

تأثیر مقیاس ماتریس روابط خویشاوندی ژنگانی بر برآورد مؤلفه‌های واریانس و درستی پیش‌بینی ارزش‌های اصلاحی

سید مهدی حسینی وردنجان^۱، محمدمهدی شریاتی^{۲*} و حسین نعیمی پور یونسی^۱
۱ و ۲. دانشجوی دکتری و استادیار ژنتیک و اصلاح نژاد دام، دانشگاه فردوسی مشهد
(تاریخ دریافت: ۱۳۹۵/۳/۴ - تاریخ پذیرش: ۱۳۹۶/۳/۲۴)

چکیده

در این پژوهش، روش برای پیش‌بینی فراسنجه‌های ناشناخته پنج مدل بهترین پیش‌بینی نااریب خطی ژنگانی (ژنومی G-BLUP) از روش بیز و نمونه‌گیری گیبس استفاده شد. در هر مدل از مقیاس‌های متفاوتی برای ماتریس G شامل استفاده از فراوانی آللی جمعیت بنیان‌گذار (*Gfoun*)، فراوانی آللی جمعیت مرجع (*Gref*)، فراوانی آللی برابر با ۰/۵ (*G05*)، یک ماتریس نرمال شده با میانگین عنصرهای قطری برابر با یک (*Gnorm*) و یک ماتریس G وزن‌شده با ماتریس A (*Gwei*)، استفاده شد. برای مقایسه نتایج از یک جمعیت دارای آمیزش تصادفی و یک جمعیت انتخاب‌شده، برای صفتی با وراثت‌پذیری ۰/۲۵ روی یک ژنگان با ۱۰۵ QTL و ۳۰۰۰ نشانگر تک نوکلئوتیدی روی سه کروموزوم استفاده شد. نتایج نشان داد، عنصرهای ماتریس‌های G در مقایسه با ماتریس A واریانس بالاتری دارند. میانگین عنصرهای قطری و غیر قطری به‌غیر از *Gwei* و *Gnorm* از عنصرهای متناظر در A بالاتر بودند. روش‌های *Gnorm*-BLUP و *G05*-BLUP در مقایسه با سه روش دیگر منجر به برآورد متورم واریانس ژنتیکی شدند که این تورم در جمعیت انتخاب‌شده کمتر بود. میانگین درستی پنج مدل G-BLUP در جمعیت تصادفی ۰/۰۸۴ بالاتر (۰/۷۳۶ در مقابل ۰/۶۵۲) از جمعیت انتخاب‌شده و میانگین اریبی ۰/۰۱۴ پایین‌تر (۰/۰۲۶ در مقابل ۰/۰۴) بود. اریبی پیش‌بینی ارزش اصلاحی حقیقی جمعیت انتخاب‌شده با استفاده از *Gwei* نزدیک به صفر ولی با *Gref* بیشتر از ۰/۰۶ بود. بیشترین درستی و کمترین اریب می‌تواند با استفاده از فراوانی آللی جمعیت مرجع که با ماتریس A مقیاس شده‌اند، به‌دست آید.

واژه‌های کلیدی: اعتبارسنجی متقابل، پیش‌بینی ژنگانی، روش بیز، فراوانی آللی، قابلیت پیش‌بینی.

Effect of scaling genomic relationship matrix on estimation of variance components and accuracy of breeding values

Sayed Mahdi Hosseini Vardanjani¹, Mohammad Mahdi Shariati^{2*} and Hossein Naeemipour Yuonesi¹

1, 2. Ph.D. Student and Assistant Professor, Animal Breeding and Genetics, Department of Animal Science, University of Mashhad, Iran

(Received: May 24, 2016 - Accepted: Jun. 14, 2017)

ABSTRACT

In this study, Bayesian approach via Gibbs sampling was used to predict unknown parameters of five equivalent Genomic Best Linear Unbiased Predictions (G-BLUP), each with different scale of G matrix by using allele frequency of founder population (*Gfoun*), allele frequency of reference population (*Gref*), allele frequency equal to 0.5 (*G05*), a normalized matrix with average diagonal coefficients equal to 1 (*Gnorm*) and a weighted G matrix with A matrix (*Gwei*). A random mating population and a selected population were used to compare results of a trait with heritability of 0.25 on a genome constructed of three chromosomes with 105 QTLs and 3000 single nucleotide polymorphisms. The results showed that higher variance existed in the elements of G matrices compared with A matrix. Average diagonal and off-diagonal elements except *Gnorm* and *Gwei* were higher than corresponding elements in A. *Gnorm*-BLUP and *G05*-BLUP methods led to inflated genetic variance in contrast other three methods and this inflation was lower in selected population. Average accuracy over 5 G-BLUP in random population was 0.084 higher than selected population (0.762 vs. 0.652) and bias was 0.041 lower (0.026 vs. 0.04). Bias of prediction of true breeding value of selected population by using *Gwei* almost was zero but with *Gref* greater than 0.06. The greatest accuracy and the smallest bias can be obtained by using allele frequency of reference population that re-scaled with A matrix.

Keywords: Allele frequency, bayesian approach, cross validation, genomic prediction, predictive ability.

* Corresponding author E-mail: shariati52@gmail.com

مقدمه

انتخاب برای صفات اقتصادی مهم در حیوانات و گیاهان به‌طور مرسوم بر پایهٔ رکوردهای پدیدگانی (فنوتیپی) افراد و رابطهٔ خویشاوندی آن‌ها انجام می‌شود، به‌طوری‌که رکوردهای پدیدگانی افراد و احتمال یکسان اجدادی بودن ژن‌ها با استفاده از شجره در قالب روش بهترین پیش‌بینی ناریب خطی ترکیب‌شده تا شایستگی ژنتیکی افراد پیش‌بینی شود (Henderson, 1984). اما پیشرفت‌های اخیر در فناوری مولکولی منجر به کشف نشانگرهای DNA، توالی‌یابی و نقشه‌یابی ژنگان (ژنوم) گونه‌های حیوان‌های مزرعه شده است. ردیابی به ارث رسیدن قطعه‌های کوچک کروموزومی در افراد با استفاده از داده‌های DNA آسان‌تر بوده و در نتیجه داده‌های دقیق‌تری در رابطه با یکسان بودن آن‌ها نسبت به آنچه از شجره به دست می‌آید، ارائه می‌دهند (Daetwyler *et al.*, 2010; Harris & Johnson, 2010).

ارزش ارثی ژنگانی حیوان‌ها را می‌توان با استفاده از جمع اثرگذاری هر یک از نشانگرها (Calus, 2010) و یا با گنجاندن یک ماتریس روابط خویشاوندی ژنگانی (G) به‌دست‌آمده از حیوان‌های تعیین نژادگان (ژنوتیپ) شده، در معادله‌های بهترین پیش‌بینی ناریب خطی (BLUP) پیش‌بینی کرد (Hayes & Goddard, 2008; Habier *et al.*, 2010; Tiezzi & Nejati-Maltecca, 2015). برای نخستین بار (Javaremi *et al.*, 1997) از یک ماتریس رابطهٔ خویشاوندی مبتنی بر نشانگرها به‌جای استفاده از یک ماتریس رابطه‌های خویشاوندی مورد انتظار (A) استفاده کردند و نشان دادند، واریانس خطای پیش‌بینی کاهش و در مقابل همبستگی بین ارزش‌های اصلاحی پیش‌بینی‌شده و حقیقی یا درستی ارزیابی افزایش می‌یابد. افزایش درستی ارزیابی‌ها با جایگزین کردن ماتریس رابطهٔ خویشاوندی تحقق‌یافته به‌جای ماتریس رابطهٔ خویشاوندی مورد انتظار، ناشی از تحت تأثیر نبودن این ماتریس از کاستی‌های شجره و یا فرض صفر بودن میانگین نمونه‌گیری مندلی است. بررسی‌های بعدی نیز با استفاده از همانندسازی نشان دادند که انتخاب ژنگانی هم از

رابطهٔ خویشاوندی ژنگانی بین افراد و هم از نبود تعادل پیوستگی بین نشانگرها و QTL برای بهبود درستی ارزش اصلاحی ژنگانی پیش‌بینی‌شده استفاده می‌کند (Habier *et al.*, 2007; Daetwyler *et al.*, 2010).

یک مسئلهٔ مهم در رابطه با تشکیل ماتریس رابطه‌های خویشاوندی ژنگانی استفاده از فراوانی آللی مناسب برای تشکیل این ماتریس است. از لحاظ تئوری برای تشکیل ماتریس رابطه‌های خویشاوندی ژنگانی باید از فراوانی آللی حیوان‌های بنیان‌گذار در شجره استفاده شود (VanRaden, 2008). اما از آنجایی‌که حیوان‌های پایه بسیار کم تعیین نژادگان می‌شوند، در عمل چنین فراوانی آللی در دست نخواهد بود. بنابراین نیاز به روش‌های دیگری برای ایجاد این ماتریس است. برخی از این روش‌ها برای تشکیل ماتریس G هنگامی‌که این ماتریس در قالب روش تک‌مرحله‌ای^۱ برای تشکیل یک ماتریس مخلوط به نام H با ماتریس رابطه‌های خویشاوندی مبتنی بر شجره (A) ترکیب می‌شود توسط محققان بررسی شده است (Forni *et al.*, 2011). در این بررسی‌های به‌طور معمول با استفاده از فراوانی‌های آللی دیگر و یا افزودن یک‌میزان ثابت به عنصرهای ماتریس G این ماتریس تشکیل می‌شود. پس با توجه به اینکه برای مقیاس کردن ماتریس رابطه‌های خویشاوندی ژنگانی روش‌های مختلف پیشنهاد شده است یک پرسش پیش می‌آید و آن اینکه در جامعه‌های مختلف و با استفاده از معادلات بهترین پیش‌بینی ناریب خطی ژنگانی (G-BLUP) کدام ساختار ماتریس G بهینه است و یا اینکه هرکدام به چه میزان از بهینه بودن انحراف دارند. بنابراین هدف از این پژوهش ارزیابی تأثیرپذیری درستی ارزش‌های اصلاحی پیش‌بینی‌شده و برآورد مؤلفه‌های واریانس با استفاده از ساختارهای متفاوت ماتریس رابطه‌های خویشاوندی ژنگانی در جمعیت‌های دارای آمیزش تصادفی و جمعیت‌های تحت تأثیر انتخاب بود.

مواد و روش‌ها

هماندسازی ژنگان و جمعیت

صفتی با واریانس پدیدگانی یک و وراثت‌پذیری ۰/۲۵ روی ژنگانی متشکل از سه کروموزوم، هر یک به طول ۱۰۰ سانتی‌مورگان همانندسازی شد. روی هر یک از این کروموزوم‌ها شمار ۱۰۰۰ نشانگر تک نوکلئوتیدی توزیع شد. موقعیت نشانگر به صورت تصادفی در نظر گرفته شد. انتخاب شمار کروموزوم و شمار نشانگرها با توجه به نیازهای محاسباتی و رویه رایج در بررسی‌های همانندسازی در انتخاب ژنگانی انجام گرفت. شمار QTL ۱۰۵ با موقعیت تصادفی روی ژنگان توزیع شد. تأثیر QTLها از یک توزیع گاما با فراسنجه شکل ۰/۴ و فراسنجه نرخ ۱/۶ نمونه‌گیری شد و اجازه داده شد تا هر QTL دارای ۲ تا ۴ آلل باشد ولی نشانگرها تنها ۲ آلل (چندشکلی تک نوکلئوتیدی) داشتند. فرآیند جهش برگشت‌پذیر برای نشانگرها $2/5 \times 10^{-3}$ و برای QTLها $2/5 \times 10^{-5}$ در نظر گرفته شد.

در آغاز برای ایجاد تعادل جهش-رانس ژنتیکی، جمعیتی با اندازه مؤثر نزدیک به ۸۰ فرد متشکل از ۴۰۰ حیوان ماده و ۲۰ حیوان نر برای صد نسل آمیزش تصادفی داده شدند که در فرآیند آن با یک اتساع تدریجی جمعیت اجازه داده شد شمار حیوان‌ها افزایش پیدا کند و به شمار ۱۰۰۰ حیوان برای داشتن کمترین حیوان مورد نیاز در جمعیت‌های کنونی برسد. این حالت برای ۴۰۰ نسل دیگر ادامه یافت و شمار حیوان‌های نر نیز در نسل آخر به ۷۰ رسانده شد. در نتیجه این فرآیند جمعیت تاریخی ساخته شد. از آخرین نسل جمعیت تاریخی (نسل ۵۰۰) حیوان‌های پایه‌گذار جمعیت‌های کنونی (نسل صفر) مشتق شدند به طوری که ۴۵۵ حیوان ماده و ۳۵ حیوان نر برای جمعیت اول و به همین شمار حیوان برای جمعیت دوم به طور تصادفی انتخاب شدند. در جمعیت اول هیچ‌گونه انتخابی صورت نگرفت و اجازه داده شد تا آمیزش تصادفی در جمعیت برقرار باشد و افراد نسل‌های بعدی به صورت تصادفی از نسل کنونی انتخاب شوند. ولی در جمعیت دوم برای همسانی با غالب برنامه‌های اصلاح نژادی، انتخاب حیوان‌ها بر پایه ارزش‌های اصلاحی بود به طوری که حیوان‌های مورد

نیاز برای نسل بعدی از آمیزش حیوان‌های دارای بالاترین ارزش اصلاحی محاسبه شده بر پایه BLUP مبتنی بر شجره در نسل کنونی انتخاب می‌شدند. همانندسازی این دو جمعیت برای ۹ نسل ادامه یافت. با این ساختار همانندسازی، میانگین عدم تعادل پیوستگی در بین تکرارها و همچنین شمار نشانگرهای در حال تفرق در بین تکرارها پس از ۵۱۰ نسل به ترتیب $0/145 \pm 0/005$ و $r^2 = 0/145 \pm 0/005$ بودند.

نشانگرها و QTLهای همانندسازی شده برای همه افراد از نسل هشت به بعد در هر یک از جمعیت‌ها (جمعیت دارای آمیزش تصادفی و جمعیت انتخاب شده) ذخیره شدند. برای هر دو جنس نر و ماده یک ارزش پدیدگانی از مجموع ارزش QTLهای آن فرد که یک انحراف محیطی دارای توزیع نرمال با میانگین صفر و واریانس یک به آن افزوده می‌شد همانندسازی و ذخیره شد. شجره نیز برای همه حیوان‌ها از نسل ۰ تا نسل ۹ به طور کامل ثبت و در فایل شجره ذخیره شد. حیوان‌های نسل ۸ و ۹ به عنوان جامعه آزمون استفاده شدند. برای همانندسازی ژنگان و جمعیت‌ها از نرم‌افزار QMSim استفاده شد (Sargolzaei & Schenkel, 2009). همانندسازی ژنگان و جمعیت‌ها ده بار تکرار و برای ارزیابی نیز همان‌طور که در ادامه خواهد آمد از اعتبارسنجی متقابل پنج طرفه استفاده شد. بنابراین برآورد مؤلفه‌های واریانس ناشی از ده تکرار و برآورد درستی ارزیابی‌ها میانگین ده بار اعتبارسنجی متقابل پنج طرفه هستند.

روش‌های آماری

مدل مختلط رایج برای پیش‌بینی ارزش‌های اصلاحی ژنگانی با استفاده از ماتریس رابطه‌های خویشاوندی ژنگانی (G-BLUP) به شکل زیر بود:

$$y = Xb + Zg + e$$

که در آن y بردار ارزش‌های پدیدگانی، b بردار اثر ثابت جنس، g بردار ارزش‌های اصلاحی ژنگانی و e نیز بردار باقی‌مانده هستند. ماتریس‌های X و Z نیز به ترتیب رکوردها را به اثر ثابت و حیوان‌ها مرتبط

ناخالصی (هتروزیگوت) و خالصی دیگر نشان داده می‌شوند. فراوانی آلل دوم در جایگاه i را با p_i نشان داده و ماتریس P را برای همه جایگاه‌ها به صورت $2\pi i$ تشکیل می‌دهیم. با کم کردن ماتریس P از M ، ماتریس Z به دست می‌آید که میانگین اثر نشانگرها را صفر می‌کند. در نهایت ماتریس رابطه‌های خویشاوندی ژنگانی با استفاده از $G = \frac{ZZ^T}{2\sum p_i(1-p_i)}$ به دست می‌آید. مقیاس‌های مختلف برای ماتریس رابطه‌های خویشاوندی ژنگانی با استفاده از p_i ‌های مختلف تعریف شدند.

در این بررسی با توجه به فراوانی آلل دوم در هر جایگاه چهار نوع مختلف ماتریس G ایجاد شد. فراوانی آلل دوم در هر جایگاه با استفاده از حیوان‌های پایه‌گذار به دست آمد و از آن برای تشکیل ماتریس رابطه‌های خویشاوندی ژنگانی G_{foun} استفاده شد. فراوانی آلل دوم در هر جایگاه با استفاده از حیوان‌های تعیین نژادگان شده در جمعیت مرجع نیز به دست آمد و ماتریس G_{ref} ایجاد شد. همچنین به‌طور فرضی فراوانی آللی 0.5 برای همه جایگاه‌ها در نظر گرفته شد و ماتریس G_{05} ساخته شد. یک ماتریس رابطه‌های خویشاوندی عادی شده (G_{norm}) با استفاده از همان فراوانی آللی جمعیت مرجع ولی با این تفاوت که مخرج با $\frac{Tracs(ZZ^T)}{n}$ جایگزین شد نیز به دست آمد. مقیاس کردن ماتریس G با استفاده از این مخرج هنگامی که همخوانی پایین است و یا شمار نسل کم است با ارائه میانگین عنصرهای قطری برابر با یک، قابل مقایسه بودن با ماتریس A را تضمین می‌کند (Forni et al., 2011). نوع دیگری از مقیاس کردن ماتریس G با استفاده از رابطه $G_{wei} = \beta G_{ref} + \alpha$ به دست آمد که ضریب‌های β و α با استفاده از رابطه زیر به دست آمدند (Christensen et al., 2012):

$$\text{mean}(\text{diag}(G))\beta + \alpha = \text{mean}(\text{diag}(A))$$

$$\text{mean}(G)\beta + \alpha = \text{mean}(A)$$

ارزیابی روش‌ها

برای سنجش قابلیت پیش‌بینی مدل‌ها در جامعه مرجع از روش اعتبارسنجی متقابل پنج طرفه استفاده شد. در این روش مجموعه حیوان‌های جامعه مرجع به

می‌سازند. فرض شد که $g \sim N(0, G\sigma_g^2)$ و $e \sim N(0, I\sigma_e^2)$ باشند که G ماتریس رابطه‌های خویشاوندی ژنگانی بوده که به شکلی که در ادامه خواهد آمد ایجاد می‌شود. همچنین $I\sigma_e^2$ و $I\sigma_g^2$ نیز به ترتیب واریانس ژنتیکی افزایشی و واریانس محیطی هستند.

مؤلفه‌های واریانس، ارزش‌های اصلاحی و دیگر فراسنجه‌های مدل به روش بیزی و با استفاده از نمونه‌گیری گیبس در محیط نرم‌افزار BGLR به دست آمدند. به این صورت که به اثر ثابت یک توزیع یکنواخت اختصاص داده شد. برای ارزش‌های اصلاحی و واریانس ژنتیکی به ترتیب یک توزیع نرمال و یک توزیع کای اسکویئر معکوس با مشخصات گرفته شد. توزیع پیشین بردار باقی‌مانده به صورت $p(g, \sigma_g^2) = N(g|0, G\sigma_g^2)\chi^{-2}(\sigma_g^2|S_g, df_g)$ در نظر گرفته شد. توزیع پیشین بردار باقی‌مانده به صورت $p(e|\sigma_e^2) = N(e|0, I\sigma_e^2)$ بود که واریانس باقی‌مانده نیز از یک توزیع کای اسکویئر معکوس با مشخصات $p(\sigma_e^2) = \chi^{-2}(\sigma_e^2|S_e, df_e)$ نمونه‌گیری می‌شد. در هر اجرا از یک زنجیره به طول ۱۵۰۰۰ نمونه استفاده شد که ۱۵۰۰ نمونه اول به‌عنوان دوره سوخته حذف می‌شدند. با آزمایش و خطا مشخص شد که همگرایی با همین شمار دور به‌خوبی به دست می‌آید و دورهای بیشتر تأثیر چشمگیری بر برآورد ناشناخته‌ها نداشتند.

تشکیل ماتریس رابطه‌های خویشاوندی ژنگانی (G)

برای تشکیل ماتریس G از روش ارائه شده توسط (VanRaden, 2008) استفاده شد. آغاز ماتریس $M_{n \times m}$ که مشخص‌کننده نژادگان نشانگرها برای افراد است را تشکیل داده که n نشان‌دهنده حیوان‌ها و m نیز شمار نشانگرهایی است که پس از مرحله کنترل کیفیت شامل حذف نشانگرهای با فراوانی آلل نادر کمتر از 0.1 و انحراف از تعادل هاردی واینبرگ در سطح 5×10^{-5} ، باقی‌مانده بودند (مرحله کنترل کیفیت به‌طور میانگین در بین تکرارها منجر به حذف $35/42 \pm 7/57$ نشانگر شد). عنصرهای ماتریس M به شیوه‌های مختلف کدگذاری می‌شوند ولی به‌طور معمول آن‌ها به‌صورت شمار رونوشت آلل دوم در هر نشانگر، یعنی 0 ، 1 و 2 ، به ترتیب برای خالصی (هموزیگوت)،

بودن عنصرهای ماتریس G در مقایسه با ماتریس A مربوط به داده‌های اضافی است که ماتریس G ارائه می‌دهد ولی بخشی دیگر ناشی از خطاهایی است که از مقیاس کردن نادرست G و استفاده نکردن از شمار کافی SNP در تشکیل این ماتریس رخ می‌دهد. در پژوهشی گزارش شده است، مقادیر عنصرهای قطری با کم شدن شمار SNP افزایش پیدا می‌کنند (Chen *et al.*, 2011). بنابراین هنگامی که شمار SNP استفاده شده در تشکیل ماتریس G کم است امکان افزایش میزان مطلق خطا وجود دارد که این می‌تواند موجب ایجاد رابطه‌های خویشاوندی دروغین شده و دلیلی برای بزرگ‌تر بودن غیرطبیعی عنصرهای قطری ماتریس‌های *Gfoun* و *Gref* باشد. به‌رحال نتایج این بررسی با نتایج گزارش شده در دیگر بررسی‌ها همسان هستند. به‌طور مثال در بررسی روی جوجه‌های گوشتی برای تشکیل ماتریس G از دو فراوانی آللی ۰/۵ و فراوانی آللی حیوان‌های جامعه مرجع و نیز شمار مختلف نشانگر SNP استفاده و گزارش شد که همواره عنصرهای ماتریس‌های G ساخته شده از ماتریس A بزرگ‌تر بود (Chen *et al.*, 2011). کمینه و بیشینه اعداد غیر قطری در ماتریس G تشکیل شده با استفاده از فراوانی آللی ۰/۵ روی خوک‌های تحت انتخاب ۰/۳۸۷ و ۱/۲۳۱ و در ماتریس A به ترتیب ۰/۱۰ و ۰/۶ گزارش شد (Forni *et al.*, 2011) که هر دو این مقادیر در G بیشتر هستند که هماهنگ با این بررسی است.

همه عنصرهای قطری و غیر قطری در همه ماتریس‌های ژنگانی واریانس بیشتری نسبت به ماتریس شجره‌ای داشتند. واریانس بیشتر عنصرهای در G نسبت به A مورد انتظار بود چراکه رابطه خویشاوندی ژنگانی منعکس‌کننده کسر حقیقی ژن‌های مشترک بین دو فرد است درحالی‌که رابطه‌های شجره‌ای تنها پیش‌بینی‌ها هستند و هنگامی که خطای پیش‌بینی صفر نباشد واریانس پیش‌بینی‌ها از متغیر پیش‌بینی شده کمتر است (Forni *et al.*, 2011). از سویی افرادی که تصور می‌شود با استفاده از شجره رابطه خویشاوندی یکسانی داشته باشند ولی در عمل با استفاده از داده‌های ژنگان رابطه خویشاوندی کمتر و یا بیشتر از آنچه مورد انتظار بود را نشان خواهند داد که خود باعث افزایش واریانس می‌شود.

پنج گروه یکسان تقسیم شدند که چهار گروه آن‌ها به‌عنوان گروه آزمون و گروه پنجم به‌عنوان گروه ارزیابی در نظر گرفته شد و این کار پنج بار تکرار شد و میانگین همبستگی بین ارزش‌های ژنتیکی پیش‌بینی شده و ارزش‌های ژنتیکی حقیقی به‌عنوان درستی هر مدل گزارش شد. برای سنجش اریبی پیش‌بینی‌ها نیز از تابعیت ارزش‌های ژنتیکی حقیقی بر ارزش‌های ژنتیکی پیش‌بینی شده استفاده شد. در این شیوه ضریب رگرسیون برابر با یک نشان‌دهنده نبود اریبی پیش‌بینی‌ها و ضریب رگرسیون کمتر از یک و بیشتر از یک به ترتیب نشان‌دهنده بیش برآورد و کم برآورد ارزش‌های ژنتیکی پیش‌بینی شده در مقایسه با ارزش‌های ژنتیکی حقیقی است.

نتایج و بحث

آماره‌های توصیفی رابطه‌های خویشاوندی بین حیوان‌ها در ماتریس‌های مختلف رابطه‌های خویشاوندی ژنگانی و ماتریس رابطه‌های خویشاوندی مبتنی بر شجره برای حیوان‌های تعیین نژادگان شده در جدول‌های ۱ و ۲ ارائه شده است. با مقایسه عنصرهای قطری و غیر قطری در ماتریس‌های مختلف رابطه‌های خویشاوندی دیده می‌شود که همواره این ضریب‌ها در جمعیت دارای آمیزش تصادفی کمتر از جمعیت با آمیزش‌های انتخابی است که دلیل آن نیز همان انتخاب افراد دارای ارزش اصلاحی بالاتر در جمعیت دوم و در نتیجه افزایش رابطه‌های خویشاوندی بین افراد در جمعیت دوم است. در حقیقت این پیامدی از انتخاب بر پایه BLUP در طول نسل‌هاست که منجر به افزایش رابطه‌های خویشاوندی و هم‌خونی بالاتر می‌شود.

مقایسه ضریب‌های ماتریس A با ماتریس‌های G نیز بزرگ‌تر بودن این ضریب‌ها در G‌های مختلف نسبت به A را نشان می‌دهد. البته تنها استثناء ماتریس *Gwei* است که ضریب‌های آن تا حدودی همسان با ماتریس A است که ناشی از نوع مقیاس کردنی است که برای تشکیل این ماتریس استفاده شد به‌طوری‌که ضریب‌های α و β اطمینان از یکسان بودن عنصرهای این ماتریس‌ها می‌دهند. بخشی از بزرگ‌تر

عنصرهای قطری و غیر قطری در ماتریس A و G05 را به ترتیب ۰/۹۹ و ۰/۸۲۹ گزارش کردند که نسبت ۱/۲ بین آنها برقرار است. این نسبت تا حدودی همسان نسبت ۱/۳ بین همین مقادیر (۰/۹۵) در برابر ۰/۷۱ به ترتیب برای ماتریس A و G05 در جمعیت انتخاب شده) در این بررسی است. در حالی که در بررسی دیگر نسبت ۱/۵ بین این دو کمیت (۰/۹۷) در برابر ۰/۶۶ به ترتیب برای ماتریس A و G05) گزارش شد (Forni et al., 2011).

از آنجایی که بررسی‌های مختلف از جمعیت‌های متفاوت با درجه‌های مختلف خویشاوندی استفاده می‌کنند بهتر است برای مقایسه کمی ضریب‌های ماتریس‌های مختلف رابطه‌های خویشاوندی به جای مقایسه مستقیم عنصرهای این ماتریس‌ها به مقایسه نسبی تفاوت عنصرهای قطری و غیر قطری این ماتریس‌ها پرداخت. تفاوت بین این عنصرها در این بررسی در جدول ۳ گزارش شده است. Chen et al. (2011) در نتایج بررسی‌های خود تفاوت بین

جدول ۱. کمینه، بیشینه، میانگین و واریانس عنصرهای قطری در ماتریس‌های مختلف رابطه‌های خویشاوندی

		A	Gfoun	Gref	G05	Gnorm	Gwei
Population 1	Minimum	1.000	1.907	1.968	1.098	0.914	0.830
	Maximum	1.267	2.685	2.761	1.690	1.281	1.610
	Average	1.012	2.089	2.154	1.221	1.000	1.012
	Variance	0.0008	0.008	0.008	0.004	0.002	0.008
Population 2	Minimum	1.000	1.924	2.211	1.103	0.904	0.815
	Maximum	1.315	2.787	3.147	1.773	1.287	1.708
	Average	1.040	2.183	2.445	1.262	1.000	1.038
	Variance	0.001	0.013	0.015	0.008	0.002	0.014

جدول ۲. کمینه، بیشینه، میانگین و واریانس عنصرهای غیر قطری در ماتریس‌های مختلف رابطه‌های خویشاوندی

		A	Gfoun	Gref	G05	Gnorm	Gwei
Population 1	Minimum	0	0.816	0.860	0.224	0.4	-0.260
	Maximum	0.765	2.084	2.150	1.220	1.000	1.009
	Average	0.032	1.109	1.157	0.449	0.537	0.032
	Variance	0.002	0.01	0.01	0.006	0.002	0.01
Population 2	Minimum	0.0008	0.844	1.052	0.25	0.43	-0.290
	Maximum	0.858	2.372	2.701	1.445	1.105	1.283
	Average	0.089	1.212	1.451	0.533	0.593	0.089
	Variance	0.003	0.02	0.02	0.01	0.003	0.02

جدول ۳. تفاوت میانگین‌های عنصرهای قطری و غیر قطری در ماتریس‌های مختلف رابطه‌های خویشاوندی

	A	Gfoun	Gref	G05	Gnorm	Gwei
Population 1	0.98	0.98	1	0.77	0.46	0.98
Population 2	0.95	0.93	0.99	0.71	0.41	0.95

برآورد شده تا حدودی همسان هستند که می‌تواند ناشی از همسانی تفاوت بین عنصرهای قطری و غیر قطری این ماتریس‌ها، به ترتیب ۰/۹۳، ۰/۹۵ و ۰/۹۹ باشد. برآورد واریانس ژنتیکی با استفاده از دو ماتریس دیگر در مقایسه با این سه ماتریس متورم^۱ است و این تورم در رابطه با Gnorm شدیدتر است. نتایج این جدول گویای یک قاعده کلی در برآورد واریانس ژنتیکی است به صورتی که این برآورد به‌طور معکوس با تفاوت بین عنصرهای قطری و

برآوردهای مؤلفه‌های واریانس و وراثت‌پذیری با استفاده از ماتریس‌های ژنگانی مختلف در جدول ۴ ارائه شده است. در جمعیت اول برآورد واریانس‌های ژنتیکی و باقی‌مانده با استفاده از Gfoun و Gref تا حدودی یکسان است. همچنین این برآوردها با Gwei نیز تفاوت بسیار کمی دارند به‌طوری‌که وراثت‌پذیری با استفاده از این سه ماتریس تنها در دامنه ۰/۲۵۱ تا ۰/۲۶۷ تغییر دارد. اگرچه برای مقیاس کردن این سه ماتریس از روش‌ها و فراوانی‌های آلی مختلف استفاده شد ولی واریانس‌های

1. Inflate

از آنجایی که مهم‌ترین تأثیر انتخاب در یک جمعیت تغییر در فراوانی آللی است و فراوانی آللی نیز به‌طور مستقیم در واریانس ژنتیکی (Falconer & Mackay, 1995) و در مقیاس ماتریس G دخالت دارد پس برآورد مؤلفه‌های واریانس تغییر چشمگیری در جمعیت انتخاب‌شده دارند. برخلاف جمعیتی که بدون انتخاب بود و هر دو فراوانی آللی استفاده‌شده با استفاده از حیوان‌های پایه‌گذار و آزمون تا حدودی یکسان بودند اما در جمعیت انتخاب‌شده تغییر این دو فراوانی منجر به برآوردهای متفاوت از واریانس ژنتیکی شد. همچنین همه واریانس‌های ژنتیکی برآوردشده در جمعیت دوم نسبت به متناظرهای آن‌ها در جمعیت اول کمتر است که این نیز ناشی از اثرگذاری کاهش واریانس بر اثرگذاری انتخاب است (Bulmer, 1976).

غیر قطری ماتریس واریانس-کوواریانس ژنتیکی به‌کار رفته در ساختار معادله‌های BLUP مرتبط است و این قاعده در هر دو جمعیت صادق است. بنابراین بالاترین واریانس ژنتیکی با استفاده از $Gnorm$ به دست آمد که در هر دو جمعیت کمترین تفاوت بین عنصرهای قطری و غیر قطری آن وجود داشت. با توجه به اینکه برآورد واریانس باقی‌مانده برای هر پنج ماتریس تا حدودی برابر است در نتیجه واریانس ژنتیکی بالاتر منجر به وراثت‌پذیری بالاتر (۰/۴۲ و ۰/۳۳) به ترتیب برای جمعیت یک و دو) برای $Gnorm$ شده است.

نتایج در جمعیت دوم به کلی تابع نوع ماتریس G به‌کار رفته است و هر کدام از G ها برآوردی متفاوت از σ_a^2 دارند. تنها برآوردهای نزدیک به هم را می‌توان برآوردهای ناشی از $Gfoun$ و $Gwei$ دانست.

جدول ۴. برآورد مؤلفه‌های واریانس \pm اشتباه استاندارد با استفاده از ماتریس‌های مختلف رابطه‌های خویشاوندی ژنتیکی

Table 4. Estimation of variance component \pm standard error by using different genomic relationship matrixes

		<i>Gfoun</i>	<i>Gref</i>	<i>G05</i>	<i>Gnorm</i>	<i>Gwei</i>
Population 1	σ_e^2	0.728 \pm 0.015	0.729 \pm 0.015	0.726 \pm 0.015	0.728 \pm 0.015	0.724 \pm 0.014
	σ_a^2	0.249 \pm 0.012	0.244 \pm 0.012	0.324 \pm 0.017	0.53 \pm 0.03	0.264 \pm 0.012
	h^2	0.255 \pm 0.011	0.251 \pm 0.011	0.31 \pm 0.014	0.42 \pm 0.016	0.267 \pm 0.011
Population 2	σ_e^2	0.777 \pm 0.012	0.778 \pm 0.012	0.776 \pm 0.012	0.778 \pm 0.012	0.770 \pm 0.012
	σ_a^2	0.173 \pm 0.03	0.156 \pm 0.03	0.224 \pm 0.04	0.386 \pm 0.06	0.182 \pm 0.03
	h^2	0.182 \pm 0.023	0.17 \pm 0.022	0.224 \pm 0.024	0.332 \pm 0.032	0.191 \pm 0.02

با این حال ضریب همبستگی به دست آمده با $Gwei$ اشتباه استاندارد پایین‌تری داشته و میزان کمینه آن بالاتر از $Gnorm$ است. به‌هر حال، همین اعداد از متناظرهای خود در جمعیت دوم دست‌کم ۸ درصد بالاتر بودند. اگرچه تفاوت بین ضریب‌های همبستگی با استفاده از ماتریس‌های ژنتیکی مختلف در جمعیت اول خیلی چشمگیر به نظر نمی‌رسد ولی مقایسه ضریب‌ها در جمعیت دوم به‌روشنی برتری $Gwei$ را در مقابل $Gfoun$ که به‌طور معمول به‌عنوان یک استاندارد استفاده می‌شود نشان می‌دهند. با این‌که به‌طور معمول این فراوانی در دسترس نیست ولی روش‌هایی برای برآورد این فراوانی‌ها پیشنهاد شده است (Gengler et al., 2007). در مقایسه با دیگر بررسی‌ها، (Aguilar et al., 2010) درستی بالاتری را با استفاده از فراوانی

جدول ۵ توصیف آماری ضریب‌های همبستگی بین ارزش‌های اصلاحی پیش‌بینی‌شده و حقیقی را نشان می‌دهد. مقایسه ضریب‌های همبستگی بین دو جمعیت نشان می‌دهد که همبستگی بین ارزش‌های اصلاحی حقیقی و پیش‌بینی‌شده با استفاده از G های مختلف همواره در جمعیت با آمیزش تصادفی بالاتر است. نتیجه کلی دیگر این است که استفاده از فراوانی آللی جمعیت مرجع برای مقیاس کردن G کمترین برتری را دارد و تفاوتی هم بین جمعیت‌های دارای آمیزش تصادفی و جمعیت‌های دارای انتخاب که به‌طور معمول در اصلاح دام با چنین جمعیت‌هایی سروکار داریم، نیست. بالاترین ضریب همبستگی در جمعیت اول با استفاده از $Gnorm$ به دست آمد که اختلاف آن با $Gwei$ در حد بسیار ناچیز ۰/۰۰۵ بود.

کردن G با A و مقیاس نکردن آن گزارش نکردند. ولی در بررسی آن‌ها استفاده از یک G مقیاس شده با A در قالب روش تک‌مرحله‌ای به‌طور میانگین برای ۱۶ صفت ۰/۳ درصد درستی بالاتری از روش تک‌مرحله‌ای که در آن G مقیاس نشده بود ارائه داد. در بررسی Forni *et al.* (2011) ضریب همبستگی بین ارزش اصلاحی پیش‌بینی‌شده (EBV) و ارزش اصلاحی پیش‌بینی‌شده ژنگانی (GEBV) با استفاده از فراوانی آلی ۰/۵ بالاتر (۰/۷۹) از استفاده از فراوانی آلی مشاهده‌شده در جمعیت مرجع و استفاده از *Gnorm* (۰/۷۸) گزارش شد (Forni *et al.*, 2011).

آلی ۰/۵ در مقایسه با استفاده از فراوانی آلی کنونی جمعیت آزمون و یا برآورد فراوانی آلی حیوان‌های بنیان‌گذار گزارش کردند (Aguilar *et al.*, 2010). در بررسی دیگر Chen *et al.* (2011) تفاوت خاصی بین استفاده از فراوانی آلی ۰/۵ برای همه جایگاه‌ها و استفاده از فراوانی آلی جمعیت مرجع برای صفات وزن بدن و گوشت سینه در جوجه‌های گوشتی مشاهده نکردند (Chen *et al.*, 2011). Gao *et al.* (2012) نیز دو مدل G-BLUP که یک اثرگذاری چندژن اضافی نیز در مدل‌ها وجود داشت را برای شانزده صفت بررسی کردند و هیچ تفاوتی بین مقیاس

جدول ۵. کمینه، بیشینه و میانگین \pm اشتباه استاندارد ضریب‌های همبستگی بین ارزش‌های اصلاحی حقیقی و پیش‌بینی‌شده

Table 5. Minimum, maximum and average \pm standard error of correlation coefficient between true and estimated breeding values

		<i>Gfoun</i>	<i>Gref</i>	<i>G05</i>	<i>Gnorm</i>	<i>Gwei</i>
Population 1	Minimum	0.594	0.628	0.601	0.637	0.685
	Maximum	0.852	0.833	0.853	0.88	0.86
	Average	0.737 \pm 0.007	0.723 \pm 0.007	0.735 \pm 0.007	0.745 \pm 0.008	0.740 \pm 0.005
Population 2	Minimum	0.384	0.377	0.376	0.415	0.436
	Maximum	0.84	0.847	0.839	0.865	0.854
	Average	0.644 \pm 0.016	0.65 \pm 0.017	0.653 \pm 0.014	0.65 \pm 0.016	0.663 \pm 0.013

اریبی پیش‌بینی‌ها با استفاده از همه مدل‌ها در جمعیت دوم بدون استثناء میزانی افزایش یافت. البته با این توضیح که در رابطه با استفاده از *Gwei* اریبی به سمت بیش برآورد^۲ ارزش‌های اصلاحی حقیقی حرکت کرده است. با این حال این بیش برآوردی تنها ۰/۰۰۹ بوده و در عمل قابل چشم‌پوشی است و همسان جمعیت اول استفاده از ماتریس A برای مقیاس کردن G بهترین نتیجه را که منجر به تولید برآوردی ناریب می‌شود در پی دارد. ارزش‌های اصلاحی ژنگانی پیش‌بینی‌شده متورم یک موضوع بسیار نگران‌کننده در ارزیابی‌های ژنگانی است که به‌طور عمده مرتبط با مخلوط کردن نادرست داده‌های حیوان‌های نژادگان‌شده و نژادگان‌نشده و دارای رکورد پدیدگانی است (Meuwissen *et al.*, 2011; Vitezica *et al.*, 2011). در چنین شرایطی به‌طور معمول ارزش‌های اصلاحی ژنگانی متورم، ارزش بیشتری به حیوان‌های جوان در مقایسه با حیوان‌های پیرتر که با روش‌های سنتی تجزیه و تحلیل شده و تعیین نژادگان

جدول ۶ عرض از مبدأ، ضریب رگرسیون و ضریب تعیین پیش‌بینی ارزش‌های اصلاحی حقیقی با استفاده از پیش‌بینی‌های ارزش اصلاحی با مدل‌های مختلف G-BLUP را در دو جمعیت نشان می‌دهد. ضریب رگرسیون در جمعیت اول بین ۱/۰۰۸ تا ۱/۰۳۳ تغییر داشت که همگی مقادیر قابل قبولی بوده و تنها میزان بسیار ناچیزی کم برآورد^۱ در پیش‌بینی ارزش اصلاحی حقیقی به چشم می‌خورد. میانگین اریبی در پنج مدل ۰/۰۲۶ بود. در این جمعیت *Gwei* پیش‌بینی‌هایی ناریب تولید می‌کند (میزان اریبی بسیار ناچیز ۰/۰۰۸ و قابل چشم‌پوشی است)، *G05* در رتبه بعدی قرار دارد، *Gfoun* و *Gnorm* به‌طور مشترک برآوردی یکسان ارائه می‌دهند و در نهایت اریب‌ترین پیش‌بینی (۰/۰۳۳) با به‌کارگیری *Gref* به‌دست می‌آید. ضریب رگرسیون مورد انتظار در مدل ارزیابی در غیاب هرگونه پیش‌انتخاب برای تعیین نژادگان کردن حیوان‌ها نزدیک به یک است (Mäntysaari *et al.*, 2010).

2. Overestimate

1. Underestimate

نشده‌اند می‌دهد. در جمعیت دوم نیز استفاده از فراوانی آلی جمعیت مرجع برای مقیاس کردن ماتریس G اریب‌ترین پیش‌بینی‌ها را تولید کرد. در بررسی داده‌های واقعی در گاو برای تشکیل ماتریس G با استفاده از فراوانی‌های آلی متفاوت استفاده شد و با به‌کارگیری این Gها در ساختار معادله‌های G-BLUP نشان داده شد، ارزش‌های اصلاحی ژنگانی پیش‌بینی‌شده برای تولید

نشده‌اند می‌دهد. در جمعیت دوم نیز استفاده از فراوانی آلی جمعیت مرجع برای مقیاس کردن ماتریس G اریب‌ترین پیش‌بینی‌ها را تولید کرد. در بررسی داده‌های واقعی در گاو برای تشکیل ماتریس G با استفاده از فراوانی‌های آلی متفاوت استفاده شد و با به‌کارگیری این Gها در ساختار معادله‌های G-BLUP نشان داده شد، ارزش‌های اصلاحی ژنگانی پیش‌بینی‌شده برای تولید

جدول ۶. عرض از مبدأ، شیب خط رگرسیون و ضریب تعیین تابعیت ارزش‌های اصلاحی حقیقی بر پیش‌بینی‌شده

Table 6. Intercept, coefficient of regression and determination coefficient of true breeding values on estimated breeding values

		<i>Gfoun</i>	<i>Gref</i>	<i>G05</i>	<i>Gnorm</i>	<i>Gwei</i>
Population 1	b0	-0.005±0.055	-0.037±0.046	0.005±0.036	-0.038±0.057	-0.009±0.018
	b1	1.031±0.033	1.033±0.033	1.025±0.033	1.031±0.033	1.008±0.03
	R ²	0.63±0.015	0.63±0.015	0.63±0.015	0.63±0.014	0.63±0.014
Population 2	b0	0.825±0.043	0.833±0.026	0.845±0.025	0.867±0.031	1.067±0.038
	b1	1.048±0.034	1.063±0.043	1.031±0.037	1.051±0.04	0.991±0.034
	R ²	0.52±0.04	0.52±0.041	0.52±0.04	0.52±0.04	0.54±0.038

b0: Intercept, b1: Coefficient of regression, R²: Determination coefficient

به‌طورمعمول فراوانی آلی جمعیت پایه به‌عنوان یک استاندارد برای مقیاس کردن ماتریس G توصیه می‌شود ولی در عمل این فراوانی در دست نبوده و برآوردهای آن‌ها باید استفاده شوند بنابراین خطای برآورد این فراوانی ممکن است موجب اریب شدن رابطه‌های خویشاوندی و در نتیجه برآورد نادرست مؤلفه‌های واریانس و ارزش‌های اصلاحی شود. ولی استفاده از ماتریس A برای تشکیل ماتریس وزن‌شده *Gwei* نه تنها این مشکل را حل می‌کند بلکه به‌مراتب نتایج درست‌تر و نااریب، حتی بهتر از *Gfoun* تولید می‌کند. استفاده از فراوانی آلی مشاهده‌شده در جمعیت مرجع *Gref* نسبت به دیگر مقیاس‌ها نتایجی اریب حتی در جمعیت‌های بدون انتخاب که احتمال نداشتن تغییر فراوانی آلی در آن‌ها وجود دارد تولید می‌کند. استفاده از فراوانی آلی ۰/۵ نیز به‌عنوان یک گزینه جایگزین دیگر در جمعیت‌های دارای انتخاب می‌تواند مطرح باشد. با این حال استنباط‌های دقیق‌تر هنگامی به دست خواهد آمد که بررسی‌هایی از این قبیل روی جامعه‌های مختلف با درجه‌های متفاوتی از گزینش و آمیزش‌های تصادفی و همچنین صفاتی با وراثت‌پذیری و شمار QTLهای مختلف انجام شود که نیاز به تحقیقات جامع‌تر و بیشتری است.

نتیجه‌گیری

بهبود ژنتیکی صفات مختلف به‌عنوان مهم‌ترین هدف برنامه‌های اصلاح نژادی در دام و طیور هنگامی فراهم است که پیش‌بینی شایستگی ژنتیکی حیوان‌ها با آنچه در عمل رخ می‌دهد بیشترین نزدیکی را داشته باشد. امروزه با دسترسی به انواع تراشه‌های تجاری DNA امکان اندازه‌گیری بهتر همانندی ژنتیکی و در نتیجه پیش‌بینی دقیق‌تر ارزش‌های ژنتیکی حیوان‌ها فراهم شده است. هدف این بررسی به دست آوردن رابطه‌های خویشاوندی در G بود که نسبت به همان نسل پایه استفاده‌شده در A تعریف‌شده باشند. چنین ماتریسی امکان مقایسه حیوان‌های نژادگان‌شده و نژادگان‌نشده ولی دارای رکورد پدیدگانی در صورتی که به‌طور جداگانه نیز تجزیه و تحلیل شوند را بهتر فراهم می‌کند. نتایج این تحقیق نشان داد، ساختار ماتریس G اثر مستقیم بر قابلیت پیش‌بینی مدل‌ها و برآورد مؤلفه‌های واریانس دارد. اگرچه قابلیت پیش‌بینی مدل‌ها در جمعیت‌های با آمیزش تصادفی به‌عنوان یک جمعیت شاهد در این بررسی خیلی تفاوت نداشت که البته هنوز هم این تفاوت بسیار کم می‌تواند مهم باشد ولی در جمعیت انتخاب‌شده، به‌عنوان جمعیت هدف یک اصلاحگر، تفاوت‌ها در برآورد اجزای واریانس بسیار محسوس بود.

REFERENCES

1. Aguilar, I., Misztal, I., Johnson, D., Legarra, A., Tsuruta, S. & 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, 743-752.
2. Bulmer, M. (1976). The effect of selection on genetic variability: a simulation study. *Genetical Research*, 28, 101-117.
3. Calus, M. (2010). Genomic breeding value prediction: methods and procedures. *Animal*, 4, 157-164.
4. Chen, C.-Y., Misztal, I., Aguilar, I., Legarra, A. & Muir, W. (2011). Effect of different genomic relationship matrices on accuracy and scale. *Journal of animal science* 89, 2673-2679.
5. Christensen, O. F., Madsen, P., Nielsen, B., Ostersen, T. & Su, G. (2012). Single-step methods for genomic evaluation in pigs. *Animal*, 6, 1565-1571.
6. Daetwyler, H. D., Pong-Wong, R., Villanueva, B. & Woolliams, J. A. (2010). The impact of genetic architecture on genome-wide evaluation methods. *Genetics*, 185, 1021-1031.
7. Falconer, D. & Mackay, T. (1995). *Introduction to Q quantitative Genetics*. Longman 19, 1.
8. Forni, S., Aguilar, I. & Misztal, I. (2011). Different genomic relationship matrices for single-step analysis using phenotypic, pedigree and genomic information. *Genetics Selection Evolution*, 43.
9. Gao, H., Christensen, O. F., Madsen, P., Nielsen, U. S., Zhang, Y., Lund, M. S. & 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, 10.1186.
10. Gengler, N., Mayeres, P. & Szydlowski, M. (2007). A simple method to approximate gene content in large pedigree populations: application to the myostatin gene in dual-purpose Belgian Blue cattle. *Animal*, 1, 21-28.
11. Habier, D., Fernando, R. & Dekkers, J. (2007). The impact of genetic relationship information on genome-assisted breeding values. *Genetics*, 177, 2389-2397.
12. Habier, D., Tetens, J., Seefried, F.-R., Lichtner, P. & Thaller, G. (2010). The impact of genetic relationship information on genomic breeding values in German Holstein cattle. *Genetics Selection Evolution*, 42, 5.
13. Harris, B. & Johnson, D. (2010). Genomic predictions for New Zealand dairy bulls and integration with national genetic evaluation. *Journal of Dairy Science*, 93, 1243-1252.
14. Hayes, B. & Goddard, M. (2008). Technical note: prediction of breeding values using marker-derived relationship matrices. *Journal of Animal Science*, 86, 2089-2092.
15. Henderson, C. (1984). *Applications of linear models in animal breeding* (University of Guelph, Guelph, ON, Canada).
16. Makgahlela, M., Strandén, I., Nielsen, U., Sillanpää, M. & Mäntysaari, E. (2013). The estimation of genomic relationships using breedwise allele frequencies among animals in multibreed populations. *Journal of Dairy Science*, 96, 5364-5375.
17. Mäntysaari, E., Liu, Z. & VanRaden, P. (2010). *Interbull validation test for genomic evaluations*. Interbull bulletin, 17.
18. Meuwissen, T., Luan, T. & Woolliams, J. (2011). The unified approach to the use of genomic and pedigree information in genomic evaluations revisited. *Journal of Animal Breeding and Genetics*, 128, 429-439.
19. Nejati-Javaremi, A., Smith, C. & Gibson, J. (1997). Effect of total allelic relationship on accuracy of evaluation and response to selection. *Journal of Animal Science*, 75, 1738-1745.
20. Sargolzaei, M. & Schenkel, F.S. (2009). QMSim: a large-scale genome simulator for livestock. *Bioinformatics*, 25, 680-681.
21. Tiezzi, F. & Maltecca, C. (2015). Accounting for trait architecture in genomic predictions of US Holstein cattle using a weighted realized relationship matrix. *Genetics Selection Evolution* 47, 24.
22. VanRaden, P. (2008). Efficient methods to compute genomic predictions. *Journal of Dairy Science*, 91, 4414-4423.
23. Vitezica, Z., Aguilar, I., Misztal, I. & Legarra, A. (2011). Bias in genomic predictions for populations under selection. *Genetics Research*, 93, 357-366.