Shahid Bahonar University of Kerman
Agricultural Biotechnology Journal
Iranian Biotechnology Society
Print ISSN: 2228-6705 Online ISSN: 2228-6500

The Study of Genetic Diversity in Khazak Native Chicken Using Whole Genome Sequencing

Elaheh Rostamzadeh Mahdabi

PhD Student, Department of Animal Science, Shahid Bahonar University of Kerman, Kerman, Iran. E-mail: E_rostamzadeh@agr.uk.ac.ir

Masood Asadi Fozi

* Corresponding author. Professor, Department of Animal Science, Shahid Bahonar University of Kerman, Kerman, Iran. E-mail: masadi@uk.ac.ir

Ali Esmailizadeh

Professor, Department of Animal Science, Shahid Bahonar University of Kerman, Kerman, Iran. E-mail: aliesmaili@uk.ac.ir

Ahmad Ayatollahi Mehrgardi

Associate Professor, Department of Animal Science, Shahid Bahonar University of Kerman, Kerman, Iran. E-mail: a_ayatmehr@yahoo.com

Seyed Hojat Masoudzadeh

MSc. of animal breeding, Department of Animal Science, Shahid Bahonar University of Kerman, Kerman, Iran. E-mail: masoudzadeh1360@yahoo.com

Abstract

Objective

Indigenous animals are considered as the national treasures and strategic sources of each country. It is essential to preserve and reproduce them. Since no artificial selection has been employed in Iranian native breeds, there is considerable diversity in their genome. Therefore, having detailed information on the animals' genome can be important for designing breeding plans in the future. The aim of this study was to identify the functional effects of single nucleotide polymorphisms and investigate the genetic diversity in the Khazak chicken ecotype using sequence information of the whole genome.

Materials and methods

In this study, 17 Khazak chicken samples were sequenced using Next Generation Sequencing (NGS) technique. After preprocessing short sequences using FastQC, Trimmomatic, BWA,

Picard software and the GATK program, single nucleotide polymorphisms were identified in GATK program using the UnifiedGenotyper tool. Samtools software was applied to calculate short reads coverage and the percentage of read alignment with the reference genome. The filtering steps of SNPs were performed using Plink software. The genetic diversity within the target population was calculated using the VCFtools program.

Results

The mean coverage depth for 17 chickens was 6.45. After preprocessing the short sequences, were identified 15,981,176 and 10,552,735 SNPs before and after the quality control, respectively. Functional annotation of the single nucleotide polymorphisms indicated that most of the SNPs are located in the intron and intergenic regions and only 2.8% of them belong to the exon region. The rate of Transition to transversion mutations (Ti / Tv) was 2.43 and the ratio of heterozygous to homozygous SNPs was 3.3 in the Khazak birds, on average.

Conclusions

The higher number of heterozygous SNPs indicates that there is a high genetic diversity in the Khazak ecotype. Like other Iranian native chickens, the ecotype has the genetic diversity that can be considered a native laying ecotype in the breeding programs.

Keywords: Genetic diversity, single nucleotide polymorphisms, Khazak native chicken.

Paper Type: Research Paper.

Citation: Rostamzadeh Mahdabi E, Asadi Fozi M, Esmailizadeh A, Ayatollahi A, Masoudzadeh H (2021) The study of genetic diversity in Khazak native chicken using whole genome sequencing. *Agricultural Biotechnology Journal* 13 (3), 91-106.

Agricultural Biotechnology Journal 13 (3), 91-106. DOI: 10.22103/jab.2021.17202.1296

Received: July 20, 2021. Accepted: August 22, 2021.

Publisher: Faculty of Agriculture and Technology Institute of Plant Production, Shahid Bahonar University of Kerman-Iranian Biotechnology Society.

© the authors

مجله بیوتکنولوژی کشاورزی
شاپا چاپی: ۶۷۰۵-۲۲۲۸ شاپا الکترونیکی: ۶۵۰۰-۲۲۲۸

بررسی تنوع ژنتیکی در مرغ بومی خزک با استفاده از داده‌های توالی‌یابی کل ژنوم

الهه رستم‌زاده مهدابی

دانشجوی دکتری ژنتیک و اصلاح نژاد دام، بخش علوم دامی، دانشکدهٔ کشاورزی، دانشگاه شهید باهنر کرمان، ایران. رایانامه: E_rostamzadeh@agr.uk.ac.ir

مسعود اسدی فوزی

*نویسنده مسئول: استاد بخش علوم دامی، دانشکدهٔ کشاورزی، دانشگاه شهید باهنر کرمان، ایران. رایانامه: masadi@uk.ac.ir

علی اسمعیلی‌زاده

استاد بخش علوم دامی، دانشکدهٔ کشاورزی، دانشگاه شهید باهنر کرمان، ایران. رایانامه: aliesmaili@uk.ac.ir

احمد آیت الهی مهرجردی

دانشیار بخش علوم دامی، دانشکدهٔ کشاورزی، دانشگاه شهید باهنر کرمان، ایران. رایانامه: a_ayatmehr@yahoo.com

سید حجت مسعودزاده

دانش آموخته مقطع کارشناسی ارشد ژنتیک و اصلاح دام، بخش علوم دامی، دانشکدهٔ کشاورزی، دانشگاه شهید باهنر کرمان، ایران. رایانامه: masoudzadeh1360@yahoo.com

تاریخ دریافت: ۱۴۰۰/۰۴/۲۹ تاریخ پذیرش: ۱۴۰۰/۰۵/۳۱

چکیده

هدف: حیوانات بومی به عنوان سرمایه ملی و ذخایر استراتژیک هر کشوری محسوب می‌شوند. حفظ و تکثیر آن‌ها از ارزش و اهمیت بسیاری برخوردار می‌باشد. با توجه به اینکه تاکنون انتخاب مصنوعی در مرغان نژاد بومی ایران انجام نشده است، تنوع قابل توجهی در ژنوم آن‌ها مشاهده می‌شود. از این رو کسب اطلاعات و شناخت دقیق‌تر ژنوم این حیوانات برای برنامه‌های اصلاح نژادی ضروری می‌باشد. بنابراین، هدف از انجام این تحقیق، شناسایی و بررسی آثار عملکردی چندشکلی‌های تک نوکلئوتیدی و همچنین بررسی تنوع ژنتیکی در ۱۷ مرغ بومی خزک با استفاده از اطلاعات توالی کل ژنوم می‌باشد.

مواد و روش‌ها: در این پژوهش، تعداد ۱۷ نمونه مرغ خزک با استفاده از روش توالی‌یابی نسل بعد (NGS) توالی‌یابی شد. بعد از انجام مراحل پیش پردازش توالی‌های کوتاه با استفاده از نرم‌افزارهای FastQC، Trimmomatic، BWA، Picard و ابزارهای برنامه GATK، در نهایت چند شـــکلی‌های تک نوکلئوتیدی با اســـتفاده از ابزار UnifiedGenotyper در برنامه GATK شناسایی شدند. جهت محاسبه میزان پوشش خوانش‌ها و درصد الاینمنت خوانش‌ها با ژنوم مرجع از نرم افزار Samtools استفاده شـد. مراحل کنترل کیفیت چندشـکلی‌های تک نوکلئوتیدی با اسـتفاده از نرم افزار Plink انجام شـد، سـپس تنوع ژنتیکی درون جمعیت مورد نظر با برنامه VCFtools محاسبه شد.

نتایج: میانگین عمق پو شش توالی‌ها برای ۱۷ مرغ خزک ۶/۴۵ بد ست آمد. بعد از پردازش توالی‌های کوتاه، تعداد ۱۵،۹۸۱،۱۷۶ و ۱۰،۵۵۲،۷۳۵ SNP به ترتیب قبل و بعد از مرحله کنترل کیفیت شناسایی شد. نتایج حاشیه‌نویسی عملکردی چند شکلی‌های تک نوکلئوتیدی نشـــان داد که اکثر SNPها در نواحی اینترون و بین ژنی قرار دارند و تنها ۲/۸ درصـــد آنها متعلق به ناحیه اگزون می‌باشند. نسبت جهش‌های انتقالی به جهش‌های متقاطع (Ti / Tv) ۲/۴۳ و نسبت SNPهای هتروزیگوت به هموزیگوت به طور میانگین در مرغان خزک ۳/۳ بدست آمد.

نتیجه‌گیری: تعداد SNPهای هتروزیگوت بالاتر، نشان دهنده‌ی تنوع ژنتیکی بالا در اکوتیپ خزک می‌باشد. این اکوتیپ مانند سایر اکوتیپ‌های بومی ایران دارای تنوع می‌باشد که می‌تواند به عنوان یک اکوتیپ تخم‌گذار بومی در برنامه‌های اصلاح نژاد حائز اهمیت باشد.

کلیدواژه‌ها: تنوع ژنتیکی، چند شکلی تک نوکلئوتیدی، مرغ بومی خزک.

نوع مقاله: پژوهشی.

استناد: رستم‌زاده الهه، اسدی فوزی مسعود، اسمعیلی‌زاده علی، آیت الهی احمد، مسعودزاده حجت (۱۴۰۰) بررسی تنوع ژنتیکی در مرغ بومی خزک با استفاده از داده‌های توالی‌یابی کل ژنوم . *مجله بیوتکنولوژی کشاورزی*، ۱۳(۳)، ۹۱-۱۰۶.

Publisher: Faculty of Agriculture and Technology Institute of Plant Production, Shahid Bahonar University of Kerman-Iranian Biotechnology Society.

© the authors

مقدمه

تکامل حیوانات اهلی طی ۳/۵ میلیون سال، با توجه به وجود تنوع ژنتیکی بسیار زیاد و همچنین سازگاری با محیط اطراف شکل گرفته است. انسان با توجه به نیازهایش و با استفاده از تنوع ژنتیکی موجود در حیوانات و گیاهان، شروع به دستکاری در تکامل آن‌ها کرده است که این فرآیند به عنوان اهلی‌سازی نیز شناخته می‌شود (Diamond 2002) مطالعات متعدد، زمان‌ها و مکان‌های

مختلفی را برای اهلی‌شدن مرغ از گونه‌های مرغ جنگلی گزارش کرده‌اند که اکثر آنها نشان می‌دهند که مرغ اهلی از مرغ جنگلی قرمز در جنوب و جنوب شرقی آسیا نشات گرفته است (Sawai et al. 2010; Tixier-Boichard et al. 2011; Peters et al. 2015; Osman et al. 2016). پرورش ماکیان در ایران و انتشار آن از طریق این کشور تاریخچه‌ای بسیار کهن دارد. ایران (پرشیا) یک امپراطوری بزرگ از قرن ۵ قبل از میلاد تا تقریبا قرن ۷ میلادی بود و از هند (دهلی) تا دریاهای سیاه و مدیترانه گسترده بود (Moazeni et al. 2016a; Mohammadifar and Mohammadabadi 2017). در آن زمان و بعد از آن، در قرون وسطی ایران در محل تقاطع راه‌ها برای حمل و نقل محصولات، از قبیل ماکیان از شرق به غرب، هم از طریق خشکی و هم از طریق دریا قرار داشت. جنگ های زیادی در حوالی ایران و کشورهای همسایه در طی این دوره‌ها نیز توسعه و گسترش جمعیت‌های ماکیان را تسهیل کرد. حفاری‌های باستان شناسی حضور ماکیان را در ایران در زمان‌های باستان تأیید کرده است (Nikbakhti et al., 2009; Moazeni et al. 2016b). بر اساس تحقیقات West and Zhou استخوان‌های یافت شده در ایران در سه منطقه وجود داشته‌اند: دو کشف در تپه یحیی (Tepe Yahya) (جنوب شرقی ایران) به ترتیب متعلق به ۳۸۰۰ تا ۳۹۰۰ قبل از میلاد و ۱۰۰۰ قبل از میلاد و دیگری در تخت سلیمان (شمال غربی ایران) متعلق به ۱۰۰۰ قبل از میلاد (Mohammadifar et al. 2014; Mohammadifar and Mohammadabadi 2018). اکثر نژادهای بومی ایران از دسته مرغ‌های آسیایی می‌باشند. کشور ایران در پهنه‌ی گسترده خود دارای آب و هوای بسیار متنوع و در نتیجه تحت عوامل محیطی گوناگون می‌باشد. عوامل محیطی مختلف موجب تشکیل انواع زیستگاه‌ها و تنوع حیوانات مناطق مختلف ایران شده است (Tavakolian 2000). نژادهای بومی به دلیل سازگاری به شرایط محیطی و مقاومت در برابر بیماری‌ها دارای اهمیت می‌باشند. مرغ بومی نقش مهمی در خانواده‌های روستایی به عنوان منبع پروتئین حیوانی با کیفیت بالا و ایجاد اشتغال و افزایش درآمد روستاییان دارد. اهمیت نژادهای بومی برای اقتصاد روستاییان در کشورهای در حال توسعه و توسعه یافته و عمدتاً در آسیا بسیار زیاد می‌باشد. مرغ‌های بومی تولید تخم پایین و سرعت رشد کندی نسبت به نژادهای تجاری دارند، اما توانایی تحمل شرایط سخت محیطی و شیوه‌های نامناسب دامپروری را دارند. مرغ‌های نژاد بومی مخزن ژنوم و ژن‌های اصلی برای اصلاح مرغ با عملکرد بالا، سازگاری با مناطق گرمسیری و مقاومت در برابر بیماری‌ها می‌باشند (Padhi 2016). به عبارت دیگر منابع ژنتیکی مرغان بومی می‌تواند زمینه را برای تولید نژادهای با تولید مناسب و سازگار با شرایط روستایی فراهم کند (Hoffmann 2005).

جامع‌ترین روش بدست آوردن اطلاعات در مورد ژنوم هر موجود زنده، تعیین ترتیب دقیق نوکلئوتیدها، معروف به توالی‌یابی، در توالی کامل DNA می‌باشد (Venter et al. 2015). ظهور توالی‌یابی نسل جدید (NGS[1]) باعث کشف میلیون‌ها واریانت ژنتیکی و اثرات عملکردی آنها در ژنوم مهره‌داران شده است (Gheyas et al. 2015b). از نشانگرهای DNA می‌توان برای برآورد تنوع ژنتیکی در داخل و بین جمعیت‌ها استفاده کرد (Boettcher et al. 2010). نشانگرهای DNA مختلفی مانند

[1]. Next Generation Sequencing

RAPDs، AFLPs و میکروساتلایت‌ها برای ارزیابی تنوع ژنتیکی مرغ استفاده شده‌اند (Lynch & Milligan 1994; Abebe et al. 2015). که به طور خاص SNPها به دلیل تراکم ژنومی بالا، برای برآورد تنوع ژنتیکی و تعیین دقیق ساختار جمعیت مورد استفاده قرار می‌گیرند (Engelsma 2012; Strillacci et al. 2017).

مرغ خزک، بومی منطقه سیستان (زابل) می‌باشد و در بعضی شهرستان‌های دیگر به ندرت یافت می‌شود. این پرنده به عنوان تیپ تخم‌گذار شناخته می شود که میانگین تخم‌مرغ تولیدی آنها ۱۲۰ عدد در سال و متو سط وزن تخم‌مرغ ۴۵/۱۵ گرم می‌با شد. این مرغ جثه کوچک با پاهای کوتاه دارد. وزن بدن در مرغ و خروس اکوتیپ خزک به طور میانگین به ترتیب ۱۲۴۵ و ۱۵۷۸ گرم می‌باشد به طوری‌که وزن پایین این اکوتیپ می‌تواند دارای این مزیت باشد که اکثر خوراک مصرفی مرغ، جهت تولید تخم مصرف می شود و خوراک کمتری را برای انرژی نگهداری م صرف می‌کند و در نتیجه با هزینه کمتری تخم‌مرغ تولید می شود (Deljoy Sarayan 2011; Tavakolian 2000). با توجه به اینکه انتخاب مصنوعی در مرغان نژاد بومی انجام نشده است می‌توان انتظار دا شت که تنوع قابل توجهی در ژنوم آنها م شاهده می شود. از این رو لزوم حفظ آنها به عنوان ذخایر ژنتیکی و سرمایه‌های ملی برای برنامه‌های ا صلاح نژادی، ک سب اطلاعات و شناخت دقیق‌تر ژنوم این حیوانات را ضروری می سازد. هدف از انجام این تحقیق، شنا سایی و برر سی آثار عملکردی چند شکلی‌های تک نوکلئوتیدی و همچنین برر سی تنوع ژنتیکی در اکوتیپ مرغ بومی خزک با استفاده از اطلاعات توالی کل ژنوم بود.

مواد و روش‌ها

در این پژوهش نمونه‌های خون مربوط به اکوتیپ خزک از ورید زیر بال گرفته شد. ا ستخراج DNA از نمونه‌های خون ۱۷ مرغ خزک در آزمای شگاه ژنتیک بخش علوم دامی دان شگاه شهید باهنر کرمان انجام گرفت. توالی‌یابی کل ژنوم با ا ستفاده از روش توالی یابی نسل بعد توسط دستگاه ایلومینا Hiseq 2000 در کشور چین انجام شد. کیفیت اولیه داده‌های خروجی توالی‌یابی توسط برنامه FastQC بررسی شد (Andrews 2010) و با استفاده از نرم افزار Trimmomatic توالی‌های کوتاه با کیفیت پایین پیرا یش[۲] شدند (Bolger et al. 2014). جهت هم‌ردیفی توالی‌های کوتاه نسبت به ژنوم مرجع مرغ (ftp://ftp.ensembl.org/pub/release-/fasta/gallus_gallus/dna) از الگوریتم MEM در نرم افزار BWA استفاده شد (Li & Durbin 2009). به منظور عملکرد بهتر، راندمان ذخیره‌سازی و پردازش داده‌ها، خروجی مرحله قبل با استفاده از نرم افزار Samtools به صورت فشرده (BAM file) تبدیل شد (Li et al. 2009). هنگام آماده‌سازی نمونه‌ها به خصوص در زمان ساخت کتابخانه‌های ژنومی، تعداد زیادی از توالی‌های تکراری از یک نمونه یکسان در طول واکنش زنجیره‌ای پلیمراز به وجود

۲. Trim

می‌آیندکه باعث فراخوانی واریانت‌های[۳] مثبت کاذب می‌شود (Ebbert et al. 2016). به همین جهت این توالی‌های کوتاه تکرار شده با استفاده از برنامه Picard (http://broadinstitute.github.io/picard) حذف شدند. بعد از انجام مراحل پیش پردازش فایل ژنوم مرجع با استفاده از نرم افزار Samtools و Picard، جهت همردیفی مجدد[۴] اطراف حذف و اضافه های کوتاه و کالیبراسون مجدد نمره کیفیت بازها از نرم افزار GATK استفاده شد (McKenna et al. 2010). از آنجایی که خوانش‌ها معمولاً کوتاه هستند و می‌توانند حاوی خطاهای توالی‌یابی باشند، یک ناحیه ژنومی بیش از یک بار توالی‌یابی می‌شود. تعداد خوانش‌های همپوشان با یک ناحیه ژنومی عمق[۵] توالی‌یابی ژنوم را مشخص می‌کند. در مطالعه حاضر، جهت محاسبه میزان پوشش یا عمق خوانش‌ها و در صد همردیفی خوانش‌ها با ژنوم مرجع از نرم‌افزار Samtools ا ستفاده شد. در نهایت چند شکلی‌های تک نوکلئوتیدی توسط ابزار UnifiedGenotyper در برنامه ی GATK فراخوانی شدند.

درتحقیق حاضر، تنوع ژنتیکی واقع بر کروموزوم‌های اتوزومی بررسی شد و بنابراین SNPs کروموزوم‌های جنسی (Z, W) و گروه‌های پیوسته[۶] حذف شدند. جهت آنالیز بعدی، لازم است تا داده‌های با کیفیت پایین شناسایی و حذف شوند. به همین منظور با استفاده از نرم افزار Plink افراد با بیش از ۱۰ درصد ژنوتیپ گمشده، SNPهای با فراوانی آللی کمیاب[۷] کمتر از ۰/۰۵ و SNPهای که آستانه‌ی P-value آنها کمتر از $۱۰^{-۷}$ بودند حذف شدند و SNPهای با ۹۰ درصد نرخ فراخوانی تعیین ژنوتیپ[۸] حفظ شدند. بعد از شناسایی چند شکلی‌های تک نوکلئوتیدی در اکوتیپ خزک، حاشیه‌نویسی[۹] عملکردی آنها با استفاده از برنامه‌ی SnpEff انجام شد (Cingolani et al. 2012). در این مطالعه، تنوع ژنتیکی درون جمعیت خزک در پنجره‌های کشویی ۵۰kb با اندازه ۲۵kb در هر مرحله در امتداد هر کروموزوم با استفاده از نرم افزار VCFtools محاسبه شد (Danecek et al. 2011).

نتایج و بحث

پس از کنترل کیفیت و پیرایش داده‌ها، به‌طور میانگین تعداد ۶۴ میلیون توالی کوتاه به ازای هر فرد باقی ماند. درصد نقشه‌یابی این توالی‌ها با ژنوم مرجع به طور میانگین ۹۷/۸۶ درصد بود که نشان‌دهنده‌ی راندمان خوب همردیفی با ژنوم مرجع می‌باشد. میانگین پوشش توالی‌ها برای ۱۷ فرد خزک ۶/۴۵ بود (جدول ۱). در ابتدا ۱۵،۹۸۱،۱۷۶ SNP از کل ژنوم فراخوانی شد که با نتایج مطالعات گذشته در مورد شناسایی چند شکلی‌های تک نوکلئوتیدی در مرغ با استفاده از توالی‌یابی کل ژنوم سازگاری دارد (Gheyas et al.

۳. Variant calling
۴. Realignment
۵ Coverage
۶. Linkage groups
۷. Minor allele frequency
۸. Genotyping rate
۹. Annotation

Agricultural Biotechnology Journal; Print ISSN: 2228-6705, Electronic ISSN: 2228-6500

2015a; Boschiero et al. 2018). پس از فیلتر داده‌ها (جدول ۲) و حذف SNPهای واقع در کروموزوم‌های جنسی، میتوکندری وگروه‌های پیوسته تعداد ۱۰،۵۵۲،۷۳۵ SNP در ۱۷ فرد باقی ماند (جدول ۱).

جدول ۱. تعداد چند شکلی‌های تک نوکلئوتیدی شناسایی شده و عمق پوشش توالی‌ها در هر حیوان

Table 1. The number of single nucleotide polymorphisms and depth of sequences coverage in each animal.

شناسه ID	تعداد SNPهای قبل از فیلتر کردن Number of SNPs before filtering	تعداد SNPهای بعد از فیلتر کردن Number of SNPs after filtering	عمق یا پوشش خوانش‌ها Reads depth or coverage
1	6196847	4097374	6.89
2	6178796	4039444	7.37
3	6142123	4058722	6.30
4	6185972	4059061	6.46
5	6007664	3941982	5.89
6	6352983	4230247	6.96
7	6265045	4130401	6.89
9	6433291	4270594	6.91
10	6149651	3993491	6.91
11	6408635	4247051	7.10
12	6123644	4048742	6.41
13	6346333	4224131	7.91
14	6197874	4138646	7.49
15	6206745	4129314	7.35
16	5819055	3811352	4.33
17	5840739	3811560	4.36
18	5647746	3672491	4.17

در جدول ۲ تعداد SNPهای حذف شده بعد انجام کنترل کیفیت در هر یک از پارامترهای اعمال شده نشان داده شده است. بیشترین تعداد SNP حذف شده مربوط به فراوانی آلل‌های کمیاب می‌باشد که برای به حداقل رساندن خطاها در مجموعه داده های توالی DNA بزرگ یا به عبارت دیگر برای دوری از نتایج مثبت کاذب این SNPها از مجموعه داده‌ها حذف می‌شوند (Linck & Battey 2019). در مطالعه حاضر هیچ فردی به دلیل ژنوتیپ‌های گمشده حذف نشد که نشان‌دهنده کیفیت خوب توالی‌یابی به ازای هر فرد می‌باشد.

جدول ۲. تعداد SNP حذف شده در آنالیز کنترل کیفیت برحسب آستانه‌ی تعریف شده

Table 2. The number of SNPs removed in the quality control analysis according to the defined threshold

پارامتر Parameter	آستانه تعریف شده defined threshold	تعداد SNPهای حذف شده number of removed SNPs
Minor allele frequencies/counts	0.05	4007969
Hardy-Weinberg equilibrium	10^{-7}	56325
Missing rate per person	0.1	0
Missing rate per SNP	0.1	844653

حاشیه‌نویسی عملکردی چند شکلی‌های تک نوکلئوتیدی نشان داد که اکثر SNPها در نواحی غیر کدکننده ژنوم از جمله نواحی اینترونی (۴۷/۵۶ درصد) و بین‌ژنی (۳۰ درصد) قرار دارند که با مطالعه‌ی Lawal et al. (2018) مطابقت دارد. ۹/۳۳ و ۹/۲۷ درصد SNPها به ترتیب در نواحی بالا دست و پایین دست قرار دارند و فقط حدود ۲/۸ درصد SNPها متعلق به نواحی اگزونی می‌باشند (جدول ۳). ازمیان حدود ۵۱۸،۵۱۱ چند شکلی تک نوکلئوتیدی در ناحیه اگزون، ۷۲/۷۷ درصد SNP ها در ناحیه بد معنی، ۲۲/۵۸ درصد در ناحیه خاموش و ۴/۶۶ درصد در ناحیه بی‌معنی قرار داشتند (جدول ۴). در مطالعه‌ای به منظور شناسایی چندشکلی‌های تک نوکلئوتیدی در مرغان بومی فارس با استفاده از داده های توالی‌یابی کل ژنوم، ۷۲/۳۳ درصد SNP ها در ناحیه بد معنی، ۲۵/۵ و ۰/۱۶درصد به ترتیب در ناحیه خاموش و بی‌معنی شناسایی شدند (Eskandari et al. 2019) که تعدادSNP های واقع در نواحی بد معنی با مطالعه ما مغایرت داشت. جهش های بد معنی در واقع جانشینی‌های ناهم‌معنی[۱۰] هستند که از جهش‌های نقطه ای ایجاد می‌شوند، جهش‌هایی در یک نوکلئوتید منفرد که منجر به تغییر آمینواسید و در نتیجه تغییر پروتئین کدگذاری شده می‌شود. جهش های بی‌معنی نیز جانشینی‌های ناهم‌معنی هستند و زمانی ایجاد می‌شوند که یک جهش در توالی DNA باعث شود که کدون کدکننده به یک کدون پایان تغییر یابد و در نتیجه ساخت پروتئین متوقف می‌شود (David 2003).

جدول ۳. تعداد اثرات چندشکلی‌های تک نوکلئوتیدی در نواحی مختلف ژنوم مرغ خزک

Table 3. Number of SNP effects in the different regions of genome

Splice_ site _acceptor	Splice_ site_ donor	Splice_ site_ region	UTR_ 3_ prime	UTR_ 5_ prime	بالادست Up Stream	پایین دست Down Stream	بین ژنی Inter Genic	اینترون Intron	اگزون Exon
5,570	5,526	41,700	98,072	2,609	1,711,652	1,706,156	5,550,430	8,750,139	518,511
(0.03%)	(0.03%)	(0.23%)	(0.54%)	(0.14%)	(9.33%)	(9.27%)	(30.14%)	(47.52%)	(2.82%)

[۱۰]. Non-synonymous

Agricultural Biotechnology Journal; Print ISSN: 2228-6705, Electronic ISSN: 2228-6500

در بررسی تنوع ساختاری در ژنوم سگ و گرگ بومی ایران نشان داده شد که از ۰/۸۱ درصد SNPهای شناسایی شده در ناحیه اگزون، حدود ۴۰ درصد آنها تغییردهنده آمینواسیدها در پروتئین هستند و حدود ۵۹/۵۵ درصد SNPها در ناحیه خاموش قرار دارند که تاثیری در تغییر آمینواسیدهای پروتئین ندارند و ۰/۴۲ درصد آنها ایجادکننده کدون پایان هستند (Amiri ghanatsaman et al. 2016)

در پژوهشی به منظور بررسی حاشیه‌نویسی عملکردی، ۱۳،۴۴۹،۳۳۹ SNP در دو لاین مرغ تخم‌گذار و گوشتی شناسایی شد که اکثر چندشکلی های تک نوکلئوتیدی در ناحیه اینترونی (۴۲/۱) و بین ژنی (۳۸/۵) قرار دارند و حدود ۱/۲ درصد SNP واقع در ناحیه اگزون قرار داشتند (Boschiero et al. 2018) که نسبت به نتایج بدست آمده در مطالعه حاضر، تعداد SNPهای واقع در ناحیه ی اگزون ژنوم کمتر بوده است. در گذشته، واریانت‌های هم‌معنی[۱۱] (خاموش) به عنوان جایگزینی کدون تعریف می‌شدند که اسیدآمینه رمزگذاری شده را تغییر نمی‌دهند اما شواهد موجود نشان می‌دهد که این واریانت‌ها به اصطلاح خاموش، می‌توانند تأثیر قابل توجهی در بیان و عملکرد پروتئین داشته باشند و نباید به عنوان واریانت‌های خاموش در نظر گرفته شود (Tsai et al. 2008; Edwards et al. 2012).

جدول ٤. تعداد اثرات چندشکلی‌های تک نوکلئوتیدی با دسته عملکردی در اکوتیپ خزک

Table 4. Number of SNP effects by functional class in the Khazak ecotype

خاموش Silent	بی معنی Nonsense	بد معنی Missense
74,687 (22.58%)	15,424 (4.66%)	240,715 (72.77%)

در پژوهشی Gheyas et al. (2015a) با استفاده از توالی‌یابی کل ژنوم جمعیت‌های مختلف مرغ (۲۴۳ مرغ از ۲۴ لاین گوشتی و تخم‌گذار)، تعداد ۱۵ میلیون SNP شناسایی کردند و گزارش دادند که ۴۳/۳۵ SNPها در نواحی اینترونی و ۵۰/۲۲ درصد آنها در نواحی بین ژنی قرار دارند و فقط حدود ۲/۲ درصد چندشکلی‌های تک نوکلئوتیدی کد کننده پروتئین می‌باشند که تا حدودی با نتایج پژوهش ما مغایرت داشت. این ممکن است به دلیل عواملی مانند تعداد و نمونه مورد بررسی، روش‌های آنالیز داده‌ها و عمق توالی‌یابی باشد.

نسبت جهش‌های انتقالی به جهش‌های متقاطع (Ti / Tv) ۲/۴۳ بدست آمد (جدول ۵) که با نتایج Boschiero (2018) et al. و Oh et al. (2016) هنگام بررسی ژنوم مرغ نژادهای مختلف قابل مقایسه بود به طوری که آنها نسبت این جهش‌ها را ۲/۳ برآورد کردند. نسبت بالاتر بدست آمده در نتایج ما نشان‌دهنده‌ی کیفیت مطلوب شناسایی SNPها می‌باشد، زیرا هرچه نسبت Ti / Tv بیشتر باشد نشان‌دهنده میزان خطای کمتر می‌باشد (Bainbridge et al. 2011; Guo et al. 2012). علاوه بر این،

[۱۱]. Synonymous

نسبت Ti / Tv مورد انتظار از واریانت‌های جدید واقعی می‌تواند در ژنوم نسبت به ناحیه مورد هدف (کل ژنوم، اگزوم، ژن‌های خاص)، گونه و محتوای CpG و GC آن ناحیه متفاوت باشد (Rosenberg et al. 2003). به عنوان مثال، در مورد اگزوم‌ها، افزایش حضور سیتوزین متیله شده در دی نوکلئوتیدهای CpG در مناطق اگزونی منجر به افزایش نسبت Ti / Tv به دلیل یک دی آمیناسیون ساده و انتقال سیتوزین متیله شده به تیمین می‌شود (Aslam et al. 2012). همچنین مشاهده شده است که محتوای GC در پرندگان و پستانداران بیشتر از بی مهرگان می‌باشد. نسبت Ti / Tv مشاهده شده در این مطالعه کمتر از یافته‌ی Aslam et al. (2012) در مورد بوقلمون می‌باشد (۲/۴۵). اما کمی بالاتر از انسان می‌باشد (۲/۲) (DePristo et al. 2011). این نسبت بالاتر به احتمال زیاد با اندازه کوچکتر ژنوم و درصد GC بالاتر در ژنوم پرندگان توضیح داده می‌شود.

شناسایی واریانت‌های هتروزیگوت دشوارتر از هموزیگوت می‌باشد. نشان داده شده است که یک پوشش توالی ۱۰X-۶ برای تشخیص ۹۹٪ از واریانت‌های مختلف (هموزیگوت و هتروزیگوت) کافی می‌باشد. با این حال، برخی از مطالعات گزارش داده اند که برای تشخیص ۹۹٪ SNPهای هتروزیگوت پوششی بالاتر x ۲۰ ضروری می‌باشد (Li et al. 2008; Sims et al. 2014; Boschiero et al. 2018). در این مطالعه با استفاده از پوشش توالی ۷x-۶ نشان داده شد که تعداد چندشکلی‌های تک نوکلئوتیدی هتروزیگوت بیشتر بود. در جدول ۶ تعداد چند شکلی‌های تک نوکلئوتیدی هموزیگوت و هتروزیگوت به ازای هر مرغ نشان داده شده است. در همه ۱۷ پرنده تعداد SNPهای هتروزیگوت بیشتر از هموزیگوت بود که با مطالعات پیشین در اکوتیپ‌های مرندی و فارس مطابقت دارد (Eskandari et al. 2019; Akbari et al. 2020). بیشترین و کمترین واریانت هتروزیگوت به ترتیب مربوط به مرغ‌های با شناسه ۱۸ (۳،۴۰۹،۳۰۵) و ۱۰ (۲،۴۳۸،۰۲۲) می‌باشد; این امر احتمالا نشان‌دهنده تنوع بیشتر در مرغ با شناسه ۱۸ و همخونی بیشتر در مرغ با شناسه ۱۰ می‌باشد. به طور میانگین نسبت SNP های هتروزیگوت به هموزیگوت در اکوتیپ خزک ۳/۳۰ برآورد شد.

در بررسی تنوع ژنتیکی با استفاده از نرم افزار VCFtools میزان تنوع ژنتیکی درون اکوتیپ خزک ۰/۰۰۱۴ ± ۰/۰۰۳۷ برآورد شد. در پژوهشی Lawal et al. (2018) تنوع ژنتیکی در نژادهای مختلف مرغ را بررسی کردند و نشان دادند که بالاترین میزان تنوع نوکلئوتیدی ژنومی درون نژاد مرغ جنگلی قرمز می‌باشد (π =۰/۰۰۵۲) و در نژادهای مرغ بومی، مرغان سریلانکا بالاترین تنوع نوکلئوتیدی (π =۰/۰۰۴۶) را دارا می‌باشند و تنوع ژنتیکی در میان مرغان عربستان سعودی (π = ۰/۰۰۳۹) و اتیوپی جارسو (π =۰/۰۰۳۶) تا حدودی با نتایج ما مطابقت داشت.

جدول ۵. تعداد SNP های از نوع انتقالی و متقاطع در ژنوم

Table 5. SNP number of transition and transversion in genome

نسبت جهش‌های انتقالی به متقاطع Transition/ transversion ratio	تعداد جهش‌های متقاطع The number of transversons mutations	تعداد جهش‌های انتقالی The number of transition mutations
2.43	23,921,271	58,258,716

Agricultural Biotechnology Journal; Print ISSN: 2228-6705, Electronic ISSN: 2228-6500

جدول ٦. تعداد چندشکلی‌های تک نوکلئوتیدی هموزیگوت و هتروزیگوت در اکوتیپ خزک

Table 6. The number of homozygous and hetrozygous SNPs in the Khazak ecotype

شناسه ID	هتروزیگوت het	هموزیگوت homo	نسبت هتروزیگوت به هموزیگوت het/homo
1	2904713	911016	3.188432
2	3315877	816415	4.061509
3	2970185	965500	3.076318
4	3262493	868991	3.754346
5	3158207	848611	3.721619
6	3094614	912957	3.38966
7	2599731	1106553	2.349396
9	2662645	991202	2.686279
10	2438022	1126265	2.164697
11	3096389	829391	3.733328
12	2993729	887630	3.372722
13	3031310	881242	3.439816
14	3031518	892941	3.394981
15	2819592	1022783	2.756784
16	3260992	871860	3.74027
17	3139233	842526	3.725978
18	3409305	719833	4.736244
Average	3011091	911512.7	3.3034

نتیجه‌گیری: با اســتفاده از داه‌های کل ژنوم، آثار عملکردی چند شــکلی‌های تک نوکلئوتیدی و تنوع ژنتیکی درون اکوتیپ خزک برای اولین بار مورد برر سی قرار گرفت. ۱۰/۵ میلیون چند شکلی تک نوکلئوتیدی در ۱۷ پرنده شنا سایی شد که حدود ۲/۸ در صد آنها کد کننده پروتئین می‌باشــند. اکوتیپ خزک مانند ســایر اکوتیپ‌های بومی کشــور دارای تنوع می‌باشــد و از آنجایی که حیوانات بومی به عنوان سرمایه ملی و ذخایر ا ستراتژیک هر کشوری محسوب می شوند مرغ خزک می‌تواند به عنوان یک اکوتیپ تخم‌گذار بومی در برنامه‌های اصلاح نژاد مرغ حائز اهمیت باشد.

سپا سگزاری: از معاونت محترم پژوهشی دانشگاه شهید باهنر کرمان به خاطر حمایت مالی در قالب اعتبار ویژه پژوهشی سال ۱۳۹۹ به نگارنده سوم مقاله برای تامین بخ شی از بودجه برای اجرای پژوهش حا ضر سپا سگزاری می شود. از آقای دکتر حامد خراتی کوپایی که در جمع آوری نمونه و مراحل توالی یابی داده ها نقش داشته اند، تشکر و قدردانی می‌شود.

منابع

اسکندری طاهره، اسمعیلی‌زاده کشکوئیه علی، اسدی فوزی مسعود (۱۳۹۷) شناسایی نشانگرهای تک نوکلئوتیدی در مرغ بومی فارس با استفاده از روش توالی‌یابی کل ژنوم. مجله بیو تکنولوژی کشاورزی ۱۰ (۱)، ۱۵۱-۱۳۹.

اکبری رسول، اسمعیلی‌زاده کشکوئیه علی، امیری قنات سامان زینب، آیت اللهی مهجردی احمد (۱۳۹۹) شناسایی تنوع ژنوم در مرغ مرندی با استفاده از روش توالی‌یابی کل ژنوم. مجله بیو تکنولوژی کشاورزی ۱۲ (۱)، ۱۶۱-۱۷۶.

امیری قنات سامان زینب، اسمعیلی‌زاده کشکوئیه علی، اسدی فوزی مسعود (۱۳۹۵) بررسی تنوع ساختاری ژنگان سگ و گرگ بومی ایران با روش توالی‌یابی کل ژنوم. مجله علوم دامی ایران ۴۷ (۲)، ۲۷۱-۲۷۷.

توکلیان جواد، (۱۳۷۸) نگرشی بر ذخایر ژنتیکی دام و طیور بومی ایران. موسسه تحقیقات علوم دامی ایران ، ایران. ۲۰۱-۱۶۹

دلجوی سرایان ج، ۱۳۹۰ .تعیین صفات کیفی تخم مرغ در مرغ بومی خزک. پایان نامه کارشناسی ارشد، دانشکده کشاورزی، دانشگاه زابل. ۴۲-۲۸.

محمدی فر آمنه، فقیه ایمانی سید علی، محمد آبادی محمد رضا، سفلایی محمد (۱۳۹۲) تاثیر ژن TGFB3 بر ارزش‌های فنوتیپی و ارثی صفات وزن بدن در مرغ بومی فارس. مجله بیوتکنولوژی کشاورزی ۵ (۴)، ۱۲۵-۱۳۶.

نیکبختی مهدی، میرزایی حمیدرضا، افشاریان شاندیز مجید و همکاران (۱۳۸۸) بررسی تنوع ژنتیکی مرغان بومی استان خراسان رضوی با استفاده نشانگرهای ریزماهواره. پژوهش‌های علوم دامی ایران ۱(۲)، ۱۹-۲۵.

References

Abebe AS, Mikko S, Johansson AM (2015) Genetic diversity of five local Swedish chicken breeds detected by microsatellite markers. PLoS One 4, 1-13.

Akbary R, Esmaeelizadeh A, Amiri Ghanatsaman Z, Ayetollahi Mehrjerdi A (2020) Identification of genome diversity in marandi chicken using whole genome sequencing method. J Agric Biotech 12,161-176 (In Persian).

Amiri Ghanatsaman Z, Esmailizadeh Koshkoiyeh A, Asadi Fozi M (2016) Study of structural diversity of genome Iranian native dog and wolf with the method whole genome sequencing. Iran J Anim Sci 47, 271-277 (In Persian).

Andrews S (2010) FastQC: a quality control tool for high throughput sequence data. Babraham Bioinformatics, Babraham Institute, Cambridge, United Kingdom.

Aslam ML, Bastiaansen JW, Elferink MG et al. (2012) Whole genome SNP discovery and analysis of genetic diversity in Turkey (Meleagris gallopavo). BMC genomics 13, 1-14.

Bainbridge MN, Wang M, Wu Y et al. (2011) Targeted enrichment beyond the consensus coding DNA sequence exome reveals exons with higher variant densities. Genome Biol 12, 1-12.

Boettcher P, Tixier-Boichard M, Toro M et al. (2010) Objectives, criteria and methods for using molecular genetic data in priority setting for conservation of animal genetic resources. Anim Genetics 41, 64-77.

Bolger AM, Lohse M, Usadel B (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30, 2114-2120.

Boschiero C, Moreira GCM, Gheyas AA et al. (2018) Genome-wide characterization of genetic variants and putative regions under selection in meat and egg-type chicken lines. BMC Genomics 19, 83.

Cingolani P, Platts A, Wang LL et al. (2012) A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff :SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. Fly 6, 80-92.

Danecek P, Auton A, Abecasis G et al. (2011) The variant call format and VCFtools. Bioinformatics 27, 2156-2158.

Deljoy Sarayan J, (2011). Determination of Egg Qualitative Traits in Khazak Native Chicken. Master thesis, Zabol University. pp. 28-42 (In Persion).

DePristo MA, Banks E, Poplin R et al. (2011) A framework for variation discovery and genotyping using next-generation DNA sequencing data. Nature genetics 43, 491-498.

Diamond J (2002) Evolution, consequences and future of plant and animal domestication. Nature 418, 700-707.

Ebbert MT, Wadsworth ME, Staley LA et al. (2016) Evaluating the necessity of PCR duplicate removal from next-generation sequencing data and a comparison of approaches. BMC Bioinformatics 17, 239.Edwards NC, Hing ZA, Perry A et al. (2012) Characterization of coding synonymous and non-synonymous variants in ADAMTS13 using ex vivo and in silico approaches. PLoS One 7, e38864.

Engelsma KA (2012) Use of SNP markers to conserve genome-wide genetic diversity in livestock. PhD thesis, Wageningen University. pp. 12-20.

Eskandari T, Esmailizadeh AK, Mohammadabadi MR, Sohrabi S (2018) Identification of single nucleotide polymorphisms in Fars native chicken using whole genome sequencing data. Agric Biotechnol J 10, 139-151 (In Persian).

Gheyas AA, Boschiero C, Eory L et al. (2015a) Functional classification of 15 million SNPs detected from diverse chicken populations. DNA Res 22, 205-217.

Gheyas AA, Boschiero C, Eory L et al. (2015b) Functional classification of 15 million SNPs detected from diverse chicken populations. DNA Res 22, 205-217.

Guo Y, Li J, Li C-I et al. (2012) The effect of strand bias in Illumina short-read sequencing data. BMC genomics 13, 1-11.

Hoffmann I (2005) Research and investment in poultry genetic resources–challenges and options for sustainable use. Worlds Poult Sci J 61, 57-70.

Lawal RA, Al-Atiyat RM, Aljumaah RS et al. (2018) Whole-genome resequencing of red junglefowl and indigenous village chicken reveal new insights on the genome dynamics of the species. Front Genet 9, 264.

Li H, Durbin R (2009) Fast and accurate short read alignment with Burrows–Wheeler transform. Bioinformatics 25, 1754-1760.

Li H, Handsaker B, Wysoker A et al. (2009) The sequence alignment/map format and SAMtools. Bioinformatics 25, 2078-2079.

Li H, Ruan J, Durbin R (2008) Mapping short DNA sequencing reads and calling variants using mapping quality scores. Genome Res 18, 1851-1858.

Linck E, Battey C (2019) Minor allele frequency thresholds strongly affect population structure inference with genomic data sets. Mol Ecol Resour 19, 639-647.

Lynch M, Milligan BG (1994) Analysis of population genetic structure with RAPD markers. Mol Ecol 3, 91-99.

McKenna A, Hanna M, Banks E et al. (2010) The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res 20, 1297-1303.

Moazeni S, Mohammadabadi MR, Sadeghi M et al. (2016a) Association between UCP Gene Polymorphisms and Growth, Brreeding Value of Growth and Reproductive Traits in Mazandaran Indigenous Chicken. Open J. Anim. Sci 6, 1-8.

Moazeni SM, Mohammadabadi MR, Sadeghi M et al. (2016b) Association of the melanocortin-3(MC3R) receptor gene with growth and reproductive traits in Mazandaran indigenous chicken. J. Livest. Sci. Technol 4, 51-56.

Mohammadifar A, Faghih Imani SA, Mohammadabadi MR, Soflaei M (2014) The effect of TGFb3 gene on phenotypic and breeding values of body weight traits in Fars native fowls. Agric. Biotechnol J 5, 125-136 (In Persian).

Mohammadifar A, Mohammadabadi MR (2018) Melanocortin-3 receptor (mc3r) gene association with growth and egg production traits in Fars indigenous chicken. Malays Appl Biol 47, 85–90.

Mohammadifar A, Mohammadabadi MR (2017) The Effect of Uncoupling Protein Polymorphisms on Growth, Breeding Value of Growth and Reproductive Traits in the Fars Indigenous Chicken. Iran J Appl Anim Sci 7, 679-685.

Nikbakhti M, Mirzaee HR, Afsharian Shandiz M, et al. (2009) Analyses of genetic variation in Khorasan indigenous chicken breed by using of microsatellite markers. Iran J Anim Sci Res 1 (2), 19-25 (In Persian).

Oh D, Son B, Mun S et al. (2016) Whole genome re-sequencing of three domesticated chicken breeds. Zool Sci 33, 73-77.

Osman SA-M, Yonezawa T, Nishibori M (2016) Origin and genetic diversity of Egyptian native chickens based on complete sequence of mitochondrial DNA D-loop region. Poult Sci 95, 1248-1256.

Padhi MK (2016) Importance of indigenous breeds of chicken for rural economy and their improvements for higher production performance. Scientifica 2016.

Peters J, Lebrasseur O, Best J et al. (2015) Questioning new answers regarding Holocene chicken domestication in China. Proc Natl Acad Sci 112, E2415-E2415.

Rosenberg MS, Subramanian S, Kumar S (2003) Patterns of transitional mutation biases within and among mammalian genomes. Mol Biol Evol 20, 988-993.

Sawai H, Kim HL, Kuno K et al. (2010) The origin and genetic variation of domestic chickens with special reference to junglefowls Gallus g. gallus and G. varius. PLoS One 5, e10639.

Sims D, Sudbery I, Ilott NE et al. (2014) Sequencing depth and coverage: key considerations in genomic analyses. Nat. Rev. Genet 15, 121.

Strillacci M, Cozzi M, Gorla E et al. (2017) Genomic and genetic variability of six chicken populations using single nucleotide polymorphism and copy number variants as markers. Anim 11, 737-745.

Tavakolian J (2000) An introduction to genetic resorces of native farm animals in Iran. Animal Science Research Center of Iran, Iran. pp. 169-201 (In Persion).

Tixier-Boichard M, Bed'hom B, Rognon X (2011) Chicken domestication: from archeology to genomics. C. R. Biol 334, 197-204.

Tsai C-J, Sauna ZE, Kimchi-Sarfaty C et al. (2008) Synonymous mutations and ribosome stalling can lead to altered folding pathways and distinct minima. J. Mol. Biol 383, 281-291.

Venter JC, Smith HO, Adams MD (2015) The sequence of the human genome. Clin. Chem 61, 1207-1208.