

PRINCÍPY NOVÝCH METÓD SEKVENOVANIA DNA

MARTIN GAPLOVSKÝ^a a KATARÍNA GAPLOVSKÁ-KYSELÁ^b

^a Katedra farmaceutickej chémie, Farmaceutická fakulta, Univerzita Komenského, Odbojárov 10, 831 04 Bratislava,

^b Katedra genetiky, Prírodovedecká fakulta, Univerzita Komenského, Mlynská dolina, pavilón B-1, 842 15 Bratislava

gaplovsky@fpharm.uniba.sk,

gaplovska-kyselaa@fns.uniba.sk

Došlo 20.7.11, prepracované 3.1.12, prijaté 20.1.12.

Kľúčové slová: sekvenovanie DNA, detekcia jednotlivých molekúl, fluorescenčne značené nukleotidy

Obsah

1. Úvod
2. Metódy prvej generácie
3. Metódy druhej generácie
4. Metódy tretej generácie
 - 4.1. Sekvenovanie syntézou
 - 4.2. Sekvenovanie priamym zobrazením
 - 4.3. Sekvenovanie pomocou nanopórov
5. Záver

1. Úvod

Spojenie technológií z oblastí biochémie, chémie, fyziky a technických odborov umožnilo v poslednom období dosiahnuť významný pokrok v určovaní primárnej štruktúry nukleových kyselín¹. Objavili sa úplne nové originálne postupy a metódy, ktorými sa proces sekvenovania DNA² podstatne zrýchlil a stal sa finančne menej náročný. Nástup týchto nových metód vytvára podmienky pre prevrátané zmeny v oblastiach, ako sú biomedicínsky výskum, diagnostika ochorení^{3,4} alebo farmakogenetika⁵. Realitou sa stávajú nové aplikácie, ako napríklad určenie kompletných genómov a detailná analýza vybraných úsekov DNA alebo štúdium expresie génov. V tejto práci chceme opísať princípy, na ktorých sú založené súčasné a nové metódy sekvenovania a naznačiť vývoj, ktorým sa bude táto oblasť výskumu v najbližších rokoch uberať.

2. Metódy prvej generácie

Dlhé obdobie bola štandardnou metódou sekvenovania DNA Sangerova metóda^{6,7}. Od jej vzniku v roku 1975

je nepretržite zlepšovaná a postupne stále viac automatizovaná. Pri polymerázou katalyzovanej syntéze komplementárnych vlákien v reakčnom roztoku obsahujúcom zmes štandardných nukleotidov a označených terminálnych nukleotidov vznikajú fragmenty DNA rozličnej dĺžky. Terminálne nukleotidy, po ich zabudovaní do reťazca, ukončia ďalšie predlžovanie komplementárneho vlákna a keďže reakcia prebieha s veľkým množstvom rovnakých templátov súčasne, vytvoria sa všetky možné dĺžky fragmentov DNA. Sekvencia študovaného úseku sa potom určí zoradením týchto fragmentov podľa ich dĺžky a odčítaním fluorescenčného signálu z terminálneho nukleotidu. V rámci Sangerovej metódy sú v súčasnosti používané automatizované zariadenia, využívajúce na separáciu a zoradenie fragmentov DNA kapilárnu elektroforézu⁸, veľmi často v paralelnom usporiadaní⁹. Využitie tejto technológie v nedávnej minulosti umožnilo skúmať genómy stále zložitejších organizmov a nakoniec úspešne dokončiť aj projekt ľudského genómu¹⁰.

Metódy založené na klasickej Sangerovej sekvenácii DNA sa považujú za metódy prvej generácie. Hlavnými obmedzeniami metód prvej generácie sú časová, materiálová a pracovná náročnosť prípravy vzoriek, najmä namnoženie a čistenie analyzovaného DNA materiálu¹¹. Ďalej je to obmedzené množstvo vzoriek, ktoré sa dajú analyzovať paralelne, v súčasnosti je možné analyzovať naraz 384 sekvencií s dĺžkou 600 až 1000 nukleotidov¹². Štandardné 96 kapilárové zariadenia vyprodukuje denne sekvenciu DNA o veľkosti približne 6 Mb (cit.¹³). Sekvenovanie prvého ľudského genómu týmto spôsobom trvalo viac ako 10 rokov a vyžiadalo si celkovo prostriedky v hodnote asi 3 miliardy dolárov⁴.

3. Metódy druhej generácie

Tieto obmedzenia, do značnej miery prekonávajú metódy druhej generácie, ktoré sa objavili okolo roku 2005. Pri podstatne nižších nákladoch sú schopné produkovať obrovské množstvo dát. V niektorých prípadoch je možné v jednej analýze paralelne sekvenovať viac ako sto miliónov krátkych úsekov. Cenou za zvýšenie výkonu je však vyššia frekvencia chybných stanovení nukleotidov, ktorá presahuje frekvenciu chybných stanovení 1/10 000 až 1/100 000 dosahovanú pri Sangerovej metóde¹³.

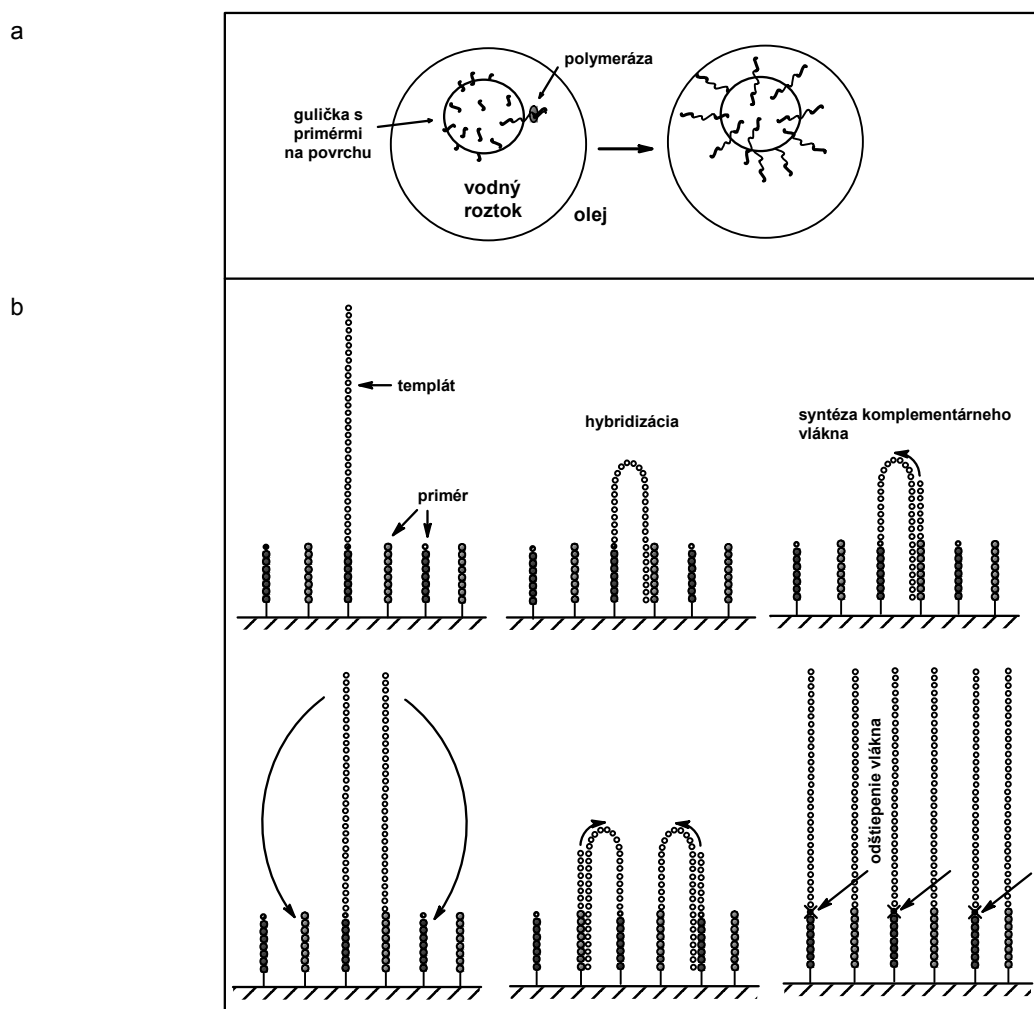
Vo všeobecnosti pozostáva sekvenovanie DNA z troch základných procesov. Sú to príprava vzorky, vlastné určenie poradia báz v jednotlivých fragmentoch a následné usporiadanie sekvencií získaných z týchto fragmentov.

Príprava vzorky zahŕňa extrakciu DNA, jej čistenie a fragmentáciu (rozbitie DNA na množstvo malých úsekov napríklad pôsobením ultrazvuku). Jednotlivé fragmenty sa

následne upravia do podoby DNA templátov vhodných pre použitie enzýmu, zvyčajne DNA polymerázy. Táto úprava sa zabezpečuje pomocou ligáz alebo transferáz, pridaním adaptérov pre ukotvenie molekuly na pevnej fáze a tiež pridaním štandardnej sekvencie vhodnej na hybridizáciu s primérom¹⁴. Nasleduje amplifikácia (zmnoženie) jednotlivých DNA templátov a vytvorenie priestorovo lokalizovaných súborov identických molekúl obsahujúcich tisíce až milióny kópií naviazaných na pevnú fázu. Na amplifikáciu sa využívajú rozličné techniky. Patrí sem napríklad emulzná polymerázová reťazová reakcia¹⁵ (obr. 1a). Polymerové guličky s množstvom štandardných primérov ukotvených na povrchu sa zmiešajú s fragmentami DNA v pomere, ktorý favorizuje maximálne 1 úsek DNA na guľičku. Následne sa vytvorí emulzia typu voda v oleji, čím sa jednotlivé guľičky oddelia a uzatvoria do mikroraktorov v kvapkách vodnej fázy emulzie. Každý fragment

sa pomocou PCR mnohonásobne zmnoží. Výsledkom sú guľičky husto pokryté namnoženou DNA.

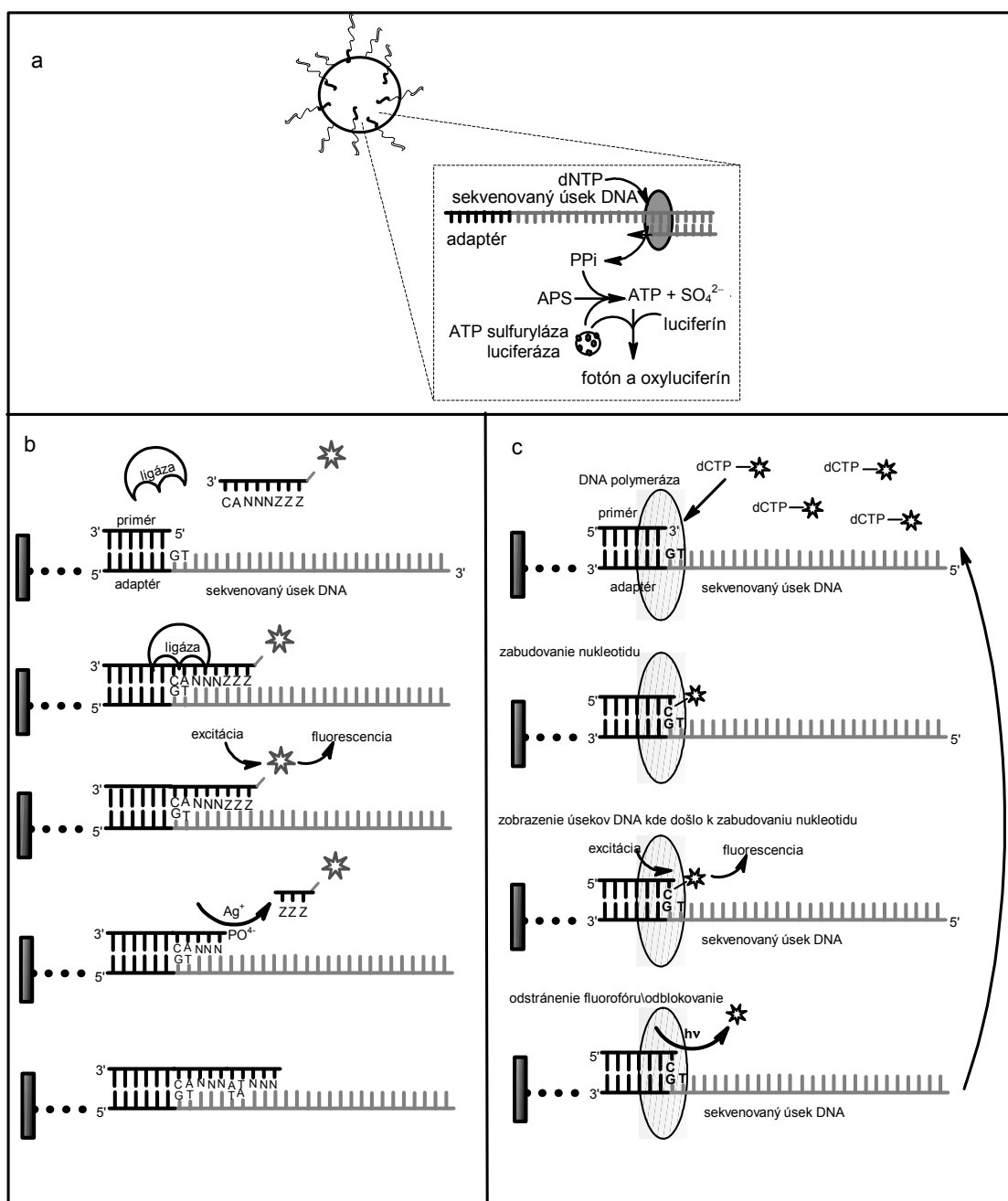
Inou metódou je amplifikácia na pevnej fáze¹⁶ (obr. 1b). Na pevnej fáze sa s vysokou hustotou imobilizujú dva typy štandardných primérov. Fragmenty DNA sa upravujú tak, že každý koniec je komplementárny s jedným z primérov. Následne sa tiež imobilizujú na pevnej fáze tak, aby vzdialenosť medzi jednotlivými fragmentmi bola dostatočná na ich optickú identifikáciu. Na voľný koniec fragmentu sa pripojí imobilizovaný primér a pomocou PCR sa DNA zreplikuje. Denaturáciou sa vlákna uvoľnia a následne znova hybridizujú s voľným primérom imobilizovaným na povrchu. Po dostatočnom počte cyklov replikácie sa komplementárne reťazce DNA odštiepia. Z jedného pôvodného fragmentu tak vznikne súbor identických fragmentov lokalizovaných na jednom mieste.



Obr. 1. Metódy prípravy priestorovo lokalizovaných súborov identických fragmentov DNA; a) emulzná PCR, b) amplifikácia na pevnej fáze

V súčasnosti je komerčne dostupných viacero sekvenovacích platforiem druhej generácie. Určenie poradia báz v rámci týchto metód je založené na rozličných typoch enzýmových reakcií. Pyrosekvenovanie^{17,18} (obr. 2a) využíva DNA polymerázu na predlžovanie komplementárneho reťazca. Pri každom zabudovaní nukleotidu sa uvoľňuje pyrofosfát, ktorý je pomocou ATP sulfurylázy v prítomnosti adenyllyl sulfátu konvertovaný na ATP. ATP

je substrátom pre enzým luciferáza, ktorý emituje svetlo vo viditeľnej oblasti. Vzniknutá bioluminescencia je v rámci určitých limitov úmerná počtu zabudovaných nukleotidov v danom kroku. V praxi sú namnožené identické úseky DNA naviazané na guľičky obklopené inými guľičkami, na ktorých sú ukotvené enzýmy ATP sulfuryláza a luciferáza. Guľičky s DNA sú umiestnené po jednej v jamkách mikrotitračnej doštičky. Tieto guľičky sú po



Obr. 2. Metódy sekvenovania druhej generácie; a) pyrosekvenovanie, b) sekvenovanie ligáciou, c) sekvenovanie s využitím reverzibilných terminátorov

stupne ponorené v reakčných roztokoch obsahujúcich vždy jeden zo štyroch nukleotidov. Prídanie správneho nukleotidu spôsobí, že enzým DNA polymeráza ho zabuduje do komplementárneho vlákna. V prípade, že v sekvencii nasleduje viac rovnakých báz za sebou, reakcia sa zastaví, až keď sa zabudujú všetky nukleotidy rovnakého typu. Pri tom sa uvoľní vždy jeden pyrofosfát na nukleotid. Pyrofosfát je ATP sulfurylázou zmenený na ATP. ATP umožňuje premenu luciferínu na oxyluciferín pomocou luciferázy, pričom sa emituje fotón. Zaznamenaná luminiscencia je priamo úmerná počtu nukleotidov zabudovaných v danom kroku.

Technológia využívajúca reverzibilné terminátory¹⁹ (obr. 2b) je založená na cyklickej prerušovanej syntéze komplementárneho vlákna pomocou polymerázy. Enzým DNA polymeráza zabuduje do reťazca vhodný fluorescenčne označený nukleotid, ktorý následne reverzibilne zablokuje predĺžovanie reťazca. Zabudovaný nukleotid sa zobrazí a identifikuje buď na základe toho, že reakcia prebehla, v prípade, že sa v reakčnom roztoku v danom kroku nachádza len jeden typ nukleotidov. Alebo na základe vlnovej dĺžky fluorescence ak sa naraz používajú všetky štyri nukleotidy. Polymeráza sa potom odblokuje odštiepením blokujúcej skupiny a fluorescenčnej značky. Vymení sa reakčný roztok a sekvenácia pokračuje ďalším cyklom.

Ďalšia pomerne rozšírená technológia pracuje na princípe postupného predĺžovania vlákna pripájaním krátkych oligonukleotidov pomocou DNA ligázy^{20,21} (obr. 2c). Na pevnej fáze sú cez univerzálny adaptér pripojené úseky DNA s naviazaným primérom. V reakčnom roztoku sa okrem enzýmu DNA ligázy nachádzajú fluorescenčne označené krátke úseky DNA (zvyčajne oktaméry) s rozličnými kombináciami báz na prvých dvoch pozíciách z 3' konca. DNA ligáza preferenčne pripojí k priméru polynukleotid, ktorý má bázy na pozícií 1 a 2 komplementárne so zodpovedajúcimi bázami sekvenovaného úseku. Po zobrazení fluorescence sa odštiepia 3 bázy vrátane značky z 5' konca. Primér sa takto predĺži o 5 báz a má na 5' konci voľný fosfát pre ďalšiu ligáciu. Po viacnásobnej ligácii (až do 10 cyklov) sa celý syntetizovaný reťazec uvoľní a spolu s reakčným roztokom odstráni. Nasleduje taký istý postup s primérom skrátenými postupne o jeden až štyri nukleotidy, čím sa dosiahne priradenie dinukleotidu ku každej pozícii skúmanej sekvencie.

Pri všetkých spomínaných metódach sa v jednotlivých súboroch identických fragmentov zisťuje poradie báz postupom, pri ktorom sa vždy po zabudovaní nukleotidu alebo nukleotidov reakcia vo všetkých sledovaných vláknach zastaví, aby sa umožnila ich identifikácia. Tým sa zároveň zaisťuje synchronizácia, ktorá je nutnou podmienkou pre vyhodnotenie údajov zaznamenaných pri použití metód založených na sledovaní súborov rovnakých molekúl.

Tieto synchronne kontrolované metódy využívajú na zastavenie reakcie, napríklad prísun len jedného zo 4 typov nukleotidov v reakčnom roztoku, prípadne modifikované nukleotidy reverzibilne blokujúce reakciu (obr. 3a–b). Po zabudovaní nukleotidu(ov) a zastavení reakcie sa prebytok

reaktantov opláchnie a pridaný nukleotid sa na základe fluorescenčného značenia alebo s využitím chemiluminescencie zobrazí. Keď je v danom cykle prítomný v reakčnom roztoku len jeden typ nukleotidu, už samotný fakt, že reakcia prebehla, je dostatočný na určenie sekvencie. Detekciu označeného nukleotidu je možné dosiahnuť aj s relatívne jednoduchou optikou pre zaznamenávanie fluorescence v úzkom rozsahu vlnových dĺžok. Ak sú v reakčnom roztoku prítomné všetky štyri typy dNTP naraz, je potrebné rozlíšiť ich napríklad zaznamenávaním fluorescence na rozličných vlnových dĺžkach alebo na základe rozdielnej doby života fluorescence. Fluorescenčné značky sa po detekcii deaktivujú chemickým alebo fyzikálnym procesom, aby nerušili detekciu v nasledujúcom cykle. V prípade technológie využívajúcej reverzibilné terminátory sa odstráni aj blokujúca skupina alebo skupiny. Následne sa znova pridá čerstvý reakčný roztok a sekvenovanie pokračuje ďalším cyklom. Sekvencie získané z veľkého množstva jednotlivých fragmentov sa následne usporiadajú do správneho poradia. Najčastejšie sa to dosiahne priradením fragmentov k referenčnému genómu²². Ak nie je referenčný genóm k dispozícii, využívajú sa počítačové algoritmy²³ umožňujúce získať genóm *de novo*.

Hlavnými obmedzeniami metód druhej generácie sú náročný proces amplifikácie fragmentov, nedokonalá synchronizácia inkorporácie nukleotidov a z toho vyplývajúce chyby určenia poradia báz. Nutnosť a zároveň náročnosť synchronizácie inkorporácie nukleotidov určuje aj maximálnu dĺžku fragmentov DNA, ktoré je možné v rámci určitej metódy sekvenovať. Zvyčajne je to okolo 20–100 báz, v prípade pyrosekvenovania až niekoľko sto báz. Akceptovateľná presnosť sa dosahuje viacnásobným paralelným sekvenovaním vzorky. Množstvo zaznamenaných údajov vytvára veľké nároky na skladovanie dát a ich správne usporiadanie, najmä pri sekvenovaní veľkého množstva krátkych úsekov.

4. Metódy tretej generácie

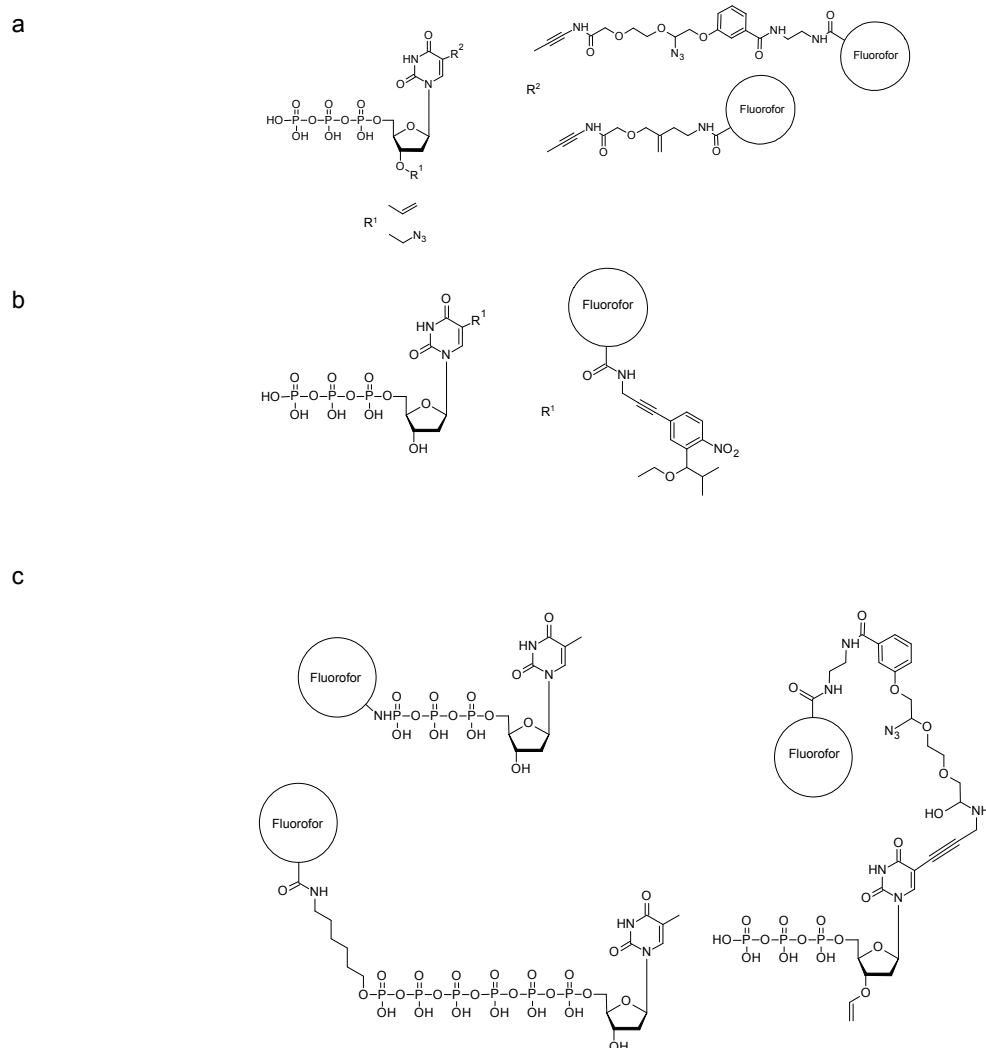
Za metódy tretej generácie sa považujú novo vyvíjané technológie, ktoré majú potenciál prekonať metódy druhej generácie v dôležitých parametroch, ako sú maximálna dĺžka fragmentov, ktoré je možné sekvenovať, spotreba reagentov alebo rýchlosť. Tieto metódy, zvyčajne na určovanie poradia báz, využívajú jednotlivé molekuly DNA bez potreby zastavenia procesu medzi jednotlivými krokmi detekcie. Patria sem aj metódy, ktoré úplne nespĺňajú toto kritérium, avšak nedajú sa už zaradiť ani medzi metódy druhej generácie. Napríklad polovodičový sekvenátor [www.iontorrent.com], ktorý využíva pole mikroskopických jamôk v polovodičivom materiáli, v ktorých sa zaznamenáva uvoľňovanie vodíkových iónov pri zabudovaní báz do DNA. Táto metóda pracuje s prirodzenými nukleotidmi a nevyužíva svetlo ani optiku, čím celkovo zjednodušuje proces sekvenovania. Stále však využíva systém oplachovania a detekcie zabudovania bázy pri cyk-

lickom striedaní 4 nukleotidov v reakčnom roztoku, vyžaduje tiež zmnoženie DNA templátu v každej jamke.

Sekvenovanie DNA založené na pozorovaní jednotlivých molekúl môže byť rozdelené zhruba do troch kategórií. Sú to technológie, pri ktorých sa pozoruje jedna molekula DNA polymerázy pri syntéze vlákna DNA. Ďalej sekvenovanie pomocou nanopórov, pri ktorom jednotlivé vlákna prechádzajú cez nanopóry alebo jednotlivé bazy sú detegované pri prechode nanopórmí. A nakoniec priame zobrazovanie jednotlivých DNA molekúl použitím pokročilých mikroskopických techník.

4.1. Sekvenovanie syntézou

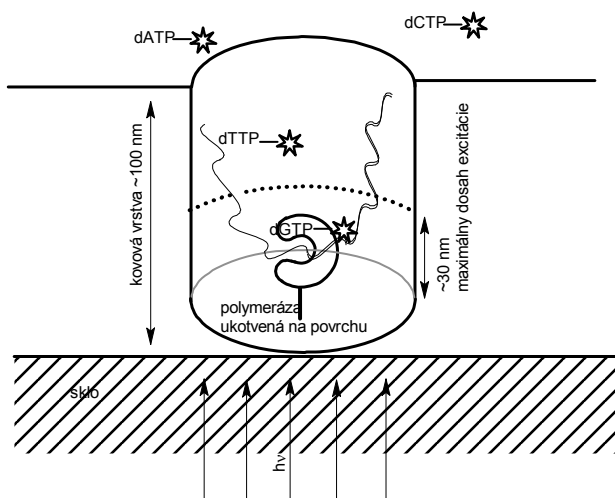
Prvé komerčne dostupné zariadenie na sekvenovanie jednotlivých molekúl²⁴ (od spoločnosti Helicos) zobrazuje predlžovanie jednotlivých molekúl DNA ukotvených v rovine na pevnom povrchu. Využíva pri tom modifikovanú polymerázu a špeciálne fluorescenčne značené nukleotidy²⁵, v ktorých je značka pripievaná k nukleotidu chemicky štiepateľnou skupinou, ktorá umožňuje sekvenovanie po krokoch (obr. 3b) Zastavovanie reakcie po jednotlivých krokoch zabudovania nukleotidov je teda stále potrebné a dĺžka sekvenovaného úseku je iba okolo 32 nukleotidov²⁶. Metóda však už nepotrebuje namnoženie DNA templátov. Chybovosť správneho určenia zabudovaného nukleotidu je stále vysoká > 5 %, ale metóda umož-



Obr. 3. **Modifikácie nukleotidov používaných pri sekvenovaní DNA;** a) nukleotidy s blokovaným 3' koncom, po zabudovaní modifikovaného nukleotidu je pre ďalší krok sekvenácie potrebné odštiepiť fluorescenčnú značku a odblokovať 3' koniec, b) nukleotidy s voľným 3' koncom, po zabudovaní modifikovaného nukleotidu je pre ďalší krok sekvenácie potrebné iba odštiepiť fluorescenčnú značku, c) nukleotidy pre sekvenovanie v reálnom čase

ňuje sekvenovať jeden úsek paralelne mnoho krát a chyby tak vylúčiť, čím sa dosiahne celková presnosť > 99 %. V jednej analýze je možné sledovať až okolo miliardy úsekov DNA. Na rozdiel od metód druhej generácie sa syntéza jednotlivých molekúl môže zaznamenávať asynchrónne, t.j. pre každé vlákno DNA zvlášť, pričom pre presnosť nie je nevyhnutné, aby zabudovanie nukleotidov v každom cykle, t.j. pred zastavením reakcie, prebehlo so 100% výťažkom. Modifikácia tejto metódy umožňuje, nahradením DNA polymerázy reverznou transkriptázou, priamo sekvenovať RNA t.j. bez jej konverzie na cDNA a bez amplifikácie. Fragmenty RNA sú polyadenylované, blokované na 3' konci a ukotvené na povrchu podobne ako je to pri sekvenovaní DNA.

Sekvenovanie jednotlivých molekúl v reálnom čase²⁷ (obr. 4) je prístup, pri ktorom sa jednotlivé molekuly DNA polymerázy pozorujú pri syntéze reťazca DNA bez prerušovania reakcie. Táto technológia musí prekonať dve zásadné prekážky. V prvom rade je potrebné v reálnom čase zaznamenať signál z jedného konkrétneho nukleotidu, ktorý je zabudovaný do štruktúry DNA, pričom tento nukleotid sa nachádza v reakčnom roztoku s množstvom ďalších fluorescenčne značených nukleotidov. Druhý problém je napojenie fluorescenčnej značky na nukleotid



Obr. 4. Sekvenovanie jednotlivých molekúl v reálnom čase. Enzym DNA polymeráza je ukotvený na pevnej fáze. Po pridaní reakčného roztoku s DNA templátmi a fluorescenčne označenými nukleotidmi sa začne syntéza komplementárneho vlákna. Pri zabudovaní nukleotidu sa zaznamená signál špecifický pre daný typ nukleotidu. Používajú sa nukleotidy s fluorofórom naviazaným cez fosfát, čím sa dosiahne jeho automatické odštiepenie pri zabudovaní do DNA. Signál z voľne difundujúcich nukleotidov, ktoré neboli zabudované je asi 1000 násobne kratší ako signál zo zabudovaného nukleotidu. Aby sa naraz nezaznamenal signál z dvoch alebo viacerých nukleotidov zmenší sa detekčný objem mikroskopu napríklad použitím tzv. zero mode waveguides, čo sú jamky v tenkej kovovej vrstve na sklenenom povrchu

takým spôsobom, aby sa značka dala po zabudovaní nukleotidu odstrániť a zároveň sa neprerušil proces syntézy a v detekčnom objeme sa nevyskytovali dve značky zároveň.

Detekcia konkrétneho nukleotidu v roztoku nukleotidov sa zabezpečí zmenšením detekčného objemu t.j. objemu roztoku, v ktorom sa zaznamenávajú emitované fotóny. Objem musí byť taký malý, aby sa v ňom štatisticky s vysokou pravdepodobnosťou v danom čase nachádzal len jeden nukleotid, za podmienky, že koncentrácia nukleotidov v roztoku je dostatočná pre správnu funkciu enzýmu.

Štandardné metódy na štúdium jednotlivých molekúl v roztokoch zahŕňajú fluorescenčnú korelačnú spektroskopiu²⁸ alebo priame pozorovanie jednotlivých molekúl v zriedených roztokoch pomocou difrakčne limitovanej optiky²⁹. Tieto metódy využívajú detekčné objemy s veľkosťou niekoľko desiatín femtolitra a na to, aby sa nimi dali sledovať jednotlivé molekuly sú potrebné piko- až nanomolárne koncentrácie fluorescenčne značených molekúl³⁰. Mnoho enzýmov však pracuje pri rádovo vyšších koncentráciách substrátov. Niektoré DNA polymerázy sú schopné pracovať aj pri koncentrácii nukleotidov okolo 10^{-6} mol l^{-1} . Fyziologické koncentrácie v bunke sú ešte o tri rády vyššie. Koncentrácia substrátu môže ovplyvniť funkciu a kinetiku enzýmu napríklad jeho alosterickou kontrolou alebo relaxáciou jeho aktívnej konformácie³¹. Na sledovanie mikromolárnych koncentrácií je teda potrebná zhruba 1000 násobná redukcia detekčného objemu v porovnaní so štandardnými technikami.

Zmenšenie detekčného objemu sa dá dosiahnuť napr. použitím TIRF (Total Internal Reflection Fluorescence) mikroskopie³², alebo tzv. Zero Mode Waveguides³³ čo sú jamky s priemerom v desiatkach nanometrov vyhlbené v tenkom (asi 100 nm) kovovom filme, ktorým je potiahnutý sklenený povrch. Malý rozmer týchto jamiek zabráňuje viditeľnému svetlu, ktoré má vlnovú dĺžku okolo 600 nm, aby cez ne preniklo. Pri ožiarení cez sklenenú vrstvu intenzita svetla v jamkách s hĺbkou exponenciálne klesá a efektívne je ožiarená len časť jamky do 30 nm od jej dna.

V týchto jamkách sú na modifikovanom sklenenom povrchu upevnené molekuly polymerázy. Na jednu jamku pripadá maximálne 1 molekula polymerázy. V reakčnom roztoku sa nachádzajú všetky štyri typy nukleotidov označené každý inou fluorescenčnou značkou. Nukleotidy sa voľne pohybujú v roztoku a doba, ktorú takto sa pohybujúci nukleotid strávi v detekčnom objeme pri dne jamky s polymerázou, sa pohybuje rádovo v mikrosekundách. Iný prípad nastane, ak je správny nukleotid zabudovaný do syntetizovaného vlákna. Tento proces trvá na rozdiel od difúzie niekoľko milisekúnd, teda približne 1000 krát dlhšie, čomu zodpovedá aj podstatne vyšší počet emitovaných fotónov a tým aj dobrý pomer signálu k šumu. Sekvenovací prístroj zaznamená záblesk, ktorého vlnová dĺžka zodpovedá zabudovanému nukleotidu. Následne je potrebné fluorescenčnú značku z detekčného objemu odstrániť. Táto druhá prekážka sa prekoná tak, že na rozdiel od metód druhej generácie, ktoré využívajú nukleotidy s fluorescenčnou značkou pripevnenou priamo na nukleotid, je pri

sekvenovaní v reálnom čase značka pripojená na fosfátový reťazec³⁴ (obr. 3c). Tento reťazec je v procese zabudovania nukleotidu do reťazca DNA prirodzeným spôsobom odštiepený, fluorescenčná značka sa uvoľní a rýchlo sa difúziou vzdiali z detekčného objemu. Následne sa môže zabudovať a detegovať ďalší nukleotid. Sekvenovanie jednotlivých molekúl v reálnom čase umožňuje paralelne sekvenovať fragmenty s priemernou dĺžkou nad 1000 bázových párov.

Odlíšny spôsob, ako prekonať problém s detekciou jedného nukleotidu v roztoku nukleotidov, je využiť FRET (fluorescenčný rezonančný prenos energie), v takom prípade je priamo na polymeráze pripevnený fluorofór, ktorý emituje fotóny po prenose energie zo svetlom excitovanej značky zabudovávaného nukleotidu, ktorý sa v momente reakcie nachádza dostatočne blízko pre FRET, to je vo vzdialenosti niekoľko nanometrov. Po zabudovaní nukleotidu a uvoľnení fluorescenčnej značky sa FRET ihneď preruší³⁵.

4.2. Sekvenovanie priamym zobrazením

Alternatívou k sekvenovaniu syntézou môže byť v budúcnosti priame zobrazenie DNA pomocou elektrónovej mikroskopie³⁶ alebo skenovacej tunelovej mikroskopie³⁷. Predpokladom pre použitie týchto metód je prichytenie natiahnutého vlákna DNA na pevný povrch a dostatočné priestorové rozlíšenie mikroskopických techník umožňujúce identifikovať jednotlivé bázy. Uvažuje sa aj s možnosťou použiť bázy označené substituentami obsahujúcimi atómy s vyššou atómovou hmotnosťou, pre ich jednoduchšiu identifikáciu [www.zsgenetics.com]. Tieto technológie majú potenciál čítať veľmi dlhé úseky DNA (až niekoľko miliónov báz) pri nízkych nákladoch.

4.3. Sekvenovanie pomocou nanopórov

Metódy využívajúce nanopóry sú založené na prechode molekuly DNA (alebo jej jednotlivých zložiek) cez membránové póry a detekcii jednotlivých báz na základe ich vplyvu na elektrický prúd. Vo vývoji je systém zložený z modifikovaného póru α -hemolizínu^{38–40}, ku ktorému je pripojená z jednej strany membrány exonukleáza a z druhej strany syntetický cyklodextrín. Tento systém je včlenený do lipidovej dvojvrstvy takým spôsobom, že pri vytvorení potenciálu na membráne, zmenou koncentrácie iónov, sa aktivuje exonukleáza, ktorá odštiepuje z DNA jednotlivé nukleotidy. Nukleotidy sú identifikované na základe charakteristického narušenia toku iónov cez nanopór. Pracuje sa na vývoji stabilnejších syntetických membrán a nanopórov, ktoré zvýšia spoľahlivosť a rýchlosť tejto metódy [www.nanoporetech.com].

Ďalšia možnosť je použitie nanopórov priamo na sekvenovanie intaktnej jednovláknovej DNA⁴¹. Na rozdiel od predchádzajúcej metódy, ktorá bola limitovaná nízkym pozdĺžnym rozlíšením póru α -hemolizínu, sa využíva proteín Porin A z *Mycobacterium smegmatis*, ktorý ma kratší úsek, ktorým sa pór otvára a uzatvára a tým aj potenciálne

vyššie rozlíšenie. Problémom je však príliš rýchly prechod vlákna DNA cez pór, prevyšujúci rýchlosť umožňujúcu spoľahlivú detekciu vplyvu jednotlivých báz na narušenie prúdu iónov v póre.

Spoločnosť IBM vyvíja nanoštrukturované sekvenovacie zariadenie schopné elektronicke detegovať nukleotidy v jednotlivých molekulách DNA⁴². Nanoštruktúry sú vlastne póry s pozdĺžne vrstveným povrchom skladajúcim sa zo striedajúcich vrstiev kovu a dielektrického materiálu (ako v tranzistore). Vhodné modulovanie prúdu v takýchto tranzistoroch umožní prechod jednotlivých molekúl DNA cez póry. Pôvodne bola táto technológia navrhnutá na základe teoretických výpočtov a molekulovo dynamických simulácií, ale IBM nedávno publikovala spôsob, ako ovládať rýchlosť, ktorou vlákno DNA prechádza cez nanopór⁴³. Ostáva vyriešiť problém odlišenia signálu jednotlivých báz od signálu susedných báz v reťazci⁴⁴.

5. Záver

Metódy sekvenovania nukleových kyselín zaznamenali za posledných 10 rokov nevídaný pokrok. Optimalizáciou prípravy vzoriek, povrchovej chémie, fluorescenčných značiek, enzýmových reakcií, optiky a elektroniky sa dosiahla hranica, keď je možné získať údaje o poradí báz aj z jednotlivých molekúl DNA. Dá sa predpokladať, že dosiahnutý technologický pokrok sa uplatní aj v iných oblastiach výskumu, mimo sekvenovania nukleových kyselín.

V súčasnosti je možné v krátkom čase sekvenovať celé genómy rozličných organizmov vrátane človeka. Paradoxne, problémom prestáva byť určenie poradia báz, ale stáva sa ním, napriek veľkým pokrokom v informatike, vyhodnocovanie, interpretácia a dokonca aj ukladanie veľkého množstva získaných údajov.

LITERATÚRA

- Fuller C. W., Middendorf L. R., Benner S. A., Church G. M., Harris T., Huang X., Jovanovich S. B., Nelson J. R., Schloss J. A., Schwartz D. C., Vezenov D. V.: *Nat. Biotechnol.* 27, 1013 (2009).
- Metzker M. L.: *Nat. Rev. Genet.* 11, 31 (2010).
- Ocana A., Pandiella A.: *Mol. Cancer* 9, 202 (2010).
- Pospíšilová S., Tichý B., Mayer J.: *Čas. Lék. Čes.* 148, 296 (2009).
- Stratton M. R., Campbell P. J., Futreal P. A.: *Nature* 458, 719 (2009).
- Sanger F., Nicklen S., Coulson A. R.: *Proc. Natl. Acad. Sci. U.S.A.* 74, 5463 (1977).
- Sanger F., Coulson A. R.: *J. Mol. Biol.* 94, 441 (1975).
- George K. S., Zhao X., Gallahan D., Shirkey A., Zareh A., Esmaeli-Azad B.: *J. Chromatogr., B: Biomed. Sci. Appl.* 695, 93 (1997).
- Shendure J. A., Porreca G. J., Church G. M.: *Curr.*

- Protoc. Mol. Biol. Chapter 7, Unit 7.1.
10. International Human Genome Sequencing Consortium: *Nature* 431, 931 (2004).
 11. Hert D. G., Fredlake C. P., Barron A. E.: *Electrophoresis* 29, 4618 (2008).
 12. Emrich C. A., Tian H., Medintz I. L., Mathies R. A.: *Anal. Chem.* 74, 5076 (2002).
 13. Kircher M., Kelso J.: *Bioessays* 32, 524 (2010).
 14. Linnarsson S.: *Exp. Cell Res.* 316, 1339 (2010).
 15. Dressman D., Yan H., Traverso G., Kinzler K. W., Vogelstein B.: *Proc. Natl. Acad. Sci. U.S.A.* 100, 8817 (2003).
 16. Fedurco M., Romieu A., Williams S., Lawrence I., Turcatti G.: *Nucleic Acids Res.* 34, e22 (2006).
 17. Ronaghi M., Uhlen M., Nyren P.: *Science* 281, 363 (1998).
 18. Margulies M., Egholm M., Altman W. E., Attiya S., Bader J. S., Bemben L. A., Berka J., Braverman M. S., Chen Y. J., Chen Z., Dewell S. B., Du L., Fierro J. M., Gomes X. V., Godwin B. C., He W., Helgesen S., Ho C. H., Irzyk G. P., Jando S. C., Alenquer M. L., Jarvie T. P., Jirage K. B., Kim J. B., Knight J. R., Lanza J. R., Leamon J. H., Lefkowitz S. M., Lei M., Li J., Lohman K. L., Lu H., Makhijani V. B., McDade K. E., McKenna M. P., Myers E. W., Nickerson E., Nobile J. R., Plant R., Puc B. P., Ronan M. T., Roth G. T., Sarkis G. J., Simons J. F., Simpson J. W., Srinivasan M., Tartaro K. R., Tomasz A., Vogt K. A., Volkmer G. A., Wang S. H., Wang Y., Weiner M. P., Yu P., Begley R. F., Rothberg J. M.: *Nature* 437, 376 (2005).
 19. Bentley D. R., Balasubramanian S., Swerdlow H. P., Smith G. P., Milton J., Brown C. G., Hall K. P., Evers D. J., Barnes C. L., Bignell H. R., Boutell J. M., Bryant J., Carter R. J., Keira Cheetham R., Cox A. J., Ellis D. J., Flatbush M. R., Gormley N. A., Humphray S. J., Irving L. J., Karbelashvili M. S., Kirk S. M., Li H., Liu X., Maisinger K. S., Murray L. J., Obradovic B., Ost T., Parkinson M. L., Pratt M. R., Rasolonjatovo I. M., Reed M. T., Rigatti R., Rodighiero C., Ross M. T., Sabot A., Sankar S. V., Scally A., Schroth G. P., Smith M. E., Smith V. P., Spiridou A., Torrance P. E., Tzonev S. S., Vermaas E.H., Walter K, Wu X, Zhang L, Alam MD, Anastasi C, Aniebo IC, Bailey D. M., Bancarz I. R., Banerjee S., Barbour S. G., Baybayan P. A., Benoit V. A., Benson K. F., Bevis C., Black P. J., Boodhun A., Brennan J. S., Bridgham J. A., Brown R. C., Brown A. A., Buermann D. H., Bundu A. A., Burrows J. C., Carter N. P., Castillo N., Chiara E., Catenazzi M., Chang S., Neil Cooley R., Crake N. R., Dada O. O., Diakoumakos K. D., Dominguez-Fernandez B., Earnshaw D. J., Egbujor U. C., Elmore D. W., Etchin S. S., Ewan M. R., Fedurco M., Fraser L. J., Fuentes Fajardo K. V., Scott Furey W., George D., Gietzen K. J., Goddard C. P., Golda G. S., Granieri P. A., Green D. E., Gustafson D. L., Hansen N. F., Harnish K., Haudenschild C. D., Heyer N. I., Hims M. M., Ho J. T., Horgan A. M., Hoschler K., Hurwitz S., Ivanov D. V., Johnson M. Q., James T., Huw Jones T. A., Kang G. D., Kerelska T. H., Kersey A. D., Khreb-tukova I., Kindwall A. P., Kingsbury Z., Kokko-Gonzales P. I., Kumar A., Laurent M. A., Lawley C. T., Lee S. E., Lee X., Liao A. K., Loch J. A., Lok M., Luo S., Mammen R. M., Martin J. W., McCauley P. G., McNitt P., Mehta P., Moon K. W., Mullens J. W., Newington T., Ning Z., Ling Ng B., Novo S. M., O'Neill M. J., Osborne M. A., Osnowski A., Ostadan O., Paraschos L. L., Pickering L., Pike A. C., Pike A. C., Chris Pinkard D., Pliskin D. P., Podhasky J., Quijano V. J., Raczy C., Rae V. H., Rawlings S. R., Chiva Rodriguez A., Roe P. M., Rogers J., Rogert Bacigalupo M. C., Romanov N., Romieu A., Roth R. K., Rourke N. J., Ruediger S. T., Rusman E., Sanches-Kuiper R. M., Schenker M. R., Seoane J. M., Shaw R. J., Shiver M. K., Short S. W., Sizto N. L., Sluis J. P., Smith M. A., Ernest Sohna J., Spence E. J., Stevens K., Sutton N., Szajkowski L., Tregidgo C. L., Turcatti G., Vandevondele S., Verhovskiy Y., Virk S. M., Wakelin S., Walcott G. C., Wang J., Worsley G. J., Yan J., Yau L., Zuerlein M., Rogers J., Mullikin J. C., Hurler M. E., McCooke N. J., West J. S., Oaks F. L., Lundberg P. L., Klenerman D., Durbin R., Smith A. J.: *Nature* 456, 53 (2008).
 20. Tomkinson A. E., Vijayakumar S., Pascal J. M., Ellenberger T.: *Chem. Rev.* 106, 687 (2006).
 21. Valouev A., Ichikawa J., Tonthat T., Stuart J., Ranade S., Peckham H., Zeng K., Malek J. A., Costa G., McKernan K., Sidow A., Fire A., Johnson S. M.: *Genome. Res.* 18, 1051 (2008).
 22. Delseny M., Han B., Hsing Y. I.: *Plant Sci.* 179, 407 (2010).
 23. Schatz M. C., Delcher A. L., Salzberg S. L.: *Genome Res.* 9, 1165 (2010).
 24. Braslavsky I., Hebert B., Kartalov E., Quake S. R.: *Proc. Natl. Acad. Sci. U.S.A.* 100, 3960 (2003).
 25. Harris T. D., Buzby P. R., Babcock H., Beer E., Bowers J.: *Science* 320, 106 (2008).
 26. Pushkarev D., Neff N. F., Quake S. R.: *Nat. Biotechnol.* 27, 847 (2009).
 27. Eid J., Fehr A., Gray J., Luong K., Lyle J., Otto G., Peluso P., Rank D., Baybayan P., Bettman B., Bibillo A., Bjornson K., Chaudhuri B., Christians F., Cicero R., Clark S., Dalal R., Dewinter A., Dixon J., Foquet M., Gaertner A., Hardenbol P., Heiner C., Hester K., Holden D., Kearns G., Kong X., Kuse R., Lacroix Y., Lin S., Lundquist P., Ma C., Marks P., Maxham M., Murphy D., Park I., Pham T., Phillips M., Roy J., Serra R., Shen G., Sorenson J., Tomaney A., Travers K., Trulson M., Vieceli J., Wegener J., Wu D., Yang A., Zaccarin D., Zhao P., Zhong F., Korlach J., Turner S.: *Science* 323, 133 (2009).
 28. Mertz J., Xu C., Webb W. W.: *Opt. Lett.* 20, 2532 (1995).
 29. Medina M. A., Schwillie P.: *Bioessays* 8, 758 (2002).
 30. Betzig E., Chichester R. J.: *Science* 262, 1422 (1993).
 31. Lerch H. P., Mikhailov A. S., Hess B.: *Proc. Natl. Acad. Sci. U.S.A.* 99, 15410 (2002).

32. Starr T. E., Thompson N. L.: *Biophys. J.* 80, 1575 (2001).
33. Levene M. J., Korlach J., Turner S. W., Foquet M., Craighead H. G., Webb W. W.: *Science* 299, 682 (2003).
34. Hardin S., Gao X., Briggs J., Willson R., Tu, S.: C. US Patent 7,329,492 (2000).
35. Schadt E. E., Turner S., Kasarskis A.: *Hum. Mol. Genet.* 19, R227 (2010).
36. Krivanek O. L., Chisholm M. F., Nicolosi V., Pennycook T. J., Corbin G. J., Dellby N., Murfitt M. F., Own C. S., Szilagy Z. S., Oxley M. P., Pantelides S. T., Pennycook S. J.: *Nature* 464, 571 (2010).
37. Blow N.: *Nat. Methods* 5, 267 (2008).
38. Clarke J., Wu H. C., Jayasinghe L., Patel A., Reid S., Bayley H.: *Nat. Nanotechnol.* 4, 265 (2009).
39. Howorka S., Cheley S., Bayley H.: *Nat. Biotechnol.* 19, 636 (2001).
40. Stoddart D., Heron A. J., Mikhailova E., Maglia G., Bayley H.: *Proc. Natl. Acad. Sci. U.S.A.* 106, 7702 (2009).
41. Derrington I. M., Butler T. Z., Collins M. D., Manrao E., Pavlenok M., Niederweis M., Gundlach J. H.: *Proc. Natl. Acad. Sci. U.S.A.* 107, 16060 (2010).
42. Polonsky S., Rossnagel S., Stolovitzky G.: *Appl. Phys. Lett.* 91, 153103 (2007).
43. Luan B., Peng H., Polonsky S., Rossnagel S., Stolovitzky G., Martyna G.: *Phys. Rev. Lett.* 104, 238103 (2010).
44. Krems M., Zwolak M., Pershin Y. V., Di Ventra M.: *Biophys. J.* 97, 1990 (2009).

M. Gaplovský^a and K. Gaplovská-Kyselá^b (^a *Department of Pharmaceutical Chemistry, Faculty of Pharmacy, Comenius University, Bratislava,* ^b *Department of Genetics, Faculty of Natural Science, Comenius University, Bratislava*): **Principles of New DNA Sequencing Methods**

Recently, a combination of biochemical, chemical, physical and engineering techniques enabled a substantial advancement of determination of primary structure of nucleic acids. New methods emerged that made DNA sequencing faster, less expensive and less labor-consuming. This review describes principles of the current and new sequencing methods as well as indicates directions of future development.