

Sveučilište u Zagrebu
Agronomski fakultet

Mario Shihabi, Katarina Turkalj

**Procjena koeficijenta inbridinga na temelju X kromosoma
istarske i paške ovce**

Zagreb, 2021.

Ovaj rad izrađen je na Zavodu za opće stočarstvo Agronomskog fakulteta Sveučilišta u Zagrebu pod vodstvom doc. dr. sc. Maje Ferenčaković i predan je na natječaj za dodjelu Rektorove nagrade u akademskoj godini 2020./2021.

Sadržaj

1. UVOD.....	1
1.1. INBRIDING.....	2
1.1.1. Djelovanje inbridinga.....	2
1.2. METODE PROCJENE INBRIDINGA.....	4
1.2.1. Procjena inbridinga putem rodovnika.....	4
1.2.2. Genomska procjena inbridinga.....	5
1.2.2.1. Markeri visoke rezolucije.....	6
1.2.2.2. Metode genomske procjene inbridinga.....	6
1.2.2.3. ROH i F_{ROH}	7
1.2.2.4. Wright-ova F statistika.....	10
1.3. UTJECAJ SPOLA NA IZRAČUN GENETSKIH PARAMETARA.....	11
1.3.1. Evolucija spolnih kromosoma.....	11
1.3.2. Važnost kromosoma X u procjeni inbridinga.....	12
1.3.3. Procjena inbridinga na spolnom kromosomu X.....	13
1.4. KARAKTERISTIKE ISTARSKE I PAŠKE OVCE.....	14
1.4.1. Istarska ovca.....	15
1.4.2. Paška ovca.....	16
2. HIPOTEZE I OPĆI I SPECIFIČNI CILJEVI RADA.....	17
2.1. HIPOTEZE.....	17
2.2. SPECIFIČNI CILJEVI RADA.....	17

3. MATERIJAL I METODE.....	18
3.1. UZORCI, GENOTIPIZACIJA I KONTROLA KVALITETE GENOTIPOVA	18
3.2. RODOVNICI I PROCJENA KOEFICIJENTA INBRIDINGA IZ RODOVNIKA .	18
3.3. PROCJENA INBRIDINGA IZ GENOTIPSKIH PODATAKA.....	19
3.4. STATISTIČKA OBRADA I VIZUALIZACIJA	20
4. REZULTATI.....	21
4.1. ANALIZA RODOVNIKA I PROCJENA INBRIDING KOEFICIJENATA NA TEMELJU RODOVNIKA.....	21
4.2. ANALIZA GENOMSKIH PODATAKA I PROCJENA INBRIDING KOEFICIJENATA NA TEMELJU GENOMSKIH PODATAKA	21
4.2.1. ROH segmenti i F_{ROH} inbriding koeficijenti.....	21
4.2.2. Wright-ovi F_{IS} koeficijenti	25
4.3. USPOREDBA KOEFICIJENATA INBRIDINGA	27
4.3.1. Koeficijenti izračunati preko rodovnika i genomski inbriding koeficijenti.....	27
4.3.2. Koeficijenti X kromosoma izračunati preko rodovnika i genomski inbriding koeficijenti X kromosoma	27
4.3.3. Koeficijenti inbridinga procijenjeni iz autosoma i iz X kromosoma.....	27
5. RASPRAVA.....	31
6. ZAKLJUČCI	36
7. ZAHVALE	37
8. POPIS LITERATURE.....	38

9. SAŽETAK.....	47
10. SUMMARY.....	48
11. ŽIVOTOPIS.....	49

1. UVOD

Zbog pozitivnog utjecaja na produktivnost i većeg profita, križanje u srodstvu se često koristi u uzgoju životinja i biljaka. Križanjem „sličnog sa sličnim“ postiže se veća ujednačenost i fiksiraju aleli za poželjna svojstva unutar populacije. Međutim, početkom 21. stoljeća, nenamjerne posljedice težnje što većoj produktivnosti domaćih životinja postale su pitanje javne i profesionalne zabrinutosti. Kod modernih pasmina, povećao se udio smrtnosti te neke jedinke pate od bolesti i defekata. Životinje nose naslijeđene nedostatke zbog inbridinga, što ih predisponira na zdravstvene probleme. Inbriding povećava udio homozigota s namjerom fiksacije korisnih alela, a povećava i udio fiksacije štetnih recesivnih alela. Navedene posljedice inbridinga davno su primijećene kod ljudi, stoga gotovo sve kulture zabranjuju križanje bliskih rođaka.

Kroz povijest, inbriding se procjenjivao isključivo putem rodovnika. Takav pristup imao je mnoge nedostatke, ali je bio najprikladniji. Napretkom bioinformatike i molekularne genetike javlja se sve veći broj metoda procjene. U posljednje vrijeme, metoda koja prednjači u izračunu inbridinga je *Runs of Homozygosity* (ROH).

Dosad se izračun inbridinga iz genomskih podataka uglavnom zasnivao samo na autosomima, što nije ispravno ako uzmemo u obzir utjecaj spolno vezanog nasljeđivanja na mnoga ekonomski važna svojstva, ali i općenito na zdravlje i preživljavanje ženskih jedinki. Poznato je kako je inbriding X kromosoma povezan s povećanjem smrtnosti kod ženki, a kako su kod sisavaca ženke homogametnog spola to dovodi do povećanja frekvencije štetnih, spolno ograničenih alela pa čak i njihove fiksacije. Sa stajališta konzervacijske genetike pak treba voditi računa kako promijenjen odnos spolova prema mužjacima uzrokuje povećan rizik od izumiranja vrste, a upravo inbriding X kromosoma može dovesti do te pojave.

1.1. INBRIDING

Inbriding se odnosi na parenje jedinki povezanih podrijetlom (Jordan i sur., 2005.; Ralls i sur., 2014.). U domaćih životinja najčešće se govori o parenju između oca i kćeri, braće i sestara ili prvih rođaka (Woodruff, 2001.) te se uspoređuje u odnosu na baznu populaciju koja se sastoji od jedinki koje nisu u srodstvu (Ballou, 1983.). Kod jedinke, pod pojmom alela identičnih po podrijetlu (*identical by descent*; IBD) smatramo 2 alela unutar lokusa koji potječu od istog alela, odnosno zajedničkog pretka. Suprotno tome, razlikujemo *identical by state* (IBS) alele koje smatramo fizički identičnima, no ne potječu od jedinstvenog zajedničkog pretka (Stevens, 2001.; Ballou, 1983.). Inbriding se kvantificira koeficijentom inbridinga.

Koeficijent inbridinga (F) numerička je procjena stupnja inbridinga jedinke (Bittles, 2001.) i opisuje se kao vjerojatnost da jedinka ima dva alela u lokusu koji su identični podrijetlom. Budući da je F vjerojatnost, odgovara vrijednostima od 0 do 1 koji predstavljaju udio autozigotnosti. Inbriding povećava vjerojatnost da su 2 alela u genotipu IBD te uvelike doprinosi fiksaciji alela i smanjuje udio heterozigotnosti. Heterozigotnost je izravno proporcionalna koeficijentu inbridinga, što uzrokuje inbriding depresiju, povećanje genetskih defekata te redistribuciju varijance (Ralls i sur., 2014.). Neizbježan je (inbriding) u malim, zatvorenim populacijama te se koeficijent inbridinga u efektivnoj populaciji povećava za $1/(2N_e)$ po generaciji (Woodruff, 2001.).

1.1.1. Djelovanje inbridinga

Inbriding povećava vjerojatnost nasljeđivanja recesivnih alela (homozigoti), omogućujući tako izražavanje njihovih pozitivnih, ali i štetnih učinaka. Potpomognut genetskim driftom i snagom selekcije, dovodi do poboljšanja ili takoreći abnormalnosti jedinki, tj. populacije. Osim kao jedan od najčešćih načina povećanja produktivnosti i profita, doprinosi stoljećima u stvaranju pasmina, linija unutar pasmina te pri ispoljavanju željenih osobina, kao

i zadržavanju istih. Međutim, dovodi i do preraspodjele genetskih varijacija između i unutar populacija te smanjenju prosjeka populacije vezanog za *fitness* svojstvo (Curik i sur., 2014.; Charlesworth i Willis, 2009.; Arcos-Burgos i Muenke, 2002.; Alvarezi sur., 2009.).

Slijedom toga, u većini populacija, inbriding rezultira padom vrijednosti svojstava vezanih za preživljavanje i reprodukciju, što se naziva inbriding depresijom. Vidljiva je u biokemiji, morfologiji, ponašanju te metabolizmu i razvoju jedinke (Ralls i sur., 2014.). Generalno, uz najjednostavniji slučaj s jednim lokusom, govori se o inbriding depresiji kao manifestaciji cjelokupnih genomskih učinaka parenja bliskih rođaka (Woodruff, 2001.). Inbriding depresija jest stohastički element zbog slučajnog uzorkovanja alela tijekom reprodukcije koja dovodi do smanjenja srednje vrijednosti nekog kvantitativnog svojstva. Pri tome se pojedinci s istim koeficijentom inbridinga, odnosno jednakom vjerojatnošću nošenja IBD alela, mogu razlikovati u razinama homozigotnosti za bilo koje svojstvo i, prema tome, mogu varirati u stupnju inbriding depresije (Ralls i sur., 2014.). Ako se isti mutirani gen naslijeđi od oba roditelja, pojedinac će izraziti promjenu u nekom obliku, pri rođenju ili kasnije u životu, ovisno o prirodi i mjestu mutacije, pridonoseći tako fenomenu inbriding depresije.

Inbriding depresija se objašnjava dvama teorijama koje počivaju na odstupanju od vrijednosti homozigota. Prva teorija jest teorija dominantnosti koja inbriding depresiju definira kroz ekspresiju fiksiranih štetnih recesivnih alela (ne dolaze do izražaja u heterozigotnom obliku) te zahtjeva superiornost heterozigota nad prosjekom homozigota (Davenport, 1908.; Jones, 1917.). Druga, overdominantna teorija govori da se inbriding depresija pojavljuje kada je heterozigot superioran u odnosu na oba homozigota, a smanjena je učestalost heterozigota (zbog inbridinga), što posljedično dovodi do smanjenog ispoljavanja overdominantnosti (Charlesworth i Charlesworth, 1987.; Charlesworth i Charlesworth, 1999.). Procjena inbriding depresije se vrši putem regresije nekog svojstva na koeficijent inbridinga pa je jasno da nam je

za preciznu procjenu potrebno precizno mjerenje fenotipa (svojstva), ali i pouzdan inbriding koeficijent.

1.2. METODE PROCJENE INBRIDINGA

Tradicionalno se koeficijent inbridinga (F) procjenjivao putem rodovnika, kao što je prikazao Wright (1922.) na populaciji goveda. S brzim razvojem proteinskih markera, mikrosatelita i u zadnje vrijeme SNP-ova širom genoma, F se sve više procjenjuje iz podataka genetskih markera (Wang, 2016.). Oba načina procjene koeficijenta inbridinga ukratko su opisani ispod.

1.2.1. Procjena inbridinga putem rodovnika

Koeficijent inbridinga putem rodovnika je moguće procijeniti pomoću analize putanje (*Path analysis*), metodom *Gene dropping*-a ili tabularnom metodom (odnosno preko matrice aditivne srodnosti) (Ballou, 1983.).

Analiza putanje je najčešća metoda koja se koristi i temelji se na praćenju putanja u rodovniku, od jedinke za koju se računa koeficijent inbridinga do svakog zajedničkog pretka roditelja. Vršiti se izračun vjerojatnosti da je jedinka naslijedila dva alela identična po porijeklu (IBD). Kod praćenja putanje se zatvara krug tako da nema vraćanja na jedinke koje su već ubrojene u tu putanju. Koeficijent inbridinga jedinke jednak je zbroju doprinosa svake putanje, uključujući koeficijent inbridinga svakog zajedničkog pretka. Koristi se sljedeća formula (1) u izračunu (Ballou, 1983.):

$$F_x = \sum^k \left(\frac{1}{2}\right)^m (1 + F_A) \quad (1)$$

gdje je:

k = broj putanja kroz sve zajedničke pretke,

m = broj jedinki u putanji,

F_A = koeficijent inbridinga zajedničkog pretka.

Sljedeća metoda je *gene dropping*, koja predstavlja simulacijsku proceduru u kojoj se hipotetski aleli dodjeljuju svakom *founder*-u u populaciji, a zatim se promatra hipotetska segregacija roditeljskih alela po Mendelu (temelj jest transmisija alela sa roditelja na potomke s vjerojatnošću omjera 50:50 tijekom mejoze) (MacCluer i sur., 1986.; Jung i sur., 2006.).

Tabularna metoda temelji se na matrici aditivne srodnosti između svih jedinki u rodovniku; tj. na principu zajedničkih alela između 2 jedinke koji se onda izražava kao prosječna srodnost između jedne jedinke i roditelja druge (Chang i sur., 1991.).

Općenito, procjena inbridinga na temelju rodovnika ima nekoliko izrazitih nedostataka, kao što su ne uzimanje u obzir inbridinga bazne populacije, odnosno smatranje da početne jedinke u rodovniku imaju inbriding jednak 0. Također, ne uzima u obzir stohastičku prirodu rekombinacije, odnosno faktor šanse (*Mendelian sampling*) (Ferenčaković i sur., 2013a.), kao i potencijalnu pristranost rezultata zbog selekcije (Curik i sur., 2002.). Nedostatak je i što dolazi do čestih pogrešaka u rodovniku zbog netočnih zapisa, krive interpretacije i krive identifikacije.

1.2.2. Genomska procjena inbridinga

U 21. stoljeću, sve većim razvojem tehnologije i genotipiziranja, pojavila se mogućnost genomske procjene inbridinga na temelju molekularne informacije. Procjena se vrši određenim metodama, na temelju markera koji predstavljaju varijabilna mjesta genoma. Genomskom procjenom inbridinga ispravljaju se navedeni nedostaci te se pruža uvid na kojim kromosomima je više zastupljen inbriding, odnosno gdje primarno djeluje direktna selekcija.

1.2.2.1. *Markeri visoke rezolucije*

Nakon dosta duge dominacije mikrosatelitskih markera na području molekularne genetike za proučavanje genoma ljudi i životinja, pojavljuje se *single nucleotide polymorphism* (SNP) (Vignal i sur., 2002.). Kao što sugerira skraćenica, SNP-ovi jesu varijante u genomu (Brody, 2016.), jedna promjena baze u sekvenci DNA, s uobičajenom zamjenom dva moguća nukleotida na danom položaju (Koopae i Koshkoiyeh, 2014.). Javljaju se u kodirajućim i nekodirajućim regijama genoma (Marwal i sur., 2004.), a kako bi se taj osnovni položaj s alternativnim sekvencama mogao uopće smatrati SNP-om, smatra se da bi najrjeđi alel trebao imati primjetnu učestalost od 1 % ili veću (ako je varijacija < 1 % radi se o točkastoj mutaciji) (Vignal i sur., 2002.).

Veliki broj SNP-ova mogu se istovremeno obrađivati pomoću tehnologija sekvenciranja sljedeće generacije od strane tvrtki poput Illumine (Caetano, 2009; Butler, 2012). Tehnologije vezane uz SNP u početku su bile intenzivne, vremenski zahtjevne i smatrane skupim procesima do danas kada su jedan od onih najautomatiziranijih, učinkovitijih i relativno jeftinih metoda (Marwal i sur., 2004.). SNP-ovi na SNP čipu predstavljaju genomsku informaciju na temelju koje je moguće procijeniti inbriding određenim metodama, koje su navedene u sljedećem poglavlju.

1.2.2.2. *Metode genomske procjene inbridinga*

U posljednjih nekoliko godina, predložene su brojne metode procjene koeficijenta inbridinga iz genomske informacije, ali ne postoji konsenzus koja je najprikladnija (Goudet i sur., 2018.). Koeficijent inbridinga može se procijeniti: pristupima najveće vjerojatnosti (*maximum likelihood*) (Milligan, 2003.; Wang, 2007.), metodama momenta (Ritland, 1996.; Purcell i sur., 2007.), iz elemenata dijagonale genomske matrice srodnosti (*Genomic relationship matrix*; GRM) (VanRaden, 2008.), iz jednostavnih mjera heterozigotnosti ili

homozigotnosti (*Homozygosity*; HOM) (Szulkin i sur., 2010.; Bjelland i sur., 2013.), na temelju genotipskih korelacija (Ackerman i sur., 2017.) te kao udio genoma unutar homozigotnih odsječaka (ROH) (McQuillan i sur., 2008.; Ferenčaković i sur., 2013a.). Alemu i sur. (2021.) usporedili su navedene metode procjene, koristeći sekvence cijelog genoma od 245 jedinki Holstein pasmine i ustanovili da svaka od njih ima prednosti i nedostatke, ovisno o cilju istraživanja. Metode povezane s homozigotnosti (F_{ROH} i F_{HOM} metoda) su bile učinkovitije u procjeni udjela genoma koji je IBD (inbriding), bez obzira na učestalost ili starost alela. Također, ustanovljena je visoka korelacija između F_{ROH} i F_{HOM} metode (0,96). Zbog prethodno navedenog te trenutne popularnosti u upotrebi, naše istraživanje je koristilo F_{ROH} metodu za procjenu inbridinga na temelju markera na spolnom kromosomu X. Također, kao indikator homozigotnosti, druga metoda koja se koristila u ovom istraživanju je F_{IS} , fiksacijski indeks iz Wright-ove F statistike.

1.2.2.3. ROH i F_{ROH}

ROH-ovi su kontinuirani, neprekidni odsječci homozigotnih genotipova bez heterozigotnosti u diploidnom stanju (Gibson i sur., 2006.). Smatra se da je glavni uzrok njihova nastajanja križanje u srodstvu i stoga su dobar alat u procjeni razine inbridinga, posebice kad nije dostupan rodovnik (Ferenčaković i sur., 2013b.). Što je veća dužina homozigotnog odsječka, to je i veća vjerojatnost da predstavlja inbriding te zato ne uzimamo u obzir odsječke kraće od određene dužine, specifične za pojedinu vrstu, koja ovisi o mjeri LD-a (*linkage disequilibrium*). Također, ako uzmemo u obzir princip rekombinacije, prema duljini homozigotnog odsječka moguće je odrediti broj generacija od zajedničkog pretka. Veoma dugi odsječci predstavljaju nedavni inbriding (16 Mb dug segment ~3 generacije), dok kraći homozigotni odsječci predstavljaju davni inbriding (1 Mb dug segment ~50 generacija) (Ferenčaković i sur., 2013a.). Očekivana duljina autozigotnog segmenta slijedi

eksponencijalnu distribuciju s prosjekom od $\frac{1}{2}g$ Morgana, gdje g predstavlja broj generacija od zajedničkog pretka (Howrigan i sur., 2011.).

Postoje dvije glavne metode za identificiranje ROH-a: promatrački algoritmi brojanja genotipova i algoritmi temeljeni na modelima (*hidden Markov model*; HMM). Prva metoda putem algoritama skenira svaki kromosom pomičući prozor određene veličine duž duljine genoma detektirajući segmente homozigotnih SNP-ova. Sam korisnik određuje ulazne parametre poput broja dozvoljenih heterozigota, dopuštenog broja *missing* SNP-ova kako bi se tolerirale eventualne pogreške i propusti genotipizacije (Curik i sur., 2014.). Takav pristup je implementiran u PLINK. Još neki od programa za procjenu ROH-a su cgaTOH softver i Germline (koristi *haplotype matching* za procjenu IBD segmenata) (Ceballos i sur., 2018.). Pristupi temeljeni na modelima koji koriste HMM su druga glavna metoda, a jedan takav postoji unutar Beagle-a i RZooROH programa (Druet i Gautier, 2017.; Ceballos i sur., 2018.). Jedan od pristupa zabilježen je i kod Leutenegger i sur. (2003.) te se temelji na prozorima, uz drugačije modeliranje IBD procesa na kromosomu putem *hidden Markov* modela kako bi se identificirali HBD segmenti. Isto tako, Vigeland (2020.) u svom radu navodi paket ribd kao jedini sa kompletnim funkcijama za izračun srodstva i koeficijenta identiteta za autosome i X kromosom.

Novija istraživanja, poput Solé i sur. (2017.), pomoću *hidden Markov* modela, obrađuju procjenu vjerojatnosti *homozygosity by descent* (HBD) za svaku poziciju markera i dijele autozigotnost u više klasa ovisno o duljini HBD segmenta (Druet i Gautier, 2017.). HMM je statistički model, razvijen kako bi se procijenila vjerojatnost HBD segmenta duž kromosoma, iskoristile informacije o markerima, frekvencijama alela i postotcima genotipskih pogrešaka. Takav model može podnijeti *whole genome sequence* skup genomske matrice srodnosti (*genomic relationship matrix*; GRM) bez informacija o rodovniku putem BLUPF90 paketa.

No, ono u čemu, između ostalog, ROH prednjači jest mogućnost razlikovanja „novijeg“ od „starijeg“ inbridinga. Proučavanjem rezultata dobivenih od strane Solé i sur. (2017.), procjene dobivene HMM-om su bile u visokoj korelaciji sa onima dobivenim putem ROH-a detektiranog PLINK-om.

McQuillan i sur. (2008.) su uveli F_{ROH} kao genomsku mjeru autozigotnosti jedinke i definirali ga kao udio genoma, koji se nalazi u ROH-u određene minimalne duljine u odnosu na ukupni genom, u području interesa. Opća formula za izračun F_{ROH} iz SNP čip podataka je $F_{ROH} = L_{ROH} / L_{AUTOSOME}$, gdje L_{ROH} je ukupna duljina svih ROH-ova u genomu jedinke, koji sadrže minimalno određeni broj uzastopnih homozigotnih SNP-ova, a $L_{AUTOSOME}$ se odnosi na duljinu autosomalnog genoma obuhvaćenog SNP-ovima na SNP čipu.

Također, bitno je istaknuti da gustoća SNP čipa koji se koristi za identifikaciju ROH segmenata ima veliki utjecaj na učinkovitost identifikacije. Purfield i sur. (2012.) istraživali su razlike u ROH segmentima identificiranim putem dva SNP čipa koji se najčešće koriste kod goveda: Illumina BovineSNP50 Genotyping BeadChip s 54 001 SNP-a (50 k), i Illumina BovineHD Genotyping BeadChip sa 777 972 SNP-a (HD). Rezultati sugeriraju da je 50 k SNP čip prikladan samo za identifikaciju ROH-ova dužih od 5 Mb. Odnosno, zbog njegove gustoće, 50 k SNP čip ima veću stopu greške zbog manje mogućnosti otkrivanja heterozigotnosti unutar ROH-a promatrane dužine (Ferenčaković i sur., 2013b.). Kao rezultat toga, analize temeljene na takvim čipovima mogu precijeniti broj segmenata kraćih od 4 Mb. Naše istraživanje je koristilo HD SNP čip kako bi se minimalizirale navedene pogreške, odnosno kako bi se mogli identificirati i ROH-ovi manjih dužina, do minimalne dužine koja je potencijalno nastala inbridingom.

1.2.2.4. Wright-ova F statistika

Koristeći koeficijent inbridinga, Wright definira „fiksacijske indekse“ te tako F statistika uključuje 3 indeksa (F_{IS} , F_{IT} , F_{ST}) (Hartl i Clark, 1997., Othman i sur., 2005.). Wrightovi (1949.) indeksi fiksacije su uvelike od koristi prilikom opisa raspodjele genetske raznolikosti između i unutar populacija, odnosno subpopulacija (Weir i Cockerham, 1984.). Nadalje, Nei 1977. godine, definira indekse fiksacije F_{IS} , F_{IT} i F_{ST} na sljedeći način:

$$F_{IT} = 1 - H_0 / H_T, (1)$$

$$F_{IS} = 1 - H_0 / H_S, (2)$$

$$F_{ST} = 1 - H_S / H_T, (3)$$

gdje H_0 označava učestalost svih heterozigota, dok H_S i H_T predstavljaju očekivane heterozigotnosti u Hardy-Weinbergovoj ravnoteži (*Hardy-Weinberg equilibrium*; HWE) ili u različitostima gena unutar subpopulacija, odnosno u ukupnoj populaciji. Prvenstveno su bili definirani u vidu korelacije dviju sjedinjenih gameta (Nei i Chesser, 1983.), a obzirom na to pokazano je (Wright, 1949.) da je korelacija između slučajnih gameta, izvučenih iz iste subpopulacije, u odnosu na ukupan broj, dana formulom $F_{ST} = (F_{IT} - F_{IS}) / (1 - F_{IS})$.

Određena razina inbridinga je neizbježna u malim subpopulacijama. Dakle, genetska diferencijacija među subpopulacijama putem F statistike može se protumačiti kao neka vrsta učinka srodstva koji proizlazi iz substrukture populacije (Hartl i Clark, 1997.). F_{IT} možemo podijeliti na F_{ST} i F_{IS} zbog Wahlund efekta, odnosno zbog inbridinga. Taj fiksacijski indeks, opisan formulom (1), prikazuje razinu inbridinga u odnosu na HW populaciju, a do promjene heterozigotnosti jedinke dovodi zbog neslučajnog parenja, odnosno podjele populacije na subpopulacije u odnosu na ukupnu populaciju. F_{IS} (2) prikazuje promjenu heterozigotnosti

jedinke zbog ne slučajnog parenja unutar subpopulacije (intervala od -1 do 1, gdje 1 kao visoka vrijednost ukazuje na visoku stopu inbridinga (Moioli i sur., 2001.). Prema Wrightu uvriježeno je za F_{ST} (3) da S označava subpopulaciju, a T ukupnu populaciju (Weir i Hill, 2002.). Ovaj indeks fiksacije objašnjava smanjenje heterozigotnosti uzrokovano podjelom na subpopulacije zbog genetskog drifta, a sam je uvijek pozitivne vrijednosti zbog toga što očekivana heterozigotnost u subpopulaciji ne može biti veća od one u ukupnoj (Hartl i Clark, 1997., Othman i sur., 2005., Weir i Cockerham, 1984.).

1.3. UTJECAJ SPOLA NA IZRAČUN GENETSKIH PARAMETARA

1.3.1. Evolucija spolnih kromosoma

U mnogim aspektima, spolni kromosom X je veoma sličan autosomu. Na to ukazuje i njegovo evolucijsko porijeklo. Dva spolna kromosoma, X i Y, divergirali su iz jedinstvenog autosoma prije ~300 milijuna godina (Lahn i Page, 1999.), još uvijek su homologni i rekombiniraju međusobno u blizini svojih krajeva, unutar dvaju pseudoautosomalnih regija. Drugdje im se evolucijski putevi razilaze. Y kromosom je izgubio najveći dio sekvence i gena te se razvio u jedinstveni uzorak ponavljajuće sekvence, u kojoj je osim na krajevima, varijabilnost uzrokovana jedino mutacijama (Skaletsky i sur., 2003.; Jobling i Tyler-Smith, 2003.). Nasuprot tome, X kromosom nije izgubio autosomalni oblik, fizički je najstabilniji jezgrin kromosom (barem kod placentalnih sisavaca) i jedini je u širem pogledu zadržao potpunu sinteničnost između miša i čovjeka (Waterston i sur., 2002.). Važno je istaknuti činjenicu da je divergencija spolnih kromosoma uzrokovala divergenciju spolova, što potvrđuje da kod sisavaca, ženke posjeduju dva spolna kromosoma X, a mušjaci po jedan spolni kromosom X i Y (Waters i sur., 2007.). Zbog takve raspodjele spolnih kromosoma, Y kromosom se ne može upotrijebiti u izračunu inbridinga (Grossman, 1990.). Međutim, jasna je mogućnost upotrebe X kromosoma u izračunima mnogih genetskih parametara kod ženskih

jedinki, kao što je inbriding, što dosad u praksi često nije bio slučaj, budući da su se istraživanja zasnivala samo na autosomima.

1.3.2. Važnost kromosoma X u procjeni inbridinga

Veličina kromosoma X, kod sisavaca, varira od 150 do 170 Mb, ovisno o vrsti, što čini ~5 % genoma. Karakterizira ga stopa rekombinacije od 0,8 cM/Mb, niska stopa mutacije (0,015 mutacija/Mb/generaciji), umjerena genetska raznolikost (0,04 %) i efektivna veličina populacije (N_e) koja iznosi $\frac{3}{4}$ u odnosu na autosome (Schaffner, 2004.). Smanjena razina genetske raznolikosti u odnosu na autosome, uzrokovana je izravno i neizravno. Izravno, nižom stopom mutacija i rekombinacija, a neizravno, povećanim djelovanjem selekcije i genetskog drifta do kojih dolazi zbog manje efektivne veličine populacije, što je posljedica niže stope rekombinacija. Niža stopa rekombinacije je posljedica mogućnosti rekombinacije na cijelom kromosomu X samo kod ženskog spola, a iznosi oko $\frac{2}{3}$ u odnosu na prosjek genoma i uzrokuje višu razinu LD-a (*linkage disequilibrium*) te samim time i brže djelovanje selekcije (Schaffner, 2004.). Pretpostavka je da selekcija vodi ka fiksaciji korisnih recesivnih mutacija u odnosu na manje poželjne oblike. Međutim, zbog LD-a, također se fiksiraju i aleli koji se nalaze u blizini korisne mutacije, a što je veći LD, to je i veća udaljenost fiksiranja od korisne recesivne mutacije.

Inbriding na kromosomu X povezan je s povećanjem smrtnosti kod ženki, što vodi ka pomaku odnosa ženskog i muškog spola prema muškom i smanjenju efektivne veličine populacije. Povećanje smrtnosti homogametnog spola je uzrokovano povećanjem frekvencije štetnih, spolno ograničenih alela. Odnosno, povećanjem X kromosomalnog koeficijenta inbridinga dolazi do fiksacije velikog broja štetnih alela koji pojedinačno imaju mali utjecaj, što rezultira većom inbriding depresijom kod ženskog spola. Nagnutost odnosa spolova prema mužjacima uzrokuje povećan rizik od izumiranja vrste (Robinson i sur., 2014.).

1.3.3. Procjena inbridinga na spolnom kromosomu X

Kod sisavaca, ženski spol posjeduje dva, a muški samo jedan X kromosom. Zato se, kod muških jedinki, X kromosom nalazi u hemizigotnom obliku (osim u blizini krajeva). Prema tome, procjenu koeficijenta inbridinga (F) na spolnom kromosomu X možemo izvršiti samo kod ženskih jedinki, gdje je F jednak vjerojatnosti da su dva homologna alela naslijeđena od roditelja, identična po porijeklu (IBD), kao i kod autosomalnog nasljeđivanja (Grossman, 1990.).

Malo istraživanja je obrađeno na temu procjene koeficijenta inbridinga na spolnom kromosomu X ili o tome kako se uspoređuje s koeficijentom inbridinga procijenjenim na autosomima (Calderón i sur., 2009.). Računanje X kromosomske matrice srodnosti moglo bi biti informativno za bolje razumijevanje arhitekture složenih osobina i za poboljšanje predviđanja istih (Druet i Legarra, 2020.), a Fernando i Grossman (1990.) iznijeli su „pravila“ za konstrukciju matrice srodnosti X kromosoma i njezinog inverza (iz podataka rodovnika). Kako navode, ova se matrica razlikuje od svoga autosomnog pandana (*numerator relationship matrix*) po tome što za matricu koja se odnosi na X kromosom dijagonalne vrijednosti za muške jedinke iznose 0,5 te ne postoji veza između muške jedinke i njegova oca (*sire*). Također navode kako je razvijena i tabularna metoda izrade matrice kovarijanci između srodnika za X kromosomalne lokuse (Fernando i Grossman, 1990.).

Prilikom izračuna X kromosomalnog koeficijenta inbridinga putem rodovnika koristi se formula u odlomku navedenom u poglavlju 1.2.1., uz manje korekcije: u putanji se uzimaju u obzir samo ženske jedinke (definirane su kao X kromosomalni koeficijent *kinship*-a između roditelja jedinke), a ovisno o broju muških jedinki (Johnston i sur., 2019.) – svaka putanja gdje se javlja dvije ili više muških jedinki se uklanja (zbog nepristranosti). Kudo i Sakaguchi (1963.) također uzimaju u obzir odnos potomak – roditelj, ovisno o tome je li zajednički predak

ženskog ili muškog spola. Nadalje, postoji i KinInbcoefX program, prema Zheng i Bourgain (2009.). Radi se o programu za C++ koji računa *kinship* X kromosoma i inbriding koeficijente u samovoljno srodnim jedinkama temeljem podataka prikupljenih o spolu i iz rodovnika. Pritom je, koeficijent srodstva X kromosoma (C_{AB}), za jedinke A i B vjerojatnost da je slučajno odabrani par alela, jedan iz A i jedan iz B ne pseudoautosomalnog lokusa na X kromosomu, identičan po porijeklu. Ova metoda koeficijent inbridinga X kromosoma za mušku jedinku uvijek definira kao 1, bez obzira na to postoji li inbriding ili ne. Koeficijent inbridinga X kromosoma za ženku definira se kao koeficijent srodstva X kromosoma između njezinih roditelja (Zheng i Bourgain, 2009.).

Procjena koristeći mjere inbridinga zasnovane na frekvenciji alela ili putem ROH-a na autosomima i X kromosomu ženskih jedinki zabilježena je u radu Yengo i sur. (2019.). Preciznost i pouzdanost detekcije ROH segmenata utječe i na njihovu daljnju upotrebu, stoga je veoma važno pronalaženje najkompetentnije metode. Nakon detekcije ROH segmenata, prije navedena formula koeficijenta inbridinga iz ROH koeficijenata (McQuillan i sur., 2008.) se može prenamijeniti na kromosom X, gdje L_{ROH} predstavlja duljinu svih ROH-ova minimalne dužine na kromosomu X, a $L_{AUTOSOME}$ predstavlja duljinu kromosoma X koja je obuhvaćena SNP-ovima na SNP čipu. Taj princip je koristilo naše istraživanje.

1.4. KARAKTERISTIKE ISTARSKE I PAŠKE OVCE

Istarska i paška ovca pripadaju autohtonim pasminama ovaca Hrvatske. Sve naše autohtone pasmine ovaca imaju manje ili više izražen lokalni karakter. Zbog dugotrajnog uzgajanja na istom području, prilagodile su se određenim klimatskim i hranidbenim uvjetima te postale otporne na bolesti. Odlikuje ih visoka otpornost, skromnost i prilagodljivost različitim uvjetima držanja te proizvodnim ciljevima. Pripadaju skupini ovaca kombiniranih proizvodnih svojstava, s tim da se ponajviše koriste za proizvodnju mesa i mlijeka.

Karakterizira ih ekstenzivni ili umjereno ekstenzivni sustav uzgoja (Pandek i sur., 2005.). Svoj potpuni rast i razvoj završavaju u dobi od 3 do 4 godine. Specifičnost im je i snažno izražen instinkt za pronalaženje hrane i na najoskudnijim terenima. U priobalju i otocima, „preživljavanje“ u surovim mediteranskim uvjetima, ljetne vrućine i često gladovanje zimi uz jake vjetrove i posolicu, formirali su ovce malog tjelesnog okvira. U unutrašnjosti Istre, karakterističan je malo krupniji tip ovaca (Šimpraga i sur., 2008.).

1.4.1. Istarska ovca

Istarska ovca je nastala i uzgaja se na području Istre, što je i vidljivo iz samog imena. Kada se govori o podrijetlu ove pasmine, pretpostavlja se da je rezultat oplemenjivanja autohtone pramenke sa uvoznim pasminama (najčešće talijanskim). Pripada skupini dugorepih pasmina i ima ispupčen profil nosne kosti u muških i ženskih grla (izraženiji je u ovnova) te postojanje rogova u većine grla (Hrvatski savez uzgajivača ovaca i koza, 2021.). Runo istarske ovce uglavnom je poluzatvoreno do otvoreno, šareno sa temeljnom crnom bojom s bijelim prugama ili bijelom s crnim, smeđim ili sivim pjegama (Mikulec i sur., 2007.). Također, Mikulec i sur. (2007.) navode kako su ovce krupnije, izduženijih ekstremiteta u odnosu na ostale hrvatske pasmine te ne pripadaju skupini kasnozrelih pasmina. Tjelesna masa izražena u kilogramima za ovce iznosi do 70, a u ovnova istarske ovce i do 100 (Hrvatski savez uzgajivača ovaca i koza, 2021.). Veličina matične populacije prema spolu (Hrvatske izvorne i zaštićene pasmine – Banka gena domaćih životinja, 2019.) broji 180 muških jedinki te 4629 ženskih jedinki, efektivna veličina populacije iznosi 693, a prema statusu ova ovca je ugrožena. Pripada skupini ovaca kombiniranih proizvodnih odlika, no u nas je prvenstveno namijenjena proizvodnji mlijeka, odnosno preradom do proizvoda poput ovčjeg sira (Hrvatski savez uzgajivača ovaca i koza, 2021.).

1.4.2. Paška ovca

Paška ovca nastala je na području istoimenog otoka gdje se i danas uzgaja. Nastala je oplemenjivanjem autohtone pramenke s ovnovima pasmine *negretti* i *bergamo* te *gentile di puglia*. Radi se o ovci čvrste i zbijene građe te širokog trupa sa jakim nogama (Pavlinić, 1936.) čija visina u ovaca iznosi 56 cm s težinom oko 30 kilograma, dok se kod ovnova mjeri visina do 63 centimetra te težina oko 50 kg (Agroportal, 2019.). Tijelo ove ovce prekriveno je zatvorenim do poluzatvorenim runom, najčešće bijele vune. Ovce su šute (bez rogova), a ovnovi imaju dobro razvijene i čvrste rogove (Agroportal, 2019.). Također, kao i istarska ovca, nije pripadnik kasnozrelih pasmina (Pavić i sur., 2005.). Navodi se efektivna veličina populacije koja iznosi 266, 70 muških jedinki te 1344 ženske jedinke kao dio matične populacije što je svrstava na status ugrožene kod autohtonih pasmina ovaca Hrvatske (Hrvatske izvorne i zaštićene pasmine – Banka gena domaćih životinja, 2019.). Njena korist, prvenstveno, je vidljiva u proizvodnji mlijeka, odnosno sira. Osim toga, dobiva se i vrlo tražena janjad izvrsne kakvoće – zahvaljujući spomenutim klimatskim uvjetima mediterana, kao i vegetacijom kojim je okružena (Mioč i sur., 1999.). Uzgoj paške ovce predstavlja način upravljanja ekstenzivnim pašnjacima, posebno važnim za održavanje flore otoka Paga.

2. HIPOTEZE I OPĆI I SPECIFIČNI CILJEVI RADA

Cilj ovog rada je koristeći podatke rodovnika istarske i paške ovce i genomske informacije dobivene iz SNP čipa srednje rezolucije (Illumina OvineSNP50 BeadChip (Illumina Inc., San Diego, CA) utvrditi razlike u procjeni inbridinga X spolnog kromosoma putem različitih metoda. Nadalje, želimo utvrditi je li moguće upotpuniti procjenu inbriding koeficijenta ženskih jedinki s informacijom sa spolnog kromosoma i tako ne zanemariti opasnost te pojave za spolno vezana svojstva i općenito za ženske jedinke.

2.1. HIPOTEZE

- 1. Ne postoji statistički značajna razlika između mjera inbridinga procijenjenih putem rodovnika i putem genomske informacije.*
- 2. Ne postoji statistički značajna razlika između inbriding koeficijenata procijenjenih na autosomu i na X spolnom kromosomu.*

2.2. SPECIFIČNI CILJEVI RADA

- 1. Obraditi i provjeriti dobivene podatke (rodovnik i genotip), i prilagoditi ih programskim paketima za procjenu inbriding koeficijenata.*
- 2. Procijeniti inbriding koeficijente iz rodovnika i genotipskih podatka (ukupni i X kromosom).*
- 3. Usporediti dobivene koeficijente inbridinga (ukupni i X kromosom).*

3. MATERIJAL I METODE

3.1. UZORCI, GENOTIPIZACIJA I KONTROLA KVALITETE GENOTIPOVA

Za potrebe ovog rada podatci o 175 ovaca (96 jedinki istarske i 76 paške ovce) preuzeti su iz Spehar i sur. (2019.). Jedinke su genotipizirane s Illumina OvineSNP50 BeadChip (Illumina Inc., San Diego, CA) koji sadrži 54241 SNP prosječne gustoće 1 SNP na svakih 50,9 kilobaza. Putem SAS 9.4 (SAS Institute Inc., 2013.) programskog paketa, prvo smo isključili sve SNP-ove koji nisu pripisani niti jednom kromosomu i one koji su pripisani Y kromosomu ili mitohondrijskoj DNK. U sljedećem koraku uklonili smo SNP-ove koji nisu udovoljavali mjerama kvalitete „GenTrain Score“ i „GenCall Score“ čija visoka vrijednost osigurava pouzdanost svakog genotipa. Uklonjeni su SNP-ovi čiji je „GenTrain Score“ bio manji ili jednak 0,4 i SNP-ovi čiji je „GenCall Score“ manji ili jednak 0,8. Tako očišćeni i pripremljeni podatci potom su uvedeni u programski paket SNP & Variation Suite v8.7.0 (Golden Helix, Inc., Bozeman, MT, www.goldenhelix.com) gdje su iz daljnje analize isključeni svi SNP-ovi koji nedostaju u više od 10 % jedinki i jedinke kojima nedostaje više od 5 % SNP-ova, te sve muške jedinke. Konačno, podatci preostalih 110 jedinki (68 jedinki istarske i 42 jedinke paške ovce) podijeljeni su u dva seta podataka, onaj koji sadrži SNP-ove s 26 autosoma (43268 SNP-ova) i onaj s X kromosoma (1199 SNP-ova).

3.2. RODOVNICI I PROCJENA KOEFICIJENTA INBRIDINGA IZ RODOVNIKA

Rodovnike ovih dviju pasmina ustupio je Centar za stočarstvo Hrvatske agencije za poljoprivredu i hranu. U rodovniku istarske ovce nalazimo 813, u rodovniku paške 660 jedinki. Oba rodovnika su rekodirana programskim paketom CFC (Sargolzaei i sur., 2006.) te potom pripremljena za programski paket ENDOG v 4.8 (Gutiérrez i Goyache, 2005.). Putem ovog programskog paketa izračunati su koeficijent inbridinga F_{PED} , te parametar kvalitete i

potpunosti rodovnika – ekvivalent kompletne generacije (engl. *Equivalent Complete Generations*) - ECG). ECG se izračunava tako da se izraz $(1/2)^n$ primjeni na svakog pretka životinje u pri čemu je n broj generacija, a potom se dobivene vrijednosti za pretke zbroje. Na primjer ako jedinka ima samo poznate roditelje onda za svakog roditelja $(1/2)^1$ iznosi 0,5 te je ECG u tom slučaju 1, kod poznatih roditelja i baka $((1/2)^2)$ i djedova $((1/2)^2)$ ECG iznosi 2, ako su bake i djedovi poznati samo s jedne strane ECG je 1,5 itd. (Maignel i sur., 1996.). Sam koeficijent inbridinga može biti beznačajna informacija ukoliko ne znamo od koliko predaka dolazi, a upravo nam ova vrijednost to govori.

Putem programskog paketa KinInbcoefX (Zheng i Bourgain, 2009.) su iz rodovnika izračunati koeficijent inbridinga X-kromosoma za ženske jedinke (F_{PED_X}).

3.3. PROCJENA INBRIDINGA IZ GENOTIPSKIH PODATAKA

Programskim paketom SNP & Variation Suite v8.7.0 (Golden Helix, Inc., Bozeman, MT, www.goldenhelix.com) procijenjen su koeficijenti inbridinga za autosome (F_{AUTO}) i posebno za spolni kromosom (F_X). Korišten je ekvivalent Wright-ovog F_{IS} koeficijenta koji prikazuje promjenu heterozigotnosti jedinke zbog ne slučajnog parenja unutar subpopulacije, a u navedenom programskom paketu se definira kao inbriding koeficijent jedinke temeljen na subpopulaciji koje je ona dio pri čemu u intervalu od -1 do 1 pozitivne vrijednosti upućuju na uzgoj u srodstvu (Golden Helix, Inc., Bozeman, MT, www.goldenhelix.com).

Također su izračunati i koeficijenti inbridinga temeljeni na ROH segmentima (F_{ROH}). Korištene duljine ROH segmenata bile su >4 Mb, > 8 Mb i >16 Mb, što je određeno prema gustoći SNP čipa (Ferenčaković i sur., 2013b). ROH segment je definiran kao niz od najmanje 15 SNP-ova, koji nisu udaljeni više od 1Mb niti su na području genoma gdje je gustoća SNP-ova manja od 1SNP svakih 100kb. Prilikom detekcije ROH segmenta za kategoriju između 4 i 8 Mb dopustili smo da nam nedostaje jedan SNP, između 8 i 16 dva SNP-a, dok smo za

kategoriju segmenata većih od 16 Mb dopustili da nam nedostaju četiri SNP-a, ali i da se u segmentu nalazi jedan heterozigotni SNP (Ferenčaković i sur., 2013.b). Sumiranjem detektiranih ROH segmenata po autosomima i spolnom kromosomu X te njihovim dijeljenjem s ukupnom dužinom genoma pokrivenog SNP-ovima, odnosno X kromosoma dobili smo koeficijente $F_{ROH>4Mb}$, $F_{ROH>8Mb}$ i $F_{ROH>16Mb}$ za autosome te $F_{ROH_X>4Mb}$, $F_{ROH_X>8Mb}$ i $F_{ROH_X>16Mb}$ za spolni kromosom X (McQuillan i sur., 2008.; Yengo i sur., 2019.).

3.4. STATISTIČKA OBRADA I VIZUALIZACIJA

Statistička obrada podataka izvršena je programskim paketom R (R Core Team,(2020.). Uz osnovne R programske pakete, za unos podataka, odnosno ispis rezultata korišteni su paketi *readxl* (Wickham i Brayan, 2019.) i *writexl* (Ooms, 2021.). Za deskriptivnu statistiku korišten je i paket *pastecs* (Grosjean i Ibanez, 2018.). Vizualizacija podataka obavljena je u programskom paketu JMP® Pro 15.2.0 (SAS Institute Inc., 2020.). Razlike u srednjim vrijednostima grupa ispitane su t-testom, jakost linerane veze između različitih koeficijenata ispitana je Pearsonovom korelacijom. U oba slučaja statistička značajnost se utvrđivala ako je $p < 0,05$.

4. REZULTATI

4.1. ANALIZA RODOVNIKA I PROCJENA INBRIDING KOEFICIJENATA NA TEMELJU RODOVNIKA

Deskriptivna statistika prosječnog koeficijenta inbridinga (F_{PED}) i koeficijenta inbridinga X kromosoma F_{PED_X} genotipiziranih ženskih jedinki populaciji istarske i paške ovce prikazana je u Tablici 4-1. Prosječni ekvivalent kompletne generacije genotipiziranih ženskih jedinki istarske ovce iznosio je 3,9192 (0,0000 – 6,7402), a paške ovce 2,3127 (0,0000 – 4,5313).

Tablica 4-1. Deskriptivna statistika vrijednosti koeficijenta inbridinga na temelju rodovnika u populaciji istarske i paške ovce.

Istarska ovca				
	Prosjek	Minimum	Maksimum	Standardna devijacija
F_{PED}	0,0624	0,0000	0,3491	0,0950
F_{PED_X}	0,0890	0,0000	0,6563	0,1683
Paška ovca				
F_{PED}	0,0168	0,0000	0,1250	0,0335
F_{PED_X}	0,0149	0,0000	0,2500	0,0494

4.2. ANALIZA GENOMSKIH PODATAKA I PROCJENA INBRIDING KOEFICIJENATA NA TEMELJU GENOMSKIH PODATAKA

4.2.1. ROH segmenti i F_{ROH} inbriding koeficijenti

Deskriptivna statistika ROH segmenata identificiranih kod 68, odnosno 42 ženske jedinke istarske i paške ovaca na autosomima (ROH) i spolnom kromosomu X (ROH_X) prikazane su u Tablici 4-2. i 4-3. Korištene dužine segmenta >4Mb, >8Mb i >16Mb odgovaraju inbridingu koji potječe od zajedničkog pretka udaljenog približno 12,5, 6,75, i 3,125 generacija.

Tablica 4-2. Deskriptivna statistika parametara detektiranih ROH segmenata kod istarske i paške ovce na autosomima (ROH). Parametri su prosječan broj SNP-ova unutar ROH-a (nSNP), prosječan broj ROH segmenata (nROH) te prosječna ukupna dužina genoma pokrivena ROH segmentima u Mb (totROH).

		Istarska ovca		Paška ovca	
		Prosjek (SD ¹)	Min ² (Max ³)	Prosjek (SD)	Min (Max)
ROH >4Mb	nSNP	2792,72 (2868,46)	0,00 (15551,00)	1803,08 (2452,45)	0,00 (8159,00)
	nROH	13,85 (10,71)	0,00 (55,00)	8,32 (9,44)	0,00 (37,00)
	totROH	167,14 (172,04)	0,00 (929,80)	107,9 (146,72)	0,00 (492,86)
ROH >8Mb	nSNP	2177,57 (2559,41)	0,00 (13628,00)	1467,63 (2201,09)	0,00 (6957,00)
	nROH	7,34 (7,16)	0,00 (35,00)	4,71 (6,46)	0,00 (23,00)
	totROH	130,57 (153,78)	0,00 (816,83)	87,37 (131,29)	0,00 (415,65)
ROH >16Mb	nSNP	1339,62 (1914,08)	0,00 (10082,00)	979,74 (1644,94)	0,00 (5150,00)
	nROH	2,87 (3,69)	0,00 (18,00)	2,08 (3,22)	0,00 (10,00)
	totROH	80,05 (114,96)	0,00 (603,73)	58,38 (98,1)	0,00 (303,92)

¹ Standardna devijacija; ² Minimum; ³ Maksimum.

Tablica 4-3. Deskriptivna statistika parametara detektiranih ROH segmenata kod istarske i paške ovce na spolnom kromosomu X (ROH_X). Parametri su prosječan broj SNP-ova unutar ROH-a (nSNP), prosječan broj ROH segmenata (nROH) te prosječna ukupna dužina genoma pokrivena ROH segmentima u Mb (totROH).

		Istarska ovca		Paška ovca	
		Prosjek (SD ¹)	Min ² (Max ³)	Prosjek (SD)	Min (Max)
ROH_X >4Mb	nSNP	176,73 (163,33)	0,00 (545,00)	82,4 (36,24)	0,00 (117,00)
	nROH	1,8 (1,26)	0,00 (5,00)	1,2 (0,45)	0,00 (2,00)
	totROH	15,08 (13,7)	0,00 (45,04)	7,98 (3,63)	0,00 (12,71)
ROH_X >8Mb	nSNP	97,27 (139,15)	0,00 (381,00)	57,6 (53,65)	0,00 (110,00)
	nROH	0,6 (0,83)	0,00 (2,00)	0,6 (0,55)	0,00 (1,00)
	totROH	8,2 (11,62)	0,00 (32,38)	5,42 (4,96)	9,59
ROH_X >16Mb	nSNP	55,93 (96,78)	0,00 (248,00)	0,00 (0,00)	0,00 (0,00)
	nROH	0,27 (0,46)	0,00 (1,00)	0,00 (0,00)	0,00 (0,00)
	totROH	4,8 (8,27)	0,00 (20,51)	0,00 (0,00)	0,00 (0,00)

¹ Standardna devijacija; ² Minimum; ³ Maksimum.

U Tablici 4-4. rezultati su deskriptivne statistike inbriding koeficijenata procijenjenih na temelju detektiranih ROH segmenata autosoma (F_{ROH}) i ROH segmenata X kromosoma (F_{ROH_X}).

Tablica 4-4. Deskriptivna statistika inbriding koeficijenata procijenjenih na temelju detektiranih ROH segmenata autosoma (F_{ROH}) i ROH segmenata X kromosoma (F_{ROH_X}) genotipiziranih ženskih jedinki istarske i paške ovce.

	Istarska ovca		Paška ovca	
	Prosjek (SD ¹)	Min ² (Max ³)	Prosjek (SD)	Min (Max)
$F_{ROH >4Mb}$	0,0632 (0,0651)	0,0000 (0,3516)	0,0369 (0,0541)	0,0000 (0,1864)
$F_{ROH >8Mb}$	0,0494 (0,0581)	0,0000 (0,3089)	0,0299 (0,0482)	0,0000 (0,1572)
$F_{ROH >16Mb}$	0,0303 (0,0435)	0,0000 (0,2283)	0,0200 (0,0358)	0,0000 (0,1149)
$F_{ROH_X >4Mb}$	0,0260 (0,0694)	0,0000 (0,3520)	0,0074 (0,0223)	0,0000 (0,0993)
$F_{ROH_X >8Mb}$	0,0142 (0,0494)	0,0000 (0,2531)	0,0050 (0,0184)	0,0000 (0,0750)
$F_{ROH_X >16Mb}$	0,0083 (0,0335)	0,0000 (0,1603)	0,0000 (0,0000)	0,0000 (0,0000)

¹ Standardna devijacija; ² Minimum; ³ Maksimum.

4.2.2. Wright-ovi F_{IS} koeficijenti

U radu su procijenjeni i ekvivalenti Wright-ovog F_{IS} koeficijenta za autosome (F) i za X kromosom (F_X) za genotipizirane ženske jedinke iz obje populacije (istarska i paška ovca). Deskriptivna statistika ovih koeficijenata prikazana je u Tablici 4-5.

Tablica 4-5. Deskriptivna statistika ekvivalenta Wright-ovog F_{IS} koeficijenta za autosome (F) i za X kromosom (F_X) za genotipizirane ženske jedinke istarske i paške ovce.

Istarska ovca				
	Prosjek	Minimum	Maksimum	Standardna devijacija
F	0,0317	-0,0572	0,3357	0,0700
F_X	0,0402	-0,1942	0,7017	0,1500
Paška ovca				
F	0,0359	-0,0299	0,1888	0,0601
F_X	0,0238	-0,1660	0,2202	0,0871

„Box-plotovi“ svih prikazanih mjera genotipiziranih ženki istarske i paške ovce prikazane su grafički u Figurama 4-1. i 4-2., pri čemu su oni koji se odnose na X kromosom neosjenčani.

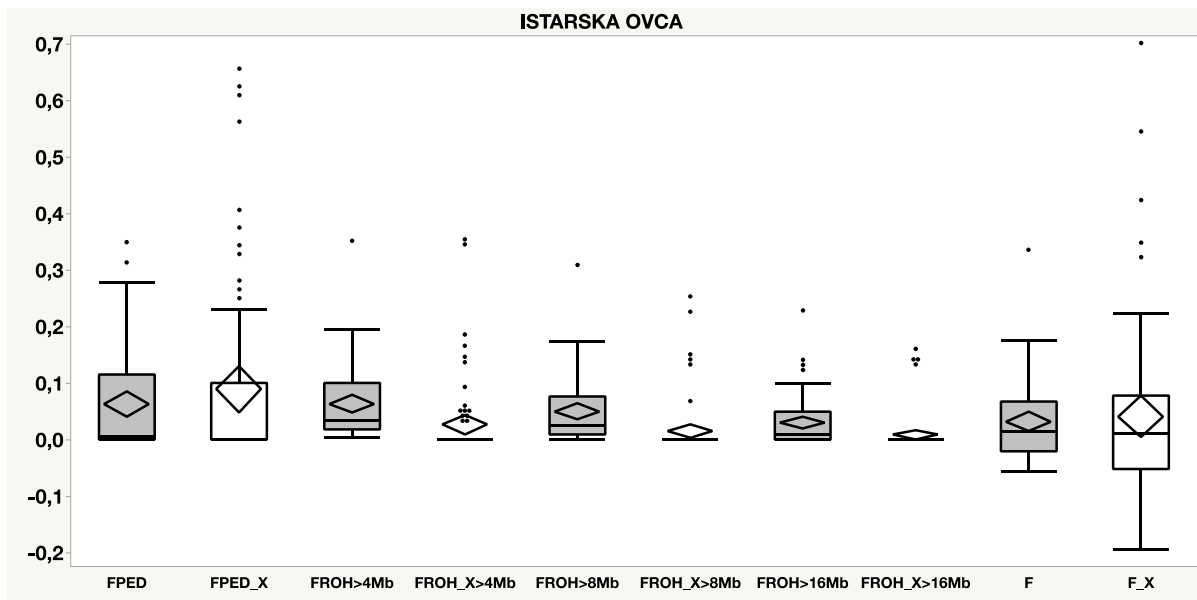


Figura 4-1. „Box-plotovi“ procijenjenih inbriding koeficijenata genotipiziranih ženskih jedinki istarske ovce. Neosjenčani „box-plotovi“ odnose se na mjere dobivene preko X kromosoma.

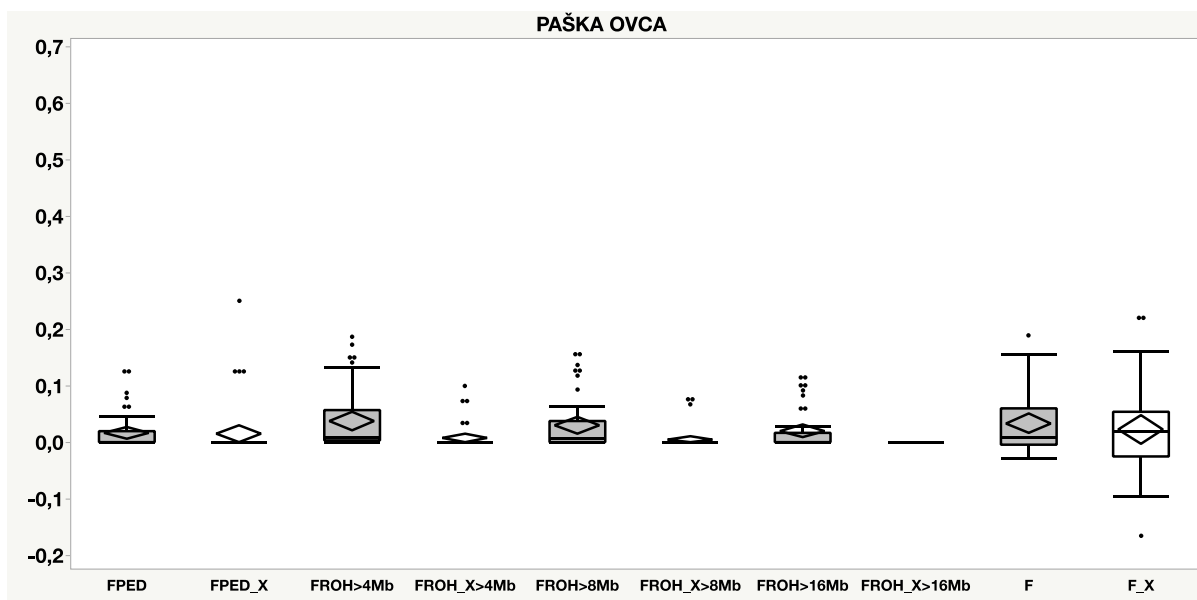


Figura 4-2. „Box-plotovi“ procijenjenih inbriding koeficijenata genotipiziranih ženskih jedinki paške ovce. Neosjenčani „box-plotovi“ odnose se na mjere dobivene preko X kromosoma.

4.3. USPOREDBA KOEFICIJENATA INBRIDINGA

4.3.1. Koeficijenti izračunati preko rodovnika i genomski inbriding koeficijenti

U Tablici 4-6. nalaze se rezultati t-testa i koeficijenti korelacije r dobiveni između između autosomalnih koeficijenata inbridinga procijenjenih na temelju rodovnika (F_{PED}), te onih procijenjenih iz genoma, a to su ekvivalent Wright-ovog F_{IS} koeficijenta (F) i tri genomska koeficijenta dobivena iz ROH segmenata različite duljine ($F_{ROH>4Mb}$, $F_{ROH>8Mb}$, $F_{ROH>16Mb}$) Koeficijenti su procijenjeni samo za genotipizirane ženske jedinke istarske i paške ovce.

4.3.2. Koeficijenti X kromosoma izračunati preko rodovnika i genomski inbriding koeficijenti X kromosoma

Usporedbe različitih mjera koeficijenta inbridinga (t-test i koeficijenti korelacije r) X kromosoma dobivenih putem rodovnika i genomske informacije nalaze se u Tablici 4-7. U pitanju su isti koeficijenti kao i u prethodnom odlomku, ali razliku čini što se oni odnose samo na X kromosom.

4.3.3. Koeficijenti inbridinga procijenjeni iz autosoma i iz X kromosoma

U nastavku su uspoređeni koeficijenti inbridinga X kromosoma i koeficijenti inbridinga autosoma. Rezultati ove usporedbe nalaze se u Tablici 4-8.

Tablica 4-6. Koeficijent korelacije r (ispod dijagonale) i razlike srednjih vrijednosti grupa ($\mu_x - \mu_y$; iznad dijagonale) između autosomalnih koeficijenata inbridinga procijenjenih na temelju rodovnika (F_{PED}), ekvivalenta Wright-ovog F_{IS} koeficijenta (F) te genomskog koeficijenta inbridinga procijenjenog iz ROH segmenata različite duljine ($F_{ROH>4Mb}$, $F_{ROH>8Mb}$, $F_{ROH>16Mb}$) genotipiziranih ženskih jedinki istarske i paške ovce.

$r \setminus \mu_x - \mu_y$	Istarska ovca					$r \setminus \mu_x - \mu_y$	Paška ovca				
	F	F_{PED}	$F_{ROH>4Mb}$	$F_{ROH>8Mb}$	$F_{ROH>16Mb}$		F	F_{PED}	$F_{ROH>4Mb}$	$F_{ROH>8Mb}$	$F_{ROH>16Mb}$
F	1\0	0,03*	0,032**	0,018	0,001	F	1\0	0,016	0,004	0,004	0,013
F_{PED}	0,37**	1\0	0,001	0,013	0,03*	F_{PED}	0,12	1\0	0,02*	0,013	0,003*
$F_{ROH>4Mb}$	0,97***	0,33**	1\0	-	-	$F_{ROH>4Mb}$	0,98***	0,13	1\0	-	-
$F_{ROH>8Mb}$	0,96***	0,32**	-	1\0	-	$F_{ROH>8Mb}$	0,96	0,12	-	1\0	-
$F_{ROH>16Mb}$	0,92***	0,26*	-	-	1\0	$F_{ROH>16Mb}$	0,93***	0,09	-	-	1\0

* p vrijednost između 0,05 i 0,01,

** p vrijednost između 0,01 i 0,001,

*** p vrijednost < 0,001.

Tablica 4-7. Koeficijent korelacije r (ispod dijagonale) i razlike srednjih vrijednosti grupa ($\mu_x - \mu_y$; iznad dijagonale) između X kromosomalnih koeficijenata inbridinga procijenjenih na temelju rodovnika (F_{PED}), ekvivalenta Wright-ovog F_{IS} koeficijenta (F) te genomskog koeficijenta inbridinga procijenjenog iz ROH segmenata različite duljine ($F_{ROH>4Mb}$, $F_{ROH>8Mb}$, $F_{ROH>16Mb}$) genotipiziranih ženskih jedinki istarske i paške ovce.

$r \setminus \mu_x - \mu_y$	Istarska ovca					$r \setminus \mu_x - \mu_y$	Paška ovca				
	F_X	F_{PED_X}	$F_{ROH_X >4Mb}$	$F_{ROH_X >8Mb}$	$F_{ROH_X >16Mb}$		F	F_X	F_{PED_X}	$F_{ROH_X >4Mb}$	$F_{ROH_X >8Mb}$
F_X	1\0	0,05	0,01	0,01	0,03	F_X	1\0	0	0,01	0,01	0,02
F_{PED_X}	0,29*	1\0	0,06*	0,08***	0,08***	F_{PED_X}	-0,12	1\0	0	0,01	0,02
$F_{ROH_X >4Mb}$	0,82***	0,16	1\0	-	-	$F_{ROH_X >4Mb}$	0,55**	-0,10	1\0	-	-
$F_{ROH_X >8Mb}$	0,71***	0,09	-	1\0	-	$F_{ROH_X >8Mb}$	0,5***	-0,08	-	1\0	-
$F_{ROH_X >16Mb}$	0,55***	0,07	-	-	1\0	$F_{ROH_X >16Mb}$	-	-	-	-	1\0

* p vrijednost između 0,05 i 0,01,

** p vrijednost između 0,01 i 0,001,

*** p vrijednost < 0,001.

Tablica 4-8. Koeficijent korelacije r i razlike srednjih vrijednosti grupa ($\mu_x - \mu_y$) između X kromosomalnih i autosomalnih koeficijenata inbridinga procijenjenih na temelju rodovnika (F_{PED} i F_{PED_X}), ekvivalenta Wright-ovog F_{IS} koeficijenta (F i F_X) te genomskih koeficijenata inbridinga procijenjenih iz ROH segmenata različite duljine ($F_{ROH_X>4Mb}$, $F_{ROH_X>8Mb}$, $F_{ROH_X>16Mb}$ i $F_{ROH>4Mb}$, $F_{ROH>8Mb}$, $F_{ROH>16Mb}$) genotipiziranih ženskih jedinki istarske i paške ovce.

Istarska ovca				Paška ovca			
Usporedba		$\mu_x - \mu_y$	r	Usporedba		$\mu_x - \mu_y$	r
F	F_X	0,009	0,49***	F	F_X	0,011	0,33*
F_{PED}	F_{PED_X}	0,027	0,84***	F_{PED}	F_{PED_X}	0,002	0,59***
F_{ROH>4Mb}	F_{ROH_X>4Mb}	0,037	0,28*	F_{ROH>4Mb}	F_{ROH_X>4Mb}	0,03**	0,06
F_{ROH>8Mb}	F_{ROH_X>8Mb}	0,035***	0,12	F_{ROH>8Mb}	F_{ROH_X>8Mb}	0,025**	0,11
F_{ROH>16Mb}	F_{ROH_X>16Mb}	0,022**	0,07	F_{ROH>16Mb}	F_{ROH_X>16Mb}	0,02***	-

* p vrijednost između 0,05 i 0,01,

** p vrijednost između 0,01 i 0,001,

*** p vrijednost < 0,001.

5. RASPRAVA

Cilj ovog rada bio je usporediti razlike u procjeni inbriding koeficijenta kada se koristi X kromosom ili klasične metode (rodovnik i genomska informacija s autosoma). S ekonomskog stajališta i stajališta konzervacijske genetike važno je uključiti informaciju sa spolnog kromosoma u procjenu inbridinga. Nadalje, niža stopa rekombinacija (0,8 cM/Mb), stopa mutacija i efektivne veličine populacije ($3/4$) u odnosu na autosome općenito čine X kromosom podložnijim inbridingu, a time i svim negativnim posljedicama (Schaffner, 2004.; Robinson i sur., 2014.). Kako bismo mogli ispitati naše nul-hipoteze, kojima smo definirali ne postojanje razlika, prije svega smo morali prikupiti podatke genotipa i rodovnika ženskih jedinki. Obzirom na to da je zastupljenost genotipizacije ženskih jedinki i dalje prilično niska, takvi se podatci teško nalaze. Ipak preuzeli smo iz Spehar i sur. (2019.), unutar kojih je bilo 110 genotipiziranih ženskih jedinki (68 jedinki istarske ovce i 42 jedinke paške ovce) za koje smo imali i rodovnik. Jedinke su genotipizirane SNP čipom srednje rezolucije (Illumina OvineSNP50 BeadChip (Illumina Inc., San Diego, CA), za koji smo postavili stroge kriterije kontrole kvalitete („GenTrain Score“ $\geq 0,4$ i „GenCall Score“ $\geq 0,8$) kako bismo smanjili mogućnost pogreške u utvrđivanju stanja SNP-ova (homozigot ili heterozigot). Ova greška može uvelike utjecati na preciznost procjene ROH segmenata koji svoju prisutnost u genomu duguju upravo inbridingu (Ferenčaković i sur., 2013b.; Curik i sur., 2014.).

Analizom rodovnika utvrdili smo kako naše genotipizirane jedinke u prosjeku imaju prilično „plitak“ rodovnik, odnosno kako je informacija o njihovim precima nedostatna. To smo utvrdili putem ekvivalenta kompletnih generacija – ECG koji prosječno za genotipizirane ženke istarske ovce iznosi 3,9192 (0,0000 – 6,7402), a za paške ovce 2,3127 (0,0000 – 4,5313).

Uobičajeno u dubljim rodovnicima možemo vidjeti veće vrijednosti (Ferenčaković i sur., 2013.) što omogućava bolje korelacije sa inbriding koeficijentima iz kraćih ROH segmanata.

Koeficijenti inbridinga na autosomima i spolnom kromosomu X analizirani su različitim metodama kod istarske i paške ovce, kako bi se utvrdile razlike između metoda i populacija te ispitala mogućnost uključivanja informacija sa spolnog kromosoma X u procjenu koeficijenta inbridinga. Usporedbom koeficijenata inbridinga procijenjenih na temelju rodovnika (F_{PED} , F_{PED_X}) i koeficijenata inbridinga procijenjenih različitim metodama na temelju genomske informacije (F_{ROH} , F , F_{ROH_X} , F_X), također je ispitana i valjanost dobivenih podataka (rodovnik i genotip). ROH segmenti sve više postaju glavni alat za procjenu koeficijenta inbridinga, no ne postoji jasan standard kako ih procijeniti (Ferenčaković i sur., 2013b.; Curik i sur., 2014.), kao ni standardna metoda (Solé i sur., 2017.) Kod odluke o metodi detekcije oslonili smo se na rad Curik i sur. (2014.) gdje su opisane razlike između dvije metode prebrajanja i Solé i sur. (2017.) koji dokazuju kako su koeficijenti dobiveni složenim HMM algoritmima u visokoj korelaciji s metodama prebrojavanja koju koriste Ferenčaković i sur. (2013b.).

U populaciji istarske ovce, analizom rodovnika, prosječni koeficijent inbridinga na autosomima (F_{PED}) iznosi 0,0624, a prosječni koeficijent inbridinga X kromosoma iznosi 0,089 (kod 63 % jedinki je jednak 0). Na temelju markera na autosomima kod istarske ovce, koeficijenti inbridinga prosječno iznose: prema $F_{ROH>4Mb}$ 0,0632, prema $F_{ROH>8Mb}$ 0,0494 (6 jedinki; $F_{ROH>8Mb} = 0$) te prema $F_{ROH>16Mb}$ 0,0303 (22 jedinke; $F_{ROH>16Mb} = 0$), uz najviše vrijednosti od 0,3515, 0,3089 i 0,2283. Nasuprot tome, prosječni koeficijenti inbridinga na temelju markera na spolnom kromosomu X iznose: $F_{ROH>4Mb}$ 0,0369 (53 jedinke; $F_{ROH>4Mb} = 0$), $F_{ROH>8Mb}$ 0,0142 (62 jedinke; $F_{ROH>8Mb} = 0$) i $F_{ROH>16Mb}$ 0,0083 (64 jedinke; $F_{ROH>16Mb} = 0$), uz najviše vrijednosti koje iznose redom: 0,3520, 0,2531 i 0,1603. Kako bi se povećala točnost procjene, korištena je i druga molekularna metoda (Wright-ova F statistika) te F iznosi 0,0317,

a F_X 0,0402. Prema rodovniku, F_{PED_X} u odnosu na F_{PED} je 1,4 puta viši, a prema genomskim informacijama F_X odnosu na F je 1,27 puta viši, dok su $F_{ROH>4Mb}$, $F_{ROH>8Mb}$ i $F_{ROH>16Mb}$ u odnosu na $F_{ROH_X>4Mb}$, $F_{ROH_X>8Mb}$ i $F_{ROH_X>16Mb}$ niži za 1,71, 3,45 i 3,65 puta. Vizualizacija raspodjele pojedinih koeficijenata inbridinga prikazana je na Figuri 4-1. Zabilježene su korelacije i razlike u prosjecima (t-test) od 0,84 i 0,027 između F_{PED_X} i F_{PED} , od 0,49 i 0,009 između F_X i F , od 0,28 i 0,037 između $F_{ROH_X>4Mb}$ i $F_{ROH>4Mb}$, od 0,12 i 0,035 između $F_{ROH_X>8Mb}$ i $F_{ROH>8Mb}$ te od 0,07 i 0,022 između $F_{ROH_X>16Mb}$ i $F_{ROH>16Mb}$. Utvrđena je značajna korelacija između F_X i F , F_{PED_X} i F_{PED} te $F_{ROH_X>4Mb}$ i $F_{ROH>4Mb}$. Uspoređujući rodovnik s molekularnim metodama, na autosomima sve korelacije su bile značajne uz najvišu korelaciju između F i F_{PED} (0,37), a značajna razlika u prosjeku zabilježena je između F i F_{PED} (0,03), F i $F_{ROH_X>16Mb}$ (0,03) te F i $F_{ROH_X>4Mb}$ (0,32). Nasuprot, na spolnom kromosomu X utvrđena je značajna korelacija između F_X i F_{PED_X} (0,29) i značajna razlika u prosjeku između F_{PED} i F_{ROH} koeficijenata inbridinga. Prema tome, uspoređujući rodovnik i ROH metodu, najmanja razlika je uočena između F_{PED} i $F_{ROH>16Mb}$, što ukazuje na nisku („plitku“) kompletност rodovnika. To u potpunosti odgovara opaženom prosječnom ECG-u od 3,9192 jer ROH segmenti > 16 Mb predstavljaju nedavni inbriding star najviše 3 generacije (Ferenčaković i sur., 2013a.). Upotrebom genomske informacije nadopunjuju se nedostaci rodovnika, pogotovo u slučaju „plitke“ kompletnosti.

U populaciji paške ovce, prema rodovniku na autosomima prosječni F_{PED} iznosi 0,0168, sa najvećom vrijednosti od 0,125. Nadalje, prosječni koeficijent inbridinga X kromosoma prema rodovniku (F_{PED_X}) je 0,0149 sa najvećom vrijednosti od 0,25. Deskriptivnom statistikom također je utvrđeno da je 70 % jedinki imalo vrijednost nula kada govorimo o F_{PED} -u, odnosno 90 % kada govorimo o F_{PED_X} -u. U odnosu na rodovnik, koeficijenti inbridinga na temelju markera na autosomima kod paške ovce prosječno iznose: $F_{ROH>4Mb}$ 0,0369 (4 jedinke; $F_{ROH>4Mb} = 0$), prema $F_{ROH>8Mb}$ 0,0299 (14 jedinke; $F_{ROH>8Mb} = 0$) te prema $F_{ROH>16Mb}$ 0,0303

(23 jedinke; $F_{ROH>16Mb} = 0$), uz najviše vrijednosti 0,1864, 0,1572 i 0,1149. Suprotno, koeficijenti inbridinga na spolnom kromosomu X iznose: prema $F_{ROH>4Mb}$ 0,0074 (37 jedinki; $F_{ROH>4Mb} = 0$), $F_{ROH>8Mb}$ 0,0050 (39 jedinki; $F_{ROH>8Mb} = 0$) te $F_{ROH>16Mb} = 0$, uz najviše vrijednosti 0,0993 i 0,075. Prema Wright-ovoj F statistici F iznosi 0,0359, a F_X 0,0238. Temeljem rodovnika, F_{PED_X} u odnosu na F_{PED} je 1,13 puta niži, a prema genomskim informacijama F_X odnosu na F je 1,51 puta niži, dok su $F_{ROH>4Mb}$, $F_{ROH>8Mb}$ i $F_{ROH>16Mb}$ u odnosu na $F_{ROH_X>4Mb}$, $F_{ROH_X>8Mb}$ i $F_{ROH_X>16Mb}$ viši za 5 – 6 puta. Utvrđena je statistički značajna korelacija između F i F_X (0,33) te F_{PED} i F_{PED_X} (0,59) i značajna razlika u prosjecima između F_{ROH} koeficijenata inbridinga. Stoga, uspoređujući rodovnik i ROH metodu, najmanja razlika je uočena između F_{PED} i $F_{ROH>16Mb}$, što također ukazuje na nisku („plitku“) kompletnost rodovnika.

Na temelju svega navedenog možemo odbaciti naše nul hipoteze te reći kako postoji statistički značajna razlika između mjera inbridinga procijenjenih putem rodovnika i putem genomske informacije kao i statistički značajna razlika između inbriding koeficijenata procijenjenih na autosomu i na X spolnom kromosomu. Unatoč tome što postoji razlika, vidimo da se neki koeficijenti slično ponašaju (postoji korelacija). Ipak, dobivene rezultate treba promatrati s oprezom. „Plitak“ rodovnik, odnosno nizak ECG nam dopušta da koeficijente dobivene preko rodovnika smatramo pouzdanim. Genomske metode ovdje su pouzdanije, ali također nailazimo na problem jer kod detekcije ROH segmenata pretpostavljamo rekombinaciju 1 cM / 1 Mb, a znamo da je rekombinacija X kromosoma oko 0,8 cM / 1 Mb, stoga krivo procjenjujemo starost segmenta, a time i starost inbridinga.

Kako bi se provela preciznija i pouzdanija analiza prvenstveno je potrebno simulirati populaciju s „dubokim“ rodovnikom i visokom gustoćom markera. Iako postoje programski paketi koji to uspješno izvode, glavni je izazov simulirati rekombinaciju X kromosoma, budući da ona generalno postoji samo kod ženskih jedinki. Simulacija koja bi uspješno riješila ovaj

problem svakako bi razjasnila obrazac rekombinacije i nastajanje ROH segmenata na X kromosomu te se nadamo da će ovaj rad biti dobra polazna točka za buduće napore u procjeni inbridinga X kromosoma.

6. ZAKLJUČCI

1. Nakon kontrole kvalitete te izbacivanja muških jedinki, podatci su se sastojali od 68 jedinki istarske ovce i 42 jedinke paške ovce te 43268 SNP-ova na autosomima i 1199 SNP-ova na spolnom kromosomu X.
2. Koeficijenti inbridinga procijenjeni iz ROH segmenata su statistički značajno veći od koeficijenata inbridinga na temelju rodovnika, u svim slučajevima.
3. Utvrđena je loša kompletnost rodovnika kod obje populacije na što ukazuje ECG od 3,9192 kod istarske i 2,3127 kod paške ovce.
4. Usporedbom koeficijenata inbridinga procijenjenih metodama temeljenim na genomskoj informaciji utvrđene su visoke korelacije te mala razlika u prosjeku, što ide u prilog njihovoj točnosti.
5. Potrebno je simulirati populacije s „dubokim” rodovnikom i visokom gustoćom markera, uzimajući u obzir specifičnost rekombinacije X kromosoma kako bi se došlo do pouzdanijih zaključaka.

7. ZAHVALE

Zahvaljujemo se:

Doc. dr. sc. Maji Ferenčaković, na stručnom vodstvu, ukazanom povjerenju, vremenu, strpljenju, savjetima i sugestijama kroz niz protokola i izrade ovoga rada. Hvala na prenesenom znanju, iskustvu koje ćemo i dalje koristiti i usmjeravati tijekom svog studiranja i rada.

Prof. dr. sc. Inu Čuriku i izv. prof. dr. sc. Vlatki Čubrić Čurik, na pruženom znanju, pomoći i konstruktivnim savjetima bez kojih ne bismo mogli izraditi ovaj rad.

Dr. sc. Mariji Špehar i Centru za stočarstvo Hrvatske agencije za poljoprivredu i hranu na ustupljenim rodovnicima istarske i paške ovce.

8. POPIS LITERATURE

1. Ackerman, M. S., Johri, P., Spitze, K., Xu, S., Doak, T. G., Young, K. i sur. (2017). Estimating seven coefficients of pairwise relatedness using population-genomic data. *Genetics* 206, 105–118.
2. Agroportal (2019). <https://www.agroportal.hr/ovcarstvo/25797> – pristup 20.06.2021.
3. Alemu, S. W., Kadri, N. K., Harland, C., Faux, P., Charlier, C., Caballero, A., Druet, T. (2021). An evaluation of inbreeding measures using a whole-genome sequenced cattle pedigree. *Heredity* 126(3) 410–423.
4. Alvarez, G., Ceballos, F.C., Quinteiro, C. (2009). The role of inbreeding in the extinction of a European royal dynasty. *PLoS One* 4, e5174.
5. Arcos-Burgos, M., Muenke, M. (2002). Genetics of population isolates. *Clinical Genetics* 61, 233–247.
6. Ballou, J. (1983). Calculating Inbreeding Coefficients from Pedigrees. <http://nature.berkeley.edu/classes/espm-103/Ballou-Inbreeding%20Coefficients.pdf>
7. Bittles, A. H. (2001). Incest, Inbreeding, and their Consequences. *International Encyclopedia of the Social & Behavioral Sciences*, 7254–7259.
8. Bjelland, D. W., Weigel, K. A., Vukasinovic, N., Nkrumah, J. D. (2013). Evaluation of inbreeding depression in Holstein cattle using whole-genome SNP markers and alternative measures of genomic inbreeding. *Journal of Dairy Science* 96, 4697–4706.
9. Brody, T. (2016). Biomarkers. *Clinical Trials*, 377–419.
10. Butler, J. M. (2012). Single Nucleotide Polymorphisms and Applications. U: *Advanced Topics in Forensic DNA Typing*, 347–369.

11. Caetano, A. R. (2009). Marcadores SNP: conceitos básicos, aplicações no manejo e no melhoramento animal e perspectivas para o futuro. *Revista Brasileira de Zootecnia*, 38(spe), 64–71.
12. Calderón, R., Aresti, U., Ambrosio, B., González-Martín, A. (2009). Inbreeding coefficients for X-linked and autosomal genes in consanguineous marriages in Spanish populations: the case of Guipúzcoa (Basque Country). *Annals of human genetics* 73(2), 184–195.
13. Ceballos, F. C., Hazelhurst, S., Ramsay, M. (2018). Assessing runs of Homozygosity: a comparison of SNP Array and whole genome sequence low coverage data. *BMC Genomics* 19(1).
14. Chang, H. L., Fernando, R. L., Grossman, M. (1991). On the principle underlying the tabular method to compute coancestry. *Theoretical and Applied Genetics* 81(2), 233–238.
15. Charlesworth, B., Charlesworth, D. (1999). The genetic basis of inbreeding depression. *Genetical Research* 74(3), 329–340.
16. Charlesworth, D., Charlesworth, B. (1987). Inbreeding Depression and its Evolutionary Consequences. *Annual Review of Ecology and Systematics* 18, 237–268.
17. Charlesworth, D., Willis, J. (2009). The genetics of inbreeding depression. *Nature Reviews Genetics* 10, 783–796.
18. Curik, I., Ferenčaković, M., Sölkner, J. (2014). Inbreeding and runs of homozygosity: A possible solution to an old problem. *Livestock Science* 166, 26–34.
19. Curik, I., Sölkner, J., Stipic, N. (2002). Effects of models with finite loci, selection, dominance, epistasis and linkage on inbreeding coefficients based on pedigree and genotypic information. *Journal of Animal Breeding and Genetics* 119, 101–115.
20. Davenport, C. B. (1908). Degeneration, albinism, and inbreeding. *Science* 28, 454–455.

21. Druet, T., Gautier, M. (2017). A model-based approach to characterize individual inbreeding at both global and local genomic scales. *Molecular Ecology* 26(20), 5820–5841.
22. Druet, T., Legarra, A. (2020). Theoretical and empirical comparisons of expected and realized relationships for the X-chromosome. *Genetics Selection Evolution* 52(1), 1–17.
23. Ferenčaković, M., Hamzić, E., Gredler, B., Solberg, T. R., Klemetsdal, G., Čurik, I., Sölkner, J. (2013a). Estimates of autozygosity derived from runs of homozygosity: empirical evidence from selected cattle populations. *Journal of Animal Breeding and Genetics* 130, 286–293.
24. Ferenčaković, M., Sölkner, J., Čurik, I. (2013b). Estimating autozygosity from highthroughput information: effects of SNP density and genotyping errors. *Genetics Selection Evolution* 45, 42.
25. Fernando, R. L., Grossman, M. (1990). Genetic evaluation with autosomal and X-chromosomal inheritance. *Theoretical and Applied Genetics* 80, 75–80.
26. Gibson, J., Morton, N. E., Collins, A. (2006). Extended tracts of homozygosity in outbred human populations. *Human Molecular Genetics* 15, 789–95.
27. Goudet, J., Kay, T., Weir, B. S. (2018). How to estimate kinship. *Molecular Ecology* 27, 4121–4135.
28. Grosjean, P., Ibanez, F. (2018). *pastecs: Package for Analysis of Space-Time Ecological Series*. R package version 1.3.21. <https://CRAN.R-project.org/package=pastecs>.
29. Grossman, M. (1990). *Quantitative Genetics with X-Chromosomal Loci*.

30. Gutiérrez, J. P., Goyache, F. (2005). A note on ENDOG: a computer program for analysing pedigree information. *Journal of Animal Breeding and Genetics*, 122: 172–176. (maknuti razmak)
31. Hartl, D. L., Clark, A. G., Clark, A. G. (1997). *Principles of population genetics* (Vol. 116). Sunderland, MA: Sinauer associates. Heterosis. *Proceedings of the National Academy of Sciences*. 3(4), 310–312.
32. Howrigan, D. P., Simonson, M. A., Keller, M. C. (2011). Detecting autozygosity through runs of homozygosity: a comparison of three autozygosity detection algorithms. *BMC Genomics* 12, 460.
33. Hrvatske izvorne i zaštićene pasmine – Banka gena domaćih životinja (2019). <https://bag.mps.hr/hrvatske-izvorne-i-zasticene-pasmine/> – pristup 19.06.2021.
34. Hrvatski savez uzgajivača ovaca i koza (2021). <http://www.ovce-koze.hr/ovcarstvo-kozarstvo/ovcarstvo-i-kozarstvo-u-rh/ovcarstvo-i-kozarstvo-u-rh/pasmine-ovaca/istarska-ovca/> – pristup 19.06.2021.
35. Jobling, M. A., Tyler-Smith, C. (2003). The human Y chromosome: an evolutionary marker comes of age. *Nature Reviews Genetics* 4, 598–612.
36. Johnston, H. R., Keats, B. J. B., Sherman, S. L. (2019). Population genetics. In *Emery and Rimoin's Principles and Practice of Medical Genetics and Genomics*, Academic Press. 359–373.
37. Jones, D. F. (1917). Dominance of linked factors as a means of accounting for heterosis. *Genetics*, 2(5), 466.
38. Jordan, D. R., Tao, Y. Z., Godwin, I. D., Henzell, R. G., Cooper, M., McIntyre, C. L. (2005). Comparison of identity by descent and identity by state for detecting genetic regions under selection in a sorghum pedigree breeding program. *Molecular Breeding* 14(4), 441–454.

39. Jung, J., Weeks, D. E., Feingold, E. (2006). Gene-dropping vs. empirical variance estimation for allele-sharing linkage statistics. *Genetic Epidemiology* 30(8), 652–665.
40. Koopae, H. K., Koshkoiyeh, A. E. (2014). SNPs genotyping technologies and their applications in farm animals breeding programs: review. *Brazilian Archives of Biology and Technology* 57(1), 87–95.
41. Kudo, A., Sakaguchi, K. (1963). A Method for Calculating the Inbreeding Coefficient II. Sex-Linked Genes. *American journal of human genetics* 15(4), 476–480.
42. Lahn, B. T., Page, D. C. (1999). Four evolutionary strata on the human Y chromosome. *Science* 286, 964–967.
43. Leutenegger, A. L., Prum, B., Génin, E., Verny, C., Lemainque, A., Clerget-Darpoux, F., Thompson, E. A. (2003). Estimation of the inbreeding coefficient through use of genomic data. *The American Journal of Human Genetics* 73(3), 516–523.
44. MacCluer, J. W., VandeBerg, J. L., Read, B., Ryder, O. A. (1986). Pedigree analysis by computer simulation. *Zoo Biology* 5(2), 147–160.
45. Maignel, L., Boichard, D., Verrier, E. (1996). Genetic variability of French dairy breeds estimated from pedigree information. *Interbull Bull*, 14, 49–54.
46. Marwal, A., Sahu, A. K., Gaur, R. K. (2014). Molecular Markers. *Animal Biotechnology*, 289–305.
47. McQuillan, R. i sur. (2008). Runs of homozygosity in European populations. *The American Journal of Human Genetics* 83, 359–372.
48. Mikulec, D., Pavić, V., Sušić, V., Mioč, B., Mikulec, K., Barać, Z., Prpić, Z., Vnučec, I. (2007.): Odlike vanjštine različitih kategorija istarskih ovaca, *Stočarstvo* 61(1), 13–22.
49. Milligan, B. G. (2003). Maximum-likelihood estimation of relatedness. *Genetics* 163, 1153–1167.

50. Mioč, B., Pavić, V., Posavi, M., Sinković, K. (1999). Program uzgoja i selekcije ovaca u Republici Hrvatskoj. HSSC, Zagreb.
51. Moioli, B., Georgoudis, A., Napolitano, F., Catillo, G., Giubilei, E., Ligda, C., Hassanane, M. (2001). Genetic diversity between Italian, Greek and Egyptian buffalo populations. *Livestock Production Science* 70(3), 203–211.
52. Nei, M. (1977). F-statistics and analysis of gene diversity in subdivided populations. *Annals of human genetics* 41(2), 225–233.
53. Nei, M., Chesser, R. K. (1983). Estimation of fixation indices and gene diversities. *Annals of human genetics* 47(3), 253–259.
54. Ooms, J. (2021). writexl: Export Data Frames to Excel 'xlsx' Format. R package version 1.4.0. <https://CRAN.R-project.org/package=writexl>.
55. Othman, O. E., Ramadan, H. A. I., Samad, M. A., Abou Mossallam, A. A., de Hondt, H. A., El Nahas, S. M. (2005). Genetic diversity between different breeds of goat reared in Egypt using microsatellite analysis. *Journal of Genetic Engineering and Biotechnology (NRC)* 3(1), 43–57.
56. Pandek, K., Mioč, B., Barać, Z., Pavić, V., Antunac, N., Prpić, Z. (2005). Mliječnost nekih pasmina ovaca u Hrvatskoj. *Mljekarstvo: časopis za unaprjeđenje proizvodnje i prerade mlijeka* 55(1), 5–14.
57. Pavić, V., Mioč, B., Barać, Z., Vnućec, I., Sušić, V., Antunac, N. i Samardžija, D. (2005). Vanjština paške ovce. *Stočarstvo*, 59 (2), 83-90. Preuzeto s <https://hrcak.srce.hr/1948> – pristup 19.06.2021.
58. Pavlinić, P. (1936). Paška ovca. Poseban otisak iz Veterinarskog arhiva, knjiga 6, Zagreb.

59. Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A. R., Bender, D. i sur. (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. *American Journal of Human Genetics* 81, 559–575.
60. Purfield, D., Berry, D., McParland, S., Bradley, D. (2012). Runs of homozygosity and population history in cattle. *BMC Genetics* 13, 70.
61. R Core Team (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.
62. Ralls, K., Frankham, R., Ballou, J. D. (2014). Inbreeding and Outbreeding. Reference Module in Life Sciences.
63. Ritland, K. (1996). Estimators for pairwise relatedness and individual inbreeding coefficients. *Genetics Research* 67, 175–185.
64. Robinson, S. P., Simmons, L. W., Kennington, W. J. (2014). Sex-biased mortality associated with inbreeding in *Drosophila melanogaster*. *BMC Evolutionary Biology* 14, 51.
65. Sargolzaei, M., Iwaisaki, H., Colleau, J. J. (2006). CFC: A tool for monitoring genetic diversity. Proc. 8th World Congr. Genet. Appl. Livest. Prod., CD-ROM Communication 27–28. Belo Horizonte, Brazil.
66. SAS Institute Inc. (2013). SAS® 9.4 Guide to Software Updates and Product Changes. Cary, NC: SAS Institute Inc.
67. SAS Institute Inc. (2020). JMP® 15 Basic Analysis. Cary, NC: SAS Institute Inc.
68. Schaffner, S. F. (2004). The X chromosome in population genetics. *Nature Reviews Genetics* 5, 43–51.
69. Skaletsky, H. i sur. (2003). The male-specific region of the human Y chromosome: a mosaic of discrete sequence classes. *Nature* 423, 825–837.

70. Solé, M., Gori, A. S., Faux, P., Bertrand, A., Farnir, F., Gautier, M., Druet, T. (2017). Age-based partitioning of individual genomic inbreeding levels in Belgian Blue cattle. *Genetics Selection Evolution* 49(1).
71. Stevens, E. L., Heckenberg, G., Roberson, E. D. O., Baugher, J. D., Downey, T. J., Pevsner, J. (2011). Inference of Relationships in Population Data Using Identity-by-Descent and Identity-by-State. *PLoS Genetics* 7(9), e1002287.
72. Szulkin, M., Bierne, N., David, P. (2010). Heterozygosity-fitness correlations: a time for reappraisal. *Evolution: An International Journal of Organic Evolution* 64, 1202–1217.
73. Šimpraga, M., Mioč, B., Vukotić, B., Radin, L. (2008). Hrvatske izvorne pasmine ovaca, osnova ekološke ovčarske proizvodnje-1. dio. *Ekološka poljoprivreda; ovčarstvo u Hrvatskoj. Veterinarska stanica* 39(1), 9–17.
74. VanRaden, P. M. (2008). Efficient methods to compute genomic predictions. *Journal of Dairy Science* 91, 4414–4423.
75. Vigeland, M. D. (2020). Relatedness coefficients in pedigrees with inbred founders. *Journal of Mathematical Biology* 81, 185–207.
76. Vignal, A., Milan, D., SanCristobal, M., Eggen, A. (2002). A review on SNP and other types of molecular markers and their use in animal genetics. *Genetics Selection Evolution*, 34(3), 275.
77. Wang, J. (2007). Triadic IBD coefficients and applications to estimating pairwise relatedness. *Genetics Research* 89, 135–153.
78. Wang, J. (2016). Pedigrees or markers: which are better in estimating relatedness and inbreeding coefficient? *Theoretical Population Biology* 107, 4–13.
79. Waters, P. D. i sur. (2007). Sex chromosomes of basal placental mammals. *Chromosoma* 116, 511–518.

80. Waterston, R. H. i sur. (2002). Initial sequencing and comparative analysis of the mouse genome. *Nature* 420, 520–562.
81. Weir, B. S., Cockerham, C. C. (1984). Estimating F-statistics for the analysis of population structure. *Evolution*, 1358–1370.
82. Weir, B. S., Hill, W. G. (2002). Estimating F-statistics. *Annual review of genetics* 36(1), 721–750.
83. Wickham, H., Bryan, J. (2019). readxl: Read Excel Files. R package version 1.3.1. <https://CRAN.R-project.org/package=readxl>
84. Woodruff, D. S. (2001). Populations, Species, and Conservation Genetics. *Encyclopedia of Biodiversity*, 811–829.
85. Wright, S. (1922). Coefficients of inbreeding and relationship. *The American Naturalist* 56, 330–338.
86. Wright, S. (1949). The genetical structure of populations. *Annals of eugenics* 15(1), 323–354.
87. Yengo, L., Wray, N. R., Visscher, P.M. (2019). Extreme inbreeding in a European ancestry sample from the contemporary UK population. *Nature Communications* 10, 3719.
88. Zheng, Q., Bourgain, C. (2009). Kininbcoef: Calculation of kinship and inbreeding coefficients. https://www.illumina.com/documents/products/datasheets/datasheet_ovinesnp50.pdf (zadnje pristupljeno 28.06.2021.).

9. SAŽETAK

Inbriding, odnosno parenje jedinki povezanih podrijetlom kvantificira se koeficijentom inbridinga (F). Srodne jedinke međusobno se križaju s ciljem fiksacije recesivnih alela što omogućuje ispoljavanje njihovih pozitivnih učinaka. Međutim, fiksacijom recesivnih alela ispoljavaju se i štetni učinci, što je primijećeno kod malih populacija domaćih životinja. S obzirom da su mnoga evolucijski i ekonomski važna svojstva djelomično određena genima na spolnom kromosomu X , neophodno ga je uključiti u procjenu koeficijenta inbridinga kod ženskih jedinki. Prema istraživanjima, X kromosomalni inbriding uz pozitivne učinke, povećava i frekvenciju štetnih spolno ograničenih alela te tako mijenja odnos spolova i povećava rizik od izumiranja vrste, tj. pasmine. Vrijednosti koeficijenta inbridinga, tradicionalno su procjenjivane na temelju rodovnika, međutim sve većom dostupnošću genomske informacije razvija se sve veći broj molekularnih metoda procjene koje nadopunjuju nedostatke. U ovom radu procijenjeni su koeficijenti inbridinga na autosomima i spolnom kromosomu X u populacijama istarske i paške ovce (68, odnosno 42 ženske jedinke) na temelju rodovnika te genomske informacije (50k SNP čip), koristeći 2 različite metode: ROH i F_{IS} . Upotrebom studentovog t -testa i Pearson-ove korelacije uspoređene su spomenute metode procjene te autosomalni i X kromosomalni koeficijenti inbridinga. Također, ispitana je kompletnost rodovnika u obje populacije. U svim slučajevima molekularni koeficijenti su statistički značajno veći od koeficijenata inbridinga na temelju rodovnika, što uz niske ECG vrijednosti ukazuje na loše kompletnosti rodovnika. Kako bi se provela preciznija i pouzdanija analiza, prvenstveno je potrebno simulirati populaciju s dubokim rodovnikom i visokom gustoćom markera što bi svakako razjasnilo obrazac rekombinacije i nastajanja ROH segmenata na X kromosomu.

Ključne riječi: *X kromosom, istarska ovca, paška ovca, inbriding, koeficijenti inbridinga*

10. SUMMARY

Inbreeding, or the mating of individuals related by ancestry, is quantified by the inbreeding coefficient (F). Related individuals interbreed with each other with the aim of fixing recessive alleles, which allows the manifestation of their positive effects. However, fixation of recessive alleles also has deleterious effects, which has been observed in small populations of domestic animals. Since many evolutionarily and economically important traits are partly determined by genes on sex chromosome X, it is necessary to include it in the evaluation of the inbreeding coefficient in females. According to research, inbreeding on the X chromosome, in addition to the positive effects, also increases the frequency of deleterious sex-specific alleles, which leads to gender bias and increases the risk of species extinction, i.e. breed. Inbreeding coefficient values have traditionally been estimated based on pedigree, but with the increasing availability of genomic information, a growing number of molecular estimation methods are being developed to complement the shortcomings. In this work, inbreeding coefficients on autosomes and sex chromosome X in Istrian and Pag sheep populations (68 and 42 females, respectively) were estimated based on pedigree and genomic information (50k SNP chip), using 2 different methods: ROH and F_{IS} . Using Student's t-test and Pearson correlation, the above evaluation methods and autosomal and X-linked inbreeding coefficients were compared. In addition, the completeness of the pedigrees in both populations was examined. In all cases, the molecular coefficients are statistically significantly higher than the pedigree-based inbreeding coefficients, indicating poor pedigree completeness at low ECG values. To perform a more accurate and reliable analysis, it is primarily necessary to simulate a population with a deep pedigree and high marker density which would certainly clarify the pattern of recombination and formation of ROH segments on the X chromosome.

Keywords: *X chromosome, Istrian sheep, Pag sheep, inbreeding, inbreeding coefficients*

11. ŽIVOTOPIS

Mario Shihabi

Rođen sam 06. ožujka 1997. godine u Zagrebu. Pohađao sam Osnovnu školu „Vidici“ u Šibeniku, a srednju školu „Medicinska i kemijska škola Šibenik“ smjer farmaceutski tehničar, završavam 2015. godine. Potom, iste godine upisujem preddiplomski studij, smjer Agroekologija na Agronomskom fakultetu Sveučilišta u Zagrebu. Tijekom 2017. godine vlasnik sam ugostiteljskog obrta u Šibeniku. U periodu od 2018. godine do danas, član sam studentske udruge „Klub studenata Agronomskog fakulteta“ u kojoj sam obnašao funkciju dopredsjednika tijekom 2019. i 2020. godine. Nakon završetka preddiplomskog studija, 2019. godine upisujem prvu godinu diplomskog studija Genetika i oplemenjivanje životinja. U periodu 2020. i 2021. godine, izabran sam u predstavničko tijelo studenata („Studentski zbor“) Agronomskog fakulteta Sveučilišta u Zagrebu. U istom periodu, član sam povjerenstva za izvanstavnne aktivnosti te fakultetskog vijeća navedene sastavnice. U slobodno bavim se velikim nogometom i futsalom koji aktivno igram u klubu „MNK Pokrovnik“. Član sam futsal ekipe Agronomskog fakulteta Sveučilišta u Zagrebu od 2015. godine te obnašam funkciju kapetana od 2018. godine do danas. Služim se engleskim jezikom B2 stupnja.

Katarina Turkalj

Rođena sam 22. listopada 1997. godine u Zagrebu. Pohađala sam Osnovnu školu „Mate Lovrak“ u Zagrebu, a srednjoškolsko obrazovanje provodim u „XVIII. gimnaziji“ koju upisujem 2012. godine, smjer jezični engleski – talijanski. Preddiplomski studij, smjer Agroekologija, upisujem 2016. godine na Agronomskom fakultetu Sveučilišta u Zagrebu. U tom periodu BS-a, sudjelujem kao demonstrator u nastavi na Zavodu za poljoprivrednu botaniku gdje sam i aktivan član Botaničke grupe, a kasnije i student edukator. Nakon završetka

preddiplomskog studija, 2019. godine upisujem prvu godinu diplomskog studija Genetika i oplemenjivanje životinja. Tada, sudjelujem i kao član Biračkog odbora na studentskim izborima te postajem zamjenikom člana u Fakultetskom vijeću. Interesi u slobodno vrijeme su mi fotografija te slikarstvo i kulinarstvo. Služim se talijanskim jezikom A2 stupnja te engleskim jezikom B2 stupnja.