedení, cenné rady a připomínky, ochotu konzultovat a trpělivost při vypracovávání této diplomové práce. Dále děkuji doc. RNDr. Radku Vrtělovi, Ph.D a celému personálu Ústavu lékařské genetiky a fetální medicíny Fakultní nemocnice Olomouc za vstřícné jednání a pomoc.

## SOUHRN

Krátké tandemové repetice (STR lokusy) se nachází u prokaryotních i eukaryotních organismů včetně člověka. Jsou tvořeny za sebou se opakujícími repeticemi DNA o délce 1 – 6 bp a tvoří úseky o celkové délce větší než 100 bp. Tyto lokusy jsou vysoce polymorfní, kodominantní a široce využívané genetické markery. Své uplatnění nachází při prenatální diagnostice, výzkumu rakoviny, genetickém mapování a populačních studiích. Velký význam mají STR lokusy při identifikaci osob a testování paternity.

Ve spádové oblasti Morava byl hodnocen výskyt alelových frekvencí ve 13 STR lokusech (D21S1414, D21S1435, D21S1446, D21S1411, penta D, D18S535, D18S51, D18S386, D13S631, D13S317, D13S305, D13S634 a XHPRT). Na základě zjištěných dat je nejpolymorfnějším zkoumaným STR lokusem D18S386. Naproti tomu nejméně polymorfní lokusy jsou D21S1446 a D13S305.

Pro stanovení pravděpodobnost otcovství a pravděpodobnost náhodné shody pro sledované muže byly místo markerů D18S535 a XHPRT využity markery D18S391 a D18S499. Získané údaje byly srovnány s výsledky zjištěnými pomocí komerčního kitu AmpF/STR Identifier. Statistickou analýzou dat nebyl mezi testovanou sadou markerů a komerčním kitem prokázán signifikantní rozdíl. U testovaných STR lokusů bylo ale celkem nalezeno více mutací než u komerčního kitu. Vyšší míra mutací může ovlivnit výsledek paternitního testu, ale nemá vliv na identifikaci osob.

## SUMMARY

Short tandem repeats (STRs) are found in prokaryotes and eucaryotes, including human. STRs are short tandemly repeated DNA sequences with 1 – 6 bp repeat unit forming series with lengths of up to 100 bp. These loci are highly polymorphic, codominant and widely used genetic markers. They are used for prenatal diagnosis, cancer research, genetic mapping and population studies. STR loci are very important to paternity testing and identification of individuals.

Allele frequencies of 13 STR loci (D21S1414, D21S1435, D21S1446, D21S1411, penta D, D18S535, D18S51, D18S386, D13S631, D13S317, D13S305, D13S634 and XHPRT) were evaluated in Moravian population. The most polymorphic STR locus is D18S386. On the other hand, the least polymorphic tested STR loci are D21S1446 and D13S305.

For paternity testing and estimating of random match probability of alleged father's genotype were used markers D18S391 and D18S499 instead of markers D18S535 and XHPRT. These data were compared with commercial AmpF/STR Identifier kit. There was found no statistically significant difference in set of tested markers and commercial kit. However, more mutations were detected in tested STR loci. Higher mutation rate could affect paternity analysis, but couldn't affect identification of individuals.

# OBSAH

1 ÚVOD .....	8
2 SOUČASNÝ STAV ŘEŠENÉ PROBLEMATIKY .....	9
2.1 Organizace lidského genomu .....	9
2.1.1 Polymorfismus .....	10
2.2 Krátké tandemové repetice .....	12
2.2.1 Mutace STR lokusů .....	12
2.2.2 Faktory ovlivňující mutace STR lokusů .....	13
2.2.3 Mechanismy mutací STR lokusů .....	14
2.3 Využití STR lokusů pro testování paternity a identifikaci osob .....	17
2.3.1 Testování paternity .....	17
2.3.1.1 Pravděpodobnost vyloučení z otcovství .....	17
2.3.1.2 Paternitní index .....	18
2.3.1.2.1 Výpočet paternitního indexu v případě, že není k dispozici genom jednoho z rodičů .....	20
2.3.1.2.1.1 Genom matky není k dispozici .....	20
2.3.1.2.1.2 Genom otce není k dispozici .....	21
2.3.1.3 Pravděpodobnost otcovství .....	21
2.3.1.4 Vliv mutací .....	22
2.3.1.5 Neautozomální markery .....	22
2.3.1.5.1 Chromozóm Y .....	22
2.3.1.5.2 Chromozóm X .....	23
2.3.1.5.3 Mitochondriální DNA .....	23
2.3.2 Identifikace osob .....	23
2.3.2.1 Frekvence genotypů a pravděpodobnost náhodné shody .....	24
2.4 Další aplikace .....	25
2.4.1 Prenatální diagnostika .....	25
2.4.2 Výzkum rakoviny .....	25
2.4.3 Genetické mapování a vazebná analýza .....	26
2.4.5 Studium evoluce a migrace .....	26
3 CÍL PRÁCE .....	28
4 MATERIÁL A METODY .....	29

4.1 Určení populačních frekvencí .....	29
4.2 Analýza paternity a pravděpodobnosti náhodné shody .....	31
5 VÝSLEDKY .....	33
5.1 Frekvence sledovaných STR lokusů v populaci .....	33
5.2 Analýza paternity a určení pravděpodobnosti náhodné shody .....	38
6 DISKUZE .....	62
7 ZÁVĚR .....	65
8 LITERATURA .....	66
9 SEZNAM POUŽITÝCH ZKRATEK A SYMBOLŮ .....	70
10 PŘÍLOHY .....	72

# 1 ÚVOD

Velký rozvoj molekulární biologie nastává s objevem polymerázové řetězové reakce (PCR). Tuto metodu navrhl v roce 1983 Kary Mullis. PCR zvýšila citlivost DNA analýzy, analýza se stala jednodušší a rychlejší. PCR umožňuje také analýzu degradovaných vzorků a skýtá potenciální možnost automatizace.

DNA analýza je velmi užitečným nástrojem pro forenzní účely, jako je identifikace osob při masových katastrofách, poskytuje velmi cenný důkaz při řešení kriminálních případů a v neposlední řadě při určování příbuzenských vztahů.

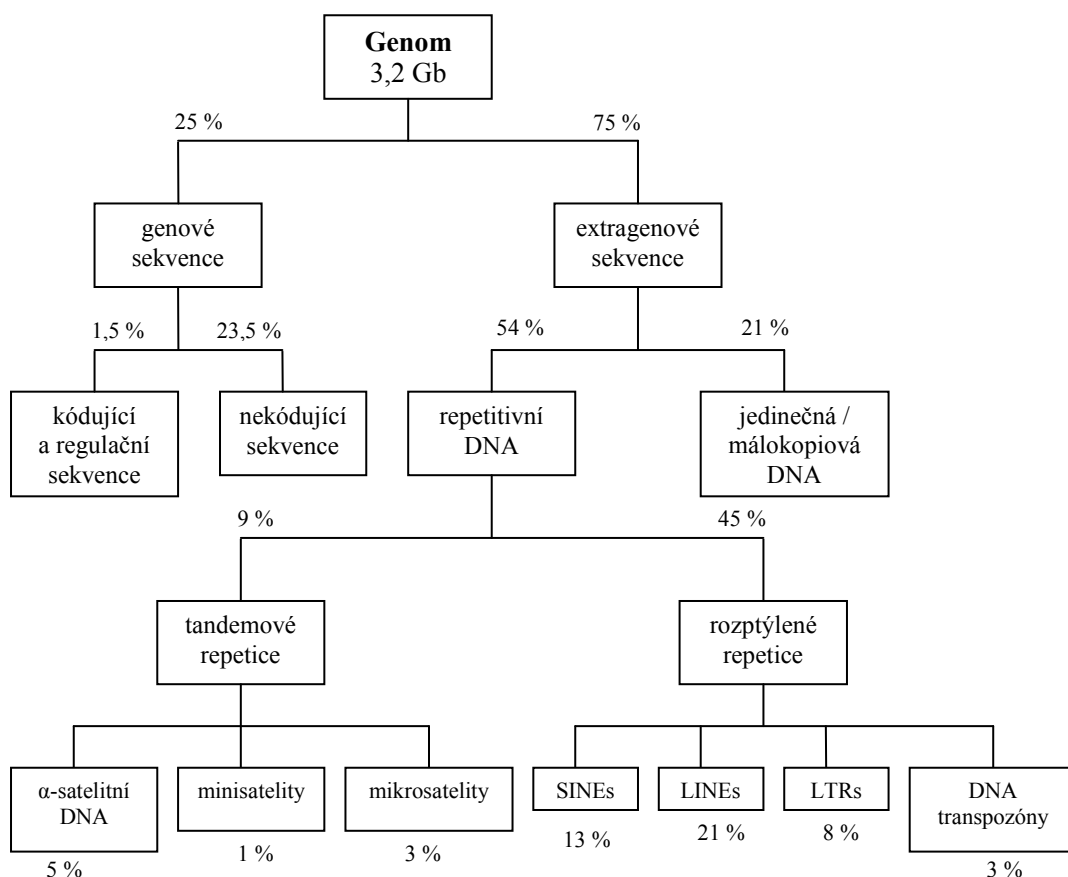
Od začátku 90. let 20. století se ve forenzní praxi využívá polymorfizmu krátkých tandemových repetitivních sekvencí. V současné době velká většina DNA analýz pro forenzní účely zahrnuje právě analýzu polymorfizmu STR lokusů (Butler, 2005; Goodwin *et al.*, 2007).



## 2 SOUČASNÝ STAV ŘEŠENÉ PROBLEMATIKY

### 2.1 Organizace lidského genomu

Velikost lidského genomu je přibližně 3,2 Gb (International Human Genome Sequencing Consortium, 2001). Genom se skládá z jedinečné a repetitivní DNA (Obr. 1). Jedinečná DNA je v genomu přítomna v jedné nebo pouze několika málo kopiích a zahrnuje geny a regulační sekvence. Lidský genom obsahuje 20 000 – 25 000 genů. Sekvence kódující přímo proteiny zahrnují asi 1,5 % genomu. Dále jedinečnou DNA tvoří genové sekvence, které sice proteiny nekódují, ale účastní se regulace (promotory, enhancery, represory, polyadenylační signál), introny, pseudogeny a genové fragmenty (Goodwin *et al.*, 2007).



Obr. 1: Schéma organizace lidského genomu (upraveno podle Goodwin *et al.*, 2007)

Přibližně 75 % celého genomu tvoří extragenová DNA. Asi 50% genomu tvoří různé typy repetitivních sekvencí. Nejčastějšími typy rozptýlených repetice jsou dlouhé roztroušené elementy (long interspred elements, LINEs) s délkou repetice nad 500 bp,

krátké roztroušené elementy (short interspred elements, SINEs) s repeticí dlouhou do 500 bp, LTR sekvence (long terminal repeats) a DNA transpozóny. Tyto sekvence jsou odvozovány transpozicí. (Goodwin *et al.*, 2007; International Human Genome Sequencing Consortium, 2001).

Tandemově uspořádané repetice tvoří 10 % genomu. Podle délky repetice se tyto sekvence dělí na  $\alpha$ -satelitní DNA, která tvoří centromerické oblasti chromozómů s repeticemi dlouhými přibližně 171 bp, minisatelitní DNA s délkou repetice 6 – 100 bp a mikrosatelitní DNA, u které je délka repetice 1 – 6 bp (Goodwin *et al.*, 2007; Näslund *et al.*, 2005).

### 2.1.1 Polymorfismus

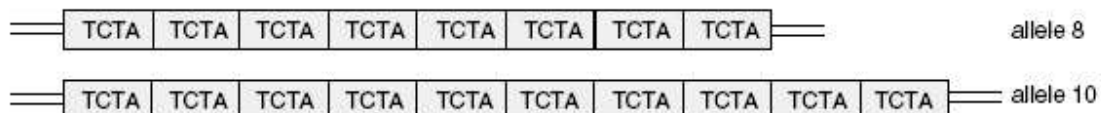
Mezi dvěma libovolnými lidmi je sekvence jaderné DNA v 99,9% totožná. Za genetickou variabilitu je tedy zodpovědný zbyvajících zlomek procenta. Odchytky v DNA se mohou vyskytovat v kódujících sekvencích a tím způsobovat tvorbu odlišných variant proteinů. Jiné varianty se nachází v extragenových úsecích či v intronech genů, neovlivňují funkci žádného genu a mohou být detekovány pouze přímou analýzou DNA (Nussbaum *et al.*, 2004).

Změny v DNA se nazývají mutace. Mutace jsou hlavním zdrojem genetické variability. Mutace představují základ evoluce, ale také způsobují genetické choroby. Každý jedinec nese ve svém genomu 100 – 200 nových mutací (Nachman *et al.*, 2000), z nichž většina leží mimo geny. Některé mutace jsou škodlivé, jiné jsou považovány za selekčně neutrální. Neutrální mutace se vyskytují mimo geny nebo v nekódujících oblastech chromozómů. V nekódujících oblastech genomu dochází k mutacím častěji, než v oblastech kódujících.

Různé varianty určité sekvence DNA na daném lokusu se nazývají alely. Vyskytuje-li se některá alela v populaci ve více než 1% případů, vzniká genetický polymorfismus. Alely s frekvencí nižší než 1% se označují jako vzácné varianty (Nussbaum *et al.*, 2004).

Polymorfismus může být sekvenční nebo délkový. K sekvenčnímu polymorfismu náleží jednodukleotidové polymorfizmy (single nucleotide polymorphism, SNP). Tento druh polymorfizmu se vyskytuje v genomu nejčastěji. Jeden SNP připadá na každých 300 – 1000 bp v genomu. Různé alely obsahují v určitém místě sekvence jiný nukleotid, např. jedna alela obsahuje guanin a druhá adenin. SNPs vznikají chybami při replikaci DNA během gametogeneze. SNPs nejsou vysoce polymorfní, ale bialelické (Hanchard, 2005).

Tandemové repetitivní sekvence v genomu tvoří délkový polymorfismus. Různé alely těchto lokusů jsou dány různým počtem opakování repetice, a tedy různou délkou sekvence. Pokud je délka repetice 1 – 6 bp, jedná se mikrosatelity, nazývané též krátké tandemové repetice (short tandem repeats, STR) (Obr. 2). Mikrosatelity tvoří alely v rozsahu 50 – 300 bp. Když je repetice dlouhá 6 – 100 bp, jedná se o minisatelity. Tyto sekvence tvoří alely v rozsahu 500 bp až 30 kb. Minisatelitové sekvence mohou být monomorfní nebo polymorfní. Polymorfní minisatelity jsou někdy označovány jako variabilní počet tandemových repetic (variable number of tandem repeats, VNTR). Předpokládaná hustota výskytu VNTRs v genomu je 9,1 VNTRs/Mb (Goodwin *et al.*, 2007; Näslund *et al.*, 2005; Vergnaud *et Denoeud*, 2000). Minisatelitové a mikrosatelitové lokusy jsou polymorfnější než SNPs.



Obr. 2: Struktura STR. Alely se od sebe liší počtem opakování repetitivní jednotky (Goodwin *et al.*, 2007).

## 2.2 Krátké tandemové repetice

Krátké tandemové repetice (short tandem repeats, STR), nazývané též mikrosatelity, se nachází u prokaryotních i eukaryotních organismů včetně člověka. V lidském genomu zaujímají asi 3 % jeho velikosti. Jsou tvořeny za sebou se opakujícími repeticemi DNA o délce 1 – 6 bp a tvoří úseky o celkové délce větší než 100 bp (Fan *et* Chu, 2007). Mnoho STRs se nachází v nekódujících oblastech genomu, v kódujících oblastech se jich nachází jen asi 8 % (Ellegren, 2000). V průměru se vyskytuje jeden STR na 2 kb (International Human Genome Sequencing Consortium, 2001). Nejvíce těchto sekvencí se nalézá na chromozómu 19 (Subramanian *et al.*, 2003).

STRs mohou být klasifikovány na různé typy. Podle délky repetitivní jednotky se STRs dělí na mono-, di-, tri-, tetra-, penta- a hexanukleotidové repetice (Fan *et* Chu, 2007), přičemž nejčastější jsou repetice dinukleotidové a nejméně je hexanukleotidových (International Human Genome Sequencing Consortium, 2001). Podle struktury repetice lze STR lokusy rozdělit na perfektní repetice (jednoduché repetice, simple repeats), které obsahují pouze jednu repetitivní jednotku a imperfektní repetice (složené repetice, compound repeats) složené z různých typů repetice (Fan *et* Chu, 2007).

Některé trinukleotidové repetice způsobují u člověka vznik hlavně neurologických chorob. Jde o vzácná onemocnění s dominantní dědičností. Příklady takových onemocnění jsou syndrom fragilního X, Huntingtonova choroba, myotonické dystrofie a některé typy spinocerebrálních ataxií. U těchto chorob dochází k abnormální expanzi trinukleotidových repetice. Tyto repetice asociované s chorobami byly identifikovány pouze u člověka. Toto vede k předpokladu, že přítomnost trinukleotidových repetice uvnitř některých genů spojených s mozkem může být jednou z příčin evoluce funkce mozku (Fan *et* Chu, 2007).

### 2.2.1 Mutace STR lokusů

Zatímco v jedinečných sekvencích DNA dochází k mutacím zřídka, a to s frekvencí asi  $10^{-9}$  na generaci. U STR lokusů je frekvence mutací o několik řádů vyšší, v rozmezí od  $10^{-6}$  po  $10^{-2}$  na generaci (Fan *et* Chu, 2007; Schlötterer, 2000). U nepatogenních STR lokusů složených z dinukleotidových repetice dochází v porovnání s lokusy tetranukleotidovými k mutacím 1,5 – 2x častěji. Frekvence mutací u trinukleotidových repetice leží mezi di- a tetranukleotidovými, zatímco trinukleotidové lokusy asociované

s nemocemi vykazují 4 – 7x vyšší frekvenci mutací než nepatogenní lokusy dinukleotidové (Chakraborty *et al.*, 1997).

Mutace v STR lokusech jsou nejčastěji popisovány pomocí krokového mutačního modelu (stepwise mutation model, SMM). Podle tohoto modelu se mikrosatelit mění pouze o jednu či malý počet jednotek. Produktem mutace je často již existující alela. Mutace v STR lokusech tedy popírají hlavní předpoklad modelu nekonečného množství alel (infinite alleles model, IAM), podle kterého každou mutací vzniká alela nová.

Předpoklady SMM jsou malé změny v počtu repetitivních jednotek, stejná pravděpodobnost zvýšení a snížení počtu jednotek, nelimitovaná velikost alel a nezávislost míry a rozsahu mutací na počtu jednotek. Tento model je v souladu s mechanismem vzniku STR mutací sklouznutím DNA polymerázy (Fan *et Chu*, 2007).

SMM má ale dva nedostatky: nepokrývá stabilní distribuci alel a nedokáže vysvětlit absenci velmi dlouhých alel v populaci. Ačkoliv je většina mutací v STR lokusech v souladu s tímto modelem, dochází k odchýlkám (Huang *et al.*, 2002; Kruglyak *et al.*, 1998). Pokud dojde v mikrosatelitu ke změně o více jednotek naráz, jde o vícekrokový mutační model (multistep mutation model, MMM), který zahrnuje mutační procesy jako SMM, ale i mutace většího rozsahu. Frekvence vícekrokových mutací se zvyšuje se zvyšující se variancí v počtu repetitivních jednotek. Pokud distribuce alel odpovídá tomuto modelu, mutace v STR lokusech může být způsobena mechanismem nerovnoměrného crossing overu. K vícekrokovým mutacím dochází u člověka až v 14 % případů (Ellergen, 2000; Kayser *et al.*, 2000).

### **2.2.2 Faktory ovlivňující mutace STR lokusů**

Mezi faktory ovlivňující mutace STR lokusů patří počet repetitivních jednotek, sekvenční motiv repetic, délka repetitivní jednotky, přiléhající sekvence, přerušení mikrosatelitu, míra rekombinace a transkripce, atd. (Schlötterer, 2000).

Jedním z klíčových faktorů ovlivňujících mutace v STR lokusech je počet repetitivních jednotek. Se zvyšujícím se počtem repetic se zvyšuje míra mutací. U kratších mikrosatelitů dochází s vyšší frekvencí k prodlužování alel, zatímco delší mikrosatelity mají tendenci se spíše zkracovat, než prodlužovat. Pokud dojde k mutaci, dlouhé mikrosatelity mutují spíše na kratší a krátké mikrosatelity mutují na delší (Huang *et al.*, 2002; Lai *et Sun*, 2003).

Míru mutací STR lokusů ovlivňuje také délka repetitivní jednotky. Nejvíce mutací nastává u dinukleotidových repetitivních jednotek, méně u trinukleotidových a ještě méně u tetranukleotidových jednotek (Chakraborty *et al.*, 1997).

Dalším faktorem je sekvenční motiv repetice. Sekvence jednotek s vysokým obsahem AT mutují rychleji než ty s vysokým obsahem GC. Sekvence s vysokým obsahem GC mohou redukovat míru sklouznutí polymerázy (Chakraborty *et al.*, 1997).

Míra STR mutací je ovlivněna i pohlavím. V mnoha lokusech dochází v germinálních buňkách k více mutacím u mužů než u žen, což může být zapříčiněno tím, že při spermatogenezi dochází k vyššímu počtu mitóz než při oogenezi (Ellegren, 2000). Vliv na míru mutací má i věk mužů.

Kromě sklouznutí polymerázy dochází v STR lokusech i k dalším mutacím, jako jsou tranzice, transverze, jednonukleotidové inserce a delece a další. Frekvence těchto změn je v porovnání se sklouznutím polymerázy nízká, ale mohou narušit sekvenci nukleotidů v mikrosatelitu a tím změnit jejich mutabilitu. (Fan *et Chu*, 2007).

Míru mutací v STR lokuse může ovlivňovat i rekombinace. Některé studie sice neobjevily žádnou asociaci mezi rekombinací a mírou mutací v mikrosatelitech (Huang *et al.*, 2002), jiné ano. Neexistuje ale důkaz pro silnou pozitivní korelaci mezi mírou rekombinace a mutacemi STR lokusů (Payseur *et Nachman*, 2000). STR lokusy nerekombinující oblasti chromozómu Y vykazují stejnou míru mutací jako autozomální lokusy, což naznačuje, že rekombinace není hlavním mechanismem vzniku variability STR lokusů.

Vliv sousedních sekvencí na mutace mikrosatelitu není zcela jasná (Fan *et Chu*, 2007; Schlötterer, 2000).

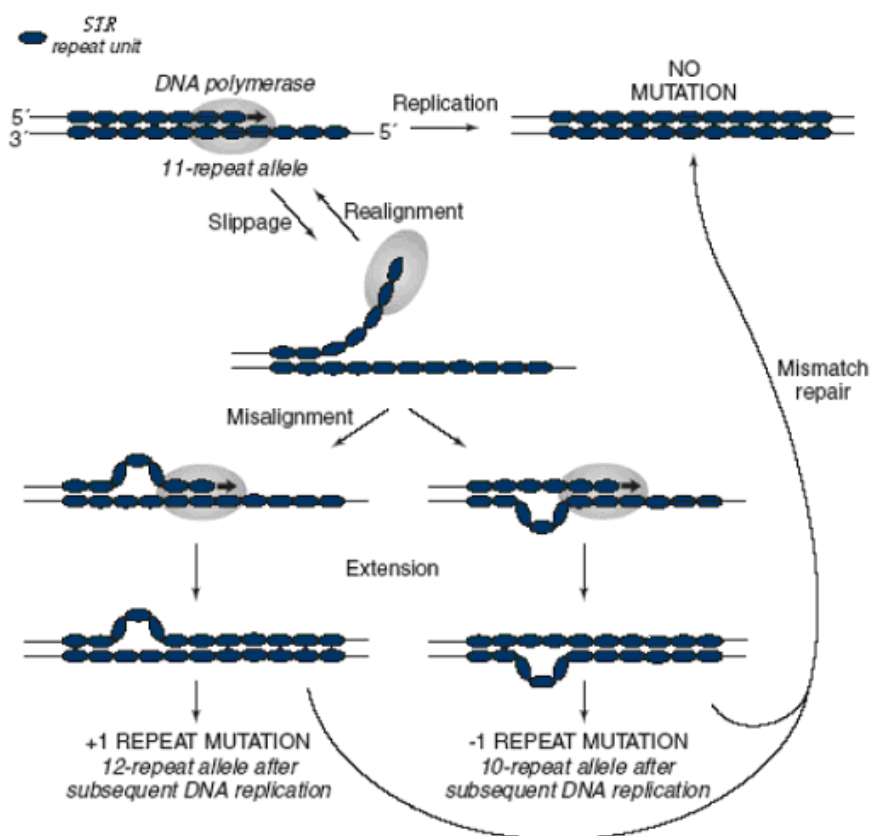
### **2.2.3 Mechanismy mutací STR lokusů**

Existují tři možné mechanismy vzniku STR mutací: nerovnoměrný crossing over v průběhu meiózy, retrotranspoziční mechanismus a sklouznutí DNA polymerázy při replikaci, což je pravděpodobně nejrozšířenější mechanismus.

Nerovnoměrný crossing over dává vznik velkým blokům satelitní DNA a je spojený s výměnou repetitivních jednotek mezi homologními chromozómy. Tento proces sice nastává mezi různými chromozómy, ale také hraje omezenou roli při mutacích STRs. Tento mechanismus může být zodpovědný za vícekrokovou mutaci STR lokusů (Huang *et al.*, 2002).

Mechanismus retrotranspozice uvažuje, že A-bohaté STR vznikají prodloužením 3' konce retrotranskriptu, podobně jako je tomu u polyadenylace mRNA. Existuje totiž asociace mezi STRs bohatými na adenin a transponovatelnými elementy, ačkoliv vysoká hustota transponovatelných elementů se nemusí vždy shodovat s vysokou hustotou výskytu STRs (Nadir *et al.*, 1996).

Nejvíce uznávaným vysvětlením mutací v STR lokusech je sklouznutí DNA polymerázy při replikaci. V důsledku sklouznutí polymerázy dochází k chybnému párování mezi templátovým a nově vznikajícím vláknem DNA o jednu či více jednotek. V místě chybného párování dochází k vytvoření smyčky (Obr. 3). Molekula DNA pak obsahuje jiný počet repetitivních jednotek. Míra sklouznutí polymerázy neodpovídá míře mutací STR lokusů. *In vivo* je totiž mnoho smyček na DNA vznikajících při posunu polymerázy rozpoznáno systémem opravujícím chybné párování. Pozorované množství STR mutací závisí na frekvenci sklouznutí polymerázy a účinnosti systému opravujícího chybné párování. Tento opravný systém redukuje míru STR mutací 100 – 1000x.



Obr. 3: Schéma sklouznutí DNA polymerázy při replikaci (Fan *et Chu*, 2007).

Na míru výskytu sklouznutí polymerázy má vliv několik faktorů, z nichž nejdůležitější je délka repetitivní jednotky. Ke sklouznutí polymerázy dochází častěji u dinukleotidových jednotek než jednotek tetranukleotidových (Kruglyak *et al.*, 1998). Mezi další faktory ovlivňující sklouznutí polymerázy patří počet, lokalizace a sekvence jednotek (Schlötterer, 1998). U člověka roste míra sklouznutí polymerázy exponenciálně se zvyšujícím se počtem repetitivních jednotek (Lai *et Sun*, 2003). Minimální počet jednotek nutných pro další expanzi je čtyři až pět u dinukleotidových STR a dvě repetice u tetranukleotidových STR.

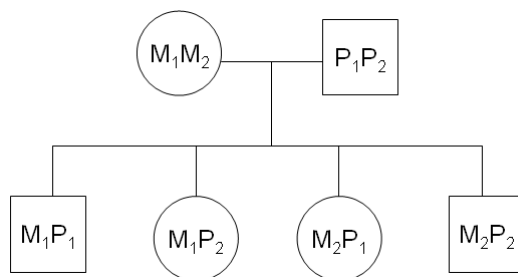


## 2.3 Využití STR lokusů pro testování paternity a identifikaci osob

### 2.3.1 Testování paternity

Testování paternity je založeno na faktu, že každý genetický marker dítěte je zděděn od jednoho z rodičů. Konečný výsledek testu paternity je stanoven porovnáním genetických markerů matky, dítěte a muže označeného za otce. Paternitní test je možný i za předpokladu, že se neanalyzuje vzorek matky, ale míra věrohodnosti výsledku je nižší.

Genetické markery využívané pro testování paternity musí vykazovat mendelistickou dědičnost. STR lokusy jsou děděny tímto způsobem. Uvažujeme-li jeden lokus a paternální alely tohoto lokusu jsou  $P_1$ ,  $P_2$  a maternální alely  $M_1$ ,  $M_2$ , potom u jejich potomků mohou vzniknout čtyři možné kombinace alel, a to  $P_1M_1$ ,  $P_1M_2$ ,  $P_2M_1$ ,  $P_2M_2$  (Obr. 4), každá s pravděpodobností 0,25. Alely markeru užívaných pro paternitní účely musí být distribuovány podle Hardy-Weinbergova zákona. Další podmínkou je, že odlišné genetické systémy jsou děděny nezávisle na sobě, tzn. jsou lokalizovány na různých chromozómech. Pokud totiž dva lokusy leží u sebe dostatečně blízko, jsou děděny společně. Frekvence mutací u těchto makreků je malá (Buckleton *et al.*, 2005; Mertens *et al.*, 1999).



Obr. 4: Příklad rodokmenu zobrazující všechny možné kombinace genotypů potomků (upraveno podle Butler, 2005).

Pro hodnocení výsledku paternitního testu se využívají tři metody, a to pravděpodobnost vyloučení z otcovství, paternitní index (PI) a pravděpodobnost otcovství. Všechny tři metody využívají znalosti frekvence alel v populaci.

#### 2.3.1.1 Pravděpodobnost vyloučení z otcovství

Genotypy matky a dítěte definují v každém lokusu paternitní alelu, protože každý marker dítěte nezděděný od matky musí být přítomen v genotypu otce.

Pravděpodobnost vyloučení z otcovství (PE; power / probability of exclusion) je pravděpodobnost vyloučení náhodného muže z populace z otcovství na základě genotypizačního výsledku matky a dítěte (Fung *et al.*, 2002). Genotyp muže označeného za otce se nebere v úvahu. Tato pravděpodobnost může být definována jako podíl mužů, kteří nenesou paternální alelu. Když jsou možné paternální alely na lokusu  $A_1, A_2, \dots, A_n$ , potom je pravděpodobnost vyloučení z otcovství pro tento lokus  $PE_1$  za předpokladu Hardy-Weinbergovi rovnováhy:

$$PE_1 = \left( 1 - \sum_{i=1}^n P(A_i) \right)^2$$

Kombinovaná pravděpodobnost vyloučení z otcovství vypočítá jako:

$$PE = 1 - \prod_1 (1 - PE_1)$$

(Buckleton *et al.*, 2005).

Průměrná pravděpodobnost vyloučení muže z otcovství ( $\overline{PE}$ ) je pravděpodobnost, s jakou dojde k vyloučení falešně nařčeného muže při použití daného genetického systému. Tato pravděpodobnost závisí na polymorfizmu, např. na počtu alel v systému a jejich distribuci v populaci (Mertens *et al.*, 1999). Průměrná pravděpodobnost vyloučení muže z otcovství pro daný lokus se vypočítá jako:

$$\overline{PE} = h^2(1 - 2hH^2)$$

kde  $h$  je heterozygotita a  $H$  homozygotita daného lokusu (Goodwin *et al.*, 2007). Pokud jsou genetické systémy nezávislé, kombinovaná průměrná pravděpodobnost vyloučení z otcovství se vypočítá podobně jako kombinovaná pravděpodobnost vyloučení z otcovství.

### 2.3.1.2 Paternitní index

Paternitní index (PI) je proměr pravděpodobností, který se skládá ze dvou konkurenčních a vzájemně se vylučujících hypotéz:

$H_0$ : označený muž je skutečným otcem dítěte

$H_1$ : označený muž není skutečným otcem dítěte.

Z předpokladu existence nějakého důkazu, typicky genotypů dítěte ( $G_C$ ), otce ( $G_{AF}$ ) a matky ( $G_M$ ), je na základě Bayesova teorému aposteriorní pravděpodobnost rovna součinu apriorní pravděpodobnosti a pravděpodobnostního poměru:

$$\frac{P(H_0 | G_C, G_M, G_{AF})}{P(H_1 | G_C, G_M, G_{AF})} = \frac{P(G_C, G_M, G_{AF} | H_0)}{P(G_C, G_M, G_{AF} | H_1)} \cdot \frac{P(H_0)}{P(H_1)}$$

Pravděpodobnostní poměr:

$$\frac{P(G_C, G_M, G_{AF} | H_0)}{P(G_C, G_M, G_{AF} | H_1)}$$

se označuje jako paternitní index (PI).

Čitatel a jmenovatel zlomku jsou podmíněny genotypem matky, dítěte a muže označeného za otce a mohou být odvozeny z kombinačního čtverce (Goodwin *et al.*, 2007).

Předpokládá se, že matka je jistá, čili je často možné určit maternální a paternální alelu (Buckleton *et al.*, 2005). Předpokládejme případ, kdy matka má genotyp  $pr$ , dítě  $pq$  a muž označený za otce  $qs$ . Pokud je tento muž otcem, musela matka dítěti předat alelu  $p$  a otec alelu  $q$ . Pokud muž otcem není, potom matka musela dítěti předat alelu  $p$  a alelu  $q$  muselo dítě zdědit od jiného muže z populace, který je nositelem alely  $q$ . Pokud je označený muž otcem, pravděpodobnost, že dítě bude nositelem genotypu  $pq$ , je 0,25. Pokud označený muž není otcem, matka předala alelu  $p$  s pravděpodobností 0,5 a šance, že jiný muž je otcem dítěte závisí na frekvenci alely  $q$  v populaci ( $q$ ). PI je potom tedy:

$$PI = \frac{0,25}{0,5q} = \frac{1}{2q}$$

Stejným způsobem lze odvodit PI pro ostatní možné kombinace (Tab. I) (Goodwin *et al.*, 2007).

PI se spočítá pro jednotlivé lokusy. Vynásobením jednotlivých PI mezi sebou získáme kombinovaný paternitní index (CPI) (Gjertson *et al.*, 2007). Obecně přijímaná minimální hodnota CPI pro označení muže za otce je rovno 100. PI 100 koreluje s pravděpodobností, že označený muž má 99 ku 1 větší šanci, že je otcem dítěte než náhodný muž z populace (Butler, 2005).

Tab. I: PI pro všechny nevyložené kombinace maternálního a paternálního genotypu (Buckleton *et al.*, 2005)

Genotyp matky	Genotyp dítěte	Genotyp muže	PI
qq	qq	qq	$\frac{1}{q}$
pq			
pp			
pr	pq		
qr			
qq	qq	pq	$\frac{1}{2q}$
pq			
qr			
pp	pq		
pr			
pr	qr		
rr			
rs	pq	qq	$\frac{1}{p+q}$
		pq	$\frac{1}{2(p+q)}$
		qr	

PI – paternitní index; p, q, r, s – jednotlivé alely na lokusu; p, q – frekvence alel p, q v populaci

### 2.3.1.2.1 Výpočet paternitního indexu v případě, že není k dispozici genom jednoho z rodičů

Hodnocení paternity se obvykle provádí na základě genotypu matky, dítěte a muže označeného za otce. V některých případech je ale třeba vyřešit spornou paternitu u dítěte i v případě, že genotyp matky nebo otce není k dispozici (Lee *et al.*, 2000; Lee *et al.*, 2001).

#### 2.3.1.2.1.1 Genom matky není k dispozici

Pokud není genotyp matky k dispozici, nelze jasně určit paternální alelu a pro určení čitatele a jmenovatele paternitního indexu nelze použít kombinační čtverec. Z otcovství je vyloučen muž, který ve svém genotypu neobsahuje ani jednu ze dvou alel, které obsahuje genotyp dítěte. (Lee *et al.*, 2000).

Předpokládáme případ, kdy genotyp dítěte je pq a genotyp muže označeného za otce je qr. Pokud je označený muž otcem, potom je paternální alela q předána s pravděpodobností 0,5. Maternální alela musí být v tomto případě p a pravděpodobnost jejího předání závisí na její frekvenci v populaci. Pokud označený muž otcem není, je

pravděpodobnost výskytu genotypu pq dítěte dána podle Hardy-Weinbergova zákona  $2pq$ . Paternitní index je potom (Buckleton *et al.*, 2005; Lee *et al.*, 2000):

$$PI = \frac{0,5p}{2pq} = \frac{1}{4q}$$

Stejným způsobem lze odvodit PI pro ostatní možné kombinace (Tab. II)

Tab. II: PI pro všechny nevyložené kombinace paternálního genotypu a genotypu dítěte (Buckleton *et al.*, 2005)

Genotyp dítěte	Genotyp muže	PI
qq	qq	$\frac{1}{q}$
qq	qp	$\frac{1}{2q}$
qp	qq	
qp	qp	$\frac{p+q}{4pq}$
qp	qr	$\frac{1}{4p}$

PI – paternitní index; p, q, r – jednotlivé alely na lokusu; p, q – frekvence alel p, q v populaci

#### 2.3.1.2.1.2 Genom otce není k dispozici

V některých případech není k dispozici genotyp otce, často pokud muž označený jako otec již zemřel. V tomto případě lze využít genetickou informaci příbuzných osob. Využita může být genetická informace rodičů tohoto muže (Serra *et al.*, 2008). Lze využít i genotypy jistých dětí tohoto muže a jejich matky (Lee *et al.*, 2001).

#### 2.3.1.3 Pravděpodobnost otcovství

Na základě Bayesova teorému může být pravděpodobnost otcovství spočítána z paternitního indexu a apriorní pravděpodobnosti. Apriorní pravděpodobnost souvisí s negenetickými důkazy (E), např. svědectvím matky (Goodwin *et al.*, 2007). Pravděpodobnost otcovství (W) je tedy rovna:

$$W = \frac{PI \cdot P(H_0|E)}{PI \cdot P(H_0|E) + [1 - P(H_p|E)]}$$

V praxi se při výpočtu pravděpodobnosti otcovství apriorní pravděpodobnost pokládá rovna 0,5 pro obě hypotézy  $H_0$  i  $H_1$ . S využitím tohoto předpokladu je pravděpodobnost otcovství rovna:

$$W = \frac{PI}{1 + PI}$$

(Goodwin *et al.*, 2007; Buckleton *et al.*, 2005; Mertens *et al.*, 1999).

#### **2.3.1.4 Vliv mutací**

Při určování otcovství musí být brány v potaz také mutace. Pokud se u dítěte a údajného otce neshodují dva STR lokusy, nemůže být tento muž ještě z otcovství vyloučen. Testováním více genetických systémů se zvyšuje šance, že bude nalezena mutace. Při STR analýze se vyšetřuje deset nebo více lokusů a nález dvou nesouhlasných lokusů mezi dítětem a biologickým otcem není neobvyklý (Butler, 2005; Mertens *et al.*, 1999).

Pokud dojde k mutaci v místě nasedání primeru, může v závislosti na tom, zda mutace nastala na 5' nebo 3' konci sekvence, vzniknout nulová alela (Buckleton *et al.*, 2005, Gjertson *et al.*, 2007).

#### **2.3.1.5 Neautozomální markery**

Ve sporných případech paternity mohou být velmi užitečné neautozomální markery (Serra *et al.*, 2008; Ayers *et al.*, 2005; Presciuttini *et al.*, 2004; Kayser *et al.*, 2000).

##### **2.3.1.5.1 Chromozóm Y**

STR lokusy na Y chromozómu jsou předávány z otce na syny. Většina chromozómu Y nepodléhá rekombinaci a proto jsou schopné charakterizovat mužskou linii rodu. Pokud dva muži nesou stejné STR lokusy na chromozómu Y, je pravděpodobnost, že nejsou příbuzní velmi malá (Kayser *et al.*, 2001; Presciuttini *et al.*, 2004).

Výsledky pro Y chromozómové markery by měly být řešeny jako haplotypy a pro výpočet pravděpodobnosti by měla být použita frekvence haplotypu (Gjertson *et al.*, 2007).

### 2.3.1.5.2 Chromozóm X

Muži jsou hemizygoti pro lokusy na chromozómu X, proto je pravděpodobnost, že náhodně vybraný muž z populace sdílí s dítětem paternální alelu nižší než případě autozomálních markerů (Ayes *et Powley*, 2005).

Využití STR lokusů na chromozómu X má velký význam v případě, že při testování paternity u potomka ženského pohlaví není k dispozici genotyp muže označeného za otce. Genetická informace matky označeného muže přináší informaci o haplotypu na chromozómu X (jež zdědil od matky), který je předáván dítěti (Serra *et al.*, 2007).

### 2.3.1.5.3 Mitochondriální DNA

Mitochondriální DNA je přenášena v bloku z generace na generaci po maternální linii. Tato DNA zůstává, kromě případů mutací, stejná z generace na generaci (Gjertson *et al.*, 2007). Testování mitochondriálních markerů se využívají v případě, že není jistá matka dítěte (Buckleton *et al.*, 2005).

## 2.3.2 Identifikace osob

STR lokusy jsou široce využívány pro identifikaci osob, hlavně v kriminálních případech a při masových katastrofách. V 90. letech 20. stol. byl v USA laboratořemi FBI vytvořen systém CODIS (the combined DNA index system), který je hojně využíván pro identifikační a paternitní účely. Tento systém testuje 13 autozomálních STR lokusů (CSF1PO, FGA, TH01, TPOX, VWA, D3S1358, D5S818, D7S820, D8S1179, D13S317, D16S539, D18S51, D21S11) a amelogenin pro určení pohlaví jedince (Butler 2006; Tamaki *et Jeffreys*, 2005). Tyto lokusy leží na různých chromozómech, s výjimkou lokusů D5S818 a CSF1PO, které oba leží na chromozómu 5 (Holt *et al.*, 2000). Tyto lokusy mají velmi nízké pravděpodobnosti náhodné shody (Tamaki *et Jeffreys*, 2005).

Ve Velké Británii a většině evropských zemí se využívá systém 10 STR lokusů, který obsahuje dva přidatné markery D2S1338 a D19S433 a osm lokusů shodných s CODIS (FGA, TH01, VWA, D3S1358, D8S1179, D16S539, D18S51, D21S11) (Butler, 2006).

Frekvence jakéhokoliv DNA profilu může být spočítána na základě znalostí alel v DNA profilu a alelových frekvencí, které jsou k dispozici v databázích (Butler, 2005). Databáze obsahující frekvence alel se tvoří na základě zjišťování výskytu alel v definované populaci. Pokud je databáze určena pro statistické hodnocení síly DNA důkazu, je pro její

konstrukci doporučováno hodnotit nejméně 200 alel (nebo 100 osob) pro každý lokus (Goodwin *et al.*, 2007).

Jední z nejsouhrnější a široce využívaným internetovým zdrojem informací o hlavních STR lokusech pro identifikaci osob je Short Tandem Repeat DNA Internet Database, neboli STRBase (<http://www.cstl.nist.gov/strbase/>), kterou spravuje The National Institute of Standards and Technology (NIST). Databáze obsahuje popis pozorovaných alel a anotovaných sekvencí pro každý STR lokus, popis dostupných komerčních kitů, sekvence primerů, literaturu a odkazy na vědce a společnosti pracující v dané oblasti (Butler, 2006; Ruitberg *et al.*, 2001).

### 2.3.2.1 Frekvence genotypů a pravděpodobnost náhodné shody

Očekávané frekvence genotypů se počítají na základě Hardy-Weinbergova zákona. Ten říká, že uvnitř populace s náhodným oplozením zůstávají frekvence genotypů pro každý lokus konstantní. Pokud je populace v Hardy-Weinbergově rovnováze, frekvence genotypů mohou být určeny z alelových frekvencí. Frekvence homozygotního genotypu může být spočítána jako  $p^2$  a frekvence heterozygotního genotypu jako  $2pq$ , kde  $p$  a  $q$  jsou frekvence alel v populaci. Celková frekvence profilu se vypočítá vynásobením jednotlivých genotypových frekvencí pro každý lokus mezi sebou (Goodwin *et al.*, 2007; Tracey, 2001).

Frekvence genotypů může být použita pro spočítání pravděpodobnosti náhodné shody ( $P_M$ , matching probability, probability of random match) někdy označované také jako pravděpodobnost identity ( $P_I$ , probability of identity), což je pravděpodobnost, že dva náhodně vybraní lidé z dané populace budou mít v testovaném lokusu identický genotyp. Tato pravděpodobnost se spočítá jako součet druhých mocnin genotypových frekvencí, kde  $n$  je počet testovaných genotypů a  $G_i$  je frekvence daného genotypu (Butler, 2005):

$$P_M = \sum_{i=1}^n G_i^2$$

Pravděpodobnost náhodné shody lze definovat také jako pravděpodobnost, že náhodně vybraný jedinec z populace bude mít identický genetický profil. Tato pravděpodobnost odpovídá frekvenci daného genotypu v populaci (Goodwin *et al.*, 2007). Pro všech 13 lokusů zahrnutých v CODIS systému je průměrná pravděpodobnost, že náhodně vybraný jedinec z populace bude mít identický genotyp jako testovaná osoba menší než  $1:10^{12}$  ((Butler, 2005).



## 2.4 Další aplikace

STR lokusy mají kromě určování otcovství a identifikace osob řadu dalších aplikací.

### 2.4.1 Prenatální diagnostika

Testování STR lokusů může být využito pro detekci nejčastějších chromozomálních abnormalit. Nejčastějšími změnami jsou trisomie chromozómů 13, 18, 21 a aneuploidie pohlavních chromozómů. Tyto chromozomální mutace jsou spojeny s Patauovým, Edwardsovým a Downovým syndromem (trisomie chromozómu 13, 18, respektive 21) a méně závažným Turnerovým (monozomie X) a Klinefelterovým (XXY) syndromem (Mann *et al.*, 2004). Pomocí STR markerů lze detekovat i parciálních trisomie a také identifikovat přítomnost dalších buněčných linií ve vzorku, která může znamenat mozaicismus či kontaminaci buňkami matky.

Prenatální diagnostika založená na STR lokusech je rychlá, jednoduchá, citlivá a přesná. Účinnost metody nezávisí na délce těhotenství. Narozdíl od klasických cytogenetických metod mohou být rodiče o výsledku testu informováni v průběhu 24 – 48 hodin od odebrání vzorku. Pomocí amplifikace nepolymorfní sekvence amelogeninového genu (AMXY) lze také určit pohlaví plodu (Cirigliano *et al.*, 2004).

### 2.4.2 Výzkum rakoviny

STR markery jsou také často využívány pro studium genetické instability v mnoha typech nádorů. Při kancerogenezi dochází k přestavbám chromozómů, při čemž dochází ke ztrátě heterozygoty u mikrosatelitů, které se nachází v zasažené oblasti (Chistiakov *et al.*, 2006). Vedle ztráty heterozygoty se v nádorových buňkách vyskytuje také mikrosatelitová instabilita. Ta je definována jako změna délky mikrosatelitu způsobená delecí nebo inzercí repetitivních jednotek v nádoru v porovnání s normální tkání způsobená defektem v opravě chybného párování (mismatch repair, MMR). Mikrosatelitová instabilita byla objevena u kolorektálních nádorů označovaných jako hereditární nepolypózní karcinom tlustého střeva (hereditary non-polyposis colon cancer, HNPCC). Později byl tento jev zjištěn u mnoha sporadicky se vyskytujících typů nádorů.

Výskyt mikrosatelitové instability může sloužit jako prognostický genetický marker. Vysoká hladina mikrosatelitové instability (tj. vyšší než 30 – 40 %) se pravidelně

vyskytuje u kolorektálních nádorů, nádorů žaludku a vaječnicků. Zjištění vysoké hladiny mikrosatelitové instability u kolorektálních nádorů, nádorů žaludku, pankreatu a pravděpodobně i jícnu je spojeno s lepší prognózou pro pacienty. Naproti tomu vysoká instabilita mikrosatelitu u nemalobuněčného karcinomu plic znamená pro pacienta prognózu horší. Mikrosatelitová instabilita může mít vliv na účinnost chemoterapie a radioterapie, hlavně u kolorektálních nádorů a rakoviny vaječnicků (Lawes *et al.*, 2003).

STR markery mohou být také využity pro určení výše rizika vzniku rakoviny. Některé STR lokusy leží uvnitř kódujících sekvencí a mají vliv na transkripci genů. Nižší počet repetitivních jednotek v prvním exonu genu pro androgenový receptor je spojen s vyšším rizikem vzniku rakoviny prostaty u mužů a také vzniku rakoviny prsu u žen. Se zvýšeným rizikem vzniku rakoviny prsu je spojen i polymorfismus v promotoru genu pro estrogenový receptor  $\alpha$ , polymorfismus v genech kódujících proteiny důležité pro metabolismus steroidních hormonů či genu pro IGF-I (Zhang *et al.*, 2007). Pro identifikaci pacientů s rizikem vzniku rakoviny může být využita i mikrosatelitová instabilita. Tato instabilita představuje časně genetické změny u vývoje adenokarcinomů jícnu či prekanceróz hlavy a krku (Lawes *et al.*, 2003).

### **2.4.3 Genetické mapování a vazebná analýza**

Mikrosatelitové markery lze využít pro genetické mapování. Tyto markery byly úspěšně využity při mapování lidských chromozómových chorob (John *et al.*, 2004; Zhang *et al.*, 2005). STR lokusy jsou vysoce polymorfnní (a vysoce informativní) a pro analýzu postačuje malé množství DNA. STR v nekódujících oblastech jsou velmi nápomocné při vytváření hustých vazebných map. Tyto mapy definují pořadí a vzdálenost lokusů podél chromozómů na základě jejich dědičnosti v rodinách či mapovacích populacích. Pokud lokusy leží na stejném chromozómu, nejsou děděny nezávisle, ale vykazují vazbu, která je tím silnější, čím jsou lokusy blíže u sebe. Sleduje se kosegregace markerů, ty jsou umístěny do vazebných skupin a na základě poměru rekombinantů se sestaví nejvíce pravděpodobné pořadí markerů uvnitř skupiny. Genetická vzdálenost se obvykle měří v centimorganech (cM) (Botstein *et al.*, 2003; Chistiakov *et al.*, 2006).

### **2.4.5 Studium evoluce a migrace**

Mikrosatelitové sekvence jsou také důležitým zdrojem informací o fylogenetických vztazích mezi lidskými populacemi a používají se pro studium populačního růstu

a migrace. Protože je ale u těchto lokusů vysoká míra mutací, jsou vhodné pro studium krátkodobé evoluce (King *et al.*, 2000; Zhivotovsky *et al.*, 2003).

Statistické nástroje používané pro zpracování dat získaných analýzou STR lokusů zachází s počtem repetitivních jednotek jako s kvantitativní proměnnou. Rekonstrukce historie populací by měla být založena na genetické informaci z celého genomu. Pro získání spolehlivého závěru o vztazích mezi blízkými příbuznými populacemi by kvůli redukci statistických chyb měla být studie založena na stovkách lokusů, které jsou nezbytné hlavně pro studium separace a expanze populací (Zhivotovsky *et al.*, 2003).

### **3 CÍL PRÁCE**

- provést hodnocení vybraných STR lokusů v populaci
- provést genotypizační analýzu určených STR lokusů a výpočet paternitního a identifikačního indexu na vybraném souboru případů
- srovnat výsledky s komerčním identifikačním kitem AmpF/STR Identifier

## 4 MATERIÁL A METODY

### 4.1 Určení populačních frekvencí

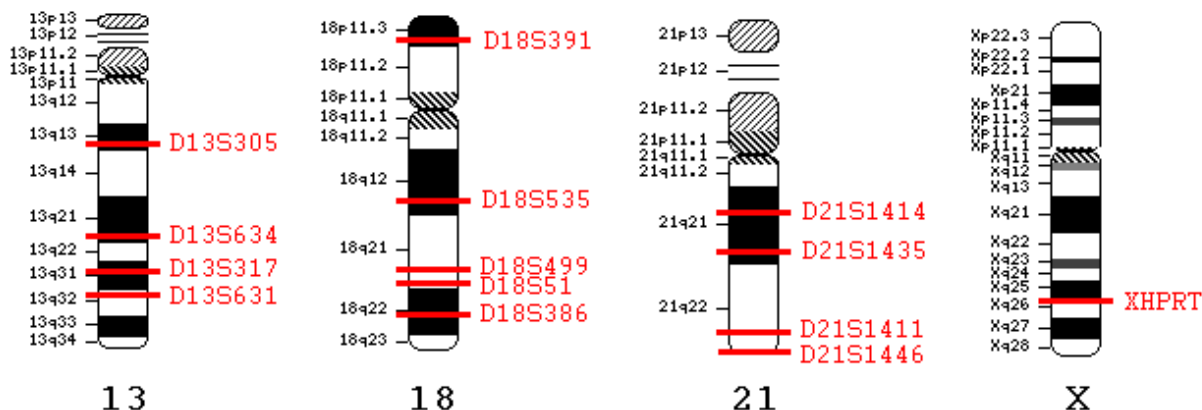
Pro určení alelových frekvencí v populaci byla využita data uložená na Ústavu lékařské genetiky a fetální medicíny Fakultní nemocnice Olomouc od roku 2002 do července 2009. Hodnocené lokusy leží na chromozómech 13, 18, 21 a X (Tab. III, Obr. 5). Celkem bylo hodnoceno 3477 alel lokusu D21S1414, 3454 alel lokusu D21S1435, 2992 alel lokusu D21S1446, 1185 alel lokusu D21S1411, 3015 alel lokusu pentaD, 1666 alel lokusu D18S535, 1548 alel lokusu D18S51, 754 alel lokusu D18S386, 1572 alel lokusu D13S631, 1198 alel lokusu D13S317, 753 alel lokusu D13S305, 620 alel lokusu D13S634 a 253 alel lokusu XHPRT.

Tab. III: Přehled studovaných STR lokusů (Cirigliano *et al.*, 2004; Mann *et al.*, 2004; NCBI Map Viewer; STRBase)

STR lokus	značení	repetice	lokalizace na chromozómu
D21S1414	FAM	(TCTA) <sub>n</sub>	21q21
D21S1435	TET	(TCTA) <sub>n</sub>	21q21
D21S1446	FAM	(TCTA) <sub>n</sub>	21q22.3
D21S1411	HEX	(TCTA) <sub>n</sub>	21q22.3
pentaD	TET	(AAAGA) <sub>n</sub>	21q22.3
D18S535	FAM	(TCTA) <sub>n</sub>	18q12.3
D18S51	FAM	(GAAA) <sub>n</sub>	18q21.33
D18S386	HEX	(GAAA) <sub>n</sub>	18q22.1
D18S391	FAM	(TCTA) <sub>n</sub>	18p11.31
D18S499	TET	(TTTA) <sub>n</sub>	18q21.32
D13S631	HEX	(TCTA) <sub>n</sub>	13q31-32
D13S317	TET	(TCTA) <sub>n</sub>	13q31
D13S305	HEX	(GAAA) <sub>n</sub> / (GGAA) <sub>n</sub>	13q13
D13S634	HEX	(GAAA) <sub>n</sub> / (GGAA) <sub>n</sub>	13q11/q21
XHPRT	FAM	(AGAT) <sub>n</sub>	Xq26.1

Velikosti alel byly posuzovány pomocí velikostního markeru TAMRA500 SizeStandart (Applied Biosystem) v programu Genescan Analysis. Některé alely zřejmě obsahují pouze část repetice. Velikost takové alely potom není celým násobkem repetice.

Nejmenší alela byla označena číslem 1, pokud velikost alely neodpovídala celému násobku repetice, jsou tyto alely označeny desetinným číslem.



Obr. 5: Lokalizace sledovaných STR lokusů (podle NCBI Map Viewer)

U sledovaných STR lokusů byly určovány následující parametry: frekvence alel v populaci, jejich pozorovaná heterozygotita ( $H_{obs}$ ) a očekávaná heterozygotita ( $H_{exp}$ ), průměrná pravděpodobnost vyloučení otcovství ( $\overline{PE}$ ), polymorfní informační obsah (PIC), pravděpodobnost náhodné shody ( $P_M$ ) a diskriminační síla ( $P_D$ ).

Parametry byly počítány podle následujících vzorců (Goodwin *et al.*, 2007), kde  $n$  je počet alel lokusu,  $p_i, p_j$  jsou frekvence jednotlivých alel a  $G_i$  je frekvence genotypu.

Očekávaná heterozygotita (expected heterozygosity,  $H_{exp}$ ) je pravděpodobnost, že dvě náhodně vybrané alely v populaci budou odlišné:

$$H_{exp} = 1 - \sum_{i=1}^n p_i^2$$

Diskriminační síla (power of discrimination,  $P_D$ ) je pravděpodobnost, že dva náhodně vybraní jedinci z populace budou mít rozdílný genotyp:

$$P_D = 1 - \sum_{i=1}^n G_i^2$$

Polymorfní informační obsah (polymorphism information content, PIC) vyjadřuje míru polymorfizmu daného lokusu.

$$PIC = 1 - \sum_{i=1}^n p_i^2 - \left( \sum_{i=1}^n p_i^2 \right)^2 + \sum_{i=1}^n p_i^4$$

Pro chromozóm X se parametry počítají následovně (Szibor *et al.*, 2003):

Průměrná pravděpodobnost vyloučení muže z otcovství pro trojici matka – dcera – označený muž:

$$\overline{PE} = 1 - \sum_{i=1}^n p_i^2 + \sum_{i=1}^n p_i^4 - \left( \sum_{i=1}^n p_i^2 \right)^2$$

Diskriminační síla lokusu pro ženu ( $P_D^f$ ):

$$P_D^f = 1 - 2 \left( \sum_{i=1}^n p_i^2 \right)^2 + \sum_{i=1}^n p_i^4$$

Diskriminační síla lokusu pro muže ( $P_D^m$ ):

$$P_D^m = 1 - \sum_{i=1}^n p_i^2$$

Polymorfní informační obsah lokusu na chromozómu X:

$$PIC = 1 - \sum_{i=1}^n p_i^2 - \sum_{i=1}^n \sum_{i < j} 2 p_i^2 p_j^2$$

kde  $n$  je počet alel lokusu a  $p_i, p_j$  jsou frekvence alel na lokusu.

## 4.2 Analýza paternity a pravděpodobnosti náhodné shody

Vzorky určené pro analýzu paternity pocházely z DNA banky Ústavu lékařské genetiky a fetální medicíny Fakultní nemocnice Olomouc, které byly již dříve pro paternitní účely analyzovány komerčním kitem. Celkem bylo analyzováno 21 trojic otec – dítě – matka a 1 dvojice otec – dítě. Všechny vzorky byly anonymizovány.

Vzorky byly amplifikovány pomocí dvou multiplex PCR. Směs primerů pro chromozóm 21 obsahovala primery pro lokusy D21S1414, D21S1435 a D21S1411. Směs primerů pro chromozómy 13 a 18 obsahovala primery pro lokusy D13S631, D13S317, D13S305, D13S634, D18S51, D18S386, D18S499, D18S391. STR lokusy D21S1446 a pentaD byly amplifikovány zvlášť. Směs pro jednu reakci obsahuje 5  $\mu$ l Master mix (TopBio), 0,3  $\mu$ l primerů, 1  $\mu$ l templátová DNA a doplněna vodou na objem 10  $\mu$ l. Vzorky byly nejdříve vystaveny predenaturaci při 95°C po dobu 10 min, poté proběhlo 29 cyklů PCR: denaturace 30 s při 94°C, nasedání primerů 1 min při 59°C, prodlužování primerů 1 min při 72°C. Poté proběhla finální extenze po dobu 10 min při teplotě 72°C a 30 min při teplotě 60°C.

DNA byla po amplifikaci analyzována pomocí kapilární elektroforézy ABI PRISM 310 Genetic Analyser (Applied Biosystem). Vzorky pro analýzu byly připraveny smícháním 1  $\mu$ l amplifikované DNA, 10  $\mu$ l formamidu a 0,5  $\mu$ l TAMRA500 SizeStandart. Vzorky byly poté denaturovány po dobu 3 min při teplotě 95°C a následně přeneseny na 3 min na led. Vlastní elektroforéza probíhala za standardních podmínek v modulu GS STR POP4 C.

Pro výpočet paternitního indexu a následně pravděpodobnosti otcovství byly využity zjištěné frekvence alel. U lokusů D18S499 a D18S391 nebyly frekvence určeny a pro výpočet byla použita nejvyšší zjištěná frekvence alely ze všech zkoumaných lokusů. U lokusu D13S317 byla u 4 studovaných rodin, u lokusu D21S1414 a D21S1411 u jedné rodiny zjištěna alela mimo rozsah alel. Pro výpočet byla použita frekvence nejméně čtené alely daného lokusu.

Paternitní index byl vypočítán podle Brennera. Pro většinu sledovaných lokusů není známa mutační rychlost. Literatura udává, že mutační rychlost u STR lokusů využívaných pro paternitní účely se pohybuje v rozmezí od  $5 \cdot 10^{-4}$  do  $7 \cdot 10^{-3}$  na generaci (Vicard *et al.*, 2008). Pro výpočet byla použita hodnota  $\mu = 3,75 \cdot 10^{-3}$ .

Pro porovnání výsledků určení pravděpodobnosti otcovství a pravděpodobnosti náhodné shody pomocí sledovaných STR lokusů a komerčního kitu AmpF/STR Identifier (Promega) byl použit párový t-test.



## **5 VÝSLEDKY**

### **5.1 Frekvence sledovaných STR lokusů v populaci**









## **5.2 Analýza paternity a určení pravděpodobnosti náhodné shody**





















































## 6 DISKUZE

STR lokusy jsou vysoce polymorfní, kodominantní a široce využívané genetické markery. Své uplatnění nachází při prenatální diagnostice, výzkumu rakoviny, genetickém mapování a populačních studiích. Velký význam mají STR lokusy při identifikaci osob a testování paternity.

U vybraných STR lokusů na chromozómech 21 (lokusy D21S1414, D21S1435, D21S1446, D21S1411, penta D), 18 (lokusy D18S535, D18S51, D18S386), 13 (lokusy D13S631, D13S317, D13S305, D13S634) a X (XHPRT) byla stanovena frekvence jednotlivých alel. Velký vliv na informativnost STR lokusu má počet alel a jejich distribuce. Více možných alel v lokusu a vyrovnanější distribuce jejich alelových frekvencí má za následek vyšší heterozygotnost. Trojice lokusů D21S1435, D21S1411 a penta D, trojice D21S1414, D18S535 a D13S317 a dvojice lokusů D21S1446 a D18S51 mají stejné počty alel, ale jejich pravděpodobnost náhodné shody se liší téměř dvojnásobně. Pokud je na několika lokusech stejný počet alel, u lokusu s lepší vyvážeností alelových frekvencí je nižší pravděpodobnost náhodné shody (Butler, 2005). Z uvedených skupin lokusů mají nejnižší pravděpodobnost náhodné shody lokusy penta D, D13S317 a D18S51.

Na základě zjištěných dat je ze souboru vybraných STR lokusů pro forenzní účely nejvhodnější lokus D18S386. U tohoto lokusu byl zjištěn nejvyšší počet alel, nejvyšší pozorovaná heterozygotita, průměrná pravděpodobnost vyloučení z otcovství a polymorfní informační obsah. Zároveň má tento lokus velmi nízkou pravděpodobnost náhodné shody, která pro tento lokus udává, že u dvou náhodně vybraných jedinců z populace je pravděpodobnost nalezení stejného genotypu na lokusu nižší než 1%. Naproti tomu nejvyšší pravděpodobnost náhodné shody, 6,1%, byla zjištěna u lokusu D21S1446. Tento lokus má také nejnižší polymorfní informační obsah. Nejnižší pozorovanou heterozygotitu a pravděpodobnost vyloučení vykazoval lokus D13S305.

Znalost frekvence alel pro jednotlivé lokusy je nezbytná pro určení paternitního indexu a pravděpodobnosti otcovství při hodnocení paternitního testu. Pro analýzu paternity a určení pravděpodobnosti náhodné shody pro identifikační účely nebyly použity markery D18S535 a XHPRT, ale byly využity markery D18S391 a D18S499.

Ačkoliv každý genetický marker dítěte je zděděn od jednoho z rodičů, výsledky paternitního testu mohou ukázat odlišnost mezi genotypem dítěte a genotypem muže označeného za otce. Se zvyšujícím se počtem testovaných lokusů roste pravděpodobnost, že bude zjištěna neshoda v genotypech (Jacewicz *et al.*, 2004). Aby mohl být muž

z otcovství vyloučen, musí být při testování 12 a více STR lokusů nalezeny více jak dvě neshody (Butler, 2005). Možným vysvětlením těchto neshod je, že došlo k mutaci. U STR lokusů využívaných pro paternitní účely a identifikaci osob dochází k mutacím v rozmezí  $5 \cdot 10^{-4}$  až  $7 \cdot 10^{-3}$  na generaci (Vicard *et al.*, 2008). Možnost mutace musí být při určování otcovství brána v potaz (Jacewicz *et al.*, 2004). Mutace v STR lokusu zjištěná u muže, který byl označen za otce dítěte, výrazně snižuje hodnotu paternitního indexu (Buckleton *et al.*, 2005). Ačkoliv mutace mohou mít dopad na výsledek paternitního testu, nemají ale vliv na identifikaci osob, protože mutace jsou v průběhu života konzistentní (Butler, 2005).

Celkem byly u komerčního kitu AmpF/STR Identifier zjištěny dva případy paternální mutace a jeden případ mutace maternální. Při testování paternity pomocí vybraných STR lokusů bylo celkově detekováno dvanáct mutací paternálních a dvě mutace maternální. Tyto výsledky jsou v souladu s faktem, že míra mutací v STR lokusech je ovlivněna pohlavím. Paternální mutace jsou mnohem častější než mutace maternální (Buckleton *et al.*, 2005; Fan *et al.*, 2007). Ačkoliv na základě výsledku párového t-testu ( $t = 0,966$ ,  $df = 21$ ,  $P = 0,05$ ) nebyl mezi pravděpodobností otcovství určenou pomocí kitu AmpF/STR Identifier a testovaným souborem STR lokusů nalezen signifikantní rozdíl, u markerů využívaných pro analýzu paternity by neměla být míra mutací příliš vysoká.

Markery používané pro testování paternity musí vykazovat mendelistickou dědičnost, musí být distribuovány podle Hardy-Weinbergova zákona a odlišné genetické systémy musí být děděny nezávisle na sobě, tzn. musí být ve vazebné rovnováze. Genetické systémy jsou na sobě nezávislé, pokud jsou lokalizovány na různých chromozómech, nebo pokud se nachází na stejném chromozómu dostatečně daleko od sebe. Podmínka nezávislosti lokusů je důležitá pro stanovení frekvence genotypu ve více STR lokusech. Pokud segregace alel do gamet jednoho lokusu není výrazně ovlivněna segregací alel jiného lokusu, celková frekvence profilu se vypočítá vynásobením jednotlivých genotypových frekvencí pro každý lokus mezi sebou. Možná vazba mezi alelami v různých STR lokusech (vazebná nerovnováha) může ovlivnit frekvenci genotypu a tím i pravděpodobnost náhodné shody (Butler, 2006; Holt *et al.*, 2000).

Podmínku nezávislosti splňují markery použité v komerčním kitu, protože lokusy jsou lokalizovány na různých chromozómech. Výjimku tvoří lokusy D5S818 a CSF1PO ležící na chromozómu 5. Mezi těmito lokusy je ale dostatečně velká vzdálenost (asi 26,5 Mb). Lokalizace lokusů na stejném chromozómu ve vzdálenosti několika Mb by neměla mít dopad na spolehlivé určení celkové frekvence genetického profilu vynásobením genotypů v jednotlivých lokusech (Butler, 2006). Testovaný soubor STR

lokusů podmínku nezávislosti zcela nesplňuje. Tento soubor obsahoval 5 STR lokusů lokalizovaných na chromozómu 21, 4 lokusy na chromozómu 18 a 4 lokusy na chromozómu 13. Přesto se pravděpodobnost náhodné shody u tohoto souboru markerů pohybovala v rozmezí  $2,12 \cdot 10^{-16}$  do  $1,63 \cdot 10^{-30}$ . Statistický srovnáním obou souborů markerů pomocí párového t-testu ( $t = 0,183$ ,  $df = 21$ ,  $P = 0,05$ ) nebyl mezi nimi prokázán signifikantní rozdíl.



## 7 ZÁVĚR

Ve spádové oblasti Morava byl ve 13 STR lokusech (D21S1414, D212S1435, D21S1446, D21S1411, penta D, D18S535, D18S51, D18S386, D13S631, D13S317, D13S305, D13S634 a XHPRT) hodnocen výskyt alelových frekvencí v populaci. Na základě zjištěných dat je nejpolymorfnějším zkoumaným STR lokusem D18S386. Naproti tomu nejméně polymorfní lokusy jsou D21S1446 a D13S305.

Pro určení pravděpodobnosti otcovství a pravděpodobnosti náhodné shody u sledovaného muže byly místo markerů D18S535 a XHPRT využity markery D18S391 a D18S499. Při porovnání dat získaných pomocí testované sady markerů a komerčního kitu AmpF/STR Identifier u stejných rodin nebyl po statistické analýze mezi použitými sadami markerů prokázán signifikantní rozdíl. U testovaných STR lokusů bylo ale celkem nalezeno více mutací než u komerčního kitu, což může mít vliv na výsledek paternitního testu, ale nemá vliv na identifikaci osob.

## 8 LITERATURA

- Ayres, K., L., Powley, W., M. (2005): Calculating the exclusion probability and paternity index for X-chromosomal loci in the presence of substructure. *Forensic Science International* 149: 201 – 203.
- Botstein, D., Risch, N. (2003): Discovering genotypes underlying human phenotypes: past successes for mendelian disease, future approaches for complex disease. *Nature Genetics supplement* 33: 228 – 237.
- Buckleton, J., Clayton, T., Triggs, C. (2005): Parentage testing. In: Buckleton, J., Triggs, C., M., Wals, S., J. (ed.): *Forensic DNA evidence interpretation*. CRP Press, USA.
- Butler, J., M. (2005): *Forensic DNA typing*. Elsevier academic press, USA.
- Butler, J., M. (2006): Genetics and genomics of core short tandem repeat loci used in human identity testing. *Journal of Forensic Science* 51 (2): 253 – 265.
- Cirigliano, V., Voglino, G., Cañadas, M., P., Marongiu, A., Ejarque, M., Ordoñez, E., Plaja, A., Massobrio, M., Todros, T., Fuster, C., Campogrande, M., Egozcue, J., Adinolfi (2004): Rapid prenatal diagnosis of common chromosome aneuploidies by QF-PCR. Assessment on 18 000 consecutive clinical samples. *Molecular Human Reproduction* 10 (11): 839 – 846.
- Ellegren, H. (2000): Heterogenous mutation processes in human microsatellite DNA sequences. *Nature Genetics* 24: 400 – 402.
- Fan, H., Chu, J.-Y. (2007): A brief review of short tandem repeat mutation. *Genomics, Proteomics & Bioinformatics* 5 (1): 7 – 14.
- Fung, W., K., Chung, Y., Wong, D. (2002): Power of exclusion revisited: probability of excluding relatives of the true father from paternity. *International Journal of Legal Medicine* 116: 64 – 67.
- Gjertson, D., W., Brenner, C., H., Baur, M., P., Carracedo, A., Guidet, F., Luque, J., A., Lessing, R., Mayr, W., R., Pascali, V., L., Prinz, M., Schneider, P., M., Morling, N. (2007): ISFG: Recommendations on biostatistics in paternity testing. *Forensic Science International: Genetics* 1: 223 – 231.
- Goodwin, W., Linacre, A., Hadi, S. (2007): *An introduction to forensic genetics*. John Wiley & Sons Ltd, Chichester.
- Hanchard, N., A. (2005): Genetic susceptibility and single-nucleotide polymorphisms. *Seminars in Fetal & Neonatal Medicine* 10: 283 – 289.
- Holt, C., L., Stauffer, C., Wallin, J., M., Lazaruk, K., D., Nguyen, T., Budowle, B., Walsh, P., S. (2000): Practical applications of genotypic surveys for forensic STR testing. *Forensic Science International* 112: 91 – 109.

- Huang, Q., Y., Xu, F., H., Shen, H., Deng, H., Y., Liu, Y., J., Liu, Y., Z., Li, J., L., Recker, R., R., Deng, H., W. (2002): Mutation patterns at dinucleotide microsatellite loci in humans. *The American Journal of Human Genetics* 70: 625 – 634.
- Chakraborty, R., Kimmel, M., Stivers, D., N., Davison, L., J., Deka, R. (1997): Relative mutation rates at di-, tri-, and tetranucleotide microsatellite loci. *The Proceedings of the National Academy of Science USA* 94: 1041 – 1046.
- Chistiakov, D., A., Hellemans, B., Volckaert, F., A., M. (2006): Microsatellites and their genomic distribution, evolution, function and applications: A review with special reference to fish genetics. *Aquaculture* 255: 1 – 29.
- International Human Genome Sequencing Concorcium (2001): Initial sequencing and analysis of the human genome. *Nature* 409: 860 – 921.
- Jacewicz, R., Berent, J., Prośniak, A., Dobosz, T., Kowalczyk, E., Szram, S. (2004): Non-exclusion paternity case with a triple genetic incompatibility. *International Congress Series* 1261: 511 – 513.
- John, S., Shephard, N., Liu, G., Zeggini, E., Cao, M., Chen, W., Vasavda, N., Mills, T., Barton, A., Hinks, A., Eyre, S., Jones, K., W., Ollier, W., Silman, A., Gibson, N., Worthington, J., Kennedy, G., C. (2004): Whole-genome scan, in a complex disease, using 11 245 single-nucleotide polymorphisms: comparison with microsatellites. *American Journal of Human Genetics* 75: 54 – 64.
- Kayser, M., Roewer, L., Hedman, M., Henke, L., Henke, J., Brauer, S., Krüger, C., Krawczak, M., Nagy, M., Dobosz, T., Szibor, R., Kniff, P., Stoneking, M., Sajantila, A. (2000): Characteristics and frequency of germline mutations at microsatellite loci from the human Y chromosome, as revealed by direct observation in father/son pairs. *The American Journal of Genetics* 66: 1580 – 1588.
- Kayser, M., Sajantila, A. (2001): Mutations at Y-STR loci: implications for paternity testing and forensic analysis. *Forensic Science International* 118: 116 – 121.
- King, J., P., Kimmel, M., Chakraborty, R. (2000): A power analysis of microsatellite-based statistics for inferring past population growth. *Molecular Biology and Evolution* 17 (12): 1859 – 1868.
- Kruglyak, S., Durrett, R., T., Schug, M., D., Aquadro, C., F. (1998): Equilibrium distributions of microsatellite repeat length resolving from a balance between slippage events and point mutations. *The Proceedings of the National Academy of Science USA* 95: 10774 – 10778.
- Lai, Y., Sun, F. (2003): The relationship between microsatellite slippage mutation rate and the number of repeat units. *Molecular Biology and Evolution* 20 (12): 2123 – 2131.
- Lawes, D., A., SenGupta, S., Boulos, P., B. (2003): The clinical importance and prognostic implications of microsatellite instability in sporadic cancer. *European Journal of Surgical Oncology* 29: 201 – 212.
- Lee, H.-S., Lee, J., W., Han, G.-R., Hwang, J.-J. (2000): Motherless case in paternity testing. *Forensic Science International* 114: 57 – 65.

- Lee, J., W., Lee, H.-S., Park, M., Hwang, J., J. (2001): Paternity determination when the alleged father's genotypes are unavailable. *Forensic Science International* 123: 202 – 210.
- Mann, K., Donaghue, C., Fox, S., P., Docherty, Z., Ogilvie C., M. (2004): Strategies for the rapid prenatal diagnosis of chromosome aneuploidy. *European Journal of Human Genetics* 12: 907 – 915.
- Mertens, G., Vandenberghe, A., Muylle, L. (1999): Paternity testing: a review of theoretical and practical issues. *Archives of Public Health* 57: 131 – 140.
- Nadir, E., Margalit, H., Gallily, T., Ben-Sasson, S., A. (1996): Microsatellite spreading in the human genome: Evolutionary mechanisms and structural implications. *The Proceedings of the National Academy of Science USA* 93: 6470 – 6475.
- Nachman, M., W., Crowell, S., L. (2000): Estimate of the mutation rate per nucleotide in humans. *Genetics* 156: 297 – 304.
- Näslund, K., Saetre, P., Salomé, J., Bergström, T., F., Jareborg, N., Jazin, E. (2005): Genome-wide prediction of human VNTRs. *Genomics* 85: 24 – 35.
- Nussbaum, R., L., McInnes, R., R., Willard, H., F. (2004): *Klinická genetika*. Triton, Praha.
- Payseur, B., A., Nachman, M., W. (2000): Microsatellite variation and recombination rate in the human genome. *Genetics* 156: 1285 – 1298.
- Presciuttini, S., Toni, C., Marroni, F., Spinetti, I., Bailey-Wilson, J., E., Domenici, R. (2004): The number of STR markers necessary to resolve relationships in deficiency paternity cases. *International Congress Series* 1261: 541 – 543.
- Ruitberg, C., M., Reeder, D., J., Butler, J., M. (2001): STRBase: a short tandem repeat DNA database for the human identity testing community. *Nucleic Acid Research* 29 (1): 320 – 322.
- Serra, A., Bento, A., M., Carvalho, M., Andrae, L., Batista, L., Oliveira, M., C., Lopes, V., Balsa, F., Corte-Real, F., Anjos, M., J. (2008): X-chromosome STR typing in deficiency paternity cases. *Forensic Science International: Genetics Supplement Series* 1: 162 – 163.
- Schlötterer, C. (1998): Genome evolution: Are microsatellites really simple sequences? *Current Biology* 8: 132 – 134.
- Schlötterer, C. (2000): Evolutionary dynamics of microsatellite DNA. *Chromosoma* 109: 365 – 371.
- Subramanian, S., Mishra, R., K., Singh, L. (2003): Genome-wide analysis of microsatellite repeats in humans: their abundance and density in specific genomic regions. *Genome Biology* 4: R13.

- Szibor, R., Krawczak, M., Hering, S., Edelmann, J., Kuhlisch, E., Krause, D. (2003): Use of X-linked markers for forensic purposes. *International Journal of Legal Medicine* 117: 67 – 74.
- Tamaki, K., Jeffreys, A., J. (2005): Human tandem repeat sequences in forensic DNA typing. *Legal Medicine* 7: 244 – 250.
- Tracey, M. (2001): Short tandem repeat-based identification of individuals and parents. *Croatian Medical Journal* 42 (3): 233 – 238.
- Vergnaud, G., Denoeud, F. (2000): Minisatellites: Mutability and genome architecture. *Genome Research* 10: 899 – 907.
- Vicard, P., Dawid, A., P., Mortera, J., Lauritzen, S., L. (2008): Estimating mutation rates from paternity casework. *Forensic Science International: Genetics* 2: 9 – 18.
- Zhang, C., Cawley, S., Liu, G., Cau, M., Gorrell, H., Kennedy, G., C. (2005): A genome-wide linkage analysis of alcoholism on microsatellite and single-nucleotide polymorphism data, using alcohol dependence phenotypes and electroencephalogram measures. *BMC Genetics* 6: S17 – S22.
- Zhang, W., Yu, Y., Y. (2007): Polymorphisms of short tandem repeat of genes and breast cancer susceptibility. *European Journal of Surgical Oncology* 33: 529 – 534.
- Zhivotovsky, L., A., Rosenberg, N., A., Feldman, M., W. (2003): Features of evolution and expansion of modern humans, inferred from genomewide microsatellite markers. *American Journal of Human Genetics* 72: 1171 – 1186.

### **Internetové zdroje:**

NCBI Map Viewer: <http://www.ncbi.nlm.nih.gov/projects/mapview/>

Short Tandem Repeat DNA Internet Database (STRBase): <http://www.cstl.nist.gov/strbase/>

## 9 SEZNAM POUŽITÝCH ZKRATEK A SYMBOLŮ

AMXY	amelogenin
bp	pár bazí (base pair)
cM	centimorgan
CODIS	the combined DNA index system
CPI	kombinovaný paternitní index
DNA	deoxyribonukleová kyselina
FBI	Federal Bureau of Investigation
Gb	gigabáze (giga base pair)
$H_{exp}$	očekávaná heterozygotita (expected heterozygosity)
HNPCC	hereditární nepolypózní karcinom tlustého střeva
$H_{obs}$	pozorovaná heterozygotita (observed heterozygosity)
IAM	model nekonečného množství alel (infinite alleles model)
IGF-I	insulin like growth factor I
kb	kilo pár bazí (kilo base pair)
LINE	dlouhé roztroušené elementy (long interspred elements)
LTR	dlouhé koncové repetice (long terminal repeats)
Mb	megabáze (mega base pair)
MMM	vícekový mutační model (multistep mutation model)
MMR	oprava chybného párování bazí (mismatch repair)
PCR	polymerázová řetězová reakce
$P_D$	diskriminační síla (power of discrimination)
$P_D^f$	diskriminační síla pro ženu
$P_D^m$	diskriminační síla pro muže
PE	pravděpodobnost vyloučení z otcovství (power / probability of exclusion)
$\overline{PE}$	průměrná pravděpodobnost vyloučení z otcovství
PI	paternitní index
$P_I$	pravděpodobnost identity (probability of identity)
PIC	polymorfni informační obsah (polymorphism information content)
$P_M$	pravděpodobnost náhodné shody (matching probability, probability of random match)
SINE	krátké roztroušené elementy (short interspred elements)
SMM	řadový mutační model (stepwise mutation model)
SNP	jednonukleotidový polymorfismus (single nucleotide polymorphism)

STR	krátké tandemové repetice (short tandem repeats)
VNTR	variabilní počet tandemových repetice (variable number of tandem repeats)
W	pravděpodobnost otcovství

## **10 PŘÍLOHY**

**STR lokusy hodnocené v populaci – velikost alel, počet zjištěných alel a jejich frekvence**





