

Razvoj novih tehnologija sekvenciranja i njihova primjena u istraživanju genoma domaćih životinja

Development of new sequencing technologies and their application in genome analysis of domestic animals

Gvozdanović, K., Čuljak, V., Margeta, P.

Poljoprivreda/Agriculture

ISSN: 1848-8080 (Online)

ISSN: 1330-7142 (Print)

DOI: <http://dx.doi.org/10.18047/poljo.21.2.11>



Poljoprivredni fakultet u Osijeku, Poljoprivredni institut Osijek

Faculty of Agriculture in Osijek, Agricultural Institute Osijek

RAZVOJ NOVIH TEHNOLOGIJA SEKVENCIRANJA I NJIHOVA PRIMJENA U ISTRAŽIVANJU GENOMA DOMAĆIH ŽIVOTINJA

Gvozdanović, K.⁽¹⁾, Čuljak, V.⁽²⁾, Margeta, P.⁽¹⁾

Stručni članak
Professional paper

SAŽETAK

Sekvenciranje i detaljno istraživanje genoma domaćih životinja započeto je sredinom prošloga stoljeća. U prvome redu to se odnosilo na razvoj metoda prve generacije sekvenciranja, odnosno Sangerovu metodu sekvenciranja. Primjena tehnologijama nove generacije u analizi genoma domaćih životinja trenutno je najzastupljenija metoda sekvenciranja životinjskoga genoma. Primjenom tih metoda dobiva se i do 100 puta više podataka u usporedbi sa Sangerovom metodom sekvenciranja. Razvoj novih tehnologija sekvenciranja od 2005. godine do danas omogućile su provođenje analiza koje uključuju RNK sekvenciranje, genotipiziranje cijeloga genoma, imunoprecipitaciju povezanu s DNK mikročipovima, detektiranje mutacija i nasljednih bolesti i sekvenciranje mitohondrijskoga genoma. Primjena novih metoda sekvenciranja u analizi genoma domaćih životinja otvara vrata prema boljem razumijevanju genetske osnove proizvodnih svojstava važnih za unaprijeđenje stočarske proizvodnje.

Ključne riječi: sekvenciranje, Sangerova metoda sekvenciranja, tehnologija sekvenciranja nove generacije, analiza genoma domaćih životinja

UVOD

Razvoj metoda sekvenciranja započeo je Frederick Sanger tijekom 1975. godine. Razvio je brzu metodu sekvenciranja DNK pomoću dideoksinukleotida i DNK polimeraze. Prva istraživanja sekvence DNK koju je proveo Sanger temeljila se na enzimatskome dideoksi sekvenciranju uz terminaciju novo sintetiziranoga lanca pomoću dideoksinukleotidnih analoga (Metzker, 2005.). Nakon toga, tijekom 1977. godine, Allan Maxam i Walter Gilbert predstavili su novu metodu sekvenciranja DNK sekvence u kojoj su DNK fragmenti obilježeni specifičnim bazama i razdvojeni gel elektroforezom. Navedena istraživanja prethodila su razvoju prvoga automatiziranoga DNK sekvencera (Pareek i Smoczynski., 2011.). Primjena je Sangerove metode dugotrajna, skupa te zahtijeva veći utrošak ljudskoga rada. S obzirom na to da je danas potreba za novim informacijama i podacima velika, potrebno je razvijati nove metode analize genoma (Brajušković, 2006.). Nedostatak informacija o genetskim sekvencama te odgovarajuće platforme za genotipizaciju visoke gustoće razlog su manje količine informacija o genima i mutacijama koje dovode do pojave određenih proizvodnih svojstava.

Procjena je da se neravnoteža povezanosti (linkage disequilibrium, LN) kod svinja proteže na nekoliko stotina kilobaza te da bi 30.000-50.000 jednonukleotidnih polimorfizama (*Single Nucleotide Polymorphism, SNP*) bilo potrebno za analizu genoma domaćih životinja. Provedbu genome-wide analiza u genomu svinja omogućit će korištenje markera visoke gustoće, dostupnost potrebnih sekvenci te ekonomski isplativo SNP genotipiziranje. Upravo je nedostatak prikladnih SNP-ova za provedbu genotipiziranja ograničavajući čimbenik za razvoj visoko paralelnih testova genotipizacije kod svinja (Wiedmann i sur., 2008.). Razvoj novih tehnologija sekvenciranja genoma otvara mogućnost njihove primjene u rekvenciranju, analizi genske ekspresije, analizi mitohondrijske RNK(miRNA), istraživanju metilacije DNK te sekvenciranju transkriptoma (Robertson i sur., 2007.; Hilier i sur., 2008.). Analizom se dobiva velika količina podataka, odnosno veliki broj kratkih sekvenci veličine 30-40 baznih

(1) Kristina Gvozdanović, mag. ing. agr. (kbudimir@pfos.hr), dr. sc. Polonca Margeta – Sveučilište Josipa Jurja Strossmayera u Osijeku, Poljoprivredni fakultet u Osijeku, Kralja Petra Svačića 1d, 31000 Osijek, Hrvatska (2) Vice Čuljak, dr. vet. med., Ministarstvo poljoprivrede, Uprava za veterinarstvo i sigurnost hrane, Veterinarski ured Osijek, Trg Lava Mirskog 3, 31000 Osijek, Hrvatska

parova po jednoj analizi. Veliki broj podataka zahtijeva učinkovite algoritme koji bi ih mogli obraditi i osigurati visoki stupanj točnosti (Li i sur., 2008.).

Iako je primjena novih tehnologija sekvenciranja bila usmjerena na analizu cijeloga genoma, različitim pristupima omogućene su analize pojedinih regija genoma. Utjecaj novih tehnologija sekvenciranja vidljiv je u području genomike. To se odnosi na detektiranje pojave varijacija i *de novo* mutacija, somatskih varijacija te rijetkih genetskih bolesti koje se mogu pojaviti u populaciji ljudi (Koboldt i sur., 2013.). Primjena novih metoda obuhvaća široko područje, od analiza varijacija sekvenci, komparativne genomike i evolucije, forenzike do biotehnologije. Tijekom zadnjih 5 godina došlo je do razvoja platformi za ultra brzo sekvenciranje DNK. Neke od tih metoda su sekvenciranje hibridizacijom (*Sequencing By Hybridization*, SBH), sekvenciranje nanoporama te sekvenciranje sintezom (*Sequencing By Synthesis*, SBS). Sekvenciranje sintezom uključuje primjeru različitih metoda s DNK polimerazama. Metzker (2005.), stoga predlaže podjelu SBS metoda u tri kategorije: Sangerovo sekvenciranje, dodatak jednoga nukleotida (Single Nucleotide Addition, SNA) i ciklički reverzibilna terminacija (Cyclic Reversible Termination, CRT) (Metzker, 2005.; Ries-Filho, 2009.). Budućnost genomike i razvoj novih biološko molekularnih metoda ima budućnost u agrogenomici, medicini i farmakogenomici (Wang i Weinshilbom, 2008., Budimir i sur., 2013.).

Cilj je ovoga rada prikaz povijesnoga razvoja metoda sekvenciranja te usporedba različitih generacije sekvenciranja, uz prikaz njihovih primjena pri analizi genoma domaćih životinja.

TEHNOLOGIJA PRVE GENERACIJE

Sangerova metoda, odnosno metoda terminacije sinteze lanca (Chain Termination Method), temelji se na mogućnosti razdvajanja jednolančanih DNK molekula koje se razlikuju u jednome nukleotidu pomoću gel elektroforeze. Sanger i sur. (1980.) sekvencirali su genom bakteriofaga ϕ X174, čija je sekvenca bila velika 5375 nukleotida. Osnovni princip Sangerove metode sekvenciranja korištenje je 2,3-dideoksinukleotid trifosfata (ddNTP) i DNK polimeraze. S obzirom na to da DNK polimeraza ne pravi razliku između dNTP-a i ddNTP-a, ddNTP-i mogu biti ugrađeni u rastući lanac DNK molekule vezanjem 5' fosfatne grupe i 3' hidroksilne grupe posljednjega nukleotida u rastućem lancu. Do terminacije sinteze rastućega lanca ne dolazi zbog odsustva 3' hidroksilne skupine na ddNTP-u, što onemogućava stvaranje fosfodiesterne veze. Specifična terminacija lanca *in vitro* omogućena je korištenjem ddNTP-a i specifičnim djelovanjem DNK polimeraze (Metzker, 2005.; Metzker, 2009.). Walter Gilbert i Alan Maxam (1977.) razvili su metodu kemijske degradacije DNK, u kojoj su fragmenti DNK degradirani te razdvojeni primjenom gel elektroforeze (Brajšković, 2006.; Pareek i Smoczynski, 2011.). Istraživanja koja su proveli Sanger te Maxam i Gilbert rezultirali su uvođenjem prvoga automatiziranoga DNK

sekvencera (Caltech). Modifikacijom Sangerove metode sekvenciranja te primjenom termostabilnih DNK polimeraza i razvojem metode lančane reakcije polimeraze (Polymerase Chain Reaction, PCR) omogućen je razvoj metode sekvenciranja termalnim ciklusima. Danas je metoda modificirana u toj mjeri da se DNK sekvence sekvenciraju u automatskim sekvencerima. Primjerice, korištenjem prvoga automatiziranoga DNK sekvencera određena je sekvenca HPRT gena (Hypoxanthineguanine Phosphoribosyl Transferase gene).

Primjenom sekvencera smanjene su pogreške koje se mogu javiti tijekom određivanja slijeda nukleotida u sekvenci. Princip rada sekvencera temelji se na određivanju zadnjega ddNTP-a, koji je obilježen fluorescentnom bojom. Razlikuju se dva tipa automatskih sekvencera, a to su automatski sekvencer, koji koristi gel u obliku ploče, i sekvencer u kojem se gel nalazi u obliku cjevčice. Prvi komercijalni DNK sekvenceri, koji su razvijeni tijekom 1996. godine, u svome su radu koristili gel u obliku ploče, dok se dvije godine kasnije razvijaju uređaji kod kojih je pločasti gel zamijenjen kapilarama punjenim polimernim matriksom. S obzirom na veliki broj podataka koji se dobivaju primjenom automatskih sekvencera, nužno je razviti i računalne programe koji obrađuju dobivene podatke (Hutchison III, 2007.).

TEHNOLOGIJA DRUGE GENERACIJE

Roche 454 sustav

Roche 454 prvi je uspješno implementiran sustav sekvenciranja nove generacije. U svome radu koristi tehnologiju pirosekvenciranja. Tehnologija pirosekvenciranja temelji se na detektiranju pirofosfata koji se oslobađaju prilikom ugradnje nukleotida (Harismendy i sur., 2009.).

Tijekom procesa analize dolazi do denaturacije DNK u jednolančani oblik te amplifikacije. Nakon toga procesa komplementaran dNTP spaja se s uzvojnicom pomoću ATP sulfurilaze, luciferase, luciferina, DNK polimeraze i adenozin 5 fosfosulfata te se oslobađa pirofosfat. Pretvorba luciferina u oksiluciferin dovodi do generiranja vidljive svjetlosti. Istovremeno, apiraza razgrađuje nesparene baze. Zadnji korak uključuje dodavanje novih dNTPa te započinjanje novogaciklusa reakcija (Foehlich, 2010.). Dužina čitanja sekvenci toga sustava kreće se od 100 do 250 baznih parova te do 200 000 baza. Razvojem sustava rasla je i dužina čitanja te količina informacija koju je sustav mogao dati. Tako je tijekom 2009. godine razvijen sustav koji je omogućavao dobivanje i do 14G informacija po čitanju (Huse i sur., 2007.; Metzker, 2009.). Roche sustav posjeduje nekoliko značajnih prednosti u odnosu na ostale sustave, a prvi je njegova brzina. Potrebno je tek 10 sati od početka sekvenciranja do njegovoga završetka. Troškovi analize relativno su visoki te iznose 12,56 \$ za 10^6 baza. Nedostatak je toga sustava mogućnost stvaranja pogrešaka koje se javljaju na polibazama, koje su duže od 6 baznih parova. Softver Roche 454 sustava (GS RunProcessor) zadužen

je za generiranje podataka za sekvenciranje, konverziju signala, korekciju signala, korekciju područja signala te normalizaciju pozadinske slike. Nakon provedene analize, sustav proizvodi niz datoteka (Standardni Flowgram Format, SFF).

AB Solid sustav

AB solid sustav (Sequencing by Oligonucleotide Ligation and Detection, SOLID) temelji se na tehnologiji sekvenciranja dviju baza, kojima je temelj sekvenciranje ligacijom. Sekvenciranje uključuje vezanje DNK s osam baznom probom, koja sadrži vezivno mjesto za prvu bazu, mjesto cijepanja za petu bazu te četiri različite fluorescentne boje, koje su vezane sa zadnjom bazom. Fluorescencija će biti zabilježena nakon vezanja DNK i komplementarnih proba (Metzker, 2009.). Nakon sekvenciranja dolazi do akumuliranja originalne obojane sekvence. Boje odgovaraju baznim parovima, a dekodiranje originalne obojane sekvence moguće je samo u slučaju poznavanja jedne baze na određenoj poziciji u sekvenci. Pogreška u bojanju dovest će do pogrešaka u dekodiranju sekvence (Mardis, 2008.). Uspješnost izvođenja metode sekvenciranja primjenom AB solid sekvencera iznosi 99,85 %. Troškovi su implementacije toga sustava visoki, zahtijevaju uvođenje računalnih sustava, klimatiziranoga prostora u kojemu se obavlja analiza podataka, osoblja osposobljenoga za analizu dobivenih podataka, brze mreže te velike memorije sustava. Troškovi sekvenciranja iznose 40 \$ za 10^9 baznih parova, dok dužina trajanja analize može potrajati i do 7 dana. Glavni su nedostaci toga načina sekvenciranja dužina trajanja analize te veličina sekvence. AB solid sustav koristi se u rekvenciranju cijeloga genoma, ciljanom rekvenciranju te analizi epigenoma. Osim toga, koristi se pri analizi RNK, profiliranju genetske ekspresije te analizi transkriptoma. Softver za taj sustav

je BioScope, a omogućava CHIP-sekvenciranje, analizu cijeloga transkriptoma te pruža okvir za obavljanje rekvenciranja. Radi u dva koraka, u prvome koraku konvertira baze u obojane sekvence, dok drugi korak uključuje uspoređivanje obojane sekvence s originalnom sekvencom, kako bi se dobili podaci o mapiranju s novim algoritmom za mapiranje.

Illumina sustav

Tehnologija Illumina sekvencera temelji se na sekvenciranju sintezom (*Sequencing By Synthesis*, SBS). Metoda se izvodi tako da se u prvome koraku radi denaturacija DNK i fiksnih proba, nakon čega slijedi amplifikacija u oblik klastera koji sadrži umnožene DNK fragmente. U reakciji sudjeluju linearizirani enzimi, 4 različita nukleotida s različitim fluorescentnim bojama (ddATP, ddGTP, ddCTP, ddTTP) te odvojiva blokirajuća grupa koja će dodavati bazu po bazu, dok će signal biti zabilježen CCD senzorom (Charge-Coupled Device, CCD). Postoji nekoliko različitih modela toga sustava. Početkom 2010. godine razvijena je Illumina HiSeq 2000. Taj je sustav jeftiniji u usporedbi s Roche 454 i solid sustavom, a cijena koštanja je 0,2 \$ za 10^6 baza. Trajanje analize je 10 dana, a njezin output je 600G. Nakon završetka analize, postotak pogreške manji je od 2%. Da bi Illumina HiSeq 2000 bio operativan, potrebno mu je osigurati softver, sustav za kontrolu programa, real-time analizator te programe za backup analize. Najnoviji razvijeni sustav je Illumina MiSeq, lansiran na tržište 2011. godine (Metzker, 2009.; Ledergerber i Dessimoz, 2011.).

Uspoređujući tri opisana sustava za analizu, dolazimo do zaključka da Illumina HiSeq 2000 ima najviše outpute uz najniže troškove, solid sustav daje najveću točnost, dok je uz primjenu Roche 454 moguće analizirati najduže sekvence.

Tablica 1. Karakteristike sekvencera tehnologije treće generacije (Liu i sur., 2012.)

Table 1. Characteristics of third generation technology sequencers (Liu et al., 2012)

Sekvencer <i>Sequencer</i>	Roche 454	SOLiD™4	Illumina HiSeq 2000	Sanger 3730xl
Tehnologija sekvenciranja <i>Sequencing technology</i>	Pirosekvenciranje <i>Pyrosequencing</i>	Sekvenciranje sintezom <i>Sequencing by synthesis</i>	Sekvenciranje ligacijom <i>Sequencing by ligation</i>	Dideoksi terminacija lanca <i>Dideoxy chain termination</i>
Dužina čitanja <i>Read length</i>	700 bp	50SE, 50PE, 101PE	50+35 bp 50+50 bp	400-900 bp
Točnost <i>Accuracy</i>	99,9%	98%	99,94%	99,999%
Izlazni podaci <i>Output data</i>	0,7G	600 Gb	120 Gb	1,9-84 kb
Vrijeme <i>Time</i>	24 h	3-10 dana	7-10 dana	20 min-3sata
Cijena uređaja <i>Instrument price</i>	500,000 \$	690,000 \$	495,000 \$	95,000 \$
Troškovi/milijun baza <i>Cost/milion bases</i>	10 \$	0,07 \$	0,13 \$	2400 \$

TEHNOLOGIJA TREĆE GENERACIJE

Tehnologija treće generacije ima dvije glavne karakteristike. Prije njezinog izvođenja nije potrebno izvođenje PCR reakcije, što značajno skraćuje vrijeme potrebno za provođenje analize. Druga karakteristika je svojstvo signala koji se bilježi u realnome vremenu, bez obzira bio on fluorescentan ili električni. Signal se prati tijekom enzimatske reakcije dodavanja nukleotida u komplementarni lanac (Ledergerber i Dessimoz, 2011.).

Pacific Bioscience razvio je jednomolekularni real-time SMRT sustav (*Single molecule real time sequencing*, SMRT) koji koristi modificirane enzime te promatra enzimatsku reakciju u trenutku njezinog odvijanja. Reakcija se odvija tako da enzim spaja nukleotide u komplementarni lanac te uklanja fluorescentnu boju prethodno vezanu s nukleotidom. Uređaj bilježi promjenu u realnome vremenu (Timp i sur., 2010.). Princip rada toga uređaja pogodan je za predviđanje strukturnih varijacija u sekvenci te za epigenetska istraživanja koja se odnose na metilaciju DNK (Lam i sur., 2012.). Prednosti su toga načina analize u tome što je priprema uzorka kratka te traje tek od 4 do 6 sati, nije potrebno provođenje PCR reakcije, što skraćuje vrijeme izvođenja reakcije, dužina čitanja iznosi 1300 baznih parova, što je značajno više nego kod tehnologije druge generacije, te trajanje analize, koje može biti do najviše 24 sata.

Nanoporno sekvenciranje predstavlja drugu metodu treće generacije sekvenciranja. Takav način sekvenciranja svoj rad temelji na iskorištavanju biološke funkcije nanopora. Nanopore se mogu naći u proteinskim kanalima uklopljene u lipidni dvosloj, olakšavajući nabijenim česticama prolazak iz stanice. Sekvenciranje se bazira na uklapanju jednolančane DNK u α -hemolizin (α HL). α -hemolizin je izoliran iz *Staphylococcus aureus* te predstavlja protein velik 33kD. Može podnijeti voltaže do 100mV te je, zbog toga, pogodan za korištenje u izgradnji bloka nanopora. Pri izvođenju metode protok se iona odvija kontinuirano, a svaki poremećaj toga toka vrlo je jednostavno otkriti. Očitavanje se temelji na razlikama između svih deoksiribonukleotid monofosfata (dNMP) (Branton i sur., 2008.). Prednosti su nanopornoga sekvenciranja u potencijalu čitanja, koji iznosi >5 kbp., jednolančane DNK uzvojnice sekvencira se kroz nanopore pomoću depolimerizacije DNK lanca, nanoporno sekvenciranje manje je temperaturno osjetljivo te vrijeme pripreme uzorka, koje je znatno kraće, a to se posebno se odnosi na korake kloniranja i amplifikacije (Clarke i sur., 2009.).

Uspoređujući metode sekvenciranja druge i treće generacije, mogu se uočiti neke značajne razlike. Metode druge generacije u svojoj analizi uključuju provedbu PCR reakcije. Metode treće generacije posjeduju potencijal za bolje iskorištavanje DNK polimeraze, posebice njezine katalitičke aktivnosti, te smanjivanje vremena potrebnoga za dobivanje rezultata analize. Prednosti su treće generacije sekvenciranja nad drugom generacijom veća propusnost, brže provođenje analize, veća dužina sekvence, izravna detekcija haplotipova te veća pre-

ciznost analize, koja omogućava otkrivanje varijacija, i mala količina potrebnog uzorka (Whiteford i sur., 2009.). Oba sustava primjenjuju tehnologiju sekvenciranja sintezom, no, za razliku od tehnologije druge generacije, koja obrađuje mnogo kopija DNK predloška, tehnologija treće generacije radi s DNK molekulama u realnome vremenu.

PRIMJENA TEHNOLOGIJE NOVE GENERACIJE SEKVENCIRANJA

Razvoj novih tehnologija sekvenciranja doveo je do povećavanja točnosti analize te snižavanja troškova njezinoga provođenja. One svoju primjenu nalaze u metagenomici, epigenomici te transkriptomici (Budimir i sur., 2014.). Podaci dobiveni primjenom tih tehnologija koriste se u stočarstvu, dijagnostici, ratarstvu i terapijama. *De novo* sekvenciranje uključuje fragmente dužine do 800 baznih parova, a važni su u rekvenciranju. Sekvenciranjem cijeloga genoma može se dobiti DNK sekvenca genoma organizma. Resekvenciranjem cijeloga genoma može se provoditi istraživanje funkcionalnih gena, no glavna prepreka takvome načinu analize visoki su troškovi njihovoga provođenja. Rješenje je problema sekvenciranje pojedinih regija, primjerice korištenjem mikročipova u istraživanju. Primjena u epigenomici je u istraživanju imunoprecipitacije kromatina i metilacije DNK. Imunoprecipitacija kromatina tehnika je imunoprecipitacije, koja se koristi u istraživanjima interakcija između proteina i DNK u stanici (Johnson i sur., 2007., Budimir i sur., 2013.). RNK sekvenciranje koristi se u analizi ekspresije mRNK. Primjena tehnologije sekvenciranja nove generacije je i u analizi transkriptoma, odnosno skupa svih RNK molekula (mRNK, tRNK, rRNK te ostalih nekodirajućih RNK) (Marioni i sur., 2008.; Fehniger i sur., 2010.). Tijekom 2009. godine objavljeno je istraživanje s rezultatima koji su pokazali da je 98% genoma svinje sekvencirano. Genom svinje veličine je $2,7 \times 10^9$ baznih parova (Bai i sur., 2012.). Sekvenciranje je i opisan pod vodstvom Konzorcija za sekvenciranje genoma svinje. Strategija sekvenciranja kombinira sekvenciranje metodom sačmarice BAC klonova i sekvenciranje cijeloga genoma (Schook i sur., 2005.; Archibald i sur., 2010.). Svinja ima veliko značenje zbog važnosti u prehrambenoj industriji, no ona je i važan model u biomedicinskim istraživanjima. Istraživanje koje su proveli Corominas i sur. (2013.) uključivalo je analizu 55 gena diferencijalno eksprimiranih u jetri, koji imaju ključnu ulogu u sastavu mišićnih lipida. Istraživanje je uključivalo križance iberijske svinje i landrasa, a metoda koja se koristila u istraživanju bila je RNK sekvenciranje. Metodu RNK sekvenciranja pri genetskim analizama svinja koristili su i drugi istraživači (Jung i sur., 2012.; Ramayo-Caldas i sur., 2012.; Esteve-Codina i sur., 2011.). Dostupno je relativno malo informacija o varijacijama sekvenci u genomu svinje. Metode pirosekvenciranja omogućit će dobivanje dovoljnoga broja informacija za dizajniranje navedene metode analize genoma. Istraživanjem koje su proveli Wiedmann i sur. (2008.) detektirano je 115,572 SNPova. Analiza varija-

cija broja kopija može unaprijediti istraživanja genetske raznolikosti, funkcionalne genomike i evolucije. Takva istraživanja već su provodena genomu čovjeka, a tijekom zadnjih godina genomima domaćih životinja: goveda, ovce, koze, kokoši, pure i svinje. Istraživanja genoma svinje uključivala su genotipiziranje i rekvenciranje genoma, odnosno metode koje se temelje na tehnologiji sekvenciranja nove generacije. Jiang i sur. (2014.) proveli su istraživanje koje je uključivalo kineske autohtone pasmine svinja, azijske divlje svinje te tri komercijalne pasmine svinja, a temeljilo se na primjeni metoda rekvenciranja. Ramos i sur. (2009.) proveli su istraživanje kojim su željeli detektirati veći broj SNPova koji bi bili korišteni u istraživanjima onih svojstava koja su važna za proizvođače. U istraživanje su bile uključene četiri konvencionalne pasmine svinja te divlja svinja. Korištene su kombinacije tehnologija Illumina i Roche 454, što je rezultiralo većim brojem GA sekvenci te dužim čitanjima.

ZAKLJUČAK

Vrhunac razvoja i primjene tehnologije sekvenciranja nove generacije bio je 2011. godine, kada se obilježavala 10. godišnjica sekvenciranja ljudskoga genoma. Tehnologija sekvenciranja nove generacije ima široki raspon primjena od genetskih analiza drevnih genoma, pružanja objašnjenja uloge RNA u procesima pojave različitih bolesti te njezine primjene u istraživanju genoma domaćih životinja. Metode sekvenciranja obuhvaćaju niz metoda koje su grupirane s obzirom na pripremu uzorka, tehnologiju sekvenciranja i način obrade podataka. Tijekom zadnjih godina došlo je do značajnoga pomaka od korištenja automatizirane Sangerove metode sekvenciranja prema novim tehnologijama koje se koriste u analizi genoma. U prvome redu to je bio razvoj tehnologija druge generacije sekvenciranja, a u zadnjih nekoliko godina i razvoj treće generacije.

LITERATURA

1. Archibald, A.L., Bolund, L., Churcher, C., Fredholm, M., Groenen, M.A.M., Harlizius, B., Lee, K.T., Milan, D., Rogers, J., Rothschild, M.F., Uenishi, H., Wang, J., Schook, L.B. (2010): Pig genome sequence - analysis and publication strategy. *BMC Genomics*, 11: 438. doi: <http://dx.doi.org/10.1186/1471-2164-11-438>
2. Bai, Y., Sartor, M., Cavalcoli, J. (2012): Current status and future perspectives for sequencing livestock genomes. *Journal of Animal Science and Biotechnology*, 3(8): 1-6. doi: <http://dx.doi.org/10.1186/2049-1891-3-8>
3. Brajušković, G. (2006.): *Genomika. Vojnosanitetski preglod*, 63(6): 604-610.
4. Branton, D., Deamer, D.W., Marziali, A. (2008): The potential and challenges of nanopore sequencing. *Nature Biotechnology*, 26(10): 1146-1153. doi: <http://dx.doi.org/10.1038/nbt.1495>
5. Budimir, K., Kralik, G., Margeta, V. (2013.): Epigenetske modifikacije genoma svinje. *Poljoprivreda/Agriculture*, 19(1): 76-80.
6. Budimir, K., Margeta, V., Kralik, G., Margeta, P. (2014.): Utjecaj polimorfizma FABP3 i LEPR gena na sadržaj intramuskularne masti u mišićnom tkivu svinja. *Poljoprivreda/Agriculture*, 20(1): 48-53.
7. Clarke, J., Wu, H.C., Jayasinghe, L., Patel, A., Reid, S., Bayley, H. (2009): Continuous base identification for single-molecule nanopore DNA sequencing. *Nature Nanotechnology*, 4(4): 265-270. doi: <http://dx.doi.org/10.1038/nnano.2009.12>
8. Corominas, J., Ramayo-Caldas, Y., Puig-Oliveras, A., Estellé, J., Castelló, A., Alves, E., Pena, R.N., Ballester, M., Folch, J.M. (2013): Analysis of porcine adipose tissue transcriptome reveals differences in de novo fatty acid synthesis in pigs with divergent muscle fatty acid composition. *BMC Genomics*, 14(1): 843. doi: <http://dx.doi.org/10.1186/1471-2164-14-843>
9. Esteve-Codina, A., Kofler, R., Palmieri, N., Bussotti, G., Notredame, C., Perez-Enciso, M. (2011): Exploring the gonad transcriptome of two extreme male pigs with RNA-seq. *BMC Genomics*, 12(1): 552. doi: <http://dx.doi.org/10.1186/1471-2164-12-552>
10. Fehniger, T.A., Wylie, T., Germino, E. (2010): Next-generation sequencing identifies the natural killer cell microRNA transcriptome. *Genome Research*, 20(11): 1590-1604. doi: <http://dx.doi.org/10.1101/gr.107995.110>
11. Foehlich, T. (2010): High-throughput nucleic acid analysis. U.S. Patent.
12. Harismendy, O., Strausberg, R.L., Wang, X., Stockwell, T.B., Beeson, K.Y., Schork, N.J., Murray, S.S., Topol, E.J., Levy, S., Frazer, K.A. (2009): Evaluation of next generation sequencing platforms for population targeted sequencing studies. *Genome Biology*, 10(3): 1-13. doi: <http://dx.doi.org/10.1186/gb-2009-10-3-r32>
13. Hillier, L.W., Marth, G.T., Quinlan, A.R., Dooling, D., Fewell, G., Barnett, D., Fox, P., Glasscock, J.I., Hickenbotham, M., Huang, W., Magrini, V.J., Richt, R.J., Sander, S.N., Stewart, D.A., Stromberg, M., Tsung, E.F., Wylie, T., Schedl, T., Wilson, R.K., Mardis, E.R. (2008): Whole-genome sequencing and variant discovery in *C. elegans*. *Nature Methods*, 5(2): 183-188. doi: <http://dx.doi.org/10.1038/nmeth.1179>
14. Huse, S.M., Huber, J.A., Morrison, H.G., Sogin, M.L., Welch, D.M. (2007): Accuracy and quality of massively parallel DNA pyrosequencing. *Genome Biology*, 8(7): 1-9. doi: <http://dx.doi.org/10.1186/gb-2007-8-7-r143>
15. Hutchison, III C.A. (2007): DNA sequencing: bench to bedside and beyond. *Nucleic Acids Research*, 35(18): 6227-6237. doi: <http://dx.doi.org/10.1093/nar/gkm688>
16. Jiang, J., Wang, J., Wang, H., Zhang, Y., Kang, H., Feng, X., Wang, J., Yin, Z., Bao, W., Zhang, Q., Liu, J.F. (2014): Global copy number analyses by next generation sequencing provide insight into pig genome variation. *BMC Genomics*, 15(1): 593. doi: <http://dx.doi.org/10.1186/1471-2164-15-593>

17. Johnson, D.S., Mortazavi, A., Myers, R.M., Wold, B. (2007): Genome-wide mapping of in vivo protein-DNA interactions. *Science*, 316(5830): 1497-1502. doi: <http://dx.doi.org/10.1126/science.1141319>
18. Jung, W.Y., Kwon, S.G., Son, M., Cho, E.S., Lee, Y., Kim, J.H., Kim, B.W., Park, D.H., Hwang, J.H., Kim, T.W. (2012): RNA-Seq approach for genetic improvement of meat quality in pig and evolutionary insight into the substrate specificity of animal carbonyl reductases. *PLoS ONE*, 7(9): 42198. doi: <http://dx.doi.org/10.1371/journal.pone.0042198>
19. Koboldt, D.C., Meltz Steinberg, K., Larson, D.E., Wilson, R.K., Mardis, E.R. (2013): The Next-Generation Sequencing Revolution and Its Impact on Genomics. *Cell*, 155(26): 27-38. doi: <http://dx.doi.org/10.1016/j.cell.2013.09.006>
20. Lam, H.Y.K., Clark, M.J., Chen, R., Chen, R., Natsoulis, G., O'Huallachain, M., Dewey, F.E., Habegger, L., Ashley, E.A., Gerstein, M.B., Butte, A.J., Ji, H.P., Snyder, M. (2012): Performance comparison of whole-genome sequencing platforms. *Nature Biotechnology*, 30(1): 78-82. doi: <http://dx.doi.org/10.1038/nbt.2065>
21. Ledergerber, C., Dessimoz, C. (2011): Base-calling for next-generation sequencing platforms. *Briefings in Bioinformatics*, 1-9. doi: <http://dx.doi.org/10.1093/bib/bbq077>
22. Li, H., Ruan J., Durbin, R. (2008): Mapping short DNA sequencing reads and calling variants using mapping quality scores. *Genome Res.*, 18: 1851-1858. doi: <http://dx.doi.org/10.1101/gr.078212.108>
23. Liu, L., Li, Y., Li, S., Hu, N., He, Y., Pong, R., Lin, D., Lu, L., Law, M. (2012): Comparison of Next-Generation Sequencing Systems. *Journal of Biomedicine and Biotechnology*, 1-11. doi: <http://dx.doi.org/10.1155/2012/251364>
24. Mardis, E.R. (2008): The impact of next-generation sequencing technology on genetics. *Trends in Genetics*, 24(3): 133-141. doi: <http://dx.doi.org/10.1016/j.tig.2007.12.007>
25. Marioni, J.C., Mason, C.E., Mane, S.M., Stephens, M., Gilad, Y. (2008): RNA-seq: an assessment of technical reproducibility and comparison with gene expression arrays. *Genome Research*, 18(9): 1509-1517. doi: <http://dx.doi.org/10.1101/gr.079558.108>
26. Maxam, A.M., Gilbert, W. (1977): A new method for sequencing DNA proceedings of the National Academy of Sciences, 74(2): 560-564.
27. Metzker, M.L. (2005): Emerging technologies in DNA sequencing. *Genome Res.*, 15: 1767-1776. doi: <http://dx.doi.org/10.1101/gr.3770505>
28. Metzker, M.L. (2009): Sequencing technologies - the next generation. *Genetics*, 11: 31-46. doi: <http://dx.doi.org/10.1038/nrg2626>
29. Pareek, C.S., Smoczynski, R. (2011): Sequencing technologies and genome sequencing. *J Appl Genetics*, 52: 413-435. doi: <http://dx.doi.org/10.1007/s13353-011-0057-x>
30. Ramayo-Caldas, Y., Mach, N., Esteve-Codina, A., Corominas, J., Castello, A., Ballester, M., Estelle, J., Ibanez-Escriche, N., Fernandez, A., Perez-Enciso, M. (2012): Liver transcriptome profile in pigs with extreme phenotypes of intramuscular fatty acid composition. *BMC Genomics*, 13(1): 547. doi: <http://dx.doi.org/10.1186/1471-2164-13-547>
31. Ramos, A.M., Crooijmans, R., Affara, N.A., Amaral, A.J., Archibald, A.L., Beever, J.E., Bendixen, C., Churcher, C., Clark, R., Dehais, P., Hansen, M.S., Hedegaard, J., Hu, Z.L., Kerstens, H.H., Law, A.S., Megens, H.J., Milan, D., Nonneman, D.J., Rohrer, G.A., Rothschild, M.F., Smith, T., Schnabel, R.D., Van Tassell, C.P., Taylor, J.F., Wiedmann, R.T., Schook, L.B., Groenen, M.A.M. (2009): Design of a High Density SNP Genotyping Assay in the Pig Using SNPs Identified and Characterized by Next Generation Sequencing Technology. *Plos ONE*, 4(8): e6524. doi: <http://dx.doi.org/10.1371/journal.pone.0006524>
32. Reis Filho, J.S. (2009): Next-generation sequencing. *Breast Cancer Research*, 11(3): 1-12. doi: <http://dx.doi.org/10.1186/bcr2431>
33. Robertson, G., Hirst, M., Bainbridge, M., Bilenky, M., Zhao, Y., Zeng, T., Euskirchen, G., Bernier, B., Varhol, R., Delaney, A., Thiessen, N., Griffith, O.L., He, A., Marra, M., Snyder, M., Jones, S. (2007): Genome-wide profiles of STAT1 DNA association using chromatin immunoprecipitation and massively parallel sequencing. *Nature Methods*, 4: 651-657. doi: <http://dx.doi.org/10.1038/nmeth1068>
34. Sanger, F., Coulson, A., Barrell, B.G., Smith, A.J.H., Roe, B.A. (1980): Cloning in single-stranded bacteriophage as an aid to rapid DNA sequencing. *Journal of Molecular Biology*, 143(2): 161-178. doi: [http://dx.doi.org/10.1016/0022-2836\(80\)90196-5](http://dx.doi.org/10.1016/0022-2836(80)90196-5)
35. Schook, L.B., Beever, J.E., Rogers, J., Humphray, S., Archibald, A., Chardon, P., Milan, D., Rohrer, G., Eversole, K. (2005): Swine Genome Sequencing Consortium (SGSC): a strategic roadmap for sequencing the pig genome. *Comp Funct Genom.*, 6: 251-255. doi: <http://dx.doi.org/10.1002/cfg.479>
36. Timp, W., Mirsaidov, U.M, Wang, D., Comer, J., Aksimentiev, A., Timp, G. (2010): Nanopore sequencing: electrical measurements of the code of life. *IEEE Transactions on Nanotechnology*, 9(3): 281-294. doi: <http://dx.doi.org/10.1109/TNANO.2010.2044418>
37. Wang, L., Weinshilboum, R.M. (2008): Pharmacogenomics: candidate gene identification, functional validation and mechanisms. *Human Molecular Genetics*, 17(2): 1-6. doi: <http://dx.doi.org/10.1093/hmg/ddn270>
38. Whiteford, N., Skelly, T., Curtis, C., Ritchie, M.E., Lohr, A., Zaranek, A.W., Abnizova, I., Brown, C. (2009): Primary data analysis for the Illumina Solexa sequencing platform. *Bioinformatics*, 25: 2194-2199. doi: <http://dx.doi.org/10.1093/bioinformatics/btp383>
39. Wiedmann, R.T., Smith, T.P.L., Nonneman, D.J. (2008): SNP discovery in swine by reduced representation and high throughput pyrosequencing. *BMC Genetics*, 9(1): 81. doi: <http://dx.doi.org/10.1186/1471-2156-9-81>

DEVELOPMENT OF NEW SEQUENCING TECHNOLOGIES AND THEIR APPLICATION IN GENOME ANALYSIS OF DOMESTIC ANIMALS

SUMMARY

Sequencing and detailed study of the genom of domestic animals began in the middle of the last century. It was primarily referred to development of the first generation sequencing methods, i.e. Sanger sequencing method. Next generation sequencing methods are currently the most common methods in the analysis of domestic animals genom. The application of these methods gave us up to 100 time more data in comparison with Sanger method. Analyses including RNA sequencing, genotyping of whole genome, immunoprecipitation associated with DNA microarrays, detection of mutations and inherited diseases, sequencing of themitochondrial genome and many others have been conducted with development and application of new sequencing methods since 2005 until today. Application of new sequencing methods in the analysis of domestic animal genome provides better understanding of the genetic basis for important production traits which could help in improving the livestock production.

Key-words: *sequencing, Sanger sequencing method, next generation sequencing, domestic animal genome analysis*

(Primljeno 22. prosinca 2014.; prihvaćeno 21. listopada 2015. - Received on 22 December 2014; accepted on 21 October 2015)