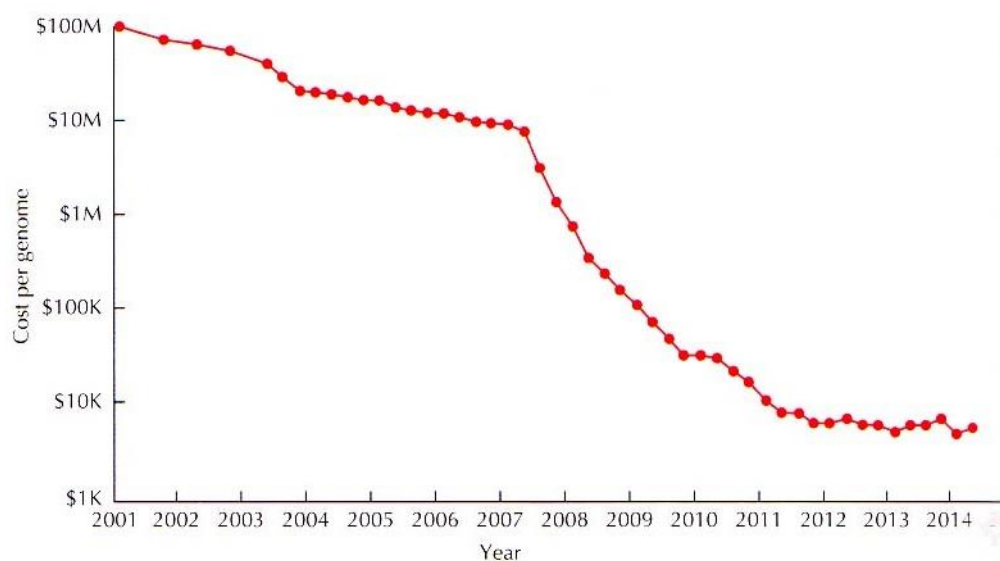


Sekvenciranje sljedeće generacije

Ljudski genom, kao i većina genoma modelnih organizama, su sekvencirani tehnikom dodavanja dideoksinukleotida koju je prvi opisao Fred Sanger 1977. godine. Automatizacija ove osnovne metode je omogućila brže i efikasnije sekvenciranje cijelog genoma, a kulminirala je uspješnim završetkom projekta ljudskog genoma 2004. godine. Međutim, unatoč robustnoj automatizaciji, sekvenciranje ovim pristupom je bilo sporo i skupo, tako da je sekvenciranje kompletnog ljudskog genoma bio veliki pothvat. Početno sekvenciranje ljudskog genoma trajalo je 15 godina uz trošak od oko 3 milijarde dolara. Počevši od 2005. godine, razvijene su nove metode sekvenciranja, nazvane 'next-generation sequencing' odnosno sekvenciranje sljedeće generacije, koje su značajno povećale brzinu i općenito smanjile troškove sekvenciranja genoma (*Slika 1*). Od 2001. godine, troškovi sekvenciranja ljudskog genoma su smanjeni više od 10.000 puta - od približno 100 milijuna na nekoliko tisuća.



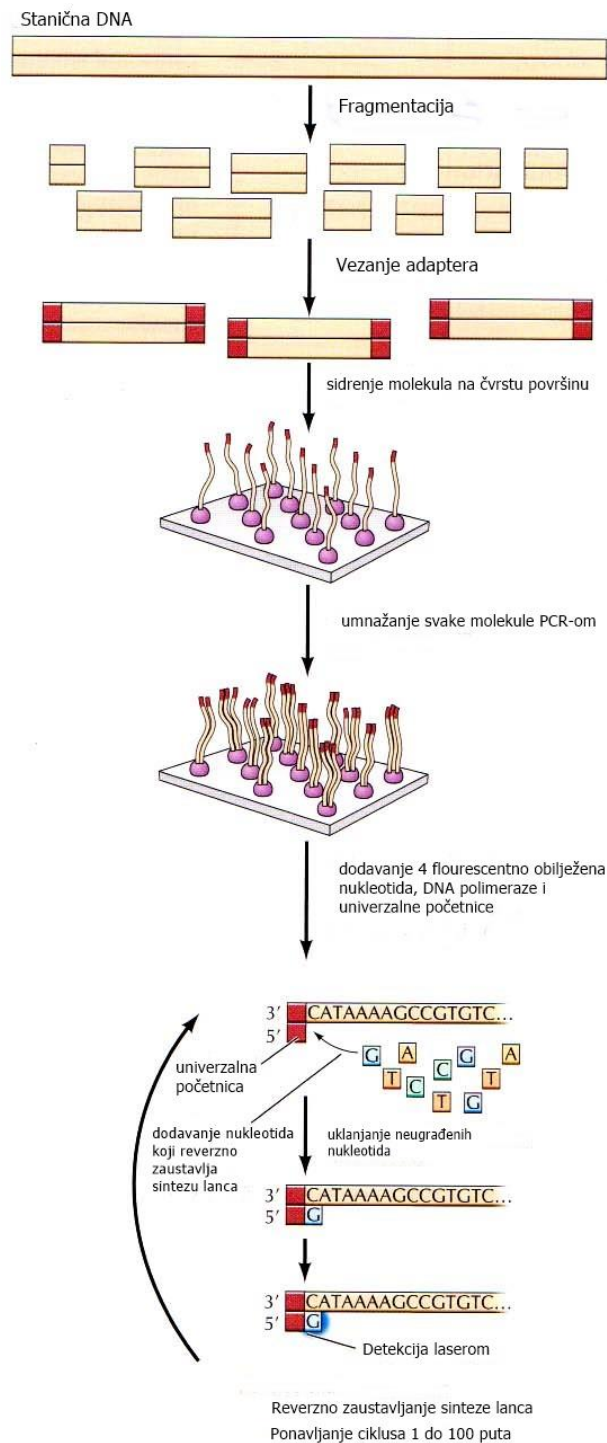
Slika 1. Od 2001. godine trošak sekvenciranja ljudskog genoma pao je s oko 100 milijuna na nekoliko tisuća dolara do 2015. godine

Brzina sekvenciranja se još više povećala, tako da je sada moguće sekvencirati čitav ljudski genom u samo nekoliko dana. Te dramatične promjene tehnologije sekvenciranja otvaraju vrata za sekvenciranje čitavih genoma velikog broja različitih pojedinaca, omogućujući nove pristupe u razumijevanju genetske osnove za mnoge bolesti koje pogađaju čovječanstvo, uključujući rak, bolesti srca, degenerativne bolesti živčanog sustava kao što su Parkinsonova i

Alzheimerova bolest. Osim toga, bolje razumijevanje našeg jedinstvenog genetskog materijala dovodi do razvoja novih, prilagođenih strategija za prevenciju i liječenje bolesti.

Sekvenciranje sljedeće generacije (koje se još naziva i masivno paralelno sekvenciranje) odnosi se na sekvenciranje DNA koje uključuje nekoliko različitih metoda u kojima se milijuni predložaka (engl. *template*) sekvenciraju istodobno u jednoj reakciji. Opća strategija tih metoda ilustrirana je na slici 2.

U prvom koraku se DNA fragmentira, zatim se dodaju aditivne sekvence (oligonukleotidi koji služe kao početnice za umnažanje i sekvenciranje), koje se vežu na krajeve svakog fragmenta. Jednolančani fragmenti DNA se zatim vežu na čvrstu površinu i umnažaju PCR-om kojim se proizvode milijuni prostorno odvojenih klastera, odnosno, grupiranih DNA fragmenata (predložaka). Milijuni predložaka se potom mogu paralelno sekvencirati pomoću lasera koji prate ugrađivanje fluorescentnih nukleotida. Sekvence dobivene iz ove velike zbirke preklapajućih fragmenata se mogu zatim sastaviti kako bi se dobila kontinuirana genomska sekvenca. Alternativno, ako je dostupna već poznata genomska sekvenca (na primjer, ljudski genom), ista se može koristiti kao referentni genom pomoću kojeg se sekvence fragmenata od određenog pojedinca mogu poravnati.



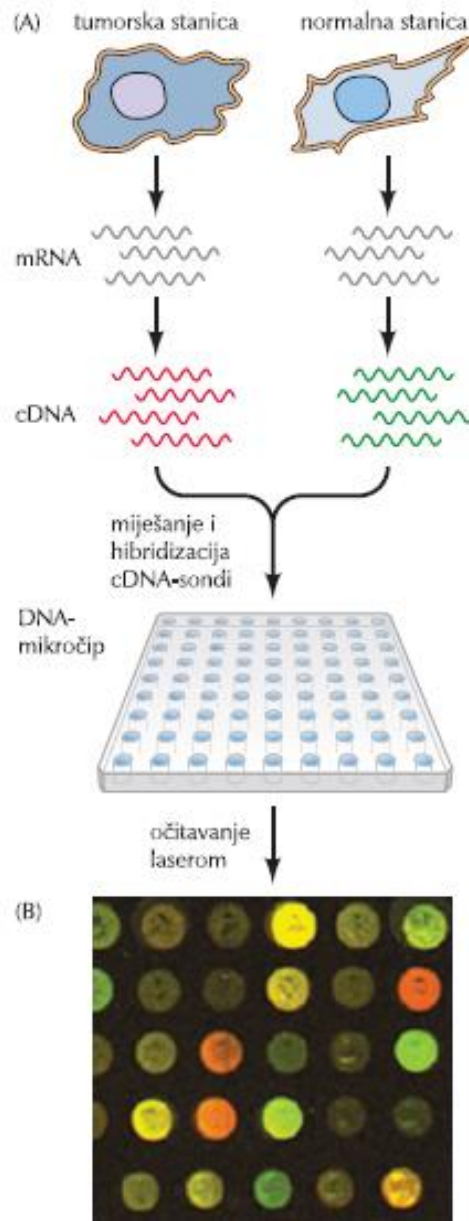
Slika 2. Sekvenciranje sljedeće generacije (sekvenciranje sintezom). Stanična DNA se fragmentira i aditivne sekvence se vežu na krajeve svakog fragmenta. Jednolančane molekule se zatim vežu na čvrstu površinu i umnažaju PCR-om, stvarajući milijune klastera molekula. Skupa se dodaju četiri fluorescentno obilježena nukleotida, koji reverzno zaustavljaju sintezu lanca, DNA polimeraza i početnica koja prepoznaje aditivnu sekvencu. Ugradnja obilježenog nukleotida u svaki klaster DNA molekula detektira se pomoću lasera. Neugrađeni nukleotidi se uklanjaju, zaustavljanje sinteze lanca je reverzno, i ciklus se ponavlja da bi se dobila sekvenca milijuna klastera istovremeno.

Prvi pojedinačni ljudski genom koji je sekvencioniran objavljen je 2007. i 2008. godine, uključujući i genome Craiga Ventera i James Watsona. Od tada, pa do danas, određene su sekvence genoma tisuća pojedinaca. Uz cijenu sekvenciranja pojedinog genoma u rasponu od nekoliko tisuća dolara, može se očekivati da će sekvenciranje pojedinačnog genoma postati dio medicinske prakse što će omogućiti da terapija bude prilagođena potrebama pojedinaca. Najbolji primjer za to je razvoj novih lijekova za liječenje raka, koji su specifično usmjereni na mutacije koje se mogu identificirati sekvenciranjem genoma karcinoma kod pojedinih pacijenata. U budućnosti možemo očekivati da će sekvenciranje genoma zdravih ljudi imati važnu ulogu u prevenciji bolesti identificiranjem gena za određenu bolest, time će se moći poduzeti i odgovarajuće mjere za prevenciju bolesti. Na primjer, sekvenciranjem genoma mogle bi se identificirati specifične genske mutacije kod žena koje imaju visoki rizik za razvoj karcinoma dojke, a koje bi se mogle spriječiti mastektomijom. Daljnji nastavak napretka u genomici dovest će do njezine veće primjene u medicini, ali će nam također pomoći da razjasnimo doprinos naših gena i drugih jedinstvenih karakteristika, kao što su atletska sposobnost ili inteligencija, i doprinijeti boljem razumijevanju interakcija između gena i okoliša koji vode do složenih ljudskih ponašanja.

Globalna analiza ekspresije gena

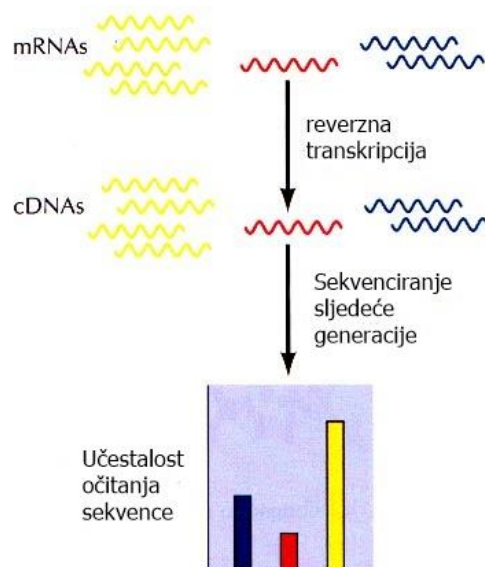
Dostupnost cijelih genomskih sekvenci je omogućila istraživačima proučavanje ekspresije gena na cjelogenomskoj razini. Danas je moguće analizirati svu RNA koja je transkribirana u stanici (transkriptom) umjesto da se analizira jedan po jedan gen. Jedna od najčešće korištenih metoda za globalnu analizu ekspresije gena je hibridizacija s DNA-mikročipovima, koja omogućuje istovremenu analizu desetina tisuća gena. DNA-mikročip se sastoji od predmetnog stakalca ili membrane na koju su, uz pomoć robotiziranih sustava, fragmenti oligonukleotida utisnuti u velikom broju malih točaka visoke gustoće (Slika 4.27, 5. izdanje knjige). U svakoj se točki na mikročipu nalazi jedan oligonukleotid. Na klasično predmetno mikroskopsko stakalce moguće je utisnuti više od 100.000 jedinstvenih sekvenci DNA, što omogućuje pripremu DNA-mikročipova koji pokrivaju cjelokupne genome. Kao što je prikazano na slici 4.27, DNA-mikročipove moguće je primijeniti za analizu ekspresije gena, primjerice za uspoređivanje gena koje eksprimiraju dva različita tipa stanica. U eksperimentima ovog tipa sintetiziraju se cDNA probe od ukupne mRNA koja je eksprimirana u dva tipa stanica (primjerice tumorskim i normalnim stanicama). Te dvije skupine cDNA obilježe se različitim fluorescentnim bojama (obično crvenom i zelenom), pomiješaju i hibridiziraju s mikročipovima na kojima se nalazi

20.000 ili više humanih gena predstavljenih kao pojedinačne točkice. DNA-mikročip se tada analizira laserskim čitačem visoke rezolucije, a relativni omjeri transkripcije određenog gena u tumorskim odnosno normalnim stanicama odgovaraju omjeru intenziteta fluorescencije u odgovarajućim točkicama na DNA-mikročipu.



Slika 4-27. **DNA-mikročipovi.** (A) Primjer usporedne analize ekspresije gena u tumorskim i normalnim stanicama. mRNA izolirana iz normalnih i tumorskih stanica koristi se kao kalup za sintezu cDNA-sondi koje se zatim obilježavaju različitim fluorescentnim bojama (primjerice crveni fluorescentni biljeg za tumorske stanice i zeleni za normalne stanice). Pomiješane dvije različite cDNA hibridiziraju s DNA-mikročipom. Na mikročipu se nalaze točkice s oligonukleotidima koje predstavljaju 10.000 ili više humanih gena. Relativna ekspresija svakoga gena u tumorskoj stanici u odnosu na njegovu ekspresiju u normalnoj stanici oslikana je omjerom crvene i zelene fluorescencije na odgovarajućim položajima na DNA-mikročipu. (B) Fotografija dijela DNA-mikročipa.

Stalni razvoj sekvenciranja sljedeće generacije omogućuje određivanje i kvantificiranje svih RNA eksprimiranih u stanici. Ovim pristupom, nazvan *RNA-seq*, stanične mRNA se izoliraju, pretvaraju u cDNA reverznom transkripcijom i podvrgavaju sekvenciranju sljedeće generacije (Slika 3). Za razliku od analize mikročipovima, *RNA-seq* otkriva cjelokupni sadržaj transkribiranih sekvenci u stanici, a ne samo onih koji hibridiziraju s probom na mikročipu. Frekvencija kojom se pojedinačne sekvence detektiraju u *RNA-seq* je proporcionalna količini RNA u stanici, tako da ova analiza određuje i ukupnu zastupljenost te identitet svih transkripcijskih sekvenci. Osjetljivost *RNA-seq* je dovoljno visoka da omogući analizu na razini stanice, tako da se mogu odrediti transkriptomi pojedinih stanica. Jedna od zanimljivosti otkrivenih analizom *RNA-seq* je ta da se transkribira mnogo više RNA od očekivane koju kodiraju protein-kodirajući geni jedne ljudske stanice.



Slika 3. ***RNA-seq***. Stanične mRNA se reverzno transkribiraju u cDNA, i podvrgavaju sekvenciranju sljedeće generacije. Rezultati prikazuju redoslijed svih mRNA u stanici. Relativna količina svake mRNA je prikazana učestalošću kojom je ta sekvenca zastupljena u ukupnom broju dobivenih sekvenci.

*Tekst preveden iz knjige "The Cell - A Molecular Approach" (Cooper GM, Hausman RE, 7.izdanje, str. 164-167)

**Tekst preveo: Ivana Gunjača, dipl. ing. mol. bio.