

Genetická genealogie – sledování rodových linií pomocí analýzy DNA

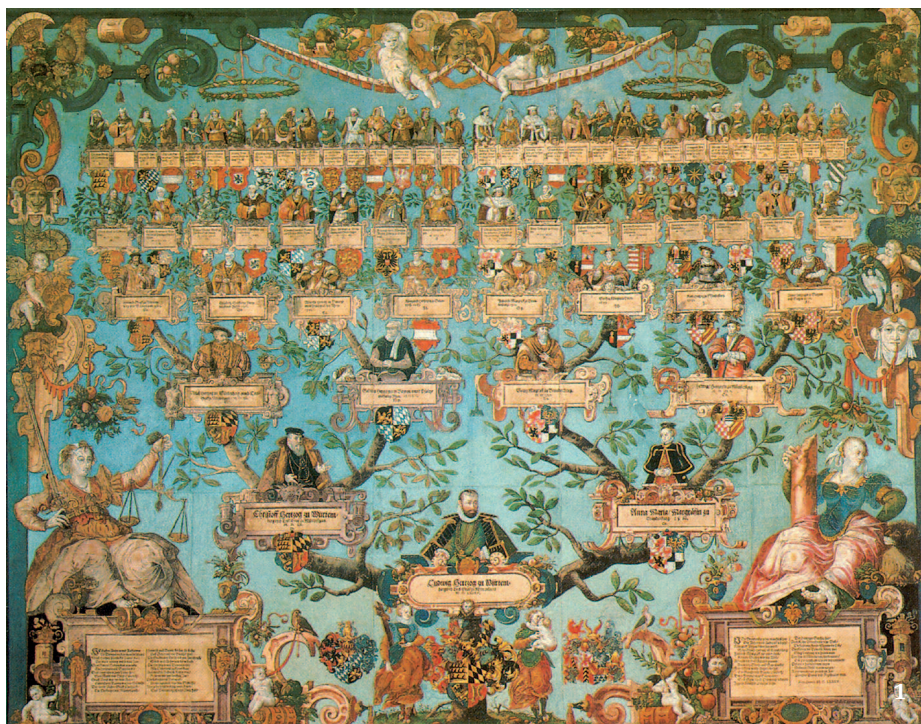
I. Chromozom Y

„Žádný člověk nemůže poznat cíl své životní pouti, pokud se nedozví, odkud přesně pochází a jak se dostal na své současné místo.“
Maya Angelou (americká básnířka)

Sir George Howard Darwin (1845–1912), syn Charlese Darwina, použil v r. 1875 naprosto jedinečný přístup analýzy příjmení pro odhad frekvence sňatků mezi bratranci a sestřenicemi. Ve svém článku publikovaném v *Journal of Statistical Society of London* Darwin vypočetl předpokládanou frekvenci sňatků osob se stejným příjmením (izonymie). Došel k závěru, že se sňatky mezi britskými bratranci a sestřenicemi pohybují v rozmezí 2,25 % až 4,5 %, přičemž vyšší společenské třídy vykazovaly vyšší výskyt manželství „v rámci rodiny“. Za zmínku určitě stojí, že jeho rodiče (Charles Darwin a Emma Wedgwood) byli také bratranec a sestřenice. G. H. Darwin ve své práci dále zkoumal, zda se v psychiatrických léčebnách nenachází oproti průměru v populaci více potomků bratřinců a sestřenic. Přestože se za svého života věnoval hlavně matematice a astronomii, můžeme ho díky jeho přínosu na poli zkoumání vztahu příjmení a příbuzenského vztahu bez jakýchkoli pochyb pokládat za otce genetické genealogie.

Samotný obor genetické genealogie je však velmi mladý. Jeho znovuzrození lze datovat na začátek 90. let 20. stol., kdy byly na chromozomu Y lokalizovány STR oblasti (Short Tandem Repeats – krátké tandemové repetice, viz dále) použitelné pro sledování mužských linií. Významný pak byl r. 1997, ve kterém Karl Skorecki

a jeho spolupracovníci z lékařské fakulty v Haifě (Izrael) publikovali v časopise *Nature* odborný článek s názvem *Y chromosomes of Jewish Priests* (Chromozomy Y u židovských kněží), ve kterém právě pomocí analýzy tohoto chromozomu potvrdili teorii o společném předkovi židovských kněží – kohenů (viz dále).



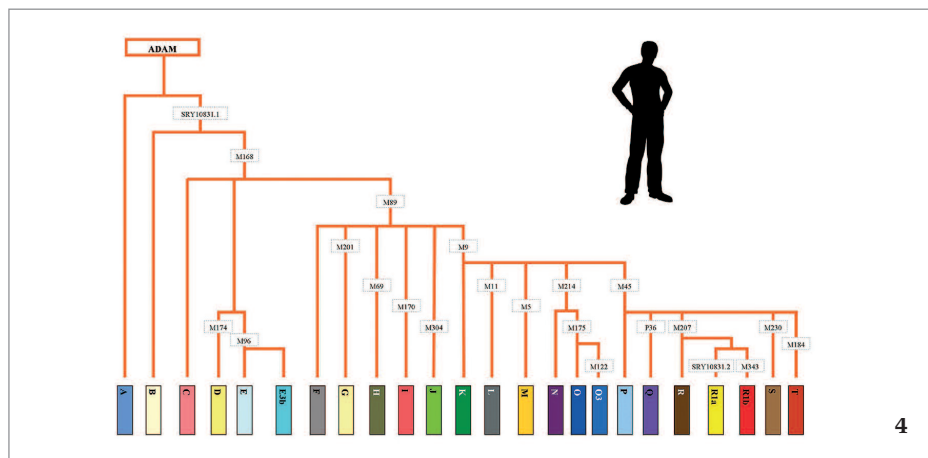
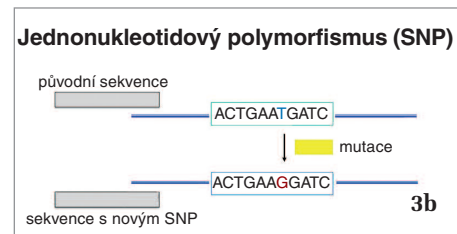
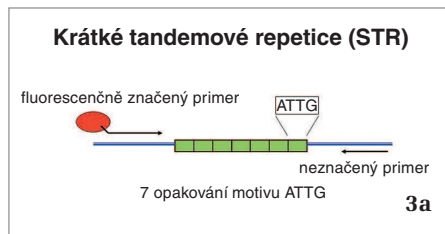
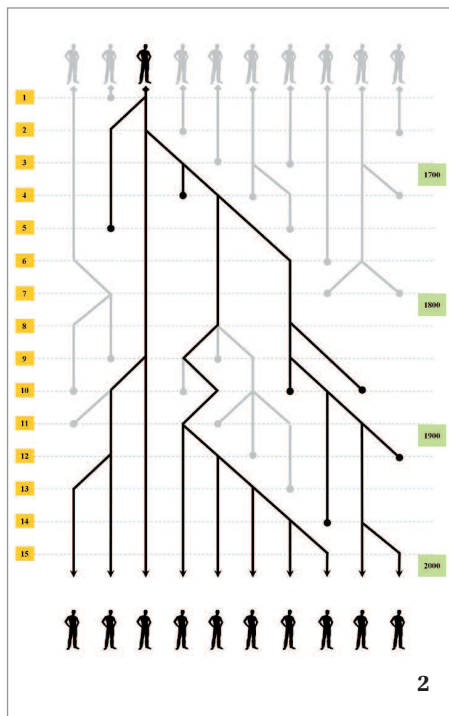
Masový rozvoj genetické genealogie způsobila společnost Family Tree DNA, která v r. 2000 jako první začala nabízet komerční testy DNA za účelem zjištění geografického původu předků a rodinných vazeb. Velkou popularitu tomuto oboru pak zajistila kniha *Sedm dcer Eviných*, kterou vydal v r. 2001 Bryan Sykes. Dalším milníkem je Genografický projekt spuštěný v r. 2005 společností National Geographic Society a firmou IBM. Cílem tohoto ojedinělého projektu je zmapovat pomocí analýzy DNA migrace lidí v historii. V České republice se genetickou genealogií zabývá např. laboratoř Forenzní DNA servis v Praze.

Jaké oblasti DNA zkoumá genetická genealogie při analýze mužských linií?

Geneticko-genealogická analýza DNA mužských linií je zaměřena na nerekombinující část chromozomu Y (NRY, Non-recombining region of the human Y chromosome) – ostatní páry chromozomů 1–22 a chromozom X si mohou mezi sebou při rekombinaci vyměňovat genetickou informaci. Tato část chromozomu Y obsahuje jadernou DNA, která se nalézá pouze u mužů a dědí se výhradně po paternální linii, tedy z otce na syna, podobně jako u nás rodné příjmení. Tento systém přenosu dědičné informace tak vytváří nepřerušovanou spojnicí mezi generacemi a může být velkým pomocníkem při rekonstrukci rodinné historie. I když představuje chromozom Y pouze 0,38 % celkové DNA obsažené v lidské buňce a je nejmenším ze všech chromozomů, obsahuje největší nerekombinující blok DNA v lidském genomu (přibližně 58 milionů párů bází) a je proto vhodným „záznamovým zařízením“ genetických změn.

Genetické studie ukázaly, že v určitém okamžiku vývoje lidstva se velikost efektivní populace snížila na necelých 3 000 jedinců, což bylo příčinou, že se do dnešních dnů zachovala pouze jedna hlavní „adamovská“ varianta Y chromozomu. Genetický „praotec Adam“ žil asi před 60–90 tisíci lety v Africe. V odborné literatuře je též někdy nazýván Y-MRCA (Most Recent Common Ancestor – nejbližší společný předek). Všichni muži v sobě mají kousek jeho DNA. Chromozom Y u dnešních mužů se však od původního „adamovského“ liší o všechny dědičné změny (mutace), které na chromozomu Y v průběhu věků vznikly. Skutečnost, že pro genetického praotce bylo zvoleno jméno Adam, je však poněkud zavádějící. V žádném případě to neznamená, že by v Africe žil jen jeden muž. Současně s genetickým Adamem žila celá řada mužů, pouze nelze nalézt jejich genetickou stopu v současné populaci. Na obr. 2 je znázorněn přenos nerekombinující části chromozomu Y generacemi. Genetická informace některých mužů uložená na chromozomu Y nemusí být předána dalším generacím. Tito muži v průběhu generací buď nemají potomky, nebo mají jen dcery, které od otce nemohou přijmout Y chromozom, a tedy ho ani nemohou předat svým potomkům.

1 Příklad grafického výstupu klasické genealogie – strom rodiny Herzog Ludwig. Z archivu autora



Mezi znaky chromozomu Y, které jsou nejčastěji zkoumány genealogickými a forenzními DNA laboratořemi, patří STR (krátké tandemové repetice, obr. 3a) a SNP (Single Nucleotide Polymorphism – jednonukleotidové polymorfismy, obr. 3b) oblasti. K analýze se mohou dále používat inserční a deleční polymorfismy (INDELs) a minisatelity (ty mají oproti STR delší motiv opakování, někdy se nazývají také variabilní počet tandemových repetit, VNTR – Variable Number of Tandem Repeats). Sada získaných hodnot může být použita ke stanovení haplotypu Y chromozomu. Z geneticko-genealogického pohledu je haplotyp přepisem informací o variabilních místech na tomto chromozomu.

Haplo skupinou se nazývají rodiny chromozomů Y příbuzných po rodové linii. Haplo skupiny chromozomu Y (zjednodušeně na obr. 4) jsou definovány rozdíly v nerekombinující části DNA. STR a SNP systémy jsou výhodné k odlišení chromozomů Y různých rodových linií. Pro tento účel je velmi důležitá rozdílná mutační rychlost STR a SNP. Binární SNP markery (jednonukleotidový polymorfismus může nabývat jen dvou hodnot, např. u mutace M286 můžeme detekovat jen báze guanin nebo adenin) mutují velmi pomalu (10^{-8} na bázi na generaci), lze je považovat za prakticky neměnné, a proto mohou velmi dobře definovat jednotlivé haplo skupiny. Např. haplo skupina D chromozomu Y (Hg D) je definována mutací M174 (obr. 4). Průměrná mutační rychlost STR je naopak mnohonásobně větší (10^{-3} na STR na generaci). Pokud bychom sledovali hypotetickou mužskou linii od zakladatele rodu po dobu 1 000 let (50 generací), tak takřka určitě nedojde na chromozomu Y ke změně haplo skupiny (charakteristické SNP mutace), ale zcela určitě dojde ke změně haplotypu. Např. u haplotypu, který se skládá ze 30 STR systémů, lze očekávat, že haplotyp zakladatele rodu a jeho praprapra...vnuků se bude po 50 generacích (1 000 let) při průměrné mutační rychlosti STR lokusů 0,2 % (pravděpodobnost změ-

ny alely STR lokusu v % vztažená na přenos genetické informace na další pokolení) lišit ve třech STR systémech. Pokud by ale průměrná mutační rychlost byla dvojnásobná (tj. 0,4 %), bude se zakladatel rodu a jeho potomci po tisících letech lišit pravděpodobně již v 6 STR systémech (obr. 5).

Slovo pravděpodobně je zde zcela na místě. Je totiž nutné si uvědomit, že mutace jsou náhodné události a mohou nastat kdykoli. K předpokládanému počtu mutací tak může dojít dříve nebo později, než je teoretický výpočet (předpokládaný počet mutací je průměrnou hodnotou, kterou dostaneme při sledování velkého počtu podobných případů).

Co lze z genealogického hlediska určit pomocí analýzy chromozomu Y?

- Zda mají různí jedinci mužského pohlaví stejného předka;
- zda jsou dvě nebo více osob mužského pohlaví se stejným, stejně znějícím nebo významově stejným příjmením spřízněny přes společného předka;
- kolik různých společných předků má daná skupina osob (např. všichni nositelé příjmení Kropáček);
- do jaké rodové linie (haplo skupiny) náleží určitý haplotyp Y chromozomu a kde tato linie vznikla (např. na Blízkém východě);
- analýza výskytu nové mutace chromozomu Y může být použita pro odhad stupně separace mezi jednotlivými muži, a to z hlediska počtu generací, které uplynuly od této mutace.

Genetika a příjmení

V řadě populací se příjmení předávají z generace na generaci. Pokud se příjmení, stejně tak jako chromozom Y předává pouze z otce na syna, lze charakteristické znaky předávaného chromozomu Y sledovat v průběhu generací. Zatímco v Číně byla rodová příjmení používána s největší pravděpodobností již před cca 5 000 lety, v Turecku byla zavedena až v r. 1934

2 Přenos chromozomu Y ve sledu generací – schematické znázornění ztráty genetické informace při nezplození potomků mužského pohlaví. Šedivé linie značí zplození dcery, černé linie syny, ukončené linie plným kroužkem značí absenci potomka. Ve žlutých rámečcích je uveden počet uplynulých generací.

3a Schéma krátkých tandemových repetit (STR). Osoby se mezi sebou mohou lišit počtem opakování. Pokud zkoumáme STR na autozomech, může být jedinec v určitém systému homozygotní (obě alely mají stejnou sekvenci), nebo heterozygotní (alely se liší sekvencí). STR markery mají poměrně vysokou identifikační schopnost.

3b Schéma jednonukleotidového polymorfismu (SNP). Jde o změnu jedné báze v určité pozici. SNP markery mají poměrně malou identifikační schopnost.

4 Zjednodušený strom haplo skupin chromozomu Y. Charakteristické SNP mutace jsou uvedeny v rámečku.

Na úplném vrcholu stromu si můžeme představit genetického „Adama“ – prapředka, do jehož DNA na Y chromozomu se v průběhu tisíciletí zaznamenávaly mutace, na základě kterých odlišujeme jednotlivé haplo skupiny.

5 Pravděpodobnost výskytu určitého počtu mutací v STR systémech vypočtená pomocí Poissonovy statistické funkce (Turner 2002). Výsledné rozložení pro 30 STR systémů, 50 generací a mutační rychlost 0,2 % (modré sloupce) a 0,4 % (červené sloupce). Se zvyšující se mutační rychlostí (počet mutací vztažený na generaci) vzrůstá pravděpodobnost výskytu určitého počtu mutací. Všechny orig. D. Vaňka

a v Mongolsku dokonce až v r. 2000. V Čechách a na Moravě začali lidé ke svým jménům přidávat příjmení přibližně ve 14. stol., nejprve to ale bylo právo vyšších společenských vrstev. Je zajímavé, že až patent Josefa II. o dědičných příjmeních

ukončil možnost, aby si lidé, jak se občas stávalo, několikrát za život příjmení změnili. Toto nařízení vlastně platí dodnes.

Myšlenka sledovat výskyt příjmení a určité varianty chromozomu Y není nová a objevila se dokonce před objevením polymorfismů DNA. Pravděpodobně první taková studie byla provedena u francouzsko-kanadské rodiny, ve které si příslušníci mužské linie po 12 generací předávali Y chromozom s translokovanou částí autozomu (Schmid 1984). Po provedení genealogické studie bylo zjištěno, že ke změně chromozomu došlo přibližně před 300 lety a všichni nositelé tohoto příjmení jsou potomky francouzského dělníka, který v r. 1665 emigroval do Kanady.

Molekulární biologie a genetika může být použita i pro ověření legend a historických příběhů. Podle bible byl Mojžíšův bratr Áron vybrán za prvního kněze (kohena). Kněžský úřad a pojmenování kohena (cohanim znamenáná v hebrejštině kněz) se přeneslo i na Áronovy děti, a tím začala tradice předávání kněžského úřadu na potomky (dcera kohena mohla být kněžem, ale toto postavení již nemohla přenést na své potomky). Porovnáním Y haplotypů osob s příjmením odvozeným od Kohena (Cohen, Kahan, Mazer, Kaplan, Katz, Cowan, Kagan, Kovacs, Kohn, Kahn, Kunha apod.) byla u více než 60 % osob židovského původu zjištěna blízká příslušnost ke stejnému charakteristickému haplotypu, který sestává ze 6 STR (Skorecki 1997). U kontrolní skupiny nežidovských osob byla shoda pouze u 15 % jedinců. Tyto výsledky byly v době svého vzniku často prezentovány jako důkaz potvrzující pravdivost bible popisující dobu před 3 000 lety, kdy byl postaven první chrám jeruzalémský a kdy kohenové měli vedoucí úlohu v židovské komunitě. Díky malému počtu použitých STR (6) však byly závěry postupem času zpochybňovány a velmi rozšířená byla teorie o větším počtu zakladatelů kněžství (kohenů).

Michael Hammer a Karl Skorecki se v r. 2009 tímto problémem dále zabývali a publikovali v časopise Human Genetics novou studii, kde se jim pomocí rozšířené sady genetických markerů (rozšířený modální haplotyp) podařilo vyřešit fylogenetické umístění charakteristického

haplotypu kohena, definovat jeho geografické rozšíření a vyvrátit tím spekulace o počátečním větším počtu zakladatelů židovského kněžství.

Dalším příkladem použití genetiky pro ověření historických domněnek je korejská studie (Kim 1999). Podle ústně předávané legendy by měla být celá korejská populace potomky jediného muže jménem Tagoon, který žil před 5 000 lety. V tomto případě genetická analýza prokázala, že korejské chromozomy Y náležejí většímu počtu linií.

V České republice probíhá v současné době projekt Genetika a příjmení, který se zaměřuje na studium vazby chromozomu Y na česká příjmení. Pro zahrnutí příjmení do projektu je nezbytné, aby se vyskytovalo v populaci s relativně nízkou četností (100–300 jedinců) a mělo pokud možno jedinečný původ (např. jméno Rotrecll může odkazovat na muže – zakladatele rodu – který žil před několika staletími na určitém místě na Moravě a nosil typický červený kabát, od něhož bylo odvozeno jeho příjmení). V ČR je dnes evidováno více než 60 200 různých příjmení s četností nad pět osob (počet těchto osob je 4 970 000). Pětici nejčastějších příjmení jsou Novák (34 199 výskytů), Svoboda (25 225), Novotný (24 288), Dvořák (22 221) a Černý (17 831), na opačném konci nalezneme s pěti výskytů v naší populaci např. příjmení Andrlíček, Hašourek, Kimlíček, Medák, Pastýř nebo Šlapnička. Celkový počet příjmení v ČR (s výskytem alespoň u jedné osoby) je pak více než 172 500. Většina příjmení s nízkou četností (1–5) patří cizincům, ale za několik generací by měl být zaznamenán jejich nárůst, a tím i rozšíření specifických chromozomů Y. Pro srovnání, ve Velké Británii je v současnosti evidováno více než 1,6 milionu různých příjmení, ale před 130 lety to bylo jen 420 000 (Jobling 2001).

Korelace výskytu určité varianty chromozomu Y a příjmení je komplikována několika faktory (King 2009):

- Některá příjmení mohou mít více než jedno centrum vzniku a pak budou nositelé tohoto příjmení patřit k různým liniím.
- Není možné vyloučit nemanželský vztah, který v případě narození potomka mužského pohlaví povede ke změně linie chromozomu Y. Frekvence výskytu ne-

manželských dětí je v současné evropské populaci na úrovni několika málo procent (1–4 %), v minulosti se podle některých autorů tento počet pohyboval na úrovni 1,3 % na generaci (Sykes 2000).

- Adopce dítěte mužského pohlaví spojená s předáním příjmení pěstouna opět zcela změnila původní rodovou Y-linii.

● Mutace v STR markerech postupem času změnil haplotyp chromozomu Y. Po dlouhé době a při použití rychle mutujících STR markerů bude téměř nemožné určit původní informace uložené v DNA dochází v tomto případě tzv. genetickým driftem (postupnou změnou relativní frekvence výskytu určité alely v populaci).

- Pokud muž přijme příjmení manželky a tato skutečnost není známa (došlo k ní např. před stovkami let), je opět do klanu určitého příjmení zanesena cizí linie chromozomu Y.

Pro studium vazby Y chromozomu na příjmení byly v ČR vybrány skupiny mužů s příjmením, které se vyskytuje u 100–300 osob. Předběžné výsledky u některých „klanů“ potvrdily jedinečné centrum jejich vzniku. Díky zpracovaným rodinným genealogiím a záznamům v Berní rule z r. 1654 lze dokonce vypočítat mutační rychlost STR markerů pro jednotlivé haploskupiny chromozomu Y a zjistit tak, v jaké době žil prapředek – zakladatel rodové Y linie. Pokud se následně vloží haplotyp Y určitého klanu do veřejně dostupné databáze (např. www.ysearch.org nebo www.yhrd.org), je možné v případě běžnějších haploskupin nalézt poměrně velké množství osob se stejným či podobným haplotypem, ale s různým příjmením. Pokud se např. haplotyp osoby nalezené v databázi shoduje s haplotypem vyšetřované osoby ve všech 17 testovaných lokusech, znamená to, že existuje 50% pravděpodobnost společného předka před 10 generacemi a 80% pravděpodobnost před 24 generacemi. Pokud se však zjistí shoda pouze v 16 ze 17 lokusů (rozdíl mezi dvěma haplotypy v jednom z testovaných lokusů nazýváme genetickou vzdáleností 1), časový interval, kdy měli tyto dva muži společného předka, se prodlužuje. V tomto případě měly porovnávané osoby s 50% pravděpodobností společného předka před 25 generacemi. Obecně pak platí, čím vyšší počet testovaných a porovnávaných lokusů, tím přesněji lze určit počet generací ke společnému předkovi.

Pomocí analýzy DNA metodami genetické genealogie můžeme zjistit zajímavé informace o mužských rodových liniích, zejména o místě jejich vzniku, ale třeba i nalézt v současné populaci „bratrance“, se kterými sdílíme nejen stejné či podobné příjmení, ale i stejný nebo podobný chromozom Y. Výsledky genealogicko-genetického zkoumání najdou uplatnění i v jiných vědních oborech, jako je např. historie, archeologie, molekulární biologie nebo forenzní genetika. V druhé části článku se budeme zabývat sledováním mateřských linií pomocí analýzy sekvence mitochondriální DNA.

Více o genetické genealogii najdete např. v časopise Journal of Genetic Genealogy nebo na webových stránkách ISOGG (International Society of Genetic Genealogy).

Pravděpodobnost výskytu určitého počtu STR mutací

