

# Identifikacija signala pozitivne selekcije u genomu pastirskih pasa

---

**Glavaš, Vjeran**

**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://um.nsk.hr/um:nbn:hr:204:518213>

*Rights / Prava:* [In copyright](#) / [Zaštićeno autorskim pravom.](#)

*Download date / Datum preuzimanja:* **2024-08-26**



*Repository / Repozitorij:*

[Repository Faculty of Agriculture University of Zagreb](#)



SVEUČILIŠTE U ZAGREBU  
AGRONOMSKI FAKULTET

**Identifikacija signala pozitivne selekcije u genomu  
pastirskih pasa**

DIPLOMSKI RAD

Vjeran Glavaš

Zagreb, Rujan, 2021.

**SVEUČILIŠTE U ZAGREBU  
AGRONOMSKI FAKULTET**

Diplomski studij:

Genetika i oplemenjivanje životinja

**Identifikacija signala pozitivne selekcije u genomu  
pastirskih pasa**

DIPLOMSKI RAD

Vjeran Glavaš

Mentor:

Izv. prof. dr. sc. Vlatka Čubrić Čurik

Zagreb, Rujan, 2021.

**SVEUČILIŠTE U ZAGREBU  
AGRONOMSKI FAKULTET**

**IZJAVA STUDENTA  
O AKADEMSKOJ ČESTITOSTI**

Ja, **Vjeran Glavaš**, JMBAG 0178110797, rođen/a 29.12.1997. u Osijek, izjavljujem da sam samostalno izradila/izradio diplomski rad pod naslovom:

**Identifikacija signala pozitivne selekcije u genomu pastirskih pasa**

Svojim potpisom jamčim:

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

U Zagrebu, dana \_\_\_\_\_

\_\_\_\_\_  
*Potpis studenta / studentice*

SVEUČILIŠTE U ZAGREBU  
AGRONOMSKI FAKULTET

**IZVJEŠĆE**

O OCJENI I OBRANI DIPLOMSKOG RADA

Diplomski rad studenta/ice **Vjeran Glavaš**, JMBAG 0178110797, naslova

**Identifikacija signala pozitivne selekcije u genomu pastirskih pasa**

obranjen je i ocijenjen ocjenom \_\_\_\_\_, dana \_\_\_\_\_.

Povjerenstvo:

potpisi:

1. izv.prof.dr.sc. Vlatka Čubrić Čurik mentor

\_\_\_\_\_

2. prof.dr.sc. Ino Čurik član

\_\_\_\_\_

3. doc.dr.sc. Maja Ferenčaković član

\_\_\_\_\_

## Zahvala

Ovim putem se zahvaljujem svojoj mentorici izv.prof.dr.sc. Vlatki Čubrić-Čurik na mentorstvu i ukazanoj pomoći tijekom pisanja ovog diplomskog rada, na svim informacijama i na svim savjetima koje je dijelila ne samo samnom, već sa svim studenticama i studentima na studiju genetike i oplemenjivanja. Od srca se želim zahvaliti i doc.dr.sc Maji Ferenčaković na nesebičnoj pomoći u provedbi analiza potrebnih za izradu rada kao i na svim konzultacijama koje je održala bez obzira na prepuni raspored. Zahvaljujem se i prof.dr.sc. Ini Čuriku na svim savjetima, pomoći i profesionalnosti koju je profesor pokazao u procesu nastanka rada. Svima prethodno spomenutima se također zahvaljujem i na strpljenju, pozitivnosti, smirenosti i na svemu što su me naučili u prethodnim godinama studiranja, svaki savjet i svaka kritika oblikovala je moje znanje i proširila moj pogled, kako na svijet tako i na obrazovanje i genetiku kao znanost. Od srca sam zahvalan!

Također se zahvaljujem i svim ostalim profesoricama, profesorima i mentorima koje sam imao čast slušati tijekom školovanja i uz čiju sam pomoć naučio puno korisnih stvari. Posebno bih se htio zahvaliti svojoj obitelji za svu podršku koju su mi pružili tijekom ovih pet godine studiranja i za svaki savjet koji su bili spremni dati. Od početka školovanja su mi pomogli da ostvarim sve što sam htio i uvijek bili uz mene u slučaju da je nešto pošlo u krivom smjeru, također im se zahvaljujem na financijskoj pomoći koju su mi pružili, bez koje ništa od ovoga ne bi bilo moguće. Hvala Vam!

Zahvalio bih se i svim kolegicama i kolegama, posebice kolegicama V. Barukčić, L. Bradić, N. Kovačević, V. Perišić, B. Reponj i K. Turkalj te kolegi M. Shihabi-u, za sve pozitivne trenutke koje smo imali tijekom školovanja i za svaku podijeljenu bilješku i za svako zajedničko učenje koje mi je pomoglo u ostvarivanju što boljih rezultata na ispitima. Nadam se da ću i u daljnjem obrazovanju i životu imati ovakvu potporu i još jednom se od srca zahvaljujem svima što su mi pružili ovo nezaboravno iskustvo.

# Sadržaj

Sažetak .....	1
Summary .....	2
<b>1. Uvod .....</b>	<b>3</b>
<b>2. Pregled literature .....</b>	<b>4</b>
2.1. Selekcija .....	4
2.2. Runs of homozygosity (ROH) .....	5
2.3. Integrated haplotype score (iHS) .....	6
2.4. Pastirske pasmine pasa .....	8
2.4.1. Balkanske pasmine pastirskih pasa .....	9
<b>2.4.1.1. Tornjak.....</b>	<b>9</b>
<b>2.4.1.2. Šarplaninac .....</b>	<b>11</b>
<b>2.4.1.3. Kraški ovčar/Kraševac .....</b>	<b>11</b>
<b>3. Cilj istraživanja .....</b>	<b>13</b>
<b>4. Materijali i metode .....</b>	<b>14</b>
4.1. iHS analiza .....	14
4.2. ROH analiza .....	15
4.3. Identifikacija signala selekcije.....	16
4.4. Anotacija gena unutar regija signala pozitivne selekcije.....	16
<b>5. Rezultati .....</b>	<b>17</b>
5.1. Rezultati iHS analize .....	17
<b>5.1.1. CanFam 1: 57.27 Mb – 61.23 Mb .....</b>	<b>18</b>
<b>5.1.2. CanFam 6: 4.07 Mb – 4.43 Mb .....</b>	<b>19</b>
<b>5.1.3. CanFam 17: 5.83 Mb – 5.84 Mb .....</b>	<b>20</b>
<b>5.1.4. CanFam 22: 6.13 Mb – 6.66 Mb .....</b>	<b>21</b>
<b>5.1.5. CanFam 25: 3.87 Mb – 4.80 Mb .....</b>	<b>22</b>
<b>5.1.6. CanFam 30: 30.91 Mb – 31.26 Mb .....</b>	<b>23</b>
5.2. Rezultati eROHi analiza .....	23
<b>5.2.1. CanFam 1: 2.36 Mb – 3.32 Mb .....</b>	<b>24</b>
<b>5.2.2. CanFam 1: 60.93 Mb – 61.02 Mb .....</b>	<b>25</b>
<b>5.2.3. CanFam 3: 31.09 Mb – 32.14 Mb .....</b>	<b>25</b>
<b>5.2.4. CanFam 5: 1.08 Mb – 2.67 Mb .....</b>	<b>26</b>
<b>5.2.5. CanFam 6: 38.13 Mb – 41.07 Mb .....</b>	<b>27</b>
<b>5.2.6. CanFan 6: 23.51 Mb – 24.40 Mb .....</b>	<b>27</b>
<b>5.2.7. CanFam 7: 1.38 Mb – 2.14 Mb .....</b>	<b>27</b>
<b>5.2.8. CanFam 8: 126.54 Kb – 3.13 Mb.....</b>	<b>28</b>
<b>5.2.9. CanFam 10: 525.75 Kb – 3.89 Mb.....</b>	<b>29</b>

5.2.10.	CanFam 10: 15.74 Mb – 17.45 Mb .....	29
5.2.11.	CanFam 11: 8.89 Mb – 10.09 Mb .....	29
5.2.12.	CanFam 14: 1.01 Mb – 2.98 Mb .....	30
5.2.13.	CanFam 15: 1.16 Mb – 2.15 Mb .....	31
5.2.14.	CanFam 16: 692.73 Kb – 2.6 Mb .....	32
5.2.15.	CanFam 17: 229.18 Kb – 1.68 Mb .....	33
5.2.16.	CanFam 18: 954.32 Kb – 4.89 Mb .....	34
5.2.17.	CanFam 20: 581.21 Kb – 2.68 Mb .....	34
5.2.18.	CanFam 21: 269.07 Kb – 2.82 Mb .....	35
5.2.19.	CanFam 22: 1.31 Mb – 3.22 Mb .....	36
5.2.20.	CanFam 25: 223.81 Kb – 3.94 Mb .....	36
5.2.21.	CanFam 30: 927.94 Kb – 2.22 Mb .....	37
5.2.22.	CanFam 30: 5.09 Mb – 6.44 Mb .....	37
5.2.23.	CanFam 34: 917.49 Kb – 1.50 Mb .....	38
6.	Rasprava .....	39
6.1.	Funkcionalnost pojedinih gena identificirani unutar iHS analiza .....	39
6.1.1.	Geni povezani sa unutar staničnim mehanizmima .....	39
6.1.2.	Geni povezani s anatomskim i histološkim razvojem organizma... ..	40
6.1.3.	Geni povezani s reprodukcijom .....	41
6.2.	Funkcionalnost pojedinih gena identificiranih unutar eROHi analiza .....	41
6.2.2.	Geni povezani s reprodukcijom .....	43
6.2.3.	Geni povezani s razvojem osjetila .....	44
6.2.4.	Geni povezani s anatomskim i histološkim razvojem organizma i njegovim normalnim funkcioniranjem .....	45
6.2.5.	Ostali geni .....	47
7.	Zaključci .....	49
8.	Popis literature .....	50
9.	Prilog A .....	57
10.	Prilog B .....	63
11.	Životopis .....	67



# Sažetak

Diplomskog rada studenta/ice **Vjeran Glavaš**, naslova

## **Identifikacija signala pozitivne selekcije u genomu pastirskih pasa**

Umjetna selekcija koja se provodi na psima (*Canis lupus familiaris*) predstavlja jedan od najzanimljivijih primjera selekcije, kako zbog velikog broja jedinki tako i zbog širok spektra varijacija unutar specifičnih morfoloških i bihevioralnih karakteristika pojedine pasmine. U ovom istraživanju korišteni su genomske podaci 123 jedinke pastirskih pasmina pasa s ukupno 107.403 SNP-ova na kojima su provedene iHS i eROHi analize u svrhu otkrivanja signala pozitivne selekcije. Metode su se temeljile na otkrivanju genomske varijacije unutar jedne velike meta-populacije od 11 različitih pastirskih pasmina pasa. Provedeno istraživanje obuhvatilo je i identifikaciju gena kao i proučavanje njihove funkcionalnosti unutar psećeg organizma. Uspješno je identificirano 29 regija pod selekcijom s pripadajućih 462 gena koji utječu na ekspresiju važnih morfoloških i bihevioralnih karakteristika.

**Ključne riječi:** selekcija, iHS, eROHi, varijacije, geni

# Summary

Of the master's thesis – student **Vjeran Glavaš**, entitled

## **Identifying signatures of positive selection in shepherd dogs' genome**

Artificial selection in dogs (*Canis lupus familiaris*) is one of the most interesting examples of selection, both because of the large number of individuals and the wide range of variation within the specific morphological and behavioral characteristics of each breed. In this study, genomic data from 123 individuals of shepherd dog breeds with a total of 107,403 SNPs was used, on which iHS and eROHi analyzes were performed to detect positive selection signals. The methods were based on the detection of genomic variation within a large meta-population of 11 different shepherd dog breeds. The studies performed included both the identification of genes and the study of their functionality within the canine organism. Twenty-nine regions that are under positive selection were identified as well as 462 genes that play a crucial role in development of certain morphological and behavioral characteristics.

**Keywords:** Selection, iHS, eROHi, Variations, Genes

# 1. Uvod

Intenzivan uzgojno selekcijski rad ljudi, kroz duži vremenski period i tijekom procesa formiranja pojedinih životinjskih pasmina, za posljedicu je imao kreiranje specifičnih regija, odnosno segmenata unutar njihova genoma koje možemo koristiti i okarakterizirati kao signale selekcije. U kojoj mjeri pozitivna selekcija doprinosi molekularnoj evoluciji, jedno je od važnijih pitanja evolucijske genetike (Booker i sur. 2017.). Klasična paradigma u modernoj genetici jest takozvana *neutralna teorija* koja tvrdi da je većina molekularnih promjena posljedica genetskog drifta, dok pozitivna selekcija igra manju ulogu u nastanku istih (Kimura, 1983.). Mogućnost identifikacije i vizualizacije signala pozitivne selekcije ključan je korak unutar populacijske genetike, ali i evolucijske biologije jer omogućava lakše kontroliranje i proučavanje pojedinih vrsta, pasmina i populacija. Identifikaciju signala selekcije moguće je obaviti pomoću raznih analiza genomskih podataka jedinki. Jedne od takvih analiza su i intra-populacijske analize u koje ubrajamo ekstremne otoke kontinuiranih segmenata homozigotnih nizova (eng: *Extreme ROH islands*; eROHi) i Integrirani rezultat haplotipova (eng: *Integrated haplotype score*; iHS), koje su provedene u svrhu pisanja ovog rada. Spomenute metode omogućuju otkrivanje regija koje nazivamo otocima kontinuiranih homozigotnih segmenata (eng: *Runs of homozygosity islands*; ROHi) kao i dokaze nedavne pozitivne selekcije unutar različitih lokusa.

Uzmemo li u obzir da je prvotna zadaća psa bila lov, a druga zadaća mu je bila čuvanje nastambe/članova obitelji/zajednice, treća je svakako bila čuvanje stada odnosno stoke (Jeger, 2017.). Pasmine koje su se koristile u tu svrhu, kao na primjer i neke balkanske pasmine u koje ubrajamo tornjaka, šarplaninaca i kraškog ovčara, morale su posjedovati specifične karakteristike na kojima se i temeljio daljnji selektivni uzgoj. Neke od tih karakteristika su krupna glava, moćne čeljusti, dobroćudnost, krupnoća tijela i sposobnost preživljavanja u izoliranim planinskim područjima pri raznim vremenskim uvjetima. Identifikacija signala pozitivne selekcije na pojedinim segmentima, odnosno unutar specifičnih genomskih regija, gdje se nalaze i geni s utjecajem na ekspresiju spomenutih svojstava, može služiti kao dokaz provođenja pozitivne selekcije kroz duži vremenski period na pojedinim pasminama pastirskih pasa. Takva saznanja mogu nam pomoći u daljnjem uzgojno selekcijskom radu, ali i u eventualnom poboljšavanju pojedinih karakteristika kod različitih pasmina.

## 2. Pregled literature

### 2.1. Selekcija

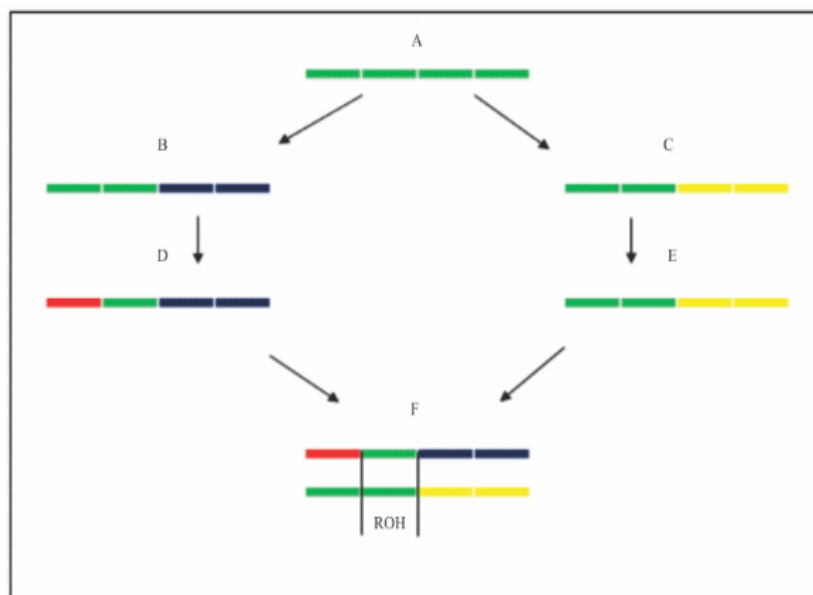
Selekcija (lat. *selectio*) podrazumijeva odabiranje ili izdvajanje jedinki odnosno pojedinaca s najboljim i najprikladnijim osobinama, za unaprijed postavljenu svrhu (Leksikografski zavod Miroslav Krleža, 2021.). Selekcija može biti prirodna ili potaknuta ljudskim radom (umjetna selekcija). Kada je riječ o umjetnoj selekciji, ona se provodi čovjekovim odabiranjem jedinki koje su u određenim karakteristikama superiornije u odnosu na ostale jedinke iz iste populacije, kao roditelje sljedećih generacija. Takav odabir jedinki temelji se na vrijednostima kvalitativnih i kvantitativnih svojstava kod istih. Cilj umjetne selekcije jest uz pomoću odabranih superiornijih jedinki, kreirati generaciju unutar koje će ostvareno potomstvo posjedovati poboljšane vrijednosti svojstava na kojima se temeljio odabir roditelja, odnosno ostvariti određeni udio genetskog napretka. Kvantitativna svojstva jesu svojstva na čiju ekspresiju djeluje veliki broj gena manjeg pojedinačnog utjecaja, okolina u kojoj životinja obitava kao i interakcija okolišnih čimbenika i genetske komponente jedinke (genotipa), dok ekspresiju kvalitativnih svojstava kod jedinke kontrolira manji broj gena većeg pojedinačnog utjecaja, a okolina ima maleni ili ne zamjetan utjecaj. Provođenje umjetne selekcije zahtijeva poznavanje nasljedne osnove i genetske vrijednosti jedinki za pojedino svojstvo koje se želi poboljšati ili ukloniti unutar sljedeće generacije, a važnu ulogu u tome igra uzgojna vrijednost. Uzgojna vrijednost (UV) jedinke definirana je kao suma prosječnih uzgojnih vrijednosti njezinih roditelja ili kao prosječna aditivna genetska vrijednost koju će jedinka naslijediti od oba roditelja. Točnost procjene UV jedinke utječe na izbor životinja i na cjelokupni uspjeh u selekcijskom radu (Jovanovac, 1997.).

Veličina, oblik i ponašanje modernog pripitomljenog psa, posljedica je dugotrajne umjetne selekcije koja traje već 14 000 godina (Akey, 2010.), a izuzetna raznolikost tjelesnih i bihevioralnih karakteristika pasa jedan je od najzanimljivijih primjera pripitomljavanja (Wang i sur. 2013.) (prema Darwin, 1868; Serpell 1996; VonHoldt, B. M. i sur. 2011; Vonholdt, B. M. i sur. 2010.). Selektivni uzgoj pasa podrazumijeva odabir jedinki za rasplod s namjerom da se proizvedu, sačuvaju ili uklone pojedine fizičke karakteristike, bihevioralne karakteristike, zdravstvena stanja ili vještine i sklonosti unutar sljedeće generacije ili da se jedinke specijaliziraju za određene poslove i namjene. Ovisno o namjeni pasmine, jedinke će posjedovati specifične morfološke i bihevioralne karakteristike koje joj omogućavaju lakše i uspješnije izvršavanje primarne svrhe. Sukladno tome, jedinke lovačkih pasmina posjedovat će izraženu brzinu, reflekse, eksplozivnost i vitko tijelo, jedinke pastirskih pasmina kondiciju, izdržljivost i sposobnost lake dresure, a jedinke obiteljskih pasmina će karakterizirati manja građa tijela i određene specifične vanjske karakteristike.

## 2.2. Runs of homozygosity (ROH)

Nizovi homozigotnih segmenata (*eng. Runs of homozygosity*, ROH) predstavljaju kontinuirane nizove homozigotnih genotipova bez heterozigotnosti u diploidnom obliku (Ferenčaković i sur. 2013.). Jednostavan prikaz nastanka ROH segmenta unutar jedinke F, prikazan je na slici 2.2.1. Takvi segmenti prikladni su za proučavanje povijesti populacije, izračun genomskog parenja u srodstvu (*eng. inbreeding*) i za dešifriranje genetske pozadine kompleksnih bolesti i svojstava kao i za identifikaciju gena koji su povezani s ekonomski važnim svojstvima. Ovakva autozigotnost javlja se unutar populacija zbog pojave raznih populacijskih fenomena kao što su parenje u srodstvu, genetski drift, usko grlo populacije ili zbog prirodne i umjetne selekcije (Falconer and Mackay, 1996; Curik i sur. 2014.). Duljina identificiranih ROH segmenata unutar populacije može se koristiti za procjenu trenutka pojave *inbreedinga*. Duži ROH segmenti su najčešće rezultat nedavnog *inbreedinga*, prilikom čega rekombinacije nisu skratile haplotipove naslijeđene od zajedničkoga pretka, dok kratki ROH segmenti s druge strane nastaju kao rezultat povijesnog *inbreedinga* (Ferenčaković i sur. 2013.). Koeficijent genomskog *inbreedinga*  $F_{ROH}$ , prvi puta su definirali McQuillan i sur. (2008). Od 2010. godine analize ROH segmenata postale su standardni pristup u istraživanju *inbreedinga* i proučavanju selekcije unutar populacija, s prvim provedenim studijama na govedima 2010. godine (Sölkner i sur. 2010.), na psima 2010. godine (Boyko i sur. 2010.), kod svinja 2012. godine (Bosse i sur. 2012.) i kod drugih životinja kao što su konji, koze i ovce narednih godina.  $F_{ROH}$  procjenjuje cjelogenomsku autozigotnost uzimajući u obzir utjecaj srodnosti između osnivačkih jedinki populacije, stohastičku prirodu rekombinacija i mutacija kao i potencijalnu pristranost koja proizlazi iz selekcije (Peripolli i sur. 2018.).

Otkrivanje signala selekcije pomoću ROH segmenata temelji se točnije na takozvanim ROHi regijama genoma. Prema Zhang-u (2015.) ROH uzorci nisu nasumično distribuirani unutar genoma, a ROHi se distribuiraju i dijele među jedinkama, što je vjerojatno rezultat selekcijskih događaja. Stoga, ROHi se mogu koristiti za identificiranje signala selekcije, jer genomske regije s istim ROH segmentima najvjerojatnije sadrže alele povezane s genetskim poboljšanjem svojstava jedinki te su u interesu uzgojnih programa. Detekcija ROH segmenata genoma može se obavljati pomoću obzervacijskih metoda, unutar kojih se koriste algoritmi kao PLINK ili SVS, pomoću takozvanih *model based* algoritama u koje se ubrajaju *Hidden Markov Model* (HMM), *Homozygosity heterogenous hidden Markov model* i *Hidden Markov model viterbi* te uz pomoću *Haplotype-matching* algoritmovi odnosno *Germline* algoritama.



Slika 2.2.1. Prikaz nastanka ROH segmenta unutar jedinke F

Izvor: SciELO <https://www.scielo.br/j/pab/a/3V55mLNpV3BffSjk7Pdb7FG/?lang=en#> - pristup (25.06.2021.)

### 2.3. Integrated haplotype score (iHS)

Integrirani rezultat haplotipova (eng: *Integrated haplotype score*; iHS) jest mjera udjela takozvane proširene homozigotnosti haplotipa (eng: *Extended haplotype homozygosity*; EHH) promatranog polimorfizma jednog nukleotida (eng: *single nucleotide polymorphism*; SNP) unutar populacije, prilikom čega se izračun vrši putem ancestralnih i deriviranih alela (Whole-Genome Homozygosity Analysis and Mapping Machina, <http://coruscant.itmat.upenn.edu/whamm/index.html>). Tako se iHS kao mjera uglavnom standardizira prema distribuciji promatranih iHS vrijednosti određenog skupa SNP-ova sa sličnim deriviranim frekvencijama alela. Kao signali pozitivne selekcije (eng. *Selective sweep*) smatraju se haplotipovi s visokim EHH vrijednostima deriviranih alela u odnosu na ancestralne alele. Gautieru i sur. 2017. (preme Sabeti i sur 2002.) EHH vrijednost za alel ciljnog ili fokusnog markera definiraju kao vjerojatnost da su dva nasumično uzorkovana kromosoma, koji nose takav marker i/ili alel, homozigotni unutar određene susjedne kromosomske regije. Ova mjera procjenjuje se iz uzorka, računajući pritom ukupnu homozigotnost kromosomske regije između ciljnog markera i drugog genetskog markera prema formuli (1):

$$EHH_{s,t}^a = \frac{1}{n_a(n_a-1)} \sum_{k=1}^{K_{s,t}^a} n_k(n_k - 1) \quad (1)$$

unutar koje izraz  $n_a$  označava ukupan broj kromosoma na kojima se nalazi ciljni alel  $a$ ,  $K_{s,t}^a$  označava broj različitih haplotipova prisutnih na kromosomima, a  $n_k$  označava broj kromosoma sa zajedničkim haplotipovima  $k$ .

Ukoliko je određeni alel pod selekcijom, odnosno ukoliko se njegova frekvencija u populaciji naglo povećava, unutar neutralnog modela očekuju se veće količine njegovih homozigotnih haplotipova. Vizualizacijom EHH vrijednosti alela pod selekcijom, moguće je primijetiti da će njegova površina ispod krivulje dijagrama biti puno veća u odnosu na druge neutralne alele odnosno alele koji nisu pod selekcijom. Kako bi se takav efekt kod alela primijetio, radi se izračun integrala EHH vrijednosti u udaljenim kromosomskim regijama (u odnosu na ciljni alel) dok one ne dosegnu postavljene granične vrijednosti (Voight i sur. 2006.). Na temelju toga, nestandardizirana iHS vrijednost se računa prema formuli (2):

$$\text{nestandardizirani } iHS = \ln\left(\frac{iHHa}{iHHd}\right) \quad (2)$$

gdje  $iHHa$  i  $iHHd$  označavaju integriranu EHH vrijednost (iHH) ovisno o tome uzima li se u obzir ancestralni ili derivirani alel.

Ukoliko je stopa propadanja EHH slična na ancestralnim i deriviranim alelima, spomenute mjere ( $iHHa$  i  $iHHd$ ) poprimaju vrijednosti od  $\approx 1$ , a nestandardizirani iHS iznosi tada  $\approx 0$ . Kako se, kao što je već spomenuto, unutar neutralnog modela novonastali aleli javljaju u nižim frekvencijama i s dužim haplotipovima u odnosu na alele visoke frekvencije, nestandardizirani iHS se prilagođava za izračun finalne statistike koja ima vrijednost prosjeka 0, a vrijednost varijance 1, bez obzira na frekvenciju alela u ciljnom SNP-u (Voight i sur. 2006.). Takva, krajnja formula (3) po Voightu i sur. (2006.) glasi:

$$iHS = \frac{\ln\left(\frac{iHHa}{iHHd}\right) - E_p\left[\ln\left(\frac{iHHa}{iHHd}\right)\right]}{SD_p\left[\ln\left(\frac{iHHa}{iHHd}\right)\right]} \quad (3)$$

iHS se kao statistička metoda otkrivanja signala pozitivne selekcije temelji na različitim razinama *linkage disequilibrium* (LD) koje se nalaze u susjednim kromosomskim regijama promatranog alela, u odnosu na ostale bazne ili alele na istim pozicijama (Haplotter Guide, <http://haplotter.uchicago.edu/instruction.html>). Unutar seta podataka odnosno unutar skupa SNP-ova, iHS se izračunava za svaki pojedinačni SNP prilikom čega je svaki od analiziranih SNP-ova, smatran ciljnim. Za svaki od tih SNP-ova vrši se mjerenje intenziteta selekcije koja može djelovati direktno na promatrani SNP ili u njegovoj blizini. No prema Voightu i sur. (2016.), preciznije jest tražiti takozvane *window*se odnosno okvire unutar koji se nalaze uzastopni SNP-ovi s ekstremnim iHS vrijednostima. Same analize iHS vrijednosti oslanjaju se na načelu da, pod utjecajem pozitivne selekcije, određeni aleli mogu vrlo brzo povisiti svoju frekvenciju unutar populacije, prilikom čega njihova dugoročna povezanost s bližim polimorfizmima neće biti pod utjecajem rekombinacija (Sabeti i sur. 2007.). Zbog pojave takozvanih *selective sweeps*-a koji uzrokuju nastanak karakterističnih signala varijacija unutar susjednih genomskih regija promatranog markera ili SNP-a (Williamson i sur. 2007.), odnosno uzrokuju pojavu ekstremno visokih ili niskih iHS vrijednosti na obližnjim SNP-ovima

ukoliko je selekcija provedena na ancestralnom alelu, u ovom radu u obzir se uzimaju SNP-ovi s pozitivnim i negativnim iHS vrijednostima, a identifikacija signala temelji se na nizu SNP-ova s  $-\log_{10} p_{iHS}$  vrijednostima  $> 5$ .

## 2.4. Pastirske pasmine pasa

Pastirski psi (eng. *Livestock guard dogs*; LGD) uglavnom su psi veće tjelesne mase (preko 45 kilograma), i iako su predani i prijateljski nastrojeni prema vlasniku, ovi psi su često agresivni prema drugim psima, drugim ljudima i predatorima kao što su vukovi, lisice, kune, mačke pa čak i prema većim predatorima kao što je medvjed. Kao posebna skupina radnih pasa, pastirski psi razvili su specifične karakteristike koje uključuju sposobnost življenja na otvorenom tijekom cijele godine, spremnost da ne uznemiravaju niti da ubijaju stoku (pa čak ni kada su gladni), visoko razvijeno razumijevanje za ponašanje i govor tijela stoke koju čuvaju te vješt pristup otkrivanju i odvratanju predatora (Barth, 2017.). Neke od najpoznatijih i najčešće korištenih pasmina su anatolski ovčar, kavkaski ovčar, sardinijski ovčar, pirinejski planinski pas, mađarski kuvasz, maremansko-abruceški ovčar, pastore della sila i tibetanski mastif. Pastirski psi se koriste već tisućljećima i to ponajviše na području Europe i Azije. Uz čuvanje stoke, mnoge jedinke koriste se i kao čuvari imanja i obitelji te kao terapijski psi za osobe s fizičkim invaliditetom jer zbog svoje velike tjelesne mase pružaju dobar oslonac i zaštitu. Kao i drugi psi, pastirski psi su društvene životinje te imaju potrebu živjeti unutar grupe (čopora) osobito s osobama i psima koje poznaju od rane dobi. Takvu naviku psi su naslijedili od neposrednih predaka, odnosno vukova, te se ona koristi za ranu socijalizaciju pasa (Rigg, 2001.). Boja krzna pastirskih pasa prilagođena je vrsti stoke koju čuvaju, pa se tako mađarski kuvasz i pirinejski planinski pas, pasmine bijelog krzna, najčešće koriste za čuvanje stada ovaca dok se anatolski ovčar najčešće koristi u svrhu čuvanja koza (slika 2.4.1.). Vjeruje se da takva prilagodba pomaže stoki da prihvati psa čuvara, pastiru da razlikuje psa od predatora, ali i predstavlja element iznenađenja ukoliko predator pristupi stadi (Rigg, 2001.).

Do nedavno je selekcija na ovim pasmina bila bazirana isključivo na vanjskim karakteristikama jedinki i na radnim sposobnostima. Ljudi su odabirali jedinke koje su imali na raspolaganju i prilagodili ih potrebnim zadaćama, kreirajući pritom populaciju unutar koje su prisutne velike varijacije u vanjskom izgledu, ali su jedinke dosljedne u radnoj spremi. Takve jedinke postepeno su standardizirane kroz planirani selektivni uzgoj (Sponenberg 2000.).





Slika 2.4.1. Jedinka pasmine mađarski kuvasz uz stado ovaca

Izvor: INFOBUDAPEST <https://info-budapest.com/hungarian-dog-breeds/> -pristup (22.06.2021.)

## **2.4.1. Balkanske pasmine pastirskih pasa**

### **2.4.1.1. Tornjak**

Gotovo izumrli potomci genetski homogenih, izvornih arhaičnih, odnosno drevnih tipova pastirskih pasa bili su temeljne jedinice u procesu kreiranja pasmine tornjak (Federation cynologique internationale, 2017.). Jedinke koje su pripadale početnoj populaciji tornjaka, bile su rasprostranjene na planinskim područjima Bosne i Hercegovine i Hrvatske te u njihovim okolnim dolinama. Prvi zapisi o pasmini datiraju još iz 1067. godine i 1374. godine, kada je pasmina tornjak po prvi puta opisana kao bosansko-hercegovačko-hrvatska pasmina. Godine 1972. na području Bosne i Hercegovine i Hrvatske počinju se provoditi razna istraživanja o povijesti pasmine i sustavno spašavanje od izumiranja. Uzgoj u čistoj krvi započeo je 1978. godine, a današnja populacija se sastoji od mnoštva čistokrvnih jedinki izabranih iz različitih generacija tijekom dužeg vremenskog perioda, te su one rasprostranjene ponajviše na području Bosne i Hercegovine i Hrvatske (FCI, 2017.). Prema klasifikaciji FCI-a (2017.), pasmina tornjak spada u drugu grupu (pinčeri i šnaučeri – molosoidne pasmine – švicarski planinski i stočarski psi) te je namijenjena za čuvanje i zaštitu stoke i zaštitu imanja, a primjer korištenja pasmine u tu svrhu prikazan je na slici (slika 2.4.1.1.1.).



Slika 2.4.1.1.1. Jedinke tornjaka korištene u svrhu čuvanja stada ovaca  
Izvor: American kennel club, <https://www.akc.org/dog-breeds/tornjak/> - pristup (22.06.2021.)

Jedinke tornjaka odlikuje veličina, okretnost, moćnost i proporcionalnost tijela, čiji je oblik gotovo kvadratičan. Tornjak je dobro usklađen i snažan pas sa dugom i gustom dlakom. Morfološke karakteristike pasmine prikazane su na slici (2.4.1.1.2.). Kada su u pitanju bihevioralne karakteristike, jedinke ove pasmine su opisane kao vrlo odvažne, poslušne, samopouzdanе i oštре u čuvanju imovine i stada koje im je povjereno. Jedinke su često nepovjerljive prema strancima, ali vrlo privržene vlasniku te ih je lako dresirati i lako uče kontrolirati i čuvati stado. Za daljnji rasplod, odnosno uzgojno selekcijski rad koriste se jedinke koje ispoljavaju sve spomenute morfološke i bihevioralne karakteristike, dok svaka jedinka koja jasno pokazuje fizičke abnormalnosti ili abnormalnosti glede ponašanja treba biti diskvalificirana iz daljnjeg uzgoja (Hrvatski kinološki savez, 2017.).



Slika 2.4.1.1.2. Jedinka pasmine Tornjak  
Izvor: American kennel club, <https://www.akc.org/dog-breeds/tornjak/> - pristup (22.06.2021.)

### 2.4.1.2. Šarplaninac

Šarplaninac, odnosno jugoslavenski šarplaninski pas se prvi puta počeo uzgajati na planinskim područjima istočne Jugoslavije točnije u predjelu Šar-planine po kojoj je pasmina i dobila ime. Podrijetlo pasmine jest i dalje kontroverzno. Postoje razni izvori koji tumače da je pasmina donesena u Europu sa područja Azije tijekom masovnih prapovijesnih migracija ljudi. Izvorni tip pasmine prisutan je isključivo u onim područjima gdje još uvijek prevladava intenzivno stočarstvo i gdje ovi psi mogu obavljati svoju izvornu ulogu kao čuvari stada na otvorenim pašnjacima (FCI, 1970.). Pasma je prvi puta priznata od strane FCI-a davne 1939. godine te je svrstana u drugu grupu, kao i tornjak.

Jedinke su robusne, čvrstog stava, mirnog temperamenta, tijelo im je proporcionalno i pokriveno gustom i grubom dlakom, zaštitnički su nastrojene i odane svome vlasniku. Prosječna visina u grebenu kod mužjaka iznosi 62 centimetara, a kod ženki 58 centimetara, dok se težina muških jedinki kreće između 35 i 45 kilograma, a ženke dosežu težine od 30 do 40 kilograma (FCI, 1970.). Selekcija se temelji na svim spomenutim karakteristikama kao i na dodatnim morfološkim karakteristikama koje su vezane uz samu građu tijela, karakteristike trupa i ekstremiteta te su prikazane i detaljno opisane na slici (slika 2.4.1.2.1.), a za daljnji rasplod koriste se samo funkcionalno i klinički zdrave jedinke, odnosno jedinke sa klasičnom konformacijom.



Slika 2.4.1.2.1. Detaljan opis poželjnih morfoloških karakteristika jedinke šarplaninac pasmine  
Izvor: mojafarma, <http://mojafarma.ba/jugoslavenski-ovcarski-pas-sarplaninac/> - pristup (22.06.2021.)

### 2.4.1.3. Kraški ovčar/Kraševac

Kraški ovčar, poznatiji i kao kraševac, jest pasmina koja postoji već nekoliko stoljeća i pripada molosoidnoj skupini pasa (FCI, 2000.). Skupinu molosoidnih pasmina karakteriziraju široka prsa, veliki stas i mišićave noge, a skupina je ime dobila po nazivu plemena *Molossi* iz kojega potječu ovi psi. Vjeruje se da su jedinke kraškog ovčara

pratili Ilire kroz njihovu migraciju na području Istre i dalmatinskih otoka te su naselile područje Slovenije. Pasma je prvi puta spomenuta davne 1689. godine, a službeno je priznata tek 1939. godine pod imenom „Ilirski ovčar“. Nakon velikog broja zasjedanja FCI-a, 1968. godine, pasmina je preimenovana u „kraški ovčar“ te je i danas poznata pod istim imenom.

Jedinke kraškog ovčara su srednje veličine, robusne, posjeduju izraženu muskulaturu trupa i ekstremiteta, uši i rep su im viseći dok je dlaka najčešće željezno siva, duga i obilna (FCI, 2000.). Umjereno oštar temperament, poslušnost, privrženost vlasniku i hrabrost su također karakteristike koje krasi ovu pasminu. Iako je pasmina u povijesnim vremenima korištena za čuvanje stoke, danas je kraški ovčar korišten ponajviše kao pas čuvar ili kao obiteljski pas. Vanjski izgled pasmine kao i primjer njezina korištenja u svrhu zaštite stoke prikazan je na slici (slika 2.4.1.3.1.). Selekcija jedinki kraškog ovčara za daljnji rasplod temelji se najčešće na morfološkim karakteristikama koje ne smiju odstupati od standarda postavljenog od strane FCI-a.



Slika 2.4.1.3.1. Jedinica kraškog ovčara korištena u svrhu čuvanja stada ovaca  
Izvor: Thezaurus. Institute for Slovenian Studies of Victoria,  
[https://thezaurus.com/karst\\_sheepdog/](https://thezaurus.com/karst_sheepdog/) - pristup (22.06.2021.)

### **3. Cilj istraživanja**

Na genomu postoje regije koje pokazuju signale pozitivne selekcije.

Cilj istraživanja jest provedbom analiza na visoko rezolutnim genetskim markerima 123 jedinke pastirskih pasmina pasa, identificirati regije genoma zajedno sa susjednim genima, koje pokazuju signale pozitivne selekcije. Unutar tih regija također će se provesti anotacija kandidat gena.

## 4. Materijali i metode

Za potrebe ovog rada genomski podatci 496 jedinki pasa i 17 jedinki vukova preuzeti su iz Janeš i sur (2021.). Od ukupno 496 jedinki pasa, njih 123 se ubraja u pastirske pasmine (anatolski ovčar; n = 6, kavkaski ovčar; n = 5, pirinejski planinski pas; n = 13, sardinijski ovčar; n = 6, mađarski kuvasz; n = 10, paremansko-abruceški ovčar; n = 14, pastore della sila; n = 14, tibentanski mastif; n = 11, kraševac; n = 10, tornjak; n = 20, šarplaninac; n = 14). Jedinke ovih 11 različitih pastirskih pasmina, unutar analiza tretirane su kao jedna meta-populacija. Korišteni genomski podaci prethodno su prošli kontrolu kvalitete.

### 4.1. iHS analiza

Za uspješan izračun vrijednosti iHS-a specifičnih haplotipova, potrebno je provesti određivanje takozvane faze (faziranje) na razini svakog pojedinačnog kromosoma, odnosno u ovom slučaju na razini svakog od 38 autosomalnih kromosoma koji se nalaze unutar genoma psa. Razdvajanje pojedinačnih kromosoma iz cjelokupnog seta podataka kao i njihovo grupiranje s pripadajućim parametrima prethodilo je utvrđivanju faze, a ono je provedeno pomoću softvera PLINK v1.9 (Purcell i sur. 2007.). Na tako uspješno razdvojenim kromosomima provedeno je faziranje unutar Shapeit softvera (Delaneau i sur. 2012.). Ovaj softver predstavlja brzu i točnu metodu procjene haplotipova unutar genotipskih podataka ili podataka sekvence, a prilikom faziranja u obzir uzima specifične parametre/funkcije s pripadajućim vrijednostima (korištenim u ovom radu), kao što su *Multi-threading* (--thread; 3), *Iteration number* (--burn; 10, --prune; 10, --main; 50), *Window size W* (--window; 1) i *Effective population size* (--effective-size; 3500) uz pomoću kojih se dobiva krajnja procjena haplotipova. Rezultat analiza Shapeit softvera jesu šest novonastalih dokumenata (.haps, .graph, .log, .sample, .ind.mm i .snp.mm) od kojih je, za daljnje analize iHS vrijednosti najvažniji .haps dokument. Takav dokument u sebi sadrži identifikacijske oznake SNP-ova, njihovu poziciju na pojedinim kromosomima, informacije vezane uz ancestralni i alternativni alel kao i oznaku kromosoma na kojemu su pozicionirani promatrani SNP-ovi. Kao što je već spomenuto u poglavlju 2.3. ovoga rada, prema Voightu i sur. (2006.) za izračun iHS vrijednosti potrebno je poznavati vrijednost integrirane EHH za haplotipove koji centriraju ancestralni alel (iHHa) i za haplotipove koji centriraju derivirani alel (iHHd). Izračun tih vrijednosti, kao i vrijednosti standardiziranog iHS-a moguće je obaviti upotrebom programskog paketa *rehh* unutar R programa (Gautier i Vitalis 2012.). Spomenuti paket zahtjeva pretvorbu .haps oblika dokumenta u takozvani *variant call format* (.vcf) unutar kojega se nalaze identificirani genetski haplotipovi, skupina alela unutar pojedinog lokusa na promatranom kromosomu kao i svi derivirani aleli pojedinog lokusa (Fairley i sur. 2020.). Unošenje .vcf dokumenata unutar R programa moguće je korištenjem „*data2haplohh*“ funkcije, koja ulazni dokument pretvara u „*haplohh*“ objekt unutar radnog direktorija programa (Klassmann i sur. 2020.). Nakon unosa svih .vcf dokumenata, u ovom slučaju 38 dokumenata

(zasebni .vcf dokument za svaki od 38 autosomalnih kromosoma), na svakome od njih provodi se „scan\_hh“ funkcija koja daje izračun integrirane EHH vrijednosti (iHH), vrijednost integrirane proširene homozigotnosti haplotipova specifične pozicije (eng: *site-specific Extended Haplotype Homozygosity*; EHHS) (iEHHS) i integrirane normalizirane EHHS vrijednosti (inES) (Klassmann i sur. 2020.). Izračunate vrijednosti za svaki od kromosoma smještaju se u jedan krajnji objekt, koristeći funkciju *rbind*. Na takvom objektu moguće je vršiti završni izračun iHS vrijednosti korištenjem *ihh2ihs* funkcije koja iHS predstavlja kao standardizirani oblik iHH vrijednosti. Dobivene vrijednosti, zajedno s postavljenim *threshold-om* ( $-\log_{10} p_{iHS} > 5$ ) vizualiziraju se na Manhattan plotu, što je prikazano i u poglavlju 5. ovog rada.

## 4.2. ROH analiza

Unutar skupa genomskih podataka 123 jedinke, ROH segmenti identificirati će se i procijeniti pomoću SNP & Variation Suite (SVS) v8.7.0 programskog paketa (Golden Helix, Inc., Bozeman, MT, [www.goldenhelix.com](http://www.goldenhelix.com)), čije se analize temelje na takozvanoj uzastopnoj (eng: *consecutive*) metodi. Prije provedbe samih analiza, program zahtjeva specifične parametre pomoću kojih će se na svakom od 38 autosomalnih kromosoma pokušati identificirati sve moguće nizove SNP-ova koji se pojavljuju u homozigotnom obliku. U okviru analiza provedenih u svrhu pisanja ovog rada, segmenti koji su identificirani kao ROH morali su sadržavati  $\geq 15$  uzastopnih homozigotnih SNP-ova prilikom čega je maksimalna dopuštena udaljenost između dva SNP-a iznosila 1.000 kb. Sama procjena segmenata provodila se na pet definiranih klasa, unutar kojih su postavljene različite dužine segmenata: 1 – 2 Mb, 2 – 4 Mb, 4 – 8 Mb, 8 – 16 Mb i segmenti veći od 16 Mb. Zbog mogućih pogreški tijekom procesa genotipizacije HD SNP čipa, u svakoj od prethodno nabrojanih klasa definira se i broj dopuštenih heterozigota kao i *missing* SNP-ova. Broj dopuštenih heterozigota i *missing* SNP-ova izračunat je prema Ferenčaković i sur. (2013b.). Kod segmenata unutar prve klase (1 – 2 Mb) dopušten je jedan heterozigotni SNP kao i četiri *missing* SNP-ova. Segmentima unutar druge klase (2 – 4 Mb) dopušteno je posjedovanje dva heterozigotna SNP-a i osam *missing* SNP-ova. U trećoj klasi, sa segmentima dužine 4 – 8 Mb, dopušteno je četiri heterozigotna SNP-a i 16 *missing* SNP-ova, dok su u četvrtoj i petoj klasi te brojke nešto veće, osam heterozigotnih SNP-ova i 32 *missing* SNP-ova unutar četvrte klase i 16 heterozigotnih SNP-ova i 64 *missing* SNP-ova u posljednjoj klasi. Prema Ferenčaković i sur. 2013. (prema Howrigan i sur. 2011.), broj i veličina ROH segmenata identificiranih unutar genotipskih podataka mogu uvelike ovisiti o specifičnim parametrima i graničnim vrijednostima postavljenim u procesu analiza sekvenci. Rezultati dobiveni unutar SVS programskog paketa daljnje su analizirani putem SAS 9.4. softvera (SAS Institute Inc 2013.), pomoću kojega je proveden izračun količine pojedinog SNP-a unutar populacije od 123 jedinke. Učestalost pojave svakog SNP-a zatim je standardizirana, te je na temelju dobivene vrijednosti izračunata p vrijednost za postojeće SNP-ove. Osnovnim programskim paketima *stats* i *base* (R Core Team), unutar R programa procijenjena je p vrijednost

na negativnoj  $\log_{10}$  skali. Vizualizacija cjelokupog skupa kromosoma kao i svakog pojedinačnog, također je izvršena unutar R programa koristeći se pritom *qqman* programskim paketom (Turner 2018.). Vizualizacija je obavljena na Manhattan plot-u (poglavlje 5.).

### **4.3. Identifikacija signala selekcije**

Identifikacija i interpretacija signala selekcije temeljila se isključivo na p vrijednostima na negativnoj  $\log_{10}$  skali ostvarenoj unutar iHS i ROH analiza. Niz SNP-ova s vrijednošću frekvencije unutar populacije iznad 5 na negativnoj  $\log_{10}$  skali, okarakterizirao se kao značajan signal selekcije kod provedbe obje analize (eROHi i iHS). Kao što je već rečeno u poglavlju 2.3., iHS analize u obzir uzimaju određene razine LD-a koji se može javiti između pojedinih markera ili SNP-ova, te je posljedično tome vrijednostima početnih pozicija pojedinih regija pod selekcijom oduzeto 1.000.000 baznih parova (Mb), dok je završnim pozicijama dodana ista vrijednost. Takav postupak proveo se i na pozicijama regija unutar eROHi analiza.

### **4.4. Anotacija gena unutar regija signala pozitivne selekcije**

Upotrebom SNP & Variation Suite (SVS) v8.7.0 programskog paketa (Golden Helix, Inc., Bozeman, MT, [www.goldenhelix.com](http://www.goldenhelix.com)) provedena je identifikacija gena unutar označenih regija genoma u korištenoj populaciji od 123 jedinke pastirskih pasmina. U procesu identifikacije, korištena je referentna sekvenca genoma psa (Canfam3.1) koja je dostupna unutar programskog paketa. Provjera funkcionalnost pojedinih identificiranih gena provedena je putem UniProt (<https://www.uniprot.org>) i GeneCards (<https://www.genecards.org>) platformi, koje sadrže sve do sada poznate i istražene funkcije gena od interesa. Opis gena čija je funkcija unutar psećeg organizma nepoznata, nije se provodio.

Sve dosadašnje poznate funkcije nekih od gena identificiranih unutar specifičnih regija pojedinih kromosoma, navedene su u poglavlju 6.



## 5. Rezultati

Sve spomenute analize identifikacije signala pozitivne selekcije provedenu su na 38 autosomala kromosoma kod 123 jedinke pastirskih pasmina pasa. Na korištenim podacima provedena je kontrola kvalitete kako bi se smanjila mogućnost pogreške u daljnjim analizama i kako bi se osigurali što točniji rezultati istih. Očišćeni set podataka sastojao se od ukupno 107.403 SNP-a unutar genoma 123 jedinke.

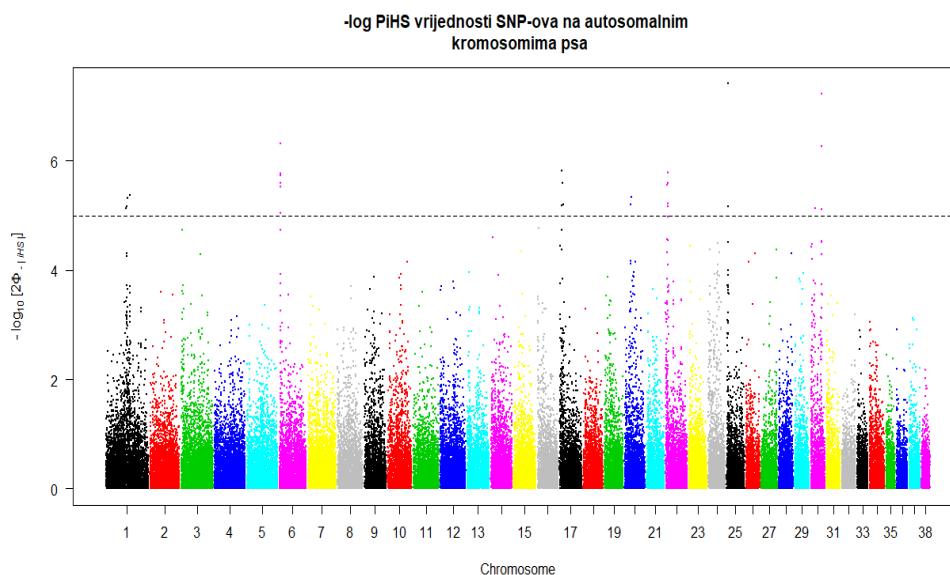
### 5.1. Rezultati iHS analize

Provedbom specifičnih analiza ustanovljene su iHS i  $-\log_{10} p_{iHS}$  vrijednosti za svaki SNP, koristeći se pritom podacima haplotipova jedinki. Na temelju ustanovljenih  $-\log_{10} p_{iHS}$  vrijednosti, ukupno 21 SNP autosomala kromosoma je identificiran kao *outlier* ( $-\log_{10} p_{iHS}$  vrijednost  $> 5$ ) i potencijalni signal pozitivne selekcije. Takvi SNP-ovi bili su distribuirani na 6 od ukupno 38 autosomala kromosoma. Kromosomi koji su unutar iHS analiza identificirani kao nositelji signala pozitivne selekcije, vizualizirani su unutar R programa na Manhattan plot-u (slika 5.1.1.). Detaljna vizualizacija pojedinih kromosoma i njihovih regija unutar kojih se nalaze nizovi SNP-ova s  $-\log_{10} p_{iHS}$  vrijednosti  $> 5$ , također je obavljena u prethodno navedenom programu. Tablica 1. obuhvaća sve kromosome na kojima je identificiran signal pozitivne selekcije, regiju koja je pod pozitivnom selekcijom (izražena u broju baznih parova), ukupan broj SNP-ova s  $-\log_{10} p_{iHS}$  vrijednošću  $> 5$  unutar regije kao i identifikacijske oznake najsignifikantnijeg SNP-a („*peak SNP*“) s pripadajućom  $-\log_{10} p_{iHS}$  vrijednosti.

Tablica 1. Rezultati iHS analize

*Chr	Regija (Mb)	Ukupan br. SNP-ova s $-\log_{10} p_{iHS} > 5$	*Peak SNP	$-\log_{10} p_{iHS}$
1	57.27– 61.23	3	TIGRP2P4883_rs8809779	5,172557976
6	4.00 – 4.44	7	BICF2G630808220	6,320486615
17	5.83 – 5.84	2	BICF2P767397	5,826612552
22	6.13 – 6.65	4	BICF2P301108	5,792661844
25	3.87 – 4.80	2	BICF2P137430	7,410584882
30	30.91 – 31.26	3	BICF2G630400340	7,231606610

\*Chr = Kromosom, \*Peak SNP = SNP/outlier s najvećom  $-\log_{10} p_{iHS}$  vrijednošću



Slika 5.1.1. Manhattan plot  $-\log_{10} p_{iHS}$  analize 38 autosomalnih kromosoma psa

### 5.1.1. CanFam 1: 57.27 Mb – 61.23 Mb

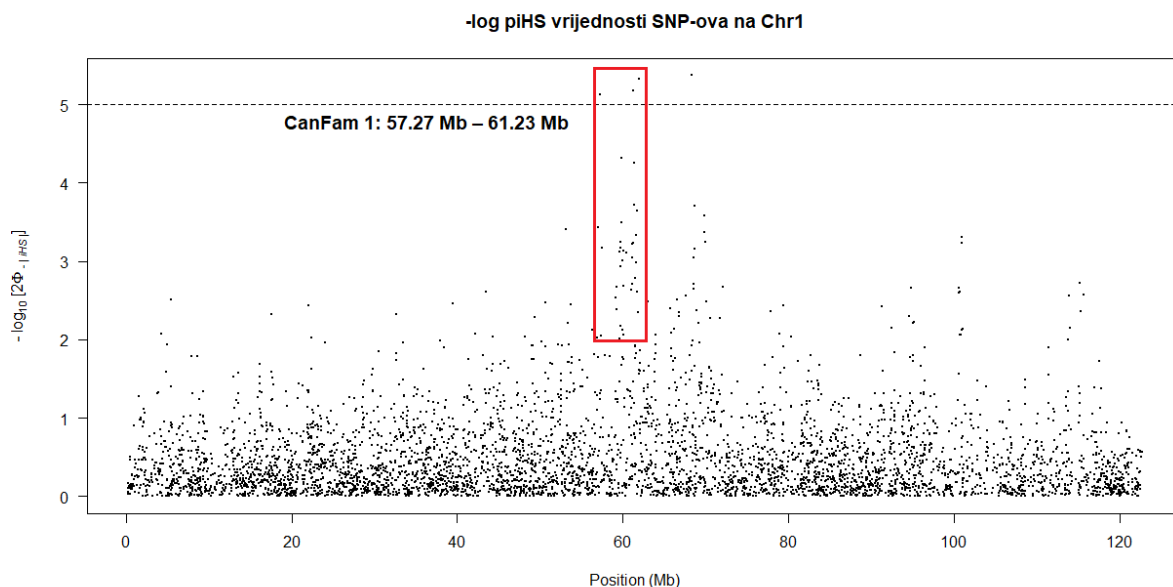
Na 1. autosomalnom kromosomu unutar genoma psa identificirana je regija pod pozitivnom selekcijom, duljine 3.96 Mb koju je moguće vidjeti na slici (slika 5.1.1.1.) Unutar specifične regije nalazi se ukupno 153 SNP-ova, ali samo njih 3 posjeduju  $-\log_{10} p_{iHS}$  vrijednosti od  $> 5$  (tablica 2.).

Tablica 2. SNP-ovi s  $-\log_{10} p_{iHS}$  vrijednosti  $> 5$  na 1. kromosomu

*SNP ID	Pozicija (Mb)	* $-\log_{10} p_{iHS}$	* $ iHS $
TIGRP2P4883_rs8809779	57274975	5,132683355	-4,482787149
BICF2P1006887	57276881	5,132683355	-4,482787149
BICF2P1166414	61233718	5,172557976	-4,502333964

\*SNP ID = identifikacijska oznaka SNP-a, \* $-\log_{10} p_{iHS}$  = p vrijednost SNP-a, \* $|iHS|$  = iHS vrijednost SNP-a

Kada su u pitanju iHS vrijednosti, SNP sa najnižom vrijednošću ( $|iHS| = -4,50233$ ) nalazi se na poziciji 61.23 Mb kromosoma i nosi oznaku BICF2P1166414, dok je TIGRP2P5414\_rs9146508 SNP sa najvećom iHS vrijednosti od 3,601596, a pozicioniran je na 59.85 Mb. Unutar regije identificirano je najviše gena, njih 34: RFX6, VGLL2, ROS1, GOPC, NEPN, NUS1, SLC35F1, CEP85L, PLN, MCM9, ASF1A, FAM184A, MAN1A1, TBC1D32, GJA1 i 19 drugih gena s još nepoznatom funkcijom.



Slika 5.1.1.1. Manhattan plot  $-\log_{10} p_{iHS}$  vrijednosti SNP-ova na 1. kromosomu

### 5.1.2. CanFam 6: 4.07 Mb – 4.43 Mb

Najjači signal pozitivne selekcije otkriven je na 6. autosomalnom kromosomu, kao što je prikazano na slici (slika 5.1.2.1.). Regija signala proteže se od 4.07 Mb do 4.41 Mb, a u njoj je pozicionirano ukupno 30 SNP-a, od kojih je 7 s  $-\log_{10} p_{iHS}$  vrijednosti iznad 5 (Tablica 3.).

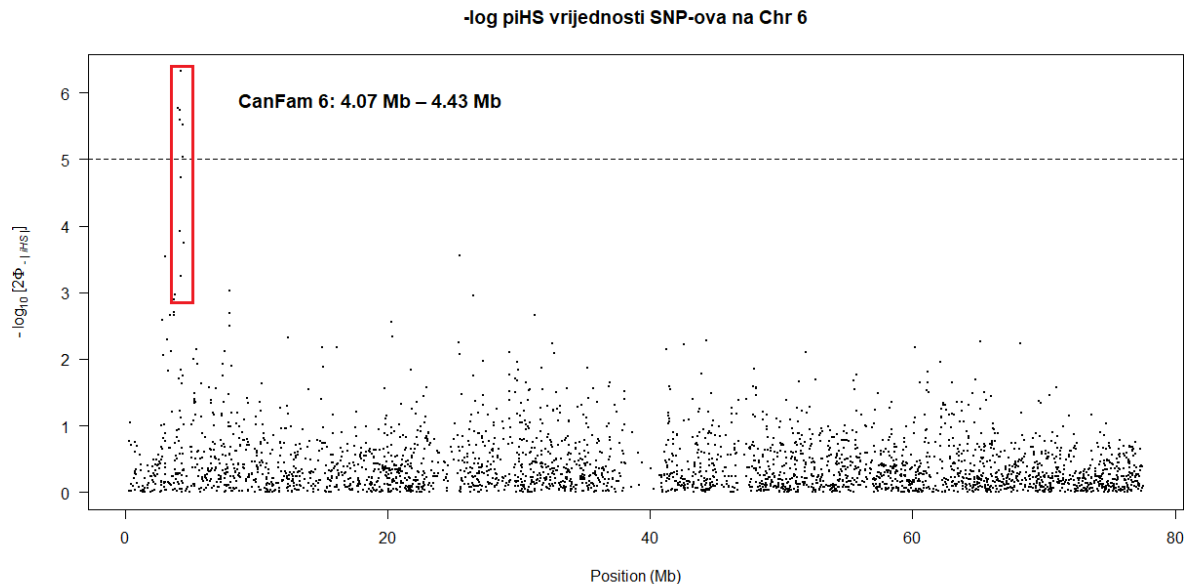
Tablica 3. SNP-ovi s  $-\log_{10} p_{iHS}$  vrijednosti > 5 na 6. kromosomu

*SNP ID	Pozicija (Mb)	* $-\log_{10} p_{iHS}$	* iHS
BICF2P560997	4071018	5,775899424	4,789097488
BICF2G630808220	4278402	6,320486615	5,034901703
BICF2P487152	4188620	5,739317785	4,772164928
BICF2P1065607	4182673	5,593247906	4,703991033
BICF2P1065608	4182767	5,593247906	4,703991033
BICF2G630808343	4416587	5,527600913	4,67305261
BICF2G630808339	4407406	5,041255382	4,437670137

\*SNP ID = identifikacijska oznaka SNP-a, \* $-\log_{10} p_{iHS}$  = p vrijednost SNP-a, \*|iHS| = iHS vrijednost SNP-a

SNP s najvećom iHS vrijednosti unutar spomenute regije jest BICF2G630808220 (|iHS|= 5,03), dok je BICF2G630808290 SNP sa najnižom iHS vrijednosti (|iHS|= -

2.26). Ovi SNP-ovi pozicionirani su na 4.2 Mb i 4.3 Mb. Unutar regije od interesa identificirano jest ukupno 6 kandidat gena: LOC102155051, LOC102155153, LOC111096184, LOC111096185, LOC111096186 i LOC111096482, koji zasad nemaju poznatu funkciju unutar psećeg genoma.



Slika 5.1.2.1. Manhattan plot  $-\log_{10} p_{iHS}$  vrijednosti SNP-ova na 6. kromosomu

### 5.1.3. CanFam 17: 5.83 Mb – 5.84 Mb

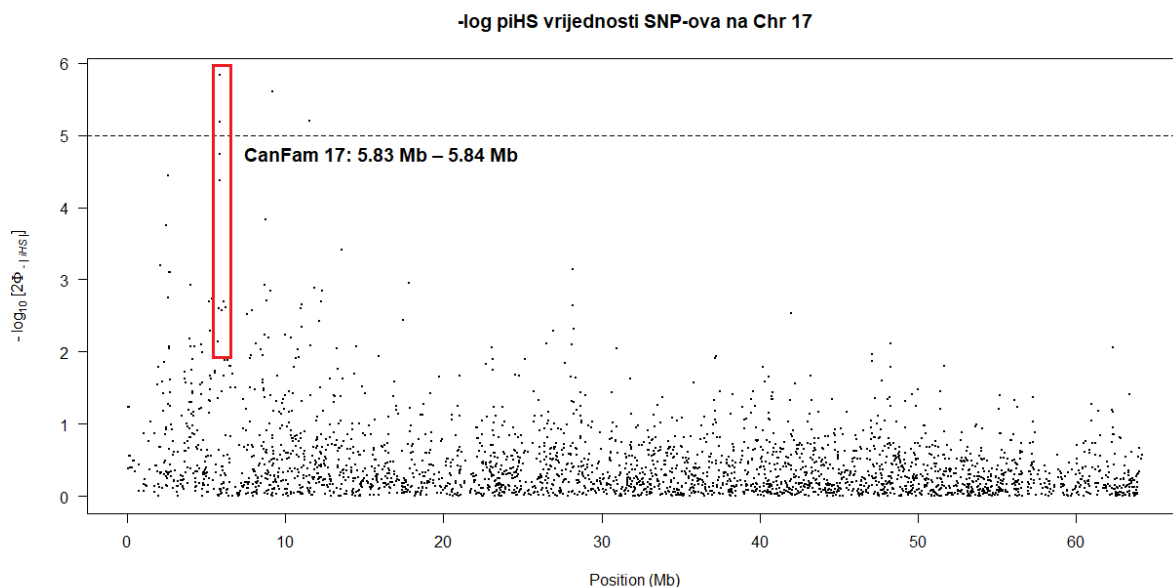
Na 7. autosomalnom kromosomu analizama iHS-a, identificiran je signal pozitivne selekcije unutar malene regije koja se proteže od 5.83 Mb do 5.84 Mb. Unutar regije pozicionirano je ukupno 4 SNP-a od kojih 2 poprimaju  $-\log_{10} p_{iHS}$  vrijednosti veće od 5. kao i najniže iHS vrijednosti, što je moguće primijetiti u tablici 4. Distribucija  $-\log_{10} p_{iHS}$  na cjelokupnom 17. kromosomu prikazana je na slici (slika 5.1.3.1.).

**Tablica 4.** SNP-ovi s  $-\log_{10} p_{iHS}$  vrijednosti  $> 5$  na 17. kromosomu

*SNP ID	Pozicija (Mb)	* $-\log_{10} p_{iHS}$	* iHS
BICF2P767397	5838018	5,826613	-4,81248
BICF2P473910	5847300	5,178165	-4,50508

\*SNP ID = identifikacijska oznaka SNP-a, \* $-\log_{10} p_{iHS}$  = p vrijednost SNP-a, \*|iHS| = iHS vrijednost SNP-a

Na 7. autosomalnom kromosomu, unutar spomenute regije identificiran je samo jedan gen LOC111090474 čija je funkcija unutar psećeg genoma još uvijek nepoznata.



Slika 5.1.3.1. Manhattan plot  $-\log_{10} p_{iHS}$  vrijednosti SNP-ova na 17. kromosomu

#### 5.1.4. CanFam 22: 6.13 Mb – 6.66 Mb

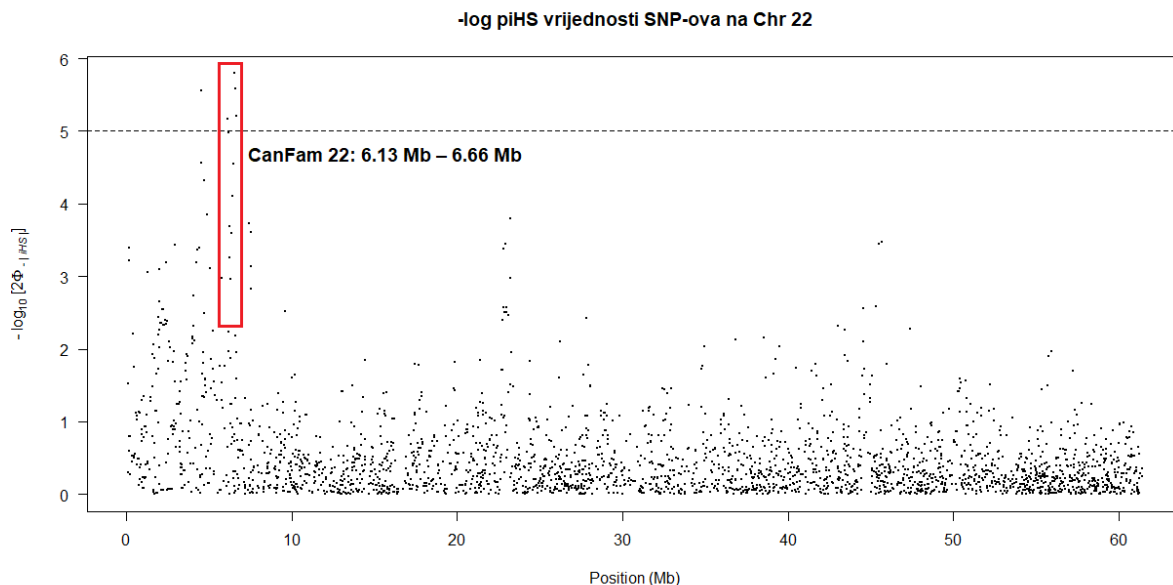
Nešto manja regija od svega 51.82 Kb, detektirana je kao signal selekcije na 22. autosomalnom kromosomu. Unutar nje pozicionirano je 4 SNP-ova s  $-\log_{10} p_{iHS}$  vrijednosti  $> 5$  te 18 ostalih SNP-ova. Najsignifikantnije  $-\log_{10} p_{iHS}$  i  $iHS$  vrijednosti, zajedno sa pripadajućim SNP-ovima prikazane su u tablici 5.

**Tablica 5.** SNP-ovi s  $-\log_{10} p_{iHS} > 5$  na 22. kromosomu

*SNP ID	Pozicija (Mb)	* $-\log_{10} p_{iHS}$	* $ iHS $
BICF2P301108	6566166	5,792662	-4,79684
BICF2G630316549	6588073	5,588421	-4,70172
BICF2G630316559	6650969	5,212498	-4,52183
BICF2G630316367	6132742	5,165878	4,499065

\*SNP ID = identifikacijska oznaka SNP-a, \* $-\log_{10} p_{iHS}$  = p vrijednost SNP-a, \* $|iHS|$  =  $iHS$  vrijednost SNP-a

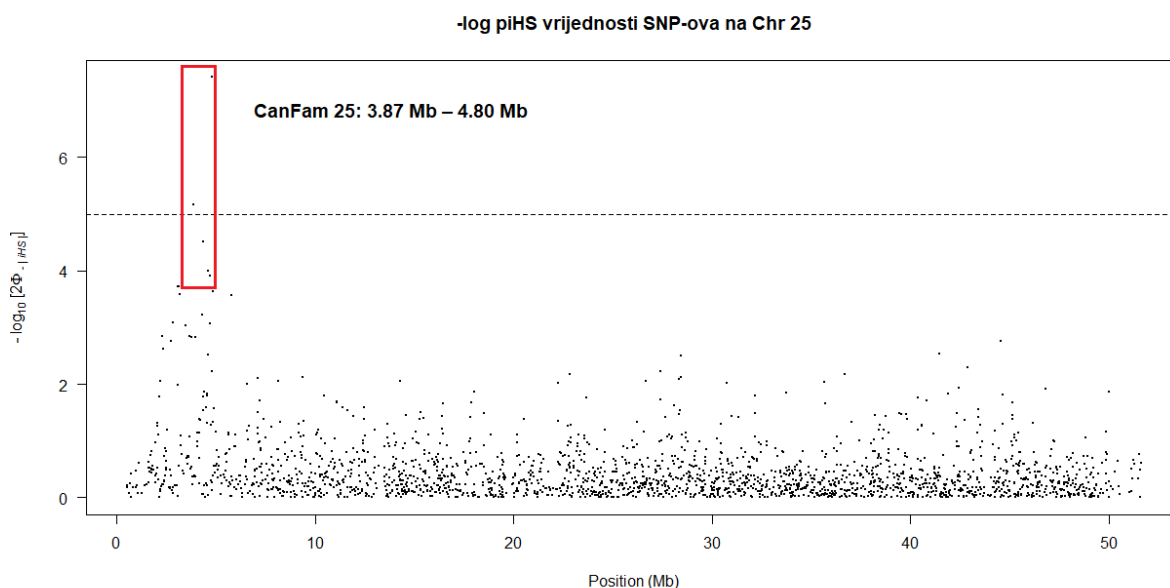
Slika 5.1.4.1. prikazuje distribuciju  $P_{iHS}$  vrijednosti za SNP-ove koji su pozicionirani na 22. autosomalnom kromosomu. Iako je regija signala vrlo mala, unutar nje uspješno je identificirano ukupno 7 gena: TSC22D1, SERP2 i ostalih 5 s nepoznatom funkcijom.



Slika 5.1.4.1. Manhattan plot  $-\log_{10} \pi_{iHS}$  vrijednosti SNP-ova na 22. kromosomu

### 5.1.5. CanFam 25: 3.87 Mb – 4.80 Mb

Pretposljednja regija koja je identificirana kao signal selekcije unutar analiza iHS-a jest regija od 3.87 Mb do 4.80 Mb na 25. kromosomu. Unutar regije nalazi se ukupno 40 SNP-ova. Dva SNP-a (BICF2P137430 i BICF2S23751649), poprimala su  $-\log_{10} \pi_{iHS}$  vrijednosti  $> 5$ . Distribucija  $-\log_{10} \pi_{iHS}$  za spomenute, ali i ostale SNP-ove prikazana je na slici (slika 5.1.5.1.). Nadalje, unutar ove regije identificirano je 6 potencijalnih kandidat gena: SERTM1, CCNA1, SPART, CCDC169, SOHLH2 i DCLK1 kao i 6 gena s nepoznatim funkcijama i nazivima.



Slika 5.1.5.1. Manhattan plot  $-\log_{10} \pi_{iHS}$  vrijednosti SNP-ova na 25. kromosomu

### 5.1.6. CanFam 30: 30.91 Mb – 31.26 Mb

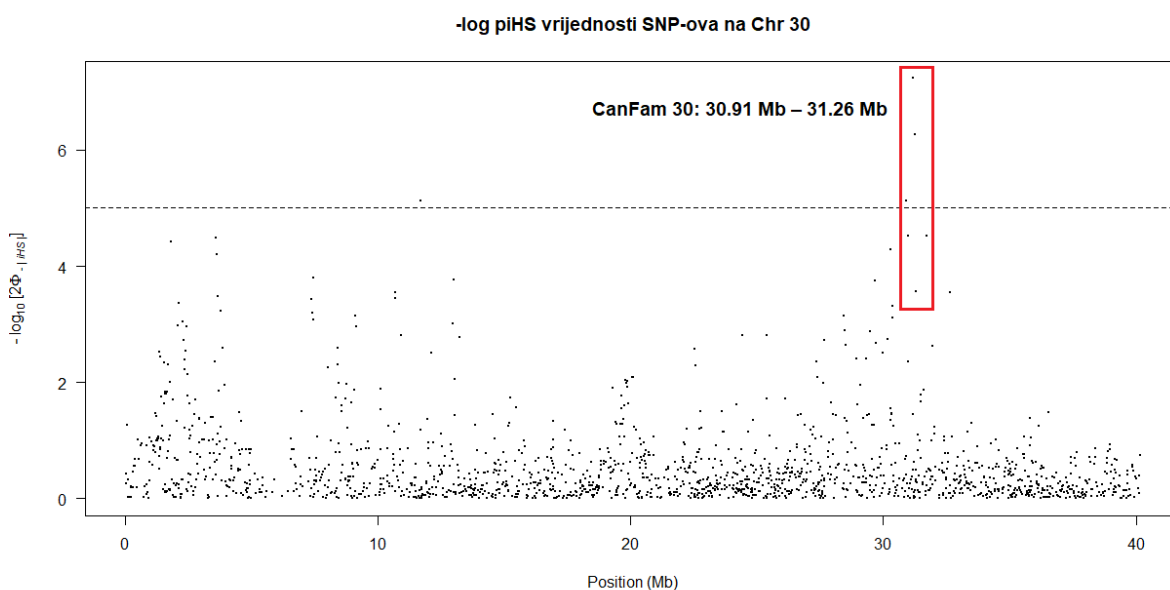
Na 30. kromosomu tri SNP-a poprimaju  $-\log_{10} p_{iHS}$  vrijednosti  $> 5$ . Identificirani SNP-ovi s takvim vrijednostima nalaze se unutar Tablice 6. dok je na slici 5.1.6.1. prikazana  $P_{iHS}$  za svaki od SNP-ova.

**Tablica 6.** SNP-ovi s  $-\log_{10} p_{iHS} > 5$  na 30. kromosomu

*SNP ID	Pozicija (Mb)	* $-\log_{10} p_{iHS}$	* $ iHS $
BICF2G630400340	31170271	7,231607	-5,42282
BICF2G630400259	31266677	6,265831	-5,01074
TIGRP2P371564_rs9117831	30917281	5,121265	-4,47718

\*SNP ID = identifikacijska oznaka SNP-a, \* $-\log_{10} p_{iHS}$  = p vrijednost SNP-a, \* $|iHS|$  = iHS vrijednost SNP-a

Unutar regije identificirana su 4 gena: SMAD6, SMAD3 i 2 gena s nepoznatom funkcijom LOC111093334 i LOC111093346.

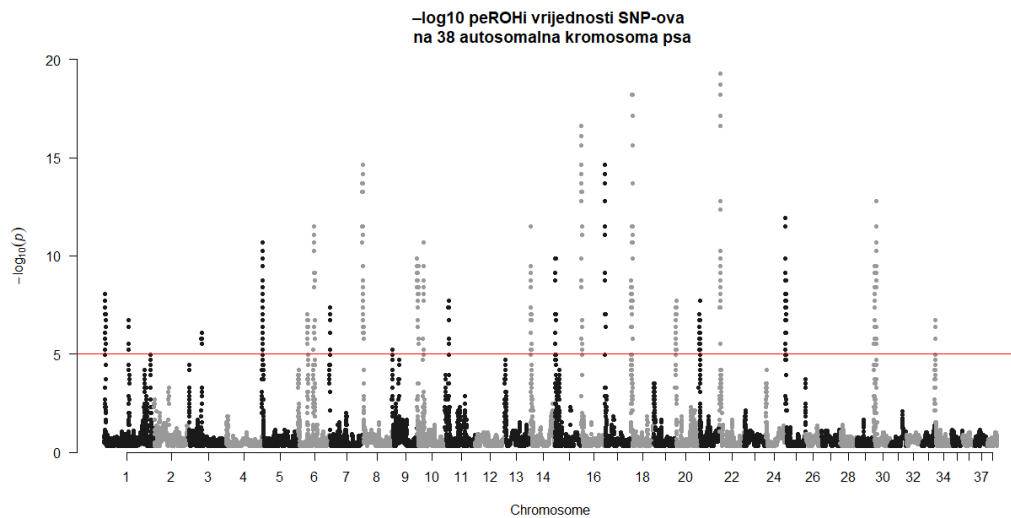


Slika 5.1.6.1. Manhattan plot  $-\log_{10} p_{iHS}$  vrijednosti SNP-ova na 30. kromosomu

## 5.2. Rezultati eROHi analiza

Svih 38 autosomnih kromosoma uključeno je u analize identifikacije ROH segmenata kao i analize učestalosti pojedinih SNP-ova unutar istih. Provedeno istraživanje koristilo je HD SNP čip za otkrivanje ROH segmenata manjih dužina, od svega 1 Mb. Nakon detekcije ROH segmenata unutar genomskih podataka, uslijedio je izračun  $-\log_{10} p_{eROHi}$  vrijednosti za svaki pojedinačni SNP unutar segmenta. Na taj

način identificirane su 23 genomske regije pod pozitivnom selekcijom, koje su raspodijeljene na 20 autosomalnih kromosoma (Slika 5.2.1.).



Slika 5.2.1. Manhattan plot  $-\log_{10} p_{eROHi}$  vrijednosti analize SNP-ova na 38 autosomalnih kromosoma psa

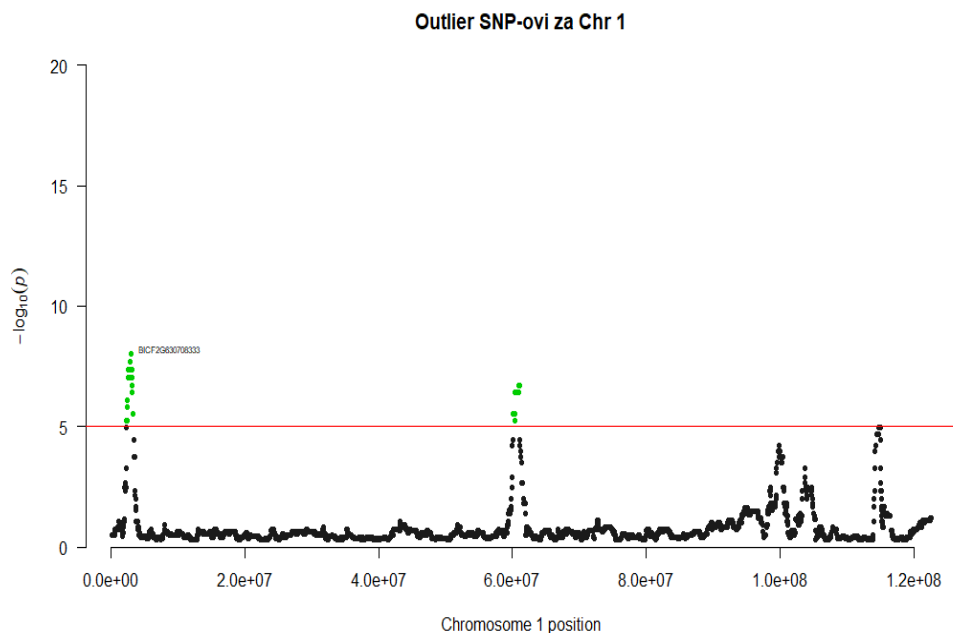
Promatranjem eROHi-a kao i identificiranih signala, moguće je primijetiti da najznačajniji SNP-ovi (SNP-ovi s najsignifikantnijom  $-\log_{10} p_{eROHi}$  vrijednošću) unutar signala predstavljaju takozvani *peak*, odnosno mjesto na kojemu se najvjerojatnije nalazi selektirani marker/alel, no zbog LD-a postoji vjerojatnost da se takav alel nalazi i u blizini *peak*-a.

Kako po Ferenčaković i sur. (2013.) duljina ROH segmenata može pružiti uvid u starost inbreedinga, segmenti identificirani u ovim analizama mogu nam pomoći u razumijevanju nedavne, ali i povijesne selekcije, bilo da je ona provedena ljudskim radom ili prirodnom prilagodbom na vanjske čimbenike. Kao što je već spomenuto u poglavlju 4.2. ovog rada, *qqman* programskim paketom obavljena je vizualizacija svih regija pod selekcijom što je moguće vidjeti u nastavku rada.

### 5.2.1. CanFam 1: 2.36 Mb – 3.32 Mb

Prvi identificirani signal nalazi se na 1. autosomalnom kromosomu. Regija od 2.36 Mb do 3.32 Mb sadrži 34 *outlier* SNP-ova, a najznačajniji od njih (BICF2G630708333;  $-\log_{10} p_{eROHi} = 8,040612532$ ) prikazan je zajedno s ostalim *outlier*-ima na slici (slika 5.2.1.1.). Spomenuti *outlier* pozicioniran je na 3.00 Mb. Unutar regije također je uspješno identificirano 15 gena: GARL1, MBP, ZNF236, C1H18orf65, ZNF516 i 10 ostalih gena s nepoznatom funkcijom.





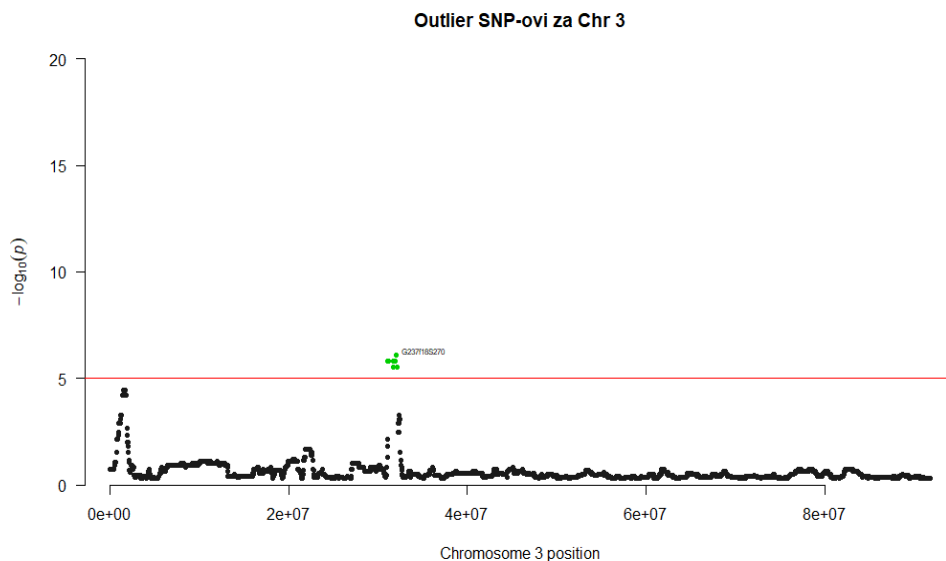
Slika 5.2.1.1. *Outlier* SNP-ovi unutar obje regije pod selekcijom na 1. autosomalnom kromosomu (bazirani na  $-\log_{10} p_{eROHi}$  vrijednosti)

### 5.2.2. CanFam 1: 60.93 Mb – 61.02 Mb

Drugi po redu od identificiranih signala također se nalazio na 1. kromosomu. Regija od 60.93 Mb do 61.02 Mb kromosoma 1. sadrži ukupno 36 *outlier* SNP-ova te je unutar nje uspješno identificiran jedan gen: GJA1. Od ukupno 36 *outlier*-a, 4 su posjedovala jednake  $-\log_{10} p_{eROHi}$  vrijednosti ( $-\log_{10} p_{eROHi} = 6,713364689$ ) te su pozicionirani na 60.93 Mb, 60.94 Mb, 61.01 Mb i 61.02 Mb (Slika 5.2.1.1.).

### 5.2.3. CanFam 3: 31.09 Mb – 32.14 Mb

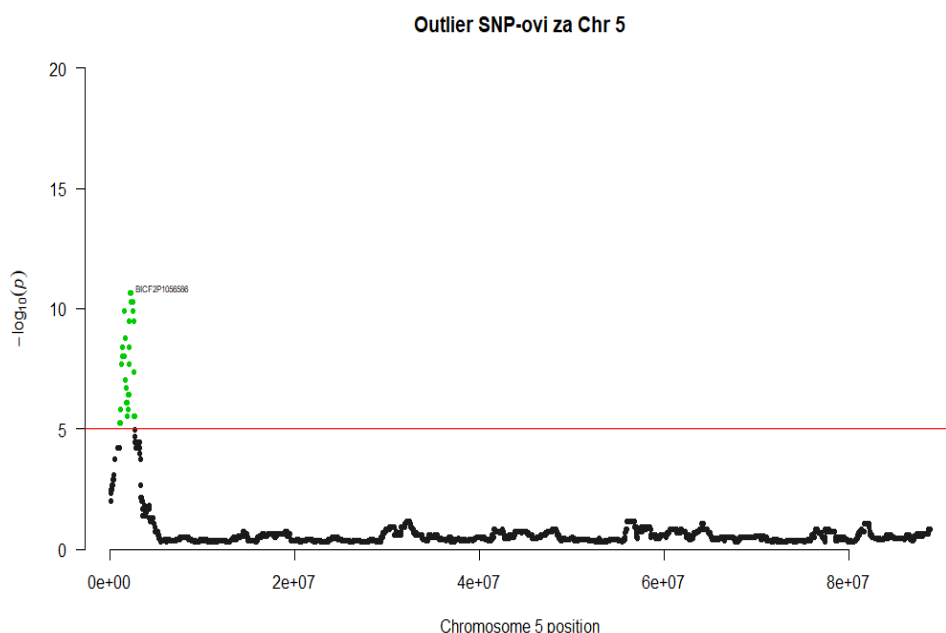
Nadalje, regija na 3. autosomalnom kromosomu (31.09 Mb – 32.14 Mb) sadrži 18 *outlier* SNP-ova (Slika 5.2.3.1.), te je kao takva okarakterizirana kao signal selekcije. Nakon provedenih eROHi analiza, unutar formirane regije signala naknadno je uspješno identificirano 19 gena. Njih jedanaest nema poznate funkcije i imena. Preostalih osam su: ANKRD31, GCNT4, MIR8901, COR6C40, COR6C30, NIPA1, NIPA2 te CYFIP1.



Slika 5.2.3.1. *Outlier* SNP-ovi na 3. autosomalnom kromosomu (bazirani na  $-\log_{10} P_{eROHi}$  vrijednosti), a najznačajniji *outlier* jest G327f18S270 ( $-\log_{10} p_{eROHi} = 6,097981412$ )

#### 5.2.4. CanFam 5: 1.08 Mb – 2.67 Mb

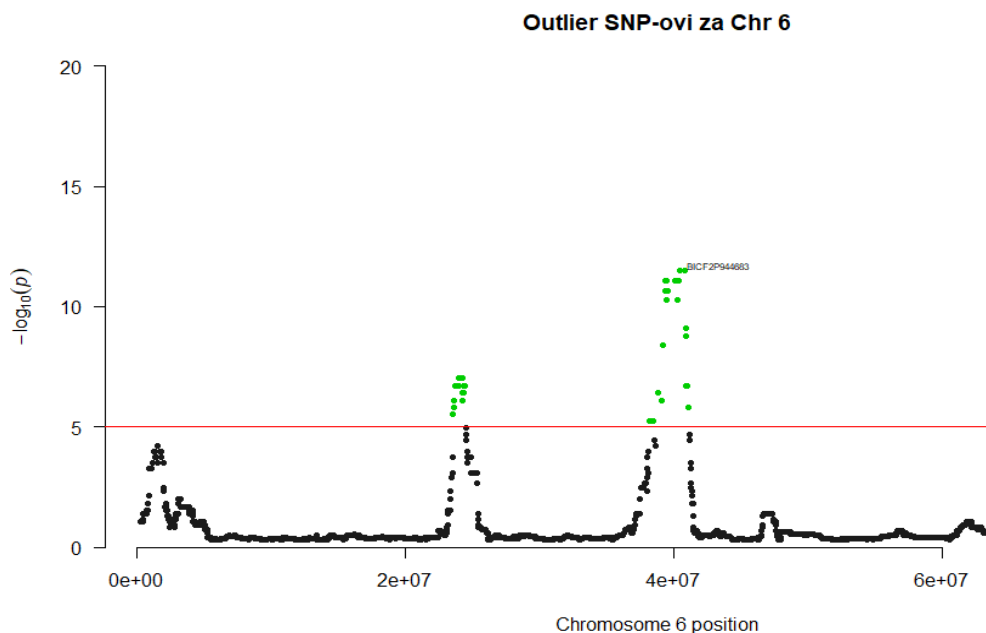
Na 5. autosomalnom kromosomu nešto manja regija od svega 1.59 Mb je identificirana kao signal selekcije (Slika 5.2.4.1.). Unutar nje pozicionirano je nešto više *outlier* SNP-ova nego na prethodnim kromosomima, njih 66, od kojih je njih 13 poprimilo iste  $-\log_{10} p_{eROHi}$  vrijednosti. Unutar regije također je identificirano 11 gena: JAM3, IGSF9B, SPATA19, OPCML i ostalih 7 s nepoznatim imenima i funkcijama.



Slika 5.2.4.1. *Outlier* SNP-ovi na 5. autosomalnom kromosomu (bazirani na  $-\log_{10} p_{eROHi}$  vrijednosti)

### 5.2.5. CanFam 6: 38.13 Mb – 41.07 Mb

Kromosom 6 jest još jedan od kromosoma unutar kojeg su identificirane 2 regije pod selekcijom. Prva regija proteže se od 38.13 Mb do 41.07 Mb i sadrži 25 *outlier* SNP-ova (Slika 5.2.5.1.). Unutar nje nalazi se poznatih 132 gena. Od kojih su najvažniji: PKMYT1, TBC1D24, CCFN, MLST8, RAB26, PKD1, NDUFB10, MRPS34, RHDF1, CACNA1H, SOX8, WDR24, MRPL28, HBM i NPRL3.



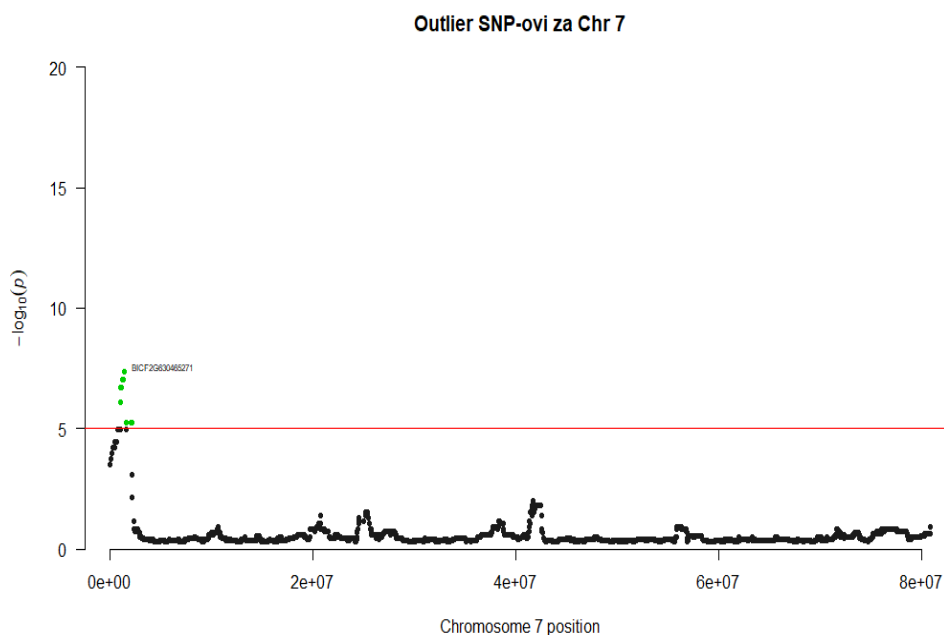
Slika 5.2.5.1. *Outlier* SNP-ovi unutar prve regije pod selekcijom na 6. autosomalnom kromosomu (bazirani na  $-\log_{10} p_{eROHi}$  vrijednosti)

### 5.2.6. CanFan 6: 23.51 Mb – 24.40 Mb

Kao što je prethodno napomenuto, još jedna regija pod selekcijom nalazi se na 6. autosomalnom kromosomu, no za razliku od prošle, ova regija je nešto manja. Unutar nje pozicionirano je 28 *outlier* SNP-ova, od kojih je najznačajniji BICF2P754953 pozicioniran na 24.15 Mb. Prikaz svih *outliera* nalazi se na slici (slika 5.2.5.1.) u poglavlju 5.2.5. Nadalje, unutar regije identificirano je 8 potencijalnih kandidat gena: DENND2D, CEPT1, DRAM2, MIR8799E-1, LRIF1 i CD53.

### 5.2.7. CanFam 7: 1.38 Mb – 2.14 Mb

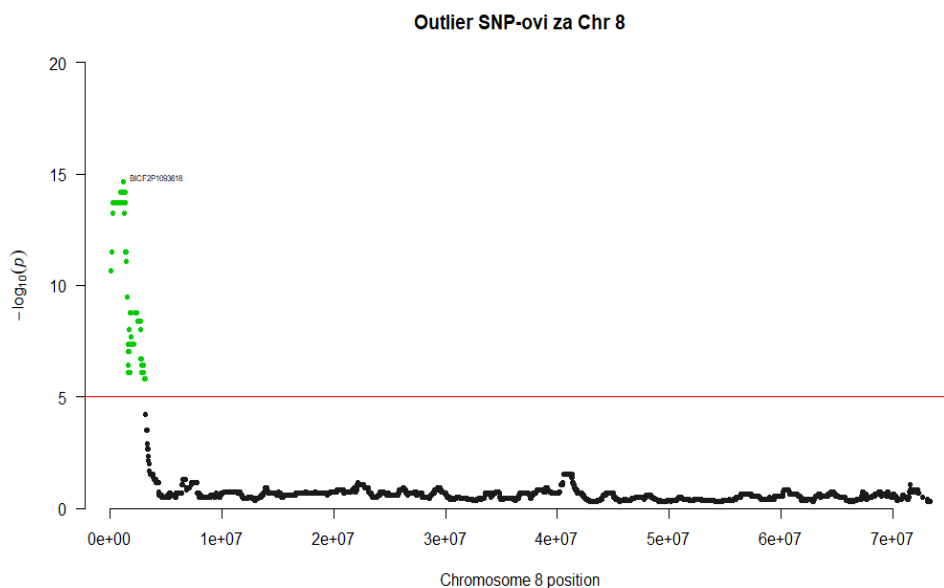
Provedbom ROH analiza i naknadnim izračunom  $-\log_{10} p_{eROHi}$  vrijednosti za SNP-ove, na 7. kromosomu identificirana je regija pod selekcijom. Unutar nje nalazi se manji broj *outlier-a* od kojih su najznačajniji BICF2G630465271, BICF2P460595 i BICF2G630465273 s  $-\log_{10} p_{eROHi} = 7,360878331$ . Svi *outlier-i* pozicionirani na 7. autosomalnom kromosomu prikazani su na slici (slika 5.2.7.1.). Dvanaest gena je također uspješno identificirano unutar regije, a oni su: LMOD1, SHISA4, IPO9, NAV1, CSRP1, TNNT2, LAD1, TNNI1, PKP1, TMEM9, ASCL5 i CACNA1S.



Slika 5.2.7.1. *Outlier* SNP-ovi na 7. autosomalnom kromosomu (bazirani na  $-\log_{10} p_{eROHi}$  vrijednosti)

### 5.2.8. CanFam 8: 126.54 Kb – 3.13 Mb

Najveći signal selekcije otkriven je na 8. kromosomu. Regija sadrži ukupno 11 *outliera* (Slika 5.2.8.1.), od kojih 11 poprima jednake  $-\log_{10} p_{eROHi}$  vrijednosti ( $-\log_{10} p_{eROHi} = 14,17240645$ ). S obzirom na veličinu regije, možemo reći da je broj identificiranih gena unutar nje malen. Ukupno je identificirano 19 gena: ZNF496, NLRP3, GCSAML, CATSPERB, TC2N, FBLN5, TRIP11, ATXN3, CPSF2, SCL24A4, RIN3, LFMN, GOLGA5, CHGA, ITPK1, DAD1, ABHD4, OXA1L i SLC7A7.



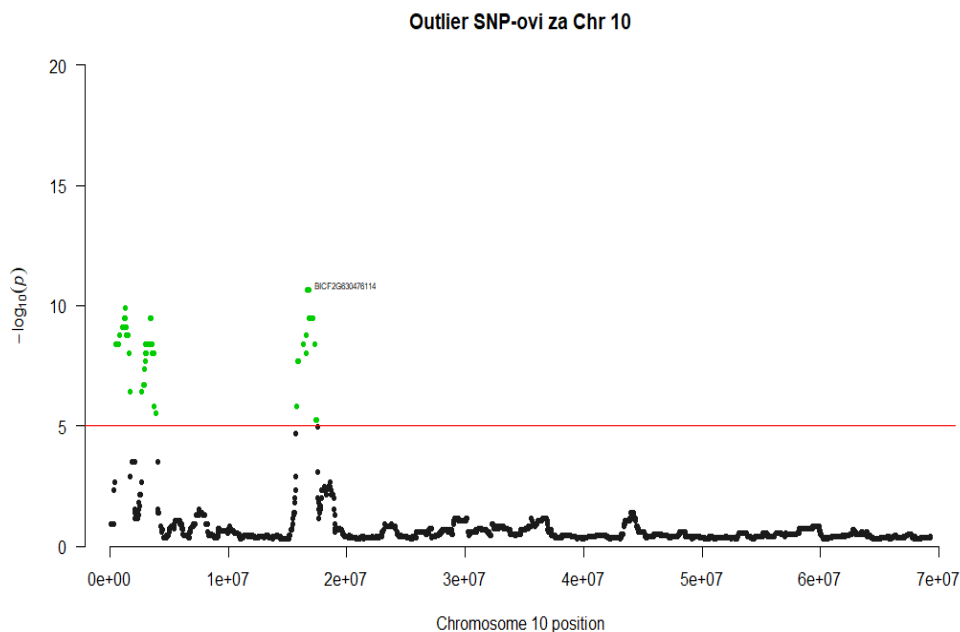
Slika 5.2.8.1. *Outlier* SNP-ovi na 8. autosomalnom kromosomu (bazirani na  $-\log_{10} p_{eROHi}$  vrijednosti)

### 5.2.9. CanFam 10: 525.75 Kb – 3.89 Mb

Ukupno 36 *outlier* SNP-ova s  $-\log_{10} p_{eROHi} > 5$  nalazi se unutar jedne od identificiranih regija pod selekcijom na 10. kromosomu. Njih 5 poprima iste vrijednosti i pozicionirani su na 1.17 Mb, 1.23 Mb, 1.27 Mb i 3.41 Mb. Također, regija sadrži 64 uspješno identificiranih gena. Neki od najvažnijih su: SLC39A5, COQ10A, IL23A, STAT2, TIMELESS, MIP, RDH16, MYO1A, STAT6, R3HDM2, INHBC, INHBE, MARS, DDIT3, AGAP2, CYP27B1 i LRIG3 (slika 5.2.10.1.).

### 5.2.10. CanFam 10: 15.74 Mb – 17.45 Mb

Druga regija pod selekcijom na 10. kromosomu proteže se od 15.74 Mb do 17.45 Mb i ukupno sadrži 18 *outlier* SNP-ova. Najznačajniji među njima jest BICF2G630476114 ( $-\log_{10} p_{eROHi} = 5,23492256$ ), koji je zajedno sa svim ostalim *outlier* SNP-ovima prve i druge regije prikazan na slici (slika 5.2.10.1.). Za razliku od prve regije, unutar ove regije uspješno je identificiran nešto manji broj gena, njih 39: KCNC2, CAPS2, GLIPR1L1, GLIPR1L2, GLIPR1, RABL2B, ACR, SHANK3, ARSA, MAPK8IP2, CHKB, CPT1B, SYCE3, KLHDC7B, ODF3B, NCAPH2, LMF2, MIOX, ADM2, SBF1, PPP6R2, DENND6B, PLXNB2, MAPK11, MAPK12, HDAC10, TUBGCP6, SELENOO, TRABD, PANX2, MOV10L1, MLC1, TILL8, IL17REL, PIM3, CRELD2, ALG12, ZBED4 i BRD1.

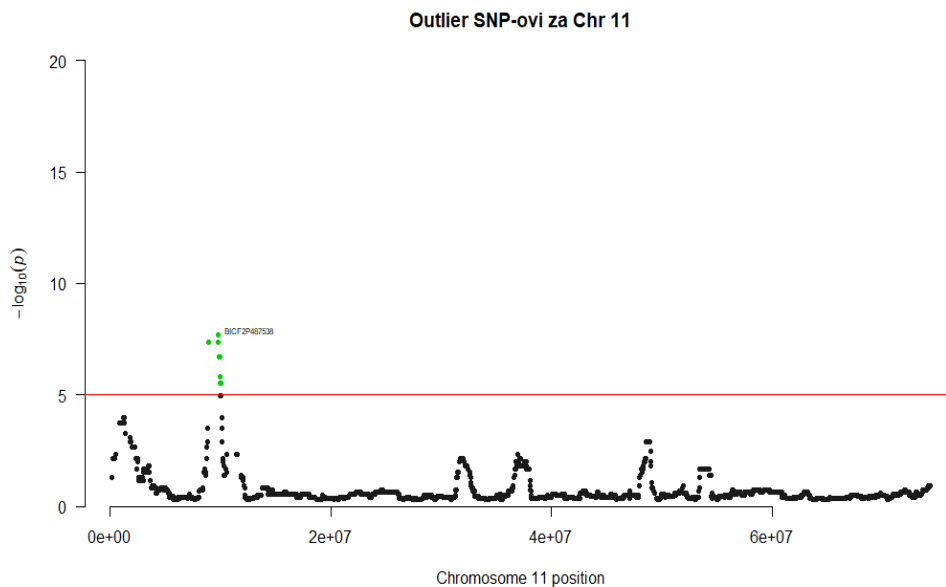


Slika 5.2.10.1. *Outlier* SNP-ovi unutar obje regije pod selekcijom 10. autosomalnom kromosomu (bazirani na  $-\log_{10} p_{eROHi}$  vrijednosti)

### 5.2.11. CanFam 11: 8.89 Mb – 10.09 Mb

Regija s najmanje *outlier* SNP-ova identificirana je na 11. kromosomu te je njezina duljina 1.2 Mb. Dvanaest SNP-ova s  $-\log_{10} p_{eROHi}$  vrijednosti  $> 5$  pozicioniralo

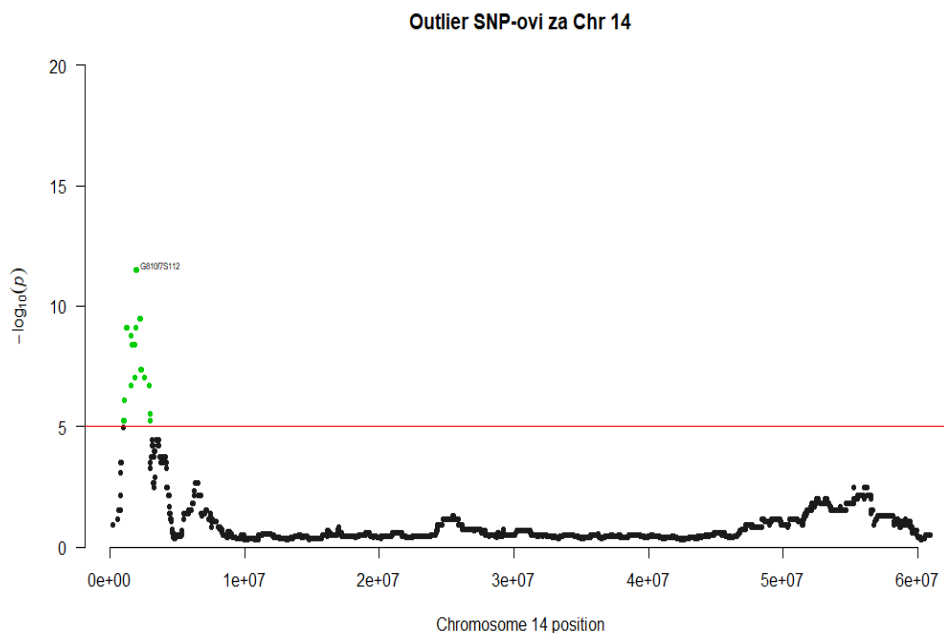
se unutar spomenute regije, a najznačajniji među njima jest BICF2P487538 ( $-\log_{10} p_{eROHi} = 7,696713$ ). Detaljnijim uvidom u regiju, uspješno je identificirano 32 gena, od čega samo jedan ima poznato ime i funkciju: PRR16. Svih 12 *outlier* SNP-ova prikazano je na slici (slika 5.2.11.1) zajedno s oznakom najznačajnijeg.



Slika 5.2.11.1. *Outlier* SNP-ovi unutar regije pod selekcijom 11. autosomalnog kromosoma (bazirani na  $-\log_{10} p_{eROHi}$  vrijednosti)

## 5.2.12. CanFam 14: 1.01 Mb – 2.98 Mb

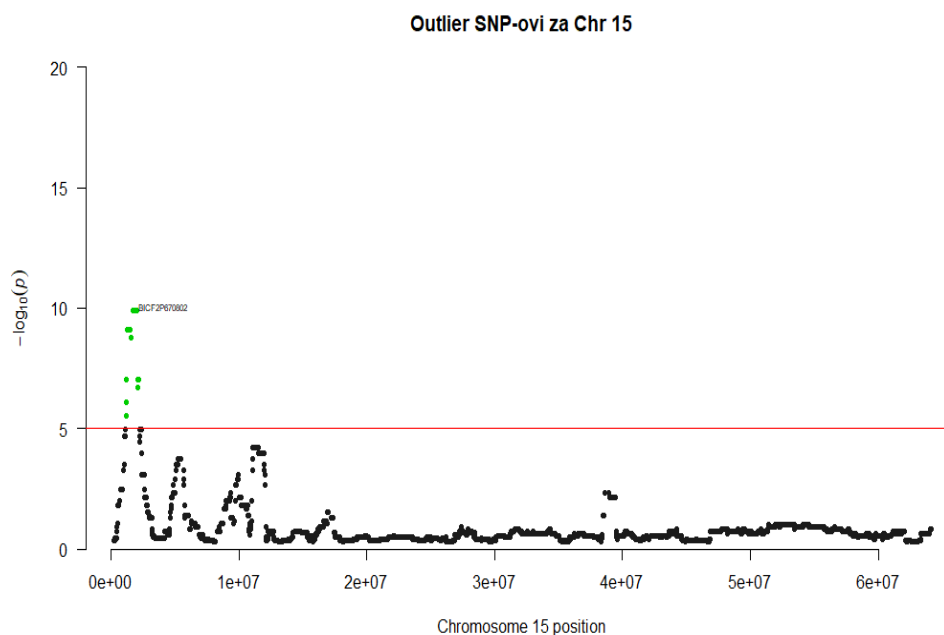
Nadalje, regija s 30 *outlier* SNP-ova identificirana je kao potencijalni signal selekcije na 14. kromosomu. G810f7S112, BICF2P27763 i BICF2P983194 SNP-ovi su koji poprimaju jednake, relativno visoke  $-\log_{10} p_{eROHi}$  vrijednosti od 11,50008288 te vjerojatno predstavljaju regiju unutar koje se nalazi uzročnik selekcije. Blisko su pozicionirani na kromosomu, a njihove točne pozicije su 1.94 Mb, 1.96 Mb te 1.99 Mb. Svi *outlier* SNP-ovi prikazani su na slici dolje (slika 5.2.12.1.). Istraživanjem regije uspješno je identificirano 13 gena: SNAP47, JMJD4, TRIM58 te skupina gena mirisnih receptora, COR2L17, COR2AV2, COR2T24, COR2T15, COR2T13, COR2T18, COR2G5, OR04C09, OR2M9 i OR2T2.



Slika 5.2.12.1. *Outlier* SNP-ovi unutar regije pod selekcijom na 14. autosomalnom kromosomu (bazirani na  $-\log_{10} p_{eROHi}$  vrijednosti)

### 5.2.13. CanFam 15: 1.16 Mb – 2.15 Mb

Analize eROHi identificirale su regiju od 1.16 MB do 2.15 Mb na 15. autosomalnom kromosomu kao mogući signal pozitivne selekcije. Regija sadrži 21 *outlier* SNP od kojih sedam poprima jednake vrijednosti te tvore niz *outlier* SNP-ova na kojemu se temeljila identifikacija signala selekcije. Uz 21 *outlier*, unutar regije uspješno su identificirana i 7 gena s poznatim nazivom i ostali bez poznate funkcije i naziva. Identificirani geni su: SLFN1, SCMH1, MIR8839, HIVEP3, FOXO6, EDN2 i CTPS1. Svi *outlier* SNP-ovi su vizualizirani i označeni (zelenom bojom) na Manhattan plot-u (slika 5.2.13.1.).

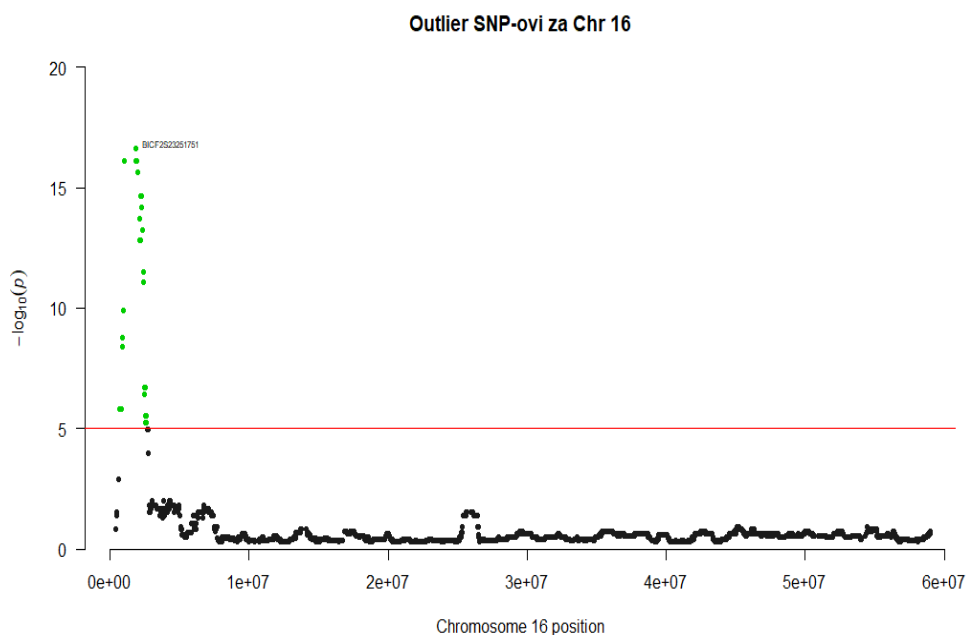


Slika 5.2.13.1. *Outlier* SNP-ovi unutar regije pod selekcijom na 15. autosomalnom kromosomu (bazirani na  $-\log_{10} p_{eROHi}$  vrijednosti)

## 5.2.14. CanFam 16: 692.73 Kb – 2.6 Mb

Potencijalni signal selekcije javio se i na 16. kromosomu, unutar regije od 692.73 Kb do 2.6 Mb. Unutar same regije nalazi se ukupno 40 *outlier* SNP-ova. Neki od njih kao što su BICF2S23251751, BICF2P280451, BICF2P684798, BICF2S231219, BICF2P1288995, BICF2P187620 i BICF2P502785 poprimaju jednake  $-\log_{10} p_{eROHi}$  vrijednosti, a na temelju njihove pozicije može se tumačiti uzročnik selekcije. Kao i do sada, svi *outlier* SNP-ovi zajedno s njihovom  $-\log_{10} p_{eROHi}$  vrijednošću vizualizirani su pomoću Manhattan plot-a (slika 5.2.14.1.). Upotrebom GenomeBrowse paketa unutar GoldenHelix softvera (Golden Helix, Inc., Bozeman, MT, [www.goldenhelix.com](http://www.goldenhelix.com)), uspješno je identificirano 11 gena u ovoj regiji: PDIA4, IGFBP1, EZH2, IGFBP3, CUL1, CNTNAP2, ADCY1 i skupinu gena *microRNA* molekule (miRNA) koji se nalaze unutar stanica mliječne žlijezde: MIR8862, MIR8853, MIR8852 te MIR8793-1.

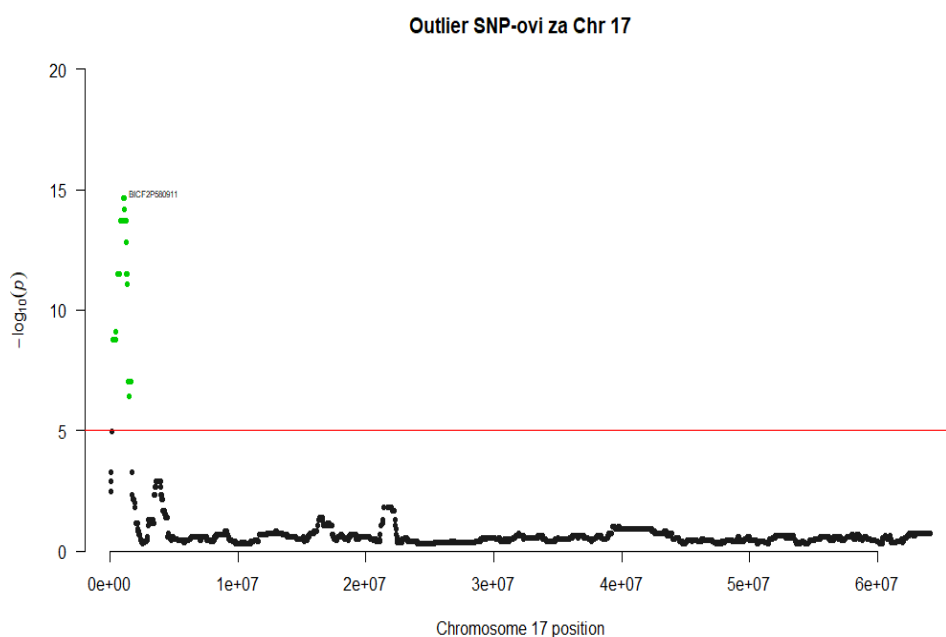




Slika 5.2.14.1. *Outlier* SNP-ovi unutar regije pod selekcijom na 16. autosomalnom kromosomu (bazirani na  $-\log_{10} p_{eROHi}$  vrijednosti)

### 5.2.15. CanFam 17: 229.18 Kb – 1.68 Mb

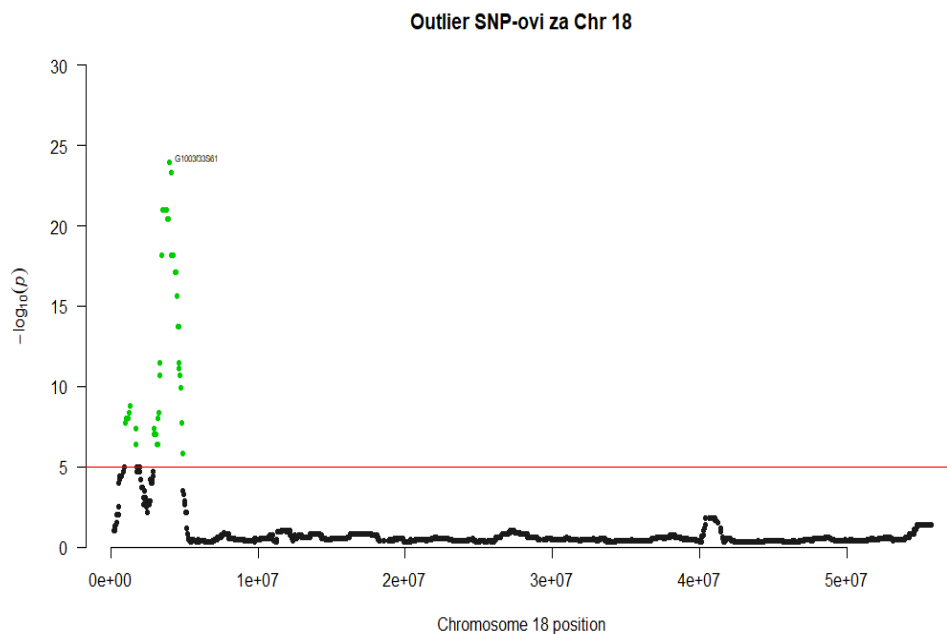
Sljedeća po redu regija pod selekcijom jest regija na 17. kromosomu, unutar koje se nalazi 27 *outlier* SNP-ova i 7 uspješno identificiranih gena. Najznačajniji SNP-ovi su BICF2P543628, BICF2S23632105, BICF2S23619658 i BICF2S23653084 s  $-\log_{10} p_{eROHi} = 13,70662$ . Identificirani geni su: TPO, TMEM8, SNTG2, SH3YL1, PXDN, MYT1L i ACP1. Regija i *outlier* SNP-ovi prikazani su na slici (slika 5.2.15.1.).



Slika 5.2.15.1. *Outlier* SNP-ovi unutar regije pod selekcijom na 17. autosomalnom kromosomu (bazirani na  $-\log_{10} p_{eROHi}$  vrijednosti)

### 5.2.16. CanFam 18: 954.32 Kb – 4.89 Mb

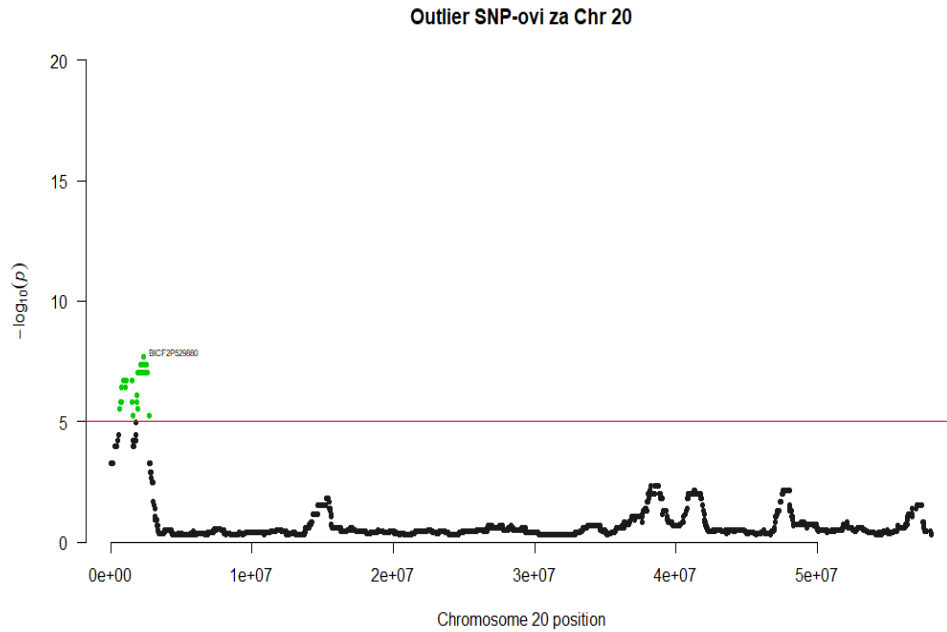
Nešto povećana regija na 18. kromosomu s većim brojem *outlier* SNP-ova, identificirana je kao signal selekcije. Ukupno 52 *outlier*-a smješteno je unutar ove regije, a najznačajniji među njima su G1003f33S61 i BICF2P960438 na pozicijama 3.98 Mb i 4.08 Mb. Vizualni prikaz regije i *outlier* SNP-ova nalazi se dalje u tekstu (slika 5.2.16.1.). Nadalje, unutar regije identificirano jest 8 gena s poznatim imenima i ostali geni s nepoznatim. Geni s poznatim imenima su: ZPBP, VWC2, SPATA48, IKZF1, GRB10, FIGNL1, DDC i COBL.



Slika 5.2.16.1. *Outlier* SNP-ovi unutar regije pod selekcijom na 18. autosomalnom kromosomu (bazirani na  $-\log_{10} p_{eROHi}$  vrijednosti)

### 5.2.17. CanFam 20: 581.21 Kb – 2.68 Mb

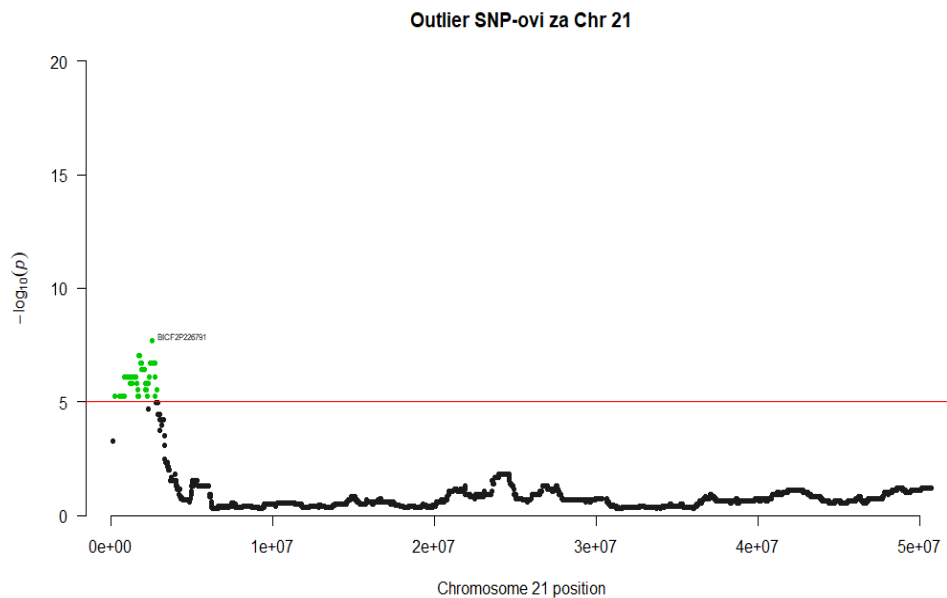
Još jedna od regija pod selekcijom, identificirana na temelju rezultata eROHi analiza, nalazi se na 20. kromosomu i sadrži 41 *outlier* SNP. Veći broj *outlier* SNP-ova ( $n = 12$ ) poprima jednake  $-\log_{10} p_{eROHi}$  vrijednosti. Unutar nje, uz veći broj *outlier*-a, pronađen je i veći broj gena, njih 21: ZXDC, UROC1, TXNRD3, TRPA1, SEC61A1, RUVBL1, RPN1, PODXL2, PLXNA1, MGLL, MCM2, KBTBD12, GATA2, EEFSEC, DNAJB8, CHST13, CHCHD6, CFAP100, C20H3orf56, C20H3ORF22 i ABTB1. Svi *outlier*-i i cjelokupna regija signala, prikazani su na slici (slika 5.2.17.1.).



Slika 5.2.17.1. *Outlier* SNP-ovi unutar regije pod selekcijom na 20. autosomalnom kromosomu (bazirani na  $-\log_{10} p_{eROHi}$  vrijednosti)

### 5.2.18. CanFam 21: 269.07 Kb – 2.82 Mb

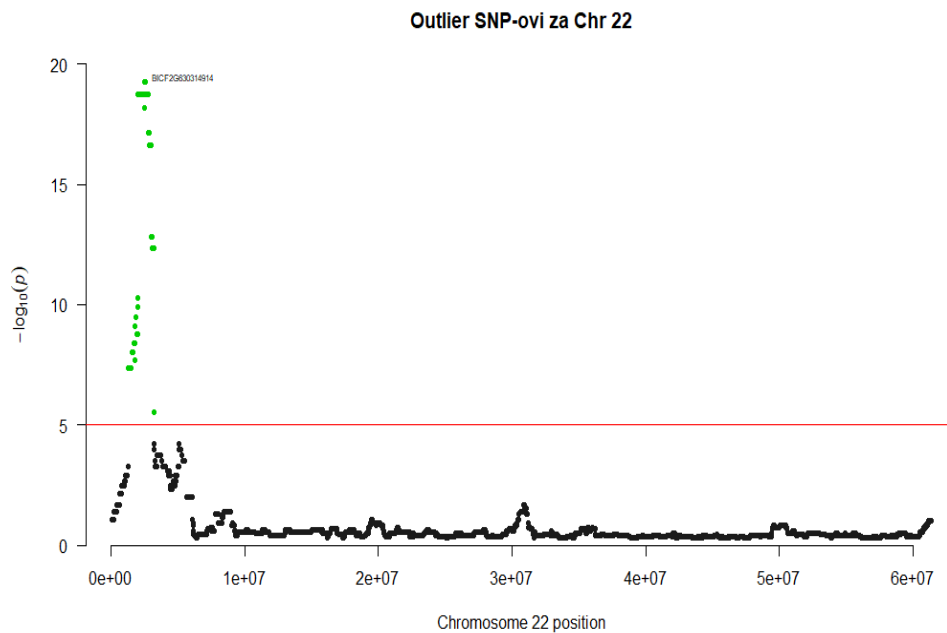
Regija koja se proteže od 269.07 Kb do 2.82 Mb na 21. kromosomu također se karakterizira signalom selekcije (slika 5.2.18.1.). Na pozicijama 2.12 Mb, 2.140 Mb, 2.147 Mb, 2.17 Mb i 2.33 Mb nalaze se *outlier* SNP-ovi s najsignifikantnijim  $-\log_{10} p_{eROHi}$  vrijednostima ( $-\log_{10} p_{eROHi} = 5,802307$ ). Iako je regija relativno velika, unutar nje identificirano je samo 3 gena: PGR, CNTN5 i ARHGAP42.



Slika 5.2.18.1. *Outlier* SNP-ovi unutar regije pod selekcijom na 21. autosomalnom kromosomu (bazirani na  $-\log_{10} p_{eROHi}$  vrijednosti)

### 5.2.19. CanFam 22: 1.31 Mb – 3.22 Mb

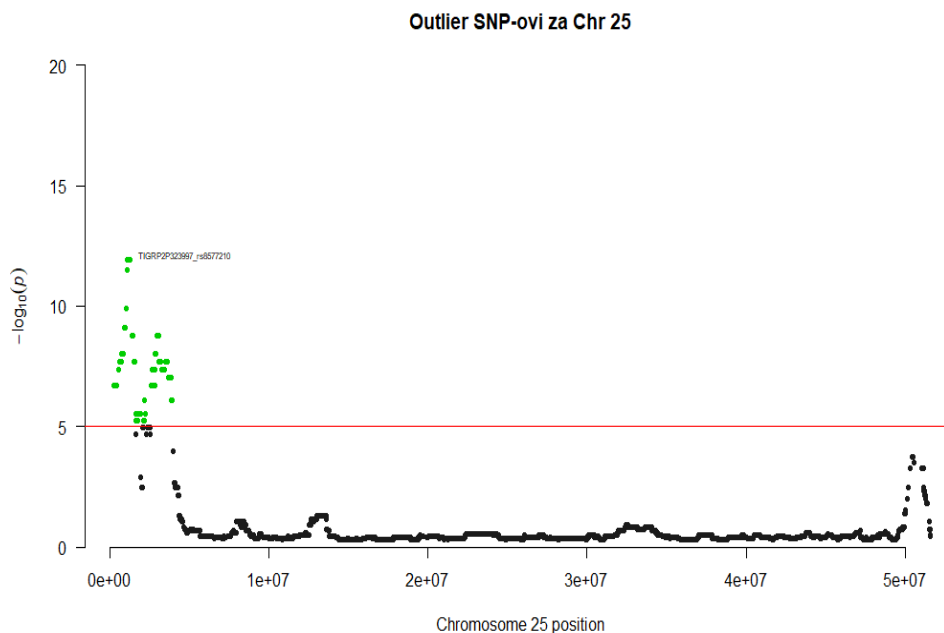
Jedna od regija s najvećim brojem *outlier* SNP-ova nalazi se na 22. kromosomu. Naime, unutar nje pozicionirano je 101 *outlier* SNP-ova, a najznačajniji među njima poprimili su vrijednosti  $-\log_{10} p_{eROHi} = 19,28023645$ . Zbog velikog broja *outliera* s ekstremnim  $-\log_{10} p_{eROHi}$  vrijednostima, regija se identificira kao signalom selekcije (slika 5.2.19.1.). Unutar nje također se nalazi i 18 poznatih gena: TRIM13, SPRYD7, SETDB2, RCBTB2, RCBTB1, RB1, PHF11, MLNR, MIR16-A, MIR15A, LPAR6, KPNA3, KCNRG, FNDC3A, EBPL, CYSLTR2, CDADC1 i CAB39L.



Slika 5.2.19.1. *Outlier* SNP-ovi unutar regije pod selekcijom na 22. autosomalnom kromosomu (bazirani na  $-\log_{10} p_{eROHi}$  vrijednosti)

### 5.2.20. CanFam 25: 223.81 Kb – 3.94 Mb

Najviše *outlier* SNP-ova (  $n = 124$ ) pozicionirano je na 25. kromosomu. Posljedično tome regija od 223.81 Kb do 3.94 Mb smatra se signalom selekcije (slika 5.2.20.1.). Šest je najznačajnijih SNP *outliera* s jednakim vrijednostima  $-\log_{10} p_{eROHi} = 11,92509$ . U regiji duljine 3.71 Mb, uspješno je identificirano 17 gena: UFM1, TRPC4, SUPT20H, STOML3, SMAD9, SLC25A15, RFXAP, PROSER1, POSTN, NHLRC3, MRPS31, LHFPL6, FREM2, FOXO1, EXOSC8, COG6 i ALG5.



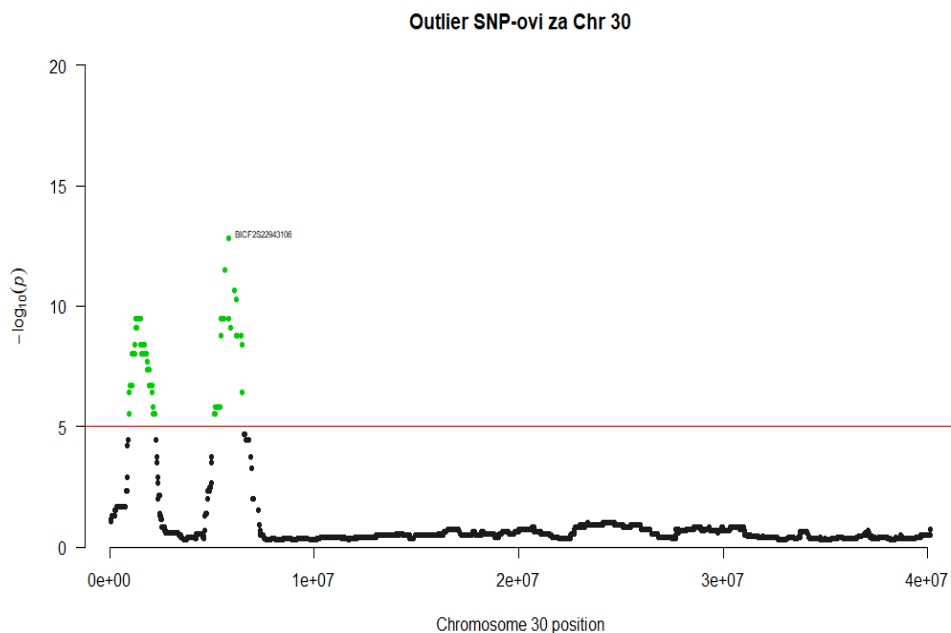
Slika 5.2.20.1. *Outlier* SNP-ovi unutar regije pod selekcijom na 25. autosomalnom kromosomu (bazirani na  $-\log_{10} p_{eROHi}$  vrijednosti)

### 5.2.21. CanFam 30: 927.94 Kb – 2.22 Mb

Prva od ukupno dvije regije koje prikazuju značajne signale selekcije na 30. kromosomu, duljine je od 927.94 Kb do 2.22 Mb i sadrži 90 *outlier* SNP-ova (slika 5.2.22.1.). Najznačajniji SNP, BICF2S22943106 ostvario je vrijednost  $-\log_{10} p_{eROHi} = 12,79954094$ . Gena s poznatim imenima koji su identificirani unutar regije su: RYR3, GREM1, FMN1, EMC7, CHRM5 i AVEN.

### 5.2.22. CanFam 30: 5.09 Mb – 6.44 Mb

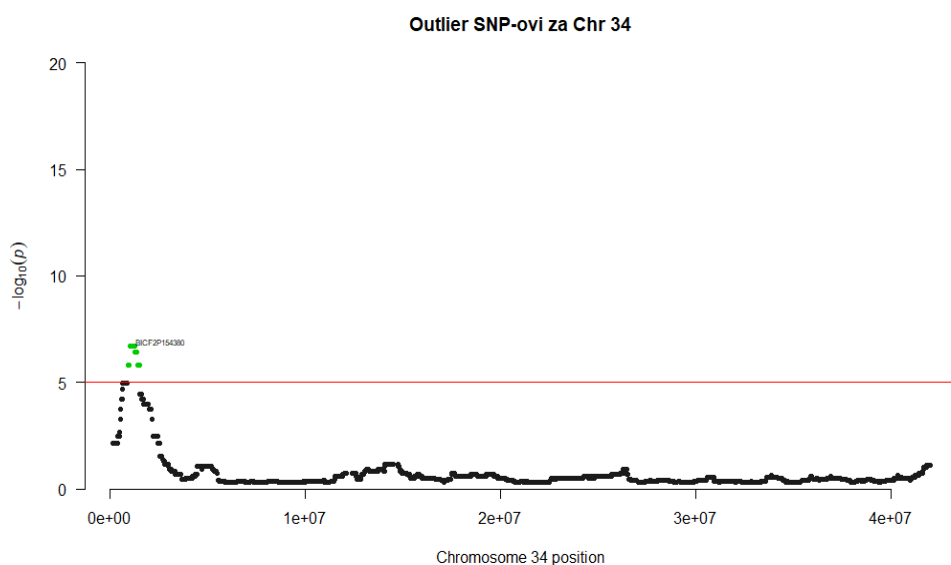
Druga regija na 30. kromosomu koja pokazuje signale selekcije sadrži 21 *outlier* SNP. Unutar regije od 1.35 Mb identificiran je samo jedan gen naziva DNAH5. Ova regija, kao i prethodna regija na istome kromosomu prikazana je na slici (slika 5.2.22.1.).



Slika 5.2.22.1. *Outlier* SNP-ovi unutar obje regije pod selekcijom na 30. autosomalnom kromosomu (bazirani na  $-\log_{10} p_{eROHi}$  vrijednosti)

### 5.2.23. CanFam 34: 917.49 Kb – 1.50 Mb

Zadnja regija identificirana kao signali selekcije putem analiza eROHi-a je regija od 917.49 Kb do 1.50 Mb na 34. kromosomu. Ukupno 37 *outlier* SNP-ova pozicionirano je kroz cijelu regiju, a najznačajnijih SNP-ova s jednakim  $-\log_{10} p_{eROHi}$  vrijednostima bilo je čak četrnaest ( $-\log_{10} p_{eROHi} = 6,713364689$ ). Cjelokupna regija unutar koje se nalaze *outlier* SNP-ovi zajedno s identifikacijskom oznakom jednog od njih, vizualizirana je na Manhattan plot-u (slika 5.2.23.1.).



Slika 5.2.23.1. *Outlier* SNP-ovi unutar regije pod selekcijom na 34. autosomalnom kromosomu (bazirani na  $-\log_{10} p_{eROHi}$  vrijednosti)

## 6. Rasprava

Kao što je već spomenuto unutar 4.4. poglavlja, u daljnjem tekstu navedene su i opisane sve poznate funkcije nekih od identificiranih gena unutar regija pod selekcijom. Opis gena i njihovih funkcija podijeljen je u 2 potpoglavlja. Potpoglavlje 6.1. sadrži opisane funkcije gena unutar regija pod selekcijom identificiranih putem iHS analiza, nakon čega slijedi potpoglavlje 6.2. s definiranim funkcijama gena unutar regija koje su identificirane kao signali selekcije na temelju eROHi analiza. Funkcije pojedinih gena istražene su pomoću UniProt (<https://www.uniprot.org>) i GeneCards (<https://www.genecards.org>) platformi kao i dostupne literature. Geni koji nisu navedeni unutar ovog poglavlja, nažalost nemaju dovoljno dostupnih informacija o njihovoj funkciji unutar psećeg genoma.

### 6.1. Funkcionalnost pojedinih gena identificirani unutar iHS analiza

#### 6.1.1. Geni povezani sa unutar staničnim mehanizmima

Prvi od identificiranih gena je RFLX6 gen. Ovaj gen pripada u grupu regulatornih faktora te unutar organizma kodira protein uključen u kontrolu transkripcije. Kod ljudi je najznačajniji u procesu reguliranja transkripcije  $\beta$  stanica (stanice Langerhansovih otočića zadužene za lučenje inzulina i amilina) kao i u reguliranju njihove funkcije. Gen nije toliko istražen kod pasa, ali je bio ključan gen unutar istraživanja koje su proveli Mahmoud i sur. 2018. godine. Istraživanje je za cilj imalo proučiti ulogu stanica ljudske koštane srži u kontroli pojave dijabetesa kod pasa. Samo istraživanje se temeljilo na takozvanom *Mitchell-Riley* sindromu koji se karakterizira neonatalnim dijabetesom s hipoplazijom gušterače, a nastaje kao posljedica mutacije RFLX6 gena.

PLN gen je također jedan od najvažnijih, odnosno interesantnijih gena identificiranih unutar regija pod selekcijom. Zadužen je za kodiranje proteina koji sudjeluju u normalnom prijenosu  $Ca^{2+}$  iona unutar mišićja, posebice unutar srčanog mišićja, kontrolom *sarco (endo) plasmic reticulum  $Ca^{2+}$ -ATPase* (SERCA2a). Unutar humanih i animalnih modela srčanih zastoja, SERCA-PLN inhibitorni kompleksi su povišeni (H. MacLennan i sur. 2003.). Uz ta saznanja, možemo reći da je zbog iznimno zahtjevne namjene pastirskih pasmina pasa, normalno funkcioniranje ovog gena vrlo važno kako bi pasmine mogle podnijeti veće radne napore i dugotrajan posao zaštite stoke.

MCM9 gen, punog naziva *minichromosome maintenance 9* unutar psećeg genoma djeluje u procesu homolognog popravljavanja greški koje se mogu javiti pri rekombinacijama.

Nadalje, TSC22D1 gen koji pripada porodici TSC22 gena unutar psećeg organizma kodira proteine koji sudjeluju u regulaciji vezanja aktivatora transkripcije na DNA molekulu, sudjeluju u pozitivnim i negativnim apoptozama te u pozitivnim regulacijama

stanične proliferacije, a samim time mutacije gena mogu dovesti do pojave tumora mliječne žlijezde (Gene Ontology Consortium: going forward 2015.).

Posljednji, gen SERP2 jest gen koji kodira takozvane *Stress-associated endoplasmic reticulum protein*, zadužene za stresni odgovor. Proteine koje kodira ovaj gen moguće je i povezati s pojavom epilepsije kod pasa. Upravo to dokazano je u istraživanju koje su proveli Nowakowska i sur. (2020.). Unutar istraživanja korištene su jedinke miševa i pasa. Imunohistokemijskim bojanjem označile su se regije mozga miševa koji su doživjeli epilepsiju kako bi uz pomoću njih potakli epileptički napadaji kod psa. Nakon toga, provedena su imunohistokemijska bojanja regija mozga psa, te je utvrđena prisutnost *Heat shock* proteina koji su inhibitori SERP2 gena.

### 6.1.2. Geni povezani s anatomskim i histološkim razvojem organizma

VGLL2 gen je također jedan od gena identificiranih unutar regije pod selekcijom. Ovaj gen kodira protein s *transcriptional enhancer factor-1* (TEF-1) interakcijskom domenom koji sudjeluje u normalnom razvoju skeletnog mišićja kod sisavaca. Ukoliko se osvrnemo na poglavlje 2.4. i na činjenicu da su pastirske pasmine pasa uglavnom većih tjelesnih proporcija, možemo reći da je VGLL2 gen najvjerojatnije jedan od gena zaduženih za ekspresiju takvih morfoloških karakteristika. Yang i sur. (2021.) proveli su istraživanje unutar kojega su utvrdili povezanost između 12 transkriptnih faktora i 167 gena (uključujući i VGLL2 gen) te su uspješno kreirali takozvanu *putative gene regulatory* mrežu koja objašnjava razvoj skeletnog mišićja sisavaca.

Kada je u pitanju razvoj mišićne mase, potrebno je izdvojiti i ROS1 gen identificiran unutar regije pod selekcijom. Gen pripada potporodici tirozin-kinaza inzulinskim receptorima te kodira integrirani membranski protein koji potiče aktivnost tirozin kinaze. Uzmemo li u obzir da su tirozin kinaze enzimi uključeni u prijenos signala unutar stanice, regulaciju njihova rasta i dr., gen se povezuje direktno s razvojem mišićja i drugih tkiva kod svih organizama, pa tako i kod pasa.

GJA1 gen je također jedan od gena koji sudjeluje u normalnom funkcioniranju srčanog mišića. Naime ovaj gen kodira proteine koji sudjeluju u procesu vezivanja  $\alpha$  i  $\beta$  tubula, kao i u razvoju i funkcioniranju srčanog mišića kod odraslih jedinki pasa (Gene Ontology Consortium: going forward 2015.).

Još jedan od gena zanimljive funkcije jest TBC1D32 gen koji sudjeluje u normalnom razvoju vida. Odnosno gen kodira protein koji dovodi do potpunog razvoja oka i očnih živaca te samim time jedinki omogućuje, u ovom slučaju, kontroliranje stoke i predatora na velikim otvorenim površinama. Mogućnost jedinke pastirske pasmine da na vrijeme uoči bilo kakvu prijetnju stadi, jest jedna od važnijih karakteristika u radnoj namjeni pasmine.

SMAD6 i SMAD3 geni pripadaju u posebnu skupinu gena naziva *mothers against decapentaplegic*, a proteini koje kodiraju kod pasa igraju važnu ulogu u normalnom razvoju aorte i srčanog krvotoka te bilo kakve mutacije unutar gena dovode do



promjena unutar srčanih zalistaka. Kod ljudi, oni moduliraju više razvojnih signalnih puteva, a njihovo međudjelovanje regulira mezodermalnu diferencijaciju i srčanu morfogenezu (Varshney i sur. 2021.).

### **6.1.3. Geni povezani s reprodukcijom**

GOPC gen kodira protein s PDZ domenom koji je kod nekih vrsta povezan s plodnošću muških jedinki. Razna istraživanja provedena na miševima, dokazala su da mutacije unutar spomenutog gena i manjak proteina koje kodira, dovode do pojave neplodnosti i poremećaja koji se naziva *Globozoospermia*. Funkcija gena kod pasa povezana je također s plodnošću, ali i s pojavom tumora. Naime Panagopoulous (preme. Charest i sur. 2003.) navodi da kod određenih vrsta, uključujući i pse, postoji mogućnost da intersticijske delecije uzrokuju fuziju GOPC gena i ROS1 gena (GOPC-ROS1) što dovodi do poremećaja u rastu stanica i pojave tumora.

## **6.2. Funkcionalnost pojedinih gena identificiranih unutar eROHi analiza**

### **6.2.1. Geni povezani s unutarstaničnim mehanizmima**

Gen MBP, ekspresira se unutar leđne moždine i ostalih 56 tkiva. Kodira takozvani *Myelin basic* protein unutar mijelina živčanih vlakana. Koncentracija osnovnog proteina mijelina (MBP) u cerebrospinalnoj tekućini koristi se kao biokemijski marker kod ljudi za procjenu demijelinizirajućih lezija u središnjem živčanom sustavu (Oji i sur. 2008.), te su na temelju toga Oji i sur. u svome istraživanju koristili isti pristup za računanje koncentracije MBP-a u cerebrospinalnoj tekućini pasa. Rezultati istraživanja dokazali su da je udio MBP-a veći kod jedinki koje su oboljele od degenerativne mijelopatije.

Gen CYFIP1 koji je pronađen unutar identificirane regije pod selekcijom na 3. kromosomu kodira *Cytoplasmic FMR1-interacting protein 1*. Protein pripada u kompleks *CYFIP1-EIF4E-FMR1* koji se veže na kapu mRNA i posreduje u translacijskoj represiji. U kompleksu *CYFIP1-EIF4E-FMR1* ova podjedinica je adapter između EIF4E i FMR1. Potiče aktivnost potiskivanja translacije FMR1 u mozgu vjerojatno posredujući njegovu povezanost s EIF4E i mRNA (The UniProt Consortium).

JAM3 gen kodira JAM protein. JAM pripada super porodici imunoglobulina: ima jednu transmembransku domenu, a smatra se da je njegov izvanstanični dio presavijen u dvije domene slične imunoglobulinu (Itoh i sur. 2001.).

CRYM gen unutar psećeg organizma kodira protein Mu-kristalin čija je enzimska funkcija određena kao ketimin reduktaza, reducirajući cikličke ketimine u reducirane oblike.

Gen NAV1 kodira protein natrijevih kanala unutar psećeg organizma. Sami protein ima ulogu u neuropatskim mehanizmima boli.

*Cysteine and glycine rich protein 1* produkt je djelovanja CSRP1 gena u psećem genomu. Gen sudjeluje u raznim procesima vezanja iona za specifične receptore kao i u razvoju i funkcioniranju mišićja, što je kod pastirskih pasmina bitna karakteristika.

TNNT2 gen kodira takozvani „Troponin T2, cardiac type“ protein koji je zapravo podjedinica troponiozina koja veže tropomiozin kao i regulacijskog kompleksa tankih niti koje potiču osjetljivost na kalcij i aktivnost mišićne aktomiozin ATP-aze.

CACNA1S gen zadužen je za kodiranje *Voltage-dependent L-type calcium channel subunit alpha* proteina koji posreduju ulasku kalcijevih iona u specifične stanice, a također su uključeni u različite procese ovisne o kalciju, uključujući mišićnu kontrakciju, oslobađanje hormona ili neurotransmitera, ekspresiju gena, pokretljivost stanica, staničnu diobu i staničnu smrt.

Stanični odgovor na estrogen, regulacija mineralizacije kostiju, transport i homeostaza kalcijevih iona i katabolički procesi vitamina D neke su on najvažnijih funkcija gena CYP27B1.

Posebna porodica gena, takozvani *Solute carrier genes* odnosno geni čiji proteini služe prijenosu tvari i molekula također su identificirani unutar pojedinih regija na većem broju gena. Specifični geni koji su identificirani u ovom slučaju su: SLC25A15, SLC16A17, ALC26A10, SLC24A4, SLC39A5 i SLC9A3R2. Sama porodica SLC proteina sastoji se od 55 genskih obitelji s najmanje 362 gena za kodiranje proteina koji posreduju u transportu širokog spektra molekula kroz stanice (Bluemel i sur. 2015.).

GCSAML gen zadužen je za kodiranje proteina uključenih u procese transporta limfocita unutar psećeg organizma.

DAD1 kodira *Dolichyl-diphosphooligosaccharide* protein. Ovaj protein je uključen u put glikozilacije proteina odnosno njihove modifikacije.

TPO gen kodira peroksidaze štitnjače, koji sudjeluju u spajanju hormonogenih tirozina u tiroglobulinu radi dobivanja hormona štitnjače tiroksina i trijodtironina (T3 i T4).

Gen zanimljivog imena TIMELESS regulator je cirkadijanih ritmova, sudjeluje u razvoju pluća i potiče stanični odgovor na hidroksi-ureu.

HBM gen kodira podjedinicu hemoglobina *mu* koja sudjeluje u transportu kisika unutar organizma i potpomaže oksidativnoj detoksikaciji.

CACNA1H gen kodira takozvanu *Calcium voltage-gated channel* podjedinicu koja sudjeluje u unosu kalcijevih iona u stanice, staničnom odgovoru na hormonske stimulanse, stvaranju akcijskog potencijala na završecima neurona, biosintetičkim procesima kortizola i adlosterona.

MARS gen djeluje u staničnom odgovoru na gladovanje, rRNA transkripciji i staničnom odgovoru na stimulans inzulina.

SNAP47 kodira *Synaptosome associated* protein koji sudjeluje u procesu egzocitoze.

IGFBP1 i IGFBP3 su geni koji kodiraju takozvane *Insulin-like growth factor-binding*“proteine, uključene u prolongiranje života inzulin faktora rasta (Igf).

ADCY1 gen kodira denilat ciklazu, protein uključen u katalizaciju formiranja molekule cAMP-a kao rezultat signalizacije G proteina.

TRPC4 gen uključen je u prijenosu magnezijevih i kalcijevih iona kroz specifične kanale unutar mišićja.

FOXO1 gen još je jedan od najzanimljivijih identificiranih gena unutar regija pod selekcijom. Naime ovaj gen sudjeluje u raznim procesima, ali najvažniji u ovom slučaju su stanični odgovor na glad i stanični odgovor na hipoksiju i hiperoksiju. Uzmemo li u obzir da su pašnjaci na kojima psi pastirskih pasmina čuvaju stada, najčešće na otvorenim ravnicama udaljenim od grada, sela i sl. te da dosta vremena psi provode vrijeme bez vlasnika, moguće je vidjeti povezanost između spomenutog gena i bivanja na otvorenom gdje je promjenjiv tlak zraka i gdje je ponekad oskudan izvor hrane.

RAB26 kodira istoimeni protein uključen u proces razvoja egzokrinog sustava i regulaciju egzocitoze.

NDUFB10 i COQ10A geni uključeni su u procese staničnog disanja, te se ekspresira unutar mišićja i 55 ostalih tkiva unutar psećeg organizma.

Još jedan važan gen koji je identificiran unutar regije pod selekcijom na 17. kromosomu jest SNTG2. Ovaj gen kodira protein distroglikan koji je dalje zadužen za stabilnost i preživljavanje stanica, mijelinizaciju perifernih živaca te djeluje kao receptor za adheziju stanica u mišićnom i drugom tkivu.

FAM98B gen zadužen je za regulaciju proliferacije stanica unutar psećeg organizma kao i za regulaciju ekspresije drugih specifičnih gena. Ekspresira se u plućima i 52 ostala tkiva.

RYR3 gen kodira proteine rijanodin receptora. Oni su zaduženi za normalno funkcioniranje kanala kalcijevih iona kroz tkiva psećeg organizma, kao i za vezanje istoga na receptore.

PDPK1 gen unutar psećeg organizma neposredno sudjeluje u nekoliko funkcija. Najvažnija uloga proteina kojeg kodira gen je djelovanje u obliku epidermalnog signalnog receptora faktora rasta, sudjelovanje u migraciji stanica i prijenosu kalcijevih iona kroz mišićje kao i poticanje rasta beta stanica gušterače.

Unutar ove skupine gena pripadaju i geni: ZNF236, ZNF516, GCNT4, ANKRD31, IGSF9B, BICDLA2, CLDN6, PAQR4, SRRM2, TEDC2, NPW, TBL3, IPO9, LAD1, RIN3, LGMN, GOLGA, ITPK1, CS, PAN2, GLS2, DTX3, NAB2, STAC3, DS9, CDK4, AVIL, JMJD4, FOXO6, PDIA4, TMEM18, DDC, RUVBL1, MCM2, DNAJB8, RFXAP, SPRED1 i EMC7 čija funkcija unutar konteksta ovog istraživanja nije značajna.

## **6.2.2. Geni povezani s reprodukcijom**

SPATA48 i SPATA19, geni identificirani unutar regija pod selekcijom 18. i 5. kromosoma, povezani su s procesom spermatogeneze kod muških jedinki pasa. *Spermatogenesis associated 19* (SPATA19) važan je gen povezan s reprodukcijom (Wang i sur. 2012.). Sami rasplod, ali i selekcija temelje se na mogućnosti začeća te se jedinke koje pokazuju abnormalnosti u svojstvima plodnosti isključuju iz daljnjeg uzgoja.

Još jedan od važnijih gena vezanih za svojstva plodnosti je ANKS4B gen identificiran unutar regije pod selekcijom na 6. kromosomu. Ovaj gen kodira takozvani

*Ankyrin repeat and sterile alpha motif domain containing 4B* protein. Ovaj protein ima središnju ulogu tijekom spermatogeneze potiskivanjem transponiranih elemenata i sprječavanjem njihove mobilizacije, što je bitno za integritet zametne linije te također može djelovati posredujući u interakcijama proteina i proteina tijekom sazrijevanja zametnih stanica (The UniProt Consortium).

Za razliku od ANKS4B gena, ZP2 gen je važan za svojstva plodnosti ženki. Ovaj gen kodira *Zona pellucida sperm-binding* protein koji je sastavni dio *zone pelucide*, izvanstaničnog matriksa koji okružuje oocite te posreduje u vezivanju spermija za jajnu stanicu.

CATSPERB gen kodira „*Cation channel sperm associated auxiliary subunit beta*“ proteine. Gen se ekspresira u testisima i druga dva tkiva mužjaka. Sami proteini sudjeluju u pokretanju spermija i osiguravaju normalan motilitet spermija.

ZPBP gen nema poznati protein koji kodira, ali vjeruje se da gen sudjeluje u kataboličkim procesima stanica psećeg organizma kao i u spajanju spermija na *zonu pellucidu*.

SCMH1 gen zajedno s proteinom kojeg kodira (*Scm polycomb group protein homolog 1*) sudjeluje u procesu spermatogeneze kod muških jedinki i kreiranju specifičnih vanjskih oznaka svake jedinke.

ABHD4 gen utječe na plodnost. Protein koji je rezultat ekspresije ovog gena sudjeluje u kapacitaciji spermija. Kapacitacija spermija opisuje niz biokemijskih i biofizičkih promjena koje zajedno čine stanicu spermija sposobnom za oplodnju jajne stanice (Houston i sur. 2018.).

PGR gen jedan je od važnijih gena u spolnom sustavu ženki. Zadužen je za normalno funkcioniranje progesteronskih receptora. Učinke progesterona posreduju dva proteina (receptora progesterona), nazvana A i B, koje kodira jedan gen i djeluju kao faktor transkripcije aktiviran *ligand-om* za regulaciju ekspresije reproduktivnih ciljnih gena (Conneely i sur. 2000.).

PKMYT1 gen čiji protein kinaza sudjeluje u procesu mejoze, igra bitnu ulogu u reprodukciji pasa.

SYCE3 uključen je također u procese spermatogeneze kao i u mejotičke procese, kao i NCAPH2 gen koji potiče vezanje kromatina, kondenzaciju kromosoma u procesu mejoze i separaciju sestrinskih kromatida.

HSD17B6, kao i RDH16 uključeni su u funkcioniranje spolnog sustava unutar kojega reguliraju aktivnost testosterona, estradiola i androsterona.

Kada je u pitanju reprodukcija i razvoj mladog organizma psa, jedan od važnijih gena koji je ujedno i uspješno identificiran unutar ovog istraživanja jest CCNF gen. Ovaj gen kodira protein ciklin koji sudjeluje u mejotičkim procesima, potiče ulazak stanica u procese mitoze, sudjeluje u razvoju placente kod skotnih ženki i dr.

### **6.2.3. Geni povezani s razvojem osjetila**

Geni COR6C40, COR6C30, COR2L17, COR2AV2, COR2T24, COR2T15, COR2T15, COR2T13, COR2T18, OR2M9, OR2T2, ORO4C09 te COR2G5 koji su

identificirani unutar regija pod selekcijom autosomalnih kromosoma pripadaju u takozvane olafaktorne gene, te su uključeni u percepciju mirisa. Mirisni receptori stupaju u interakciju s mirisnim molekulama u nosu kako bi pokrenuli neuronski odgovor koji pokreće percepciju mirisa (Gene Ontology Consortium: going forward 2015.). Broj i vrstu mirisnih receptora prisutnih kod pojedinog psa diktiraju pasmina, genetika i trening (Jenkins i sur. (preme Gerritsen i sur. 2015.)). Za psa koji radi, aktivno njuškanje tijekom pretraživanja "nos dolje, rep gore" pozicije i učinkovita lokalizacija izvora mirisa ključni su za obavljanje zadanog posla (Jenkins i sur. 2018.). Za pastirske pasmine pasa, normalno funkcioniranje olafaktornih gena i mirisnih receptora jedan je od važnijih čimbenika u uspješnom čuvanju stada ili imovine od predatora i drugih ljudi.

LRIG3 gen zadužen je za morfogenezu kostiju uha, prilikom čega direktno utječe na kvalitetu percepcije zvuka.

MIP gen je najvjerojatnije odgovoran za regulaciju osmolarnosti leće oka. Potpomaže interakcijama između homotetramera iz susjednih membrana koje mogu stabilizirati stanične spojeve u jezgri očne leće.

DRAM2, gen identificiran u prvoj regiji pod selekcijom na 6. kromosomu, zadužen je za regulaciju autofagoze te za vizualnu percepciju koja je pastirskim pasminama vrlo bitna.

MYO1A kodira protein miozin, koji unutar organizma potiče vezanje aktinskih niti, potiče motoričku aktivnost te je zaslužan za percepciju zvuka kod pasa.

#### **6.2.4. Geni povezani s anatomskim i histološkim razvojem organizma i njegovim normalnim funkcioniranjem**

Gen TRIPP11 zadužen je za kodiranje proteina *Thyroid hormone receptor interactor 11* koji sudjeluje u diferencijaciji hondrocita uključenoj u morfogenezi i endohondralnoj osifikaciji. Kako su pastirske pasmine uglavnom pasmine veće tjelesne mase i robusnije građe, česti su zdravstveni problemi vezani za kosti i pokretljivost. Istraživanje različite ekspresije gena uključenih u formiranje endohondralne kosti kod pasa malih i velikih pasmina mogla bi pružiti novi uvid u razvoj ploča rasta i identificirati novi načini regeneracije kosti i hrskavice (Teunissen i sur. 20017.).

PRR16 gen zadužen je za kodiranje *Proline rich 16* proteina koji sudjeluje u kontroli veličina stanica u tkivu psa.

SMAD9 je također jedan od gena koji su važni za normalni razvoj organizma. Ovaj gen sudjeluje u morfogenezi anatomskih struktura te mijenjanju puteva beta receptora faktora rasta.

Stanični odgovor na vitamin K, adhezija stanica kao i razvoj tkiva su procesi koje jednim dijelom kontrolira ekspresija POSTN gena.

FREM2 gen uključen je u procese stanične adhezije, komunikacije stanica, morfogenezu prstiju kod embrija, razvoj oka, bubrega, jetre i unutarnjeg uha kao i u proces morfogeneze epitela.

Gen COBL u psećem organizmu uključen je u normalni razvoj jetre i probavnog trakta.

FMN1 gen, identificiran u drugoj regiji pod selekcijom na 30. kromosomu, kodira protein formin. Formin igra ulogu u morfogenezi prednjih i stražnjih ekstremiteta psa, morfogenezi skeleta, ekspresiji gena kao i u vezanju mikrotubula.

HIVEP3 gen kodira takozvane *HIVEP zinc finger 3* zadužene za diferencijaciju mišićnih stanica.

MYT1L je također jedan od identificiranih gena na 17. kromosomu, kodira mijelinske transkriptorne faktore koji su uključeni u razvoj neurona i njihovo normalno funkcioniranje.

EDN2 gen je koji kodira endoteli-2 protein. Protein svoju funkciju ispoljava unutar tkiva glatkog mišićja psa, gdje potiče kontrakciju krvnih žila i kontrakciju samog mišićja. Također je jedan od važnijih proteina u kontroli krvnog tlaka kod pasa.

MMP25 gen zadužen za razvoj skeleta. Protein koji kodira ima vezna mjesta za ione kalcija i magnezija te sudjeluje u kolagenskim kataboličkim procesima kao i u organizaciji ekstracelularnog matriksa.

Uz sve do sada nabrojane gene koji imaju funkciju u razvoju i funkcioniranju mišićja, još jedan od takvih gena je i TNNI1 gen, koji kodira troponin I1 protein. Protein igra važnu ulogu u kontrakciji mišića, razvoju skeletnog mišićja kao i srčanog mišića.

STAT6 gen povezujemo sa proliferacijom stanica mliječne žlijezde i morfogenezom iste te s aktiviranjem termoregulacije za vrijeme hladnijih temperatura zraka.

AGAP2 sudjeluje u razvoju alveola unutar mliječne žlijezde.

PLXNB2 gen uključen u anatomske i histološke razvoj organizma, djeluje na razvoj mozga, proliferaciju neuroblasta, transport i oblik stanica i migraciju neurona.

TUBGCP6 gen zadužen je za formiranje kralježnice.

RHBDF1 je također jedan od gena uključenih u rast i diferencijaciju stanica, funkcioniranje faktora rasta i kataboličke procese.

SOX8 gen ima velike varijacije funkcija u psećem organizmu, a neke od najvažnijih u pogledu razvoja organizma su: razvoj masnog tkiva i diferencijacija masnih stanica, sazrijevanje stanica, intrauterini razvoj embrija, diferencijacija osteoblasta, regulacija ekspresije gena, regulacija koncentracije pojedinih hormona, diferencijacija stanica skeletnog mišićja, spermatogeneza i morfogeneza mokraćnog sustava.

NPRL3 gen zajedno sa podjedinicom *Gator1* koju kodira, sudjeluje u morfogenezi aorte i srčanog mišićja, razvoju usne šupljine i regulaciji TOR signala.

Protein iz porodice TBC1, kojega kodira TBC1D24 gen sudjeluje u normalnom razvoju živčanog sustava psa, odnosno potiče rast dendrita i aksona, njihovu obnovu te sudjeluje u degradaciji neurona prilikom oksidativnog stresa.

Gen NTN3 kodira protein netrin koji sudjeluje u morfogenezi animalnih organa i tkiva, razvoju dendrita i motoričkim procesima.

KREMEN2 kodira istoimeni protein koji potiče razvoj ekstremiteta i druge stanične funkcije.

Policistin, protein koji je uključen u razvoj bubrega, kod psa je kodiran od strane PKD1 gena na 6. kromosomu.

RPL3L, MRPL28 i MRPS34 geni zajedno sa proteinima koje kodiraju uključeni su u građu ribosoma.

R3HDM2 je još jedan od gena uključenih u kontrakciju skeletnog mišićja na način da kontrolira prolaz kalcijevih iona unutar i van stanica.

MLST8 gen kodira podjedinicu beta G proteina koja sudjeluje u takozvanom TOR signalnom putu. Ciljna molekula rapamicina (mTOR) unutar organizma sisavac je Ser-Thr kinaza koja funkcionira kao središnji element u signalnom putu uključenom u kontrolu rasta i proliferacije stanica. Aktivnost mTOR-a ne kontroliraju samo aminokiseline, već i hormoni i faktori rasta koji aktiviraju protein kinazu (Harris i sur. 2003.). Uz aktivnost protein kinaze veže se i WDR24 gen.

### 6.2.5. Ostali geni

OPCML gen kodira takozvane *Opioid binding* proteine unutar genoma psa. Razna istraživanja provedena su na ovim proteinima, no najvažnije proveli su Keates i sur. 1999. godine. Naime unutar istraživanja promatrala se mogućnost korištenja spomenutih proteina u liječenju artritisa kod pasa. Ideja je bila da se primjenom malih doza opioidnih sredstava može smanjiti osjećaj boli i nelagode kod jedinki koje pate od artritisa. Takvom primjenom opoida, došlo je do značajnog povećanja (100 puta) koncentracije imunoloških stanica u sinovijalnoj tekućini koje je bilo popraćeno 50-strukim povećanjem gustoće specifičnih [(3) H] -morfin vezujućih proteina u zglobnom i periartikularnom tkivu (Keates i sur. 1999.) Uloga ovih proteina, točnije OPCML gena u kontroli i liječenju artritisa može uvelike pomoći pastirskim i drugim radnim skupinama pasa.

PANX2 gen koji kodira protein paneksin, protein uključen u odgovor na ishemiju. Potpuna globalna ishemija miokarda ili nulti protok koronarnih arterija u pasa rezultira nizom dobro definiranih promjena koje počinju kada se metabolizam unutar miokarda pretvori iz aerobnog u anaerobni. Ti se procesi nastavljaju sve dok se ne javi nekroza miokarda (Jennings i sur. 1988.).

VWA3A jedan je od gena identificiranih unutar regije pod selekcijom na 6. kromosomu. Kodira takozvani *von Willebrand factor A*. Von Willebrandov faktor veliki je multimerni glikoprotein koji se sintetizira u različitim stanicama i ima različite uloge u hemostazi. Mutacije unutar VWA3A gena uzrokuju pojavu *Von Willebrand-ove* bolesti kod pasa. Von Willebrandova bolest opisana je u preko 50 pasmina pasa. Osobina je najraširenija kod dobermanskog pinča, pembroke velškog korgija, erdel terijera, škotskog terijera i šetlandskog ovčara, ali su teško pogođene jedinke ili obitelji identificirane u mnogim čistokrvnim i križanim pasminama (Brooks i sur. 1992.).

DNAH5, jedini gen identificiran unutar regije pod selekcijom 34. kromosoma, kodira proteine koji tvore takozvani *Dynein axonemal* lanac. Proteini ovog gena imaju raznolike funkcije unutar psećeg organizma. Neke od njih su: determinacije simetrije trupa, uspostava lokalizacije pojedinih staničnih elemenata, pokretljivost spermatozoida, razvoj srca i pokretljivost mikrotubula.

GREM1 gen koji kodira gremlin 1 protein, igra ulogu u 35 različitih funkcija. Neke od najvažnijih za normalno funkcioniranje psećeg organizma su: signalizacija stanica, morfogeneza stanica, pozitivna kontrola angiogeneze, pozitivna kontrola proliferacije stanica, pozitivna regulacija aktivnosti telomera, pozitivna regulacija diferencijacije i funkcije srčanih stanica i dr. U ovim funkcijama sudjeluje i RB1 gen.

TRIM58 gen je unutar psećeg organizma direktno povezan s imunološkim sustavom i ekspresijom drugih gena. Ponajviše se ekspresira u slezeni. Uz njega, ulogu u imunološkom odgovoru ima i TRIM13 gen.

TRPA1, jedan od identificiranih gena na 20. kromosomu izuzetno je zanimljiv kada su u pitanju pastirske pasmine pasa. Naime pasmine kao takve priviknute su na rad i život na otvorenim prostorima uz razne vremenske uvjete. Ovaj gen, sudjeluje u staničnom odgovoru na hladnoću kao i u termocepciji, odnosno percepciji temperature. Uz te funkcije, također je zadužen za stanični odgovor na bol i aktivaciju kanala kalcijevih iona u mišićju.

CLDN6 kodira protein koji nazivamo Klaudin. Ovaj protein zadužen je za strukturalnu molekularnu aktivnost, aktivnost receptora virusnih čestica u organizmu psa kao i za adheziju stanica.

IL23A jedan je od gena uključenih u razvoj i funkcioniranje imunološkog sustava psa. Sudjeluje u proliferaciji T-limfocita, imunom odgovoru na infekciju gram negativnim bakterijama, upalnom odgovoru, imunom odgovoru na infekciju virusnim česticama, nekrozi tumornih stanica, stvaranju antitijela i dr. S obzirom da pastirski psi većinu vremena provode na otvorenom gdje su izloženi raznim mikroorganizmima iz okoliša, moguće je objasniti identifikaciju ovog gena u regijama pod selekcijom, kao posljedicu prilagodbe jedinki na okolišne uvjete.

INHBC gen kodira C podjedinicu beta inhibina dok INHBE gen kodira E podjedinicu beta inhibina. Aktivini i inhibini su među 33 člana obitelji TGF- $\beta$  i prvi su put opisani kao regulatori lučenja folikularno stimulirajućeg hormona (FSH) i eritropoeze. Međutim, aktivini i inhibini od tada su uključeni u različite biološke procese, u rasponu od ranih faza embrionalnog razvoja, do visoko specijaliziranih funkcija u terminalno diferenciranim stanicama i tkivima (Namwanje i sur. 2016.)

Još jedan od gena s veliki brojem funkcija jest GLI1, koji unutar psećeg organizma sudjeluje u spermatogenezi, razvoju pluća, proliferaciji stanica srčanog mišića, regeneraciji jetre, odgovoru na ranjavanje, diferencijaciji osteoblasta, DNA replikaciji i dr.

DDIT3 gen i proteini koje kodira sudjeluju u specifikaciji unutarnjih i vanjskih osi, maturaciji krvnih žila, staničnom odgovoru na gladovanje, percepciji zvuka i dr.



## 7. Zaključci

1. Provedbom analiza *Integrated haplotype score-a* (iHS) na genomskim podacima 123 jedinke pastirskih pasmina pasa koji su se sastojali od 107 403 SNP-a, uspješno je identificirano 6 kromosomskih regija pod selekcijom s pripadajućih 60 gena.
2. Provedbom analiza *Extreme runs of homozygosity* (eROHi) vrijednosti na istim genomskim podacima, identificirano je 23 kromosomskih regija pod selekcijom s pripadajuća 402 gena.
3. Izvršena je i funkcionalna analiza gena unutar pojedinog signala, prilikom čega su uočeni geni važni za ekspresiju raznih morfoloških, ali i bihevioralnih svojstava kod promatranih pastirskih jedinki.
4. Ova saznanja mogu nam poslužiti u razumijevanju prilagodbe pastirskih pasmina na okolišne čimbenike i na radnu namjenu koja im je dodijeljena te pružiti detaljan uvid u provedbu prirodne i umjetne selekcije na specifična svojstva poput izražene muskulature, izraženog osjeta mirisa i vida, kao i na određene biološke i reproduktivne karakteristike.

## 8. Popis literature

1. Barth B. (2017). How to Choose a Livestock Guard Dog <https://modernfarmer.com/2017/09/choose-livestock-guard-dog/>-pristup (02.08.2021.)
2. Bluemel, J., Korte, S., Schenck, E., & Weinbauer, G. (Eds.). (2015). The nonhuman primate in nonclinical drug development and safety assessment. Academic Press.
3. Bluemel, J., Korte, S., Schenck, E., & Weinbauer, G. (Eds.). (2015). The nonhuman primate in nonclinical drug development and safety assessment. Academic Press.  
PMCID: PMC5873274  
PMID: 28681971
4. Booker, T.R., Jackson, B.C. & Keightley, P.D. Detecting positive selection in the genome. BMC Biol 15, 98 (2017). <https://doi.org/10.1186/s12915-017-0434-y> - pristup (07.06.2021.)
5. Bosse M., Megens HJ., Madsen O., Paudel Y., Frantz LAF., Schook LB., et al. Regions of homozygosity in the porcine genome: Consequence of demography and the recombination landscape. PLoS Genet. 2012;8:e1003100.
6. Brooks, M., Dodds, W. J., & Raymond, S. L. (1992). Epidemiologic features of von Willebrand's disease in Doberman pinschers, Scottish terriers, and Shetland sheepdogs: 260 cases (1984-1988). Journal of the American Veterinary Medical Association, 200(8), 1123–1127.
7. Boyko AR., Quignon P., Li L., Schoenebeck JJ., Degenhardt JD., Lohmueller KE., et al. A simple genetic architecture underlies morphological variation in dogs. PLoS Biol. 2010;8:e1000451.
8. Charest, A., Lane, K., McMahon, K., Park, J., Preisinger, E., Conroy, H., & Housman, D. (2003). Fusion of FIG to the receptor tyrosine kinase ROS in a glioblastoma with an interstitial del(6)(q21q21). Genes, chromosomes & cancer, 37(1), 58–71. <https://doi.org/10.1002/gcc.10207> -pristup (29.08.2021.)
9. Conneely, O. M., & Lydon, J. P. (2000). Progesterone receptors in reproduction: functional impact of the A and B isoforms. Steroids, 65(10-11), 571-577.
10. Curik, I., Ferenčaković, M., Sölkner, J. (2014). Inbreeding and runs of homozygosity: a possible solution to an old problem. Livest. Sci. 166, 26–34. doi: 10.1016/j.livsci.2014.05.034

11. Darwin, C. *The Variation of Animals and Plants under Domestication* (1868.)
12. Delaneau, O., Marchini, J., & Zagury, J. F. (2011). A linear complexity phasing method for thousands of genomes. *Nature methods*, 9(2), 179–181. <https://doi.org/10.1038/nmeth.1785> -pristup (15.08.2021.)
13. Eileen K. Jenkins, Mallory T. DeChant and Erin B. Perry  
When the Nose Doesn't Know: Canine Olfactory Function Associated With Health, Management, and Potential Links to Microbiota (2018) <https://www.frontiersin.org/article/10.3389/fvets.2018.00056>-pristup (01.09.2021.)
14. Falconer, D. S., Mackay, T. F. C. (1996). *Introduction to quantitative genetics* (Harlow: Addison Wesley Longman).
15. FCI (2017.) Federation cynologique internationale. <http://www.fci.be/en> - pristup 22.06.2021.
16. Ferenčaković et al. *Genetics Selection Evolution* 2013, 45:42. <http://www.gsejournal.org/content/45/1/42> -pristup (22.7.2021.)
17. Gabr, M. M., Zakaria, M. M., Refaie, A. F., Ismail, A. M., Khater, S. M., Ashamallah, S. A., Azzam, M. M., & Ghoneim, M. A. (2018). Insulin-producing Cells from Adult Human Bone Marrow Mesenchymal Stromal Cells Could Control Chemically Induced Diabetes in Dogs: A Preliminary Study. *Cell transplantation*, 27(6), 937–947. <https://doi.org/10.1177/0963689718759913> - pristup (30.08.2021.)
18. Gautier, Mathieu & Vitalis, Renaud. (2012). Rehh An R package to detect footprints of selection in genome-wide SNP data from haplotype structure. *Bioinformatics* (Oxford, England). 28. 1176-7. 10.1093/bioinformatics/bts115.
19. Gerritsen R, Hank R. *K9 Scent Training: A Manual for Training Your Identification, Tracking, and Detection Dog*. Vancouver, BC: Brush Education, Inc (2015). URL: <https://vetbooks.ir/k9-scent-training-a-manual-for-training-your-identification-tracking-and-detection-dog/> -pristup (17.08.2021.)
20. Grzmil P., Konietzko J., Boehm D., Hoelter S., M, Aguilar A., Javaheri A., Kalaydjiev S., Adler T., Bolle I., Adham I., Dixkens C., Wolf S., Fuchs H., Gailus-Durne V., Wurst W., Ollert M., Busch D., Schulz H., Hrabe de Angelis M., Burfeind P.: Targeted Disruption of the Mouse *Npal3* Gene Leads to Deficits in Behavior, Increased IgE Levels, and Impaired Lung Function. *Cytogenet Genome Res* 2009;125:186-200. doi: 10.1159/000230003

21. Haplotter Guide, <http://haplotter.uchicago.edu/instruction.html>
22. Harris, T. E., & Lawrence, J. C. (2003). TOR signaling. *Science's STKE*, 2003(212), re15-re15.
23. HKS (2017.) Hrvatski kinološki savez. <https://web.hks.hr/> - pristup 22.06.2021.
24. Houston, B. J., Nixon, B., Martin, J. H., De Iuliis, G. N., Trigg, N. A., Bromfield, E. G., ... & Aitken, R. J. (2018). Heat exposure induces oxidative stress and DNA damage in the male germ line. *Biology of reproduction*, 98(4), 593-606.
25. Howrigan, D.P., Simonson, M.A. & Keller, M.C. Detecting autozygosity through runs of homozygosity: A comparison of three autozygosity detection algorithms. *BMC Genomics* 12, 460 (2011). <https://doi.org/10.1186/1471-2164-12-460> - pristup (22.08.2021.)
26. Janeš, M.; Zorc, M.; Ferenčaković M.; Curik, I.; Dovč, P.; Cubrik-Curik, V. Genomic Characterization of the Three Balkan Livestock Guardian Dogs. *Sustainability* **2021**, *13*, 2289. <https://doi.org/10.3390/su13042289> -pristup (02.06.2021.)
27. Jennings RB., Reimer KA., Steenbergen C. Complete global myocardial ischemia in dogs. *Crit Care Med.* 1988 Oct;16(10):988-96. doi: 10.1097/00003246-198810000-00010. PMID: 3048898.
28. Joshua M. Akey, Alison L. Ruhe, Dayna T. Akey, Aaron K. Wong, Caitlin F. Connelly, Jennifer Madeoy, Thomas J. Nicholas, Mark W. Neff: Tracking footprints of artificial selection in the dog genome *Proceedings of the National Academy of Sciences* Jan 2010, 200909918; DOI: 10.1073/pnas.0909918107
29. Jovanovac S. (1997.) Opće stočarstvo. Poljoprivredni fakultet, Osijek
30. Keates, H. L., Cramond, T., & Smith, M. T. (1999). Intraarticular and periarticular opioid binding in inflamed tissue in experimental canine arthritis. *Anesthesia and analgesia*, 89(2), 409–415. <https://doi.org/10.1097/00000539-199908000-00030> -pristup (02.09.2021.)
31. Kimura M. The neutral theory of molecular evolution. Cambridge University Press; 1983.

32. MacLennan, D., Kranias, E. Phospholamban: a crucial regulator of cardiac contractility. *Nat Rev Mol Cell Biol* 4, 566–577 (2003). <https://doi.org/10.1038/nrm1151> -pristup (29.08.2021.)
33. McQuillan, R., Leutenegger, A. L., Abdel-Rahman, R., Franklin, C. S., Pericic, M., Barac-Lauc, L., Smolej-Narancic, N., Janicijevic, B., Polasek, O., Tenesa, A., Macleod, A. K., Farrington, S. M., Rudan, P., Hayward, C., Vitart, V., Rudan, I., Wild, S. H., Dunlop, M. G., Wright, A. F., Campbell, H., ... Wilson, J. F. (2008). Runs of homozygosity in European populations. *American journal of human genetics*, 83(3), 359–372. <https://doi.org/10.1016/j.ajhg.2008.08.007> -pristup (20.06.2021.)
34. Masahiko Itoh, Hiroyuki Sasaki, Mikio Furuse, Harunobu Ozaki, Toru Kita, Shoichiro Tsukita; Junctional adhesion molecule (JAM) binds to PAR-3 : a possible mechanism for the recruitment of PAR-3 to tight junctions . *J Cell Biol* 6 August 2001; 154 (3): 491–498. doi: <https://doi.org/10.1083/jcb.200103047> -pristup (02.09.2021.)
35. Namwanje, M., & Brown, C. W. (2016). Activins and inhibins: roles in development, physiology, and disease. *Cold Spring Harbor perspectives in biology*, 8(7), a021881.
36. Nowakowska M., Gualtieri F., von Rüden E., Hansmann F., Baumgärtner W., Tipold A., Potschka H., Profiling the Expression of Endoplasmic Reticulum Stress Associated Heat Shock Proteins in Animal Epilepsy Models, *Neuroscience*, Volume 429 (2020), Pages 156-172, <https://doi.org/10.1016/j.neuroscience.2019.12.015>. –pristup (01.09.2021.)
37. Oji, T., Kamishina, H., Cheeseman, J. A., & Clemmons, R. M. (2007). Measurement of myelin basic protein in the cerebrospinal fluid of dogs with degenerative myelopathy. *Veterinary clinical pathology*, 36(3), 281–284. <https://doi.org/10.1111/j.1939-165x.2007.tb00225.x> -pristup (20.08.2021.)
38. Oksenberg, N., & Ahituv, N. (2013). The role of AUTS2 in neurodevelopment and human evolution. *Trends in genetics : TIG*, 29(10), 600–608. <https://doi.org/10.1016/j.tig.2013.08.001> -pristup (02.09.2021.)
39. Package : PLINK [version]  
 Authors : Shaun Purcell, Christopher Chang  
 URL: [www.cog-genomics.org/plink/1.9/](http://www.cog-genomics.org/plink/1.9/) -pristup (13.07.2021.)
40. Pasji život. Pasmine po namjeni - pastirski psi: čuvari stada, ovčari, govedari.  
 Autor: Rujana Jeger. <http://rujanajeger.com/index.php/pasmine-pasa/item/579->

[pasmine-po-namjeni-pastirski-psi-cuvari-stada-ovcari-govedari](#) - pristup (06.06.2021.)

41. Peripolli, E., Stafuzza, N.B., Munari, D.P. et al. Assessment of runs of homozygosity islands and estimates of genomic inbreeding in Gyr (*Bos indicus*) dairy cattle. *BMC Genomics* 19, 34 (2018). <https://doi.org/10.1186/s12864-017-4365-3> -pristup (29.07.2021.)
42. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, Maller J, Sklar P, de Bakker PIW, Daly MJ & Sham PC (2007) PLINK: a toolset for whole-genome association and population-based linkage analysis. *American Journal of Human Genetics*, 81.
43. qqman: an R package for visualizing GWAS results using Q-Q and manhattan plots. Stephen D. Turner. bioRxiv 005165; doi: <https://doi.org/10.1101/005165> -pristup: (24.07.2021.)
44. R Core Team (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing
45. Riggs R. (2001). Livestock guarding dogs: their current use world wide. Department of Zoology, University of Aberdeen, Tillydrone Avenue [http://www1.nina.no/lcie\\_new/pdf/634994135320630456\\_IUCN%20CSG%20Occasional%20Papers%20Rigg%20LGDs.pdf](http://www1.nina.no/lcie_new/pdf/634994135320630456_IUCN%20CSG%20Occasional%20Papers%20Rigg%20LGDs.pdf) –pristup (02.08.2021.)
46. Sabeti P. C., Reich D. E., Higgins J. M., Levine H. Z. P., Richter D. J., others, 2002 Detecting recent positive selection in the human genome from haplotype structure. *Nature* 419: 832–837.
47. Sabeti, P., Varilly, P., Fry, B. et al. Genome-wide detection and characterization of positive selection in human populations. *Nature* 449, 913–918 (2007). <https://doi.org/10.1038/nature06250> -pristup (29.07.2021.)
48. SAS Institute Inc 2013. SAS/ACCESS® 9.4 Interface to ADABAS: Reference. Cary, NC: SAS Institute Inc.
49. Selekcija. Hrvatska enciklopedija, mrežno izdanje. Leksikografski zavod Miroslav Krleža, 2021. <http://www.enciklopedija.hr/Natuknica.aspx?ID=55278> – pristup (06.06.2021.)
50. Serpell, J. . *The Domestic Dog: Its Evolution, Behaviour, and Interactions with People*. 284, Cambridge University Press (1996) .

51. Sölkner J., Ferenčaković M., Fürst C., Čurik I. Genomic metrics of individual autozygosity, applied to a cattle population. In Proceedings of the 61st Annual Meeting of the European Association for Animal Production: 23–27 August 2010; Heraklion. 2010.
52. Sponenberg D.P. (2000). Livestock guard dogs: what is a breed, and why does it matter?"  
<http://www.mindspring.com/~skocher/Sponenberg.htm> -pristup (02.08.2021.)
53. Susan Fairley, Ernesto Lowy-Gallego, Emily Perry, Paul Flicek, The International Genome Sample Resource (IGSR) collection of open human genomic variation resources, *Nucleic Acids Research*, Volume 48, Issue D1, 08 January 2020, Pages D941–D947, <https://doi.org/10.1093/nar/gkz836> -pristup (22.08.2021.)
54. Teunissen, M., Riemers, F. M., van Leenen, D., Groot Koerkamp, M., Meij, B. P., Alblas, J., Penning, L. C., Miranda-Bedate, A., & Tryfonidou, M. A. (2018). Growth plate expression profiling: Large and small breed dogs provide new insights in endochondral bone formation. *Journal of orthopaedic research : official publication of the Orthopaedic Research Society*, 36(1), 138–148.  
<https://doi.org/10.1002/jor.23647> -pristup (20.08.2021.)
55. The Gene Ontology Consortium, Gene Ontology Consortium: going forward, *Nucleic Acids Research*, Volume 43, Issue D1, 28 January 2015, Pages D1049–D1056, <https://doi.org/10.1093/nar/gku1179> -pristup (01.09.2021.)
56. The UniProt Consortium  
UniProt: the universal protein knowledgebase in 2021  
*Nucleic Acids Res.* 49:D1 (2021)
57. The Year of the Rat: The Rat Genome Database at 20: a multi-species knowledgebase and analysis platform.  
Smith JR, Hayman GT, Wang SJ, Laulederkind SJF, Hoffman MJ, Kaldunski ML, Tutaj M, Thota J, Nalabolu HS, Ellanki SLR, Tutaj MA, De Pons JL, Kwitek AE, Dwinell MR, Shimoyama ME. *Nucleic Acids Res.* 2019 Nov 12. pii: gkz1041. doi: 10.1093/nar/gkz1041.
58. Varshney A., Chahal G., Santos L., Stolper J., C. Hallab J., T. Nim H., Nikolov M., Yip A., Ramialison M.,  
Human Cardiac Transcription Factor Networks,  
Editor(s): Olaf Wolkenhauer, *Systems Medicine*, Academic Press (2021)  
Pages 429-453,  
<https://doi.org/10.1016/B978-0-12-801238-3.11597-1>. –pristup (20.08.2021.)

59. Voight F. B., Kudravalli S. , Wen X., Pritchard K.J.  
A Map of Recent Positive Selection in the Human Genome
60. Vonholdt, B. M. et al. Genome-wide SNP and haplotype analyses reveal a rich history underlying dog domestication. *Nature* 464, 898–902 (2010) .
61. VonHoldt, B. M. et al. A genome-wide perspective on the evolutionary history of enigmatic wolf-like canids. *Genome Res.* 21, 1294–1305 (2011) .
62. Wang, J.Y., Lan, J., Zhao, J. et al. Molecular characterization, polymorphism and association of porcine SPATA19 gene. *Mol Biol Rep* 39, 9741–9746 (2012).  
<https://doi.org/10.1007/s11033-012-1839-x> -pristup (02.09.2021.)
63. WHAMM! v0.14 Whole-Genome Homozygosity Analysis and Mapping Machina 2008. URL: <http://coruscant.itmat.upenn.edu/whamm/ihs.html>
64. Williamson SH, Hubisz MJ, Clark AG, Payseur BA, Bustamante CD, et al. (2007) Localizing Recent Adaptive Evolution in the Human Genome. *PLOS Genetics* 3(6): e90. <https://doi.org/10.1371/journal.pgen.0030090> -pristup (17.06.2021.)
65. Yang, Y., Fan, X., Yan, J., Chen, M., Zhu, M., Tang, Y., Liu, S., & Tang, Z. (2021). A comprehensive epigenome atlas reveals DNA methylation regulating skeletal muscle development. *Nucleic acids research*, 49(3), 1313–1329.  
<https://doi.org/10.1093/nar/gkaa1203> -pristup (02.09.2021.)
66. Zhang Q., Guldbbrandtsen B., Bosse M., Lund MS., Sahana G. Runs of homozygosity and distribution of functional variants in the cattle genome. *BMC Genomics*. 2015;16:542.



## 9. Prilog A

#Potrebni paketi za iHS (integrated haplotype score) metode:

```
library(rehh) #Extended haplotype homozygosity testovi za otkrivanje selekcije
```

```
library(rehh.data)
```

```
library(vcfR) #Unos i korištenje .vcf file-ova (Variant call format)
```

```
library(data.table)
```

#Odabir radnog direktorija:

```
setwd("C:/Users/Vjerman/Desktop/sve za diplomski/vcfzar")
```

#Unos podataka (.vcf file-ova ta svaki od 38 autosomalnih kromosoma)

```
LGD_pasmine_CHR1 <- data2haplohh (hap_file = "LGD-pasmine_chr1.vcf",  
polarize_vcf = FALSE)
```

```
LGD_pasmine_CHR2 <- data2haplohh (hap_file = "LGD-pasmine_chr2.vcf",  
polarize_vcf = FALSE)
```

```
LGD_pasmine_CHR3 <- data2haplohh (hap_file = "LGD-pasmine_chr3.vcf",  
polarize_vcf = FALSE)
```

```
LGD_pasmine_CHR4 <- data2haplohh (hap_file = "LGD-pasmine_chr4.vcf",  
polarize_vcf = FALSE)
```

```
LGD_pasmine_CHR5 <- data2haplohh (hap_file = "LGD-pasmine_chr5.vcf",  
polarize_vcf = FALSE)
```

```
LGD_pasmine_CHR6 <- data2haplohh (hap_file = "LGD-pasmine_chr6.vcf",  
polarize_vcf = FALSE)
```

```
LGD_pasmine_CHR7 <- data2haplohh (hap_file = "LGD-pasmine_chr7.vcf",  
polarize_vcf = FALSE)
```

```
LGD_pasmine_CHR8 <- data2haplohh (hap_file = "LGD-pasmine_chr8.vcf",  
polarize_vcf = FALSE)
```

```
LGD_pasmine_CHR9 <- data2haplohh (hap_file = "LGD-pasmine_chr9.vcf",  
polarize_vcf = FALSE)
```

```
LGD_pasmine_CHR10 <- data2haplohh (hap_file = "LGD-pasmine_chr10.vcf",  
polarize_vcf = FALSE)
```

```
LGD_pasmine_CHR11 <- data2haplohh (hap_file = "LGD-pasmine_chr11.vcf",  
polarize_vcf = FALSE)
```

```
LGD_pasmine_CHR12 <- data2haplohh (hap_file = "LGD-pasmine_chr12.vcf",  
polarize_vcf = FALSE)
```

```
LGD_pasmine_CHR13 <- data2haplohh (hap_file = "LGD-pasmine_chr13.vcf",  
polarize_vcf = FALSE)
```

```
LGD_pasmine_CHR14 <- data2haplohh (hap_file = "LGD-pasmine_chr14.vcf",  
polarize_vcf = FALSE)
```

```
LGD_pasmine_CHR15 <- data2haplohh (hap_file = "LGD-pasmine_chr15.vcf",  
polarize_vcf = FALSE)
```

```
LGD_pasmine_CHR16 <- data2haplohh (hap_file = "LGD-pasmine_chr16.vcf",  
polarize_vcf = FALSE)
```

```

LGD_pasmine_CHR17 <- data2haplohh (hap_file = "LGD-pasmine_chr17.vcf",
polarize_vcf = FALSE)
LGD_pasmine_CHR18 <- data2haplohh (hap_file = "LGD-pasmine_chr18.vcf",
polarize_vcf = FALSE)
LGD_pasmine_CHR19 <- data2haplohh (hap_file = "LGD-pasmine_chr19.vcf",
polarize_vcf = FALSE)
LGD_pasmine_CHR20 <- data2haplohh (hap_file = "LGD-pasmine_chr20.vcf",
polarize_vcf = FALSE)
LGD_pasmine_CHR21 <- data2haplohh (hap_file = "LGD-pasmine_chr21.vcf",
polarize_vcf = FALSE)
LGD_pasmine_CHR22 <- data2haplohh (hap_file = "LGD-pasmine_chr22.vcf",
polarize_vcf = FALSE)
LGD_pasmine_CHR23 <- data2haplohh (hap_file = "LGD-pasmine_chr23.vcf",
polarize_vcf = FALSE)
LGD_pasmine_CHR24 <- data2haplohh (hap_file = "LGD-pasmine_chr24.vcf",
polarize_vcf = FALSE)
LGD_pasmine_CHR25 <- data2haplohh (hap_file = "LGD-pasmine_chr25.vcf",
polarize_vcf = FALSE)
LGD_pasmine_CHR26 <- data2haplohh (hap_file = "LGD-pasmine_chr26.vcf",
polarize_vcf = FALSE)
LGD_pasmine_CHR27 <- data2haplohh (hap_file = "LGD-pasmine_chr27.vcf",
polarize_vcf = FALSE)
LGD_pasmine_CHR28 <- data2haplohh (hap_file = "LGD-pasmine_chr28.vcf",
polarize_vcf = FALSE)
LGD_pasmine_CHR29 <- data2haplohh (hap_file = "LGD-pasmine_chr29.vcf",
polarize_vcf = FALSE)
LGD_pasmine_CHR30 <- data2haplohh (hap_file = "LGD-pasmine_chr30.vcf",
polarize_vcf = FALSE)
LGD_pasmine_CHR31 <- data2haplohh (hap_file = "LGD-pasmine_chr31.vcf",
polarize_vcf = FALSE)
LGD_pasmine_CHR32 <- data2haplohh (hap_file = "LGD-pasmine_chr32.vcf",
polarize_vcf = FALSE)
LGD_pasmine_CHR33 <- data2haplohh (hap_file = "LGD-pasmine_chr33.vcf",
polarize_vcf = FALSE)
LGD_pasmine_CHR34 <- data2haplohh (hap_file = "LGD-pasmine_chr34.vcf",
polarize_vcf = FALSE)
LGD_pasmine_CHR35 <- data2haplohh (hap_file = "LGD-pasmine_chr35.vcf",
polarize_vcf = FALSE)
LGD_pasmine_CHR36 <- data2haplohh (hap_file = "LGD-pasmine_chr36.vcf",
polarize_vcf = FALSE)
LGD_pasmine_CHR37 <- data2haplohh (hap_file = "LGD-pasmine_chr37.vcf",
polarize_vcf = FALSE)
LGD_pasmine_CHR38 <- data2haplohh (hap_file = "LGD-pasmine_chr38.vcf",
polarize_vcf = FALSE)

```

```
#Skeniranje genomskih podataka uz pomocu scan_hh funkcije:
LGD_pasmine_CHR1.scan<-scan_hh(LGD_pasmine_CHR1,
discard_integration_at_border = FALSE)
LGD_pasmine_CHR2.scan<-scan_hh(LGD_pasmine_CHR2,
discard_integration_at_border = FALSE)
LGD_pasmine_CHR3.scan<-scan_hh(LGD_pasmine_CHR3,
discard_integration_at_border = FALSE)
LGD_pasmine_CHR4.scan<-scan_hh(LGD_pasmine_CHR4,
discard_integration_at_border = FALSE)
LGD_pasmine_CHR5.scan<-scan_hh(LGD_pasmine_CHR5,
discard_integration_at_border = FALSE)
LGD_pasmine_CHR6.scan<-scan_hh(LGD_pasmine_CHR6,
discard_integration_at_border = FALSE)
LGD_pasmine_CHR7.scan<-scan_hh(LGD_pasmine_CHR7,
discard_integration_at_border = FALSE)
LGD_pasmine_CHR8.scan<-scan_hh(LGD_pasmine_CHR8,
discard_integration_at_border = FALSE)
LGD_pasmine_CHR9.scan<-scan_hh(LGD_pasmine_CHR9,
discard_integration_at_border = FALSE)
LGD_pasmine_CHR10.scan<-scan_hh(LGD_pasmine_CHR10,
discard_integration_at_border = FALSE)
LGD_pasmine_CHR11.scan<-scan_hh(LGD_pasmine_CHR11,
discard_integration_at_border = FALSE)
LGD_pasmine_CHR12.scan<-scan_hh(LGD_pasmine_CHR12,
discard_integration_at_border = FALSE)
LGD_pasmine_CHR13.scan<-scan_hh(LGD_pasmine_CHR13,
discard_integration_at_border = FALSE)
LGD_pasmine_CHR14.scan<-scan_hh(LGD_pasmine_CHR14,
discard_integration_at_border = FALSE)
LGD_pasmine_CHR15.scan<-scan_hh(LGD_pasmine_CHR15,
discard_integration_at_border = FALSE)
LGD_pasmine_CHR16.scan<-scan_hh(LGD_pasmine_CHR16,
discard_integration_at_border = FALSE)
LGD_pasmine_CHR17.scan<-scan_hh(LGD_pasmine_CHR17,
discard_integration_at_border = FALSE)
LGD_pasmine_CHR18.scan<-scan_hh(LGD_pasmine_CHR18,
discard_integration_at_border = FALSE)
LGD_pasmine_CHR19.scan<-scan_hh(LGD_pasmine_CHR19,
discard_integration_at_border = FALSE)
LGD_pasmine_CHR20.scan<-scan_hh(LGD_pasmine_CHR20,
discard_integration_at_border = FALSE)
LGD_pasmine_CHR21.scan<-scan_hh(LGD_pasmine_CHR21,
discard_integration_at_border = FALSE)
```

```

LGD_pasmine_CHR22.scan<-scan_hh(LGD_pasmine_CHR22,
discard_integration_at_border = FALSE)
LGD_pasmine_CHR23.scan<-scan_hh(LGD_pasmine_CHR23,
discard_integration_at_border = FALSE)
LGD_pasmine_CHR24.scan<-scan_hh(LGD_pasmine_CHR24,
discard_integration_at_border = FALSE)
LGD_pasmine_CHR25.scan<-scan_hh(LGD_pasmine_CHR25,
discard_integration_at_border = FALSE)
LGD_pasmine_CHR26.scan<-scan_hh(LGD_pasmine_CHR26,
discard_integration_at_border = FALSE)
LGD_pasmine_CHR27.scan<-scan_hh(LGD_pasmine_CHR27,
discard_integration_at_border = FALSE)
LGD_pasmine_CHR28.scan<-scan_hh(LGD_pasmine_CHR28,
discard_integration_at_border = FALSE)
LGD_pasmine_CHR29.scan<-scan_hh(LGD_pasmine_CHR29,
discard_integration_at_border = FALSE)
LGD_pasmine_CHR30.scan<-scan_hh(LGD_pasmine_CHR30,
discard_integration_at_border = FALSE)
LGD_pasmine_CHR31.scan<-scan_hh(LGD_pasmine_CHR31,
discard_integration_at_border = FALSE)
LGD_pasmine_CHR32.scan<-scan_hh(LGD_pasmine_CHR32,
discard_integration_at_border = FALSE)
LGD_pasmine_CHR33.scan<-scan_hh(LGD_pasmine_CHR33,
discard_integration_at_border = FALSE)
LGD_pasmine_CHR34.scan<-scan_hh(LGD_pasmine_CHR34,
discard_integration_at_border = FALSE)
LGD_pasmine_CHR35.scan<-scan_hh(LGD_pasmine_CHR35,
discard_integration_at_border = FALSE)
LGD_pasmine_CHR36.scan<-scan_hh(LGD_pasmine_CHR36,
discard_integration_at_border = FALSE)
LGD_pasmine_CHR37.scan<-scan_hh(LGD_pasmine_CHR37,
discard_integration_at_border = FALSE)
LGD_pasmine_CHR38.scan<-scan_hh(LGD_pasmine_CHR38,
discard_integration_at_border = FALSE)

```

#rbind funkcija povezuje sve prethodno dobivene objekte unutar jednog zasebnog objekta:

```

LGD_pasmine_scan_IHH<-
rbind(LGD_pasmine_CHR1.scan,LGD_pasmine_CHR2.scan,LGD_pasmine_CHR3.s
can,LGD_pasmine_CHR4.scan,LGD_pasmine_CHR5.scan,LGD_pasmine_CHR6.sc
an,LGD_pasmine_CHR7.scan,LGD_pasmine_CHR8.scan,LGD_pasmine_CHR9.sca
n,LGD_pasmine_CHR10.scan,LGD_pasmine_CHR11.scan,LGD_pasmine_CHR12.s
can,LGD_pasmine_CHR13.scan,LGD_pasmine_CHR14.scan,LGD_pasmine_CHR1
5.scan,LGD_pasmine_CHR16.scan,LGD_pasmine_CHR17.scan,LGD_pasmine_CH

```

```
R18.scan,LGD_pasmine_CHR19.scan,LGD_pasmine_CHR20.scan,LGD_pasmine_
CHR21.scan,LGD_pasmine_CHR22.scan,LGD_pasmine_CHR23.scan,LGD_pasmin
e_CHR24.scan,LGD_pasmine_CHR25.scan,LGD_pasmine_CHR26.scan,LGD_pas
mine_CHR27.scan,LGD_pasmine_CHR28.scan,LGD_pasmine_CHR29.scan,LGD_p
asmine_CHR30.scan,LGD_pasmine_CHR31.scan,LGD_pasmine_CHR32.scan,LGD
_pasmine_CHR33.scan,LGD_pasmine_CHR34.scan,LGD_pasmine_CHR35.scan,L
GD_pasmine_CHR36.scan,LGD_pasmine_CHR37.scan,LGD_pasmine_CHR38.sca
n)
```

```
#Izračun -log10 pihs vrijednosti iz ihh vrijednosti:
wgscansve_LGD_pasmine.ihs <- ihh2ihs(LGD_pasmine_scan_IHH)
```

```
#Vizualizacija rezultata (u boji):
```

```
freqbinplot(wgscansve_LGD_pasmine.ihs)
distribplot(wgscansve_LGD_pasmine.ihs$ihs$IHS, xlab = "iHS", main="Genom-wide
distribution \n (LGD pasmine)")
distribplot(wgscansve_LGD_pasmine.ihs$ihs$IHS, xlab = "iHS",qqplot = TRUE,
main="Normal Q-Q plot (LGD pasmine)")
```

```
#qqplot=TRUE -> pravi se qqplot umjesto distribution density krivulje
```

```
manhattanplot(wgscansve_LGD_pasmine.ihs, main = "Inverted haplotype score (LGD
pasmine)")
manhattanplot(wgscansve_LGD_pasmine.ihs,pval = TRUE,threshold=c(5),main = "P
vrijednosti SNP-ova na autosomalnim \n kromosomima LGD pasmina (iHS) ")
```

```
#Vizualizacija rezultata (crno-bijeli Manhattan plot):
```

```
palette(c("black", "gray"))
manhattanplot(wgscansve_LGD_pasmine.ihs,
              pval = TRUE,
              threshold = c(5),
              cex = 0.6,
              main = "iHS signali pozitivne selekcije na autosomima \n kod LGD pasmina")
```

```
#Kreiranje Manhattan plot-ova po zasebnim kromosomima na kojima SNP-ovi prelaze
trehsold:
```

```
#Kromosom 1 kod LGD pasmina:
```

```
manhattanplot(wgscansve_LGD_pasmine.ihs,
              pval = TRUE, cr.col = "black",
              threshold = c(5), chr.name = "1",
              main = "p vrijednost SNP-ova na Chr1")
```

```
#Kromosom 6 kod LGD pasmina:
```

```
manhattanplot(wgscansve_LGD_pasmine.ihs,
```

```
pval = TRUE, cr.col = "black",  
threshold = c(5), chr.name = "6",  
main = "p vrijednost SNP-ova na Chr 6")
```

```
#Kromosom 17 kod LGD pasmina:  
manhattanplot(wgscansve_LGD_pasmine.ihs,  
pval = TRUE, cr.col = "black",  
threshold = c(5), chr.name = "17",  
main = "p vrijednost SNP-ova na Chr 17")
```

```
#Kromosom 25 kod LGD pasmina:  
manhattanplot(wgscansve_LGD_pasmine.ihs,  
pval = TRUE, cr.col = "black",  
threshold = c(5), chr.name = "25",  
main = "P vrijednost SNP-ova na Chr 25")
```

```
#Kromosom 30 kod LGD pasmina:  
manhattanplot(wgscansve_LGD_pasmine.ihs,  
pval = TRUE, cr.col = "black",  
threshold = c(5), chr.name = "30",  
main = "P vrijednost SNP-ova na Chr 30")
```

```
#Pokretanje paketa potrebnog za izvoz podataka u .xlsx obliku:  
library(writexl)
```

```
#Izvoz podataka u obliku excel tablice:  
write_xlsx(wgscansve_LGD_pasmine.ihs,"C:\\Users\\Vjeran\\Desktop\\sveza  
diplomski\\iHS_LGD_pasmine.xlsx")
```

## 10. Prilog B

```
#Pokretanje paketa potrebnih za izračun  $-\log_{10} p_{eROHi}$  vrijednosti:
library(qqman)
library(animation)

#Unos podataka s izračunatim Z_ROH-om (iz SAS-a):

Froh1_sve<-read.table(file='FROH1_SVE_PASMINE.txt',header=TRUE)

#Izracun P-value:

Froh1_sve$pvalue_R <- pnorm(-abs(Froh1_sve$Z_ROH1_sve))

#Transformacija (-log10 skala):

Froh1_sve$minus_log10 <- -log10(Froh1_sve$pvalue_R)

#Vizualizacija:

manhattan (Froh1_sve, bp= "bp", chr= "chr", p= "pvalue_R", snp="snp_name", logp=
TRUE, ylim=c(0,20), suggestiveline =FALSE, genomewideline = 5)

#Kreiranje zasebnog objekta sa svim SNP-ovima koji su ostvarili vrijednosti  $p_{eROHi}>5$ :
SNPovi<-subset(froh_chr1_sve,froh_chr1_sve$minus_log10>5)

#Vizualizacija zasebnih kromosoma na kojima SNP-ovi prelaze treshold:
#(uz pomocu objekta "SNPovi" na plot-u označavamo sve outlier SNP-ove);
#naredba "annotatePval" na plot-u vizualizira najznačajniji SNP;

# Chr 1
froh_chr1_sve<-subset(Froh1_sve,Froh1_sve$chr==1)
manhattan (froh_chr1_sve, bp= "bp", chr= "chr", p= "pvalue_R", snp="snp_name",
logp= TRUE, ylim=c(0,20), main="Outlier SNP-ovi za Chr 1",suggestiveline =FALSE,
genomewideline = 5, highlight = zadnje$snp_name,annotatePval = 0.005)

# Chr 3
froh_chr3_sve<-subset(Froh1_sve,Froh1_sve$chr==3)
manhattan (froh_chr3_sve, bp= "bp", chr= "chr", p= "pvalue_R", snp="snp_name",
logp= TRUE, ylim=c(0,20), main="Outlier SNP-ovi za Chr 3",suggestiveline =FALSE,
genomewideline = 5, highlight = zadnje$snp_name,annotatePval = 0.005)

# Chr 5
froh_chr5_sve<-subset(Froh1_sve,Froh1_sve$chr==5)
```

```
manhattan (froh_chr5_sve, bp= "bp", chr= "chr", p= "pvalue_R", snp="snp_name",  
logp= TRUE, ylim=c(0,20), main="Outlier SNP-ovi za Chr 5",suggestiveline =FALSE,  
genomewideline = 5, highlight = zadnje$snp_name,annotatePval = 0.005)
```

```
# Chr 6
```

```
froh_chr6_sve<-subset(Froh1_sve,Froh1_sve$chr==6)  
manhattan (froh_chr6_sve, bp= "bp", chr= "chr", p= "pvalue_R", snp="snp_name",  
logp= TRUE, ylim=c(0,20), main="Outlier SNP-ovi za Chr 6",suggestiveline =FALSE,  
genomewideline = 5, highlight = zadnje$snp_name,annotatePval = 0.005)
```

```
# Chr 7
```

```
froh_chr7_sve<-subset(Froh1_sve,Froh1_sve$chr==7)  
manhattan (froh_chr7_sve, bp= "bp", chr= "chr", p= "pvalue_R", snp="snp_name",  
logp= TRUE, ylim=c(0,20), main="Outlier SNP-ovi za Chr 7",suggestiveline =FALSE,  
genomewideline = 5, highlight = zadnje$snp_name,annotatePval = 0.005)
```

```
# Chr 8
```

```
froh_chr8_sve<-subset(Froh1_sve,Froh1_sve$chr==8)  
manhattan (froh_chr8_sve, bp= "bp", chr= "chr", p= "pvalue_R", snp="snp_name",  
logp= TRUE, ylim=c(0,20), main="Outlier SNP-ovi za Chr 8",suggestiveline =FALSE,  
genomewideline = 5, highlight = zadnje$snp_name,annotatePval = 0.005)
```

```
# Chr 9
```

```
froh_chr9_sve<-subset(Froh1_sve,Froh1_sve$chr==9)  
manhattan (froh_chr9_sve, bp= "bp", chr= "chr", p= "pvalue_R", snp="snp_name",  
logp= TRUE, ylim=c(0,20), suggestiveline =FALSE, genomewideline = 5, highlight =  
zadnje$snp_name,annotatePval = 0.005)
```

```
# Chr 10
```

```
froh_chr10_sve<-subset(Froh1_sve,Froh1_sve$chr==10)  
manhattan (froh_chr10_sve, bp= "bp", chr= "chr", p= "pvalue_R", snp="snp_name",  
logp= TRUE, ylim=c(0,20), main="Outlier SNP-ovi za Chr 10",suggestiveline =FALSE,  
genomewideline = 5, highlight = zadnje$snp_name,annotatePval = 0.005)
```

```
# Chr 11
```

```
froh_chr11_sve<-subset(Froh1_sve,Froh1_sve$chr==11)  
manhattan (froh_chr11_sve, bp= "bp", chr= "chr", p= "pvalue_R", snp="snp_name",  
logp= TRUE, ylim=c(0,20), main="Outlier SNP-ovi za Chr 11",suggestiveline =FALSE,  
genomewideline = 5, highlight = zadnje$snp_name,annotatePval = 0.005)
```

```
# Chr 14
```

```
froh_chr14_sve<-subset(Froh1_sve,Froh1_sve$chr==14)
```



```
manhattan (froh_chr14_sve, bp= "bp", chr= "chr", p= "pvalue_R", snp="snp_name",  
logp= TRUE, ylim=c(0,20), main="Outlier SNP-ovi za Chr 14",suggestiveline =FALSE,  
genomewideline = 5, highlight = zadnje$snp_name,annotatePval = 0.005)
```

```
# Chr 15
```

```
froh_chr15_sve<-subset(Froh1_sve,Froh1_sve$chr==15)  
manhattan (froh_chr15_sve, bp= "bp", chr= "chr", p= "pvalue_R", snp="snp_name",  
logp= TRUE, ylim=c(0,20), main="Outlier SNP-ovi za Chr 15",suggestiveline =FALSE,  
genomewideline = 5, highlight = zadnje$snp_name,annotatePval = 0.005)
```

```
# Chr 16
```

```
froh_chr16_sve<-subset(Froh1_sve,Froh1_sve$chr==16)  
manhattan (froh_chr16_sve, bp= "bp", chr= "chr", p= "pvalue_R", snp="snp_name",  
logp= TRUE, ylim=c(0,20), main="Outlier SNP-ovi za Chr 16",suggestiveline =FALSE,  
genomewideline = 5, highlight = zadnje$snp_name,annotatePval = 0.005)
```

```
# Chr 17
```

```
froh_chr17_sve<-subset(Froh1_sve,Froh1_sve$chr==17)  
manhattan (froh_chr17_sve, bp= "bp", chr= "chr", p= "pvalue_R", snp="snp_name",  
logp= TRUE, ylim=c(0,20), main="Outlier SNP-ovi za Chr 17",suggestiveline =FALSE,  
genomewideline = 5, highlight = zadnje$snp_name,annotatePval = 0.005)
```

```
# Chr 18
```

```
froh_chr18_sve<-subset(Froh1_sve,Froh1_sve$chr==18)  
manhattan (froh_chr18_sve, bp= "bp", chr= "chr", p= "pvalue_R", snp="snp_name",  
logp= TRUE, ylim=c(0,30), main="Outlier SNP-ovi za Chr 18",suggestiveline =FALSE,  
genomewideline = 5, highlight = zadnje$snp_name,annotatePval = 0.005)
```

```
# Chr 20
```

```
froh_chr20_sve<-subset(Froh1_sve,Froh1_sve$chr==20)  
manhattan (froh_chr20_sve, bp= "bp", chr= "chr", p= "pvalue_R", snp="snp_name",  
logp= TRUE, ylim=c(0,20), main="Outlier SNP-ovi za Chr 20",suggestiveline =FALSE,  
genomewideline = 5, highlight = zadnje$snp_name,annotatePval = 0.005)
```

```
# Chr 21
```

```
froh_chr21_sve<-subset(Froh1_sve,Froh1_sve$chr==21)  
manhattan (froh_chr21_sve, bp= "bp", chr= "chr", p= "pvalue_R", snp="snp_name",  
logp= TRUE, ylim=c(0,20), main="Outlier SNP-ovi za Chr 21",suggestiveline =FALSE,  
genomewideline = 5, highlight = zadnje$snp_name,annotatePval = 0.005)
```

```
# Chr 22
```

```
froh_chr22_sve<-subset(Froh1_sve,Froh1_sve$chr==22)
```

```
manhattan (froh_chr22_sve, bp= "bp", chr= "chr", p= "pvalue_R", snp="snp_name",  
logp= TRUE, ylim=c(0,20), main="Outlier SNP-ovi za Chr 22",suggestiveline =FALSE,  
genomewideline = 5, highlight = zadnje$snp_name,annotatePval = 0.005)
```

```
# Chr 25
```

```
froh_chr25_sve<-subset(Froh1_sve,Froh1_sve$chr==25)  
manhattan (froh_chr25_sve, bp= "bp", chr= "chr", p= "pvalue_R", snp="snp_name",  
logp= TRUE, ylim=c(0,20), main="Outlier SNP-ovi za Chr 25",suggestiveline =FALSE,  
genomewideline = 5, highlight = zadnje$snp_name,annotatePval = 0.005)
```

```
# Chr 30
```

```
froh_chr30_sve<-subset(Froh1_sve,Froh1_sve$chr==30)  
manhattan (froh_chr30_sve, bp= "bp", chr= "chr", p= "pvalue_R", snp="snp_name",  
logp= TRUE, ylim=c(0,20), main="Outlier SNP-ovi za Chr 30",suggestiveline =FALSE,  
genomewideline = 5, highlight = zadnje$snp_name,annotatePval = 0.005)
```

```
# Chr 34
```

```
froh_chr34_sve<-subset(Froh1_sve,Froh1_sve$chr==34)  
manhattan (froh_chr34_sve, bp= "bp", chr= "chr", p= "pvalue_R", snp="snp_name",  
logp= TRUE, ylim=c(0,20), main="Outlier SNP-ovi za Chr 34",suggestiveline =FALSE,  
genomewideline = 5, highlight = zadnje$snp_name,annotatePval = 0.005)
```

```
# Export podataka
```

```
write_xlsx(Froh1_sve,"C:\\Users\\Vjerman\\Desktop\\sveza  
diplomski\\eROHi_sve_LGD_pasmine.xlsx")
```

## 11. Životopis

Vjeran Glavaš rođen je 29. prosinca 1997. godine u Osijeku, Republika Hrvatska, od oca Davora i majke Sande rođene Vučak.

Prvi i drugi razred osnovne škole pohađao je u Meci u Osnovnoj školi Dardi, podružnica Meca, dok je ostale razrede pohađao u Osnovnoj školi Darda u Dardi gdje je 2012. godine i završio osnovno školovanje s vrlo dobrim uspjehom.

Srednju Poljoprivrednu i veterinarsku školu u Osijeku pohađao je u vremenu od 2012. do 2016. te stekao zvanje veterinarski tehničar.

Tijekom srednjoškolskog obrazovanja sudjeluje na državnom natjecanju u disciplini Veterina, održanog u Zadru 22.04.2016. te osvaja 2. mjesto.

Sveučilišni preddiplomski studij Animalne znanosti pohađao je od 2016. do 2019. godine kada je obranom završnog rada na temu „Svojstva za selekciju u proizvodnji mlijeka“ stekao akademski naziv Sveučilišni prvostupnik inženjer animalnih znanosti.

Student je 2. godine sveučilišnog diplomskog studija Genetika i oplemenjivanje životinja, na Sveučilištu grada Zagreba, Agronomski fakultet.

Odlično se služi engleskim jezikom u govoru i pismu. Hobii su mi briga o životinjama, crtanje i fotografija. Tijekom dvije godine studiranja bio je tutor mlađim generacijama. Zna se odlično služiti računalima.