

Sveučilište u Zagrebu
Agronomski fakultet

Iva Katić, Martina Jašinski, Tea Glunčić, Zdravko Begić

ANALIZA VISOKO REZOLUTNE GENETSKE INFORMACIJE U FILOGENEZI
MITOGENOMA GOVEDA

Zagreb, 2017.

Ovaj rad izrađen je u Laboratoriju za bioinformatiku Agronomskog fakulteta pri Zavodu za opće stočarstvo, Svetošimunska cesta 25, Zagreb, pod stručnim vodstvom doc. dr. sc. Maja Ferenčaković i predan je na natječaj za dodjelu Rektorove nagrade u akademskoj godini 2016./2017.

SADRŽAJ RADA

1. UVOD	1
1.1. TAKSONOMIJA GOVEDA.....	2
1.1.1. Bos taurus	2
1.1.2. Bos indicus	3
1.2. MITOHONDRIJ.....	3
1.3. MITOHONDRIJSKA DNA.....	4
1.4. MITOHONDRIJSKE HAPLOGRUPE GOVEDA.....	5
1.4.1. Haplogrupa T.....	7
1.4.2. Haplogrupa Q	7
1.4.3. Haplogrupa R	8
1.4.4. Haplogrupa P.....	8
1.5. POLIMORFIZAM JEDNOG NUKLEOTIDA.....	9
1.6. GENOTIPIZIRANJE	10
1.7. KOMERCIJALNI SNP ČIPOVI.....	10
1.8. WRIGHTOVA STATISTIKA	12
1.8.1. Primjena „F“ statistike	13
1.9. GOVEĐI GENOM UMD 3.1 (ASSEMBLY).....	13
2. HIPOTEZA I SPECIFIČNI CILJEVI RADA.....	15
2.1. HIPOTEZE.....	15
2.2. SPECIFIČNI CILJEVI RADA.....	15
3. MATERIJAL I METODE.....	16
3.1. PORAVNAVANJE I IZBOR NUKLEOTIDNIH SLJEDOVA	16
3.2. IZRAČUN FST KOEFICIJENTA	16
3.3. ODREĐIVANJE MUTACIJA PRISUTNIH NA SNP ČIPU.....	17
4. REZULTATI.....	18
4.1. NUKLEOTIDNI SLJEDOVI I DETEKTIRANE HAPLOGRUPE	18
4.2. FST KOEFICIJENTI NUKLEOTIDNIH SLJEDOVA	18
4.3. FST ZNAČAJNE MUTACIJE NA KOMERCIJALNOM SNP ČIPU.....	21
5. RASPRAVA.....	22
6. ZAKLJUČCI	25
7. ZAHVALE	26
8. POPIS LITERATURE	27
9. SAŽETAK.....	34
10. SUMMARY	35

11. ŽIVOTOPISI.....	36
12. PRILOZI.....	38

1. UVOD

Govedarstvo je jedna od najvažnijih grana stočarstva i poljoprivrede. Ona je temelj ukupne stočarske proizvodnje i kao takva je od višestruke gospodarske važnosti. Obuhvaća uzgoj goveda, bikova, jakova i zebua. Postoje tri proizvodna tipa: mliječna, mesna i kombinirana goveda. Čovjek je domesticirao govedo prije 10 000 godina i on je potomak divljeg *Bos primigeniusa*. Današnje moderne pasmine goveda se razvrstavaju u tri glavne grupe: zebuine, sanga i taurine.

Mitohondrijska DNA (mtDNA) je dvolančana kružna molekula koja se nasljeđuje s majčine strane. Pomoću mtDNA moguće su razne analize koje nam mogu poboljšati uvid u filogenezu te postanak razvoja goveda i pomoću nje se mogu pratiti evolucijski procesi nastajanja. Uz pomoć lančane reakcije polimeraze dobiva se slijed nukleotida mtDNA te se dobiva informacija kako o varijabilnim tako i o konzerviranim mjestima slijeda nukleotida mtDNA.

Pozicija baze i slijed alternativne genomske DNA smatra se polimorfizmom jednog nukleotida (engl. Single nucleotide polymorphism; SNP) ako je učestalost alela 1 % ili više u promatranoj populaciji. Postoje SNP čipovi visoke, srednje i niske rezolucije. Za dobivanje rezultata koristila se Wright-ova „F“ statistika koja nam je omogućila izdvajanje mutacija koje razdvajaju haplogrupe i sub-haplogrupe mitogenoma goveda. Kod goveda postoje haplogrupe i njihove sub-haplogrupe (I, I1, I2, P1, Q1, Q2, R1, R2, T1, T2, T3, T4, T5). Od navedenih haplogrupa najčešće se pojavljuje haplogrupa T, a rijetke haplogrupe su P, R i Q.

Cilj ovog rada je utvrditi prisutnost mutacija (SNP-ova) značajnih za razdvajanje haplogrupa mitogenoma goveda prisutnih na komercijalnom SNP čipu goveda.

1.1. TAKSONOMIJA GOVEDA

Domaće govedo nalazi svog pretka u divljem auroh govedu (*Bos primigenius*), vrsti koja je do 17. stoljeća prirodno obitavala na području sjeverne Afrike, Euroazije i Indije (van Vuure, 2005.). Danas je uvriježeno mišljenje kako se domestikacija domaćeg goveda dogodila u doba neolitika u dva navrata, na području „Plodnog polumjeseca“ (jugozapadna Azija) pri čemu nastaje *Bos taurus*, te na području južne Azije gdje nastaje *Bos indicus* (Loftus i sur., 1994.), a ostale na području Afrike prije više od 1000 godina (Chan i sur., 2010.).

Slika 1. prikazuje današnje moderne pasmine goveda koje se razvrstavaju u tri glavne grupe: zebuini (1A), sanga (1B) i taurini (1C), pri čemu sanga predstavlja križance između taurina i zebuina.



Slika 1. Tri glavne grupe modernih goveda: A) zebuini; B) sanga; C) taurini

Preuzeto s: <http://en.wikipedia.org/wiki/Cattle>

1.1.1. *Bos taurus*

Bos taurus ili europsko govedo je podvrsta pripitomljenog goveda koji je podrijetlom iz Bliskog istoka. Genetska istraživanja dokazuju kako je današnje moderno govedo nastalo od 80 jedinki auroha. Aurosi su prije 10 500 godina naseljavali gornje dijelove Mezopotamije u blizini sela Çayönü u jugoistočnoj Turskoj i Dja'de el Mughara u sjevernom dijelu Iraka (Bollongino i sur., 2012.).

1.1.2. Bos indicus

Bos indicus ili indijsko govedo koje je poznato pod imenom zebu ili grbavo govedo. Zebu govedo je podvrsta domaćeg goveda podrijetlom iz južne Azije. Specifična karakteristika ovog goveda je njegova masna grba na hrptu, spuštene uši i veliki obješeni podbradak. Evolucijski je govedo dobro adaptirano na topliju klimu te se upravo ta osobina iskorištava u stvaranju novih pasmina tropskog i subtropskog pojasa. Danas se ponajviše koriste kao radna snaga za vuču u poljoprivredi te za proizvodnju mlijeka i mesa. Danas je poznato 75 zebu pasmina ponajviše na području Afrike, južne Azije i Južne Amerike. Najpoznatije pasmine zebu goveda su: Gir, Kankrej, Braman, Nelore, Ongole, Baggara, Kangavam i mnoga druga lokalna goveda (<https://sh.wikipedia.org/wiki/Zebu>).

1.2. MITOHONDRIJ

Mitohondriji su stanične organele u eukariotskim stanicama. Razlikuju se od ostalih staničnih organela zbog posjedovanja vlastite DNA koja kodira tRNA, rRNA i neke od proteina. Ključna uloga mitohondrija je proizvodnja energije i time su odgovorni za najveći dio upotrebljive energije koja je oslobođena razgradnjom ugljikohidrata i masnih kiselina (Cooper i sur., 2010.).

Mitohondriji su kuglasti ili valjkasti stanični organeli (Slika 2.) obavijeni dvostrukom membranom. Vanjska membrana je glatka i manje površine, dok je unutarinja membrana veća i naborana. Nabori unutarnje membrane nazivaju se kriste i dijele mitohondrij na dva dijela: matriks i međumembranski prostor. U matriksu se nalaze DNA, RNA i ribosomi (Ančić i sur., 2008). Matriks ispunjava unutrašnjost mitohondrija i sadrži nekoliko enzima koji konvertiraju produkte metabolizma, lipida i proteina kroz Krebsov ciklus do ugljičnog dioksida i vode, uz

oslobađanje energije u obliku adenozin trifosfata (ATP). Tipična eukariotska stanica sadrži oko 2 000 mitohondrija (Cooper i sur., 2010.).



Slika 2. Mitohondrij

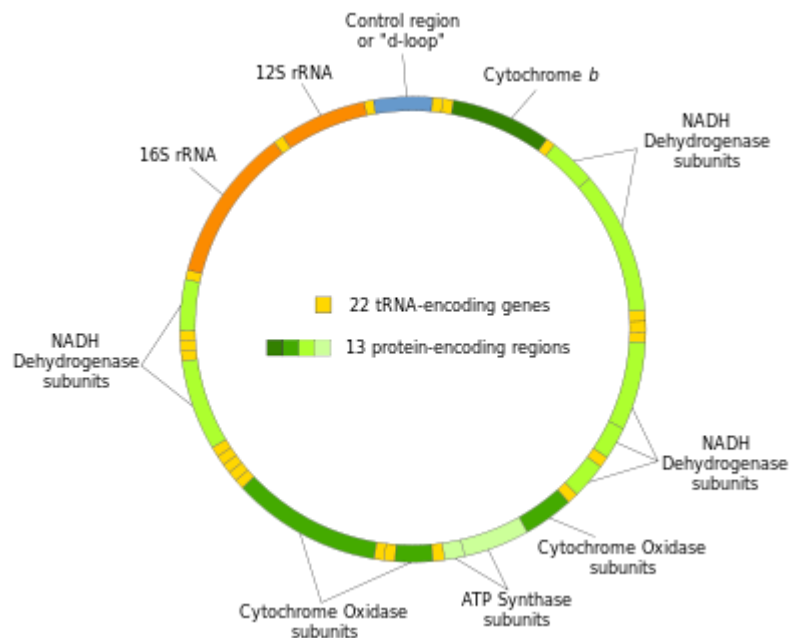
Preuzeto s <http://www.instrukcije-poduka.com/instrukcije-iz-biologije-mitohondriji.html>

1.3. MITOHONDRIJSKA DNA

Mitohondrijska DNA (mtDNA) je gola kružna dvolančana DNA (Slika 3.) koja se replicira neovisno o jezgrinom genomu. Kod čovjeka mtDNA sadrži 16 569 parova baza (engl. „base pairs“, bp) i kodira 37 gena, od čega 13 za proteine mitohondrija koji sudjeluju u sastavu enzima oksidativne fosforilacije, transporta elektrona i ATP sintaze, 22 tRNA i dvije rRNA (Cooper i sur., 2010.). Sastoji se od kodirajućeg i ne kodirajućeg dijela. U kodirajućem dijelu su sve kodne sekvence u dodiru jedna s drugom bez introna, a ne kodirajući lanac je zamjenska petlja (D-loop) (Holt i sur., 2000., Yang i sur., 2002., Zeviani i sur., 2004.).

Mitohondrijska DNA sisavaca nasljeđuje se jednostrano i bez rekombinacija po majčinoj strani te relativno brzo mutira. MtDNA ima mnogo viši nivo varijacija nego DNA jezgre. Brzina mutacije za mtDNA iznosi 2 % na milijun godina tj. 10^{-8} nukleotidnih mjesta po godini što znači da evoluirá 10 puta brže nego nuklearna DNA (Brown i sur., 1982.). Upravo zbog ovih

osobina mtDNA se smatra korisnim genetskim markerom za proučavanje porijekla, genetske raznolikosti i diferencijacije visoko srodnih vrsta i jedinki unutar vrsta.



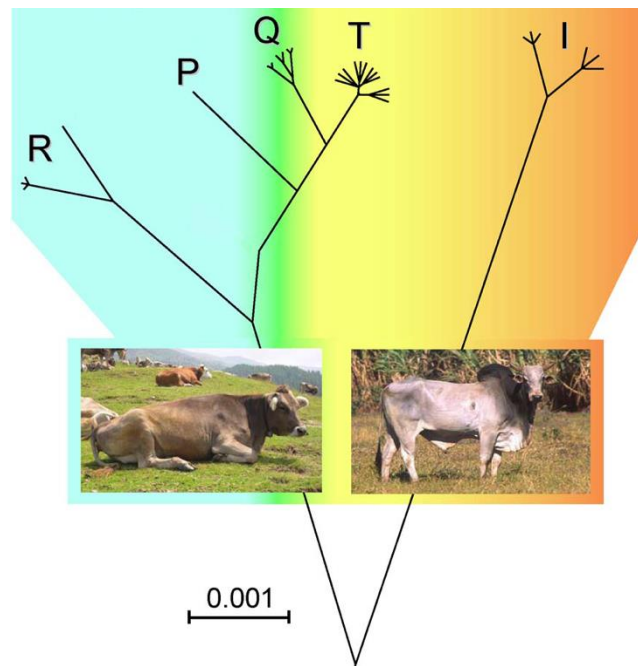
Slika 3. Mitohondrijska DNA

Preuzeto s <https://www.google.hr/search?q=mitohondrij&client=firefox>

1.4. MITOHONDRIJSKE HAPLOGRUPE GOVEDA

Vjeruje se kako je današnje moderno govedo nastalo na području Bliskog istoka. Novija istraživanja dokazuju kako je nastalo spajanjem dviju neovisnih i geografski drugačijih domestikacija. Obje su se dogodile u jugozapadnoj Aziji, jedna na području Plodnog polumjeseca i uključivala je auroha čija je mtDNA pripadala haplogrupi T (*Bos taurus*) sa sub-haplogrupama T1, T2, T3, T4 i T5, a druga na području nizine Indus gdje su se grbave zebu pasmine miješale s *Bos primigenius* pasminama, te pripadaju haplogrupama I1 i I2, *Bos indicus* (Achilli i sur., 2008., Achilli i sur., 2009.). Haplogrupa T je makro-haplogrupa jer nju nalazimo u najvećem broju te je široko rasprostranjena. No, osim nje, pronađene su još tri haplogrupe: O, P i R koje su dosta rijetke (Slika 4). Smatra se kako je domestikacija haplogrupe Q počela

na Bliskom istoku, a za haplogrupe P i R se smatra kako potječu od europskih populacija divljeg goveda aurohsena. Postoje istraživanja koja dokazuju o mogućnosti domestikacije *Bos primigeniusa* u Italiji (Bonfiglio i sur., 2012.).



Slika 4. Filogeneza mtDNA goveda
Preuzeto iz (Achilli i sur., 2009.).

1.4.1. Haplogrupa T

Makro-haplogrupa T se sastoji od dviju sestrinskih sub-haplogrupa T1'2'3 i T5. T1'2'3 je formirana iz prijašnjih definiranih haplogrupa T1, T2 i T3. Sub-haplogrupe T1 i T3 su definirane sa samo jednom poznatom mutacijom u kontrolnoj regiji. Sub-haplogrupa T2 je bolje okarakterizirana (Achilli i sur., 2008.). T haplogrupa je najviše rasprostranjena.

Goveda haplogrupe T1 su od posebnog interesa radi toga što posjeduju osobite filogeografske značajke. Goveda ove haplogrupe su većinom prisutna na srednjem istoku i kod anatolijskih pasmina, ali je dosta česta i u pasminama južne Europe (Portugal, Španjolska, Grčka i Italija), te je fiksirana u afričkom govedu. T1 je dosta istraživana haplogrupa radi mogućih ispitivanja neovisne domestikacije u Africi i njezin utjecaj u formiranju iberijskih i kreolskih pasmina (Bonfiglio i sur., 2012.). T2 i T3 u znatnim frekvencijama nalazimo na Bliskom istoku, ali T3 haplogrupa je pretežito vezana za područje Europe. Haplogrupa T4 se najvećim dijelom nalazi u istočnoj Aziji i to ponajviše u Japanu (Mannen i sur., 2004.). Postoje još haplogrupe T5 i T6. Sve navedene sub-haplogrupe se dalje granaju i na sub-subhaplogrupe.

1.4.2. Haplogrupa Q

Pretpostavka je kako se goveda haplogrupe Q mogu naći u nekoliko europskih država, te Egiptu, Turskoj i Kini. Pronađena je i u talijanskim pasminama. Haplogrupa Q se dijeli u još dvije sub-haplogrupe Q1 i Q2. Haplogrupe Q i T su prisutne u istim auroh populacijama ili u istom geografskom području (Bonfiglio i sur., 2010.).

1.4.3. Haplogrupa R

Haplogrupa R je jedna od rijetkih haplogrupa i pojavljuje se samo u talijanskim pasminama. Nije pronađena nigdje drugdje u svijetu. Nema nikakve zajedničke karakteristike s haplogrupama T i Q. Iako je haplogrupa Q pronađena u Italiji, vjeruje se kako je tijekom domestikacije tekao kod svake haplogrupe na različiti način. Postavlja se pitanje, zašto je haplogrupa R pronađena samo u Italiji? Jedan od odgovora su klimatske promjene koje su utjecale na Europu pri kraju pleistocena. Italija je bila jedna od migrantskih zemalja Europe tijekom zadnjeg ledenog doba. Kasnije, njezino rasprostranjivanje na sjever Europe je bilo onemogućeno zbog Alpi. Haplogrupa R se sastoji od dviju linija sub-haplogrupe R1 i R2 (Bonfiglio i sur., 2010.).

1.4.4. Haplogrupa P

Analize mtDNA i DNA iz starih kostiju daju nam uvid u demografsku strukturu izumrlih životinja. Ostatci auroha iz tercijarnog doba dokazuju njegov ishodišni put iz Indije prije 1.5 do 2 milijuna godina (Mona i sur., 2010.). Smatra se kako je *Bos primigenius* u pleistocenu naseljavao područje Europe, sjeverne Afrike i male Azije. Najstariji ostatci dokazuju njegovo postojanje prije 275 000 godina. Kako je već navedeno, postoje tri sub-populacije auroha: indijski - *Bos primigenius nomadicus* koji je predak današnjih zebu goveda - *Bos indicus*, sjevernoafrički auroh - *Bos primigenius africanus* i euroazijski auroh - *Bos primigenius primigenius* (Zeyland i sur., 2013.).

Većina analiziranih ostataka auroha upućuje kako goveda iz sjeverne i centralne Europe, te mala skupina talijanskih auroha pripadaju haplogrupi P koja uključuje jedinke čije su kosti pronađene u Engleskoj. Ovoj haplogrupi pripada i korejsko moderno govedo. Početak

širenja haplogrupe P datira od prije 16 000 do 36 000 godina (Achilli i sur., 2008., Mona i sur., 2010.).

1.5. POLIMORFIZAM JEDNOG NUKLEOTIDA

Polimorfizam jednog nukleotida (engl. single nucleotide polymorphism; SNP) predstavlja promjenu samo jedne baze u DNA sekvenci s uobičajenom alternativom od dva moguća nukleotida na toj određenoj poziciji. Pozicija baze i slijed alternativne genomske DNA smatra se SNP-om ako je učestalost alela 1 % ili više u promatranoj populaciji. Iako, bilo koje od četiri moguće nukleotidne baze mogu biti prisutne. SNP-ovi su obično bialelni. Jedan od razloga za to je niska frekvencija jedne nukleotidne supstance na podrijetlo SNP-a, a procjenjuje se kako je između 1×10^{-9} i 5×10^{-9} po nukleotidu (Li i sur., 1981. i Martinez-Arias i sur., 2001.).

Dakle, vjerojatnost da će se dvije neovisne baze promijeniti na jednom mjestu je jako niska. Drugi razlog tome je pristranost kod mutacija, što je dovelo do pojavljivanja dva SNP tipa. Mehanizmi mutacije mogu rezultirati u tranziciji: purin - purin (A, G) ili pirimidin - pirimidin (C, T) ili transverziji purin – pirimidin ili pirimidin – purin. Dvostruko je više prijelaza transverzije nego translacije (Vignal i sur., 2002.).

Iako su opisani brojni pristupi otkrivanja SNP-ova, uključujući i neke koji se trenutno koriste za genotipiziranje, glavne su one koje se temelje na usporedbi lokus – specifična sekvenca, generirana različitim kromosomima. Najjednostavnije je, kod targetiranja definirane regije, napraviti određivanje slijeda nukleotida tj. sekvenciranje produkata PCR-a dobivenog kod različitih individua. Međutim, ovaj pristup zahtjeva visoki financijski izdatak zbog potrebe za lokus specifičnim početnicama. Ograničen je na regiju za koju su dostupni podatci sekvence. Nastaje diploidan slijed u kojem nije uvijek lako razlikovati sekvencirane baze i polimorfizme zbog pojave dvostrukih pikova, kao što se očekuje kod heterozigota. Različiti pristupi se temelje

na usporedbama sekvenci koje su dobivene iz kloniranih fragmenata i to se može smatrati važnim za razvoj SNP karata genoma (Vignal i sur., 2002.).

1.6. GENOTIPIZIRANJE

Postoje mnoge tehnike genotipiziranja SNP-ova. Jedna od glavnih značajki većine tehnika, osim onih koje se baziraju na izravnoj hibridizaciji, je odvajanje u dva koraka: 1.) stvaranje produkata alel-specifične molekulske reakcije i 2.) razdvajanje i detekcija alel specifičnih produkata za njihovu identifikaciju. Neke od tehnika su rezanje restriksijskih enzima, proširenje početnica, testiranje oligonukleotidnog vezanja (OLA) i pirosekvenciranje. Mnoge od tih metoda danas se nude kao komercijalni kitovi (Vignal i sur., 2002.).

1.7. KOMERCIJALNI SNP ČIPOVI

U prvoj polovici 21. stoljeća dogodio se veliki napredak u stvaranju genomsko-seleksijskih alata. „High Density“ (HD) SNP čipovi povećavaju točnost procjena uzgojnih vrijednosti, predviđaju fenotipska obilježja već u ranoj dobi jedinke, skraćuju generacijske intervale i poboljšavaju kontrolu uzgoja u srodstvu. Predviđanje cijeloga genoma (engl. „Whole Genome Prediction“; WGP) uz pomoć komercijalnih SNP čipova, srednje do visoko gustih rezolucija, transformirale su uzgoj domaćih životinja i biljaka (Abaci i sur., 2016.).

Generalno se razlikuju tri vrste SNP čipova. SNP čipovi niske rezolucije (low density), srednje rezolucije (medium density) SNP-ova i visoke rezolucije (high density). U slučaju goveda najviše se koristi Illumina BovineSNP50 Genotyping BeadChip (Illumina Inc., San Diego, CA) koji sadrži 54 001 SNP-ova i Illumina BovineHD Genotyping BeadChip (Illumina Inc., San Diego, CA) koji sadrži 777 972 SNP-ova (Ferenčaković i sur., 2013.).

Trenutno se Illumina BovineSNP50 Genotyping BeadChip (Slika 5.) koristi za otkrivanje genoma kod mliječnih krava. Illumina BovineHD Genotyping BeadChip je stvoren 2010. godine sa ciljem preciznijih, gušćih i ravnomjernije raspoređenih SNP-ova s genoma, a uključuje i SNP-ove s Y kromosoma i mitohondrijske SNP-ove (Su i sur., 2011.).

Na tržištu se mogu naći i čipovi niske rezolucije (low density) za genotipiziranje goveda, te se najčešće koristi 3K BovineChip koji sadrži 2 900 SNP-ova koji su jednako raspoređeni te sadrži 82 SNP-a za određivanje pasmina. Uz pomoć 3K BovineChip-a može se genotipizirati više jedinki uz manji trošak, te su korisni za otkrivanje podrijetla proizvoda specifičnih za pasmine. Drugi čipovi manje rezolucije koji se još koriste su: 95 SNP PercentageChip za identifikaciju i korekciju pedigre genoma, te 384 SNP koji se koristi za početno pregledavanje krava (https://aipl.arsusda.gov/publish/WC9/WC9_10_cvt_jbc.ppt).

Jedan od većih nedostataka SNP čipova je njihova cijena. Stoga je Illumina stvorila i čipove manje rezolucije i to Bovine LD Genotyping BeadChip 7K koji ima manju cijenu i dobra je alternativa 50K i precizniji je od 3K čipa (Wang i sur., 2016.).

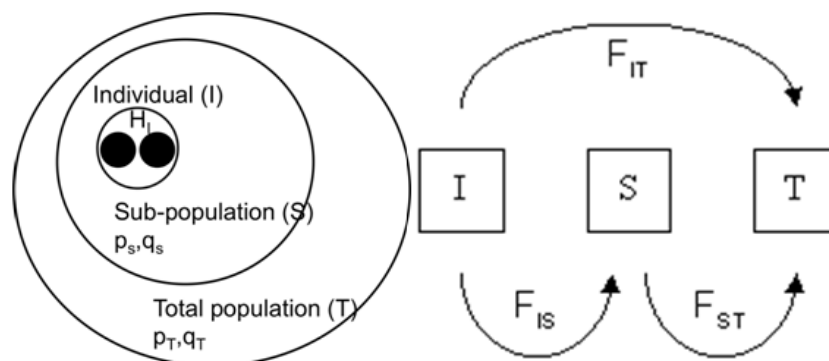


Slika 5. Illumina BovineSNP50 Genotyping BeadChip i Illumina BovineHD Genotyping BeadChip

Preuzeto s <https://www.illumina.com/>

1.8. WRIGHTOVA STATISTIKA

Sewall Wright je 1922. godine definirao inbriding koeficijente za opisivanje distribucije genetičke raznolikosti unutar i između fragmentiranih populacija. Wright je definirao tri različita koeficijenta: FIS, FST i FIT (Slika 6.) koji predstavljaju tzv. „F“ statistiku. FIS koeficijent je prosječno odstupanje od Hardy Weinberg ekvilibrijuma (HWE) unutar populacije ili korelacija između alela unutar jedinki u odnosu na populaciju kojoj pripadaju. FST koeficijent je proporcija genetske raznolikosti uzrokovane razlikama u frekvencijama alela između populacija ili korelacija između alela unutar populacije u odnosu na cijelu populaciju. Statistika je povezana s varijancom frekvencije alela između populacija i stupnja srodnosti između jedinki unutar populacija. Male vrijednosti FST-a označuju sličnost frekvencije alela unutar svake populacije, dok veće vrijednosti označuju različitost frekvencije alela (Holsinger i Weir, 2009.). FIT koeficijent kvantificira odstupanje frekvencija genotipova u kombiniranom uzorku od HWE ili korelacija između alela unutar jedinki u odnosu na kombinirani uzorak (Wright, 1922.).



Slika 6. Odnos Wright-ove statistike
Preuzeto s <https://en.wikipedia.org/wiki/F-statistics>

1.8.1. Primjena „F“ statistike

Wright-ovi koeficijenti inbridinga nalaze svoju primjenu kod:

1. procjene stope migracije
2. zaključivanja o demografskoj povijesti – populacijske ili parne procjene FST-a mogu pružiti uvid u demografsku povijest kada su procjene dostupne s puno lokusa
3. identifikacije genomskih regija pod selekcijom, lokus – specifične procjene
4. forenzike i asocijacijskog mapiranja - korekcija za vjerojatnost podudaranja bazirana na frekvenciji alela iz veće populacije u koju pripada sub-populacija alela (Holsinger i Weir, 2009.).

Wright-ova statistika, a posebno FST, daje nam uvid u evolucijske procese koji utječu na strukturu genetske varijabilnosti unutar i između populacija te se najviše koriste kao deskriptivna statistika u populacijskoj i evolucijskoj genetici. FST identificira regije u genomu koje su bile pod selekcijom.

1.9. GOVEĐI GENOM UMD 3.1 (ASSEMBLY)

Svaka eukariotska stanica ima različiti broj kromosoma. Goveđa stanica ih posjeduje 60. Isto tako, svaki genom ima svoju dužinu te goveđi genom sadrži oko 3 000 000 000 parova baza. Referentni genom u ovom radu je UMD 3.1., Bos taurus pasmine hereford koji je uveden u NCBI pod lokusom AY526085. Dužina referentnog mitohondrijskog genoma iznosi 16 338 parova baza. UMD Bos taurus 3.1 assembly sastavljen je u Centru za bioinformatiku i računalnu biologiju (CBCB) na sveučilištu „Maryland“ u Americi. U javnosti je dostupan od prosinca 2009. godine. Za sekvenciniranje cijelog genoma koristila se kombinacija BAC-by-BAC koja

čini pretežno do 11 000 000 ponavljanja i metoda Shotgun za sekvenciranje cijelog genoma koja sadrži približno 24 000 000 ponavljanja. Dužina UMD-a 3.1 iznosi 2.65 Gb (http://www.ensembl.org/Bos_taurus/Info/Annotation).

2. HIPOTEZA I SPECIFIČNI CILJEVI RADA

Cilj ovog rada je analizirati mtDNA SNP-ove dostupne na Illumina BovineHD Genotyping BeadChip (Illumina Inc., San Diego, CA) i utvrditi njihovu pogodnost za određivanje glavnih haplogrupa mitogenoma goveda. Također, želimo utvrditi važnost polimorfizama mitogenoma za razdvajanje haplogrupa na temelju Wright-ovog FST koeficijenta. Time, osim što utvrđujemo SNP-ove koji razdvajaju haplogrupe, isto tako i predlažemo moguća polimorfna mjesta koja nisu prisutna na SNP čipu, a bila bi od koristi za filogenetska istraživanja.

2.1. HIPOTEZE

1. Mutacije specifične za razdvajanje haplogrupa mitogenoma goveda prisutne su na komercijalnom Illumina BovineHD Genotyping BeadChip (Illumina Inc., San Diego, CA) SNP čipu i mogu se koristiti u filogenezi mitogenoma goveda.

2.2. SPECIFIČNI CILJEVI RADA

1. Identificirati mutacije mitogenoma goveda poravnavanjem i analizom sljedova nukleotida cijele mtDNA iz banke gena.

2. Izračunati Wright-ov FST koeficijent za sva polimorfna mjesta na sljedovima nukleotida cijele mtDNA predstavnika haplogrupa mitogenoma goveda i utvrditi njihovu značajnost za razdvajanje spomenutih haplogrupa.

3. Utvrditi prisutnost mutacija (SNP-ova) značajnih za razdvajanje haplogrupa mitogenoma goveda prisutnih na komercijalnom Illumina BovineHD Genotyping BeadChip (Illumina Inc., San Diego, CA) SNP čipu

3. MATERIJAL I METODE

3.1. PORAVNAVANJE I IZBOR NUKLEOTIDNIH SLJEDOVA

Ukupno 252 nukleotidna slijeda cijele mtDNA (Prilog Tablica S1) preuzeto je iz banke gena (<https://www.ncbi.nlm.nih.gov>). Preuzeti nukleotidni sljedovi su potom poravnati i editirani uz pomoć programskog paketa MEGA 7 (Kumar i sur., 2016.) i online putem paketa “Clustal Omega” (<http://www.ebi.ac.uk/Tools/msa/clustalo/>). Za poravnavanje korišteni su referentni nukleotidni sljedovi V00654 (referentni slijed mtDNA) i NC_006853.1 (referentni slijed mtDNA u UMD3.1).

Kako su po svojoj prirodi SNP-ovi na SNP čipu bialelni, tako su iz skupa nukleotidnih sljedova uklonjene sve insercije, delecije, dvoznačne pozicije, te pozicije na kojima je prisutno više od dva alela. Jednako tako uklonjena je i ponavljajuća sekvenca koja je uobičajeno prisutna u mtDNA goveda, no ne koristi se u filogenezi.

Preuzeti nukleotidni sljedovi sastavni su dio programskog paketa MitoToolPy (Peng i sur., 2015.) specijaliziranog alata za analizu mtDNA goveda, peradi, pasa, konja, koza, svinja, ovaca i jakova. Programski paket koristi navedene nukleotidne sljedove (Tablica S1) kao referentne za određivanje mtDNA haplogrupa.

3.2. IZRAČUN FST KOEFICIJENTA

Daljnja obrada nukleotidnih sljedova imala je za cilj prilagoditi iste programskim paketima koji se uobičajeno koriste za analizu SNP čip polimorfizama. Vlastitom skriptom za programski paket SAS 9.3 (SAS Institute Inc.) nukleotidni sljedovi su pretvoreni u ulazne podatke za programski paket PLINK v1.9 (Purcell i sur. 2007.) čija je namjena analiza genomskih podataka dobivenih putem komercijalnih SNP čipova. Putem PLINK programskog

paketa paketi su pretvoreni u standardni oblik za analizu SNP-ova, odnosno u .ped i .map format. Format .ped sadrži u sebi podatke o analiziranoj jedinci te o genotipu u bialelnom formatu, dok .map format sadrži informaciju o samom SNP-u, kao što su ime SNP-a, kromosom, genomska i fizička pozicija.

Standardizirani podatci su potom ponovno uvedeni u programski paket SAS 9.3 gdje se radio izračun FST koeficijenta. FST koeficijenti izračunati su na način da su sekvence iz pojedinih haplogrupa svrstane u posebne grupe koje su pri analizi međusobno uspoređene svaka sa svakom. Za taj izračun je korištena procedura PROC ALLELE uz naredbu POP koja uspoređuje populacije (u ovom slučaju grupe) i dodatnu opciju INDIVLOCI koja omogućava izračun FST koeficijenta za svaki marker posebno između analiziranih grupa. Dobiveni FST koeficijenti odmah analizirani i pristupilo se njihovoj detekciji na SNP čipu.

3.3. ODREĐIVANJE MUTACIJA PRISUTNIH NA SNP ČIPU

Nakon što su markeri sa značajnom vrijednošću FST-a identificirani i izdvojeni pripisana im je originalna pozicija s referentnog nukleotidnog slijeda NC_006853.1. Korištenjem informacije o pozicijama SNP-ova, preuzete iz mape referentnog genoma goveda za SNP čipove UMD 3.1, utvrđivalo se jesu li mutacije sa značajnom vrijednošću FST koeficijenta prisutne u komercijalnom Illumina BovineHD Genotyping BeadChip (Illumina Inc., San Diego, CA). Pri ovom postupku koristio se programski paket SAS 9.3.

4. REZULTATI

4.1. NUKLEOTIDNI SLJEDOVI I DETEKTIRANE HAPLOGRUPE

Poravnavanje 252 nukleotidna slijeda cijele mtDNA rezultiralo je setom konačne dužine od 16 463 bazna para. Kako su u setu bile prisutne insercije, delecije, dvoznačne pozicije i ponavljajuća regija njihovo uklanjanje dalo je set ukupne dužine od 16 312 baznih parova. Prema rezultatima dobivenim programskim paketom MitoToolPy (Peng i sur., 2015.) dobiveno je 13 glavnih haplogrupa i sub-haplogrupa (Tablica S1). Iako se mogla razaznati podjela na sub-subhaplogrupe u ovom radu je naglasak zadržan na haplogrupama i sub-haplogrupama i to redom I, II, I2, P1, Q1, Q2, R1, R2, T1, T2, T3, T4, T5.

Naknadnom analizom utvrđeno je kako se na 12 pozicija nalaze mutacije koje su tri ili četiri alelne te su uklonjene. Riječ je o pozicijama 1282, 3533, 3839, 5145, 8971, 13037, 13070, 13352, 14110, 15079, 16027 i 16034. Konačni set za daljnje analize sadržavao je 16 300 parova baza.

4.2. FST KOEFICIJENTI NUKLEOTIDNIH SLJEDOVA

Standardiziranjem nukleotidnih sljedova i njihovom analizom u programskim paketima PLINK v1.9 (Purcell i sur., 2007.) i SAS 9.3 (SAS Institute Inc.) dobiven je set prikladan za izračun FST koeficijenata. Ukupno je detektirano 922 mutacije čija je FST vrijednost bila jednaka 1 ($FST = 1$). Uspoređivanje svake sa svakom haplogrupom je dalo točan broj mutacija sa značajnom FST vrijednošću između pojedinih haplogrupa (Tablica 1). Najviše mutacija koje razdvajaju haplogrupe nađeno je između haplogrupa T1 i T3 (544), a najmanje između haplogrupa Q1 i Q2 (28). Ukoliko se promotre regije mtDNA koje sadržavaju najviše mutacija sa značajnom vrijednošću FST koeficijenta (Tablica 2.) tada najviše mutacija nalazimo u D-loop regiji (136), a najmanje na regijama koje kodiraju tRNA-Met, tRNA-Ala, OL, tRNA-Ser

I tRNA-Thr (1). Na regijama koje kodiraju tRNA-Phe, tRNA-Val, tRNA-Trp i tRNA-Arg nije opažena niti jedna mutacija sa značajnom vrijednošću FST koeficijenta.

Tablica 1. Broj mutacija sa značajnom vrijednošću FST koeficijenta. Iznad dijagonale nalazi se broj mutacija na ukupnom setu, dok je ispod dijagonale broj mutacija detektiranih na SNP čipu.

Haplogrupa	I	I1	I2	P1	Q1	Q2	R1	R2	T1	T2	T3	T4	T5
I	0	75	80	194	190	169	192	189	446	274	465	193	185
I1	7	0	36	242	249	232	232	230	495	325	512	258	246
I2	4	3	0	237	247	230	232	227	491	325	507	256	244
P1	12	15	12	0	82	62	116	112	335	164	359	92	77
Q1	13	16	13	5	0	28	122	117	323	149	347	73	57
Q2	10	13	10	2	3	0	101	96	311	131	333	54	38
R1	15	18	15	9	12	9	0	33	368	202	393	131	116
R2	12	15	12	6	9	6	3	0	362	197	387	124	109
T1	29	32	29	21	21	21	27	25	0	373	544	410	305
T2	17	19	16	9	9	8	16	13	22	0	396	136	124
T3	30	33	30	24	24	24	29	26	29	24	0	328	326
T4	12	18	15	7	6	5	14	11	18	8	21	0	326
T5	13	16	13	5	4	3	12	9	18	6	21	21	0

Tablica 2. Pozicije i dužine kodirajućih i nekodirajućih regija mtDNA prema referentnom nukleotidnom slijedu NC_006853.1 te broj mutacija sa značajnom vrijednošću FST koeficijenta detektiranih na cijelom setu i na SNP čipu.

Ime regije/gena	Početna pozicija (bp)	Završna pozicija (bp)	Dužina (bp)	Broj mutacija	Broj mutacija na SNP čipu
D loop	1	363	363	30	1
tRNA-Phe	365	430	66	0	0
12s-rRNA	431	1385	955	32	1
tRNA-Val	1386	1452	67	0	0
16S rRNA	1453	3023	1571	51	2
tRNA-Leu	3024	3098	75	4	0
ND1	3101	4056	956	50	5
tRNA-Ile	4057	4125	69	4	0
tRNA-Gln	4125	4194	70	4	0
tRNA-Met	4197	4265	69	1	0
ND2	4266	5307	1042	56	3
tRNA-Trp	5308	5374	67	0	0
tRNA-Ala	5376	5444	69	1	0
tRNA-Asn	5446	5518	73	2	0
O _L	5519	5549	31	1	0
tRNA-Cys	5551	5617	67	4	0
tRNA-Tyr	5618	5685	68	3	0
COI	5687	7231	1545	76	4
tRNA-Ser	7229	7299	71	1	0
tRNA-Asp	7304	7372	69	7	0
COII	7374	8057	684	31	2
tRNA-Lys	8061	8127	67	4	0
ATP8	8129	8329	201	17	1
ATP6	8290	8970	681	35	0
COIII	8970	9750	781	40	2
tRNA-Gly	9754	9822	69	4	0
ND3	9823	10168	346	17	2
tRNA-Arg	10170	10238	69	0	0
ND4 (L)	10239	10535	297	17	0
ND4	10536	11906	1371	78	2
tRNA-His	11907	11976	70	2	0
tRNA-Ser2	11977	12036	60	4	0
tRNA-Leu2	12038	12108	71	2	0
ND5	12109	13929	1821	118	9
ND6	13930	14440	511	33	1
tRNA-Glu	14441	14509	69	3	1
CYTB	14514	15653	1140	79	5
tRNA-Thr	15658	15726	69	1	0
tRNA-Pro	15727	15791	65	4	0
D loop	15792	16338	547	106	5

4.3. FST ZNAČAJNE MUTACIJE NA KOMERCIJALNOM SNP ČIPU

Nakon analize cijelog seta nuklotidnih sljedova cijele mtDNA i utvrđivanja mutacija sa značajnom vrijednošću FST koeficijenta iste su uspoređene s mapom referentnog genoma goveda za SNP čipove UMD 3.1. Utvrđeno je da se od 922 mutacije njih 46 nalazi na komercijalnom Illumina BovineHD Genotyping BeadChip (Illumina Inc., San Diego, CA). Najviše SNP-ova značajnih za razdvajanje haplogrupa uočeno je između haplogrupa I1 i T3 (33), a najmanje između haplogrupa P1 i Q2 (2). Od 46 prisutnih mutacija najviše ih se nalazilo na području koje kodira gen *ND5*.

5. RASPRAVA

Mitohondrijska DNA se već dulji niz godina koristi kao alat za rekonstrukciju genealogije rodova i praćenje evolucije domaćih životinja s majčine strane (Wang i sur. 2014.). Komercijalni SNP čipovi po svojoj prirodi sadrže mutacije koje se mogu koristiti i često se koriste u filogenezi (Petroldi i sur., 2010.). Iako se SNP čipovi najviše koriste za asocijacijske studije cijelog genoma onu mogu predstavljati koristan izvor informacija i kod istraživanja evolucije (Makina i sur. 2014.). Na komercijalnom Illumina BovineHD Genotyping BeadChip (Illumina Inc., San Diego, CA) čipu prisutno je ukupno 344 SNP-a mtDNA, a do sada nismo zapazili rad koji bi koristio tu informaciju u filogenezi goveda ili bilo koje druge vrste.

Analiza poravnanih i editiranih nukleotidnih sljedova nakon izračuna pokazala je kako glavne haplogrupe i sub-haplogrupe razdvaja ukupno 922 polimorfna bialelna mjesta. Zaključak da upravo ti polimorfizmi utječu na razdvajanje haplogrupa donosi se na temelju vrijednosti Wright-ovog FST koeficijenta. Naime, FST koeficijent je proporcija genetske raznolikosti uzrokovane razlikama u frekvencijama alela između populacija ili korelacija između alela unutar populacije u odnosu na cijelu populaciju. Male vrijednosti FST-a označuju sličnost frekvencije alela unutar svake populacije, dok veće vrijednosti označuju različitost frekvencije alela (Holsinger i Weir, 2009.). Stoga su ovdje razmatrane samo maksimalne moguće vrijednosti FST-a odnosno $FST = 1$. Pronađena 922 polimorfna mjesta gdje je $FST = 1$, bila su neravnomjerno raspoređena po genomu. To je uobičajeno kod mtDNA gdje se gotovo uvijek najveći broj polimorfizama (mutacija) nalazi upravo na D-loop regiji. Vukašinović 2014. godine u svom diplomskom radu analizira 968 nukleotidnih sljedova s različitih dijelova mtDNA (od čega 260 cijelih sljedova) i nalazi 912 polimorfni mjesta, odnosno 499 informativnih pozicija (pojavljuju se kod više od jedne jedinice). Korištenjem 252 nukleotidna slijeda koja se nalaze unutar programskog paketa MitoToolPy (Peng i sur., 2015.) specijaliziranog za analizu mtDNA goveda postignuta je detekcija više informativnih mutacija.

Tomu je tako ponajviše jer smo upotrebom biranih nukleotidnih sljedova uspjeli obuhvatiti sve haplogrupe i sub-haplogrupe goveda detektirane do trenutka pisanja ovog rada. Vukašinović 2014. godine također nalazi najveći broj polimorfizama (139) (Tablica 3.) na D-loop regiji dok je ovdje u toj regiji detektirano 136. Iako je kontrolna regija mtDNA najčešće analizirana mtDNA regija kod goveda (Achilli i sur., 2009.), ona ne kodira gene. Stoga je bilo za očekivati manji broj SNP-ova prisutnih na SNP čipu u toj regiji. Od spomenuta 344 SNP-a prisutna na Illumina BovineHD Genotyping BeadChip (Illumina Inc., San Diego, CA) čipu njih samo 12 se nalazi u kontrolnoj regiji dok je ostatak u kodirajućim regijama. Budući da to čini ispod 10 % polimorfizama koji uobičajeno razdvajaju haplogrupe D-loop regije, može se naslutiti kako SNP čip nije pogodan za korištenje u navedene svrhe. Nadalje, od navedenih 12 postojećih SNP-ova na D-loop regiji, ovdje smo detektirali samo 6 onih koji imaju značajnu vrijednost F_{ST} koeficijenta i mogu poslužiti u razdvajanju populacija, odnosno u ovom slučaju haplogrupa. Pronalazak mutacija sa značajnom vrijednošću F_{ST} koeficijenta u kodirajućim regijama (njih 786), potvrđuje navode drugih autora (npr. Torroni i sur., 2006.) kako analize samo malog dijela same mtDNA, kao što je kontrolna regija, mogu biti neadekvatne. Osim što se gubi dio informacije vezane za filogenezu jedinke, jednako tako se gubi i mogućnost proučavanja polimorfizama kodirajućih regija mtDNA koje mogu dati informaciju kako o mitohondrijskim bolestima, tako i o drugim utjecajima tih polimorfizama na fenotip jedinke. Kako su komercijalni SNP čipovi upravo i orijentirani na asocijacijske studije tako ne čudi činjenica da su SNP-ovi prisutni na njima češći u kodirajućim regijama.

Analiza broja mutacija koje odvajaju haplogrupe pokazala je kako najveći broj mutacija odvajaju haplogrupe T1 i T3 (544). Iako rezultat na početku djeluje nelogičan, budući da su pitanju sestrinske haplogrupe, treba imati na umu ekstremno velik broj sub-subhaplogrupa i još nižih kategorija (Olivieri i sur., 2015.) unutar T1 i T3 sub-haplogrupa. Taj broj je uvjetovan i daleko najvećom prisutnošću uzoraka općenito u T haplogrupi budući da je ona specifična za

moderna goveda *Bos taurus*. Treba primijetiti da velik broj mutacija razdvaja i *Bos taurus* od *Bos indicus*-a, ili preciznije sub-haplogrupe I, I1 i I2 od T1, T2 i T3 (Tablica 2.). To je očekivan rezultat obzirom na porijeklo ovih haplogrupa. Na slično stanje nailazimo i kada promatramo SNP-ove sa značajnom F_{ST} vrijednošću sa SNP čipa. Ponovno vidimo da najveći broj SNP-ova razdvaja upravo taurine od zebruina. Isto tako sub-haplogrupa T3 ima ukupno najveći broj SNP-ova koji je razdvaja od svih ostalih. Ovakav rezultat je ponovno očekivan budući da se najveći broj modernih goveda nalazi upravo u sub-haplogrupi T3 (Olivieri i sur., 2015.), pa tako postoji i pristranost SNP čipa koji je stvoren ponajviše na temelju komercijalnih i modernih pasmina.

Kako se od 344 SNP-a prisutna na komercijalnom Illumina BovineHD Genotyping BeadChip (Illumina Inc., San Diego, CA) čipu, njih samo 46 može tretirati kao SNP-ovi koji mogu poslužiti razdvajanju populacija, tako nije izgledno da se ovaj alat može koristiti u analizi filogeneze mtDNA haplogrupa goveda. Trebalo bi konstruirati takav SNP čip koji bi sadržavao veći broj polimorfizama značajnih za filogenezu mtDNA i na taj način postupak genotipizacije učiniti još učinkovitijim, a dobivene podatke iskoristivijim na više razina.

6. ZAKLJUČCI

1. Poravnavanje 252 nukleotidna slijeda cijele mtDNA rezultiralo je setom konačne dužine od 16 463 bazna para.
2. Prema rezultatima dobivenim programskim paketom MitoToolPy (Peng i sur., 2015.) dobiveno je 13 glavnih haplogrupa i sub-haplogrupa I, I1, I2, P1, Q1, Q2, R1, R2, T1, T2, T3, T4, T5.
3. Standardiziranjem nukleotidnih sljedova i njihovom analizom u programskim paketima PLINK v1.9 (Purcell i sur., 2007.) i SAS 9.3 (SAS Institute Inc.) detektirane su 922 mutacije čija je FST vrijednost bila jednaka 1 ($FST = 1$).
4. Najviše mutacija koje razdvajaju haplogrupe nađeno je između haplogrupa T1 i T3 (544), a najmanje između haplogrupa Q1 i Q2 (28).
5. Regije mtDNA s najviše mutacija sa značajnom vrijednošću FST koeficijenta (Tablica 2.) je D-loop regija (136).
6. Od detektiranih 922 mutacije njih 46 nalazi se na komercijalnom Illumina BovineHD Genotyping BeadChip (Illumina Inc., San Diego, CA).
7. Najviše SNP-ova značajnih za razdvajanje haplogrupa uočeno je između haplogrupa I1 i T3 (33), a najmanje između haplogrupa P1 i Q2 (2).
8. Nije izgledno da se Illumina BovineHD Genotyping BeadChip (Illumina Inc., San Diego, CA) čip može koristiti u analizi filogeneze mtDNA haplogrupa goveda.
9. Buduće generacije SNP čipova trebale bi sadržavati veći broj polimorfizama značajnih za filogenezu mtDNA i na taj način postupak genotipizacije učiniti još učinkovitijim..

7. ZAHVALE

Zahvaljujemo se:

Doc. dr. sc. Maji Ferenčaković na stručnom vodstvu, savjetima i sugestijama, te na strpljenju pri izradi ovoga rada.

Prof. dr. sc. Ini Čuriku i prof. dr. sc. Vlatki Čubrić Čurik na pruženom znanju bez kojeg ne bismo mogli izraditi ovaj rad.

8. POPIS LITERATURE

Abaci, S. H., Önder, H. & Spurlock, D. M. (2016.). Determining the relationships between genomic and phenotypic breeding values.

Achilli, A., Bonfiglio, S., Olivieri, A., Malusà, A., Pala, M., Kashani, B. H. & Bandelt, H. J. (2009.). The multifaceted origin of taurine cattle reflected by the mitochondrial genome. *PLoS One*, 4(6), e5753.

Achilli, A., Olivieri, A., Pellecchia, M., Uboldi, C., Colli, L., Al-Zahery, N. & Battaglia, V. (2008.). Mitochondrial genomes of extinct aurochs survive in domestic cattle. *Current Biology*, 18(4), R157-R158.

Ančić, V., Bogut, I. & Đumljija, S. (2008.). From molecule to organism.

Bollongino, R., Burger, J., Powell, A., Mashkour, M., Vigne, J. D. & Thomas, M. G. (2012.). Modern taurine cattle descended from small number of Near-Eastern founders. *Molecular biology and evolution*, mss092.

Bonfiglio, S., Achilli, A., Olivieri, A., Negrini, R., Colli, L., Liotta, L. & Ferretti, L. (2010.). The enigmatic origin of bovine mtDNA haplogroup R: sporadic interbreeding or an independent event of *Bos primigenius* domestication in Italy? *PLoS One*, 5(12), e15760.

Bonfiglio, S., Ginja, C., De Gaetano, A., Achilli, A., Olivieri, A., Colli, L. & Penedo, M. C. T. (2012.). Origin and spread of *Bos taurus*: new clues from mitochondrial genomes belonging to haplogroup T1. *PLoS One*, 7(6), e38601.

Brown W.M., Prager E.M., Wang A. & Wilson A.C. (1982.). Mitochondrial DNA sequences of primates: tempo and mode of evolution. *Journal of molecular evolution* 18, 225-39.

Chan E.K., Nagaraj S.H. & Reverter A. (2010.). The evolution of tropical adaptation: comparing taurine and zebu cattle. *Animal genetics* 41, 467-77.

Cooper G.M., Hausman R.E., Lauc G. & Zoldoš V. (2010.). Stanica: molekularni pristup. Medicinska naklada.

Ferenčaković, M., Sölkner, J. & Curik, I. (2013.). Estimating autozygosity from high-throughput information: effects of SNP density and genotyping errors. *Genetics Selection Evolution*, 45(1), 42.

Holsinger, K. E. & Weir, B. S. (2009.). Genetics in geographically structured populations: defining, estimating and interpreting F_{ST} . *Nature Reviews Genetics*, 10(9), 639-650.

Holt, I. J., Lorimer, H. E. & Jacobs, H. T. (2000.). Coupled leading-and lagging-strand synthesis of mammalian mitochondrial DNA. *Cell*, 100(5), 515-524.

Kumar, S., Stecher, G. & Tamura, K. (2016.). MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets; *Molecular Biology and Evolution* 33:1870-1874.

Li, W. H., Gojobori, T. & Nei, M. (1981.). Pseudogenes as a paradigm of neutral evolution. *Nature*, 292(5820), 237-239.

Loftus R., MacHugh D., Bradley D., Sharp P. & Cunningham P. (1994.). Evidence for two independent domestications of cattle. *Proceedings of the National Academy of the United States of America* 91, 2757 - 61.

Makina, S. O., Muchadeyi, F. C., van Marle-Köster, E., MacNeil, M. D. & Maiwashe, A. (2014.). Genetic diversity and population structure among six cattle breeds in South Africa using a whole genome SNP panel.

Mannen, H., Kohno, M., Nagata, Y., Tsuji, S., Bradley, D. G., Yeo, J. S. & Amano, T. (2004.). Independent mitochondrial origin and historical genetic differentiation in North Eastern Asian cattle. *Molecular phylogenetics and evolution*, 32(2), 539-544.

Martínez-Arias, R., Calafell, F., Mateu, E., Comas, D., Andrés, A. & Bertranpetit, J. (2001.). Sequence variability of a human pseudogene. *Genome research*, 11(6), 1071-1085.

Mona, S., Catalano, G., Lari, M., Larson, G., Boscato, P., Casoli, A. & Bertorelle, G. (2010.). Population dynamic of the extinct European aurochs: genetic evidence of a north-south differentiation pattern and no evidence of post-glacial expansion. *BMC evolutionary biology*, 10(1), 83.

Olivieri, A., Gandini, F., Achilli, A., Fichera, A., Rizzi, E., Bonfiglio, S. & Lancioni, H. (2015.). Mitogenomes from Egyptian Cattle Breeds: New clues on the origin of haplogroup Q and the early spread of *Bos taurus* from the Near East. *PloS one*, 10(10), e0141170.

Peng, M. S., Fan, L., Shi, N. N., Ning, T., Yao, Y. G., Murphy, R. W., Wang, W. Z. & Zhang, Y. P. (2015.). DomeTree: a Canonical Toolkit for Mitochondrial DNA Analyses in Domesticated Animals. *Molecular Ecology Resources*. doi: 10.1111/1755-0998.12386.

Pertoldi, C., Tokarska, M., Wójcik, J. M., Kawalko, A., Randi, E., Kristensen, T. N. & Bendixen, C. (2010.). Phylogenetic relationships among the European and American bison and seven cattle breeds reconstructed using the BovineSNP50 Illumina Genotyping BeadChip. *Acta Theriologica*, 55(2), 97-108.

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M., Bender, D., Maller, J., Sklar, P., de Bakker, P., Daly, M. J. & Sham, P. C. (2007.). PLINK: A Tool Set for Whole-Genome and Population-Based Linkage Analyses. *American Journal of Human Genetics*, 81.

Su, G., Brøndum, R. F., Ma, P., Guldbbrandtsen, B., Aamand, G. P. & Lund, M. S. (2011.). Genomic prediction using high-density SNP markers in Nordic Holstein and Red. *Interbull Bulletin*, (44).

Torrioni A., Achilli A., Macaulay V., Richards M. & Bandelt H.-J. (2006.). Harvesting the fruit of the human mtDNA tree. *Trends in Genetics* 22, 339-45.

Van Eenennaam, A. L. (2011.). Improving EPD accuracy by combining EPD information with DNA test results. *Proc. of Applied Reproductive Strategies in Beef Cattle*. Joplin, Missouri.

Van Vuure C. (2005.). *Retracing the Aurochs: History, Morphology and Ecology of an Extinct Wild Ox*. Coronet Books Incorporated.

Vignal, A., Milan, D., SanCristobal, M. & Eggen, A. (2002.). A review on SNP and other types of molecular markers and their use in animal genetics. *Genetics Selection Evolution*, 34(3), 275-306.

Vukašinović Zoran (2014.). *Diplomski rad, Analiza polimorfnosti sekvence mitohondrijske DNA goveda*, Agronomski fakultet

Wang G. D., Xie H. B., Peng M. S., Irwin D, Zhang Y. P. (2014.). Domestication genomics: evidence from animals. *Annual Review of Animal Biosciences*, 2, 65–84

Wang, Y., Lin, G., Li, C. & Stothard, P. (2016.). Genotype Imputation Methods and Their Effects on Genomic Predictions in Cattle. *Springer Science Reviews*, 4(2), 79-98.

Wright, S. (1922.). Coefficients of inbreeding and relationship. *Am. Nat.* 56:330–338.

Yang, M. Y., Bowmaker, M., Reyes, A., Vergani, L., Angeli, P., Gringeri, E. & Holt, I. J. (2002.). Biased incorporation of ribonucleotides on the mitochondrial L-strand accounts for apparent strand-asymmetric DNA replication. *Cell*, 111(4), 495-505.

Zeviani, M., & Di Donato, S. (2004.). Mitochondrial disorders. *Brain*, 127(10), 2153-2172.

Zeyland, J., Wolko, Ł., Bocianowski, J., Szalata, M., Słomski, R., Dzieduszycki, A. M. & Lipiński, D. (2013.). Complete mitochondrial genome of wild aurochs (*Bos primigenius*) reconstructed from ancient DNA. *Polish journal of veterinary sciences*, 16(2), 265-273.

<http://en.wikipedia.org/wiki/Cattle> (zadnje pristupljeno 27. travnja 2017.).

http://www.ensembl.org/Bos_taurus/Info/Annotation (zadnje pristupljeno 27. travnja 2017.).

<http://www.google.hr/search?q=mitohondrij&client=firefox> (zadnje pristupljeno 27. travnja 2017.).

<http://www.illumina.com/> (zadnje pristupljeno 27. travnja 2017.).

<http://www.instrukcije-poduka.com/instrukcije-iz-biologije-mitohondriji.html> (zadnje pristupljeno 27. travnja 2017.).

https://aipl.arsusda.gov/publish/WC9/WC9_10_cvt_jbc.ppt (zadnje pristupljeno 27. travnja 2017.).

<https://en.wikipedia.org/wiki/F-statistics> (zadnje pristupljeno 27. travnja 2017.).

<https://sh.wikipedia.org/wiki/Zebu> (zadnje pristupljeno 27. travnja 2017.).

9. SAŽETAK

Mitohondrijska DNA uobičajen je alat kod proučavanja filogeneze rodova i evolucije modernog goveda s majčine strane. Do slijeda nukleotida mtDNA najčešće se dolazi putem metode lančane reakcije polimerazom, pri čemu se dobiva cijeli slijed koji sadrži i varijabilna i konzervirana mjesta. Komercijalni SNP čipovi po svojoj prirodi sadrže samo varijabilna mjesta te se ponekad koriste u filogenezi nuklearne DNA, no nije nam poznata upotreba varijabilnih mjesta mtDNA prisutnih na SNP čipu u istu svrhu. Stoga je cilj ovog rada bio utvrditi prisutnost mutacija (SNP-ova) značajnih za razdvajanje haplogrupa mitogenoma goveda prisutnih na komercijalnom Illumina BovineHD Genotyping BeadChip i provjeriti mogućnost njihove upotrebe u filogenezi goveda. Analizom 252 nukleotidna slijeda cijelog mitogenoma goveda, preuzeta iz banke gena utvrdili smo 13 glavnih haplogrupa i sub-haplogrupa. Izračunom Wright-ovog F_{ST} koeficijenta između haplogrupa i sub-haplogrupa pronađeno je ukupno 922 pozicije čija vrijednost F_{ST} koeficijenta upućuje na činjenicu da omogućavaju razdvajanje haplogrupa i sub-haplogrupa. Usporedbom pozicije mutacija sa značajnom vrijednošću F_{ST} koeficijenta na nukleotidnim sljedovima sa SNP-ovima prisutnima na SNP čipu, pronađeno je samo 46 SNP-ova koji bi potencijalno mogli poslužiti u filogenezi haplogrupa mitogena. Obzirom na tako nizak broj, nije izgledno da se komercijalni Illumina BovineHD Genotyping BeadChip može koristiti u analizi filogeneze mtDNA haplogrupa goveda. Sljedeće generacije SNP čipova trebalo bi konstruirati na način da sadrže veći broj polimorfizama značajnih za filogenezu mtDNA i na taj način postupak genotipizacije učiniti još učinkovitijim, a dobivene podatke iskoristive na više razina.

Ključne riječi: mitogenom goveda, SNP čip, haplogrupa, Wright-ov F_{ST} koeficijent, filogeneza

10. SUMMARY

Mitochondrial DNA is a common and useful tool for analysis of maternal lineages of the modern cattle breeds, as well as for tracing evolutionary history from a matrilineal perspective. The usual way to obtain mtDNA is through the method of polymerase chain reaction, where resulting sequence has both variable and conserved sites. On the other side, commercial SNP chips are constructed to detect only variable sites and they are sometimes used for phylogeny analysis. However, we are not aware of a study that uses mtDNA polymorphic sites from the SNP chip in order to investigate maternal lineages. Consequently, the aim of this study was to detect the presence of mtDNA SNPs on commercial Illumina BovineHD Genotyping BeadChip, and to check if they are usable for phylogenetic analyses of cattle. Analysis of 252 whole mtDNA sequences from gene bank resulted in 13 main haplogroups and sub-haplogroups. By calculating Wright's F_{ST} coefficient between those haplo and sub-haplogroups we detected 922 polymorphic sites with the value of F_{ST} significant for their divergence. By comparing positions of detected polymorphic sites with the positions of the SNPs on the commercial SNP chip we found only 46 potential candidates with significant F_{ST} value. Therefore Illumina BovineHD Genotyping BeadChip is most likely not usable for the analysis of mtDNA phylogeny in cattle. New generations of SNP chips should incorporate more polymorphic sites that are significant for haplogroup divergence. This way genotyping would be even more effective and obtained data could be used for more purposes.

Key words: bovine mitogenome, SNP chip, haplogroup, Wright's F_{ST} coefficient, phylogeny

11. ŽIVOTOPISI

Iva Katić

Rođena sam 02. rujna 1994. u Zagrebu. Pohađala sam OŠ „Fran Krsto Frankopan“ do sedmog razreda, te sam zbog preseljenja sedmi i osmi razred završila u OŠ „Špansko Oranice“. 2009. godine upisujem Prirodoslovnu školu „Vladimir Prelog“, smjer Ekološki tehničar u kojoj sam bila 4 godine predsjednica razreda i predsjednica vijeća 2 godine. Agronomski fakultet upisujem 2013. godine, te 2016. godine postajem prvostupnica agroekologije. 2016. godine upisujem prvu godinu diplomskog studija Genetika i oplemenjivanje životinja.

Aktivno se bavim športom trinaest godina, u zboru sam pjevala 6 godina i bila sam pet godina član folklornog ansambla. Radim u športskom dućanu Decathlon, dio sam internet agencije Luxury islands & retreats travel agency, te ljeti radim na recepciji u poslovnoj jedinici BCI - Business Center International. Poznavanje engleskog jezika B razine.

Martina Jašinski

Rođena sam 06. kolovoza 1994. godine u Zagrebu. Pohađala sam OŠ „Davorin Trstenjak“ i nakon toga sam upisala Prirodoslovnu školu „Vladimir Prelog“, smjer Ekološki tehničar. Maturirala sam 2013. godine, te nakon srednje škole upisujem preddiplomski studij Agronomskog fakulteta u Zagrebu, smjer Agroekologija. Nakon završetka preddiplomskog studija, 2016. godine upisujem diplomski studij Genetika i oplemenjivanje životinja. Bavim se aktivno jahanjem 13 godina, sudjelujem na natjecanjima iz dresurnog i preponskog jahanja. Radim u športskom dućanu Decathlon. Poznavanje engleskog jezika B razine.

Tea Glunčić

Rođena sam 14. lipnja 1994. godine u Zagrebu. Pohađala sam OŠ „Davorin Trstenjak“ te nakon toga upisala sam 15. gimnaziju (MIOC), matematičko-informatički smjer. Nakon toga upisujem preddiplomski studij Agroekologije na Agronomskom fakultetu u Zagrebu. Godine 2016. postala sam sveučilišna prvostupnica inženjerka agroekologije. Na jesen 2016. godine upisujem diplomski studij Genetika i oplemenjivanje životinja. Rekreativno se bavim plesom i pilatesom dugi niz godina. Poznavanje engleskog i njemačkog jezika B razine.

Zdravko Begić

Rođen sam 23. studenog 1975. godine u Zagrebu. Pohađao sam srednjoškolsko obrazovanje Birotehnika, smjer Upravni referent. Radio sam u ugostiteljstvu sve do odluke upisa preddiplomskog studija akademske godine 2012./2013. godine, te stječem diplomu prvostupnik inženjer animalnih znanosti. Nakon završetka preddiplomskog studija nastavljam s diplomskim studijem na Agronomskom fakultetu smjer Genetika i oplemenjivanje životinja akademska godina 2016./2017. Od ožujka 2005. godine ponosan sam otac Begić Tomasa. Poznavanje engleskog jezika B razine.

12. PRILOZI

Tablica S1. Popis preuzetih nuklotidnih sljedova cijele mtDNA goveda i pripadajuće haplogrupe prema radu (Peng i sur., 2015.).

Redni broj	NCBI oznaka	Haplogrupa	Redni broj	NCBI oznaka	Haplogrupa
1	V00654	T3	46	DQ124393.1	T2
2	AY526085	T1	47	DQ124394.1	T3
3	AB074962.1	T4	48	DQ124395.1	T3
4	AB074963.1	T4	49	DQ124396.1	T2
5	AB074964.1	T4	50	DQ124397.1	T3
6	AB074965.1	T3	51	DQ124398.1	T3
7	AB074966.1	T3	52	DQ124399.1	T1
8	AB074967.1	T3	53	DQ124400.1	T4
9	AB074968.1	T3	54	DQ124401.1	T4
10	AY126697.1	I1	55	DQ124402.1	T3
11	AY526085.1	T1	56	DQ124403.1	I
12	AY676855.1	T3	57	DQ124404.1	T3
13	AY676856.1	T2	58	DQ124405.1	T3
14	AY676857.1	T3	59	DQ124406.1	T3
15	AY676858.1	T3	60	DQ124407.1	T3
16	AY676859.1	T3	61	DQ124408.1	T3
17	AY676860.1	T3	62	DQ124409.1	T3
18	AY676861.1	T3	63	DQ124410.1	T3
19	AY676862.1	T3	64	DQ124411.1	T3
20	AY676863.1	T3	65	DQ124412.1	T4
21	AY676864.1	T3	66	DQ124413.1	T3
22	AY676865.1	T3	67	DQ124414.1	T3
23	AY676866.1	T3	68	DQ124415.1	T3
24	AY676868.1	T3	69	DQ124416.1	T3
25	AY676869.1	T3	70	DQ124417.1	T3
26	AY676871.1	T3	71	DQ124418.1	T3
27	AY676872.1	T3	72	EU177815.1	T3
28	AY676873.1	T3	73	EU177816.1	T3
29	DQ124372.1	T4	74	EU177817.1	T3
30	DQ124373.1	T3	75	EU177818.1	T3
31	DQ124374.1	T3	76	EU177819.1	T3
32	DQ124375.1	T4	77	EU177820.1	T3
33	DQ124376.1	T3	78	EU177821.1	T3
34	DQ124377.1	T4	79	EU177822.1	T3
35	DQ124380.1	T3	80	EU177823.1	T3
36	DQ124382.1	T3	81	EU177824.1	T3
37	DQ124383.1	T2	82	EU177825.1	T3
38	DQ124384.1	T3	83	EU177826.1	T3
39	DQ124385.1	T3	84	EU177827.1	T3
40	DQ124386.1	T3	85	EU177828.1	T3
41	DQ124387.1	T3	86	EU177829.1	T3
42	DQ124388.1	T3	87	EU177830.1	T3
43	DQ124390.1	T3	88	EU177831.1	T3
44	DQ124391.1	T3	89	EU177832.1	T3

45	DQ124392.1	T4	90	EU177833.1	T3
----	------------	----	----	------------	----

NCBI „National Center for Biotechnology Information“ (<https://www.ncbi.nlm.nih.gov/>)

Tablica S1. Popis preuzetih nuklotidnih sljedova cijele mtDNA goveda i pripadajuće haplogrupe prema radu (Peng i sur., 2015.). - nastavak

Redni broj	NCBI oznaka	Haplogrupa	Redni broj	NCBI oznaka	Haplogrupa
91	EU177834.1	T3	136	GU947020	T1
92	EU177835.1	T3	137	GU947021	T3
93	EU177836.1	T3	138	HM045018	T3
94	EU177837.1	T3	139	HQ025805	T2
95	EU177838.1	T3	140	HQ184030	Q2
96	EU177839.1	T3	141	HQ184031	Q2
97	EU177841.1	T1	142	HQ184032	Q2
98	EU177842.1	T1	143	HQ184033	Q2
99	EU177843.1	T1	144	HQ184034	Q1
100	EU177844.1	T1	145	HQ184035	Q1
101	EU177846.1	T1	146	HQ184036	Q1
102	EU177847.1	T1	147	HQ184037	Q1
103	EU177848.1	T1	148	HQ184038	Q1
104	EU177849.1	T2	149	HQ184039	Q1
105	EU177850.1	T2	150	HQ184040	R1
106	EU177851.1	T2	151	HQ184041	R1
107	EU177852.1	T2	152	HQ184042	R1
108	EU177853.1	T2	153	HQ184043	R1
109	EU177854.1	T2	154	HQ184044	R1
110	EU177855.1	T2	155	HQ184045	R2
111	EU177856.1	T2	156	JN817298	T1
112	EU177857.1	T2	157	JN817299	T1
113	EU177858.1	T2	158	JN817300	T1
114	EU177859.1	T2	159	JN817301	T1
115	EU177860.1	T2	160	JN817302	T1
116	EU177861.1	T2	161	JN817303	T1
117	EU177862.1	T5	162	JN817304	T1
118	EU177863.1	T5	163	JN817305	T1
119	EU177864.1	T5	164	JN817306	T1
120	EU177865.1	T5	165	JN817307	T1
121	EU177866.1	Q1	166	JN817308	T1
122	EU177867.1	Q1	167	JN817309	T1
123	EU177868.1	I1	168	JN817310	T1
124	EU177870	I2	169	JN817311	T1
125	FJ971080	Q2	170	JN817312	T1
126	FJ971082	Q1	171	JN817313	T1
127	FJ971083	Q1	172	JN817314	T1
128	FJ971084	R1	173	JN817315	T1
129	FJ971085	R1	174	JN817316	T1
130	FJ971086	R1	175	JN817317	T1
131	FJ971087	R2	176	JN817318	T1
132	FJ971088	I1	177	JN817319	T1
133	GQ129207	T3	178	JN817320	T1
134	GQ129208	T3	179	JN817321	T1

135	GU947019	T3	180	JN817322	T1
-----	----------	----	-----	----------	----

NCBI „National Center for Biotechnology Information“ (<https://www.ncbi.nlm.nih.gov/>)

Tablica S1. Popis preuzetih nuklotidnih sljedova cijele mtDNA goveda i pripadajuće haplogrupe prema radu (Peng i sur., 2015.). - nastavak

Redni broj	NCBI oznaka	Haplogrupa	Redni broj	NCBI oznaka	Haplogrupa
181	JN817323	T1	217	KC153977	T3
182	JN817324	T1	218	KF163061	T1
183	JN817325	T1	219	KF163063	T1
184	JN817326	T1	220	KF163064	T1
185	JN817327	T1	221	KF163065	T1
186	JN817328	T1	222	KF163066	T1
187	JN817329	T1	223	KF163067	T1
188	JN817330	T1	224	KF163069	T1
189	JN817331	T1	225	KF163070	T1
190	JN817332	T1	226	KF163071	T1
191	JN817333	T1	227	KF163072	T1
192	JN817334	T1	228	KF163073	T1
193	JN817335	T1	229	KF163075	T1
194	JN817336	T1	230	KF163076	T1
195	JN817337	T1	231	KF163077	T1
196	JN817338	T1	232	KF163078	T1
197	JN817339	T1	233	KF163079	T1
198	JN817340	T1	234	KF163080	T1
199	JN817341	T1	235	KF163081	T1
200	JN817342	T1	236	KF163082	T1
201	JN817343	T1	237	KF163083	T1
202	JN817344	T1	238	KF163084	T1
203	JN817345	T1	239	KF163085	T1
204	JN817346	T1	240	KF163086	T1
205	JN817347	T1	241	KF163089	T1
206	JN817348	T1	242	KF163090	T1
207	JN817349	T1	243	KF163091	T1
208	JN817350	T1	244	KF163092	T1
209	JN817351	T1	245	KF163093	T1
210	JQ437479	P1	246	KF163094	T1
211	JQ967333	T3	247	KF926377	T3
212	KC153972	T3	248	KT184456	T2
213	KC153973	T1	249	KT184470	T1
214	KC153974	T3	250	NC_005971	I1
215	KC153975	T1	251	NC_013996	P1
216	KC153976	T3	252	EU177869.1	I2

NCBI „National Center for Biotechnology Information“ (<https://www.ncbi.nlm.nih.gov/>)