

Moderní metody sekvenování DNA

Termínu sekvenování DNA, které se také označuje jako sekvenace nebo někdy sekvenování DNA, odpovídá asi nejlépe české čtení DNA. Pomocí této metody určujeme pořadí nukleotidů v molekule deoxyribonukleové kyseliny. Pro sekvenování DNA, které patří v současnosti mezi nejrozšířenější způsoby analyzování biologického materiálu, bylo vyvinuto poměrně velké množství různých metod a technik. Používané moderní technologie pro sekvenování DNA můžeme rozdělit na tradiční Sangerovu metodu versus metody sekvenování tzv. druhé generace. Právě metody sekvenování druhé generace způsobily do jisté míry revoluci v moderní biologii, a to jak z pohledu vědeckých otázek, na něž lze hledat odpovědi, tak z pohledu zcela nových technik výzkumné práce. Následující odstavce shrnují co nejpřístupnější formou informace, jak sekvenování DNA funguje i k čemu se dnes může použít. Pro lepší pochopení jsou na webové stránce Živa u článku uvedeny odkazy na instruktážní videa, která celý proces zobrazují v animované podobě.

Sangerovo sekvenování

Informace v molekule DNA je zapsána pomocí čtyř typů nukleotidových bází neboli nukleotidů – adeninu, cytozinu, thyminu a guaninu – pro něž se používají zkratky A, C, T, G. Molekula DNA, která je především řetězcem těchto nukleotidů, se v buňkách nachází ve dvouřetězcové formě. Jednotlivá vlákna DNA jsou podélně spojena a párování nukleotidů má přísné pravidlo označované jako komplementarita: adenin se vždy páruje s thyminem, zatímco partnerem cytozinu je vždy guanin. Díky vzájemné komplementaritě obou řetězců DNA jsou organismy schopné pomnožit (zkopírovat) si vlastní genetickou informaci. Během procesu zvaného replikace se totiž dvouřetězcová DNA rozdělí na dvě jednoduchá vlákna a protein DNA polymeráza vytvoří ke každému oddělenému řetězci (templátu) nový komplementární řetězec tak, aby se A po každé párovalo s T a C vždy s G. Tento popis představuje samozřejmě jen velmi zjednodušenou verzi celého procesu, neboť replikace DNA je extrémně složitý proces a sled reakcí obstarává kromě DNA polymerázy i řada dalších proteinů. Pro naše potřeby však tato zjednodušená představa postačí.

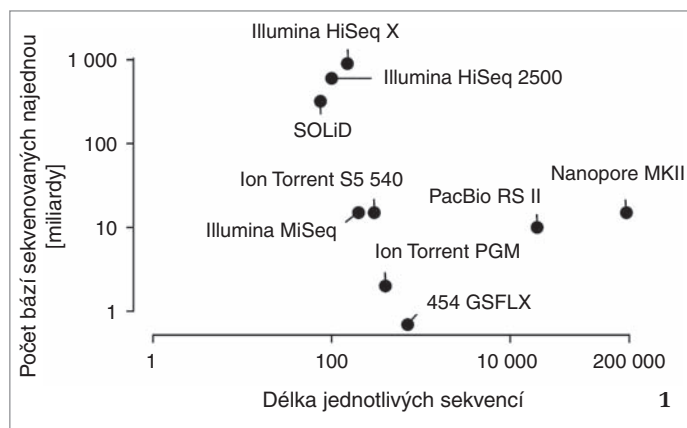
Naprostá většina technik sekvenování DNA se zakládá na metodách zjišťujících pořadí postupně přidávaných nukleotidových bází při syntéze nového řetězce DNA, který vzniká podle stávající předlohy – templátu. V r. 1977 objevil Frederick Sanger (1918–2013) se svými kolegy metodu sekvenování DNA, jež umožnila čtení velmi dlouhých řetězců, a to až 1 000 bází na jednu sekvenační reakci. Dnes se metodě říká podle jejího objevitele Sangerovo sekvenování. A Fredericku Sangerovi přinesl tento objev již druhou Nobelovu cenu. Zařadil se tak po bok dalších dvěma vědcům, Marii Curie-Sklodovské a Johnu Bardeenovi, kterým se za jejich vědecké objevy podařilo získat toto nejvyšší možné ocenění dvakrát.

Sangerova metoda se stala víceméně jedinou běžně používanou metodou sekvenování na mnoho let. Opakovaně je při ní syntetizován nový řetězec DNA podle stejného templátu, avšak při každém opakování dochází k přerušení syntézy nového řetězce tím, že se do něj náhodně přiřadí modifikovaná nukleotidová báze (dideoxynukleotid), která nedovolí DNA polymeráze pokračovat v syntéze nově vznikající

cího vlákna. Proces si můžeme představit jako postupné řetězení kostek stavebnice, kdy nové kostky jsou vybírány z krabice, která kromě normálních obsahuje i kostky pozměněné. Ty sice vypadají stejně jako běžná součást stavebnice, a dokonce je lze navázat na kostky v řetězci, ale na ně samotné již žádnou další kostku napojit nelze. Jejich zařazením se tedy stavba zastaví. Při Sangerově sekvenování však na sobě každá modifikovaná báze (v našem příkladu kostka, která ukončí postupné řetězení) navíc nese specifickou fluorescenční barvu, pomocí níž poznáme, o jakou bázi jde. Náhodným přerušováním syntézy nově vznikajícího komplementárního řetězce DNA (podle identického templátu) vznikají různě dlouhé molekuly. Při dostatečném počtu opakování vznikne směs molekul, které délkou odpovídají všem možným pozicím – ve směsi nalezneme molekuly dlouhé 10, 11, 12, 13, 14 až 1 000 bází. Na konci každé takové molekuly se nachází fluorescenčně označený dideoxynukleotid (nefunkční ukončovací kostka stavebnice), a protože každé z těchto čtyř modifikovaných písmen „DNA abecedy“ má jinou fluorescenční barvu, snadno poznáme, o který nukleotid jde. Na základě velikosti se molekuly DNA (zakončené dideoxynukleotidem) seřadí od nejkratší po nejdelší

1 Graf ukazující rozdíly v délce jednotlivých sekvencí a počtu sekvenovaných bází pro metody sekvenování druhé a třetí generace. Pro dnes nejpoužívanější metody Illumina a Ion Torrent (blíže v textu) ukazuje graf několik různých verzí daných technologií. Je důležité zmínit, že jak Illumina MiSeq, tak Ion Torrent PGM zastupují tzv. stolní sekvenátory, neboť jsou poměrně levné (stojí řádově několik milionů Kč) a lze je používat i v rámci jediné laboratoře. Ostatní sekvenační přístroje se používají především v rámci větších sekvenačních center. Orig. M. Kolísko, upraveno podle <https://flxlexblog.wordpress.com/2014/06/11/developments-in-next-generation-sequencing-june-2014-edition/>

2 Sběrání jednotlivých buněk. Izolace jediné buňky druhu *Pseudotriconympha* sp., což je doposud nekultivovatelný prvok žijící v trávicím traktu termitů – sběr jednotlivých buněk tedy představuje jedinou možnost, jak získat sekvenční data z tohoto organismu. Buňka je „nasáta“ do skleněné mikropipety a přenesena např. do mikroskopu. Foto P. J. Keeling



a následně se s využitím jejich odlišného fluorescenčního značení přečte výsledná sekvence, tedy pořadí bází.

Mezi hlavní výhody Sangerova sekvenování patří značná délka čtených úseků DNA, které se dají sekvenovat jedinou reakcí, a také vysoká přesnost čtení, při které vzniká jen minimum chyb (špatně přečtených písmen).

Hlavní metodické úskalí této techniky spočívá v tom, že v rámci celého procesu dochází k sekvenování pouze jednoho úseku DNA (na rozdíl od metod druhé generace – viz dále). Další nevýhodou je poměrně vysoká cena a nízká rychlost, což se projeví zejména v situaci, kdy je potřeba zjistit sekvenci velkého množství DNA (např. celého genomu nebo malého množství genů, ale u velkého počtu různých organismů).

Metody druhé generace

Sangerovo sekvenování se běžně používá dodnes, ale v posledním desetiletí se objevily nové sekvenační metody tzv. druhé generace (Next Generation Sequencing, NGS). Většina těchto postupů opět využívá syntézu DNA podle templátu, ale na rozdíl od Sangerovy metody jsou schopny detekovat přidávání bází jednu po druhé a zároveň sekvenovat tisíce až miliony rozdílných molekul DNA najednou.

Jejich hlavní nevýhodou oproti Sangerovu sekvenování je krátká maximální délka výsledných sekvencí, která se dnes obvykle pohybuje zhruba od 100 až po 500 bází (Sangerova metoda nabízí až 1 000 bází). Nevýhodou představuje i menší přesnost a častější chyby při čtení DNA.

U všech těchto metod je DNA nejdříve „nastříhána“ na relativně krátké části a na jejich konce je přilepen adaptér – velmi krátká molekula DNA o přesně dané sekvenci. Slouží k následnému navázání (přichycení) sekvenovaného úseku DNA na pevný povrch. Takto upravená DNA se říká sekvenační knihovna. Po uchycení DNA pomocí adaptéru na povrchu, na kterém bude docházet k sekvenaci, je každý řetězec DNA namnožen, čímž vznikne skupina neboli klastr identických molekul DNA koncentrovaných v jednom místě (sekvenace adaptéru je komplementární ke krátkým řetězcům DNA uchyceným na sekvenačním povrchu). Tato koncentrace posílí výsledný signál (viz dále), což umožní jeho zachycení kamerou, neboť signál z pouhé jedné molekuly DNA by nebyl dostatečně silný.

● 454 sekvenování a Ion Torrent

První prakticky použitelnou a komerčně úspěšnou metodou sekvenování druhé generace bylo tzv. 454 sekvenování. Objevil ho Jonathan Rothberg a publikoval v r. 2005, v době, kdy byl na rodičovské dovolené. Následné uvedení metody do praxe představovalo obrovský vědecký i komerční úspěch. Nejmodernější verze dokáže analyzovat více než milion molekul DNA najednou a délka každé jednotlivé sekvence se pohybuje okolo 700 až 1 000 bází (1 000 se uvádí jako maximum). Během 454 sekvenování je molekula DNA nejdříve přichycena na malou „kuličku“, na jejímž povrchu se postupně namnoží, až kuličku zcela pokryjí identické molekuly DNA.

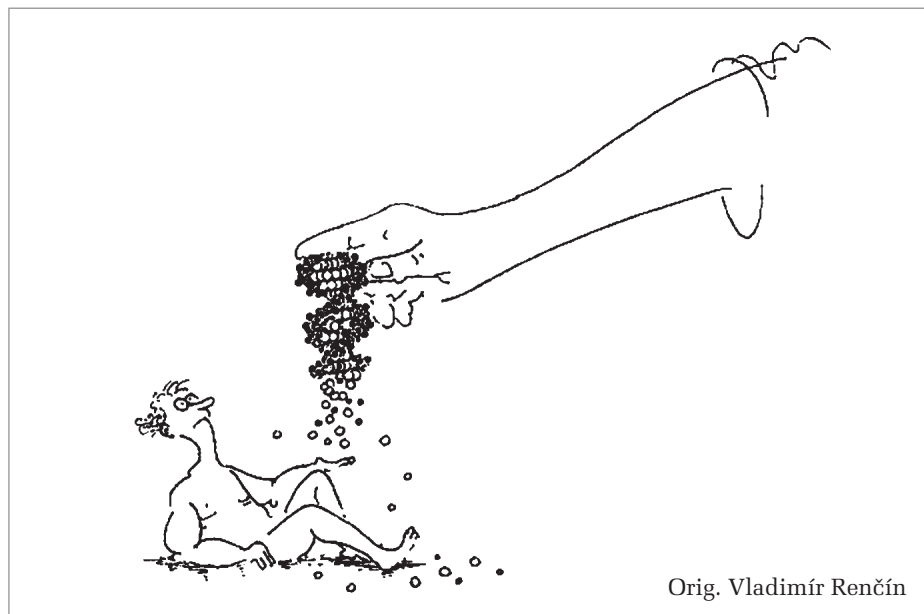
Kulička i s DNA je následně vložena do jedné z milionů komůrek destičky, kde probíhá sekvenační reakce na principu pyrosekvenování. Pyrosekvenování se zakládá na skutečnosti, že během vložení každé nové báze do rostoucího řetězce DNA se uvolní molekula zvaná pyrofosfát (proto pyrosekvenování). Uvolněný pyrofosfát se posléze stane součástí několika na sebe navazujících enzymatických reakcí, na jejichž konci čeká enzym luciferáza (nazvat enzym po pánu pekel je pro humor molekulárních biologů dost příznačné). Ten vydá světelný záblesk, jenž lze zachytit vysoce citlivou kamerou. Při 454 sekvenování je v určitém momentě přidán do reakční směsi vždy pouze jeden typ báze a v okamžiku, kdy je tato báze vložena do rostoucího řetězce DNA, dojde přes uvolněný pyrofosfát a luciferázu ke světelnému záblesku. Pokud je do rostoucího řetězce DNA zařazeno několik stejných bází za sebou, např. když DNA molekula templátu obsahuje sekvenci AAA a je tedy přidáno třikrát T, vyzáří se třikrát více světla než v případě přiřazení jednoho T. Kamera snímá celou destičku a podle toho, která komůrka se rozsvítí, pozná, kde proběhlo přidání báze, a podle intenzity světla kolik bází bylo přidáno najednou (viz obr. 3). Čtyři typy nukleotidů jsou do směsi přidávány jeden po druhém a mezi jednotlivými kroky dochází k odstranění přebytečných nukleotidů. Tím se zajistí, že v reakční směsi je vždy přítomen jediný typ nukleotidu. Počítač následně analyzuje záznam a podle světelných signálů z každé komůrky vytvoří výsledné sekvence odpovídající templátu DNA v jednotlivých komůrkách.

Na podobném principu jako pyrosekvenování stojí i sekvenování metodou Ion Torrent. Ion Torrent však neměří množství uvolněného pyrofosfátu po připojení nové báze do DNA řetězce pomocí luciferázy, ale pH v reakční směsi. Opět podle intenzity změny pH lze poznat, kolik bází bylo přiřazeno (pH roztoku se mění s každou přidanou bází o 0,02 jednotky). Právě proto, že jak 454, tak Ion Torrent sekvenování spoléhají na sílu signálu k zjištění informace, kolik bází bylo přidáno najednou, mají obě tyto metody problém se čtením delších

homopolymerních řetězců složených z jediného typu nukleotidů, např. sekvence AAAAAAAAAA. V takovém případě však ani jedna z metod nebude schopna přesně rozlišit, zda bylo přidáno 9, 10 či 11 adeninů. Nejčastější čtecí chybou 454 i Ion Torrent sekvenování je právě chybějící báze, nebo naopak báze navíc (četnost chyb u obou přístupů se uvádí kolem 1 %).

● Illumina

V současnosti nejpoužívanější metodou sekvenování je komerčně velice úspěšná technologie od společnosti Illumina. Dnes stojí na prvním místě v počtu najednou sekvenovaných bází (až 900 miliard). Nevýhoda spočívá v poněkud kratších sekvencích pohybujících se pouze kolem stovky bází. Při sekvenování metodou Illumina jsou pomocí adaptéru jednotlivé nastříhané molekuly DNA přichyceny na malou destičku. Každá molekula DNA se pak opakovaně namnoží, až na destičce vznikne mozaika milionů klastrů, přičemž každou skupinu tvoří vzájemně identické molekuly. Vlastní sekvenační proces pak využívá podobného mechanismu jako Sangerovo sekvenování, kdy jsou do rostoucího řetězce zařazeny báze s navázanou fluorescenční barvou (každé písmeno má specifickou barvu), které syntézu zastaví. Oproti Sangerovu sekvenování je tato blokáce syntézy vratná. Po přečtení nově přidané báze vysoce citlivou kamerou dojde k enzymatickému odstranění jak fluorescenčního značení, tak blokující části molekuly a může proběhnout další kolo reakce, tedy přidání další báze. Kamera v každém jednotlivém kole syntézy řetězce DNA snímá signál z celé destičky a podle rozdílné fluorescence pozná, jaké písmeno bylo přidáno u každé z milionů skupin. Počítač pak opět analyzuje záznam krok po kroku a podle toho, jak se mění fluorescenční signál v rámci každé skupiny (skládající se z identických molekul DNA), zrekonstruuje přesnou sekvenci molekul DNA v příslušné skupině. Illumina má velmi vysokou přesnost čtení DNA (uvádí se mezi 99 až 99,9 %). Nejčastější chybou je špatně přečtená báze, záměna jednoho nukleotidu, např. jako když v textu omylem přečteme jiné písmeno.



Orig. Vladimír Renčín

Tab. 1 Shrnutí parametrů diskutovaných sekvenačních technik. Data převzata z publikace S. Goodwina a kol. (2016)

Technologie	Délka sekvencí	Počet sekvencí (v milionech)	Počet sekvenovaných bází (v milionech)	Přesnost čtení
454	400–1000	1	až 700	99 %
Ion Torrent	200–400	0,4–80	až 1500	99 %
Illumina	36–300	12–4000	až 900 000	99,9 %
SOLiD	50–75	700–1400	až 320 000	99,9 %
PacBio	20 000	0,055	až 1000	87 %
Nanopore	až 200 000	0,1	až 1500	88 %
Sangerova metoda	~ 1000	sekvence jedné molekuly DNA	0,001	99,9 %

● SOLiD sekvenování

Jako poslední metodu sekvenování druhé generace uvedeme SOLiD sekvenování (Sequencing by Oligonucleotide Ligation and Detection), která je dnes relativně vzácně používanou metodou. Tato technologie na rozdíl od všech předchozích nespolehá na enzym DNA polymerázu, nýbrž na enzym zvaný ligáza, která umí připojit části jednořetězcových molekul DNA k stávajícímu řetězci DNA. Zjednodušeně lze říci, že při SOLiD sekvenování se k templátu přidávají kusy DNA, tzv. sondy, které začínají všemi možnými dvojkombinacemi čtyř základních nukleotidů, tedy 16 různých sond. Každá sonda také nese jednu ze čtyř fluorescenčních značení, což znamená, že čtyři různé dvojkombovací nukleotidů jsou označeny stejnou fluorescenční značkou. V každém kroku pak enzym ligáza připojí k rostoucímu novému řetězci sondu nesoucí dvojkombovací nukleotidů odpovídající templátové DNA a snímač přečte její fluorescenční značení, které je poté odstraněno a může se připojit další sonda. Aby došlo k přečtení kompletní sekvence, je jedna templátová molekula čtena opakovaně, ale „začátek“ čtení se vždy posune o jeden nukleotid, a každá báze je tak přečtena několikrát. Z kombinace znalosti sekvence adaptéru, kterým sekvenovaná DNA začíná, a výsledného signálu čtyř fluorescenčních barev, jak jdou po sobě v jednotlivých čteních, lze odvodit výslednou DNA sekvenci. SOLiD sekvenování má podobný výstup jako Illumina a produkuje rovněž krátké sekvence (maximálně 100 bází). SOLiD má problémy se čtením palindromických úseků (sekvencí shodných u obou komplementárních řetězců), jež mohou vytvářet smyčku v templátové DNA, která je pak nepřístupná pro nasednutí sondy.

Metody třetí generace

Je důležité zmínit, že v posledních několika letech se začaly využívat i metody sekvenování třetí generace. Na rozdíl od metod druhé generace (454 nebo Illumina) není DNA templát před sekvenováním nijak namnožen, a tak dochází ke čtení signálu z jediné (původní) molekuly. Zatím jsou komerčně dostupné dva typy sekvenování třetí generace, avšak vzhledem k technickým omezením se ještě neuplatňují v masovém měřítku. Je však vysoce pravděpodobné, že v příštích několika letech budou patřit mezi naprosto standardní techniky.

Společnost Pacific Biosciences již uvedla první komerčně úspěšnou metodu třetí generace (PacBio), která k detekci výsledné sekvence také využívá fluorescenční zna-

čené nukleotidy. Novinkou je vysoká citlivost, jež umožňuje v reálném čase zaznamenávat zařazení byť jediného nukleotidu do (jediného) řetězce DNA.

Druhá metoda byla uvedena společností Oxford Nanopore. Zde je jednořetězcová molekula DNA protahována mikroskopickým pórem na syntetické membráně. Protože každá DNA báze má trochu jiný tvar, dochází při protahování k odlišnému „ucpání“ póru a citlivé snímače přístroje dokážou zjistit, jak výrazně je pór v danou chvíli „zaplněn“, a tedy jaká báze v daný okamžik membránou prochází.

Obě metody sekvenování třetí generace jsou schopné přečíst 10 i více tisíc bází v rámci jedné analyzované molekuly DNA a produkuje tak velmi dlouhé sekvence, což je výhodné zejména pro sekvenování genomů (souboru veškeré genetické informace organismu zapsané v DNA). Na druhou stranu mají ale obě metody vysokou frekvenci chyb, v současnosti se pohybuje okolo 10 až 15 %. Významnou předností sekvenátorů Oxford Nanopore je velikost. Jde o malý kapselní přístroj, který se přes USB dá připojit k notebooku a nabízí tak příruční využití kdekoli na světě. Konkrétním příkladem jeho použití bylo sekvenování viru Ebola přímo v terénu během epidemie krváivé horečky způsobené tímto virem v západní Africe v r. 2015.

Využití metod sekvenace druhé generace

● Genomika a transkriptomika

Nové sekvenační technologie do jisté míry změnilly tvář moderní biologie, a to nejen díky množství dat, které lze relativně levně získat v řádu hodin až několika málo dnů. Významnou roli sehrála i jejich cena, která se mnohonásobně snížila právě díky čtení milionu molekul DNA najednou. V současnosti je možné poměrně snadno sekvenovat celé genomy organismů, a to v rámci jediné laboratoře, namísto konsorcií desítek a stovek pracovišť, jak tomu bylo v minulosti (viz např. Živa 2016, 5: 203–206). Sekvenování genomů se tak dočkalo obrovského rozmachu, např. fylogenetika (fylogeneze založená na analýze jednoho nebo několika genů) se postupně mění na fylogenomiku (fylogenezi založenou na analýze celých genomů, viz např. Živa 2016, 4: 175–178). Jen pro zajímavost lze uvést, že sekvenování lidského genomu pomocí tradiční Sangerovy metody stálo několik miliard dolarů, trvalo zhruba 10 let a spolupracovala na něm celá plejáda laboratoří a vědců. Dnes bychom za podobný projekt zaplatili „pouze“ desítky tisíc dolarů a znamenal by práci několika šikovných výzkumníků. Díky tomu již můžeme

poměrně snadno sekvenovat kompletní genom z víceméně jakéhokoli organismu, z něhož jsme schopni získat dostatek DNA. Největší překážkou sekvenování genomů se stal proces počítačového skládání výsledných sekvencí genomů. Nevýhoda většiny sekvenačních technologií, jako je 454 nebo nejčastěji používaná Illumina, spočívá v produkci sekvencí dlouhých pouze několik stovek bází. Tím vzniká obtížně řešitelný problém, jak tyto krátké úseky správně poskládat do kompletní sekvence genomu. Jde vlastně o neuvěřitelně složitý puzzle s miliony dílků a s nejasným výsledným obrázkem postrádajícím jakékoli vědné hranice či linie. Proto se vkládají velké naděje do uplatnění metod třetí generace, které naopak dokážou přečíst velice dlouhé úseky DNA, a tím značně usnadnit výsledné skládání sekvencí do celkových genomů. V přirovnání k puzzle si můžeme představit skládanku pro menší děti s výrazně většími dílky, z nichž se mnohem snáze složí obrázek. Oba principy lze pochopitelně různě kombinovat a využívat přednosti toho kterého postupu.

Další metodou, která se dočkala rozmachu díky sekvenování druhé generace, je sekvenování transkriptomů (viz Živa 2016, 2: 61–63 a 3: 104–106). Při této metodě se místo kompletního genomu zjišťuje sekvence pouze aktivních genů, tedy genů, které jsou v buňkách v danou chvíli přepisovány do mediátorové RNA (mRNA) a překládány do proteinů. Při sekvenování transkriptomů se nejdříve získá veškerá mRNA z daného organismu (nebo jen z určité tkáně, orgánu apod.) a přepíše se do DNA molekuly zvané copy DNA (cDNA). Teprve tato DNA je následně sekvenována. Výhoda sekvenace transkriptomů oproti celým genomům spočívá v získání sekvencí genů bez balastní (nepotřebné) nekódující části genomové DNA. Nekódující DNA totiž často tvoří podstatnou část genomu organismu a sekvence jednotlivých genů je proto potřeba v této záplavě pracně hledat, což se ne vždy spolehlivě daří.

Pomocí sekvenování transkriptomů také můžeme studovat odlišnou expresi (míru přepisu) jednotlivých genů v závislosti na vnějších nebo vnitřních podmínkách. Zpravidla porovnáváme transkriptomy získané z organismu nacházejícího se ve dvou či více různých „stavech“ – např. pěstovaného při různých teplotách nebo ve zdravém či nemocném stavu apod. Porovnáním přítomnosti a četnosti sekvencí jednotlivých genů lze určit, které geny jsou v příslušné „fázi“ aktivnější, a tedy nejspíše zodpovídají za reakci daného organismu na tento „stav“, třeba na změnu teploty nebo onemocnění.

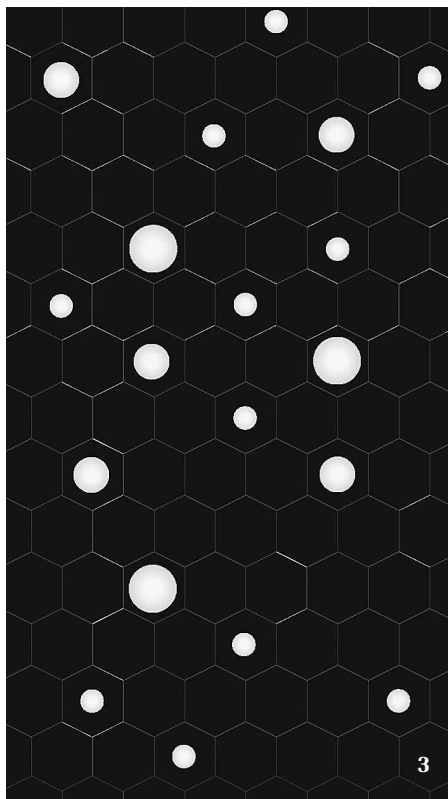
Další novinkou vzešlou z dílny sekvenování druhé generace je sekvenování „exomu“. Místo celého genomu nebo v danou chvíli aktivních genů (transkriptomů) sekvenujeme pouze kódující část genomu, tedy geny jako takové, bez intronů (Pozn.: U většiny eukaryotických organismů se většina genů skládá z exonů a intronů. Exony představují kódující část genu, zatímco introny nic nekódují, a proto jsou před přeložením do patričního genomového produktu vystřiženy.). Samozřejmě předem potřebujeme velmi dobře znát genom daného organismu a hranice jednotlivých

kódujících částí příslušných genů. Tato metoda se často používá při klinických studiích některých geneticky podmíněných chorob. V takovém případě lze porovnat exomy jedinců trpících poruchou a jedinců zdravých, a následně tak identifikovat mutace, které jsou pravděpodobně zodpovědné za nástup nemoci. Právě cenová dostupnost a vysoká efektivita metod sekvenování druhé generace dnes umožňuje zkoumat a rozpoznat příčiny řady vzácných a dříve málo studovaných genetických poruch.

● Environmentální sekvenční metody

Tyto metody se zaměřují na sekvenování DNA získané přímo z prostředí (environmentální vzorky) s cílem dozvědět se např. co nejvíce o druhovém složení (vybrané skupiny) organismů ve studovaném ekosystému. Nové sekvenční metody otevřely dosud zamčenou bránu, která umožňuje studovat zatím zcela nepřístupné skupiny, zejména nekultivovatelné mikroorganismy. Mikrobiální jednobuněčné organismy jsou všudypřítomné a hrají významnou, často i klíčovou roli ve fungování ekosystémů. Jejich studium však máme značně ztížené tím, že naprostou většinu zatím neumíme kultivovat v laboratořích, a tak nám jejich existence mnohdy zůstává utajena. Sekvenování DNA izolované přímo z prostředí tak představuje jediný možný přístup, jak se o těchto organismech alespoň něco dozvědět (viz článek na str. 118–120 tohoto čísla Živy).

Jednou z nejstandardnějších metod environmentálního sekvenování je tzv. sekvenování ampliconů. Pracujeme při něm pouze s úsekem DNA a nikoli s veškerou DNA přítomnou ve vzorku. DNA se izoluje z reprezentativní části studovaného ekosystému, např. ze vzorku vodního sloupce nebo zeminy. U takto získaného materiálu, jenž obsahuje DNA všech organismů nacházejících se ve vybraném vzorku prostředí, se pak pomocí PCR namnoží určitý zvolený úsek (amplifikuje, proto označení amplicon). Nejčastěji jde o gen pro ribozomovou RNA (pro malou či velkou podjednotku ribozomu, anglicky SSU – small subunit nebo LSU – large subunit), který má tu výhodu, že se univerzálně vyskytuje u všech buněčných organismů, v rámci populace příslušného druhu je identický, současně se však sekvence genu mezi jednotlivými organismy mírně liší. Tímto amplifikačním procesem (PCR) získáme velké množství molekul DNA kódujících sice pouze jeden gen (úsek), ale ze všech, nebo alespoň z naprosté většiny organismů přítomných ve vzorku odebraném z přirozeného prostředí. Když je tento mix molekul jednoho genu sekvenován, analýza získaných sekvenčních dat ukáže druhovou diverzitu studovaného prostředí (protože každý druh má specifickou sekvenční SSU/LSU). Navíc podle počtu kopií genů náležejících druhu lze s určitou dávkou nejistoty stanovit i početnost příslušného druhu ve studovaném prostředí. Jako příklad použití těchto metod z poslední doby uvedme studium mikrobiomů (blíže viz Živa 2015, 3: 106–107). Mikrobiom je společenství mikrobiálních organismů, především bakterií žijících v asociaci s hostitelem. Nejznámějším příkladem je střevní mikroflóra



3 Ilustrace signálu při 454 sekvenování. V „rozsvícených“ komůrkách došlo k přidání určité nukleotidové báze (při sekvenaci 454 se vždy přidává pouze jeden typ báze). Síla signálu určí, kolik bází bylo přidáno najednou. Nepřipojené nukleotidy se posléze odstraní a sekvenční proces pokračuje přidáním jiného nukleotidu. Orig. M. Kolísko

různých živočichů včetně člověka. Studium mikrobiomů se v posledních letech rozrostlo do víceméně samostatného vědního oboru a jednou z metod nejčastěji používaných k jejich stanovení se stala právě sekvenace ampliconů. Výzkum a srovnání diverzity mikrobiomů např. ukazuje rozdíly v druhovém složení a početnosti jednotlivých druhů bakterií mezi jedinci zdravými i trpícími různými zdravotními obtížemi (od alergií až po rakovinu střeva). Tato metoda může také napomáhat při řešení specifických úloh, jako je třeba rozlišení jednotlivých druhů potravy (kořisti) z obsahu natráveniny z těla studovaného predátora (viz Živa 2017, 1: 32–34).

● Sekvenování jednotlivých buněk

Environmentální metody slouží zejména k získávání informací o diverzitě organismů ve studovaném ekosystému, ať je tímto ekosystémem světový oceán, nebo střeva mušky octomilky. Jedním z hlavních důvodů pro použití environmentálních metod je nemožnost kultivovat naprostou většinu mikroorganismů v laboratorních podmínkách. Ovšem někdy nám nestačí znát pouze sekvenční jednoho genu, ale chceme se o příslušném organismu dozvědět více. A právě pro tyto případy je určena metoda sekvenování jediné buňky, která je do jisté míry komplementární k environmentálnímu přístupu. Není nutné studovaný mikroorganismus kultivovat, stačí ho pouze „odchytit“ z prostředí a sekvenovat

(obr. 2). Vzhledem k nepatrnému množství jakéhokoli materiálu (včetně DNA či RNA) izolovaného z jediné buňky je nutné nukleové kyseliny nejdříve enzymaticky namnožit, aby se podobně jako v předchozích případech vytvořily kopie existujících molekul DNA/RNA. To s sebou bohužel nese riziko, že některé úseky budou kopírovány častěji než jiné, což znamená, že výsledek bude reprezentovat jenom část skutečného genomu studované buňky.

Dnes sekvenování genomů z jednotlivých buněk používáme především pro bakterie, které mají obecně menší genomy (nejčastěji v milionech bází) a jednodušší strukturu. U eukaryotických organismů s komplexními a velkými genomy (v desítkách, stovkách a více milionech bází) se prozatím zdá být, až na několik výjimek, sekvenování genomů jednotlivých buněk až tou nejkratší možností. V posledních letech se však objevily metody pro sekvenování transkriptomů, tedy, jak jsme uvedli výše, pouze aktivních genů z jednotlivých buněk. Právě tyto metody se dnes používají jak pro studium eukaryotických mikroorganismů, tak např. pro studium rakoviny, kdy se zkoumá, jak se od sebe liší buňky v rámci jednoho nádoru.

Několik poznámek na konec

Popsané sekvenční metody se stále vyvíjejí a mění, některé zanikají, jiné se naopak objevují. Např. podpora 454 sekvenování byla ukončena v r. 2016. Uvedený přehled a příklady nových sekvenčních metod také nejsou vyčerpávající a jsou do určité míry ovlivněny zájmy a zkušenostmi autora tohoto článku. Kromě zmíněných metod existuje řada dalších využití moderních sekvenčních technologií, které nejen obohacují naše biologické znalosti, ale často mají i značný význam pro klinický výzkum.

Příchův obrovského množství sekvenčních dat také poněkud změnil biologickou práci. Není to tak dávno, co většina biologů buď prováděla laboratorní experimenty, nebo pracovala v terénu. Díky novým a levnějším sekvenčním technologiím začíná dnes některým oborům dominovat práce u počítače zaměřená na analýzu těchto dat. Není neobvyklé, že doktorandský student stráví první půlrok generováním dat a posléze několik let jejich analýzou. Bioinformatici, tedy lidé schopní analyzovat velké soubory sekvenčních dat, jsou stále žádanějšími spolupracovníky, a mnozí terénní nebo experimentální biologové se postupně tyto metody doučují a mění se v bioinformatiky. Zavedení sekvenčních metod druhé generace od základu změnilo biologii, jak jsme ji znali v minulých desetiletích, a bude zajímavé pozorovat, jaké nové metodiky změni biologii v blízké budoucnosti.

Na str. 120 tohoto čísla Živy najdete barevná schémata ukazující základní principy tří zde popisovaných metod (Sangerova metoda, 454 sekvenování a Illumina).

Použitou literaturu a webové adresy názorných animací detailně zobrazujících sekvenční metody druhé generace uvádíme na webové stránce Živy.