



## توليدات دامی

دوره ۲۱ ■ شماره ۴ ■ زمستان ۱۳۹۸

صفحه‌های ۴۳۰-۴۱۹

### ارزیابی صحت ارزش‌های اصلاحی ژنومی برخی صفات اقتصادی جمعیت گاوهای هلستاین ایران با استفاده از SNPها و بلوک‌های هاپلوتایپی

بهزاد رجبی مراند<sup>۱</sup>، حسین مرادی شهربابک<sup>۲\*</sup>، مصطفی صادقی<sup>۳</sup>، رستم عبدالهی آرپناهی<sup>۴</sup>  
۱. دانشجوی دکتری، گروه علوم دامی، پردیس کشاورزی و منابع طبیعی، دانشگاه تهران، کرج، ایران.  
۲. استادیار، گروه علوم دامی، پردیس کشاورزی و منابع طبیعی، دانشگاه تهران، کرج، ایران.  
۳. دانشیار، گروه علوم دامی، پردیس کشاورزی و منابع طبیعی، دانشگاه تهران، کرج، ایران.  
۴. استادیار، گروه علوم دامی، پردیس ابوریحان، دانشگاه تهران، پاکدشت، ایران.  
تاریخ دریافت مقاله: ۱۳۹۸/۰۳/۲۸ تاریخ پذیرش مقاله: ۱۳۹۸/۰۵/۲۸

#### چکیده

هدف پژوهش حاضر، بررسی صحت ارزش‌های اصلاحی ژنومی (GEBV) به‌دست‌آمده برای دو صفت اقتصادی مهم تولید شیر و امتیاز سلول‌های بدنی با استفاده از SNPها و بلوک‌های هاپلوتایپی مبتنی بر LD به‌کمک دو روش آماری GBULP و بیز B بود. اطلاعات ژنوتیپی مربوط به ۱۶۵۴ راس گاو نر که با استفاده از تراشه‌هایی با تراکم‌های مختلف تعیین ژنوتیپ شده بودند مورد استفاده قرار گرفت. صحت ارزش‌های اصلاحی به‌دست‌آمده هنگام استفاده از SNPها حاکی از برتری بیز B نسبت به GBULP بود، به‌طوری‌که برای صفت تولید شیر و امتیاز سلول‌های بدنی صحت ارزیابی‌ها با استفاده از GBULP به‌ترتیب برابر ۵۴/۳ و ۴۳/۷ درصد و برای روش بیز B به‌ترتیب برابر ۵۷/۷ و ۴۴/۲ درصد بود. برای صفت تولید شیر، صحت پیش‌بینی‌های حاصل با استفاده از بلوک‌های هاپلوتایپی در هر دو روش آماری، بالاتر از صحت حاصل هنگام استفاده از SNPها بود درحالی‌که برای صفت امتیاز سلول‌های بدنی، این افزایش صحت، زمانی‌که از روش آماری GBULP استفاده شد مشهود بود ولی هنگام استفاده از روش بیز B این برتری تنها زمانی به‌دست آمد که مقدار آماره  $r^2$  مورد استفاده برای تشکیل بلوک‌ها بالاتر از ۰/۲ بود. براساس نتایج حاصل، سطح بهینه آماره  $r^2$  برای تشکیل بلوک‌های هاپلوتایپی بستگی به نوع صفت و وراثت‌پذیری آن دارد ولی در مجموع، استفاده از آماره  $r^2$  بیش‌تر از ۰/۲ جهت تشکیل بلوک‌های هاپلوتایپی، می‌تواند منجر به افزایش صحت ارزش‌های اصلاحی ژنومی در هر دو صفت در مقایسه با استفاده از SNPها شود.

**کلیدواژه‌ها:** انتخاب ژنومی، بلوک‌های هاپلوتایپی مبتنی بر LD، صحت پیش‌بینی، عدم تعادل پیوستگی، نشانگرهای SNP.

### Evaluation of prediction accuracy of the genomic breeding values of some economic traits in Iranian Holstein cattle using SNP markers and haplotype blocks

Behzad Rajabi Marand<sup>1</sup>, Hossien Moradi Shahrabak<sup>2\*</sup>, Mostafa Sadeghi<sup>3</sup> and Rostam Abdollahi-Aarpanahi<sup>4</sup>

1. Ph.D. Student, Department of Animal Sciences, University College of Agriculture & Natural Resources, University of Tehran, Karaj, Iran

2. Assistant Professor, Department of Animal Sciences, University College of Agriculture and Natural Resources, University of Tehran, Karaj, Iran.

3. Associate Professor, Department of Animal Sciences, University College of Agriculture and Natural Resources, University of Tehran, Karaj, Iran.

4. Assistant Professor, Department of Animal and Poultry Science (Aburairhan), University of Tehran, Pakdasht, Iran.

Received: June 18, 2019

Accepted: August 19, 2019

#### Abstract

The aim of current study was to evaluate the accuracy of genomic breeding values (GEBV) for two important economical traits of milk yield and somatic cell score using SNP markers and LD-based haplotype blocks (haploblocks) by two statistical methods of GBULP and Bayes B. The data set consisted of 1654 bulls genotyped with different marker densities. When SNPs were used, the accuracy of breeding values obtained by Bayes B was better than GBLUP. In other words, for milk yield and somatic cell score traits, the prediction accuracy of GBLUP was 0.54 and 0.44 and by Bayes B was 0.58 and 0.44, respectively. For milk yield, the prediction accuracy of using haploblocks in both statistical methods was higher than the prediction accuracy using SNPs, while for the somatic cell score, this increase was more pronounced when GBLUP was used. However, when Bayes B was used this superiority was only obtained when the  $r^2$  statistic used to build the haploblocks was higher than 0.2. The results showed that the optimum level of  $r^2$  for building haploblocks depends on the trait type and its heritability. As a result, using  $r^2$  statistic more than 0.2 for building haploblocks can increase the accuracy of breeding values for both traits compared to SNP markers.

**Keywords:** Genomic Selection, LD-based haplotype blocks, Linkage Disequilibrium, Prediction Accuracy, SNP Markers.

## مقدمه

انتخاب ژنومی شکلی از انتخاب به کمک نشانگر است که با استفاده از نقشه‌های نشانگری با تراکم زیاد، به خصوص SNPهایی که کل ژنوم را دربرمی‌گیرند و اغلب در عدم تعادل پیوستگی (LD) با QTLهای مجاور خود هستند، ارزش ژنتیکی کل را پیش‌بینی می‌کند [۱۰، ۱۴]. علی‌رغم پیشرفت‌هایی که با استفاده از انتخاب ژنومی در طی چندین سال گذشته حاصل شده است، این نوع انتخاب چالش‌های خاص خودش را دارد. به بیان ساده، در انتخاب ژنومی با استفاده از SNPها، احتمال این‌که SNPهای مختلف نزدیک QTL و یا در LD کامل با آن‌ها باشند بسیار کم است که از دلایل آن می‌توان به تراکم محدود نشانگرهای مورد استفاده در تعیین ژنوتیپ حیوانات، پایین بودن احتمال سببی بودن و نیز پایین بودن میزان آگاهی دهندگی این واریانت‌ها اشاره کرد [۳]. علاوه بر آن، در مدل‌هایی که از SNPها استفاده می‌کنند، معمولاً SNPها به صورت کاملاً مستقل از یکدیگر در نظر گرفته می‌شوند و امکان در نظر گرفتن اثر وقوع نوترکیبی در مدل‌های ارزیابی وجود ندارد، درحالی‌که نوترکیبی می‌تواند باعث تغییر فاز پیوستگی بین نشانگر و QTL شود.

یکی از مفاهیم مهمی که در انتخاب ژنومی مطرح است، هاپلوتایپ است. در واقع هاپلوتایپ، گروهی از آلل‌های کنار هم یا مجموعه‌ای از چند شکلی‌های تک نوکلئوتیدی (SNP) بر روی یک کروموزوم هستند که به دلیل وجود عدم تعادل ناشی از پیوستگی (LD)، تمایل دارند که به صورت متصل به هم، از فردی به فرد دیگر منتقل شوند. یکی از بحث‌هایی که در رابطه با بلوک‌های هاپلوتایپی (هاپلو بلاک‌ها) مطرح است، استفاده از آن‌ها به جای SNPها در ارزیابی‌های ژنومی می‌باشد، به طوری‌که در مطالعات مختلف، از هاپلوبلاک‌ها برای پیش‌بینی ارزش‌های اصلاحی استفاده شده است [۳، ۴، ۱۹ و ۲۲].

از مزایای مهم استفاده از هاپلو بلاک‌ها در مقایسه با SNPها، می‌توان به این نکته اشاره کرد که یک QTL شاید در عدم تعادل ناشی از پیوستگی (LD) کامل با یک SNP نباشد ولی همین QTL می‌تواند در LD کامل با یک هاپلوبلاک (که چندین نشانگر در کنار هم هستند) باشد [۶]. علاوه بر آن، مزیت دیگر هاپلوتایپ‌ها در مقایسه با SNPها، توانایی بالای آن‌ها در شناسایی جهش‌ها می‌باشد. به عبارت دیگر، در مطالعات مربوط به اصلاح نژاد دام، SNPها را عموماً دو آللی در نظر می‌گیرند و حتی زمانی‌که جهش اتفاق می‌افتد ممکن است که فراوانی آلل‌ها نسبتاً بدون تغییر باقی بماند، اما زمانی‌که از هاپلوتایپ‌ها استفاده می‌شود جهش‌های اتفاق افتاده در جایگاه‌های مختلف باعث ایجاد تغییرات عمده‌ای در فراوانی هاپلوتایپ‌ها می‌شوند [۵]. بنابراین در صورتی‌که از هاپلوبلاک‌ها استفاده شود می‌توان اثر SNPها را به صورت همزمان برآورد کرد که در نتیجه منجر به افزایش صحت ارزش‌های اصلاحی ژنومی گردید.

تاکنون مطالعه‌ای در زمینه استفاده از SNPها و مقایسه آن با روش‌های ارزیابی مبتنی بر بلوک‌های هاپلوتایپی با استفاده از داده‌های واقعی، بر جمعیت گاوهای هلشتاین ایران انجام نگرفته است. بنابراین هدف اصلی مطالعه حاضر، محاسبه و مقایسه صحت ارزش‌های اصلاحی ژنومی به دست آمده با استفاده از بلوک‌های هاپلوتایپی و SNPها در ارزیابی‌های ژنومی بود.

## مواد و روش‌ها

این مطالعه با همکاری و مشارکت مرکز اصلاح نژاد و بهبود تولیدات دامی کشور انجام گرفت. در این مطالعه از اطلاعات ژنوتیپی مربوط به ۱۹۴۰ راس گاو هلشتاین که با تراکم‌های مختلف (۱۷ نوع تراشه مختلف با تراکم ۷k تا ۱۴۰k) توسط شرکت‌های تابعه مرکز اصلاح نژاد دام

جدول ۱. خلاصه نتایج حاصل از کنترل کیفیت اطلاعات

ژنومی	
۱۹۴۰	تعداد کل حیوانات
۲۸۶	حیوانات حذف شده براساس ویرایش ابتدایی اطلاعات ژنوتایپی
۰	حیوانات با بیش از ۵ درصد ژنوتیپ از دست رفته
۱۶۵۴	حیوانات باقیمانده
۵۴۰۰۱	تعداد نشانگر استفاده شده (SNP)
۱۶۶۴	تعداد نشانگرها با موقعیت فیزیکی نامشخص
۱۱۵۲	نشانگرهای موجود بر روی کروموزوم‌های جنسی
۵۱۱۸۵	نشانگرهای باقیمانده قبل از ایمپوتیشن
	کنترل کیفیت نشانگرهای ایمپوت شده
۱۶۷۷۷	نشانگرهای حذف شده براساس HWE کم‌تر از $10^{-6}$
۲۴۹	نشانگرهای حذف شده براساس MAF کم‌تر از ۱ درصد
۳۴۱۵۹	تعداد نشانگرهای باقیمانده

در این پژوهش، دو صفت تولید شیر و شمارش سلول‌های پیکری (SCC) به دلیل وراثت پذیری متفاوت و قرارگرفتن در اهداف اصلاح نژادی، مورد مطالعه قرار گرفتند. از امتیاز سلول‌های بدنی (SCS) به جای شمارش سلول‌های بدنی (SCC) استفاده شد که از طریق رابطه (۱) به دست آمد [۱]:

$$SCS = \log_2 \frac{SCC}{100000} + 3 \quad (\text{رابطه ۱})$$

اطلاعات فنوتیپی استفاده شده برای برآورد اثرات SNPها و صحت ارزش‌های اصلاحی ژنومی به صورت ارزش‌های اصلاحی غیررگرسیون گاوهای نر (DRP) بود که از طریق ارزش‌های اصلاحی کلاسیک (EBV) و تعداد دختران آن گاو نر به دست آمد [۲۰]. به همین منظور ابتدا با استفاده از مدل تابعیت تصادفی ارزش‌های اصلاحی کلاسیک برآورد شده و سپس به کمک ارزش‌های اصلاحی حاصل، مقادیر DRP گاوهای نر برای هر صفت به تفکیک با استفاده از رابطه (۲) به دست آمد [۲۰]:

کشور تهیه شده بود، استفاده گردید. پس از ویرایش ابتدایی داده‌ها، در مجموع اطلاعات ژنوتیپی مربوط به ۲۸۶ راس دام به دلیل ایراداتی از قبیل تکراری بودن در اثر تعیین ژنوتیپ با تراکم‌های مختلف، عدم وجود شماره بین‌المللی حیوان ژنوتیپ شده در فایل اطلاعات، خطا در نحوه ژنوتایپینگ (تکرار جایگاه‌های ژنوتایپ شده به دفعات زیاد) و عدم وجود اطلاعات شجره‌ای کنار گذاشته شد. در نهایت اطلاعات ژنوتیپی مربوط به ۱۶۵۴ راس گاو نر که با استفاده از تراشه‌هایی با تراکم‌های مختلف تعیین ژنوتیپ شده بودند مورد استفاده قرار گرفت.

با توجه به این که برای انجام آنالیزهای ژنومی به تراشه‌هایی با تراکم‌های همگن نیاز بود بنابراین لازم بود که کلیه تراشه‌ها به یک تراکم واحد ایمپوت شوند. برای این منظور تراشه با تراکم ۵۴۰۰۱ به عنوان تراشه مرجع در نظر گرفته شد. برای ایمپوت صحیح، ابتدا یک کنترل کیفیت اولیه بر روی تراشه ۵۴۰۰۱ انجام گرفت و SNPهایی که در این تراشه دارای موقعیت فیزیکی نامشخص و یا بر روی کروموزوم‌های جنسی X و Y بودند، حذف شدند. در نتیجه با حذف این SNPها تراشه ۵۴۰۰۱ به ۵۱۱۸۵ کاهش یافت و سپس کلیه تراشه‌ها با استفاده از عمل ایمپوتیشن و به کمک نرم‌افزار FImpute (نسخه ۲,۲) [۱۸] به تراکم ۵۱۱۸۵ رسیدند.

بعد از انجام ایمپوتیشن، کیفیت داده‌های ژنومی ایمپوت شده، با استفاده از فاکتورهای از قبیل تعداد حیوانات ژنوتایپ شده به ازای هر SNP، تعادل هاردی-وینبرگ (HWE) و حداقل فراوانی آلی (MAF) انجام گرفت. برای کنترل کیفیت داده‌های ژنومی، از برنامه PLINK (نسخه ۱,۰۷) [۱۶] استفاده شد. در جدول ۱ خلاصه نتایج حاصل از کنترل کیفیت اطلاعات ژنومی مورد استفاده آورده شده است.

برای دو جایگاه دو آللی با آلل‌های  $A_1/A_2$  و  $B_1/B_2$  مقدار آماره  $r^2$  به صورت رابطه ۳ تعریف می‌شود [۱۱، ۱۳]:

$$r^2 = \frac{D^2}{p_{A_1} p_{A_2} p_{B_1} p_{B_2}} \quad (\text{رابطه ۳})$$

در رابطه فوق

$D = P(A_1 B_1)P(A_2 B_2) - P(A_1 B_2)P(A_2 B_1)$  می‌باشد. مقدار  $r^2$  همبستگی بین دو جایگاه را نشان می‌دهد که مقدار آن بین صفر و ۱ تغییر می‌کند. مقدار  $r^2$  برابر با ۱ نشان‌دهنده وجود LD کامل بین دو جایگاه و مقدار  $r^2$  برابر با صفر بیانگر استقلال کامل دو جایگاه از یکدیگر است. در این مطالعه، بلوک‌های هاپلوتایپی برای هر کروموزوم به صورت مجزا تشکیل شد. برای این منظور از مقادیر  $r^2$  برابر ۰/۱، ۰/۱۵، ۰/۲، ۰/۲۵ و ۰/۳ برای تشکیل بلوک‌های هاپلوتایپی استفاده شد. به عبارت دیگر در صورتی که مقدار LD بین SNPها بالاتر از آستانه تعریف شده ( $r^2$ ) بود آن گروه به عنوان بلوک هاپلوتایپی در نظر گرفته شد. استفاده از LD برای تشکیل بلوک‌های هاپلوتایپی باعث جلوگیری از گروه‌بندی تصادفی نشانگرها شده و در نتیجه باعث کاهش تعداد واریانت‌ها در هر بلوک و در نهایت باعث کاهش تعداد متغیرهای مورد نیاز برای پیش‌بینی‌های ژنومی می‌گردد [۶]. طول بلوک‌ها بستگی به نوع کروموزوم و آستانه‌ی در نظر گرفته شده داشت ولی به طور متوسط اندازه هر بلوک بین ۲ الی ۸ نشانگر بود.

در این مطالعه، از دو روش آماری GBLUP و Bیز برای پیش‌بینی ارزش‌های اصلاحی حیوانات استفاده شد. برای این منظور سناریوهای متعددی با استفاده از (۱) استفاده از کل SNPها، (۲) استفاده از بلوک‌های هاپلوتایپی و SNPهای بلوک‌بندی‌نشده، (۳) استفاده از بلوک‌های هاپلوتایپی با ۵ آستانه تعریف‌شده برای  $r^2$  انجام گرفت. در مجموع، ۱۱ آنالیز مختلف برای هر صفت صورت گرفته و پیش‌بینی‌های ژنومی برای هر یک از این

$$DRP = PA + \frac{(EBV - PA)}{Rel_{DD}} \quad (\text{رابطه ۲})$$

$$DE_{prg} = \frac{Rel_{EBV}}{1 - Rel_{EBV}} - \frac{Rel_{PA}}{1 - Rel_{PA}}$$

$$Rel_{DD} = \frac{DE_{prg}}{DE_{prg} + 1}$$

در این رابطه،  $Rel_{EBV}$ ،  $Rel_{PA}$  و  $Rel_{DD}$  به ترتیب بیانگر قابلیت اعتماد ارزش‌های اصلاحی برآوردشده، میانگین والدین و انحراف دختران بوده و  $DE_{prg}$  نشان‌دهنده معادل دختر است که با استفاده از اطلاعات نتاج به دست آمد. به دلیل اندازه کوچک جمعیت، به جای استفاده از تاریخ تولد برای تشکیل جمعیت‌های مرجع و تأیید از روش 5-fold-Cross Validation استفاده شد و میانگین این ۵ تکرار برای هر صفت گزارش شد. اندازه جمعیت‌های مرجع و تأیید برای هر صفت در جدول ۲ نشان داده شده است. با توجه به این که تعداد افراد دارای رکورد برای صفات مختلف یکسان نبودند، بنابراین تعداد افراد موجود در جمعیت‌های مرجع و تأیید براساس نوع صفت متغیر بود.

جدول ۲. اندازه جمعیت‌های مرجع و تأیید مورد استفاده

برای پیش‌بینی‌های ژنومی

صفت		
جمعیت	تولید شیر	امتیاز سلول‌های بدنی
مرجع	۹۸۳	۶۸۰
تایید	۲۴۶	۱۷۰
کل	۱۲۲۹	۸۵۰

برای اندازه‌گیری میزان LD جهت بلوک‌بندی هاپلوتایپ‌ها از سه آماره  $D$ ،  $D'$  و  $r^2$  استفاده می‌شود. با توجه به این که آماره  $r^2$  نسبت به سایر آماره‌ها کم‌تر به فراوانی آللی وابسته است [۱۸]، بنابراین در این مطالعه، جهت تشکیل بلوک‌های هاپلوتایپی از آماره  $r^2$  استفاده شد.

## تولیدات دامی

ارزیابی صحت ارزش‌های اصلاحی ژنومی برخی صفات اقتصادی جمعیت گاوهای هلستاین ایران با استفاده از SNPها و بلوک‌های هاپلوتایپی

در این رابطه،  $M$ ، نشان‌دهنده تعداد نشانگرها،  $p_i$ ، فراوانی آلل بزرگ،  $2p_i$ ، بیانگر میانگین و  $2p_i(1-p_i)$ ، انحراف استاندارد می‌باشد. در این روش اثر نشانگرها از توزیع نرمال تبعیت می‌کنند به عبارت دیگر نشانگرها به صورت تصادفی و با واریانس یکسان در نظر گرفته می‌شوند.

در روش بیز  $B$  فرض بر این است که تنها یکسری از جایگاه‌های ژنی وجود دارد که کل واریانس ژنتیکی صفت را توجیه می‌کند و در نتیجه بخش اعظمی از جایگاه‌های ژنی نقشی در ایجاد واریانس ژنتیکی صفت ندارند. به عبارت دیگر فرض اصلی این روش این است که بسیاری از SNPها در محل‌هایی از ژنوم قرار دارند که فاقد QTL است و بر صفت تأثیر ندارند درحالی‌که فقط بخش اندکی از SNPها در عدم تعادل پیوستگی با QTL بوده و در نتیجه دارای اثر می‌باشند. مدل آماری مورد استفاده در روش بیز  $B$  به صورت رابطه ۷ بود.

$$y_j = \mu + \sum_{i=1}^m X_{ij} \beta_i \delta_i e_j \quad (\text{رابطه ۸})$$

در رابطه فوق،  $y$ ، نشان‌دهنده فنوتیپ حیوان (DRP)،  $\mu$ ، میانگین کل،  $m$ ، تعداد نشانگر،  $X$ ، نشان‌دهنده ژنوتیپ نشانگر در جایگاه  $i$  مربوط به ژامین حیوان،  $\beta_i$ ، اثر جایگزینی آلی در جایگاه  $i$  و  $\delta_i$ ، متغیر اختصاصی این مدل بوده و نشان‌دهنده عدم حضور (با احتمال  $\pi$ ) یا حضور (با احتمال  $1-\pi$ ) جایگاه  $i$  در مدل است (رابطه ۹).

$$\delta_{i=1} \Rightarrow \beta_i \sim N(0, \sigma_i^2) \quad (\text{رابطه ۹})$$

$$\delta_{i=0} \Rightarrow \beta_i = 0$$

در روش بیز  $B$  اثرات نشانگری با احتمال  $\pi$  برابر صفر و با احتمال  $1-\pi$  از توزیع  $t$  به دست می‌آید. برای محاسبه اثر نشانگرها از الگوریتم MCMC استفاده شد. به این منظور از یک زنجیره با طول ۵۰۰۰۰ تکرار استفاده شد که

آنالیزها انجام و با یکدیگر مقایسه شدند. مدل آماری ۴ در روش GBLUP استفاده شد [۲۱].

$$y = 1\mu + Zg + e \quad (\text{رابطه ۴})$$

در این رابطه،  $y$ ، بردار فنوتیپ تصحیح‌شده (DRP) حیوانات جمعیت مرجع؛  $\mu$ ، میانگین کل؛  $u$ ، بردار ارزش‌های اصلاحی ژنومی و  $Z$ ، ماتریس ضرایب با ابعاد  $n \times m$  است ( $n$  تعداد حیوانات و  $m$  تعداد نشانگرها) که مشاهدات را به بردارهای ارزش‌های اصلاحی مرتبط می‌کند؛  $e$ ، بردار اثرات تصادفی باقیمانده و  $1$ ، بردار یکه می‌باشد. در این مدل فرض می‌شود که  $g \sim N(0, G\sigma_g^2)$

و  $e \sim N(0, R\sigma_e^2)$ ، ماتریس خویشاوندی ژنومی؛  $R$ ، ماتریس قطری؛  $\sigma_e^2$ ، واریانس باقیمانده و  $\sigma_g^2$ ، واریانس ژنتیکی افزایشی می‌باشد. در روش بلوک‌های هاپلوتایپی، ممکن است برای هر بلوک هاپلوتایپی بیش از یک متغیر وجود داشته باشد، زیرا در هر بلوک هاپلوتایپی ممکن است بیش از دو SNP وجود داشته باشد. در این حالت ماتریس ژنوتیپی یا  $Z$  برای هر بلوک هاپلوتایپی بیش از یک ستون خواهد داشت و دارای ابعاد  $n \times q$  ( $n$ : تعداد حیوانات و  $q$  تعداد کل بلوک‌های هاپلوتایپی می‌باشد) خواهد بود که در آن اعداد ۰، ۱ و ۲ به این معنی خواهند بود که هر فرد چند کپی از هر بلوک هاپلوتایپی موجود در هر ستون را دارا می‌باشد. برای برآورد اثر نشانگرها معادلات مختلط هندرسون به صورت رابطه ۵ تشکیل شد.

$$\hat{g} = [I + G^{-1}\alpha]^{-1}[y] \quad (\text{رابطه ۵})$$

در این رابطه،  $\alpha$  باز رابطه ۶ محاسبه شد.

$$\alpha = \left( \frac{1-h^2}{h^2} \right) \quad (\text{رابطه ۶})$$

و  $G$  ماتریس خویشاوندی ژنومی است که از طریق رابطه ۷ به دست آمد [۲۳]:

$$G_{ijk} = \frac{1}{M} \sum_{i=1}^M \frac{(x_{ij} - 2p_i)(x_{ik} - 2p_i)}{2p_i(1-p_i)} \quad (\text{رابطه ۷})$$

## تولیدات دامی

آمارى اختلاف معنی داری با هم دارند. از نرم افزار BGLR (نسخه ۱/۰/۸) [۸] برای اجرای روش های GBLUP و بیز B، از نرم افزار Haploview (نسخه ۴/۲) [۲] برای تشکیل بلوک های هاپلوتایپی و از نرم افزار REMLF90 برای برآورد وراثت پذیری و مؤلفه های واریانس برای هر صفت استفاده شد.

### نتایج و بحث

تعداد و انواع بلوک های هاپلوتایپی به همراه تعداد SNP های بلوک بندی نشده برای ۳۴۱۵۹ جایگاه نشانگری در جدول ۳ نشان داده شده است. زمانی که مقدار آستانه برای تشکیل بلوک های هاپلوتایپی پایین (۰/۱) است، در این صورت تعداد زیادی از SNP ها این شانس را دارند که در داخل بلوک ها قرار گیرند. بنابراین در این شرایط، بلوک هایی تشکیل خواهد شد که طول آن ها بزرگ، ولی تعداد آن ها کم تر می باشد و چون اکثر نشانگرها در داخل بلوک ها قرار گرفته اند، بنابراین تعداد SNP های بلوک بندی نشده کم تر خواهد بود. در مقابل، به موازات افزایش حد آستانه آماره  $t^2$  بلوک های هاپلوتایپی شکسته شده و در نتیجه اندازه بلوک ها کوتاه تر ولی تعداد آن ها بیش تر خواهد شد. بنابراین تعداد SNP هایی که نتوانسته اند حد آستانه لازم را کسب کرده و در داخل بلوک قرار بگیرند افزایش می یابد.

جدول ۳. تعداد و انواع بلوک های هاپلوتایپی به همراه تعداد

SNP های بلوک بندی نشده برای ۳۴۱۵۹ نشانگر

آماره	تعداد بلوک های هاپلوتایپی	تعداد آلل های هاپلوتایپی	تعداد SNP های بلوک بندی نشده
$t^2$	هاپلوتایپی	هاپلوتایپی	هاپلوتایپی
۰/۱	۴۴۷۴	۱۸۳۹۶	۲۹۷۲
۰/۱۵	۴۹۸۲	۱۹۶۴۳	۳۸۴۶
۰/۲	۵۸۸۱	۲۱۱۵۴	۵۰۲۳
۰/۲۵	۶۴۲۷	۲۲۵۳۰	۶۲۸۶
۰/۳	۶۷۵۳	۲۴۴۶۶	۸۸۴۵

۲۰۰۰۰ نمونه اول به عنوان دوره قلق گیری حذف شده و ۳۰۰۰۰ نمونه باقیمانده برای برآورد میانگین توزیع پسین هر پارامتر استفاده گردید. صحت ارزیابی های ژنومی هر یک از مدل ها برای هر یک از روش ها (SNP ها و بلوک های هاپلوتایپی) از همبستگی بین DRP و ارزش اصلاحی ژنومی محاسبه می شود. به کمک اطلاعات در دسترس، مقدار DRP محاسبه شده و از طریق ضریب همبستگی پیرسون بین GEBV و DRP صحت برآورد ارزش های اصلاحی ژنومی رابطه ۱۰ به دست آمد [۷].

$$r_{(GEBV, TBV)} = \frac{Cor(GEBV, DRP)}{Cor(DRP, TBV)} \quad (\text{رابطه ۱۰})$$

$$= \frac{Cor(GEBV, DRP)}{\sqrt{r_{DRP}^2}}$$

جهت بررسی وجود اختلاف آماری بین استفاده از SNP ها و بلوک های هاپلوتایپی، قابلیت اعتماد پیش بینی ها با استفاده از آزمون هتلینگ [۱۲] مقایسه شد. در صورتی  $r_{GEBV}$  مربوط به پیش بینی اول برابر  $r_{GEBV}$  پیش بینی دوم است که  $Cor(GEBV, DRP)$  مدل اول برابر  $Cor(GEBV, DRP)$  مدل دوم باشد. اگر  $GEBV, i = Cor(DRP, GEBV[\text{model } i])$  بوده و  $GEBV, j = Cor(DRP, GEBV[\text{model } j])$  باشد آماره  $t$  استفاده شده برای آزمون فرض صفر  $(H_0: GEBV, i = GEBV, j)$  به صورت رابطه ۱۱ خواهد بود.

$$t = (r_{GEBV, i} - r_{GEBV, j}) \sqrt{\frac{n(1+r_{ij})}{2D}} \quad (\text{رابطه ۱۱})$$

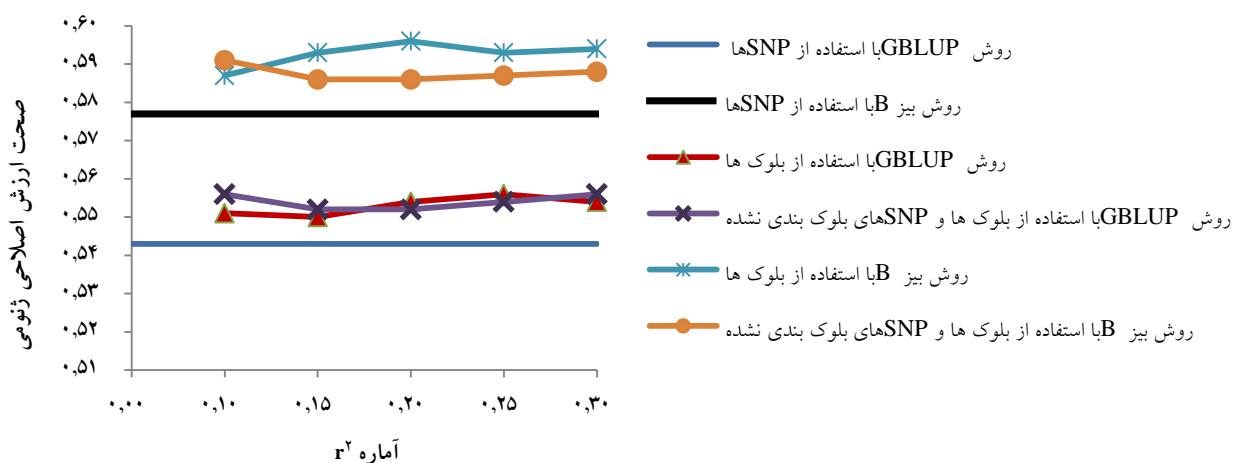
در رابطه فوق،  $t$  نشان دهنده همبستگی مشاهده شده،  $n$  نشان دهنده درجه آزادی  $(n = N - 3)$  با  $N$  مشاهده و  $D$ ، ترانواده ماتریس همبستگی بین DRP و GEBV برای مدل های  $i$  و  $j$  می باشد. اگر مقدار  $P$  محاسبه شده کم تر از سطح خطای مورد نظر  $(\alpha = 0.05)$  باشد در این صورت فرض  $H_0$  رد خواهد شد. بنابراین نتیجه گیری می شود که با سطح احتمال  $\alpha$  همبستگی های به دست آمده به لحاظ

ارزیابی صحت ارزش‌های اصلاحی ژنومی برخی صفات اقتصادی جمعیت گاوهای هلستاین ایران با استفاده از SNPها و بلوک‌های هاپلوتایپی

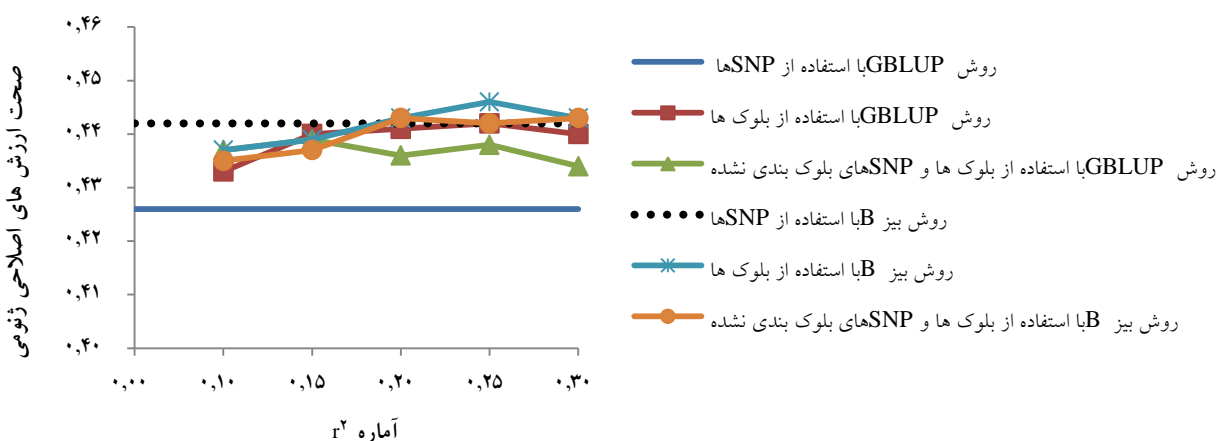
می‌تواند باعث کاهش مدت زمان محاسبات در تراکم‌های بالاتر نشانگری شود.

برای مقادیر مختلف  $r^2$ ، صحت ارزش‌های اصلاحی ژنومی ( $r_{GEBV,DRP}$ ) صفات مورد مطالعه با دو روش GBLUP و Bیز B و با استفاده از SNPها و بلوک‌های هاپلوتایپی به‌عنوان متغیر مستقل محاسبه و با یکدیگر مقایسه شدند (شکل‌های ۱ و ۲).

افزایش تعداد بلوک‌ها، منجر به افزایش تعداد آلل‌های هاپلوتایپی مشاهده‌شده در جمعیت گردیده است. با این وجود، تعداد آلل‌های هاپلوتایپی تشکیل شده کم‌تر از تعداد SNPهای استفاده شده (۳۴۱۵۹ نشانگر) بود که حاکی از آن است که استفاده از LD برای تشکیل بلوک‌های هاپلوتایپی، می‌تواند منجر به کاهش تعداد متغیرها جهت انجام آنالیزهای ژنومی گردد که همین امر



شکل ۱. صحت ارزش‌های اصلاحی ژنومی ( $r_{GEBV,DRP}$ ) به‌دست آمده با استفاده از GBLUP و Bیز B برای صفت تولید شیر براساس SNPها و بلوک‌های هاپلوتایپی تشکیل شده براساس مقادیر مختلف آماره  $r^2$



شکل ۲. صحت ارزش‌های اصلاحی ژنومی ( $r_{GEBV,DRP}$ ) به‌دست آمده با استفاده از GBLUP و Bیز B برای صفت امتیاز سلول‌های بدنی بر اساس SNPها و بلوک‌های هاپلوتایپی تشکیل شده براساس مقادیر مختلف آماره  $r^2$

## تولیدات دامی

دوره ۲۱ ■ شماره ۴ ■ زمستان ۱۳۹۸

روش‌های GBLUP و بیز B، به ترتیب برابر ۰/۶۴ و ۰/۶۸ و برای تعداد سلول‌های سوماتیک شیر به ترتیب برابر ۰/۴۹ و ۰/۴۹ بود [۱۵] که بالاتر از مقادیر به دست آمده در مطالعه حاضر بود. علت این امر را می‌توان به استفاده از EBV به جای DRP به عنوان متغیر پاسخ و در نتیجه بیش برآورد صحت ارزش‌های اصلاحی ژنومی در آن مطالعه نسبت داد.

جهت بررسی معنی‌داری تفاوت مشاهده شده بین روش‌های مختلف آماری و همچنین مقایسه استفاده از بلوک‌های هاپلوتایپی به جای SNPها، صحت ارزش‌های اصلاحی به دست آمده از طریق سناریوهای مختلف، با استفاده از آزمون هتلینگ آزمون شدند. برای صفات تولید شیر و امتیاز سلول‌های بدنی، زمانی که از SNPها استفاده شد اختلاف معنی‌داری بین دو روش آماری مشاهده شد ( $p < 0.05$ ). در مقابل، وقتی از بلوک‌های هاپلوتایپی و همچنین بلوک‌های هاپلوتایپی به همراه SNPهای بلوک‌بندی نشده استفاده شد تنها برای صفت تولید شیر اختلاف معنی‌داری بین دو روش آماری مشاهده گردید و برای صفت امتیاز سلول‌های بدنی تفاوتی بین دو روش آماری مورد استفاده مشاهده نشد (جدول ۴).

برای صفت تولید شیر، صحت پیش‌بینی‌های حاصل با استفاده از بلوک‌های هاپلوتایپی در هر دو روش آماری، بالاتر از صحت حاصل هنگام استفاده از SNPها بود، در حالی که برای صفت امتیاز سلول‌های بدنی، این افزایش صحت، زمانی که از روش آماری GBLUP استفاده شد مشهود بود ولی هنگام استفاده از روش بیز B این برتری تنها زمانی به دست آمد که مقدار آماره  $r^2$  مورد استفاده برای تشکیل بلوک‌ها بالاتر از ۰/۲ بود. علت این برتری را می‌توان به دلیل وارد شدن بخشی از اثرات متقابل بین آلل‌ها به مدل هنگام استفاده از بلوک‌های هاپلوتایپی به جای SNPها نسبت داد که باعث افزایش در صحت ارزش‌های اصلاحی ژنومی شده است.

صحت ارزش‌های اصلاحی به دست آمده هنگام استفاده از SNPها حاکی از برتری بیز B نسبت به GBLUP بود به طوری که برای صفت تولید شیر و امتیاز سلول‌های بدنی صحت ارزیابی‌ها با استفاده از GBLUP به ترتیب برابر ۵۴/۳ و ۴۳/۷ درصد و برای روش بیز B به ترتیب برابر ۵۷/۷ و ۴۴/۲ درصد بود. در مطالعه‌ای که بر روی گاوهای هلستاین ایران انجام گرفت صحت پیش‌بینی‌های ژنومی برای صفت تولید شیر برای

جدول ۴. مقادیر P آزمون t هتلینگ برای مقایسه صحت ارزش‌های اصلاحی به دست آمده با استفاده از GBLUP و بیز B

آماره $r^2$	SNPها		بلوک‌های هاپلوتایپی		بلوک‌ها و SNPهای بلوک‌بندی نشده	
	تولید شیر	امتیاز سلول‌های بدنی	تولید شیر	امتیاز سلول‌های بدنی	تولید شیر	امتیاز سلول‌های بدنی
-	۰/۰۲۸*	۰/۰۴۴*	-	-	-	-
۰/۱	-	-	۰/۰۱۵*	۰/۶۳۴	۰/۰۱۳*	۰/۸۲۷
۰/۱۵	-	-	۰/۰۱۱*	۰/۹۲۲	۰/۰۲۳*	۰/۸۹۲
۰/۲	-	-	۰/۰۰۴*	۰/۸۷۴	۰/۰۱۲*	۰/۲۴۴
۰/۲۵	-	-	۰/۰۱۳*	۰/۶۶۵	۰/۰۱۴*	۰/۳۱
۰/۳	-	-	۰/۰۰۷*	۰/۷۴۹	۰/۰۱۵*	۰/۱۱۲

\* نشان‌دهنده معنی‌داری در سطح ۵ درصد ( $p < 0.05$ ).



ارزیابی صحت ارزش‌های اصلاحی ژنومی برخی صفات اقتصادی جمعیت گاوهای هلشتاین ایران با استفاده از SNPها و بلوک‌های هاپلوتایپی

تفاوت‌های مشاهده‌شده تنها در برخی سطوح  $r^2$  معنی‌دار بود، ولی برای صفت تولید شیر این تفاوت در اکثر سطوح  $r^2$  معنی‌دار بود (جدول ۵).

در مقابل، زمانی که از روش آماری بیز B استفاده شد برای صفت تولید شیر نتایج نسبتاً مشابهی با روش GBLUP به‌دست آمد ولی برای صفت امتیاز سلول‌های بدنی تفاوت معنی‌داری بین استفاده از SNPها و بلوک‌های هاپلوتایپی در سطوح مختلف  $r^2$  مشاهده نگردید (جدول ۶).

برای مقایسه صحت ارزش‌های اصلاحی به‌دست‌آمده با SNPها و بلوک‌های هاپلوتایپی، مقادیر P آزمون هتلینگ محاسبه گردید. برای هر دو صفت مورد مطالعه، زمانی که از GBLUP استفاده شد تنها در صورتی که مقدار آماره  $r^2$  مورد استفاده برای تشکیل بلوک‌های هاپلوتایپی بالاتر از ۰/۲ بود بلوک‌های هاپلوتایپی نسبت به SNPها تفاوت معنی‌داری نشان داد. درحالی‌که وقتی از بلوک‌های هاپلوتایپی به‌همراه SNPهای بلوک‌بندی‌نشده استفاده گردید. علی‌رغم این‌که برای صفت امتیاز سلول‌های بدنی

جدول ۵. مقادیر P آزمون t هتلینگ به‌منظور مقایسه صحت ارزش‌های اصلاحی به‌دست‌آمده با SNPها و بلوک‌های هاپلوتایپی با

#### استفاده از روش GBLUP

آماره $r^2$	بلوک‌های هاپلوتایپی		بلوک‌های هاپلوتایپی و SNPهای بلوک‌بندی‌نشده	
	تولید شیر	امتیاز سلول‌های بدنی	تولید شیر	امتیاز سلول‌های بدنی
۰/۱	۰/۱۲۷	۰/۳۸۶	۰/۰۳*	۰/۰۶۸
۰/۱۵	۰/۱۷۸	۰/۱۶۳	۰/۰۸۷	۰/۰۳۲*
۰/۲	۰/۰۴۵*	۰/۰۶۵	۰/۰۶۳	۰/۱۴۷
۰/۲۵	۰/۰۴۲*	۰/۰۴۶*	۰/۰۴۶*	۰/۰۳۹*
۰/۳	۰/۰۵*	۰/۰۴۹*	۰/۰۳۱*	۰/۱۱۹

\* نشان‌دهنده معنی‌داری در سطح ۵ درصد ( $p < 0.05$ ).

جدول ۶. مقادیر P آزمون t هتلینگ به‌منظور مقایسه صحت ارزش‌های اصلاحی به‌دست‌آمده با SNPها و بلوک‌های هاپلوتایپی با

#### استفاده از روش بیز B

آماره $r^2$	بلوک‌های هاپلوتایپی		بلوک‌های هاپلوتایپی و SNPهای بلوک‌بندی‌نشده	
	تولید شیر	امتیاز سلول‌های بدنی	تولید شیر	امتیاز سلول‌های بدنی
۰/۱	۰/۰۸۵	۰/۷۹۶	۰/۰۴*	۰/۵۸۱
۰/۱۵	۰/۰۴۳*	۰/۸۳۶	۰/۰۸۱	۰/۶۴۳
۰/۲	۰/۰۲۸*	۰/۹۲۵	۰/۰۷۷	۰/۹۵۲
۰/۲۵	۰/۰۴۷*	۰/۷۱۲	۰/۰۵۸	۰/۹۸۸
۰/۳	۰/۰۳۴*	۰/۹۳۷	۰/۰۵۵	۰/۸۹۴

\* نشان‌دهنده معنی‌داری در سطح ۵ درصد ( $p < 0.05$ ).

## تولیدات دامی

دوره ۲۱ ■ شماره ۴ ■ زمستان ۱۳۹۸

طول ثابت با تعداد SNP مشخص برای هر هاپلوتایپ تعریف کرده‌اند که این طول ثابت به صورت تعداد مشخصی SNP در کنار هم که تشکیل یک بلوک هاپلوتایپی را می‌دهند تعریف می‌شود [۳، ۲۲]. برخی دیگر از پژوهش‌گران برای تعریف بلوک، از بین ۱۰ جایگاه ژنی پشت سر هم تنها اولین جایگاه ژنی را در نظر گرفته و بلوک‌ها را تعریف کرده‌اند [۱۹]. درحالی‌که در تعریف بلوک‌های هاپلوتایپی این نگرانی وجود دارد که ایجاد بلوک‌های هاپلوتایپی ممکن است منجر به افزایش تعداد متغیرها شود زیرا در صورتی که SNP‌ها را به صورت تصادفی گروه‌بندی شود در این صورت تعداد متغیرها می‌تواند به شدت افزایش یابد.

در تحقیق حاضر از LD برای تشکیل هر بلوک هاپلوتایپی استفاده شد. این روش علاوه بر کاهش تعداد متغیرها باعث عدم کاهش اطلاعات به دست آمده نیز می‌شود [۹]. با تعریف حداقل مقدار LD بین SNP‌ها می‌توان آن‌ها را در گروه‌هایی دسته‌بندی کرد که این گروه‌ها طول ثابتی ندارند و به دلیل وجود LD نسبتاً قوی، تعداد واریانت‌ها به ازای هر بلوک هاپلوتایپی در مقایسه با زمانی‌که بلوک‌های هاپلوتایپی براساس تعداد ثابتی از SNP‌های نزدیک هم تعریف می‌شوند، به صورت قابل ملاحظه‌ای کاهش یافته و در نتیجه تعداد بلوک‌های هاپلوتایپی مورد نیاز برای ارزیابی‌های ژنومی کاهش می‌یابد. به عبارت دیگر، بلوک‌های مبتنی بر LD براساس حداقل LD مورد نیاز بین نشانگرها تعریف می‌شوند. در واقع، در یک گروه از نشانگرها که به صورت متوالی و پشت سر هم قرار گرفته‌اند اگر میزان LD بین هر جفت نشانگر بالاتر از مقدار LD مدنظر باشد، در این صورت این گروه به عنوان یک بلوک هاپلوتایپی در نظر گرفته می‌شود. به دلیل وجود LD بین نشانگرها در بلوک‌های هاپلوتایپی، تعداد ترکیبات مربوط به SNP‌ها یا به عبارتی تعداد آلل‌های هاپلوتایپی کم خواهد بود. بنابراین می‌توان

نتایج به دست آمده در این مطالعه با نتایجی که بر روی جمعیت گاوهای هلشتاین کشورهای اسکانندیناوی (دانمارک، نروژ، سوئد و فنلاند) گزارش شد [۶] مطابقت داشت. در مطالعه مذکور، نشان داده شد که استفاده از بلوک‌های هاپلوتایپی به جای SNP‌ها باعث افزایش معنی‌داری در صحت پیش‌بینی ارزش اصلاحی ژنومی برای صفت پروتئین شیر که دارای معماری ژنتیکی و وراثت‌پذیری مشابه صفت تولید شیر است می‌شود، درحالی‌که برای صفت ورم پستان که دارای وراثت‌پذیری پایینی است تفاوت معنی‌داری بین استفاده از بلوک‌های هاپلوتایپی در مقایسه با SNP‌ها مشاهده نگردیده است.

صحت ارزش‌های اصلاحی به دست آمده نشان داد که برای صفاتی که وراثت‌پذیری بالایی دارند (صفت تولید شیر با  $h^2=0/3$ ) استفاده از بلوک‌های هاپلوتایپی به جای SNP‌ها می‌تواند باعث افزایش صحت پیش‌بینی‌ها شود. در مقابل برای صفات با وراثت‌پذیری پایین (صفت امتیاز سلول‌های بدنی با  $h^2=0/14$ ) تفاوت چندانی بین استفاده از بلوک‌های هاپلوتایپی و SNP‌ها مشاهده نگردید. با توجه به این‌که اطلاعات فنوتیپی استفاده شده در این مطالعه، به صورت DRP بود بهبود پیش‌بینی‌ها برای صفت تولید شیر را می‌توان به صحت بالای داده‌های استفاده شده برای این صفت در مقایسه با صفت امتیاز سلول‌های بدنی نسبت داد. هنگامی‌که از آزمون هتلینگ استفاده شد اختلافات مشاهده شده بین روش‌های آماری و نوع متغیر استفاده شده به لحاظ آماری تأیید شد و برای صفت تولید شیر، برتری استفاده از بلوک‌ها به جای SNP‌ها در هر سناریو ثابت شد ولی برای صفت امتیاز سلول‌های بدنی تنها در صورتی‌که مقدار آماره  $r^2$  مورد استفاده برای تشکیل بلوک‌ها بالاتر از ۰/۲ بود بلوک‌های هاپلوتایپی نسبت به SNP‌ها تفاوت معنی‌داری نشان دادند.

برخی پژوهش‌گران بلوک‌های هاپلوتایپی را به صورت

- simulation of data, validation, reporting, and benchmarking. *Genetics*. 193: 347-365.
- De los Campos G and Perez PR (2012) Bayesian Generalized Linear Regression BGLR. R Package. <http://bglr.rforge.rproject.org/>
  - Gabriel SB, Schaffner SF, Nguyen H, Moore JM, Roy J, Blumenstiel B, Higgins J, DeFelice M, Lochner A, Faggart M and Liu-Cordero SN (2002) The structure of haplotype blocks in the human genome. *Science* 296(5576): 2225-2229.
  - Goddard ME, Hayes BJ and Meuwissen TH (2010) Genomic selection in livestock populations. *Genetics research*, 92(5-6): 413-421.
  - Hill WG and Robertson A (1968) Linkage disequilibrium in finite populations. *Theoretical and Applied Genetics* 38(6): 226-231.
  - Hotelling H (1940) The selection of variates for use in predictions with some comments on the problem of nuisance parameters. *AnnMath Stat* 11: 271-283.
  - McRae AF, McEwan JC, Dodds JC, Wilson T, Crawford AM, Slate J (2002) Linkage disequilibrium in domestic sheep. *Genetics* 160: 1113-1122.
  - Meuwissen TH, Hayes BJ and Goddard ME (2001) Prediction of total genetic value using genome-wide dense marker maps. *Genetics*, 157(4): 1819-1829.
  - Mohammadi Y, Shariati MM, Zerehdaran S, Razmkabir M, Sayyadnejhad MB and Zandi MB (2016) Comparison of accuracy of genomic breeding value estimation for production traits in Iranian Holstein cattle by parametric and nonparametric methods. *Journal of Animal Production* 18(1): 1-11. (in Persian)
  - Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR and Bender D (2007) PLINK: a tool set for whole-genome association and population-based linkage analyses. *American Journal of Human Genetics* 81: 559-575.
  - Sargolzaei M, Schenkel FS, Jansen JB and Schaeffer LR (2008) Extent of linkage disequilibrium in Holstein cattle in North America. *J. Dairy Sci.* 91: 2106-2117.
  - Sargolzaei M, Chesnais JP and Schenkel FS (2014) A new approach for efficient genotype imputation using information from relatives. *BMC Genomics* (15): 478.
  - Schrooten C, Schopen GCB, Parker A, Medley A and Beatson P (2013) Across-breed genomic evaluation based on bovine high density genotypes and phenotypes of bulls and cows. In *Proc Assoc Advmt Anim Breed Genet* (20): 138-141.

چنین استنباط کرد که بلوک‌های هاپلوتایپی مبتنی بر LD می‌تواند باعث کاهش تعداد متغیرهای مورد استفاده به‌عنوان پیش‌بینی‌کننده در مدل‌ها شود. در مطالعه حاضر ۵ سطح مختلف آماره  $r^2$  برای تشکیل بلوک‌های هاپلوتایپی ارزیابی شده و با روش‌های آماری GBLUP و بیز B مقایسه شدند. نتایج نشان داد که سطح بهینه آماره  $r^2$  برای تشکیل بلوک‌های هاپلوتایپی بستگی به نوع صفت و وراثت‌پذیری آن دارد ولی در مجموع، استفاده از آماره  $r^2$  بیش‌تر از ۰/۲ جهت تشکیل بلوک‌های هاپلوتایپی، می‌تواند منجر به افزایش صحت ارزش‌های اصلاحی ژنومی در هر دو صفت (با وراثت‌پذیری متفاوت) در مقایسه با استفاده از SNPها شود.

### تعارض منافع

هیچ‌گونه تعارض منافع توسط نویسندگان وجود ندارد.

### منابع

- Ali AKA, Shook GE (1980) An optimum transformation for somatic cell concentration in milk. *J Dairy Sci* 63: 487-490.
- Barrett J, Fry B, Maller J and Daly M (2004) Haploview analysis and visualization of LD and haplotype maps. *Bioinformatics* 21(2): 263-265.
- Boichard D, Guillaume F, Baur A, Croiseau P, Rossignol MN, Boscher MY, Druet T, Genestout LUCIE, Colleau JJ, Journaux L and Ducrocq V (2012) Genomic selection in French dairy cattle. *Animal Production Science* 52(3): 115-120
- Calus MPL, De Roos APW and Veerkamp RF (2008) Accuracy of genomic selection using different methods to define haplotypes. *Genetics* 178(1): 553-561.
- Curtis D, North BV, Sham PC (2001) Use of an artificial neural network to detect association between a disease and multiple marker genotypes. *Ann Hum Genet* 65: 95-107.
- Cuyabano, B.C., Su, G. and Lund, M.S (2014) Genomic prediction of genetic merit using LD-based haplotypes in the Nordic Holstein population. *BMC genomics* 15(1): 1171.
- Daetwyler HD, Calus MPL, Pong-Wong R, De Los Campos G and Hickey JM (2013) Genomic prediction in animals and plants:

20. VanRaden PM, Van Tassell CP, Wiggans GR, Sonstegard TS, Schnabel RD, Taylor JF and Schenkel FS (2009) Invited review: Reliability of genomic predictions for North American Holstein bulls. *J Dairy Sci* 92: 16-24.
21. VanRaden PM (2008) Efficient methods to compute genomic predictions. *Journal of dairy science* 91(11): 4414-4423.
22. Villumsen TM, Janss L and Lund MS (2008) The importance of haplotype length and heritability using genomic selection in dairy cattle. *Journal of Animal Breeding and Genetics* 126(1): 3-13.
23. Yang J, Benyamin B, McEvoy BP, Gordon S, Henders AK, Nyholt DR, Madden PA, Heath AC, Martin NG, Montgomery GW and Goddard ME (2010) Common SNPs explain a large proportion of the heritability for human height. *Nature genetics* 42(7): 565-569.