Identifikacija signala pozitivne selekcije u genomu hrvatskih ovaca

Shihabi, Mario

Master's thesis / Diplomski rad

2021

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:204:179823

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
University of Zagreb
Faculty of Agriculture

IDENTIFIKACIJA SIGNALA POZITIVNE SELEKCIJE U GENOMU HRVATSKIH OVACA

DIPLOMSKI RAD

Mario Shihabi

Zagreb, srpanj 2021.

Sveučilište u Zagrebu
Agronomski fakultet
University of Zagreb
Faculty of Agriculture

Diplomski studij:

Genetika i oplemenjivanje životinja

IDENTIFIKACIJA SIGNALA POZITIVNE SELEKCIJE U GENOMU HRVATSKIH OVACA

DIPLOMSKI RAD

Mario Shihabi

Mentor:

Prof. dr. sc. Ino Čurik

Zagreb, srpanj 2021.

Sveučilište u Zagrebu
Agronomski fakultet

University of Zagreb
Faculty of Agriculture

IZJAVA STUDENTA

O AKADEMSKOJ ČESTITOSTI

Ja, **Mario Shihabi**,JMBAG 0178107784, rođen 06.03.1997. u Zagrebu, izjavljujem da sam samostalno izradio diplomski rad pod naslovom:

IDENTIFIKACIJA SIGNALA POZITIVNE SELEKCIJE U GENOMU HRVATSKIH OVACA

Svojim potpisom jamčim:

- da sam 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 upoznat s odredbama Etičkog kodeksa Sveučilišta u Zagrebu (Čl. 19).

U Zagrebu, dana _____ _____

Potpis studenta

Sveučilište u Zagrebu
Agronomski fakultet

University of Zagreb
Faculty of Agriculture

IZVJEŠĆE

O OCJENI I OBRANI DIPLOMSKOG RADA

Diplomski rad studenta **Mario Shihabi**, JMBAG 0178107784,naslova

IDENTIFIKACIJA SIGNALA POZITIVNE SELEKCIJE U GENOMU HRVATSKIH OVACA

obranjen je i ocijenjen ocjenom _____ , dana _____ .

Povjerenstvo:			potpisi:
1.	prof. dr. sc. Ino Čurik	mentor	_____
2.	izv. prof. dr. sc. Vlatka Čubrić Čurik	član	_____
3.	doc. dr. sc. Maja Ferenčaković	član	_____

Zahvala

Zahvaljujem se:

Prof. dr. sc. Inu Čuriku, prvenstveno na pristupačnosti u bilo koje doba dana i noći te nesebičnom širenju znanja. Hvala što ste me usmjerili na genetiku tijekom studiranja/života. Hvala na vremenu, stručnom vodstvu, povjerenju, strpljenju, savjetima i sugestijama kroz niz protokola i izrade ovoga rada, kao i tijekom cijelog diplomskog studija, a i prije.

Izv. prof. dr. sc. Vlatki Čubrić Čurik, na pruženom znanju, pomoći i konstruktivnim savjetima tijekom diplomskog studija. Hvala na ustupljenim genotipovima iz projekta ANAGRAMS („Primjena NGS metoda u procjeni genomske varijabilnosti preživača") za potrebe ovog rada.

Doc.dr.sc. Maji Ferenčaković na nesebičnoj pomoći, strpljenju te dostupnošću u bilo koje doba dana i noći. Hvala na savjetima i sugestijama pri izradi ovoga rada te općenito tijekom cijelog diplomskog studija. Hvala na smirenošću i usmjeravanju tijekom niza situacija (kad bi bilo potrebno, a bilo bi).

Doc.dr.sc. Vladimiru Brajkoviću na ogromnoj količini vremena utrošenoj na nas studente. Hvala na skriptama, savjetima, sugestijama te pomoći tijekom izrade rada i diplomskog studija. Hvala što ste uvijek pristupačni i dobronamjerni.

Hvala svim prijateljima što su mi ispunili studentske dane. Hvala svim članovima KSA na druženjima, zabavama i svim agro raspravama tijekom studiranja, a prvenstveno Svenu i Dinu što su me u to uvalili. Hvala Antunu i Skeji na niz konstruktivnih agronomskih i bioloških rasprava. Hvala mojoj braći Amiru i Fuadu na podršci tijekom studiranja. Hvala mom mačku što je istrpio sa mnom srednjoškolsko i fakultetsko obrazovanje. Na kraju, hvala majci, Anki i ocu, Kamelu na podršci, strpljenju i povjerenju tijekom obrazovanja, pogotovo kad bi to trebalo.

SADRŽAJ

1. UVOD 1
2. PREGLED LITERATURE 3
2.1. Dosadašnja saznanja 3
2.1.1. Autosomi 3
2.1.2. Spolni kromosom X 4
2.2. Signali pozitivne selekcije 6
2.3. Identifikacija signala pozitivne selekcije 7
2.3.1. Markeri visoke rezolucije 7
2.3.2. Metode 8
2.3.2.1. ROH 8
2.3.2.2. iHS 9
2.4. Karakteristike autohtonih pasmina ovaca Hrvatske 11
3. MATERIJALI I METODE 13
3.1. Uzorak i genotipizacija 13
3.2. Priprema podataka za identifikaciju 13
3.2.1. Kontrola kvalitete 13
3.2.2. Autosomi 14
3.2.3. Spolni kromosom X 14
3.3. iHS 14
3.4. ROH 15
3.5. Identifikacija regija pod selekcijom 16
3.6. Anotacija gena 16
4. REZULTATI I RASPRAVA 17
4.1. Identificirani signali pozitivne selekcije i kandidat geni 17
4.1.1. Autosomi 17
4.1.1.1. iHS 17
4.1.1.1.1. OAR 2: 121.34 Mb – 123.95 Mb 19
4.1.1.1.2. OAR 6: 26.82 Mb – 42.29 Mb 20
4.1.1.1.3. OAR 10: 45.02 Mb – 46.53 Mb 22

4.1.1.1.4. OAR 18: 2.15 Mb – 3.86 Mb 23
4.1.1.1.5. OAR 23: 25.47 Mb – 26.78 Mb 24
4.1.1.2. eROHi 25
4.1.1.2.1. OAR 2: 111.81 Mb – 115.17 Mb 27
4.1.1.2.2. OAR 2: 121.3 Mb – 124.03 Mb 28
4.1.1.2.3. OAR 6: 32.71 Mb – 39.18 Mb 29
4.1.1.2.4. OAR 10: 36.35 Mb – 39.91 Mb 30
4.1.1.2.5. OAR 10: 42.11 – 45.58 Mb 31
4.1.1.2.6. OAR 10: 70.87 Mb – 71.93 Mb 31
4.1.1.3. Usporedba 32
4.1.2. Spolni kromosom X 33
4.1.1.2. iHS 33
4.1.1.3. eROHi 35
4.1.1.3.2. OAR X: 62.77 Mb - 75.21 Mb 37
4.1.1.4. Usporedba 38
6. ZAKLJUČCI 40
7. POPIS LITERATURE 41

SAŽETAK

Diplomskog rada studenta **Mario Shihabi**, naslova

IDENTIFIKACIJA SIGNALA POZITIVNE SELEKCIJE U GENOMU HRVATSKIH OVACA

Ovčarstvo je jedan od najvažnijih stočarskih sektora u mediteranskom području Hrvatske odakle potječe većina hrvatskih autohtonih pasmina. Prisutni signali pozitivne selekcije posljedica su adaptacije na oskudne okolišne čimbenike, ali i odabira (selekcije) ljudi. U ovom istraživanju je uz autosome bio uključen i spolni kromosom X (kod ženskih jedinki), zbog utjecaja spolno vezanog nasljeđivanja na ekonomski i evolucijski važna svojstva. Cilj rada bio je identificirati signale pozitivne selekcije putem dvije metode, iHS i eROHi, koristeći uzorak veličine 102 jedinke iz 4 autohtone pasmine ovaca Hrvatske i 606 006 SNP-ova (Infinium® HD SNP BeadChip), na kojem je prije analize obavljena kontrola kvalitete. Obje metode temelje se na analizi genomskih varijacija unutar jedne velike meta-populacije, koja ovdje predstavlja hrvatske autohtone pasmine ovaca. Uočeni su mnogi ekonomski i evolucijski važni geni te je tako razjašnjena prisutnost pojedinih signala u populaciji. Za precizniju i pouzdaniju analizu potrebno je povećati uzorak te uključiti još metoda identifikacije.

Ključne riječi: Hrvatske autohtone pasmine ovaca, Signali pozitivne selekcije, Kandidat geni, Autosomi, Spolni kromosom X.

ABSTRACT

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

DETECTION OF POSITIVE SELECTION SIGNATURES IN THE GENOME OF CROATIAN SHEEP

Sheep farming is one of the most important livestock sectors in the Mediterranean region of Croatia, where most of the Croatian autochthonous breeds originate from. The present positive selection signatures are a consequence of adaptation to scarce environmental factors, but also of artificial selection. In addition to autosomes, sex chromosome X (in females) was included in this study because of the known influence of sex-linked inheritance on many economically and evolutionarily important traits. The aim of this study was to identify positive selection signatures by two methods, iHS and eROHi, using a sample size of 102 individuals from 4 Croatian autochthonous breeds and 606 006 SNPs (Infinium® HD SNP BeadChip) on which quality control was performed before analysis. Both methods are based on the analysis of genomic variation within a large meta-population, which here represents Croatian autochthonous sheep breeds. Many economically and evolutionarily relevant genes were identified, highlighting the presence of multiple signatures in the population. For a more accurate and reliable analysis, it is necessary to increase the sample size and include additional identification methods.

Keywords: Croatian autochthonous sheep breeds, Positive selection signatures, Candidate genes, Autosomes, Sex chromosome X.

1. UVOD

Domestikacija i naknadna selekcija životinja na ekonomska ili morfološka svojstva mogu ostaviti različite signale na genomu populacije. Ta selekcija, u kombinaciji s prirodnom prilagodbom lokalnim okolišnim uvjetima, rezultirala je s preko tisuću različitih pasmina ovaca koje se fenotipski razlikuju (Kijas i sur. 2012.). Autohtone pasmine ovaca raširene su po cijelom svijetu i dobro su prilagođene različitim okolišnim uvjetima. Pasmine balkanskih ovaca (populacije) smatraju se prepoznatljivom svjetskom skupinom koja je prilagođena njihovom specifičnom geografskom i uzgojnom okruženju. Prema Ciani i sur. (2020.), Balkan i Italija (odnosno balkanske pasmine ovaca, zajedno s talijanskim pasminama) su regije čvorišta iz kojih su se ovce razišle po ostatku Europe. Dakle, razumijevanje regija genoma i susjednih gen kandidata koji su odgovorni za lokalnu prilagodbu ovaca na Balkanu, može pridonijeti uspjehu mnogih genomskih analiza. Također, uzgoj ovaca jedan je od najvažnijih stočarskih sektora u mediteranskom području Hrvatske. Stoga je identifikacija genomskih regija koje pokazuju signale pozitivne selekcije i od ekonomske važnosti. Uz navedeno, identifikacija ovih regija omogućila bi i bolje upravljanje varijabilnosti i zaštite ovih populacija.

Općenito, pozitivna selekcija povećava frekvenciju poželjnog alela te vodi ka njegovoj fiksaciji unutar populacije (Stainton i sur. 2015.). Kao rezultat, polimorfizam na području selekcije je smanjen u populaciji. Odnosno, dolazi do redukcije u genetskoj varijabilnosti oko poželjne mutacije (*selection sweep*), a udaljenost redukcije ovisi o selektivnom pritisku i starosti selekcije (broj generacija). Takvi jedinstveni otisci pronađeni u genomu jedinki, nazivaju se signalima pozitivne selekcije (Fay i Wu 2000., Jensen i sur. 2016.).

Napretkom bioinformatike i molekularne genetike, tj. sve većom dostupnošću genomskih informacija, javlja se sve veći broj statističkih metoda za identifikaciju signala pozitivne selekcije na genima i genomu. Metode koje se opsežno koriste, koristeći informacije iz jedne populacije (unutar populacijske metode), su *integrated haplotype score* (iHS) prema Voight i sur. (2006.) te *extreme Runs Of Homozigosity islands* (eROHi) prema Mészáros i sur. (2015.). iHS metoda se temelji na *linkage disequilibrium* teoriji i može otkriti regije s brzo povećanom frekvencijom deriviranih ili ancestralnih alela na području selekcije

(Voight i sur. 2006.). U odnosu na iHS metodu, *Runs of homozigosity* (ROH) su kontinuirani, neprekidni odsječci homozigotnih genotipova bez heterozigotnosti u diploidnom stanju (Gibson i sur. 2006.) i smatraju se signalima pozitivne selekcije, ako su visoko prisutni u populaciji.

Kao korak prema ciljevima iz prvog odlomka, ovo istraživanje je koristilo prethodno navedene dvije komplementarne metode za identifikaciju regija genoma, zajedno sa susjednim kandidat genima, koje pokazuju signale pozitivne selekcije kod 101 jedinke, od 4 autohtone pasmine ovaca Hrvatske (Dalmatinska pramenka, Dubrovačka ruda, Istarska ovca, Paška ovca). Identifikacija signala se vršila na temelju visoko rezolutnih genetskih markera (Ovine Infinium® HD SNP BeadChip), uz prethodno obavljenu kontrolu kvalitete. Zbog sličnih okolišnih uvjeta uzgoja, sve 4 pasmine (populacije) su tretirane kao jedna velika meta-populacija, koja predstavlja autohtone pasmine ovaca Hrvatske. Regije genoma koje pokazuju signale pozitivne selekcije najvjerojatnije su posljedica adaptacije na oskudne okolišne čimbenike, ali i odabira (selekcije) ljudi. U istraživanju je, uz autosome, bio uključen i spolni kromosom X, zbog utjecaja spolno vezanog nasljeđivanja na mnoga važna svojstva (u ekonomskom i evolucijskom pogledu). Kandidat geni koji su potencijalno uzrokovali statistički značajne signale selekcije određeni su koristeći informacije sa SNP & Variation Suite (SVS) v8.7.0 programskog paketa (Golden Helix, Inc., Bozeman, MT, www.goldenhelix.com).

2. PREGLED LITERATURE

2.1. Dosadašnja saznanja

Genom ovaca sastoji se od 27 parova kromosoma (26 autosoma i spolni kromosom: XX ili XY), ukupne približne veličine oko 2,61 Gb (Bahcall 2014.). Nakon uspješnog razvoja genomskih metoda i alata za identifikaciju signala selekcije kod goveda i ljudi, zbog ekonomske važnosti, povećava se i broj srodnih istraživanja na ovcama. Dosad, balkanske pasmine ovaca na razini genoma, nikad nisu analizirane na signale selekcije. Prije ulaska u samu analizu, u ovom poglavlju su ukratko navedena dosadašnja saznanja (istraživanja) o identificiranim signalima pozitivne selekcije kod ovaca. Također, opisane su karakteristike i važnost uključivanja spolnog kromosoma X u identifikaciju signala pozitivne selekcije.

2.1.1. Autosomi

Potrage za intervalima genoma, genima i polimorfizmima, koji značajno utječu na ekonomski važna svojstava kod ovaca, rezultirali su otkrivanjem signala pozitivne selekcije povezanih s bojom dlake (Fariello i sur. 2014.), metabolizmom i taloženjem masti (Moradi i sur. 2012.), prirastom i proizvodnjom mesa (Purfield i sur. 2017.), prinosom mlijeka (Gutiérrez-Gil i sur. 2014., Manunza i sur. 2016., Moioli i sur. 2013.), reproduktivnim svojstvima (Kijas i sur. 2012., Liu i sur. 2016.), vunom (Demars i sur. 2014.), otpornosti na parazite (Mcrae i sur. 2014.), prisutnosti/odsutnosti rogova (Kijas i sur. 2012., Liu i sur. 2016.), itd.

Nadalje, temperaturni režim jedan je od ključnih klimatskih čimbenika koji utječu na adaptaciju, odnosno preživljavanje i uspješno uzgajanje domaćih životinja (Yurchenko i sur. 2019.). Različite forme mnogih gena su favorizirane u različitim okolinama. Na primjer, prema Kim i sur. (2016.) u genomima egipatskih autohtonih pasmina ovaca dolazi do promjena u više gena koji utječu na adaptaciju na vruće sušno okruženje, uključujući gene koji sudjeluju u melanogenezi, veličini i razvoju tijela, energetskom i probavnom

metabolizmu, kao i u živčanom te autoimunom odgovoru. Također, Lv i sur. (2014.) identificirali su 17 gena koji su povezani s klimatskim odabirom pri promatranju skupa nativnih populacija ovaca iz širokog raspona geografskih područja. Devet gena sudjelovalo je izravno u energetskim i regulatornim aktivnostima (npr. TBC1D12 i FBXO8) ili kodiranim enzimskim aktivatorima (npr. THY1), dok je osam dodatnih gena bilo uključeno u endokrine i autoimune procese (npr. EDNRB, NMUR1, PRL, IL12RB1 i ACVR2A). Signali adaptacije, uzrokovani temperaturom i sunčevom svjetlošću, povezani su s TBC1D12 genom (Lv i sur. 2014.). Komparativna genomika pruža dodatne dokaze o povezanosti gena i aklimacije. Na primjer, TRPM8 gen utječe na toplinski osjećaj kod miševa (Bautista i sur. 2007.), a isti gen je povezan s tolerancijom na hladnoću kod ovaca (Kijas i sur. 2012., Liu i sur. 2016., Fariello i sur. 2014.).

2.1.2. Spolni kromosom X

X kromosom i autosomi razlikuju se u nekoliko aspekata njihove molekularne evolucije i populacijske genetike, uključujući efektivnu veličinu populacije (Ne), stopu mutacije, stopu rekombinacije, osjetljivost na genetski drift, signale selekcije i demografiju (Schaffner 2004., Vicoso i Charlesworth 2006., Hedrick 2007., Keinan i Reich 2010.), koji svi različito utječu na genetsku raznolikost X kromosoma u usporedbi s autosomima (Begun i Whitley 2000., Casto i sur. 2010., Heyer i Segurel 2010.). Veličina kromosoma X, kod sisavaca, varira od 150 do 170 Mb, ovisno o vrsti, što čini ~5 % genoma. Karakterizira ga stopa rekombinacije od 0,8 cM/Mb, niska stopa mutacije (0,015 mutacija/Mb/generaciji), umjerena genetska raznolikost (0,04 %) i efektivna veličina populacije (Ne) koja iznosi ¾ u odnosu na autosome (Schaffner 2004.). Također, smatra se specijaliziranijim od autosoma i ima važnu ulogu u evoluciji ljudi i životinja.

Kod sisavaca, ženski spol posjeduje dva, a muški po jedan X i Y kromosom. Stoga, identifikacija signala pozitivne selekcije na spolnom kromosomu X se može izvršiti samo kod ženskih jedinki, jer se kod muških jedinki X kromosom nalazi u hemizigotnom obliku (osim u blizini krajeva).

Smanjena razina genetske raznolikosti X kromosoma u odnosu na autosome, uzrokovana je izravno i neizravno. Izravno, nižom stopom mutacija i rekombinacija, a neizravno, povećanim djelovanjem selekcije i genetskog drifta zbog manje efektivne veličine populacije, što je posljedica niže stope rekombinacija. Niža stopa rekombinacije je posljedica mogućnosti rekombinacije na cijelom kromosomu X samo kod ženskog spola, a iznosi oko 2/3 u odnosu na prosjek genoma i uzrokuje višu razinu *linkage disequilibrium*-a. Istraživanja su pokazala da selekcija na X kromosomu smanjuje genomsku raznolikost u većoj mjeri od one na autosomima (19% -26%) (McVicker i sur. 2009.).

Uz prethodno, X kromosom ima visoku gustoću gena i zato može biti dobra podloga za identifikaciju signala selekcije (Zhu i sur. 2015.). Ma i sur. (2014.), proučavali su signale selekcije na X kromosomu kod svinja i utvrdili da su geni relevantni za kvalitetu mesa, reprodukciju i imunološki sustav pronađeni u potencijalnim selekcijskim regijama. Moradi i sur. (2012.), ispitivanjem tanko i masno repnih populacija ovaca, utvrdili su povećanu homozigotnost u selekcijskim regijama u korist pasmina s masnim repom na kromosomima 5 i X. Uzimajući u obzir spol, pritisak selekcije na X kromosom je viši od onog na autosomima, što ukazuje da su geni na X kromosomu pod izravnijom i učinkovitijom selekcijom (Disteche 2006., Graves 2006.). Na X kromosomu detektiran je niz selektivnih gena povezanih s fenotipovima abnormalnog ponašanja, poput hipoaktivnosti, smanjenog dugotrajnog potenciranja, abnormalnog rasta dlake te abnormalnog razvoja srca i klijetki (Chen i sur. 2018.). Ti su fenotipovi najvjerojatnije svojstva koja su bila pod snažnom selekcijom tijekom domestikacije (Kerje i sur. 2004., Carneiro i sur. 2014., Jensen 2014., Medugorac i sur. 2017.). Nadalje, prema Chen i sur. (2018.) X kromosom pokazao je više ROH-ova, veću prosječnu veličinu ROH-a i veću pokrivenost ROH-ovima, od autosoma kod 943 ženske jedinke iz 68 populacija ovaca. Zbog svega navedenog, jasna je važnost uključivanja spolnog kromosoma X u ispitivanje signala pozitivne selekcije kod ženskih jedinki.

2.2. Signali pozitivne selekcije

Pojava novog, poželjnog alela može biti uzrokovana mutacijom, migracijom iz druge subpopulacije ili može biti posljedica već postojećeg rijetkog alela nakon promjene okoline (prethodno neutralnog). Selekcija značajno utječe na promjenu frekvencija određenih alela ovisno o selekcijskom pritisku. Ako je poželjni alel snažno selektiran, neizbježno je da se i frekvencija povezanih (bliskih), neutralnih (ili slabo selektiranih) varijanti povećava (Stephan 2016.). Smith i Haigh (1974.) su opisali ovaj proces i nazvali ga genetskim *hitchhiking*-om. *Hitchhiking* efekt može drastično smanjiti genetsku varijabilnost na mjestu selekcije ili u blizini, što se naziva „*selective sweep*" (Berry i sur. 1991.). Odnosno, rezultat *hitchhiking* efekta su regije s povećanom homozigotnošću (signali pozitivne selekcije). To često dovodi do pojave različitih haplotipova u populacijama koje su bile izložene različitim selekcijskim pritiscima (Lewontin i Krakauer 1973.).

Signali pozitivne selekcije, odnosno *selective sweeps* koji proizlaze iz novog poželjnog alela nazivaju se *hard sweeps*, a oni koji proizlaze iz poželjnog, već postojećeg alela (prethodno neutralnog) ili iz nekoliko novih mutacija na istom području, nazivaju se *soft sweeps* (Hermisson i Pennings 2005.).

Glavna razlika između *soft* i *hard selective sweep* je u očekivanom broju različitih haplotipova koji posjeduju poželjnu mutaciju ili mutacije, a time i u očekivanom broju haplotipova koji *hitchhike*-u do značajne frekvencije tijekom *selective sweep*-a te koji ostaju u populaciji kad nastupi fiksacija. *Hard sweep* je događaj u kojem jedan haplotip sadrži selektivno poželjan alel i raste u frekvenciji, dok u *soft sweep*-u nekoliko haplotipova sadrže poželjan alel ili alele i mogu istodobno rasti u frekvenciji (Harris i sur. 2018.). Ova ključna razlika rezultira različitim očekivanjima u frekvencijama te veličini *linkage disequilibrium*-a, a time i u čestim statističkim testovima koji se temelje na tim parametrima (Jensen i sur. 2014.).

2.3. Identifikacija signala pozitivne selekcije

2.3.1. Markeri visoke rezolucije

Dugo godina genetsko mapiranje je bilo ograničeno na identificiranje putem tradicionalnih fenotipskih markera. To je uključivalo gene koji kodiraju lako uočljive karakteristike poput krvnih grupa ili oblika sjemena. Zbog kompleksnosti većine svojstava, potaknut je razvoj genetskih markera koji bi mogli identificirati genetske karakteristike koje nije lako uočiti (Pierce 2012.). Dosad, zabilježeno je mnogo različitih tipova genetskih markera za proučavanje genoma. Nakon dosta duge dominacije mikrosatelita, genetski marker koji u posljednje vrijeme prednjači u upotrebi je *single nucleotide polymorphism* (SNP) (Vignal i sur. 2002.). Kao što sugerira skraćenica, SNP je jedno nukleotidna varijabilnost te se smatra najpolimorfnijim markerom na genomu. Odnosno, SNP je bilo koja pozicija u genomu, u kojoj se javljaju dvije (ili više) nukleotidne varijante pri kojima je frekvencija rjeđe varijante > 1 % u populaciji (< 1 %; točkasta mutacija). Postoje ukupno četiri različite moguće nukleotidne varijante (4 dušične baze: adenin, timin, citozin i gvanin). SNP-ovi se javljaju u kodirajućim i nekodirajućim regijama genoma (Marwal i sur. 2004.)

SNP genotipizacija jedna je od najčešćih metoda za proučavanje genetskih varijacija i signala selekcije zato što su SNP-ovi genetski stabilni, veoma brojčani u genomu i prikladni za automatiziranu analizu velike propusnosti (Saravanan i sur. 2021.). U početku je bila intenzivna, vremenski zahtjevna i smatrana skupim procesom do danas kada je jedna od najautomatiziranijih, učinkovitijih i relativno jeftinih metoda (Marwal i sur. 2004.). Najčešći alat za genotipizaciju putem SNP-ova je “SNP čip” koji se sastoji od tisuće ili stotine tisuća markera raspoređenih po genomu. Pomoću SNP-ova na SNP čipu očitava se genomska informacija jedinki. Kod ovaca najčešće se koriste OvineSNP50 Bead Chip s 54 241 SNP-ova i Ovine Infinium® HD SNP BeadChip sa 606 006 SNP-ova (Kijas i sur. 2014.). Zbog veće gustoće pokrivenosti genoma SNP-ovima (veći broj SNP-ova), jedinke za ovo istraživanje su genotipizirane s Ovine Infinium® HD SNP BeadChip. Različite statističke metode razvijene su za otkrivanje signala pozitivne selekcije pomoću SNP podataka (Qanbari i Simianer 2014., Saravanan i sur. 2021.).

2.3.2. Metode

2.3.2.1. ROH

Karakterizaciju ROH segmenata omogućuju SNP podaci širom genoma. ROH-ovi su kontinuirani nizovi homozigotnih genotipova koji nastaju unutar jedinke kada dva haplotipa dijele zajedničkog pretka (Ferenčaković i sur. 2013.), odnosno kada se naslijede identični haplotipovi od oba roditelja (McQuillan i sur. 2008.). Opseg i učestalost ROH-ova pružaju informacije kako o podrijetlu same jedinke, tako i o populaciji u cjelini. Smatra se da je glavni uzrok njihova nastajanja križanje u srodstvu i stoga su dobar alat u procjeni razine inbridinga, posebice kad nije dostupan rodovnik (Ferenčaković i sur. 2013b.). Također, s obzirom na to da pozitivna selekcija je često karakterizirana lokalnim smanjenjem raznolikosti haplotipova, raspodjela ROH-ova u genomu daje informacije o genomskim regijama koje su potencijalno bile izložene nedavnom i/ili davnom selektivnom pritisku (Pryce i sur. 2014., Kim i sur. 2013.). Zato se ROH smatra učinkovitom metodom za identifikaciju signala pozitivne selekcije (Mastrangelo i sur. 2017., He i sur. 2020.). Ako uzmemo u obzir princip rekombinacije, prema duljini homozigotnog odsječka moguće je odrediti broj generacija od zajedničkog pretka (Ferenčaković i sur. 2013.). Veoma dugi odsječci predstavljaju nedavni selektivni pritisak (16 Mb dug segment ~3 generacije), dok kraći homozigotni odsječci predstavljaju davni selektivni pritisak (1 Mb dug segment ~50 generacija).

Nadalje, bitno je istaknuti da gustoća SNP čipa koji se koristi za identifikaciju ROH segmenata ima veliki utjecaj na učinkovitost identifikacije. Zbog njegove gustoće, 50 k SNP čip ima veću stopu greške zbog manje mogućnosti otkrivanja heterozigotnosti unutar ROH-a promatrane dužine (Ferenčaković i sur. 2013.). Kao rezultat toga, analize temeljene na takvim čipovima mogu precijeniti broj segmenata kraćih od 4 Mb (Purfield i sur. 2012.).

Postoje dva glavna pristupa za identificiranje ROH-a: promatrački algoritmi za brojanje genotipa i algoritmi temeljeni na modelima (*hidden Markov model*; HMM). Metode oba pristupa su implementirane u veliki broj programa. Prvi pristup je primarno stvoren za genotipske podatke te ne istražuje sve dostupne informacije iz sekvenciranja populacije.

Mnoge metode se zasnivaju na tome (*sliding window* metoda; Purcell i sur. 2007., *consecutive* metoda; Marras i sur. 2015.). Određuju se ulazni parametri poput broja dozvoljenih heterozigota i dopuštenog broja *missing* SNP-ova kako bi se tolerirale eventualne pogreške te propusti genotipizacije (Curik i sur. 2014.). Nasuprot tome, HMM je statistički model, razvijen kako bi se procijenila vjerojatnost *homozygosity by descent* (HBD) segmenta duž kromosoma, koristeći informacije o markerima, frekvencijama alela i postotcima genotipskih pogrešaka. Prema Solé i sur. (2017.), rezultati ROH procjene putem ova dva pristupa, su u visokoj korelaciji.

2.3.2.2. iHS

iHS metoda se temelji na proširenoj homozigotnosti haplotipa (*extended haplotype homozygosity*; EHH), izračunatoj za ancestralne i derivirane alele na svakom ciljnom SNP-u unutar populacije (Voight i sur. 2006.). EHH, kao funkcija udaljenosti, mjeri propadanje identiteta haplotipova koji na jednom kraju sadrže određeni ciljni (*core*) alel. Za svaki alel, homozigotnost haplotipa započinje s 1 i propada do 0 s povećanjem udaljenosti od ciljnog mjesta. Kada alel brzo raste u frekvenciji zbog snažne selekcije, tada je karakteristična visoka razina homozigotnosti haplotipa koja je mnogo duža nego što se očekivalo u neutralnom modelu. Stoga, u grafovima EHH u odnosu na udaljenost, područje pod EHH će obično biti puno veće za selektirani alel nego za neutralni alel. Da bi se uhvatio ovaj učinak, izračunava se integral opaženog propadanja EHH od određenog ciljnog alela, dok EHH ne dosegne 0,05. Prema Voight i sur. (2006.), ovaj integrirani EHH (iHH) (zbrojen u oba smjera od ciljnog SNP-a) označava se sa iHH_A ili iHH_D, ovisno o tome izračunava li se s obzirom na ancestralni ili derivirani ciljni alel.

Za iHS test vrijedi sljedeća formula (Voight i sur. 2006.):

$$nestandardizirani\ iHS = \ln\left(\frac{iHH_A}{iHH_D}\right).$$

Ako je stopa propadanja EHH slična na ancestralnim i deriviranim alelima, onda vrijedi $iHH_A/iHH_D \approx 1$, a time i nestandardizirani iHS ≈ 0. Velike negativne iHS vrijednosti ukazuju na neobično duge haplotipove koji sadrže derivirani alel, a velike pozitivne vrijednosti ukazuju

na duge haplotipove koji sadrže ancestralni alel. Budući da su u neutralnim modelima aleli niske frekvencije uglavnom mlađi i povezani s duljim haplotipovima od alela više frekvencije, nestandardizirani iHS se prilagođava (standardizira) kako bi se dobila konačna formula koja ima srednju vrijednost 0 i varijancu 1, bez obzira na frekvenciju alela na ciljnom SNP-u (Voight i sur. 2006.):

$$iHS = \frac{\ln\left(\frac{iHH_A}{iHH_D}\right) - E_\rho\left[\ln\left(\frac{iHH_A}{iHH_D}\right)\right]}{SD_\rho\left[\left(\frac{iHH_A}{iHH_D}\right)\right]}.$$

Očekivanje i standardna devijacija od $\ln(iHH_A/iHH_D)$ procjenjuju se iz empirijske raspodjele na SNP-ovima čija frekvencija p od deriviranog alela odgovara frekvenciji na ciljnom SNP-u. iHS je konstruiran tako da ima približno standardnu normalnu raspodjelu i stoga su veličine iHS signala iz različitih SNP-ova izravno usporedive, bez obzira na frekvencije alela na tim SNP-ovima. Budući da je iHS standardiziran pomoću empirijske raspodjele na cijelom genomu, on pruža mjeru koliko su neobični haplotipovi oko određenog SNP-a u odnosu na genom u cjelini (Voight i sur. 2006.).

iHS se izračunava za svaki SNP redom, tretirajući svaki SNP kao ciljni SNP. U svakom SNP-u mjeri snagu dokaza selekcije koja djeluje na ili u blizini tog SNP-a. Međutim, umjesto da svaki SNP se tretira zasebno, preciznije je tražiti prozore uzastopnih SNP-ova s ekstremnim iHS vrijednostima. To je zato što signali selekcije teže stvaranju nakupina ekstremnih iHS vrijednosti po cijeloj regiji selekcije, dok su pod neutralnim modelom ekstremne iHS vrijednosti raspršene ujednačenije (Voight i sur. 2006.).

U principu, najveće negativne iHS vrijednosti koje ukazuju da je derivirani alel porastao u frekvenciji, su najzanimljivije. Međutim, *selective sweep* također može proizvesti i velike pozitivne iHS vrijednosti na obližnjim SNP-ovima ako je ancestralni alel pod selekcijom. Iz tog razloga, u ovom istraživanju su tretirane i krajnje pozitivne i krajnje negativne iHS vrijednosti kao potencijalno zanimljive.

Za ovu metodu, potrebno je prvo procijeniti genetske mape visoke rezolucije na temelju *linkage disequilibrium*-a uzoraka te se zatim koriste procijenjene genetske udaljenosti pri izračunavanju iHS vrijednosti ((Voight i sur. 2006.).

iHS metoda ima ograničene mogućnosti u otkrivanju niskofrekventnih signala selekcije i signala selekcije koji su vrlo blizu fiksaciji (Voight i sur. 2006.).

2.4. Karakteristike autohtonih pasmina ovaca Hrvatske

Republika Hrvatska posjeduje 9 autohtonih (izvornih) pasmina ovaca: lička pramenka, dalmatinska pramenka, dubrovačka ruda, istarska ovca, paška ovca, creska ovca, rapska ovca, krčka ovca i cigaja. Sve naše autohtone pasmine ovaca imaju manje ili više izražen lokalni karakter. Zbog dugotrajnog uzgajanja na istom području, prilagodile su se određenim klimatskim i hranidbenim uvjetima te postale otporne na bolesti. Odlikuje ih visoka otpornost, skromnost i prilagodljivost različitim uvjetima držanja te proizvodnim ciljevima. Pripadaju skupini ovaca kombiniranih proizvodnih svojstava, s tim da se ponajviše koriste za proizvodnju mesa i mlijeka. Svoj potpuni rast i razvoj završavaju u dobi od 3 do 4 godine. Specifičnost im je i snažno izražen instinkt za pronalaženje hrane i na najoskudnijim terenima. U priobalju i otocima, preživljavanje u prilično surovim mediteranskim uvjetima, ljetne vrućine i često gladovanje zimi uz jake vjetrove i posolicu, formirali su ovce malog tjelesnog okvira. U unutrašnjosti Istre, Lici, Slavoniji i Baranji, karakterističan je malo krupniji tip ovaca. Jedina autohtona pasmina koja se uzgaja u Slavoniji i Baranji je Cigaja (Šimpraga i sur. 2008.).



Slika 2-1. Autohtone pasmine ovaca Hrvatske (A = Paška ovca, B = Dubrovačka ruda, C = Istarska ovca, D = Dalmatinska pramenka) (Hrvatske izvorne i zaštićene pasmine – Banka gena domaćih životinja (mps.hr)).

Tablica 2-1. Broj uzgajivača, veličina matične populacije, Ne[1] i status ugroženosti kod autohtonih pasmina ovaca Hrvatske, zabilježeni 2019-te godine (Hrvatske izvorne i zaštićene pasmine – Banka gena domaćih životinja (mps.hr)).

Pasmina	Broj uzgajivača	Broj muških jedinki	Broj ženskih jedinki	Ne	Status
Paška ovca	51	180	4629	693	Potencijalno ugrožena
Creska ovca	9	34	877	131	Ugrožena
Krčka ovca	6	10	468	39	Ugrožena
Istarska ovca	18	70	1344	266	Ugrožena
Rapska ovca	21	18	557	70	Ugrožena
Dalmatinska pramenka	105	278	10810	1084	Nije ugrožena
Lička pramenka	54	380	10657	1468	Nije ugrožena
Dubrovačka ruda	26	61	751	226	Ugrožena
Cigaja	13	34	1056	132	Ugrožena

[1]Ne = efektivna veličina populacije.

3. MATERIJALI I METODE

3.1. Uzorak i genotipizacija

Kao uzorak, korišteni su genotipovi iz projekta ANAGRAMS („Primjena NGS metoda u procjeni genomske varijabilnosti preživača"). Ukupno 102 jedinke od četiri autohtone pasmine ovaca Hrvatske (Dalmatinska pramenka; n = 26, Dubrovačka ruda; n = 26, Istarska ovca; n = 25, Paška ovca; n = 25) su genotipizirane putem Ovine Infinium® HD SNP BeadChip, koji sadrži 606 006 SNP-ova. Sve četiri pasmine su tretirane kao jedna velika meta-populacija.

3.2. Priprema podataka za identifikaciju

3.2.1. Kontrola kvalitete

Za kontrolu kvalitete genotipizacije, korišteni su softver SAS 9.4. (SAS Institute, Cary, NC) i softver PLINK v1.9 (Chang i sur. 2015., Purcell i sur. 2007.). S ciljem zadržavanja samo SNP-ova na autosomima (26) i spolnom kromosomu X u uzorku, putem SAS 9.4. softvera, isključeni su svi SNP-ovi na mitohondrijskoj DNA, Y kromosomu te oni koji nisu pripisani kromosomima. Nadalje, u istom softveru, koristeći Illumina parametre kontrole (*GenTrainScore* manji od 0,4 i *GenCallScore* manji ili jednak 0,8) isključeni su SNP-ovi upitne kvalitete. Idući korak kontrole kvalitete proveo se putem PLINK v1.9 softvera: isključeni su SNP-ovi koji nisu uspješno genotipizirani kod 10 % jedinki ili više (funkcija *--geno 0.1*), a zatim su isključene jedinke kojima je nedostajalo 5 % genotipova ili više (funkcija --mind 0.05). Konačno, pomoću istog softvera se vršilo izbacivanje duplih SNP-ova, koristeći funkciju *--list-duplicate-vars* (uz parametre *ids-only* i *suppress-first)* za određivanje duplikata, pa koristeći funkciju *--exclude* za njihovo izbacivanje iz uzorka.

3.2.2. Autosomi

Nakon prethodno obavljene kontrole kvalitete, uzorak (set podataka) se filtrirao na autosome i spolni kromosom X, putem PLINK v1.9 softvera (funkcija *--chr*). Za identifikaciju signala pozitivne selekcije na autosomima, preostalo je 457 429 SNP-ova genotipiziranih kod 101 jedinke (zbog loše genotipizacije, kontrolom kvalitete izbačena je jedna jedinka Dalmatinske pramenke).

3.2.3. Spolni kromosom X

Koristeći funkciju *--check-sex* (PLINK v1.9 softver), u uzorku (101 jedinka) je određen spol jedinki. Zatim, putem funkcije *--remove* izbačene su muške jedinke. Konačni set podataka za identifikaciju signala pozitivne selekcije na spolnom kromosomu X sastojao se od 21 748 SNP-ova genotipiziranih kod 59 ženskih jedinki (Dalmatinska pramenka; n = 13, Dubrovačka ruda; n = 20, Istarska ovca; n = 13, Paška ovca; n = 13).

3.3. iHS

Prema formuli iz poglavlja 2.3.2. (Voight i sur. 2006.), iHS je definiran kao omjer logaritama vrijednosti integrirane EHH za haplotipove koji centriraju ancestralni alel (iHH_A) i vrijednosti integrirane EHH za haplotipove koji centriraju derivirani alel (iHH_D). Kako bi se navedeno izračunalo, prethodno je nužno određivanje „faze“ (haplotipova) na razini svakog pojedinačnog kromosoma (autosoma i spolnog kromosoma X). Faziranje, koje je prethodilo iHS analizi izvršeno je putem Shapeit softvera (Delaneau i sur. 2008.). Nadalje, izračun iHH_A i iHH_D vrijednosti te zatim izračun standardizirane iHS vrijednosti za svaki autosomalni i X kromosomalni SNP, proveo se koristeći R programski paket rehh (Gautier i Vitalis 2012.). Za unos ulaznih podataka, navedeni paket zahtjeva vcf format.

Unutar rehh paketa, standardizacija se provela u svakom binu (0,0.25) zasebno zbog frekvencijske ovisnosti očekivanih iHS vrijednosti pod neutralnošću. Implicitna pretpostavka ovog pristupa je da u svakom binu dominiraju neutralni markeri.

Standardizirane iHS vrijednosti približno slijede standardnu Gaussovu krivulju, ako se markeri razvijaju pretežno neutralno. Stoga je praktično vrijednostima dodijeliti p vrijednost u odnosu na nultu hipotezu neutralne evolucije. P vrijednost na negativnoj $\log_{10}$ skali je dodijeljena objema stranama raspodjele, jer nas zanimaju i velike pozitivne i velike negativne iHS vrijednosti. Vizualizacija podataka je također obavljena unutar R programskog paketa rehh (Gautier i Vitalis 2012.).

3.4. ROH

U oba seta podataka zasebno, na svakoj jedinki procijenjeni su ROH-ovi pomoću SNP & Variation Suite (SVS) v8.7.0 programskog paketa (Golden Helix, Inc., Bozeman, MT, www.goldenhelix.com) koji koristi uzastopnu (*consecutive*) metodu. Za unos ulaznih podataka, navedeni softver zahtjeva datoteke u ped i map Plink formatu. Odabrani algoritam je dizajniran za pronalaženje uzastopnih homozigotnih SNP-ova, tako da kontinuirano na čitavom kromosomu, ispituje svaki mogući niz koji odgovara parametrima koje je postavio korisnik.

Kako bi se segment smatrao ROH-om, morao je imati 15 ili više uzastopnih homozigotnih SNP-ova, s razmacima od najviše 1000 kb između dva SNP-a. Nadalje, prilikom ROH procjene, definirano je pet različitih klasa po dužini: 1-2, 2-4, 4-8, 8-16 i >16 mb. Zbog stope greške genotipizacije HD SNP čipa (0,25 %), u svakoj pojedinoj klasi dozvolio se određeni broj heterozigota i *missing* SNP-ova, koji su izračunati prema Ferenčaković i sur. (2013b.). Dozvoljeno je sljedeće: ROH dužine 1-2 mb; jedan heterozigotni SNP i četiri *missing* SNP-a, ROH dužine 2-4 mb; dva heterozigotna SNP-a i osam *missing* SNP-ova, ROH dužine 4-8 mb; četiri heterozigotna SNP-a i 16 *missing* SNP-ova, ROH dužine 8-16 mb; osam heterozigotnih SNP-ova i 32 *missing* SNP-a te ROH dužine >16 mb; 16 heterozigotnih SNP-ova i 64 *missing* SNP-a.

Nakon ROH procjene, sa SAS 9.4. softverom izračunat je broj u kojem se svaki pojedini SNP nalazi unutar ROH-a u populaciji. Ta učestalost je zatim standardizirana u istom softveru, kako bi se mogla procijeniti p vrijednost za svaki pojedini SNP te potom transformirati na negativnu $\log_{10}$ skalu. P vrijednosti na negativnoj $\log_{10}$ skali su procijenjene u osnovnim R programskim paketima stats i base (R Core Team 2020.). Vizualizacija je obavljena u R programskom paketu qqman (Turner 2018.).

3.5. Identifikacija regija pod selekcijom

Značajnost identificiranih signala pozitivne selekcije se temeljila na p vrijednosti na negativnoj $\log_{10}$ skali od iHS i ROH metode, koje su prvo promatrane zasebno, pa komparativno. Prema frekvenciji SNP-a unutar ROH-a u populaciji, nizovi *outlier*-a, tj. SNP-ova s p vrijednosti na negativnoj $\log_{10}$ skali >5 na autosomima te >3 na spolnom kromosomu X, smatrani su značajnim signalima pozitivne selekcije (eROHi). Prema iHS vrijednosti SNP-a, nizovi *outlier*-a (minimalno 3 *outlier*-a unutar 500 kb), tj. SNP-ova s p vrijednosti na negativnoj $\log_{10}$ skali >5 na autosomima te >3 na spolnom kromosomu X, smatrani su značajnim signalima pozitivne selekcije. Nadalje, prilikom definiranja regija pod pozitivnom selekcijom, poziciji prvog i posljednjeg *outlier*-a unutar svakog signala, oduzeto je, odnosno nadodano 500 kb zbog *linkage disequilibrium*-a.

3.6. Anotacija gena

Nakon identificiranja značajnih regija pod pozitivnom selekcijom, anotacija gena provedena je koristeći informacije sa SNP & Variation Suite (SVS) v8.7.0 programskog paketa (Golden Helix, Inc., Bozeman, MT, www.goldenhelix.com) na temelju lokacija na OAR4.0 referentnom genomu ovaca. Funkcionalna analiza kandidat gena je provedena putem UniProt (https://www.uniprot.org) i GeneCards (https://www.genecards.org) platformi. U pretraživanju funkcije gena, uz genomsku informaciju ovaca (UniProt), korištene su i genomske informacije drugih vrsta, uključujući čovjeka (GeneCards) i govedo (UniProt).

4. REZULTATI I RASPRAVA

4.1. Identificirani signali pozitivne selekcije i kandidat geni

Ovo istraživanje bilo je podijeljeno u dvije analize (dva seta podataka). U prvoj analizi, identificirali su se signali pozitivne selekcije na temelju SNP-ova na autosomima, a u drugoj analizi signali pozitivne selekcije na temelju SNP-ova na spolnom kromosomu X kod ženskih jedinki. Prije početka analiza, prethodno je obavljena kontrola kvalitete na istom setu podataka. Kontrola kvalitete je obavljena kako bi se smanjila mogućnost pogreške u utvrđivanju stanja SNP-ova (homozigot ili heterozigot), što uvelike može utjecati na rezultate korištenih metoda.

4.1.1. Autosomi

Na očišćenom setu podataka od 101 jedinke i 457 429 autosomalnih SNP-ova, detektirani su signali pozitivne selekcije, koristeći dvije unutar-populacijske metode, iHS i eROHi, s ciljem povećanja preciznosti identifikacije.

4.1.1.1. iHS

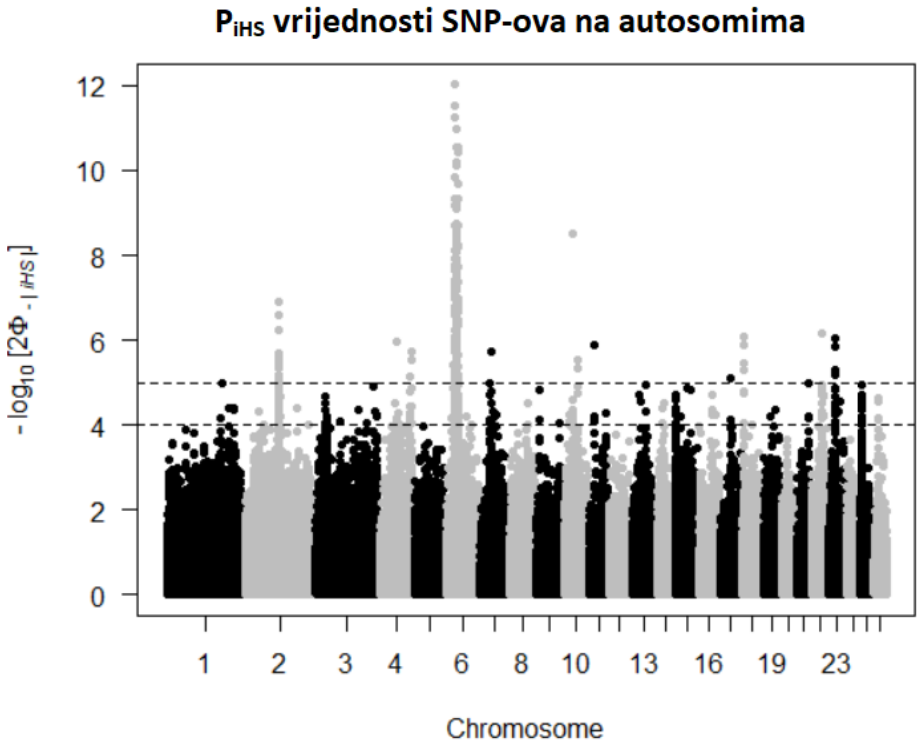
iHS i pripadajuće p vrijednosti, izračunate su za svaki SNP koristeći haplotipove. Na temelju njih, uočeno je 249 SNP *outlier*-a na razini autosoma (p vrijednost na negativnoj $\log_{10}$ skali >5). Zatim, slijedeći pristup niza *outlier*-a, utvrđeno je pet genomskih regija pod pozitivnom selekcijom (Tablica 4-1.), koje ukupno sadrže 237 značajnih SNP-ova. Zbog niskih P_{iHS} vrijednosti SNP-ova u njihovoj blizini (unutar 500 kb), 12 SNP *outlier*-a nisu spadali pod identificirane signale pozitivne selekcije. Kako bi vizualizirali raspodjelu regija pod pozitivnom selekcijom, Slika 4-1. prikazuje raspodjelu P_{iHS} vrijednosti autosomalnih SNP-ova u populaciji. Ukupno je identificirano 104 gena koji su se preklapali s pet regija pod

pozitivnom selekcijom. U nastavku, svaki od identificiranih signala (regija) pozitivne selekcije je opisan pojedinačno, uz pripadajuće gene i njihovu funkcionalnu analizu.

Tablica 4-1. Signali pozitivne selekcije identificirani iHS metodom na autosomima populacije.

Chr[1]	Pozicija (Mb)	nOutlier[2]	Peak SNP[3]	SNP pozicija (Mb)	P_{iHS}	Broj gena
2	121.34 – 123.95	15	rs420083082	122.21	6.92	12
6	26.82 – 42.29	211	rs424374580	33.35	12.01	71
10	45.02 – 46.53	3	rs419671628	46.04	5.53	1
18	2.15 – 3.86	4	rs421860218	2.65	6.1	5
23	25.47 – 26.78	4	rs411647206	26.13	6.06	15

[1]Chr = Kromosom, [2]nOutlier = Broj značajnih SNP-ova unutar pojedinog signala, [3]Peak SNP = SNP s najvećom P_{iHS} vrijednosti unutar pojedinog signala.

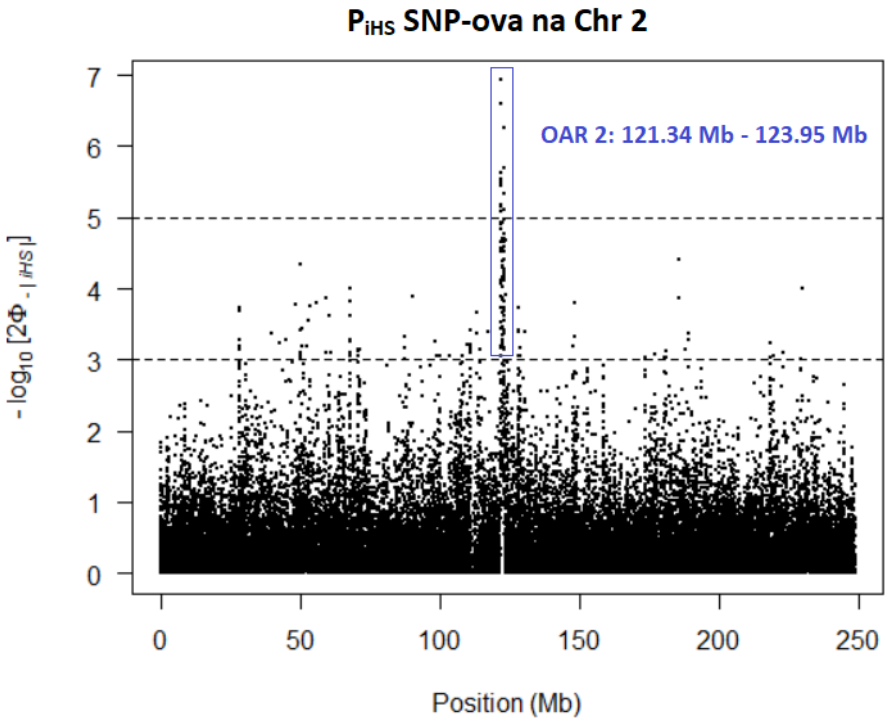


Slika 4-1. Raspodjela P_{iHS} vrijednosti SNP-ova na autosomima u populaciji.

4.1.1.1.1. OAR 2: 121.34 Mb – 123.95 Mb

Na kromosomu 2, detektiran je snažan signal pozitivne selekcije, duljine 2.61 Mb, koji se proteže od 121.34 Mb do 123.95 Mb. Sastoji se od ukupno 222 SNP-a, od kojih su 15 *outlier*-i. Najznačajniji *outlier*-i bili su rs420083082 na poziciji 122.21 Mb, sa |iHS| = -5.29 te rs407432408 na poziciji 123.43 Mb, sa |iHS| = 5.01. Nadalje, unutar ove regije identificirano je 12 potencijalnih kandidat gena: LOC105608872, ZSWIM2, LOC105611807, FAM171B, ITGAV, ZC3H15, LOC101120641, LOC101122056, FSIP2, LOC105608888, LOC101121406 i ZNF804A.

Geni LOC105608872, LOC105611807, LOC101120641, LOC105608888 i LOC101121406 još uvijek nemaju definiranu funkciju te dodijeljeno adekvatno ime. Također, novitet je i gen LOC101122056, međutim barem je poznato da kodira transmembranski protein koji sudjeluje u unutar staničnom transportu kod ovaca. Zatim, gen ZSWIM2 pozitivno regulira apoptozu te omogućuje aktivnost ubikvitin proteinske ligaze, a za gen FAM171B dosad nisu zabilježene funkcije kod ovce, goveda i čovjeka. Gen ITGAV ima ključnu ulogu u adipogenskoj diferencijaciji matičnih stanica masnog tkiva (Morandi i sur. 2016.), a akumulacija masnog tkiva često je karakteristika povećane tjelesne težine kod ovaca. Gen ZC3H15 pozitivno regulira aktivnost GTP-aze. Potrebno je istaknuti gen FSIP2 koji značajno utječe na reprodukciju, sudjelujući u sklopu i organizaciji flagelarnog aksonema spermija, odnosno sklopa mikrotubula i pridruženih proteina koji čine jezgru flagele spermija (odgovorna za kretanje spermija). Konačno, gen ZNF804A pozitivno regulira razvoj projekcije neurona i pozitivno regulira ekspresiju gena. Prema Gorssen i sur. (2021.), signal selekcije u ovoj regiji je također pronađen kod sljedećih mesnih pasmina; Rouge de l'Ouest te njemački, škotski i novo zelandski Texel, što ponajprije potencijalno ukazuje na selekciju određenih varijanti unutar ITGAV i FSIP2 gena.

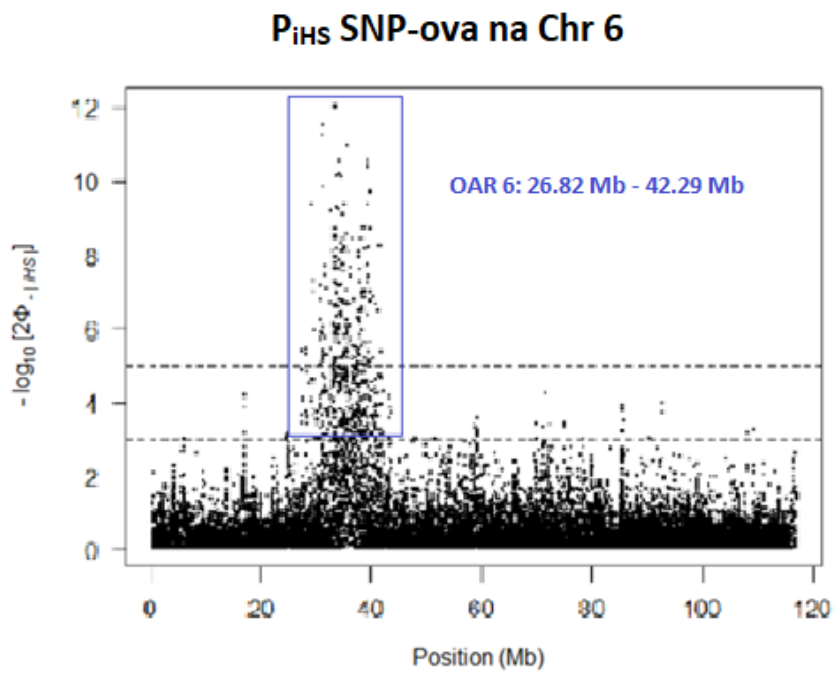


Slika 4-2. Raspodjela P_{iHS} vrijednosti SNP-ova na kromosomu 2.

4.1.1.1.2. OAR 6: 26.82 Mb – 42.29 Mb

Najsnažniji (najveća prosječna P_{iHS} vrijednost) signal pozitivne selekcije dužine 15.47 Mb, detektiran je na kromosomu 6, u području od 26.82 Mb do 42.29 Mb. Signal se sastojao od 2799 SNP-ova, od kojih su 211 *outlier*-i. Najznačajniji *outlier*-i bili su: rs424374580 na poziciji 33,35 Mb, sa |iHS| = -7.14, rs406425399 na poziciji 31.43 Mb, sa |iHS| = -6.98 i rs418558099 na poziciji 30.94 Mb, sa |iHS| = -6.89. Unutar navedene regije identificirano je 71 kandidat gena: PDHA2, UNC5C, LOC105615441, LOC105609581, TRNAK-CUU, LOC106991233, BMPR1B, LOC105615443, PDLIM5, HPGDS, LOC106991234, SMARCAD1, LOC101102565, LOC101116832, LOC101102814, ATOH1, GRID2, LOC106991211, TRNAS-AGA, CCSER1, TRNAW-CCA, LOC105615447, LOC106991208, MMRN1, SNCA, GPRIN3, TIGD2, FAM13A, HERC3, NAP1L5, PYURF, PIGY, HERC5, HERC6, PPM1K, ABCG2, PKD2, SPP1, MEPE, IBSP, LOC101103815, TRNAA-CGC, LAP3, MED28, FAM184B, LOC105615455, LOC105608051, DCAF16, NCAPG, LCORL, LOC105615456, LOC105608050, LOC105608049, LOC101104580, TRNASTOP-UCA, LOC106991224, SLIT2, LOC101122950, PACRGL, KCNIP4, LOC105608045, LOC106991210, TRNAS-GGA, LOC105611897, LOC106991209, LOC105615458, LOC101104829, TRNAW-CCA, ADGRA3, LOC105611900 i LOC101103396.

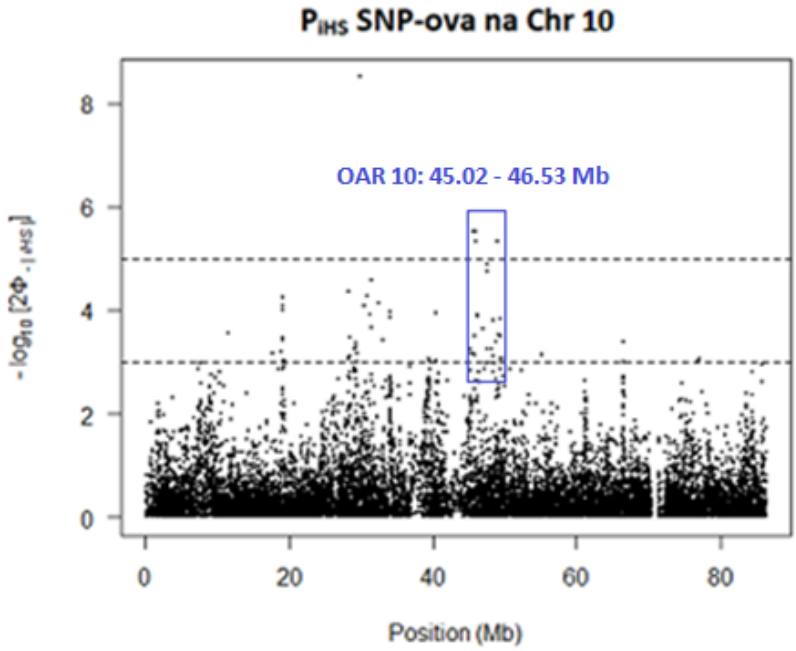
Geni naziva LOC (28) te TRNAK-CUU, TRNAS-AGA, TRNAW-CCA, TRNAA-CGC, TRNASTOP-UCA, TRNAS-GGA i TRNAW-CCA (7) nemaju zabilježenu funkciju. Nadalje, najznačajniji *outlier*-i su se nalazili unutar gena GRID2, CCSER1 i FAM13A. Gen GRID2 kodira membranski protein koji sudjeluje u transportu iona te je bitan u razvoju živčanog sustava i regulaciji apoptoze neurona. Prema Abo-Ismail i sur. (2018.), varijante u CCSER1 genu su povezane s iskoristivosti hranidbe kod goveda mesnog tipa, a kod ovaca Yurchenko i sur. (2019.) pronašli su signal pozitivne selekcije povezan s navedenim genom u autohtonim ruskim pasminama (Volgograd i Krasnoyarsk te bajkalske pasmine). Zatim, u ovoj regiji pronađeni su mnogi kandidat geni povezani s proizvodnjom mlijeka i prirastom (13 gena: SNCA, FAM13A, HERC6, ABCG2, PKD2, SPP1, MEPE, IBSP, LAP3, FAM184B, DCAF16, NCAPG i LCORL), koji su također zabilježeni kod navedenih ruskih pasmina. Inače, zabilježen je snažan signal pozitivne selekcije u području gena ABCG2, SPP1, LAP3, NCAPG, LCORL, PKD2, IBSP i MEPE kod mnogo pasmina domaćih ovaca, uključujući većinu europskih (Rochus i sur. 2018.) i autohtonih kineskih pasmina (Wei i sur. 2015.). Major gen u ovoj regiji povezan sa svojstvima proizvodnje mlijeka je ABCG2. SPP1 gen je povezan s postotkom i prinosom mliječnih proteina, prinosom mlijeka i regulacijom laktacije u mliječnih goveda (de Mello i sur. 2012.). Navedeno ukazuje da je ova regija pod pozitivnom selekcijom u našoj populaciji najviše zbog odabira (selekcije) ljudi prema vrijednostima svojstava relevantnih u proizvodnji mlijeka i sira te zatim zbog utjecaja na prirast i iskoristivost hranidbe (kombinirana proizvodnja mlijeko – meso).



Slika 4-3. Raspodjela P_{iHS} vrijednosti SNP-ova na kromosomu 6

4.1.1.1.3. OAR 10: 45.02 Mb – 46.53 Mb

Malena regija od 45.02 Mb do 46.53 Mb na kromosomu 10, duljine 1.51 Mb, detektirana je kao signal pozitivne selekcije od 90 SNP-ova (samo 3 *outlier*-a). Samo jedan gen, DACH1, je identificiran unutar navedene regije. Važan je za stanični ciklus te se smatra transkripcijskim faktorom koji je uključen u organogenezu. Nadalje, signal pozitivne selekcije povezan s navedenim genom navode Tao i sur. (2021.) u pasmini Luzhong te ga ističu kao jednog od kandidat gena povezanih s veličinom legla kod ovaca.

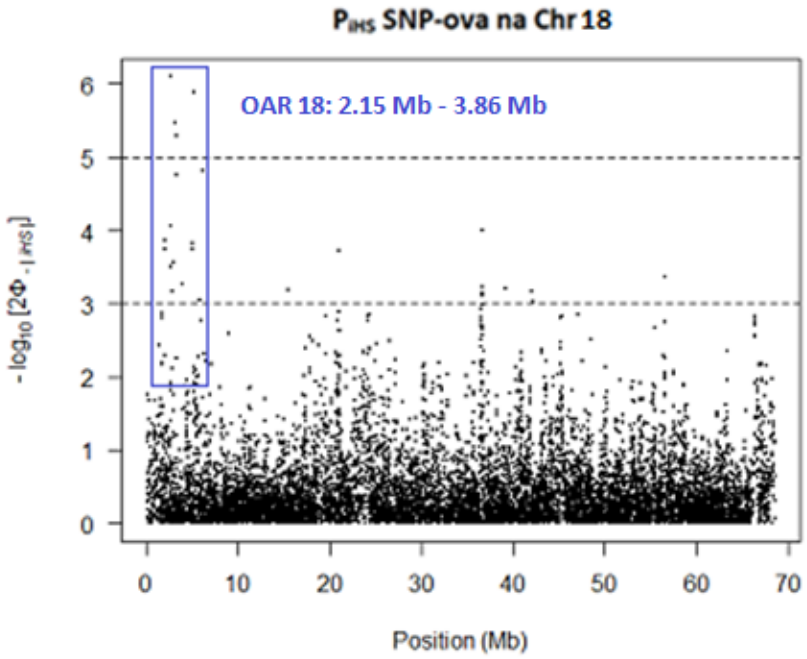


Slika 4-4. Raspodjela P_{iHS} vrijednosti SNP-ova na kromosomu 10.

4.1.1.1.4. OAR 18: 2.15 Mb – 3.86 Mb

Na kromosomu 8, detektirana je regija pod pozitivnom selekcijom, raspona 1.71 Mb u području od 2.15 Mb do 3.86 Mb. Regija sadrži 129 SNP-ova od kojih su četiri *outlier*-i. Unutar je identificirano pet gena; TRNAE-UUC, LOC106990481, GABRB3, GABRA5 i GABRG3.

Od toga, TRNAE-UUC i LOC106990481 nemaju zabilježenu funkciju. Geni GABRB3, GABRA5 te GABRG3 imaju važnu ulogu u formiranju ligand kloridnog kanala koji je komponenta heteropentamernog receptora za GABA (glavni inhibitorni neurotransmiter u mozgu). Stoga, ovi geni između ostalog utječu na istraživačko ponašanje, inervaciju, motoričko, reproduktivno te socijalno ponašanje, što ukazuje da su visoki udjeli pojedinih varijanti u navedenim genima u populaciji, potencijalno rezultat adaptacije.

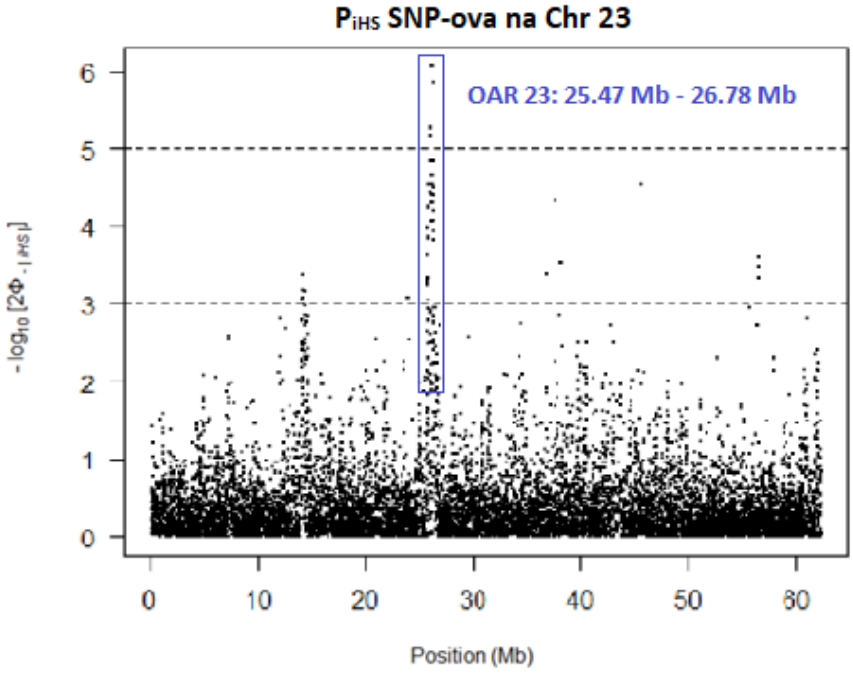


Slika 4-5. Raspodjela P_{iHS} vrijednosti SNP-ova na kromosomu 18.

4.1.1.1.5. OAR 23: 25.47 Mb – 26.78 Mb

Zadnja identificirana iHS regija na autosomima nalazila se na kromosomu 23, u području od 25.47 Mb do 26.78 Mb (1.31 Mb). Signal je sadržavao 67 SNP-ova od kojih su četiri *outlier*-a. Unutar je identificirano 15 kandidat gena, što ako usporedimo s prethodnim signalima, ukazuje na visoku gustoću gena u ovoj regiji. Identificirani su sljedeći kandidat geni; TRAPPC8, LOC105604435, TRNAG-CCC, B4GALT6, TTR, DSG2, LOC106990157, DSG3, LOC105604437, DSG4, DSG1, DSC1, DSC2, DSC3 i TRNAC-GCA.

Geni LOC105604435, TRNAG-CCC, LOC106990157, LOC105604437 i TRNAC-GCA su novitet i još nemaju zabilježenu funkciju. Nasuprot tome, geni DSG2, DSG3, DSG4, DSG1, DSC1, DSC2 i DSC3 značajno utječu na kvalitetu i kvantitetu vune. Stoga, evidenta je povezanost ove regije sa svojstvima proizvodnje vune. Prema tome, prisutnost ovog signala u populaciji potencijalno je posljedica oplemenjivanja naših priobalnih i otočnih pasmina s merino uvoznim ovnovima (najčešće talijanskim), kroz povijest. Signal selekcije povezan s ovim genima, također je identificiran u ruskim pasminama u radu Yurchenko i sur. (2019.).



Slika 4-6. Raspodjela P_{iHS} vrijednosti SNP-ova na kromosomu 23.

4.1.1.2. eROHi

ROH detekcija te zatim učestalost SNP-a unutar ROH-a i pripadajuće p vrijednosti, određene su za svaki autosomalni SNP.

Ovo istraživanje je koristilo HD SNP čip kako bi se mogli precizno identificirati i ROH-ovi manjih dužina (do 1 Mb), što prema Purfield i sur. (2012.) 50k SNP čip nije u stanju, zbog niske gustoće pokrivenosti SNP-ovima. Kod odluke o metodi ROH detekcije oslonio sam se na rad Curik i sur. (2014.) gdje su opisane razlike između dvije metode prebrajanja te na rad Solé i sur. (2017.) koji dokazuju kako su koeficijenti dobiveni složenim HMM algoritmima u visokoj korelaciji s metodama prebrojavanja koju koriste Ferenčaković i sur. (2013b.). Od metoda prebrojavanja, *consecutive* metoda je odabrana u odnosu na *sliding window* (PLINK) metodu, zato što *sliding window* metoda može potencijalno identificirati umjetne ROH-ove te krivo procijeniti segmente duže od postavljene dužine *window*-a (Ferenčaković i sur. 2013b.).

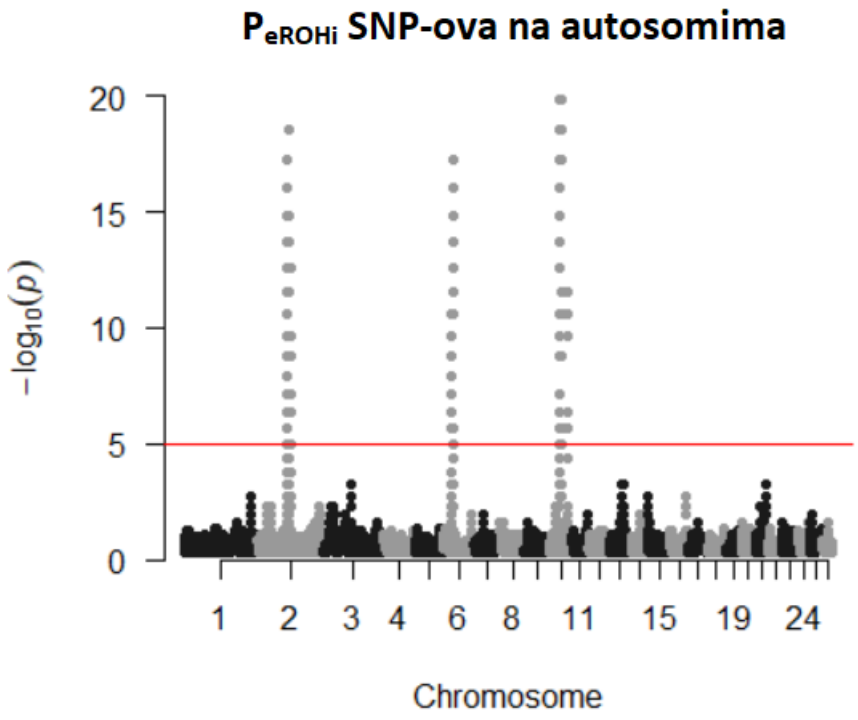
Nakon prvog koraka, SNP-ovi s p vrijednosti na negativnoj $\log_{10}$ skali >5, smatrani su značajnim te potom slijedeći pristup niza *outlier*-a, utvrđeno je šest genomskih regija pod pozitivnom selekcijom (eROHi; Tablica 4-2.), koje ukupno sadrže 2352 *outlier*-a. Unutar njih, ukupno je identificirano 110 gena. Prilikom promatranja eROHi identificiranih signala, najznačajniji SNP-ovi unutar signala predstavljaju tzv. pik, odnosno mjesto na kojem ili u blizini kojeg se najvjerojatnije nalazi uzročnik selekcije. Također, prema broju *outlier*-a, odnosno prema duljini pojedinog eROHi signala možemo promatrati starost inbridinga. Prema tome, veći signali su vjerojatno posljedica nedavnog odabira (selekcije) ljudi, a manji posljedica adaptacije ili odabira ljudi, kroz povijest.

Kao i kod iHS analize, u nastavku je opisan svaki identificirani signal pozitivne selekcije pojedinačno, uz pripadajuće gene. Slika 4-2. prikazuje raspodjelu P_{eROHi} vrijednosti SNP-ova na autosomima u populaciji.

Tablica 4-2. Signali pozitivne selekcije identificirani eROHi metodom na autosomima u populaciji.

Chr[1]	Pozicija (Mb)	nOutlier[2]	Peak SNP[3]	SNP pozicija (Mb)	P_{eROHi}	Broj gena
2	111.81 – 115.17	313	rs418759610	112.55	17.24	25
2	121.3 – 124.03	241	rs426247174	122.56	56.18	12
6	32.71 – 39.18	1038	rs409328107	37.16	17.24	37
10	36.35 – 39.91	270	rs419315841	37.9	45.32	19
10	42.11 – 45.58	466	rs422158607	43.54	24.08	4
10	70.87 – 71.93	24	rs400310253	70.5	11.59	13

[1]Chr = Kromosom, [2]nOutlier = Broj značajnih SNP-ova unutar pojedinog signala, [3]Peak SNP = SNP s najvećom P_{eROHi} vrijednosti unutar pojedinog signala.



Slika 4-7. Raspodjela P_{eROHi} SNP-ova na autosomima u populaciji.

4.1.1.2.1. OAR 2: 111.81 Mb – 115.17 Mb

eROHi metodom, prvi identificiran signal pozitivne selekcije nalazio se na kromosomu 2, u području od 111.81 Mb do 115.17 Mb, raspona 3.36 Mb. Signal se sastojao od 313 *outlier*-a od kojih je najznačajniji SNP bio rs418759610 na poziciji 112.55 Mb (nalazio se kod 31 % jedinki). Unutar ove regije identificirano je 25 kandidat gena; LOC101116291, LGSN, OCA2, LOC101102534, TRNAE-CUC, NIPA1, NIPA2, CYFIP1, TUBGCP5, CCDC115, IMP4, PTPN18, LOC101104460, LOC105611782, AMER3, ARHGEF4, FAM168B, PLEKHB2, LOC105608693, LOC101106144, LOC101106402, LOC106990902, LOC105608704, LOC105608706 i LOC101106660.

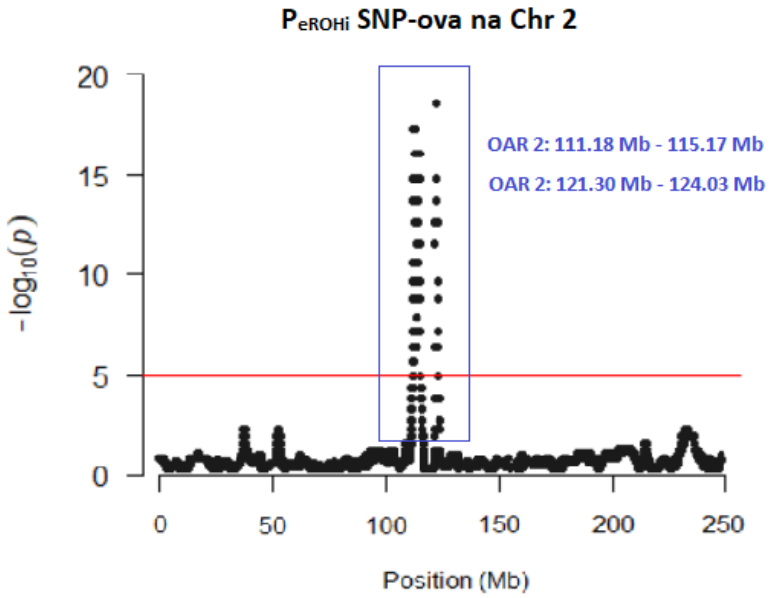
Geni naziva LOC (11) te TRNAE-CUC nemaju zabilježenu funkciju. Geni LGSN i OCA2 su povezani s pigmentacijom, odnosno varijacijom u boji dlake kod ljudi i domaćih životinja (Han i sur. 2008., Li i sur. 2014.). Saravanan i sur. (2021.) pronašli su signal povezan s prethodnim genima te signal povezan s genima ARHGEF4, FAM168B i PLEKHB2 u Changthangi pasmini ovce. Gen PLEKHB2 je značajno povezan sa svojstvima kvalitete mesa, a također je naznačen kao relevantan u Deccani pasmini od strane Leal-Gutierrez i sur.

(2020.). Općenito, Purfield i sur. (2017.) identificirali su QTL na konformaciju i težinu trupa koji se djelomično preklapa s ovom regijom. Nadalje, prema Grossen i sur. (2021.) navedeni signal pozitivne selekcije je također pronađen u sljedećim pasminama; njemački, škotski i novo zelandski Texel, Soay (primitivna pasmina) i Ronderib Afrikaner. Prethodno ukazuje da je u našoj populaciji najvjerojatnije regija pod selekcijom zbog značajnog utjecaja na proizvodnju mesa.

4.1.1.2.2. OAR 2: 121.3 Mb – 124.03 Mb

Na kromosomu 2, identificiran je još jedan signal pozitivne selekcije, od 121.3 Mb do 124.03 Mb (2.73 Mb). Sastojao se od 241 *outlier*-a, a najznačajniji SNP bio je rs426247174 na poziciji 122.56 Mb. Navedeni SNP se nalazio unutar ROH-a kod 53 % jedinki u populaciji. Prema prosječnim p vrijednostima unutar regije, ova regija se smatra najsnažnijim eROHi signalom u ovoj populaciji. Unutar nje pronađeno je 12 kandidat gena; LOC105608872, ZSWIM2, LOC105611807, FAM171B, ITGAV, ZC3H15, LOC101120641, LOC101122056, FSIP2, LOC105608888, LOC101121406 i ZNF804A.

Funkcionalna analiza gena, kao i dosadašnja saznanja o signalu prethodno su opisana u poglavlju 4.1.1.1.1.

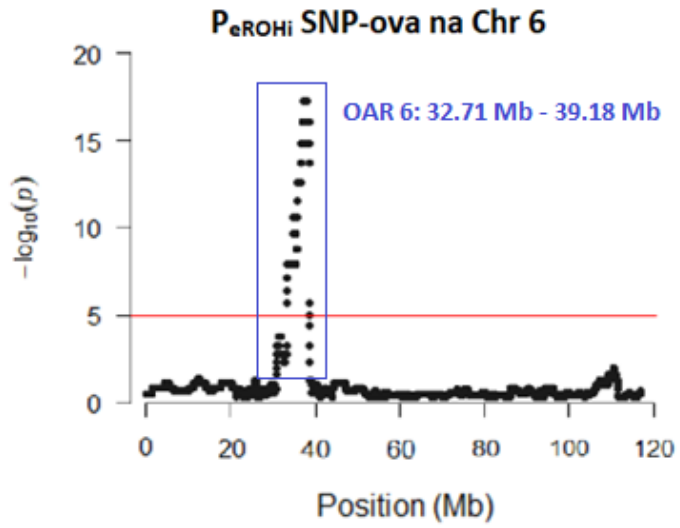


Slika 4-8. Raspodjela P_{eROHi} SNP-ova na kromosomu 2.

4.1.1.2.3. OAR 6: 32.71 Mb – 39.18 Mb

Na kromosomu 6, od 32.71 Mb do 39.18 Mb (6.47 Mb) identificiran je signal pozitivne selekcije od 1038 *outlier*-a. Najznačajniji *outlier* bio je SNP rs409328107 (37.16 Mb), koji se nalazio unutar ROH-a kod 31 % jedinki. Unutar ove regije detektirano je 37 gen kandidata; CCSER1, TRNAW-CCA, LOC105615447, LOC106991208, MMRN1, SNCA, GPRIN3, TIGD2, FAM13A, HERC3, NAP1L5, PYURF, PIGY, HERC5, HERC6, PPM1K, ABCG2, PKD2, SPP1, MEPE, IBSP, LOC101103815, TRNAA-CGC, LAP3, MED28, FAM184B, LOC105615455, NCAPG, LOC105608051, DCAF16, LCORL, LOC105615456, LOC105608050, LOC105608049, LOC101104580, TRNASTOP-UCA i LOC106991224.

Funkcionalna analiza gena te dosadašnja saznanja o ovom signalu prethodno su opisana u poglavlju 4.1.1.1.2. Duljina ovog signala, dodatno ukazuje na nedavni odabir (selekciju) ljudi, zbog utjecaja ovog signala na ekonomski važna svojstva (primarno na svojstva proizvodnje mlijeka i sira).



Slika 4-9. Raspodjela P_{eROHi} SNP-ova na kromosomu 6.

4.1.1.2.4. OAR 10: 36.35 Mb – 39.91 Mb

Prva od tri eROHi identificirane regije pod pozitivnom selekcijom na kromosomu 10, protezala se od 36.35 Mb do 39.91 Mb (3.56 Mb). Čini ju 270 *outlier*-a od kojih je najznačajniji bio rs419315841, na 37.9 Mb (31 % jedinki). Unutar nje identificirano je 19 kandidat gena; ZMYM2, LOC101119317, PSPC1, LOC105611677, LOC106991381, MPHOSPH8, LOC106990345, PARP4, TRNAG-UCC, CENPJ, LOC105611678, RNF17, LOC101120091, ATP12A, LOC105614068, LOC101120586, LOC101120503, LOC105613168 i PCDH9.

Od toga, geni naziva LOC (10) te TRNAG-UCC nemaju navedenu funkciju. Gen ZMYM2 omogučuje vezanje iona cinka te može djelovati kao faktor transkripcije. Gen PSPC1 između ostalog sudjeluje u regulaciji cirkadijanog ritma, aktivaciji imunološkog odgovora te ritmičkim procesima, a prema Myojin i sur. (2004.) regulira nuklearne događaje tijekom spermatogeneze te tako značajno utječe na plodnost/reprodukciju. Nadalje, gen PARP4 sudjeluje u staničnoj smrti, upalnom odgovoru i regulaciji aktivnosti telomeraze. Gen CENPJ kodira centrosomalni protein koji sudjeluje u formiranju centriola i vrlo je važan za funkcije centrosoma kao što su mitoza, pokretljivost ili unutarstanična komunikacija. Također, značajno uteče na probavu i metabolizam. Zatim, gen RNF17 sudjeluje u razvoju spermatida i tako značajno utječe na reprodukciju, a gen ATP12A katalizira hidrolizu ATP-a izmjenom H^+ i K^+ iona kroz plazmatsku membranu (odgovoran za apsorpciju kalija u raznim tkivima). Kim i sur. (2016.), usporedbom egipatske Barki pasmine i britanskih pasmina, navode gen PCDH9 kao kandidat gen adaptacije na sušna okruženja. Dok nasuprot, Mastrangelo i sur. (2019.) usporedbom fenotipa navedenih pasmina te uzevši u obzir saznanja kod ljudi (Wang i sur. (2011.) PCDH9 navode kao kandidat gen na pretilost), ukazuju da se potencijalno radi o gen kandidatu na taloženje masti. Signal u ovoj regiji također je pronađen u indijskoj Garole te talijanskoj Merinizzata pasmini (Gorssen i sur. 2021.). Prema svemu navedenom, ovaj signal je najvjerojatnije prisutan u našoj populaciji zbog utjecaja na reprodukciju i adaptaciju.

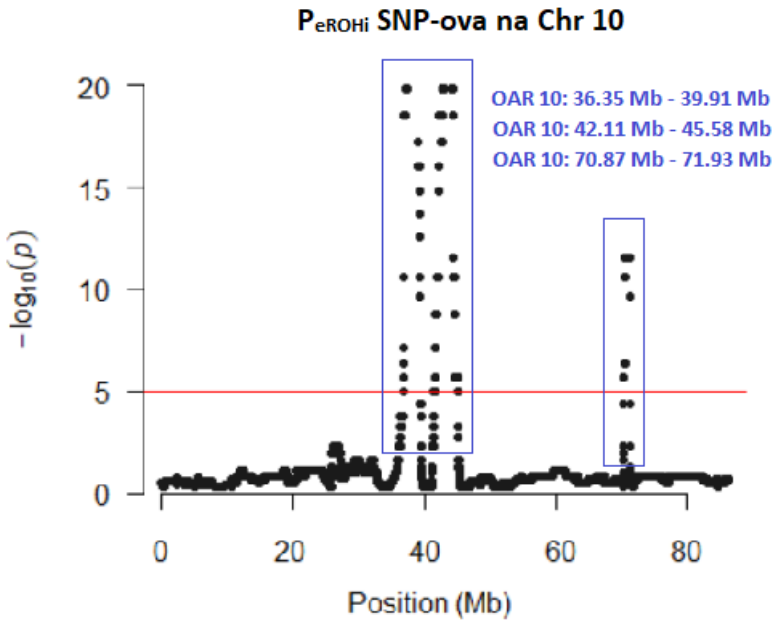
4.1.1.2.5. OAR 10: 42.11 – 45.58 Mb

Drugi signal na kromosomu 10, pronađen je u regiji od 42.11 Mb do 45.58 Mb (3.47 Mb) te se sastojao od 466 *outlier*-a. Najznačajniji *outlier* bio je rs422158607 na poziciji 43.54 Mb (31 % jedinki), a unutar ove regije identificirano je četiri kandidat gena; LOC101121526, TRNAW-CCA, LOC101121784 i KLHL1.

Prva tri navedena gena nemaju definiranu funkciju. Nasuprot tome, gen KLHL1 sudjeluje u neurološkom razvoju te tako između ostalog, utječe na lokomotorno ponašanje. Prema Sweet-Jones i sur. (2021.), navedeni gen je identificiran i naglašen kao najbolje rangiran unutar signala kod velških autohtonih pasmina. Prisutnost ovog signala u populaciji je također potencijalno rezultat adaptacije.

4.1.1.2.6. OAR 10: 70.87 Mb – 71.93 Mb

Zadnji pronađen signal na kromosomu 10 (70.87 Mb – 71.93 Mb), bio je duljine 1.06 Mb i sastojao se od samo 24 *outlier*-a (najznačajniji outlier; rs400310253, na poziciji 70.5 Mb). Najznačajniji *outlier* se nalazio unutar ROH-a kod 26 % jedinki. U ovoj regiji je pronađeno 13 gena; LOC105601928, LOC105613864, LOC105616214, LOC105614203, LOC105616213, LOC101108321, LOC105616212, LOC106991389, LOC105601884, LOC101109105, LOC101109370, LOC101107612 i LOC101109626. Međutim, niti jedan od navedenih gena nema definiranu funkciju. Stoga, jedino što možemo zaključiti o ovom signalu je da duljina signala ukazuje na davni inbriding.



Slika 4-10. Raspodjela P_{eROHi} SNP-ova na kromosomu 10.

4.1.1.3. Usporedba

Usporedbom signala pozitivne selekcije identificiranih iHS metodom, sa signalima identificiranim eROHi metodom na autosomima u populaciji, identificirane su dvije regije koje su se preklapale; OAR 2: 121.34 Mb – 123.95 Mb i OAR 6: 32.71 Mb – 39.18 Mb. Stoga, s visokom preciznošću se može utvrditi da su navedene regije nositelji pojedinih karakteristika unutar proučavane populacije, na što ukazuje i njihova funkcionalna analiza, odnosno geni koji se nalaze unutar njih. Međutim, ako uzmemo u obzir činjenicu da iHS metoda ima ograničene mogućnosti u identifikaciji signala selekcije koji su vrlo blizu fiksaciji (Voight i sur. 2006.), detekcija signala koristeći komparativno ove dvije metode može promaknuti pojedine signale. Odnosno, u slučaju da je neki signal nositelj pojedine karakteristike specifične za našu populaciju, nalazio bi se kod svih jedinki. Takav signal je moguće identificirati ROH analizom, ali iHS metodom to nije moguće (temelji se na odnosu iHH_A i iHH_D; minimalni *minor allele frequency* (MAF) = 0,05). U ovoj analizi, iHS metodom na autosomima nije bilo moguće procijeniti vrijednost kod 10 666 SNP-ova. Kako bi se provela preciznija i pouzdanija analiza, odnosno ispravili prethodni nedostatci, prvenstveno je

potrebno uključiti još metoda identifikacije signala pozitivne selekcije, a potom i povećati uzorak.

4.1.2. Spolni kromosom X

Na očišćenom setu podataka od 59 ženskih jedinki i 21 748 SNP-ova, koristeći iste metode kao kod autosoma, detektirani su signali pozitivne selekcije na spolnom kromosomu X. Identifikaciju signala pozitivne selekcije na spolnom kromosomu X moguće je izvršiti samo kod ženskih jedinki, zato što se X kromosom, kod muških jedinki, nalazi u hemizigotnom obliku (osim u blizini krajeva).

S ekonomskog stajališta i stajališta konzervacijske genetike važno je uključiti informaciju sa spolnog kromosoma u identifikaciju signala pozitivne selekcije. Također, X kromosom je pod višim selekcijskim pritiskom u odnosu na autosome (Disteche 2006., Graves 2006., McVicker i sur. 2009., Chen i sur. 2018.) i ima visoku gustoću gena (Zhu i sur. 2015.) te zato može biti dobra podloga za identifikaciju signala selekcije.

Kako bi se korigirao viši selekcijski pritisak, snizila se razina značajnosti u odnosu na autosome, odnosno X kromosomalni SNP-ovi s p vrijednosti na negativnoj $\log_{10}$ skali>3 su smatrani značajnim.

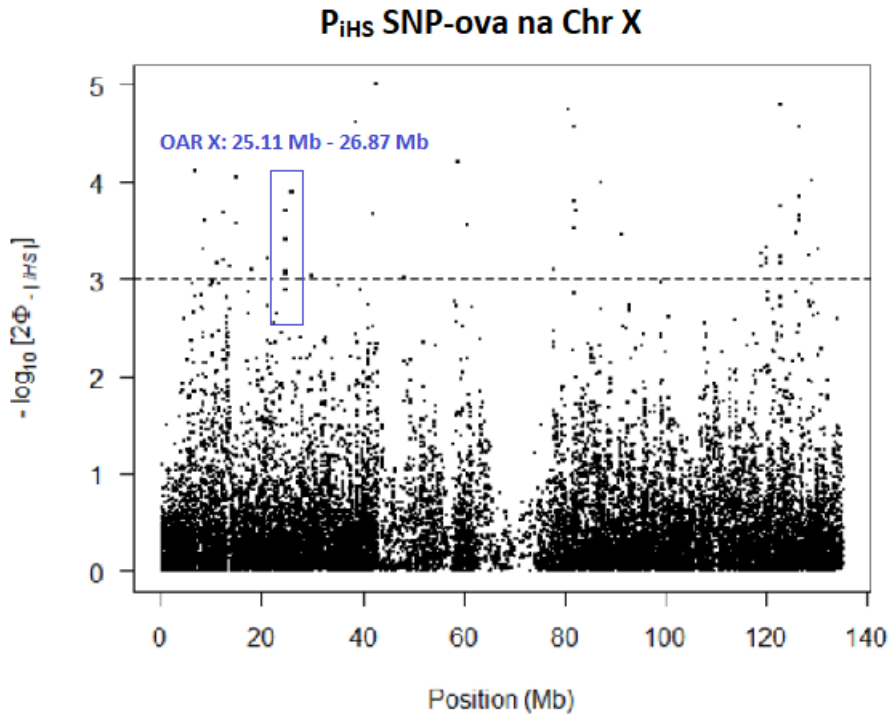
4.1.1.2. iHS

iHS pristupom, uočeno je 59 SNP *outlier*-a, međutim utvrđena je samo jedna genomska regija pod pozitivnom selekcijom (Tablica 4-3.), koja je opisana ispod. Zbog niskih P_{iHS} vrijednosti SNP-ova u njihovoj blizini (unutar 500 kb), 53 *outlier*-a su zanemareni. Slika 4-11. prikazuje raspodjelu P_{iHS} vrijednosti SNP-ova na spolnom kromosomu X.

Tablica 4-3. Signali pozitivne selekcije identificirani iHS metodom na spolnom kromosomu X u populaciji.

Chr[1]	Pozicija (Mb)	nOutlier[2]	Peak SNP[3]	SNP pozicija (Mb)	P_{Ihs}	Broj gena
X	25.11 – 26.87	6	rs427403757	26.37	3.9	3

[1]Chr = Kromosom, [2]nOutlier = Broj značajnih SNP-ova unutar pojedinog signala, [3]Peak SNP = SNP s najvećom P_{iHS} vrijednosti unutar pojedinog signala.



Slika 4-11. Raspodjela P_{iHS} SNP-ova na spolnom kromosomu X.

4.1.1.2.1. OAR X: 25.11 Mb - 26.87 Mb

Jedini iHS identificiran signal pozitivne selekcije na spolnom kromosomu X, nalazio se u području od 25.11 Mb do 26.87 Mb (1.76 Mb). Sastojao se od 194 SNP-a, od kojih su 6 *outlier*-i. Najveće |iHS| vrijednosti unutar signala su uočene na poziciji 25.61 Mb, a unutar ove regije pronađeno je tri kandidat gena; LOC101116885, IL1RAPL1 i LOC105605662.

Gen LOC101116885 kodira protein koji je sastavni dio ribosoma, a gen LOC105605662 nema zabilježenu funkciju. Preostali gen IL1RAPL1 sudjeluje u prijenosu signala, a Gambino i sur. (2009.) ističu njegovu uključenost u kontrolu inhibitornih mreža tijekom cerebralnog razvoja te umiješanost u kognitivno oštećenje vezano za spolni kromosom X. Zatim, kod

čovjeka je povezan sa sustavom memorije hipokampusa i sposobnošću učenja. Prema tome, glavnim uzročnikom prisutnosti ovog signala kod ženskih jedinki smatra se gen ILRAPL1 (potencijalni rezultat adaptacije).

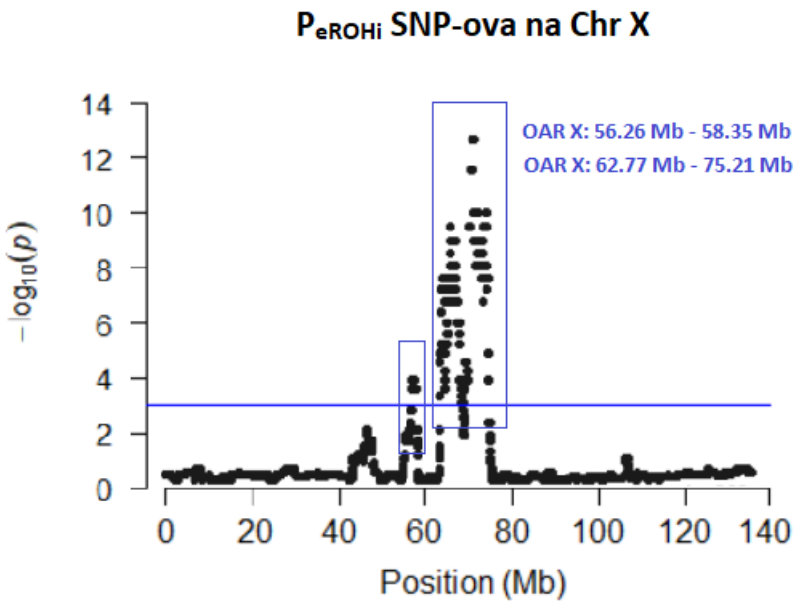
4.1.1.3. eROHi

Kao i kod autosoma, na spolnom kromosomu X izvršena je ROH detekcija te zatim učestalost svakog pojedinog SNP-a unutar ROH-a uz pripadajuće p vrijednosti. Potom, određeni su *outlier*-i te pristupom niza *outlier*-a utvrđene su dvije genomske regije pod pozitivnom selekcijom (Tablica 4-4.), koje sadrže ukupno 620 *outlier*-a. Navedene regije su opisane u sljedeća dva poglavlja, uz pripadajuće gene, a unutar njih ukupno je identificirano 102 gena. Slika 4-12. prikazuje raspodjelu P_{eROHi} vrijednosti SNP-ova na spolnom kromosomu X u populaciji.

Tablica 4-4. Signali pozitivne selekcije identificirani eROHi metodom na spolnom kromosomu X u populaciji.

Chr[1]	Pozicija (Mb)	nOutlier[2]	Peak SNP[3]	SNP pozicija (Mb)	P_{eROHi}	Broj gena
X	56.26 – 58.35	44	rs403393381	57.29	3.93	6
X	62.77 – 75.21	576	rs402113599	71.06	12.68	96

[1]Chr = Kromosom, [2]nOutlier = Broj značajnih SNP-ova unutar pojedinog signala, [3]Peak SNP = SNP s najvećom P_{eROHi} vrijednosti unutar pojedinog signala.



Slika 4-12. Raspodjela P_{eROHi} SNP-ova na spolnom kromosomu X.

4.1.1.3.1. OAR X: 56.25 Mb - 58.35 Mb

Na spolnom kromosomu X, prva identificirana eROHi regija nalazila se u području od 56.25 Mb do 58.35 Mb (2.1 Mb). Sastojala se od 44 *outlier*-a, a najznačajniji *outlier*, pronađen unutar ROH-a kod 46 % jedinki, bio je rs403393381 na poziciji 57.29 Mb. Unutar ove regije identificirano je šest kandidat gena; LOC101119962, AR, OPHN1, YIPF6, STARD8 i EFNB1.

Gen LOC101119962 nema naznačenu funkciju. Gen AR, između ostalog, sudjeluje u signalnom putu receptora androgena i estrogena te u staničnom odgovoru na podražaj sterodinih hormona. Gen OPHN1 ima niz funkcija koji je značajan u biološkim ritmovima, diferencijaciji, transkripciji te regulaciji transkripcije. Nadalje, gen YIPF6 sudjeluje u razvoju crijevnih epitelnih stanica, a gen STARD8 u prijenosu signala. Konačno, gen EFNB1 sudjeluje u diferencijaciji, razvoju neurološkog sustava, specifikaciji embrija te pozitivnoj regulaciji proliferacije T stanica. Pregledom literature nije uočena prisutnost ove regije kod drugih pasmina, međutim dosad je zabilježen malen broj istraživanja koji su proučavali signale pozitivne selekcije na spolnom kromosomu X kod ovaca. Stoga, jedino što možemo utvrditi je sljedeće; prema funkcionalnoj analizi gena, jasan je utjecaj ovog signala na niz općenitih, a

i spolno ograničenih svojstava, ali ne možemo sa sigurnošću razjasniti njegovu prisutnost u populaciji.

4.1.1.3.2. OAR X: 62.77 Mb - 75.21 Mb

Druga eROHi identificirana regija na spolnom kromosomu X, sastojala se od 600 SNP-ova, od kojih su 576 *outlier*-i. Prema prosječnoj p vrijednosti navedeni signal je puno snažniji od prethodnog. Protezao se od 62.77 Mb do 75.21 Mb, duljine 12.44 Mb. 24 ne značajna SNP-a, sa graničnim vrijednostima, nalazila su se u području od 68.59 Mb – 69.19 Mb. Najznačajniji outlier bio je rs402113599 na poziciji 71.06 Mb (čak 80 % jedinki). Unutar ove regije detektirano je 96 kandidat gena; LOC105605507, LOC101112542, MIR374B, MIR374A, ZCCHC13, LOC101103953, TRNAW-CCA, SLC16A2, RLIM, LOC101113206, KIAA2022, LOC105605508, ABCB7, LOC106990333, LOC101104721, LOC101115345, UPRT, ZDHHC15, MAGEE2, ATP7A, LOC106991989, LOC101116886, MAGT1, ATRX, LOC105605514, LOC105605513, LOC106990342, FGF16, LOC101105221, LOC101105476, LOC101105726, PBDC1, LOC101105984, PGK1, TAF9B, LOC101106237, LOC106991969, LOC105605515, LOC105605685, ZCCHC5, LPAR4, P2RY10, LOC105605516, LOC101106999, GPR174, ITM2A, LOC101119617, TBX22, LOC101119881, LOC106991970, LOC101108029, FAM46D, BRWD3, LOC105605518, LOC105605688, HMGN5, SH3BGRL, LOC105605689, LOC101108813, LOC101109075, LOC105601886, LOC101109338, LOC101109596, LOC101120893, LOC101109864, LOC106991982, LOC101121151, POU3F4, LOC106991971, CYLC1, LOC105605690, RPS6KA6, LOC101110913, LOC105605691, HDX, LOC101111166, LOC105605519, LOC101111431, LOC101122577, LOC101111692, LOC105605520, APOOL, LOC101123097, ZNF711, POF1B, TRNAC-GCA, LOC101112957, CHM, DACH2, LOC106990654, LOC101103041, LOC106990655, LOC105605524, LOC101103296, LOC101113993 i LOC101114248.

Geni naziva LOC (58) te MIR374B, MIR374A, TRNAW-CCA i TRNAC-GCA nemaju zabilježenu funkciju. Također, neku posebno bitnu funkciju nemaju ZCCHC13, KIAA2022, ABCB7, UPRT, ZDHHC15, MAGT1, FGF16, PBDC1, PGK1, TAF9B, ZCCHC5, LPAR4, P2RY10, GPR174, ITM2A, TBX22, SH3BGRL, ZNF711 i DACH2. Funkcije preostalih gena su opisane u

nastavku. SLC16A2 gen sudjeluje u metaboličkom procesu staničnih aminokiselina i hormona štitnjače, negativnoj regulaciji proliferacije živčanih prekursorskih stanica te transportu preko krvno-moždane barijere. Gen RLIM sudjeluje u nasumičnoj inaktivaciji X kromosoma, a MAGEE2 se smatra tumorskim antigenom. Gen ATP7A pozitivno regulira biosintezu melanina. Zatim, kod gena ATRX zabilježen je niz funkcija koje sudjeluju u organizaciji kromosoma, popravku DNA, razvoju prednjeg mozga i spermatogenezi. Gen HMGN5 sudjeluje u organizaciji kromatina, stimulira ekspresiju gena i inhibira apoptozu. Gen APOOL kodira komponentu kompleksa MICOS, velikog proteinskog kompleksa unutarnje membrane mitohondrija koji ima presudnu ulogu u održavanju kristalnih spojeva, arhitekturi unutarnje membrane i stvaranju mjesta kontakta s vanjskom membranom. Gen POF1B sudjeluje u organizaciji aktinskih niti, sklopu dvostaničnog čvrstog spoja i morfogenezi epitelnih stanica, dok gen CHM pozitivno regulira aktivnost GTP-aze.

Pik signala se protezao od 70.3 Mb do 71.9 Mb, što prvotno indicira na gene POU3F4, CYLC1, RPS6KA6 i HDX. Gen POU3F4 ima važnu ulogu u morfogenezi pužnice, diferencijaciji neurona prednjeg mozga, negativnoj regulaciji apoptoze mezenhimskih stanica te kod osjetilne percepcije zvuka. Gen CYLC2 kodira protein koji je strukturni dio citoskeleta, a gen RPS6KA6 sudjeluje u fosforilaciji proteina te u odgovoru na oštećenja DNA i prijenosu signala. Gen HDX nema naznačenu neku posebnu funkciju osim regulacije transkripcije.

Prema Zhu i sur. (2020.), signali selekcije u području SLC16A, RLIM, RBBP4, KIAA2022, ABCB7, MAGT1, ATRX, FGF16, HMGN, CHM i DACH2 (13) gena, pronađeni su kod kineskih autohtonih pasmina. Sve navedeno ukazuje da postoji niz relevantnih gena unutar ove regije, kao i u prošlom signalu, no ne može se u potpunosti razjasniti prisutnost ovog signala u populaciji.

4.1.1.4. Usporedba

Usporedbom signala pozitivne selekcije identificiranih iHS metodom, sa signalima identificiranim eROHi metodom na spolnom kromosomu X kod ženskih jedinki, nije se preklapala niti jedna regija. Međutim, ne smije se zanemariti da u ovoj analizi iHS metodom na spolnom kromosomu X nije bilo moguće procijeniti vrijednost kod 701 SNP-ova. Kao i kod

autosoma, za precizniju i pouzdaniju analizu potrebno je povećati uzorak te uključiti još metoda identifikacije signala pozitivne selekcije. Nadam se da će ovaj rad poslužiti kao polazna točka za buduće napore u identifikaciji signala pozitivne selekcije na spolnom kromosomu X te da će identificirani autosomalni i X kromosomalni signali doprinijeti budućim istraživanjima prilikom usporedbe vlastitih rezultata, s ciljem postizanja boljeg (točnijeg) zaključka.

6. ZAKLJUČCI

1. Nakon kontrole kvalitete, uzorak za identifikaciju signala pozitivne selekcije na autosomima, sastojao se od 101 jedinke (Dalmatinska pramenka; n = 25, Dubrovačka ruda; n = 26, Istarska ovca; n = 25, Paška ovca; n = 25) i 457 429 SNP-ova.
2. Nasuprot tome, za identifikaciju signala pozitivne selekcije na spolnom kromosomu X, nakon izbacivanja muških jedinki iz očišćenog seta, uzorak se sastojao od 59 ženskih jedinki (Dalmatinska pramenka; n = 13, Dubrovačka ruda; n = 20, Istarska ovca; n = 13, Paška ovca; n = 13) i 21 748 SNP-ova.
3. iHS metodom na autosomima je identificirano pet signala, a eROHi metodom šest signala pozitivne selekcije, uz pripadajućih 104, odnosno 110 gena.
4. iHS metodom na spolnom kromosomu X je identificiran jedan signal, a eROHi metodom dva signala pozitivne selekcije, uz pripadajućih tri, odnosno 102 gena.
5. Izvršena je funkcionalna analiza gena unutar signala; uočeni su geni koji utječu na mnoga ekonomski i evolucijski važna svojstva te je tako razjašnjena (više ili manje) prisutnost pojedinih signala u populaciji.
6. Promatrajući komparativno rezultate iHS i eROHi metoda, identificirana su dva signala pozitivne selekcije; OAR 2: 121.34 Mb – 123.95 Mb i OAR 6: 32.71 Mb – 39.18 Mb.
7. Identifikacija signala pozitivne selekcije koristeći komparativno iHS i eROHi metodu može promaknuti pojedine signale (visoko frekventne) zbog nedostataka iHS metode te je zato za precizniju i pouzdaniju analizu potrebno uključiti još metoda identifikacije, a potom i povećati uzorak.
8. Ovaj rad će potencijalno poslužiti kao polazna točka za buduće napore u identifikaciji signala pozitivne selekcije na spolnom kromosomu X te će identificirani signali (na autosomima i X kromosomu) doprinijeti budućim istraživanjima pri usporedbi vlastitih rezultata.

7. POPIS LITERATURE

1. Abo-Ismail, M. K., Lansink, N., Akanno, E., Karisa, B. K., Crowley, J. J., Moore, S. S., ..., Plastow, G. S. (2018). Development and validation of a small SNP panel for feed efficiency in beef cattle. Journal of animal science 96(2), 375–397.
2. Bahcall, O. (2014). Sheep genome. Nature Genetics 46, 665.
3. Bautista, D. M., Siemens, J., Glazer, J. M., Tsuruda, P. R., Basbaum, A. I., Stucky, C. L., ..., Julius, D. (2007). The menthol receptor TRPM8 is the principal detector of environmental cold. Nature 448(7150), 204–208.
4. Begun, D. J., Whitley, P. (2000). Reduced X-linked nucleotide polymorphism in Drosophila simulans. Proceedings of the National Academy of Sciences of the United States of America 97(11), 5960–5965.
5. Berry, A. J., Ajioka, J. W., Kreitman, M. (1991). Lack of polymorphism on the Drosophila fourth chromosome resulting from selection. Genetics 129(4), 1111–1117.
6. Biscarini, F., Cozzi, P., Gaspa, G., Marras, G. (2018). detectRUNS: detect runs of homozygosity and runs of heterozygosity in diploid genomes.
7. Carneiro, M., Rubin, C. J., Di Palma, F., Albert, F. W., Alföldi, J., Barrio, A. M., ..., Andersson, L. (2014). Rabbit genome analysis reveals a polygenic basis for phenotypic change during domestication. Science 345(6200), 1074–1079.
8. Casto, A. M., Li, J. Z., Absher, D., Myers, R., Ramachandran, S., Feldman, M. W. (2010). Characterization of X-linked SNP genotypic variation in globally distributed human populations. Genome biology 11(1), 1–19.
9. Chang, C. C., Chow, C. C., Tellier, L. C., Vattikuti, S., Purcell, S. M., Lee, J. J. (2015). Second-generation PLINK: rising to the challenge of larger and richer datasets. Gigascience 4(1), s13742–015.
10. Chen, Z. H., Zhang, M., Lv, F. H., Ren, X., Li, W. R., Liu, M. J., ..., Li, M. H. (2018). Contrasting patterns of genomic diversity reveal accelerated genetic drift but reduced directional selection on X-chromosome in wild and domestic sheep species. Genome biology and evolution 10(5), 1282–1297.

11. Ciani, E., Mastrangelo, S., Da Silva, A., Marroni, F., Ferenčaković, M., Ajmone-Marsan, P., ..., Lenstra, J. A. (2020). On the origin of European sheep as revealed by the diversity of the Balkan breeds and by optimizing population-genetic analysis tools. Genetics Selection Evolution 52, 1–14.
12. Curik, I., Ferenčaković, M., Sölkner, J. (2014). Inbreeding and runs of homozygosity: A possible solution to an old problem. Livestock Science 166, 26–34.
13. Delaneau O., Coulonges C., Zagury J. (2008). Shape-IT: new rapid and accurate algorithm for haplotype inference. BMC Bioinformatics 9, 540.
14. Demars, J., Cano, M., Drouilhet, L., Plisson-Petit, F., Bardou, P., Fabre, S., ..., Allain, D. (2017). Genome-wide identification of the mutation underlying fleece variation and discriminating ancestral hairy species from modern woolly sheep. Molecular biology and evolution 34(7), 1722–1729.
15. Disteche, C. M. (2006). Dosage compensation of the active X chromosome in mammals. Nature genetics 38(1), 47–53.
16. Fariello, M. I., Servin, B., Tosser-Klopp, G., Rupp, R., Moreno, C., San Cristobal, M., ..., International Sheep Genomics Consortium. (2014). Selection signatures in worldwide sheep populations. PloS one 9(8), e103813.
17. Fay, J. C., Wu, C. I. (2000). Hitchhiking under positive Darwinian selection. Genetics 155(3), 1405–1413.
18. Ferenčaković, M., Sölkner, J., Čurik, I. (2013b). Estimating autozygosity from highthroughput information: effects of SNP density and genotyping errors. Genetics Selection Evolution 45, 42.
19. Gambino, F., Kneib, M., Pavlowsky, A., Skala, H., Heitz, S., Vitale, N., ..., Humeau, Y. (2009). IL1RAPL1 controls inhibitory networks during cerebellar development in mice. European Journal of Neuroscience 30(8), 1476–1486.
20. Gautier, M., Vitalis, R. (2012). rehh: an R package to detect footprints of selection in genome-wide SNP data from haplotype structure. Bioinformatics 28(8), 1176–1177.
21. Gibson, J., Morton, N. E., Collins, A. (2006). Extended tracts of homozygosity in outbred human populations. Human Molecular Genetics 15, 789–95.
22. Gorssen, W., Meyermans, R., Janssens, S., Buys, N. (2021). A publicly available repository of ROH islands reveals signatures of selection in different livestock and pet species. Genetics Selection Evolution 53(1), 1–10.

23. Graves, J. A. M. (2006). Sex chromosome specialization and degeneration in mammals. Cell 124(5), 901–914.
24. Gutiérrez-Gil, B., Arranz, J. J., Pong-Wong, R., García-Gámez, E., Kijas, J., Wiener, P. (2014). Application of selection mapping to identify genomic regions associated with dairy production in sheep. PloS one 9(5), e94623.
25. Han, J., Kraft, P., Nan, H., Guo, Q., Chen, C., Qureshi, A., ..., Hunter, D. J. (2008). A genome-wide association study identifies novel alleles associated with hair color and skin pigmentation. PLoS genetics 4(5), e1000074.
26. Harris, A. M., Garud, N. R., DeGiorgio, M. (2018). Detection and classification of hard and soft sweeps from unphased genotypes by multilocus genotype identity. Genetics 210(4), 1429–1452.
27. He, S., Fu, J., Jiang, W., Cheng, Y., Chen, J., Guo, Z. (2020). BloTISRT: Blockchain-based Threat Intelligence Sharing and Rating Technology. In Proceedings of the 2020 International Conference on Cyberspace Innovation of Advanced Technologies (pp. 524-534).
28. Hedrick, P. W. (2007). Sex: differences in mutation, recombination, selection, gene flow, and genetic drift. Evolution 61(12), 2750–2771.
29. Hermisson, J., Pennings, P. S. (2005). Soft sweeps: molecular population genetics of adaptation from standing genetic variation. Genetics 169(4), 2335–2352.
30. Heyer, E., Segurel, L. (2010). Looking for signatures of sex-specific demography and local adaptation on the X chromosome. Genome Biology 11(1), 203.
31. Jensen P. (2014). Behavior genetics and the domestication of animals. Annual Review of Animal Biosciences 2, 85–104.
32. Jensen, J. K., Welk, B. A., Williams, R. E. A., Sosa, J. M., Huber, D. E., Senkov, O. N., ..., Fraser, H. L. (2016). Characterization of the microstructure of the compositionally complex alloy Al1Mo0. 5Nb1Ta0. 5Ti1Zr1. Scripta Materialia 121, 1–4.
33. Keinan, A, Reich, D. (2010). Can a sex-biased human demography account for the reduced effective population size of chromosome X in nonAfricans? Molecular Biology and Evolution 27(10), 2312–2321.
34. Kerje, S., Sharma, P., Gunnarsson, U., Kim, H., Bagchi, S., Fredriksson, R., ..., Andersson, L. (2004). The Dominant white, Dun and Smoky color variants in chicken

are associated with insertion/deletion polymorphisms in the PMEL17 gene. Genetics 168(3), 1507–1518.

35. Kijas, J. W., Lenstra, J. A., Hayes, B., Boitard, S., Porto Neto, L. R. P., San Cristobal, M. i sur. (2012). Genome-wide analysis of the world's sheep breeds reveals high levels of historic mixture and strong recent selection. PLoS biology 10(2), e1001258.
36. Kijas, J. W., Porto-Neto, L., Dominik, S., Reverter, A., Bunch, R., McCulloch, R., ..., International Sheep Genomics Consortium. (2014). Linkage disequilibrium over short physical distances measured in sheep using a high-density SNP chip. Animal genetics 45(5), 754–757.
37. Kim, E. S., Cole, J. B., Huson, H., Wiggans, G. R., Van Tassell, C. P., Crooker, B. A., ..., Sonstegard, T. S. (2013). Effect of artificial selection on runs of homozygosity in US Holstein cattle. PloS one 8(11), e80813.
38. Kim, E. S., Elbeltagy, A. R., Aboul-Naga, A. M., Rischkowsky, B., Sayre, B., Mwacharo, J. M., Rothschild, M. F. (2016). Multiple genomic signatures of selection in goats and sheep indigenous to a hot arid environment. Heredity 116(3), 255–264.
39. Leal-Gutiérrez, J. D., Elzo, M. A., & Mateescu, R. G. (2020). Identification of eQTLs and sQTLs associated with meat quality in beef. BMC genomics 21(1), 1-15.
40. Lewontin, R. C., Krakauer, J. (1973). Distribution of gene frequency as a test of the theory of the selective neutrality of polymorphisms. Genetics 74(1), 175–195.
41. Li, M. H., Tiirikka, T., Kantanen, J. (2014). A genome-wide scan study identifies a single nucleotide substitution in ASIP associated with white versus non-white coat-colour variation in sheep (Ovis aries). Heredity 112(2), 122–131.
42. Liu, Z., Ji, Z., Wang, G., Chao, T., Hou, L., Wang, J. (2016). Genome-wide analysis reveals signatures of selection for important traits in domestic sheep from different ecoregions. BMC genomics 17(1), 1–14.
43. Lv, F. H., Agha, S., Kantanen, J., Colli, L., Stucki, S., Kijas, J. W., ..., Ajmone Marsan, P. (2014). Adaptations to climate-mediated selective pressures in sheep. Molecular biology and evolution 31(12), 3324–3343.
44. Ma, Y., Zhang, H., Zhang, Q., Ding, X. (2014). Identification of selection footprints on the X chromosome in pig. PloS one 9(4), e94911.

45. Manunza, A., Cardoso, T. F., Noce, A., Martínez, A., Pons, A., Bermejo, L. A., ..., Amills, M. (2016). Population structure of eleven Spanish ovine breeds and detection of selective sweeps with BayeScan and hapFLK. Scientific reports 6(1), 1–10.
46. Marras, G., Gaspa, G., Sorbolini, S., Dimauro, C., Ajmone-Marsan, P., Valentini, A., ..., Macciotta, N. P. (2015). Analysis of runs of homozygosity and their relationship with inbreeding in five cattle breeds farmed in Italy. Animal genetics 46(2), 110–121.
47. Marwal, A., Sahu, A. K., Gaur, R. K. (2014). Molecular Markers. Animal Biotechnology, 289–305.
48. Mastrangelo, S., Tolone, M., Sardina, M. T., Sottile, G., Sutera, A. M., Di Gerlando, R., Portolano, B. (2017). Genome-wide scan for runs of homozygosity identifies potential candidate genes associated with local adaptation in Valle del Belice sheep. Genetics Selection Evolution 49(1), 1–10.
49. Mastrangelo, S., Bahbahani, H., Moioli, B., Ahbara, A., Al Abri, M., Almathen, F., ... & Ciani, E. (2019). Novel and known signals of selection for fat deposition in domestic sheep breeds from Africa and Eurasia. PLoS One 14(6), e0209632.
50. McQuillan, R., Leutenegger, A. L., Abdel-Rahman, R., Franklin, C. S., Pericic, M., Barac-Lauc, L., ..., Wilson, J. F. (2008). Runs of homozygosity in European populations. The American Journal of Human Genetics 83(3), 359–372.
51. Mcrae, K. M., McEwan, J. C., Dodds, K. G., Gemmell, N. J. (2014). Signatures of selection in sheep bred for resistance or susceptibility to gastrointestinal nematodes. BMC Genomics 15(1), 1–13.
52. McVicker, G., Gordon, D., Davis, C., Green, P. (2009). Widespread genomic signatures of natural selection in hominid evolution. PLoS Genetics 5(5), e1000471.
53. Medugorac, I., Graf, A., Grohs, C., Rothammer, S., Zagdsuren, Y., Gladyr, E., ..., Capitan, A. (2017). Whole-genome analysis of introgressive hybridization and characterization of the bovine legacy of Mongolian yaks. Nature genetics 49(3), 470.
54. Mészáros, G., Boison, S. A., Pérez O´Brian, A. M., Ferenčaković, M., Curik, I., da Silva, M. V. B., Utsunoomiya, Y. T., Garcia, J. F., Solkner, J. (2015). Genomic analysis for managing small and endangered populations: a case study in Tyrol Grey cattle. Frontiers in Genetics 6, 173.
55. Moioli, B., Scatà, M. C., Steri, R., Napolitano, F., Catillo, G. (2013). Signatures of selection identify loci associated with milk yield in sheep. BMC genetics 14(1), 1–7.

56. Moradi, M. H., Nejati-Javaremi, A., Moradi-Shahrbabak, M., Dodds, K. G., McEwan, J. C. (2012). Genomic scan of selective sweeps in thin and fat tail sheep breeds for identifying of candidate regions associated with fat deposition. BMC genetics 13(1), 1–15.
57. Morandi, E. M., Verstappen, R., Zwierzina, M. E., Geley, S., Pierer, G., Ploner, C. (2016). ITGAV and ITGA5 diversely regulate proliferation and adipogenic differentiation of human adipose derived stem cells. Scientific reports 6(1), 1–14.
58. Myojin, R., Kuwahara, S., Yasaki, T., Matsunaga, T., Sakurai, T., Kimura, M., ..., Kurihara, Y. (2004). Expression and functional significance of mouse paraspeckle protein 1 on spermatogenesis. Biology of reproduction 71(3), 926-932.
59. Pierce, B. A. (2012). Genetics: a conceptual approach. Macmillan.
60. 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. Genetics Selection Evolution 46(1), 1–14.
61. Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A., Bender, D., ..., Sham, P. C. (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. The American journal of human genetics 81(3), 559–575.
62. Purfield, D., Berry, D., McParland, S., Bradley, D. (2012). Runs of homozygosity and population history in cattle. BMC Genetics 13, 70.
63. Purfield, D. C., McParland, S., Wall, E., Berry, D. P. (2017). The distribution of runs of homozygosity and selection signatures in six commercial meat sheep breeds. PLoS One 12(5), e0176780.
64. Qanbari, S., Simianer, H. (2014). Mapping signatures of positive selection in the genome of livestock. Livestock Science 166, 133-143.
65. R Core Team (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing.
66. Saravanan, K. A., Panigrahi, M., Kumar, H., Bhushan, B., Dutt, T., Mishra, B. P. (2021). Genome-wide analysis of genetic diversity and selection signatures in three Indian sheep breeds. Livestock Science 243, 104367.
67. Schaffner, S. F. (2004). The X chromosome in population genetics. Nature Reviews Genetics 5, 43–51.

68. Smith, J. M., Haigh, J. (1974). The hitch-hiking effect of a favourable gene. Genetics Research 23(1), 23–35.
69. Solé, M., Gori, A. S., Faux, P., Bertrand, A., Farnir, F., Gautier, M., Druet, T. (2017). Age-based partitioning of individual genomic inbreeding levels in Belgian Blue cattle. Genetics Selection Evolution 49(1).
70. Stainton, J. J., Haley, C. S., Charlesworth, B., Kranis, A., Watson, K., Wiener, P. (2015). Detecting signatures of selection in nine distinct lines of broiler chickens. Animal Genetics 46(1), 37–49.
71. Stephan, W. (2016). Signatures of positive selection: from selective sweeps at individual loci to subtle allele frequency changes in polygenic adaptation. Molecular ecology 25(1), 79–88.
72. Sweet-Jones, J., Lenis, V. P., Yurchenko, A. A., Yudin, N., Swain, M., & Larkin, D. (2021). Genotyping and Whole-Genome Resequencing of Welsh Sheep Breeds Reveal Candidate Genes and Variants for Adaptation to Local Environment and Socioeconomic Traits. Frontiers in Genetics 12, 903.
73. Šimpraga, M., Mioč, B., Vukotić, B., Radin, L. (2008). Hrvatske izvorne pasmine ovaca, osnova ekološke ovčarske proizvodnje-1. dio. Ekološka poljoprivreda; ovčarstvo u Hrvatskoj. Veterinarska stanica 39(1), 9–17.
74. Tao, L., He, X. Y., Wang, F. Y., Pan, L. X., Wang, X. Y., Gan, S. Q., ..., Chu, M. X. (2021). Identification of genes associated with litter size combining genomic approaches in Luzhong mutton sheep. Animal Genetics.
75. Turner S. D. (2018). qqman: an R package for visualizing GWAS results using Q-Q and manhattan plots. Journal of Open Source Software 3(25), 731.
76. Vicoso, B, Charlesworth, B. (2006). Evolution on the X chromosome: unusual patterns and processes. Nature Reviews Genetics 7(8), 645–653.
77. 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.
78. Voight, B. F., Kudaravalli, S., Wen, X., Pritchard, J. K. (2006). A map of recent positive selection in the human genome. PLoS Biology 4, e72.

79. Wang, K., Li, W. D., Zhang, C. K., Wang, Z., Glessner, J. T., Grant, S. F., ..., Price, R. A. (2011). A genome-wide association study on obesity and obesity-related traits. PloS one 6(4), e18939.
80. Yurchenko, A. A., Deniskova, T. E., Yudin, N. S., Dotsev, A. V., Khamiruev, T. N., Selionova, M. I., ..., Larkin, D. M. (2019). High-density genotyping reveals signatures of selection related to acclimation and economically important traits in 15 local sheep breeds from Russia. BMC genomics 20(3), 294.
81. Zhu, C., Fan, H., Yuan, Z., Hu, S., Zhang, L., Wei, C., ..., Du, L. (2015). Detection of selection signatures on the X chromosome in three sheep breeds. International journal of molecular sciences 16(9), 20360–20374.
82. Zhu, C., Li, M., Qin, S., Zhao, F., Fang, S. (2020). Detection of copy number variation and selection signatures on the X chromosome in Chinese indigenous sheep with different types of tail. Asian-Australasian journal of animal sciences 33(9), 1378.