

## صحت پیش‌بینی‌های ژنومی برای صفات مرتبط با شیر در گاو نژاد نجدی

- سید مهدی حسینی وردنجانی  
دانشجوی دکتری ژنتیک و اصلاح نژاد دام دانشگاه فردوسی مشهد
- محمد مهدی شریعتی (نویسنده مسئول)  
استادیار گروه علوم دامی دانشگاه فردوسی مشهد.
- حسین مرادی شهر بابک  
استادیار گروه علوم دامی پردیس کشاورزی و منابع طبیعی کرج، دانشگاه تهران
- مجتبی طهمورث پور  
استاد گروه علوم دامی دانشگاه فردوسی مشهد

تاریخ دریافت: بهمن ۱۳۹۶ تاریخ پذیرش: اردیبهشت ۱۳۹۷

شماره تماس نویسنده مسئول: ۰۹۱۳۳۹۱۵۳۰۶

Email: mm.shariati@um.ac.ir

شناسه دیجیتال (DOI): 10.22092/asj.2018.115893.1549

### چکیده

هدف از این پژوهش، ارزیابی عملکرد انتخاب ژنومی برای صفات تولید شیر، درصد چربی و درصد پروتئین در گاو نژاد نجدی در گله‌های ایستگاهی و اقماری با استفاده از مدل‌های آماری مختلف بود. از ارزش‌های اصلاحی سنتی بدست آمده از یک مدل رگرسیون تصادفی با استفاده از اطلاعات شجره‌ای و فنوتیپی هر صفت بین سال‌های ۱۳۶۹ تا ۱۳۹۵ به عنوان متغیر پاسخ استفاده شد. با استفاده از طرح ۱۰ بار تکرار جمعیت آموزش-آزمون و چهار مدل بهترین پیش‌بینی نااریب خطی ژنومی مقیاس شده با فراوانی آلی مشاهده شده (GBLUP) و فراوانی آلی ۰/۵ (G05BLUP)، بیز A و بیز B قابلیت پیش‌بینی‌ها، ارزیابی شدند. نتایج نشان داد، GBLUP عملکرد بهتری نسبت به G05BLUP برای تولید شیر (۰/۴۱۱ در مقابل ۰/۳۸۵) داشت ولی عملکرد G05BLUP برای درصد چربی (۰/۲۵۷ در مقابل ۰/۳۰۲) و درصد پروتئین (۰/۳۶۳ در مقابل ۰/۳۸۸) بهتر بود. صحت برآورد ارزش اصلاحی تولید شیر و درصد چربی با استفاده از بیز A و بیز B به ترتیب به ۰/۳۷۱ و ۰/۳۵۳ کاهش و ۰/۳۲۹ و ۰/۳۱۴ افزایش یافتند. برای درصد پروتئین روش‌های بیزی و GBLUPs صحت مشابه داشتند. در بین تمام روش‌ها و صفات، بیز A برای پروتئین با ۰/۱۴ و G0BLUP برای تولید شیر با ۰/۷۱ به ترتیب کمترین و بیشترین اریب پیش‌بینی بصورت انحراف از یک را داشتند. صحت پیش‌بینی‌ها با استفاده از گله‌های اقماری علاوه بر گله ایستگاهی بین ۰/۰۱ تا ۰/۰۹ بسته به روش و صفت افزایش یافت ولی اریب نیز از قبل بیشتر بود. در نتیجه، صحت پیش‌بینی ژنومی برای صفات تولید شیر در گاو نجدی متوسط، ولی با توجه به اندازه کوچک جمعیت مناسب هستند که کاربرد انتخاب ژنومی برای این نژاد را ممکن می‌سازد.

Animal Science Journal (Pajouhesh & Sazandegi) No 122 pp: 93-104

### The accuracy of genomic predictions for milk related traits in Najdi cattle breed.

By: 1: Sayed Mahdi Hosseini-Vardanjani , PhD student of animal breeding and genetics, Department of Animal Science, Ferdowsi University of Mashhad, Iran

2\*: Mohammad Mahdi Shariati , Assistant Professor of animal breeding and genetics, Department of Animal Science, Ferdowsi University of Mashhad, Iran

3: Hossein Moradi Shahrebabak, Assistant Professor of animal breeding and genetics, Department of Animal Science, University College of Agriculture and Natural Resources, University of Tehran, Iran

4: Mojtaba Tahmoorespur , Professor of animal breeding and genetics, Department of Animal Science, University of Mashhad, Iran

**Received: January 2018**

**Accepted: May 2018**

The objective of this study was to evaluate the performance of genomic selection for milk yield, fat percent and protein percent in Najdi cattle breed from station and Agmari herds using different statistical models. Traditional estimated breeding values obtained with a random regression model using pedigree and phenotypic information for each trait from 1990 to 2016 were used as response variables. Predictability was evaluated by using a 10 replicated Training-Testing scheme and considering four models including genomic best linear unbiased prediction scaled with observed allele frequency (GBLUP), and allele frequency of 0.5 (G05BLUP), BayesA and BayesB. The results showed that GBLUP had better performance than G05BLUP for milk yield (0.411 vs 0.385), but performance of G05BLUP was better for fat percent (0.257 vs 0.302) and protein percent (0.363 vs 0.388). The accuracy of breeding values of milk yield decrease to 0.371 and 0.353, and accuracy of fat percent increased to 0.329 and 0.314 with BayesA and BayesB, respectively. Bayesian methods had same accuracy as GBLUPs for protein percent. Across traits and methods BayesA with 0.14 for protein percent, and G05BLUP with 0.71 for milk yield had smallest and highest bias as deviate from 1, respectively. Accuracy of prediction using Agmari herds in addition of station herd increased 0.01 to 0.09 depending on method and trait, but biases also were more than before. In conclusion, accuracy of genomic prediction of milk traits in Najdi cattle breed are moderate but suitable considering small size of population which makes genomic selection feasible in this breed.

**Key words:** Genomic selection, statistical model, variable selection, genetic markers, accuracy

#### مقدمه

کاهش فاصله نسلی با انتخاب در هنگام تولد تسریع می‌نماید (Schaeffer, ۲۰۰۶). علاوه بر این، هزینه‌های تعیین ژنوتیپ نسبت به هزینه‌های آزمون نتاج و همچنین ارزش اقتصادی حیوانات نیز نسبتاً کم است (Schaeffer, ۲۰۰۶). در کنار این عوامل، در سال‌های اخیر انواع تراشه‌های نشانگری برای تعیین ژنوتیپ حیوانات برای چندین هزار نشانگر با قیمت نسبتاً پایین در دسترس قرار گرفته‌اند. در نتیجه، انتخاب ژنومی به عنوان یک فناوری جدید با استفاده از اطلاعات نشانگرهای مترکم توزیع

انتخاب ژنومی<sup>۱</sup> (GS) با استفاده از نشانگرهای مولکولی و داده‌های فنوتیپی، چشم انداز جدیدی را در برنامه‌های اصلاح نژادی حیوانات مزرعه به ویژه گاو شیری که در آن انتخاب گاوهای نر برتر برای صفات تولید شیر مشروط به طی شدن دوره‌ی آزمون نتاج<sup>۲</sup> است باز کرده است. انتخاب ژنومی پیشرفت ژنتیکی در گاو شیری را از طریق افزایش صحت برآورد ارزش اصلاحی گاوهای نر در هنگام تولد، افزایش شدت انتخاب و

<sup>1</sup> Genomic selection

<sup>2</sup> Progeny test

در جنوب و جنوب غرب کشور است. این نژاد یک نژاد شیری-گوشتی بوده و به عنوان یک منبع خوب تولید شیر و گوشت نقش مؤثری در تأمین پروتئین حیوانی و اقتصاد روستایی خوزستان ایفا می‌کند و از این رو برای نگهداری و حفظ آن تلاش می‌شود. تحمل درجه حرارت بالا، گوساله‌زایی خوب، آسان‌زایی و کم بودن موارد ابتلا به بیماری‌های دستگاه تناسلی، قدرت تحرک زیاد، مقاوم بودن در برابر ابتلا به انگل‌های داخلی و خارجی، سازش با شرایط آب و هوایی و تغذیه‌ای نامطلوب از خصوصیات بارز گاو نجدی است (نظری و همکاران، ۱۳۸۶). علاوه بر این خصوصیات، می‌توان با طرح‌ریزی و اجرای برنامه‌های اصلاح نژادی، صفات مرتبط با تولید شیر این نژاد را نیز افزایش داد. بنابراین، هدف از مطالعه حاضر بررسی امکان طرح‌ریزی انتخاب ژنومی در گاو نژاد نجدی برای صفات تولید شیر، درصد چربی و درصد پروتئین با استفاده از روش‌های مختلف پیش‌بینی ژنومی بود.

## ۱- مواد و روش‌ها

### ۱-۱- داده‌های ژنوتیپی

تعداد ۲۱۱ رأس گاو نجدی از هر دو جنس نر و ماده از ایستگاه پرورش و اصلاح نژاد گاو نجدی و گله‌های اقماری این نژاد در خوزستان (به ترتیب هر کدام ۹۸ و ۱۱۳ حیوان) خون‌گیری و سپس تعیین ژنوتیپ شدند. برای تعیین ژنوتیپ حیوانات از تراشه با تراکم ۴۰k استفاده شد که در کل منجر به خوانش تعداد ۳۰۱۰۶ نشانگر چندشکلی تک‌نوکلئوتیدی شد. مراحل کنترل کیفیت داده‌های ژنوتیپی با استفاده از نرم افزار PLINK نسخه ۰.۷.۱ (Purcell و همکاران، ۲۰۰۷) و با توجه به معیارهای فراوانی آلل نادر (۰/۰۱)، نرخ خوانش نشانگر (۰/۲)، و انحراف از تعادل هاردی واینبرگ ( $10^{-5}$ ) ویرایش شدند (Su و همکاران، ۲۰۱۲). همچنین، نشانگرهای روی کروموزوم‌های جنسی و نشانگرهای با موقعیت ناشناخته بر روی ژنوم نیز حذف شدند. ژنوتیپ‌های گم‌شده با استفاده از نرم افزار Beagle نسخه ۲.۳.۳ (Browning and Browning، ۲۰۰۹) به صورت کروموزوم به کروموزوم ایمپوت شدند. خلاصه آماری مراحل کنترل کیفیت

شده در سطح ژنوم می‌تواند برای اصلاح نژاد گاو شیری به طور فوری سودمند باشد و از اینرو در حال حاضر بسیاری از کشورهای دنیا از این روش برای اصلاح نژاد حیوانات مزرعه استفاده می‌کنند (Hoze و همکاران، ۲۰۱۴).

یک عامل کلیدی برای انتخاب ژنومی موفق توانایی پیش‌بینی ارزش‌های اصلاحی ژنومی<sup>۳</sup> با صحت زیاد است که به عواملی مختلف بستگی دارد. از مهمترین این عوامل می‌توان به تعداد حیوانات تعیین ژنوتیپ شده در جمعیت مرجع که برای برآورد اثرات نشانگرها استفاده می‌شوند و روش آماری مورد استفاده اشاره کرد (Lourenco و همکاران، ۲۰۱۴). تقریباً هیچ راه حلی در رابطه با افزایش تعداد حیوانات جمعیت مرجع در مورد نژادهای بومی نظیر نژادهای بومی موجود در کشور وجود ندارد. البته استفاده از یک جمعیت مرجع چند نژادی شامل حیوانات تعیین ژنوتیپ شده از یک نژاد مشابه در کشورهای مختلف و یا چند نژاد در یک کشور، به عنوان یک پیشنهاد برای افزایش اندازه جمعیت مرجع مطرح شده است ولی در کل نتایج همواره امیدوار کننده نبوده‌اند (Lund و همکاران، ۲۰۱۴). اما در رابطه با روش آماری مورد استفاده، یک مدل آماری مناسب امکان افزایش صحت انتخاب را به وجود خواهد آورد. روش‌های آماری متعددی برای برآورد اثرات نشانگرها در جمعیت مرجع ارائه شده است که در کل، گروهی ارزش اصلاحی حیوانات و گروه دیگر اثر نشانگرهای ژنتیکی را به عنوان متغیر تصادفی در نظر می‌گیرند. در روش بهترین پیش‌بینی ناریب خطی ژنومی، برای تمام نشانگرها یک توزیع پیشین نرمال با یک واریانس ثابت فرض می‌گردد (Nejati-Javaremi و همکاران، ۱۹۹۷؛ VanRaden و همکاران، ۲۰۰۸). در حالی که در روش‌های مبتنی بر برآورد اثر نشانگر مانند انواع روش‌های بیز با انتخاب یک توزیع پیشین انعطاف پذیرتر امکان وجود تعداد کمی نشانگر با اثر زیاد و در مقابل تعداد زیادی نشانگر با اثر کوچک و یا حتی صفر فراهم می‌شود (Habier و همکاران، ۲۰۱۱).

گاو نجدی شاخص‌ترین توده جمعیتی گاو در خوزستان بوده که از سوی فائو نیز به همین نام شناخته شده است. زیستگاه این نژاد

<sup>3</sup> Genomic breeding values

برای بررسی طبقه‌بندی جمعیت یک تجزیه و تحلیل مؤلفه‌های اصلی<sup>۴</sup> (PCA) بر روی ماتریس روابط خویشاوندی ژنومی (G) ساخته شده بر اساس روشی که در ادامه خواهد آمد نیز انجام شد.

و ژنوتیپ‌های حذف شده و باقیمانده به تفکیک در جدول یک آورده شده است. از آنجایی که حیوانات از دو منبع یعنی گله موجود در ایستگاه پرورش و اصلاح نژاد گاو نجدی و همچنین گله‌های اقماری این نژاد در خوزستان نمونه‌گیری شده بودند، پس

جدول ۱- تعداد نشانگر حذف شده بعد از هر مرحله کنترل کیفیت داده‌های ژنوتیپی به تفکیک گله

معیار	موقعیت ناشناخته	کروموزوم‌های جنسی	نرخ خوانش نشانگر	MAF	HW-D	نشانگرهای باقیمانده
ایستگاهی	۲۸۶	۱۵۳۴	۲۹۸۳	۸۲۷	۴۵۱	۲۴۲۳۷
ایستگاهی+اقماری	۲۸۶	۱۵۳۴	۲۸۷۳	۶۷۷	۱۹۹۹	۲۲۹۶۹

MAF: حداقل فراوانی آللی، HW-D: عدم تعادل هاردی-واینبرگ

## ۲-۱- داده‌های فنوتیپی

در این مطالعه از اطلاعات شجره‌ای و رکوردهای مربوط به سه صفت شیر تولیدی، درصد چربی شیر و درصد پروتئین شیر که طی سال‌های ۱۳۶۹ تا ۱۳۹۵ در ایستگاه پرورش و اصلاح نژاد گاو نجدی واقع در خوزستان جمع آوری شده بودند، استفاده شد. مجموعه اطلاعات شامل شماره حیوان، پدر و مادر حیوان، سال زایش، جنس گوساله، تیپ تولد، سن مادر هنگام زایش و رکوردهای صفات ذکر شده در ماه‌های مختلف بودند. ابتدا داده‌ها با استفاده از نرم افزار Fox Pro 2.6 و Excel 2010 ویرایش و برای ادامه آنالیز آماده‌سازی شدند. به منظور شناسایی اثر عوامل ثابت مؤثر بر صفات مورد بررسی و بنابراین منظور کردن آنها در مدل، آنالیز کمترین مربعات با استفاده از رویه GLM در نرم افزار SAS انجام شد. مؤلفه‌های واریانس- کوواریانس برای صفات مورد بررسی با استفاده از یک مدل رگرسیون تصادفی و روش حداکثر درست‌نمایی محدود شده در نرم افزار DMU برآورد شدند (Madsen و همکاران، ۲۰۰۶). معیار همگرایی برای برآورد تمام پارامترها شامل مؤلفه‌های واریانس- کوواریانس و ارزش‌های اصلاحی در آنالیزها<sup>۱۰</sup> در نظر گرفته شد. در نهایت ارزش‌های اصلاحی سنتی بدست آمده به عنوان متغیر پاسخ در آنالیزهای ژنومی بعدی استفاده شدند (Hayes و همکاران، ۲۰۰۹؛ Forni و همکاران، ۲۰۱۱).

## ۳-۱- روش‌های آماری

ارزش‌های اصلاحی ژنومی برای سه صفت مورد بررسی با استفاده از چهار مدل مختلف شامل GBLUP با استفاده از دو مقیاس مختلف برای تشکیل ماتریس روابط خویشاوندی ژنومی (GBLUP و G05BLUP)، بیز A و بیز B پیش‌بینی شدند. بهترین پیش‌بینی نااریب خطی ژنومی برای پیش‌بینی مستقیم ارزش‌های اصلاحی ژنومی با استفاده از مدل زیر انجام شد:

$$y = 1\mu + Sp + Zg + e$$

در این مدل  $y$  بردار متغیر پاسخ،  $\mu$  میانگین،  $p$  بردار متغیر کمکی مؤلفه اصلی اول،  $g$  بردار ارزش‌های اصلاحی ژنومی و  $e$  نیز بردار باقیمانده‌ها هستند. ماتریس طرح  $Z$  رکوردها را به  $g$  ارتباط می‌دهد،  $S$  حاوی بردار ویژه مؤلفه اصلی اول و  $1$  نیز برداری با عنصرهای یک برای اثر میانگین است. برای اثرات ژنتیکی و محیطی و میانگین به ترتیب توزیع پیشین  $g \sim N(0, G\sigma_g^2)$ ،  $e \sim N(0, I\sigma_e^2)$  و  $\mu \sim Flat(0, \sigma_\mu^2)$  در نظر گرفته شد که ماتریس  $G$  ماتریس روابط خویشاوندی ژنومی،  $\sigma_g^2$ ،  $\sigma_e^2$  و  $\sigma_\mu^2$  واریانس‌های ژنتیکی افزایشی، باقیمانده و میانگین هستند. ماتریس روابط خویشاوندی ژنومی بر اساس VanRaden و همکاران (۲۰۰۸) به صورت  $G = \frac{ZZ^T}{2\sum p_i(1-p_i)}$  تعریف شد که  $p_i$  فراوانی مشاهده شده آلل دوم در جایگاه  $i$  است. عنصرهای ماتریس  $Z_{n \times m}$  نیز به ترتیب  $0 - 2p_i$ ،  $1 - 2p_i$  و

آنها به عنوان متغیر پاسخ بود. همچنین برای بررسی اثر افزایش تعداد حیوانات بر قابلیت پیش‌بینی مدل‌ها از مجموعه دومی از داده‌ها شامل مجموع داده‌های ایستگاهی و گله‌های اقماری استفاده شد که متغیر پاسخ برای گله‌های اقماری ارزش پیش‌بینی شده برای آنها بر اساس بهترین مدل پیش‌بینی نااریب خطی ژنومی در مرحله قبل بود.

قابلیت پیش‌بینی مدل‌ها با استفاده از روش تکرار جمعیت به گروه آموزش و گروه آزمون<sup>۴</sup> و محاسبه همبستگی بین مقادیر ارزش‌های اصلاحی ژنومی پیش‌بینی شده و ارزش اصلاحی سنتی در گروه آزمون ارزیابی شد. به این شکل که در هر تکرار به طور تصادفی ۸۰ درصد افراد کل جمعیت به عنوان گروه آموزش و ۲۰ درصد باقیمانده به عنوان گروه آزمون استفاده شدند و این رویه ۱۰ بار تکرار شد که در نهایت منجر به ۱۰ برآورد از صحت پیش‌بینی برای هر مدل می‌شد. این روش به عنوان یک روش قوی‌تر از روش اعتبارسنجی متقابل<sup>۵</sup> مخصوصاً زمانی که اندازه جمعیت کوچک باشد و نیاز به تعداد بیشتری برآورد صحت پیش‌بینی باشد پیشنهاد شده است (de Los Campos, ۲۰۱۲). همچنین برای ارزیابی اریب پیش‌بینی‌ها هم از تابعیت ارزش اصلاحی سنتی بر ارزش اصلاحی ژنومی پیش‌بینی شده استفاده شد. در این شیوه ضریب رگرسیون برابر با ۱ نشان دهنده عدم اریبی پیش‌بینی‌ها و انحراف از ۱ به عنوان اریبی مدل شناخته می‌شود. آنالیزها با استفاده از نرم افزار BGLR نسخه ۰.۵ انجام شد (Pérez and de Los Campos, ۲۰۱۴). در هر اجرا یک زنجیره به طول ۵۰۰۰۰ نمونه گیبس در نظر گرفته شده که ۱۰۰۰۰ دور اول به عنوان دوره‌های سوخته<sup>۶</sup> حذف شدند و پس از آن هر پنجمین نمونه برای استنباط پارامترها گرفته شدند.

## ۲- نتایج و بحث

یکی از نگرانی‌های عمده در تمام روش‌های پیش‌بینی ژنومی مشکل جدا کردن طبقه‌بندی جمعیت از اثرات واقعی QTL است. به طور معمول برای منظور کردن طبقه‌بندی جمعیت در

$2 - 2p_i$  برای ژنوتیپ هموزیگوس  $aa$ ، هتروزیگوس  $Aa$  و  $AA$  هموزیگوس بودند. همچنین،  $n$  تعداد حیوانات و  $m$  تعداد نشانگرها هستند. در این بررسی نوع دیگری از ماتریس روابط خویشاوندی ژنومی نیز ساخته و در آنالیزهای بعدی از آن استفاده شد که به جای استفاده از فراوانی آللی مشاهده شده در حیوانات تعیین ژنوتیپ شده از فراوانی آللی  $0.5 (G05)$  برای تمام جایگاه‌ها برای مقیاس کردن ماتریس روابط خویشاوندی ژنومی استفاده شد (Forni و همکاران، ۲۰۱۱).

مدل مورد استفاده برای پیش‌بینی ژنومی با روش‌های بیز  $A$  و بیز  $B$  در شکل ماتریسی به صورت زیر بود:

$$y = 1\mu + Sp + Wa + e$$

در این مدل بردار اثرات نشانگرها و  $W$  ماتریس کوواریت‌های ژنوتیپی به صورت صفر، ۱ و ۲ به ترتیب برای ژنوتیپ هموزیگوس  $aa$ ، هتروزیگوس  $Aa$  و  $AA$  هموزیگوس هستند. سایر اجزای مدل نیز در بخش قبل معرفی شده‌اند. در بیز  $A$  توزیع حاشیه‌ای اثرات نشانگرها به صورت یک توزیع  $t$  مقیاس شده بود، ولی در بیز  $B$  اثرات نشانگرها از یک توزیع مخلوط با یک احتمال چگال برای صفر و یک توزیع  $t$  مقیاس شده نمونه‌گیری شدند. همان‌طور که توسط Gianola و همکاران (۲۰۰۹) نشان داده شده، این معادل این است که فرض کنیم که اثر نشانگر در جایگاه  $j$  دارای توزیع نرمال با میانگین صفر و واریانس نامعلوم خاص آن جایگاه  $\sigma_j^2$  است که به این واریانس نیز یک توزیع پیشین کای اسکویر معکوس مقیاس شده با پارامتر مقیاس  $S_\alpha^2$  و درجه آزادی  $\nu_\alpha$  اختصاص داده می‌شود. ساختار بندی بیز  $B$  به صورت بیز  $A$  بوده و تنها تفاوت معرفی یک پارامتر  $\pi$  است که احتمال پیشین صفر بودن اثر نشانگرها است. پارامتر  $\pi$  به عنوان یک ناشناخته تلقی و به آن یک توزیع بتا  $Beta(p_0, \pi_0)$  اختصاص داده شد (Pérez and de Los Campos, ۲۰۱۴).

در نهایت ارزش اصلاحی ژنومی هر حیوان با استفاده از  $GBV_i = \sum_{j=1}^m w_{ij} \hat{\alpha}_j$  بدست آمد. در روش‌های بیز دو مجموعه داده به طور مجزا مورد استفاده قرار گرفت. مجموعه اول تنها شامل حیوانات ایستگاهی و استفاده از ارزش اصلاحی سنتی

<sup>4</sup> Replicate TRN-TST experiments

<sup>5</sup> Cross validation

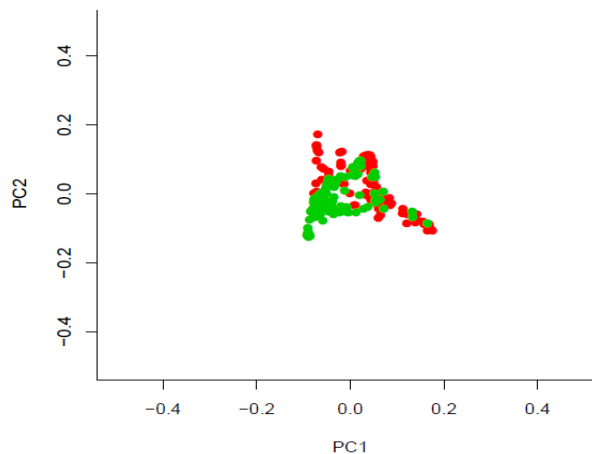
<sup>6</sup> Burn-in

مؤلفه اصلی دهم تنها ۱/۴۳ واریانس کل را تشریح می‌کرد. بنابراین، در این بررسی برای برآورد اثرات نشانگرها و یا ارزش‌های اصلاحی ژنومی مستقیم از مؤلفه اصلی اول به عنوان یک متغیر کمکی برای در نظر گرفتن ساختار متنوع جمعیت استفاده شد. عدم تصحیح ساختار متفاوت جمعیت و پیوستگی خانواده‌ای منجر به برآورد اریب اثرات نشانگرها شده که به نوبه خود می‌تواند منجر به بیش برآورد صحت پیش‌بینی شود (Tiezzi و همکاران، ۲۰۱۵).

پیش‌بینی‌های ژنومی، از یک اثر افزایشی پلی‌ژنی و یا از مؤلفه‌های اصلی در مدل استفاده می‌شود (Price و همکاران، ۲۰۱۰). نتایج آنالیز مؤلفه‌های اصلی بر روی ماتریس روابط خویشاوندی ژنومی (G) در قالب شکل ۱ و جدول ۲ آورده شده است. نتایج نشان داد که اگر چه وجود زیرگروه‌ها در میان افراد در گله‌های مختلف وجود ندارد ولی تنوع قابل توجهی در بین حیوانات وجود دارد. مؤلفه اصلی اول ۱۲/۵۶ درصد تغییرات کل را تشریح می‌کرد و سهم مؤلفه‌های بعدی به شدت کاهش نشان داد به طوری که

جدول ۲- توصیف آماری تجزیه و تحلیل مؤلفه‌های اصلی ماتریس روابط خویشاوندی ژنومی

مؤلفه اصلی	۱	۲	۳	۴	۵	۶	۷	۸	۹	۱۰
مقدار ویژه	۲۱/۷۴	۵/۳۸	۵/۱۷	۳/۹۶	۳/۵۴	۳/۱۹	۲/۹۱	۲/۷۷	۲/۵۳	۲/۴۸
واریانس (%)	۱۲/۵۶	۳/۱۱	۲/۹۹	۲/۲۹	۲/۰۴	۱/۸۴	۱/۶۸	۱/۵۹	۱/۴۶	۱/۴۳
واریانس تجمعی	۱۲/۵۶	۱۵/۶۷	۱۸/۶۶	۲۰/۹۵	۲۲/۹۹	۲۴/۸۴	۲۶/۵۲	۲۸/۱۲	۲۹/۵۸	۳۱/۰۱



شکل ۱- پراکنش بردارهای ویژه مؤلفه اصلی اول (PC1) در مقابل مؤلفه اصلی دوم (PC2). رنگ قرمز و سبز به ترتیب حیوانات ایستگاهی و گله‌های اقماری

نسبت به G05BLUP بهتر عمل کرد و در رابطه با دو صفت دیگر صحت پایین‌تری داشت. توانایی پیش‌بینی صحیح ارزش‌های اصلاحی افراد در روش بهترین پیش‌بینی ناریب خطی به ساختار ماتریس واریانس-کوواریانس بکار برده شده بستگی دارد. برتری استفاده از فراوانی آللی ۰/۵ برای تشکیل ساختار ماتریس واریانس-کوواریانس در دیگر مطالعات هم گزارش شده است.

صحت و اریب (به صورت انحراف از یک) پیش‌بینی‌های ژنومی برای سه صفت مورد بررسی بر اساس دو مدل GBLUP و G05BLUP در شکل دو نشان داده شده است. به طور میانگین بالاترین و پایین‌ترین صحت در بین صفات به ترتیب برای تولید شیر و درصد چربی شیر بدست آمد. در بین مدل‌ها نیز تفاوت رتبه بین مدل‌ها مشاهده شد بطوری که GBLUP تنها برای تولید شیر

همکاران، ۲۰۰۹)، ولی این محققین از تعداد بیشتری حیوان در جمعیت مرجع استفاده می‌کردند در حالی که در مطالعه‌ای دیگر با استفاده از SNP-BLUP و حدود ۶۰۰ حیوان در جمعیت مرجع همبستگی مشابه‌تری با تحقیق حاضر یعنی ۰/۴۵ و ۰/۲۹ برای تولید پروتئین و درصد پروتئین گزارش شد (Hayes و همکاران، ۲۰۰۹).

ضرایب رگرسیون اریب بودن پیش‌بینی‌ها را نشان دادند، هر دو روش بهترین پیش‌بینی نااریب خطی ژنومی به طور مشترک کم برآوردهایی از مقادیر ارزش‌های اصلاحی سنتی بودند. رتبه‌بندی مدل‌ها در تولید پیش‌بینی نااریب دقیقاً مشابه پیش‌بینی با صحت بالا نبوده و روند متفاوتی داشتند. انقباض Shrinkage کمتر یا بیشتر کوواریانس ژنتیکی بین حیوانات می‌تواند منجر به برآوردهای منقبض شده با درجات متفاوت از ارزش‌های ارثی و یا اثر نشانگرها شود که منجر به برآورد ضریب رگرسیون متفاوت برای مدل تابعیت شود. با این حال صحت و اریب دو اصطلاح کاملاً متفاوت بوده و هر کدام جنبه‌های مختلفی از قابلیت پیش‌بینی یک مدل را ارائه می‌دهند. مشابه با تحقیق حاضر استفاده از دو مدل GBLUP و GBLUP به اضافه یک اثر پلی‌ژن در مدل منجر به صحت ۰/۴۳ و ۰/۴۲۸ برای تولید شیر شد. همچنین، اریب مدل با صحت بالاتر ۴ درصد بالاتر از مدل دیگر بود (Gao و همکاران، ۲۰۱۲). ضرایب رگرسیون ۱/۷۵، ۱/۵ و ۱/۱۲ نیز به ترتیب برای صفات درصد چربی، درصد پروتئین و تولید شیر با استفاده از GBLUP توسط Tiezzi و همکاران (۲۰۱۵) گزارش شده است.

صحت بالاتری با استفاده از فراوانی آللی ۰/۵ در مقایسه با استفاده از فراوانی آللی فعلی جمعیت مرجع و یا برآورد فراوانی آللی حیوانات بنیانگذار توسط Aguilar و همکاران (۲۰۱۰) گزارش شده است. در مطالعه‌ی Forni و همکاران (۲۰۱۱) ضریب همبستگی بین ارزش اصلاحی پیش‌بینی شده سنتی و ارزش اصلاحی پیش‌بینی شده‌ی ژنومی با استفاده از فراوانی آللی ۰/۵ بالاتر از استفاده از فراوانی آللی مشاهده شده در جمعیت مرجع گزارش شد. حسینی وردنجان و همکاران (۱۳۹۶) نیز اثر استفاده از فراوانی آللی مشاهده شده و فراوانی برابر با ۰/۵ را در جمعیت‌های تحت انتخاب و بدون انتخاب بررسی و گزارش کردند که استفاده از فراوانی آللی ۰/۵ در هر دو جمعیت منجر به افزایش صحت پیش‌بینی‌ها می‌شود. در این مطالعه صحت بدست آمده برای تولید شیر و درصد پروتئین تا حد زیادی شبیه به هم بود. روش‌های BLUP در سطح نشانگرها (SNP-BLUP) و یا در سطح افراد (G-BLUP) وزن یکسانی را برای تمام نشانگرها فرض می‌کنند. به نوعی این روش همسانی برای بهترین پیش‌بینی خطی سنتی با استفاده از ماتریس A تحت فرضیه بی‌نهایت ژن هر کدام با اثر بسیار کوچک است. به طوری که در ماتریس G میانگین رابطه خویشاوندی بین افراد در کل ژنوم در نظر گرفته می‌شود پس برای تمام صفات در جمعیت یکسان است و به معماری ژنتیکی صفات بستگی ندارد. به هر حال، در مقایسه با تحقیق حاضر مقادیر بالاتری برای همبستگی بین ارزش‌های اصلاحی ژنومی و ارزش‌های اصلاحی سنتی با استفاده از GBLUP برای تولید شیر، درصد چربی و درصد پروتئین به ترتیب ۰/۵۹، ۰/۶۲ و ۰/۶۱ گزارش شده است (Hayes و



شکل ۲- نمودار صحت (سمت چپ) و اریب (سمت راست) پیش‌بینی ژنومی سه صفت تولید شیر، درصد چربی و درصد پروتئین با استفاده از ماتریس‌های مختلف روابط خویشاوندی ژنومی

خواهد شد. چرا که آزمایش‌های نقشه‌یابی جایگاه‌های صفات کمی (QTL) وجود ژن‌های عمده اثری را گزارش کرده‌اند (Hayes and Goddard, ۲۰۰۱). مطالعات نشان داده‌اند که پیش‌بینی‌های GBLUP برعکس روش‌های بیز بیشتر بر پیوستگی درون خانواده‌ای و نمونه‌گیری مندلی استوار بوده و LD بین نشانگرها و QTL در درجه دوم اهمیت قرار دارد (Tiezzi و همکاران، ۲۰۱۵). Daetwyler و همکاران (۲۰۱۰)، صحت بالاتری را با استفاده از بیز B نسبت به GBLUP برای صفاتی که تعداد QTL در مقایسه با تعداد قطعات مستقل کروموزومی کمتر بود گزارش کردند در حالی که این برتری برای صفات پلی ژن وجود نداشت. همچنین گزارش شده است زمانی که صفات تحت تأثیر تعداد زیادی QTL با اثرات کوچک قرار دارند GBLUP دقت پیش‌بینی بالاتری نسبت به روش بیز B دارد (Clark و همکاران ۲۰۱۱). روند مشابهی توسط Coster و همکاران (۲۰۱۰) و همچنین Li and Sillanpää (۲۰۱۲) گزارش شده است. بسیاری از مطالعات با استفاده از داده‌های واقعی نیز گزارش کرده‌اند که برای اکثر صفات روش GBLUP به خوبی روش‌های دارای پتانسیل انتخاب متغیر بیزی عمل می‌کند (Hayes و همکاران، ۲۰۰۹).

صحت ارزش‌های اصلاحی ژنومی پیش‌بینی شده با دو روش بیز A و B و استفاده از دو مجموعه داده در جدول ۳ نشان داده شده است. با استفاده از داده‌های ایستگاهی صحت ارزش‌های اصلاحی برای تولید شیر، درصد چربی شیر و درصد پروتئین در مقایسه با روش‌های GBLUP، به ترتیب کاهش (۰/۰۶)، افزایش (۰/۰۷) و تقریباً ثابت بودند. بیشترین افزایش صحت برای درصد چربی شیر با استفاده از روش بیز A بدست آمد که اختلاف آن با بیز B بسیار کم بود. همچنین پتانسیل دو روش بیز برای پیش‌بینی ارزش‌های اصلاحی حیوانات برای درصد پروتئین مشابه بود که هیچ تفاوتی با روش‌های GBLUP نیز نداشتند (تفاوت‌ها در حدود ۰/۰۲ هستند). به طور معمول انواع روش‌های بیز بکار برده شده در پیش‌بینی ژنومی که پتانسیل انتخاب متغیر را دارند بهتر از مدل‌های فاقد پتانسیل انتخاب متغیر مثل روش بهترین پیش‌بینی نااریب خطی ژنومی (GBLUP) عمل کرده‌اند (Guo و همکاران، ۲۰۱۴). با این حال، مزیت روش‌های بیزی نسبت به GBLUP بیشتر بستگی به معماری ژنتیکی صفت از جمله تعداد جایگاه‌های صفت کمی (QTLs) و هم چنین تراکم نشانگر دارد. فرض وزن یکسان برای تمام نشانگرها باعث کاهش صحت پیش‌بینی‌ها زمانی که بخش بزرگی از واریانس ژنتیکی توسط برخی نشانگرها تشریح می‌شود،



جدول ۳- صحت و اریب پیش‌بینی‌های ژنومی با روش‌های بیزی و دو مجموعه حیوانات ایستگاهی و کل حیوانات

حیوانات ایستگاهی+اقماری				حیوانات ایستگاهی				صفت
صحت		اریب		صحت		اریب		
بیز B	بیز A	بیز B	بیز A	بیز B	بیز A	بیز B	بیز A	
۱/۵۱	۱/۷۲	۰/۴۳۴	۰/۴۵۶	۱/۱۸	۱/۲۴	۰/۳۵۳	۰/۳۷۱	شیر تولیدی
۱/۲۴	۱/۷۶	۰/۳۲۴	۰/۳۶۸	۱/۴۹	۱/۳۷	۰/۳۱۴	۰/۳۲۹	درصد چربی
۱/۷۹	۱/۵۳	۰/۴۵۷	۰/۴۱۱	۱/۲۷	۱/۱۴	۰/۳۷۵	۰/۳۶۷	درصد پروتئین

انتخاب ژنومی استفاده می‌شود. ولی به طور ضمنی پیش فرض این کار برآورد کاملاً قابل اعتماد<sup>۷</sup> این ارزش‌ها در جمعیت است. این قابلیت اعتماد در داده‌های چند میلیونی با دقت بالای رکورددرداری نسبت به مجموعه داده‌های رکورددرداری شده برای جمعیت‌های بومی و اندازه جمعیت کوچک دست یافتنی‌تر است. با این حال بهتر است که رکوردهای فنوتیپی خام و یا رکوردهای فنوتیپی تصحیح شده به طور مستقیم برای برآورد ارزش‌های اصلاحی ژنومی استفاده شود. ولی ماهیت حیوانات تعیین ژنوتیپ شده در این مطالعه امکان استفاده از رکوردهای فنوتیپی را نمی‌داد چرا که صفات بررسی شده محدود به جنس بودند در حالی که در این مطالعه از هر دو جنس استفاده شد. همچنین برخی از حیوانات ماده نیز به سن تولید رکوردد رسیدند. بنابراین این بسیار مهم است که شیب رگرسیون برآورد شده به طور مستقیم به عنوان یک کم برآورد ارزش اصلاحی شناخته نشود بلکه این امکان وجود دارد که تعداد رکوردها و ساختار ماتریس A و سایر پارامترهای BLUP سنتی برآوردهای منبسط‌تری از ارزش‌های ارثی ارائه کرده باشند.

### ۳- نتیجه‌گیری

انتخاب ژنومی به دلیل پتانسیل بالا برای افزایش رشد ژنتیکی به عنوان یک روش استاندارد در اصلاح نژاد گاوهای شیری تبدیل شده است. با این حال لازمه اجرای موفق آن بکارگیری مدل‌های پیش‌بینی است که بتوانند با استفاده از ژنوتیپ‌های نشانگرهای گسترده شده در ژنوم، به درستی ارزش‌های اصلاحی ژنومی را برآورد کنند. در این پژوهش قابلیت پیش‌بینی ۴ مدل مختلف رایج

با اضافه شدن حیوانات اقماری به جمعیت ایستگاهی، صحت پیش‌بینی‌ها افزایش یافت که این افزایش بسته به صفت و روش استفاده شده متفاوت بود. با افزایش تعداد حیوانات به واسطه افزایش داده برای برآورد اثر هر نشانگر امکان افزایش صحت وجود دارد. گزارش شده است که صحت پیش‌بینی‌های ژنومی به شدت به تعداد حیوانات تعیین ژنوتیپ شده بستگی دارد به طوری که گاهی این ارتباط به طور خطی توصیف شده است (VanRaden و همکاران، ۲۰۰۹؛ Calus، ۲۰۱۰؛ Daetwyler و همکاران، ۲۰۱۰). همچنین این امکان وجود دارد که صحت ارزش‌های اصلاحی ژنومی در جمعیت‌های مرجع کوچک شبیه ارزیابی‌های BLUP سنتی بر اساس شجره و عملکرد خود فرد و یا فرزندانش باشد (Lourenco و همکاران، ۲۰۱۴).

در این پژوهش ما اثر استفاده از یک ماتریس باقیمانده ناهمگن در ساختار مدل‌های بیز هنگامی که از کل داده‌ها استفاده می‌شد را نیز برای در نظر گرفتن این مطلب که رکوردهای حیوانات گله‌های اقماری در حقیقت پیش‌بینی شده‌اند نه واقعی، بررسی کردیم. وزن رکوردها هم متناسب با میزان صحت مدل بهترین پیش‌بینی نااریب خطی ژنومی بود که بیشترین صحت را برای آن صفت ارائه کرده بود. با این حال نتایج یکسان بود و یا تفاوت‌ها بسیار اندک بودند و بنابراین استفاده از ماتریس باقیمانده ناهمگن منجر به افزایش صحت ارزیابی‌ها نشد.

نکته مهم دیگر اینکه در این مطالعه از ارزش‌های اصلاحی سنتی به عنوان متغیر پاسخ استفاده شد که به طور معمول هم در مطالعات

<sup>7</sup> Reliable

- Aguilar, I., Misztal, I., Johnson, D., Legarra, A., Tsuruta, S. and Lawlor, T. (2010). Hot topic: A unified approach to utilize phenotypic, full pedigree, and genomic information for genetic evaluation of Holstein final score. *Journal of Dairy Science*, 93(2), 743-752.
- Browning, B. L. and Browning, S. R. (2009). A unified approach to genotype imputation and haplotype-phase inference for large data sets of trios and unrelated individuals. *The American Journal of Human Genetics*, 84(2), 210-223.
- Calus, M. (2010). Genomic breeding value prediction: methods and procedures. *Animal*, 4(02), 157-164.
- Clark, S. A., Hickey, J. M. and Van der Werf, J. H. (2011). Different models of genetic variation and their effect on genomic evaluation. *Genetics Selection Evolution*, 43(1), 18.
- Coster, A., Bastiaansen, J. W., Calus, M. P., van Arendonk, J. A. and Bovenhuis, H. (2010). Sensitivity of methods for estimating breeding values using genetic markers to the number of QTL and distribution of QTL variance. *Genetics Selection Evolution*, 42(1), 9.
- Daetwyler, H. D., Pong-Wong, R., Villanueva, B. and Woolliams, J. A. (2010). The impact of genetic architecture on genome-wide evaluation methods. *Genetics*, 185(3), 1021-1031.
- de los Campos, G. Genome enabled prediction methods: Laboratory. *University of Alabama at Birmingham, Birmingham, Alabama* <http://www-personal.une.edu/~jvanderw/GdlCHandouts.pdf>.
- Forni, S., Aguilar, I. and Misztal, I. (2011). Different genomic relationship matrices for single-step analysis using phenotypic, pedigree and genomic information. *Genetics Selection Evolution*, 43(1), 1.
- Gianola, D., de los Campos, G., Hill, W. G., Manfredi, E. and Fernando, R. (2009). Additive genetic variability and the Bayesian alphabet. *Genetics*, 183(1), 347-363.

در انتخاب ژنومی، هر کدام با فرضیه‌های متفاوت برای اثرات نشانگرها برای بکارگیری در انتخاب ژنومی در گاو بومی نژاد نجدی بررسی شدند. دو مدل بهترین پیش‌بینی نااریب خطی ژنومی در رابطه با صفت تولید شیر صحت بالاتری را داشتند در حالی که کم‌ترین صحت را برای درصد چربی ارائه می‌کردند. در حالی که دو مدل بیز بررسی شده پتانسیل بهبود صحت را برای صفت تولید چربی ارائه کردند ولی در کل میانگین صحت پیش‌بینی برای صفت درصد چربی با هر دو روش BLUP و بیز نسبت به دو صفت دیگر (تولید شیر و درصد پروتئین) کم‌تر بود. همچنین، اثر اضافه شدن حیوانات گله‌های اقماری با رکورد پیش‌بینی شده به عنوان راهکاری برای افزایش جمعیت نیز نشان داد که صحت پیش‌بینی‌ها افزایش می‌یابد ولی این افزایش بسته به حساسیت مدل‌ها و صفت متفاوت بودند. بنابراین امکان طرح‌ریزی انتخاب ژنومی در این نژاد با برآورد ارزش‌های اصلاحی ژنومی صفت تولید شیر با روش بهترین پیش‌بینی نااریب خطی و استفاده از روش‌های بیز برای صفت درصد چربی و پروتئین وجود خواهد داشت. اگر چه داشتن پیش‌بینی‌های ژنومی دقیق در جمعیت کوچک این نژاد مشکل است ولی نتایج امیدوار کننده هستند.

#### منابع

- حسینی وردنجانی، س. م.، شریعتی، م. م.، و نعیمی پور، ح. (۱۳۹۶). تأثیر مقیاس ماتریس روابط خویشاوندی ژنگانی بر برآورد مؤلفه‌های واریانس و درستی پیش‌بینی ارزش‌های اصلاحی. *مجله علوم دامی ایران*، دوره ۴۸، شماره ۲، صص ۱۹۷-۲۰۶.
- نظری، م.، بیگی نصیری، م. ت.، و فیاضی، ج. (۱۳۸۶). ارزیابی قابلیت‌های ژنتیکی و فنوتیپی صفات تولید شیر، چربی و درصد چربی گاوهای شیری نجدی با استفاده از مدل حیوانی تک و دو صفتی. *مجله دامپزشکی ایران*، دوره سوم، شماره ۴، صص ۷۹-۷۳.

- Lund, M. S., Su, G., Janss, L., Guldbbrandtsen, B. and Brøndum, R. F. (2014). Genomic evaluation of cattle in a multi-breed context. *Livestock Science*, 166, 101-110.
- Madsen, P., Sørensen, P., Su, G., Damgaard, L. H., Thomsen, H. and Labouriau, R. (2006). *DMU-a package for analyzing multivariate mixed models*. Paper presented at the 8th World Congress on Genetics Applied to Livestock Production.
- Nejati-Javaremi, A., Smith, C. and Gibson, J. (1997). Effect of total allelic relationship on accuracy of evaluation and response to selection. *Journal of animal science*, 75(7), 1738-1745.
- Pérez, P. and de Los Campos, G. (2014). Genome-wide regression & prediction with the BGLR statistical package. *Genetics*, 114.164442.
- Price, A. L., Zaitlen, N. A., Reich, D. and Patterson, N. (2010). New approaches to population stratification in genome-wide association studies. *Nature reviews genetics*, 11(7), 459.
- Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A., Bender, D. and Daly, M. J. (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. *The American Journal of Human Genetics*, 81(3), 559-575.
- Schaeffer, L. (2006). Strategy for applying genome-wide selection in dairy cattle. *Journal of Animal Breeding and Genetics*, 123(4), 218-223.
- Su, G., Brøndum, R. F., Ma, P., Guldbbrandtsen, B., Aamand, G. P. and Lund, M. S. (2012). Comparison of genomic predictions using medium-density (~ 54,000) and high-density (~ 777,000) single nucleotide polymorphism marker panels in Nordic Holstein and Red Dairy Cattle populations. *Journal of dairy science*, 95(8), 4657-4665.
- Tiezzi, F. and Maltecca, C. (2015). Accounting for trait architecture in genomic predictions of US Holstein cattle using a weighted realized relationship matrix. *Genetics Selection Evolution*, 47(1), 24.
- Guo, G., Zhao, F., Wang, Y., Zhang, Y., Du, L. and Su, G. (2014). Comparison of single-trait and multiple-trait genomic prediction models. *BMC genetics*, 15(1), 30.
- Gao, H., Christensen, O. F., Madsen, P., Nielsen, U. S., Zhang, Y., Lund, M. S. and Su, G. (2012). Comparison on genomic predictions using three GBLUP methods and two single-step blending methods in the Nordic Holstein population. *Genetics Selection Evolution*, 44(1), 8.
- Habier, D., Fernando, R. L., Kizilkaya, K. and Garrick, D. J. (2011). Extension of the Bayesian alphabet for genomic selection. *BMC bioinformatics*, 12(1), 1.
- Hayes, B., and Goddard, M. E. (2001). The distribution of the effects of genes affecting quantitative traits in livestock. *Genetics Selection Evolution*, 33(3), 209-230.
- Hayes, B. J., Bowman, P. J., Chamberlain, A. C., Verbyla, K. and Goddard, M. E. (2009). Accuracy of genomic breeding values in multi-breed dairy cattle populations. *Genetics Selection Evolution*, 41(1), 51.
- Hayes, B. J., Bowman, P. J., Chamberlain, A. and Goddard, M. (2009). Invited review: Genomic selection in dairy cattle: Progress and challenges. *Journal of dairy science*, 92(2), 433-443.
- Hoze, C., Fritz, S., Phocas, F., Boichard, D., Ducrocq, V. and Croiseau, P. (2014). Efficiency of multi-breed genomic selection for dairy cattle breeds with different sizes of reference population. *Journal of dairy science*, 97(6), 3918-3929.
- Li, Z. and Sillanpää, M. J. (2012). Overview of LASSO-related penalized regression methods for quantitative trait mapping and genomic selection. *Theoretical and Applied Genetics*, 125(3), 419-435.
- Lourenco, D., Misztal, I., Tsuruta, S., Aguilar, I., Ezra, E., Ron, M., Shirak, A. and Weller, J. (2014). Methods for genomic evaluation of a relatively small genotyped dairy population and effect of genotyped cow information in multiparity analyses. *Journal of dairy science*, 97(3), 1742-1752.

