

استراتژی‌های تعیین ژنوتیپ انتخابی گاوهای ماده در برنامه‌های انتخاب ژنومی با شبیه‌سازی جمعیت گاو شیری

النا دهنوی^{۱*}، سعید انصاری مهباری^۲، فلاویو شنکل^۳، مهدی سرگلزایی^۴،^۵

تاریخ دریافت: ۹۵/۱۱/۶ تاریخ پذیرش: ۹۶/۹/۸

^۱ دانش‌آموخته دکتری گروه علوم دامی، دانشکده کشاورزی، دانشگاه صنعتی اصفهان

^۲ دانشیار گروه علوم دامی، دانشکده کشاورزی، دانشگاه صنعتی اصفهان

^۳ استاد مرکز بهبود ژنتیکی دام (CGIL)، دانشگاه گوئلف، انتاریو، کانادا

^۴ دانشیار مرکز بهبود ژنتیکی دام (CGIL)، دانشگاه گوئلف، انتاریو، کانادا

^۵ شرکت سیمکس، انتاریو، کانادا

*مسئول مکاتبه: Email: e.dehnavi@ag.iut.ac.ir

چکیده

زمینه مطالعاتی: افزودن اطلاعات ژنوتیپی و فنوتیپی گاوهای ماده به جمعیت مرجع، راهبردی سودمند در افزایش میزان صحت و کاهش اریب برآورد ارزش اصلاحی ژنومی ارزیابی شده است. هدف: مطالعه حاضر تأثیر بکارگیری داده‌های ژنومی گاوهای ماده در جمعیت مرجع توسط استراتژی‌های متفاوت انتخاب، تعداد افراد ژنوتیپ شده و وراثت‌پذیری بر صحت و اریبی ارزش اصلاحی ژنومی را در جمعیت تأیید مورد بررسی قرار داد. روش کار: پس از ایجاد جمعیت اولیه، جمعیت‌های ثانویه و اصلی به صورت روش چهار مسیری مشابه به‌نژادی گاو شیری شبیه‌سازی شد. سپس تعداد متفاوتی از گاو ماده (۲۵۰۰، ۵۰۰۰ و ۱۰۰۰۰ ماده) جهت افزودن به جمعیت مرجع حاوی ۵۰۰۰ گاو نر برتر تعیین ژنوتیپ گردید. دو سطح وراثت‌پذیری متوسط (۰/۳) و پایین (۰/۰۵) به‌طور مستقل در این مطالعه در نظر گرفته شد. نتایج: انتخاب گاوهای ماده جهت ژنوتیپ کردن با در نظر گرفتن صحت برآورد ارزش اصلاحی (AH و ATT)، ضمن صحت بالاتر، میانگین مربعات خطا و اریب کمتری را نسبت به انتخاب بدون در نظر گرفتن صحت (H و TT) نشان دادند. انتخاب گاوهای ماده با روش ATT، به‌طور کلی برای دو صفت منجر به بیشترین افزایش صحت برآورد ارزیابی ژنومی در محدوده ۰/۱۲۳ تا ۰/۲۱۵ شد. به دنبال آن روش‌های R و TT به ترتیب بیشترین افزایش صحت ارزیابی ژنومی را داشتند. بدین صورت افزایش صحت در محدوده ۰/۱۱۷ تا ۰/۲۰۴ در روش R و ۰/۱۱۳ تا ۰/۱۹۶ در روش TT مشاهده شد. تعیین ژنوتیپ گاوهای ماده با حداکثر ارزش اصلاحی (AH و H) منجر به کمترین صحت و بیشترین خطا و اریب گردید. با افزایش تعداد گاوهای ماده ژنوتیپ شده، MSE و اریب کمتر شد. متوسط ضریب رگرسیون TBV بر روی DGV در همه روش‌های مورد مطالعه از ۱/۱۲ با ۲۵۰۰ گاو ماده به ۱/۰۸ با ۱۰۰۰۰ ماده برای صفت با وراثت‌پذیری متوسط رسید. نتیجه‌گیری نهایی: در میان روش‌های مورد بررسی در این مطالعه افراد موجود در دو دامنه توزیع ارزش اصلاحی با حداکثر صحت کاندیداهای خوبی از افراد جامعه جهت تعیین ژنوتیپ و قرارگیری در جمعیت مرجع به‌نظر می‌رسند و اطلاعات بهتری را برای پیش‌بینی افراد نسل بعد با هر نوع سطح ژنتیکی فراهم می‌کنند.

واژگان کلیدی: ارزیابی ژنومی، تعیین ژنوتیپ انتخابی گاو ماده، جمعیت تأیید، جمعیت مرجع

مقدمه

استفاده از طرح‌های آمیزش که در آن روابط بین حیوانات در سطح ژنوم اندازه‌گیری و کنترل می‌شود جلوگیری از افزایش همخونی و نقص‌های ژنتیکی را بهتر میسر می‌کند (پریس و همکاران ۲۰۱۲ ب و ون‌رادن و همکاران ۲۰۱۱). گاوهای ماده همراه با داده‌های ژنومی در کارایی تکنولوژی‌های تولیدمثلی نقش مؤثری می‌توانند ایفا کنند (پریس و همکاران ۲۰۱۲ آ؛ توماسن و همکاران ۲۰۱۶). بدلیل اهمیت استفاده از گاوهای ماده در جمعیت مرجع در کنار گاوهای نر در دهه اخیر محققین سعی نموده‌اند در مطالعات مختلف نحوه شناسایی و استفاده از ژنوتیپ آن‌ها را مورد بررسی قرار دهند (توماسن و همکاران ۲۰۱۴؛ ویگن و همکاران ۲۰۱۲؛ هوگ و همکاران ۲۰۱۱). در واقع به‌منظور کاهش هزینه‌های ژنوتیپ‌کردن، استفاده از روش‌های تعیین ژنوتیپ انتخابی از همان ابتدا مطرح شده است (لندر و بوستن ۱۹۸۹). روش تعیین ژنوتیپ انتخابی اولین بار توسط لندر و بوستن (۱۹۸۹) معرفی شد. داروآسی و سلر در سال ۱۹۹۲ این روش را برای یک نشانگر منفرد چسبیده به QTL استفاده کردند. طبق تعریف در این روش‌ها یک گروه انتخابی از گاوها جهت توالی‌یابی DNA معرفی می‌شوند (انصاری و همکاران ۲۰۰۸). مطالعات نشان داده‌اند که با ژنوتیپ بخشی از افراد جامعه به‌جای کل آن همچنان می‌توان QTLها را با قدرت بالایی تشخیص داد (داروآسی و سلر ۱۹۹۲). کینگهورن (۱۹۷۹) شاخص احتمال ژنوتیپ جهت انتخاب بهینه افراد به‌کمک آنالیز تفرق از طریق شجره را جهت تعیین ژنوتیپ ارائه نمود. مدل‌های تعیین ژنوتیپ می‌تواند به صورت گروه به گروه (کینگهورن ۱۹۷۹) یا فرد به فرد (مکرسن ۲۰۰۴) باشد. به‌رحال، هدف همه روش‌های مذکور تشخیص بهتر ژن‌های مؤثر بر صفات کمی و کاهش هزینه‌های ژنوتیپ بوده است (انصاری و همکاران ۲۰۰۸). امروزه نیز استفاده از این روش‌ها و همچنین ایمپوتیشن^۷ (کالوس و همکاران ۲۰۱۴) راهی در جهت

امروزه ظهور داده‌های مولکولی در تکنولوژی‌های ژنومی و تعیین توالی DNA، ابزار جدید و مؤثری در افزایش بازدهی و بهره‌وری حیوانات اهلی محسوب می‌شود (گودارد و هیز ۲۰۰۸ و هیز و همکاران ۲۰۰۹). انتخاب ژنومی یکی از روش‌های رایج و قابل اجرا در سال‌های اخیر در برخی کشورها بوده است (گودارد و هیز ۲۰۰۸، موسین و همکاران ۲۰۰۱ و هیز و همکاران ۲۰۰۹). برنامه‌های انتخاب ژنومی، یک معادله پیش‌بینی ارزش توالی‌های DNA را در جمعیت مرجع با استفاده توأم از داده‌های ژنوتیپی و فنوتیپی فراهم می‌کند. این معادله پیش‌بینی سپس می‌تواند برای برآورد ارزش‌های اصلاحی حیوانات ژنوتیپ شده و فاقد داده فنوتیپی (جمعیت تایید) استفاده شود (موسین و همکاران ۲۰۰۱). بنابراین، تصمیم‌گیری نحوه انتخاب افراد به‌منظور بهبود گله‌های گاو شیری می‌تواند بر روی حیوانات جوان و حتی در بدو تولد، با دقتی بالاتر از دقت متوسط ارزش اصلاحی والدین انجام گیرد. این روش می‌تواند کاربرد قابل توجهی در طراحی برنامه‌های اصلاحی داشته باشد چون به جای انتظار برای رکوردهای دختران یک گاو نر که روندی ۵-۶ ساله دارد، نرهای جوان بدون فرزند می‌توانند به عنوان پدران گله استفاده شوند (شفر و ویگل ۲۰۱۲). بکارگیری اطلاعات ژنومی ماده‌ها زمینه جدید و مناسبی در عصر حاضر می‌باشد (پریس و داتولر ۲۰۱۲؛ پریس و همکاران ۲۰۱۲ آ و ب). در واقع تعیین ژنوتیپ گاوهای ماده می‌تواند قابلیت اطمینان انتخاب ژنومی را با افزایش تعداد حیوانات موجود در جمعیت مرجع افزایش دهد (هوغ و همکاران ۲۰۱۱). با اطلاعات ژنومی ماده‌ها می‌توان مادران ممتاز و تلیسه‌های جایگزین گله را شناسایی و از اصل و نسب آن‌ها اطمینان حاصل کرد (پریس و داتولر ۲۰۱۲ و پریس و همکاران ۲۰۱۲ آ).

^۶Genotype probability index (GPI)^۷Segregation analysis^۸Imputation^۱Reference population or Training set^۲Validation population or testing set^۳Parentage^۴Selective genotyping

جهش-رانش برسد و سپس طی ۱۰ نسل دیگر تعداد افراد افزایش یافت و به ۴۰۱۰۰ فرد (۱۰۰ نر و ۴۰،۰۰۰ ماده در نسل آخر) رسید. در این نسل (نسل ۴۰۱۰ ام) عدم تعادل لینکاژی در مناطق با فواصل کوتاه ($Mb\ 0/5$) مطابق با آنچه بروی داده واقعی گاو شیری در این فواصل دیده شده بود، ایجاد شد (شکل ۲).

جمعیت ثانویه (اخیر)

به منظور ایجاد عدم تعادل لینکاژی در مناطق با فواصل بلند ($Mb\ 5-0/5$)، ۲۵ نسل دیگر با آمیزش افراد براساس حداکثر ارزش اصلاحی و همچنین با در نظر گرفتن حداقل نرخ همخونی بین هر جفت آمیزش انجام شد. جمعیت ثانویه براساس جمعیت گاو شیری به صورت چهارمسیره و شامل دو طبقه تجاری و هسته در نظر گرفته شد. در هر نسل ۲۰ نر ممتاز با ۲۰۰ مادر گاو نر تحت عنوان گله هسته آمیزش داده شد. در گله تجاری نیز ۱۰۰ گاو نر فعال با ۲۰،۰۰۰ ماده تجاری آمیزش داده شد. نرخ تولیدمثلی گاوهای ماده در کل این جمعیت برابر با یک فرزند با احتمال برابر برای هر جدسیت تعیین شد. در هر نسل ۵۰۰ نر جوان براساس متوسط ارزش اصلاحی والدین انتخاب و وارد برنامه آزمون نتاج شد. سپس ۱۰ و ۵۰ گاو نر از ۵۰۰ نر پروف شده به ترتیب برای جایگزینی با نرهای طبقه ممتاز و تجاری انتخاب شد. نیمه دیگر نرها در هر دو طبقه فوق از نسل قبلی شان بودند. نرخ جایگزینی برای ماده‌ها در هر دو طبقه ۰/۳۳ در نظر گرفته شد (دهنوی و همکاران ۲۰۱۸).

جمعیت اصلی

از آنجایی که در جمعیت ثانویه و اصلی تعداد افراد و در نتیجه فراوانی‌های آلی متفاوت بود، جمعیت اصلی شامل گروه‌های تجاری و ممتاز برای ۶ نسل ایجاد شد تا اثر انتخاب بر واریانس ژنتیکی افزایشی مشاهده شود (بولمر، ۱۹۷۱).

در جمعیت اصلی نرخ تولیدمثلی و جایگزینی گاوهای ماده در دو طبقه متفاوت بود. نرخ تولید مثلی گاوهای

کاهش هزینه‌های تعیین ژنوتیپ جمعیت مرجع می‌باشد و عمده هزینه‌های مذکور از ژنوتیپ کردن همه افراد متولد شده جمعیت هدف نشأت می‌گیرد (بولیگن و همکاران ۲۰۱۲).

تاکنون مطالعاتی بر روی انتخاب حیوانات براساس ارزش اصلاحی یا انحراف تولیدی صورت گرفته است (احسانی و همکاران ۲۰۱۰؛ بولیگن و همکاران ۲۰۱۲؛ مونرو و همکاران ۲۰۱۲). مطالعات انجام گرفته در زمینه استراتژی‌های انتخاب، عمدتاً به ارزش صفات (ارزش اصلاحی یا انحراف تولیدی) توجه داشته‌اند (احسانی و همکاران ۲۰۱۰؛ بولیگن و همکاران ۲۰۱۲؛ مونرو و همکاران ۲۰۱۲). مطالعه حاضر روش‌های مختلف پیش‌انتخاب گاوهای ماده جهت تعیین ژنوتیپ و استفاده در جمعیت مرجع را همراه با در نظر گرفتن تأثیر صحت ارزش اصلاحی مورد بررسی و ارزیابی قرار داده است. همچنین تأثیر بکارگیری داده‌های ژنومی گاوهای ماده در جمعیت مرجع تحت استراتژی‌های مختلف برنامه‌های اصلاح نژادی گاو شیری بر صحت و آریبی ارزش اصلاحی ژنومی جمعیت تأیید مورد بررسی قرار گرفته است.

مواد و روش‌ها

ساختار جمعیت

در این مطالعه شبیه‌سازی سه جمعیت متوالی شامل جمعیت اولیه (تاریخی)، جمعیت ثانویه (اخیر) و جمعیت اصلی برای ایجاد جمعیتی مشابه با جمعیت گاو شیری با نرم‌افزار QMSim انجام شد (سرگلزایی و شنکل ۲۰۰۹). پارامترهای شبیه‌سازی جمعیت در این مطالعه براساس دهنوی و همکاران (۲۰۱۸) می‌باشد.

جمعیت اولیه (تاریخی)

ابتدا جمعیت اولیه با اندازه موثر ۱۶۰۰ حیوان غیرخویشاوند با جفت شدن تصادفی آبازسازی شد. این ساختار جمعیتی برای ۴۰۰۰ نسل تکرار شد تا به تعادل

[†]Random union of gametes

[‡]Historical population

[§]Recent population

ایجاد صحت از پیش تعیین شده برای هر یک از گاوهای نر و ماده شبیه‌سازی شد.

$$V = \begin{pmatrix} var & cov \\ cov & var \end{pmatrix}$$

$$cholesky(V) = CC'$$

$$X = \begin{pmatrix} TBV/\sqrt{var} & NRND \\ \vdots & \vdots \end{pmatrix} \times C'$$

$$= \begin{pmatrix} TBV & PseudoPhenotype \\ \vdots & \vdots \end{pmatrix}$$

که V ماتریس (کو) واریانس TBV ، $NRND$ اعداد تصادفی نرمال‌آمی با شد. C ماتریس پایین مثلثی حاصل از تجزیه چالسکی و C' ترانهاد ماتریس C است. ماتریس نهایی X شامل دو ستون TBV و شبه فنوتیپ‌آمی با شد. شبه فنوتیپ شبیه‌سازی شده به‌عنوان ارزش اصلاحی غیررگرسیون^۴ در نظر گرفته شد. مقادیر ارزش اصلاحی بدست آمده از معادلات فوق برای گروه مرجع استفاده شد. معادلات مختلط براساس قابلیت اطمینان مربوطه وزندهی شدند. برای صفات با وراثت‌پذیری پایین (۰/۰۵) قابلیت اطمینان ۷۵ و ۳۰ درصد به‌ترتیب برای گاوهای نر و ماده در نظر گرفته شد. برای صفات با وراثت‌پذیری متوسط (۰/۳) قابلیت اطمینان ۹۰ و ۶۰ درصد به‌ترتیب برای گاوهای نر و ماده در نظر گرفته شد (دهنوی و همکاران ۲۰۱۸).

دو سطح وراثت‌پذیری ۰/۳ و ۰/۰۵ نظیر صفات تولیدی و عملکردی به‌طور مستقل در نظر گرفته شد. برای ارزیابی ژنومی، حیوانات نسل پنجم و ششم جمعیت اصلی به‌ترتیب به‌عنوان جمعیت مرجع و جمعیت تأیید استفاده شد. همه افراد در جمعیت تأیید، فرزندان کاندیداهای انتخابی در جمعیت مرجع بودند. برای بررسی تأثیر روش انتخاب گاوهای ماده جهت تعیین ژنوتیپ و قرارگیری در جمعیت مرجع بر صحت، خطا و اریب پیش‌بینی‌های ژنومی، تعداد متفاوتی از گاوهای ماده (۲۵۰۰، ۵۰۰۰ و ۱۰۰۰۰) به جمعیت مرجع حاوی ۵۰۰۰ گاو نر برتر افزوده شد.

استراتژی‌های انتخاب برای تعیین ژنوتیپ

ماده در طبقه تجاری و ممتاز به‌ترتیب، ۱ و ۲۰ فرزند در نظر گرفته شد (توماسن و همکاران ۲۰۱۶). در طبقه تجاری ۰/۳۳ از دختران تولیدی حاصل از هر دو گروه تجاری و ممتاز به صورت تصادفی انتخاب و جایگزین ماده‌های پیشین این گروه شد. در طبقه ممتاز بعد از برداشت ۲۰ تخمک در طول یک سال یا کمتر از هر گاو برتر، آن حیوان حذف شد. بنابراین همه افراد نسل بعدی این طبقه از فرزندان این افراد و همچنین فرزندان گروه تجاری با ارزش اصلاحی بالاتر در نظر گرفته شد. نرخ جایگزینی نرها در هر دو طبقه ۵۰ درصد بود.

با توجه به اینکه دقت برآورد انتخاب ژنومی کمتر از آزمون نتاج می‌باشد، تعداد نرهای معادل با استفاده از روش کامرون، $PEV = \frac{1-r^2}{N}$ ، محاسبه شد (N تعداد افراد و r^2 قابلیت اطمینان جمعیت) (کامرون ۱۹۹۷). در این مطالعه قابلیت اطمینان ۹۰ و ۷۰ درصد به‌ترتیب برای برنامه‌های آزمون نتاج در جمعیت ثانویه و انتخاب ژنومی در جمعیت اصلی در نظر گرفته شد. بدین صورت ۶۰ و ۳۰۰ نر به‌ترتیب در طبقه ممتاز و تجاری استفاده شد. بنابراین، در هر نسل در طبقه ممتاز، ۶۰ نر برتر با ۶۰۰ گاو ماده ممتاز آمیزش یافت تا ۱۲۰۰۰ فرزند با احتمال مساوی برای هر جنس ایجاد کنند و در طبقه تجاری، با آمیزش ۳۰۰ گاو نر فعال با ۳۰۰۰۰ گاو ماده، ۳۰۰۰۰ فرزند با احتمال ۵۰ درصد برای هر جنس ایجاد گردید.

انتخاب براساس صحت از پیش تعیین شده ۰/۸۳۶ (معادل قابلیت اطمینان انتخاب ژنومی ۰/۷) صورت گرفت (ونرادن و همکاران ۲۰۰۹؛ ویگن و همکاران ۲۰۱۰). از آنجایی‌که اطلاعات گاو نر پروف شده براساس تعداد زیاد رکوردهای دخترانش می‌باشد، لذا قابلیت اطمینان بالاتر از اطلاعات گاو ماده براساس رکوردهای خود و فرزنداناش دارد. به منظور بازتاب این تفاوت در قابلیت اطمینان، برداری هم‌بسته با ارزش اصلاحی حقیقی برای

^۳Pseudo-phenotype

^۴De-regressed EBV

^۱Predefined accuracy

^۲Normal random number deviate (NRND)

بازگشتی در نظر گرفته شد (سرگلزایی و شنکل ۲۰۰۹). در داده‌های شبیه‌سازی همانند داده واقعی حاصل از توالی یابی کل ژنوم همه آلل‌ها با هر مقدار فراوانی آلی سهم یکسانی در برآورد دارند. در حالی که در تراشه‌ها، اسنیپ‌هایی که فراوانی آلی بالایی در جمعیت دارند انتخاب می‌شوند. در اسنیپ‌چیپ‌ها، اسنیپ‌های که فراوانی آلی کمیاب^۱ بالایی دارند و بهتر در جمعیت تفرق می‌یابند گنجانده می‌شوند و اسنیپ‌هایی که فراوانی MAF پایین‌تری در جمعیت دارند در تراشه قرار نمی‌گیرند. این عمل نوعی اریب به نام اریب نمونه‌گیری^{۱۰} در تراشه‌ها ایجاد می‌کند که حاصل انتخاب اسنیپ‌ها می‌باشد. برای معرفی اریب نمونه‌گیری به داده شبیه‌سازی، رابطه $\left(\frac{MAF}{v_1}\right)^{v_2}$ در برنامه QMSim اضافه شد. مقادیر v_1 و v_2 اعداد ثابت برای محاسبه تقریب می‌باشد. این روش از همه اسنیپ‌های تفرق یافته حاصل شده و اسنیپ‌های مورد نظر را با احتمال $\left(\frac{MAF}{0.2}\right)^{0.5}$ به صورت یکنواخت انتخاب می‌کند (سرگلزایی و شنکل ۲۰۰۹).

معیارهای ارزیابی

همبستگی پیرسون، میانگین مربعات خطا^{۱۱} این ارزش اصلاحی حقیقی و ارزش ژنومی مستقیم^{۱۲} ضریب رگرسیون TBV بر روی DGV به ترتیب به عنوان معیار صحت، خطا و اریب ارزیابی‌های ژنومی اندازه‌گیری شد. برای بررسی نرخ عدم تعادل لینکاژی، پارامتر r^2 در فواصل مختلف جمعیت مرجع برآورد شد و با مقادیر محاسبه شده بر روی داده واقعی گاو نر آمریکای شمالی مقایسه شد (هیل و رابرتسون ۱۹۶۸).

کدهای مورد نیاز برای محاسبات و آنالیز در محیط R نوشته شده شد. برای ارزیابی ژنومی روش snpBLUP

روش انتخاب گاوهای ماده جهت ژنوتیپ کردن به صورت انتخاب تصادفی (R)، بالاترین ارزش اصلاحی (H)، افراد کران بالای ارزش اصلاحی با حداکثر صحت EBV (AH)^{۱۳}، افراد دو کران بالا و پایین توزیع ارزش اصلاحی (TT)^{۱۴} و افراد دو کران بالا و پایین توزیع ارزش اصلاحی با در نظر گرفتن حداکثر صحت EBV (ATT)^{۱۵} در نظر گرفته شد. برای انجام روش دوم تا پنجم ابتدا ۲۵ و ۱۲/۵ درصد افراد جمعیت اصلی در کران بالا و دو کران بالا و پایین توزیع ارزش اصلاحی به ترتیب انتخاب شدند (شکل ۱). سپس برای روش‌های H و TT به ترتیب افراد با حداکثر ارزش اصلاحی و افراد با حداکثر و حداقل ارزش اصلاحی انتخاب شدند. برای روش‌های AH و ATT از میان ۲۵ و ۱۲/۵ درصد افراد انتخابی در کران بالا و دو کران بالا و پایین، افرادی که بیشترین صحت ارزش اصلاحی را نیز داشتند، انتخاب گردیدند (شکل ۱). صحت ارزش اصلاحی به صورت نزدیکی مقدار EBV به TBV در نظر گرفته شد. بنابراین، افرادی که توان دوم تفاوت EBV از TBV پایین‌تری داشتند به عنوان افرادی که صحت بالاتری دارند در گروه‌های مربوطه انتخاب شدند.

ژنوم

در این مطالعه ژنومی به طول ۲۴۹۶ ساننتی‌مورگان^{۱۶} حاوی ۲۹ کروموزوم اتوزوم با طول متفاوت و مشابه با اندازه کروموزوم‌های گاو (بو مانوا و همکاران ۲۰۱۰) شبیه‌سازی شد. تعداد ۵۰۰۰۰ نشانگر دوآلی و ۷۵۰ QTL چندآلی (۲، ۳ و ۴ آلل) با فراوانی آلی اولیه برابر با توزیع یکنواخت به‌طور تصادفی در ژنوم پخش شد. اثر QTL با توزیع گاما (با پارامتر شکل ۰/۴ و پارامتر مقیاس ۱/۶۶) نمونه‌گیری شد. نرخ جهش برای SNPها و QTL $10^{-4} \times 1$ در هر ساننتی‌مورگان تحت مدل جهش

^{۱۱}Recurrent mutation model

^{۱۲}Whole sequencing genome

^{۱۳}Minor allele frequency (MAF)

^{۱۴}Ascertainment bias

^{۱۵}Mean square error (MSE)

^{۱۶}Direct genomic value (DGV)

^۱Random

^۲Highest

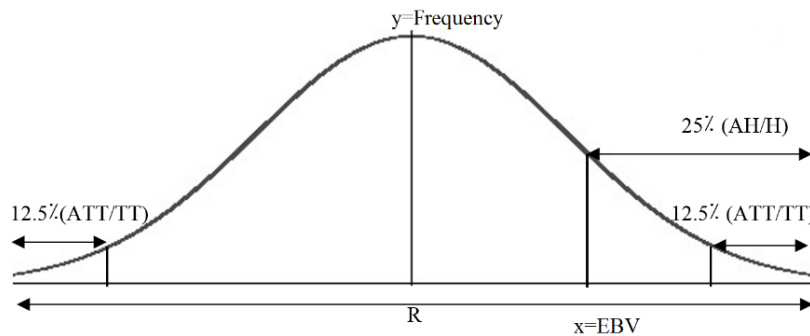
^۳Accurately highest

^۴Two tailed

^۵Accurately two tailed

^۶Centimorgan (cM)

موجود در نرم‌افزار SNP1101 بکار گرفته شد (سرگلزایی ۲۰۱۴). شبیه‌سازی برای ۲۰ بار تکرار گردید.

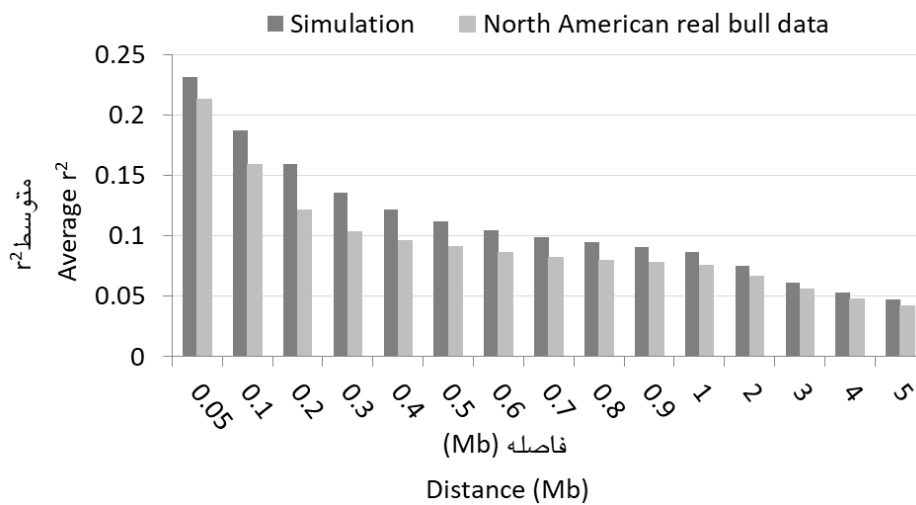


شکل ۱- توزیع ارزش اصلاحی و استراتژی‌های مختلف تعیین ژنوتیپ انتخابی گاوهای ماده
Figure1- EBV distribution and different cow selective genotyping strategies

و بلند تشکیل شده است (شکل ۲). در این مطالعه هر کروموزوم با طول متفاوت و مشابه با آن چه در داده واقعی گاو وجود دارد شبیه‌سازی شد (بومانوا و همکاران ۲۰۱۰).

نتایج و بحث

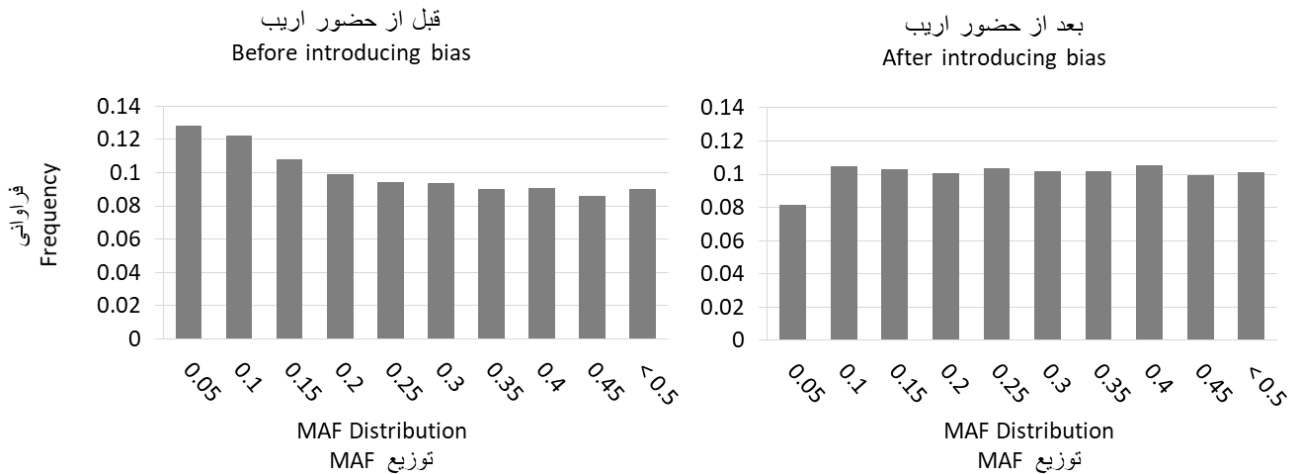
بررسی مقادیر r^2 در جمعیت شبیه‌سازی شده نشان داد که عدم تعادل لینکاژی به طور مناسب در فواصل کوتاه



شکل ۲- متوسط r^2 در فواصل مختلف داده شبیه‌سازی شده و داده گاو نر آمریکای شمالی
Figure2- Average r^2 for different marker distances (Mb) from simulation data and north American real data

در نتیجه جایگاه‌های ژنی در کروموزوم‌های طویل نسبت به کروموزوم‌های کوتاه کمتر همراه یکدیگر به ارث می‌رسند. متعاقباً کروموزوم‌های طویل مقدار LD کمتری در فواصل کم نسبت به کروموزوم‌های کوتاه دارند (بومانوا و همکاران ۲۰۱۰).

در نظر گرفتن طول متفاوت کروموزوم‌ها در کنار تعداد آن‌ها یکی از موضوعات مهم در شبیه‌سازی ژنوم می‌باشد. قطعات بلند کروموزوم در مقایسه با قطعات کوتاه، بیشتر در معرض وقوع نوترکیبی قرار می‌گیرند (بومانوا و همکاران ۲۰۱۰ و سیامن و همکاران ۲۰۰۴).

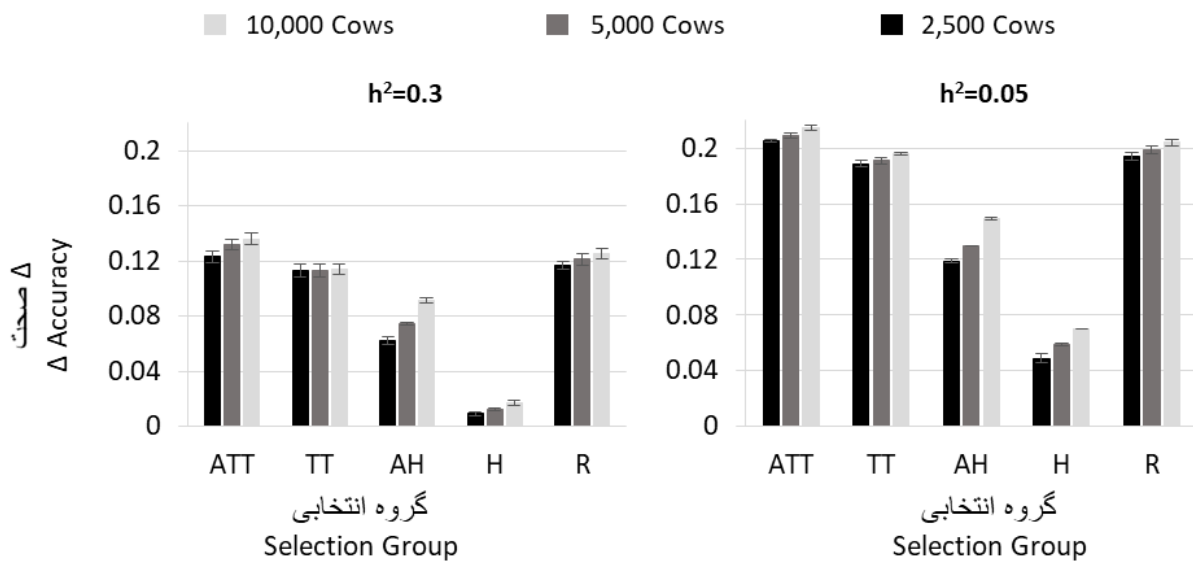


شکل ۳- توزیع فراوانی آلل کم قبل و بعد از معرفی اربیب نمونه‌گیری به داده شبیه‌سازی

Figure 3- MAF distribution before and after introducing ascertainment bias in the simulated dataset

رکوردبرداری از فرزندانش و قرار گیری در جمعیت مرجع را از دست می‌دهد. از آنجایی که گاوهای نر همواره یکسان انتخاب شده بودند، تفاوت صحت ارزیابی (Δ صحت) بعد از افزودن گاوهای ماده انتخابی با روش‌های مختلف نشان داده شد (شکل ۴). مشابه با مطالعات قبلی (توماسن و همکاران ۲۰۱۴؛ لورنس و همکاران ۲۰۱۴؛ ویگن و همکاران ۲۰۱۲؛ هوگ و همکاران ۲۰۱۱) افزودن اطلاعات گاوهای ماده به جمعیت مرجع حاوی گاوهای نر باعث تفاوت صحت ارزیابی جمعیت تأیید نسبت به حالتی شد که فقط گاوهای نر در جمعیت مرجع بود. به‌طور کلی در میان همه روش‌های انتخاب، Δ صحت جمعیت تأیید بین ۰/۰۰۹ تا ۰/۱۳۶ و ۰/۰۴۹ تا ۰/۲۱۵ به‌ترتیب برای صفات با وراثت‌پذیری ۰/۳ و ۰/۰۵ متغیر بود.

در مطالعه حاضر اربیب نمونه‌گیری با انتخاب جایگاه‌های در حال تفرق برطبق توزیع MAF در داده واقعی گاونر آمریکای شمالی به داده شبیه‌سازی وارد شد (شکل ۳). اربیب نمونه‌گیری در اثر طراحی تراشه‌های اسنیپ پدیدار می‌گردد. اسنیپ‌ها با MAF پایین در تراشه اسنیپ قرار نمی‌گیرند (کانر و همکاران ۲۰۰ و نیلسن ۲۰۰۴). اربیب نمونه‌گیری از جمله مواردی است که در نظر گرفتن آن برای شبیه‌سازی درست یک تراشه اسنیپ مهم می‌باشد. برطبق نتایج مطالعه حاضر صحت ارزیابی‌های ژنومی بستگی به روش انتخاب گاوهای ماده جمعیت مرجع جهت تعیین ژنوتیپ داشت (شکل ۴). در همه روش‌های انتخاب جمعیت مرجع ابتدا حاوی ۵۰۰۰ گاو نر برتر بود و سپس تعداد متفاوتی گاو ماده به آن اضافه می‌شد. گاو نر با سطح ژنتیکی پایین به علت حذف از گله شانس



شکل ۴- دلتا صحت جمعیت تأیید به کمک گروه‌های انتخابی مختلف جمعیت مرجع

Figure 4- Δ Accuracy of validation population using different selection groups of reference population

گروه‌های انتخابی مختلف شامل R: انتخاب افراد به‌طور تصادفی، H: انتخاب افراد دارای بیشترین ارزش اصلاحی، AH: انتخاب افراد کران بالای ارزش اصلاحی با حداکثر صحت EBV، TT: انتخاب افراد دو کران بالا و پایین توزیع ارزش اصلاحی و ATT: انتخاب افراد دو کران بالا و پایین توزیع ارزش اصلاحی با در نظر گرفتن حداکثر صحت EBV می‌باشد. Δ صحت، تفاوت صحت ارزیابی ژنومی جمعیت تأیید به کمک جمعیت مرجع حاوی گاو ماده و نر از صحت ارزیابی ژنومی جمعیت تأیید به کمک جمعیت مرجع حاوی فقط گاو نر می‌باشد.

Different selection groups involve in individuals that are selected R: randomly, H: based on the highest breeding values, AH: from the most accurate breeding values from top distribution tail, TT: from two tails of the distribution of breeding values, and ATT: from the most accurate breeding values from two distribution tails. Δ Accuracy is subtraction of accuracy of validation population when cows and bulls are included in the reference population from accuracy of validation population when only bulls are included in the reference population.

انتخابی R از ۰/۱۱۷ تا ۰/۱۲۵ و از ۰/۱۹۴ تا ۰/۲۰۴ برای صفات با نرخ توارث به ترتیب ۰/۳ و ۰/۰۵ برای جمعیت مرجع حاوی تعداد متفاوت ماده مشاهده گردید. انتخاب گاوهای ماده با بیشترین ارزش اصلاحی صرف نظر از در نظر گرفتن صحت آن (H و AH) منجر به کمترین افزایش صحت جمعیت تأیید شد. روش انتخاب AH منجر به تغییر دلتا صحت از ۰/۰۶۲ تا ۰/۰۹۱ در صفت با نرخ توارث ۰/۳ و از ۰/۱۱۹ تا ۰/۱۴۹ در صفت با نرخ توارث ۰/۰۵ شد. این نرخ در روش H به محدوده ۰/۰۰۹ تا ۰/۰۱۷ برای صفت با نرخ توارث ۰/۳ و محدوده ۰/۰۴۹ تا ۰/۰۷ برای صفت با نرخ توارث ۰/۰۵ رسید.

نتایج مطالعه حاضر مطابق با نتایج بولیگن و همکاران (۲۰۱۲) و مونترو و همکاران (۲۰۱۲) و مغایر با نتایج احسانی و همکاران (۲۰۱۰) بود. بولیگن و همکاران (۲۰۱۲) کمترین همبستگی TBV و GEBV را برای

در روش‌هایی که صحت برآورد ارزش اصلاحی نیز در نظر گرفته شد (AH و ATT) نسبت به روش‌هایی که آن را در نظر نگرفتند (H و TT)، تفاوت صحت ارزیابی ژنومی بیشتری مشاهده شد. در میان روش‌های مختلف، انتخاب گاوهای ماده در دو دامنه بالا و پایین توزیع ارزش اصلاحی که صحت بالایی نیز داشتند (ATT)، منجر به بیشترین Δ صحت ارزیابی ژنومی شد. دلتا صحت جمعیت تأیید در روش ATT از ۰/۱۲۳ تا ۰/۱۳۶ برای صفت با نرخ توارث ۰/۳ و از ۰/۲۰۵ تا ۰/۲۱۵ برای صفت با نرخ توارث ۰/۰۵ برای تعداد متفاوت گاو ماده در جمعیت مرجع متغیر بود. به دنبال آن انتخاب تصادفی (R) و سپس انتخاب گاوهای ماده با بیشترین و کمترین ارزش اصلاحی بدون در نظر گرفتن صحت ارزش اصلاحی (TT) به ترتیب بیشترین افزایش صحت ارزیابی ژنومی را داشتند. دلتا صحت جمعیت تأیید در گروه

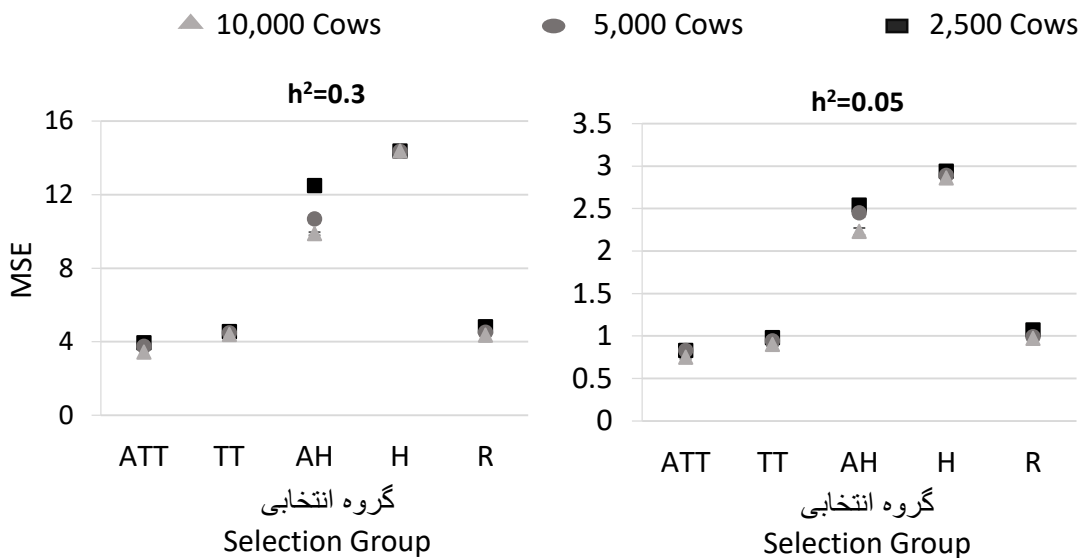
نشد (H و TT) از میانگین مربعات خطا کمتری برخوردار بودند (شکل ۵).

مقدار MSE جمعیت تأیید برای صفت با وراثت پذیری ۰/۳ در روش انتخاب ATT از ۳/۴۴ تا ۳/۹۴ متغیر بود که این نرخ در روش TT به محدوده ۴/۳۹ تا ۴/۵۵ رسید. معیار فوق برای صفت با وراثت‌پذیری ۰/۰۵ در روش ATT و TT به ترتیب در محدوده ۰/۷۵ تا ۰/۸۳ و ۰/۹ تا ۰/۹۸ برای گروه‌های مختلف جمعیت مرجع مشاهده گردید. MSE جمعیت تأیید در روش انتخاب R برای دو صفت با وراثت‌پذیری ۰/۳ و ۰/۰۵ به ترتیب در محدوده ۴/۳۵ تا ۴/۸ و ۰/۹۷ تا ۱/۰۷ برای گروه‌های مختلف جمعیت مرجع متغیر بود. مقدار این معیار برای روش AH، ۹/۸۷ تا ۱۲/۴۸ و ۲/۲۳ تا ۲/۵۴ به ترتیب برای دو صفت با وراثت‌پذیری ۰/۳ و ۰/۰۵ بود که به مقادیر ۱۴/۳۴ تا ۱۴/۳۷ و ۲/۸۶ تا ۲/۹۴ برای دو صفت فوق در روش H رسید. همچنین تعداد افراد انتخابی نیز در MSE تأثیرگذار بود. با افزایش تعداد گاو ماده (از ۲۵۰۰ به ۱۰۰۰۰ حیوان) MSE کاهش یافت.

در مطالعه مونترو و همکاران (۲۰۱۲) نیز با افزایش تعداد گاوهای ماده در جمعیت مرجع، MSE کاهش یافت. بولیگن و همکاران (۲۰۱۲) برای روش‌های انتخاب تصادفی، انتخاب براساس مقادیر دو کران بالا و پایین توزیع ارزش اصلاحی و براساس حداقل ارتباط خویشاوندی کمترین MSE نسبت به انتخاب براساس بیشترین ارزش اصلاحی مشاهده کردند.

روش‌های انتخاب براساس کمترین و یا بیشترین ارزش اصلاحی گزارش کردند. آن‌ها بیشترین همبستگی TBV و GEBV را برای سه استراتژی شامل انتخاب به‌طور تصادفی، انتخاب براساس مقادیر دو کران بالا و پایین توزیع ارزش اصلاحی و انتخاب براساس حداقل ارتباط خویشاوندی، نسبت به انتخاب براساس بیشترین ارزش اصلاحی مشاهده کردند (بولیگن و همکاران ۲۰۱۲). براساس مطالعه مونترو و همکاران (۲۰۱۲) صرف نظر از اینکه معیار انتخاب ارزش اصلاحی باشد یا انحراف تولید، انتخاب براساس دو آستانه حداقل و حداکثر ارزش اصلاحی یا انحراف تولید، بیشترین صحت ارزیابی ژنومی را به دنبال داشت. انتخاب براساس بیشترین ارزش اصلاحی یا انحراف تولید کمترین صحت ارزیابی انتخاب ژنومی را در پی داشت (مونترو و همکاران ۲۰۱۲). طبق مطالعه احسانی و همکاران (۲۰۱۰) انتخاب افراد به‌طور تصادفی و براساس ارزش اصلاحی متوسط به ترتیب منجر به بیشترین و کمترین قابلیت اطمینان در نتایج پیش‌بینی ارزش ژنومی حیوانات در جمعیت تأیید شد. تفاوتی بین انتخاب افراد براساس بیشترین و کمترین ارزش اصلاحی از نقطه نظر مقادیر قابلیت اطمینان نتایج ارزیابی ژنومی در جمعیت تأیید مشاهده نشد (احسانی و همکاران ۲۰۱۰).

باتوجه به اینکه مطالعات قبلی (بولیگن و همکاران ۲۰۱۲؛ مونترو و همکاران ۲۰۱۲) میانگین مربعات خطا را جهت بررسی خطای گروه‌های مختلف تعیین ژنوتیپ انتخابی برآورد و گزارش کردند، در مطالعه حاضر نیز این معیار محاسبه و مقایسه گردید (شکل ۵). همچنین با مقایسه روند معکوس بین مقادیر MSE با مقادیر دلتا صحت در گروه‌های انتخابی مختلف می‌توان از برآورد درست صحت در روش‌های متفاوت انتخاب اطمینان حاصل کرد. استراتژی‌های انتخاب که صحت ارزش اصلاحی نیز جهت انتخاب در نظر گرفته شد (AH و ATT)، نسبت به استراتژی‌هایی که صحت ارزش اصلاحی در نظر گرفته



شکل ۵- میانگین حداقل مربعات (MSE) جمعیت تأیید به کمک گروه‌های انتخابی مختلف جمعیت مرجع

Figure 5- Mean squared error (MSE) of validation population using different selection groups of reference population

گروه‌های انتخابی مختلف شامل R: انتخابی افراد به‌طور تصادفی، H: انتخاب افراد دارای بیشترین ارزش اصلاحی، AH: انتخاب افراد کران بالا و پایین ارزش اصلاحی با حداکثر صحت EBV، TT: انتخاب افراد دو کران بالا و پایین توزیع ارزش اصلاحی و ATT: انتخاب افراد دو کران بالا و پایین توزیع ارزش اصلاحی با در نظر گرفتن حداکثر صحت EBV می‌باشد.

Different selection groups involve in individuals that are selected R: randomly, H: based on the highest breeding values, AH: from the most accurate breeding values from top distribution tail, TT: from two tails of the distribution of breeding values, and ATT: from the most accurate breeding values from two distribution tails.

به‌کمک جمعیت مرجع حاوی فقط ۵۰۰۰ گاو نر برتر نیز در جدول ۱ ارائه شده است. مقادیر دلتا رگرسیون (Δb) تفاوت ضریب رگرسیون TBV بر روی DGV جمعیت تأیید به‌کمک جمعیت مرجع حاوی گاو ماده و نر نسبت به جمعیت مرجع حاوی فقط گاو نر می‌باشد. مقادیر دلتا رگرسیون (Δb) TBV بر روی DGV حاکی از تأثیر استفاده از داده‌های ژنوتیپی و فنوتیپی گاوهای ماده در کنار گاو نر در جمعیت مرجع می‌باشد. افزودن اطلاعات ژنومی گاوهای ماده در همه استراتژی‌های مورد بررسی باعث بهبود اریب در مقایسه با استفاده از اطلاعات فقط گاو نر در جمعیت مرجع شد (جدول ۱). میزان اریب بسته به روش انتخاب گاو ماده جهت تعیین ژنوتیپ متفاوت بود. استراتژی‌های انتخاب که صحت ارزش اصلاحی نیز جهت انتخاب در نظر گرفته شد (AH و ATT)، نسبت به استراتژی‌هایی که صحت ارزش اصلاحی در نظر گرفته نشد (H و TT) از اریب کمتری

با افزایش تعداد گاوهای ماده تعیین ژنوتیپ شده در جمعیت مرجع ضریب رگرسیون TBV بر روی DGV جمعیت تأیید نیز کاهش یافت (جدول ۱). ضریب رگرسیون TBV بر روی DGV در گروه‌های انتخابی ATT، TT، AH، H و R برای صفت با وراثت‌پذیری ۰/۳ به‌ترتیب از ۱/۰۷، ۱/۰۷، ۱/۱۶، ۱/۲۰ و ۱/۰۹ با ۲۵۰۰ گاو ماده به مقادیر ۱/۰۵، ۱/۰۵، ۱/۱۲، ۱/۱۳ و ۱/۰۶ با ۱۰۰۰۰ گاو ماده تعیین ژنوتیپ شده در جمعیت مرجع رسید. ضریب رگرسیون برای صفت با وراثت‌پذیری ۰/۵ به‌ترتیب از ۱/۱۳، ۱/۱۴، ۱/۱۷، ۱/۲۰ و ۱/۱۵ در گروه‌های انتخابی ATT، TT، AH، H و R با ۲۵۰۰ گاو ماده در جمعیت مرجع به ۱/۱۳، ۱/۱۳، ۱/۱۶ و ۱/۱۴ با ۱۰۰۰۰ گاو ماده در جمعیت مرجع رسید (جدول ۱). همچنین به منظور مقایسه بهتر روش‌های انتخاب گاو ماده مقادیر اریب به‌صورت انحراف از ضریب رگرسیون TBV بر روی DGV جمعیت تأیید

برخوردار بودند (جدول ۱). البته این تفاوت در روش انتخاب، براساس دو آستانه حداکثر و حداقل کمتر بود یا وجود نداشت (جدول ۱).

جدول ۱- ضریب رگرسیون ($\pm SE$) و دلتا رگرسیون (Δb) TBV بر روی DGV در جمعیت تأیید به‌کمک گروه‌های انتخابی مختلف جمعیت مرجع

Table 1- Regression coefficient ($\pm SE$) and delta regression coefficient (Δb) of validation population using different selection groups of reference population

h ²	گاو ماده Cow	گروه‌های انتخابی Selection groups									
		ATT		TT		AH		H		R	
		b($\pm SE$)	Δb	b($\pm SE$)	Δb	b($\pm SE$)	Δb	b($\pm SE$)	Δb	b($\pm SE$)	Δb^{***}
0.3	2,500	1.068 (0.002)	-0.154	1.067 (0.002)	-0.155	1.156 (0.003)	-0.066	1.199 (0.004)	-0.023	1.089 (0.002)	-0.133
	5,000	1.059 (0.002)	-0.163	1.056 (0.002)	-0.166	1.133 (0.003)	-0.089	1.172 (0.004)	-0.05	1.066 (0.002)	-0.156
	10,000	1.053 (0.002)	-0.169	1.053 (0.002)	-0.169	1.121 (0.003)	-0.101	1.128 (0.003)	-0.094	1.059 (0.002)	-0.163
0.05	2,500	1.128 (0.003)	-0.086	1.137 (0.004)	-0.077	1.167 (0.003)	-0.047	1.199 (0.003)	-0.015	1.147 (0.002)	-0.067
	5,000	1.123 (0.003)	-0.091	1.123 (0.003)	-0.091	1.158 (0.002)	-0.056	1.179 (0.003)	-0.035	1.136 (0.002)	-0.078
	10,000	1.128 (0.003)	-0.086	1.127 (0.002)	-0.087	1.159 (0.002)	-0.055	1.165 (0.002)	-0.049	1.136 (0.002)	-0.078

* جمعیت مرجع حاوی تعداد متفاوت گاو ماده ژنوتیپ شده (۲,۵۰۰، ۵,۰۰۰ و ۱۰,۰۰۰ حیوان) می‌باشد.

** گروه‌های انتخابی مختلف شامل R: انتخابی افراد به‌طور تصادفی، H: انتخاب افراد دارای بیشترین ارزش اصلاحی، AH: انتخاب افراد کران بالای ارزش اصلاحی با حداکثر صحت EBV، TT: انتخاب افراد دو کران بالا و پایین توزیع ارزش اصلاحی و ATT: انتخاب افراد دو کران بالا و پایین توزیع ارزش اصلاحی با در نظر گرفتن حداکثر صحت EBV می‌باشد.

*** Δb تفاوت ضریب رگرسیون TBV بر روی DGV جمعیت تأیید به‌کمک جمعیت مرجع حاوی گاو ماده و نر نسبت به جمعیت مرجع حاوی فقط گاو نر می‌باشد.

* The reference population includes different numbers of genotyped cows (2,500, 5,000 and 10,000 animals).

** Different selection groups involve in individuals that are selected R: randomly, H: based on the highest breeding values, AH: from the most accurate breeding values from top distribution tail, TT: from two tails of the distribution of breeding values, and ATT: from the most accurate breeding values from two distribution tails.

*** Δb is subtraction of regression coefficient of TBV on DGV for validation population when cows and bulls are included in the reference population from regression coefficient of TBV on DGV for validation population when only bulls are included in the reference population.

دو کران بالا و پایین توزیع ارزش اصلاحی منجر به بیشترین همبستگی پیش‌بینی‌های ژنومی با رکوردهای فنوتیپی گردید.

در گله‌های تجاری، افراد برای یک یا چند صفت برای چندین نسل انتخاب می‌شوند. انصاری و همکاران (۲۰۰۸) با شبیه‌سازی یک جمعیت گاو شیری برای ۲۵ سال نشان دادند که تعیین ژنوتیپ انتخابی در برنامه‌های کاربردی اصلاح گاو شیری و کاهش هزینه‌های تعیین ژنوتیپ با حداقل کاهش پاسخ به انتخاب در مقایسه با استفاده از تمام افراد در جمعیت مرجع همراه است.

مطابق با تحقیق حاضر مونتر و همکاران (۲۰۱۲) در بین روش‌های انتخاب، بیشترین اریب را برای انتخاب براساس بالاترین ارزش اصلاحی و یا انحراف تولید گزارش کردند. جنگو و همکاران (۲۰۱۷) روش‌های انتخاب مختلف شامل انتخاب همه گاوهای ماده، انتخاب ۳۰، ۴۰ و ۵۰ درصد از گاوهای ماده به‌طور تصادفی، گاوهای ماده دارای بیشترین ارزش اصلاحی، گاوهای ماده با بیشترین و کمترین ارزش اصلاحی، گاوهای ماده دو کران بالا و پایین توزیع ارزش اصلاحی مورد بررسی قرار دادند. بر طبق مطالعه آن‌ها در میان روش‌های انتخاب ۵۰ درصد گاوهای ماده، انتخاب افراد

حذف گاوهای با سطح ژنتیک پایین امکان دسترسی به اطلاعات آن‌ها و فرزندانشان را محدود می‌کند. همچنین در گله‌های تجاری اغلب گاوداران تمایل به تعیین ژنوتیپ گاوهای ممتاز را دارند. اما باید توجه نمود که بر طبق مطالعه حاضر و مطالعات قبلی (بولیگن و همکاران ۲۰۱۲ و مونتر و همکاران ۲۰۱۲) استفاده از هر دو دامنه افراد جامعه (حداقل و حداکثر ارزش اصلاحی) صحت بیشتر و خطا و اریب کمتری را به همراه دارد. نتایج مطالعه حاضر نشان داد که در نظر گرفتن صحت ارزش اصلاحی بدست آمده نیز مهم می‌باشد. در صورت عدم دسترسی یا امکان محاسبه ارزش اصلاحی، انتخاب تصادفی جایگزین مناسبی به نظر می‌رسد. در واقع افراد حاصل از این دو روش انتخاب، نمونه درستی از جامعه به نظر می‌رسند و اطلاعات بهتری را از جایگاه هم ژن‌های مطلوب و هم نامطلوب به دست می‌دهند. مطالعات نشان دادند که انتخاب ۵۰ درصد گاوهای ماده از دو کران بالا و پایین جامعه، حدود ۸۸ درصد اطلاعات همه افراد جامعه را می‌تواند پوشش دهد (جنگو و همکاران ۲۰۱۷).

استفاده از افراد با بالاترین ارزش اصلاحی کمترین صحت و بیشترین میانگین مربعات خطا و اریب در ارزیابی ژنومی جمعیت تأیید را به همراه داشت. حتی استفاده از معیار صحت ارزش‌های اصلاحی هم نتوانست کمکی در افزایش صحت ارزیابی ژنومی و یا کاهش خطا و اریب این روش در مقایسه با سایر روش‌های انتخاب کند. افراد با بالاترین ارزش اصلاحی در واقع سهم زیادتری در ژن‌های مفید جامعه دارند. استفاده از آن‌ها به تنهایی برای برآورد ارزش اصلاحی ژنومی همه افراد نسل بعد شاید گزینه مناسبی نباشد. زیرا افراد نسل بعد افرادی هستند که لزوماً همه آن‌ها نمی‌توانند حامل ژن‌های برتر باشند.

سیاسگزاری

از بخش فناوری اطلاعات (IT) مرکز بهبود دام (CGIL)، دانشگاه گوئلف برای اجازه دسترسی به ابررایانه مرکز جهت شبیه‌سازی و آنالیز داده قدردانی می‌شود. از

مطالعه حاضر دو سطح وراثت‌پذیری ۰/۳ و ۰/۵ را به صورت مستقل مورد بررسی قرار داد. به‌طور کلی همه گروه‌های انتخابی در صفت با وراثت‌پذیری ۰/۵ منجر به دلتا صحت بالاتر و MSE کمتر نسبت به صفت با وراثت‌پذیری ۰/۳ شدند. صفات با وراثت‌پذیری پایین سهم واریانس افزایشی کمتری نسبت به صفات با وراثت‌پذیری بالا دارند. استفاده از ماتریس روابط ژنومی کمک بیشتری در بیان سطح ژنتیکی افراد می‌کند. بولیگن و همکاران (۲۰۱۲) در کنار شبیه‌سازی صفات به طور مستقل، دو صفت با همبستگی ژنتیکی ۰/۵ را در نظر گرفتند و پنج روش تعیین ژنوتیپ انتخابی مختلف را روی یک صفت اجرا کردند و نتایج را بر روی صفت همبسته در جمعیت تأیید مورد بررسی قرار دادند. انتخاب افراد در دو دامنه توزیع ارزش اصلاحی و یا انحراف تولیدی صفت اول منجر به صحت بالاتر و MSE پایین‌تر صفت همبسته گردید. نتایج انتخاب برای صفت همبسته مشابه با نتایج انتخاب بروی یک صفت منفرد بود (بولیگن و همکاران ۲۰۱۲).

با توجه به اینکه با نرم‌افزار QMSim که در مطالعه حاضر مورد استفاده قرار گرفت امکان شبیه‌سازی بیش از یک صفت وجود نداشت، در جمعیت مورد مطالعه دو صفت به صورت مستقل و بدون کواریانس بین آن‌ها شبیه‌سازی شدند. بدیهی است که با افزایش تعداد صفات، بازدهی اقتصادی روش‌های تعیین ژنوتیپ انتخابی بخصوص برای صفات با همبستگی کم، پایین می‌آید. در هر صورت، تعیین ژنوتیپ انتخابی در صفاتی ترکیبی^۱ مانند LPI، TPI یا NM می‌تواند مورد استفاده قرار گیرد. به‌هرحال، اجرای آنالیز دو یا چندصفتی در مطالعات بعدی و مقایسه با نتایج آنالیز تک‌صفتی توصیه می‌شود. جنگو و همکاران (۲۰۱۷) رکوردهای نر و ماده را به صورت دو صفت در مدل دوصفتی ارزیابی کردند. بر طبق مطالعه آنها استفاده از رکوردهای نر و ماده به صورت مدل دوصفتی در مقایسه با استفاده از اطلاعات نر و ماده به صورت مدل تک‌صفتی تفاوت چندانی در نتایج پیش‌بینی‌های ژنومی نداشته است.

^۱Informative

^۱Composite traits

شبکه گاو شیری کانادا (CDN)، گوئلف، جهت دسترسی به داده آمریکای شمالی برای محاسبه برخی پارامترهای مورد نیاز جهت شبیه‌سازی جمعیت تشکر می‌گردد.

منابع مورد استفاده

- Ansari-Mahyari S, Sørensen AC, Lund MS, Thomsen H and Berg P, 2008. Across-family marker-assisted selection using selective genotyping strategies in dairy cattle breeding schemes. *Journal of Dairy Science* 91:1628–1639.
- Bohmanova J, Sargolzaei M and Schenkel F, 2010. Characteristics of linkage disequilibrium in North American Holsteins. *BMC. Genomics* 11 (1):1-11.
- Boligon AA, Long N, Albuquerque LG, Weigel KA, Gianola D and Rosa GJ, 2012. Comparison of selective genotyping strategies for prediction of breeding values in a population undergoing selection. *Journal of Animal Science* 90: 4716-22.
- Bulmer MG, 1971. The Effect of Selection on Genetic Variability. *The American Naturalist* 105, 201-211.
- Cameron ND, 1997. Selection indices and prediction of genetic merit in animal breeding. CAB international. Wallingford, 208 Pp.
- Calus MP, Bouwman AC, Hickey JM, Veerkamp RF and Mulder HA, 2014. Evaluation of measures of correctness of genotype imputation in the context of genomic prediction: a review of livestock applications. *Animal* 8:1743-53.
- Darvasi A and Soller M, 1992. Selective genotyping for determination of linkage between a marker locus and a quantitative trait locus. *Theoretical and Applied Genetics*. 85:353–359.
- Dehnavi E, Mahyari SA, Schenkel FS and Sargolzaei M, 2018. The effect of using cow genomic information on accuracy and bias of genomic breeding values in a simulated Holstein dairy cattle population. *Journal of Dairy Science* 101(6): 5166-5176.
- Ehsani A, Janss L and Christensen OF, 2010. Effects of selective genotyping on genomic prediction. Pp 444, *Proceeding of 9th World Congress Genetic Applied Livestock Production*. Leipzig, Germany (Abstr).
- Goddard ME and Hayes BJ, 2008. Review Article: Genomic selection. *Journal of Animal Breeding and Genetic* 124: 323–330.
- Hayes BJ, Bowman PJ, Chamberlain AJ and Goddard ME, 2009. Genomic selection in dairy cattle: progress and challenges. *Journal of Dairy Science* 92(2): 433–443.
- Hill WG and Robertson A, 1968. Linkage disequilibrium in finite populations. *Theoretical and Applied Genetics*. 38:226–231.
- Jenko J, Wiggans GR, Cooper TA, Eaglen SAE, Luff WGdeL, Bichard M, Pong-Wong R and JA Woolliams. 2017. Cow Genotyping Strategies for Genomic Selection in a Small Dairy Cattle Population. *Journal of Dairy Science* 100 (1): 439–52.
- Jensen-Seaman MI, Furey TS, Payseur BA, Lu Y, Roskin KM, Chen CF, Thomas MA, Haussler D and Jacob HJ, 2004. Comparative recombination rates in the rat, mouse, and human genomes. *Genome Research*. 14: 528-38.
- Jiménez-Montero JA, Gonzalez-Recio O and Alenda R, 2012. Genotyping strategies for genomic selection in small dairy population. *Animal* 6:1216–1224.
- Kinghorn BP, 1997. An Index of Information Content for Genotype Probabilities Derived from Segregation Analysis. *Genetics* 145(2): 479-483.
- Kuhner MK, Beerli P, Yamato J and Felsenstein J, 2000. Usefulness of single nucleotide polymorphism data for estimating population parameters. *Genetics* 156: 439-447.
- Lander E and Botstein D, 1989. Mapping Mendelian factors underlying quantitative traits using RFLP linkage maps. *Genetics* 121:185–199.
- Lourenco DA, Misztal I, Tsuruta S, Aguilar I, Ezra E, Ron M, Shirak A and Weller JI, 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: 1742-52.

- Macrossan PE, 2004. Strategies to Minimize DNA Testing Costs for Research and Development Programs Involving Pedigreed Populations. PhD Thesis, University of New England, Australia.
- Mc Hugh N, Meuwissen TH, Cromie AR and Sonesson AK, 2011. Use of female information in dairy cattle genomic breeding programs. *Journal of Dairy Science* 94(8):4109-4118.
- Meuwissen T, Hayes B and Goddard ME, 2001. Prediction of total genetic value using genome-wide dense marker maps. *Genetics*. 157(4): 1819-1829.
- Nielsen R, 2004. Population genetic analysis of ascertained SNP data. *Human Genomics* 1: 1-7.
- Pryce JE and Daetwyler H, 2012. Designing dairy cattle breeding schemes under genomic selection: a review of international research. *Animal Production Science* 52(3): 107-114.
- Pryce JE, Hayes BJ and Goddard ME, 2012 a. Genotyping dairy females can improve the reliability of genomic selection for young bulls and heifers and provide farmers with new management tools. *Proceeding of 38th ICAR Conference, Cork, Ireland.*
- Pryce JE, Hayes BJ and Goddard ME, 2012 b. Novel strategies to minimize progeny inbreeding while maximizing genetic gain using genomic information. *Journal of Dairy Science* 95:377-388.
- Sargolzaei M, 2014. SNP1101 User's Guide. Version 1.0.
- Sargolzaei M. and FS Schenkel, 2009. QMSim: a large-scale genome simulator for livestock. *Bioinformatics*, 25: 680-681.
- Scheffers JM and Weigel KA, 2012. Genomic selection in dairy cattle: Integration of DNA testing into breeding programs. *Animal Frontiers*. 2(1):4-9.
- Thomasen JR, Willam A, Egger-Danner C and Sørensen AC, 2016. Reproductive technologies combine well with genomic selection in dairy breeding programs. *Journal of Dairy Science* 99(2):1331-1340.
- Thomasen, JR, Sørensen AC, Lund MS and Guldbbrandtsen B, 2014. Adding cows to the reference population makes a small dairy population competitive. *Journal of Dairy Science* 97(9):5822-5832.
- VanRaden PM, Olson KM, Null DJ and Hutchison JL, 2011. Harmful recessive effects on fertility detected by absence of homozygous haplotypes. *Journal of Dairy Science*, 94:6153-61.
- VanRaden PM, Van Tassell C, Wiggans G, Sonstegard T, Schnabel R, Taylor J and Schenkel F, 2009. Invited review: Reliability of genomic predictions for North American Holstein bulls. *Journal of Dairy Science* 92(1):16-24.
- Wiggans GR, Cooper TA, VanRaden PM and Silva MV, 2010. Increased reliability of genetic evaluations for dairy cattle in the United States from use of genomic information. Page 476. *Proceeding of 9th World Congress Genetic Applied Livestock Production, Leipzig, Germany, German Society for Animal Science, Leipzig, Germany (Abstr).*
- Wiggans GR, Vanraden PM and Cooper TA, 2012. Technical note: adjustment of all cow evaluations for yield traits to be comparable with bull evaluations. *Journal of Dairy Science* 95(6):3444-3447.

Cow selective genotyping strategies for genomic selection programs in some simulated dairy cattle

E Dehnavi^{1*}, S Ansari Mahyari², F S Schenkel³ and M Sargolzaei^{4, 5}

Received: January 25, 2017 Accepted: November 29, 2017

¹PhD Graduate, Department of Animal Science, College of Agriculture, Isfahan University of Technology, Isfahan, Iran

²Associate Professor, Department of Animal Science, College of Agriculture, Isfahan University of Technology, Isfahan, Iran

³Professor, Centre for Genetic Improvement of Livestock (CGIL), University of Guelph, Guelph, ON, Canada

⁴Associated professor, Centre for Genetic Improvement of Livestock (CGIL), University of Guelph, Guelph, ON, Canada

⁵Semex Alliance, Guelph, ON, Canada

*Corresponding author: Email: e.dehnavi@ag.iut.ac.ir

Introduction: Adding genotypic and phenotypic information from cows to the reference population has the potential to increase the accuracy and decrease selection bias of genomic estimated breeding values (GEBV) in dairy cattle populations (Mc-Hugh et al. 2011; Thomasen et al. 2014; Wiggans et al. 2010). However, the cost of genotyping limits its use to an informative subset of animals (Boligon et al. 2012) and hence looking for an optimal strategy is crucial herein. Now a day, selective genotyping strategies (Boligon et al. 2012; Jiménez-Montero et al. 2012) and imputation methods (Calus et al. 2014) are some suggestive tools to reduce genotyping costs. Therefore, the aim of this study was to assess the effect of inclusion of cow genomic data in the reference population for genomic prediction under scenarios encompassing different selection strategies, number of genotyped cows and heritability on accuracy and bias of genomic breeding values when compared to the use of only bulls in the reference population.

Material and methods: A dairy cattle population was simulated by mimicking real linkage disequilibrium (LD) extent and population structure using QMSim software (Sargolzaei and Schenkel, 2009). This population mimicked a 4-pathway dairy cattle selection program with the use of artificial insemination (AI) technology, followed by a population with the use of genomic selection (GS). Within the GS reference population, pseudo-phenotypes were simulated for bulls and cows by calculating different predefined accuracies based on de-regressed EBV of Holstein cattle in Canada. That reflected difference between accuracy of EBV from a bull and accuracy of EBV from a cow which is based on his daughters' records and her own and daughters' records, respectively. Parameters of population simulation in this study were based on Dehnavi et al (2018). Two traits with a heritability of 0.30 and 0.05 (representing production and functional traits, respectively) were simulated independently. The reference population consisted of 5000 top selected bulls plus either 2500, 5000 or 10000 cows. Cows were selected randomly (R), based on the highest breeding values (H), the most accurate breeding values from top distribution tail (AH), two tails of the distribution of breeding values (TT), or the most accurate breeding values from two distribution tails (ATT) (Fig 1). The simulation was repeated 20 times for each scenario. Genomic predictions were computed using snpBLUP method in SNP1101 software (Sargolzaei, 2014). Pearson correlation, mean square error (MSE) between true breeding value (TBV) and direct genomic value (DGV), and coefficient regression of TBV on DGV were calculated as measurement criteria for accuracy, error and bias of genomic predictions, respectively.

Results and discussion: The extent of linkage disequilibrium in the simulated AI population showed a similar pattern that observed in North American Holstein cattle (Fig 2). The ascertainment bias was

introduced according to the minor allele frequency (MAF) distribution observed in North American Holstein data (Fig 3). The ATT scenario resulted in higher accuracy of GEBV (Fig 4), lower MSE (Fig 5) and bias (Table 1) compared to other scenarios. This scenario led to 0.123 to 0.215 gain in accuracy (Δ accuracy) compared to the use of bulls only for traits with heritability equal to 0.30 and 0.05, respectively (Fig 4). R and TT scenarios followed ATT. Gains in accuracy were 0.117 to 0.204 and 0.113 to 0.196 for scenarios R and TT, respectively (Fig 4). Regardless of EBV accuracy, selecting cows with high breeding values (H, AH) led to the lowest accuracy of GEBV (Fig 4) and the highest MSE (Fig 5) and bias (Table 1). Overall the observed gains in accuracy for the two traits varied from 0.062 to 0.149 and 0.009 to 0.07 for scenarios AH and H, respectively (Fig 4). An increase in the number of cows decreased both MSE (Fig 5) and bias (Table 1). The average regression coefficient of TBV on DGV across all scenarios varied from 1.12 for 2500 cows to 1.08 for 10000 cows for the trait with a heritability of 0.30 (Table 1). These coefficients ranged from 1.16 to 1.14 for 2500 to 10000 cows for the trait with heritability of 0.05 (Table 1). Similar to these results, using Guernsey breed Jenko et al (2017) showed when only half of population was genotyped, genotyping cows with phenotypes in extremes was superior up to 8-10 folds in accuracy for yield traits than genotyping cows at random or genotyping cows with upper tail phenotypes. Genotyping cows with tail phenotypes can cover on average 88% of the difference between the scenario where all the cows were genotyped or only half of them were genotyped at random (Jenko et al. 2017). The present study showed that sampling cows from the most accurate EBVs on extremes seemed to be more informative on all SNPs, both favorite and unfavorable alleles in the population.

Conclusion: The inclusion of cows in the reference population increased the accuracy of the genomic predictions across all scenarios and decreased bias of them as well. However, sampling cows from the most accurate EBVs from two distribution tails seemed to be more useful to precisely and accurately predict genomic breeding values of young animals than the other sampling strategies that were investigated in this study.

Keywords: Cows selective genotyping, Genomic evaluation, Reference population, Validation population.