Iranian Journal of Animal Science
Vol 51, No 3, Autumn 2020 (221-229)
DOI: 10.22059/ijas.2020.308210.653796

علوم دامی ایران
دورۀ ۵۱، شمارۀ ۳، پاییز ۱۳۹۹ (ص ۲۲۹-۲۲۱)

مطالعه پیوستگی سراسری ژنوم با صفات تولید شیر و معیار سلول‌های بدنی گاوهای هلشتاین ایران

رستم پهلوان[۱]، محمد مرادی شهربابک[۲*]، اردشیر نجاتی جوارمی[۲] و رستم عبداللهی آرپناهی[۳]
۱ و ۲. دانشجوی دکتری و استاد، گروه علوم دامی، پردیس کشاورزی و منابع طبیعی دانشگاه تهران، کرج، ایران
۳. محقق، دانشگاه فلوریدا، آمریکا

(تاریخ دریافت: ۱۳۹۹/۵/۲۵ - تاریخ پذیرش: ۱۳۹۹/۷/۲۸)

چکیده

در پرورش گاوشیری هدف اصلی افزایش سود است. افزایش بیماری ورم‌پستان از مشکلات اصلی پرورش گاوشیری با کاهش طول‌عمر تولیدی و افزایش هزینه‌ها، سبب تحمیل زیان‌های اقتصادی جدی به این صنعت شده‌است. این مطالعه باهدف بررسی معماری ژنتیکی و شناسایی نواحی ژنومی مرتبط با صفات تولیدشیر و تعداد سلول‌های بدنی به‌عنوان یکی از شاخص‌های ارزیابی کیفیت شیر و معیاری غیرمستقیم از بیماری ورم‌پستان، انجام شده‌است. برای بررسی ارتباط و شناسایی پنجره‌های ژنومی اصلی، از روش مطالعه پیوستگی سراسری ژنوم یک‌مرحله‌ای با داده‌های ژنومی ۱۹۳۸ راس گاو نر استفاده شد. نتایج براساس واریانس ژنتیکی افزایشی پنجره‌های ۱/۵ مگابازی از SNPهای مجاور ارائه شده است. پنجره‌هایی که بیش از ۱٪ واریانس را کنترل می‌کنند به‌عنوان نواحی ژنومی اصلی و جهت یافتن ژن‌های کاندیدا استفاده شدند. تعداد ۱۱ پنجره ژنومی اصلی روی ۹ کروموزوم اتوزوم، حاوی ۹٤ ژن کاندیدا حدود ۲۰٪ واریانس ژنتیکی معیار سلول‌های بدنی را توصیف می‌کردند. بیشترین واریانس مربوط به پنجره‌ای روی کروموزوم ١٤ (۳/۸۵٪) بود. تعداد ٦ پنجره اصلی روی ٦ کروموزوم شامل ۸۹ ژن کاندیدا، مرتبط با تولیدشیر بودند. این پنجره‌ها ۸/۸٪ واریانس و مهم‌ترین پنجره روی کروموزوم ۱۰، حدود ۲/۰۸٪ واریانس را کنترل می‌نمودند. درخصوص مقدار چربی و پروتئین شیر، به‌ترتیب تعداد ۹ منطقه روی ۷ کروموزوم، حدود ۱۵/٦٪ و ۹ پنجره روی ۸ کروموزوم، حدود ۱۰/٦٪ از واریانس را توصیف می‌کردند. نتایج نشان داد که ٤ ناحیه ژنومی اثر پلیوتروپی دارند. یافته‌های این تحقیق می‌تواند به‌عنوان منبع اطلاعاتی مهمی در ارزیابی‌های ژنومی صفات تولید شیر و تعداد سلول‌های بدنی باشد.

واژه‌های کلیدی: پنجره‌های ژنومی، ژن‌های کاندیدا، معیار سلول‌های بدنی، هلشتاین.

Genome-wide association study for milk production and somatic cell score traits in Iranian Holstein cattle

Rostam Pahlavan[1], Mohammad Moradi-Shahrbabak[2*], Ardeshir Nejati-Javaremi[2] and Rostam Abdolahi-Arpanahi[3]
1, 2. Ph.D. Candidate in Animal Breeding and Genetics and Professor, Department of Animal Science, College of Agriculture and Natural Resources, University of Tehran, Karaj, Iran
3. Researcher, Department of Animal Science, University of Florida, USA
(Received: Aug. 15, 2020 - Accepted: Oct. 19, 2020)

ABSTRACT

The main objective of dairy farmers is to maximize their profit. Increased incidence of mastitis in farms is one of the health problems, causing in serious economic losses as a consequence of treatment costs and reduction of production and longevity. The objective of this study was to evaluate the genetic architecture and associated genomic regions with milk production and somatic cell score (SCS) as an indirect measure of mastitis and the quality of raw milk. Thus, an SNP data set from 1938 Holstein bulls were used in a single-step genome-wide association study. The proportion of additive genetic variance (agv) for each of 1.5-Mb genomic window (adjacent SNPs) was used to identify informative genomic regions, accounting for more than 1% of the agv. A total of 11 significant windows over 9 *bovine* autosomes were found for the SCS. A peak on BTA14 explained the largest proportion of variance (3.85%). These regions together, explained 20% of agv and harbored 94 candidate genes. For milk yield, we identified 6 informative windows across 6 chromosomes, and a peak on BTA10 explained 2.08% of agv. These regions, explained 8.8% of the agv and sheltered 89 candidate genes. For the fat yield, 9 significant windows were identified on 7 chromosomes and explained 15.6% of agv, and 9 windows contained 87 candidate genes on 8 bovine autosomes were associated with milk protein yield (10.6% of agv). Four genomic regions had a pleiotropic effect. These findings can be an important source of information in genomic evaluation of dairy cattle.

Keywords: Candidate gene, genomic windows, Holstein, somatic cell score.

* Corresponding author E-mail: moradim@ut.ac.ir

مقدمه

رشد جمعیت جهان و اهمیت شیر در تأمین احتیاجات غذایی، سبب شده تا بهبود عملکرد دام‌های اهلی به‌خصوص گاو شیری از نظر صفات مهم اقتصادی نظیر تولید شیر، به‌عنوان بخش مهمی از اهداف پرورش و برنامه‌های اصلاح نژادی در جهان تبدیل شود. در این خصوص بهبود قابل‌توجهی در عملکرد تولیدی گاوهای شیری صورت گرفته است. ورم پستان به‌واسطه عوامل محیطی، مدیریتی و ضعف مقاومت و ایمنی (بیشتر اکتسابی) حیوان نسبت به عوامل بیماری‌زا، یکی از مهم‌ترین بیماری‌های عفونی بوده که سبب زیان‌های اقتصادی جدی به صنعت گاو شیری و مشکلات مرتبط با کیفیت لبنیات گاو شیری در سراسر دنیا می‌شود. به‌دلیل هزینه‌های بالای ناشی از ورم پستان در صنعت، در سال‌های اخیر، مقاومت به این عارضه به‌عنوان یکی از اهداف اصلاح نژادی مهم از جنبه‌های اقتصادی و آسایش دام مدنظر واقع شده است. کشورهای محدودی به‌طور مستقیم و روتین، وقوع ورم پستان را رکوردگیری می‌نمایند و اغلب انتخاب مستقیم برای مقاومت به ورم پستان انجام نمی‌شود. در این خصوص، تعداد (شمارش) سلول‌های بدنی (سلول‌های سوماتیک) در شیر یا تبدیل لگاریتمی آن (SCS) به‌علت واریانس ژنتیکی بالاتر، رکوردگیری ساده‌تر و همبستگی بسیار بالا و مثبت آن با وقوع ورم پستان، به‌عنوان یک معیار از ورم پستان استفاده می‌شود. تحقیقات رابطه مثبت و نامطلوبی (Antagonism) را بین سلول‌های سوماتیک و صفات تولیدی از جمله تولید شیر، نشان می‌دهد (Carlén *et al*., 2004; Heringstad *et al*., 2008). تعداد سلول‌های بدنی، یکی از شاخص‌های مهم ارزیابی کیفیت و سلامتی شیر خام است و به‌منظور قیمت‌گذاری مدنظر قرار می‌گیرد. افزایش آن، سبب کاهش کیفیت فرآوری و کاهش کیفیت شیر خام به‌دلیل تغییر در ترکیبات شیر (چربی، پروتئین، لاکتوز و اسیدیته) می‌شود.

توسعه روش‌های توالی‌یابی و تولید تراشه‌های متراکم SNP، استفاده از پویش سرتاسر ژنوم (GWAS) برای یافتن واریانت‌های ژنتیکی وابسته به صفات اقتصادی گاو شیری را ممکن ساخت (Meredith *et al*., 2012; Peñagaricano *et al*., 2012) و به‌عنوان یک ابزار بسیار مؤثر در بررسی معماری ژنتیکی صفات پیچیده در راستای ارزیابی ژنومی دام‌های اهلی مطرح شده است. در یک مطالعه GWAS در گاو شیری هلشتاین آمریکا، مناطق ژنومی و SNPهای معنی‌دار مؤثر بر تعداد سلول‌های بدنی و تولید شیر گزارش نمودند (Cole *et al*., 2011). روی گاوهای هلشتاین شمال اروپا (Sahana *et al*., 2014) و برزیل (Iung *et al*., 2019) نیز SNPهای معنی‌داری مرتبط با صفت تولید شیر، سلول‌های سوماتیک و ورم پستان مشاهده و گزارش شده است. در مطالعه (Wang *et al*., 2015) روی گاوهای هلشتاین چین، تعداد ۴۸ SNP در سراسر ژنوم با تغییرات تعداد سلول‌های سوماتیک مرتبط بودند که عمده آن‌ها روی کروموزوم ۱۴ قرار داشتند. اگرچه مطالعات مختلف، SNPها و ژن‌های مؤثر بر سلول‌های سوماتیک و وقوع ورم پستان را گزارش کرده‌اند؛ ولی اغلب آن‌ها متفاوت بودند و SNPهای مشابه بین تحقیقات بسیار کم بوده است که دلایل آن را، شرایط محیط، نوع مدیریت گاوداری‌ها شامل صنعتی و نیمه‌صنعتی، تفاوت در عوامل بیماری‌زای بومی و پاسخ میزبان به آنها، پس زمینه ژنتیکی جمعیت موردبررسی و غیره می‌دانند که به مقدار زیادی بر ارتباط واریانت‌های ژنتیکی و ژن‌ها در طول ژنوم با فنوتیپ مؤثر هستند (Chen *et al*., 2015; Sahana *et al*., 2014; Wang *et al*., 2015). بنابراین هدف از این مطالعه بررسی ارتباط SNPهای سطح ژنوم با صفات تعداد سلول‌های سوماتیک، تولید شیر، چربی و پروتئین به‌منظور یافتن ژن‌ها یا نواحی ژنومی و کروموزومی شناخته‌شده یا جدید مرتبط با توارث این صفات به‌صورت جداگانه و مشترک بین صفات در گاوهای نر هلشتاین دارای اطلاعات نشانگری و رکورد دختری در داخل کشور، می‌باشد.

مواد و روش‌ها

اطلاعات فنوتیپ و جمعیت مورداستفاده

در این تحقیق از رکوردهای تصحیح‌شده دوبار دوشش ۳۰۵ روز صفات تولید شیر، چربی و پروتئین شیر و SCC تعداد ۸۴۰۹۷۲ رأس گاو ماده زایش اول هلشتاین ایران مربوط به سال‌های ۱۳۹۸-۱۳۸۲؛ موجود در بانک

اطلاعات مرکز اصلاح نژاد و بهبود تولیدات دامی کشور، استفاده شد. خلاصه آماری و تعداد داده‌های مورداستفاده پس از ویرایش و حذف داده‌های خارج از محدوده $\mu \pm 3\sigma$ از فایل ارقام (جدول ۱) به‌تفکیک ارائه شده است. در فایل شجره نیز تعداد کل حیوانات، پدرها و مادرها به‌ترتیب برابر ۹۳۴۲۱۶، ۱۰۱۲۵ و ۵۶۶۲۹۶ رأس دام بود.

برای نرمال‌نمودن تعداد سلول‌های بدنی از تبـدیل لگاریتمی (SCS) با استفاده از (رابطه ۱) اسـتفاده شـد (Alam *et al.*, 2015):

$$SCS = Log_2(SCC/100000) + 4 \quad (۱)$$

مدل مورداستفاده برای ارزیابی ژنتیکی و محاسبه ارزش‌های اصلاحی (ژنومی) صفات موردمطالعه به شرح (رابطه ۲) بود (Zhou *et al.*, 2019):

$$y = \boldsymbol{X}b + \boldsymbol{Z}g + \mathrm{e} \quad (۲)$$

در این رابطه، $\boldsymbol{y}$، بردار رکوردهای فنوتیپی تصحیح‌شده دوبار دوشش ۳۰۵ روز صفات تولید شیر، چربی و پروتئین شیر و تعداد سلول‌های بدنی گاوهای ماده زایش اول؛ $\boldsymbol{b}$، بردار اثر عوامل ثابت شامل آثار ثابت گله- سال- فصل زایش (HYS) و سن زایش اول (متغیر کمکی) بودند؛ $\boldsymbol{g}$، بردار مقادیر ارزش‌های اصلاحی (ژنومی)؛ $\boldsymbol{X}$ و $\boldsymbol{Z}$، ماتریس ضرایب که مشاهدات را به‌ترتیب به بردارهای اثر عوامل ثابت و مقادیر ارزش‌های اصلاحی مرتبط می‌کند و **e**: بردار اثر عوامل باقیمانده است. همچنین برای آماده‌سازی و ویرایش داده‌ها از نرم‌افزارهای Visual Foxpro 9.0، R 3.5 و SAS 9.0 استفاده شد.

اطلاعات ژنوتیپ، کنترل کیفیت و استنباط SNPها (Imputation)

داده‌های ژنوتیپ تراشه‌های SNP اسپرم‌های گاو نر هلشتاین موجود در بانک اطلاعات مرکز اصلاح نژاد و بهبود تولیدات دامی کشور اخذ گردید. ژنوتیپ‌ها از ۵ نوع آرایه SNP مختلف از 30K تا 140K متشکل شده بودند. تعداد دام در تراشه‌ها بدین‌ترتیب بود: ۱- تراشه 30k برابر ۳۵ راس؛ ۲- تراشه 76k برابر ۴۰۱ راس؛ ۳- تراشه 139k برابر ۴۱۴ راس؛ ۴- تراشه 50k (۵۴۶۰۱ عدد SNP) برابر ۳۶۹ راس؛ ۵- تراشه 50k (۵۴۰۰۱ عدد SNP) برابر ۷۴۳ راس. برای مشابه نمودن مجموعه داده‌های مختلفی که از تراشه‌های با تراکم مختلف تهیه می‌شوند، استنباط SNPها انجام شد. در این تحقیق به‌منظور یکسان نمودن تراکم SNPها به تراشه مرجع 50K از روش استنباط و برنامه FIMPUTE (Sargolzaei *et al.*, 2011) استفاده گردید. کنترل کیفیت این داده‌ها در دو مرحله پیش از استنباط SNPها با نرم‌افزار PLINK (Purcell *et al.*, 2007) بر اساس معیارهای ذیل انجام شد: ۱- حذف حیوانات با بیش از ۱۰ درصد ژنوتیپ ازدست‌رفته و حذف SNPهایی که در بیش از ۱۰ درصد افراد وجود نداشتند؛ ۲- حداقل فراوانی آللی (MAF) کمتر از ۰/۰۱ درصد؛ ۴- انحراف از تعادل هاردی-واینبرگ ($P<10^{-6}$). درنهایت تعداد ۴۲۷۱۰ SNP و ۱۹۳۸ رأس دام از فیلترها عبور نمودند.

GWAS به روش یک مرحله‌ای (ssGWAS) در قالب پنجره‌های ۱/۵ مگابازی و جستجوی ژن‌های کاندیدا

در GWAS به‌کمک روش ژنومی یک‌مرحله‌ای (Zhou *et al.*, 2019) به‌طور همزمان از منابع اطلاعاتی شامل کلیه اطلاعات ژنوتیپ، شجره و فنوتیپ افراد دارای اطلاعات ژنوتیپ و بدون اطلاعات ژنوتیپ استفاده می‌شود. در این روش از ترکیب ماتریس روابط خویشاوندی ژنومی بر اساس اطلاعات SNPها (**G**) و ماتریس روابط خویشاوندی شجره‌ای (**A**) و تشکیل ماتریس ترکیبی جدید **H**، برای پیش‌بینی ارزش‌های اصلاحی ژنومی و سپس محاسبه اثر هر SNP استفاده می‌شود.

جدول ۱. خلاصه آماری و تعداد داده‌های مورداستفاده برای هر یک از صفات موردبررسی

Table 1. Summary statistics and number of records for the analyzed traits

Trait	No. of records	Mean (SD)	Min	Max	Heritability*
Milk yield	632039	8459(1592)	2503	16178	**0.30**
Fat yield	440638	268.01(61.70)	55	600	**0.22**
Protein yield	404306	260.51(43.66)	56	540	**0.24**
SCS	**395359**	**3.25(1.11)**	**1.14**	**9.47**	**0.08**

* Without genomic data.

* بدون در نظر گرفتن اطلاعات و روابط خویشاوندی ژنومی.

ماتریس **H** به‌عنوان ماتریس روابط خویشاوندی ژنومی ترکیبی بوده و معکوس آن عبارت است از $\mathbf{A}^{-1}$، نشان‌دهنده معکوس ماتریس روابط خویشاوندی شجره‌ای؛ $\mathbf{G}^{-1}$، معکوس ماتریس خویشاوندی ژنومی بر اساس اطلاعات SNPها و $\mathbf{A}_{22}^{-1}$ معکوس ماتریس روابط خویشاوندی بر اساس اطلاعات شجره مربوط به حیواناتی که اطلاعات ژنوتیپ آن‌ها در دسترس است (Misztal *et al.*, 2014). با عنایت به متفاوت بودن اثر هر SNP، از GWAS یک مرحله‌ای تصحیح‌شده مکرر (iterative Weighted single step GWAS= WssGWAS) استفاده شد (Wang, H *et al.*, 2012)، که در آن ماتریس خویشاوندی ژنومی و آثار SNP بر اساس اهمیت نسبی روی صفت، تصحیح و در نهایت سبب افزایش صحت می‌شوند (Zhang *et al.*, 2016; Zhou *et al.*, 2019). وزن‌دادن به SNPها در هر مرحله تکرار (iteration) براساس روش پیشنهادی (VanRaden, 2008) تحت عنوان روش غیرخطی A (Nonlinear A) انجام شد. هنگامی‌که نتایج دومرحله متوالی یکسان شوند، توقف و نتایج مورد استفاده واقع می‌شوند (در این تحقیق تکرار سوم). نتایج این تحقیق براساس واریانس ژنتیکی توصیف‌شده توسط پنجره‌های (SNP window) ۱/۵ مگابازی از SNPهای مجاور (Moving windows) ارائه شده است (Han & Peñagaricano, 2016). به عبارت دیگر روش ویندوزی (پنجره) مورد استفاد در این تحقیق، روش *n* Mb window of adjacents SNPs (Misztal *et al.*, 2014) می‌باشد و مقدار n برابر ۱/۵ مگاباز منظور شده شده است (Han & Peñagaricano, 2016). مقدار واریانس ژنتیکی توصیف شده در یک توسط هر پنجره (ناحیه ژنومی) ۱/۵ مگابازی از (رابطه ۳) حساب می‌شود (Zhou *et al.*, 2019):

$$\frac{Var(u_i)}{\sigma_u^2} * 100 = \frac{Var(\sum_{j=1}^{B} Z_j\, s_j)}{\sigma_u^2} * 100 \qquad (۳)$$

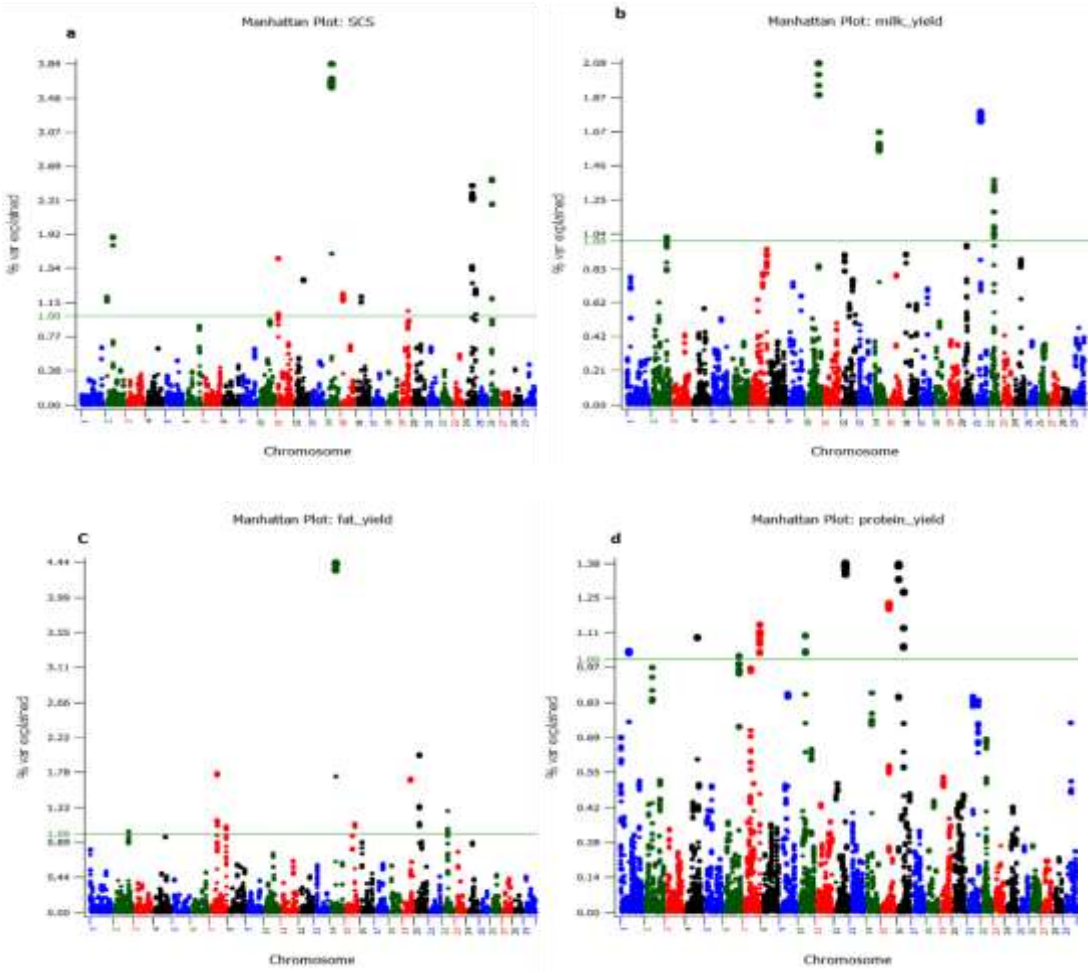
در این رابطه، u_i ارزش ژنتیکی i امین ناحیه ژنومی تحت مطالعه؛ B تعداد کل SNPهای مجاور در ناحیه ژنومی و S_j اثر j امین SNP در i امین ناحیه است. پنجره‌هایی که یک یا بیش از یک درصد واریانس ژنتیکی را دربرداشتند، به‌عنوان پنجره‌ها و نواحی ژنومی اصلی در نظر گرفته شدند و جهت شناسایی ژن‌های کاندیدا منظور می‌شوند (Han & Peñagaricano, 2016). جهت انجام آنالیزهای GWAS یک مرحله‌ای از نرم‌افزارهای خانواده Blupf90 (Misztal *et al.*, 2014) و برای رسم نمودار مطالعه پیوستگی (Manhattan Plot) از نرم‌افزار SNPEVG2 (Wang *et al.*, 2012) استفاده شد. تعیین ژن‌های کاندیدای بالقوه (Potential Candidate genes) بر اساس موقعیت ابتدا و انتهای هر پنجره غیرهمپوشان (Peñagaricano *et al.*, 2012)، با استفاده از ابزار Biomart تعبیه‌شده در بانک اطلاعاتی Ensembl انجام شد. ژنوم اسمبلی UMD نسخه ۳/۱ گاوی به‌عنوان مرجع در نظرگرفته شد.

نتایج و بحث

در این تحقیق ارتباط SNPهای سطح ژنوم با صفات تعداد سلول‌های بدنی (SCS)، تولید شیر، چربی و پروتئین با روش WssGWAS انجام و نمودار منهتن سهم واریانس ژنتیکی هر پنجره به تفکیک صفات در شکل ۱ ارائه شده است.

صفت تعداد سلول‌های بدنی

در شکل a-۱، نتایج ارتباط پنجره‌های SNP با صفت تعداد سلول‌های بدنی (SCS) مشاهده می‌شود. تعداد ۱۱ منطقه ژنومی اصلی مرتبط با تغییرات صفات SCS به‌ترتیب روی کروموزوم‌های ۲، ۱۱، ۱۲، ۱۴، ۱۵، ۱۶، ۱۹، ۲۴ و ۲۶ یافت شدند. این نواحی ژنومی با تعداد متوسط SNP ۲۷، در مجموع حدود ۲۰ درصد واریانس ژنتیکی افزایشی را توصیف می‌کردند. مهمترین پنجره‌های ژنومی (جدول ۲) به‌ترتیب روی کروموزوم شماره ۱۴ با کنترل حدود ۳/۸۵ درصد، کروموزوم شماره ۲۶ با کنترل حدود ۲/۵۵ درصد و کروموزوم شماره ۲۴ (۳۵/۰۲-۳۶/۴۹: ۲۴) با کنترل حدود ۲/۴۷ درصد از واریانس ژنتیکی افزایشی صفت بودند. در یک تحقیق روی گاو هلشتاین، یک ناحیه ژنومی در موقعیت ۳۵-۳۴/۸ مگابازی (معادل ۴۶/۹۲-۴۶/۷۲ سانتی مورگان) روی کروموزوم ۱۴ مؤثر و معنی‌دار گزارش شده است (Daetwyler *et al.*, 2008). در یک تحقیق در سال ۲۰۱۱ روی گاو هلشتاین آمریکا، SNPهای معنی‌داری بر توارث صفت SCS روی کروموزوم ۱۴ در موقعیت‌های ۳۸/۲ و ۳۰/۴ مگابازی گزارش شده است (Cole *et al.*, 2011).



شکل ۱. نمودارهای منهتن ارتباط و درصد واریانس ژنتیکی افزایشی توصیف‌شده حاصل از پنجره‌های ۱/۵ مگابازی SNP. خط سبز رنگ نشان‌دهنده آستانه ۱٪ واریانس ژنتیکی افزایشی می‌باشد. a: صفت SCS. b: صفت تولید شیر. c: صفت مقدار چربی. d: صفت مقدار پروتئین.

Figure 1. Manhattan plot for the proportion of additive genetic variance explained by each of 1.5-Mb genomic windows associated with the SCS (a) milk (b), fat (c) and protein (d) yield. Green line represents informative genomic regions accounting for more than 1% of the additive genetic variance.

در یک تحقیق منطقه ژنومی روی کروموزوم شماره ۲۴ در موقعیت ۶۲-۶۰ مگاباز، حدود ۰/۸۳ درصد از واریانس ژنتیکی افزایشی صفت را با در نظر گرفتن استرس گرمایی توصیف می‌نمود (Sigdel *et al.*, 2019). یکی از اهداف اصلی GWAS، شناسایی ژن‌های کاندیدا در نواحی ژنومی اصلی مرتبط با فنوتیپ صفت است (جدول ۲). منطقه ژنومی روی کروموزوم شماره ۱۴، شامل ۳ ژن کاندید با اسامی *C14H8orf34*، *PREX2* و *SULF1* بود. ژن *SULF1* در جدا نمودن گروه‌های فسفات از هپارین فسفات و ژن *PREX2* در ریخت‌زایی (توسعه) سلول‌های عصبی نقش دارند. ناحیه ژنومی کروموزوم شماره ۲۶ شامل ۴ ژن با اسامی *A1CF*، *ASAH2*، *PRKG1* و *SGMS1* بود که از میان آنها *PRKG1* در تنظیم فعالیت کانال کلسیم و در مسیر فعال‌سازی پلاکت‌ها نقش دارد. ژن‌های خانواده *ACE* روی کروموزوم شماره ۱۹ در تراوش‌های کلیوی و توازن و تعادل الکترولیت ها نقش دارد. مشابه نتایج ما، در یک تحقیق متاآنالیز، ژن *CASP4* (موثر بر پاسخ های التهابی) روی کروموزوم ۱۵، ژن‌های *LYPLAL1* و *TGFB2* روی کروموزوم ۱۶ به‌عنوان ژن‌های کاندیدای احتمالی مؤثر بر SCS مطرح شده است (Chen *et al.*, 2015). این ژن‌ها در محدوده ۲ مگابازی SNPهای معنی‌دار در مقالات مورد بررسی در متاآنالیز بوده‌اند.

جدول ۲. پنجره‌های ۱/۵ مگاباز از SNPهای مجاور با بیش از دو درصد واریانس ژنتیکی افزایشی مؤثر بر صفت SCS و ژن‌های کاندیدا

Table 2. 1.5-Mb genomic windows associated with the SCS, accounting for more than 1% of the additive genetic variance and related candidate genes

Chromosome	Start (Mb)	End (Mb)	Variance explained (%)	Potential Candidate Genes
14	32.00	33.40	3.84	*C14H8orf34, PREX2, SULF1*
26	7.52	9.00	2.54	*A1CF, ASAH2, PRKG1, SGMS1*
24	35.02	36.49	2.47	*ADCYAP1, CETN1, CLUL1, COLEC12, ENOSF1, MIR544B2, ROCK1, THOC1, TYMS, USP14, YES1*

صفت تولید شیر

در شکل b-۱، نتایج ارتباط پنجره‌های SNP با صفت تولید شیر مشاهده می‌شود. تعداد ۶ ناحیه ژنومی ۱/۵ مگاباز از SNPهای مجاور غیرهمپوشان شناسایی شدند که موقعیت کروموزومی و ژن‌های کاندیدای آنها مطابق (جدول ۳) است. این مناطق در مجموع حدود ۸/۸ درصد از واریانس ژنتیکی افزایشی و از میان آنها، ناحیه ژنومی روی کروموزوم شماره ۱۰ بیشترین مقدار و حدود ۲/۰۸ درصد از واریانس و پس از آن ناحیه ژنومی روی کروموزوم شماره ۲۱ حدود ۱/۷۹ درصد از واریانس را توصیف می‌کرد. ژن *ACOT* روی کروموزوم ۱۰ در تعادل اکسیداتیو اسیدهای چرب موثر است. در تحقیق (Meredith *et al.*, 2012) در گاوهای هلشتاین ایرلند، SNPهای معنی‌داری روی کروموزوم ۲۱ در موقعیت ۲۴/۹ مگاباز و در کروموزوم ۲۲ در موقعیت ۲۰/۵ مگاباز گزارش شده است. ناحیه ژنومی کروموزوم شماره ۲۰ شامل ۳ ژن با اسامی *CCDC115*، *HCN1* و *MRPS30* (مؤثر در آپوپتوز یا از بین رفتن سلول)، حدود ۰/۹۸ درصد از واریانس ژنتیک افزایشی صفت را کنترل می‌کنند. این جایگاه در تحقیقات زیادی شناسایی و گزارش شده است. در یک تحقیق در سال ۲۰۱۹ در گاوهای هلشتاین آمریکا، SNPهای معنی‌داری روی کروموزوم ۲۰ در موقعیت‌های ۳۰/۰۳، ۲۹/۳ و ۲۹/۴ مگاباز (Jiang *et al.*, 2019) بر تولید شیر و در گاوهای هلشتاین کانادا SNPهای معنی‌داری روی همین کروموزوم در موقعیت ۳۱/۳-۲۹/۳ مگاباز بر تداوم شیردهی (Do *et al.*, 2017) شناسایی شده است. هر دو تحقیق ژن *HCN1* را به‌عنوان ژن کاندیدای مؤثر بر صفات موردمطالعه گزارش نمودند (Do *et al.*, 2017; Jiang *et al.*, 2019). در گاوهای هلشتاین و نژادهای دیگر شیری منطقه اسکاندیناوی و ایرلند،SNPهای معنی‌داری روی کروموزوم ۲۰ در موقعیت ۳۱/۳-۳۰ مگاباز بر توارث تولید شیر شناسایی و گزارش شده است (Kadri *et al.*, 2015; Meredith *et al.*, 2012).

صفت تولید چربی شیر

بر اساس شکل c-۱، تعداد ۹ ناحیه ژنومی مرتبط با تغییرات صفات تولید چربی شیر به‌ترتیب روی کروموزوم‌های ۲، ۷، ۱۴، ۱۵، ۱۹، ۲۰ و ۲۲ یافت شدند. این نواحی در مجموع حدود ۱۵/۶ درصد از واریانس ژنتیکی را توصیف می‌کنند. ناحیه ژنومی روی کروموزوم شماره ۱۴ (۳۲/۷۰- ۳۱/۲۴ مگاباز) بیشترین مقدار و حدود ۴/۴۴ درصد از واریانس را توصیف می‌نمود. ژن‌های کاندیدای موجود در این ناحیه عبارت از *ARMC1*، *CRH*، *DNAJC5B*، *MTFR1*، *RRS1* و *TRIM55* بودند. *DNAJC5B* عضو خانواده ژن‌های *HSP40* است که در توسعه و زنده‌مانی سلول نقش دارند. مشخص شده که ژن *MTFR1* یک پروتئین میتوکندریایی تولید می‌کند که شکافت میتوکندری را زیاد می‌کند و ژن *CRH* در متابولیسم چربی دخالت دارد (Wathes *et al.*, 2012). در یک تحقیق روی گاوهای هلشتاین کانادا، SNPهای معنی‌داری در این ناحیه ژنومی شناسایی شده است (Nayeri *et al.*, 2016). پس از آن ناحیه ژنومی روی کروموزوم شماره ۲۰ (۱۹/۲۳- ۱۷/۷۴ مگاباز)، حدود ۲ درصد از واریانس ژنتیکی افزایشی را کنترل می‌نماید. در یک تحقیق در گاوهای هلشتاین ایرلند، SNPهای معنی‌داری روی کروموزوم شماره ۷ در موقعیت ۳۸/۷ مگاباز، کروموزوم شماره ۱۹ در موقعیت ۴۸/۷ مگاباز، کروموزوم شماره ۲۰ در موقعیت ۱۸/۹ مگاباز و کروموزوم شماره ۲۲ در موقعیت ۲۵/۳ مگاباز، بر توارث تولید چربی شیر شناسایی و گزارش شده است (Meredith *et al.*, 2012).

جدول ۳. نواحی ژنومی ۱/۵ مگابازی*، میزان واریانس توصیف‌شده و ژن‌های کاندیدای مرتبط با صفت تولید شیر

Table 3. Informative 1.5-Mb genomic windows* associated with the milk yield and related candidate genes

Chromosome	Start (Mb)	End (Mb)	Variance explained (%)	Potential Candidate Genes
10	83.57	85.06	2.08	*ACOT2, ACOT4, ACOT6, DCAF4, DPF3, HEATR4, NUMB, PAPLN, PSEN1, RBM25, RGS6, RIOX1, ZFYVE1*
21	28.42	29.92	1.79	*CHRNA7, OTUD7A, PCSK6, SNRPA1, TARSL2, TJP1, TM2D3*
14	32.07	33.40	1.66	*C14H8orf34, PREX2, SULF1*
22	23.68	25.14	1.37	*CNTN4, CNTN6*
2	108.12	109.60	1.02	*CHPF, GMPPA, INHA, OBSL1, SLC4A3, SPEG, STK11IP, TMEM198*

* پنجره‌های ۱/۵ مگاباز از SNPهای مجاور که آستانه ۱ و بیش از یک درصد واریانس ژنتیکی افزایشی مؤثر بر صفت را گذرانده باشند.

* Genomic windows accounting for equal or more than 1% of the additive genetic variance.

QTL و SNPهایی مؤثر بر مقدار چربی شیر روی کروموزوم شماره ۱۹ در موقعیت ۴۹-۴۷ مگاباز روی گاو شیری کانادا گزارش شده است (Oliveira *et al.*, 2019). در گاوهای هلشتاین آمریکا، مناطق ژنومی و SNPهای معنی‌داری روی کروموزوم۲ در موقعیت ۱۰۹ مگاباز و روی کروموزوم ۱۵ در موقعیت ۶۱/۳ مگاباز (به‌همراه ژن کاندید *MPPED2*)، شناسایی شده است (Jiang *et al.*, 2019).

صفت تولید پروتئین شیر

بر اساس شکل d-۱، تعداد ۹ منطقه ژنومی اصلی (کنترل بیش از ۱ درصد از واریانس ژنتیک افزایشی) مرتبط با تغییرات صفات تولید پروتئین شیر به‌ترتیب روی کروموزوم‌های ۱، ۴، ۶، ۷، ۱۰، ۱۲، ۱۵ و ۱۶ یافت شدند. این نواحی در مجموع حدود ۱۰/۶ درصد از واریانس ژنتیکی را توصیف می‌کنند. پنجره ژنومی روی کروموزوم شماره ۱۲ (۶۹/۷۹- ۶۸/۳۳ مگاباز) بیشترین مقدار و حدود ۱/۳۹ درصد از واریانس را توصیف می‌کرد. این ناحیه شامل ۴ ژن کاندیدا با اسامی *DCT*، *GPC6*، *GPR180* و *TGDS* می‌باشد. در گاو شیری آمریکا، روی کروموزوم ۱۲ در موقعیت ۶۸/۱ مگاباز (Sun *et al.*, 2014) و در گاو هلشتاین آلمان در موقعیت ۶۸/۸ و ۶۸/۹ مگاباز (Strucken *et al.*, 2012) SNPهای معنی‌داری گزارش شده‌اند. پس از آن پنجره ژنومی روی کروموزوم شماره ۱۶ (۲/۸۷- ۱/۳۹ مگاباز)، حدود ۱/۳۸ درصد از واریانس ژنتیکی افزایشی را کنترل می‌نماید. در این ناحیه ژن *PIK3C2B* از خانواده ژن‌های *PIK3* در مسیر متابولیسم کربوهیدرات‌ها مؤثرند. در گاو شیری نژاد مونتبلیارد فرانسه در موقعیت ۳/۲-۱/۸ مگاباز (Sanchez *et al.*, 2019) و در گاو شیری استرالیا در موقعیت‌های ۱/۵ و ۶۸/۹ مگابازی کروموزوم ۱۶ SNPهای معنی‌داری گزارش شده است (Benedet *et al.*, 2019). در خصوص کروموزوم‌های دیگر، (Jiang *et al.*, 2019) SNP معنی‌داری روی کروموزوم ۴ (در موقعیت ۸۱/۷ مگاباز) و روی کروموزوم ۱۰ (در موقعیت ۳۴/۷۵ مگاباز) گزارش نموده‌اند.

مقایسه کلی صفات

با توجه به داده‌ها و اطلاعات موجود و مبتنی بر مقایسه مجموع واریانس ژنتیکی توصیف شده توسط صفات و تعداد نواحی ژنومی اصلی، می‌توان گفت که علاوه بر تأیید توارث پلی‌ژنیک (کمی)، تأثیر ژن‌های عمده اثر بر صفات SCS و مقدار چربی شیر در مقایسه با صفات تولید شیر و مقدار پروتئین شیر بیشتر است. این نتایج متفاوت از نتایج مطالعه روی گاوهای هلشتاین کانادا در مورد صفت تعداد سلول‌های بدنی (SCS) (Oliveira *et al.*, 2019b) بود؛ که این تفاوت می‌تواند ناشی از تفاوت در تعداد SNP در پنجره ژنومی و آستانه معنی‌داری باشد.

نواحی همپوشان و مشترک بین صفات می‌تواند بیانگر ارتباط یا همبستگی ژنتیکی (پلیوتروپی یا پیوستگی) بین صفات باشد. با مقایسه نتایج، می‌توان تعداد ۴ ناحیه مشترک بین بعضی صفات را شناسایی نمود که عبارتند از: ۱- ناحیه ژنومی روی کروموزوم ۱۴ در موقعیت ۳۳-۳۱ مگاباز مؤثر بر صفات تولید شیر، مقدار چربی شیر و تعداد سلول‌های بدنی (SCS)؛ ۲- ناحیه ژنومی روی کروموزوم ۲۲ در موقعیت ۲۵/۱-۲۲/۴ مگاباز مؤثر بر صفات تولید شیر و مقدار چربی شیر؛ ۳- ناحیه ژنومی روی کروموزوم ۷ در موقعیت ۱۱۱/۶۷-۱۱۰/۱۸ مگاباز مؤثر بر صفات

مقدار چربی و پروتئین شیر (این ناحیه حدود ۰/۹۵ درصد از واریانس ژنتیکی صفت تولید شیر را کنترل می‌نماید)؛ ۴- ناحیه ژنومی روی کروموزوم ۱۹ در موقعیت ۴۸/۵-۴۶/۵ مگاباز مؤثر بر صفات مقدار چربی شیر و تعداد سلول‌های بدنی (SCS).

نتیجه‌گیری کلی

در این تحقیق نواحی ژنومی و پنجره‌های SNP با اهمیت و مرتبط با صفات اقتصادی تعداد سلول‌های بدنی (SCS)، تولید شیر، چربی و پروتئین شیر، با استفاده از روش WssGWAS بررسی شد. این تحقیق با هدف شناسایی کروموزوم‌ها، پنجره‌های ژنومی و ژن‌های کاندیدای (شناخته‌شده و جدید) مؤثر انجام شده است که می‌تواند به‌عنوان منبع اطلاعاتی برای انتخاب و اعتبارسنجی SNPها و استفاده آن‌ها در ارزیابی‌های ژنومی گاوهای هلشتاین ایران یا برای مطالعات بعدی پویش ژنومی در نظر گرفته شوند. همچنین ناحیه ژنومی روی کروموزوم ۱۴ و در موقعیت ۳۳-۳۱ مگاباز می‌تواند به‌عنوان یک منطقه ژنومی عمده اثر بر توارث صفات تعداد سلول‌های بدنی و تولید شیر، مقدار چربی شیر در نظر گرفته شده و در صورت تأیید در مطالعات دیگر، می‌تواند پیشنهاد شود.

سپاسگزاری

از مرکز اصلاح نژاد و بهبود تولیدات دامی کشور به‌خاطر در اختیار قراردادن داده‌های داده‌های موردنیاز این تحقیق، تشکر و قدردانی می‌گردد.

REFERENCES

1. Alam, M., Cho, C. I., Choi, T. J., Park, B., Choi, J. G., Choy, Y. H., Lee, S. S. & Cho, K. H. (2015). Estimation of genetic parameters for somatic cell scores of holsteins using multi-trait lactation models in Korea. *Asian-Australasian Journal of Animal Sciences, 28*(3), 303-310.
2. Benedet, A., Ho, P., Xiang, R., Bolormaa, S., De Marchi, M., Goddard, M. & Pryce, J. (2019). The use of mid-infrared spectra to map genes affecting milk composition. *Journal of Dairy Science, 102*(8), 7189-7203.
3. Carlén, E., Strandberg, E. & Roth, A. (2004). Genetic parameters for clinical mastitis, somatic cell score, and production in the first three lactations of Swedish Holstein cows. *Journal of Dairy Science, 87*(9), 3062-3070.
4. Chen, X., Cheng, Z., Zhang, S., Werling, D. & Wathes, D. C. (2015). Combining genome wide association studies and differential gene expression data analyses identifies candidate genes affecting mastitis caused by two different pathogens in the dairy cow. *Open Journal of Animal Sciences, 5*(4), 358-393.
5. Cole, J. B., Wiggans, G. R., Ma, L., Sonstegard, T. S., Lawlor, T. J., Crooker, B. A., Van Tassell, C. P., Yang, J., Wang, S., Matukumalli, L. K. & Da, Y. (2011). Genome-wide association analysis of thirty one production, health, reproduction and body conformation traits in contemporary U.S. Holstein cows. *BMC Genomics, 12*(1), 408.
6. Daetwyler, H. D., Schenkel, F. S., Sargolzaei, M. & Robinson, J. A. B. (2008). A genome scan to detect quantitative trait loci for economically important traits in Holstein cattle using two methods and a dense single nucleotide polymorphism map. *Journal of Dairy Science, 91*(8), 3225-3236.
7. Do, D., Bissonnette, N., Lacasse, P., Miglior, F., Sargolzaei, M., Zhao, X. & Ibeagha-Awemu, E. (2017). Genome-wide association analysis and pathways enrichment for lactation persistency in Canadian Holstein cattle. *Journal of Dairy Science, 100*(3), 1955-1970.
8. Han, Y. & Peñagaricano, F. (2016). Unravelling the genomic architecture of bull fertility in Holstein cattle. *BMC genetics, 17*(1), 143.
9. Heringstad, B., Sehested, E. & Steine, T. (2008). Correlated selection responses in somatic cell count from selection against clinical mastitis. *Journal of Dairy Science, 91*(11), 4437-4439.
10. Iung, L., Petrini, J., Ramírez-Díaz, J., Salvian, M., Rovadoscki, G., Pilonetto, F., Dauria, B., Machado, P., Coutinho, L. & Wiggans, G. (2019). Genome-wide association study for milk production traits in a Brazilian Holstein population. *Journal of Dairy Science, 102*(6), 5305-5314.
11. Jiang, J., Ma, L., Prakapenka, D., VanRaden, P. M., Cole, J. B. & Da, Y. (2019). A large-scale genome-wide association study in US Holstein cattle. *Frontiers in Genetics, 10*, 412.
12. Kadri, N. K., Guldbrandtsen, B., Lund, M. S. & Sahana, G. (2015). Genetic dissection of milk yield traits and mastitis resistance quantitative trait loci on chromosome 20 in dairy cattle. *Journal of Dairy Science, 98*(12), 9015-9025.

13. Meredith, B. K., Kearney, F. J., Finlay, E. K., Bradley, D. G., Fahey, A. G., Berry, D. P. & Lynn, D. J. (2012). Genome-wide associations for milk production and somatic cell score in Holstein-Friesian cattle in Ireland. *BMC Genetics, 13*(1), 21.
14. Misztal, I., Wang, H., Aguilar, I., Legarra, A., Tsuruta, S., Lourenço, D., Fragomeni, B., Zhang, X., Muir, W. & Cheng, H. (2014). GWAS using ssGBLUB. In: Proceedings of *10th World Congress.* Genet. Appl. Livest. Prod., Vancouver, British Columbia, Canada.
15. Nayeri, S., Sargolzaei, M., Abo-Ismail, M. K., May, N., Miller, S. P., Schenkel, F., Moore, S. S. & Stothard, P. (2016). Genome-wide association for milk production and female fertility traits in Canadian dairy Holstein cattle. *BMC Genetics, 17*(1), 75. doi:10.1186/s12863-016-0386-1
16. Oliveira, H. R., Cant, J., Brito, L., Feitosa, F., Chud, T., Fonseca, P., Jamrozik, J., Silva, F., Lourenco, D. & Schenkel, F. (2019a). Genome-wide association for milk production traits and somatic cell score in different lactation stages of Ayrshire, Holstein, and Jersey dairy cattle. *Journal of Dairy Science, 102*(9), 8159-8174.
17. Oliveira, H. R., Lourenco, D. A. L., Masuda, Y., Misztal, I., Tsuruta, S., Jamrozik, J., Brito, L. F., Silva, F. F. & Schenkel, F. S. (2019b). Application of single-step genomic evaluation using multiple-trait random regression test-day models in dairy cattle. *Journal of Dairy Science, 102*(3), 2365-2377.
18. Peñagaricano, F., Weigel, K. A. & Khatib, H. (2012). Genome-wide association study identifies candidate markers for bull fertility in Holstein dairy cattle. *Animal Genetics, 43*, 65-71.
19. Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A., Bender, D., Maller, J., Sklar, P., De Bakker, P. I. & Daly, M. J. (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. *The American Journal of Human Genetics, 81*(3), 559-575.
20. Sahana, G., Guldbrandtsen, B., Thomsen, B., Holm, L. E., Panitz, F., Brøndum, R. F., Bendixen, C. & Lund, M. S. (2014). Genome-wide association study using high-density single nucleotide polymorphism arrays and whole-genome sequences for clinical mastitis traits in dairy cattle. *Journal of Dairy Science, 97*(11), 7258-7275.
21. Sanchez, M.-P., Ramayo-Caldas, Y., Wolf, V., Laithier, C., El Jabri, M., Michenet, A., Boussaha, M., Taussat, S., Fritz, S. & Delacroix-Buchet, A. (2019). Sequence-based GWAS, network and pathway analyses reveal genes co-associated with milk cheese-making properties and milk composition in Montbéliarde cows. *Genetics Selection Evolution, 51*(1), 34.
22. Sargolzaei, M., Chesnais, J. & Schenkel, F. (2011). FImpute-An efficient imputation algorithm for dairy cattle populations. *Journal of Dairy Science, 94*(1), 421.
23. Sigdel, A., Abdollahi-Arpanahi, R., Aguilar, I. & Peñagaricano, F. (2019). Whole Genome Mapping Reveals Novel Genes and Pathways Involved in Milk Production Under Heat Stress in US Holstein Cows. *Frontiers in Genetics, 10*, 928.
24. Strucken, E. M., Bortfeldt, R. H., de Koning, D. J. & Brockmann, G. A. (2012). Genome-wide associations for investigating time-dependent genetic effects for milk production traits in dairy cattle. *Animal Genetic, 43*(4), 375-382.
25. Sun, C., VanRaden, P. M., Cole, J. B. & O'Connell, J. R. (2014). Improvement of prediction ability for genomic selection of dairy cattle by including dominance effects. *PLoS One, 9*(8), e103934.
26. VanRaden, P. M. (2008). Efficient methods to compute genomic predictions. *Journal of Dairy Science, 91*(11), 4414-4423.
27. Wang, H., Misztal, I., Aguilar, I., Legarra, A. & Muir, W. (2012). Genome-wide association mapping including phenotypes from relatives without genotypes. *Genetics Research, 94*(2), 73-83.
28. Wang, S., Dvorkin, D. & Da, Y. (2012). SNPEVG: a graphical tool for GWAS graphing with mouse clicks. *BMC Bioinformatics, 13*(1), 319.
29. Wang, X., Ma, P., Liu, J., Zhang, Q., Zhang, Y., Ding, X., Jiang, L., Wang, Y., Zhang, Y., Sun, D., Zhang, S., Su, G. & Yu, Y. (2015). Genome-wide association study in Chinese Holstein cows reveal two candidate genes for somatic cell score as an indicator for mastitis susceptibility. *BMC genetics, 16*(1), 111.
30. Wathes, D. C., Clempson, A. M. & Pollott, G. E. (2012). Associations between lipid metabolism and fertility in the dairy cow. *Reproduction, Fertility and Development, 25*(1), 48-61.
31. Zhang, X., Lourenco, D., Aguilar, I., Legarra, A. & Misztal, I. (2016). Weighting strategies for single-step genomic BLUP: an iterative approach for accurate calculation of GEBV and GWAS. *Frontiers in genetics, 7*, 151.
32. Zhou, C., Li, C., Cai, W., Liu, S., Yin, H., Shi, S., Zhang, Q. & Zhang, S. (2019). Genome-wide association study for milk protein composition traits in a Chinese Holstein population using a single-step approach. *Frontiers in genetics, 10*(72).