Research on Animal Production, Vol.14, No. 41, Autumn 2023 p: 98-104

"Research Paper"

Identification and Classification of Genome Variants of Khuzestan River Buffalos Using Sequencing Technology

Seyed Milad Hosseini[1], Hossein Moradi Shahrbabak[2] and Mohammad Moradi Shahrbabak[3]

1- Ph.D. Student, Agriculture and Natural Resources Campus, University of Tehran, Karaj, Iran, (Corresponding author: milad.hosseiny88@gmail.com)
2- Assistant Professor, Animal Science Department, Agriculture and Natural Resources Campus, University of Tehran, Karaj, Iran
3- Professor, Animal Science Department, Agriculture and Natural Resources Campus, University of Tehran, Karaj, Iran
Received: 7 May, 2023 Accepted: 13 September, 2023

Extended Abstract

Introduction and Objective: The rapid development of sequencing technologies and bioinformatic tools has made it possible to sequence the entire genome of many species, providing a starting point for exploring the genome-wide genetic diversity of organisms. The buffalo is one of the most important domesticated animals in the world and has great economic value for humans, being used for milk, meat, leather and most importantly as labor. Khuzestani buffalo are among the most important buffalo in Iran. These animals are scattered in western and southwestern Iran. The largest population of this breed is in the Khuzestan province and they are fully adapted to the geographical features of this region

Material and Methods: In this study, whole genome sequencing of the Khuzestan buffalo was performed in paired-end format using Illumina technology. Quality control of the readings obtained met all quality parameters and processing of the readings showed that all samples were of high quality. The BWA-MEM software package was then used to match the data to the reference genome. Genome variants were obtained with the Freebayes software. The final filtering of the variant file then took place. This filtering serves to increase the quality of the analysis and to sort out inferior variants.

Results: The alignment percentage for all samples was over 97.44% and also the coverage in the sequenced samples ranged from x 4.3 to x 9.12, indicating reasonable quality of the reads. Also, in this study we identified 76,676,529 million variants fairly evenly distributed on the chromosomes of the Khuzestan buffalo. The highest and the lowest number of variants were observed on chromosome 1 chromosome 28, repectively. In this study, the total number of translocation and inversion mutations in the genome level was 298,193,017 and 135,694,466, respectively. In addition, the ratio of transitions to transversion was estimated at 2.1989 (Ts/Tv).

Conclusion: The presence of genetic diversity in the native population of Iranian buffalos are valuable parameters that can be studied to evaluate the genetic potential of economic traits in these populations in order to design suitable breeding programs.

Keywords: Buffalo, Khuzestan, Variant, Whole genome sequencing

[Downloaded from rap.sanru.ac.ir on 2025-04-05]
[DOR: 20.1001.1.22518622.1402.14.41.10.5]
[DOI: 10.61186/rap.14.41.98]

دانشگاه علوم کشاورزی و منابع طبیعی ساری

"مقاله پژوهشی"

شناسایی و طبقه‌بندی واریانت‌های ژنومی گاومیش‌های رودخانه‌ای خوزستان با استفاده از فناوری توالی‌یابی

سید میلاد حسینی[۱]، حسین مرادی شهربابک[۲] و محمد مرادی شهربابک[۳]

۱- دانشجوی دوره دکتری، پردیس کشاورزی و منابع طبیعی دانشگاه تهران، کرج، ایران، (نویسنده مسوول: milad.hosseiny88@gmail.com)
۲- استادیار گروه علوم دامی، پردیس کشاورزی و منابع طبیعی دانشگاه تهران، کرج، ایران
۳- استاد گروه علوم دامی، پردیس کشاورزی و منابع طبیعی دانشگاه تهران، کرج، ایران
تاریخ دریافت: ۱۴۰۲/۲/۱۷ تاریخ پذیرش: ۱۴۰۲/۶/۲۲
صفحه: ۹۸ تا ۱۰۴

چکیده مبسوط

مقدمه و هدف: توسعه سریع فناوری‌های توالی‌یابی کل ژنوم و ابزارهای بیوانفورماتیک، نقطه شروعی برای کشف تنوع ژنتیکی گسترده در سطح کل ژنوم موجودات را فراهم می‌کند. گاومیش از جمله حیوانات اهلی مهم در سطح جهانی است که ارزش اقتصادی زیادی برای جوامع انسانی دارد که برای تولید محصولاتی مثل شیر، گوشت، چرم و نیز به‌عنوان نیروی کار از آن‌ها استفاده می‌شود. گاومیش‌های نژاد خوزستانی یکی از مهمترین نژاد گاومیش در ایران هستند که در غرب و جنوب‌غربی ایران پراکنده شده‌اند. بیشترین جمعیت این نژاد در استان خوزستان وجود دارد که با شرایط جغرافیایی این منطقه کاملاً سازگار شده‌اند. در این مطالعه، شناسایی تنوع ژنتیکی در سطح کل ژنوم گاومیش‌های نژاد خوزستانی با استفاده از تکنیک توالی‌یابی انجام گرفت.

مواد و روش‌ها: کنترل کیفیت خوانش‌های توالی‌یابی شده با استفاده از پارامترهای مرتبط با کیفیت به‌وسیله‌ی نرم‌افزار FastQC (v0.12.1) انجام شد. ارزیابی خوانش‌های قطعات توالی‌یابی نشان داد که همه‌ی نمونه‌ها از کیفیت بالایی برخوردار بودند. برای هم ردیفی داده‌ها با ژنوم مرجع از بسته نرم‌افزاری BWA-MEM استفاده شد. در ادامه واریانت‌های ژنومیک با استفاده از نرم‌افزار Freebayes به‌دست آمد. سپس، فیلتراسیون نهایی روی فایل واریانت‌ها انجام گرفت. این فیلتراسیون به‌منظور افزایش کیفیت آنالیز و حذف واریانت‌های با کیفیت پایین انجام گرفت.

یافته‌ها: درصد همردیفی برای همه‌ی نمونه‌ها بالاتر از ۹۷/۴۴ درصد و میزان همپوشانی در نمونه‌های توالی‌یابی شده بین ۴/۳ x تا ۱۲/۹ x بود که نشان دهنده‌ی کیفیت مناسب خوانش‌ها است. همچنین، در این مطالعه تعداد ۷۶.۶۷۶.۵۲۹ میلیون واریانت شناسایی شده است که به‌طور نسبتاً مساوی در بین کروموزوم‌های گاومیش‌های خوزستانی توزیع شده بودند. بیشترین و کمترین تعداد واریانت‌ها به‌ترتیب روی کروموزوم شماره ۱ و ۲۸ مشاهده شد. در این مطالعه، تعداد جهش‌های جابه‌جایی و معکوس شناسایی شده در سطح ژنوم به‌ترتیب برابر با ۲۹۸.۱۹۳.۰۱۷ و ۱۳۵.۶۹۴.۴۶۶ و نسبت جهش‌های جابه‌جایی به معکوس به‌میزان ۲/۱۹۸۹ (Ts/Tv) برآورد شد.

نتیجه‌گیری: وجود تنوع ژنتیکی در جمعیت گاومیش‌های بومی ایران، پارامترهای باارزشی هستند که با مطالعه‌ی آن‌ها می‌توان پتانسیل ژنتیکی صفات اقتصادی در این جمعیت‌ها را به‌منظور طراحی برنامه‌های مناسب اصلاح‌نژادی مورد ارزیابی قرار داد.

واژه‌های کلیدی: توالی‌یابی کل ژنوم، خوزستانی، گاومیش، واریانت

مقدمه

گاومیش رودخانه‌ای یکی از حیوانات مهم برای تولیدکنندگان و گله‌داران کوچک و منبع مهمی برای بازارهای تخصصی در کشورهای در حال توسعه است. گاومیش‌های اهلی، در سرتاسر جهان پراکنده شده‌اند و در ۱۲۹ کشور عمدتاً، گرمسیری و نیمه‌گرمسیری یافت می‌شوند. این حیوانات به اقتصاد روستایی به‌ویژه در آسیا، و در بسیاری از مناطقی که بوفالوها از گاو اهمیت بیشتری دارند، کمک می‌کنند. جمعیت گاومیش‌های جهان بیش از ۲۰۷ میلیون راس می‌باشد (FAOSTAT; Scherf & Pilling, 2015). دو نوع گاومیش رودخانه‌ای اهلی وجود دارد، گاومیش رودخانه‌ای (*Bubalus bubalis bubalis*، 2n = 50) که پراکنش جهانی دارد، اما نوع غالب آن در غرب از هند تا اروپا است. نوع دوم، بوفالو مردابی است (*Bubalus bubalis carabanensis*، 2n = 48) که بیشتر در کشورهای آسیای شرقی، به‌ویژه از هند تا چین، اندونزی و فیلیپین یافت می‌شود. بیشتر از ۹۷ درصد از جمعیت گاومیش‌های جهان در قاره‌ی آسیا قرار دارند. در آفریقا، گاومیش‌های رودخانه‌ای اهلی تنها در مصر با جمعیت حدود ۴ میلیون راس و همچنین در موزامبیک یافت می‌شوند (Borghese et al., 2022; Iamartino et al., 2017). زمان واگرایی دوگونه‌ی گاومیش باتلاق و رودخانه‌ای در تعدادی از مطالعات از ۱۰هزار سال تا ۱/۷ میلیون سال تخمین زده شده است. جمعیت گاومیش در ایران، ۲۲۴ هزار رأس است. جمعیت گاومیش کشور از ۲۲۲ هزار رأس در سال ۲۰۲۰ به ۲۲۴ هزار رأس در سال ۲۰۲۱ رسیده است. این رشد بیشتر در مناطق روستایی و به‌صورت گله‌های کوچک بوده است که این مهم نشان دهنده‌ی افزایش نقش این حیوان سودمند در اقتصاد روستایی است. در ایران هدف از پرورش گاومیش به‌طور عمده به‌منظور تولید شیر بوده و گوشت در درجه بعدی اهمیت قرار دارد. در مجموع ۲/۵۷ درصد از تولیدات گوشت و شیر کشور به گاومیش اختصاص دارد (۱۵۸ هزار تن شیر (۱/۴۰ درصد) و ۱۰/۴۸ هزار تن گوشت (۱/۱۷ درصد) (AGRI., 2021). جمعیت گاومیش‌ها حدود ۱۱/۱ درصد جمعیت گاوها در جهان است، اما، در مقایسه با سایر گونه‌های اهلی شده، تعداد انسان‌های بیشتری برای امرار معاش خود به گاومیش‌ها وابسته هستند (Kierstein et al., 2004). گاومیش‌های آبی خوزستانی در غرب و جنوب‌غربی ایران پراکنده شده‌اند. بیشترین جمعیت از این نژاد در استان خوزستان وجود دارد. خوزستان دارای آب و هوای خیلی گرم با رطوبت بالا در تابستان و زمستان‌های سرد است. دمای هوا در زمستان تا ۵ درجه سانتیگراد و در تابستان تا ۴۵ درجه می‌رسد

[Downloaded from rap.sanru.ac.ir on 2025-04-05]
[DOR: 20.1001.1.22518622.1402.14.41.10.5]
[DOI: 10.61186/rap.14.41.98]

و این نژاد برای این شرایط آب و هوایی آداپته شده است. در دو دهه‌ی گذشته، حدود ۲ درصد افزایش سالانه در جمعیت گاومیش‌های جهان مشاهده شده است و این تعداد از گاومیش‌ها حدود ۱۳ درصد از تولید جهانی شیر را به‌خود اختصاص می‌دهند (Rehman et al., 2019). به‌طور کلی، تنوع‌های ژنتیکی مانند واریانت‌های تکنوکلئوتیدی[1] و حذف و اضافه‌های کوتاه[2] جزء مهمترین تنوع‌های سطح ژنوم هستند و محدوده این تغییرات می‌تواند از یک جفت باز نوکلئوتیدی تا جایگزینی‌های بزرگ در حد کروموزوم، متفاوت باشد (Alkan et al., 2011). روش‌های متعددی برای شناسایی واریانت‌های سطح ژنوم در دسترس می‌باشد. یکی از آنها روش‌های سیتوژنتیکی است که از تکنیک‌هایی مانند تهیه کاریو تایپ، رنگ‌آمیزی کروموزومی و روش‌های بر پایه تکنیک FISH[3] بهره می‌بردند. گرچه این روش‌ها هنوز در بسیاری از ازمایشگاه‌های کلینیکی و تحقیقاتی انجام می‌شوند، اما، دقت و حساسیت زیادی برای واریانت‌های ساختاری کوچکتر ندارند و نمی‌توانند آن‌ها را با صحت بالا شناسایی کنند. از این‌رو برای شناسایی واریانت‌های کوچکتر با حساسیت و وضوح بالا، روش‌های مولکولی نوین‌تر گسترش پیدا کرده‌اند (Krausz et al., 2011). اما، امروزه با گسترش نسل جدید فناوری‌های توالی‌یابی پر برونداد NGS، امکان توالی‌یابی همه ژنوم‌های بزرگ فراهم شده است (Lupski & Stankiewicz, 2005). این تکنولوژی با مزیت‌هایی مانند قابلیت تفکیک بالاتر، پوشش گسترده‌تر ژنوم، زمان کم توالی‌یابی و قیمت مناسب، انقلابی را در زمینه تحقیقات ژنومیک بوجود آورده است. به‌کمک این فناوری می‌توان در مدت زمان بسیار کمی نسبت به روش‌های پیشین، با ارائه صدها میلیون خوانش توالی از یک ژنوم، تصویر آشکاری از خصوصیات ژنومی به نمایش گذاشت (Metzker, 2010). به‌دلیل اهمیتی که چندشکلی‌های تک نوکلئوتیدی و حذف و اضافه‌های کوتاه در مطالعات ژنتیکی دارند و با توجه به وارثت‌پذیر بودن و فراوانی بالای واریانت‌ها در سطح ژنوم حیوانات اهلی، برخی از واریانت‌ها می‌توانند با ژن‌های مرتبط با صفات اقتصادی، تولیدی و سلامتی دام‌ها در ارتباط باشند و شناسایی و معرفی این تنوع‌ها در شناسایی عوامل ژنتیکی موثر در خصوصیات مهم و اقتصادی، همچنین، افزایش بهره‌وری، سلامت و تولید حیوانات اهلی بسیار مفید باشد (Koopaee & Koshkoiyeh, 2014).

مواد و روش‌ها

داده‌های خام این مطالعه مربوط به پروژه‌ای که توسط کنسرسیوم بین‌المللی ژنوم گاومیش با هماهنگی دپارتمان کشاورزی ایالات متحده آمریکا[4] (USDA) و موسسه پارک علم Padano (PTP) ایتالیا به‌منظور کامل نمودن توالی‌یابی و سرهم‌سازی ژنوم گاومیش انجام شد، به‌دست آمد. از دیگر اهداف این پروژه، شرح نویسی[5]، کشف SNPها و تنوع‌های ساختاری در ژنوم گاومیش‌های رودخانه‌ای و باتلاقی و همچنین، تجزیه و تحلیل تنوع ژنتیکی موجود در جمعیت‌های گاومیش‌های موجود در جهان بوده است. در این پروژه، ۳۹۰ حیوان از ۳۰ جمعیت مختلف در سرتاسر جهان شرکت کردند. ژنوم تمام حیوانات در موسسه (PTP) ایتالیا تعیین توالی شد. در ایران، مجری طرح شرکت دانشگاهی و دانش‌بنیان توسعه کشت و دام نو اندیش البرز بوده که طی قراردادی که با مجریان این طرح در کشور ایتالیا به سرپرستی دکتر Jhone Williams از دانشگاه مریلند بسته شد، همکاری ایران با این کنسرسیوم بین‌المللی آغاز و در پی آن توالی کل ژنوم گاومیش‌های ایران با استفاده از تکنیک توالی‌یابی با برونده بالا یا نسل دوم توالی‌یابی[6](NGS)، انجام شد. استخراج DNA از نمونه‌های خون با روش استخراج نمکی و از نمونه‌های بافت با روش کلروفورم ایزوآمیل الکل انجام شد. توالی‌یابی کل ژنگان به‌صورت Paired-End توسط شرکت ایلومینا Hiseq 2500 تعیین و انجام گرفت. در این پروژه از اطلاعات توالی‌یابی شده ۵ راس گاومیش نژاد خوزستانی استفاده شد.

کنترل کیفیت داده‌ها

سنجش کیفیت داده‌ها توسط نرم‌افزار(FastQC (v0.11.5 انجام شد (Andrews, 2010). برای ویرایش داده‌ها از نرم‌افزار Trimmomatic استفاده شد (Bolger et al., 2014). این نرم‌افزار به‌عنوان ابزاری کارآمد با پیش‌پردازش‌های مؤثر و منطبق با داده‌های pairedend است و برای داده‌های توالی‌یابی نسل جدید دستگاه شرکت ایلومینا بهینه‌سازی شده است. وظایف این نرم‌افزار شامل حذف آداپتورها و حذف یا ویرایش خوانش‌های بی‌کیفیت می‌باشد.

همردیفی و مکان‌یابی خوانش‌ها بر روی ژنوم مرجع

برای هم‌ردیفی داده‌ها با ژنوم مرجع گاو (UMD3.1) از بسته نرم‌افزاری BWA-MEM استفاده شد که نسبت به الگوریتم‌های دیگر دارای سرعت پردازش بالاتری می‌باشد. سپس با استفاده از بسته نرم‌افزاری samtools فایل خروجی با فرمت sam را به bam تبدیل می‌کند (Li et al., 2009). برای به‌دست آوردن درصد همردیفی و همپوشانی از دستورflagstat و depth در نرم‌افزار SamTools استفاده شد. فایل واریانت‌های ژنومیک با استفاده از نرم‌افزار freebayes در سامانه برخط (https://github.com/freebayes/freebayes) freebayes به دست آمد (Garrison & Marth, 2012).

مستندسازی واریانت‌های ژنومی

به‌منظور حاشیه‌نگاری، محاسبه شمار هموزیگوت‌ها و هتروزیگوت‌ها در چند ریختی‌های تک نوکلئوتیدی و حذف و اضافه‌های کوچک، و همچنین شمار جهش‌های جابه‌جایی و معکوس در چند ریختی های تک نوکلئوتیدی، از برنامه SnpEff (http://pcingola.github.io/SnpEff/) software version 5.1 استفاده شد. این برنامه از طریق مستندسازی، آثار واریانت‌ها را روی ژن‌ها پیش‌بینی می‌کند.

[1] Single nucleotide variations
[2] Insertions and Deletions
[3] Fluorescent in Situ Hybridization
[4] United States Department of Agriculture
[5] Annotation
[6] Next generation sequencing

[Downloaded from rap.sanru.ac.ir on 2025-04-05]
[DOR: 20.1001.1.22518622.1402.14.41.10.5]
[DOI: 10.61186/rap.14.41.98]

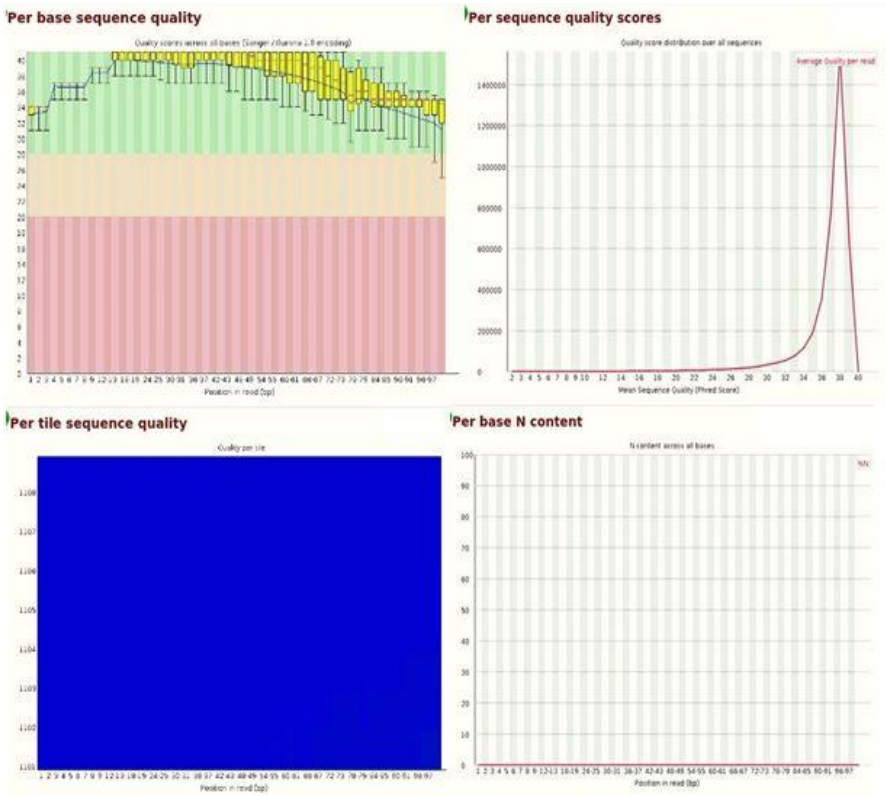
نتایج و بحث

واریانت‌های ژنتیکی نوعی تنوع هستند که در اثر تغییرات مواد ژنتیکی موجودات می‌تواند به نتاج منتقل شود و منجر به تنوع ژنتیکی در سطوح مختلف می‌شود. انواع مختلفی از تنوع ژنتیکی در ژنوم موجودات وجود دارد، از وارونگی کروموزومی میکروسکوپی[1] تا جهش تک نوکلئوتیدی[2]. امروزه با توسعه علم ژنومیک، اطلاعات تنوع ژنتیکی قابل مطالعه مانند پلی‌مورفیسم تک نوکلئوتیدی (SNP)، درج و حذف قطعات کوچک (Indel)، تنوع ساختاری (SV) و تنوع تعداد کپی (CNV) جامع‌تر شده است (Cheng et al., 2014; Shao et al., 2018). ما در این تحقیق کل ژنوم ۵ راس از گاومیش‌های نژاد خوزستانی را با استفاده از فناوری توالی‌یابی ایلومینا مورد بررسی قرار دادیم. همچنین، نوع توالی‌یابی نمونه‌ها به‌صورت دوسرخوانش[3] بود. خوانش‌های به‌دست آمده از توالی‌یابی دوسرخوانش به‌وضوح قدرت تفسیر صحیح نواحی مشکل‌ساز ژنوم، از جمله تکرارهای نامناسب، و تشخیص واریانت‌ها را افزایش می‌دهد. با افزایش اندازه و پیچیدگی ژنوم‌ها، خوانش‌های دوسر کارآمدتر از خواندن‌های یک‌سر[4] هستند، زیرا این روش تنها به مکان‌یابی یک باز از هر جفت باز خوانده شده در ژنوم نیاز دارد تا اکثر خوانش‌ها را به‌درستی مورد چینش قرار دهد (Hillier et al., 2008). با توجه به نتایج کنترل کیفیت اولیه‌ی خوانش‌های به‌دست آمده، تمامی خوانش‌ها برای همه‌ی ۵ نمونه‌ی گاومیش‌ها از کیفیت بالایی برخوردار بودند. میانگین طول خوانش‌های کوتاه برای همهٔ داده‌ها ۱۰۱ بود. درصد همردیفی برای هر ۵ نمونه بالای ۹۷/۴۴ درصد بود که نشان دهنده‌ی کیفیت بالای خوانش‌های کوتاه است. جزئیات سنجش کیفیت توالی‌یابی توسط نرم‌افزار FastQC، در شکل ۱ قابل مشاهده است (شکل ۱). میزان همپوشانی در نمونه‌های توالی‌یابی شده بین ۴/۳ x تا ۱۲/۹ x بود. اطلاعات تکمیلی مربوط به توالی‌یابی در جدول ۱ آورده شده است (جدول ۱).

به‌وسیله‌ی این خوانش‌های کوتاه با کیفیت، همردیفی با ژنوم مرجع انجام شد و واریانت‌های سطح ژنوم گاومیش‌های نژاد خوزستانی محاسبه شد. در نهایت ما تعداد ۷۶.۶۷۶.۵۲۹ میلیون واریانت شناسایی کردیم که از اینها، تعداد ۵۸,۰۳۹,۸۰۹ SNP، تعداد ۶.۲۹۲.۵۷۵ Indels، تعداد ۱۰,۶۱۷,۱۹۵ MNP و تعداد ۱,۷۲۶,۹۲۶ MIXED بودند (جدول ۲). این تعداد از واریانت‌ها به‌طور نسبتاً مساوی در بین کروموزوم‌های گاومیش‌های خوزستانی توزیع شده بود. بیشترین تعداد واریانت‌ها در کروموزوم شماره ۱ مشاهده شد و همچنین کمترین تعداد واریانت‌ها روی کروموزوم شماره ۲۸ توزیع شده بود. در این مطالعه کل جهش‌های جابه‌جایی[5] و معکوس[6] شناسایی شده در سطح کل ژنوم گاومیش‌های خوزستانی به‌ترتیب ۲۹۸.۱۹۳.۰۱۷ و ۱۳۵.۶۹۴.۴۶۶ بود. همچنین، نسبت جهش‌های جابه‌جایی به معکوس ۲/۱۹۸۹ (Ts/Tv) برآورد شد. جایگزینی نوکلئوتیدها می‌تواند به‌صورت جهش جابه‌جایی که جایگزینی یک پورین به‌جای باز نوکلئوتیدی پورین یا یک باز نوکلئوتیدی پیریمیدین به‌جای پیریمیدین است یا به‌صورت جهش معکوس که جایگزینی یک باز نوکلئوتیدی پورین با باز نوکلئوتیدی پیریمیدین یا برعکس است رخ دهد. همانطور که در مطالعه‌ای دیگر نیز گزارش شده است، نرخ بالاتر شاخص ts/tv نشان دهنده‌ی کیفیت بیشتر فراخوانی SNPها است (Bakhtiarizadeh & Alamouti, 2020). علاوه بر این تعداد جهش‌های هموزایگوت در همه‌ی نمونه‌ها بیشتر از جهش‌های هتروزایگوت بود. میزان جهش‌های هتروزایگوت به هموزایگوت (Het/Hem) نیز ۰/۵۵۲ برآورد شد. مستندسازی نتایج نشان داد که از تعداد کل واریانت‌های کشف شده در سطح ژنوم گاومیش‌های خوزستانی، تعداد ۵۴.۱۵۱.۱۳۷ (۶۲/۱۴۷ درصد) در مناطق اینترژنیک، تعداد ۲۴.۱۳۸.۳۶ (۲۷/۷۰۳درصد) در مناطق اینترون‌ها، تعداد ۳.۸۸۹.۹۹۸ (۴/۴۶۴ درصد) در مناطق پایین‌دست ژنی، تعداد ۳.۷۰۸.۲۲۵ (۴/۲۵۶ درصد) در مناطق بالادست ژنی و تعداد ۸۹۷.۷۹۱ (۱۳/۰۳ درصد) در اگزون‌ها شناسایی شدند. کمتر از ۰/۵ درصد از واریانت‌ها نیز در دیگر مناطق واقع شده بودند. همانطور که گزارش شد بیشترین تعداد واریانت‌ها در مناطق اینترژنیک و کمترین تعداد واریانت‌ها در اگزون‌ها واقع شده بودند (شکل ۲).

[1] Microscopic chromosome inversion
[2] Single nucleotide mutation
[3] Paired-end
[4] Single-end
[5] Transition
[6] Transversion

[Downloaded from rap.sanru.ac.ir on 2025-04-05]
[DOR: 20.1001.1.22518622.1402.14.41.10.5]
[DOI: 10.61186/rap.14.41.98]



شکل ۱- خروجی نرم‌افزار FastQC مربوط به سنجش کیفیت توالی‌یابی. چپ بالا: کیفیت توالی به‌ازاء هر باز، چپ پایین: کیفیت توالی به ازاء هر تایل، راست بالا: امتیاز کیفیت توالی، راست پایین: محتوای باز خوانده نشده

Figure 1. Output of FastQC software related to sequencing quality. Upper left: per base sequence quality, lower left: per tile sequence quality, upper right: per sequence quality score, lower right: per base N content

جدول ۱- خلاصه‌ی اطلاعات مربوط به توالی‌یابی کل ژنوم گاومیش‌های رودخانه‌ای خوزستان

Table1. Summary of whole-genome sequencing information related to the Khuzestan river buffalos

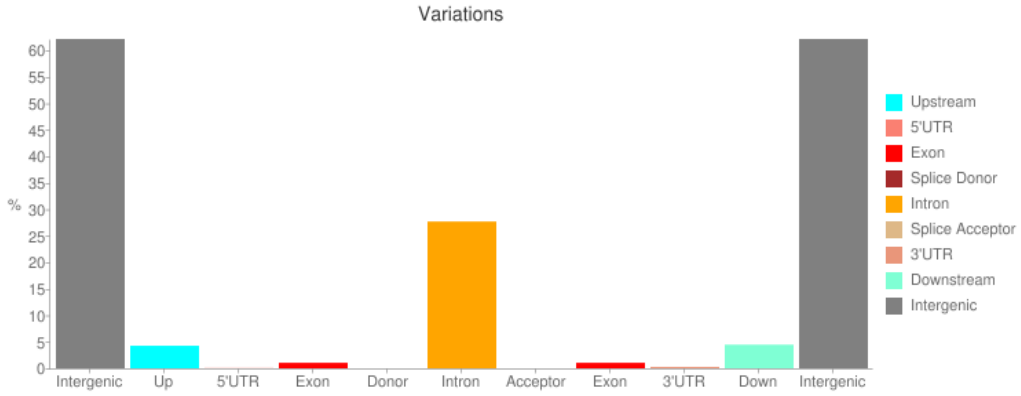
درصد خوانش‌های جفت شده صحیح Properly paired (%)	تعداد خوانش‌های جفت شده صحیح properly paired	خوانش ۲read	خوانش ۱read	خوانش‌های جفت شده در توالی یابی Paired in sequencing	تعداد خوانش‌های مکان‌یابی شده mapped	درصد خوانش‌های مکان‌یابی شده (%) Mapped	خوانش‌های تائید شده در کنترل کیفیت QC- passed reads	
88/75	123082080	69342643	69342643	138685286	137565213	98/28	139976876	نمونه۱ Sample1
87/66	217291552	123933124	123933124	247866248	243824711	97/44	250233580	نمونه۲ Sample2
87/99	112621082	64000000	64000000	128000000	126932078	98/26	129179214	نمونه ۳ Sample3
88/27	141803472	80322575	80322575	160645150	159582606	98/44	162118133	نمونه ۴ Sample4
87/29	111127822	63655606	63655606	127311212	126280842	98/29	128476954	نمونه ۵ Sample5

جدول ۲- تعداد اثرات تنوع‌ها بر اساس شدت

Table 2. Number of variant effects by impact

نوع شدت اثر Type	تعداد Count	درصد Percent
زیاد HIGH	11,055	0.013%
کم LOW	646,635	0.742%
ملایم MODERATE	286,153	0.328%
ملایم کننده MODIFIER	86,190,230	98.917%

[Downloaded from rap.sanru.ac.ir on 2025-04-05]
[DOR: 20.1001.1.22518622.1402.14.41.10.5]
[DOI: 10.61186/rap.14.41.98]



شکل ۲- نمودار درصد پراکنش واریانت‌ها در نواحی مختلف ژنوم گاومیش‌های رودخانه‌ای خوزستان
Figure 2 . The graf of the percentage distribution of variants in different regions of the genome of Khuzestan river buffalos

در این مطالعه تعداد کل جهش‌های خاموش[1] (هم‌معنی) ۵۱۹.۸۶۱ (۷۱/۳۷ درصد)، بی‌معنی[2] ۲۰۷.۱۱۴ (۲۸/۴۳ درصد) و بد معنی[3] ۱.۴۲۷ (۰/۱۹ درصد) به‌دست آمد. در مطالعه‌ای که توسط ژو و همکاران در سال ۲۰۲۰ انجام شد، نشان داده شد که تغییرات تک نوکلئوتیدی غیرمترادف به‌طور بالقوه به هتروزیس و افزایش مقاومت به بیماری‌ها در طول اهلی شدن کمک می‌کنند و ما باید روی SNPهای غیر مترادف و حذف و اضافه‌های کوچک در مناطق کدکننده برای تجزیه و تحلیل بعدی ژن های جهش یافته تمرکز کنیم (Zhou et al., 2020). در این مطالعه، تعداد ۶.۲۹۲.۵۷۵ حذف و اضافه‌ی کوچک با حداقل طول ۱، حداکثر طول ۲۷ جفت باز و میانگین ۱/۴ جفت باز شناسایی شد. حذف و اضافه‌های کوچک به‌عنوان ایجادکننده‌ی تفاوت‌های فنوتیپی در میان گونه‌های مختلف شناخته می‌شوند و همچنین، حضور این حذف و اضافه‌های کوچک در سطح ژنوم می‌تواند قالب خوانش ترجمه ژنوم را تغییر دهد (Abyzov et al., 2011). نتایج آنالیزها نشان داد که بیشتر از ۹۸ درصد واریانت‌ها شدت ملایم کننده[4] داشتند، این نوع از واریانت‌ها تاثیر روی مناطق غیر کدکننده دارند. همچنین کمترین مقدار برای واریانت‌های با شدت زیاد[5] ثبت شد (۰/۰۱۳ درصد) که این نوع از واریانت‌ها دارای اثر مخربی بر روی مسیرهای پروتئین‌سازی و آنزیم‌ها دارند (جدول ۳) (Shirasawa et al., 2013).

جدول ۳- تعداد واریانت‌ها بر اساس نوع
Table 3. Number of variants based on type

نوع واریانت Type	تعداد واریانت Total
SNP	58,039,809
MNP	10,617,195
INS	3,145,321
DEL	3,147,254
MIXED	1,726,936
Total	76,676,515

نتیجه‌گیری کلی

وجود تنوع و استعدادهای ژنتیکی در توده‌های گاومیش بومی ایران، سرمایه‌های ژنی باارزشی هستند که با مطالعه‌ی آن می‌توان تنوع‌های ژنتیکی را شناسایی کرد و پتانسیل ژنتیکی صفات اقتصادی آنها را به‌منظور طراحی برنامه‌های مناسب اصلاح‌نژادی، به‌طور دقیق تعیین کرد. ما در این تحقیق با استفاده از توالی‌یابی کل ژنوم گاومیش‌های نژاد خوزستانی ایران که با تکنولوژی ایلومینا انجام شد، تنوع‌های ژنومی کل ژنوم را شناسایی و معرفی کردیم. تجزیه وتحلیل داده‌های توالی‌یابی نسل جدید تنوع‌های کل ژنوم گاومیش‌ها را به‌خوبی نشان داد و تنوع‌ها تقسیم‌بندی و معرفی شدند. حذف و اضافه‌های کوچک و چندشکلی‌های چند نوکلئوتیدی در جمعیت‌های مختلف منبعی ارزشمند در تحقیقات ژنتیکی هستند که می‌توانند در مکان‌های مختلف ژنومی در کنترل ویژگی‌های مهم اقتصادی موثر باشند. علاوه بر این، شناسایی SNPها در مقیاس‌های بزرگ، می‌توانند در مطالعات ارتباط ژنومی و نیز به‌منظور توسعه آرایه‌های SNP با چگالی بالا برای نژاد گاومیش‌های ایرانی مورد استفاده قرار گیرند.

[1] Silent
[2] Missense
[3] Nonsense
[4] Modifier Impact
[5] High Impact

[Downloaded from rap.sanru.ac.ir on 2025-04-05]
[DOR: 20.1001.1.22518622.1402.14.41.10.5]
[DOI: 10.61186/rap.14.41.98]

منابع

Abyzov, A., Urban, A. E., Snyder, M., & Gerstein, M. (2011). CNVnator: an approach to discover, genotype, and characterize typical and atypical CNVs from family and population genome sequencing. *Genome research*, *21*(6), 974-984.

AGRI. (2021). http://amar.maj.ir

Alkan, C., Coe, B. P., & Eichler, E. E. (2011). Genome structural variation discovery and genotyping. *Nature reviews genetics*, *12*(5), 363-376 .

Andrews, S. (2010). FastQC: a quality control tool for high throughput sequence data. Available online. *Retrieved May*, *17*, 2018 .

Bakhtiarizadeh, M. R., & Alamouti, A. A. (2020). RNA-Seq based genetic variant discovery provides new insights into controlling fat deposition in the tail of sheep. *Scientific Reports*, *10*(1), 13525 .

Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, *30*(15), 2114-2120 .

Borghese, A., Chiariotti, A., & Barile, V. L. (2022). Buffalo in the World: Situation and Perspectives. In *Biotechnological Applications in Buffalo Research* (pp. 3-31). Springer .

Cheng, Z., Lin, J., Lin, T., Xu, M., Huang, Z., Yang, Z., . . . Zheng, J. (2014). Genome-wide analysis of radiation-induced mutations in rice (Oryza sativa L. ssp. indica). *Molecular BioSystems*, *10*(4), 795-805 .

FAOSTAT, F. Agriculture Data 2019 Available online: http://www. fao. org/faostat/en/# data. *QC (accessed on 10 May 2021* .

Garrison, E., & Marth, G. (2012). Haplotype-based variant detection from short-read sequencing. *arXiv preprint arXiv:1207.3907* .

Hillier, L. W., Marth, G. T., Quinlan, A .R., Dooling, D., Fewell, G., Barnett, D., . . . Huang, W. (2008). Whole-genome sequencing and variant discovery in C. elegans. *Nature methods*, *5*(2), 183-188 .

Iamartino, D., Nicolazzi, E. L., Van Tassell, C. P., Reecy, J. M., Fritz-Waters, E. R., Koltes ,J. E., Ajmone-Marsan, P. (2017). Design and validation of a 90K SNP genotyping assay for the water buffalo (Bubalus bubalis). *PloS one*, *12*(10), e0185220 .

Kierstein, G., Vallinoto, M., Silva, A., Schneider, M. P., Iannuzzi, L., & Brenig, B. (2004 .(Analysis of mitochondrial D-loop region casts new light on domestic water buffalo (Bubalus bubalis) phylogeny. *Molecular phylogenetics and evolution*, *30*(2), 308-324 .

Koopaee, H. K., & Koshkoiyeh, A. E. (2014). SNPs genotyping technologies and their applications in farm animals breeding programs. *Brazilian Archives of Biology and Technology*, *57*, 87-95 .

Krausz, C., Chianese, C., Giachini, C., Guarducci, E., Laface, I., & Forti, G. (2011). The Y chromosome-linked copy number variations and male fertility .*Journal of endocrinological investigation*, *34*, 376-382 .

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., . . . Subgroup, G. P. D. P. (2009). The sequence alignment/map format and SAMtools. *bioinformatics*, *25*(16), 2078-2079 .

Lupski ,J. R., & Stankiewicz, P. (2005). Genomic disorders: molecular mechanisms for rearrangements and conveyed phenotypes. *PLoS genetics*, *1*(6), e49 .

Metzker, M. L. (2010). Sequencing technologies–the next generation. *Nature reviews genetics*, *11*(1), 31-46 .

Rehman, S. U., Shafique, L., Yousuf, M. R., Liu, Q., Ahmed, J. Z., & Riaz, H. (2019). Spectrophotometric calibration and comparison of different semen evaluation methods in Nili-Ravi buffalo bulls. *Pak. Vet. J*, *39*(4), 2074-7764 .

Scherf, B. D., & Pilling ,D. (2015). The second report on the state of the world's animal genetic resources for food and agriculture .

Shao, X., Hu, C., Sheng, O., Bi, F., Deng, G., Yang, Q., & Yi, G. (2018). Genomewide variations of triploid banana (AAA group)'grand Nain'by wholegenome resequencing. *Plant Physiol J*, *54*(4), 581-593 .

Shirasawa, K., Fukuoka, H., Matsunaga, H., Kobayashi, Y., Kobayashi, I., Hirakawa, H., Tabata, S. (2013). Genome-wide association studies using single nucleotide polymorphism markers developed by re-sequencing of the genomes of cultivated tomato. *DNA research*, *20*(6), 593-603 .

Zhou, L., Wang, C., Gao, X., Ding, Y., Cheng, B., Zhang, G., Zhang, L. (2020). Genome-wide variations analysis of sorghum cultivar Hongyingzi for brewing Moutai liquor. *Hereditas*, *157*, 1-11 .

[Downloaded from rap.sanru.ac.ir on 2025-04-05]
[DOR: 20.1001.1.22518622.1402.14.41.10.5]
[DOI: 10.61186/rap.14.41.98]