



تحلیل کارایی اطلاعات ژنومی آرایه‌های ژنومی اختصاصی گاومیش در مطالعات ارتباطی کل ژنوم (GWAS) و انتخاب ژنومی (GS) در گاومیش‌های ایران

مهدی مخبر^{۱*}، محمد مرادی شهر بابک^۲، مصطفی صادقی^۳، حسین مرادی شهر بابک^۲ و جان ویلیامز^۳

۱. عضو هیأت علمی گروه علوم دامی دانشگاه ارومیه، ارومیه، ایران

۲. عضو هیأت علمی گروه مهندسی علوم دامی دانشگاه تهران، کرج، ایران

۳. عضو هیأت علمی گروه علوم دام و دامپزشکی دانشگاه ادلاید، استرالیا

* ایمیل نویسنده مسئول: m.mokhber@urmia.ac.ir

چکیده

دقت مطالعات پیوستگی ارتباطی کل ژنوم و موفقیت در برنامه‌های انتخاب ژنومی سطوح عدم تعادل لینکاژی (LD) در طول ژنوم و تدارم فاز LD در بین جمعیت‌ها بستگی دارد. در مطالعه‌ی حاضر میانگین LD بین نشانگرهای SNP مجاور و روند زوال LD با افزایش فاصله‌ی بین نشانگرها در گاومیش‌های ایران محاسبه شده و کارایی آن در مطالعات گاومیش‌های بومی ایران مورد تحلیل قرار گرفت. میانگین آماره‌ی r^2 برای نشانگرهای مجاور گاومیش‌های آذری، خوزستانی و مازندرانی به ترتیب برابر ۰/۲۶، ۰/۲۸ و ۰/۳۲ به دست آمد. میزان LD بین نشانگرهای SNP جفتی با افزایش فاصله‌ی فیزیکی کاهش یافت. این میزان برای نژادهای آذری، خوزستانی و مازندرانی در فاصله‌ی ۱ Kb به ترتیب برابر ۰/۵۴۰، ۰/۵۶۴ و ۰/۵۶۹ بود که با افزایش فاصله‌ی فیزیکی به ۱ Mb به ۰/۰۱۳، ۰/۰۲۶ و ۰/۰۹۴ کاهش یافت. نتایج نشان داد که آرایه‌های ژنومی اختصاصی گاومیش (Axiom Buffalo Genotyping 90K Array) برای مطالعات مبتنی بر آنالیز ژنوم از قبیل GWAS و GS کاربرد دارد. بعلاوه اطلاعات به دست آمده از کاربرد آرایه‌ی مذکور برای ارزیابی ژنومی توأم دو نژاد آذری و خوزستانی مناسب است.

واژه‌های کلیدی: گاومیش، عدم تعادل لینکاژی، آرایه‌های ژنومی ۹۰K اختصاصی گاومیش

مقدمه

توسعه‌ی فناوری‌های توالی‌یابی DNA و دستیابی به اطلاعات صدها هزار جایگاه نشانگر چندشکلی تک نوکلئوتیدی (SNP) امکان ارزیابی کل ژنوم و شناسایی مناطق ژنومی مرتبط با صفات کمی و انتخاب ژنومی را به جای استفاده از برنامه‌ی ارزیابی‌های اصلاح نژاد سنتی فراهم آورده است و از این راه با امکان شناسائی و انتخاب دام‌های جوان سبب کاهش فاصله بین نسل‌ها و در نتیجه تسریع در بهبود ژنتیکی حیوانات اهلی شده است. در روش ارزیابی ژنومی که اثر تمام نشانگرهای ژنتیکی موجود در سراسر ژنوم به‌طور همزمان برآورد می‌شوند، فرض اصل وجود عدم تعادل پیوستگی بین حداقل یک نشانگر با یکی از جایگاه‌های نیم‌مؤثر بر صفت می‌باشد (Meuwissen et al., 2001). عدم تعادل پیوستگی به‌عنوان ارتباط غیر تصادفی میان آلل‌های جایگاه‌های مختلف درون یک جمعیت تعریف می‌شود که به‌طور عمومی ناشی از ارتباط فیزیکی بین جایگاه‌های ژنی است (Khatkar et al., 2008). عدم تعادل پیوستگی بین نشانگرها و جایگاه‌های ژنی کنترل کننده صفات کمی (QTL) در برآورد اثر نشانگری جمعیت مرجع در برآورد ارزش اصلاحی جمعیت تأیید استفاده می‌شود به همین دلیل صحت انتخاب ژنومی تحت تأثیر میزان و پایداری فاز عدم تعادل پیوستگی بین نشانگرها و جایگاه صفت کمی قرار می‌گیرد. هرچه این پایداری بین جمعیت مرجع و تأیید بالاتر باشد صحت انتخاب ژنومی بیشتر می‌شود. در هر صورت با افزایش تعداد نسل بین جمعیت مرجع و تأیید ضرایب رگرسیونی برآورد شده در جمعیت مرجع، ارزیابی جمعیت‌های تأیید با فاصله نسلی دور محدود خواهد شد و محاسبه‌ی بایستی ضرایب رگرسیونی مجدداً محاسبه شود (Meuwissen et al., 2001). شناخت عدم تعادل پیوستگی در سطح ژنوم اطلاعات مهمی در مورد پیدا کردن ساده‌تر ژن‌ها و جهش‌های مرتبط با صفات کمی در بررسی‌های پویا مطالعات پیوستگی در سطح ژنوم (GWAS)، استنباط و توزیع کراسینگ اور، بررسی تنوع و ساختار جمعیتی و تشخیص مناطق ژنومی که تحت انتخاب قرار گرفته‌اند را فراهم می‌کند (Mokry et al., 2014). الگوی عدم تعادل پیوستگی به‌طور گسترده در انواع



حیوانات اهلی از قبیل گاو (Khatkar et al., 2008; McKay et al., 2007) و گومیش (Borquis et al., 2014; Cardoso et al., 2015)، بز (Brito et al., 2015) و گوسفند (Al-Mamun et al., 2015) صورت گرفته است. در مطالعه‌ی حاضر، کارایی اطلاعات ژنومی به دست آمده از آرایه‌های ژنومی اختصاصی گومیش در مطالعات ارتباطی کل ژنوم (GWAS) و انتخاب ژنومی (GS) در گومیش‌های ایران مورد بررسی و تحلیل قرار گرفته است. پژوهش حاضر نخستین مطالعه در زمینه بررسی ساختارهای عدم تعادل لینکاژی در سطح کل ژنوم گومیش‌های آبی کشور می‌باشد. نتایج این پژوهش می‌تواند در تبیین استراتژی‌های آتی مراکز اصلاح نژادی گومیش کشور مفید باشد.

مواد و روش‌ها

مراحل نمونه‌گیری، استخراج DNA توسط شرکت دانش بنیان نواندیش البرز و با همکاری مرکز اصلاح نژاد جهاد کشاورزی کشور و دانشگاه تهران انجام شده و تعیین ژنوتیپ ۴۱۲ راس گومیش (شامل ۲۶۲ رأس آذری، ۱۲۳ راس خوزستانی و ۲۷ رأس مازندرانی) در موسسه‌ی تحقیقاتی پادانوی ایتالیا تکمیل شد. ویرایش داده‌ها با استفاده از نرم‌افزار Plink انجام گرفت (Purcell et al., 2007). بعد از کنترل کیفیت و فیلتراسیون، SNP‌های مربوط به کروموزوم‌های جنسی از آنالیزهای بعدی کنار گذاشته شدند. نسبت کم ژنوتیپ‌های از دست رفته با استفاده از نرم‌افزار BEAGLE ایمپوت گردیدند (Browning & Browning, 2007). سپس هرکدام از جمعیت‌ها با استفاده از نرم‌افزار SnpLD (Sargolzaei M, University of Guelph, Canada) به‌طور مجزا ویرایش شده و مقادیر آماره عدم تعادل پیوستگی r^2 بین جفت بازها در فواصل تعیین شده و با استفاده از فرمول (Hill & Robertson, 1968) به دست آمد. در نهایت میانگین r^2 برای فواصل مختلف تعیین شده و همبستگی r^2 ‌های مربوط به جمعیت‌های مختلف در محیط R محاسبه شد.

نتایج و بحث

پس از انجام مراحل مختلف غربالگری، در نهایت ۶۱۱۶۶، ۶۰۱۹۸ و ۵۷۱۵۱ نشانگر SNP اتوزومی به ترتیب برای نژادهای آذری، خوزستانی و مازندرانی جهت محاسبه r^2 استفاده شد. میانگین فاصله‌ی بین نشانگرهای مجاور برای گومیش‌های آذری، خوزستانی و مازندرانی به ترتیب برابر ۴۱، ۴۲ و ۴۳ کیلوگفت باز بود. با این میزان فاصله برای هر نژاد، میانگین آماره r^2 برای نشانگرهای مجاور گومیش‌های آذری، خوزستانی و مازندرانی به ترتیب برابر ۰/۲۶، ۰/۲۸ و ۰/۳۲ به دست آمد. این مقادیر با میزان r^2 گزارش شده برای گومیش‌های شیری برزیل که با استفاده از آرایه‌های ژنومی اختصاصی گومیش به دست آمده بود (۰/۲۹)، مطابقت داشت (Cardoso et al., 2015). همچنین این مقادیر با نتایج مطالعه‌ی Borquis et al. (2014) (۲۰۱۴) مطابقت داشت. در این پژوهش میانگین r^2 برای نشانگرهای SNP مجاور برابر ۰/۲۹ (۰/۳۴ - ۰/۰۷) گزارش شد، البته در این پژوهش گومیش‌های برزیل با استفاده از آرایه‌های با تراکم بالای گاو (۷۰۰k) تعیین ژنوتیپ شده بودند (Borquis et al., 2014).

میانگین زوال LD با افزایش فاصله برای تمام نژادهای مورد بررسی کاهش یافت. شدت این کاهش در فواصل کمتر از ۳۰۰ کیلوگفت باز، بالا بود و در فواصل بالاتر از ۱ مگا جفت باز شدت کاهش بسیار پایین بود. به طوری که، نتایج نشان داد که میزان LD بین نشانگرهای SNP جفتی با افزایش فاصله‌ی فیزیکی بین نشانگرها از ۰/۵۴، ۰/۵۶۴ و ۰/۵۶۹ در فاصله‌ی ۱ کیلوگفت باز به ۰/۰۱۳، ۰/۰۲۶ و ۰/۰۹۴ در فاصله‌ی ۱ مگا جفت باز برای نژادهای آذری، خوزستانی و مازندرانی کاهش یافت. بنابراین در مقایسه‌ی LD بین جمعیت‌ها، در تمامی فواصل بالاترین میزان پیوستگی مربوط به نژاد مازندرانی و پایین‌ترین میزان مربوط به نژاد آذری بود. مقایسه‌ی زوال LD نژادهای مورد بررسی نشان داد که الگوی زوال LD دو نژاد آذری و خوزستانی در مقایسه با نژاد مازندرانی به یکدیگر شباهت بیشتری دارند. تفاوت کم الگوی زوال LD بین دو نژاد می‌تواند در اثر منشاء مشترک این دو جمعیت و یا اختلاط ژنتیکی آن‌ها باشد (Brito et al., 2015). با توجه به شرایط پرورشی نژاد مازندرانی که بخش عمده‌ای از جمعیت آن در منطقه محافظت شده‌ی میانکاله پرورش داده می‌شود که تبادل ماده‌ی ژنتیکی در آن نسبتاً پایین است، الگوی LD کاملاً متفاوت نژاد مازندرانی می‌تواند مربوط به محیط ایزوله‌ی این نژاد از نژادهای دیگر کشور باشد.

همبستگی ارزش‌های r^2 نشان دهنده‌ی درجه‌ی ارتباط ژنتیکی بین نژادها است (Daetwyler et al., 2010). مقادیر عددی تداوم فاز LD بین سه نژاد گومیش کشور برای فواصل ژنتیکی کوتاه‌تر با افزایش فاصله این مقادیر کاهش نشان داد. به طوری که تداوم فاز LD



برای فواصل مارکری کمتر از 200 Kb برای همه‌ی نژادهای مورد بررسی بالاتر از 0/9 تعیین شد و این میزان در فواصل بالاتر از 1 مگا جفت باز از 0/71 بین نژادهای آذری و خوزستانی تا 0/1 بین نژادهای خوزستانی و مازندرانی متغییر بود. نتایج نشان می‌دهد که آرایه-های ژنومی اختصاصی گاومیش (Axiom Buffalo Genotyping 90K Array) برای مطالعات مبتنی بر آنالیز ژنوم از قبیل GWAS و GS کاربرد دارد. بعلاوه آرایه‌ی مذکور برای ارزیابی ژنومی توأم دو نژاد آذری و خوزستانی مناسب است.

منابع

- Al-Mamun, H.A., Clark, S., Kwan, P. and Gondro, C. 2015. Genome-wide linkage disequilibrium and genetic diversity in five populations of Australian domestic sheep. *Genetics Selection Evolution*, 47(1), 1.
- Borquis, R.R.A., Baldi, F., de Camargo, G.M.F., Cardoso, D.F., Santos, D.J.A., Lugo, N.H., Sargolzaei, M., Schenkel, F.S., Al-buquerque, L.G. and Tonhati, H. 2014. Water buffalo genome characterization by the Illumina BovineHD BeadChip. In: *Genet Mol Res*, 13: 4202-4215.
- Brito, L.F., Jafarikia, M., Grossi, D.A., Kijas, J.W., Porto-Neto, L.R., Ventura, R.V., Salgorzaei, M. and Schenkel, F.S. 2015. Characterization of linkage disequilibrium, consistency of gametic phase and admixture in Australian and Canadian goats. *BMC genetics*, 16:67.
- Cardoso, D., Aspilcueta-Borquis, A., Santos, D., Hurtado-Lugo, H.N., De Camargo, G., Scalez, D., De Albuquerque, L., Tonhati, H. 2015. Study of Linkage Disquisition in Brazilian Dairy Buffaloes. *Proceedings, 10th World Congress of Genetics Applied to Livestock Production*.
- Daetwyler, H.D., Pong-Wong, R., Villanueva, B. and Woolliams, J.A. 2010. The impact of genetic architecture on genome-wide evaluation methods. *Genetics*, 185(3):1021-1031.
- Hill, W. G. and Robertson, A. 1968. Linkage disequilibrium in finite populations. *Theoretical and Applied Genetics*, 38: 226–231.
- Khatkar, M.S., Nicholas, F.W., Collins, A.R., Zenger, K.R., Cavanagh, J.A., Barris, W., Schnabel, R.D., Taylor, J.F. and Raadsma, H.W. 2008. Extent of genome-wide linkage disequilibrium in Australian Holstein-Friesian cattle based on a high-density SNP panel. *BMC genomics*, 9(1):187.
- McKay, S.D., Schnabel, R.D., Murdoch, B.M. and et. al. 2008. An assessment of population structure in eight breeds of cattle using a whole genome SNP panel. *BMC Genetics*, 9:37.
- Mokry, F., Buzanskas, M., de- Alvarenga -Mudadu, M. and et al. 2014. Linkage disequilibrium and haplotype block structure in a composite beef cattle breed. *BMC Genomics*, 15(Suppl 7): S6.
- Purcell, S., Neale, B., Todd-Brown, K. and et al. (2007). PLINK: a toolset for whole-genome association and population-based linkage analysis. *Am J Hum Genet*, 81: 559–575.
- Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A. and et al. 2007. PLINK: a tool set for whole-genome association and population-based linkage analyses. *American journal of human genetics*, 81(3):559-575.
- Sargolzaei, M., Iwaisaki, H. and Colleau, J.J. 2006. CFC – A tool for monitoring genetic diversity, Comm 27-28 in proceeding of the 8th World conger. *Genet. Appl. Livest. Prod Belo Horizonte, Brazil*.
- The R Project for Statistical Computing: Free software environment for statistical computing and graphics. [http:// www.r-project.org/](http://www.r-project.org/).



Survey of application of buffalo genomic arrays in genome wide association study (GWAS) and genomic selection (GS) in Iranian Buffalo

Mahdi Mokhber^{1*}, Mohammad Moradi-Shahrbabak², Mostafa Sadeghi², Hossein Moradi-Shahrbabak² and John Williams³

1. Department of Animal Science, Faculty of Agricultural Science, Urmia University, Urmia, Iran.

2. Department of Animal Science, Faculty of Agricultural Science and Engineering, University College of Agriculture and Natural Resources (UTCAN), University of Tehran, Karaj, Iran.

3. Davies Research Centre, School of Animal and Veterinary Sciences, University of Adelaide, Australia.

* Corresponding Author's Email: m.mokhber@urmia.ac.ir

Abstract

Accuracy of genome-wide association studies (GWAS) and successful application of genomic selection (GS) is depend on level of linkage disequilibrium (LD) across the genome and the persistence of the LD phase across populations. In the present research, LD between adjacent SNPs and LD decay according to physical distance between SNPs were calculated in three Iranian buffaloes. Average r^2 between adjacent SNP pairs across all chromosomes was 0.26, 0.28 and 0.32 for Azeri, Khuzestani and Mazandarani breeds, respectively. The LD between bi-allelic SNPs decreased with increasing physical distance between markers from 0.540, 0.564 and 0.569 at marker distance 1 kb to 0.013, 0.026 and 0.094 at marker distance 1Mb for AZI, KHU and MAZ, respectively. The persistence of LD phase between all populations with marker distances less than 200 Kb was higher than 0.9 and decreased with increasing marker distance. These results indicate that SNP dense is enough for GWAS and implementation of GS. For multi-breed genomic evaluation, used SNP panels is suitable for AZI-KHU buffalo breeds.

Keywords: Buffalo, Linkage disequilibrium, Axiom Buffalo Genotyping 90K Array