UNIVERZA V LJUBLJANI

BIOTEHNIŠKA FAKULTETA

Jana OBŠTETER

**UČINKOVITA UPORABA GENOMSKIH INFORMACIJ V REJSKIH PROGRAMIH ZA MAJHNE POPULACIJE GOVEDI**

DOKTORSKA DISERTACIJA

Doktorski študij Bioznanosti

Ljubljana, 2020

UNIVERZA V LJUBLJANI

BIOTEHNIŠKA FAKULTETA

Jana OBŠTETER

**UČINKOVITA UPORABA GENOMSKIH INFORMACIJ V REJSKIH PROGRAMIH ZA MAJHNE POPULACIJE GOVEDI**

DOKTORSKA DISERTACIJA

Interdisciplinarni doktorski študijski program Bioznanosti

**EFFICIENT USE OF GENOMIC INFORMATION IN BREEDING PROGRAMS FOR SMALL CATTLE POPULATIONS**

DOCTORAL DISERTATION

Interdisciplinary doctoral study programme of Biosciences

Ljubljana, 2020

Doktorsko delo je zaključek Interdisciplinarnega doktorskega študijskega programa Bioznanosti. Delo je bilo opravljeno na Katedri za …

Senat Biotehniške fakultete je za mentorja doktorskega dela dela imenoval izr. prof. dr. Gregorja Gorjanca in za somentorja dr. Janeza Jenka*.*

Recenzent/ka:

Komisija za oceno in zagovor:

Predsednik: prof. dr. *oz. ustrezni naziv* Ime PRIIMEK

Univerza v Ljubljani, Biotehniška fakulteta, Oddelek za zootehniko

Član: prof. dr. Ime PRIIMEK

Univerza v Ljubljani, Biotehniška fakulteta, Oddelek za zootehniko

Član: prof. dr. Ime PRIIMEK

Univerza v Ljubljani, Biotehniška fakulteta, Oddelek za zootehniko

Datum zagovora:

KLJUČNA DOKUMENTACIJSKA INFORMACIJA

ŠD Dd

DK

KG kvantitativna genetika, genomika, znanost o živalih, genomska selekcija, učinkovitost, govedo

AV OBŠTETER, Jana, Master of Science

SA GORJANC, Gregor (mentor), JENKO, Janez (somentor)

KZ SI-1000 Ljubljana, Jamnikarjeva 101

ZA Univerza v Ljubljani, Biotehniška fakulteta, Interdisciplinarni doktorski študijski program Bioznanosti

LI 2020

IN UČINKOVITA UPORABA GENOMSKIH INFORMACIJ V REJSKIH PROGRAMIH ZA MAJHNE POPULACIJE GOVEDI

TD Doktorska disertacija

OP IV, št. str., št. pregl., št. sl., št. pril., št. vir.

IJ sl/en

JI sl/en

AI *Izvleček naj vsebuje približno 200 besed (ne sme preseči te strani).*

KEY WORDS DOCUMENTATION

ND Dd

DC

CX quantitative genetics, genomics, animal breeding, selection

AU OBŠTETER, Jana, Master of Science

AA GORJANC, Gregor (supervisor), JENKO, Janez (co-advisor)

PP SI-1000 Ljubljana, Jamnikarjeva 101

PB University of Ljubljana, Biotechnical Faculty, Interdisciplinary doctoral study programme of Biosciences

PY 2020

TI EFFICIENT USE OF GENOMIC INFORMATION IN BREEDING PROGRAMMES FOR SMALL CATTLE POPULATIONS

DT Doctoral disertation

NO IV, št. p., št. tab., št. fig., št. ann., št. ref.

LA sl/en

Al sl/en

AB *Abstract (pribl. 200 besed, ne sme preseči te strani)*

KAZALO VSEBINE

|  |  |  |
| --- | --- | --- |
|  |  |  |
|  | KLJUČNA DOKUMENTACIJSKA INFORMACIJA | II |
|  | KEY WORDS DOCUMENTATION | III |
|  | KAZALO VSEBINE |  |
|  | KAZALO PREGLEDNIC |  |
|  | KAZALO SLIK |  |
|  | KAZALO PRILOG |  |
|  | SEZNAM GESEL *opis ključnih besed, izbor naredi kandidat s pomočjo mentorja* |  |
|  | OKRAJŠAVE IN SIMBOLI |  |
| **1** | **UVOD** | 1 |
| **2** | **PREGLED OBJAV** | 2 |
| 2.1 | NASLOV PODPOGLAVJA | 3 |
| **2.1.1** | **Naslov podpoglavja** | 4 |
| 2.1.1.1 | Naslov podpoglavja |  |
| **…** |  | **…** |
| **3** | **MATERIAL IN METODE** |  |
| 3.1 | MATERIAL |  |
| 3.2 | METODE |  |
| **4** | **REZULTATI** *(ali skupno poglavje: REZULTATI Z RAZPRAVO*  *po dogovoru z mentorjem)* |  |
| **5** | **RAZPRAVA** |  |
| **6** | **SKLEPI** |  |
| **7** | **POVZETEK** |  |
| **8** | **VIRI**  **ZAHVALA** |  |
|  | **PRILOGE** |  |

KAZALO ZNANSTVENIH DEL

Obšteter J., Jenko J., Hickey J. M., Gorjanc G. 2019. Efficient use of genomic information for sustainable genetic improvement in small cattle populations. Journal of Dairy Science, 102, 11: 9971-9982.

KAZALO SLIK

KAZALO PRILOG

Priloga A: Licenčna pogodba za uporabo članka v tiskani in elektronski verziji disertacije 1

Priloga B: Licenčna pogodba za uporabo članka v tiskani in elektronski verziji disertacije 2

Priloga C: Licenčna pogodba za uporabo članka v tiskani in elektronski verziji disertacije 3

OKRAJŠAVE IN SIMBOLI

# **1** UVOD S PREDSTAVITVIJO PROBLEMATIKE, CILJEV IN HIPOTEZ/ZNANSTVENIH VPRAŠANJ

## 1.1 Genetsko izboljševanje populacij

### **1.1.1 Osnove dedovanja in kvantitativnih lastnosti**

Večino lastnosti in bolezni, če ne vse, imajo genetsko komponento. V diploidnih dvospolnih populacijah dobi potomec polovico genetskega materiala od matere in polovico od očeta. Vendar se potomci istih staršev med seboj razlikujejo, saj biološki procesi in evolucijske sile iz generacije v generacijo spreminjajo zaporedje DNA. Glavna vira genetske raznovrstnosti sta procesa mutacije in rekombinacije, ki ustvarjata razlike med posamezniki. Mutacije so posledica netočnega podvajanja molekule DNA. Za samo evolucijo so pomembne mutacije v spolnih celicah, ki se prenesejo na naslednje generacije. Druga sila v spolnih populacijah je rekombinacija. Pri procesu nastajanja spolnih celic v mejozi prihaja do rekombinacije starševskih kromosomov, ki ustvarja genetsko raznovrstnost. K slednji pripomore tudi naključna segregacija starševskih kromosomov v gamete.

Genetski ustroj organizma imenujemo njegov genotip. V populaciji nas zanimajo mesta na genomu oz. lokusi, na katerih se posamezniki med seboj razlikujejo, t.i. polimorfna mesta. Večina polimorfnih mest ima dve različici oz. dva alela, t.j. so bialelni. To je v skladu z modelom neskončnega števila mest (angl. i*nfinite site model*) , ki pravi, da vsako mesto na genomu mutira natanko enkrat (Kimura, 1969). Na takšnih mestih so posamezniki lahko homozigotni (imajo dve kopiji istega alela) ali heterozigotni (imajo po eno kopijo vsakega alela). Ti lokusi se dedujejo po Mendelskih zakonih, zato frekvence genotipov v populaciji sledijo Hardy‑Weinbergovem ravnovesju. Po pričakovanjih je frekvenca obeh alelov v populaciji enaka, pričakovane frekvence genotipov pa so ¼ za homozigote za referenčni alel, ½ za heterozigote in ¼ za homozigotne za alternativni alel.

Genotip posameznika delno ali v celoti pogojuje posameznikov fenotip, ki je morfološka, fiziološka ali opažena lastnost. Nekatere izmed izraženih lastnosti so v celoti nadzorovane z genotipom in so pod vplivom enega ali majhnega števila genov. Te lastnosti so diskretne in sledijo Mendelskemu vzorcu dedovanja ter jih imenujmo Mendelske oz. kvalitativne lastnosti. Druge lastnosti so le delno nadzorovane z genotipom posameznika, delno pa z okoljskimi vplivi. Fenotipska vrednost je tako seštevek vpliva genov in okoljskega odklona (Falconer in Mackay, 1996). Te lastnosti imajo kontinuirano porazdeljene fenotipske vrednosti, ki sledijo normalni porazdelitvi, in jih imenujemo kvantitativne oz. kompleksne lastnosti. Kvantitativne lastnosti so nadzorovane z velikim številom genov, t.j. so poligenske. Lokuse z vplivom na fenotip imenujemo vzročni oz. kvantitativni lokusi (angl. *quantitative trait loci*, QTL), pri čemer pa lahko imajo nekateri lokusi večji (merljiv) učinek na fenotip.

Pri kvantitativnih lastnostih nas ne zanima učinek posameznih genov na fenotip, ampak skupni učinek vseh genov (genetska vrednost). Med genetske učinke sodijo aditivni učinki, kjer se učinki alelov na fenotip seštevajo, in pa neaditivni učinki, ki so posledica interakcije. Na fenotip lahko vpliva interakcija alelov na lokusu oz. dominanca, interakcija alelov na različnih lokusih oz. epistaza, ali pa interakcija genotipa z okoljem. Le aditivni učinki se dedujejo, zato si za namen selekcije in odbire staršev prizadevamo iz fenotipskih vrednosti izluščiti aditivno genetsko vrednost. Povezano s tem merimo tudi delež fenotipske variance, ki ga lahko pojasnimo z aditivnimi vplivi, kar imenujemo dedljivostni delež oz. heritabiliteta.

### 1.1.2 Selekcija in zgodovina selekcijskih ciljev v govedoreji

V kmetijstvu genetsko izboljšujemo populacije rastlin in živali. Glavna gonilna sila genetskega izboljšanja populacije je selekcija. Ključen korak selekcije je odbira genetsko najboljših posameznikov za starše naslednje generacije. Identifikacija teh posameznikov je težavna zaradi več virov variabilnosti. Med temi so genetska variabilnost za lastnost, velik okoljski vpliv, segregacija in rekombinacija pri potomcih že identificiranih elitnih živali ter vključenost različnih organizacij v selekcijsko delo. Uspeh selekcije merimo z doseženim genetskim napredkom (ΔG),

1. ; [1]

kjer je *∆G* letni genetski napredek, merjen v enotah standardnega odklona na leto, *i* intenzivnost selekcije, *r* točnost selekcije, *σa* genetska standardna deviacija za lastnost v populaciji in *L* generacijski interval v letih (Lush, 1945).

V zadnjem stoletju (50-60 letih) smo z učinkovito selekcijo, pocenitvijo proizvodnje in razvojem tehnologije naredili ogromen napredek v izboljšanju produktivnosti gospodarsko pomembnih rastlin in živali kot odgovor na svetovno povečanje potreb po hrani. V kmetijstvu izvajamo selekcijo na sklop izbranih lastnosti. Slednje ponavadi izberemo na podlagi njihove gospodarske pomembnosti oz. ekonomske vrednosti, vendar pa morajo zadostiti tudi drugim kriterijem. Lastnosti, na katere odbiramo morajo imeti zadostno genetsko variabilnost in heritabiliteto, biti jasno definirane, imeti nizko ceno merjenja ter biti redno merjene. Uporabimo lahko tudi posredne (angl. *proxy*) lastnosti, če imajo visoko genetsko korelacijo z gospodarsko pomembno lastnostjo ter imajo nižji strošek merjenja, višjo heritabiliteto, ali pa jo lahko merimo bolj zgodaj v življenju kot gospodarsko pomembno lastnost (Shook, 1989).

V govedoreji se je nabor lastnosti za selekcijo skozi zgodovino precej spreminjal (Miglior in sod., 2017). Začetki selekcije pri mlečnem govedu segajo v konec 19. in začetek 20. stoletja. Glaven cilj selekcije je bil velika prireja mleka in maščob. V ta namen so pričeli z merjenjem prireje mleka velikega števila čistopasemskih krav, vključevanjem v različne registre, kasneje pa tudi z merjenjem vsebnosti maščobe. Selekcionisti so že zelo zgodaj med selekcijske cilje vključili tudi lastnosti zunanjosti - delno zaradi njihove povezave s prirejo mleka in delno zaradi večanja pomembnosti samega izgleda živali za razstave in pasemske standarde. Kasneje so rejci prepoznali vrednost mleka kot vira beljakovin, zato so pričeli tudi z rednim merjenjem in selekcijo na vsebnost beljakovin. Šele v drugi polovici 20. stoletja se je oblikoval bolj trajnosten selekcijski cilj kot odgovor na večanje problemov z zdravjem in plodnostjo ter skrbi glede dobrobita živali. Za izboljšanje dobičkonosnosti programov in zdravja živali so konec 20. stoletja med selekcijske cilje vključili dolgoživost, plodnost, težavnost telitve, lastnosti zdravja; ter lastnosti, ki olajšajo delo z živalmi, kot npr. delovna sposobnost. V 21. stoletju so razvoj tehnologij, metodologij kot tudi uvedba genomske selekcije omogočili spremljanje in novih lastnosti in vključitev teh med selekcijske cilje. Razvoj in pocenitev tehnologije sta omogočila visoko zmogljivo zbiranje fenotipskih podatkov (angl. *high-thoughput phenotyping*), genomska selekcija pa je omogočila selekcijo na lastnosti z nizko heritabiliteto, lastnosti izraženih le pri enem spolu ali pozno v življenju in lastnosti, katerih merjenje je drago ali težavno. Te nove lastnosti težijo k izboljšanju učinkovitost živali, njihovega dobrobita, vpliva živinoreje na okolje ter izboljšanju hranilne vrednosti in tehnoloških lastnosti živalskih produktov. Veliko rejskih programov v selekcijo tako že vključuje učinkovitost izrabe krme (angl. *feed efficiency*), izpuste metana, odpornost na vročinski stres, zdravje parkljev, imunski odgovor oz. odpornost na okužbe in bolezni, sestavo mleka, koagulacijske lastnosti mleka in lastnosti pomembne pri uporabi reprodukcijskih tehnologij kot npr. embrio transfer (Miglior in sod., 2017).

## 1.2 GENETSKI NAPREDEK

### 1.2.1 Razvoj testov

Ugotovitev, da je za izboljšanje naslednje generacije pomembna posameznikova genetska in ne fenotipska vrednost so vodile do razvoja sistematičnih testov za oceno genotipa živali. Začetki sistematične selekcije segajo v 18. stoletje do Sira Roberta Bakewella, ki je pričel z načrtnim zbiranjem podatkov ter načrtnim in nadzorovanim parjenjem. Postavil je temelje testa potomcev oz. progenega testa, pri katerem testiramo skupino potomcev in na podlagi rezultatov izberemo najboljše posameznike. Natančni rodovniški in fenotipski podatki so tako postali temelj za razvoj učinkovite selekcije.

#### 1.2.1.2 Klasična selekcijska shema

Razlog za uvedbo progenega testa je bila ugotovitev, da medtem ko rezultat posameznika meri njegov fenotip, rezultati njegovih potomcev merijo njegov genotip (Rendel in Robertson, 1950). Progeni test je bil še posebno zanimiv za lastnosti, ki se izrazijo samo pri enem spolu, kot npr. prireja mleka. Že zgodaj so razumeli, da število potomcev vpliva na točnost samega testiranja (Lush, 1935). Razumeli so tudi, da je potrebno vzpostaviti ravnovesje med točnostjo testa, ki povečuje genetski napredek, in trajanjem testa, ki podaljšuje generacijski interval in zavira genetski napredek (Dickerson in Hazel, 1943). Skozi leta se je progeni test v reji mlečne govedi vse bolj formaliziral (Rendel in Robertson, 1950) do t.i. klasične selekcijske sheme (slika 1). Rejci na podlagi ocene zunanjosti izberejo novorojena ženskih telet za plemenske telice (matere mater), ki jih osemenijo s plemenskimi biki. Po pridobitvi lastnih fenotipskih vrednosti in napovedi plemenskih vrednosti so najboljše krave izbrane kot bikovske matere (matere očetov). Te so osemenjene z elitnimi plemenskimi biki, potomci teh načrtnih parjenj pa so kandidati za naslednjo generacijo plemenskih bikov. Moški potomci načrtnih parjenj se na podlagi povprečja staršev, sorodnosti in pregleda zunanjosti uhlevijo za lastni test, ki traja približno eno leto. Na podlagi rezultatov lastnega testa in povprečja staršev se najboljšim bikcem odvzame manjša doza semena za namene progenega testa. Biki imajo za čas osemenjevanja do porabe doz semena status mladega bika, po porabi semena pa pridobijo status čakajočega bika. S čakajočimi biki se praviloma ne osemenjuje. Po rojstvu potomcev in pridobitvi njihovih fenotipskih vrednosti (5-6 let), se bikom napove plemenska vrednosti, najboljši izmed mladih pa so proglašeni za plemenske bike, ki se lahko uporabljajo v večjem obsegu. Odbrani plemenski biki se uporabljajo za osemenjevanje vseh krav v populaciji in nastopajo kot očetje mater, le najboljši pa se uporabijo za osemenjevanje bikovskih mater in tako nastopajo kot očetje očetov.

Progeno testiranje omogoča zelo točno odbiro, ki pa je pogojena z dolgim generacijskim intervalom kar zavira genetski napredek. Prav tako progeno testiranje zahteva velik finančni vložek za nakup živali, vzpostavitev in vzdrževanje infrastrukture za rejo živali, osebje, hrano za živali, odvzem in hranjenje semena ter osemenitev ženskih živali.

Selekcijski parametri se bistveno razlikujejo med moško in žensko potjo selekcije pri govedi, zato sta Rendel in Robertson, 1950, definirala štiri poti selekcije pri govedu: oče-sin, oče-hči, mati-sin in mati-hči. Najstrožja selekcija se vrši v selekciji očetov, saj je vpliv moških živali zaradi velikega števila potomcev, ki jih dosežemo preko umetnega osemenjevanja, znatno večji od vpliva ženskih živali, ki imajo v povprečju le enega potomca letno.

### 1.2.3 Genomska selekcijska shema

Vse do začetka 21. stoletja je progeni test pri govedu veljal kot zlati standard. Nizka točnost odbire mladih živali brez lastnih fenotipskih vrednosti in fenotipskih vrednosti potomcev je onemogočala skrajšanje generacijskega intervala. Leta 2000 so Meuwissen, Goddard in Hayes predstavili idejo t.i. genomske selekcije, ki je zaradi načina napovedi plemenskih vrednosti omogočila zgodnejšo odbiro staršev z relativno visoko točnostjo. Genomska selekcijska shema pri mlečnem govedu (slika 2) prinaša največje spremembe v poti selekcije očetov. Takoj po rojstvu genomsko testiramo moška teleta, ki so ponavadi potomci načrtnih parjenj in ostalih parjenj z visokim povprečjem staršev. Izmed teh takoj ob spolni zrelosti odberemo najboljše kot plemenske bike za širšo uporabo v umetnem osemenjevanju. Zaradi nižjih točnosti odbire plemenskih bikov je praksa genomske selekcije odbira večjega števila očetov, s čimer se omili posledice retrospektivno slabe izbire plemenskega bika (Pryce, 2010). Genomsko informacijo lahko uporabimo tudi za odbiro mater, tako da genomsko testiramo telice in odberemo najboljše za uporabo v načrtnih parjenjih (matere očetov) ali pa splošnem osemenjevanju (matere mater). Gonilo večjega genetskega napredka je skrajšan generacijski interval, saj lahko ženske živali odberemo že pred prvo zaključeno laktacijo (Schaeffer, 2006).

Čeprav je točnost odbire plemenskih bikov nižja z genomskim kot progenim testiranjem, je genomska selekcija podvojila genetski napredek pri govedu (Schaeffer, 2006; Wiggans in sod., 2017). Glavni razlog je občutno skrajšanje generacijskega intervala predvsem v selekcijskih poteh oče-oče (iz sedem na manj kot tri leta) in oče-hči (iz sedem na pet let) (García-Ruiz in sod., 2016). Povečanju genetskega napredka z genomsko selekcijo sta prisostvovala tudi: i) večja točnost zgodnjih selekcijskih odločitev odbire živali brez lastnih fenotipskih vrednosti in fenotipskih vrednosti potomcev; in ii) povečanje intenzivnost selekcije, čemur sta botrovala pocenitev in slrajšanje testiranja. Uvedba genomske selekcije je povečala selekcijski diferencial za vse lastnosti, še posebno za lastnosti z nizko heritabiliteto kot sta zdravje in plodnost (García-Ruiz in sod., 2016).

### 1.2.2 Razvoj statističnih metod za napoved plemenskih vrednosti

Sočasno z razvojem testov je potekal tudi razvoj statističnih metod za napoved genetskih vrednosti. Največji napredek se je zgodil v začetku 20. stoletja, ko je Fisher leta 1918 predstavil infinitezimalni model, ki kontinuirano variabilnost lastnosti pojasni s predpostavko velikega števila genov z malim učinkom. Ideja je vodila do definicije plemenske vrednosti, ki jo je kot prvi leta 1975 napovedal Henderson. Plemenska vrednost je vrednost posameznika glede na povprečno vrednost njegovih potomcev in je dvakratnik odstopanja njegovih potomcev od populacijskega povprečja (Falconer in Mackay, 1996). Plemenska vrednost je seštevek povprečne plemenske vrednosti staršev (angl. *parent average*, PA), saj vsak posameznik prejme polovico genetskega materiala od vsega izmed staršev, in učinka naključne segregacije in rekombinacije starševskih genomov, t.i. učinek Mendelskega vzorčenja (angl. *Mendelian sampling term,* MST) (Kennedy in sod., 1988). Henderson je plemenske vrednosti napovedal kot najboljše linearne nepristranske napovedi (angl. *Best Unbiased Linear Prediction*, BLUP) (Henderson, 1975). »Najboljše« pomeni, da je minimizirana variancanapake napovedi (angl. *prediction error variance*), »nepristranske« pa, da je povprečje ocenjenih plemenskih vrednosti za žival resnična plemenska vrednost živali.

### 1.2.2.1 Napoved plemenskih vrednosti na podlagi rodovniških podatkov

Henderson, 1975, je plemenske vrednosti napovedal po metodi mešanih modelov s sistematičnimi vplivi in plemenskimi vrednostmi živali kot naključnimi vplivi:

*y* =**u + Xb**+**Za**+**e;** [2]

kjer je **X** matrika dogodkov, **b** je vektor ocen za sistematske vplive, **Z** je matrika dogodkov, ki povezuje naključni vpliv živali z opazovanji, **a** je vektor plemenskih vrednosti z normalno porazdelitvijo s povprečjem nič in varianco **A**\**σ2a*, e pa je vektor ostankov z normalno porazdelitvijo s povprečjem 0 in varianco *σ2e*. Pri tem *σ2a* predstavlja genetsko varianco za lastnost, **A** rodovniško matriko sorodstva in *σ2e* varianco za ostanek. Mešani model za napoved plemenskih vrednosti ima nekaj predpostavk. Prva je infinitezimalni model, ki pravi da **y**, **a** in **e** sledijo večrazsežni normalni porazdelitvi ter da lastnosti nadzoruje veliko število genov, ki imajo zelo majhen aditiven učinek ter niso vezani (Fisher 1918; Bulmer 1980). Druga predpostavka modela je, da poznamo genetske in okoljske variance in kovariance v populaciji osnovalcev (angl. *founder population*). Tretja predpostavka je, da model lahko upošteva selekcijo če vsebuje vse lastnosti, na katere smo odbirali (Mrode, 2005).

Napoved plemenskih vrednosti na podlagi rodovniških podatkov je povezala podatke vseh posameznikov z rodovniško matriko sorodstva **A**. Rodovniška matrika sorodstva **A** odraža pričakovano aditivno genetsko sorodstvo med posamezniki (Wright, 1921). Koeficienti v matriki **A** so dvakratniki koeficienta sorodstva (angl. *coancestry*), ki je verjetnost da dva posameznika podedujeta identičen gen po izvoru. Sama napoved plemenskih vrednosti zahteva inverzo matrike **A**. Direktno invertiranje matrike ima kubično kompleksnost, zato je Henderson leta 1976 razvil rekurzivno metodo za pridobitev inverze **A-1**.

Plemenske vrednosti lahko napovemo po modelu očeta, pri čemer uporabimo lastne in fenotipske vrednosti potomcev, ali po modelu živali, kjer uporabimo lastne kot tudi fenotipske vrednosti vseh sorodnikov. Točnost napovedi plemenskih vrednosti na podlagi rodovniških podatkov je močno odvisna od lastne fenotipske vrednosti oz. vrednosti potomcev, narašča pa z večjo heritabiliteto in večjim številom fenotipskih vrednosti živali in njenih tesnih sorodnikov (Mrode, 2005). Plemenska vrednosti za mlade živali brez lastne fenotipske vrednosti je tako povprečje plemenskih vrednosti staršev, točnost katere je zaradi rekombinacije in segregacije majhna. Poseben izziv predstavljajo lastnosti, ki so pogojene s spolom živali, npr. lastnosti mlečnosti, saj so za točno odbiro očetov potrebne fenotipske vrednosti njegovih potomk. Selekcija na podlagi progenega testa je posledično zaznamovana z visoko točnostjo odbire očetov, hkrati pa tudi z dolgim generacijskim intervalom, ki omejuje genetski napredek.

#### 1.2.2.2 Napoved plemenskih vrednosti na podlagi genomskih podatkov

Znižanje cene genotipizacije in razmah genotipizacijske tehnologije je omogočilo obsežnejšo genotipiziranje živali. Najprej so bili podatki o genotipu za napoved plemenskih vrednosti uporabljeni v t.i. selekciji s pomočjo molekularnih označevalcev (angl. *marker assisted selection*, MAS). Model za MAS poleg poligenske komponente (plemenske vrednosti) vključuje tudi sistematičen vpliv genotipa na znanih kvantitativnih vzročnih lokusih oz. vezanih označevalcih za določeno lastnost (Fernando in Grossman, 1989). MAS ni prinesla celostne rešitve problema, saj vzročnih lokusov za določeno lastnost ne poznamo, poleg tega pa ti lokusi pojasnijo majhen delež genetske variance za lastnost. Razlog za slednje je veliko število vzročnih lokusov z majhnim vplivom, težave pri natančni oceni velikost vpliva vzročnih lokusov in dejstva, da uporabljeni genetski označevalci označevalci ponavadi niso vzročni, ampak so v vezavnem neravnovesju z vzročnim lokusom (Collard in Mackill, 2008).

Meuwissen, Hayes in Goddard (2001) so predlagali uporabo velikega števila genetskih označevalcev iz celotnega genoma za napoved t.i. genomskih plemenskih vrednosti. Genomske plemenske vrednosti lahko napovemo po dveh ekvivalentnih modelih, in sicer modelu SNP-BLUP in genomskem BLUP-u (GBLUP). Model za SNP-BLUP je

*y* =**u + Xb**+**Zg**+**e; [**3]

kjer je **g** vektor aditivnih učinkov zamenjave alelov za uporabljene genetske označevalce, matrika **Z** pa matrika dogodkov, ki povezuje učinke označevalcev s fenotipskimi vrednostmi. Vektor **Zg** je tako vektor plemenskih vrednosti. Model za GBLUP je

*y* =**u + Xb**+**Wa**+**e; [4]**

kjer je **a** vektor plemenskih vrednosti, **W** pa matrika dogodkov, ki povezuje fenotipske in plemenske vrednosti. Genomske plemenske vrednosti **g** prihajajo iz porazdelitve N ~ (0, **G** *σ2a*). V modelu za napoved plemenskih vrednosti [2] rodovniško sorodstva **A** nadomesti matrika centriranih in standardiziranih genomskih koeficientov sorodstev **G**, ki odraža realizirano sorodstvo. Genomski koeficienti sorodstva bolj točno ocenijo resnično aditivno sorodstvo kot rodovniški zaradi natančnejše ocene učinka Mendelskega vzorčenja. Kot pri napovedi na podlagi rodovniških podatkov, je predpostavka obeh genomskih modelov infinitezimalni model. Ta predpostavlja, da imajo vsi geni in posledično vsi uporabljeni genetski označevalci vpliv na fenotip. Model tudi predpostavi, da vpliv vsakega označevalca sledi normalni porazdelitvi s povprečjem nič in varianco, ki je apriorij enaka za vse označevalce.

Genomske plemenske vrednosti lahko napovemo tudi z metodami Bayesovske statistike. Ti modeli predpostavijo, da varianca označevalcev ni konstantna ampak prihaja iz apriorij porazdelitve, in/ali da vplivi označevalcev niso normalno porazdeljeni (Meuwissen in sod., 2001). Vsak izmed Bayes modelov določi drugačno apriorij porazdelitev oz. mešanico porazdelitev za varianco in vplive označevalcev. BayesA predpostavi, da imajo genetski označevalci različne variance, ki sledijo invertirani hi-kvadrat porazdelitvi, njihovi vplivi pa sledijo Studentovi razdelitvi (Meuwissen in sod., 2001). Slednja ima debele repe, kar pomeni da so večji vplivi označevalcev bolj verjetni. BayesB sledi BayesA modelu z dodatno predpostavko, da ima le del označevalcev učinek na lastnosti, preostali označevalci pa na lastnosti na vplivajo (varianca in učinek sta nič) (Meuwissen in sod., 2001). BayesC predpostavi, da imajo apriorij vsi označevalci enako varianco, učinki pa sledijo normalni porazdelitvi. Prav tako BayesC dovoljuje možnost, da π del označevalcev nima učinka (BayesCπ) (Habier in sod., 2011).

Napoved genomskih plemenskih vrednosti zahteva referenčno oz. trening populacijo, ki jo sestavljajo genotipizirane živali s fenotipskimi vrednostmi. V referenčni populaciji ocenimo učinke genetskih označevalcev, na podlagi katerih lahko napovemo genomske plemenske vrednosti tudi mladim živalim brez fenotipske vrednosti (Meuwissen in sod., 2001). Točnost te napovedi je višja od točnosti povprečja staršev v klasični selekciji, kar je ena izmed glavnih prednosti genomske selekcije (Schaeffer, 2006; Pryce in Daetwyler, 2012). Ta točnost nam omogoča, da kot starše naslednje generacije mlade živali brez lastne fenotipske vrednosti, s čimer močno skrajšamo generacijski interval.

Genomska selekcija prinaša tudi spremembe v selekcijskih ciljih. Ker je zbiranje fenotipov potrebno le v referenčni populaciji, genomska selekcija izboljša učinkovitost selekcije na lastnosti, ki so težje merljive ali pa je merjenje drago, lastnosti z nizko heritabiliteto in lastnosti, ki se izrazijo pozno v življenju živali. Točnost genomskih plemenskih vrednosti narašča z večjo heritabiliteto, velikostjo referenčne populacije, številom fenotipskih vrednosti genotipiziranih živali, manjšo genetsko razdaljo med referenčno in napovedno populacijo ter z manjšanjem števila efektivnih neodvisnih kromosomskih fragmentov (Daetwyler in sod., 2008; Goddard, 2009; Habier in sod., 2010; Clark in sod, 2011; Goddard in sod., 2011). Število neodvisnih kromosomskih segmentov narekuje vzorec vezavnega neravnovesja med genetskimi označevalci in vzročnimi lokusi, pri čemer manjše število pomeni tesnejšo vezavo. Rekombinacija, mutacija, migracija, naključni genetski tok (angl. g*enetic drift*) in selekcija spreminjajo vzorec vezavnega neravnovesja tekom generacij, kar zmanjšuje točnost genomske napovedi. Zaradi slednjega je ključno, da referenčno populacijo redno posodabljamo z dodajanjem živali iz mlajših generacij, s čimer povečamo sorodnost in zmanjšamo genetsko razdaljo med referenčno in napovedno populacijo (Pszczola in sod., 2012).

### 1.2.2.3 **Združena napoved plemenskih vrednosti v enem koraku**

Leta 2009 so Legarra in sod. predstavili matriko sorodstva **H**, ki združuje rodovniško in genomsko informacijo. V matriki **H** je genomska informacija prenesena na kovarianco med negenotipiziranimi posamezniki. Čeprav je matrika **H** zapletena, ima preprosto inverzo **H-1**:

[5]

kjer je **A22** rodovniška matrika za genotipizirane živali, **G\*** pa prilagojena matrika G. Združevanje matrike **A** in **G** zahteva določene prilagoditve, saj matriki **A** in **G** nista kompatibilni. Rodovniška matrika ponavadi vključuje starejše živali, medtem ko so genotipizirane živali predvsem živali zadnjih generacij. Posledično matriki nimata iste genetske baze: rodovniška matrika predpostavlja povprečje plemenskih vrednosti nič le v bazni populaciji, genomska matrika pa v celotni genotipizirani populaciji, ki je ponavadi tudi mnogo generacij mlajša od rodovniške baze. Zaradi te razlike v obravnavani populaciji matriki **A** in **G** predpostavita različno genetsko varianco za lastnost. Genomska matrika je odvisna od frekvence alelov, ki pa se zaradi naključnega genetskega toka in selekcije fiksirajo, kar zmanjša genetsko varianco. Nekompatibilnost matrik v prvem koraku rešimo z dodatkom koeficienta *a* diagonali matrike **G**, pri čemer je *a* razlika med povprečnim rodovniškim in genomskim koeficientom sorodstva med genotipiziranimi živalmi. S tem izenačimo skalo za rodovniško in genomsko napoved. V drugem koraku posodobljeno matriko **G**\* iz prvega koraka pomnožimo s koeficientom *b.* Ta predstavlja razmerje med genetsko variance trenutne populacije in populacije osnovalcev in je ponavadi manjši od 1 (Legarra in sod., 2018 – NOTES). Ko v modelu za napoved plemenskih vrednosti matriko **G** zamenjamo z matriko **H**, dobimo t.i. GBLUP v enem koraku (angl. s*ingle-step GBLUP*) (Aguilar in sod., 2010). Slednji združi vse fenotipske, rodovniške in genotipske podatke in tako dodatno poveča točnost genomske napovedi (Gray in sod., 2012; Lourenco in sod., 2015).

Kot pri napovedi na podlagi rodovniških podatkov, je najbolj zahtevna operacija napovedi na plemenskih vrednosti na podlagi genomskih podatkov (v enem koraku) kreiranje in invertiranje matrike **G**, ki ima kubično kompleksnost s številom genotipov. Raziskovalci so zato razvili pristope za zmanjšanje računalniškega bremena. Misztal in sod. (2014) so razvili t.i. algoritem za testirane in mlade (angl. *Algorithm for Proven and Young*, APY). Algoritem ustvari redko inverzo matrike **G-1** preko direktne inverze le malega števila živali, t.i. jedrnih (angl. *core*) živali,in uporabe genomske rekurzije na preostalih ne‑jedrnih (angl. *non-core*) živalih. Računalniško breme lahko zmanjšamo tudi z redukcijo dimenzionalnosti matrike **G** preko singularne razgradnje vrednosti (angl. s*ingular value decomposition*) (Ødegård in sod., 2018).

## 1.3 Genetska variabilnost

Genetska variabilnost je pomembna tako za doseganje genetskega napredka za gospodarsko pomembne lastnosti kot tudi za normalno delovanje živali. Genetsko variabilnost merimo z aleli, ki so prisotni v populaciji. V populacijah najdemo dva tipa procesov, ki spreminjata genetsko variabilnost. Prvi so sistematični procesi, ki frekvence alelov spreminjajo v pričakovani smeri in količini. Mednje sodijo selekcija, migracija in mutacija. Drugi so disperzni procesi, ki se pojavljajo v majhnih populacijah kot posledica vzorčenja. Pri teh lahko predvidimo količino spremembe, ne pa tudi smer. Mednje sodita naključni genetski tok in inbriding (Falconer in Mackay, 1996).

V rejskih programih je doseganje genetskega napredka in ohranjanje genetske variabilnosti pogosto v konfliktu. Maksimiranje genetskega napredka je ponavadi povezano z izgubo genetske variabilnosti, ki je posledica visoke intenzivnosti selekcije očetov, ko-selekcije družinskih članov, predvsem z BLUP plemenskimi vrednostmi, in skrajšanja generacijskega intervala z genomsko selekcijo. Čeprav je doseganje genetskega napredka glavni cilj selekcije, si mora ta za dolgoročni napredek prizadevati tudi za ohranjanje genetske variabilnosti. Iz enačbe za genetski napredek [1] je namreč razvidno, da manjšanje genetske variabilnost manjša tudi genetski napredek.

### 1.3.1 Mere za genetsko variabilnost in selekcija

Selekcija dosega genetski napredek s spreminjanjem frekvence alel in posledično genetske variabilnosti v populaciji. S selekcijo si prizadevamo za spreminjanje frekvence alelov na vzročnih lokusih za gospodarsko pomembne lastnosti. S tem spreminjamo tako imenovano genetsko varianco, ki vključuje aditivno genetsko varianco in varianco za interakcije. Aditivna genetska varianca je varianca plemenskih vrednosti in sestoji iz dveh delov. Prvi je genska varianca, ki je aditivna genetska varianca v odsotnosti vezavnega neravnovesja in je tako odvisna le od frekvence alelov (in koeficienta dominance). Če predpostavimo infinitezimalni model, ostaja genska varianca s selekcijo nespremenjena. Druga komponenta genetske variance je prispevek vezavnega neravnovesja, ki odraža kovarianco vplivov alelov na različnih lokusih. Izguba genetske variabilnost za gospodarsko pomembne lastnosti je neizbežna, vendar ni problematična, če frekvence spreminjamo v pravo smer. Velikokrat je fiksacija koristnih alelov na vzročnih lokusih tudi zaželena.

Vendar pa selekcija zaradi nenaključnega parjenja in večje možnosti za odbiro sorodnikov povečuje parjenje v sorodstvu oz. inbriding (Woolliams in sod., 2015). Koeficient inbridinga *F* je lastnost diploidnih organizmov in je definiran kot verjetnost, da sta v posamezniku oba alela na nevtralnem mestu na genomu (brez pritiska selekcije) identična po izvoru (Wright, 1921). Povečevanje inbridinga v populaciji je lahko problematično tako za gospodarske lastnosti, še posebno pa za funkcionalne lastnosti, ki so pomembne za normalno delovanje živali. Slednjih s selekcijo sicer ne želimo izboljšati, vendar tudi ne poslabšati. Inbriding povečuje homozigotnost v populaciji. Slednje lahko povzroči depresijo zaradi inbridinga. Ta zmanjša fenotipsko vrednost za gospodarsko pomembne in funkcionalne lastnostmi (Falconer in Mackay, 1996). Povečanje homozigotnosti pomeni tudi izgubo variabilnosti na mestih za lastnosti, ki so trenutno nevtralne, vendar pa v prihodnosti lahko postanejo gospodarsko pomembne. Inbriding je lahko tudi posledica naključnega genetskega toka, ki lahko poveča frekvenco recesivnih škodljivih alelov in izražanje genetskih defektov (Charlesworth and Charlesworth, 1999). Poleg tega naključni genetski tok povzroči tudi naključne spremembe v vrednosti tako za funkcionalne kot tudi gospodarsko pomembne lastnosti ter razpad ugodnih neaditivnih genetskih interakcij.

Inbriding lahko ocenimo na podlagi rodovnika ali genomskih informacij. Inbriding na podlagi rodovniških podatkov oceni pričakovan inbriding, točnost pa močno zavisi na kakovosti in globini rodovnika. Posledično lahko z rodovniško oceno podcenimo resničen inbriding v populaciji zaradi nepravilnega ali nepopolnega rodovnika ter predpostavke o nesorodni prvi generaciji rodovnika (npr. Lutaaya in sod., 1999; Oliehoek in sod., 2009). Inbriding lahko ocenimo tudi na podlagi genomskih podatkov: i) na podlagi genomskih koeficientov sorodstva; ii) iz deleža genoma, ki ga pokrivajo nizi homozigotnosti (angl. r*uns of homozygosity*, ROH); ali iii) detekcije autozigotnih odsekov oz. delov, ki so homozigotni po izvoru (angl. *homozygostiy by descent,* HBD). Genomski koeficienti inbridinga merijo realiziran inbriding in so bolj točni, saj zajamejo učinek Mendelskega vzorčenja iter niso odvisni od točnosti in globine rodovnika. Prav tako lahko z merjenjem homozigotnosti ali autozigotnosti razločimo med nedavnim in starodavnim inbridingom, pri čemer je nedaven inbriding bolj škodljiv (McQuill in sod., 2008; Ferenčaković in sod, 2012).

Samo spremljanje koeficienta inbridinga *F* v starših in njihovih potomcih ni zadosten kriterij za trajnostno upravljanje s populacijo, saj se vpliv paritvenih odločitev na koeficient inbridinga izrazi z zamikom dveh generaciji, ko se ustvarijo zanke v rodovniku (Woolliams in sod., 2015). Poleg tega je *F* v populaciji relativen na bazno generacijo, za katero predpostavimo, da ni inbridirana in sorodna. Posledično izbira bazne generacije vpliva na ocenjen *F* v populaciji (Woolliams in sod, 1998). Čeprav je inbriding neizbežen v končno velikih populacijah, je pomembno, da uravnavamo stopnjo, s katero koeficient inbridinga narašča (*ΔF*). Stopnja inbridinga določa efektivno velikost populacije (Falconer in Mackay, 1996). V selekciji je upravljanje stopnje inbridinga pomembno, da čimbolj učinkovito uporabimo razpoložljivo genetsko variabilnost za ustvarjanje genetskega napredka oz. maksimiramo genetski napredek ob omejitvi stopnje inbridinga (Woolliams in sod., 2015).

Tudi v populacijah govedi so raziskovalci odkrili naraščanje stopnje inbridinga ter depresijo zaradi inbridinga (npr. Wiggans in sod., 1995; Thompson in sod., 1999; Doekes in sod., 2019). Pri govedi je povečevanje inbridinga neizbežno zaradi večgeneracijske direktne intenzivne selekcije, velike variance reproduktivnega uspeha in napovedi plemenskih vrednosti po metodi BLUP v kombinaciji s selektivno odbiro (angl. *truncation selection*) (povzeto v Howard in sod., 2017). Uvedba genomske selekcije je močno skrajšala generacijske intervale, kar je povečalo letno stopnjo inbridinga (Doekes, 2018; Makanjuola in sod., 2020).

Za učinkovito upravljanje stopnje inbridinga moramo tako nadzorovati prispevke posameznikov preko generacij in varianco prispevkov. Idejo genetskih prispevkov sta leta 1958 razvila James in McBride. Leta 2015 jih je Woolliams tudi matematično formaliziral kot delež genov, ki jih posameznik prispeva v prihodnji genski sklad. Dolgoročni prispevki posameznikov so povezani z genetskim napredkom (Woolliams in sod., 1999):

ΔG = Σ*ri ai*; [6]

kjer je *ri* dolgoročen prispevek posameznika *i*, *ai* pa njegov učinek Mendelskega vzorčenja. Dolgoročni prispevki so na drugi strani povezani tudi s stopnjo inbridinga (Wray in Thompson, 1990):

ΔF = ¼ Σ*ri2.* [7]

Preko genetskih prispevkov lahko tako upravljamo tako z genetskim napredkom kot tudi inbridingom in genetsko variabilnostjo. Ker izboljšanje enega izmed teh ciljev pomeni poslabšanje drugega, je ohranjanje ravnotežja ključnega pomena.

Hkratna optimizacija genetskega napredka in genetske variabilnosti ideja je implementirana v selekciji z optimiziranimi prispevki (angl. o*ptimal contribution selection,* OCS) (Meuwissen 1997; Grundy et al. 1998; Woolliams in sod., 2015). Selekcija z optimiziranimi prispevki je nasprotje selektivni odbiri, ki le maksimira genetski napredek. Ker je koeficient inbridinga posameznika enak povprečnemu koeficientu sorodstva njihovih staršev, lahko stopnjo inbridinga nadzorujemo preko skupinskega sorodstva izbranih staršev (Woolliams in sod., 2015). Skupinsko sorodstvo staršev definiramo kot njihovo povprečno sorodstvo uteženo z njihovimi prispevki k naslednji generaciji. Za samo optimizacijo so potrebne dodatne omejitve, ki izhajajo iz praktičnega vidika parjenja: zagotoviti moramo, da so prispevki vseh posameznikov ne-negativni, da je vsota moških kot tudi ženskih posameznikov ½ ter da je vsota prispevkov vseh posameznikov enaka 1 (Woolliams in sod., 2015).

Na drugi strani lahko genetsko variabilnost določene populacije povečamo z migracijo. V populacijah govedi to vršimo preko uvoza genetskega materiala predvsem očetov.

## 1.4 STOHASTIČNE SIMULACIJE

Selekcija je dolgotrajen in drag postopek z mnogo spremenljivkami, vpliva katerih še ne poznamo. Slednje onemogočajo deterministične izračune genetskih parametrov, kot so genetski napredek in genetska variabilnost. Testiranje različnih modelov v praksi je zaradi dolgih generacijskih intervalov, števila živali in potrebnih infrastrukturnih kapacitet nemogoče. Računalniške simulacije nam omogočajo hitro, poceni in ponovljeno testiranje rejskih programov, statističnih metod, predpostavk kvantitativnih modelov ter analizo kratko-, srednje- in dolgoročnih posledic selekcije. Z računalniškimi simulacijami lahko poustvarimo tako naključne kot sistematične procese – od kreiranja nove generacije posameznikov, do selekcije in parjenja izbranih staršev.

### 1.4.1 Simulacija genoma

Simulacija genomov in vzročnih mehanizmov je izjemno kompleksna zaradi različnih virov variabilnosti, različnih možnih demografskih zgodovin populacije ter negotovosti glede parametrov genoma (Daetwyler in sod., 2013). Za samo simulacijo moramo najprej oceniti genetske parametre kot so velikost in število kromosomov, zgodovinska efektivna velikost populacije, vzorec vezavnega neravnovesja, stopnje mutacij ter stopnje rekombinacij, ki se pri različnih vrstah oz. populacijah razlikujejo. Na podlagi ocenjenih parametrov simuliramo genotipe oz. haplotipe populacije osnovalcev. V praksi so se uveljavile tri glavne metode simulacije genomov. Prva je metoda na podlagi ponovnega vzorčenja (angl. *methods based on resampling*), ki uporabi resnične podatke genomskega zaporedja ter rodovnik za simulacijo genomov trenutne populacije. Ta metoda najbolje posnema resnične frekvence alelov in vezavno neravnovesje, ne omogoča pa vpeljave novih mutacij ali genetskih značilnosti. Druga je metoda na podlagi koalescentne teorije, s katero simuliramo genom nazaj v čas. Metoda identificira in združi (angl. *coalesce*) skupne prednike danega vzorca neznanih genotipov s stohastičnim procesom, pri čemer upošteva mutacijo, rekombinacijo in migracijo. Tretja metoda je metoda »naprej v čas«, ki je tudi najlažja za implementacijo in najbolj fleksibilna. Ta metoda ustvarja populacijo s časom glede na navedene genetske in demografske parametre.

### 1.4.2 Simulacija fenotipa

Za simulacijo genetskih in fenotipskih vrednosti moramo simulirati delovanje genov. Izmed simuliranih lokusov izberemo množico z učinkom, učinke pa vzorčimo iz izbrane porazdelitve. Porazdelitev učinkov je lahko enostavna, npr. normalna, lahko pa je zelo kompleksna ter vključuje nelinearne učinke. Naprednejši simulatorji omogočajo simulacijo enostavnih lastnosti z aditivnimi učinki kot tudi lastnosti z učinki dominance, epistaze in interakcije genotip-okolje (Gaynor in sod., 2020). Na podlagi učinkov genov izračunamo pravo aditivno genetsko oz. plemenso vrednost posameznikov, fenotipske vrednosti pa simuliramo z dodatkom naključnega odklona za ostanek h genetski vrednosti. Odkloni ostankov so vzorčeni iz normalne porazdelitve s povprečjem vplivov 0 in varianco, ki je enaka varianci za ostanek. Varianca za ostanek je določena na način, da se doseže nastavljena heritabiliteta za lastnost.

### 1.4.3 Simulacija populacije

Iz osnovnega nabora haplotipov s ponavljanjem vzorčimo haplotipe oz. genotipe za prvo generacijo populacije. Temu sledi odbira staršev in parjenje za kreiranje nove generacije posameznikov ter njihovih genomov, pri čemer upoštevamo naključna procesa rekombinacije in segregacije. Začetno velikost in strukturo populacije vzpostavimo s t.i. izgorevanjem (angl. *burn-in*), kjer določeno število generacij starše izbiramo in parimo naključno. Temu sledi simulacija selekcije z načrtno odbiro staršev. Kriterij za odbiro staršev je lahko povprečje staršev, klasična plemenska vrednost, genomska plemenska vrednosti, prava genetska vrednost ali fenotipska vrednost. (Faux in sod., 2016).

Simulacijski programi omogočajo simulacijo rejskih programov različnih kompleksnosti. Simuliramo lahko enostavne rejske programe z diskretnimi generacijami in malim številom rejskih odločitev (npr. samo odbira staršev iz zadnje generacije). Lahko pa simuliramo kompleksne rejske programe s prekrivajočimi generacijami, velikim številom rejskih odločitev, večstopenjskim testiranjem živali itd. V simulaciji lahko uporabimo tudi selekcijske parametre povzete po dotičnem rejskem programu, kot npr. število novorojenih živali, število živali v testu potomcev, število odbranih bikov itd. To nam omogoča simulacijo specifične populacije s pripadajočimi genetskimi parametri in testiranje hipotez na sistemu, ki je primerljiv z našim tarčnim resničnim sistemom.

## 1.5 MAJHNE POPULACIJE

Majhne populacije so zaznamovane z različnimi genetskimi, selekcijskimi in praktičnimi posebnostmi ter omejitvami. Te so predstavljale izziv že v klasični selekciji, v genomski selekciji pa so bolj izrazite ter hitro povečujejo razkol med majhnimi in večjimi populacijami. Čeprav je genomska selekcija že nekaj let uveljavljena tehnologija v večjih populacijah, se majhne populacije še vedno spopadajo z izzivi tako pri inicializaciji kot tudi kontinuiranem izvajanju genomske selekcije. Omejitve majhnih populacij se izražajo v doseganju genetskega napredka in ohranjanju genetske variabilnosti, izhajajo pa iz pomanjkanja tako živali, trajnostne strategije kot tudi finančnih sredstev. Manjše število živali genetski napredek omejuje na dva načina. Prvič, majhno število živali pomeni manjše število testiranih selekcijskih kandidatov in posledično manjšo intenzivnost selekcije. Slednje se je še posebno izrazilo z uvedbo genomske selekcije, saj ta omogoča cenejše in hitrejše testiranje velikega števila selekcijskih kandidatov. Drugič, v klasični selekciji majhne populacije dosegajo manjše točnosti testa potomcev zaradi manjšega števila potomcev s fenotipskimi vrednosti. V genomski selekciji imajo majhne populacije težave s sestavo zadostne referenčne populacije, ki zahteva posameznike z visokimi točnostmi lastnih plemenskih vrednosti ali pa veliko število posameznikov. Nezadostna referenčna populacija vodi do manjše točnosti napovedi, kar omejuje prednost genomske pred klasično selekcijo (Thomasen in sod., 2014).

Težavo nezadostne referenčne populacije lahko rešujemo s sodelovanjem v mednarodnih konzorcij in uporabo mednarodne enopasemske referenčne populacije (Jorjani, 2012; Špehar in sod., 2013; Schöpke in Swalve, 2016) ali pa uporabo večpasemske referenčne populacije. Uspeh slednjih zavisi od genetske sorodnosti populacij ali pasem. Čeprav oba pristopa povečata točnost genomske napovedi v majhnih populacijah, je združevanje znotraj pasme bolj učinkovito kot združevanj pasem (Hozé in sod., 2014; Lund in sod., 2014). Alternativa referenčni populaciji progeno testiranih bikov je tudi referenčna populacija ženskih živali. Zaradi manjše količine fenotipskih podatkov na kravo v primerjavi s progeno testiranimi biki je relativen prispevek ene krave k referenčni populaciji manjši kot enega progeno testiranega bika (de Roos, 2011). Vendar so raziskave pokazale, da lahko z dodatkom ženskih podatkov v referenčno populacijo majhne populacije postanejo konkurenčne (Thomasen in sod., 2014). V primerjavi s samo moško referenčno populacijo dodatek ženski poveča točnost napovedi ter zmanjša stopnjo inbridinga. Raziskave so pokazale tudi, da referenčna populacija hčera s fenotipskimi podatki omogoča večjo točnost napovedi kot referenčna populacija njihovih očetov (Buch in sod., 2012) ter da lahko točnost genomske napovedi povečamo s genotipizacijo živali z lastnimi fenotipskimi vrednostmi (van Grevenhof in sod., 2012).

Rejski programi v majhnih populacijah morajo posvečati posebno pozornost tudi ohranjanju genetske variabilnosti. Ta zagotavlja dolgoročen genetski napredek, kar omogoča konkurenčnost in upravičuje obstoj nacionalnih rejskih programov. V majhnih populacijah je učinek naključnega genetskega toka močnejši in ima za posledico naključno nihanje frekvenc alelov. To nihanje lahko privede do naključne fiksacije ali izgube alela ter povečane homozigotnosti in inbridinga (Falconer in Mackay. 1996).

Majhne populacije lahko tudi vnašajo genetsko variabilnost preko uvoza genetskega materiala, s čimer zmanjšujejo stopnjo inbridinga. V rejskih programih za govedi populacije uvažajo predvsem očete. Uvoz povečuje sorodnost znotraj pasme oz. globalne populacije govedi, vendar pa znižuje sorodnost znotraj majhnih populacij. Majhne populacije uvažajo očete tudi za večanje genetskega trenda, saj so ti genetsko superiorni zaradi bolj učinkovite selekcije v večjih populacijah. Vendar pa je uspeh uvoza odvisen od prisotnosti interakcije med genotipom in okoljem, ki lahko povzroči prerangiranje živali med okolji. Interakcija med genotipom in okoljem nakazuje, da se genotipi razlikujejo v genetskem potencialu za prilagajanje na spremembe v okolju (Falconer, 1952). Ponavadi jo merimo kot genetsko korelacijo med lastnostjo v različnih okoljih, pri čemer korelacija manj kot ena nakazuje prisotnost interakcije genotip-okolje (Falconer in Mackay, 1996). Predhodne študije so pokazale, da je za maksimiranje genetskega napredka v klasični selekcije potrebna korelacija med okolji večja od 0.7 oz. 0.8 (Smith in Banos, 1991), za dolgoročno sodelovanje pa več kot 0.9 (Cao in sod., 2020). Študije so tudi pokazale, da je za dolgoročno sodelovanje v genomski selekciji potrebna nekoliko nižja korelacija kot v klasični selekciji, in sicer med 0.85 in 0.875 (Cao in sod., 2020). Čeprav je uvoz koristna in nujna strategija za majhne, z večanjem odstotka uvoza manjšamo pomen nacionalne selekcije.

Majhne populacije morajo tako oblikovati trajnostno strategijo, ki bo omogočila i) učinkovito doseganje genetskega napredka za konkurenčnost na mednarodnem trgu; ii) učinkovito upravljanje z genetsko variabilnostjo, ki bo omogočala tako kratko- kot tudi dolgoročni genetski napredek; iii) ekonomsko učinkovitost, iv) učinkovito ohranjanje lokalnega pomena selekcije.

Ker si omenjeni cilji nasprotujejo, je njihovo uravnoteženje težavno za vsak rejski program, še posebno pa v majhnih populacijah. Strategije, ki bi maksimirale genetski napredek pogosto niso optimalne z vidika ohranjanja genetske variabilnosti ali pa ekonomske učinkovitost. Na primer, da bi ohranile konkurenčnost na mednarodnem trgu, morajo majhne populacije najprej nadoknaditi zaostanek v genetskem napredku. Kratkoročno je to najlažje doseči s povečanjem intenzivnost selekcije, ki pa je z vidika genetske variabilnosti problematična zaradi že tako majhnega števila razpoložljivih staršev. Prav tako bi genetski napredek lahko maksimirali z uvedbo genomske selekcije z veliko referenčno populacijo, vendar to z ekonomskega vidika ni najbolj učinkovita ali pa celo možna strategija. Zaradi pomanjkanja infrastrukture in opreme morda z ekonomskega vidika tudi ni optimalno spremljanje vseh lastnosti in rodovnika (Biscarinni in sod., 2015).

## 1.6 raziskovalni cilji in hipoteze

### 1.6.1 Raziskovalni cilji

Splošen cilj doktorske naloge je bil predstaviti učinkovito rabo genomske informacije v majhnih rejskih programih za govedo. Za svoje delo smo si zadali naslednje specifične cilje:

* razvoj orodja za stohastično simulacijo specifičnega rejskega programa za govedo s prekrivajočimi se generacijami;
* parametrična analiza slovenskih rejskih programov za mlečne pasme govedi;
* priprava izvedljivih genomskih scenarijev, ki so alternativa klasičnemu selekcijskemu scenariju v majhnih populacijah;
* primerjava genetskih trendov in dolgoročne učinkovitosti klasičnega in različnih genomskih selekcijskih scenarijev;
* analiza razporeditve finančnih sredstev med selekcijskimi aktivnostmi v obstoječih slovenskih rejskih programih;
* analiza učinkovitosti razporeditve finančnih sredstev v predlaganih genomskih selekcijskih scenarijih v majhnih populacijah govedi na primeru slovenske populacije;
* primerjava genetskih trendov in dolgoročne učinkovitosti rejskih programov z različnim odstotkom uvoza;
* analiza prispevkov nacionalne selekcije in uvoza h genetskemu trendu v testiranih scenarijih.

### 1.6.2 Raziskovalne hipoteze

V svojih raziskavah smo preverjali naslednje hipoteze:

* genomski scenariji omogočijo večji genetski napredek kot klasični scenarij;
* genomski scenariji porabijo več genetske variance kot klasični scenarij;
* način uporabe genomske informacije ne vpliva na dolgoročno učinkovitost selekcije;
* prerazporeditev finančnih sredstev v selekcijski shemi lahko poveča genetski napredek;
* prispevek nacionalne selekcije h genetskemu napredku je odvisen od interakcije genotip-okolje;
* prispevek nacionalne selekcije h genetskemu napredku pada z večanjem odstotka uvoza.

# RAZPRAVA IN SKLEPI

## RAZPRAVA

Majhne populacije so pri izvajanju selekcije omejene zaradi genetskih in praktičnih posebnosti. Zaradi majhnega števila živali ter omejenih finančnih sredstev za testiranje selekcijskih kandidatov ne morejo izvajati zelo intenzivne selekcije. Prav tako majhno število živali in potomcev po očetu omejuje točnost tako klasične kot tudi genomske selekcije. Omejitve glede števila živali in točnosti vplivajo tudi na genetsko variabilnost v populaciji. Navkljub vsem posebnostim morajo majhne populacije zagotavljati tako kratko- kot dolgoročni genetski napredek, da ostanejo konkurenčne na mednarodnem trgu in upravičijo obstoj domače selekcije.

Genomska selekcija je uveljavljena tehnologija v velikih populacijah. Nasprotno se veliko majhnih populacij še vedno spopada z njeno uvedbo, predvsem zaradi pomanjkanja finančnih sredstev, skepticizma uporabnikov in pomanjkanja strokovne podpore. Zamuda v uvedbi genomske selekcije povečuje razkol med velikimi in majhnimi populacijami. Zaradi razlike v genetskem napredku majhne populacije tako povečujejo uvoz tujega genetskega materiala in s tem zmanjšujejo pomen domače selekcije.

Klasična in genomska selekcija se razlikujeta tudi v načinu doseganja genetskega napredka in financiranja (prioritet) rejskih akcij. Klasični program implementira progeni test. Ta zaradi visokih stroškov omejuje intenzivnost selekcije in ima za posledico dolg generacijski interval, hkrati pa omogoča visoko točnost odbire. Slednja je tako glavno gonilo napredka v klasični selekciji. Rejski programi s klasično selekcijo večino finančnih sredstev tako namenijo zbiranju fenotipskih vrednosti, ki je nerazdružljivo povezano s progenim testom. Genomska selekcija pa dosega genetski napredek predvsem preko močno skrajšanega generacijskega intervala, večje intenzivnosti selekcije moških živali in večje točnosti odbire mladi živali brez lastnih fenotipskih vrednosti. Zbiranje fenotipskih podatkov je še vedno ključnega pomena za obnavljanje referenčne populacije, vendar pa ni več intrinzično povezano s selekcijo. Genomska selekcija mora tako uravnotežiti investiranje v zbiranje fenotipskih podatkov in genotipizacijo živali za obnavljanje referenčne populacije in selekcijo. Razlika v investicijskih prioritetah med obema programoma predstavlja problem za prehod na genomsko selekcijo, še posebno v rejskih programih z omejenimi sredstvi ali pa v državah, kjer so različne rejske aktivnosti financirane iz različnih virov.

Majhne populacije morajo tako stremeti k visoki učinkovitosti selekcije v vseh pogledih. V nadaljevanju smo tako razpravo razdelili na: i) učinkovitost v doseganja genetskega napredka; ii) učinkovitost ohranjanja genetske variabilnosti; in iii) ekonomske učinkovitosti.

Genetski napredek

Da bi ohranile konkurenčnost na mednarodnem trgu, morajo majhne populacije dosegati genetski napredek primerljiv z velikimi populacijami oz. globalnim genetskim napredkom. Če hočejo ujeti genetski napredek velikih populacij, morajo zaostriti enega od parametrov, ki povečujejo genetski napredek: intenzivnost, točnost ali pa dodatno skrajšati generacijski interval. V naši raziskavi (poglavje X.x) smo s stohastično simulacijo testirali vpliv števila odbranih očetov in leta uporabe na genetski napredek.