

## اهمیت استفاده از شجره و نقش افراد ژنوتیپ نشده بر لزوم به کارگیری مجدد انتخاب ژنومی در صفات آستانه‌ای دودویی

یوسف نادری\*

دانشیار، گروه علوم دامی، دانشگاه آزاد اسلامی، واحد آستارا، آستارا، ایران  
تاریخ دریافت: ۱۴۰۰/۱۱/۱۴ - تاریخ پذیرش: ۱۴۰۰/۱۲/۱۴

### چکیده

مدل‌های پیش‌بینی ژنومی تک مرحله‌ای با توجه به عملکرد بالا، هزینه‌های بالای تعیین ژنوتیپ برخی از حیوانات و برآورد ارزش اصلاحی ژنومی همزمان برای حیوانات ژنوتیپ شده و نشده، تبدیل به ابزاری غالب در ارزیابی ژنومی دام‌های اهلی شدند. هدف از تحقیق حاضر، بررسی نقش روابط خویشاوندی ژنومی در جمعیت‌های مرجع و تأیید بر عملکرد روش‌های چند مرحله‌ای و تک مرحله‌ای GBLUP و BayesB بود، که در این مسیر اهمیت افراد ژنوتیپ نشده نسل‌های بعد از انتخاب و نقش آنها در لزوم یا عدم لزوم به کارگیری مجدد انتخاب ژنومی در صفات آستانه‌ای دودویی نیز بررسی شد. بدین منظور، جمعیت‌های ژنومی برای تعداد مختلف جایگاه‌های صفات کمی (۳۳ و ۵۳۴) با دو توزیع متفاوت آثار ژنی (نرمال و گاما) روی ۲۹ کروموزوم شبیه‌سازی شدند. برای آنالیز داده‌های شبیه‌سازی شده چهار مدل آماری شامل GBLUP، SS-GBLUP، BayesB و SS-BayesB به کار گرفته شد. بر اساس نتایج بدست آمده، صحت پیش‌بینی ژنومی تحت تأثیر ارتباط بین جمعیت‌های مرجع و تأیید قرار گرفت و برای مدل‌های ژنومی چند مرحله‌ای در مقایسه با تک مرحله‌ای به طور معنی‌داری بیشتر بود. هنگامی که صفات تحت تأثیر تعداد کم QTL بودند، روش‌های تک مرحله‌ای و چند مرحله‌ای BayesB در مقایسه با GBLUP و SS-GBLUP صحت ارزیابی بالاتری نشان دادند. در توزیع نرمال اثر QTL و صفات تحت کنترل تعداد زیاد QTL، روش‌های BayesB (به‌ویژه روش چند مرحله‌ای BayesB) کمترین میزان صحت ژنومی را نشان دادند. از آنجایی که افزایش فاصله نسل یک چالش اساسی در انتخاب ژنومی صفات آستانه‌ای می‌باشد، در نتیجه استفاده از روش‌های تک مرحله‌ای، عدم لزوم به کارگیری مجدد تعیین ژنوتیپ جمعیت مرجع را در طی نسل‌های اولیه انتخاب میسر می‌سازند. با این حال، استفاده از روش تک مرحله‌ای BayesB و سودمند بودن آن در ارزیابی ژنومی، منوط به شناسایی معماری ژنتیکی صفت است.

واژه‌های کلیدی: اطلاعات شجره‌ای، روابط خویشاوندی ژنومی، فاصله نسلی، BayesB، GBLUP.

## The importance of pedigree information and the role of un-genotyped individuals on the necessity of repeating the genomic selection of binary threshold traits

Yousef Naderi\*

Associate Professor, Department of Animal Science, Astara Branch, Islamic Azad University, Astara, Iran  
(Received: Feb. 3, 2022 - Accepted: Mar. 5, 2022)

### ABSTRACT

Single-step genomic prediction models became the prevailing tool in genetic evaluations of livestock due to high genomic performance, high genotyping costs for some animals, and utilizing both genotyped and un-genotyped animals. The objective of the current study was to investigate the effect of genetic relationships between the training and validation populations on the performance of single and multi-step GBLUP and BayesB, which in this regard the importance of un-genotyped individuals in generations after genomic selection and their roles on the necessity of reusing genomic selection in binary threshold traits was investigated. For this purpose, genomic populations were simulated to reflect the different number of QTL (33 and 534) with two different distributions of QTL (normal and gamma) on 29 chromosomes. Four statistical models (GBLUP, SS-GBLUP, BayesB and SS-BayesB) were implemented to analyze the simulated data. According to the results, GEBV accuracy was influenced by the relationships between the training and validation populations more significantly for multi-step models than for single-step models. Multi-step and single-step BayesB methods showed an obvious advantage over GBLUP and SSGBLUP when the trait is controlled by fewer QTLs. For the normal QTL effects distribution, BayesB methods (especially multi-step BayesB) obtained the lower accuracy than that of more QTLs affecting the trait. Generally, since the increasing generation distance is a major challenge in the genomic selection of threshold traits, therefore the use of single-step methods makes possible reuse of genomic selection during the early generations of selection unnecessary. However, using single-step BayesB methods and its advantages in genomic evaluation depends on the identification of trait genetic architecture.

**Keywords:** BayesB, GBLUP, Genetic Relationships, Generation Distance, Pedigree Information.

\* Corresponding author E-mail: y.naderi@iau-astara.ac.ir; yousefnaderi@gmail.com

### مقدمه

با توجه به افزایش دسترسی به تراشه‌های چند شکلی نوکلئوتیدی (SNP)، انتخاب ژنومی به سرعت در اصلاح ژنتیکی حیوانات اهلی مورد استفاده قرار گرفت و به عنوان یک روش مناسب جایگزین روش سنتی (BLUP) شد (Lee et al., 2017). وجود همبستگی منفی بین صفات پیوسته با صفات آستانه‌ای از جمله مقاومت به بیماری‌ها، سخت‌زایی و صفات رفتاری، منجر به کاهش شایستگی ژنتیکی حیوان خواهد شد (Naderi & Sadeghi, 2019). پیشرفت قابل توجه در افزایش سوددهی اقتصادی در برنامه‌های اصلاح نژادی مدرن نیازمند فهم بهتر و ورود مستقیم به صفات با بروز فنوتیپی گسسته می‌باشد، که دلیل این امر را می‌توان ماهیت آستانه‌ای این نوع صفات، اثرپذیری به وسیله ژن‌های متعدد و عدم تطابق با توزیع نرمال و وراثت مندلی دانست (Sadeghi et al., 2018).

استفاده مفید از انتخاب ژنومی و کاربرد آن در ارزیابی ارزش‌های اصلاحی حیوانات همگام با توسعه فناوری تعیین ژنوتیپ، به فهم بالای روش‌های آماری مورد استفاده وابسته است. روش‌های آماری مختلف براساس رابطه بین دو متغیر مستقل (ژنوتیپ) و وابسته (فنوتیپ)، آثار نشانگرها را برآورد می‌کنند. در این مسیر عملکرد هر یک از روش‌های آماری با توجه به ساختار ژنتیکی صفت در جمعیت‌های مختلف متغیر است و بررسی کارایی روش‌های مختلف آماری را در جمعیت‌های مختلف الزامی می‌سازد (Clark et al., 2012). بهترین پیش‌بینی نااریب خطی ژنومی (GBLUP) و روش‌های Bayes از متداول‌ترین روش‌های آماری چند مرحله‌ای در ارزیابی ژنومی هستند که امروزه نسخه‌های آستانه‌ای آن‌ها گسترش یافته است (González-Recio & Forni, 2011; VanRaden & Sullivan, 2010) و به ترتیب از ساختار واریانس-کواریانس و رگرسیون مستقیم فنوتیپ به نشانگرها برای رسیدن به تابعی از روابط ژنومی بهره می‌گیرد (Sadeghi et al., 2018).

با توجه به اینکه نسبت محدودی از حیوانات یک جمعیت که دارای اطلاعات شجره‌ای و فنوتیپی هستند، حاوی اطلاعات ژنومی می‌باشند، استفاده از روش‌های

چند مرحله‌ای در ارزیابی دارای محدودیت است. بنابراین استفاده از روش‌های تک مرحله‌ای که در آن همزمان از اطلاعات ژنومی، فنوتیپی و شجره‌ای جهت برآورد ارزش‌های اصلاحی حیوانات ژنوتیپ شده و نشده استفاده می‌شود، نقش برجسته‌ای در پیشبرد اهداف اصلاحی ایفا می‌کند (Legarra et al., 2009). مطالعات مختلف در مورد ارزیابی ژنومی جوجه گوشتی، گاو شیری و گوشتی نشان از عملکرد بالای صحت روش‌های تک مرحله‌ای داشت (Aguilar et al., 2010; Chen et al., 2011; Liu et al., 2014). مطالعات نامبرده نشان می‌دهند که در روش‌های تک مرحله‌ای GBLUP (ssGBLUP) صحت ارزیابی ارزش اصلاحی پیش‌بینی شده برای افراد ژنوتیپ شده و ژنوتیپ نشده در مقایسه با روش GBLUP که فقط قادر به پیش‌بینی صحت ژنومی برای افراد ژنوتیپ شده است، بیشتر می‌باشد (Christensen OF & Lund, 2010). در مطالعه‌ای دیگر، روش ssGBLUP با بهره‌گیری از ماتریس روابط خویشاوندی (اطلاعات شجره‌ای و نشانگری) عملکرد مشابه و یا بالاتری در مقایسه با روش‌های چند مرحله‌ای نشان داد (Lourenco et al., 2015). همچنین در تحقیقی دیگر، عملکرد بالای صحت ارزیابی ژنومی در روش‌های تک مرحله‌ای بیزی نسبت به روش‌های چند مرحله‌ای بیزی گزارش شده است (Fernando et al., 2016).

با توجه به مطالب گفته‌شده، هدف از تحقیق حاضر بررسی عملکرد روش‌های تک‌مرحله‌ای GBLUP، BayesB و چند مرحله‌ای در ترکیبات مختلف ژنومی از جمله دو توزیع مختلف آثار ژنی، تعداد QTL در جمعیتی با روابط خویشاوندی ژنومی متفاوت، در نسل‌های مختلف می‌باشد. همچنین اهمیت استفاده از شجره و نقش افراد ژنوتیپ نشده بر لزوم یا عدم لزوم به کارگیری مجدد انتخاب ژنومی ناشی از کاهش عدم تعادل پیوستگی و ضعیف شدن ارتباط افراد جمعیت مرجع و تایید در نسل‌های مختلف مورد بررسی قرار گرفت.

### مواد و روش‌ها

به منظور شبیه‌سازی از نرم‌افزار QMSim استفاده شد (Sargolzaei & Schenkel et al., 2009). نخست،

وراثت‌پذیری پایین صفات آستانه‌ای (از جمله صفات بیماری و تولید مثلی) وراثت‌پذیری صفت ۰/۱ در نظر گرفته شد. با توجه به اینکه صفات آستانه‌ای بیشتر وابسته به عوامل محیطی بوده و در طی زمان توسط برخی جهش‌های ناشی از تغییر نوکلئوتیدی ایجاد شده‌اند در نتیجه جهت عدم حذف این نشانگرها، نرخ فراوانی آللی کمیاب برای این صفات مقدار پایین ۰/۱ در نظر گرفته شد.

با توجه به شبیه‌سازی اولیه جمعیت بر پایه فنوتیپ پیوسته در نرم‌افزار QMSim تغییراتی در فایل فنوتیپ خروجی این نرم افزار جهت ایجاد فنوتیپ آستانه‌ای دودویی اعمال شد. به طوریکه فنوتیپ پیوسته حیوانات جمعیت مرجع به عنوان متغیر پاسخ ( $y$ ) از طریق عوامل ثابت یا مستقل (اثر نسل  $x_1$  و اثر جنسیت  $x_2$ ) در نرم افزار R آنالیز واریانس شد و اثر عوامل ثابت و تاثیرگذار بر فنوتیپ تصحیح گردید و در نهایت مقادیر باقی‌مانده جهت بررسی نرمال بودن و کد دهی به متغیر پاسخ آستانه‌ای مورد استفاده قرار گرفت. به طوری که برای شبیه‌سازی فنوتیپ آستانه‌ای دودویی در جمعیت مرجع، ابتدا باقی‌مانده‌ها از بیشترین به کم‌ترین مرتب شدند. در نهایت آستانه یک برای حیوانات پایین‌تر از میانگین باقی‌مانده‌ها و آستانه صفر برای حیوانات بالاتر از میانگین باقی‌مانده‌ها در نظر گرفته شد.

### روش‌های آماری

#### GBLUP

جهت اجرای مدل‌های مختلط خطی تعمیم یافته در آنالیز صفات آستانه‌ای با استفاده از روش GBLUP، الگوریتم میانگین اطلاعات بیشترین درست‌نمایی محدودشده (AI-REML) از بسته نرم افزاری DMU استفاده شد. مدل آماری استفاده‌شده در روش GBLUP به صورت رابطه (۱) بود:

$$\text{logit}(\pi_r) = \log\left[\frac{\pi_r}{1-\pi_r}\right] = \emptyset + \gamma_r \quad (1)$$

در رابطه (۱)،  $\pi_r$  احتمال شیوع آستانه (در اینجا ۰/۵) برای حیوان  $r$ ،  $\emptyset$  اثر میانگین کل، و  $\gamma_r$  اثر تصادفی حیوان است. آثار تصادفی حیوانات با

یک جمعیت پایه ۲۰۰۰ راسی برای ۱۰۰۰ نسل شبیه‌سازی شد. برای ایجاد عدم تعادل پیوستگی در جمعیت فوق الذکر، پس از شبیه‌سازی جمعیت پایه، تعداد افراد جمعیت از طریق ایجاد یک گلوگاه ژنتیکی به ۲۰۰ رأس برای ۱۰۰ نسل (در نسل‌های ۸۰۱ تا ۹۰۰) کاهش یافت. سپس در آخرین جمعیت پایه، بعد از ۱۰۰ نسل (از نسل ۹۰۱ تا نسل ۱۰۰۰) تعداد افراد جمعیت به ۱۰۰۰ رأس افزایش داده شد. برای ایجاد جمعیت مرجع و تأیید، همه افراد آخرین نسل جمعیت پایه برای تولید مثل در جمعیت، مورد استفاده قرار گرفتند که در این بین ۳۰ رأس نر در نظر گرفته شد تا منعکس‌کننده‌ی نسبت نر به ماده‌ی موجود در گله‌های گاو شیری باشد و بتوان اثر تکنیک تلقیح مصنوعی بر نسبت نر به ماده را تقلید کرد (Yin *et al.*, 2014). نوع سیستم تلاقی تصادفی بود و برای ۱۰ نسل دیگر (تا نسل ۱۰۱۰) جمعیت تکثیر شد. شانس تلاقی در همه‌ی حیوانات برابر (در هر دو جنس) و یک فرزند برای هر زایش در نظر گرفته شد. درصد جایگزینی برای نرها و ماده‌ها به ترتیب ۵۰ و ۲۰ درصد بود. در جمعیت اخیر، انتخاب حیوانات برتر برای نسل بعد براساس ارزش اصلاحی بالا و معیار حذف بر اساس ارزش اصلاحی پایین و سن صورت گرفت. نشانگرها به صورت دو آللی و به صورت تصادفی برای هر یک از کروموزوم‌ها (۲۹ کروموزوم در دامنه‌ی ۴۲ تا ۱۵۸ سانتی مورگان) توزیع شدند. با توجه به اندازه کل ژنوم گاو به ازای هر کروموزوم تعداد متفاوت نشانگر (دامنه‌ی ۱۸۰ تا ۶۵۳ با توجه به نوع و اندازه کروموزوم) با هدف تولید پنل‌های ۱۰K شبیه‌سازی شد (Kappes *et al.*, 1997).

جهت شبیه‌سازی جمعیت واقعی و با توجه به پایگاه داده مربوط به گاو تعداد QTLها با توجه به دو نوع صفت آستانه‌ای بیماری مانند حساسیت به بیماری تنفسی (با تعداد QTL ۵۳۴) و اسهال (با تعداد QTL ۳۳) در طول کروموزوم‌ها و با توجه نوع کروموزوم و نقشه ژنتیکی آنها توزیع شدند. اثر QTLها از دو توزیع نرمال و گاما شبیه‌سازی شد. نرخ جهش برای نشانگرها و QTLها در هر جایگاه و در هر نسل  $2/5 \times 10^{-5}$  فرض شد (Bermann *et al.*, 2020). همچنین با توجه به

در اینجا  $\sigma_e^2$ ،  $\sigma_a^2$  و  $\sigma_g^2$  به ترتیب واریانس باقی‌مانده‌ها، نشانگرها و پلی ژنی هستند.

### BayesB

مدل کلی روش Bayes در انتخاب ژنومی (Meuwissen & González-Recio, 2001) و مدل آستانه‌ای آن (Forni, 2011) به صورت ذیل است:

$$y_i = \mu + \sum_{j=1}^m X_{ij} b_j + e_i \quad (5)$$

در معادله (5)،  $y_1$  بردار مشاهدات شامل افراد تعیین ژنوتیپ شده،  $\mu$  عرض از مبدا،  $X_{ij}$  ژنوتیپ نشانگرها،  $b$  اثر نشانگری و  $e$  اثر باقیمانده هستند. توزیع چگالی احتمال توام پیشین مجهولات رابطه فوق به صورت زیر است:

$$(6)$$

$$P(\mu, b, \sigma^2 | df, s, \omega) \sim \left\{ \prod_{i=1}^m P(b_i | \theta_{b_i}, \sigma^2) P(\theta_{b_i} | \omega) \right\} X^{-2} (\sigma^2 | df, s)$$

در رابطه (6)، توزیع پیشین یکنواخت برای میانگین،  $\chi^2(\sigma^2 | df, s)$  توزیع کای اسکور معکوس مقیاس‌دار برای واریانس باقی مانده با درجه آزادی  $df$  و پارامتر مقیاس برابر با  $S$  در نظر گرفته شد.  $P(b_j | \theta_{b_j}, \sigma^2)$  توزیع پیشین ز امین نشانگر را نشان می‌دهد که در آن  $\theta_{b_j}$  یک برداری از پارامترها است که نوع توزیع پیشین در نظر گرفته برای آثار نشانگری را مشخص می‌کند.  $P(\theta_{b_j} | \omega)$  توزیع پیشین در نظر گرفته شده برای  $\theta_{b_j}$  و  $\omega$  پارامتری است که این نوع توزیع را مشخص می‌کند. توزیع پسین ناشناخته‌های مدل به صورت رابطه (7) است:

$$P(\mu, b, \sigma^2 | y, df, s, \omega) \sim \prod_{i=1}^m N\left(y_i \mid \mu + \sum_{j=1}^m X_{ij} b_j, \sigma^2\right) \times \left\{ \prod_{j=1}^m P(\theta_{b_j} | \omega) \right\} X^{-2} (\sigma^2 | df, s) \quad (7)$$

در روش BayesB آثار نشانگری با احتمال  $\pi$  برابر با صفر و با احتمال  $1-\pi$  با استفاده از توزیع  $t$  به دست می‌آیند. به این صورت که:

$$b_j | df_b, S_b, \pi \sim \text{IID} \begin{cases} 0 \text{ with probability } \pi \\ t(df_b, S_b) \text{ with probability } 1 - \pi \end{cases} \quad (8)$$

در این مطالعه میزان  $\pi$  برابر با ۰/۹۹ در نظر گرفته شد. با در نظر گرفتن این مقدار صحت برآورد نشانگرها بالا رفته و چون از تمامی نشانگرها در مدل استفاده نمی‌شود به علت آزمون‌ها چندگانه همزمان مسئله بیش برآورد نیز حل می‌شود و اثر سایر نشانگرها را بر اساس توزیع پیشین منقبض می‌کند.

در نظر گرفتن روابط ژنومی بین حیوانات و براساس داده‌های نشانگری در مدل لحاظ شد. ماتریس روابط خویشاوندی ژنومی ( $G$ ) بر اساس مدل پیشنهادی (VanRaden, 2008) و با استفاده از نرم‌افزار ماتریس  $G$  (Su & Madsen, 2013) و با افزودن مقدار ۰/۰۱ به عناصر قطری این ماتریس ساخته شد. در این مدل اثر نشانگرها به صورت تصادفی با واریانس یکسان در نظر گرفته شد.

### SS-GBLUP

مدل کلی روش SS-GBLUP (Legarra et al., 2009) به صورت رابطه (۲) است:

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} \beta + \begin{bmatrix} Z_1 & 0 \\ 0 & Z_2 \end{bmatrix} \begin{bmatrix} g_1 \\ g_2 \end{bmatrix} + e \quad (2)$$

در اینجا  $y_1$  و  $y_2$  به ترتیب بردار فنوتیپ برای افراد ژنوتیپ نشده و ژنوتیپ شده،  $\beta$  ماتریس عوامل ثابت،  $X_1$  و  $X_2$  به ترتیب ماتریس طرح آثار ثابت برای افراد ژنوتیپ نشده و شده،  $Z_1$  و  $Z_2$  به ترتیب ماتریس طرح برای افراد ژنوتیپ نشده و شده و  $\hat{\alpha} + \varepsilon = g_1 = A_{12} A_{22}^{-1} T_2 \hat{\alpha}$  و  $g_2 = T_2 \hat{\alpha}$  که در اینجا  $T_2$  ماتریس ژنوتیپ مشاهده شده مقیاس‌دار افراد ژنوتیپ شده است.  $\hat{\alpha}$  بردار آثار نشانگرهای برآورد شده و  $A_{ij}$  زیرماتریس از ماتریس خویشاوندی شجره‌ای در ارتباط با  $g_1$  و  $g_2$  می‌باشد. ماتریس واریانس کواریانس  $g_1$  و  $g_2$  (H) به صورت رابطه (۳) تعریف می‌شود (Legarra et al., 2009):

$$H = \begin{bmatrix} A_{11} + A_{12} A_{22}^{-1} (G_2 - A_{22}) A_{22}^{-1} A_{12} & A_{12} A_{22}^{-1} G_2 \\ G_2 A_{22}^{-1} A_{12} & G_2 \end{bmatrix} \quad (3)$$

$G_2$  ماتریس خویشاوندی ژنومی برای افراد ژنوتیپ شده است. به طور کلی آثار نشانگرهای برآورد شده توزیع نرمال  $(N(0, I\sigma_a^2))$  و باقیمانده ( $e$ ) توزیع نرمال چند متغیره مفروض هستند. معادلات مختلط (MME) برای آثار نشانگرهای روش SS-GBLUP به صورت رابطه (۴) است:

$$\begin{bmatrix} X'X & X'W & X'Z_1 \\ W'X & W'W + I \frac{\sigma_e^2}{\sigma_a^2} & W'Z_1 \\ Z_1'X & Z_1'W & Z_1'Z_1 + A_{11} \frac{\sigma_e^2}{\sigma_a^2} \end{bmatrix} \begin{bmatrix} \beta \\ \hat{\alpha} \\ \hat{\varepsilon} \end{bmatrix} = \begin{bmatrix} X'y \\ W'y \\ Z_1'y_1 \end{bmatrix} \quad (4)$$

کاهش یافت. این میزان کاهش برای روش‌های چند مرحله‌ای GBLUP و BayesB معنی‌دار بود. با این حال حساسیت روش GBLUP در مقایسه با روش BayesB به افزایش فاصله نسلی (کاهش ارتباط بین جمعیت مرجع و تایید) بیشتر بود. همچنین روش تک‌مرحله‌ای GBLUP (ss-GBLUP) در مقایسه با روش ss-BayesB افت صحت بیشتری با افزایش فاصله نسل نشان داد. به طوری که با افزایش فاصله نسلی (از ۱۰۰۹ به ۱۰۱۰) تفاوت معنی‌داری بین صحت ژنومی روش ss-BayesB برای سطوح مختلف QTL مشاهده نشد. این در حالی بود که با افزایش فاصله نسل (از ۱۰۰۹ به ۱۰۱۰) برای سناریوهای با تعداد پایین QTL، صحت ژنومی روش ss-GBLUP کاهش معنی‌داری نشان داد. اما در مجموع این روش در مقایسه با روش‌های چند مرحله‌ای حساسیت کمتری به افزایش فاصله نسل نشان داد.

#### اثر مدل آماری بر صحت ارزیابی ژنومی

شکل ۲ نمودار صحت پیش‌بینی ارزیابی ژنومی را برای هریک از مدل‌های GBLUP، BayesB، ss-GBLUP و ss-BayesB به تفکیک نسل نشان می‌دهد. در سناریوهای با تعداد پایین QTL همواره عملکرد روش ss-BayesB نسبت به روش ss-GBLUP و روش BayesB نسبت به GBLUP بیشتر بود. در مجموع در سناریوهای با تعداد پایین QTL روش ss-BayesB بهترین عملکرد را داشت و با افزایش فاصله نسل این عملکرد مثبت مشهودتر بود. در تعداد بالای QTL و در توزیع نرمال آثار QTL عملکرد روش‌های تک مرحله‌ای و چند مرحله‌ای GBLUP در مقایسه با روش‌های متناظر آنها برای BayesB بالاتر بود. با این حال با افزایش فاصله نسل، برتری روش ss-GBLUP در مقایسه با روش چند مرحله‌ای GBLUP و روش ss-BayesB در مقایسه با BayesB مشهودتر و معنی‌دار بود. نتایج نشان از برتری روش‌های بیزی بر روش GBLUP در صفات تحت کنترل تعداد پایین QTL در توزیع‌های مختلف QTL داشت، این در حالی بود که برتری مشهودی بین روش‌های تک‌مرحله‌ای و چندمرحله‌ای GBLUP در مقایسه با روش‌های متناظر آنها برای روش‌های بیزی در صفات تحت با تعداد بالای QTL با توزیع گامای آثار مشاهده نشد.

#### مدل ss-BayesB

در روش تک مرحله‌ای BayesB ( Fernando *et al.*, 2014) آثار نشانگرها از طریق رابطه (۸) و معادلات مختلط آن به صورت رابطه ۹ است:

$$\begin{bmatrix} \hat{\beta} \\ \hat{\alpha} \\ \hat{\epsilon} \end{bmatrix} = \begin{bmatrix} XX & XW & XZ_1 \\ WX & WW + D^{-1}\sigma_e^2 & WZ_1 \\ Z_1'X & Z_1'W & Z_1'Z_1 + A^{-1}\sigma_a^2 \end{bmatrix}^{-1} \begin{bmatrix} X'y \\ W'y \\ Z_1'y_1 \end{bmatrix} \quad (9)$$

در اینجا **D** ماتریس قطری است.

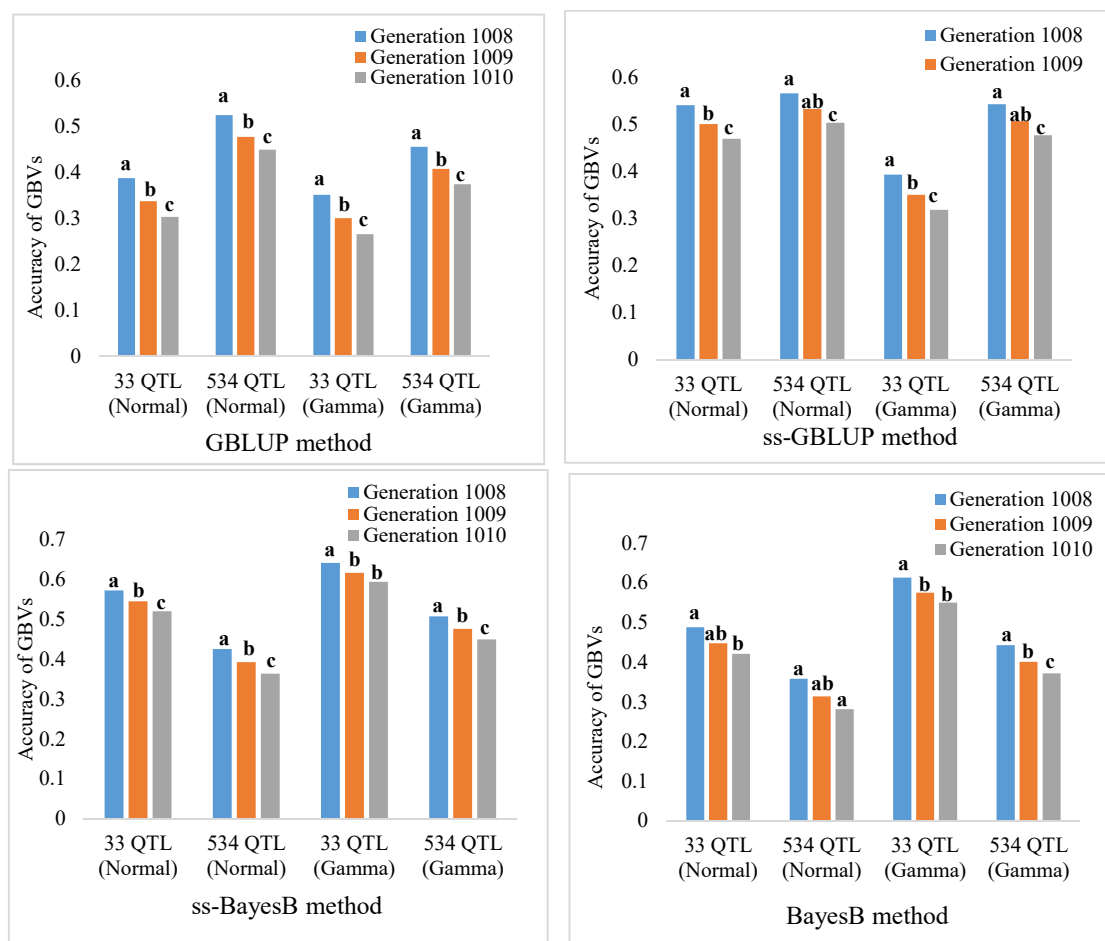
برای روش‌های تک مرحله‌ای و چند مرحله‌ای افراد نسل‌های ۱۰۰۴ تا ۱۰۰۷ به عنوان جمعیت مرجع (که دارای هر دو رکورد فنوتیپی و ژنوتیپی) در نظر گرفته شدند. علاوه بر این، با توجه به اینکه در مدل‌های تک‌مرحله‌ای به‌طور همزمان اطلاعات ژنومی، فنوتیپی و شجره‌ای جهت برآورد ارزش‌های اصلاحی حیوانات ژنوتیپ شده و نشده مورد استفاده قرار می‌گیرد، افراد نسل‌های ۱۰۰۱ تا ۱۰۰۳ به عنوان افراد دارای اطلاعات شجره‌ای و فنوتیپی نیز به اطلاعات افراد نسل ۱۰۰۴ الی ۱۰۰۷ اضافه گردید. در ادامه افراد نسل‌های ۱۰۰۸، ۱۰۰۹ و ۱۰۱۰ نیز به عنوان جمعیت تایید در روش‌های تک مرحله‌ای و چند مرحله‌ای در نظر گرفته شدند.

#### نتایج

##### اثر ارتباط بین جمعیت مرجع و تایید (طی سه نسل) بر صحت ژنومی

در این تحقیق صحت ارزش اصلاحی ژنومی در طی سه نسل مختلف ۱۰۰۸، ۱۰۰۹ و ۱۰۱۰ به عنوان جمعیت تایید با استفاده از حیوانات نسل‌های ۱۰۰۴ تا ۱۰۰۷ به عنوان جمعیت مرجع و همچنین افراد نسل‌های ۱۰۰۱ تا ۱۰۰۳ به عنوان افراد شجره‌ای دارای اطلاعات فنوتیپی و بدون اطلاعات ژنوتیپی (فقط برای مدل‌های تک مرحله‌ای) با استفاده از روش‌های تک مرحله‌ای و چند مرحله‌ای GBLUP و BayesB محاسبه شد.

در شکل ۱ نمودار صحت ژنومی در دو توزیع مختلف نرمال و گاما برای دو سطح مختلف QTL با روش GBLUP، ss-GBLUP، BayesB و ss-BayesB نشان داده شده است. به طور کلی صحت ارزیابی پیش‌بینی ژنومی با افزایش فاصله نسل در جمعیت مرجع و جمعیت تایید،



شکل ۱. اثر تعداد و توزیع متفاوت QTL بر صحت پیش‌بینی ارزش اصلاحی ژنومی در روش‌های GBLUP، ss-GBLUP، BayesB و ss-BayesB.

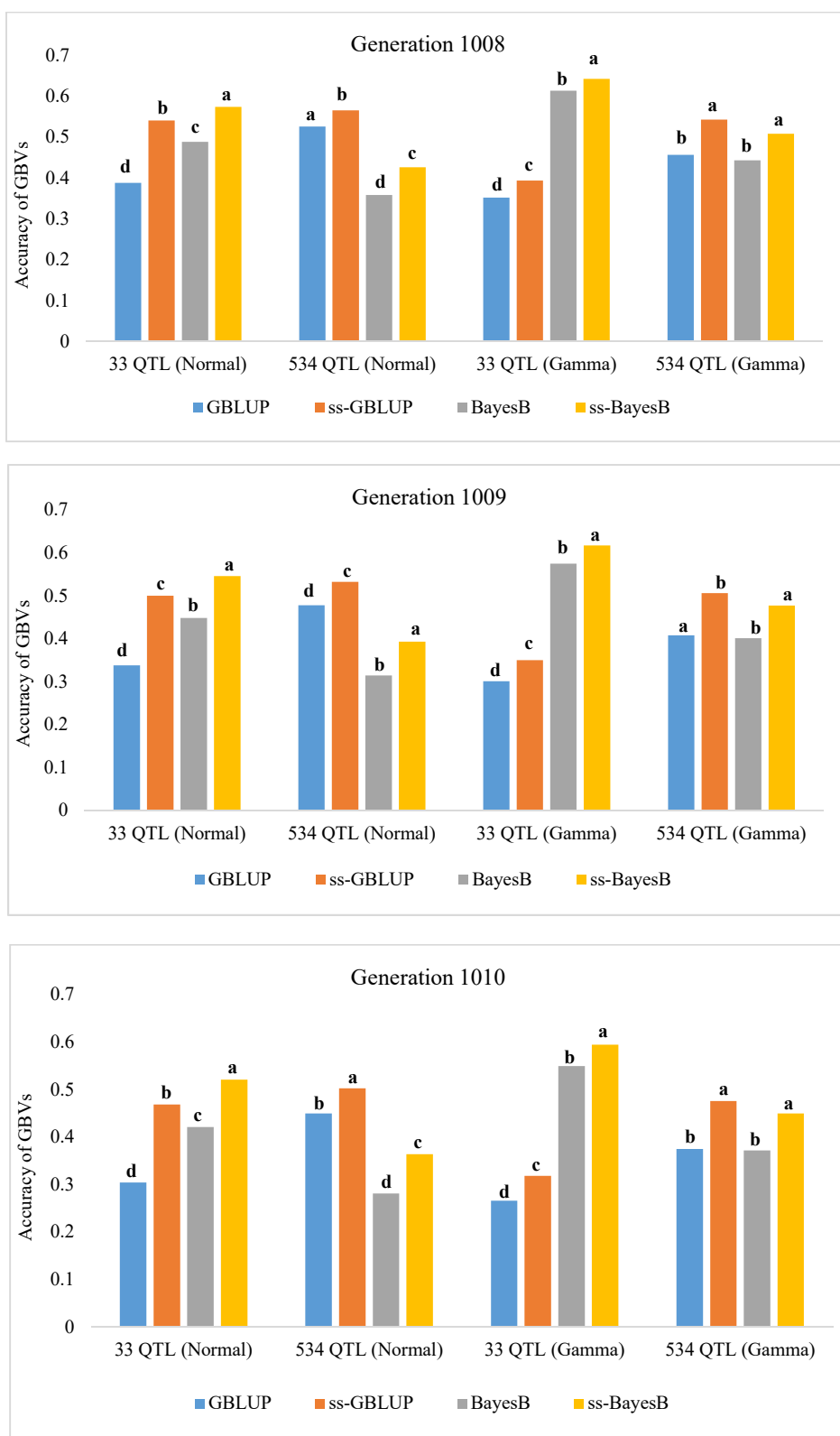
Figure 1. Effect of different number and distribution of QTL on accuracy of GBV prediction by GBLUP, ssGBLUP, BayesB and ss-BayesB methods.

#### اثر توزیع آثار ژنی بر صحت ژنومی

شکل ۳ نمودارهای اثر دو توزیع متفاوت ژنی (نرمال و گاما) بر صحت ژنومی مدل‌های GBLUP، BayesB، ss-GBLUP و ss-BayesB را به تفکیک نسل نشان می‌دهند. توزیع متفاوت آثار ژنی در روش GBLUP برای هر دو سطح QTL و برای ss-GBLUP برای صفات تحت تاثیر تعداد زیاد QTL اثر معنی‌داری روی صحت ژنومی نداشت. با این حال زمانی که توزیع آثار ژنی نرمال در نظر گرفته شد صحت ارزیابی ژنومی روش‌های GBLUP بیشتر بود. این در حالی بود که روش‌های BayesB و ss-BayesB حساسیت بالایی به نوع توزیع آثار ژنی از خود نشان دادند و عملکرد بالایی در توزیع آثار ژنی گاما از خود نشان دادند.

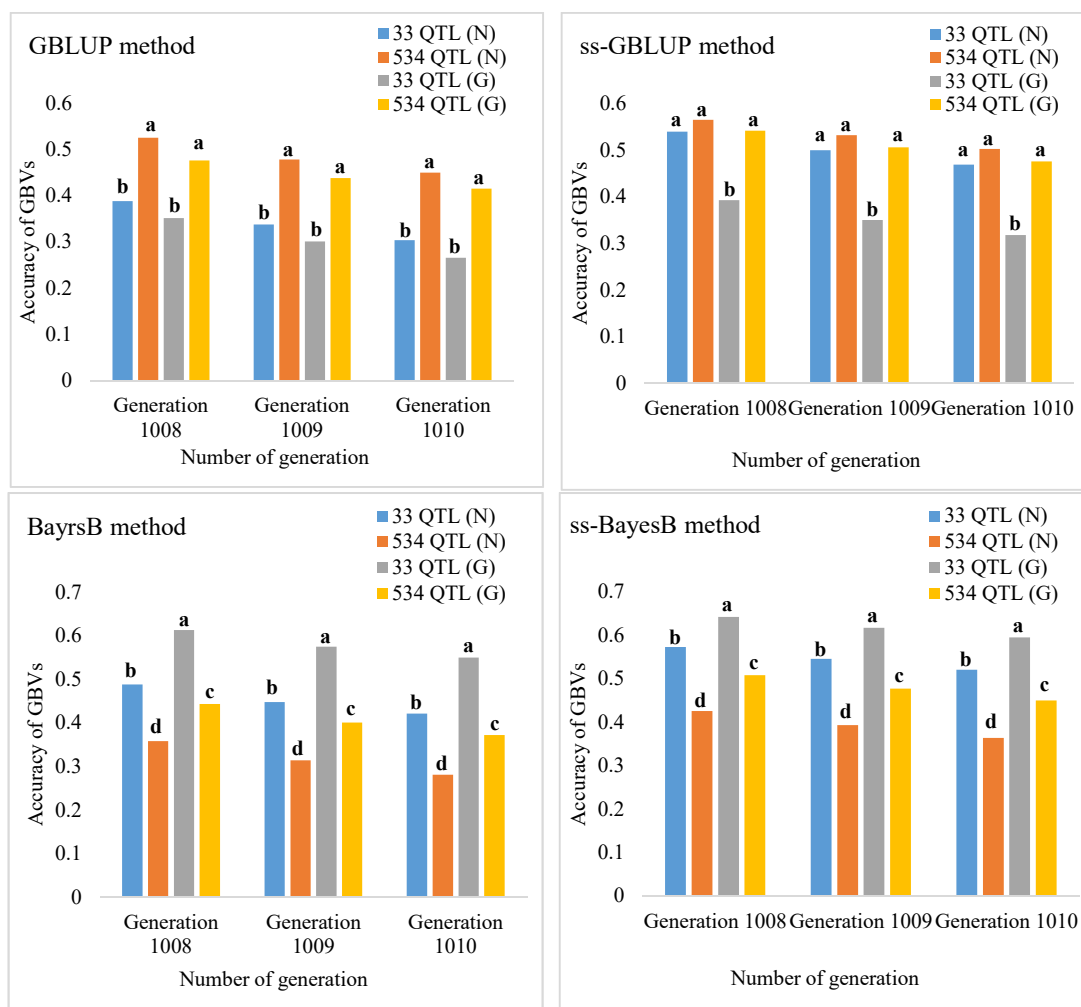
#### اثر تعداد متفاوت QTL بر صحت ژنومی

شکل ۳ نمودارهای اثر دو سطح متفاوت QTL بر صحت ژنومی هر یک از مدل‌های GBLUP، BayesB، ss-GBLUP و ss-BayesB به تفکیک نسل را نشان می‌دهند. به طور کلی تعداد QTL اثر معنی‌داری بر صحت ژنومی هر یک از مدل‌های BayesB نداشت. به طوری که روش‌های تک مرحله‌ای و چند مرحله‌ای BayesB حساسیت بسیار بالایی در مقایسه با روش GBLUP و به‌ویژه با روش ss-GBLUP از خود نشان دادند و با افزایش تعداد QTL صحت ژنومی برای روش‌های بیزی کاهش معنی‌داری نشان داد. این در حالی بود که برای توزیع نرمال آثار QTL روش ss-GBLUP مدلی با ثبات و انعطاف کمتری به نوسانات تعداد QTL از خود نشان داد.



شکل ۲. اثر تعداد و توزیع متفاوت QTL بر صحت پیش‌بینی ارزش اصلاحی ژنومی روش‌های مختلف برای نسل‌های ۱۰۰۸، ۱۰۰۹ و ۱۰۱۰.

Figure 2. Effect of different number and distribution of QTL on accuracy of GEBVs for different methods in generations 1008, 1009 and 1010.



شکل ۳. اثر تعداد و توزیع متفاوت QTL بر صحت پیش‌بینی ارزش اصلاحی ژنومی روش‌های مختلف GBLUP, ssGBLUP, BayesB و ss-BayesB در نسل‌های مختلف.

Figure 3. Effect of different number and distribution of QTL on accuracy of GBVs by GBLUP, ssGBLUP, BayesBand ss-BayesB method.ss-BayesB in different generations.

(Zhou *et al.*, 2018). با این حال میزان اثر ارتباط بین جمعیت مرجع و تایید در روش‌های مختلف انعکاس‌های مختلفی را در پی داشت که به تفسیر در ادامه به آنها پرداخته خواهد شد.

اثر ارتباط بین جمعیت مرجع و تایید بر لزوم به کارگیری مجدد انتخاب ژنومی

صحت پیش‌بینی ژنومی با افزایش فاصله نسلی به‌ویژه در مدل‌های چند مرحله‌ای کاهش یافت. این میزان کاهش صحت برای روش GBLUP نسبت به روش BayesB بیشتر بود. همچنین با افزایش فاصله نسلی صحت پیش‌بینی ژنومی روش ss-GBLUP در مقایسه

## بحث

در این مطالعه به بررسی افزایش فاصله نسل بین جمعیت مرجع و تایید، مقایسه‌ی روش‌های مختلف آماری شامل روش‌های تک مرحله‌ای و چند مرحله‌ای BayesB و GBLUP و تاثیر آن بر صحت پیش‌بینی ژنومی پرداخته شد. در این راستا و طی یک مطالعه شبیه‌سازی به بررسی صحت ارزیابی ژنومی در طی سه نسل متوالی (۱۰۰۸، ۱۰۰۹ و ۱۰۱۰) پرداخته شد. بطور کلی صحت پیش‌بینی ژنومی حاصل از هر کدام از مدل‌های مورد مطالعه با افزایش فاصله نسلی بین جمعیت مرجع و تایید کاهش یافت. این نتایج مطابق با گزارش‌های سایر محققین بود (Kang *et al.*, 2017; )



بعد از دو نسل با کاهش شدید صحت روبرو شده بودند (در نسل ۱۰۰۹) که این افت صحت برای مدل‌های تک مرحله‌ای (به‌ویژه ss-bayes) بعد از سه نسل اتفاق افتاد. در نتیجه روش‌های تک مرحله‌ای با به کارگیری افراد شجره‌ای حاوی اطلاعات فنوتیپی، لزوم به کارگیری مجدد انتخاب ژنومی طی دو نسل را منتفی می‌سازد.

#### اثر مدل آماری بر صحت پیش‌بینی ژنومی

در این مطالعه از دو روش تک مرحله‌ای و چند مرحله‌ای GBLUP و BayesB برای برآورد صحت پیش‌بینی ژنومی در طی نسل‌های مختلف استفاده شد. نتایج نشان داد که مدل‌های Bayes در صفاتی که تحت تاثیر تعداد کمی QTL باشند، عملکرد بهتری نشان می‌دهند. برعکس صحت ارزیابی ژنومی در صفاتی که تحت تاثیر تعداد بالای QTL باشند در مدل‌های GBLUP عملکرد بهتری را نشان داد. این نتایج با گزارش‌های سایر محققان در مورد آنالیز داده‌های ژنومی واقعی و شبیه‌سازی شده مطابقت دارد (Zhou et al., 2018; Hayes et al., 2009). مدل‌های Bayes و به‌ویژه مدل ss-Bayes و در مقایسه با روش‌های GBLUP و ss-GBLUP، در تعداد پایین QTL، صحت پیش‌بینی ژنومی بالاتری از خود نشان دادند. این میزان برتری در ss-BayesB به توزیع غیر نرمال اثر نشانگرها یا همان توزیع t نشانگرها به عنوان پیش فرض های این مدل برمی‌گردد، که در مطالعه حاضر این مفروضات با ساختار نشانگری با تعداد پایین QTL منطبق می‌باشد. در این راستا مطالعات شبیه‌سازی نشان از برتری روش‌های تک مرحله‌ای و چند مرحله بیز هنگامی که صفات توسط تعداد کم QTL کنترل می‌شود، داشتند (Zhou et al., 2018; Zhan et al., 2016). هنگامی که صفات تحت تاثیر تعداد زیاد QTL کنترل می‌شدند، روش‌های Bayes نسبت به روش‌های GBLUP صحت پایین تری دارند. دلیل این امر را می‌توان اینگونه عنوان کرد که در روش‌های Bayes مورد مطالعه مقدار  $0/99\pi$  فرض شد در نتیجه فقط به یک درصد نشانگرها اثر غیر صفر اختصاص داده شد ( $100 = 0/01 \times 10000$ ) که این

با ss-Bayes افت بیشتری داشت. با توجه به اینکه در افراد نسل ۱۰۰۸ در مقایسه با نسل‌های ۱۰۰۹ و ۱۰۱۰ دارای ارتباط نزدیک ژنتیکی با جمعیت مرجع بودند. و این ارتباط ژنتیکی منجر به افزایش صحت پیش‌بینی ژنومی در نسل ۱۰۰۸ در مقایسه با نسل‌های ۱۰۰۹ و ۱۰۱۰ شد. در نتیجه دلیل کاهش صحت ژنومی با افزایش فاصله نسل در روش‌های بر پایه GBLUP نسبت به روش‌های بیز را می‌توان به تاثیرپذیری بیشتر روش‌های GBLUP به ارتباط ژنتیکی دو جمعیت (تایید و مرجع) عنوان کرد (Habier et al., 2007). مطالعات در زمینه اثر ارتباط جمعیت مرجع و تایید در گاوهای نر هلشتاین آلمان نشان داد که هنگامی که ارتباط بین جمعیت مرجع و تایید کاهش پیدا کند صحت پیش‌بینی ژنومی برای نمره سلول‌های سوماتیک شیر کاهش می‌یابد (Habier et al., 2010). در یک مطالعه شبیه‌سازی (Zhou et al., 2018) نشان داده شد که روش ss-GBLUP در مقایسه با ss-Bayes حساسیت بیشتری به تغییرات فاصله نسل دارد.

مطالعات در مورد آنالیز پیش‌بینی ژنومی جمعیت گوسفند نشان داد که ساختار جمعیت و ارتباط فامیلی عنصری حیاتی در افزایش صحت پیش‌بینی ژنومی می‌باشد (Daetwyler et al., 2012). که تاییدی بر نتایج این تحقیق می‌باشد. نتایج تحقیق حاضر نشان می‌دهد که در روش‌های تک مرحله‌ای، با توجه به اینکه از اطلاعات فنوتیپی و شجره‌ای نسل‌های قبل (نسل‌های ۱۰۰۸ و ۱۰۰۹) در برآورد ارزش‌های اصلاحی نسل بعد (۱۰۱۰) استفاده می‌کنند، منجر به کاهش ناچیز صحت ارزیابی ژنومی در مقایسه با روش‌های چندمرحله‌ای شده که به نوبه خود، عدم لزوم به کارگیری مجدد انتخاب ژنومی برای حداقل یک نسل از منجر می‌شود. به طور مثال در برآورد ارزش اصلاحی نسل ۱۰۰۹، روش‌های تک مرحله‌ای بیز و GBLUP از اطلاعات شجره‌ای افراد ۱۰۰۸ در مدل استفاده می‌کنند، در نتیجه برآورد بهتری نسبت به مدل‌های چند مرحله‌ای که از این اطلاعات استفاده نمی‌کنند در پی خواهد داشت. در مطالعه حاضر روش‌های چند مرحله‌ای BayesB و به‌ویژه GBLUP

کردند (Naderi & Sadeghi, 2019). در روش‌های بیز و هنگامی که صفت تحت کنترل تعداد محدود QTL می‌باشد تنها تعداد محدودی نشانگر در ارتباط به QTLها واقعی اثر غیر صفر دارند ولی هنگامی که صفت تحت تاثیر تعداد بالای QTL باشد روش‌های GBLUP اجازه می‌دهند که تمامی نشانگرها اثر غیر صفر داشته باشند در نتیجه توافق بهتری بین آثار برآورد شده نشانگرها و آثار QTL واقعی برقرار می‌شود. مطالعات متعدد شبیه‌سازی حاکی از برتری روش‌های Bayes نسبت به روش GBLUP بوده که به نظر می‌رسد یکی از دلایل در نظر گرفتن تعداد پایین QTL در این مطالعات بود (Solberg *et al.*, 2008; Meuwissen *et al.*, 2001). در تایید نتایج این تحقیق، مطالعات به ترتیب حساسیت بالا و پایین روش‌های BayesB و روش GBLUP را به تغییرات QTL گزارش کردند.

#### اثر توزیع آثار ژنی بر صحت پیش‌بینی ژنومی

توزیع آثار ژنی ممکن است بر صحت ارزیابی ژنومی موثر باشد و باعث تفاوت‌هایی در بین روش‌های مختلف آماری شود. دلیل این امر را می‌توان به هر دو عامل تفاوت توزیع‌ها در پارامترهای شاخص تشکیل‌دهنده شکل توزیع و خصوصیات روش‌های آماری مورد نظر مربوط دانست. در نتیجه، شناسایی روش دارای بالاترین صحت ژنومی در هر توزیع می‌تواند به پیشرفت ژنتیکی مربوط باشد. در این مطالعه دو توزیع مختلف آثار ژنی (گاما و نرمال) بررسی شد. نتایج نشان داد که در صفات متأثر از تعداد کم QTL با توزیع گامای آثار ژنی روش‌های بیزی و در صفات متأثر از تعداد بالای QTL با توزیع نرمال آثار ژنی روش‌های GBLUP بهترین عملکرد را از خود نشان دادند. در جمعیت‌های با توزیع گامای آثار ژنی، یک نسبت کمی از QTLها دارای اثر بزرگ و اکثر QTLها اثر جزئی و نزدیک به صفر دارند که این فرضیه در مدل‌های GBLUP و BayesB به ترتیب آثار نامطلوب و مطلوب دارد. به طور کلی، در توزیع گاما اثر ژنی در مقایسه با توزیع نرمال، آثار QTLها تراکم بیشتری در نزدیکی صفر دارد در نتیجه، گرفتن اثر QTLها از این توزیع منجر به

میزان محدودیت در مدل BayesB منجر به ایجاد یک محدودیت در به دام انداختن تعداد ۵۳۴ QTL شد، در نتیجه مقدار صحت کاهش یافت. مطالعات در این زمینه نشان از عدم برتری مدل‌های بیز در مقایسه با روش GBLUP در صفات متأثر از تعداد بالای QTL داشتند (Zhou *et al.*, 2018; Kang *et al.*, 2017). به طور کلی استفاده از روش‌های بیزی در ارزیابی ژنومی منوط به شناسایی صفت و درک تعداد QTL کنترل‌کننده صفت است. در نتیجه می‌توان عنوان کرد که سودمند بودن هر کدام روش‌های بیز وابسته به تعداد QTL کنترل‌کننده صفت است.

#### اثر تعداد QTL بر صحت پیش‌بینی ژنومی

در این مطالعه، هر یک از روش‌های آماری مورد مطالعه با افزایش تعداد QTL عملکرد متفاوتی را نشان دادند. به طور کلی حساسیت مدل‌های تک مرحله‌ای و چند مرحله‌ای GBLUP در مقایسه با روش‌های بیز به تغییرات QTL کمتر بود. دلیل این امر را می‌توان به ماهیت GBLUP در استفاده از روابط خویشاوندی ژنومی میان جمعیت مرجع و تأیید به جای عدم تعادل پیوستگی بین QTL و نشانگرها عنوان کرد. بنابراین هنگامی که تعداد QTL افزایش یا کاهش پیدا می‌کند، روابط خویشاوندی ژنومی تغییر نکرده و در نتیجه صحت پیش‌بینی ژنومی نوسانات کمی از خود نشان می‌دهد (Zhou *et al.*, 2018). در مورد حساسیت بالای روش‌های بیز به تغییرات تعداد QTL، مطالعات شبیه‌سازی نشان داد که روش بیز نسبت به روش GBLUP حساسیت بالایی به تغییرات QTL نشان داده و با افزایش تعداد QTL صحت ژنومی کاهش یافت که دلیل این امر را می‌توان به توزیع محدود واریانس ژنتیکی بر تعداد زیادی QTL دانست که در نتیجه سهم هر QTL در ارزش ژنتیکی کل کاهش یافته، و قدرت مدل بیز (با توجه به پیش فرض‌های اولیه) در پیش‌بینی ارزش‌های اصلاحی کاهش می‌یابد. همچنین روش‌های تک مرحله‌ای و ss-BayesB از برخی خصوصیات نظیر انتخاب متغیر سود می‌برند و مفروضات آن با تعداد QTL کم سازگارتر است، در نتیجه در تعداد اندک QTL، روش‌های بیز بهتر عمل

چالش اساسی در انتخاب ژنومی به‌ویژه برای صفات آستانه‌ای با وراثت‌پذیری پایین می‌باشد، استفاده از روش‌های تک‌مرحله‌ای با به‌کارگیری افراد شجره‌ای حاوی اطلاعات فوتیپی، منجر به عدم لزوم به کارگیری مجدد انتخاب ژنومی در طی نسل‌های اولیه انتخاب می‌شود. همچنین با توجه به تاثیرپذیری کمتر روش تک مرحله‌ای BayesB به ارتباط ژنتیکی (دو جمعیت مرجع و تایید) در مقایسه با روش تک مرحله‌ای GBLUP، برآورد مجدد اثر نشانگرها (به طبع آن لزوم به کارگیری مجدد انتخاب) با ریسک کمتری همراه خواهد بود. با این حال استفاده از روش ss-BayesB و سودمند بودن آن در ارزیابی ژنومی منوط به شناسایی صفت و درک تعداد و توزیع آثار QTL‌های کنترل‌کننده صفت است.

ایجاد تعداد کمتری QTL با اثر بزرگ و تعداد بیشتری QTL با اثر کم شده و به دنبال آن باعث افزایش صحت پیش‌بینی ژنومی در روش BayesB (به‌ویژه در حضور تعداد کم QTL) شد.

#### نتیجه‌گیری کلی

نتایج این مطالعه نشان داد، با کاهش ارتباط بین جمعیت مرجع و تایید، کارایی روش‌های ارزیابی ژنومی به علت تغییر الگوی LD بین نشانگرها و QTL‌ها کاهش می‌یابد. در روش‌های تک مرحله‌ای در مقایسه با روش‌های چندمرحله‌ای به‌علت استفاده از شجره (دارای رکورد فنوتیپ) کاهش در صحت ارزیابی ژنومی در جمعیت تایید، کمتر بود. با توجه به اینکه با افزایش فاصله نسل از جمعیت مرجع یک

#### REFERENCES

1. Aguilar, I., Misztal, I., & Tsuruta, S. (2010). Genetic trends of milk yield under heat stress for US Holsteins. *Journal of Dairy Science*, 93, 1754-1758.
2. Bermann, M., Legarra, A., Hollifield, M.K., Masuda, Y., Lourenco, D., & Misztal, I. (2021). Validation of single-step GBLUP genomic predictions from threshold models using the linear regression method: An application in chicken mortality. *Journal of Animal Breeding and Genetics*, 138(1), 4-13.
3. Chen, C. Y., Misztal, I., Aguilar, I., Tsuruta, S., Meuwissen, T. H. E., Aggrey, S. E., ... & Muir, W. M. (2011). Genome-wide marker-assisted selection combining all pedigree phenotypic information with genotypic data in one step: An example using broiler chickens. *Journal of Animal Science*, 89, 23-28.
4. Christensen, O.F., & Lund, M.S. (2010). Genomic prediction when some animals are not genotyped. *Genetics Selection Evolution*, 42, 2.
5. Clark, S.A., Hickey, J.M., Daetwyler, H.D., & van der Werf, J.H. (2012). The importance of information on relatives for the prediction of genomic breeding values and the implications for the makeup of reference data sets in livestock breeding schemes. *Genetics Selection Evolution*, 44, 4.
6. Daetwyler, H. D., Kemper, K. E., Van Der Werf, J. H. J., & Hayes, B. J. (2012) Components of the accuracy of genomic prediction in a multi-breed sheep population. *Journal of Animal Science* 90: 3375-3384.
7. Fernando, R.L., Dekkers, J.C., & Garrick, D.J. (2014). A class of Bayesian methods to combine large numbers of genotyped and non-genotyped animals for whole-genome analyses. *Genetics Selection Evolution*, 46, 50.
8. González-Recio, O., & Forni, S. (2011). Genome-wide prediction of discrete traits using Bayesian regressions and machine learning. *Genetics Selection Evolution*, 43, 7.
9. Habier, D., Fernando, R., & Dekkers, J. (2007). The impact of genetic relationship information on genome-assisted breeding values. *Genetics*, 1 (77), 2389-2397.
10. 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.
11. Hayes, B.J., Visscher, P.M., & Goddard, M.E. (2009). Increased accuracy of artificial selection by using the realized relationship matrix. *Genetics Research*, 91, 47-60.
12. Kang, H., Zhou, L., Mrode, R., Zhang, Q., & Liu, J. (2017). Incorporating the single-step strategy into a random regression model to enhance genomic prediction of longitudinal traits. *Heredity*, 119, 459.

13. Kappes, S.M., Keele, J.W., Stone, R.T., McGraw, R.A., Sonstegard, T.S., Smith, T., Lopez-Corrales, N.L., & Beattie, C.W. (1997). A second-generation linkage map of the bovine genome. *Genome Research*, 7, 235-249.
14. Lee, J., Cheng, H., Garrick, D., Golden, B., Dekkers, J., Park, K., Lee, D., & Fernando, R. (2017). Comparison of alternative approaches to single-trait genomic prediction using genotyped and non-genotyped Hanwoo beef cattle. *Genetics Selection Evolution*, 49, 2.
15. Legarra, A., Aguilar, I., & Misztal, I. (2009). A relationship matrix including full pedigree and genomic information. *Journal of Dairy Science*, 92, 4656-4663.
16. Liu, Z., Goddard, M., Reinhardt, F., & Reents, R. (2014). A single-step genomic model with direct estimation of marker effects. *Journal of Dairy Science*, 97, 5833-5850.
17. Lourenco, D., Tsuruta, S., Fragomeni, B., Masuda, Y., Aguilar, I., Legarra, A., Bertrand, J., Amen, T., Wang, L., & Moser, D. (2015). Genetic evaluation using single-step genomic best linear unbiased predictor in American Angus. *Journal of Animal Science*, 93, 2653-2662.
18. Madsen, P., & Jensen, J. (2010). *A Users Guide to DMU*. A Package for Analysing Multivariate Mixed Models, Version 6.
19. Meuwissen, T., Hayes, B., & Goddard, M. (2001). Prediction of total genetic value using genome-wide dense marker maps. *Genetics*, 157, 1819-1829.
20. Naderi, Y., & Sadeghi, S. (2019). Assessment of the genomic prediction accuracy of discrete traits with imputation of missing genotypes. *Animal Science Papers and Reports*, 37, 149-168.
21. Sadeghi, S., Rafat, S.A., & Alijani, S. (2018). Evaluation of imputed genomic data in discrete traits using random forest and Bayesian threshold methods. *Acta Scientiarum Animal Sciences*, 40, e39007.
22. Sargolzaei, M., & Schenkel, F.S. (2009). QMSim: a large-scale genome simulator for livestock. *Bioinformatics*, 2 (5), 680-681
23. Solberg, T., Sonesson, A., & Woolliams, J. (2008). Genomic selection using different marker types and densities. *Journal of Animal Science*, 86, 2447-2454.
24. Su, G., & Madsen, P. (2013). User's Guide for GMATRIX version 2, a Program for Computing Genomic Relationship Matrix.
25. VanRaden, P.M. (2008). Efficient methods to compute genomic predictions. *Journal of Dairy Science*, 91, 4414-4423.
26. VanRaden, P.M., & Sullivan, P.G. (2010). International genomic evaluation methods for dairy cattle. *Genetics Selection Evolution*, 42, 7.
27. Yin, T., Pimentel, E., Borstel, U.Kv., & König, S. (2014). Strategy for the simulation and analysis of longitudinal phenotypic and genomic data in the context of a temperature× humidity-dependent covariate. *Journal of Dairy Science*, 97, 2444-2454.
28. Zhang, X., Lourenco, D., Aguilar, I., Legarra, A., & Misztal, I. (2016). Weighting strategies for single-step genomic BLUP: an iterative approach for accurate calculation of GEBV and GWAS. *Frontiers in Genetics*, 7, 151.
29. Zhou, L., Mrode, R., Zhang, S., Zhang, Q., Li, B., & Liu, J-F. (2018). Factors affecting GEBV accuracy with single-step Bayesian models. *Heredity*, 120, 100.