

## اثر افزایش تعداد ماده‌های جمعیت مرجع و نشانگرهای امپوت شده بر قابلیت اعتماد پیش‌بینی ژنومی گاوهای جرسی

داود کریمی<sup>۱\*</sup> - مجتبی طهمورث‌پور<sup>۲</sup> - محمد دادپسند<sup>۳</sup> - علی اصغر اسلمی نژاد<sup>۴</sup> - موزر ساندو لوند<sup>۵</sup>

تاریخ دریافت: ۱۳۹۲/۵/۲

تاریخ پذیرش: ۱۳۹۳/۶/۲

### چکیده

هدف از این تحقیق، ارزیابی صحت حاصل از امپوت گوی نشانگری کم تراکم (بطور میانگین در ۱۰ تکرار ۶۶۴۰ نشانگر) به گوی نشانگری تراکم متوسط در جمعیت شبیه‌سازی شده گاو جرسی با استفاده از نرم‌افزار FImpute و ارزیابی قابلیت اعتماد ارزش اصلاحی ژنومی برآورد شده برای دو صفت تولید شیر ( $h^2=0/40$ ) و باروری ( $h^2=0/04$ ) بر اساس تعداد مختلف ماده‌های جمعیت مرجع بود. نهمصد رأس گاو نر که براساس گوی نشانگری تراکم متوسط ( $50K$ ) تعیین ژنوتیپ شده بودند، به منظور ارزیابی صحت امپوت تعداد مختلف ماده‌های جمعیت مرجع بکار گرفته شدند. نه سناریو برای پیش‌بینی ارزش اصلاحی ژنومی ۳۰۰۰ حیوان تصادفی انتخاب شده جمعیت آزمون (۱۰۰۰ حیوان در هر نسل) طراحی و ارزیابی شد. میانگین صحت امپوت ژنوتیپ ماده‌ها به گوی تراکم متوسط ۹۸/۶۴ درصد بود. با توجه به میزان بالای نرخ امپوت نشانگرها تنها یک درصد اختلاف بین ارزش اصلاحی ژنومی پیش‌بینی شده حاصل از نشانگرهای ۵۰K مقایسه با نشانگرهای امپوت شده وجود داشت. با توجه به این اختلاف جزئی و از آنجا که هزینه تعیین ژنوتیپ یک حیوان بر مبنای گوی تراکم متوسط دو برابر گوی کم تراکم است، توصیه می‌شود که تمام نرهای پروف‌شده نسل مرجع براساس گوی ۵۰K تعیین ژنوتیپ شده و همه ماده‌ها به گوی تراکم متوسط امپوت شوند تا علاوه بر دستیابی به صحت بالای پیش‌بینی ارزش اصلاحی ژنومی برای هر دو صفت، بتوان با کاهش چشمگیر هزینه تعیین ژنوتیپ، تعداد حیوانات بیشتری را در جمعیت مرجع بکار برد که خود منجر به افزایش اطلاعات فنوتیپی، شجره‌ای و ژنتیکی و در نهایت صحت پیش‌بینی ژنومی می‌شود.

**کلمات کلیدی:** ارزیابی ژنومی، جمعیت مرجع، امپوت، تراکم نشانگر، گاو شیری.

### مقدمه

پذیر است (۱). در اصلاح نژاد گاوهای شیری، استفاده از آرایه‌ای شامل ۵۰۰۰۰ نشانگر که به صورت تجاری در دسترس است، فرصتی برای انتخاب مؤثر با استفاده از نشانگرهای متراکم را فراهم می‌آورد. انتخاب ژنومی بر اساس ارزش اصلاحی که بطور مستقیم از گوی نشانگری<sup>۲</sup> متراکم و گسترده در سطح ژنوم برآورد می‌گردد، بنا نهاده شده است (۱۹). در انتخاب ژنومی بر اساس روش<sup>۳</sup> GBLUP فرض می‌شود که نشانگرها در عدم تعادل پیوستگی ژنی<sup>۴</sup> (LD) با QTLها می‌باشند، و میزان اثر هر QTLی که باید برآورد شود، خیلی جزئی است. آنها نشان دادند که تنها با استفاده از اطلاعات نشانگرها می‌توان ارزش اصلاحی را با صحت ۰/۸۵ پیش‌بینی کرد (۱۹). پیش-بینی ارزش اصلاحی ژنومی در مقایسه با روش کلاسیک BLUP،

در دهه گذشته، کاربرد اطلاعات ژنتیک مولکولی در اصلاح نژاد دامها، بسیار مورد توجه قرار گرفته است. هر گونه پیشرفت در بهبود تولید شیر حیوانات بستگی به استفاده از روش های اصلاح نژادی دارد که با مشخص نمودن میزان وراثت پذیری، همبستگی ژنتیکی، بکار گیری روش های مناسب انتخاب و برآورد پیشرفت ژنتیکی امکان

۱- دانشجوی دکتری ژنتیک و اصلاح دام، گروه علوم دامی دانشکده کشاورزی دانشگاه فردوسی مشهد،

۲- استاد گروه علوم دامی دانشکده کشاورزی دانشگاه فردوسی مشهد،

۳- استادیار گروه علوم دامی دانشکده کشاورزی دانشگاه شیراز،

۴- دانشیار گروه علوم دامی دانشکده کشاورزی دانشگاه فردوسی مشهد،

۵- رییس مرکز ژنتیک کمی و ژنومیکس گروه زیست‌شناسی مولکولی و ژنتیک دانشگاه آرهوس دانمارک.

(\* - نویسنده مسئول: (Email: karimi.davood@gmail.com)

2- Marker panel

3- Genomic BLUP

4- Linkage disequilibrium

این روش‌ها، BLUP ژنومی (GBLUP) است که برای تمام اثرات نشانگرها واریانس یکسانی را فرض می‌کند (۱۹). این روش در ارزیابی ژنومی با استفاده از نشانگرهای تراکم متوسط<sup>۸</sup> بسیار معمول بوده است (۱۱، ۱۵، ۱۹ و ۳۱)، و بدلیل سادگی در مقایسه با بسیاری از مدل‌های پیچیده و زمان‌گیر، نیازهای محاسباتی آن نسبتاً کم است (۱۲). تا به امروز استفاده از ماده‌ها در جمعیت مرجع محدود به نژادهای بزرگ گاو شیری نظیر هلشتاین بوده و استفاده از ماده‌های تجاری در جمعیت مرجع گاوهای جرسی گزارش نشده است. در نژادهای کوچک گاو شیری مانند جرسی، به علت کوچک بودن جمعیت مرجع (کمبود نرهای آزمون نتاج شده) در مقایسه با نژاد هلشتاین، استفاده از ماده‌ها و تعیین ژنوتیپ آنها برای افزایش اطلاعات حیوانات جمعیت مرجع و پیش‌بینی دقیق‌تر ارزش اصلاحی حیوانات جمعیت آزمون ضروری می‌نماید. از طرفی به منظور کاهش هزینه‌های تعیین ژنوتیپ ماده‌ها، از الگوی ژنوتیپی کم تراکم و امپيوت آن به الگوی تراکم متوسط استفاده می‌شود.

اهداف این پژوهش، مقایسه پیش‌بینی‌های ژنومی حاصل از نشانگرهای امپيوت‌شده به سطح تراکم متوسط و نشانگرهای تراکم متوسط با بکارگیری مدل خطی BLUP ژنومی (GBLUP) و بررسی اثر تعداد متفاوت ماده‌های نسل مرجع بر صحت ارزش اصلاحی حیوانات نسل آزمون بود.

## مواد و روش‌ها

### شبیه‌سازی

در این پژوهش یک جمعیت گاو شیری بر اساس ساختار جمعیت گاوهای جرسی دانمارک و با استفاده از برنامه شبیه‌ساز اصلاحی ADAM (۲۳) در ۱۰ تکرار شبیه‌سازی گردید. ساختار داده‌ها برای انحراف عملکرد دختران<sup>۹</sup> (DYD) بر اساس مشاهدات ۲۰۰ دختر برای گاوهای نر با صفات دارای وراثت‌پذیری بالا (۰/۴۰) و وراثت-پذیری کم (۰/۰۴) شبیه‌سازی شد. در ابتدا ۱۰۰۰ نسل با اندازه جمعیت مؤثر<sup>۱۰</sup> ۴۰۰ راس که شامل ۲۰۰ نر و ۲۰۰ ماده بود برای تشکیل جامعه پایه<sup>۱۱</sup> ایجاد شد. تمام جایگاه‌های ژنی در نسل نخست دارای آلل‌های مجزا<sup>۱۲</sup> به صورت ۱ و ۲ با فراوانی آللی ۰/۵ بودند. برای دستیابی به LD در جمعیت، در ۱۰۰۰ نسل اول آمیزش‌ها بصورت تصادفی صورت گرفت. نرخ جهش<sup>۱۳</sup>  $1/8 \times 10^{-7}$  در هر جایگاه و هر نسل بود. پس از تشکیل جامعه پایه، در نهایت تعداد گاوهای نر به

منجر به افزایش دو تا سه برابری پیشرفت ژنتیکی و کاهش ۹۲ درصدی هزینه‌ها به علت کاهش فاصله نسل در مقایسه با آزمون نتاج می‌شود (۲۷).

برنامه‌های اصلاحی مبتنی بر انتخاب ژنومی، معادله پیش‌بینی را در جمعیت مرجع<sup>۱</sup> که دارای اطلاعات ژنوتیپی و فنوتیپی هستند، برآورد می‌کند. سپس از این معادله برای پیش‌بینی ارزش ژنومی مستقیم<sup>۲</sup> (DGV) حیوانات نسل آزمون<sup>۳</sup> که اطلاعات فنوتیپی آنها در دسترس نیست، مورد استفاده قرار می‌گیرد (۱۹). صحت پیش‌بینی DGV تابعی از تعداد افراد دارای رکوردهای فنوتیپی و اطلاعات ژنوتیپی برای تخمین اثرات مارکرها (۵)، تراکم نشانگرهای تعیین ژنوتیپ شده، LD بین QTL و نشانگر (۲۹) و وراثت‌پذیری صفت (۱۲) است.

موزر و همکاران (۲۱) گزارش کردند که استفاده از الگوهای نشانگری کم تراکم<sup>۴</sup> (۳K و ۶K) منجر به کاهش شدید قابلیت اعتماد پیش‌بینی ژنومی می‌شود (۲۱). یک راهکار مناسب بسط دادن نشانگرهای کم تراکم به سطح ۵۰۰۰۰ نشانگر<sup>۵</sup> (۵۰K) با روش امپيوت<sup>۶</sup> و کاهش هزینه‌ها و امکان تعیین ژنوتیپ تعداد بیشتری گاوهای نر و ماده است. امپيوت، در واقع، یک روش آسانی برای پیش‌بینی نشانگرهای یک الگوی کم تراکم و ارتقاء آن به الگوی تراکم بالاتر، با استفاده از اطلاعات نشانگرهای الگوی تراکم متوسط و تراکم بالا است. چندین روش برای امپيوت نشانگرهای SNP<sup>۷</sup> وجود دارد، این روش‌ها یا متکی بر پیوستگی ژنتیکی بین خویشاوندان (۵) و یا براساس LD مبنی بر اطلاعات جمعیت می‌باشند (۲۸). همچنین امکان ترکیب هر دو روش و استفاده از مزایای آنها نیز وجود دارد (۱۳ و ۲۰).

صحت امپيوت نشانگرها و GEBV حیوانات جوان با افزایش اندازه جمعیت مرجع افزایش می‌یابد (۹ و ۲۵). افزایش اندازه جمعیت مرجع منجر به افزایش تعداد حیوانات تعیین ژنوتیپ شده گردیده و در نتیجه باعث افزایش هزینه انتخاب ژنومی می‌شود. افزایش تعداد حیوانات تعیین ژنوتیپ نشده به منظور پیش‌بینی (امپيوت) ژنوتیپ آنها یک روش ارزان برای افزایش صحت GEBV است. به علاوه، هزینه‌های نسبتاً زیاد تعیین ژنوتیپ، تاکنون باعث شده است که گاوهای نر آزمون نتاج شده در جمعیت مرجع استفاده شوند.

تاکنون، شیوه‌های متفاوتی از به کارگیری اطلاعات ژنومی در برآورد ارزش اصلاحی ارائه شده است (۱۴، ۱۶، ۱۹ و ۲۲). یکی از

- 1- Reference population
- 2- Direct genomic value
- 3- Test population
- 4- Low density
- 5- Medium density
- 6- Imputation
- 7- Single nucleotide polymorphism

- 8- Medium density
- 9- Daughter yield deviation
- 10- Effective population size
- 11- Base population
- 12- Segregating alleles

۶۰ رأس کاهش یافت تا اینکه با آمیزش تصادفی با ۱۰۰ گاو ماده (۲ نتاج برای هر گاو ماده) در هر نسل، پس از ۲۰ نسل جمعیت نژاد گاوهای شیری مورد نظر شبیه‌سازی شد. برای تمام حیوانات حاضر در نسل‌های ۱ تا ۱۷، اطلاعات شجره‌ای، فنوتیپی (DYD برای گاوهای نر آزمون نتاج شده)، ارزش اصلاحی واقعی<sup>۱</sup> (TBV) و ژنوتیپی شبیه‌سازی گردید. نسل‌های ۱۸، ۱۹ و ۲۰ شامل حیواناتی بود که هیچ رکورد فنوتیپی از آنها وجود نداشت.

به منظور تطابق هرچه بیشتر پارامترهای جمعیت شبیه‌سازی با واقعیت، ژنوم شبیه‌سازی شده دارای ۳۰ کروموزوم هر یک به طول ۱۰۰ سانتی‌مورگان بود (طول ژنوم ۳۰ مورگان). جایگاه‌های ژنی نشانگرها (فایل ژنوم اصلی) که بطور متوسط ۱۵۹۴۰۰ نشانگر (میانگین ۱۰ تکرار) بود، بطور مساوی با فاصله ۰/۰۱۸ سانتی‌مورگان در سراسر ژنوم قرار گرفتند. بر اساس فایل ژنوم اصلی، برای هر حیوان الگوهای نشانگری ۶K و ۵۰K استخراج گردید. ارزش اصلاحی واقعی (TBV)، بر اساس توزیع نرمال با میانگین صفر و واریانس یک، برای اثر همه QTLها شبیه‌سازی شد. ارزش اصلاحی واقعی هر حیوان از مجموع اثر همه QTLها (۵۱۲۳ QTL به طور میانگین در ۱۰ تکرار) و با فرض مستقل بودن آنها به دست آمد.

## سناریوها

به منظور دستیابی به اهداف این پژوهش، مراحل شبیه‌سازی فوق برای دو سطح وراثت‌پذیری (۰/۴۰ و ۰/۰۴) در ۱۰ تکرار انجام شد. شش سناریو با تعداد مختلف ماده‌های جمعیت مرجع در دو سطح مختلف تراکم نشانگری (تراکم متوسط و امپیوت شده‌ها) در نظر گرفته شد. در سناریوی اول، GEBV حیوانات جوان نسل آزمون با استفاده از جمعیت مرجعی که شامل تمام گاوهای نر آزمون نتاج شده (شامل ۹۰۰ رأس) حاضر در نسل ۱ تا ۱۵ بود، برآورد گردید. در این تحقیق همه ۹۰۰ گاو نر آزمون نتاج شده حاضر در جمعیت مرجع برای الگوی نشانگری تراکم متوسط (۵۰K) تعیین ژنوتیپ شده‌اند. اطلاعات فنوتیپی گاوهای نر بر اساس DYD آنها و اطلاعات فنوتیپی گاوهای ماده رکوردهای فنوتیپی آنها بوده است.

در سناریوی دوم، به منظور بررسی نقش گاوهای ماده در برآورد GEBV حیوانات نسل آزمون، ۱۰۰۰ گاو ماده که بر اساس الگوی نشانگری تراکم متوسط (۵۰K) تعیین ژنوتیپ شده‌اند به جمعیت مرجع که شامل ۹۰۰ گاو نر بود، اضافه شدند. از آنجا که حدوداً هزینه تعیین ژنوتیپ یک حیوان با الگوی نشانگری تراکم متوسط دو برابر هزینه تعیین ژنوتیپ با الگوی نشانگری کم تراکم است (۱۰)، سناریوهای ۳ تا ۸ دارای تعداد مشابهی از اندازه جمعیت مرجع اما با تراکم‌های متفاوت الگوهای نشانگری است. بطور مثال، در سناریو

سوم، جمعیت مرجع متشکل از ۹۰۰ گاو نر ۵۰K بعلاوه ۲۰۰۰ گاو ماده امپیوت شده به ۵۰K است. از آنجا که سناریو سوم در مقایسه با سناریو دوم، تعداد گاوهای ماده دو برابر می‌باشد، اما بعلت تعیین ژنوتیپ آنها با الگوی تراکم کم (۶K)، هزینه تعیین ژنوتیپ اضافی پرداخت نمی‌گردد، در حالیکه اندازه جمعیت مرجع برای تعداد ماده‌ها دو برابر می‌شود. از طرفی بمنظور بررسی صحت GEBV حاصله از ماده‌های امپیوت شده به ۵۰K در مقایسه با صحت ماده‌های ۵۰K، سناریوی چهارم، دارای ساختار مشابه سناریو سوم است با این تفاوت که همه ۲۹۰۰ حیوان جمعیت مرجع (۹۰۰ گاو نر بعلاوه ۲۰۰۰ گاو ماده) برای الگوی تراکم نشانگری ۵۰K تعیین ژنوتیپ شدند. سناریو پنجم متشکل از ۴۹۰۰ حیوان است که بجز گاوهای نر که ۵۰K هستند، گاوهای ماده جمعیت مرجع به ۵۰K امپیوت شده‌اند. در سناریو ششم همه ۴۹۰۰ حیوان بر اساس الگوی نشانگری ۵۰K تعیین ژنوتیپ شدند. در سناریوهای هفتم و هشتم تعداد حیوانات جمعیت مرجع ۸۹۰۰ بود با این تفاوت که ۸۰۰۰ گاو ماده در سناریو هفتم به ۵۰K امپیوت شدند. در نهایت، سناریو نهم از ۱۶۰۰۰ گاو ماده امپیوت شده به ۵۰K و ۹۰۰ گاو نر ۵۰K تشکیل شده بود. بعلاوه، از بین حیوانات نسل ۱۸ تا ۲۰ (نسل آزمون)، ۳۰۰۰ حیوان بطور تصادفی (۱۰۰۰ در هر نسل) انتخاب شدند و بر اساس الگوی ۵۰K تعیین ژنوتیپ گردیدند. همچنین برای نزدیک بودن ماده‌های جمعیت مرجع با حیوانات جمعیت آزمون و بالاتر بودن GEBV، گاوهای ماده نسل‌های ۱۳ تا ۱۷ بطور تصادفی انتخاب و در ساختار جمعیت مرجع بکار رفتند.

## امپیوت نشانگرها

پیش‌بینی (امپیوت) نشانگرهای مفقود<sup>۲</sup> با استفاده از نرم‌افزار FImpute نسخه ۲/۲ (۲۶) برای ژنوم هر فرد بصورت جداگانه صورت گرفت. در مقایسه با سایر نرم‌افزارهای امپیوت نشانگرها، FImpute بدلیل اینکه هم از اطلاعات پیوستگی میان خانواده‌ها و هم اطلاعات LD موجود در جمعیت استفاده می‌کند از صحت بالایی (بیش از ۹۵ درصد) در پیش‌بینی نشانگرها برخوردار است، در ضمن در مقایسه با سایر روش‌ها بسیار سریع‌تر و مستلزم نیاز محاسباتی بسیار کمتری است (۱۷). در تمام سناریوهایی که گاوهای ماده با استفاده از نشانگرهای ۶K تعیین ژنوتیپ شده بودند (سناریوهای ۳، ۵، ۷ و ۹)، از اطلاعات نشانگرهای ۹۰۰ گاو نر آزمون نتاج شده (که در فرایند امپیوت، جمعیت معرف<sup>۳</sup> خوانده می‌شوند) موجود در جمعیت مرجع به عنوان منبع اطلاعات برای امپیوت ماده‌ها (ماده‌ها در فرایند امپیوت، جمعیت هدف<sup>۴</sup> خوانده می‌شوند) استفاده شد. امپیوت ژنوم

2- Missing marker

3- Training population

4- Target population

1- True breeding value

ششم انجام گرفت. از آنجا که نرهای آزمون نتاج شده در مقایسه با ماده‌ها فرزندان بیشتری داشتند، در نتیجه از آنالیز آماری وزنی<sup>۳</sup> استفاده شد.

قابلیت اعتماد GEBV بر اساس توان دوم همبستگی بین GEBV و TBV حیوانات نسل‌های ۱۸، ۱۹ و ۲۰ (نسل آزمون) محاسبه گردید ( $R^2_{GEBV, TBV}$ ).

## نتایج و بحث

### صحت امپيوت نشانگرها

در این مطالعه ۱۰ تکرار با موفقیت شبیه‌سازی گردید. از آنجا که در هر تکرار ۹ سناریو بررسی شد، میزان صحت روش امپيوت و قابلیت اعتماد GEBV برای هر سناریو توسط میانگین ۱۰ تکرار محاسبه شد. از میان ۹ سناریو، روش امپيوت و پیش‌بینی نشانگرهای مفقود برای ۴ سناریو در نظر گرفته شد. میزان میانگین خطای پیش‌بینی نشانگرها بین ۱/۳ تا ۱/۳۹ درصد برای تعداد مختلف حیوانات جمعیت هدف متغیر بود. در تمام مراحل امپيوت نشانگرهای ۶K به ۵۰K، تنها از اطلاعات نشانگری تمام گاوهای نر معرف (۹۰۰ گاو نر ۵۰K) آزمون نتاج شده‌ای که در جمعیت مرجع بودند برای پیش‌بینی نشانگرهای تعداد مختلف گاوهای ماده اضافه شده به جمعیت مرجع استفاده شد. با ثابت بودن تعداد نرها (جمعیت معرف) و با افزایش تعداد ماده‌ها (جمعیت هدف) از ۲۰۰۰ به ۱۶۰۰۰ صحت امپيوت ۰/۰۹ درصد افزایش یافت. داسونیل و همکاران (۷) با بررسی دو جمعیت گاوهای شیری کشورهای اسکاندیناوی و فرانسه گزارش کردند که با افزایش تعداد حیوانات جمعیت معرف از طریق اطلاعات حیوانات پروژه ژنومیکس اتحادیه اروپا نرخ خطای امپيوت نشانگرها با استفاده از نرم‌افزارهای PHASEBOOK (۸) و Beagle (۳)، در کشورهای اسکاندیناوی از ۵/۵ درصد به ۴ درصد و در فرانسه از ۳/۹ درصد به ۲/۱ درصد کاهش یافت. آنها علت نرخ کمتر خطای امپيوت را در بین گاوهای شیری فرانسه استفاده از نشانگرهای بیشتر بخاطر نحوه ویرایش نشانگرها و داشتن ژنوم متراکم‌تر دانستند (۷). در مطالعه دیگری، ژانگ و دروت (۳۳) نشان دادند که تعداد حیوانات جمعیت معرف و تعداد نشانگرها در الگوی کم تراکم بر نرخ خطای امپيوت اثر می‌گذارد.

نرخ خطای امپيوت همچنین تحت تأثیر ارتباط بین حیوانات جمعیت معرف و جمعیت هدف قرار دارد. در جمعیت گاوهای شیری کشورهای اسکاندیناوی مشاهده شد که میانگین نرخ خطای امپيوت به این موضوع که آیا پدر حیوانات نسل هدف در جمعیت معرف حضور دارد یا خیر بستگی دارد. حضور پدر حیوانات نسل هدف در

برای تک تک کروموزومها جداگانه صورت گرفت و با توجه به اینکه تقریباً ۸۳ درصد ژنوم ماده‌ها امپيوت گردید، فرآیند امپيوت ژنوم برای تمام سناریوها و تکرارها بدون اینکه نشانگری مفقود بماند، با موفقیت انجام شد. نشانگرهای امپيوت شده و ۵۰K بر اساس معیار حداقل فراوانی آلی<sup>۱</sup> (MAF) بزرگتر از ۰/۰۱ و میانگین میزان فراخوان ژنی بیش از ۶۰ درصد ویرایش گردیدند.

### محاسبه نرخ خطای امپيوت آلل‌ها

برای محاسبه نرخ خطای امپيوت آلل‌ها، پس از اتمام فرآیند امپيوت نشانگرها توسط FImpute، با استفاده از زبان برنامه‌نویسی FORTRAN F90، یک فایل با تراکم نشانگری ۵۰K برای ماده‌هایی که امپيوت شده بودند از طریق فایل ژنوم اصلی آنها ایجاد شد. سپس با مقایسه فایل اصلی ۵۰K و فایل امپيوت شده، اگر آلل‌های نشانگر امپيوت شده و اصلی مشابه بودند، تعداد خطا برابر صفر محاسبه می‌گردید و در صورتیکه آلل نشانگر فایل اصلی ۵۰K هموزیگوت می‌بود و در فرآیند امپيوت به اشتباه هتروزیگوت و یا بالعکس امپيوت شده بود، عدد یک و زمانیکه آلل نشانگر اصلی و امپيوت شده، هموزیگوت متضاد بودند عدد ۲ در نظر گرفته شد. نرخ خطا از مجموع تعداد کل نشانگرهای اشتباه امپيوت شده (اعداد غیر صفر) تقسیم بر تعداد نشانگرهای امپيوت شده ضربدر تعداد حیوانات امپيوت شده محاسبه گردید.

### پیش‌بینی ژنومی

پیش‌بینی DGV با استفاده از مدل GBLUP (۲۹) انجام شد. مدل GBLUP بصورت زیر بود:

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

که در اینجا،  $y$  بردار مشاهدات (DYD برای نرهای آزمون نتاج شده و رکوردهای فنوتیپی ماده‌ها)،  $\mu$  میانگین کل، 1 برداری متشکل از اعداد ۱،  $g$  بردار DGV،  $Z$  ماتریس طرح برای  $g$  و  $e$  بردار باقیمانده می‌باشد. فرض می‌شود که  $g$  دارای توزیع نرمال با میانگین صفر و واریانس  $G\sigma_g^2$  و  $e$  دارای توزیع نرمال با میانگین صفر و واریانس  $D\sigma_e^2$  است، در حالیکه  $G$  ماتریس روابط ژنومی،  $\sigma_g^2$  واریانس ژنتیکی افزایشی ژنومی،  $D$  یک ماتریس قطری و  $\sigma_e^2$  واریانس خطا (باقیمانده) است. ماتریس روابط ژنومی بصورت  $G = \frac{MM'}{\sum 2p_i q_i}$  است که عناصر ستون  $n$ م ماتریس  $M$ ،  $0 - 2p_i$ ،  $1 - 2p_i$  و  $2 - 2p_i$  به ترتیب برای ژنوتیپ‌های  $A_1A_1$ ،  $A_1A_2$  و  $A_2A_2$  و  $q_i$  فراوانی آلل  $A_1$  و  $p_i$  فراوانی آلل  $A_2$  هستند (۳۲). تجزیه آماری مختلط چند صفتی<sup>۲</sup> با نرم‌افزار DMU (۱۹) نسخه

1- Minor allele frequency

2- Multivariate mixed model

3- Weighted analysis

### پیش‌بینی ژنومی

قابلیت اعتماد  $GEV$  ( $R^2_{GEV,TBV}$ ) حیوانات جمعیت آزمون که از توان دوم همبستگی بین  $GEV$  و  $TBV$  برآورد گردید، در جدول ۱ ( $h^2=0/40$ ) و جدول ۲ ( $h^2=0/04$ ) نشان داده شده است. قابلیت اعتماد  $GEV$  برای حیوانات جمعیت آزمون، بدون توجه به وراثت‌پذیری صفت، با افزایش تعداد ماده‌های جمعیت آزمون بدلیل افزایش منابع اطلاعاتی اعم از اطلاعات فنوتیپی، سجره‌ای و ژنوتیپی افزایش می‌یابد. براساس جداول ۱ و ۲، در تمام سناریوها، بالاترین میزان صحت  $GEV$  برای حیوانات نسل ۱۸ بدست آمد، زیرا ارتباط ژنتیکی نزدیکتری بین آنها و جمعیت آزمون برقرار است و با زیاد شدن فاصله نسلی بین حیوانات جمعیت آزمون و مرجع، قابلیت اعتماد کاهش می‌یابد. ویگل و همکاران (۳۲) با بررسی امکان پیش‌بینی نشانگرهای حیوانات تعیین ژنوتیپ نشده در جمعیت گاوهای جرسی، نشان دادند که تفاوت بین قابلیت اعتماد  $GEV$  بدست آمده از حیواناتی که با الگوی ۵۰K تعیین ژنوتیپ شده‌اند در مقایسه با حیوانات امپیوت شده، در صفات با وراثت‌پذیری پایین، بیشتر بود (۳۲). یافته‌های این تحقیق با مشاهدات ویگل و همکاران (۳۲) همخوانی دارد. در نتیجه برای دستیابی به صحت بالاتر  $GEV$  در صفات با وراثت‌پذیری کم، نیاز به رکوردهای فنوتیپی بیشتر و الگوهایی با تراکم بالاتر ژنوتیپی می‌باشد.

جمعیت معرف، ارتباط ژنتیکی نزدیکتری را بین این دو جمعیت ایجاد می‌کند که منجر به کاهش نرخ خطای امپیوت می‌شود. بر اساس نتایج حاصله از مطالعه داسونیل و همکاران (۷)، برای جلوگیری از کاهش صحت روش امپیوت، همه گاوهای نر نسل مرجع باید برای الگوی ۵۰۰۰۰ نشانگری تعیین ژنوتیپ شوند. بر همین اساس، در این تحقیق، به علت تعیین ژنوتیپ ۵۰۰۰۰ نشانگری گاوهای نر جمعیت معرف و اینکه بر اساس ساختار جمعیت شبیه‌سازی شده، ارتباطات خویشاوندی نزدیکی بین حیوانات جمعیت معرف و هدف وجود دارد، صحت امپیوت حیوانات هدف سناریوهای ۳، ۵، ۷ و ۹ که به ترتیب دارای ۲۰۰۰، ۴۰۰۰، ۸۰۰۰ و ۱۶۰۰۰ گاو ماده بودند، ۹۸/۶۲، ۹۸/۶۱، ۹۸/۶۵ و ۹۸/۷ درصد بدست آمد. صحت زیاد بدست آمده از روش امپیوت نشان می‌دهد که ارتباط ژنتیکی نزدیکتری بین حیوانات جمعیت معرف و هدف وجود داشته و انتخاب الگوهای کم تراکم ۶K برای ماده‌ها و تراکم متوسط ۵۰K برای نرهای پروف شده و بکارگیری نرم‌افزار مناسب برای پیش‌بینی نشانگرها همگی از عوامل اثرگذار بر صحت امپیوت هستند. مقایسه نرخ خطای امپیوت بر اساس مطالعات مختلف دشوار است به دلیل اینکه ارتباط خویشاوندی بین جمعیت‌های مرجع و آزمون متفاوت است و همچنین بخاطر اینکه تعداد حیوانات جمعیت مرجع و تعداد نشانگرها تغییر می‌کند.

جدول ۱- مقایسه صحت ارزش ارثی ژنومی برآورد شده سناریوهای مختلف برای صفت تولید شیر ( $h^2=0/4$ )

صحت امپیوت	قابلیت اعتماد $GEV$		نسل	سناریو
	امپیوت شده‌ها	۵۰K		
-	-	۰/۳۷۹	۱۸	۹۰۰ گاو نر
	-	۰/۳۱۷	۱۹	
	-	۰/۲۷۹	۲۰	
-	-	۰/۴۲۰	۱۸	۹۰۰ گاو نر + ۱۰۰۰ گاو ماده
	-	۰/۳۷۶	۱۹	
	-	۰/۳۲۷	۲۰	
۹۸/۶۱	۰/۴۶۶	۰/۴۶۸	۱۸	۹۰۰ گاو نر + ۲۰۰۰ گاو ماده
	۰/۴۲۳	۰/۴۲۵	۱۹	
	۰/۳۷۵	۰/۳۷۷	۲۰	
۹۸/۶۲	۰/۵۳۳	۰/۵۳۷	۱۸	۹۰۰ گاو نر + ۴۰۰۰ گاو ماده
	۰/۵۰۸	۰/۵۱۱	۱۹	
	۰/۴۶۰	۰/۴۶۴	۲۰	
۹۸/۶۵	۰/۶۱۴	۰/۶۲۲	۱۸	۹۰۰ گاو نر + ۸۰۰۰ گاو ماده
	۰/۵۸۱	۰/۵۸۷	۱۹	
	۰/۵۵۰	۰/۵۵۸	۲۰	
۹۸/۷۰	۰/۶۸۷	-	۱۸	۹۰۰ گاو نر + ۱۶۰۰۰ گاو ماده
	۰/۶۶۷	-	۱۹	
	۰/۶۴۵	-	۲۰	

نشان داد که اندازه مؤثر جامعه یک عامل کلیدی در قابلیت اعتماد پیش‌بینی ژنومی است (۹). با توجه به نتایج به دست آمده (جدول ۱ و ۲) با افزایش اندازه مؤثر جمعیت مرجع و به دلیل اضافه شدن اطلاعات شجره‌ای، فنوتیپی و ژنوتیپی برای پیش‌بینی ارزش اصلاحی ژنومی، افزایش قابل توجهی در سناریوهای مختلفی که اندازه جمعیت مرجع افزایش می‌یافت، مشاهده شد. در این پژوهش، فرض بر این بوده است که به علت استفاده از مدل خطی مختلط اثرات تمام SNPها از توزیع نرمال پیروی می‌کنند که هر SNP واریانس یکسانی را در واریانس ژنتیکی کل ایجاد می‌کند. هیس و همکاران (۱۲) و ون‌رادن و همکاران (۳۱) با بکارگیری چنین مدلی و مقایسه آن با روش‌های دیگری که اثر SNPها از توزیع نرمال پیروی نمی‌کنند، صحت پیش‌بینی ژنومی تقریباً مشابهی را گزارش کردند (۱۲ و ۳۱). سو و همکاران (۲۹) نیز با استفاده از روش‌های GBLUP و بی‌زی در پیش‌بینی ارزش اصلاحی ژنومی و ارزیابی قابلیت اعتماد این روش‌ها، ۰/۵ درصد قابلیت اعتماد بیشتری از مدل بی‌زی در مقایسه با GBLUP بدست آوردند هر چند که در جمعیت گاوهای شیری قرمز هیچ‌گونه تفاوتی بین روش GBLUP و بی‌زی مشاهده نگردید (۲۹).

استفاده از نشانگرهای ۵۰K در مقایسه با امپيوت شده‌ها منجر به افزایش جزئی در قابلیت اعتماد DGV در هر دو صفت می‌شود. بطور متوسط، قابلیت اعتماد DGV بر اساس نشانگرهای ۵۰K حدود یک درصد بالاتر از DGV بدست آمده از نشانگرهای امپيوت شده بود. بر اساس صحت GEBV بدست آمده در صفات با وراثت‌پذیری بالا، هر چه اندازه جمعیت مرجع افزایش یابد، تفاوت بین GEBVهای بدست آمده از نشانگرهای ۵۰K در مقایسه با امپيوت شده‌ها بیشتر می‌شود (جدول ۱).

پژوهش‌های انجام شده به منظور امپيوت نشانگرها از ۳K به ۵۰K نشان دادند که نرخ خطای امپيوت آلی جزئی، حتی زمانی که حیوانات جمعیت مرجع با الگوی ۵۰K تعیین ژنوتیپ و حیوانات جمعیت آزمون به الگوی ۵۰K امپيوت شده باشند، منجر به کاهش چشمگیر قابلیت اعتماد پیش‌بینی ژنومی می‌شود. بر اساس نتایج حاصله از پژوهش‌های انجام‌شده در کشورهای اسکاندیناوی، فرانسه و آلمان، هر یک درصد خطای آلی امپيوت نشانگرها به طور میانگین منجر به از دست رفتن ۱/۳ درصد از قابلیت اعتماد ارزش اصلاحی ژنومی می‌شود (۴ و ۷).

گودارد (۹) در بررسی نقش اندازه مؤثر جمعیت حیوانات مرجع

جدول ۲- مقایسه صحت ارزش اصلاحی ژنومی برآورد شده سناریوهای مختلف برای صفت باروری ( $h^2=0/04$ )

صحت امپيوت	GEBV قابلیت اعتماد		نسل	سناریو
	امپيوت شده‌ها	K50		
-	-	۰/۲۴۲	۱۸	۹۰۰ گاو نر
	-	۰/۲۰۰	۱۹	
	-	۰/۱۶۴	۲۰	
-	-	۰/۲۶۰	۱۸	۹۰۰ گاو نر + ۱۰۰۰ گاو ماده
	-	۰/۲۲۹	۱۹	
	-	۰/۱۹۱	۲۰	
۹۸/۶۱	۰/۲۸۵	۰/۲۸۹	۱۸	۹۰۰ گاو نر + ۲۰۰۰ گاو ماده
	۰/۲۴۵	۰/۲۵۳	۱۹	
	۰/۲۰۷	۰/۲۱۵	۲۰	
۹۸/۶۲	۰/۳۲۷	۰/۳۲۹	۱۸	۹۰۰ گاو نر + ۴۰۰۰ گاو ماده
	۰/۲۹۳	۰/۲۹۶	۱۹	
	۰/۲۵۰	۰/۲۵۴	۲۰	
۹۸/۶۵	۰/۳۷۱	۰/۳۷۴	۱۸	۹۰۰ گاو نر + ۸۰۰۰ گاو ماده
	۰/۳۴۱	۰/۳۴۹	۱۹	
	۰/۲۸۶	۰/۲۹۵	۲۰	
۹۸/۷۰	۰/۴۵۶	-	۱۸	۹۰۰ گاو نر + ۱۶۰۰۰ گاو ماده
	۰/۴۳۴	-	۱۹	
	۰/۳۷۸	-	۲۰	

حضور ماده‌ها در جمعیت مرجع در مقایسه با حضور نرها می‌برند (۲). ارزش اطلاعات فنوتیپی استفاده شده برای برآورد اثرات نشانگرها بطور نسبی برای ماده‌ها در مقایسه با نرها، با افزایش وراثت‌پذیری صفت افزایش می‌یابد. صحت پیش‌بینی ارزش اصلاحی با افزایش تعداد نسل‌های بین حیوانات مرجع و جمعیت آزمون کاهش می‌یابد (۱۹). ویگل و همکاران (۳۲) با استفاده از ۱۴۴۶ گاو نر جرسی بعنوان جمعیت مرجع برای صفاتی نظیر تولید شیر و درصد پروتئین که بر اساس الگوی ۵۰K تعیین ژنوتیپ شده بودند، گزارش کردند که توانایی پیش‌بینی DGV با استفاده از SNP ۲۹۴۲ که با فرآیند امپوت به ۵۰K ارتقاء می‌یابد، تقریباً ۹۷ درصد نسبت به زمانی است که از ژنوتیپ‌های واقعی استفاده می‌شود (۳۲). نتایج بدست آمده از این پژوهش به دلیل استفاده از نشانگرهای ۶K در مقایسه با ۲۹۴۲ SNP، صحت بیش از ۹۷ درصد را نشان می‌دهد.

### نتیجه‌گیری

قابلیت اعتماد کم پیش‌بینی‌های ژنومی جمعیت‌های کوچک گاو شیری، در اینجا گاوهای جرسی دانمارک، چالش همیشگی پیش روی محققان بوده است تا با استفاده از تکنولوژی‌های جدید راهی برای جبران آن پیدا شود. نتایج به دست آمده در گذشته نشان می‌دهد امکان اینکه بتوان با استفاده بیشتر از گاوهای نر جوان و برنامه‌های اصلاحی کارآمدتر، قابلیت اعتماد کم پیش‌بینی‌های ژنومی را جبران کرد، محدود بوده است. در نتیجه پیدا کردن راههایی برای افزایش این قابلیت اعتماد از مسایل مهمی می‌باشد که در جمعیت‌های نژادهای کوچک گاو شیری بسیار مورد توجه و تمرکز بوده است. به طور خلاصه، امپوت الگوهای نشانگری کم تراکم به الگوهای نشانگری تراکم متوسط علیرغم کاهش هزینه‌ها، موجب پیشرفت انتخاب ژنومی می‌شود. نتایج به دست آمده از این تحقیق آشکارانوعی افزایش قابلیت اعتماد پیش‌بینی‌های ژنومی و افزایش پیشرفت ژنتیکی در ازای بکار بردن ماده‌های جمعیت مرجع را نشان می‌دهد. از طرفی با استفاده از روش امپوت نشانگرها و راهکار مناسب برای انتخاب حیوانات جمعیت معرف می‌توان ماده‌های جمعیت مرجع را برای کاهش هزینه‌ها با الگوی کم تراکم تعیین توالی نمود و با استفاده از نرم‌افزارهای موجود از جمله FImpute و امپوت نشانگرها با دقت بالا (بیش از ۹۸ درصد) به الگوی تراکم بالاتر، برای دستیابی به دقت و قابلیت اعتماد ارزش اصلاحی ژنومی بیشتر اقدام نمود. افزایش تعداد حیوانات دارای رکورد فنوتیپی و تعیین ژنوتیپ شده، منجر به افزایش صحت پیش‌بینی ژنومی در کل جمعیت و بنابراین افزایش پیشرفت ژنتیکی ناشی از انتخاب ژنومی می‌شود.

باخ و همکاران (۲) در بررسی افزایش تعداد حیوانات جمعیت مرجع نشان دادند که بالاترین صحت DGV زمانی حاصل می‌شود که جمعیت مرجع از گاوهای نر آزمون نتاج شده، ماده‌ها و گاوهای نر جوان تشکیل شده باشد. جمعیت مرجعی که از گاوهای نر آزمون نتاج شده و ماده‌ها تشکیل شده باشد در مقایسه با جمعیت‌های مرجع شامل نرهای آزمون نتاج شده و گاوهای نر جوان، صحت بالاتری از DGV را حاصل می‌آورد. کمترین صحت برآورد DGV زمانی حاصل می‌شود که جمعیت مرجع تنها از گاوهای نر تشکیل شده باشد (۲). در تمام سناریوها صحت DGV برای صفات با وراثت‌پذیری بالاتر در مقایسه با صفات با وراثت‌پذیری پایین، بالاتر بود (جدول‌های ۱ و ۲). از آنجاکه هزینه تعیین ژنوتیپ یک حیوان با الگوی ۵۰K دو برابر الگوی ۶K است، لذا با توجه به سناریوهای ۳ و ۲ و یکسان بودن هزینه تعیین ژنوتیپ برای هر دو سناریو و با توجه به اینکه رکوردهای فنوتیپی ماده‌ها در دسترس باشد، با استفاده از امپوت، صحت GEBV با افزایش تعداد ماده‌ها از ۱۰۰۰ به ۲۰۰۰ به میزان متوسط ۰/۸۹ درصد برای  $h^2=0/4$  و ۰/۹۲ درصد برای  $h^2=0/4$  افزایش می‌یابد (جدول ۱ و ۲). همچنین با مقایسه سناریوهای ۸ و ۹، درحالی‌که در سناریو ۸، ۸۹۰۰ (۹۰۰ گاو نر و ۸۰۰۰ گاو ماده) حیوان با الگوی ۵۰K تعیین ژنوتیپ شده‌اند در برابر سناریو ۹ (۹۰۰ گاو نر ۵۰K و ۱۶۰۰۰ گاو ماده ۶K) که ماده‌های جمعیت مرجع به ۵۰K امپوت شده‌اند، با صرف هزینه یکسانی برای تعیین ژنوتیپ، بطور متوسط برای وراثت‌پذیری بالا ۰/۸۶ درصد و برای وراثت‌پذیری پایین ۰/۸۰ درصد قابلیت اعتماد بیشتری از DGV حیوانات نسل آزمون بدست می‌آید (جدول ۱ و ۲).

بر اساس جداول ۱ و ۲، با مقایسه دو به دو سناریوهای ۳ و ۴، ۵ و ۶ و ۷ و ۸ که شامل تعداد یکسانی از حیوانات نسل مرجع هستند، اما الگوی تعیین ژنوتیپ یکسانی برای ماده‌ها ندارند، به علت صحت بالای امپوت (۹۸/۶۱ تا ۹۸/۷ درصد)، اختلاف جزئی در قابلیت اعتماد GEBV برای حیوانات جمعیت آزمون مشاهده شد. در نتیجه، می‌توان ماده‌ها را در جمعیت مرجع برای الگوی کم تراکم تعیین ژنوتیپ نمود و با استفاده از امپوت، نشانگرهای کم تراکم آنها را به تراکم متوسط (۵۰K) ارتقا داد تا علاوه بر صرف هزینه‌های کمتر، بتوان ارزش اصلاحی حیوانات نسل آزمون را با صحت بالایی پیش‌بینی نمود.

بر اساس نتایج بدست آمده از این پژوهش، استفاده از ماده‌ها در جمعیت مرجع منجر به افزایش صحت DGV می‌شود. باخ و همکاران (۲) با بررسی ارزش ماده‌ها در جمعیت مرجع برای انتخاب ژنومی، نشان دادند که اگر رکوردهای فنوتیپی کمتری در دسترس باشد، استفاده از ماده‌ها در جمعیت مرجع اثر بزرگی بر معادله پیش‌بینی ژنومی دارد. به علاوه صفات با وراثت‌پذیری بالاتر سود بیشتری از

## تشکر و سپاس‌گزاری

مؤلفین از تمامی اعضاء گروه ژنتیک کمی و ژنومیکس دانشگاه آرهوس دانمارک، بخصوص آندرس کریستین سورنسن و گوشنگ سو به دلیل همکاری صادقانه نهایت تشکر و قدردانی را بعمل می‌آورند.

همچنین از آقای مهدی سرگلزهی (دانشگاه گوتلف کانادا) که در ارسال و به کارگیری نرم افزار FImpute نهایت همکاری را مبذول داشتند، تشکر و قدردانی می‌شود.

## منابع

- ۱- سید دخت، ع.، ع. اسلمی نژاد، و م. طهمورث پور. ۱۳۹۱. آنالیز ژنتیکی صفت تولید شیر گاوه‌های هلستاین استان تهران با استفاده از مدل روز آزمون. نشریه پژوهش‌های علوم دامی ایران. ۴: ۱۷۴-۱۶۸.
- 2- Buch, L. H., M. K. Sørensen, P. Berg, L. D. Pedersen, and A. C. Sørensen. 2012. Genomic selection strategies in dairy cattle: strong positive interaction between use of genotypic information and intensive use of young bulls. *J. Anim. Breed. Gene.* 129(2):138-51.
- 3- Browning, B. L., and S. R. Browning. 2008. A unified approach to genotype imputation and haplotype-phase inference for large data sets of trios and unrelated individuals. *Am. J. Hum. Genet.* 84:210-223.
- 4- Chen, J., Z. Liu, F. Reinhardt, and R. Reents. 2011. Reliability of genomic prediction using imputed genotypes for German Holsteins: Illumina 3K to 54K bovine chip. The 2011 Interbull Open Meeting, Stavanger, Norway. Interbull, Uppsala, Sweden.
- 5- Daetwyler, H. D., B. Villanueva, and J. A. Woolliams. 2008. Accuracy of predicting the genetic risk of disease using a genome-wide approach. *PLoS ONE* 3:e3395.
- 6- Daetwyler, H. D., G. R. Wiggans, B. J. Hayes, J. A. Woolliams, and M. E. Goddard. 2010. Imputation of missing genotypes from sparse to high density using long-range phasing. Manuscript 539 in Proc. World Congress of Genetics Applied to Livestock Production, Leipzig, Germany. [www.wcgalp2010.org](http://www.wcgalp2010.org).
- 7- Dasonneville, R., R. F. Brøndum, T. Druet, S. Fritz, F. Guillaume, B. Guldbandsen, M. S. Lund, V. Ducrocq, and G. Su. 2011. Effect of imputing markers from a low-density chip on the reliability of genomic breeding values in Holstein populations. *J. Dairy Sci.* 94:3679-3686.
- 8- Druet, T., and M. Georges. 2010. A hidden Markov model combining linkage and linkage disequilibrium information for haplotype reconstruction and quantitative trait locus fine mapping. *Genetics*, 184:789-798.
- 9- Goddard, M. 2009. Genomic selection: prediction of accuracy and maximization of long term response. *Genetica* 136, 245-257.
- 10- Habier, D., R. L. Fernando, and J. C. Dekkers. 2009. Genomic selection using low-density marker panels. *Genetics* 182:343-353.
- 11- Harris, B. L., and D. L. Johnson. 2010. Genomic predictions for New Zealand dairy bulls and integration with national genetic evaluation. *J. Dairy Sci.* 93:1243-1252.
- 12- Hayes, B. J., P. J. Bowman, A. J. Chamberlain, and M. E. Goddard. 2009. Invited review: Genomic selection in dairy cattle: Progress and challenges. *J. Dairy Sci.* 92:433-443.
- 13- Johnston, J., G. Kistemaker. 2011. Comparison of different imputation methods. Interbull open meeting. Stavanger, Norway.
- 14- Kolbehdari, D., L. R. Schaeffer, and J. A. B. Robinson. 2007. Estimation of genome wide haplotype effect in half sib designs. *J. Anim. Breed. Genet.* 124:356-361.
- 15- Liu, Z. T., F. R. Seefried, F. Reinhardt, S. Rensing, G. Thaller, and R. Reents. 2011. Impacts of both reference population size and inclusion of a residual polygenic effect on the accuracy of genomic prediction. *Genet. Sel. Evol.* 43:19.
- 16- Long, N., D. Gianola, G. J. M. Rosa, K. A. Weigel, and S. Avendano. 2007. Machine learning classification procedure for selecting SNPs in genomic selection: Application to early mortality in broilers. *J. Anim. Breed. Genet.* 124:377-389.
- 17- Ma, P., R. F. Brøndum, Q. Zhang, M. S. Lund, and G. Su. 2013. Comparison of different methods for imputing genome-wide marker genotypes in Swedish and Finnish Red Cattle. *J. Dairy Sci.* 96:4666-4677.
- 18- Madsen P., and J. Jensen. 2007. DMU: A user's Guide. A Package for Analyzing Multivariate Mixed Models. Version 6, Release 4.7. [http://dmu.agrsci.dk/dmuv6\\_guideR4-6-7.pdf](http://dmu.agrsci.dk/dmuv6_guideR4-6-7.pdf) Accessed Nov. 15.
- 19- Meuwissen, T. H. E., B. J. Hayes, and M. E. Goddard. 2001. Prediction of total genetic value using genome-wide dense marker maps. *Genetics*. 157:1819-1829.
- 20- Meuwissen, T., and M. Goddard. 2010. Accurate prediction of genetic values for complex traits by whole-genome



- resequencing. *Genetics*. 185:623–631.
- 21- Moser, G., M. S. Khatkar, B. J. Hayes, and H. W. Raadsma. 2010. Accuracy of direct genomic values in Holstein bulls and cows using subsets of SNP markers. *Genet. Sel. Evol.* 42:37.
  - 22- Muir, W. M. 2007. Comparison of genomic and traditional BLUP- estimated breeding value accuracy and selection response under alternative trait and genomic parameters. *J. Anim. Breed. Genet.* 124:342–355.
  - 23- Pedersen, L. D., A. C. Sørensen, M. Henryon, S. Ansari-Mahyari, and P. Berg. 2009. ADAM: A computer program to simulate selective breeding schemes for animals. *Livestock sci.* 121(2-3): 343-344.
  - 24- Pryce, J. E., M. E. Goddard, H. W. Raadsma, and B. J. Hayes. 2010. Deterministic models of breeding scheme designs that incorporate genomic selection. *J. Dairy Sci.* 93:5455–5466.
  - 25- Pszczola, M., A. Mulder, and M. P. L. Calus. 2010. Effect of enlarging the reference population with (un) genotyped animals on the accuracy of genomic selection in dairy cattle. *J. Dairy Sci.* 94:431-441.
  - 26- Sargolzaei, M., J. P. Chesnais, and F. S. Schenkel. 2011. FImpute- An efficient imputation algorithm for dairy cattle populations. *J. Dairy Sci.* 94(E-Suppl. 1):421. (Abstr).
  - 27- Schaeffer, L. R. 2006. Strategy for applying genome-wide selection in dairy cattle. *J. Anim. Breed. Genet.* 123:218–223.
  - 28- Scheet, P., and M. Stephens. 2006. A fast and flexible statistical model for large-scale population genotype data: Applications to inferring missing genotypes and haplotypic phase. *Am. J. Hum. Genet.* 78:629–644.
  - 29- Su, G., R. F. Brøndum, P. Ma, B. Guldbandsen, G. P. Aamand, and M. S. Lund. 2012. Comparison of genomic predictions using medium-density (~54,000) and high-density (~777,000) single nucleotide polymorphism marker panels in Nordic Holstein and Red Dairy Cattle populations. *J. Dairy Sci.* 95:4657–4665.
  - 30- VanRaden, P. M. 2008. Efficient methods to compute genomic predictions. *J. Dairy Sci.* 91:4414–4423.
  - 31- VanRaden, P. M., C. P. Van Tassell, G. R. Wiggans, T. S. Sonstegard, R. D. Schnabel, J. F. Taylor, and F. S. Schenkel. 2009. Invited review: Reliability of genomic predictions for North American Holstein bulls. *J. Dairy Sci.* 92:16–24.
  - 32- Weigel, K. A., C. P. Van Tassell, J. R. O’Connell, P. M. VanRaden, and G. R. Wiggans. 2010. Prediction of unobserved single nucleotide polymorphism genotypes of Jersey cattle using panels and population-based imputation algorithms. *J. Dairy Sci.* 93:2229– 2238.
  - 33- Zhang, Z., and T. Druet. 2010. Marker imputation with low-density marker panels in Dutch Holstein cattle. *J. Dairy Sci.* 93:5487–5494.