



## تولیدات دامی

دوره ۲۰ ■ شماره ۳ ■ پاییز ۱۳۹۷  
صفحه‌های ۳۷۵-۳۸۷

### بررسی اثر متقابل ژنوتیپ و محیط با جانمی داده ژنومی شبیه‌سازی شده با استفاده از مدل‌های حیوانی مختلف

یوسف نادری\*

استادیار، گروه علوم دامی، دانشگاه آزاد اسلامی، واحد آستارا، آستارا، ایران

تاریخ پذیرش مقاله: ۱۳۹۷/۰۵/۲۱

تاریخ وصول مقاله: ۱۳۹۷/۰۲/۲۳

#### چکیده

هدف از این تحقیق ارزیابی مدل‌های تک-صفتی و چند-صفتی در سناریوهای مختلف ژنومی با در نظر گرفتن جانمی جهت برآورد صحت پیش‌بینی ژنومی و تشخیص وجود اثر متقابل ژنوتیپ و محیط ( $G \times E$ ) بود. داده‌های ژنومی با تعداد متفاوت جایگاه‌های صفات کمی (۹۰ و ۹۰۰) و سطوح مختلف عدم تعادل پیوستگی (کم و زیاد = LD) برای تراکم ۵۰K شبیه‌سازی شدند. سپس به‌طور تصادفی ۹۰ درصد نشانگرها حذف و در مرحله بعد این نشانگرها از طریق نرم‌افزار Flmpure (نسخه ۲/۲) جانمی شدند. میانگین صحت جانمی در سناریوهای با LD زیاد و کم به ترتیب ۰/۹۷۶ و ۰/۹۴۳ بود. در همه سناریوهای شبیه‌سازی شده تفاوت جزئی بین صحت ژنومی داده‌های اصلی و جانمی مشاهده شد. صحت ژنومی با کاهش سطح LD، وراثت‌پذیری و همبستگی ژنتیکی بین صفات کاهش یافت. استفاده از مدل چند-صفتی نسبت به مدل تک-صفتی باعث افزایش صحت ژنومی شد. سطح LD و همبستگی ژنتیکی بین محیط‌های مختلف، در صورت وجود اثر متقابل محیط و ژنوتیپ نقش مهمی را ایفا کردند. لحاظ کردن اثر متقابل ژنوتیپ و محیط و تأثیر آن بر افزایش صحت ژنومی از یک طرف و جانمی تراشه‌های با تراکم کم به تراکم زیاد (خصوصاً در سناریوهای با LD بالا) جهت کاهش هزینه‌های ژنومی از طرف دیگر، می‌تواند راه حل مناسب و کاربردی جهت بهبود انتخاب ژنومی باشد.

**کلیدواژه‌ها:** صحت جانمی، صحت ژنومی، عدم تعادل پیوستگی، مدل حیوانی چند-صفتی، همبستگی ژنتیکی.

## مقدمه

شبیه‌سازی شده مورد بررسی قرار گرفته است که علاوه بر شبیه‌سازی معماری‌های ژنتیکی مختلف، عوامل دیگری نیز بر صحت پیش‌بینی تأثیر می‌گذارند که می‌توان به وراثت‌پذیری صفت، تعداد افراد و تعداد نسل‌های جمعیت مرجع و تأیید، نوع و تراکم نشانگرها، نوع صفت مورد مطالعه و روش مورد استفاده برای پیش‌بینی ارزش‌های اصلاحی ژنومی اشاره کرد [۲۰].

اگرچه صحت بالای ارزیابی ژنومی در برآورد ارزش‌های اصلاحی مهمترین عامل در پیش‌بینی ژنومی می‌باشد با این حال جنبه‌های اقتصادی این امر را نمی‌توان نادیده گرفت. در این راستا محققان سعی بر آن دارند تا با استفاده از راهکارهایی منطقی از جمله جانهی (Imputation) هزینه‌های گزاف ارزیابی ژنومی کوتاه‌مدت را کاهش دهند. جانهی تراشه‌های کم تراکم به تراکم بالا راهکار مساعدی در زمینه بهبود جنبه‌های اقتصادی ژنومیک بوده است. این تکنیک به محقق اجازه می‌دهد علاوه بر کاهش هزینه‌های توالی‌یابی، برآورد قابل قبولی از صحت پیش‌بینی ژنومی نیز حاصل شود [۴]. مطالعات در زمینه جانهی نشان داد که داده‌های ژنومی با صحت جانهی بالا می‌توانند سطوح مشابهی و معقولی از صحت پیش‌بینی ژنومی در قیاس با داده‌های اصلی ایجاد کرده‌اند [۱۰]. همچنین جانهی پنل‌های کم تراکم به تراکم بالا با صحت بالای جانهی علاوه بر مقرون به صرفه بودن، امکان انتخاب تعداد زیادی افراد جوان را برای توالی‌یابی ژنوم افزایش داده و در جهت پیشبرد انتخاب ژنومی یک گام رو به جلو محسوب می‌شود [۷].

در دهه اخیر، برخی از مطالعات روی اثر متقابل  $G \times E$  تمرکز کرده‌اند که صرفاً با اطلاعات فنوتیپی و شجره‌ای این اثر متقابل مورد بررسی قرار گرفته است [۱۳]. زمانی که عوامل محیطی بر بروز ژنتیکی صفات تأثیرگذار باشند، وارد کردن اطلاعات ژنومی دقت و صحت انتخاب

اصلاح‌گران باید ژنوتیپی را انتخاب کنند که در شرایط محیطی خاص نسبت به ژنوتیپ‌های دیگر عملکرد بهتری داشته باشد. به عبارت دیگر در برنامه‌های اصلاح‌نژادی، لازم است اثر متقابل ژنوتیپ و محیط ( $G \times E$ ) نیز مورد بررسی قرار گیرد. اثر متقابل  $G \times E$  زمانی اتفاق می‌افتد که ژنوتیپ‌های مختلف در محیط‌های متفاوت به‌طور یکسان تحت تأثیر قرار نمی‌گیرند. وجود اثر متقابل به تفاوت کنترل ژنتیکی صفات در محیط‌های مختلف ناشی از همبستگی ژنتیکی وابسته است [۹]. مطالعات مختلفی در زمینه بررسی اثر متقابل  $G \times E$  انجام گرفته است و اثر متقابل قابل توجهی در صنعت پرورش گاو شیری تشخیص داده شده است [۱۳]. همچنین گزارش شده است که یک ژن در محیط‌های مختلف می‌تواند اثرات متفاوتی از خود نشان دهد و به عبارت دیگر می‌تواند بروز ژنتیکی متفاوتی در محیط‌های مختلف داشته باشد [۲].

امروزه با پیشرفت فناوری‌های "اومیکس" علاقه به سمت تشریح معماری ژنتیکی صفات مهم تولیدی و اقتصادی بیشتر شده است. معماری ژنتیکی صفات از یک سری پارامترهایی تشکیل شده است که تنوع ژنتیکی صفت را در داخل و بین جمعیت‌ها توجیه می‌کنند. این پارامترها شامل تعداد QTL، مکان آنها روی ژنوم، توزیع فراوانی جایگاه‌های ژنی، توزیع اثرات جایگاه‌های ژنی، الگوی عدم تعادل پیوستگی (LD) در بین جایگاه‌ها، مقدار اثرات افزایشی، غالبیت و اپیستازی می‌باشند. شناخت معماری ژنتیکی کاربردهایی در زمینه‌های مختلف دارد. به عنوان مثال می‌توان به درک بهتری از فرآیندهای تکاملی حفظ تنوع و تفاوت‌های ژنتیکی بین جمعیتی رسید و نیز ژن‌هایی را شناسایی کرد که در افزایش تولید و نیز درمان بیماری‌های خاصی مفید باشند. در سال‌های اخیر معماری‌های ژنتیکی مختلف با استفاده از داده‌های

## تولیدات دامی

متقابل  $G \times E$  انجام گرفته است، لذا در این تحقیق، شبیه‌سازی ژنومی به‌صورتی انجام گرفته است که با تشکیل جمعیت‌های مرجع و تأیید با دارا بودن ارتباط ژنتیکی لازم در محیط‌های مختلف بتوان اثر متقابل  $G \times E$  و تفاوت بیان ژنتیکی در محیط‌های مختلف را منعکس گردد. در نتیجه برای پی بردن به صحت پیش‌بینی‌های ژنومی ساختارهای جمعیتی مختلف، سناریوهای مختلفی از معماری‌های ژنتیکی شبیه‌سازی شد.

### مواد و روش‌ها

جمعیت‌ها با استفاده از نرم‌افزار QMSim (نسخه ۱/۱) شبیه‌سازی شدند [۲۱]. ابتدا یک جمعیت اولیه ۱۵۴۰ رأسی طی ۱۰۰۰ نسل شبیه‌سازی شد. سپس این جمعیت پایه در دو مسیر جهت تولید LD پایین و بالا شبیه‌سازی و تکثیر یافت. در مسیر اول، جمعیت اولیه ۱۰۰ نسل دیگر (تا نسل ۱۱۰۰) تکثیر یافت و جمعیتی با LD پایین شبیه‌سازی شد. در مسیر دوم، تعداد افراد جمعیت اولیه از طریق ایجاد یک گلوگاه ژنتیکی به ۱۵۴ رأس در نسل ۱۰۵۰ کاهش یافت. سپس این تعداد افراد (۱۵۴ رأس) برای ۵۰ نسل دیگر (تا نسل ۱۱۰۰) تکثیر یافته و به تعداد اولیه خود یعنی ۱۵۴۰ رأس برگشت داده شدند. برای ایجاد جمعیت مرجع و تأیید، همه افراد (۱۵۴۰ رأس) آخرین نسل جمعیت پایه برای تولید مثل در جمعیت حاضر مورد استفاده قرار گرفتند که در این بین ۴۰ رأس نر در نظر گرفته شد تا منعکس‌کننده نسبت نر به ماده موجود در گله‌های گاو شیری باشد تا بتوان اثر تکنیک تلقیح مصنوعی بر نسبت نر به ماده را تقلید کرد.

نوع سیستم تلاقی تصادفی بود و برای ۱۰ نسل دیگر جمعیت تکثیر شد. شانس تلاقی در همه حیوانات برابر (در هر دو جنس) و یک فرزند برای هر زایش در نظر گرفته شد. درصد جایگزینی برای نر و ماده به ترتیب ۸۰ و

حیوانات را افزایش می‌دهد [۲]. در دسترس بودن اطلاعات ژنوتیپی به همراه اطلاعات فنوتیپی رکوردگیری شده در طول زمان، اجرای مدل‌های چند-صفتی مبتنی بر اطلاعات ژنومی را نیز میسر می‌سازد. با اجرای این چنین مدل‌هایی، ارزش‌های اصلاحی ژنومی حیوانات برای زمان‌ها و مکان‌های مختلف برآورد خواهد شد. این نوع ارزش‌های اصلاحی مخصوصاً برای صفاتی که به‌ندرت ثبت می‌شوند و یا برای توصیف‌گرهای محیطی که اطلاعاتی از آنها در فایل داده می‌باشد دارای اهمیت بیشتری خواهد بود [۱۵]. تحقیقات نشان دادند که گاوهای نر نژاد هلشتاین در کشورهای مختلف ظهور فنوتیپی متفاوتی دارند و تفاوت بیان ژن در کشورهای مذکور دیده شد که نشان می‌دهد محیط‌های مختلف تأثیرات مختلفی بر ژنوتیپ می‌گذارند و تفاوت رتبه‌بندی گاوهای نر در کشورهای تونس و لوکزامبورگ وجود دارد [۱۳]. همچنین گزارش شد که شاخص دما-رطوبت محاسبه شده برای مناطق مختلف ایران در ارتباط با مقدار تولید شیر است و دما و رطوبت محیط می‌تواند بر پارامترهای ژنتیکی تأثیرگذار باشد [۳].

با وجود مطالعات ژنومی گسترده با استفاده از داده‌های شبیه‌سازی شده با معماری‌های ژنتیکی مختلف و برآورد صحت پیش‌بینی‌های ژنومی آن‌ها با کمک روش‌های آماری متفاوت، مطالعات محدودی در بخش بررسی اثر متقابل  $G \times E$  انجام گرفته است. با این حال تحقیقاتی در سال‌های اخیر انجام شده که توانسته‌اند در جهت بررسی اثر متقابل بین  $G \times E$ ، داده‌های ژنومی مناسبی را شبیه‌سازی کنند [۱۷ و ۲۹]. وقتی QTLها با محیط اثر متقابلی داشته باشند، یک مدل با در نظر گرفتن  $G \times E$  می‌تواند توان تشخیص اثرات QTL را افزایش دهد و صحت پیش‌بینی ژنومی را افزایش دهد [۱۷]. با توجه به اینکه مطالعات ژنومی کمی در زمینه بررسی اثر

تعداد ۹۰ QTL - سطح بالا LD؛ سناریو ۲: تعداد ۹۰۰ QTL - سطح بالا LD؛ سناریو ۳: تعداد ۹۰ QTL - سطح پایین LD؛ سناریو ۴: تعداد ۹۰۰ QTL - سطح پایین LD در تحقیق حاضر شبیه سازی شد. نرخ جهش برای نشانگرها و QTLها در هر جایگاه و در هر نسل  $2/5 \times 10^{-5}$  فرض شد. فراوانی آللی اولیه برای نشانگرها ۰/۵ و توزیع اثرات QTLها، گاما فرض شد. در هر نسل و هر جایگاه کل میزان واریانس افزایشی توسط QTL توجیح شد (جدول ۱).

۲۰ درصد در نظر گرفته شد. انتخاب حیوانات برتر برای نسل بعد براساس ارزش اصلاحی صورت گرفت. نشانگرها به صورت دو آللی و به صورت فواصل یکسان در بین ۳۰ جفت کروموزوم به طول ۱۰۰ سانتی مورگان توزیع شدند. به ازای هر کروموزوم ۱۶۶۷ نشانگر برای پنل های ۵۰K شبیه سازی شد. دو سطح مختلف QTL (۹۰ و ۹۰۰) شبیه سازی شد که به صورت تصادفی در طول کروموزومها توزیع شدند. در مجموع چهار سناریو (سناریو ۱:

جدول ۱. ساختار شجره و معماری ژنومی شبیه سازی شده

عدم تعادل پیوستگی پایین	عدم تعادل پیوستگی بالا	ساختار جمعیت
		جمعیت اولیه
۱۰۰۰ (۱۵۴۰)	۱۰۰۰ (۱۵۴۰)	فاز اول تعداد نسل (تعداد افراد)
خیبر	بله	گلوگاه
۵۰ (۱۵۴۰)	۵۰ (۱۵۴)	فاز دوم تعداد نسل (تعداد افراد)
۱۰۰ (۱۵۴۰)	۱۰۰ (۱۵۴۰)	فاز سوم تعداد نسل (تعداد افراد)
۱۵۴۰		تعداد حیوانات در نسل آخر
		جمعیت اخیر
۴۰		تعداد نرهای در نسل اخیر
۱۵۰۰		تعداد ماده‌ها در نسل اخیر
۱۰		تعداد تکثیری جمعیت اخیر بعد از نسل ۱۱۰۰
۱		تعداد نتایج به ازای هر زایش
۰/۵		احتمال نر بودن نتاج
ارزش اصلاحی / تصادفی		انتخاب / طرح آمیزش
٪۸۰		نرخ جایگزینی برای نرها
٪۲۰		نرخ جایگزینی برای ماده‌ها
سن بالا حیوان و ارزش اصلاحی پایین		معیار حذف
		ژنوم
۳۰		تعداد کروموزوم
۱۰۰		طول هر کروموزوم (سانتی مورگان)
۳ یا ۳۰		تعداد QTL به ازای هر کروموزوم
گاما (۰/۴)		اثر آلل های QTL
۱۶۶۷		تعداد نشانگر به ازای هر کروموزوم
$2/5 \times 10^{-5}$		نرخ جهش در نشانگر و QTLها

محیط از طریق جمع اثرات QTL‌های هر حیوان در محیط مدنظر محاسبه شد و با اضافه کردن اثر باقیمانده نمونه‌گیری شده از توزیع نرمال به TBV، فنوتیپ حیوانات تولید شد. برای سه محیط، وراثت‌پذیری صفت مدنظر به ترتیب برابر با ۰/۱۰، ۰/۲ و ۰/۳۰ در نظر گرفته شد. ۷۵۰ رأس گاو نسل آخر به‌طور تصادفی به سه محیط مختلف تخصیص داده شدند که در واقع در هر محیط حدود ۲۵۰ حیوان رکورددار قرار گرفت. با توجه به این معیار تشکیل جمعیت، برای هر گاو نر به‌طور متوسط ۶/۲۵ نتاج در هر محیط تخصیص داده شد. با محاسبه همبستگی بین ارزش‌های اصلاحی واقعی حیوانات در محیط‌های مختلف، همبستگی ژنتیکی صفات در محیط‌های مختلف محاسبه شد و همچنین از میانگین رکورد به‌زای هر نر در محیط‌های مختلف، همبستگی فنوتیپی بین محیط‌های مختلف محاسبه شد (جدول ۲).

#### جدول ۲. وراثت‌پذیری (قطری) و همبستگی ژنتیکی (بالا

قطری) و فنوتیپی (پایین قطری) و انحراف معیار آنها

(داخل پرانتز) برای سه محیط

محیط	یک	دو	سه
یک (وراثت‌پذیری ۰/۱)	۰/۱	۰/۷۹ (۰/۰۴)	۰/۴۷ (۰/۱۱)
دو (وراثت‌پذیری ۰/۲)	۰/۳۶ (۰/۱۳)	۰/۲	۰/۵۹ (۰/۰۷)
سه (وراثت‌پذیری ۰/۳)	۰/۲۵ (۰/۱۷)	۰/۲۹ (۰/۱۵)	۰/۳

مدل حیوانی سه-صفتی برای آنالیز رکوردهای موجود در سه محیط با در نظر گرفتن این مورد که فنوتیپ‌های اندازه‌گیری شده در محیط‌های مختلف به‌عنوان صفات همبسته هستند [۷] مورد استفاده قرار گرفت. مدل ژنومی سه-صفتی (رابطه ۱) به‌صورت زیر در نظر گرفته شد [۱۵]:

$$\begin{bmatrix} y_1 \\ \vdots \\ y_3 \end{bmatrix} = \begin{bmatrix} I_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & I_3 \end{bmatrix} \begin{bmatrix} \mu_1 \\ \vdots \\ \mu_3 \end{bmatrix} + \begin{bmatrix} Z_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & Z_3 \end{bmatrix} \begin{bmatrix} g_1 \\ \vdots \\ g_3 \end{bmatrix} + \begin{bmatrix} e_1 \\ \vdots \\ e_3 \end{bmatrix} \quad (1)$$

کنترل کیفی با استفاده از برنامه PreGSf90 (نسخه ۲۰۰۲) انجام گرفت [۱]. نشانگرهایی که حداقل فراوانی آلی کمتر ۰/۰۱ داشتند، حذف گردید. بعد از شبیه‌سازی جمعیت‌های اصلی با تراکم ۵۰K، با استفاده از کپی‌برداری از نسخه اصلی داده‌ها، جهت ایجاد جمعیت جانهای با کمک برنامه‌نویسی در نرم‌افزار R و به‌طور تصادفی اقدام به حذف ۹۰ درصد نشانگرها آخرین نسل جمعیت پایه نموده و در مرحله بعدی از طریق برنامه Flmpute (نسخه ۲/۲) اقدام به جانهای و پیش‌بینی نقاط گم‌شده از طریق روابط فامیلی و الگوریتم‌های برپایه جمعیت شد [۲۲]. در نتیجه دو سری داده اصلی و جانهای شده برای ارزیابی مراحل تحقیق استفاده شد. صحت جانهای از طریق همبستگی داده‌های اصلی و جانهای برای نشانگرها برای سناریوهای مختلف شبیه‌سازی شده مورد ارزیابی قرار گرفت.

در راستای شبیه‌سازی ساختار داده‌های ایجادشده، با استفاده از کدنویسی در نرم‌افزار R، تغییرات در فایل‌های خروجی نرم‌افزار QMSim (نسخه ۱/۱) اعمال گردید. ژنوتیپ مربوط به دختران گاوهای قرار گرفته در نسل اخیر (آخرین نسل) برای آنالیز ژنومی مورد استفاده قرار گرفت و فایل شجره شامل کل حیوانات موجود در ۱۰ نسل اخیر بود. دختران گاوهای موجود در فایل ژنوتیپ، نتاج ۴۰ رأس گاو نر بودند. QTL‌ها (۹۰ یا ۹۰۰) به‌طور تصادفی در پنج گروه قرار گرفتند و در مرحله‌ی بعد، گروه‌های یک تا سه، دو تا چهار و سه تا پنج در محیط‌های یک، دو و سه قرار گرفتند تا صرفاً در این محیط‌ها بیان شوند. این کار برای تقلید بیان ژنی متفاوت در سه محیط مختلف انجام گرفت. به‌علت همپوشانی گروه‌های QTL‌ها در محیط‌های مختلف، همبستگی ژنتیکی بین محیط‌های مختلف ایجاد گردید.

ارزش‌های اصلاحی واقعی (TBV) حیوانات در هر

$$\mathbf{G} = 0.95 \times \mathbf{G}^* + 0.05 \times \mathbf{A}_{22} \quad (6)$$

نتایجی که از این طریق از مدل سه صفتی به دست آمد، با نتایج به دست آمده از مدل حیوانی تک صفتی مقایسه شد. در مدل حیوانی تک صفتی به این صورت فرض شد که همه رکوردهای موجود در محیطهای مختلف به عنوان یک صفت هستند. برای مدل تک صفتی ساختار مدل و فرضیات مشابهی با مدل سه صفتی در نظر گرفته شد با این تفاوت که یک اثر ثابت برای مشخص کردن محیطهای مختلف در مدل قرار گرفت. آنالیزهای ژنومی با برنامه AIREMLF90 (نسخه ۲۰۰۲) انجام گرفت که در این برنامه از الگوریتم اطلاعات متوسط (AI) در روش حداکثر درستنمایی محدود شده (REML) استفاده می شود. همبستگی بین TBV و GEBV برآورد شده برای داده های اصلی و جهانی برای جمعیت تأیید از طریق مدل های مختلف، به عنوان معیار ارزیابی صحت مورد استفاده قرار گرفت. ارزیابی ها با استفاده از ۱۰ تکرار شبیه سازی برای هر سناریو انجام گرفت و میانگین و انحراف معیار صحت پیش بینی های ژنومی برای هر یک از سناریوها گزارش شد.

### نتایج

جدول ۳ میانگین صحت جهانی برای هریک از سناریوهای شبیه سازی شده از طریق همبستگی داده های اصلی و جهانی (با نرخ حذف ۹۰ درصد) را نشان می دهد. با توجه به نتایج صحت جهانی برای سناریوهای LD بالا بیشتر از سناریوهای با LD پایین بود و میانگین صحت جهانی برای سناریوهای با LD بالا و پایین به ترتیب ۰/۹۷۶ و ۰/۹۴۳ بود. بیشترین و کمترین میزان صحت جهانی به ترتیب برای سناریوی با تعداد بالای QTL و LD در محیط سوم (با وراثت پذیری بالا) و سناریوی با سطح پایین QTL و LD در محیط اول (با

که در این مدل:  $y_i$ ، مشاهدات مربوط به صفت  $i$  برابر با یک تا سه؛  $\mathbf{I}$ ، ماتریس واحد؛  $\mu_i$ ، میانگین صفت  $i$ ؛  $\mathbf{Z}_i$ ، ماتریس طرح برای اثر  $g_i$ ؛  $g_i$ ، بردار ارزش های اصلاحی ژنومی برای حیوانات در محیط  $i$  (GEBV) و  $e_i$ ، نیز بردار اثرات باقیمانده تصادفی می باشد. در این مدل فرضیات زیر در نظر گرفته شد (رابطه ۲، ۳ و ۴):

$$g_i \sim N(0, \mathbf{H} \otimes \mathbf{T}), e_i \sim N(0, \mathbf{I} \otimes \mathbf{R}) \quad (2)$$

$$\mathbf{T} = \begin{bmatrix} \sigma_{g1}^2 & \sigma_{g12} & \sigma_{g13} \\ \sigma_{g21} & \sigma_{g2}^2 & \sigma_{g23} \\ \sigma_{g31} & \sigma_{g32} & \sigma_{g3}^2 \end{bmatrix} \quad (3)$$

$$\mathbf{R} = \begin{bmatrix} \sigma_{e1}^2 & 0 & 0 \\ 0 & \sigma_{e2}^2 & 0 \\ 0 & 0 & \sigma_{e3}^2 \end{bmatrix} \quad (4)$$

در مدل ژنومی تک مرحله ای بهترین پیش بینی کننده ناریب خطی، ماتریس  $\mathbf{H}$  یک ماتریسی متشکل از ماتریس روابط خویشاوندی شجره ای ( $\mathbf{A}$ ) و ماتریس روابط ژنومی ( $\mathbf{G}$ ) می باشد که شامل حیوانات ژنوتایپینگ نشده نیز است و  $\mathbf{T}$  و  $\mathbf{R}$  به ترتیب ماتریس های واریانس-کواریانس برای اثرات ژنتیک افزایشی و باقیمانده برای سه صفت مختلف هستند. معکوس ماتریس  $\mathbf{H}$  (رابطه ۵) و ماتریس  $\mathbf{G}$  (رابطه ۶) به صورت زیر تعریف شد:

$$\mathbf{H}^{-1} = \mathbf{A}^{-1} + \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{G}^{-1} - \mathbf{A}_{22}^{-1} \end{bmatrix} \mathbf{G}^* \quad (5)$$

$$= \frac{\mathbf{Z}\mathbf{Z}'}{2 \times \sum_{k=1}^m p_k(1-p_k)}$$

که در این جا،  $\mathbf{Z} = (\mathbf{M} - 2p_k)$  و ماتریس  $\mathbf{M}$  برابر با تعداد حیوانات در تعداد نشانگر ( $m$ ) است و  $p_k$  حداقل فراوانی آلی  $k$  امین SNP می باشد. با توجه به اینکه ابعاد ماتریس  $\mathbf{G}^*$  برابر با ابعاد ماتریس روابط خویشاوندی حیوانات ژنوتایپینگ شده ( $\mathbf{A}_{22}$ ) است. نسبتی از ماتریس  $\mathbf{A}_{22}$  در فرمول محاسباتی ماتریس روابط خویشاوندی قرار گرفت تا بتوان ماتریس معکوس تشکیل داد:

## تولیدات دامی

بررسی اثر متقابل ژنوتیپ و محیط با جانهای داده ژنومی شبیه‌سازی شده با استفاده از مدل‌های حیوانی مختلف

جدول ۴. صحت (انحراف معیار) پیش‌بینی‌های ژنومی برای حیوانات موجود در سه محیط مختلف با استفاده از مدل حیوانی تک-صفتی

محیط (وراثت‌پذیری)			سناریو نوع داده
۳ (۰/۳۰)	۲ (۰/۲۰)	۱ (۰/۱۰)	
۰/۵۳۶(۰/۰۳)	۰/۵۳۹(۰/۰۲)	۰/۵۱۴(۰/۰۶)	۱ جانهای
۰/۵۸۰(۰/۰۲)	۰/۵۶۰(۰/۰۴)	۰/۵۳۸(۰/۰۳)	۱ اصلی
۰/۵۰۸(۰/۰۵)	۰/۵۰۴(۰/۰۵)	۰/۴۸۷(۰/۰۶)	۲ جانهای
۰/۵۴۸(۰/۰۷)	۰/۵۴۶(۰/۰۵)	۰/۵۱۲(۰/۰۵)	۲ اصلی
۰/۴۳۹(۰/۰۶)	۰/۴۳۱(۰/۰۳)	۰/۴۱۱(۰/۰۴)	۳ جانهای
۰/۴۷۶(۰/۰۴)	۰/۴۷۹(۰/۰۴۸)	۰/۴۵۶(۰/۰۲۹)	۳ اصلی
۰/۴۴۶(۰/۰۶)	۰/۴۲۹(۰/۰۳)	۰/۳۸۹(۰/۰۴)	۴ جانهای
۰/۵۱۶(۰/۰۶)	۰/۵۱۱(۰/۰۳)	۰/۵۰۱(۰/۰۵)	۴ اصلی

سناریو ۱: تعداد QTL ۹۰- سطح بالا LD؛ سناریو ۲: تعداد QTL ۹۰۰- سطح بالا LD؛ سناریو ۳: تعداد QTL ۹۰- سطح پایین LD؛ سناریو ۴: تعداد QTL ۹۰۰- سطح پایین LD.

زمانی‌که رکورد حیوانات در محیط‌های مختلف به‌عنوان صفات مختلف در نظر گرفته شد، مدل حیوانی ژنومی سه-صفتی مورد استفاده قرار گرفت و نتایج آن در جدول ۵ ارائه شده است. در مقایسه با مدل حیوانی تک-صفتی، صحت ارزش‌های اصلاحی ژنومی برآوردشده توسط مدل سه-صفتی برای همه سناریوها در هر سه محیط افزایش داشت. به‌طورکلی، صحت پیش‌بینی با در نظر گرفتن LD پایین در ژنوم، کاهش محسوسی نشان داد. همانند آنالیز تک-صفتی در مدل حیوانی با اثر متقابل، حساسیت صحت ژنومی سناریوهای با LD پایین نسبت به LD بالا به تغییرات تراکم نشانگری بیشتر بود. همچنین صفات با وراثت‌پذیری بالا، تعداد پایین QTL برای داده‌های اصلی دارای صحت پیش‌بینی بالایی بودند. بنابراین، برای همه سناریوهای مربوط به محیط سوم که دارای بیشترین وراثت‌پذیری (۰/۳۰) بود، بالاترین میزان صحت پیش‌بینی GEBV به‌دست آمد.

وراثت‌پذیری پایین) مشاهده شد. به‌طورکلی دامنه تغییرات میانگین صحت جانهای در بازه بین ۰/۹۳۹ تا ۰/۹۷۹ بود.

جدول ۳. صحت جانهای (انحراف معیار) در سناریوهای مختلف با ۹۰ درصد حذف نشانگرها برای حیوانات موجود در سه محیط مختلف

سناریو	صحت جانهای		
	$h^2=0/3$	$h^2=0/2$	$h^2=0/1$
۱	۰/۹۷۴(۰/۰۱۲)	۰/۹۷۸(۰/۰۱۱)	۰/۹۷۵(۰/۰۱۲)
۲	۰/۹۷۹(۰/۰۱۰)	۰/۹۷۲(۰/۰۱۲)	۰/۹۷۶(۰/۰۱۰)
۳	۰/۹۴۲(۰/۰۱۲)	۰/۹۴۳(۰/۰۱۲)	۰/۹۳۹(۰/۰۱۳)
۴	۰/۹۴۵(۰/۰۱۲)	۰/۹۴۲(۰/۰۱۱)	۰/۹۴۶(۰/۰۱۳)

سناریو ۱: تعداد QTL ۹۰- سطح بالا LD؛ سناریو ۲: تعداد QTL ۹۰۰- سطح بالا LD؛ سناریو ۳: تعداد QTL ۹۰- سطح پایین LD؛ سناریو ۴: تعداد QTL ۹۰۰- سطح پایین LD.

با در نظر گرفتن صفت در محیط‌های مختلف به‌عنوان یک صفت، ارزش‌های اصلاحی ژنومی برای حیوانات دارای رکورد‌های فنوتیپی و ژنوتیپی با استفاده از مدل تک-صفتی برای سناریوهای مختلف پیش‌بینی شد که صحت پیش‌بینی ژنومی برای سناریوهای مختلف در جدول ۴ گزارش شده است. به‌طورکلی، صحت‌های ارزش‌های اصلاحی ژنومی با افزایش وراثت‌پذیری و LD بیشتر شد. از طرفی، صحت ارزیابی GEBV برای سناریوهای اصلی بیشتر از صحت ارزیابی سناریوهای جانهای بود. به‌طورکلی حساسیت صحت ژنومی سناریوهای با LD پایین نسبت به LD بالا به تغییرات تراکم نشانگری بیشتر بود. افزایش تعداد QTL از ۹۰ به ۹۰۰، اثر بسیار جزئی بر صحت ارزیابی در برخی سناریوهای نشان داد. با در نظر گرفتن محیط‌های مختلف، سناریوی دارای LD پایین و تراکم QTL پایین با جانهای نشانگرها کم‌ترین صحت پیش‌بینی را داشت.

جدول ۵. صحت (انحراف معیار) پیش‌بینی‌های ژنومی برای حیوانات موجود در سه محیط مختلف با استفاده از مدل

حیوانی سه- صفتی

سناریو	محیط (وراثت‌پذیری)		
	۱ (۰/۱۰)	۲ (۰/۲)	۳ (۰/۳۰)
۱ جانهای	۰/۵۳۶(۰/۰۴)	۰/۶۲۱(۰/۰۴)	۰/۶۴۸(۰/۰۲)
۱ اصلی	۰/۵۶۲(۰/۰۴)	۰/۶۳۳(۰/۰۳)	۰/۶۷۱(۰/۰۳)
۲ جانهای	۰/۵۱۱(۰/۰۳)	۰/۵۸۸(۰/۰۴)	۰/۵۳۴(۰/۰۴)
۲ اصلی	۰/۵۲۹(۰/۰۴)	۰/۶۰۳(۰/۰۳)	۰/۶۷۷(۰/۰۵)
۳ جانهای	۰/۴۶۱(۰/۰۳)	۰/۵۴۲(۰/۰۴)	۰/۵۴۵(۰/۰۳)
۳ اصلی	۰/۵۰۸(۰/۰۳)	۰/۵۷۶(۰/۰۲)	۰/۶۰۳(۰/۰۲)
۴ جانهای	۰/۴۵۲(۰/۰۴)	۰/۵۱۵(۰/۰۶)	۰/۵۴۹(۰/۰۴)
۴ اصلی	۰/۵۱۴(۰/۰۳)	۰/۵۷۱(۰/۰۴)	۰/۶۱۴(۰/۰۴)

سناریو ۱: تعداد QTL ۹۰- سطح بالا LD؛ سناریو ۲: تعداد QTL ۹۰۰- سطح بالا LD؛ سناریو ۳: تعداد QTL ۹۰- سطح پایین LD؛ سناریو ۴: تعداد QTL ۹۰۰- سطح پایین LD

بحث

در این تحقیق، جانهای با صحت بالا برای سناریوهای با LD بالا (دامنه ۰/۹۷۲ تا ۰/۹۷۹) و سناریوهای با LD پایین (دامنه ۰/۹۳۹ تا ۰/۹۴۶) مشاهده شد. با توجه به این که نوع معماری ژنومی و اندازه مؤثر جمعیت از فاکتورهای مؤثر بر صحت جانهای است [۲۶]. مطالعات نشان دادند که صحت پایین جانهای در برخی از مناطق ژنوم (کمتر از ۰/۶) به ماهیت ژنوم و سطح بسیار پایین LD در این مناطق بستگی داشت [۲۵]. مطالعات اخیر نشان داد که تفاوت در معماری ژنومی و اندازه مؤثر جمعیت از عوامل مؤثر بر صحت جانهای در تراشه ۵K به ۵۰K است [۲۶]. سایر مطالعات [۲۴] گزارش کردند که جانهای ۳K به تراشه‌های با تراکم بسیار بالا می‌تواند منجر به نتایج قابل قبولی از صحت شود. گزارشات در مورد گاوهای هلشتاین آلمان [۱۹] نشان دادند که میزان LD از مهمترین فاکتورهای مؤثر بر صحت جانهای می‌باشند.

در این تحقیق، تعداد متفاوتی از نشانگرها در سناریوهای مختلف استفاده شدند با این حال تفاوت محسوسی بین صحت پیش‌بینی ژنومی داده‌های ۵۰k و جانهای خصوصاً هنگامی که میزان صحت جانهای بالا بود (در سناریوهای با LD بالا) مشاهده نشد. مطالعات [۱۹] در ارزیابی مدل‌های آماری با استفاده از داده‌های با و بدون جانهای به این نتیجه رسیده‌اند که توانایی پیش‌بینی ژنومی هنگام استفاده از جانهای ژنوتیپ بهبود بخشیده و این استراتژی را جهت کاهش هزینه‌ها در برنامه‌های انتخاب ژنومی توصیه نموده‌اند. نتایج در مورد گاوهای جرسی نشان داد که جانهای یک تراشه با تراکم پایین به تراشه ۵۰k هنگامی که صحت جانهای بالا باشد منجر به بهبود صحت ژنومی خواهد شد. در نتیجه توجه ویژه به صحت جانهای در مطالعات ژنومی بسیار حائز اهمیت خواهد بود [۲۷]. در مطالعه حاضر با توجه به صحت بالای جانهای در سناریوهای با LD بالا و به طبع آن عدم اختلاف محسوس در داده‌های جانهای و اصلی، استفاده از جانهای امری اقتصادی و مقرون به صرفه بود.

مطالعات روی جمعیت موش با نرخ جانهای ۷۵ و ۵۰ درصد به ترتیب صحت جانهای ۰/۹۴ و ۰/۹۸ را به همراه داشت، و استفاده از داده‌های جانهای نسبت به داده‌های اصلی، صرفه اقتصادی را به همراه داشت با این حال همیشه جانهای منجر به بهبود پیش‌بینی ژنومی نمی‌شود و شایستگی جانهای ژنوتیپ را وابسته به درجه ارتباط بین جمعیت تست و مرجع، تعداد نشانگرها، معماری ژنومی صفت و نوع روش آماری در پیش‌بینی اثرات نشانگرها دانستند [۱۰]. همچنین کاهش صحت پیش‌بینی ارزش‌های اصلاحی ژنومی داده‌های جانهای (با نرخ حذف ۸۸ درصد) نسبت به داده‌های واقعی با استفاده از مدل بیز حدود ۰/۰۶ گزارش شد [۷]. تحقیقات دیگر [۱۸] نشان داد که افزایش تعداد نشانگر می‌تواند اثر مبهمی بر صحت

تولیدات دامی



نظر به عنوان یک فاکتور مهم در معماری ژنتیکی صفت، می‌تواند بر صحت پیش‌بینی ژنومی اثرگذار باشد. برای صفات تولیدی با وراثت‌پذیری متوسط (۰/۳)، صحت پیش‌بینی بالاتری نسبت به صفات با وراثت‌پذیری پایین (۰/۰۵) گزارش شد و به ترتیب برابر با ۰/۴۱ و ۰/۶۷ بود که نشان‌دهنده وجود رابطه مستقیم و قوی بین صحت پیش‌بینی و وراثت‌پذیری صفت است [۵]. در مطالعه حاضر، در بین همه سناریوهای اجرا شده، صحت پیش‌بینی ژنومی به دست‌آمده برای محیط‌های سوم به دلیل وراثت‌پذیری بالا نسبت به دیگر محیط‌ها بیشتر بود. بالا بودن وراثت‌پذیری یعنی بیشتر بودن نسبت واریانس ژنتیکی به واریانس فنوتیپی و به معنای بیشتر بودن نقش ژن‌های با بیان افزایشی در ایجاد پراکنش در صفت است که باعث برآورد صحیح‌تر تأثیرات نشانگرها می‌شود. در نتیجه هرچه وراثت‌پذیری صفت بیشتر باشد، فنوتیپ فرد به ارزش ژنتیکی فرد نزدیک‌تر بوده و در نتیجه اثر نشانگرها و به دنبال آن ارزش‌های اصلاحی ژنومی افراد به طور صحیح‌تر پیش‌بینی می‌شود [۱۱].

در این تحقیق افزایش LD از طریق تأثیر مثبت در برآورد صحت جانمایی و در حله دوم از طریق تأثیر در برآورد ارزش‌های اصلاحی ژنومی بر صحت ژنومی تأثیر گذار بود. بررسی‌ها نشان می‌دهد که وجود LD قوی بین نشانگرهای مجاور در بررسی مطالعات ژنوم انسان نقش اساسی ایفا می‌کند [۱۶]. مطالعات اخیر وجود LD در بین نشانگرها را عاملی تأثیرگذار در بهبود صحت پیش‌بینی ارزش‌های اصلاحی ژنومی عنوان کردند. آنها نشان دادند که وجود LD بین نشانگر و QTL منبع اصلی اطلاعات بوده و نقش عمده‌ای در صحت پیش‌بینی ارزش‌های اصلاحی ژنومی ایفا می‌کنند [۲۳].

صحت پیش‌بینی‌های بالاتر مدل‌های حیوانی سه-صفتی نسبت به تک-صفتی نشان می‌دهد که لحاظ کردن

پیش‌بینی ژنومی داشته باشد. زیرا تعداد بیشتر نشانگر برای تعداد خاصی از فنوتیپ، می‌تواند اثر هم‌خطی و نتایج متفاوت به همراه داشته باشد. با این حال، زمانی که تراکم نشانگر دو برابر می‌شود صحت پیش‌بینی ژنومی بیشتر می‌شود، به طوری که تراکم نشانگری بالا باعث افزایش LD بین نشانگر و QTL می‌شود و در نتیجه تعداد نشانگر بیشتر می‌تواند نسبت بالاتری از واریانس ژنتیکی صفت مدنظر را توجیه کند [۱۱]. به طور کلی عدم اختلاف محسوس میانگین صحت‌های پیش‌بینی ژنومی برای سناریوهای ۵۰K اصلی نسبت به سناریوهای جانمایی، در راستای اهداف اقتصادی تحقیقات اخیر بود.

مطالعات زیادی در مورد تأثیر معماری ژنتیکی بر صحت پیش‌بینی‌های ژنومی انجام شد و به این نتیجه رسیده‌اند که روش GBLUP برای تعداد QTL‌های مختلف صحت پیش‌بینی‌های یکسانی را نتیجه می‌دهد. اما زمانی که از روش بیزی استفاده شود بیشترین صحت ارزیابی به دلیل توزیع محدود واریانس ژنتیکی بر تعداد زیادی QTL برای حالتی بود که QTL‌های کمتری صفت مورد علاقه را کنترل می‌کنند و با افزایش تعداد QTL صحت پیش‌بینی ژنومی کاهش یافت [۲۸]. همچنین برخی مطالعات گزارش کردند که در صورت استفاده از روش GBLUP، تأثیر معنی‌داری بین صحت پیش‌بینی ژنومی حاصل از تعداد QTL‌های متفاوت مشاهده نمی‌شود. به عنوان مثال، زمانی که تعداد QTL برابر با ۱۰۰ و ۱۰۰۰ عدد بود، صحت پیش‌بینی ژنومی به ترتیب برابر با ۰/۳۷ و ۰/۳۸ برآورد شد [۸]. با توجه به نتایج به دست‌آمده در مطالعات حاضر مبتنی بر روش REML، صحت پیش‌بینی‌های ژنومی یکسانی بین سناریوهای با تعداد QTL متفاوت (۹۰ و ۹۰۰) برای هر دو سری داده اصلی و جانمایی به دست آمد.

مطالعات نشان داده‌اند که وراثت‌پذیری صفت مورد

مقابلی با محیط داشته باشند، یک مدل با لحاظ کردن اثر مقابل  $G \times E$  می‌تواند توان تشخیص اثرات QTL را افزایش دهد و مدل چند-صفتی می‌تواند عملکرد بهتری داشته باشد [۱۷].

همبستگی ژنتیکی به‌عنوان معیاری مناسب برای بهبود صحت ارزیابی ژنومی با مدل چند-صفتی [۱ و ۶] و مدل‌های چند-ژنومی [۲۸] مورد استفاده قرار گرفته است. آنها نشان داده‌اند که صحت ارزیابی ژنومی ممکن است توسط عواملی مثل روابط خویشاوندی و سطح LD تحت تأثیر قرار گیرد. برای جمعیت مرجع شامل حیواناتی از محیط‌ها و کشورهای مختلف، مدل ژنومی چند-صفتی برای بررسی اثر مقابل  $G \times E$  مورد استفاده قرار گرفت تا از این طریق بتوان صحت پیش‌بینی ژنومی را افزایش داد [۱۵]. برای صفت تولید شیر در گاوهای شیری، همبستگی ژنتیکی برآورد شده در استرالیا و اکثر کشورهای نیمکره شمالی حدود ۰/۸۰ گزارش شده است [۱۲] و به دلیل وجود اثر مقابل بین کشور استرالیا و سایر کشورها، گزارش کرده‌اند که بهتر است رکوردهای فنوتیپی و ژنوتیپی نتاج گاوهای نر از سایر کشورها نیز در جمعیت مرجع قرار گیرند تا ارباب احتمالی از طریق آنالیز چند-صفتی رفع گردد.

در این تحقیق، اثرات QTL‌های شبیه‌سازی شده در محیط‌هایی برابر صفر بود اما در محیط‌های دیگر غیر صفر بود. به‌عبارت دیگر، اثر مقابل  $G \times E$  می‌تواند در سطح آلی (بیان متفاوت ژن در محیط‌های مختلف) وجود داشته و مورد بررسی قرار گیرد [۱۷] و زمانی که QTL با محیط دارای اثر مقابل است، مدل شامل اثر مقابل  $G \times E$  در مقایسه با مدل بدون اثر مقابل  $G \times E$  دارای قدرت بالاتری جهت تشخیص QTL خواهد بود. بعضی از QTL‌ها ممکن است در محیطی تشخیص داده شوند اما در محیط دیگر شناسایی نشوند، یا اگر شناسایی شوند

اثر مقابل  $G \times E$  در مدل حیوانی چند-صفتی نسبت به مدل حیوانی تک-صفتی، صحت ارزیابی را بهبود می‌بخشد. مدل حیوانی چند-صفتی برای پیش‌بینی صحت ژنومی در مطالعات اخیر مورد استفاده قرار گرفته است. نتایج این تحقیقات نشان داد که هنگام استفاده از مدل حیوانی چند-صفتی اطلاعات بیشتری به‌طور هم‌زمان نسبت به مدل‌های تک-صفتی در نظر گرفته می‌شود، در نتیجه باعث افزایش صحت پیش‌بینی ژنومی می‌شود. همانند فرآیند ارزیابی‌های ژنتیکی کلاسیک (بدون استفاده از داده‌های ژنومی)، مدل چند-صفتی می‌تواند صحت پیش‌بینی ژنومی را از طریق استفاده هم‌زمان از اطلاعات صفات همبسته افزایش دهد [۶]. در مطالعه دیگری با لحاظ کردن یک توصیف‌گر محیطی به‌عنوان عامل تأثیرگذار بر فنوتیپ، صفات همبسته به‌عنوان صفات تکرار شده در زمان در نظر گرفته شدند و به این نتیجه رسیده‌اند که وقتی از اطلاعات رکوردها در تمام زمان‌ها (مکان‌ها) استفاده شود، صحت پیش‌بینی بیشتری حاصل می‌شود [۲۹].

مطالعات در مورد اثر وراثت‌پذیری در مدل‌های حیوانی مختلف نشان داد که برای صفات با وراثت‌پذیری پایین (۰/۱۰)، در زمان استفاده از مدل حیوانی چند-صفتی نسبت به مدل حیوانی تک-صفتی، صحت پیش‌بینی حدود ۲۰ درصد افزایش پیدا می‌کند [۱۴]. در نتیجه، مدل چند-صفتی می‌تواند در تشخیص اثر مقابل  $G \times E$  برای صفات تولیدی در محیط‌های مختلف نیز کاربرد داشته باشد [۱۵]. استفاده از اطلاعات فنوتیپی موجود در محیط‌های مختلف به‌عنوان صفات متفاوت همبسته در نظر گرفته می‌شود که یکی از امتیازات و کاربردهای مهم مدل حیوانی چند-صفتی محسوب می‌شود و به این ترتیب، مدل چند-صفتی می‌تواند صحت پیش‌بینی ژنومی را افزایش دهد [۶]. وقتی QTL‌ها اثر

## تولیدات دامی

- and their effect on genomic evaluation. *Genetics Selection Evolution* 43: 18.
9. Falconer D and Mackay T (1996) Introduction to quantitative genetics. Longman, Harlow, UK. Introduction to quantitative genetics 4<sup>th</sup> ed Longman, Harlow, UK. pp. 183-184.
  10. Felipe VP, Okut H, Gianola D, Silva MA and Rosa GJ (2014) Effect of genotype imputation on genome-enabled prediction of complex traits: an empirical study with mice data. *BMC genetics* 15: 149.
  11. Goddard ME and Hayes BJ (2009) Mapping genes for complex traits in domestic animals and their use in breeding programmes. *Nature Reviews Genetics* 10: 381-391.
  12. Haile-Mariam M, Pryce J, Schrooten C and Hayes B (2015) Including overseas performance information in genomic evaluations of Australian dairy cattle. *Journal of dairy science* 98: 3443-3459.
  13. Hammami H, Rekik B, Bastin C, Soyeurt H, Bormann J, Stoll J and Gengler N (2009) Environmental sensitivity for milk yield in Luxembourg and Tunisian Holsteins by herd management level. *Journal of dairy science* 92: 4604-4612.
  14. Hayashi T and Iwata H (2013) A Bayesian method and its variational approximation for prediction of genomic breeding values in multiple traits. *BMC bioinformatics* 14: 34.
  15. Hayes BJ, Daetwyler HD and Goddard ME (2016) Models for genome  $\times$  environment interaction: Examples in livestock. *Crop Science* 56: 2251-2259.
  16. Ke X, Hunt S, Tapper W, Lawrence R, Stavrides G, Ghorji J, Whittaker P, Collins A, Morris AP and Bentley D (2004) The impact of SNP density on fine-scale patterns of linkage disequilibrium. *Human Molecular Genetics* 13: 577-588.
  17. Lillehammer M, Ødegård J and Meuwissen TH (2007) Random regression models for detection of gene by environment interaction. *Genetics Selection Evolution* 39: 105.
  18. Muir W (2007) Comparison of genomic and traditional BLUP-estimated breeding value accuracy and selection response under alternative trait and genomic parameters. *Journal of Animal Breeding and Genetics* 124: 342-355.
  19. Mulder H, Calus M, Druet T and Schrooten C (2012) Imputation of genotypes with low-density chips and its effect on reliability of direct genomic values in Dutch Holstein cattle. *Journal of dairy science* 95: 876-889.

دارای اثری متفاوت خواهند بود. بنابراین، مدل شامل اثر متقابل  $G \times E$  می‌تواند بیان ژنی متفاوت و تنوع ژنتیکی موجود در محیط‌های مختلف را به نحو بهتری تشخیص دهد.

نتایج تحقیق حاضر نشان داد که لحاظ کردن اثر متقابل  $G \times E$  و تأثیر آن بر افزایش صحت ژنومی و انتخاب ژنوتیپ‌های با شایستگی بالا برای محیط‌های مختلف، از یک طرف و جانمی تراشه‌های با تراکم کم به تراکم زیاد (خصوصاً در سناریوهای با LD بالا) جهت کاهش هزینه‌های ژنومی از طرف دیگر، می‌تواند راه حل مناسب و کاربردی جهت بهبود انتخاب ژنومی باشد.

#### منابع

1. Aguilar I, Misztal I, Legarra A and Tsuruta S (2011) Efficient computation of the genomic relationship matrix and other matrices used in single-step evaluation. *Journal of Animal Breeding and Genetics* 128: 422-428.
2. Bohlouli M, Alijani S, Javaremi AN, König S and Yin T (2017) Genomic prediction by considering genotype  $\times$  environment interaction using different genomic architectures. *Annals of Animal Science* 17: 683-701.
3. Bohlouli M, Shodja J, Alijani S and Eghbal A (2013) The relationship between temperature-humidity index and test-day milk yield of Iranian Holstein dairy cattle using random regression model. *Livestock Science* 157: 414-420.
4. Browning SR (2008) Missing data imputation and haplotype phase inference for genome-wide association studies. *Human genetics* 124: 439-450.
5. Calus M, De Haas Y, Pszczola M and Veerkamp R (2013) Predicted accuracy of and response to genomic selection for new traits in dairy cattle. *Animal* 7: 183-191.
6. Calus MP and Veerkamp RF (2011) Accuracy of multi-trait genomic selection using different methods. *Genetics Selection Evolution* 43: 26.
7. Chen L, Li C, Sargolzaei M and Schenkel F (2014) Impact of genotype imputation on the performance of GBLUP and Bayesian methods for genomic prediction. *PLoS One* 9: e101544.
8. Clark SA, Hickey JM and Van der Werf JH (2011) Different models of genetic variation

20. Pimentel EC, Wensch-Dorendorf M, König S and Swalve HH (2013) Enlarging a training set for genomic selection by imputation of un-genotyped animals in populations of varying genetic architecture. *Genetics Selection Evolution* 45: 12.
21. Sargolzaei M and Schenkel FS (2009) QMSim: a large-scale genome simulator for livestock. *Bioinformatics* 25: 680-681.
22. Sargolzaei M, Chesnais J and Schenkel F (2011) FImpute-An efficient imputation algorithm for dairy cattle populations. *J Dairy Sci* 94: 421.
23. Sun X, Fernando R and Dekkers J (2016) Contributions of linkage disequilibrium and co-segregation information to the accuracy of genomic prediction. *Genetics Selection Evolution* 48: 77.
24. Toghiani S, Aggrey S and Rekaya R (2016) Multi-generational imputation of single nucleotide polymorphism marker genotypes and accuracy of genomic selection. *animal* 10: 1077-1085.
25. VanRaden P, Null D, Sargolzaei M, Wiggans G, Tooker M, Cole J, Sonstegard T, Connor E, Winters M and van Kaam J (2013) Genomic imputation and evaluation using high-density Holstein genotypes. *Journal of dairy science* 96: 668-678.
26. Ventura RV, Miller SP, Dodds KG, Auvray B, Lee M, Bixley M, Clarke SM and McEwan JC (2016) Assessing accuracy of imputation using different SNP panel densities in a multi-breed sheep population. *Genetics Selection Evolution* 48: 71.
27. Weigel K, de Los Campos G, Vazquez A, Rosa G, Gianola D and Van Tassell C (2010) Accuracy of direct genomic values derived from imputed single nucleotide polymorphism genotypes in Jersey cattle. *Journal of dairy science* 93: 5423-5435.
28. Wientjes YC, Veerkamp RF, Bijma P, Bovenhuis H, Schrooten C and Calus MP (2015) Empirical and deterministic accuracies of across-population genomic prediction. *Genetics Selection Evolution* 47: 5.
29. Yin T, Pimentel E, Borstel UKv and 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.



## Animal Production

(College of Abouraihan – University of Tehran)

Vol. 20 ■ No. 3 ■ Autumn 2018

### Investigation of genotype × environment interaction with considering imputation in simulated genomic data via different animal models

Yousef Naderi\*

Assistant professor, Department of Animal Science, Islamic Azad University, Astara Branch, Astara, Iran.

Received: June 13, 2018

Accepted: August 12, 2018

#### Abstract

The objective of this study was to evaluate single-trait and multiple-trait animal models with considering imputation in simulated genomic data to estimate the accuracies of genomic prediction across various genomic scenarios and to detect genotype × environment ( $G \times E$ ) interaction. Genomic data were simulated to reflect variations in number of QTL (90 and 900) and linkage disequilibrium (LD = low and high) using 50K SNP panel. Afterwards, 90 percent of the markers randomly removed and imputation was performed using FImpute software (version 2.2). The average accuracy of imputation for scenarios with high and low LD was 0.976 and 0.943, respectively. In all scenarios, negligible difference on the genomic accuracies was evident, when original genotypes and imputed genotypes were compared. The genomic accuracy reduced with decreasing the LD, heritability and the genetic correlation among the traits. Comparing to single-trait animal model, using multiple-trait animal model increased genomic accuracy. The level of LD and genetic correlation across environments play important roles providing genotype × environment interaction exists. On the one hand, considering genotype × environment interaction and its effect on increasing of genomic accuracy and imputation of low to high density marker panels (especially high LD scenarios) to reduce of the cost of genomic evaluation on the other hand could be a suitable and practical approach to improve genomic selection application.

**Keywords:** Genomic accuracy, Genomic correlation, Imputation accuracy, Linkage disequilibrium, Multiple-trait animal model.