



Efficiency of genotyping by sequencing technology and its comparison with single nucleotide polymorphism array in an F₂ chickens population

Narjes Gorgani Firouzjah ¹, Rasoul Vaez Torshizi ^{2✉}, Alireza Ehsani ³, Ali Akbar Masoudi ⁴, Peymane Davoodi ⁵, Hossein Bani Saadat ⁶

1. Department of Animal Science, Faculty of Agriculture, Tarbiat Modares University, Tehran, Iran. E-mail: narges.gorgani@yahoo.com
2. Corresponding Author, Department of Animal Science, Faculty of Agriculture, Tarbiat Modares University, Tehran, Iran. E-mail: rasoult@modares.ac.ir
3. Department of Animal Science, Faculty of Agriculture, Tarbiat Modares University, Tehran, Iran. E-mail: alireza.ehsani@modares.ac.ir
4. Department of Animal Science, Faculty of Agriculture, Tarbiat Modares University, Tehran, Iran. E-mail: masoudia@modares.ac.ir
5. Department of Animal Science, Faculty of Agriculture, Tarbiat Modares University, Tehran, Iran. E-mail: pey.davoodi@gmail.com
6. Department of Animal Science, Faculty of Agriculture, Tarbiat Modares University, Tehran, Iran. E-mail: h.bany1990@gmail.com

Article Info

Article type:

Research Article

Article history:

Received: 19 February 2024

Received in revised form: 21 April 2024

Accepted: 30 April 2024

Published online: 21 December 2024

Keywords:

Genotyping by Sequencing (GBS), Single Nucleotide Polymorphism (SNP) Array, F₂ chicken population.

ABSTRACT

For researches, there is a wide variety of available technologies to collect molecular information in the field of chicken genomics. These technologies, which consist of single-nucleotide polymorphism (SNP) arrays and genotyping by sequencing (GBS), depending on the goals of the study, can have different applications. The aim of this study was to compare the results of markers genotyped by two technologies, namely, 60 K SNP BeadChip and genotyping by sequencing, using data collected on F₂ chicken population resulting from a reciprocal crosses between a native bird of Urmia and a fast-growing commercial Arian line. In genotyping by GBS, 882,918 SNPs were identified, of which 815,613 SNPs (92.40%) were located on chromosomes 1 to 28. In 60 K SNP array, the number of SNPs for each sample were 51347, which were distributed on chromosomes 1 to 28. The GBS data identified more markers and haplotype blocks than the 60 K SNP array. The rate of linkage disequilibrium (LD) in the physical distances of 10, 100 and 1000 kbp in GBS was less than that of SNP array. The large variety of SNPs in the GBS resulted in a uniform population structures and kinship. Also, in addition to the high performance for identifying single nucleotide polymorphisms, the technology of GBS also reduced the costs of the genotyping for each sample, therefore, it seems that the use of genotyping by sequencing technology could be a suitable alternative method to the 60 K SNP BeadChip array technology for genome-wide association studies in chicken population.

Cite this article: Gorgani Firouzjeh, N. Vaez Torshizi, Ehsani, A. R., Masoudi, A., Davoodi, P. & Bani Saadat, H. (2024). Efficiency of genotyping by sequencing technology and its comparison with single nucleotide polymorphism array in an F₂ chickens population. *Iranian Journal of Animal Science*, 55 (4), 693-711. DOI: <https://doi.org/10.22059/ijas.2024.372827.653999>



© The Author(s).

DOI: <https://doi.org/10.22059/ijas.2024.372827.653999>

Publisher: The University of Tehran Press.

Extended Abstract

Introduction

Nowadays, researchers in the field of agricultural genomics have a wide variety of technologies to collect molecular genetic information. These technologies, which consist of 60k single nucleotide polymorphism (SNP) BeadChip array and Genotyping by Sequencing (GBS), depending on the goals of the researcher's study, can have different applications. The high cost of single nucleotide polymorphism array technology per sample

limits their use in research applications, especially for the large-sized populations. The aim of this study was to evaluate the efficiency of the data obtained from the genotyping by sequencing (GBS) technology in an F₂ chickens population and compare the results with the data obtained from a 60k SNP BeadChip array.

Materials & Methods

A F1 chicken population was created from a reciprocal crosses between a fast-growing Arian commercial broiler line and an Urmia native fowls. The F1 population were intercrossed randomly to produce 312 F₂ chickens in eight half-sib families from five hatches. The Genomic DNA of these birds were extracted from blood samples and then genotyped by a commercial Illumina chicken SNP BeadChip panel in collaboration with Aarhus University, Denmark. Also, 368 birds, including 312 birds from the F₂ generation, 37 birds from the F1 generation, and 19 birds from the base population generation were genotyped using genotyping by sequencing technology by The Elshire Group Ltd., New Zealand.

Results

Using genotyping by sequencing technology, 882,918 SNPs were detected, of which 815,613 SNP were located on chromosomes 1 to 28 and 67,305 SNPs on autosomal chromosomes 30 to 33 and sex chromosomes Z and W. For 60k SNP array, 51347 SNP markers, located on chromosomes 1 to 28, were genotyped for each sample. The analysis of the missing genotype in 60k SNP array technology indicated that almost all the birds had all single nucleotide polymorphisms. These results for the missing genotype in the genotyping by sequencing technology were about 54%, which could be reduced by using imputation of missing data. The rate of linkage disequilibrium (LD) in the physical distances of 10 kbp, 100 kbp and 1000 kbp in GBS technology was less than that of SNP array. By increasing the physical distances, the rate of LD decreased. The number of haplotype blocks in the genotyping by sequencing technology (167396) were substantially higher than the 60 kb single nucleotide polymorphism array (6783).

Conclusion

In the present study, the diversity of SNPs in the genotyping by sequencing data was very high, creates an acceptable homogeneity of population structure and kinship structures. Since haplotypes blocks have a higher linkage disequilibrium with quantitative trait loci (QTL), it is expected that the genotyping by sequencing technology improve genome-wide association studies (GWAS) due to its higher haplotype blocks. Also, in addition to the high performance in identifying single nucleotide polymorphisms, the technology of GBS also reduced the costs of the genotyping for each sample, therefore, it seems that the use of genotyping by sequencing technology could be a suitable alternative method to the 60k SNP BeadChip array technology for genome-wide association studies in chicken population.



کار آیی فن آوری تعیین ژنوتیپ از طریق توالی یابی و مقایسه آن با آرایه‌ی چندشکلی تک‌نوکلئوتیدی در یک جمعیت مرغ نسل دو

نرجس گرگانی فیروزجاه^۱ | رسول واعظ ترشیزی^۲ | علیرضا احسانی^۳ | مسعودی علی اکبر^۴ | پیمان‌ه داودی^۵ | حسین بانی سعادت^۶

۱. گروه علوم دامی، دانشکده کشاورزی، دانشگاه تربیت مدرس، تهران، ایران. رایانامه: narges.gorgani@yahoo.com

۲. نویسنده مسئول، گروه علوم دامی، دانشکده کشاورزی، دانشگاه تربیت مدرس، تهران، ایران. رایانامه: rasoult@modares.ac.ir

۳. گروه علوم دامی، دانشکده کشاورزی، دانشگاه تربیت مدرس، تهران، ایران. رایانامه: alireza.ehsani@modares.ac.ir

۴. گروه علوم دامی، دانشکده کشاورزی، دانشگاه تربیت مدرس، تهران، ایران. رایانامه: masoudia@modares.ac.ir

۵. گروه علوم دامی، دانشکده کشاورزی، دانشگاه تربیت مدرس، تهران، ایران. رایانامه: pey.davoodi@gmail.com

۶. گروه علوم دامی، دانشکده کشاورزی، دانشگاه تربیت مدرس، تهران، ایران. رایانامه: h.bany1990@gmail.com

چکیده

اطلاعات مقاله

پژوهش‌گران حوزه ژنوم در بخش طيور انواع گسترده‌ای از فناوری‌ها را برای جمع‌آوری اطلاعات ژنتیکی در اختیار دارند. این فناوری‌ها که عبارت از آرایه‌های چندشکلی تک‌نوکلئوتیدی مرغ و تعیین ژنوتیپ مبتنی بر توالی‌یابی هستند، بسته به اهداف مطالعه می‌تواند استفاده‌های مختلفی داشته باشند. هدف از مطالعه حاضر، مقایسه‌ی نتایج حاصل از توالی‌یابی یک جمعیت F2 حاصل از تلاقی دو طرفه پرنده‌های بومی ارومیه و یک لاین تجاری گوشتی آرین با استفاده از فن‌آوری‌های توالی‌یابی و آرایه K60 بود. در فن‌آوری توالی‌یابی، ۸۸۲۹۱۸ نشان‌گر شناسایی شد که ۸۱۵۶۱۳ آن‌ها (۹۲/۴۰ درصد) مربوط به کروموزوم‌های ۱ تا ۲۸ بودند. در این فن‌آوری، کروموزوم ۱ بیشترین و کروموزوم W کم‌ترین تعداد نشان‌گر را داشتند. در فن‌آوری آرایه، تعداد ۵۱۳۴۷ نشان‌گر بر روی کروموزوم‌های ۱ تا ۲۸ پراکنده بودند و کروموزوم ۱ بیشترین و کروموزوم ۱۶ کم‌ترین تعداد نشان‌گر را نشان داد. داده‌های حاصل از فن‌آوری توالی‌یابی، تعداد نشان‌گر و بلوک‌های هاپلوتایپی بیشتری نسبت به آرایه‌ی ۶۰ کیلوبازی شناسایی کرد. میزان عدم تعادل پیوستگی در فواصل فیزیکی ۱۰ کیلوباز، ۱۰۰ کیلوباز و ۱۰۰۰ کیلوباز در فن‌آوری توالی‌یابی نسبت به آرایه‌ی ۶۰ کیلوباز کم‌تر بود. تنوع زیاد نشان‌گرها در داده‌های ژنوتیپ‌کردن از طریق توالی‌یابی موجب شد ساختارهای جمعیتی و خویشاوندی یکنواختی ایجاد شود. همچنین، علاوه بر عملکرد بالا در شناسایی نشان‌گرها، فن‌آوری ژنوتیپ‌کردن از طریق توالی‌یابی، هزینه‌های تعیین ژنوتیپ به‌ازای هر نمونه را نیز کاهش داد، بنابراین، به نظر می‌رسد استفاده از فن‌آوری ژنوتیپ‌کردن از طریق توالی‌یابی بتواند جایگزین مناسبی برای فن‌آوری آرایه‌ی K60 در مطالعات پویا ژنوم در جمعیت‌های مرغ شود.

نوع مقاله:

مقاله پژوهشی

تاریخ دریافت: ۱۴۰۲/۱۱/۳۰

تاریخ بازنگری: ۱۴۰۳/۰۲/۰۲

تاریخ پذیرش: ۱۴۰۳/۰۲/۱۱

تاریخ انتشار: ۱۴۰۳/۱۰/۰۱

کلیدواژه‌ها:

ژنوتیپ‌کردن از طریق توالی‌یابی، چندشکلی تک‌نوکلئوتیدی، مرغ های جمعیت F2.

استناد: گرگانی فیروزجاه، نرجس؛ واعظ ترشیزی، رسول؛ احسانی، علیرضا؛ مسعودی، علی اکبر؛ داودی، پیمان‌ه و بانی سعادت، حسین (۱۴۰۳). کارایی فن‌آوری تعیین ژنوتیپ از طریق توالی‌یابی و مقایسه آن با آرایه‌ی چندشکلی تک‌نوکلئوتیدی در یک جمعیت مرغ نسل دو. *نشریه علوم دامی ایران*، ۵۵ (۴)، ۶۹۳-۷۱۱. DOI: <https://doi.org/10.22059/ijas.2024.372827.653999>

ناشر: مؤسسه انشارات دانشگاه تهران.



© نویسندگان.

DOI: <https://doi.org/10.22059/ijas.2024.372827.653999>

مقدمه و پیشینه پژوهش

نیاز روز افزون جامعه به فرآورده‌های پروتئینی با منشأ حیوانی، استفاده از روش‌های نوین جهت افزایش محصول را ضروری کرده است. در میان این فرآورده‌ها، گوشت مرغ در تغذیه‌ی بشر از اهمیت ویژه‌ای برخوردار است. عمده پیشرفت ژنتیکی صفات مهم اقتصادی در جوجه‌های گوشتی با استفاده از روش‌های کلاسیک انتخاب و بر اساس اطلاعات فنوتیپی و شجره حاصل شده است. برای افزایش تولید علاوه بر روش‌های نوین مدیریت، بهداشت و تغذیه، استفاده از روش‌های علمی اصلاح‌نژاد به همراه تکنیک‌های مولکولی ضروری است. امروزه انتخاب به کمک نشان‌گرهای مولکولی برای افزایش بازده انتخاب و بهبود عملکرد تولیدی مدنظر است و بنابراین تلفیقی از روش‌های مرسوم انتخاب و روش‌های جدید مولکولی در آینده اصلاح‌نژاد طیور مورد توجه قرار گرفته است (Emara and Kim, 2003).

امروزه پژوهش‌گران حوزه ژنومی کشاورزی انواع گسترده‌ای از نوآوری‌ها را برای جمع‌آوری اطلاعات ژنتیکی در اختیار دارند. از جمله‌ی آن‌ها فرآیند تعیین ژنوتیپ است که کاربردهای زیادی در کشاورزی، از جمله غربالگری جمعیت‌های اصلاحی دارد. با غربالگری ژنتیکی می‌توان تصمیمات آگاهانه‌تری اتخاذ و بازگشت سرمایه را تسریع کرد. بکارگیری ابزارهای تعیین ژنوتیپ که هزینه کمتری به ازای هر نمونه دارند، امکان غربالگری ژنتیکی متداول را در جمعیت‌های بزرگ فراهم می‌کند که در این صورت، با بازگشت سریع سرمایه، هزینه‌ی استفاده از فن‌آوری نیز جبران می‌شود (Illumina, 2014). اگرچه ابداع آرایه‌های چندشکلی تک‌نوکلئوتیدی با تراکم بالا، بخشی از مشکلات مربوط به فرآیند تعیین ژنوتیپ را حل کرده، اما مشکلات مربوط به فرآیند شناسایی نشان‌گرها همچنان باقی مانده است (Davey et al., 2011). روش‌های توالی‌یابی نسل جدید (NGS¹) به طور گسترده برای شناسایی نشان‌گرهای ژنتیکی در سطح ژنوم و تعیین ژنوتیپ استفاده می‌شوند. در سال‌های اخیر، پیشرفت‌های بسیار زیادی در فن‌آوری روش‌های توالی‌یابی نسل جدید ایجاد شده و هزینه آن به طور چشم‌گیری کاهش یافته است. با تکامل سریع فن‌آوری‌های توالی‌یابی، روش‌های تعیین ژنوتیپ مبتنی بر توالی‌یابی (NGG²) با هزینه‌های کم‌تر توسعه یافته و مطالعات در سطح ژنوم بر روی حیوانات، گیاهان و سیستم‌های بیولوژی، که پیش از آن امکان‌پذیر نبود، را میسر ساخته است. مطالعه‌ی Elshire و همکاران (۲۰۱۱)، فن‌آوری ژنوتیپ کردن از طریق توالی‌یابی (GBS³)، که مبتنی بر روش‌های توالی‌یابی نسل جدید و بکارگیری آنزیم‌های محدودگر برای هدف‌گیری بخش‌هایی از ژنوم بود، را توسعه داده‌اند. یکی از مزیت‌های این فن‌آوری، کم‌تر بودن هزینه تعیین ژنوتیپ کردن در مقایسه با آرایه‌های چندشکلی تک‌نوکلئوتیدی است. راه‌اندازی ژنوتیپ کردن از طریق توالی‌یابی برای هر گونه، شامل انتخاب آنزیم یا آنزیم‌های محدودگر مناسب برای تولید کتابخانه ژنومی و بهینه‌سازی نسبت DNA نمونه به آدپتورهای مورد استفاده در تکثیر قطعات ژنومی است. مراحل تهیه آرایه‌های چندشکلی تک‌نوکلئوتیدی شامل مراحل پیچیده شناسایی اولیه چندشکلی تک‌نوکلئوتیدی، ارزیابی تنوع و انتخاب چندشکلی تک‌نوکلئوتیدی است. بعد از این مراحل، مجموعه‌ای از نشان‌گرهای چندشکلی تک‌نوکلئوتیدی با کیفیت بالا بر روی پلات فرم‌های تعیین ژنوتیپ بسیار مترکم قرار می‌گیرند (Glaubits et al., 2014).

در فن‌آوری ژنوتیپ کردن از طریق توالی‌یابی، شناسایی و تعیین ژنوتیپ نشان‌گرها به طور همزمان انجام می‌شود، بنابراین ارزیابی ناشی از نمونه‌گیری کم‌تر است (Thomson, 2014; Davey et al., 2011; Glaubitz et al., 2014). ارزیابی ناشی از نمونه‌گیری، هنگامی ایجاد می‌شود که داده‌های نشان‌گری، حاصل از یک نمونه تصادفی از چند شکلی‌ها در جمعیت مورد مطالعه نباشد. بنابراین، استفاده از پانل‌های طراحی شده از نمونه‌های محدود برای شناسایی و انتخاب چندشکلی‌های تک‌نوکلئوتیدی در تهیه آرایه‌های چندشکلی‌های تک‌نوکلئوتیدی، این نوع ارزیابی را در نتایج حاصل از این روش ایجاد می‌کند (Albrechtsen et al., 2010; Heslot et al., 2013). برای مثال، تراشه‌ی ۶۰K مرغ شرکت ایلومینا، که به صورت تجاری

1Next-Generation Sequencing (NGS)

2Next-Generation Sequencing-Based Genotyping

3Genotyping by Sequencing

4Ascertainment Bias

در دسترس است، با استفاده از دو سویه‌ی گوشتی و دو سویه‌ی تخمگذار طراحی شده است (Liao et al., 2015). از آن جایی که این نشان‌گرها مخصوص جمعیتی هستند که در آن شناسایی شده‌اند، بنابراین تعیین ژنوتیپ جمعیت‌های جدید، به سمت آلل‌های موجود در جمعیت مرجع اریب خواهد بود (Davey et al., 2011)، که ممکن است منجر به اریبی در برآورد پارامترهای ژنتیک جمعیت، توزیع فراوانی آللی و عدم تعادل ناشی از پیوستگی (LD) شود (Heslot et al., 2013). این مشکل به ویژه در مطالعه جمعیت‌های نژادهای بومی به‌وضوح مشاهده می‌شود. در فن آوری ژنوتیپ کردن از طریق توالی یابی نیز ممکن است اریبی جزئی وجود داشته باشد که ناشی از فاکتورهای مانند هم‌ترازی با ژنوم مرجع است (Glaubitz et al., 2014).

در مطالعه حاضر، هزینه ژنوتیپ کردن از طریق توالی یابی به ازای هر پرند به‌طور قابل ملاحظه‌ای کم‌تر از هزینه برای فن آوری آرایه ۶۰K بود. پایین بودن هزینه فن آوری ژنوتیپ کردن از طریق توالی یابی در مقایسه با فن آوری‌های آرایه ۶۰K و ۶۰۰K در مطالعات دیگر محققین نیز گزارش شده است (Pertille و همکاران، ۲۰۱۶؛ Negro و همکاران، ۲۰۱۹). با وجود این که فن آوری تعیین ژنوتیپ از طریق توالی یابی، ژنوتیپ کردن انواع گونه‌ها را با هزینه‌ی کم امکان‌پذیر کرده است، اما اغلب سطح بالایی از داده‌های گم‌شده حاصل از آن، منجر به از دست رفتن اطلاعات زیادی می‌شود (Elshire et al., 2011) که ضعف اصلی فن آوری ژنوتیپ کردن از طریق توالی یابی است. جانهی راه حل مناسبی برای افزایش صحت اطلاعات حاصل از فن آوری ژنوتیپ کردن از طریق توالی یابی است (Huang et al., 2014). ضعف دیگر فن آوری ژنوتیپ کردن از طریق توالی یابی، عدم یکنواختی در بازنمایی نمونه‌ها است. به عبارت دیگر، این فناوری از نظر تعداد خوانش، عمق خوانش و کیفیت داده‌ها، از نمونه‌ای به نمونه دیگر متفاوت است (Elshire et al., 2011). مشکل دیگر فن آوری‌های مبتنی بر توالی یابی نسل جدید، ابهام در تشخیص هتروزیگوت‌ها است که گاهی اوقات ممکن است به اشتباه هموزیگوت خوانده شوند. این خطا ممکن است ناشی از پایین بودن عمق خوانش و یا وقوع جهش در جایگاه برش آنزیم‌های محدودگر باشد. در داده‌های حاصل از فن آوری ژنوتیپ کردن از طریق توالی یابی، علاوه بر نشان‌گرهای چندشکلی‌های تک‌نوکلئوتیدی دو آللی، واریانت‌های دیگری مانند جهش‌های حذف و اضافه و ریزماهورها نیز وجود دارند که بسیاری از نرم‌افزارهای رایج قادر به تحلیل این گونه واریانت‌ها نیستند. علاوه بر این، روش‌های تحلیل داده‌های ژنوتیپ کردن از طریق توالی یابی اگرچه به‌طور مداوم در حال پیشرفت هستند، هنوز به اندازه روش‌های حاصل از آرایه‌ها کارآمد نیستند (Illumina, 2014). اگرچه امروزه آرایه‌های چندشکلی تک‌نوکلئوتیدی تنها روش غربالگری جمعیت‌ها نیستند، اما همچنان به عنوان یک گزینه بسیار مناسب محسوب می‌شوند.

با وجود محدودیت‌هایی که در فن آوری ژنوتیپ کردن از طریق توالی یابی وجود دارد، همزمان با افزایش کمیت و کیفیت اطلاعات حاصل از توالی یابی و نیز کاهش هزینه به ازای هر نمونه، ژنوتیپ کردن از طریق توالی یابی به یک فن آوری متداول و پرتعداد در پژوهش‌های حوزه ژنومیک کشاورزی تبدیل شده است (He et al., 2014). واریانت‌های بدست آمده از ژنوتیپ کردن از طریق توالی یابی، برای مطالعات پویش کل ژنوم، تهیه نقشه‌های ژنتیکی و هاپلوتیپ، آنالیز پیوستگی، تنوع ژنتیکی و فیلوژنی بکار می‌روند (He et al., 2014). ژنوتیپ کردن از طریق توالی یابی در گونه‌های گیاهی متنوع برای اهداف مختلف مانند تهیه نقشه ژنتیکی در گراس علفی (Fiedler et al., 2015)، مطالعات تکاملی در جو (Mascher et al., 2016) و مطالعات پویش ژنوم در سورگوم (Sukumaran et al., 2016) مورد استفاده قرار گرفته است. Huang و همکاران (۲۰۱۴) کاربرد ژنوتیپ کردن از طریق توالی یابی را در گیاه یولاف هگزاپلوئید بررسی کردند و ۴۵۱۱۷ جایگاه بر روی نقشه توافقی آن شناسایی کردند. در مطالعات پویش کل ژنوم، حداقل تراکم یک نشان‌گر در هر ۲ تا ۲/۸ سانتی‌مورگان نیاز است. نتایج این محققین نشان داد که نشان‌گرهای حاصل از فن آوری ژنوتیپ کردن از طریق توالی یابی در اغلب نواحی کروموزومی تراکم مورد نیاز را داشتند. براین اساس پیشنهاد شد که این روش، مفید و قدرتمند برای انجام مطالعات ژنومی در یولاف است.

اگرچه فن آوری ژنوتیپ کردن از طریق توالی‌یابی در گونه‌های گیاهی نسبت به گونه‌های جانوری بیشتر مورد استفاده قرار گرفته است اما در سال‌های اخیر این فن آوری برای مطالعات گونه‌های جانوری نیز مورد استفاده قرار گرفته است. در مطالعه‌ای، Donato و همکاران (۲۰۱۳) با استفاده از فن آوری ژنوتیپ کردن از طریق توالی‌یابی، تعداد ۴۷ گاو از هفت نژاد مختلف را تعیین ژنوتیپ کردند. در مجموع ۶۳۶۹۷ چندشکلی تک‌نوکلئوتیدی شناسایی شد. بعد از حذف چندشکلی‌های تک‌نوکلئوتیدی که نرخ خوانش کم‌تر از ۷۰ درصد داشتند، ۵۱۴۱۴ چندشکلی تک‌نوکلئوتیدی بر روی کروموزوم‌های اتوزومی با میانگین فاصله ۴۸/۱ کیلو جفت باز، ۱۱۴۳ چندشکلی تک‌نوکلئوتیدی بر روی کروموزوم X با میانگین فاصله ۱۳۰/۳ کیلو جفت باز و ۱۹۱ چندشکلی تک‌نوکلئوتیدی با موقعیت ژنومی نامشخص وجود داشت. میانگین تراکم چندشکلی تک‌نوکلئوتیدی بر روی کروموزوم‌های اتوزومی، همبستگی بالایی با اندازه کروموزوم را نشان دادند، به طوری که تراکم نشان‌گرها روی کروموزوم‌های کوچک‌تر، بیشتر بود. در تحقیق دیگری که توسط Pertille و همکاران (۲۰۱۶) انجام شد، تعداد ۴۶۲ قطعه جوجه با استفاده از فن آوری ژنوتیپ کردن از طریق توالی‌یابی بر روی پلت‌فرم ایلومینا ژنوتیپ شدند. پس از این که پارامترهای ویرایش اعمال شدند، تعداد ۱۳۴۵۲۸ چندشکلی‌های تک‌نوکلئوتیدی در جمعیت مورد مطالعه شناسایی شد. همچنین، Haobin و همکاران (۲۰۲۰) از فن آوری ژنوتیپ کردن از طریق توالی‌یابی برای مطالعات پویا کل ژنوم در صفات رنگ‌دانه‌های پوست در ۲۹۱ مرغ تخمگذار (در دو دسته ۲۱۶ مرغ سفید و ۷۵ مرغ سیاه) استفاده و ۴۵۸۳۰۴ چندشکلی‌های تک‌نوکلئوتیدی با کیفیت بالا شناسایی کردند. تعداد ۶۸ چندشکلی تک‌نوکلئوتیدی معنی‌دار مرتبط با رنگ‌دانه‌های پوست، که بر روی کروموزوم ۲۰ قرار داشتند و با ۱۰ ژن مختلف در ارتباط بودند، شناسایی شدند.

هزینه بالای فن آوری آرایه‌های چندشکلی تک‌نوکلئوتیدی به ازای هر نمونه، استفاده از آن‌ها را در کاربردهای تحقیقاتی، به‌خصوص برای جمعیت‌های با اندازه بزرگ، محدود می‌کند (Illumina, 2014). بنابراین، فن آوری ژنوتیپ کردن از طریق توالی‌یابی می‌تواند جایگزین مناسبی برای این نوع فناوری‌های پرهزینه باشد، اما لازم است کارایی آن از بعد تفاوت تعداد چندشکلی‌ها، توزیع آن‌ها روی کروموزوم‌های مختلف، وفور آل‌های کم‌یاب و عدم تعادل ناشی از پیوستگی مورد مقایسه قرار گیرد. با توجه به در دسترس بودن اطلاعات تراشه ۶۰K و داده‌های حاصل از تعیین ژنوتیپ از طریق توالی‌یابی، هدف مطالعه حاضر، بررسی کارایی داده‌های حاصل از فن آوری ژنوتیپ کردن از طریق توالی‌یابی و مقایسه داده‌های آن با داده‌های حاصل از فن آوری آرایه‌ی چندشکلی تک‌نوکلئوتیدی ۶۰K در یک جمعیت آمیخته مرغ بود.

روش شناسی پژوهش

در این پژوهش، به‌منظور انجام مقایسه فن آوری‌های ژنوتیپ کردن از طریق توالی‌یابی و آرایه‌های چندشکلی تک‌نوکلئوتیدی، از یک جمعیت F_2 حاصل از تلاقی دو طرفه ۷۹ پرنده بومی آذربایجان غربی (۴۰ مرغ و ۳۹ خروس) و ۴۹ پرنده یک لاین تجاری گوشتی آرین (۲۶ مرغ و ۲۳ خروس) استفاده شد. پرنده‌گان سویه آرین از ایستگاه پرورش و اصلاح نژاد مرغ لاین آرین واقع در بابل کنار و پرنده‌گان بومی از ایستگاه پرورش و اصلاح نژاد مرغ بومی استان آذربایجان غربی واقع در شهرستان ارومیه، با سن مشابه تقریباً ۱۳ هفته، در اسفند ۱۳۹۰ به مرکز تحقیقات پرنده‌گان دانشکده کشاورزی دانشگاه تربیت مدرس انتقال داده شدند. چگونگی ایجاد نسل‌های پایه، اول و دوم و مدیریت پرورش پرنده‌ها در هر نسل در مطالعه‌ی مقصودی و همکاران (۲۰۱۴) شرح داده شده است. به‌طور خلاصه، از آمیزش دو طرفه پرنده‌گان نسل پایه‌ی بومی و لاین آرین، جمعیت F_1 و از آمیزش متقابل پرنده‌های افراد نسل F_1 ، و طی پنج دوره‌ی جوجه‌کشی، جمعیت F_2 ایجاد شد. جوجه‌های نسل F_2 در مجموع، شامل نه خانواده برادر- خواهر ناتنی بودند. این جوجه‌ها از ۱ تا ۷ روزگی به صورت گروهی و از ۸ روزگی تا پایان دوره در قفس‌های انفرادی و در شرایط محیطی و تغذیه‌ای یکسان پرورش داده شدند. در نسل F_2 ، رکوردهای فنوتیپی مختلفی از جمله صفات وزن تولد تا ۱۲ هفتگی، مصرف خوراک هفتگی، متابولیت‌های پلاسمای خون، صفات ایمنی، و صفات لاشه در پایان هفته ۱۲ ثبت گردید (Maghsoudi et al., 2014).

برای تعیین ژنوتیپ با استفاده از تراشه‌ی ۶۰K، از ۳۱۲ قطعه پرنده نسل F₂ نمونه خون تهیه و DNA ژنومی تمام نمونه‌ها استخراج شدند. پس از اطمینان از کیفیت نمونه‌های DNA، غلظت هر یک از نمونه‌های DNA با ۵۰ نانوگرم در هر میکرولیتر، با آب دو بار تقطیر رقیق و در نهایت DNA هر یک از افراد با همکاری دانشگاه آرهوس دانمارک بوسیله‌ی پانل تجاری Illumina 60k SNP chip شرکت کاب تعیین ژنوتیپ شدند. جزئیات بیشتر روش استخراج DNA، در مطالعه جانروح علی‌آباد و همکاران (۲۰۱۶) ارائه شده است. همچنین، با استفاده از فن آوری ژنوتیپ‌کردن از طریق توالی‌یابی، ۳۶۸ پرنده شامل ۳۱۲ پرنده از نسل F₂ و ۳۷ پرنده از نسل F₁ و ۱۹ پرنده از نسل والدین تعیین ژنوتیپ شدند. ارزیابی ابتدایی ژنوتیپی با استفاده از فن آوری ژنوتیپ‌کردن از طریق توالی‌یابی طبق دستورالعمل ارائه شده توسط مؤسسه الشایر (The Elshire Group Ltd.) نیوزیلند در آزمایشگاه گروه علوم دامی دانشگاه تربیت مدرس انجام شد. DNAهای استخراج شده‌ی تمام پرندگان بر روی ژل آگارز یک درصد تعیین کیفیت شدند. ده درصد نمونه‌های DNA پس از تأیید کیفیت، توسط موسسه الشایر به‌وسیله‌ی آنزیم EcoRI، برش داده شدند تا اطمینان حاصل شود که نمونه‌های DNA برای تشکیل کتابخانه ژنومی کیفیت لازم را دارند. نمونه‌های برش داده شده نیز بر روی ژل آگارز یک درصد تعیین کیفیت شدند. پس از تأیید تصاویر، نمونه‌های DNA ۳۶۸ پرنده برای توالی‌یابی به موسسه الشایر کشور نیوزیلند ارسال شدند.

نمونه‌های DNA استخراج شده پرندگان نسل F₂ و F₁ به همراه والدین، پس از تعیین کمیت (۱۰۰ ng/μl) و کیفیت با استفاده از آنزیم برشی ApeKI (با جایگاه برشی GCWGC) جهت کاهش اندازه قطعات ژنوم مورد برش واقع شدند. DNA ژنومی هضم شده پرندگان با غلظت (۳/۶ ng/μl) با آدپتورها مخلوط شدند و سپس با استفاده از پرایمرهایی که در برگیرنده توالی‌های مکمل آدپتورهای اتصالی هستند، قطعات تکثیر شدند. با تکثیر نمونه‌های مخلوط شده، کتابخانه توالی‌یابی تشکیل و واکنش‌های توالی‌یابی DNA آغاز شد. توالی هر پرنده با توجه به بارکد آن قابل شناسایی بود. مؤسسه الشایر از پلت فرم X Ten برای توالی‌یابی استفاده نمود. توالی‌یابی X Ten خوانش‌های ۱۵۰ جفت باز با انتهای جفت شده (150 bp paired end) تولید می‌کند. تعیین کیفیت خوانش‌ها در نرم‌افزار نسخه ۵ TASSEL (Bradbury et al., 2007) انجام و توالی‌های آدپتوری از خوانش‌ها حذف شدند. سپس انتصاب خوانش‌ها به ژنوم مرجع با استفاده از نرم‌افزار bowtie2 انجام شد. شناسایی چندشکلی‌های تکنوکلوتیدی در نرم‌افزار نسخه ۵ TASSEL رویه GBSV2 (Bradbury et al., 2007) با پارامترهایی نظیر فراوانی آلل کم‌یاب مساوی و کم‌تر از ۰/۰۱ و حداقل پوشش جایگاه ۰/۰۱ صورت پذیرفت. از آنجایی که سطح بالایی از داده‌های گم‌شده از فراخوانی SNPها بدست آمد، برای افزایش صحت اطلاعات حاصل از فن آوری ژنوتیپ‌کردن از طریق توالی‌یابی، با استفاده از ویرایش 5.1 نرم‌افزار Beagle (Pook et al., 2020) جانهمی انجام شد. کنترل کیفیت داده‌های حاصل از تراشه ۶۰K با پارامترهای فراوانی آلل کم‌یاب کم‌تر از ۰/۰۵، عدم تعادل هاردی و واینبرگ در سطح احتمال کم‌تر از 10⁻⁶ و نرخ خوانش کم‌تر از ۹۵ درصد انجام شد. در پژوهش حاضر، چندشکلی‌های تکنوکلوتیدی حاصل از فن آوری ژنوتیپ‌کردن از طریق توالی‌یابی با چندشکلی‌های تکنوکلوتیدی تراشه ۶۰K از نظر تعداد بررسی شدند. در هر دو فن آوری، تجزیه و تحلیل ژنوتیپ‌های گم‌شده، عدم تعادل پیوستگی برای هر کروموزوم بصورت جداگانه و در فواصل فیزیکی 10Kbp، 100Kbp و 1000Kbp و بلوک‌های هاپلوتایپی چندشکلی‌های تکنوکلوتیدی، هیستوگرام توزیع فراوانی آلل کم‌یاب و گراف‌های آنالیز ساختارهای جمعیتی و خویشاوندی برای هر دو فن آوری مورد مطالعه، با استفاده از نرم‌افزار PLINK نسخه ۱،۹ (Purcell et al., 2007) در محیط نرم‌افزار R برآورد شد. سپس تمامی نتایج حاصل از آنالیز داده‌های حاصل از دو فن آوری با هم مقایسه شدند و مورد تجزیه و تحلیل قرار گرفتند.

یافته‌های پژوهش و بحث

کنترل کیفیت نمونه‌های ارسالی به موسسه الشایز نیوزیلند نشان داد که ۹۳ درصد نمونه‌ها کیفیت لازم برای تولید داده را داشتند و ۷ درصد باقیمانده (۱۷ پرنده) اطلاعات کمی را تولید کردند. از آنجایی که توالی‌یابی X Ten، خوانش‌های ۱۵۰ جفت باز با انتهای جفت شده تولید می‌کند، دو فایل Fastq، یکی برای خوانش اول و دیگری برای خوانش دوم تولید شد. در نهایت با استفاده از تحلیل خوانش‌های دو فایل Fastq با استفاده از نرم‌افزار نسخه ۵ TASSEL (Bradbury et al., 2007)، ۸۸۲۹۱۸ جایگاه چندشکلی تک‌نوکلئوتیدی با میانگین فاصله نشان‌گری ۱/۳ کیلو جفت باز شناسایی شدند که ۸۱۵۶۱۳ چندشکلی تک‌نوکلئوتیدی مربوط به کروموزوم‌های یک تا ۲۸ و ۶۷۳۰۵ چندشکلی تک‌نوکلئوتیدی مربوط به کروموزوم‌های ۳۰ تا ۳۳ و کروموزوم‌های جنسی Z و W بودند. در فن‌آوری آرایه چندشکلی تک‌نوکلئوتیدی ۶۰K در هر نمونه، برای کروموزوم‌های یک تا ۲۸، ۵۱۳۴۷ نشان‌گر چندشکلی‌های تک‌نوکلئوتیدی با میانگین فاصله نشان‌گری ۱۳/۳ کیلو جفت باز ژنوتیپ شد. مشاهده شدن چندشکلی‌های تک‌نوکلئوتیدی در کروموزوم‌های ۳۰ تا ۳۳ در روش فن‌آوری تعیین ژنوتیپ از طریق توالی‌یابی نشان می‌دهد که این روش قادر به مشخص کردن تعداد بیشتری کروموزوم و در نتیجه شناسایی نشان‌گرهای زیادتری در مقایسه با روش فن‌آوری آرایه‌ی چندشکلی تک‌نوکلئوتیدی است. مقایسه فواصل نشان‌گرها در داده‌های حاصل از دو فن‌آوری نشان می‌دهند که در فن‌آوری ژنوتیپ کردن از طریق توالی‌یابی نشان‌گرها تراکم بالاتری دارند و به خوبی در سراسر ژنوم گسترده شده‌اند. تعداد چندشکلی‌های تک‌نوکلئوتیدی، نسبت چندشکلی‌های تک‌نوکلئوتیدی، تعداد چندشکلی‌های تک‌نوکلئوتیدی مشترک و موقعیت فیزیکی چندشکلی‌های تک‌نوکلئوتیدی مشترک دو فن‌آوری ژنوتیپ کردن از طریق توالی‌یابی و آرایه‌ی چندشکلی تک‌نوکلئوتیدی ۶۰K در جدول ۱ آرایه شده است. بیشترین تعداد چندشکلی تک‌نوکلئوتیدی در هر دو فناوری در کروموزوم شماره یک و کم‌ترین تعداد چندشکلی‌های تک‌نوکلئوتیدی در فن‌آوری‌های ژنوتیپ کردن از طریق توالی‌یابی و آرایه‌ی چندشکلی تک‌نوکلئوتیدی ۶۰K به ترتیب در کروموزوم‌های شماره ۳۲ و ۱۶ مشاهده شد. در هر دو فناوری، بیشترین چندشکلی‌های تک‌نوکلئوتیدی مشترک مربوط به کروموزوم شماره یک بود. با وجود این که تعداد SNP‌های شناسایی شده در فن‌آوری تعیین ژنوتیپ از طریق توالی‌یابی از فن‌آوری تراشه ۶۰K بیشتر بود اما چندشکلی‌های مشترک بین آن‌ها بسیار کم و در برخی کروموزوم‌ها مانند کروموزوم‌های ۶، ۷، ۹، ۱۳، ۱۶، ۲۲، ۲۵، ۳۰، ۳۱، ۳۲، ۳۳، Z و W صفر بود (جدول ۱).

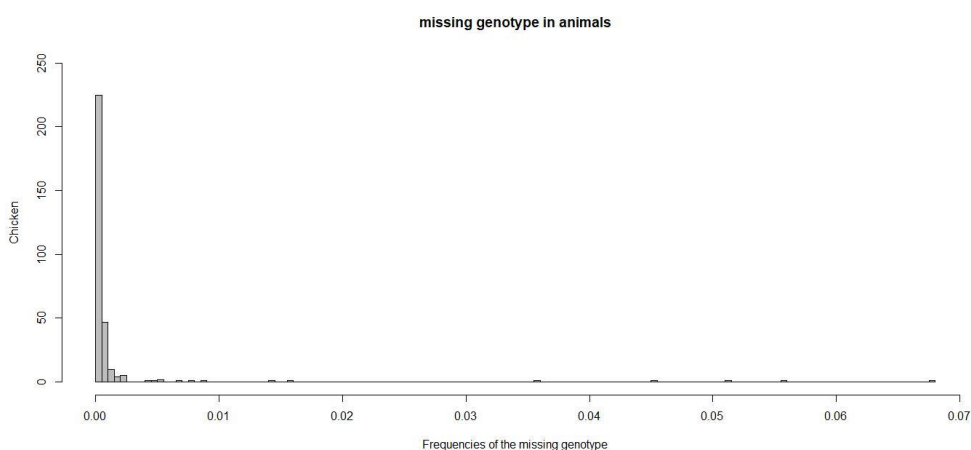
جدول ۱. آمار توصیفی شمارش SNP‌ها در فن‌آوری‌های ژنوتیپ کردن از طریق توالی‌یابی و آرایه چندشکلی تک‌نوکلئوتیدی ۶۰K

شماره کروموزوم	طول کروموزوم	تعداد SNP در GBS	تعداد SNP در SNP Array	GBS/SNPA	تعداد SNP مشترک	موقعیت فیزیکی چندشکلی‌های تک‌نوکلئوتیدی مشترک
۱	۱۹۷۶۰۸۳۸۶	۱۱۷۹۲۷	۸۳۰۳	۱۴	۱۰	۶۵۳۷۶۳۲۸، ۲۴۶۱۶۶۵۰، ۲۸۱۳۷۱۹، ۱۳۴۳۲۵۹۳۹، ۱۳۵۰۳۰۰۳۵، ۸۵۸۰۳۰۵۲، ۱۳۴۷۲۰۴۷۴، ۱۶۴۱۶۳۸۵۴، ۹۲۵۷۲۴۱۳، ۴۹۲۹۸۸۹۰
۲	۱۴۹۶۸۲۰۴۹	۸۱۱۲۲	۶۳۵۵	۱۳	۲	۶۵۳۸۹۰۱۲، ۷۵۵۷۰۷
۳	۱۱۰۸۳۸۴۱۸	۶۴۰۶۸	۴۷۳۹	۱۴	۲	۳۹۲۲۸۵۷۹، ۱۶۲۵۷۶۳
۴	۹۱۳۱۵۲۴۵	۶۱۵۲۹	۳۸۷۲	۱۶	۳	۴۶۱۹۴۰۰۷، ۴۶۲۳۱۵۲۹، ۳۵۵۹۰۳۴۷
۵	۵۹۸۰۹۰۹۸	۴۶۴۶۷	۲۵۴۲	۱۸	۱	۱۱۳۲۸۹۹
۶	۳۶۳۷۴۷۰۱	۳۱۹۲۶	۱۹۹۵	۱۶	-	-
۷	۳۶۷۴۲۳۰۸	۲۹۷۳۱	۲۰۸۹	۱۴	-	-
۸	۳۰۲۱۹۴۴۶	۲۷۷۹۳	۱۶۳۶	۱۷	۱	۲۰۹۶۴۷۱۲

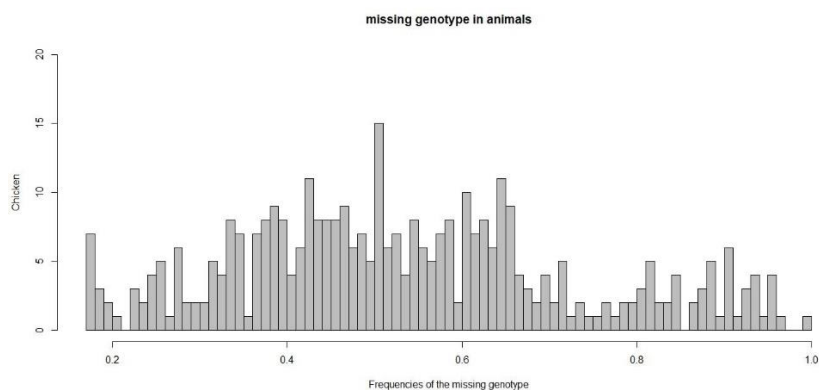
شماره کروموزوم	طول کروموزوم	تعداد SNP در GBS	تعداد SNP در SNP Array	GBS/SNPA	تعداد SNP مشترک	موقعیت فیزیکی چندشکلی‌های تک-نوکلئوتیدی مشترک
۹	۲۴۱۵۳۰۸۶	۲۶۷۴۴	۱۳۶۶	۲۰	-	-
۱۰	۲۱۱۱۹۸۴۰	۲۲۷۷۶	۱۵۵۳	۱۵	۱	۱۹۹۰۳۳۳۸
۱۱	۲۰۲۰۰۰۴۲	۱۸۹۸۱	۱۵۳۱	۱۲	۱	۷۸۳۳۱۳
۱۲	۲۰۳۸۷۲۷۸	۲۴۷۴۵	۱۵۵۹	۱۶	۱	۱۰۹۰۱۴۴۲
۱۳	۱۹۱۶۶۷۱۴	۲۵۲۸۶	۱۳۷۱	۱۸	-	-
۱۴	۱۶۳۱۹۳۰۸	۲۶۷۶۴	۱۱۷۹	۲۳	۱	۱۳۱۵۹۴۸۵
۱۵	۱۳۰۶۲۱۸۴	۱۷۴۶۷	۱۲۲۲	۱۴	۱	۱۰۲۰۴۴۸۴
۱۶	۲۸۴۴۶۰۱	۵۲۳۶	۲۴	۲۱۸	-	-
۱۷	۱۰۷۶۲۵۱۲	۲۰۶۰۴	۹۹۴	۲۱	۴	۸۸۳۳۴۷۰, ۳۹۷۵۸۸۵, ۸۶۲۵۵۶, ۸۰۸۸۴۷۴
۱۸	۱۱۳۷۳۱۴۰	۲۳۳۱۵	۱۰۴۸	۲۲	۵	۴۲۹۸۳۸۴, ۸۶۰۹۴۹, ۴۶۱۳۴۹۷ ۹۱۰۴۰۲۶, ۹۰۸۸۶۹۹
۱۹	۱۰۳۳۳۲۱۲	۱۸۳۳۱	۹۷۳	۱۹	۲	۴۷۴۳۳۹۱, ۵۱۷۶۰۷
۲۰	۱۳۸۹۷۲۸۷	۲۳۶۳۳	۱۸۱۵	۱۳	۲	۷۸۸۱۷۰۰, ۳۷۷۰۲۶
۲۱	۶۸۴۴۹۷۹	۱۱۵۶۷	۹۰۱	۱۳	۱	۶۲۸۲۵۵۱
۲۲	۵۴۵۹۴۶۲	۵۸۴۱	۴۳۲	۱۴	-	-
۲۳	۶۱۴۹۵۸۰	۱۵۶۴۷	۷۲۴	۲۲	۳	۸۷۵۷۳۹, ۸۴۹۶۲۷, ۳۸۴۰۷۵۲
۲۴	۶۴۹۱۲۲۲	۱۴۱۳۲	۸۵۳	۱۷	۱	۶۰۰۴۲۴۵
۲۵	۳۹۸۰۶۱۰	۷۶۲۲	۲۱۱	۳۶	-	-
۲۶	۶۰۵۵۷۱۰	۱۵۶۶۳	۷۷۶	۲۰	۳	۸۰۰۵۳۷, ۳۳۴۱۴۳۳, ۵۰۶۳۴۲۸, ۱۸۳۷۴۸۶
۲۷	۸۰۸۰۴۳۲	۱۶۹۳۷	۵۷۶	۲۹	۲	۳۸۶۷۴۶۰, ۲۷۲۳۹۰۵
۲۸	۵۱۱۶۸۸۲	۱۳۳۰۹	۷۰۸	۱۹	۱	۳۹۱۷۸۲۷
۳۰	۱۸۱۸۵۲۵	۱۳۴۵	-	-	-	-
۳۱	۶۱۵۳۰۳۴	۸۰۲۵	-	-	-	-
۳۲	۷۲۵۸۳۱	۶۹۶	-	-	-	-
۳۳	۷۸۲۱۶۶۶	۱۹۲۹۷	-	-	-	-
W	۶۸۱۳۱۱۴	۱۱۲۸	-	-	-	-
Z	۸۲۵۲۹۹۲۱	۳۷۲۶۴	-	-	-	-

فراوانی ژنوتیپ‌های گم‌شده، از تقسیم تعداد چندشکلی‌های تک‌نوکلئوتیدی گم‌شده پرنده‌ها بر تعداد کل چندشکلی‌های تک‌نوکلئوتیدی در هر فن‌آوری محاسبه شد. هیستوگرام ژنوتیپ گم‌شده برای داده‌های حاصل از آرایه‌ی چندشکلی تک-نوکلئوتیدی ۶۰K و ژنوتیپ کردن از طریق توالی‌یابی قبل از جانمایی و ژنوتیپ کردن از طریق توالی‌یابی بعد از جانمایی به ترتیب در شکل‌های ۱، ۲ و ۳ نشان داده شده است. در فن‌آوری آرایه چندشکلی تک‌نوکلئوتیدی ۶۰K اکثر پرنده‌گان تمام چندشکلی‌های تک‌نوکلئوتیدی را داشتند که نشان می‌دهد این روش مقدار کمی داده از دست رفته ایجاد می‌کنند (شکل ۱). برای فن‌آوری ژنوتیپ کردن از طریق توالی‌یابی، فراوانی ژنوتیپ‌های گم‌شده حدود ۵۴ درصد بود (شکل ۲). اگرچه مقدار زیاد داده‌های از دست رفته‌ی ناشی از پوشش کم توالی‌یابی یکی از کاستی‌های فن‌آوری ژنوتیپ کردن از طریق توالی‌یابی است اما می‌توان با استفاده از روش‌های جانمایی و تخمین ژنوتیپ‌ها، میزان داده‌های از دست رفته را به طور چشم‌گیری کاهش داد (Glaubitz *et al.*, 2014). برای این داده‌ها، نتایج جانمایی ژنوتیپ‌های گم‌شده نشان داد که تعداد اندکی از پرنده‌گان برای جایگاه‌هایی با

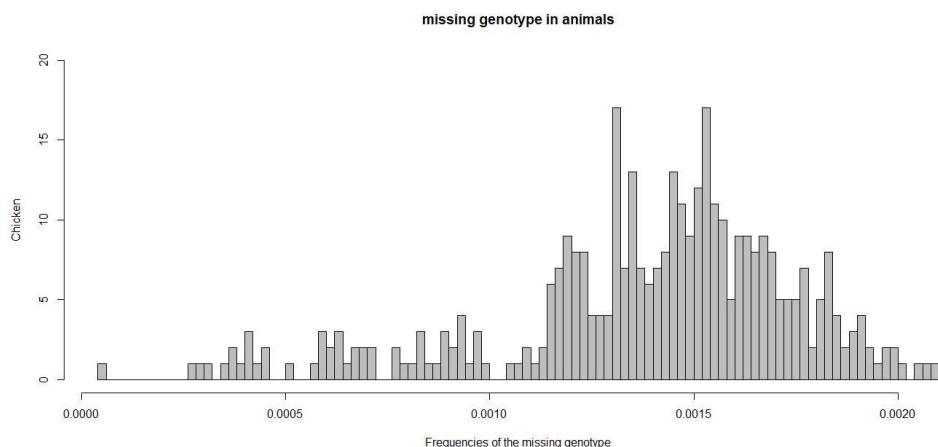
فراوانی کم (فراوانی کم تر از ۰/۰۰۵)، ژنوتیپ از دست رفته داشتند (شکل ۳). وجود ژنوتیپ‌های گم‌شده در استفاده از فن‌آوری ژنوتیپ کردن از طریق توالی‌یابی در مطالعات دیگر محققین نیز گزارش شده است (Wang *et al.*, 2020; Loureiro *et al.*, 2020). در مطالعه Horvath و همکاران (۲۰۲۰) بر روی کلزا ۲۹۰۹۷۲ نشان‌گر شناسایی شد که ۳۶/۸ درصد ژنوتیپ آن‌ها گم‌شده بود و با روش جان‌هی، تعداد آن‌ها به کم‌تر از ۱۳ درصد کاهش یافت. میانگین عدم تعادل ناشی از پیوستگی حاصل از داده‌های آرایه‌ی چندشکلی تک‌نوکلئوتیدی ۶۰K و ژنوتیپ کردن از طریق توالی‌یابی در فواصل فیزیکی 10Kbp، 100Kbp و 1000Kbp در جدول ۲ و عدم تعادل پیوستگی هر دو فن‌آوری برای هر کروموزوم در شکل ۴ نشان داده شده است. عدم تعادل پیوستگی به ارتباط غیرتصادفی بین آلل‌ها در جایگاه‌های ژنی متفاوت گفته می‌شود و به این مفهوم است که پیوستگی ژن‌ها تصادفی نیست. برای برآورد میزان عدم تعادل پیوستگی، از مربع ضریب همبستگی بین دو ژنگاه (r^2) استفاده شد. ارزش r^2 از صفر برای یک جفت جایگاه بدون عدم تعادل پیوستگی تا یک در یک جفت جایگاه با عدم تعادل پیوستگی کامل متغیر است.



شکل ۱. هیستوگرام داده‌های گم‌شده در فن‌آوری آرایه‌ی چندشکلی تک‌نوکلئوتیدی ۶۰K. محور X فراوانی ژنوتیپ‌های گم‌شده و محور Y تعداد پرندگان است.



شکل ۲. هیستوگرام داده‌های گم‌شده در فن‌آوری ژنوتیپ کردن از طریق توالی‌یابی قبل از جان‌هی. محور X فراوانی ژنوتیپ‌های گم‌شده و محور Y تعداد پرندگان است.



شکل ۳. هیستوگرام داده‌های گم‌شده در فن آوری ژنوتیپ کردن از طریق توالی یابی بعد از جانهی. محور X فراوانی ژنوتیپ‌های گم‌شده و محور Y تعداد پرندگان است.

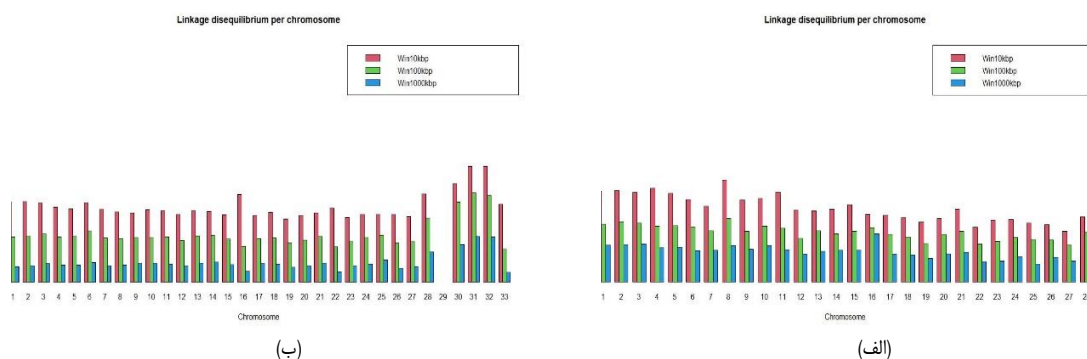
جدول ۲. آماره‌های عدم تعادل ناشی از پیوستگی به تفکیک فواصل فیزیکی^۱

آرایه‌ی چندشکلی تک‌نوکلئوتیدی			تعیین ژنوتیپ از طریق توالی یابی			طول کروموزوم
حد اکثر	حداقل	میانگین (انحراف معیار)	حد اکثر	حداقل	میانگین (انحراف معیار)	
۰/۳۳۵	۰/۱۶۷	۰/۲۴۳ (۰/۰۴۵)	۰/۲۹۰	۰/۲۰۷	۰/۲۳۷ (۰/۰۲۱)	۱۰ کیلوباز
۰/۲۱۰	۰/۱۲۱	۰/۱۶۳ (۰/۰۲۴)	۰/۲۰۹	۰/۱۱۶	۰/۱۴۵ (۰/۰۱۷)	۱۰۰ کیلوباز
۰/۱۵۶	۰/۰۵۸	۰/۰۹۹ (۰/۰۲۲)	۰/۱۰۰	۰/۰۳۴	۰/۰۵۷ (۰/۰۱۲)	۱۰۰۰ کیلوباز

^۱ برای مقایسه میانگین‌های دو روش، فقط از کروموزوم‌های ۱ تا ۲۸ استفاده شد.

به طور کلی، تعادل ناشی از پیوستگی با افزایش طول کروموزوم برای هر دو روش تحت مطالعه کاهش داشت. این کاهش برای روش تعیین ژنوتیپ از طریق توالی یابی بیشتر از روش آرایه‌ی چندشکلی بود (جدول ۲). برای فاصله ۱۰ کیلوباز، تفاوت مقدار عدم تعادل ناشی از پیوستگی در داده‌های حاصل از فن آوری آرایه‌ی ۶۰K (۰/۲۴۳) و داده‌های تعیین ژنوتیپ از طریق توالی یابی (۰/۲۳۷) اندک بود. این مقادیر در فواصل ۱۰۰ کیلوباز و ۱۰۰۰ کیلوباز برای آرایه‌ی چندشکلی بیشتر از تعیین ژنوتیپ از طریق توالی یابی بود، که نشان می‌دهد در فواصل فیزیکی بیشتر، تراشه‌های چندشکلی در شناسایی جایگاه‌های صفات کمی، بهتر از داده‌های توالی یابی عمل می‌کند. در فن آوری ژنوتیپ کردن از طریق توالی یابی، تعداد چندشکلی‌های تک‌نوکلئوتیدی در یک فاصله فیزیکی معین بسیار بیشتر از فن آوری آرایه‌ی ۶۰K است (جدول ۱). مطالعات نشان داده است با افزایش تراکم نشان‌گرها، میزان نوترکیبی نیز افزایش پیدا می‌کند (Groenen *et al.*, 2009; Negro *et al.*, 2019). در مطالعه‌ی Negro و همکاران (۲۰۱۹) تنوع تعداد نشان‌گرها بر روی کروموزوم شماره ۳ را با استفاده از سه فن آوری آرایه‌های چندشکلی تک-نوکلئوتیدی ۵۰K و ۶۰K و ژنوتیپ کردن از طریق توالی یابی مقایسه کردند و نشان دادند که با افزایش تنوع تعداد نشان‌گرها، میزان نوترکیبی نیز افزایش می‌یابد. کاهش عدم تعادل پیوستگی در فن آوری تعیین ژنوتیپ از طریق توالی یابی ممکن است به علت افزایش میزان نوترکیبی در نتیجه تعداد بیشتر چندشکلی‌های تک‌نوکلئوتیدی آن باشد. و همچنین در نتیجه‌ی تعداد زیاد نشان‌گرها، میزان LDهای ناقص نیز افزایش و در نهایت میزان LD کاهش می‌یابد. با وجود این، با توجه به این که در مطالعات پویس ژنومی، جستجوی ژن‌های هم‌بسته با چندشکلی‌های تک‌نوکلئوتیدی معنی‌دار در فواصل ۵۰۰ کیلوباز مطالعه می‌شود (Brodie *et al.*, 2016)، نتایج مطالعه‌ی حاضر نشان می‌دهد داده‌های تعیین ژنوتیپ از طریق توالی یابی نیز می‌توانند به‌عنوان

ابزار مناسب در این نوع مطالعات مورد استفاده قرار گیرد. در سرتاسر کروموزوم‌ها، کروموزوم‌های ۳۰ تا ۳۴، که فقط در داده‌های فن‌آوری ژنوتیپ‌کردن از طریق توالی‌یابی مشاهده شدند، بیشترین مقادیر عدم تعادل ناشی از پیوستگی را داشتند. برای کروموزوم‌های ۱ تا ۲۸، که در هر دو فناوری وجود داشتند، الگوی مشابهی برای عدم تعادل ناشی از پیوستگی در فاصله‌های فیزیکی مورد مطالعه مشاهده نشد. برای نمونه، بیشترین و کم‌ترین میزان عدم تعادل پیوستگی در فاصله فیزیکی ۱۰ کیلو باز، در فن‌آوری ژنوتیپ‌کردن از طریق توالی‌یابی، به ترتیب در کروموزوم‌های ۲۸ (۰/۲۹۰) و ۱۹ (۰/۲۰۷)، و در فن‌آوری آرایه چندشکلی تک‌نوکلئوتیدی ۶۰K، به ترتیب در کروموزوم‌های ۸ (۰/۳۳۵) و ۲۷ (۰/۱۶۶) مشاهده شد (شکل ۴).



شکل ۴. عدم تعادل پیوستگی در فواصل فیزیکی 10Kbp، 100Kbp و 1000Kbp در داده‌های حاصل از آرایه‌ی چندشکلی تک‌نوکلئوتیدی ۶۰K (الف)، و داده‌های حاصل از ژنوتیپ‌کردن از طریق توالی‌یابی (ب). محور X شماره کروموزوم و محور Y میزان LD در هر کروموزوم است.

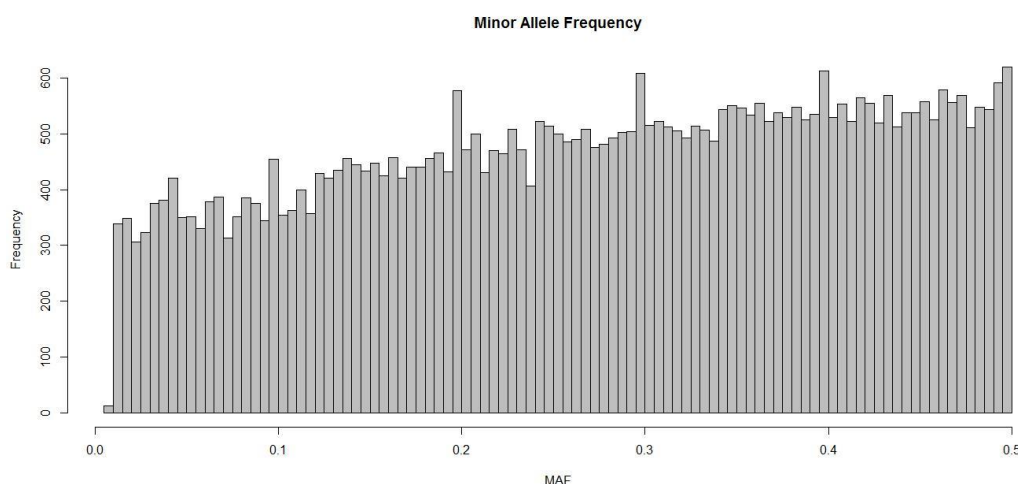
بلوک‌های هاپلوتیپی چندشکلی‌های تک‌نوکلئوتیدی، گروهی از چندشکلی‌های تک‌نوکلئوتیدی هستند که با یکدیگر پیوستگی بالایی دارند و در نتیجه انتظار آن است که توارث آن‌ها هم‌زمان صورت گیرد (Gabriel *et al.*, 2002). نتایج بلوک‌های هاپلوتیپی چندشکلی‌های تک‌نوکلئوتیدی داده‌های حاصل از دو فن‌آوری ژنوتیپ‌کردن از طریق توالی‌یابی و آرایه‌ی چندشکلی تک‌نوکلئوتیدی ۶۰K برای هر یک از کروموزوم‌های جمعیت مورد مطالعه در جدول ۳ نشان داده شده است. با استفاده از فن‌آوری ژنوتیپ‌کردن از طریق توالی‌یابی، تعداد ۱۶۷۳۹۶ بلوک هاپلوتیپی شناسایی شد. بیشترین تعداد چندشکلی‌های تک‌نوکلئوتیدی درون بلوک، روی کروموزوم شماره ۷ با تعداد ۲۰ چندشکلی تک‌نوکلئوتیدی با طول ۷۰ جفت باز بود. در فن‌آوری آرایه‌ی چندشکلی تک‌نوکلئوتیدی ۶۰K، تعداد ۶۶۵۰ بلوک هاپلوتیپی مشاهده شد. در این روش، بیشترین تعداد چندشکلی‌های تک‌نوکلئوتیدی درون بلوک، روی کروموزوم شماره ۲۸ با تعداد ۲۱ چندشکلی تک‌نوکلئوتیدی با طول ۷۰ جفت باز بود. در مطالعه‌ی حاضر، تعداد بلوک‌های هاپلوتیپی در فن‌آوری ژنوتیپ‌کردن از طریق توالی‌یابی نسبت به آرایه‌ی چندشکلی تک‌نوکلئوتیدی ۶۰K بسیار بیشتر بود. از آنجایی که هاپلوتیپ‌ها نسبت به نشان‌گرهای مجزا می‌توانند عدم تعادل پیوستگی بالاتری با جایگاه ژنومی صفات کمی داشته باشند انتظار می‌رود بلوک‌های هاپلوتیپی حاصل از فن‌آوری ژنوتیپ‌کردن از طریق توالی‌یابی، مطالعات پویش ژنوم را بهبود بخشد (Sehgal *et al.*, 2020). در بررسی Khanyile و همکاران (۲۰۱۵)، ساختار هاپلوتیپی برای داده‌های آرایه‌ی چندشکلی تک‌نوکلئوتیدی ۶۰K سه گروه مرغ بومی از مناطق مالاوی، آفریقای جنوبی و زیمبابوه را بررسی و به ترتیب، تعداد ۶۴۹، ۲۱۰۴ و ۲۴۴۲ بلوک‌های هاپلوتیپی را گزارش کردند. آگاهی از بلوک‌های هاپلوتیپی اطلاعات مهمی از مطالعات پویش ژنومی و انتخاب ژنومی ارائه می‌دهد. همچنین بلوک‌های هاپلوتیپی می‌توانند در شناسایی مناطق ژنومی تحت انتخاب طی تکامل مورد استفاده قرار گیرند (Franssen *et al.*, 2017).

جدول ۳. تعداد بلوک‌های هاپلوتیپی در فن آوری‌های ژنوتیپ کردن از طریق توالی یابی و آرایه چند شکلی‌های تک نوکلئوتیدی ۶۰K

شماره کروموزوم	تعداد بلوک در فن آوری چند شکلی - طریق توالی یابی	تعداد بلوک در فن آوری ژنوتیپ کردن از طریق توالی یابی	تعداد بلوک در فن آوری چند شکلی - های تک نوکلئوتیدی ۶۰K
۱	۲۲۶۷۲	۹۲۹	
۲	۱۵۷۲۸	۷۱۰	
۳	۱۲۷۴۵	۵۵۶	
۴	۱۱۸۸۷	۴۴۶	
۵	۸۸۵۶	۲۶۰	
۶	۶۳۳۱	۲۵۰	
۷	۵۷۱۴	۲۷۰	
۸	۵۲۴۸	۲۱۹	
۹	۵۱۹۲	۲۰۳	
۱۰	۴۱۵۰	۲۴۳	
۱۱	۳۵۰۳	۲۴۰	
۱۲	۴۶۷۷	۲۳۰	
۱۳	۴۷۴۰	۲۰۳	
۱۴	۵۰۴۸	۱۷۷	
۱۵	۳۲۵۸	۲۰۱	
۱۶	۷۹۰	۱	
۱۷	۳۸۰۱	۱۵۱	
۱۸	۴۲۸۵	۱۵۳	
۱۹	۳۳۵۳	۱۵۵	
۲۰	۴۴۱۹	۲۸۲	
۲۱	۲۱۵۷	۱۴۰	
۲۲	۱۰۰۱	۴۸	
۲۳	۲۸۴۸	۱۱۵	
۲۴	۲۶۸۷	۱۲۴	
۲۵	۱۴۲۶	۳۱	
۲۶	۲۸۶۹	۱۲۵	
۲۷	۳۰۶۶	۷۸	
۲۸	۲۶۵۲	۱۱۰	
۳۰	۲۹۸	-	
۳۱	۱۶۹۰	-	
۳۲	۱۴۱	-	
۳۳	۳۷۵۰	-	
W	.	-	
Z	۶۳۸۶	-	

توزیع فراوانی آللی کم‌یاب (MAF) در داده‌های حاصل از تعیین ژنوتیپ از طریق توالی یابی و آرایه‌ی چندشکلی تک-نوکلئوتیدی ۶۰K نیز با استفاده از برنامه PLINK 1.9 (Purcell et al., 2007) در نرم‌افزار R نسخه 4.3.1 بررسی و هیستوگرام توزیع آن‌ها به ترتیب در شکل‌های ۵ و ۶ آرایه شده است. هیستوگرام فراوانی آللی کم‌یاب در داده‌ی آرایه‌ی چندشکلی تک‌نوکلئوتیدی ۶۰K نشان داد که فراوانی آلل‌های کم‌یاب کم‌تر از ۰/۱ کم‌تر از سایر فراوانی‌ها بود (۱۳ درصد کل داده‌ها). در این فن آوری، تعداد بسیار کمی از داده‌های چندشکلی‌های تک‌نوکلئوتیدی، فراوانی آللی نزدیک به صفر و بیشتر

آن‌ها فراوانی بین ۰/۴ تا ۰/۵ داشتند. این در حالی است که برای داده‌های ژنوتیپ کردن از طریق توالی‌یابی، فراوانی آلل‌های کم‌یاب کم‌تر از ۰/۱، بیشتر از سایر فراوانی‌ها بود (۲۰ درصد کل داده). در طراحی آرایه‌ها، ابتدا نشان‌گرهای چندشکلی‌های تک‌نوکلئوتیدی در چندین جمعیت مختلف ارزیابی شده و تنها چندشکلی‌های تک‌نوکلئوتیدی که فراوانی اللی بالاتر از یک درصد دارند در آرایه وارد می‌شوند. در مقابل، روش‌های توالی‌یابی نظیر ژنوتیپ کردن از طریق توالی‌یابی، واریانت‌هایی در سرتاسر طیف فراوانی اللی را آشکار می‌سازند. مطالعات متعددی تفاوت توزیع فراوانی آلل‌های کم‌یاب در فن‌آوری‌های مختلف را گزارش کرده است. در پژوهش Romay و همکاران (۲۰۱۳) با استفاده از فن‌آوری ژنوتیپ کردن از طریق توالی‌یابی تعداد ۶۸۱۲۵۷ چندشکلی تک‌نوکلئوتیدی شناسایی شد که بیش از نیمی از این چندشکلی‌های تک‌نوکلئوتیدی از دسته آلل‌های کم‌یاب بودند. در مطالعه‌ی دیگری، این تفاوت برای فراوانی آلل‌های کم‌یاب کم‌تر از ۰/۰۵ حدود ۲۲ درصد برای فن‌آوری ژنوتیپ کردن از طریق توالی‌یابی در مقایسه با ۶ درصد برای آرایه‌ی ۵۰k بود (Negro *et al.*, 2019). نرم‌افزارهای موجود، توان آماری لازم برای تجزیه و تحلیل آلل‌های کم‌یاب را ندارند. با وجود این، پژوهش‌های متعددی استفاده از این الل‌ها برای توجیه وراثت‌پذیری از دست رفته، به‌عنوان پدیده‌ای که در آن بخش بزرگی از واریانس ژنتیکی یک صفت توسط جایگاه‌های شناسایی شده در مطالعات پویش کل ژنوم تفسیر نمی‌شود، را توصیه کرده‌اند (Gibson, 2012). با این‌که تعداد آلل‌های کم‌یاب، در داده‌های حاصل از فن‌آوری ژنوتیپ کردن از طریق توالی‌یابی، فراوان است، اما در پژوهش‌هایی که با استفاده از داده‌های حاصل از این فن‌آوری، مطالعات پویش کل ژنوم انجام شده، نتایج موفق‌تری گزارش شده است، زیرا با وجود این‌که در اکثر پژوهش‌ها، تمام واریانت‌های کم‌تر از ۰/۱ حذف می‌شوند، اما تعداد بسیار زیاد واریانت‌های باقیمانده، مطالعات پویش ژنوم را بسیار مناسب کرده است (Pingxian *et al.*, 2019; Haobin *et al.*, 2020).

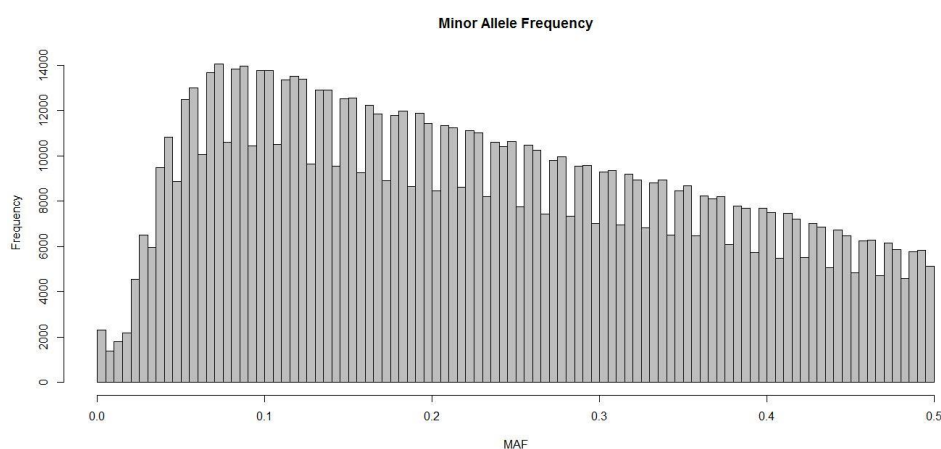


شکل ۵. هیستوگرام فراوانی MAF داده‌های حاصل از آرایه چندشکلی تک‌نوکلئوتیدی ۶۰K. محور X فراوانی آلل‌های کم‌یاب و محور Y تعداد نشان‌گرها است.

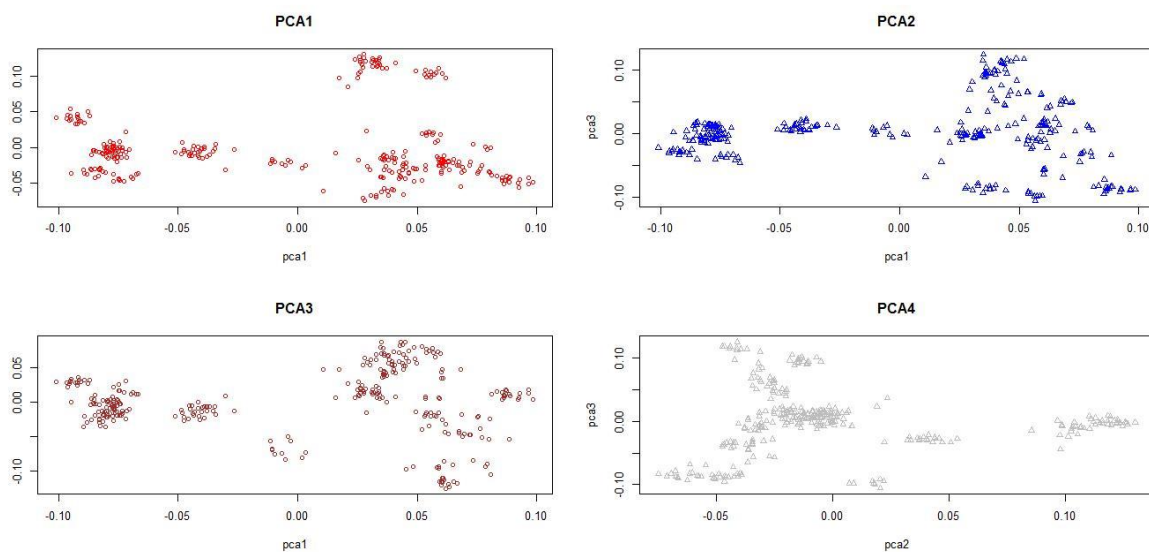
ساختارهای جمعیتی و خویشاوندی با افزایش داده‌های مثبت کاذب می‌تواند نتایج مطالعات پویش ژنومی را با مشکل مواجه کند. روش تجزیه و تحلیل مولفه‌های اصلی (PCA^۱) آزمونی را برای وجود ساختار جمعیتی در داده‌های ژنتیکی بر اساس همبستگی ژنتیکی بین افراد فراهم می‌کند (Patterson *et al.*, 2006). گراف‌های تجزیه و تحلیل مولفه‌های اصلی برای داده‌های آرایه‌ی چندشکلی تک‌نوکلئوتیدی ۶۰K و ژنوتیپ کردن از طریق توالی‌یابی، که با استفاده از برنامه PLINK 1.9 (Purcell *et al.*, 2007) در نرم‌افزار R انجام شد، برای نمایش بهترین تفکیک‌ها بر اساس مولفه‌های اصلی، به ترتیب، در

1 Principle Component Analysis

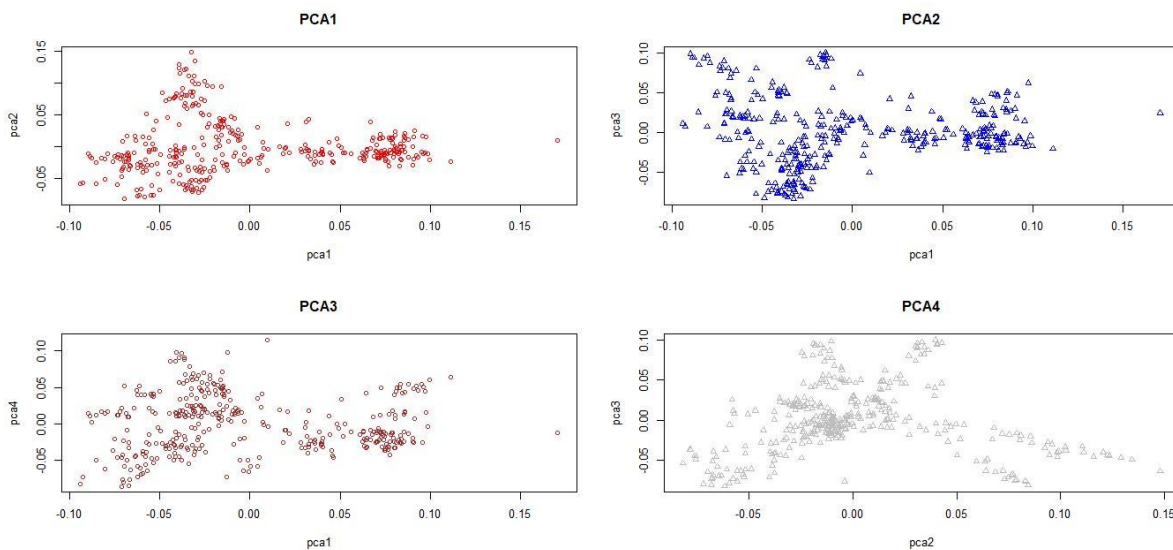
شکل‌های ۷ و ۸ ارایه شده است. به‌طور کلی، میزان واریانس قابل توصیف حاصل از چهار مولفه‌ی اصلی اول در فناوری ژنوتیپ‌کردن از طریق توالی‌یابی (۶۶/۴ درصد) اندکی کمتر از فناوری آرایه (۷۰/۰ درصد) بود. علاوه بر این، نتایج مولفه‌های اصلی در آرایه‌ی چندشکلی تک‌نوکلئوتیدی ۶۰K نشان داد که جمعیت‌ها به شکل توده‌ای در چندین کلاستر تفکیک شده‌اند. از آن جایی که فراوانی آلی و ژنوتیپی بین زیر جمعیت‌ها متفاوت از یکدیگر است و ممکن است اثر آلل یا ژنوتیپی، که ویژگی یک زیر جمعیت است، به طور کاذب معنی‌دار شود (Pritchard *et al.*, 2000)، لازم است داده‌ها در مطالعات پویس کل ژنوم برای این اثر تصحیح شوند. برای داده‌های ژنوتیپ‌کردن از طریق توالی‌یابی، نتایج مولفه‌های اصلی نشان از تنوع زیاد چندشکلی-های تک‌نوکلئوتیدی در این روش دارد، بنابراین، به علت یکنواختی قابل قبول مشاهده شده، نیاز به تصحیح داده‌ها برای مطالعات پویس کل ژنوم نیست.



شکل ۶. هیستوگرام فراوانی MAF داده‌های حاصل از ژنوتیپ‌کردن از طریق توالی‌یابی. محور X فراوانی آلل‌های کم‌یاب و محور Y تعداد نشان‌گرها است.



شکل ۷. گراف‌های تفکیک مولفه‌های اصلی در داده‌های حاصل از آرایه‌ی چندشکلی تک‌نوکلئوتیدی ۶۰K



شکل ۸. گرافهای تفکیک مولفه‌های اصلی در داده‌های حاصل از ژنوتیپ کردن از طریق توالی‌یابی

نتیجه‌گیری و پیشنهادها

در مطالعه‌ی حاضر، تعداد نشان‌گرها و بلوک‌های هاپلو تیپ‌یابی داده‌های حاصل از فن‌آوری ژنوتیپ کردن از طریق توالی‌یابی نسبت به آرایه‌ی چندشکلی تک‌نوکلئوتیدی ۶۰K بسیار بیشتر بود. میزان عدم تعادل پیوستگی در فن‌آوری آرایه‌ی چندشکلی تک‌نوکلئوتیدی ۶۰K بیشتر از فن‌آوری تعیین ژنوتیپ از طریق توالی‌یابی بود و در هر دو روش این میزان تحت تاثیر طول کروموزوم قرار داشت به طوری که با افزایش فواصل فیزیکی، عدم تعادل ناشی از پیوستگی کاهش داشت. علاوه بر این، برخلاف فن‌آوری آرایه‌ی چندشکلی تک‌نوکلئوتیدی، واریانت‌های با فراوانی کمتر از ۰/۸ در فن‌آوری ژنوتیپ کردن از طریق توالی‌یابی بسیار زیاد بود. با وجود این که استفاده از روش جانپی بسیاری از واریانت‌های گم‌شده را شناسایی کرد، اما با حذف این نشان‌گرها نیز، تعداد واریانت‌های باقیمانده‌ی حاصل از این فن‌آوری بسیار بیشتر از آرایه‌ی چندشکلی تک‌نوکلئوتیدی ۶۰K بود. برای داده‌های ژنوتیپ کردن از طریق توالی‌یابی، تنوع چندشکلی‌های تک‌نوکلئوتیدی بسیار بیشتر از داده‌های آرایه‌ی ۶۰K بود. برای مطالعه‌ی حاضر، با توجه به این که این تنوع، ساختارهای جمعیتی و خویشاوندی یکنواختی را ایجاد می‌کند در مطالعات پویس ژنوم با استفاده از داده‌های ژنومی، داده‌های ژنوتیپ کردن از طریق توالی‌یابی نیازی به تصحیح نیست. همزمان با افزایش کمیت و کیفیت اطلاعات حاصل از توالی‌یابی و نیز کاهش هزینه‌ها به‌ازای هر نمونه، ژنوتیپ کردن از طریق توالی‌یابی تبدیل به یک فن‌آوری متداول و پرترفدار در پژوهش‌های حوزه ژنومی کشاورزی شده است. بنابراین با توجه به مجموع نتایج حاصل از مطالعه حاضر، پیشنهاد می‌شود از داده‌های حاصل از فن‌آوری ژنوتیپ کردن از طریق توالی‌یابی برای مطالعات پویس ژنوم صفات متفاوت در جمعیت‌های حیوانات اهلی (مرغ) استفاده شود.

تشکر و قدردانی

بدین‌وسیله از پروفسور Just Jensen از دانشگاه آرهوس دانمارک برای تأمین هزینه‌های تعیین ژنوتیپ نشان‌گرها در فن‌آوری آرایه‌ی چندشکلی تک‌نوکلئوتیدی ۶۰K، از شرکت تجاری کاب به دلیل در اختیار قرار دادن این تراشه، و از حمایت مالی دانشگاه تربیت مدرس در تعیین ژنوتیپ داده‌ها از طریق توالی‌یابی (GBS)، که توسط شرکت الشایر نیوزیلند انجام شد، قدردانی می‌شود.

منابع

جوانروح علی‌آباد، علی؛ واعظ ترشیزی، رسول؛ مسعودی، علی‌اکبر و احسانی، علیرضا (۱۳۹۵). مطالعه پویس کل ژنومی صفات کیفی گوشت در جمعیت F₂ حاصل از مرغ بومی آذربایجان غربی و سویه گوشتی آرین. *تولیدات دامی*-دانشگاه تهران. دوره ۱۸-شماره ۴-صفحه‌های ۶۹۷-۷۰۹.

مقصودی، علی؛ واعظ ترشیزی، رسول؛ مسعودی، علی‌اکبر و کریمی ترشیزی، محمدمیر (۱۳۹۲). دوشکلی جنسی در پاسخ ایمنی سلولی و خونی دو سویه از طیور تجاری لاین آرین و بومی آذربایجان غربی. *مجله علوم دامی ایران*-دانشگاه تهران. دوره ۴۴-شماره ۴-صفحه‌های ۳۶۷-۳۷۴.

REFERENCES

- Albrechtsen, A., Nielsen, F. C., & Nielsen, R. (2010). Ascertainment biases in SNP chips affect measures of population divergence. *Molecular Biology Evolution*, 27(11), 2534–2547.
- Bradbury, P. J., Zhang, Z., Kroon, D. E., Casstevens, T. M., Ramdoss, Y. & Buckler, E. S. (2007). TASSEL: software for association mapping of complex traits in diverse samples. *Bioinformatics*. 23(19), 2633-2635.
- Brodie, A., Azaria, J. R., & Ofran, Y. (2016). How far from the SNP may the causative genes be? *Nucleic Acids Research*, 44(13), 6046-54.
- Davey, J. W., Hohenlohe, P. A., Etter, P. D., Boone, J. Q., & Catchen, J. M. (2011). Genome-wide genetic marker discovery and genotyping using next-generation sequencing. *Nature Reviews Genetics*, 12, 499–510.
- Donato, M., Peters, S. O., Mitchell, S. E., Hussain, T., & Imumorin, I. G. (2013). Genotyping-by-sequencing (GBS): a novel, efficient and cost-effective genotyping method for cattle using next-generation sequencing. *PLoS One*, 8(5), e62137.
- Elshire, R. J., Glaubitz, J. C., Sun, Q., Poland, J. A., Kawamoto, K., Buckler, E. S., & Mitchell, S. E. (2011). A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species, *PLoS One*, 6(5), e19379.
- Emara, M. G., & Kim, H. (2003). Genetic markers and their application in poultry breeding. *Poultry Science*, 82, 952–957.
- Fiedler, J. D., Lanzatella, C., Okada, M., Jenkins, J., Schmutz, J., & Tobias, C. M. (2015). High-density single nucleotide polymorphism linkage maps of lowland switchgrass using genotyping-by-sequencing. *The Plant Genome*, 8(2).
- Franssen S. U., Barton N. H., Schlotterer C. (2017). Reconstruction of Haplotype-Blocks Selected during Experimental Evolution. *Molecular Biology and Evolution*, 34(1):174-184.
- Gabriel, S. B., Schaffner, S. F., Nguyen, H., Moore, J. M., Roy, J., Blumenstiel, B., Higgins, J., Defelice, M., Lochner, A., Ggart, M., Liu-cordero, S. N., Rotimi, C., Adeyemo, A., Cooper, R., Ward, R., Lander, E. S., Daly, M. J., & Altshuler, D. (2002). The structure of haplotype blocks in the human genome. *Science*, 296, 2225-2229.
- Gibson, G. (2012). Rare and common variants: twenty arguments. *Nature Reviews Genetics*, 13, 135-145.
- Glaubitz, J. C., Casstevens, T. M., Lu, F., Harriman, J., Elshire, R. J., Sun, Q., & Buckler E. S. (2014). TASSEL-GBS: A High Capacity Genotyping by Sequencing Analysis Pipeline, *PLoS ONE*, 9(2), e90346.
- Groenen M. A., Wahlberg P., Foglio M., Cheng H. H., Megens H. J., Crooijmans R.P., Besnier F., Lathrop M., Muir W. M., Wong G. K., Gut I., & Andersson L. (2009). A high-density SNP-based linkage map of the chicken genome reveals sequence features correlated with recombination rate. *Genome Research*, 19(3):510-9.
- Haobin, H., Xiaoliang, W., Caiyun, Z., Yingying, T., Wenwei, L., Xia, C., Zhigang, X., Junfeng, Y., & Changsuo, Y. (2020). Genomic analysis of GBS data reveals genes associated with facial pigmentation in Xinyang blue-shelled layer. *Archives Animal Breeding*, 63, 483–491.
- Horvath D. P., Stamm M., Talukder Z. I., Fiedler J., Horvath A. P., Horvath G. A., Chao W. S.,

- Anderson J.V. (2020). A New Diversity Panel for Winter Rapeseed (*Brassica napus* L.) Genome-wide association studies. *Agronomy*, 10(12):2006.
- He, J., Zhao, X., Laroche, A., Lu, Z. X., Liu, H., & Li, Z. (2014). Genotyping-by-sequencing (GBS), an ultimate marker assisted selection (MAS) tool to accelerate plant breeding. *Frontiers in Plant Science*, 5, 484.
- Heslot, N., Rutkoski, J., Poland, J., Jannink, J. L., & Sorrells, M. E. (2013). Impact of marker ascertainment bias on genomic selection accuracy and estimates of genetic diversity. *PLoS One*, 8(9), e74612.
- Huang, Y. F., Poland, J. A., Wight, C. P., Jackson, E. W., & Tinker, N. A. (2014). Using genotyping-by-sequencing (GBS) for genomic discovery in cultivated oat. *PLoS ONE*, 9(7), e102448.
- Illumina. (2014). sequence-based genotyping brings agrigenomics to a crossroads. *Application Spotlight: Agrigenomics*.
- Javanrouh-Aliabad, A., Vaez Torshizi, R., Masoudi, A. A., & Ehsani, A. (2016). Genome-wide association study for meat quality traits in an F2 intercross between Azerbaijan native chickens and Ariyan broiler line. *Animal Production*. 18(4), 697-709 (in Persian).
- Khanyile, K. S., Dzomba, E. F., & Muchadeyi, F. C. (2015). Haplo-block structure of Southern African village chicken populations inferred using genome-wide SNP data. *Genetics and Molecular Research*, 14(4), 12276-12287.
- Lander, E. S., & Schork, N. J. (1994). Genetic dissection of complex traits. *Science*, 265, 2037-2048.
- Liao, R., Wang, Z., Chen, Q., Tu, Y., Chen, Z., Wang, Q., Yang, C., Zhang, X., & Pan, Y. (2015). An efficient genotyping method in chicken based on genome reducing and sequencing. *PLoS One*, 10(8), e0137010.
- Loureiro, L. O., Engstrom, M. D. & Lim, B. K. (2020). Optimization of genotype by sequencing data for phylogenetic purposes. *MethodsX*, 20(7):100892.
- Maghsoudi, A., Vaez Torshizi, R, Masoudi, A. A., & Karimi Torshizi, M. A. (2014). Sexual dimorphism in cellular and humoral immune responses in two strains of commercial Iranian Arian vs Western Azarbaijan native fowls. *Iranian Journal of Animal science*, 44, 367-374 (in Persian).
- Mascher, M., Schuenemann, V. J., Davidovich, U., Marom, N., Himmelbach, A., Hübner, S., Korol, A., David, M., Reiter, E., Riehl, S., Schreibe, M., Vohr, S. H., Green, R. E., Dawson, I. K., Russell, J., Kilian, B., Muehlbauer, G. J., Waugh, R., Fahima, T., Weiss, J. K. E., & Stein, N. (2016). Genomic analysis of 6,000-year-old cultivated grain illuminates the domestication history of barley. *Nature Genetics*, 48(9):1089-1093.
- Negro, S. S., Millet, E. J., Madur, D., Bauland, C., Combes, V., Welcker, C., Tardieu, F., Charcosset, A. and Nicolas, S. (2019). Genotyping-by-sequencing and SNP-arrays are complementary for detecting quantitative trait loci by tagging different haplotypes in association studies. *BMC Plant Biology*, 19, 318.
- Patterson, N., Price, A. L., & Reich, D. (2006). Population structure and eigenanalysis. *PLoS Genetetics*, 2(12), e190.
- Pértille, F., Guerrero-Bosagna, C., Henrique da Silva, V., Boschiero, C., de Ribamar da Silva Nunes, J., Corrêa Ledur, M., Jensen, P., & Coutinho, L. L. (2016). High-throughput and cost-effective chicken genotyping using next- generation sequencing. *Scientific Reports*, 6, 26929.
- Pértille, F., Moreira1, G. C. M., Zanella, R., de Ribamar da Silva Nunes, J., Boschiero, Clarissa., Rovadoscki1,G. A., Mourão, G. B., Ledur, M. C., & Coutinho, L. L. (2017). Genome-wide association study for performance traits in chickens using genotype by sequencing approach. *Nature*, 7, 41748.
- Pingxian, W. u., Kai, W., Jie, Z., Dejuan, C., Qiang, Y., Xidi, Y., Yihui, L., Bo, F., Anan, J., Linyuan, Shen., Weihang, X., Yanzhi, J., Li, Z., Yangshuang, Z., Xu, X., Xuwei, L., & Guoqing, T., (2019). GWAS on imputed whole-genome resequencing from genotyping-by-sequencing data for farrowing interval of different parities in pigs. *Frontiers in Genetics*, 10:1012.
- Pook, T., Mayer, M., Geibel, J., Weigend, S., Cavero, D., Schoen, C. C. & Simianer, H. (2020). Improving imputation quality in BEAGLE for crop and livestock data. *G3 (Bethesda)*, 7:10(1),

177-188.

- Pritchard, J. K., Stephens, M., Rosenberg, N.A., & Donnelly, P. (2000). Association mapping Whole-genome association and population-based linkage analyses. *The American Journal of Human Genetics*, 81, 559-575.
- Purcell, S., Neale, B. Todd-Brown, K., Thomas, L., Ferreira, M. A. R., Bender, D., Maller, J., Sklar, P., de Bakker, P. I. W., Daly, M. J., & Sham, P. C. (2007). PLINK: a tool set for in structured populations. *American Journal of Human Genetics*, 67, 170-81.
- Romay, M. C., Millard, M. J., Glaubitz J. C., Peiffer, J. A., Swarts, K. L., Casstevens, T. M., Elshire, R. J., Acharya, C. B., Mitchell, S. E., Flint-Garcia, S. A., McMullen, M. D., Holland, J. B., Buckler, E., & Gardne, C. A. (2013). Comprehensive genotyping of USA national maize inbred seed bank. *Genome Biology*, 14:R55.
- Sehgal, D., Mondal, S., Crespo-Herrera, L., Velu, G., Juliana, P., Huerta-Espino, J., Shrestha, S., Poland, J., Singh, R. & Dreisigacker, S. (2020). Haplotype-based, genome-wide association study reveals stable genomic regions for grain yield in CIMMYT spring bread wheat. *Frontier Genetics*, 3:11, 589490.
- Sukumaran, S., Li, X., Li, X., Zhu, C., Bai, G., Perumal, R., Tuinstra, M. R., Prasad, P. V., Mitchell, S. E., Tesso, T. T., & Yu, J. (2016) QTL mapping for grain yield, flowering time, and stay-green traits in sorghum with genotyping-bysequencing markers. *Crop Science*, 56, 1429-1442.
- Thomson, M. J. (2014). High-throughput SNP genotyping to accelerate crop improvement. *Plant Breeding Biotechnology*, 2, 195–212.
- Wang, X., Su, G., Hao, D., Lund, M. S. & Kadarmideen, h. N. (2020). Comparisons of improved genomic predictions generated by different imputation methods for genotyping by sequencing data in livestock populations. *Journal of Animal Science and Biotechnology*, 11, 3.
- Zhao, F. P., Wang, G. K., Zeng, T., Wei, C. H., Zhang, L., Wang, H. H., Zhang, S. Z., Liu, R. Z., Liu, Z., & Du, L. X. (2014). Estimations of genomic linkage disequilibrium and effective population sizes in three sheep populations. *Livestock Science*, 170, 22–29.