Sekvenciranje nove generacije

Vučenović, Dunja

Undergraduate thesis / Završni rad

2013

Degree Grantor / Ustanova koja je dodijelila akademski / stručni stupanj: **University of Zagreb, Faculty of Science / Sveučilište u Zagrebu, Prirodoslovno-matematički fakultet**

Permanent link / Trajna poveznica: https://urn.nsk.hr/urn:nbn:hr:217:013519

Rights / Prava: In copyright / Zaštićeno autorskim pravom.

Download date / Datum preuzimanja: **2024-03-29**



Repository / Repozitorij:

Repository of the Faculty of Science - University of Zagreb







SVEUČILIŠTE U ZAGREBU

PRIRODOSLOVNO – MATEMATIČKI FAKULTET

BIOLOŠKI ODSJEK

SEKVENCIRANJE NOVE GENERACIJE

NEXT GENERATION SEQUENCING

SEMINARSKI RAD

Dunja Vučenović

Preddiplomski studij molekularne biologije

(Undergraduate Study of Molecular Biology)

Mentor: Prof. dr. sc. Kristian Vlahoviček

Zagreb, 2013.

Sadržaj:

1. Uvod 2
2. Sekvenceri druge generacije 3
   2.1. 454 pirosekvenciranje 3
   2.2. Ilumina analizator genoma 6
   2.3. AB SOLiD sustav 7
3. Sekvenceri treće generacije 9
   3.1. PacBio 9
   3.2 Heliscope 10
   3.3 Ion Torrent Personal Genome Machine 12
4. Sekvenceri četvrte generacije – Nanopore sekvenceri 13
   4.1. Oxford Nanopore 13
   4.2. „Solid-state“ sekvenciranje nanoporama 15
5. Zaključak 16
6. Literatura 17
7. Sažetak 18
8. Summary 19

1. Uvod

Još od Averyjeva otkrića iz 1944., kojim je pokazano da je DNA molekula koja nosi genetički materijal, preko otkrića strukture DNA (Watson i Crick, 1953.) i određivanja centralne dogme molekularne biologije (Crick, 1970.), pokušavaju se razviti metode koje će informacije pohranjene u genetičkom kodu moći primijeniti te na taj način biti od koristi društvu. Upravo tehnologija sekvenciranja DNA u velikoj bi mjeri mogla pomoći medicini i biologiji u brojnim primjenama poput: kloniranja, razmnožavanja, pretraživanja promjena gena koje uzrokuju razvoj bolesti, evoluciji i drugim. U proteklih trideset godina, DNA sekvenciranje doživjelo je značajan napredak te je osnova genomske ere koju karakteriziraju velike količine podataka o genomima. Potrebno je osvrnuti se na razvoj tehnika sekvenciranja kako bismo usporedili i uvidjeli razlike između različitih metoda sekvenciranja nove generacije.

Prije govora o metodama sekvenciranja nove generacije, recimo nešto i o prvoj korištenoj metodi sekvenciranja, Sangerovoj dideoksi metodi. Razvijena 1977., ova metoda se temelji na terminaciji elongacije lanca DNA, korištenjem fluorescentno obilježenih dideoksinukleotida (ddNTPa). U nastaloj smjesi obilježenih DNA različitih duljina, boja biljega određuje vrstu krajnje baze. Sekvenca DNA molekule određuje se elektroforetskim razdvajanjem visoke razlučivosti nastale smjese te određivanjem fluorescentne boje uz pomoć lasera.

Ova metoda zahtjeva mnogo vremena i prostora, prostor za odvijanje reakcije te kapilarnu cijev ili gel za određivanje duljine DNA. Posljedično, vrlo malo fragmenata se može paralelno sekvencirati. Ljuski genom ima oko 3 milijarde baza, što možemo predstaviti s 6 milijuna fragmenata od 500 baznih parova. Ako bismo čitav genom sekvencirali ovom metodom to bi vrlo dugo trajalo jer se samo 100-ak fragmenata može paralelno sekvencirati. Također, ova metoda je znatno skuplja od metoda kasnije razvijenih.

2. Sekvenceri druge generacije

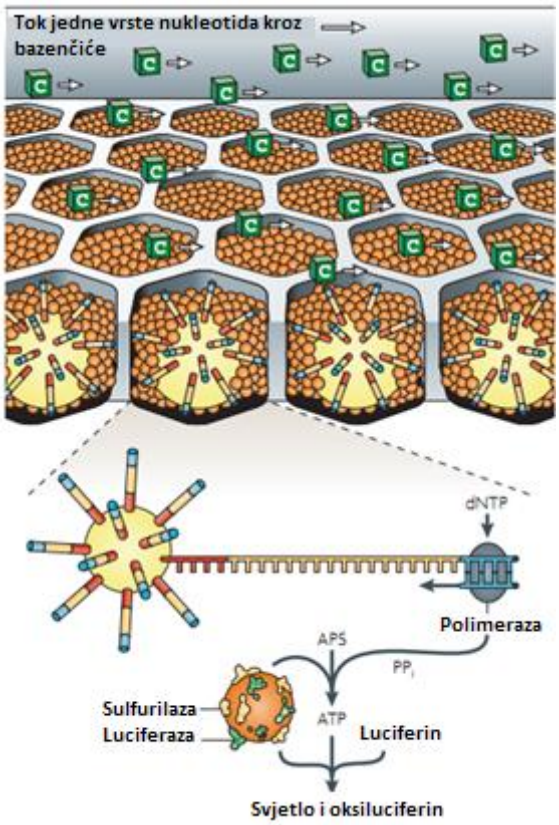
Velika potreba za povoljnim sekvenciranjem potaknula je razvoj visokoprotočne tehnologije sekvenciranja, koja paralelnim sekvenciranjem omogućuje stvaranje milijuna sekvenci odjednom (Hall N 2007). Nazvane još i „masivno paralelne“, ove metode se značajno razlikuju od Sangerova kapilarnog sekvenciranja i iako se međusobno razlikuju po načinu sekvenciranja, dijele neka svojstva. Prvenstveno, preparatorni koraci obimom su manji i jednostavniji od onih u Sangerovu sekvenciranju. Metode sekvenciranja nove generacije počinju pripremom zbirke, ligacijom sintetičkih DNA molekula (adaptori) na krajeve fragmenata koji će se sekvencirati. Sve platforme zahtijevaju umnažanje fragmenata zbirke na podlozi, uz pomoć polimeraze. Amplifikacija je potrebna za proizvodnju dovoljne količine signala koju će detektor prepoznati pri sekvenciranju. Ovaj korak uzrokuje i najveće greške u određivanju slijeda nukleotida, jer polimeraze nisu nikad u potpunosti točne. Posljednje, ovi instrumenti sekvenciraju slijedom ponovljenih reakcija koje se automatski izvode i detektiraju nukleotid. Sve platforme međusobno se razlikuju po specifikacijama reakcije sekvenciranja, što predočava izniman razvoj kemije, molekularne biologije i inženjerstva, koji su potrebni za simultano generiranje informacije o slijedu tisuća do milijuna DNA molekula. U daljnjem tekstu iznijet ću svojstva pojedinih platformi druge generacije sekvenciranja.

2.1. 454 pirosekvenciranje

Roche 454 prvi je komercijalno dostupan sekvencer, baziran na metodi sekvenciranja nove generacije. 454 sekvencer pokrenuo je razvoj nove generacije sekvenciranja, predloživši rješenja za tri tada najveća problema visokoprotočnog sekvenciranja: pripreme knjižnica i predložaka te samog sekvenciranja. Koncept sekvenciranja pomoću sinteze, čiji je predstavnik i pirosekvenciranje, predlaže dva načina sekvenciranja. S jedne strane, ideja sekvenciranja cikličkom reverzibilnom terminacijom, tj. cikličkom dodavanju fluorescentno obilježenih nukleotida, njihovoj detekciji i naposljetku odstranjivanju fluorescentne baze. S druge strane, tehnologije sekvenciranja pomoću detekcije pirofosfata, otpuštenih pri inkorporaciji nukleotida u rastući DNA lanac koji će, naposljetku, uzrokovati otpuštanje svjetla od strane enzima luciferaze. Druga ideja je izabrana za 454 platformu, jer se smatralo da je direktna inkorporacija neobilježenih nukleotida efikasnija od ponovljenih ciklusa inkorporacije, detekcije i cijepanja. Samo pirosekvenciranje u znanosti je bilo prisutno desetak godina prije razvoja 454 platforme,

no nije se smatralo kompetentnim zbog prekratkih duljina sekvenciranih fragmenata, iz kojeg razloga se uglavnom koristilo u detekciji promjena u jednom nukleotidu.

Uzorak DNA pocijepa se u fragmente, kojima se na oba kraja dodaju adaptorske sekvence koje pomažu identifikaciju i sekvenciranje. Za umnožavanje DNA fragmenata koristi se emulzijski PCR (lančana reakcija polimerazom) [emPCR] (Berka 2005.). Jednolančani DNA fragment hibridizacijski je vezan za promotor na kuglicama za umnažanje. Postupak je optimiziran tako da se za jednu kuglicu veže samo jedna vrsta fragmenta. Ti fragmenti se u kuglicama umnažaju te, nakon što dosegnu kritičan broj molekula, kuglice se raspuknu oslobađajući do deset milijuna kopija DNA fragmenta u bazenčiće na optičkom stakalcu (Slika 1). Uz kopije fragmenta, u bazenčiću se nalaze i enzimi potrebni za elongaciju i detekciju odgovarajuće baze (surfurilaza, luciferaza, DNA polimeraza, apiraza). U bazenčiće se strujanjem dovodi po jedna vrsta deoksinukleotida. Ukoliko je doveden odgovarajući deoksinukleotid (dNTP), doći će do njegove ugradnje u rastući lanac i otpuštanja pirofosfata, koji se prevodi do ATP-a uz pomoć enzima surfurilaze. Luciferaza koristi nastali ATP za proizvodnju svjetlosti koja se detektira kao signal ugradnje nukleotida u fragment DNA. U protivnom, ako dovedeni dNTP nije odgovarajući, enzim apiraza razgraditi će dovedeni dNTP (Foehlich i sur, 2010.), a u bazenčić će se dovesti iduća vrsta dNTP-a.



Slika 1. Pirosekvenciranje korištenjem Roche platforme.
Nakon unošenja DNA fragmenata u bazenčiće s kuglicama koji će umnožiti fragmente dodaje se pojedina vrsta nukleotida, zajedno sa surfurilazom i luciferazom. Ispod bazenčića se nalazi kamera koja detektira svjetlost nastalu nakon ugradnje odgovarajućeg nukleotida.
(Metzker, 2010.)

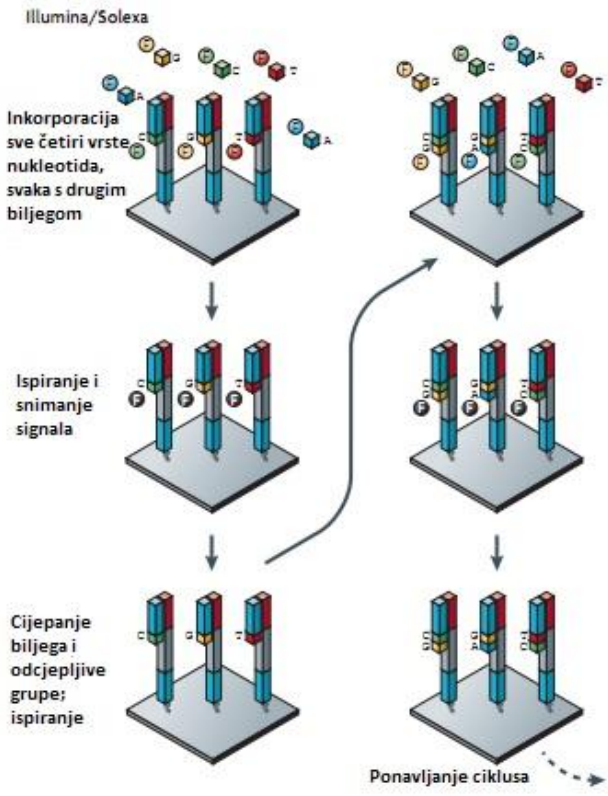
Zbog činjenice da se pirosekvenciranje temelji na detekciji otpuštene svjetlosti kad se novi nukleotid ugradi u rastući lanac, nema potrebe za fizičkim odvajanjem reakcija da bi se utvrdila sljedeća baza u lancu, za razliku od npr. elektroforeze, gdje je potrebna određena duljina reakcijskog prostora za precizno razlikovanje različitih nukleotida, što omogućuje smanjivanje

reakcijskog volumena na prostor dostatan za proizvodnju opazive svjetlosti. Također, zbog otpuštanja svjetlosti, pirosekvenciranje omogućava paralelno sekvenciranje.

Glavni nedostatak 454 tehnologije jest loše prepoznavanje homopolimera (uzastopnih sljedova istog nukleotida). Duljina homopolimera procjenjuje se intezitetom emitirane svjetlosti. U usporedbi s ostalim sekvencerima nove generacije, glavna prednost 454 platforme je *read-length* (454 može sekvencirati i 700 baza duge fragmente). Trenutno je 454 platforma najskuplja vrsta sekvencera nove generacije, a koristi se ukoliko je potrebna što veća duljina sekvenciranih fragmenata za npr. *de novo* sekvenciranje ili metagenomiku.

2.2. Ilumina analizator genoma

Ovaj sekvencer također se koristi tehnologijom sekvenciranja pomoću sinteze. Zbirke sekvenci s adaptorskim djelovima se denaturiraju u jednolančane DNA molekule, koje se učvrste na podlogu, nakon čega se te molekule umnože pomoću premošćavajućeg PCR-a kako bi se pripremile nakupine koje koriste identične fragmente. Početnice iz oba smjera povežu s podlogom pomoću fleksibilne spojnice, pa će svi amplikoni, koji nastaju od jedne molekule, tijekom umnažanja fragmenata ostati imobilizirani i grupirani na jednom mjestu na matrici. Ova platforma temelji se na donekle nekonvencionalnom premošćavajućem PCR-u, koristeći se Bst polimerazom i formamidom kao denaturirajućim sredstvom (Shendure 2008.). Nastale nakupine fragmenata mogu sadržavati oko tisuću amplikona, a nekoliko milijuna nakupina može se nalaziti i razlikovati od ostalih na svakoj od osam odijeljenih staza na pločici. Nakon generiranja nakupina, amplikoni su jednolančani i hibridizirani s početnicom koja ukazuje na regiju od interesa.



Slika 2. Shematski prikaz rada Illumina sekvencera. (Metzker, 2010.)

Svaki ciklus određivanja nukleotida sastoji se od simultanog dodavanja mješavine četiri modificirana deoksinukleotida, od kojih je svaka vrsta specifično obilježena fluorescentnim biljegom koji se može pocijepati. Ovi nukleotidi na 3' hidroksilnoj skupini imaju odcjepljivi privjesak koji dopušta ugradnju samo jednog nukleotida u polinukleotidni lanac u jednom ciklusu sekvenciranja. Nakon ugradnje jednog nukleotida, fragmenti se snimaju kamerom da bi se uočio fluorescentni biljeg ugrađenog nukleotida.

Nakon određivanja ugrađenog nukleotida, posljednjem nukleotidu cijepa se fluorescentni biljeg i odcjepljivi privjesak, što omogućava ugradnju novog nukleotida i novi ciklus sekvenciranja (Slika 2).

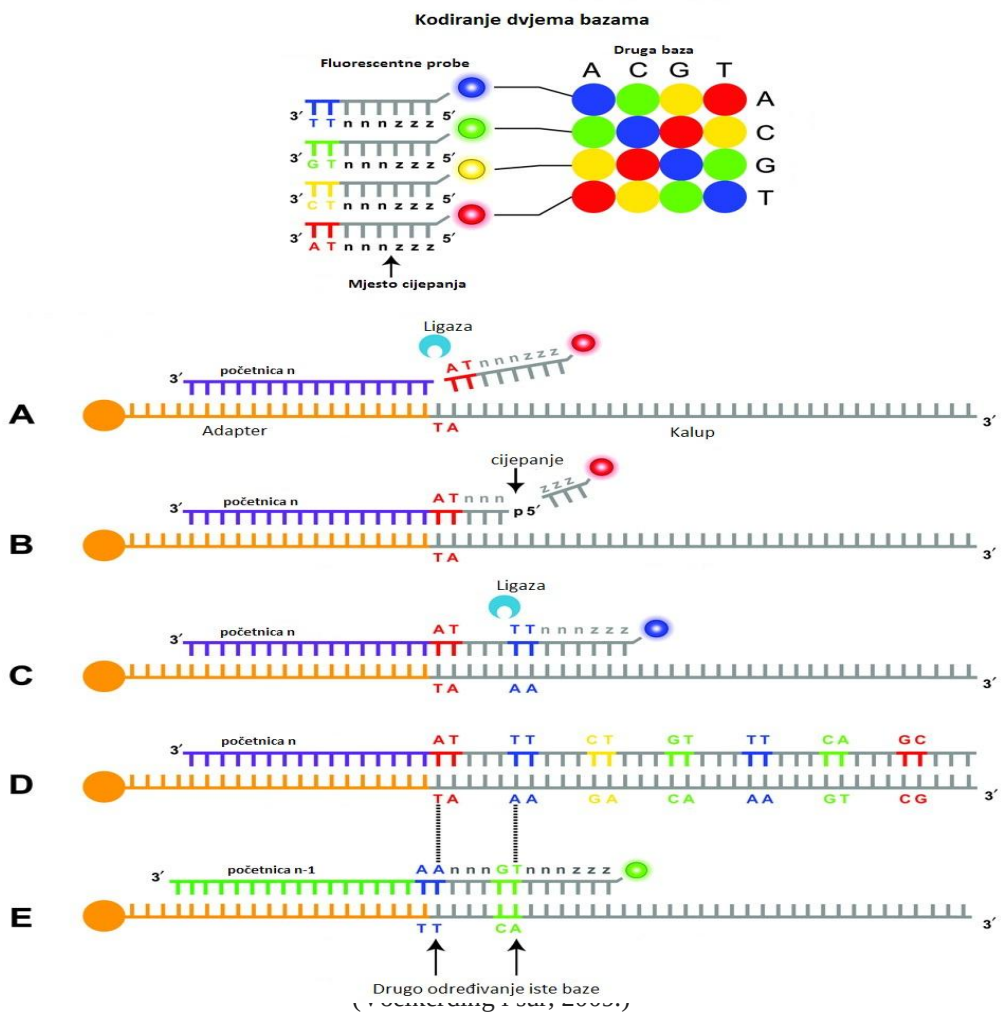
Trenutno, Illumina/Solexa analizator genoma dominira tržištem sekvencera nove generacije, vjerojatno i zbog činjenice da je cijena sekvenciranja ovom metodom najniža. Kako pločice sadrže osam odijeljenih staza, osam različitih zbirki DNA fragmenata može se paralelno sekvencirati. Duljine fragmenata do 36 parova baza rutinski se sekvenciraju. Moguće su i veće duljine, ali nivo grešaka značajno raste s duljinom fragmenta. Duljina fragmenata je određena mnogim faktorima koji uzrokuju miješanje signala, poput neuspješnog cijepanja signala ili odcjepljivog privjeska. Glavna vrsta pogreške je supstitucija, s najčešćim greškama u određivanju baze nakon gvanina (Dohm 2008.).

2.3. AB SOLiD sustav

SOLiD (Sequencing by Oligo Ligation Detection = sekvenciranje pomoću detekcije spajanja oligonukleotida) sekvencer koristi tehnologiju istovremenog određivanja dvaju nukleotida, temeljivši se na sekvenciranju pomoću ligacije. I ova metoda se, kao i prijašnje, temelji na sekvenciranju tijekom sinteze, ali za razliku od prijašnjih metoda, umjesto DNA polimeraze, u ovoj je metodi sekvenciranja ključni enzim ligaza. Pomoću emulzijskog PCR-a umnažaju se fragmenti koji će se sekvencirati i vežu za podlogu. Fragmenti se sekvenciraju pomoću probi s 8 baza, koje se spajaju s rastućim lancem. Probe se sastoje od dviju baza koje se spajaju s rastućim lancem (prve dvije baze), mjesta iza kojeg se cijepa (peta baza), i četiri različite fluorescentne boje koje su vezane za zadnju bazu. Fluorescentni signal će se opaziti ukoliko je proba komplementarna s lancem predloškom, a nestaje nakon cijepanja posljednjih triju baza probe (Slika 3).

S fragmentima vezanim za podlogu hibridiziraju početnice, dodaju se probe koje se vežu za početnice ukoliko su im prve dvije baze komplementarne s fragmentom (Slika 3, A). Nakon uspješnog vezivanja, dolazi do zadržavanja fluorescentnog signala karakterističnog za prve baze, kojeg detektiraju kamere za sve vrste signala. Poslije se proba cijepa iza pete baze (Slika 3, B), čime nestaje signala, a dolazi do vezanja nove probe, koja je komplementarna sa sljedeća dva nukleotida (Slika 3, C). Na taj način se odrede baze sa šeste i sedme pozicije. U idućem ciklusu odredit će se baze s jedanaestog i dvanaestog mjesta. Ovakva metoda omogućava određivanje svakih pet baza. Nakon završetka produljivanja početnice probama, početnica se odvaja od fragmenta i dodaje druga početnica (Slika 3, E), koja je u odnosu na prijašnju pomaknuta za

jedno mjesto, tako da će se sad odrediti baze s druge i treće pozicije fragmenta. Nakon pet promjena početnica, svaka baza u fragmentu bit će određena u dvjema neovisnim ligacijskim reakcijama. Ovakva metoda se naziva i kodiranje dvjema bazama (Mckernan i sur., 2009).



Kodiranje dvjema bazama moćna je metoda, namijenjena jasnom razlikovanju grešaka nastalih mjerenjem od polimorfizama jednog nukleotida. Kombinacija korištenja ligaze, izmjene početnica i kodiranja dvjema bazama pridonose niskoj stopi pogrešno određenih nukleotida u fragmentima, što je ujedno i najveća prednost ove platforme.

Duljina sekvenciranih fragmenata početno je iznosila 35 parova baza. Daljnjim razvojem ove metode, povećala se i preciznost sekvenciranja te duljina fragmenata, koja trenutno iznosi 85 baza. Najveći nedostatak ove metode predstavlja kratka duljina fragmenata, zbog čega se ova metoda uglavnom koristi pri resekvenciranju, istraživanju transkriptoma i epigenoma.

3. Sekvenceri treće generacije

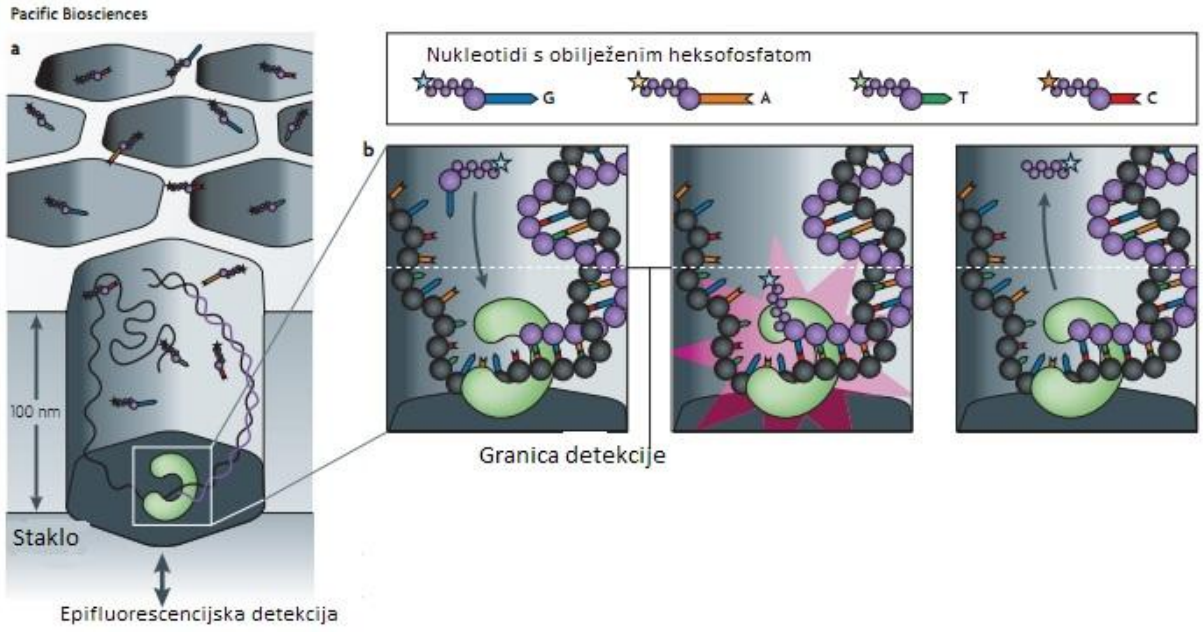
Sekvenceri prve i druge generacije zaslužni su za revolucionaran razvoj genomike, potičući impresivan broj znanstvenih postignuća poput određivanja genoma izumrlih životinja (mamut i neandertalac), standardnih modelnih organizama te pojedinačnih ljudskih genoma. Ipak, mnoge primjene sekvenciranja i aspekti genomike izvan su dosega ovih platformi, što zbog kratkih sekvenciranih fragmenata, što zbog smanjene točnosti zbog umnažanja fragmenata. S tim idejama na umu, znanstvenici su razvili platforme koje će sekvencirati dulje fragmente bez potrebe za njihovim umnažanjem. Metode sekvenciranja pojedinih molekula u realnom vremenu mogu se podijeliti u dvije grupe, one u kojima se promatra DNA polimeraza koja sintetizira pojedine molekule DNA u realnom vremenu te metode koje određuju slijed baza direktnim snimanjem DNA fragmenata naprednim mikroskopskim tehnikama. Svaka od ovih tehnologija predstavlja novi pristup sekvenciranju, ali imaju prednosti i nedostatke u odnosu na ostale. U nastavku poglavlja osvrnut ću se upravo na specifičnosti pojedinih sekvencera treće generacije.

3.1. PacBio

Tvrtka Pacific Biosciences razvila je metodu sekvenciranja pojedinih molekula u realnom vremenu (SMRT), vrstu sekvenciranja pomoću sinteze. Prednost ove vrste sekvenciranja pomoću sinteze u odnosu na postojeće jest ušteda na vremenu i reagensima, jer nije potrebno umnožavati fragmente. Ova platforma također koristi fluorescentno obilježene nukleotide, ali za razliku od prijašnjih, ovi nukleotidi obilježeni su na γ-fosfatu, dijelu nukleotida koji se prirodno cijepa pri elongaciji DNA molekule, te zbog ovako dizajniranih nukleotida nije potreban korak u kojem se cijepa fluorescentni biljeg. SMRT sekvenciranje ne zahtijeva ponavljanje koraka u kojima se postupno dodaju obilježeni nukleotidi, već je polimeraza, koja ima znatno veću procesivnost od onih korištenih u platformama druge generacije, sposobna provesti elongaciju nativno, kao što bi funkcionirala u organizmu iz kojeg je izolirana, čime se proces sekvenciranja ubrzava do 30 000 puta.

Za promatranje inkorporacije obilježenih nukleotida koristi se nanoprotonska struktura, nazvana „zero-more waveguide“, koja se sastoji od staklene podloge presvučene 100 nm debelim slojem aluminija, u kojem se nalaze cilindrični utori od 70-ak nm u promjeru (Slika 4). Aluminij se dodatno tretira kako DNA polimeraza ne bi asocirala s bočnim plohama utora. Ovo je bitan korak, jer ne postoji način da se manipulira polimerazom i dostavi je na pravo mjesto, već se

podloga uroni u otopinu polimeraza, koje se onda zalijepe za dno utora. Zbog malih dimenzija utora, valna duljina vidljive svjetlosti ne može u potpunosti prodrijeti u utor, pa je većina volumena utora u tami. Svjetlost uspije obasjati donjih 10 nm utora, u kojima se nalazi polimeraza. Obasjavanje polimeraze neophodno je za ekscitaciju obilježenih nukleotida koji se ugrađuju u lanac DNA, no zbog ograničene količine svjetlosti, bilo bi nemoguće razlikovati vrste ugrađenih nukleotida, stoga PacBio sekvenciranje koristi optičku prizmu preko koje snimaju podlogu s polimerazama. Na taj način, emitirana svjetlost razdvoji se u zrake različitih boja, koje se nalaze na različitim lokacijama, ovisno o vrsti nukleotida.



Slika 4. Shematski prikaz „zero-more waveguide“ strukture za koju je zahvaćena polimeraza. Ugradnjom odgovarajućeg nukleotida dolazi do emisije svjetlosti. (Metzker, 2010.)

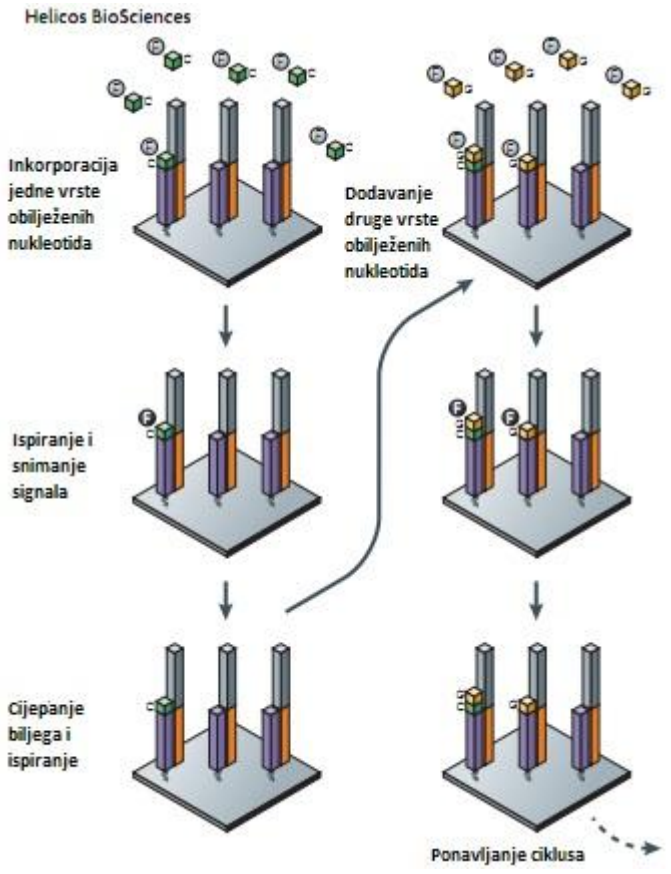
Prosječna duljina fragmenata sekvencirana PacBio platformom iznosi 600 nukleotida, a najdulje sekvence imaju do 3 000 nukleotida. Tako duge sekvence olakšavaju sastavljanje fragmenata iz regija s repetitivnim sljedovima u genomu.

3.2 Heliscope

Helicos sekvencer prvi je sekvencer na tržištu koji omogućava sekvenciranje pojedinih molekula. Također se temelji na cikličkom određivanju inkorporiranih nukleotida u procesu sekvenciranja pomoću sinteze, ali ova platforma ne zahtijeva klonsku amplifikaciju DNA molekula uz pomoć PCR uređaja, što omogućava preciznije sekvenciranje, jer se izbjegavaju

pogreške pri umnažanju fragmenata. Ova platforma ima elemente metoda sekvenciranja druge i treće generacije; ne koristi PCR, ali se koristi detekcijom emitirane fluorescencije i stupnjevitom sintezom pojedine vrste baza.

Jednolančana zbirka fragmenata pripremi se nasumičnim cijepanjem DNA molekula i dodavanjem poliadeninskih sljedova. Takvi fragmenti vežu se za podlogu hibridizacijom poliadeninskog repa sa slijedom timina koji su učvršćeni za podlogu. U svakom ciklusu, DNA polimeraza i jedna vrsta fluorescentno obilježenih nukleotida dodaju se na podlogu s nasumičnim jednolančanim fragmentima. U onim fragmentima u kojima se ugradio nukleotid, uočava se fluorescentni biljeg koji detektiraju kamere. Nakon ispiranja podloge od prijašnjeg nukleotida, cijepa se fluorescentni biljeg s posljednjeg nukleotida, jer isti sprječava ugradnju novih baza, te se na podlogu doda druga vrsta nukleotida. Nakon stotinjak ovakvih ciklusa, sekvenciraju se fragmenti od otprilike 25 parova baza.



Slika 5. Shematski prikaz rada Helicos sekvencera (Metzker, 2010.)

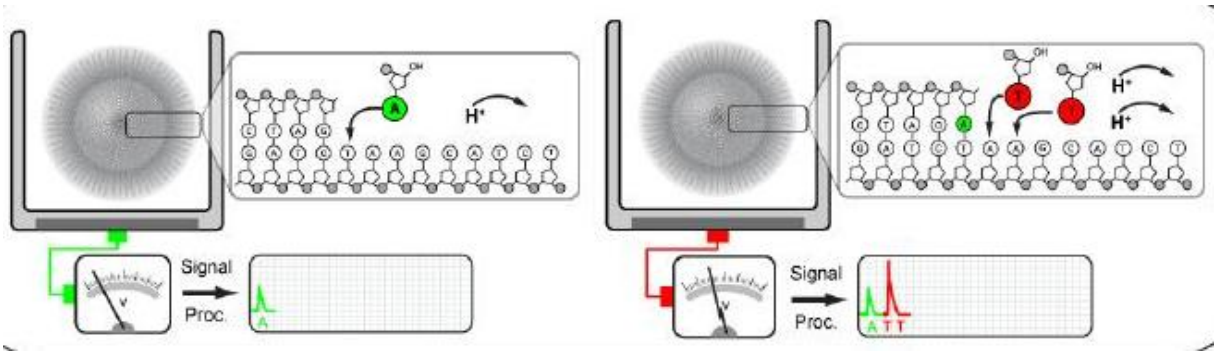
Poput sekvenciranja pomoću 454 platforme, sekvenciranje ovom metodom je asinkrono, nekim će se fragmentima slijed brže odrediti, a nekima sporije. Postoji i šansa da se nukleotid ne ugradi u slijed, bez obzira na to što je to pripadajuća baza. Kako se ova metoda koristi sekvenciranjem pojedinih molekula, a ne klonova, asinkorono sekvenciranje ne predstavlja problem i ne uzrokuje greške. Kako ovi, obilježeni nukleotidi ne sadrže privjesak koji blokira elongaciju, definiranje homopolimera predstavlja problem. Zbog sekvenciranja pojedinih molekula, problem se može ublažiti limitirajući stopu ugradnje nukleotida. Također, određeno je da uzastopna ugradnja obilježenih nukleotida u homopolimerima uzrokuje oslabljene interakcije, čime se može odrediti broj inkorporiranih nukleotida.

Točnost sekvencirajna moguće je značajno poboljšati koristeći se dvosmjernim sekvenciranjem u kojem, nakon određivanja slijeda nukleotida u fragmentu, isti se kopiraju te se njihove kopije sekvenciraju. Na taj način dobijemo dvije sekvence istog fragmenta, ali u suprotnim orijentacijama. Problem predstavlja i ugradnja neobilježenih nukleotida ili onih koji ne fluoresciraju, zbog čega najveću vrstu grešaka predstavljaju delecijske greške. Koristeći se dvosmjernim sekvenciranjem, ova platforma postaje najpreciznija od svih sekvencera druge generacije.

3.3 Ion Torrent Personal Genome Machine

Ion Torrent Personal Genome Machine prvi je sekvencer koji za svoj rad ne treba kompleksne optičke sustave, jer za određivanje ugrađenog nukleotida ne detektira otpuštene fotone svjetlosti, već oslobođene protone. Naime, pri ugradnji nukleotida u rastući lanac, dolazi do oslobađanja pirofosfata i protona koji detektiraju senzori osjetljivi na protone. Ova platforma uvelike pojednostavljuje i pojeftinjuje instrumente za sekvenciranje koji se, umjesto optičkih komponenti, sastoji od elektroničke ploče koja komunicira s čipom, mikroprocesora za obradu signala i sistema za kontrolu protoka reagensa kroz čip.

DNA se fragmentira i veže za adapterske sekvence, takvi fragmenti se umnože i nanesu na čip, tako da samo jedna vrsta fragmenata uđe u jedan senzorički utor. Nukleotidi pojedine vrste u utore se dostavljaju postupno, u ciklusima. Kad se u utoru nalazi odgovarajući nukleotid, dolazi do njegove ugradnje pomoću polimeraze vezane u utoru, što uzrokuje hidrolizu nukleotid trifosfata te otpuštanje jednog protona (Slika 6). Otpuštanje protona uzrokuje promjenu pH vrijednosti otopine u utoru proporcionalno broju ugrađenih nukleotida (0,02 pH jedinice po jednom nukleotidu (Rothberg i sur. 2011.). Promjenu pH detektira senzor na dnu utora, koji taj signal pretvara u promjenu napona struje na tom dijelu čipa. Nakon dovođenja jedne vrste nukleotida, čip se ispire kako prijašnja vrsta nukleotida ne bi zaostala u utorima. Male dimenzije utora dopuštaju difuziju u i iz utora, čime se eliminira potreba za enzimatskim uklanjanjem reagensa (Marguiles 2005.)



Slika 6. Ugradnja nukleotida uzrokuje otpuštanje protona, a samim time i snižavanje pH vrijednosti okoline. Pri ugradnji dva nukleotida uočava se jači signal zbog većeg smanjenja pH.

Prosječna duljina sekvenciranih fragmenata iznosi 100 do 200 nukleotida. Ova tehnologija ima znatno veću preciznost od BioPac tehnologije, usprkos činjenici da se tijekom pripreme fragmenata dva puta vrši amplifikacija fragmenata, što zbog nepreciznosti polimeraze uzrokuje greške pri sekvenciranju. Smanjena preciznost je primjetna u sekvencama koje imaju velik AT sadržaj te u homopolimerima koji imaju više od osam uzastopnih nukleotida iste vrste (Quail i sur. 2012.).

4. Sekvenceri četvrte generacije – Nanopore sekvenceri

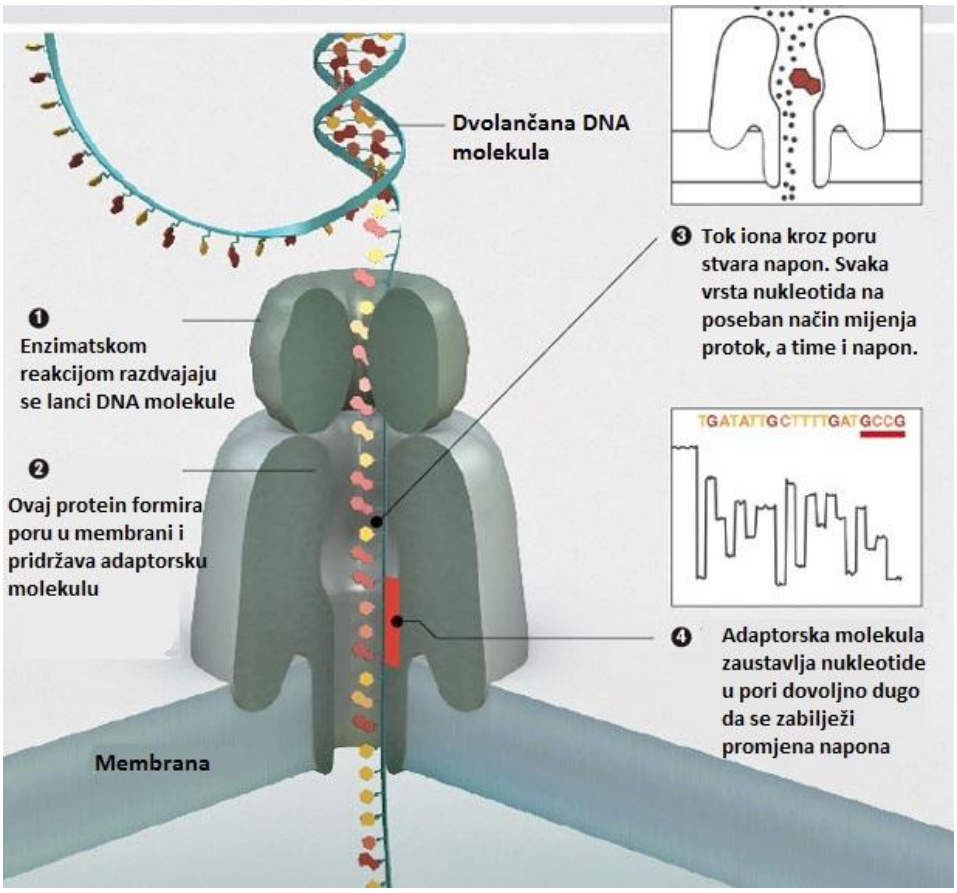
Fundamentalno različita vrsta sekvenciranja od do sad opisanih temelji se na nanoporama. Detekcija pojedinih baza, predviđa se, moguća je praćenjem promjene provodljivosti kroz nanoporu u membrani. Pore imaju otvor neznatno veći od promjera dvolančane molekule DNA koji iznosi 4 nm. Razlike u kemijskoj strukturi pojedinih nukleotida bi u teoriji uzrokovala opažljivu promjenu u toku iona kroz poru. Općenito, metode sekvenciranja nanoporama se mogu podijeliti u dvije grupe: biološke i „solid-state“ metode. Tehnologija sekvenciranja još je u razvoju i predstavlja interesantnu potencijalnu četvrtu generaciju sekvencera. Ime četvrta generacija je predložena jer ne bi koristile optičku detekciju ugrađenih nukleotida kao ni ponavljajuće cikluse ispiranja reagensa.

4.1. Oxford Nanopore

Tvrtka Oxford Nanopore Technologies razvila je metodu sekvenciranja pomoću enzima alfa-hemolizina koji premošćuje dva sloja membrane i uzrokuje lizu DNA (biološka metoda).

Membrana odjeljuje provodljive otopine u kojima se nalaze elektrode, zbog kojih postoji konstantan tok iona kroz pore. Nukleinske kiseline iz otopina mogu se provesti kroz nanopore, a strukturalni elementi molekula određuju se kao promjene u protoku iona. Pojedina vrsta nukleotida prolaskom kroz poru uzrokuje karakterističnu promjenu protoka, što omogućuje precizno određivanje slijeda nukleotida.

Sekvenciranje nanoporama može se podijeliti u dvije različite metode. Prva metoda se temelji na prolasku intaktnog DNA polimera kroz poru, a sekvenciranje se vrši u realnom vremenu kako DNA translocira kroz poru (Slika 7). Druga metoda koristi se prolazom pojedinih nukleotida kroz nanoporu, a koji se postepeno oslobađaju iz DNA lanca pomoću enzima egzonukleaze. Duljine fragmenata, koje ove metode mogu sekvencirati, daleko su veće od duljina sekvenciranih drugim metodama i iznose desetke kilobaza, no velike probleme zadaje smanjena točnost ovih metoda. Pogreške se događaju uslijed difuzije oslobođenog nukleotida, koji u tom slučaju ne će proći kroz poru, ili ukoliko se lanac DNA translocira za više od jedne baze, čime se izostave pojedine baze.



Slika 7. Shematski prikaz rada sekvencera koji se koriste nanoporama sekvencera (Schafer, 2012.)

Prednosti ove platforme leže u činjenici da nije potreban PCR za umnažanje DNA fragmenata, već se oni mogu direktno sekvencirati kako prolaze kroz poru. Također, ova metoda ne iziskuje kompleksne optičke sustave ni kemijsko obilježavanje nukleotida. Usprkos svim prednostima, sekvenciranje nanoporama još uvijek nije dovoljno komercijalno dostupno, rutinizirano i ekonomizirano da bi se moglo natjecati s ostalim metodama sekvencranja nove generacije.

4.2. „Solid-state" sekvenciranje nanoporama

Druga vrsta metoda sekvenciranja nanoporama temelji se na industrijski proizvedenim nanoporama. Ova tehnologija smatra se pravom sljedećom generacijom sekvenciranja jer nadilaze sve probleme koji se susreću kod bioloških nanopora poput stabilnosti ili teške manipulacije. Umjetne nanopore se proizvode u materijalima poput silicijevu nitridu, siliciju ili metalnim oksidima, a u novije vrijeme i u grafenu. Grafen je nov materijal koji se može proizvesti do debljine jednog atoma i predstavlja najtanju moguću membranu s visokom električnom provodljivošću. No za razliku od bioloških nanopora koje mogu razlikovati pojedine nukleotide, umjetne alternative nanopora u tome još posustaju zbog problema u kreiranju pora odgovarajućih veličina koje će omogućiti jak i točan signal pri prolazu pojedinog nukleotida uz što manje šumove okoline. Zbog navedenog, odgovarajuće nanopore s preciznom strukturom koje će omogućiti sekvenciranje bez obilježavanja nukleotida se još razvijaju.

5. Zaključak

Tehnologija i financiranje razvoja novih metoda sekvenciranja raste kao nikad do sad. Kao što je pokazano u ovom radu, razvilo se nekoliko bitno različitih načina sekvenciranja DNA molekula kroz sve generacije sekvencera. Svaka metoda ima svoje prednosti i nedostatke, tako da za odabir odgovarajuće tehnologije sekvenciranja potrebno je dobro poznavati primjenu podataka dobivenih sekvenciranjem. Od početaka razvoja metoda sekvenciranja nove generacije, cijena sekvenciranja pojedinih genoma nezaustavljivo pada i vjerojatno će uskoro doseći cijenu od 1000$ po genomu (von Bubnoff, 2008), čime će se znatno omogućiti razvoj personalizirane medicine. Doba osobnih genoma započeto razvojem sekvenciranja nove generacije predstavlja značajnu prekretnicu u genetičkim istraživanjima. Nije poznato koja će metoda sekvenciranja, i hoće li, prevladati u budućim istraživanjima, ali je gotovo sigurno da će daljnje smanjenje cijene sekvenciranja, kao i povećanje brzine i preciznosti uzrokovati da sekvenciranje postane temeljna metoda u svim aspektima bioloških znanosti.

6. Literatura

1. Berka, J. Chen, Y. J. Leamon J.H. i sur. Bead emulsion nucleic acid amplification. U.S. Patent Application, 2005.
2. von Bubnoff A. Next-generation sequencing: the race is on. Cell, 132:721–723. 2008.
3. Crick F. Central Dogma of Molecular Biology. Nature 227:561-563; 1970.
4. Dohm, J. C., Lottaz, C., Borodina, T. & Himmelbauer, H. Substantial biases in ultra-short read data sets from high-throughput DNA sequencing. Nucleic Acids Res. 36: e105, 2008.
5. Fuller CW, Middendorf LR, Benner SA, Church GM, Harris T, Huang X, Jovanovich SB, Nelson JR, Schloss JA, Schwartz DC, Vezenov DV. The challenges of sequencing by synthesis. Nat. Biotechnol. 27:1013–1023. 2009.
6. Foehlich i sur. High-throughput nucleic acid analysis. U.S. Patent, 2010.
7. Hall N. Advanced sequencing technologies and their wider impact in microbiology. J. Exp. Biol. 210:1518–25; 2007.
8. Margulies, M. i sur. Genome sequencing in microfabricated high-density picolitre reactors. Nature 437:376–380; 2005.
9. McKernan KJ. Sequence and structural variation in a human genome uncovered by short-read, massively parallel ligation sequencing using two-base encoding. Genome Res. 19:1527–1541; 2009.
10. Metzker ML. Sequencing technologies – the next generation. Nat. Rev. Genet; 11:31–46; 2010.
11. Rothberg JM. i sur. An integrated semiconductor device enabling non-optical genome sequencing. Nature 475, 2011.
12. Sanger F, Nicklen S, Coulson AR. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. 74 (12): 5463–7; 1977.
13. Sanger F, Coulson AR: A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase. J. Mol. Biol. 94 (3): 441–8; 1975.
14. Shendure J, Hanlee J: Next-generation DNA sequencing. Nat. Biotechnology 26: 1135 – 1145; 2008.
15. Watson JD, Crick FHC. A Structure for Deoxyribose Nucleic Acid. Nature 171:737-738; 1953.

7. Sažetak

Informacije pohranjene u DNA sekvenci sadrži vrijedne podatke o tom organizmu, i znajući sekvencu cijelog genoma, u mogućnosti smo razumjeti procese koji se odvijaju u tom biću. Tijekom proteklih nekoliko godina, masivno paralelne metode DNA sekvenciranja su postale široko dostupne te se cijena sekvenciranja značajno snizila. Značajno brži i jeftiniji načini sekvenciranja značajno mijenjaju ne samo brojne projekte uključene u sekvenciranje genoma, već mijenjaju tok bioloških znanosti i aplikacije informacija dobivenih sekvenciranjem. Prve razmijene metode masivno paralelnog sekvenciranja su metode koje su se bazirale na klonalnoj amplifikaciji fragmenata DNA sekvence. Umnažanjem fragmenata, sekvenceri primaju mnogo intenzivniji signal o ugrađenoj bazi, no sam proces amplifikacije podložan je brojnim greškama zbog čega je i finalna sekvenca manje precizno određena. Sljedeća, treća, generacija sekvencera koristi se samo jednom molekulom fragmenta i sekvenciranje se provodi u realnom vremenu. Ove platforme su često preciznije od metoda prijašnje generacije i mogu se koristiti za kvantitativna istraživanja. Najnovije i najperspektivnije platforme su one temeljene na sekvenciranju pomoću nanopora. Zbog malih dimenzija prostora u kojem se odvija određivanje slijeda baza, velike količine DNA molekula bi se mogle sekvencirati u jednoj reakciji, na jednom stroju. Ove metode su još uvijek u razvoju zbog brojnih problema s kojima su se susreli kreatori ovih sekvencera pri radu u ovako malim dimenzijama.

Ovaj rad detaljnije opisuje svaku od metoda masivno paralelnog sekvenciranja, navodeći prednosti i nedostatke svake od platformi kao i njihovu moguću primjenu.

8. Summary

Information stored in DNA sequence contains valuable data about that organism, and by knowing ones genome, we will be able to understand processes happening in that being. Over the past few years, massively parallel DNA sequencing platforms have become widely available dramatically reducing the cost of DNA sequencing. The fast and low-cost sequencing approaches not only change the landscape of genome sequencing projects but also usher in new opportunities for sequencing in various applications. First developed massively-parallel sequencers were those that used clonally amplified fragments of DNA. By amplifying the fragments, sequencers would receive much stronger signal of incorporated nucleotide, but amplification leads to many errors causing less accurate final sequence. Later developed, third generation of sequencers uses single molecule templates and sequencing is usually done in real time. These methods are more accurate but also they can be used in quantitative applications. The newest and the most promising platforms are ones based on nanopore sequencing. Because of small dimensions of sequencing space, enormous amount of DNA could be sequenced in just one reaction, on one sequencer. These methods are still under development because of lot of problems designers of these machines encountered when doing in such small space.

This thesis introduces into the high-throughput sequencing technologies, advantages and disadvantages of every platform and their biological application.