

Co nového v přírodních vědách**Lidský genom:
kam kráčí současný výzkum?**

DAVID ŠMAJS

Uveřejnění pilotní sekvence lidského genomu (v roce 2001) a „úplné“ dokončení jeho sekvence (v roce 2004) otevřelo problém interpretace ohromného množství získaných dat. Nutno podotknout, že sekvenci lidského genomu stále neznáme úplně dokonale, stále je v genomu přes 300 „mezer“, tedy míst, pro která přesnou sekvenci neznáme, a navíc je v „definitivní“ verzi sekvence více než deset tisíc jednonukleotidových chyb. Žádné lidské poznání však není dokonalé a s přibývajícimi sekvenačními projekty se „definitivní“ sekvence stále zpřesňuje.

Brzy po dokončení projektu sekvenace* lidského genomu bylo zřejmé, že se samotnou sekvencí lidského genomu nevystačíme (mimochodem je zajímavé, že se totéž týká také všech dosud sekvenovaných genomů jiných organismů) a že je třeba sekvenovat genomy dalších příbuzných organismů a také dalších a dalších lidí. Jen tak získáme představu o tom, co je v genomech pouhým informačním šumem a co jsou kritické změny rozhodující o zdraví a nemoci a také o podstatě lidství.

Současný výzkum lidského genomu probíhá tedy nejméně třemi hlavními směry:

1. Mapování lidského genomu z hlediska identifikace funkčních oblastí.
2. Definice rozdílů v genomu různých lidí.
3. Porovnání sekvence lidského genomu a dalších příbuzných genomů, tj. např. genomů šimpanze nebo neandertálce.

**1. Mapování lidského genomu
z hlediska identifikace funkčních oblastí**

Jedním ze základních směrů studia lidského genomu je přesné mapování míst v genomu, která mají nějakou biologickou funkci. To je nezbytným předpokladem k pochopení struktury i funkce genomu jako celku. Jde zejména o nalezení a lokalizování všech genů kódujících bílkoviny, genů kódujících různé typy RNA, celé řady sekvencí s regulační funkcí (např. promotorů, sekvencí využívaných pro regulaci transkripce atp.) a o nalezení míst v genomu, která se podílejí na funkcích dalších (např. na replikaci DNA). S trochou nad-

* Pod pojmem sekvenace genomu rozumíme experimentální postup, který vede k určení pořadí (sekvence) všech bází v deoxyribonukleotidové kyselině (DNA) daného organismu. Genom je definován jako úplná genetická informace daného organismu, kterou lze vyjádřit jako kompletní sekvenci DNA.

sázky lze říci, že právě tohle je to, co jsme si od projektu sekvenace lidského genomu slibovali hned od jeho počátku. Zdá se však, že k tomuto cíli bude třeba ještě ujit značný kus cesty, obzvláště proto, že pouhý pilotní projekt (systematicky mapující pouhé 1 % lidského genomu) přinesl velká překvapení. Jak totiž ukázal projekt ENCODE (*Encyclopedia of DNA Elements*), transkripčně aktivní jsou také podstatné části genomu (molekuly DNA), o kterých se dosud soudilo, že jsou nefunkční a jsou pouhým „balastem“. Přitom tato DNA tvoří okolo 98 % veškeré lidské DNA. Znamená to tedy, že i když RNA kódovaná touto „nefunkční“ DNA není přepisována do bílkovin, tvoří se v takovém množství a na tak rozsáhlé části DNA, že nějakou její funkci lze oprávněně očekávat. Obecně lze předpokládat, že tato negenová DNA se podílí na regulaci genové exprese (na regulaci zapínání a vypínání genů). Zatím však pro toto tvrzení nemáme žádné důkazy a bylo by poctivější přiznat, že nemáme ani tušení, jakou funkci může tato RNA mít.

Dalším příkladem projektu, který si tentokrát klade za cíl zmapovat genetické změny vyskytující se u nádorových onemocnění, je projekt *The Cancer Genome Atlas* (TCGA). Jde o vysokokapacitní sekvenaci (tj. sekvenaci téměř celého genomu) mnoha nádorových vzorků jednoho typu nádoru od mnoha různých pacientů. Nalezené genetické změny, které se budou vyskytovat ve vzorcích z většiny pacientů, pak bude možno dát do souvislosti s vývojem příslušného typu nádoru. Tohle vše je navíc plánováno pro mnoho typů nádorů. Jde tedy o jakýsi frontální útok na odhalení genetického pozadí nádorových onemocnění.

2. Definice rozdílů v genomu různých lidí

Publikace pilotní sekvenace lidského genomu (v roce 2001) se opírala o sekvenaci vzorků DNA několika dárců, a bylo tedy možné odhalit i odlišnosti v genomové sekvenci těchto individuálních (a anonymních) dárců. To ukázalo na potřebu definovat místa v genomu člověka, jimiž se různí jedinci navzájem liší. Pro zmapování míst, ve kterých se lidé mezi sebou navzájem liší, vznikl mezinárodní projekt HapMap. Vzorky DNA pro projekt HapMap pocházely od 269 lidí z Afriky, Japonska, Číny a USA. Bylo identifikováno okolo 10 milionů míst, ve kterých se lidé čtyř různých populací liší nejčastěji. (To přibližně odpovídá shodě 99,9 % mezi kterýmikoliv dvěma osobami.) Tato místa se označují jako SNP (jednonukleotidové polymorfismy). Platí přitom, že jednotlivé SNP se dědí po určitých blocích (haplotypech). Z toho vyplývá možnost definovat genom individuálního člověka jako kombinaci určitých haplotypů a pro zjištění genotypu konkrétní osoby tedy není třeba mapovat všech 10 milionů jeho genomových míst, ale stačí genotypovat jen 300.000 – 600.000 klíčových SNP. Možnost genotypovat individuální jedince bude mít zcela zásadní význam pro hledání korelace mezi konkrétním genotypem a výskytem chorob, které jsou (alespoň zčásti) geneticky podmíněné. Přitom je zřejmé, že genetická komponenta je jedním z rizikových faktorů téměř všech chorob, včetně onemocnění kardiovaskulárních, nádorových, metabolických, autoimunitních a dalších. Dobrým příkladem použití těchto genomových me-

tož je objev polymorfismu v genu kódujícím jednu bílkovinu – komplementový faktor H (CFH). Jak ukázaly pokusy asociačních studií, genový polymorfismus vedoucí k záměně tyrosinu v pozici 402 proteinu CFH za histidin vede k výraznému zvýšení rizika věkem podmíněné makulární degenerace sítnice. Toto onemocnění postihuje ve vyspělých zemích 10-15 % populace starší 60 let a je nejčastější příčinou slepoty ve stáří. Na základě znalostí získaných pomocí genomiky je poměrně snadné navrhnout diagnostický test pro detekci aminokyseliny v pozici 402 proteinu CFH a tím zjistit i riziko makulární degenerace. Do budoucna lze snad i nalézt látky, které mohou aktivitu komplementu selektivně ovlivňovat.

Druhou možností jak zjistit sekvenci lidského genomu konkrétní osoby (pacienta) je rovnou příslušný genom sekvencovat. Loni vyšla práce popisující kompletní genom první konkrétní osoby. Jedná se o genom J. Graiga Ventera, zakladatele firmy Celerá Genomics Corporation, která konkurovala mezinárodnímu projektu sekvenace lidského genomu a která částečně veřejně publikovala výsledky sekvenace lidského genomu v roce 2001. Zatím ovšem jde o proceduru, která je příliš zdouhavá a drahá. Teprve budoucnost ukáže, zdali bude kompletní vyšetření genomové sekvence pacienta patřit do baterie možných diagnostických testů.

Jaké si tedy slibujeme výhody z mapování SNP? Porovnáním genetického profilu pacientů s určitým typem onemocnění a genetického profilu zdravých osob bude možné zjistit, které genetické varianty (SNP) přispívají k vývoji určitého typu onemocnění. Přesné mapování SNP přispívajících k vývoji onemocnění kardiovaskulárních, nádorových, metabolických, autoimunitních atp. bude možno použít pro odhadnutí genetického rizika u konkrétního pacienta a event. preventivně upravit faktory prostředí, které přispívají k vývoji té které choroby (dieta, životospráva, preventivní podávání určitých léků atp.). Navíc tyto studie umožní předpovídat biologickou funkci celé řady genů, o jejichž úloze v životě buňky a celého organismu zatím nemáme nejmenší tušení. Identifikace a detailní popis proteinů, které se liší od proteinů zdravých jedinců a které vedou k predispozici k určité chorobě, umožní vývoj nových druhů léčiv, které budou rozpoznávat a ovlivňovat přímo tyto cíle.

Dlouho se také ví, že odpověď na léky a obecně na léčbu se individuálně liší od jedince k jedinci, přičemž tato variabilita je z velké části podmíněna geneticky. Do budoucna tedy očekáváme, že terapie bude takřka „šitá na míru“ pro každého z nás.

3. Porovnání sekvence lidského genomu a dalších příbuzných genomů, tj. např. genomů šimpanze nebo neandertálce

Přečtení genomu šimpanze překvapilo vědeckou komunitu tím, jak je genom šimpanze příbuzný s genomem člověka; genomy člověka a šimpanze vykazují asi 99% shodu v sekvencně homologických (příbuzných) částech obou

genomů, celkově je pak shoda (i se započtením sekvencí unikátních pro člověka i šimpanze) asi 96 %. Navíc téměř třetina genů (29 %) kóduje zcela shodné proteiny jak u člověka, tak u šimpanze a drtivá většina ostatních kódovaných proteinů se liší jen záměnou jedné aminokyseliny. Pro představu je příbuznost na genomové úrovni mezi člověkem a šimpanzem asi 10 krát větší než příbuznost mezi myší a potkanem. Nutno však podotknout, že neexistuje přímý vztah mezi fenotypovou odlišností organismů a jejich odlišností sekvencí (genotypovou). Sekvenční odlišnost člověka i šimpanze tak zhruba odpovídá odlišnosti dvou fenotypově velmi podobných druhů myší, myši domácí (*Mus musculus*) a myši středozevní (*Mus spretus*). Tyto údaje jsou velmi zajímavé v souvislosti s různými onemocněními, jejichž průběh se u lidí a nejbližších primátů liší (viz tabulka 1). Poznání genetických příčin odlišného průběhu onemocnění by umožnilo lépe nahlédnout do podstaty těchto onemocnění a do budoucna by to umožnilo cíleně vyvíjet jejich odpovídající terapie.

Tabulka 1. Některé rozdíly mezi lidmi a dalšími primáty ve výskytu a průběhu některých chorob

Výskyt chorobného stavu nebo komplikací	Lidé	Primáti
<i>Potvrzené rozdíly</i>		
progrese HIV infekce v AIDS	častá	vzácná
symptomatologie chřipky A	středně těžká až závažná	lehká
komplikace u hepatitidy B/C	středně těžké až závažné	lehké
citlivost k malárii (<i>P. falciparum</i>)	citliví	rezistentní
<i>Předpokládané rozdíly</i>		
rozvoj morbus Alzheimer	kompletní	částečný
koronární ateroskleróza	častá	vzácná
karcinomy	časté	vzácné

S nástupem nové technologie sekvencování DNA – pyrosekvence – se otevřela možnost sekvencovat genomy již vyhynulých organismů. Kromě sekvence genomu mamuta je v běhu projekt sekvenace nám nejbližšího druhu, člověka neandertálského (*Homo neanderthalensis*). Neandertálci sice vyhynuli přibližně před 30.000 tisíci lety, z jejich kostních pozůstatků je však možné získat dost DNA pro sekvenaci jejich genomu. Zvláště vhodnou k tomuto účelu byla shledána kost Vi-80, která byla nalezena roku 1980 v jeskyni Vindija v Chorvatsku; její stáří je datováno na asi 38 tisíc let. Sekvencování takto starých vzorků je komplikováno faktem, že DNA je degradována na krátké kousky a že její nukleotidové báze podléhají změnám, které vedou k jejich falešnému čtení. I přes tyto komplikace se po přečtení prvního milionu párů bází nean-

dertálce podařilo odhadnout sekvenční příbuznost se současným člověkem. Tato podobnost představuje 96 procentní shodu změn mezi současným člověkem a neandertálcem vzhledem ke změnám popsaným mezi člověkem a šimpanzem. Takto komplikovaně bylo třeba vyjádřit sekvenční příbuznost mezi člověkem a neandertálcem právě proto, že časem dochází na bazích DNA ke změnám, které vedou k jejich chybnému čtení. Podle těchto odhadů vychází sekvenční příbuznost mezi současným člověkem a neandertálcem vyšší než 99,5 % a tomu odpovídá odhadovaný výskyt posledního společného předka obou před 520 tisíci lety. Existence společného předka člověka a šimpanze se pak předpokládá v době před 6,5 miliony let. Genetické rozdíly mezi současným člověkem (*Homo sapiens*) a vybranými organismy na celogenomové úrovni a odhadovaná evoluční vzdálenost od posledního společného předka jsou uvedeny v tabulce 2. Je překvapivé, jak vysoká je sekvenční příbuznost i značně odlišných druhů savců; přitom přibližná sekvenční příbuznost na úrovni DNA odpovídá naší intuitivní představě o příbuznosti organismů.

Tabulka 2. Genetické rozdíly mezi současným člověkem (Člověk moudrý, *Homo sapiens*) a vybranými organismy na celogenomové úrovni a odhadovaná evoluční vzdálenost od posledního společného předka

Druh	Odhadovaná doba uplynulá od výskytu společného předka	Přibližná příbuznost na úrovni DNA
Člověk moudrý (<i>Homo sapiens</i>), kterýkoliv jiný jedinec		~ 99,9 %
Člověk neandertálský (<i>Homo neanderthalensis</i>)	0,5 mil. let	~ 99,5 %
Šimpanz učenlivý (<i>Pan troglodytes</i>)	6 mil. let	~ 96 %
Makak rhesus (<i>Macaca mulata</i>)	25 mil. let	~ 93 %
Potkan obecný (<i>Rattus norvegicus</i>)	75 mil. let	~ 40 % genomu je sekvenčně příbuzných (nikoliv shodných)

Je zajímavé sledovat, jak se s každým dalším přečteným genomem vždy ukáže, že by bylo potřeba – pro lepší pochopení rozdílů mezi genomem člověka a genomem jeho nejbližších příbuzných – sekvencovat genom další, takřkajíc kontrolní. A tak jsme po přečtení lidského genomu svědky sekvencování stále dalších a dalších genomů, přičemž jsme od pochopení genomových rozdílů mezi organismy – a tedy i podstaty rozdílů mezi člověkem a primáty –

stále daleko. Nápadně to připomíná příklady z jiných odvětví vědy, kdy pochopení určitého fenoménu předchází extenzivní fáze akumulace dat a výsledků včetně všeobecného tápání. Po této fázi pak skokově přichází rychlý vývoj a následující období bývá zpravidla velmi vědecky plodné. Ale na to si musíme v genomice ještě chvíli počkat.

Redakci Universitas došlo

pokračování ze str. 6

- Mapa Moravy od J. A. Komenského.** Masarykova univerzita v Brně, 2007
- X. mezinárodní kolokvium o regionálních vědách.** Sborník příspěvků z kolokvia v Pavlově 20. – 22. června 2007. Masarykova univerzita v Brně, 2007
- Německý jazyk a literatura ve výzkumu a výuce.** Editoři: Tomáš Kaňa, Hana Peloušková. Masarykova univerzita v Brně, 2007
- Nezaměstnanost – technologické a sociální proměny práce.** Božena Buchtová, ed. Sborník příspěvků. Masarykova univerzita v Brně, 2007
- PFAMEI 2007.** Právní a finanční aspekty mezinárodní ekonomické integrace. Sborník abstraktů příspěvků I. ročníku mezinárodní konference. Masarykova univerzita v Brně, 2007
- Raport Annuel. Résumé.** Défenseur public des droits République tchèque 2006. Masaryk University Brno, 2007
- Role nejvyšších soudů v evropských ústavních systémech – čas pro změnu?** Ed. Vojtěch Šimíček. Masarykova univerzita v Brně, 2007
- Sborník Abstrakt. XV. Frejtkovy dny s mezinárodní účastí.** Brno, 14. – 15. června 2007. Masarykova univerzita v Brně, 2007
- SCO 2007.** Sborník 4. ročníku konference o elektronické podpoře výuky. Eds. Petr Sojka, Martin Kvida. Brno, 30. a 31. května 2007. Masarykova univerzita v Brně, 2007
- Sborník prací Filozofické fakulty Brněnské univerzity. Řada archeologická, M 10-11.** Masarykova univerzita v Brně, 2007
- Sborník prací Filozofické fakulty Brněnské univerzity. Řada hudebněvědná, H 41.** Masarykova univerzita v Brně, 2007
- Sborník prací Filozofické fakulty Brněnské univerzity. Řada literárněvědné slavistiky, X 10.** Masarykova univerzita v Brně, 2007
- Scripta. Geology. Volume 35.** Masaryk University Brno, 2007
- Spiritualita.** Ed.: Hana Babyrádová, Jiří Havlíček. Masarykova univerzita v Brně, 2007
- Spisy Právnické fakulty Masarykova univerzity v Brně. Regulace, deregulace, autoregulace.** Ed.: Josef Bejček. Masarykova univerzita v Brně, 2007
- Spisy Právnické fakulty Masarykovy univerzity. Introduction to ict law.** Ed. Radim Polčák. Masarykova univerzita v Brně, 2007
- Petr Suchý, Strategie zadržování v americké zahraniční politice. Vznik a vývoj koncepce 1945-1953.** Masarykova univerzita v Brně, 2007
- Třináct let po. Brněnské texty k slovakistice IX.** Eds.: Ivo Pospíšil, Miloš Zelenka, Anna Zelenková. Masarykova univerzita v Brně, 2007
- Veřejná a soukromá řešení dopadů živelních pohrom v ČR. Konference s mezinárodní účastí 6. 6. – 7. 6. 2007.** Masarykova univerzita v Brně, 2007
- Visions of Canada. Canadian Studies in Europe. Volume 6.** Masaryk University Brno, 2007
- Výchova ke zdraví II. Sborník prací studentů a učitelů katedry rodinné výchovy a výchovy ke zdraví Pdf MU.** Ed.: Marie Havelková, Jitka Reissmannová. Masarykova univerzita v Brně, 2007
- Výuka ekonomických předmětů na právnických fakultách 2007.** Sborník z konference. Masarykova univerzita v Brně, 2007