Detekcija visoko heterozigotnih regija u genomu goveda

Mercvajler, Mario

Master's thesis / Diplomski rad

2017

Degree Grantor / Ustanova koja je dodijelila akademski / stručni stupanj: **University of Zagreb, Faculty of Agriculture / Sveučilište u Zagrebu, Agronomski fakultet**

Permanent link / Trajna poveznica: https://urn.nsk.hr/urn:nbn:hr:204:298634

Rights / Prava: In copyright / Zaštićeno autorskim pravom.

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



Repository / Repozitorij:

Repository Faculty of Agriculture University of Zagreb







SVEUČILIŠTE U ZAGREBU
AGRONOMSKI FAKULTET

DETEKCIJA VISOKO HETEROZIGOTNIH REGIJA U GENOMU GOVEDA

DIPLOMSKI RAD

Mario Mercvajler

Zagreb, rujan 2017.

SVEUČILIŠTE U ZAGREBU
AGRONOMSKI FAKULTET

Diplomski studij:
Genetika i oplemenjivanje domaćih životinja

DETEKCIJA VISOKO HETEROZIGOTNIH REGIJA U GENOMU GOVEDA

DIPLOMSKI RAD

Mario Mercvajler

Mentor: Doc. dr.sc. Maja Ferenčaković

Zagreb, rujan 2017.

SVEUČILIŠTE U ZAGREBU
AGRONOMSKI FAKULTET

IZJAVA STUDENTA

O AKADEMSKOJ ČESTITOSTI

Ja, **Mario Mercvajler**,JMBAG 0178087866, rođen 21.05.1992. u Wangen im Allgau,

izjavljujem da sam samostalno izradio diplomski rad pod naslovom:

DETEKCIJA VISOKO HETEROZIGOTNIH REGIJA U GENOMU GOVEDA

Svojim potpisom jamčim:

- da sam jedina autorica/jedini autor ovoga diplomskog rada;
- da su svi korišteni izvori literature, kako objavljeni tako i neobjavljeni, adekvatno citirani ili parafrazirani, te popisani u literaturi na kraju rada;
- da ovaj diplomski rad ne sadrži dijelove radova predanih na Agronomskom fakultetu ili drugim ustanovama visokog obrazovanja radi završetka sveučilišnog ili stručnog studija;
- da je elektronička verzija ovoga diplomskog rada identična tiskanoj koju je odobrio mentor;
- da sam upoznata/upoznat s odredbama Etičkog kodeksa Sveučilišta u Zagrebu (Čl. 19).

U Zagrebu, dana _____ _____

Potpis studenta

SVEUČILIŠTE U ZAGREBU
AGRONOMSKI FAKULTET

IZVJEŠĆE
O OCJENI I OBRANI DIPLOMSKOG RADA

Diplomski rad studenta **Mario Mercvajler**, JMBAG 0178087866,naslova

DETEKCIJA VISOKO HETEROZIGOTNIH REGIJA U GENOMU GOVEDA

obranjen je i ocijenjen ocjenom _____ , dana _____ .

Povjerenstvo:		potpisi:
1. Doc. dr. sc. Maja Ferenčaković	mentor	_____
2. Izv. prof. dr. sc. Vlatka Čubrić Čurik	član	_____
3. Prof. dr. sc. Ino Čurik	član	_____

Zahvala

Zahvaljujem se:

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

Golden Helixu iz SAD-a na ustupljenoj licenci za programski paket SNP & Variation Suite (v8.4.0 Win64; Golden Helix, Bozeman, MT, USA www.goldenhelix.com) koji mi je uvelike pomogao.

Univ. prof. dipl. ing. dr. rer. nat. Johannu Sölkneru s "Universität für Bodenkultur" Beč, Austrija na ustupljenim genotipovima.

Članovima stručnog povjerenstva na sugestijama.

SADRŽAJ

1 UVOD 1

1.1 Cilj istraživanja 2

2 PREGLED LITERATURE 3

2.1 Uzgoj u srodstvu 3

2.2 Inbriding depresija 4

2.3 “Runs of Homozigosity“ 4

2.4 Koeficijent uzgoja u srodstvu 5

2.5 “Heterozygosity Rich Regions” 6

2.6 “Copy number variations“ - CNV 7

2.7 Polimorfizam jednog nukleotida 7

2.8 Karakteristike Pinzgauer goveda 9

3 MATERIJAL I METODE 11

3.1 Kontrola kvalitete genotipova 11

3.2 Procjena genetskih parametara 11

3.3 Detekcija HRR segmenata i regija 11

3.4 Određivanje CNV-a 12

3.5 Validacija HRR segmenata 12

3.6 Detekcija gena u HRR segmentima 12

4 REZULTATI 14

4.1 Kontrola kvalitete genotipova i procjena genetskih parametara 14

4.2 Detekcija CNV i HRR. 16

4.3 Mapiranje gena 27

5 RASPRAVA 33

6 ZAKLJUČCI 35

7 POPIS LITERATURE 36

Sažetak

Diplomskog rada studenta **Mario Mercvajler**, naslova

DETEKCIJA VISOKO HETEROZIGOTNIH REGIJA U GENOMU GOVEDA

U posljednje vrijeme velik interes znanstvenika izazivaju "Runs of Homozygosity" (ROH) segmenti - kontinuirani dijelovi DNA nukleotidnog slijeda bez heterozigotnosti u diploidnom obliku. Nastaju kao posljedica uzgoja u srodstvu koji vodi do inbriding depresije, odnosno negativno utječe na "fitness" i proizvodna svojstva. Suprotno pojavi ROH-ova, u genomu se mogu naći „Heterozygosity rich regions" (HRR) regije s visokim postotkom heterozigotnosti. Na uzorku od 120 Pinzgauer bikova genotipiziranih HD SNP čipom detektirano je 147126 HRR. Prosječna pokrivenost genoma HRR-ovima iznosila je 0,99% (0,57-1,13%). Detektirana su i 92 česta HRR-a (proporcija veća od 0,25) gdje je nađen ukupno 91 gen (78 s opisanom funkcijom). Funkcije nađenih gena su važne za biološke procese i pojave bolesti te imaju utjecaj na "fitness" jedinke. Ovo istraživanje, za razliku od prijašnjih istraživanja, preciznije detektira same HRR-ove i zasigurno će bit polazna točka za buduća istraživanja u ovom području.

Ključne riječi: "Heterozygosity rich regions ", "Copy number variation", SNP podaci, govedo

Summary

Of the master's thesis – student **Mario Mercvajler,** entitled

DETECTION OF HETEROZYGOSITY RICH REGIONS IN CATTLE GENOME

Recently, scientists have been greatly interested in Runs of Homozygosity (ROH) segments - continuous parts of the DNA nucleotide sequence without heterozygosity in diploid state. They arise as a result of inbreeding that leads to inbreeding depression, ie negatively affecting the "fitness" and production properties. Contrary to the appearance of ROH, regions with a high percentage of heterozygosity can also be found in the genome - Heterozygosity Rich Regions (HRR). On the sample of 120 Pinzgauer bulls genotyped with HD SNP chip 147126 HRR were detected. The HRR average genome coverage was 0.99% (0.57-1.13%). 92 common HRR (proportion greater than 0.25) were detected, where a total of 91 genes (78 with the described function) were found. The functions of the discovered genes are important for biological processes and the occurrence of the disease and have an effect on the "fitness" of the individual. This research, unlike previous ones, detects HRRs more precisely and will certainly be the starting point for future research in this area.

Keywords:„Heterozygosity rich regions“, Copy number variation, SNP data, cattle

1 UVOD

Govedarstvo, kao jedna od ključnih grana u stočarskoj proizvodnji, susreće se s mnogim izazovima, a jedan od glavnih je kontrola uzgoja u srodstvu ili inbridinga i njegovih štetnih posljedica. Inbriding ne može se u potpunosti izbjeći u govedarstvu, budući da su današnje moderne pasmine goveda stvorene korištenjem ove metode. Jednako tako, zbog visokog stupnja selekcije i upotrebom umjetnog osjemenjivanja putem kojeg jedan bik poželjnih osobina može biti otac tisućama teladi, uvijek postoji i opasnost od nekontroliranog uzgoja u srodstvu. Posljedice inbridinga predstavljaju velike ekonomske gubitke jer inbriding depresija može negativno utjecati na proizvodna svojstva, zdravlje i reprodukciju životinja.

Inbriding depresija se do sada uvijek procjenjivala putem regresije nekog svojstva na koeficijent uzgoja u srodstvu dobiven putem rodovnika (F_{PED}). Takav način imao je mnoge nedostatke koje nove metode procjene uzgoja u srodstvu nastoje izbjeći (Keller i sur., 2011.). Velik broj metoda javlja se razvojem molekularne genetike i sve većom dostupnošću genomskih podataka. Metoda koja se u posljednje vrijeme uspješno koristi za izračun koeficijenta uzgoja u srodstvu iz genomskih podataka je "Runs of Homozygosity" (ROH) metoda.

ROH su kontinuirani (neprekinuti) dijelovi DNA nukleotidnog slijeda bez heterozigotnosti u diploidnom obliku (Gibson i sur., 2006.), koji se mogu koristiti za izračun koeficijenta uzgoja u srodstvu (F_{ROH}) (McQuillan i sur.,2008.). F_{ROH} je preciznija mjera autozigotnosti genoma jedinke jer nije osjetljiv na greške i dubinu rodovnika te uzima u obzir stohastičku prirodu nasljeđivanja (Keller i sur., 2011.). Jednako tako ROH segmenti i F_{ROH} mogu se uspješno koristiti za procjenu inbriding depresije (Bjelland i sur.,2013.), otkrivanje regija koje uzrokuju depresiju (Keller i sur., 2012; Pryce i sur., 2014; Saura i sur., 2015.), ali i za otkrivanje uzročnih recesivnih mutacija (Drögemülleri sur., 2011.).

Suprotno pojavi ROH segmenata, u genomu se mogu naći "Runs of Heterozygosity" segmenti (Williams i sur., 2016.) koji se opisuju kao dijelovi DNA nukleotidnog slijeda bez homozigotnosti u diploidnom obliku. Ferenčaković i sur.,(2016) za razliku od Williams i sur., (2016) zbog same prirode nastanka heterozigotnih segmenata odbacuju takvu definiciju i segmente nazivaju "Heterozygosity Rich Regions" (HRR). Takve bi regije, za razliku od ROH regija, mogle sadržavati lokuse koji doprinose boljem preživljavanju, plodnosti i drugim "fitness" svojstvima (McParland i sur., 2009.;

Ferenčaković i sur., 2016). Mogućnost detektiranja gena koji doprinose boljem "fitnessu" neke populacije od interesa je i za proizvodnju i za očuvanje populacija. Pojava heterozigotnih regija u malim lokalnim populacijama mogla bi isto upućivati na prirodnu selekciju u cilju očuvanja "fitnessa" (Williams i sur., 2016.).

ROH segmenti i HRR regije se najčešće otkrivaju pomoću SNP-ova u genotipu. Tu može doći do prvih problema, jer se mogu detektirati male ROH ili HRR regije koje mogu bit posljedica lošeg genotipiziranja odnosno nedovoljno dobrog pokrivanja haplotipa od strane SNP-ova. Osim navedenog problema može doći i do niza drugih problema kao, naprimjer, pojava rijetkih alela. Algoritmi za otkrivanje ROH, a samim time i HRR regija temelje se na stupnju sličnosti haplotipova na homolognim kromosomima u diploidnoj jedinki (Howrigan i sur., 2011.). Zbog postojanja dijelova genoma koji se ponavljaju (engl. Copy Number Variatiation; CNV) i zbog velike srednje udaljenosti između markera ("*mean intermarker distances"*; IMD) može doći do pogrešnog signala od genotipa te prikaza homozignotih segmenata kada oni to nisu (Howrigan i sur., 2011.). Takvi segmenti mogu bit okarakterizirani kao ROH segmenti od strane algoritma, što oni zapravo nisu, te su posljedica postojanja CNV ili IMD u tom dijelu genoma. Isti ovi problemi vrijede i za detekciju HRR.

1.1 Cilj istraživanja

Cilj ovoga rada je proširiti istraživanje Ferenčaković i sur., (2016.) gdje su prvi puta mapirani geni u HRR. Za razliku od prije navedenog istraživanja, nastojati će se preciznije odrediti HRR na način da se detektirane regije provjere na CNV i IMD, a tek onda provede provjera gena u regiji i njihova povezanost s "finess" svojstvima.

2 PREGLED LITERATURE

2.1 Uzgoj u srodstvu

Uzgoj u srodstvu ili inbriding se odnosi na parenje roditelja koji dijele jednog ili više zajedničkih pretaka. Stoga je concept inbridinga usko povezan sa srodnosti te oboje imaju ključnu ulogu u evolucijskoj genetici (Charlesworth i Charlesworth, 1987.) i oplemenjivanju životinja (Kristensen i Sorensen, 2005.; Pirchner, 1985.). Inbriding značajno smanjuje genetsku varijabilnost unutar populacija, dovodi do promjena u efektivnoj veličini populacije, genetskom driftu, strukturi populacije. Inbriding mijenja frekvencije genotipova povećavajući homozigotnost na štetu heterozigotnosti ostavljajući frekvencije alela nepromijenjenima. Prema Curik i sur., (2014.) inbriding može dovesti do preraspodjele genetskih varijacija između i unutar populacija (Fernandez i sur., 1995.), smanjenja prosjeka populacije za svojstva usko povezanih sa "fitnessom" (Charlesworth i Willis, 2009.) i povećane učestalosti homozigotnih recesivnih alela (Burgos i Muenke, 2002; Alvarezi sur., 2009.) te smanjenja u homeostazi (Lerner, 1954.). Curik i sur., (2014.) opisuju inbriding kao dvosjekli mač, koji zajedno s genetskim driftom i snagom selekcije, može dovesti do poboljšanja ili deformacije jedinki ili cijele populacije.

Negativne posljedice inbridinga primijećene su vrlo rano u ljudskoj povijesti, pa tako gotovo sve kulture brane i strogo osuđuju incest, budući da se kod potomaka iz takvih zajednica javlja puno veća vjerojatnost za pojavu raznih defekata i abnormalnosti. Unatoč tome, inbriding se često koristi u uzgoju bilja i životinja. Uz selekciju i križanja, uzgoj u srodstvu je treći najčešći način kojim se povećava produktivnost i profit te se kao takav koristi stoljećima za stvaranje novih pasmina, linija unutar pasmina i brzo ustaljivanje željenih osobina. Nasuprot koristima od uzgoja u srodstvu kojima se brzo ustaljuju poželjna svojstva, postoji i negativna strana. Obzirom da inbriding povećava proporciju homozigota, time povećava i proporciju štetnih recesivnih homozigota. Većina recesivnih alela u populaciji je vrlo rijetka i malo je vjerojatno da su dvije nesrodne jedinke heterozigotni nositelji za određeni lokus. S druge strane, bliski rođaci posjeduju velik dio gena koji potječu od zajedničkog pretka (identični su po porijeklu, engl. Identical by descent – IBD) pa je veća vjerojatnost da su rijetki štetni aleli prisutni kod oba roditelja. U tom slučaju sparivanje jedinki u srodstvu dovodi do smanjenog biološkog "fitnessa" te smanjene viabilnosti u određenoj populaciji, odnosno do inbriding depresije.

2.2 Inbriding depresija

Štetna posljedica uzgoja u srodstvu naziva se inbriding depresija. Inbriding depresija može uzrokovati vrlo vidljive poremećaje poput letalnih fenotipova, vidljivih abnormalnosti i bolesti, no većinom su u pitanju ne tako očita stanja. Pritom se misli na smanjenje prosjeka raznih kvantitativnih svojstava, na usporavanje rasta i razvoja, produktivnosti i općenito viabilnosti jedinke s visokim inbriding koeficijentom (Charlesworth i Willis, 2009.).

Genetska osnova inbriding depresije najčešće je objašnjena pomoću dvije glavne hipoteze. Prva je dominantna hipoteza (Davenport, 1908.; Bruce, 1910.; Keeble i Pellew, 1910.; Jones, 1917.), koja objašnjava da je inbriding depresija uzrokovana ekspresijom štetnih recesivnih alela kod homozigotnih jedinki. Budući da inbriding povećava učestalost homozigota, recesivni štetni aleli koji su skriveni kod heterozigota, će se postupno više isticati. Druga hipoteza je overdominantna hipoteza (Davenport, 1908.; Bruce, 1910.), gdje su heterozigoti superiorniji od oba homozigota te je smanjena frekvencija heterozigota uslijed inbridinga koja dovodi do smanjenja mogućnosti izražavanja overdominantnosti (Charlesworth i Charlesworth, 1987. i 1999.). Kao primjer navode se heterozigotne jedinke koje će imati veći "fitness" u usporedbi s oba homozigota zbog bolje sposobnosti da se prilagode okolišnim uvjetima. Obje hipoteze imaju slične temeljne mehanizme koje uključuju odstupanja od prosjeka homozigota. Naime, dominantnost zahtijeva da heterozigoti budu superiorniji od prosjeka homozigota, dok overdominantnost zahtijeva da heterozigoti budu superiorniji od oba homozigota.

Inbriding depresija se većinom procjenjuje putem regresije nekog svojstva na koeficijent uzgoja u srodstvu dobiven putem rodovnika. Takav način imao je mnoge nedostatke koje nove metode procjene uzgoja u srodstvu nastoje izbjeći (Keller i sur., 2011.). Velik broj metoda javlja se razvojem molekularne genetike i sve većom dostupnošću genomskih podataka. Metoda koja se u posljednje vrijeme uspješno koristi za izračun koeficijenta uzgoja u srodstvu iz genomskih podataka je "Runs of Homozygosity" (ROHom) metoda.

2.3 "Runs of Homozigosity"

Jedinka je homozigotna za određeni marker ukoliko su oba alela na tom markeru identična. U genomu jedinke mogu se naći homozigotne regije u kojima ne opažamo

heterozigotnost, stoga sve genetske varijacije unutar takvih regija posjeduju dva identična alela. Broman i Weber (1999.) su prvi prepoznali duge homozigotne segmente u ljudskom genomu, te ih nazvali „Runs of homozygosity“ (ROH), a njihovu pojavu povezalis autozigotnošću. Postojanje ROH segmenata objasnili su prijenosom istog kromosomskog segmenta s roditelja na potomka koji su ga pak naslijedili od zajedničkog pretka. Rekombinacije i drugi procesi mogu prekinuti kromosomske segmente kroz generacije. Budući da rekombinacije s vremenom prekidaju duge kromosomske segmente, dužina homozigotnih segmenata djelomično ovisi o vremenu od posljednjeg zajedničkog pretka roditelja. Stoga je broj segregacija do zajedničkog pretka niži za duge segmente u usporedbi s kratkim homozigotnim segmentima. Gibson i sur., (2006.) definiraju ROH segmente kao kontinuirane (neprekinute) dijelove DNK nukleotidnog slijeda bez heterozigotnosti u diploidnom obliku. Njihovu pojavu u genomu objašnjavaju također kao posljedicu uzgoja u srodstvu no napominju i druge mogućnosti kao što su „linkage disequilibrium“, uniparentalne disomije (posebne vrste mutacija koje karakterizira prisutnost homolognih kromosoma istog roditeljskog podrijetla u diploidnom kariotipu), heterozigotne delecije (gubitak dijela homolognog kromosoma) i selekcija. Ipak, svi nabrojani mehanizmi nastajanja homozigotnih segmenata (osim disomije), ne mogu proizvesti duge ROH segmente. Njih može proizvesti samo inbriding (Broman i Weber, 1999.; Gibson i sur., 2006.).

2.4 Koeficijent uzgoja u srodstvu

Stupanj uzgoja u srodstvu se mjeri pomoću koeficijenta inbridinga (*F*). Koeficijent inbridinga predstavlja kvantitativnu mjeru inbridinga odnosno vjerojatnost da su 2 homologna gena u jednoj jedinki identični po porijeklu tj. da su autozigotni (Malecot, 1948.).

Koeficijent inbridinga većinom se izračunava preko rodovnika (F_{PED}). Takav pristup ima mnoge nedostatke. On ne uspijeva “uhvatiti” srodnosti između osnivača populacije već ih drži za nesrodne što ne mora biti istina. Također,F_{PED} predstavlja očekivani udio genoma koji je IBD te ne uzima u obzir stohastičku prirodu rekombinacije.

Budući da su ROH izravni proizvod inbridiga oni se mogu koristiti i za procjenu koeficijenta inbridinga. McQuillan i sur., (2008.) prvi definiraju novi genomski koeficijent inbridinga (F_{ROH}) i definiraju ga kao proporciju autosomalnog genoma koji je pokriven ROH segmentima u ukupnom autosomalnom genomu pokrivenom SNP-ovima. F_{ROH} potom istražuju Keller i sur., (2011.) i zaključuju kako je superiorniji od

drugih mjera inbridinga što uključuje i one dobivene iz rodovnika. F_{ROH} je mjera osjetljiva na daleki inbriding te omogućuje izravnu kvantifikaciju autozigotnosti. Uzima u obzir stohastičku prirodu nasljeđivanja na individualnoj razini, može se koristiti ukoliko nisu dostupni podaci iz rodovnika te omogućuje razdvajanje inbridinga na kromosomskoj razini. F_{ROH} pokazuje bolju preciznost u procjeni razine inbridinga, inbriding depresije, a time i bolje praćenje i mogućnost očuvanja populacije.

Osim za kvantificiranje inbridinga ROHom segmenti imaju i druge primjene. Mogu se koristiti za uspješnu procjenu inbriding depresije (Bjelland i sur.,2013.), otkrivanje regija koje uzrokuju depresiju (Keller i sur., 2012; Pryce i sur., 2014; Saura i sur., 2015.), ali i za otkrivanje uzročnih recesivnih mutacija (Drögemülleri sur., 2011.).

2.5 "Heterozygosity Rich Regions"

Suprotno pojavi ROH segmenata, u genomu se mogu naći "Runs of Heterozygosity" segmenti koji se opisuju kao dijelovi DNA nukleotidnog slijeda bez homozigotnosti u diploidnom obliku. Izraz "Runs of Heterozygosity" prvi koriste Williams i sur., (2016.). Ova grupa autora analizirala je heterozigotnost u Chillingham govedu, kao pasmini koja je u prijašnjim istraživanjima preko mikrosatelita pokazala izuzetnu homozigotnost. Koristeći genotipove dobivene putem seta polimorfizama jednog nukleotida, (engl. Single nucleotide polymorphism chip – SNP chip) potvrdili su nedostatak varijabilnosti ove pasmine. Iako su ROH segmenti činili 95% genoma, nađene su i regije koje su strogo heterozigotne. Autori smatraju kako bi te regije mogle sadržavati lokuse koji doprinose "fitnessu" i kao takve su važne za daljnje proučavanje. Ferenčaković i sur. (2016) odbacuju naziv "Runs of Heterozygosity" te uvode novi naziv za visoko heterozigotne regije te ih nazivaju „Heterozygosity rich regions*"* odnosno skraćeno HRR. Također isti autori definiraju HRR-ove kao regije s visokim postotkom heterozigotnosti te su iste zastupljene u malom postotku unutar pasmine. Istraživanje su provodili na pasmini Pinzgauer, gdje su detektirali HRR te kasnije mapirali gene u tim regijama

Odnos između genetske raznolikosti i "fitnessa" predstavlja jedno od važnih pitanja u različitim područjima evolucijske biologije. Važnost razumijevanja odnosa genetske raznolikosti i "fitnessa" je važno za procjenu evolucijskog potencijala populacije te za istovremeno predviđanje smanjenja njihove genetske varijabilnosti.

Korelacija heterozigotnosti i "fitnessa" se koristi u proučavanju njihovog odnosa na individualnoj razini kod različitih organizama (Coltman i Slate, 2003.). Većina studija

koja su se bavila proučavanjem korelacije heterozigotnosti i ”fitnessa” došli su do zaključka da postoji linearna, pozitivna veza između mjera individualne heterozigotnosti i svojstava povezana “fitnessom”.

2.6 “Copy number variations“ - CNV

CNV-ovi se koriste kao genetski markeri ili kandidat geni u genetskom mapiranju (Wang i sur.,2008.). CNV definiramo kao kromosonski segment duljine od najmanje 1kb čije kopije variraju u usporedbi s referentnim genom (Feuk i sur.,2006.). Vjerojatno je da će značajan dio CNV-ova imati funkcionalne posljedice zbog djela aleteracije ili preklapanja gena, poremećaja gena, posljedice samog položaja ili štetnosti samog gena (Beckmann i sur.,2007, Estivill i Armengol,2007.).

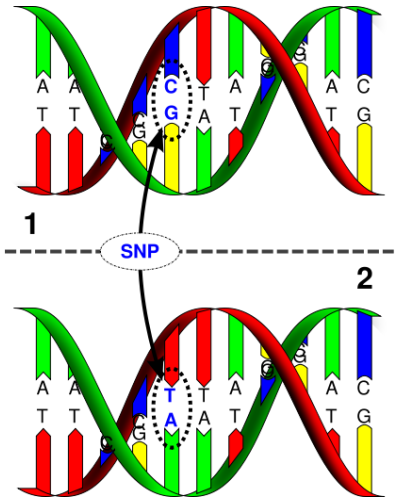
Razvijeno je više tehnologija u svrhu otkrivanja delecija ili preklapanja u genomu sisavaca (Carter,2007.). Većina metoda ovisi o analiziranju uzoraka koje ovise o intezitetu signala koje se pojavljuje kroz genom. CNV-ovi se tradicionalno detektiraju pomoću “Comparative genomic hybridization” (CGH) tehnologije (Sebat i sur.,2004, Iafrate i sur., 2004., Locke i sur, 2006., Redon i sur., 2006.). Druge platforme mogu koristiti oligonukleotidna polja za otkrivanje CNV-ova. Budući da ove tehnologije ne ovise o SNP-ovima, tehnologije mogu pokriti cijeli genom i ustvrditi veću preciznost granice CNV-ova (Wang i sur., 2006.).

Zbog sve većeg interesa u istraživanju cijelog genoma kod sisavaca, počeli su se sve više koristiti SNP-ovi u svrhu otkrivanja i analiziranja CNV-ova. Tehnologije mjere intenzitet signala svim alelima dobivenim od SNP-a i analiziraju signale od svih SNP-ova u genomu da bi dobili podatke o CNV-ovima (Peiffier i sur., 2006., Komura i sur., 2006.). U novije vrijeme, kako bi se povećala pokrivenost SNP markera u otkrivanju CNV-ova, proizvođači SNP markera kao što su Affymetrix i Illumina, uključili su ne polimorfne markere u SNP genotipizaciju, osobito u regijama gdje znamo da postoje CNV-ovi (Wang i sur., 2006.).

2.7 Polimorfizam jednog nukleotida

U selekciji goveda se uzgojni rad do nedavno isključivo temeljio na uporabi metoda iz kvantitativne genetike, gdje je temelj selekcijskog napretka bila isključivo varijanca fenotipskih odnosno proizvodnih obilježja. U zadnjih deset godina došlo je do snažnog razvoja genomike koja omogućuje otkrivanje gena pomoću kojih možemo vidjeti utjecaj tog gena na određena gospodarski značajna svojstva, odnosno korelaciju genotipa i

fenotipa za svojstva od interesa. U tom procesu koristimo razne genetske markere od kojih se sve više koriste polimorfizmi jednog nukleotida (eng. Single Nucleotide Polymorphism - SNP). Ti markeri označavaju promjenu samo jedne nukleotidne baze u DNK molekuli odnosno promjenu jednog od četiri nukleotida (adenin, timin, gvanin i citozin). SNP kao najpolimorfniji biljeg na genomu predstavlja najzanimljiviji pristup u genotipizaciji. Kada se na jednom mjestu u genomu ili unutar para kromosoma jedne jedinke nalaze dva različita nukleotida, tada se zapravo radi o njihovom polimorfizmu. Kao primjer navodimo dva DNK fragmenta iz različitih jedinki, AAGCCTA i AAGCTTA (Slika 1.), koji sadrže razliku u jednom nukleotidu. Vidi se razlika u dva različita alela, u prvom slučaju je citozin, a u drugom timin.



Slika 1. SNP (http://knowgenetics.org/snps/)

Najčešći alat za genotipizaciju putem SNP-ova je "SNP čip" putem kojeg se dobivaju na tisuće ili stotine tisuća markera raspoređenih po genomu. Kod goveda najviše se koriste Illumina BovineSNP50 Genotyping BeadChip s 54 001 SNP-ova i Illumina BovineHD Genotyping BeadChip 777 972 SNP-ova (Slika 2.).



Slika 2. Illumina BovineHD BeadChip

2.8 Karakteristike Pinzgauer goveda

Pinzgauer govedo (Slika 3.) je jedno od posljednjih autohtonih pasmina na području Alpa, koje ima veliku sposobnost prilagođavanja i otpornost na vanjske utjecaje sa čvrstoćom i sposobnošću prilagođavanja. Zahvaljujući istančanom majčinskom instinktu i dobronamjernom temperamentu sve se češće koristi u sustavu krava-tele. Pasmina je kombiniranih svojstava, izvanredne kvalitete mesa, finih vlakana, sočnosti te dobre mramoriranosti mesa koja u laktaciji daje oko 6.000 kg godišnje. Pasmina se drži slobodnim načinom te je 2010. godine broj goveda iznosio oko 49.100. Koeficijent inbridinga za pasminu je u razdoblju od 2010 do 2013 bio F=0.0405, dok je efektivna veličina populacije za baznu populaciju 79 jedinki. (Fegg, 2014.).



Slika 3. Pinzgauer govedo (http://www.hofwirtschaft-ellgass.de/unsere-weiderinder.html)

3 MATERIJAL I METODE

3.1 Kontrola kvalitete genotipova

Ukupno je putem Illumina BovineHD GenotypingBeadChip (777972 SNP-a) genotipiziran 121 bik austrijske Pinzgauer pasmine. DNA je izolirana iz sperme dobivene tijekom redovnog uzimanja ejakulata u stanicama za umjetno osjemenjivanje. Kontrola kvalitete genotipova vršila se uz pomoć programskog paketa SAS 9.4. i putem programskog paketa PLINK v1.07. (Purcell et al., 2007). Isključeni su svi SNP-ovi koji nisu pripisani kromosomima, oni koji su pripisani spolnim kromosomima i mitohondrijskoj DNA. Korištenjem Illumina parametara kontrole ("GenTrainScore" manji ili jednak 0.4 i "GenCallScore" manji ili jednak 0.7) isključeni su SNP-ovi upitne kvalitete. Zadnji korak kontrole kvalitete isključivao je SNP-ove koji su bili prisutni u manje od 10% jedinki, te jedinke kojima je nedostajalo više od 5% SNP-ova.

3.2 Procjena genetskih parametara

Genetski parametri procijenjeni i razmatrani u radu bili su broj polimorfnih SNP-ova, broj monomorfnih SNP-ova, opažena heterozigotnost (HET_{OBS}) koja je izražena kao proporcija heterozigotnih jedinki za pojedini SNP i zatim uprosječena za pojedini kromosom ili cijeli genom te očekivana heterozigotnost (HET_{EXP}) koja je dobivena putem frekvencije alela. Uz navedeno procijenjen je koeficijent inbridinga F_{IS} koji je definiran kao 1-(HET_{OBS}/HET_{EXP}) i jednak je Wright-ovom (Wright, 1949) fiksacijskom indeksu čije su vrijednosti između -1 i +1. Navedeni parametri procijenjeni su putem programskog paketa PLINK v1.07 i programskog paketa SNP & Variation Suite (v8.4.0 Win64; Golden Helix, Bozeman, MT, USA www.goldenhelix.com).

3.3 Detekcija HRR segmenata i regija

HRR segmenti su detektirani uz pomoć programskog paketa SNP & Variation Suite. Kako bi se detektirao niz heterozigotnih SNP-ova prvo je pripremljen set podataka u kojem je zamijenjen status svakog SNP-a na način da su heterozigoti proglašeni homozigotima i obratno. To je učinjeno kako bi algoritam implementiran u SNP & Variation Suite, a koji detektira homozigotne segmente, na taj način detektirao heterozigotne segmente. HRR segmenti su definirani kao niz od najmanje 50 heterozigotnih SNP-ova minimalne duljine od 1000bp. Gustoća SNP-ova u takvom segmentu morala je iznositi barem jedan SNP na 50bp. Također je dozvoljeno da

segment sadrži najviše dva negenotipizirana SNP-a i četiri homozigotna SNP-a (Williams et al., 2016), ali oni nisu smjeli biti jedan za drugim (Ferenčaković et al., 2013). HRR regija definirana je kao mjesto na genomu gdje se HRR segment nalazi u više od 25% jedinki.

3.4 Određivanje CNV-a

Vrijednosti Log R ratio (LRR) parametra izdvojeni su iz Illumina BovineHD finalreport dokumenta dobivenog kao rezultat genotipizacije i potom pripremljeni za analizu uz pomoć programskog paketa SNP & Variation Suite. Korištenjem algoritma "*multivariate copy number analysis method*" (CNAM) algoritma navedenog programskog paketa i korekcijom ("wave detection" wave factor valuet hreshold = 0.05) nad autosomalnim kromosomima putem referentnog goveđeg genoma UMD3.1 ukupno je za analizu preostalo 89 jedinki. Maksimalni broj segmenata dozvoljen na 10 000 SNP-ova bio je 20, a minimalni broj SNP-ova po segmentu je bio tri. Maksimalna vrijednost uparenih segmenata bila je 0.005 (2000 permutacija po paru). CNV segmentima je onda prema radu Zhou i sur. (2016) dodijeljen status "-1" za gubitak "0" za neutralni i "+1" za dobitak segmenta.

3.5 Validacija HRR segmenata

Kako bi se utvrdilo da li je određeni HRR segment zaista regija s visokom heterozigotnošću procijenjeni su određeni parametri. Procijenjena je proporcija heterozigotnih jedinki za svaki marker, proporcija jedinki s CNV-om za svaki marker u odnosu na CNV status (dobitak, gubitak ili oboje), te srednja udaljenost između markera (engl: mean intermarker distances; IMD) čija je gornja granica za detekciju "outlier" vrijednosti postavljena na 9.24kb.

HRR regije koje se poklapaju sa CNV regijama i IMD im je veći od 9.24kb smatraju se prividnim regijama te ne predstavljaju prava područja povećane heterozigotnosti. Analiza i vizualizacija dobivenih parametara učinjena je vlastitom ustupljenom R skriptom autora Yurija Tanija Utsunomiya i Wilsona Nandolo koju sam prilagodio HRR-ovima.

3.6 Detekcija gena u HRR segmentima

Segmenti koji su se pojavljivali u više od 25% jedinki i koji nisu bili prividne već prave regije analizirani su na gene prisutne u njima ili njihovoj blizini (± 1 Mb). Geni su

detektirani programskim paketom SNP & Variation Suite, a njihova funkcija je pretražena putem online baza podataka http://www.uniprot.org i http://www.genecards.org (posljednji pristup 25.09.2017). Genom je mapiran koristeći genetsku mapu UMD 3.1.1. (http://bovinegenome.org/?q=node/61).

4 REZULTATI

4.1 Kontrola kvalitete genotipova i procjena genetskih parametara

Nakon kontrole kvalitete genotipova, set podataka sastojao se od 611102 SNP-ova na autosomalnim kromosomima ukupne duljine od 2507812473 bp i 120 jedinki.

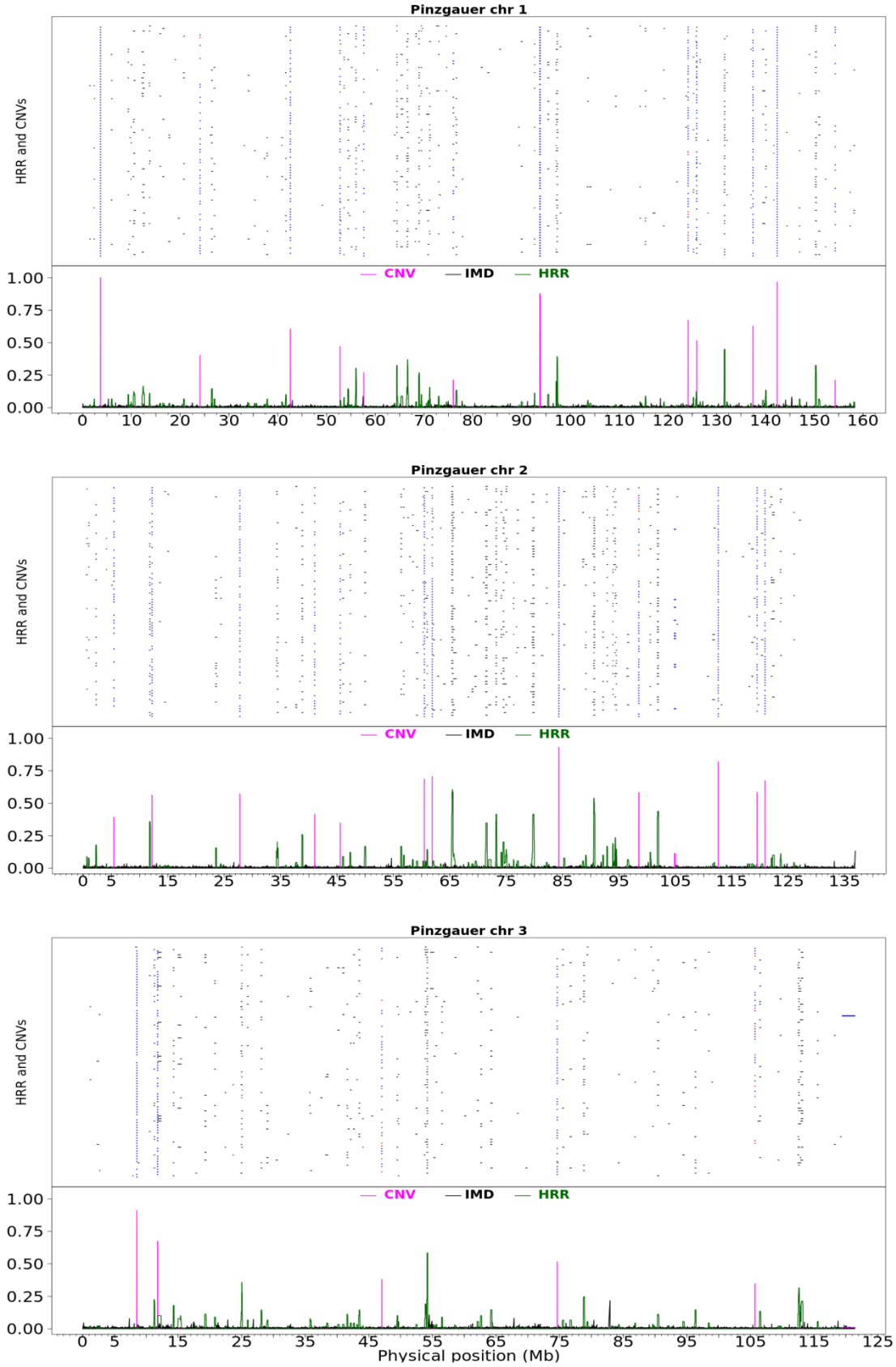
Procijenjeni ukupni genetski parametri za ovu populaciju iznosili su: broj polimorfnih SNP-ova; 603076 (98,7%), broj monomorfnih SNP-ova; 8026 (1,3%), prosječna opažena heterozigotnost (HET_{OBS}); 0,346 (0,320 – 0,363), očekivana heterozigotnost (HET_{EXP}); 0,341 (0,329 - 0,342). Procijenjeni koeficijent inbridinga F_{IS} iznosio je -0,0133 (-0,0638 – 0.0644). Također je pobrojen broj monomorfnih i polimorfnih SNP-ova i procijenjena je HET_{OBS} za svaki kromosom, što je prikazano u Tablici 1. Najveća vrijednost HET_{OBS} bila je prisutna na kromosomu 27 (0,365), a najniža na kromosomu 2 (0,316).

Tablica 1. Genetski parametri populacije Pinzgauer goveda za svaki autosomalni kromosom

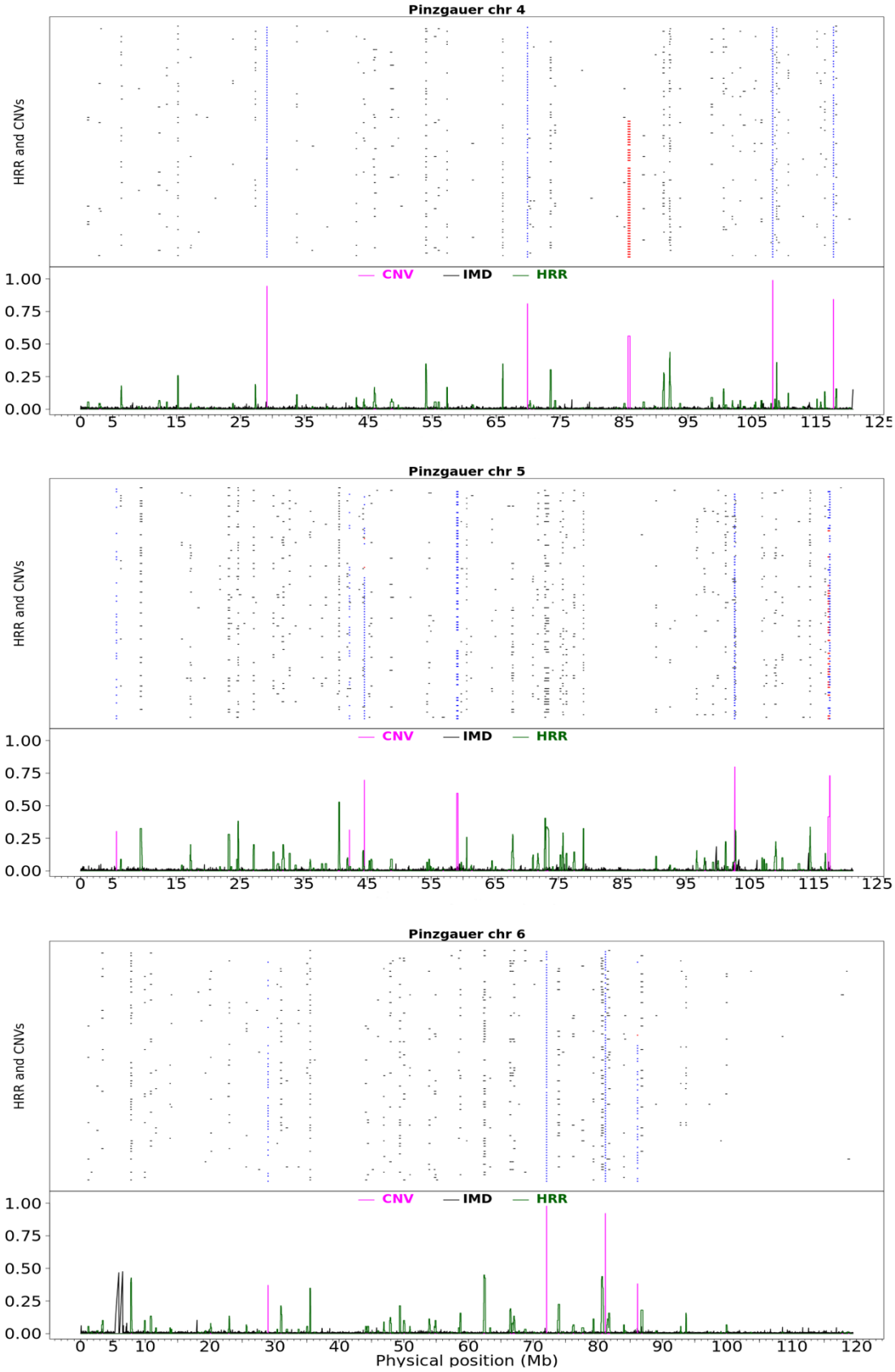
Kromosom	Ukupno SNP-ova	Broj polimorfnih SNP-ova	Broj monomorfnihSNP-ova	Opažena heterozigotnost
1	38842	38321	521	0,340
2	33150	32673	477	0,316
3	29606	29266	340	0,349
4	29315	28947	368	0,347
5	28766	28399	367	0,347
6	30149	29771	378	0,332
7	27150	26812	338	0,349
8	24650	23898	752	0,329
9	25761	25536	225	0,354
10	26063	25815	248	0,352
11	27422	27105	317	0,341
12	21828	21570	258	0,352
13	17728	17310	418	0,338
14	18703	18340	363	0,325
15	20695	20453	242	0,349
16	20061	19779	282	0,345
17	19036	18791	245	0,351
18	16460	16278	182	0,356
19	16066	15873	193	0,359
20	18723	18487	236	0,353
21	17484	17303	181	0,343
22	15568	15458	110	0,354
23	12997	12862	135	0,354
24	15483	15272	211	0,356
25	10976	10878	98	0,349
26	13272	13147	125	0,362
27	11425	11319	106	0,365
28	11211	11099	112	0,358
29	12512	12314	198	0,344

4.2 Detekcija CNV i HRR.

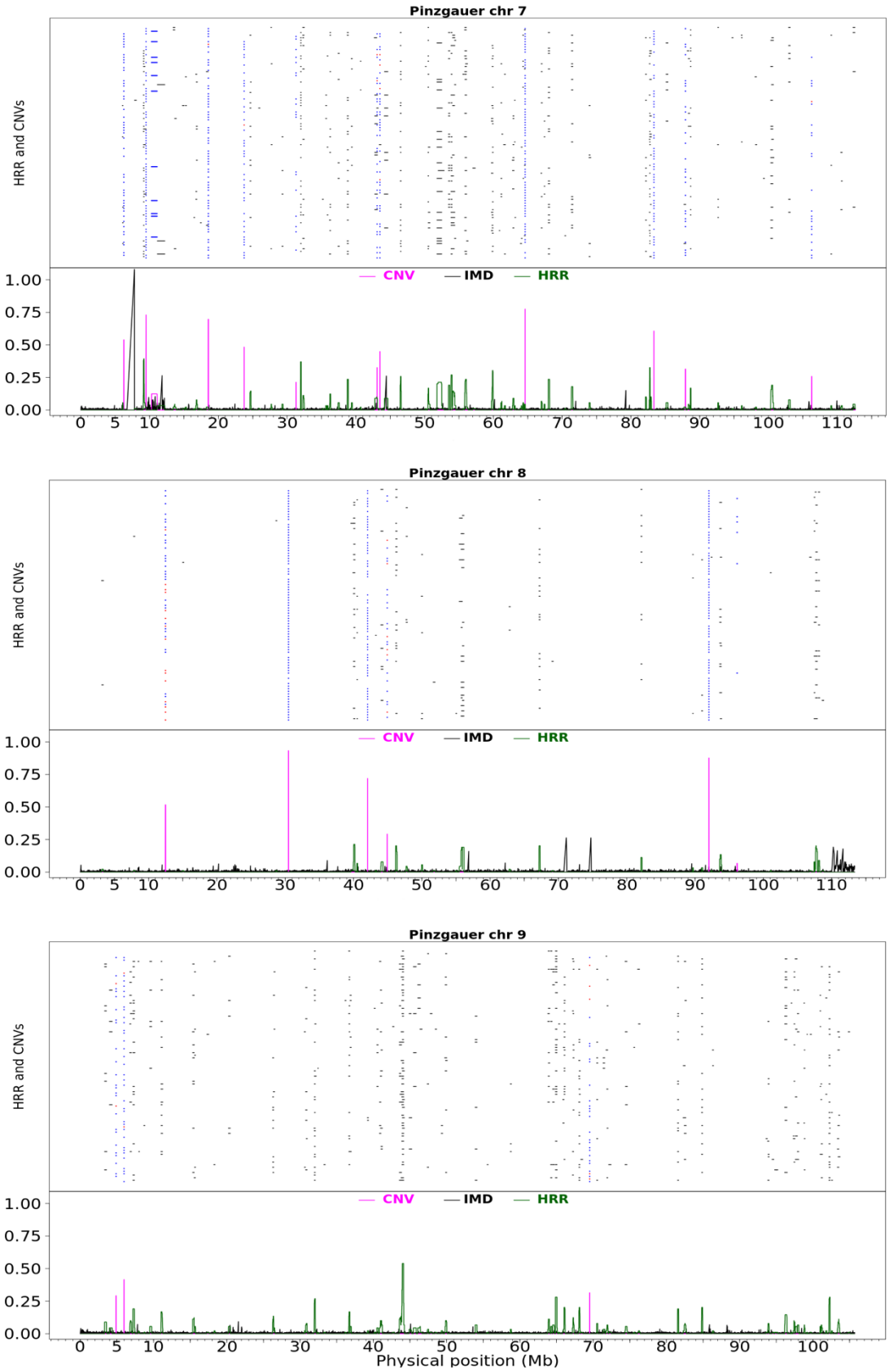
U genomu Pinzgauer goveda detektirano je ukupno 14126 HRR segmenata. Najveći segment, po duljini koju pokriva na genomu, iznosio je 1386964bp i nalazio se na kromosomu 21, a najkraći je iznosio 58072 bp i nalazio se na kromosomu 10. Kod detekcije CNV-ova nađeno je ukupno 184 regije. Najviše regija pronađeno je na kromosomu broj 2 i ukupno ih je 13, dok ih je najmanje pronađeno na kromosomu broj 16 gdje se nalazi samo jedna CNV regija. Najduža CNV regija nalazi se na kromosomu broj 11 i iznosi 4536065 bp, dok je najmanja regija na kromosmu broj 20 koja iznosi 1169 bp. U niže prikazanim slikama vidljive su sve nađene regije po kromosomu (Slika 4).



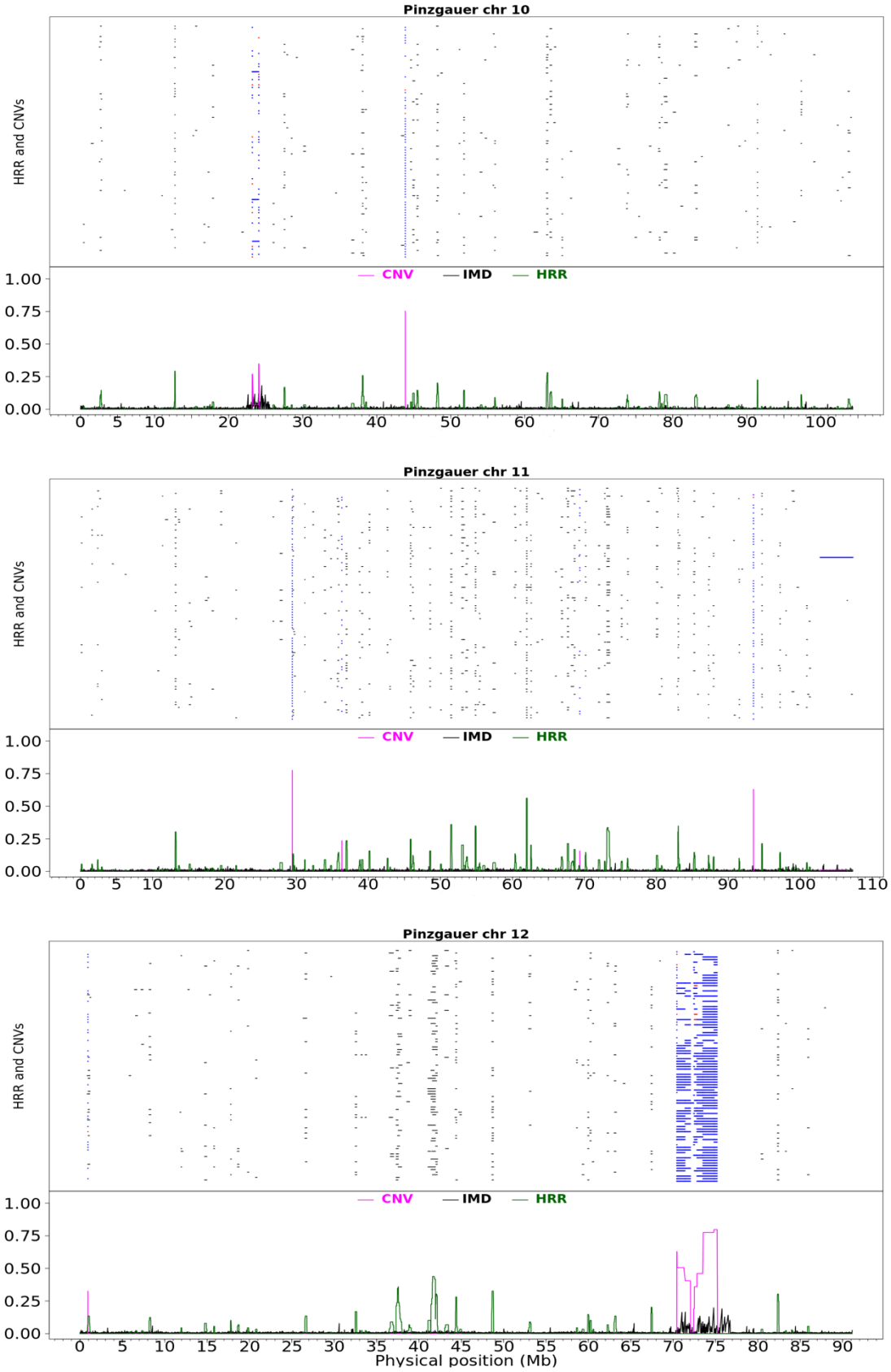
Slika 4. Položaj HRR (zeleno), CNV(magenta) i IMD (crno) na svakom kromosomu.



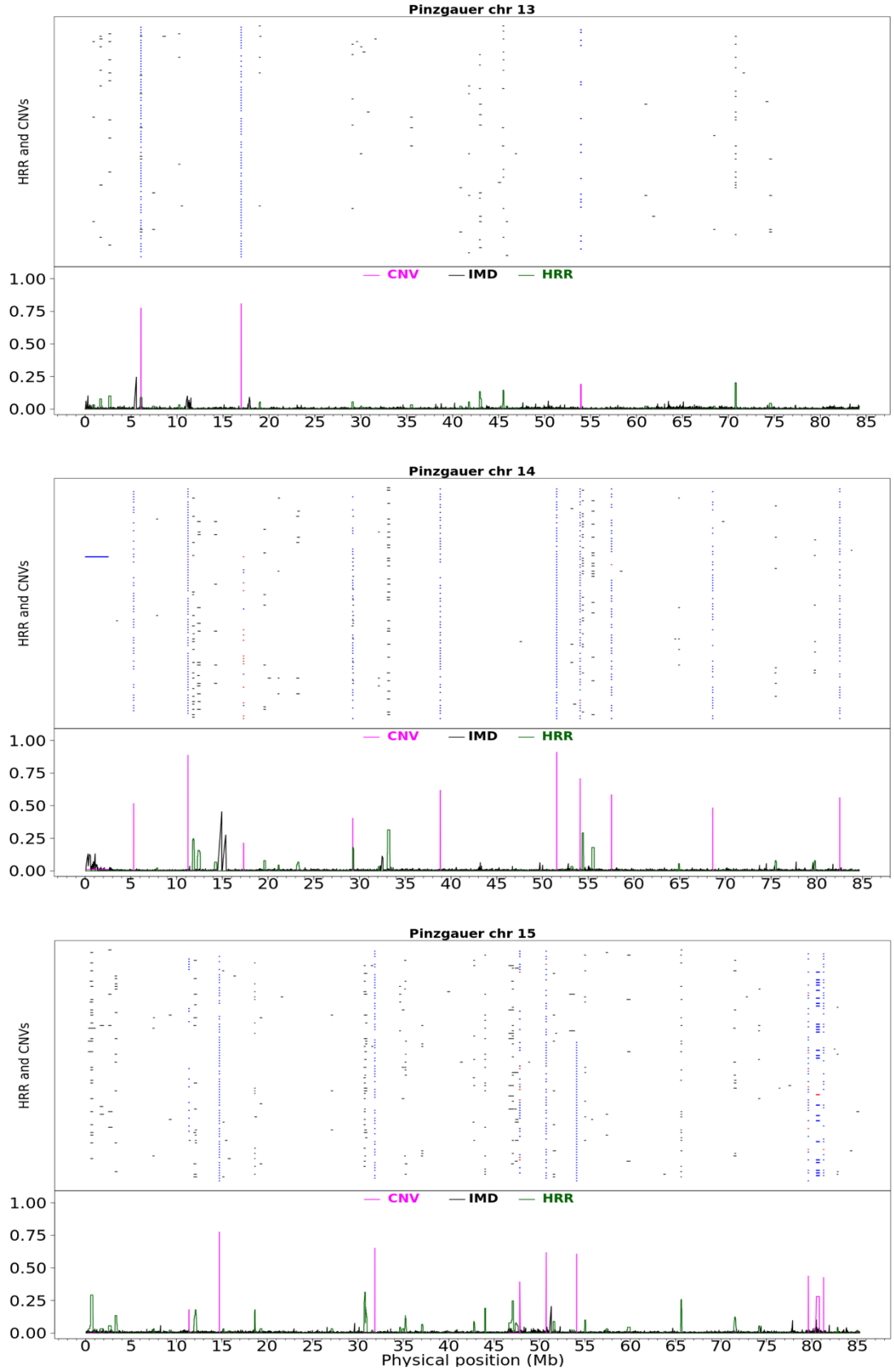
Slika 4. Položaj HRR (zeleno), CNV(magenta) i IMD (crno) na svakom kromosomu - NASTAVAK



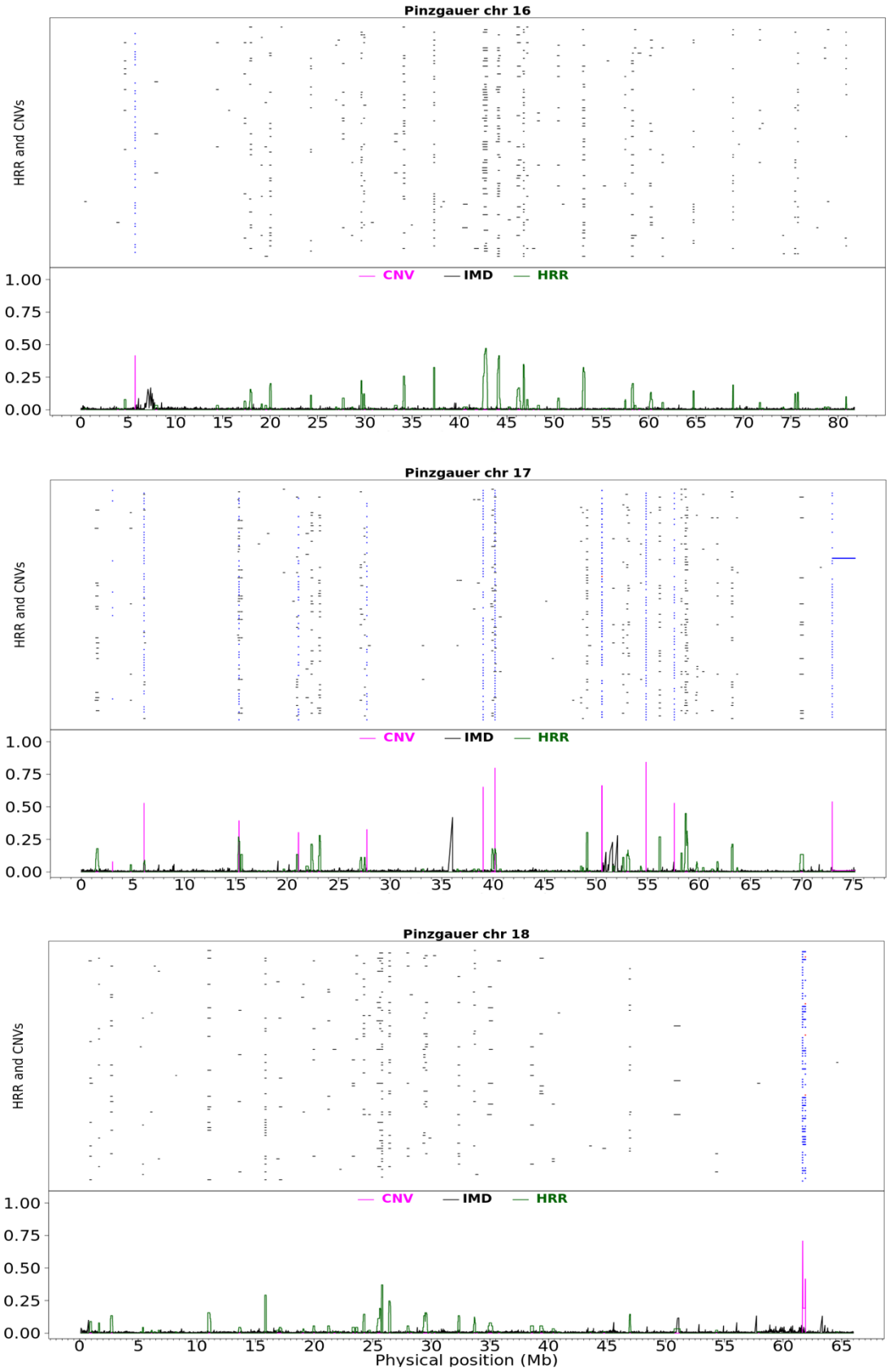
Slika 4. Položaj HRR (zeleno), CNV(magenta) i IMD (crno) na svakom kromosomu-NASTAVAK



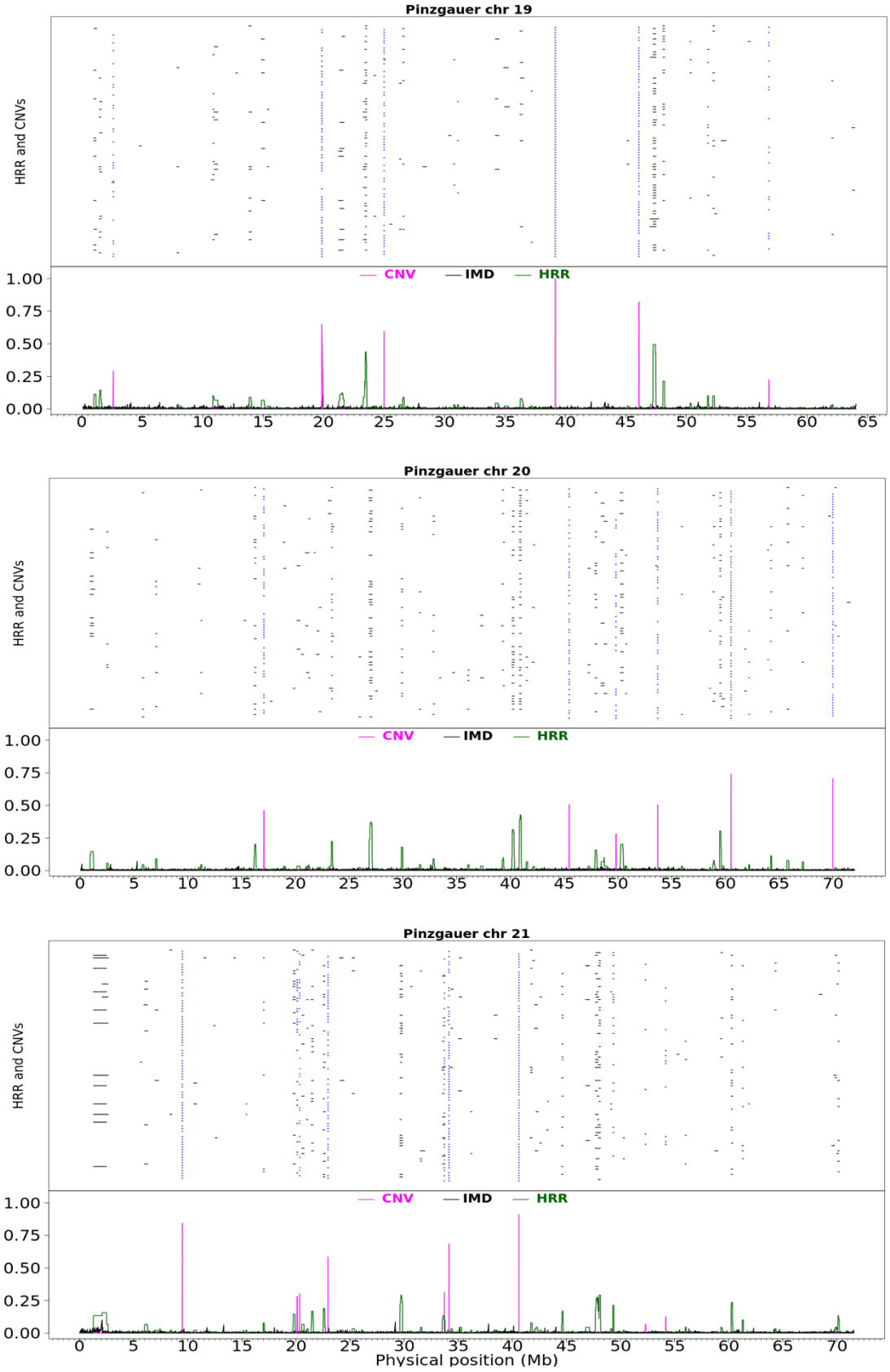
Slika 4. Položaj HRR (zeleno), CNV(magenta) i IMD (crno) na svakom kromosomu - NASTAVAK



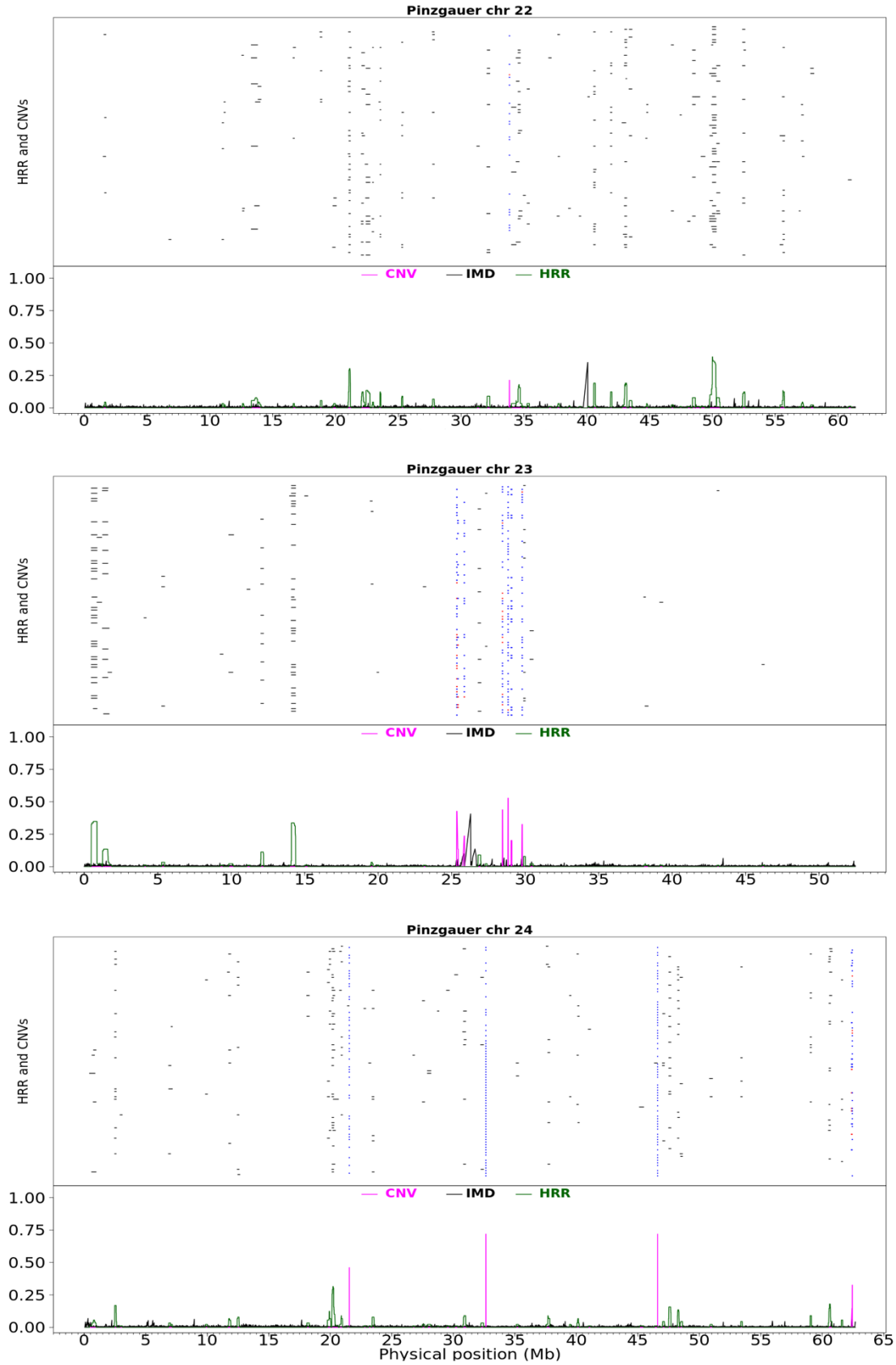
Slika 4. Položaj HRR (zeleno), CNV(magenta) i IMD (crno) na svakom kromosomu - NASTAVAK



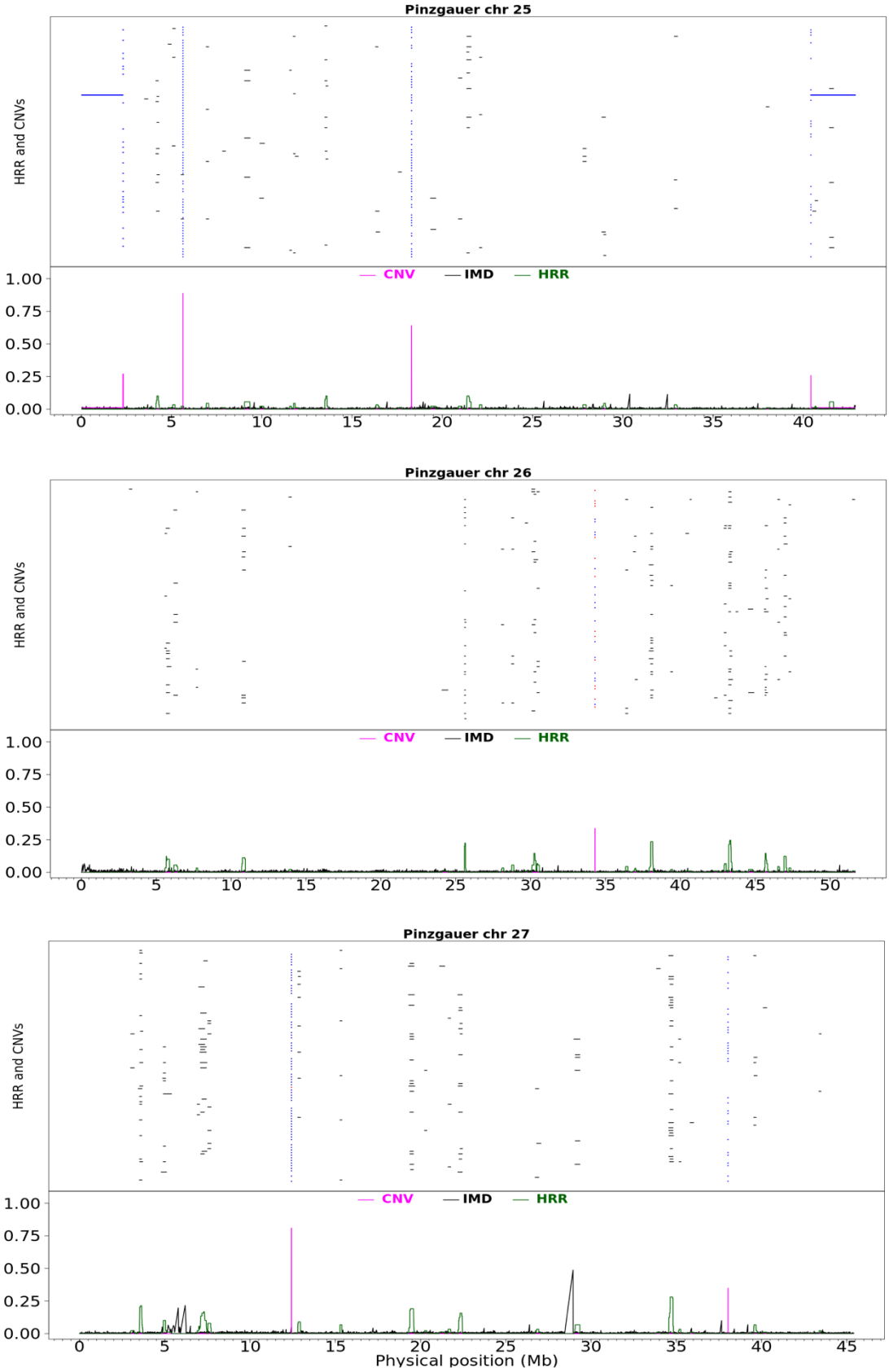
Slika 4. Položaj HRR (zeleno), CNV(magenta) i IMD (crno) na svakom kromosomu - NASTAVAK



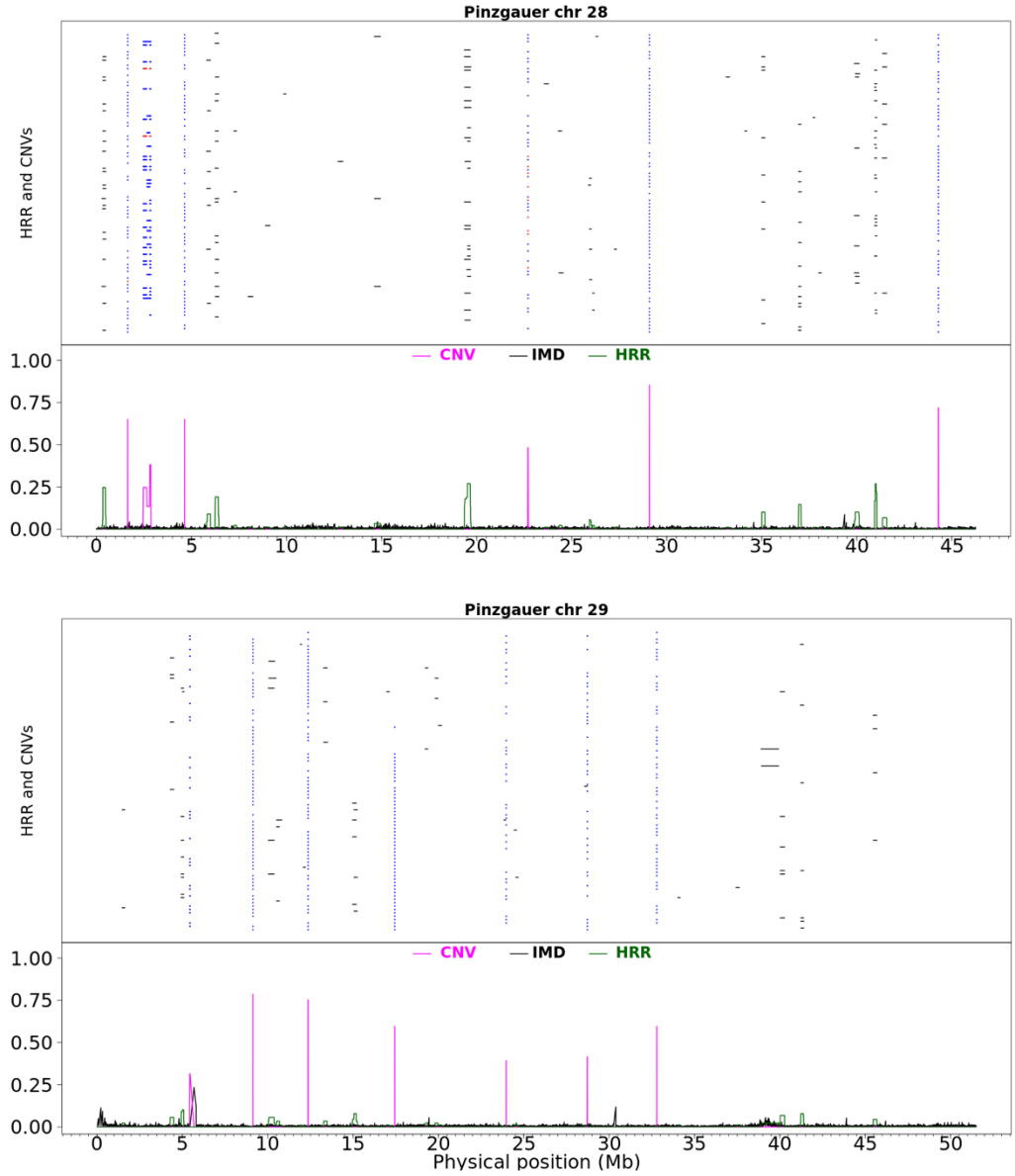
Slika 4. Položaj HRR (zeleno), CNV(magenta) i IMD (crno) na svakom kromosomu - NASTAVAK



Slika 4. Položaj HRR (zeleno), CNV(magenta) i IMD (crno) na svakom kromosomu - NASTAVAK



Slika 4. Položaj HRR (zeleno), CNV(magenta) i IMD (crno) na svakom kromosomu - NASTAVAK



Slika 4. Položaj HRR (zeleno), CNV(magenta) i IMD (crno) na svakom kromosomu-NASTAVAK

4.3 Mapiranje gena

Obzirom da smo odredili kako se HRR-ovi moraju nalaziti u najmanje 25 % jedinki te se moraju odvojiti od prividnih regija, odnosno CNV regije i tamo gdje je IMD veći od 9.24kb, dobili smo 92 HRR-ova na 25 kromosoma. U Tablici 2. prikazane su točne pozicije regija na kromosomima te geni i njihove funkcije preuzete iz online baza podataka. Na kromosomima broj 18, 19, 24, 26 i 27 nalazila se samo jedna regija, na kromosomima broj 14, 22, 23 i 28 nalaze se dvije regije, na kromosomima broj 6, 10, 17 i 20 nalaze se tri regije, na kromosomima broj 3, 9, 12 i 15 nalaze se četiri regije, na kromosu broj 4 nalazi se pet regija, na kromosmima broj 7, 11 i 16 nalaze se šest regija, na kromosomu broj 1 nalazi se šest regija, na kromosomu broj 2 nalazi se osam regija, dok se na kromosomu broj 5 nalazi najveći broj regija, odnosno 11 regija. U prethodno navedene 92 regije ukupno je pronađen 91 gen te 26 regija u kojima nije bio pronađen niti jedan gen. Od 91-og zabilježenog gena, njih 78 ima poznatu i opisanu funkciju, dok se za ostale za sada samo zna da su kodirajući geni.

Tablica 2. HRR prisutne u više od 25% jedinki te geni i njihove funkcije prisutni u njihovoj blizini (±1 Mb).

Chr	Početak regije (bp)	Kraj regije (bp)	Gen	Početak gena (bp)	Kraj gena (bp)	Uloga gena*
1	56000221	56081198	nema gena			
	64387526	64483139	IGSF11	64357122	64488105	Regulira rast
	66581806	66682945	POLQ	66552011	66659954	Jedan od 43 gena koji kontroliraju stanični ciklus kromosomske replikacije
	68909418	69053048	ROPN1	68954525	68989623	Reguliranje proteinske fosforilacije
	97252608	97391703	EIF5A2	97308271	97324962	Vezanje ribosoma
	131546556	131702250	FAIM	131548459	131564836	Djeluje kao indukcijska molekula koja sudjeluje u kontroli Fas rezistencije koju stvara površinsko djelovanje Ig-a na B stanice
			ESYT3	131669531	131722942	Vezivanje lipida
	150258078	150404172	DOPEY2	150169936	150293609	Transport golgijevog aparata
			CHAF1B	150360988	150386586	Vezanje histona
2	11763208	11899039	nema gena			
	38765083	38949731	ERMN	38780333	38785787	Važna uloga u razvoju i strukturi centralnog živčanog sustava
	65318164	65574548	GPR39	65526345	65527741	Regulacija tjelesne težine, gastrointestinalne mobilnosti i hormonske sekrecije
	71382630	71622122	STEAP3	71380830	71424474	Ovaj gen kodira protein dvostruke membrane koji djeluje kao transporter željeza
			DBI	71561192	71566372	Gen kodira protein koji upravlja hormonima
	73164909	73311480	GLI2	72977209	73168370	Igra ulogu tijekom embriogeneze
	79683737	79954328	GLS	79798261	79852401	Utječe na aktivnost glutaminaze
	90526660	90723853	ALS2	90664833	90712586	Gen djeluje kao faktor razmjene nukleinskog gvanina za malu GTPazu
	101783863	102037104	nema gena			
3	25018412	25054981	nema gena			
	54166354	54261102	GBP6	54220013	54261102	Aktivacija GTPaze i vezivanje GTP
	78737414	78905736	WDR78	78736778	78825880	Kodirajući gen, funkcija nepoznata
			INSL5	78846970	78852272	Uloga u aktiviranju hormona i vezivanje receptora za G-protein
	112552923	112695981	CSMD2	112558420	112898742	Vezan uz pojavu shizofrenije i tumorski supresor za rak debelog crijeva

*Funkcije gena preuzete s UniProt (http://www.uniprot.org) i/ili GeneCards (http://www.genecards.org).

Tablica 2. HRR prisutne u više od 25% jedinki te geni i njihove funkcije prisutni u njihovoj blizini (±1 Mb).– NASTAVAK

Chr	Početak regije (bp)	Kraj regije (bp)	Gen	Početak gena (bp)	Kraj gena (bp)	Uloga gena*
4	53954394	54097830	FOXP2	54087329	4164075	Povezivanje DNA
	65972562	66081655	MINDY4	65865704	65984140	Vjerojatno regulira hidrolizu koja uklanja ubiqintin iz protein
			INMT	66002072	66002203	Kodirajući gen, funkcija nepoznata
			CRHR2	66062161	66090181	Regulira aktiviranje receptora povezanog s G-proteinom
	73441682	73595915	ZNF804B	73326980	73897041	Kodira protein
	91185383	92201419	bta-mir-592	92126804	92126899	Kodirajućii gen, funkcija nepoznata
	108829046	108898309	TPK1	108539467	108922091	Katalizira fosforilaciju tiamina u tiamin pirofosfatu
5	9344250	9568720	PPP1R12A	9374685	9511533	Regulator proteinske fosfataze
	23143535	23405113	NUDT4	23348068	23364470	Regulira promet fosfataze
			UBE2N	23370705	23411877	Vezivanje RNA i ATP
	24675037	24762252	nema gena			
	40514553	40638405	MUC12	40558541	40597587	Gen kodira integralni membranski glikoprotein
	60558838	60620821	SNRPF	60556051	60564297	Vezivanje RNA
	67763575	67866565	NT5DC3	67791379	67850986	Kodirajući gen, funkcija nepoznata
	72842100	73477233	nema gena			
	75670953	75712321	nema gena			
	78869758	78949892	DENND5B	78850999	78966082	Gen povezan s aktivnosti kalcijevih kanala
	102723706	102834308	F1MNY8	102726402	102729724	Kodirajući gen, funkcija nepoznata
	114415026	114509640	PACSIN2	114357560	114475163	Aktivnost kinaze
			TTLL1	114494264	114530242	Vezivanje ATP i aktivacija ligaze
6	7741623	7879233	SEC24D	7651891	7761965	Kodira protein
			METTL14	7797201	7837774	Regulator stabilnosti i metilacija mRNA
	62369529	62602853	TMEM33	62464346	62478760	Kodira protein. Strukturni sastojak nuklearnog pora

*Funkcije gena preuzete s UniProt (http://www.uniprot.org) i/ili GeneCards (http://www.genecards.org).

Tablica 2. HRR prisutne u više od 25% jedinki i geni prisutni u njima – NASTAVAK

Chr	Početak regije (bp)	Kraj regije (bp)	Gen	Početak gena (bp)	Kraj gena (bp)	Uloga gena*
7	9108203	9196559	CASP14	9119591	9123207	Gen igra središnju ulogu u fazi izvođenja apoptoze stanica
			CCDC105	9168858	9178475	Kodirajući gen, funkcija nepoznata
	31967906	32046915	PRDM6	31950213	32058001	Uključen u regulaciju vaskularnih glatkih mišićnih stanica (VSMC) kontraktilnih proteina
	53848286	53993012	PCDHB4	53908354	53910744	Vezuje kalcijev ion
	59856761	59941335	PPP2R2B	59782768	60099412	Regulator proteinske fosfataze
	67994100	68179832	CNOT8	68039858	68057466	Vezanje nukleinske kiseline i vezanje RNA
			MRPL22	68106044	68141498	Strukturni dio ribosoma
	82735110	82788615	WWC1	82703348	82782761	Utječe na pamćenje i vezivanje kinaze
9	31963353	32049723	nema gena			
	43914249	44122586	CRYBG1	43886256	43937667	Kodirajući gen, funkcija nepoznata
	64858840	65060198	nema gena			
	102266051	102333567	nema gena			
10	12724329	12779487	MEGF11	12721325	12807326	Kodirajući gen, funkcija nepoznata
	38062137	38158782	F1MUY5	38111923	38135944	Gen utječe na vezivanje kolesterola
			CDAN1	38138863	38151656	Negativno utječe na DNA replikaciju
	62982611	63109546	nema gena			
11	13165746	13280543	ZNF638	13155560	13235444	Povezuje sekvence bogate citidinom u dvolančanu DNA. Povezuje RNA.
	51428588	51581473	nema gena			
	54827608	54950019	CTNNA2	54723190	55906462	Sudjeluje u vezivanju kadherina i aktivan u molekularnoj strukturi
	61932165	62057011	WDPCP	61970467	61994479	Aktivan u razvoju probavnog sustava, uspostavljanju lokalizacije proteina i razvoju dišnog susta
	67628380	67802114	GFPT1	67684390	67728959	Sudjeluje u vezivanju ugljikohidratnog derivata
	73122042	73470176	EPT1	73134172	73173557	Katalizira biosintezu fosfatidiletanolamina iz CDP-etanolamina. Samim time ima važnu ulogu u formiranju i održavanju vezikularnih membrana
			ADGRF3	73199746	73208635	Regulira aktiviranje receptora povezanog s G-proteinom
			RAB10	73335436	73400455	Ativacija GTPaze i vezivanje GTP

*Funkcije gena preuzete s UniProt (http://www.uniprot.org) i/ili GeneCards (http://www.genecards.org).

Tablica 2. HRR prisutne u više od 25% jedinki i geni prisutni u njima – NASTAVAK

Chr	Početak regije (bp)	Kraj regije (bp)	Gen	Početak gena (bp)	Kraj gena (bp)	Uloga gena*
12	37454520	37677261	nema gena			
	41418270	42171539	F1ME41	41907041	41907707	Kodirajući gen, funkcija nepoznata
	44354593	44454167	KLHL1	44295888	44616940	Utječe na ponašanje jedinke
	82300472	82476982	nema gena			
14	33066228	33304950	MCMDC2	33074314	33113841	Utječe na spermatogenezu i na vezivanje DNA i ATP
			CSPP1	33243588	33342863	Pozitivno utječe na citokenzu
	54349231	54477039	CSMD3	53901591	54429251	Kodirajući gen, funkcija nepoznata
15	547375	787417	OR9G1	577327	578564	Regulira aktiviranje receptora povezanog s G-proteinom
	30681395	30815282	PVRL1	30743667	30758246	Vezanje molekule stanične adhezije i aktiviranje receptora
	46991080	47132412	RM17	47004019	47005492	Strukturni sastojak ribosoma
			DCHS1	47042652	47064122	Vezanje kalcijevog iona
			RRP8	47082236	47085798	Vezanje metiliranog histona i vezanje RNA
			DNHD1	47108929	47171016	Kodira protein i sudjeluje u aktivinosti mikrotubula
	65579912	65642976	ABTB2	65548291	65735623	Sudjeluje u aktivnosti heterodimerizacije proteina
16	29610459	29696935	TMEM63A	29624572	29655286	Vezanje nukleinske kiseline
			LEFTY2	29663583	29666431	Regulacija apoptatskog procesa i aktivan faktor rasta
	37272286	37390354	nema gena			
	42507253	42926478	PLOD1	42595152	42625615	Membranski homodimerni protein lokaliziran u cisternama endoplazmatskog redikuluma
			NPPB	42700409	42701796	Regulira veličinu krvnih žila i aktivira hormone
	44002361	44228052	DFFA	43994268	44004665	Negativno utječe na završnu fazu apoptaze
			KIF1B	44080103	44212899	Regulira transport mitohondrija
	46737873	46834399	nema gena			
	53038694	53224856	TMEM82	53065303	53068821	Kodirajući gen, funkcija nepoznata
			DDI2	53131205	53133055	Vezivanje proteina

*Funkcije gena preuzete s UniProt (http://www.uniprot.org) i/ili GeneCards (http://www.genecards.org).

Tablica 2. HRR prisutne u više od 25% jedinki i geni prisutni u njima – NASTAVAK

Chr	Početak regije (bp)	Kraj regije (bp)	Gen	Početak gena (bp)	Kraj gena (bp)	Uloga gena*
17	23111933	23238541	nema gena			
	49059196	49199342	GLT1D1	49197829	49313590	Kodirajući gen, funkcija nepoznata
	58651296	58878879	nema gena			
18	25743634	26469905	CNGB1	25894772	25945638	Aktivnost kationskog kanala
			TEPP	25970368	25977551	Kodirajući gen, funkcija nepoznata
19	47263074	47485185	EFCAB3	47474735	47507698	Vezivanje kalcijevog iona
20	26942991	27139796	nema gena			
	40180680	41015855	ADAMTS12	39913085	40304028	Vezanje cink iona
			TARS	40349625	40373064	Vezanje ATP
	59469483	59579936	nema gena			
21	29655208	29798591	nema gena			
	47747607	48148919	PIWIL1	47726941	47761637	Vezanje mRNA i proteinske kinaze
			FZD10	47935477	47937432	Aktivnost receptora za G-protein
22	21072799	21166727	nema gena			
	49997321	50270025	DOCK3	49953523	50134786	Važna uloga u središnjem živčanom sustavu
23	479600	868795	KHDRBS2	184673	864355	Uključen u vezanje nukleinske kiseline i aktivnost heterodimerizacije proteina
	14099529	14380782	LRFN2	14309128	14351087	Kodirajući gen, funkcija nepoznata
24	20179970	20272221	nema gena			
26	38024570	38151002	nema gena			
27	34643544	34808620	IDO1	34686577	34699480	Gen igra važnu ulogu kod odgovora organizma na stres
28	19515397	19668175	JMJD1C	19393375	19663523	Gen je u interakciji s receptorima horomona štitnjače i sudjeluje u vezanju kromatinske DNA

*Funkcije gena preuzete s UniProt (http://www.uniprot.org) i/ili GeneCards (http://www.genecards.org).

5 RASPRAVA

Pinzgauer je alpska pasmina dvostruke namjene, za proizvodnju mlijeka i mesa koja je do Drugog svjetskog rata imala važnu ulogu u zemljama Austrougarske monarhije. Nakon Drugog svjetskog rata pasmina gubi važnost i smanjuje im se broj. Fegg (2014.) u svom diplomskom radu putem rodovnika procjenjuje F_{PED} od 4% i efektivnu veličinu populacije od 79 jedinki. Ferenčaković i sur. (2013.) na setu podataka koji se za samo 2 životinje razlikuje od seta koji smo analizirali u ovom istraživanju procjenjuje F_{PED} od 1,9%.

Genetski parametri procijenjeni u ovom radu pokazuju visoku polimorfnost SNP markera od 98,7% dobivenih putem Illumina BovineHD Genotyping BeadChip. Iako za stvaranje ovog SNP čipa nije korištena ova pasmina (http://www.illumina.com/Documents/products/datasheets/datasheet_bovineHD.pdf), ovako visoka polimorfnost markera odbacuje mogućnost pristranosti ovog alata prema pasminama korištenim u njegovom stvaranju. Najveća vrijednost HET_{OBS} bila je prisutna na kromosomu 27 (0,365), a najniža na kromosomu 2 (0,316) (Tablica 1). Williams i sur. (2014.) u sličnom istraživanju, ali na populaciji koja je izuzetno homozigotna (Chillingham) najveću vrijednost HET_{OBS} bilježe na kromosomu 12, a najnižu na kromosomu 15. Ni kromosomska ni ukupna HET_{EXP}(0,342) kod Pinzgauer goveda ne odstupa od istraživanja na drugim pasminama gdje se nalazila između 0.25–0.34 (prema Williams i sur., 2014.). Procijenjeni koeficijent inbridinga F_{IS} imao je negativan predznak (-0,0133) što upućuje na izostanak uzgoja u srodstvu. Negativan predznak mogao bi značiti i probleme u genotipizaciji, no zbog izvršene kontrole kvalitete i visoke polimorfnosti to ovdje nije slučaj. Ferenčaković i sur. (2013.) koristeći F_{ROH} nalaze niske vrijednosti inbridiga (1,4 – 6,2%) naspram drugih pasmina.

Analiza HRR segmenata rezultirala je s 14126 segmenta. Najveći segment, po duljini koju pokriva na genomu, iznosio je 1016036 bp i nalazio se na kromosomu broj 4, a najkraći je iznosio 36569 bp i nalazio se na kromosomu broj 3. Daljnjom analizom CNV regija utvrđeno je ukupno 184 CNV regije na genomu. Najveći segment bio je na kromosomu broj 11 i iznosio je 4536065 bp dok je najkraći segment bio na 20. kromosomu i iznosio je 1169 bp. Ferenčaković i sur.,(2016.) su u sličnom istraživanju dobili identične rezultate što se tiče HRR-ova kod Pinzgauer goveda, dok CNV regije nisu istraživali, tako da nažalost ne možemo usporediti dobivene rezultate.

Iako su HRR regije dosta male u njima najčešće nalazimo gene. Pronalaskom prividnih regija, odnosno CNV regija i IMD veći od 9.24kb, uspjelo se navedene regije odvojiti od pravih HRR-ova pri čemu je detektirano ukupno 92 regije na 25 kromosoma. Od ukupno detektirane

92 regije, njih 26 nisu sadržavale gene. U regijama je ukupno zabilježen 91 gen. Poznatu i opisanu funkciju ima 78 gena (Tablica 2.). Proučavajući funkcije tih gena na http://www.uniprot.org i http://www.genecards.org (posljednji pristup 25.09.2017.) zaključuje se kako pronađeni geni imaju važne funkcije u biološkim procesima. Pronađeno je nekoliko regija koje su izrazito bitne za sami metabolizam jedinke, a koje uključuju gene koji su izravno povezani s bitni procesima u jedinki kao što su aktivacija GTPaze, vezanje ATP, aktivacija receptora za G protein i kromosomsku replikaciju. Na kromosomu broj 3 imamo regiju u kojoj je mapiran gen *GBP6* koji je izravno uključen u aktivaciju GTPaze i vezivanje GTP, te gen *INSL5* koji je uključen u vezivanje receptora za G protein. Kromosom broj 14 sadrži gen *MCMDC2* koji je uključen u spermatogenezu. Nađeno je još niz regija i gena koji utječu na bitne procese unutar organizma, a kao primjer navodim gen *ERM* na kromosomu broj 2 koji kod ljudi ima važnu ulogu u razvoju i strukturi centralnog živčanog sustava. Također navodim gen *ROPN1* na kromosomu broj 1 koji regulira proteinske fosforilacije. Prisutnost gena s velikim utjecajem na važne biološke procese u regijama koje pokazuju visoku heterozigotnost upućuju na mehanizam selekcije u tim regijama (Van Raden i sur., 2011.).

Ovaj rad se nadovezuje na istraživanje Ferenčaković i sur., (2016.) te preciznije nastoji detektirati HRR-ove. Spomenuto istraživanje detektiralo je 11 HRR-ova u 9 kromosoma, dok je ovo istraživanje s preciznijm parametrima i izbacivanjem prividnih regija detektiralo ukupno 92 HRR-ova u 25 kromosoma. Ferenčaković i sur., (2016.) detektirali su ukupno 21 gen što je u usporedbi s našim istraživanjem znatno manje, obzirom da je u ovom istraživanju detektiran ukupno 91 gen u HRR-ovima. Pretpostavka je da geni u ovim regijama povoljno utječu na "fitness" svojstva. Potrebna je daljnja analiza HRR-ova i optimizacija njihove detekcije budući da postoji potencijal HRR pristupa kao jednostavnog alata u detekciji gena koji su odgovorni za "fitness" svojstva, kako kod komercijalnih pasmina tako i kod malih lokalnih pasmina ili divljih vrsta. Svako uspješno detektiranje posebnosti u takvim populacijama može doprinijeti njihovom genetskom očuvanju i boljem razumijevanju genetske arhitekture ugroženih divljih vrsta i kod lokalnih pasmina koje nisu toliko izložene selekciji. Ovo istraživanje, za razliku od prijašnjih istraživanja, precizinije detektira same HRR-ove i zasigurno će bit daljna polazna točka za buduća istraživanja u ovom području.

6 ZAKLJUČCI

1. U genomu Pinzgauer goveda ukupno je nađeno 14126 HRR-ova koje nisu ravnomjerno raspoređene po genomu, već postoje mjesta gdje su učestalije
2. Ukupno je pronađena 91 regija s učestalim HRR segmentima (prisutnim u više od 25% jedinki) na 25 kromosoma
3. U tih 92 regija ukupno je pronađen 91 gen
4. Od 91-og zabilježenog gena, njih 78 ima poznatu i opisanu funkciju
5. Funkcije zabilježenih gena su važne za biološke procese i pojave bolesti te imaju utjecaj na "fitness" jedinke
6. Potrebna je daljnja analiza HRR regija budući da postoji potencijal HRR pristupa kao jednostavnog alata u detekciji gena koji su odgovorni za "fitness" svojstva

7 POPIS LITERATURE

Alvarez G., Ceballos F.C. & Quinteiro C. (2009). The role of inbreeding in the extinction of a European royal dynasty. *PLoS ONE* **4**, e5174.

Arcos-Burgos, M. & Muenke, M. (2002). Genetics of population isolates. *Clin Genet***61,** 233-247.

Beckmann J.S., Estivill X. &Antonarakis S.E. (2007.). Copy number variants and genetic traits: closer to the resolution of phenotypic to genotypic variability.*Nat. Rev. Genet*.**8**,639–646.

Bjelland D., Weigel K., Vukasinovic N. & Nkrumah J. (2013). Evaluation of inbreeding depression in Holstein cattle using whole-genome SNP markers and alternative measures of genomic inbreeding. *J Dairy Sci* **96**, 4697-706.

Broman K.W. & Weber J. L. (1999). Long Homozygous Chromosomal Segments in Reference Families from the Centre d'Étude du Polymorphisme Humain. *Am J Hum Genet* **65**, 1493-500.

Bruce A. B. (1910). The Mendelian theory of heredity and the augmentation of vigor. *Science* **32**, 627-628.

Carothers A. D., Rudan I., Kolcic I., Polasek O., Hayward C., Wright A. F., et al. (2006). Estimating human inbreeding coefficients: comparison of genealogical and marker heterozygosity approaches. *Ann Hum Genet**70,*** 666-676.

Carter N.(2007.) Methods and strategies for analyzing copy number variation using DNA microarrays.*Nat. Genet***39**,S16–S21.

Charlesworth D. & Charlesworth B. (1987). Inbreeding Depression and its Evolutionary Consequences. *Annu Rev Ecol Syst* **18**, 237-268.

Charlesworth B. & Charlesworth D. (1999). The genetic basis of inbreeding depression. Charlesworth D. & Willis J. (2009). The genetics of inbreeding depression. *Nat Rev Genet* **10**, 783-796.

Coltman D. & Slate J. (2003). Microsatellite measures of inbreeding: a meta-analysis. *Evolution***57**, 971-983.

Crow J.F., (1954). Breeding structure of populations. II. Effective population number. In: Kempthorne O., Bancroft T. A., Gowen J. W., Lush J. L. (Eds.), Statistics and Mathematics in Biology, Iowa State University Press, Ames, USA, pp. 543–556.

Curik I., Ferenčaković M. & Sölkner J. (2014). Inbreeding and runs of homozygosity: A possible solution to an old problem. *Livest Sci* **166**, 26-34.

Davenport C. B. (1908). Degeneration, albinism, and inbreeding. *Science* **28**, 454-455.

Drögemüller C., Reichart U., Seuberlich T., Oevermann A., Baumgartner M., KühniBoghenbor K., et al. (2011). An Unusual Splice Defect in the Mitofusin 2 Gene (MFN2) Is Associated with Degenerative Axonopathy in Tyrolean Grey Cattle. *PLoS ONE* **6**, e18931.

Estivill X. & Armengol L. (2007.). Copy number variants and common disorders: filling the gaps and exploring complexity in genome-wide association studies.*PLoS Genet*.**3**,1787–1799.

Fegg T. (2014). Modellierung alternativer Zuchtprogramme zur Optimierung des langfristigen Zuchtfortschrittes bei der Rasse Pinzgauer.Master thesis, Universität für Bodenkultur, Wien.

Ferenčaković M.,Banadinović M.,Mercvajler.,Khayat-Zadeh N.,Meszaros G.,Cubrik-Curik V.,Curik I. & Solkner J.,(2016.). Mapping heterozygosity rich regions in Austrina Pinzgauer cattler.*Acta agriculturae Slovenica*,**5**, 41-4.

Ferenčaković M., Hamzić E., Gredler B., Solberg T. R., Klemetsdal G., Curik I. & Sölkner J. (2013). Estimates of autozygosity derived from runs of homozygosity: empirical evidence from selected cattle populations. *J Anim Breed Genet***130**, 286-93.

Fernandez S. R. Zhang Y.& Parsons C. M. (1995). Dietary formulation with cottonseed meal on a total amino acid versus a digestible amino acid basis. *Poultry Sci*, **74,** 1168-1179.

Feuk L., Carson A.R. & Scherer S:W (2006.) Structural variation in the human genome. *Nat. Rev. Genet*, **7**, 85–97.

Gibson J., Morton N. & Collins A. (2006). Extended tracts of homozygosity in outbred human populations. *Hum Mol Genet* **15**, 789-95.

Howrigan D.P., Simonson M.A. &Keller M.C.,(2011.). Detecting autozygosity through runs of homozygosity: A comparison of three autozygosity detection algorithms.*Genomics BioMed Central*,**12**, 460.

Iafrate A.J., Feuk L., Rivera M.N., Listewnik M.L., Donahoe P.K., Qi Y., Scherer S.W. &Lee C.(2004.).Detection of large-scale variation in the human genome.*Nat. Genet*.**36**,949–951.

Jones D. F. (1917). Dominance of linked factors as a means of accounting for heterosis. *Genetics* **2**, 466-479.

Keeble F. & Pellew C. (1910). The mode of inheritance of stature and of time of flowering in Peas Pisum sativum. *J Genet* **1**, 869-876.

Keller M. C., Visscher P. M. & Goddard M. E. (2011). Quantification of inbreeding due to distant ancestors and its detection using dense single nucleotide polymorphism data.*Genetics***189**, 237-249.

Keller M.C., Simonson M.A., Ripke S., Neale B.M., Gejman P.V., Howrigan D.P., Lee S.H., Lencz T., Levinson D.F., Sullivan P.F. & The Schizophrenia Psychiatric Genome-Wide Association Study, C. (2012). Runs of homozygosity implicate autozygosity as a schizophrenia risk factor. *PLoS Genet* **8**, e1002656.

Komura D., Shen F., Ishikawa S., Fitch K.R., Chen W., Zhang J., Liu G., Ihara S., Nakamura H., Hurles M.E. & et al.(2006.).Genome-wide detection of human copy number variations using high-density DNA oligonucleotide arrays.*Genome Res*.**16**,1575–1584.

Kristensen T. & Sorensen A. (2005). Inbreeding - lessons from animal breeding, evolutionary biology and conservation genetics. *Anim Sci* **80**, 121-133.

Lerner I.M. (1954). Genetic Homeostasis. Oliver and Boyd, London, UK.

Locke D.P., Sharp A.J., McCarroll S.A., McGrath S.D., Newman T.L., Cheng Z., Schwartz S., Albertson D.G., Pinkel D., Altshuler D.M. &et al. (2006).Linkage disequilibrium and heritability of copy-number polymorphisms within duplicated regions of the human genome.*Am. J. Hum. Genet*.**79,**275–290.

Malécot G. (1948). Les mathématiques de l'hérédité. Masson, Paris, France.

McParland L. C., Collinson M. E., Scott A. C.& Campbell G. (2009). The use of reflectance values for the interpretation of natural and anthropogenic charcoal assemblages. *Archaeol Anthropol Sci***1,** 249-261.

McQuillan R., Leutenegger A., Abdel-Rahman R., Franklin C., Pericic M., Barac-Lauc L., Smolej-Narancic N., Janicijevic B., Polasek O., Tenesa A., Macleod A., Farrington S., Rudan P., Hayward C., Vitart V., Rudan I., Wild S., Dunlop M., Wright A., Campbell H. & Wilson J. (2008).Runs of homozygosity in European populations. *Am J Hum Genet* **83**, 359-372.

Peiffer D.A., Le J.M., Steemers F.J., Chang W., Jenniges T., Garcia F., Haden K., Li J., Shaw C.A., Belmont J. &et al.(2006.).High-resolution genomic profiling of chromosomal aberrations using Infinium whole-genome genotyping.*Genome Res*.**16**,1136–1148.

Pirchner F. (1985). Genetic structure of population. 1. Closed populations or matings among related individuals. In: Chapman, A.B. (Ed.), General and Quantitative Genetics, Elsevier, New York, USA, pp. 227–250.

Pryce J. E., Haile-Mariam M., Goddard M. E. & Hayes B. J. (2014). Identification of genomic regions associated with inbreeding depression in Holstein and Jersey dairy cattle. *Genet Sel Evol* **46**, 71.

Redon R., Ishikawa S., Fitch K.R, Feuk L., Perry G.H., Andrews T.D., Fiegler H., Shapero M.H., Carson A.R., Chen W. & et al.(2006.).Global variation in copy number in the human genome.Nature.**444**,444–454.

Saura M., Fernández A., Varona L., Fernández A. I., de Cara M. Á. R., Barragán C. & Villanueva B. (2015). Detecting inbreeding depression for reproductive traits in Iberian pigs using genome-wide data. *Genet Sel Evol* **47**, 1-9.

Sebat J., Lakshmi B., Troge J., Alexander J., Young J., Lundin P., Maner S., Massa H., Walker M., Chi M., &et al.(2004.).Large-scale copy number polymorphism in the human genome.*Science*.**305**,525–528.

Van Raden P. M., Olson K. M., Null D. J. & Hutchison J. L. (2011). Harmful recessive effects on fertility detected by absence of homozygous haplotypes. *J Dairy Sci***94**, 6153–6161.

Williams J. L., Hall S. J. G., Del Corvo M., Ballingall K. T., Colli L., Ajmone Marsan P., Biscarini F. (2016). Inbreeding and purging at the genomic Level: the Chillingham cattle reveal extensive, non-random SNP heterozygosity. *Anim Genet* **47**, 19-27.

8 ŽIVOTOPIS

MARIO MERCVAJLER

Rođen sam 21.05.1992. godine u Wangen im Allgäu, Savezna Republika Njemačka. Pohađao sam O.Š. dr. Franjo Tuđman u Šarengradu. Potom upisujem Opću gimnaziju u Iloku, te maturiram 2011. godine. Nakon Opće gimnazije upisujem Agronomski fakultet Sveučilišta u Zagrebu, preddiplomski smjer "Agrarna ekonomika". Zatim upisujem diplomski smjer "Genetika i oplemenjivanje životinja". Aktivno se bavim nogometom. Honorarno već 3 godine radim za tvrtku "Synergy sports technology", te sam aktivan u nekoliko udruga. Prisustvovao sam konferenciji "Snaga hrvatske hrane" u Lisinskom 2015 godine. Sada sam 2. godina diplomskog studija, te pripremam diplomski rad na "Zavodu za opće stočarstvo" pod mentorstvom doc.dr.sc Maje Ferenčaković.